Submitted to the Annals of Applied Probability

CENTRAL LIMIT THEOREMS FOR SIR EPIDEMICS AND PERCOLATION ON CONFIGURATION MODEL RANDOM GRAPHS

By Frank Ball,

University of Nottingham

We consider a stochastic SIR (susceptible → infective → recovered) epidemic defined on a configuration model random graph, in which infective individuals can infect only their neighbours in the graph during an infectious period which has an arbitrary but specified distribution. Central limit theorems for the final size (number of initial susceptibles that become infected) of such an epidemic as the population size n tends to infinity, with explicit, easy to compute expressions for the asymptotic variance, are proved assuming that the degrees are bounded. The results are obtained for both the Molloy-Reed random graph, in which the degrees of individuals are deterministic, and the Newman-Strogatz-Watts random graph, in which the degrees are independent and identically distributed. The central limit theorems cover the cases when the number of initial infectives either (a) tends to infinity or (b) is held fixed as $n \to \infty$. In (a) it is assumed that the fraction of the population that is initially infected converges to a limit (which may be 0) as $n \to \infty$, while in (b) the central limit theorems are conditional upon the occurrence of a large outbreak (more precisely one of size at least $\log n$). Central limit theorems for the size of the largest cluster in bond percolation on Molloy-Reed and Newman-Strogatz-Watts random graphs follow immediately from our results, as do central limit theorems for the size of the giant component of those graphs. Corresponding central limit theorems for site percolation on those graphs are also proved.

1. Introduction. There has been considerable work in the past two decades on models for the spread of epidemics on random networks; see, for example, the recent book Kiss et al. (2017). The usual paradigm is that individuals in a population are represented by nodes in a random graph and infected individuals are able to transmit infection only to their neighbours in the graph. The graph is often constructed using the configuration model (see, for example, van der Hofstad (2016), Chapter 7), which allows for

AMS 2000 subject classifications: Primary 60K35, 92D30, 05C80; secondary 60F05, 60J27, 91D30

Keywords and phrases: Bond and site percolation, central limit theorem, configuration model, density dependent population process, random graph, SIR epidemic, size of epidemic

an arbitrary but specified degree distribution. The most-studied type of epidemic model is the SIR (susceptible $\to$ infective $\to$ recovered) model. In this model individuals are classified into three types: susceptibles, infectives and recovered. If a susceptible individual is contacted by an infective then it too becomes an infective and remains so for a time, called its infectious period, that is distributed according to a non-negative random variable I having an arbitrary but specified distribution. An infective individual recovers at the end of its infectious period and is then immune to further infection. During its infectious period, an infective contacts its susceptible neighbours in the graph independently at the points of Poisson processes having rate λ. The graph is assumed to be static and the population closed (i.e. there are no births or deaths), so eventually the epidemic process terminates. The final size of the epidemic is the number of initial susceptibles that are infected during its course. The final size is a key epidemic statistic, not only as a measure of the impact of an epidemic but also in an inferential setting, since often it can be observed more reliably than the precise temporal spread. The main aim of this paper is to develop central limit theorems for the final size of an SIR epidemic on configuration model graphs as the population size $n \to \infty$.

The configuration model, which was introduced by Bollobás (1980), is a model for random graphs with a given degree sequence. There are two distinct approaches for constructing configuration model graphs with a given degree distribution as $n \to \infty$. In both approaches, individuals are assigned a number of half-edges, corresponding to their degree, and then these half-edges are paired uniformly at random. In Molloy and Reed (1995), the degrees of individuals are prescribed deterministically whilst in Newman et al. (2001) they are i.i.d. (independent and identically distributed) copies of a random variable D, that describes the limiting degree distribution. We refer to the former as the MR random graph and to the latter as the NSW random graph. Subject to suitable conditions on the degree sequences in the MR model and D in the NSW model, law of large number limits for SIR epidemics on the two graphs are the same. That is not the case for central limit theorems as, for finite n, there is greater variability in the degrees of individuals in the NSW model than in the MR model; indeed, in the NSW model, such variability is of the same order of magnitude as the variability in the epidemic. Thus, though the asymptotic means are the same, the asymptotic variances in the central limit theorems for final size are greater for the epidemic on the NSW random graph.

There have been numerous studies, some fully rigorous and some heuristic, of SIR epidemics on configuration model networks making various as-

sumptions concerning the infectious period random variable I. For example, assuming I is constant, Andersson (1998) derives a law of large numbers for the final size of an epidemic on an NSW random graph when a strictly positive fraction of the population is initially infected in the limit as $n \to \infty$ and Britton et al. (2007) obtain a similar result for epidemics on an MR random graph initiated by a single infective. In the latter case, a large outbreak is possible only if the basic reproduction number $R_0 > 1$; see (2.8) in Section 2.4. In a highly influential paper, Newman (2002) uses heuristic percolation arguments to obtain a number of results, including the fraction of the population infected by a large outbreak, for SIR epidemics on NSW random graphs with I having an arbitrary but specified distribution. Several authors have studied the case when I has an exponential distribution, so the model becomes Markovian. Decreusefond et al. (2012) obtain a law of large numbers type result for the epidemic process on an NSW random graph with a strictly positive fraction initially infected, which yields a rigorous proof of the deterministic approximation of Volz (2008) (see also Miller (2011) and Miller et al. (2012)). Bohman and Picollelli (2012) obtain law of large numbers results for both the process and final size of an epidemic with one initial infective on an MR random graph with bounded degrees. Janson et al. (2014) obtain similar results under weaker conditions on the degree sequences considering the cases when the fraction initially infected, in the limit as $n \to \infty$, is either strictly positive or zero (assuming of course there is at least one initial infective). In the latter case, the limiting "deterministic" process involves a random time translation reflecting the time taken for the number of infectives to reach order n; a similar result is obtained by Barbour and Reinert (2013) assuming a bounded degree sequence and an arbitrary but specified distribution for I.

There has been very little work to date on central limit theorems for SIR epidemics on configuration model networks. A functional central limit theorem for the SI epidemic (in which $\mathrm{P}(I = \infty) = 1$, so infectives remain infectious forever) on an MR random graph with unbounded degrees is obtained by KhudaBukhsh et al. (2017), who note that their method is not straightforward to extend to an SIR model. Assuming that I follows an exponential distribution and bounded degrees, Ball et al. (2019) use an effective degree approach (Ball and Neal (2008)) and density dependent population processes (Ethier and Kurtz (1986), Chapter 11) to obtain functional central limit theorems for SIR epidemics on MR and NSW random graphs, in which susceptible individuals can also drop their edges to infective neighbours. They also conjecture central limit theorems for the final size of such epidemics (and hence as a special case for the final size of standard SIR

epidemics, without dropping of edges), assuming either a strictly positive fraction or a constant number of initial infectives (in which case the central limit theorem is conditional on the occurrence of a large outbreak). However, the arguments are not fully rigorous and the result for a constant number of initial infectives is based purely on the existence of equivalent results for other (non-network) SIR epidemic models. Another limitation of Ball et al. (2019) is the assumption that I is exponentially distributed, which is unrealistic for most real-life diseases.

In the present paper, we address these shortcomings and derive fully rigorous central limit theorems for the final size of SIR epidemics on MR and NSW random graphs having bounded degrees, when the infectious period I follows an arbitrary but specified distribution. We consider the cases when the limiting fraction of the population is (i) strictly positive and (ii) zero. For the latter we treat the situations where the number of initial infectives either (i) is held fixed independent of n or (ii) tends to ∞ as $n \to \infty$. The mean parameter ρ in the central limit theorems, which coincides with the corresponding law of large numbers limit, depends on the solution z of a non-linear equation (see (2.2) and (2.9) in Section 2.4). Given z, the variance parameter in the central limit theorems is fully explicit and hence easy to compute.

If I is constant, say $\mathrm{P}(I = 1) = 1$ then the above SIR model is essentially bond percolation with probability $\pi = 1 - \mathrm{e}^{-\lambda}$ and if $\mathrm{P}(I = \infty) = \pi = 1 - \mathrm{P}(I = 0)$ then it is closely related to site percolation with probability π; see, for example, Durrett (2007), page 15, and Janson (2009a). Central limit theorems for the size of the giant component (largest cluster) of bond percolation on the MR and NSW random graphs follow immediately from our results. (Corresponding theorems for site percolation are also obtained using our methodology.) Further, setting $\pi = 1$ yields central limit theorems for the giant component of those graphs (Remark 2.6); cf. Barbour and Röllin (2019) who obtain a central limit theorem for the giant component of the MR random graph, and Ball and Neal (2017) and Janson (2018) who derive respectively the asymptotic variance and a central limit theorem for the giant component of MR and NSW random graphs, all allowing for unbounded degrees.

The proofs involve constructing the random graph and epidemic on it simultaneously, modifying the infection mechanism so that when a susceptible is infected it decides which of its half-edges it will try to infect along (its remaining half-edges become recovered half-edges), with the times of those infection attempts (relative to the time of infection of the susceptible) being realisations of i.i.d. exponential random variables. The distribution of

the final outcome of the epidemic, and hence also its final size, is invariant to this modification. The process describing the evolution of the numbers of susceptibles of different degrees, infective half-edges and recovered half-edges is an asymptotically density dependent population process (Ethier and Kurtz (1986), Chapter 11, and Pollett (1990)). The asymptotic distribution of the final outcome of the epidemic is studied by considering a boundary crossing problem for a random time-scale transformation of that process. The proofs extend, at least in principle, to SIR epidemics and percolation on extensions of the configuration model that include fully-connected cliques (Trapman (2007), Gleeson (2009), Ball et al. (2010) and Coupechoux and Lelarge (2014)), though explicit calculation of the asymptotic variances may be difficult.

The remainder of the paper is organised as follows. The MR and NSW random graphs are defined in Section 2.1 and the SIR epidemic model is described in Section 2.2. The main central limit theorems (Theorems 2.1-2.3 for SIR epidemics and Theorem 2.7 for percolation) are stated in Section 2.4, together with some remarks giving comparisons of their variance parameters and applications to giant components indicated above. Some numerical illustrations, which show that the central limit theorems can yield good approximations even for relatively small graphs, are given in Section 3. The proofs are given in Section 5. They make extensive use of asymptotically density dependent population processes and in particular require a version of the functional central limit theorem for such processes to include asymptotically random initial conditions. For ease of reference, the required results for such processes are collected together in Section 4. Some brief concluding comments are given in Section 6. Calculation of the asymptotic variances for the central limit theorems is lengthy, though straightforward, so this and a few other details are deferred to an appendix.

1.1. *Notation.* All vectors are row vectors and $^\top$ denotes transpose. With the dimension being obvious from the context, I denotes an identity matrix and $\mathbf{0}$ and $\mathbf{1}$ denotes vectors all of whose elements are 0 and 1, respectively. For $x \in \mathbb{R}$, the usual floor and ceiling functions are denoted by $\lfloor x \rfloor$ and $\lceil x \rceil$, respectively. Thus $\lfloor x \rfloor$ is the greatest integer $\leq x$ and $\lceil x \rceil$ is the smallest integer $\geq x$. For a positive integer k, the kth derivative of a real-valued function f is denoted by $f^{(k)}$. The cardinality of a set A is denoted by $|A|$. Sums are zero if vacuous. We use $\xrightarrow{p}$, $\xrightarrow{a.s.}$ and $\xrightarrow{D}$ to denote convergence in probability, convergence almost sure and convergence in distribution, respectively.

Further, $\mathrm{U}(0,1)$ denotes a uniform random variable on $(0,1)$; $\mathrm{Exp}(1)$ de-

notes an exponential random variable with mean 1; $\mathrm{N}(0,\sigma^2)$ denotes a univariate normal random variable with mean 0 and variance σ^2; and $\mathrm{N}(\mathbf{0},\Sigma)$ denotes a multivariate normal random variable with mean vector $\mathbf{0}$ and variance matrix Σ, whose dimension again is obvious from the context. For a positive integer n and $p \in [0,1]$, $\mathrm{Bin}(n,p)$ denotes a binomial random variable with n trials and success probability p. Also, if I is a non-negative random variable and $\lambda \in (0,\infty)$ then $\mathrm{Bin}(n, 1-\mathrm{e}^{-\lambda I})$ denotes a mixed-Binomial random variable obtained by first sampling I_1 from the distribution of I and then, given I_1, sampling independently from $\mathrm{Bin}(n, 1-\mathrm{e}^{-\lambda I_1})$. Similarly, if I is a non-negative random variable and D is a non-negative integer-valued random variable, then $\mathrm{Bin}(D, 1-\mathrm{e}^{-\lambda I})$ denotes a mixed-Binomial random variable, where the realisations of D and I are independent. Thus, if $X \sim \mathrm{Bin}(D, 1-\mathrm{e}^{-\lambda I})$, then

$$\mathrm{P}(X=k) = \sum_{d=k}^{\infty} \mathrm{P}(D=d)\mathrm{E}\left[\binom{d}{k}(1-\mathrm{e}^{-\lambda I})^k \mathrm{e}^{-(d-k)\lambda I}\right] \qquad (k=0,1,\dots).$$

Note that we allow the possibility $D=0$.

2. Model and main results.

2.1. *Random graph.* Consider a population of n individuals labelled $1,2,\dots,n$. For $i=1,2,\dots,n$, let $D_i^{(n)}$ denote the degree of individual i. We assume that $0 \le D_i^{(n)} \le d_{\max}$ for all i, i.e. that there is a maximum degree $d_{\max}$. In the MR random graph the degrees are prescribed, while in the NSW random graph $D_1^{(n)}, D_2^{(n)}, \dots, D_n^{(n)}$ are i.i.d. copies of a random variable D having probability mass function given by $\mathrm{P}(D=i) = p_i$ $(i=0,1,\dots,d_{\max})$. Under both models, the network (random graph) is formed by attaching $D_i^{(n)}$ half-edges to individual i, for $i=1,2,\dots,n$, and then pairing up the $D_1^{(n)} + D_2^{(n)} + \cdots + D_n^{(n)}$ half-edges uniformly at random to give the edges in the random graph, which we denote by $\mathcal{G}^{(n)}$. In the NSW model, if $D_1^{(n)} + D_2^{(n)} + \cdots + D_n^{(n)}$ is odd there is a left-over stub, which is ignored. (Of course in the MR model the prescribed degrees can be chosen so that $D_1^{(n)} + D_2^{(n)} + \cdots + D_n^{(n)}$ is even.)

We are interested in asymptotic results as the number of individuals $n \to \infty$. In the MR random graph, for $i=0,1,\dots,d_{\max}$, let $v_i^{(n)} = \sum_{k=1}^{n} 1_{\{D_k^{(n)}=i\}}$ be the number of individuals having degree i. We assume that

$$\lim_{n\to\infty} \sqrt{n}\left(n^{-1}v_i^{(n)} - p_i\right) = 0 \qquad (i=0,1,\dots,d_{\max}). \tag{2.1}$$

Note that (2.1) implies $\lim_{n\to\infty} n^{-1}v_i^{(n)} = p_i$ $(i = 0, 1, \dots, d_{\max})$.

In both models the random graph may have some imperfections, specifically self-loops and multiple edges, but they are sparse in the network as $n \to \infty$; more precisely, the number of such imperfections converges in distribution to a Poisson random variable as $n \to \infty$ (Durrett (2007), Theorem 3.1.2, Janson (2009b) and Janson (2014)). It follows that law of large numbers results continue to hold if the graph is conditioned on being simple. However, that is not necessarily the case for convergence in distribution; see Janson (2010), Remark 1.4, and Barbour and Röllin (2019), Remark 2.5. Very recently, Janson (2019) gives conditions under which a switching construction can be used to transfer results on convergence in distribution to graphs conditioned on being simple. In Section 5.7, we outline how they can be used to show that our central limit theorems continue to hold when the random graph is conditioned on being simple.

2.2. *SIR epidemic.* An SIR epidemic, denoted by $\mathcal{E}^{(n)}$, is constructed on the above network as follows. Initially, at time $t = 0$, a number of individuals are infective and the remaining individuals are susceptible. (Precise statements concerning the initial infectives are made later.) Distinct infectives behave independently of each other and of the construction of the network. Each infective remains infectious for a period of time that is distributed according to a random variable I, having an arbitrary but specified distribution, after which it becomes recovered. During its infectious period, an infective contacts its neighbours in the network independently at the points of Poisson processes each having rate λ. Thus the individual-to-individual infection rate is λ and the probability that a given neighbour is contacted is $p_I = 1 - \phi(\lambda)$, where $\phi(\theta) = \mathrm{E}[\exp(-\theta I)]$ $(\theta \geq 0)$ is the Laplace transform of I. If a contacted individual is susceptible then it becomes an infective, otherwise nothing happens. The epidemic ends when there is no infective individual in the population.

For $t \geq 0$, let $X_i^{(n)}(t)$ be the number of degree-i susceptible individuals at time t $(i = 0, 1, \dots, d_{\max})$ and let $Y^{(n)}(t)$ be the total number of infectives at time t. Let $\tau^{(n)} = \inf\{t \geq 0 : Y^{(n)}(t) = 0\}$ be the time of the end of the epidemic. Then $T_i^{(n)} = X_i^{(n)}(0) - X_i^{(n)}(\tau^{(n)})$ is the total number of degree-i susceptibles that are infected by the epidemic. Let $T^{(n)} = \sum_{i=0}^{d_{\max}} T_i^{(n)}$ be the total number of susceptibles infected by the epidemic, i.e. the final size of the epidemic. We are primarily interested in the asymptotic distribution of $T^{(n)}$ as $n \to \infty$.

The proofs allow for the possibility that $p_I = 1$, i.e. $\mathrm{P}(I = \infty) = 1$. In that case the set of individuals that are infected during the epidemic $\mathcal{E}^{(n)}$

comprises all individuals in the components of $\mathcal{G}^{(n)}$ that contain at least one initial infective. Thus central limit theorems for the size of the giant component in MR and NSW random graphs follow immediately from our results; see Remark 2.6 below.

2.3. *Bond and site percolation.* In bond percolation on $\mathcal{G}^{(n)}$, each edge in $\mathcal{G}^{(n)}$ is deleted independently with probability $1-\pi$, while in site percolation $\mathcal{G}^{(n)}$, each vertex (together with all incident edges) is deleted independently with probability $1-\pi$ (Janson (2009a)). Interest is often focused on the size, $C^{(n)}$ say, of the largest connected component in the resulting graph.

2.4. *Main results.* For $i = 0, 1, \ldots, d_{\max}$, let $a_i^{(n)}$ be the number of degree-i initial infectives in the epidemic $\mathcal{E}^{(n)}$ and let $a^{(n)} = \sum_{i=0}^{d_{\max}} a_i^{(n)}$ denote the total number of initial infectives. In the epidemic on the MR random graph, we assume that $a_0^{(n)}, a_1^{(n)}, \ldots, a_{d_{\max}}^{(n)}$ are prescribed. In the epidemic on the NSW random graph, we assume that $a^{(n)}$ is prescribed and that the $a^{(n)}$ initial infectives are chosen by sampling uniformly at random without replacement from the n individuals in the population.

Let $\epsilon^{(n)} = n^{-1}a^{(n)}$ and $\epsilon_i^{(n)} = n^{-1}a_i^{(n)}$ $(i = 0, 1, \ldots, d_{\max})$. Suppose that $\epsilon^{(n)} \to \epsilon$ as $n \to \infty$ and that $\lim_{n\to\infty} \sqrt{n}(\epsilon^{(n)} - \epsilon) = 0$. For the epidemic on the MR random graph, suppose further that, for $i = 0, 1, \ldots, d_{\max}$, there exists ϵ_i such that $\lim_{n\to\infty} \sqrt{n}(\epsilon_i^{(n)} - \epsilon_i) = 0$. For the epidemic on the NSW random graph, let $\epsilon_i = \epsilon p_i$ $(i = 0, 1, \ldots, d_{\max})$. Let $\mu_D = \mathrm{E}[D]$, $\sigma_D^2 = \mathrm{var}(D)$ and, for $s \in [0, 1]$, $f_D(s) = \sum_{i=0}^{d_{\max}} p_i s^i$ and $f_{D_\epsilon}(s) = \sum_{i=0}^{d_{\max}} (p_i - \epsilon_i)s^i$. Let $q_I = 1 - p_I = \phi(\lambda)$ be the probability that an infective fails to contact a given neighbour and $q_I^{(2)} = \phi(2\lambda)$ be the probability that a given infective fails to contact two given neighbours.

The first theorem concerns the case $\epsilon > 0$, so in the limit as $n \to \infty$ a strictly positive fraction of the population is initially infective. Let $T_{\mathrm{MR}}^{(n)}$ and $T_{\mathrm{NSW}}^{(n)}$ denote the final size of the epidemic $\mathcal{E}^{(n)}$ on the MR and NSW random graphs, respectively. Let z be the unique solution in $[0, 1)$ of

$$z - q_I = \mu_D^{-1} p_I f_{D_\epsilon}^{(1)}(z) \tag{2.2}$$

and

$$\rho = 1 - \epsilon - f_{D_\epsilon}(z). \tag{2.3}$$

THEOREM 2.1. *Suppose that $\epsilon \in (0, 1)$, $p_i\epsilon_i > 0$ for at least one $i > 0$ and that $p_1 - \epsilon_1 > 0$ if $p_I = 1$. Then, as $n \to \infty$,*

$$\sqrt{n}\left(n^{-1}T_{\mathrm{MR}}^{(n)} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \sigma_{\mathrm{MR}}^2)$$

and

$$\sqrt{n}\left(n^{-1}T^{(n)}_{\mathrm{NSW}}-\rho\right)\xrightarrow{D}\mathrm{N}(0,\sigma^2_{\mathrm{NSW}}),$$

where

$$\begin{aligned}\sigma^2_{\mathrm{MR}} &= h(z)^2\left\{\left[(p_Iq_I+2(z-q_I)^2\right]\mu_D-p_I^2\left[f^{(1)}_{D_\epsilon}(z^2)+z^2f^{(2)}_{D_\epsilon}(z^2)\right]\right\}\\ &\quad+h(z)\left[2p_Izf^{(1)}_{D_\epsilon}(z^2)-(z-q_I)\mu_D\right]+1-\epsilon-\rho-f_{D_\epsilon}(z^2)\\ &\quad+\left(q_I^{(2)}-q_I^2\right)h(z)^2\left[f^{(2)}_D(1)-f^{(2)}_{D_\epsilon}(z)\right],\end{aligned}\tag{2.4}$$

with

$$h(z)=\frac{p_I^{-1}(q_I-z)}{1-p_I\mu_D^{-1}f^{(2)}_{D_\epsilon}(z)},\tag{2.5}$$

and

$$\begin{aligned}\sigma^2_{\mathrm{NSW}} &= \frac{\rho(1-\epsilon-\rho)}{1-\epsilon}-h(z)(z-q_I)\left(1-2\epsilon+\frac{2\epsilon\rho}{1-\epsilon}\right)\mu_D\\ &\quad+h(z)^2\left\{\left[p_Iq_I-2z(z-q_I)\right]\mu_D+(z-q_I)^2\left(\sigma_D^2+\frac{1-2\epsilon}{1-\epsilon}\mu_D^2\right)\right\}\\ &\quad+\left(q_I^{(2)}-q_I^2\right)h(z)^2\left[f^{(2)}_D(1)-(1-\epsilon)f^{(2)}_D(z)\right],\end{aligned}\tag{2.6}$$

with

$$h(z)=\frac{p_I^{-1}(q_I-z)}{1-p_I\mu_D^{-1}(1-\epsilon)f^{(2)}_D(z)}.\tag{2.7}$$

The second theorem concerns the case when the number of initial infectives is held fixed as $n\to\infty$, so $\epsilon=0$. More specifically, in the epidemic on the MR random graph, we assume that $a_i^{(n)}=a_i$ $(i=0,1,\ldots,d_{\max})$ for all $n\geq a=\sum_{i=1}^{d_{\max}}a_i$ and in the epidemic on the NSW random graph, we assume that $a^{(n)}=a$ for all $n\geq a$. It is well known that, for large n, the process of infectives in the early stages of such an epidemic can be approximated by a Galton-Watson branching process, $\mathcal{B}$ say, in which, except for the initial generation, the offspring distribution is $\mathrm{Bin}(\tilde{D}-1,1-\mathrm{e}^{-\lambda I})$, where $\tilde{D}$ and I are independent and $\tilde{D}$ has the size-biased degree distribution $\mathrm{P}(\tilde{D}=k)=\mu_D^{-1}kp_k$ $(k=1,2,\ldots,d_{\max})$; see, for example, Ball and Sirl (2013). This offspring distribution has mean

$$R_0=\mu_{\tilde{D}-1}p_I=\left(\mu_D+\mu_D^{-1}\sigma_D^2-1\right)p_I,\tag{2.8}$$

where $\mu_{\tilde{D}-1} = \mathrm{E}[\tilde{D} - 1]$. The quantity R_0 is called the basic reproduction number of the epidemic.

For $\mathcal{E}^{(n)}$, we say that a major outbreak occurs if and only if the event $G^{(n)} = \{T^{(n)} \geq \log n\}$ occurs. Now $\lim_{n\to\infty} \mathrm{P}(G^{(n)}) = \mathrm{P}(B = \infty)$, where B is the total progeny (not including the initial generation) of the branching process $\mathcal{B}$ (cf. Theorem 5.1). Thus, in the limit as $n \to \infty$, a major outbreak occurs with non-zero probability if and only if $R_0 > 1$.

Suppose that that $R_0 > 1$. Now let z be the unique solution in $[0, 1)$ of

$$z - q_I = \mu_D^{-1} p_I f_D^{(1)}(z) \tag{2.9}$$

and $\rho = 1 - f_D(z)$. Note that if $p_I \in (0, 1)$, or $p_I = 1$ and $p_1 > 0$, then $z > 0$.

THEOREM 2.2. *Suppose that $R_0 > 1$ and, if $p_I = 1$ then $p_1 > 0$. Then, as $n \to \infty$,*

$$\sqrt{n}\left(n^{-1}T_{\mathrm{MR}}^{(n)} - \rho\right) | G^{(n)} \xrightarrow{D} \mathrm{N}(0, \tilde{\sigma}_{\mathrm{MR}}^2)$$

and

$$\sqrt{n}\left(n^{-1}T_{\mathrm{NSW}}^{(n)} - \rho\right) | G^{(n)} \xrightarrow{D} \mathrm{N}(0, \tilde{\sigma}_{\mathrm{NSW}}^2),$$

where $\tilde{\sigma}_{\mathrm{MR}}^2$ is given by (2.4) and (2.5), with f_{D_ϵ} replaced by f_D, and $\tilde{\sigma}_{\mathrm{NSW}}^2$ is obtained by setting $\epsilon = 0$ in (2.6) and (2.7).

The next theorem concerns the case when $\epsilon = 0$ but the number of initial infectives $a^{(n)} \to \infty$ as $n \to \infty$.

THEOREM 2.3. *Suppose that $R_0 > 1$, $\epsilon = 0$, $\sum_{i=1}^{d_{\max}} a_i^{(n)} \to \infty$ as $n \to \infty$ and $p_1 - \epsilon_1 > 0$ if $p_I = 1$. Then, as $n \to \infty$,*

$$\sqrt{n}\left(n^{-1}T_{\mathrm{MR}}^{(n)} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \tilde{\sigma}_{\mathrm{MR}}^2)$$

and

$$\sqrt{n}\left(n^{-1}T_{\mathrm{NSW}}^{(n)} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \tilde{\sigma}_{\mathrm{NSW}}^2),$$

where $\tilde{\sigma}_{\mathrm{MR}}^2$ and $\tilde{\sigma}_{\mathrm{NSW}}^2$ are as in Theorem 2.2.

REMARK 2.4. *Note that $q_I^{(2)} = q_I^2$ if I is almost surely constant, otherwise $q_I^{(2)} > q_I^2$ by Jensen's inequality. Also $f_D^{(2)}(1) - f_{D_\epsilon}^{(2)}(z) > 0$, so as one would expect on intuitive grounds, if p_I is held fixed, the asymptotic variance σ_{MR}^2 is smallest when the infectious period is constant. A similar comment holds for $\sigma_{\mathrm{NSW}}^2, \tilde{\sigma}_{\mathrm{MR}}^2$ and $\tilde{\sigma}_{\mathrm{NSW}}^2$.*

REMARK 2.5. *Although it is not transparent from* (2.4) *and* (2.6)*, it is seen easily from the proof that, again as one would expect on intuitive grounds,* $\sigma^2_{\mathrm{NSW}} \geq \sigma^2_{\mathrm{MR}}$ *and* $\tilde{\sigma}^2_{\mathrm{NSW}} \geq \tilde{\sigma}^2_{\mathrm{MR}}$*, with strict inequalities unless the support of the degree random variable* D *is concentrated on a single point (when the two models are identical); see Appendix B.4.4.*

REMARK 2.6. *In the setting of Theorem 2.2, if* $a^{(n)} = 1$ *and* $p_I = 1$ *then with probability tending to* 1 *as* $n \to \infty$*, the event* $G^{(n)}$ *occurs if and only if the initial infective belongs to the giant component of* $\mathcal{G}^{(n)}$*. Thus setting* $p_I = 1$ *in Theorem 2.2 yields central limit theorems for giant components of MR and NSW random graphs.*

The final theorem is concerned with percolation. Let R_0 be given by (2.8) with $p_I = \pi$. Let $C^{(n)}_{\mathrm{MR,B}}$ and $C^{(n)}_{\mathrm{MR,S}}$ denote the size of the largest connected component after bond and site percolation, respectively, on an MR graph; define $C^{(n)}_{\mathrm{NSW,B}}$ and $C^{(n)}_{\mathrm{NSW,S}}$ analogously for the NSW graph. Then, if $R_0 > 1$, for each of these four choices for $C^{(n)}$, there exists $\epsilon > 0$ such that $\lim_{n\to\infty} \mathrm{P}(C^{(n)} \geq \epsilon n) = 1$; see Janson (2009a), Theorems 3.5 and 3.9, which also give law of large number limits for the $C^{(n)}$ under weaker conditions than here. The following theorem gives associated central limit theorems.

THEOREM 2.7. *Suppose that* $\pi \in (0,1)$ *and* $R_0 > 1$*. Let* z *be the unique solution in* $(0,1)$ *of*

$$z - 1 + \pi = \mu_D^{-1} \pi f_D^{(1)}(z) \tag{2.10}$$

and $\rho = 1 - f_D(z)$*. Then, as* $n \to \infty$*,*

$$\sqrt{n}\left(n^{-1}C^{(n)}_{\mathrm{MR,B}} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \sigma^2_{\mathrm{MR,B}}), \tag{2.11}$$

$$\sqrt{n}\left(n^{-1}C^{(n)}_{\mathrm{NSW,B}} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \sigma^2_{\mathrm{NSW,B}}), \tag{2.12}$$

$$\sqrt{n}\left(n^{-1}C^{(n)}_{\mathrm{MR,S}} - \pi\rho\right) \xrightarrow{D} \mathrm{N}(0, \sigma^2_{\mathrm{MR,S}}), \tag{2.13}$$

and

$$\sqrt{n}\left(n^{-1}C^{(n)}_{\mathrm{NSW,S}} - \pi\rho\right) \xrightarrow{D} \mathrm{N}(0, \sigma^2_{\mathrm{NSW,S}}), \tag{2.14}$$

where

(2.15)
$$\sigma^2_{\mathrm{MR,B}} = h(z)^2\left\{\left[\pi(1-\pi)+2(z-1+\pi)^2\right]\mu_D - \pi^2\left[f_D^{(1)}(z^2)+z^2 f_D^{(2)}(z^2)\right]\right\} + h(z)\left[2\pi z f_D^{(1)}(z^2) - (z-1+\pi)\mu_D\right] + 1 - \rho - f_D(z^2)$$

and

(2.16)
$$\sigma^2_{\mathrm{NSW,B}} = \rho(1-\rho) - h(z)(z-1+\pi)\mu_D + h(z)^2\left\{\left[\pi(1-\pi)-2z(z-1+\pi)\right]\mu_D + (z-1+\pi)^2\left(\sigma_D^2+\mu_D^2\right)\right\},$$

with

$$h(z) = \frac{\pi^{-1}(1-\pi-z)}{1-\pi\mu_D^{-1}f_D^{(2)}(z)};$$

(2.17)
$$\sigma^2_{\mathrm{MR,S}} = \pi^2\left\{\sigma^2_{\mathrm{MR,B}} + \pi(1-\pi)h(z)^2\left[f_D^{(2)}(1) - f_D^{(2)}(z)\right]\right\}$$

(2.18)
$$+\pi(1-\pi)[\rho - 2h(z)(1-z)\mu_D]$$

and $\sigma^2_{\mathrm{NSW,S}}$ *is given by* (2.17)*, with* $\sigma^2_{\mathrm{MR,B}}$ *replaced by* $\sigma^2_{\mathrm{NSW,B}}$*.*

3. Illustrations. In this section, we illustrate the central limit theorems in Section 2.4 by using simulations to explore their applicability to graphs with finite n. We also investigate briefly the dependence of the limiting variances in the central limit theorems on the degree distribution, graph type and infectious period distribution. We consider four degree distributions:

(i) $D \equiv d$, i.e. D is constant with $p_d = 1$;
(ii) $D \sim \mathrm{Po}(\mu_D)$, i.e. D is Poisson with mean μ_D;
(iii) $D \sim \mathrm{Geom}(p)$, i.e. $p_k = (1-p)^k p$ $(k = 0, 1, \dots)$;
(iv) $D \sim \mathrm{Power}(\alpha, \kappa)$, i.e. $p_k = ck^{-\alpha}\mathrm{e}^{-\frac{k}{\kappa}}$ $(k = 1, 2, \dots)$, where $\alpha, \kappa \in (0, \infty)$ and the normalising constant $c = \mathrm{Li}_\alpha(\mathrm{e}^{-\frac{1}{\kappa}})$, with Li_α being the polylogarithm function.

The fourth distribution is a power law with exponential cut-off (see, for example, Newman (2002)) that has been used extensively in the physics literature. Note that, with $\theta = \mathrm{e}^{-\frac{1}{\kappa}}$ and $\beta = \mathrm{Li}_\alpha(\theta)$, $f_D(s) = \beta^{-1}\mathrm{Li}_\alpha(\theta s)$, $f_D^{(1)}(s) = \frac{1}{\beta s}\mathrm{Li}_{\alpha-1}(\theta s)$ and $f_D^{(2)}(s) = \frac{1}{\beta s^2}[\mathrm{Li}_{\alpha-2}(\theta s) - \mathrm{Li}_{\alpha-1}(\theta s)]$, which enables R_0 and the asymptotic means and variances in the central limit theorems to be calculated. Distributions (ii)-(iv) have unbounded support, so

do not satisfy the conditions of our theorems. The asymptotic distributions in this section are calculated under the assumption that the theorems still apply.

Each simulation consists of first simulating one graph, by simulating $D_1^{(n)}, D_2^{(n)}, \ldots, D_n^{(n)}$ and pairing up the half-edges, and then simulating one epidemic, or percolation process, on it. For an MR graph with (limiting) degree random variable D on n nodes, the degrees are given by $D_i^{(n)} = \inf\{d \in \mathbb{Z}_+ : F_D(d) > i/(n+1)\}$ $(i = 1, 2, \ldots, n)$, where F_D is the distribution function of D. For heavy-tailed D, in particular, this choice of MR degrees converges faster with n to the intended D than one based on rounding np_i $(i = 0, 1, \ldots)$ to nearest integers. Two choices of infectious period distribution are used in the simulations: (i) $I \equiv 1$ (i.e. $\mathrm{P}(I = 1) = 1$) and (ii) I is 0 or ∞, matched to have a common p_I. Note that these yield the minimum and maximum asymptotic variances for fixed p_I. The parameters of the degree distributions are chosen so that $\mu_D = 5$; hence the choice of $(\alpha, \kappa) = (1, 13.796)$ for the power law with exponential cut-off distribution.

We first consider epidemics on an NSW network in which a fraction $\epsilon = 0.05$ of the population is initially infective. Table 1 shows estimates of $\rho_n = n^{-1}\mathrm{E}[T_{\mathrm{NSW}}^{(n)}]$ and $\sigma_n^2 = n^{-1}\mathrm{var}[T_{\mathrm{NSW}}^{(n)}]$ for epidemics with $p_I = 0.3$ and various population size n, based on 100,000 simulations for each set of parameters, together with the corresponding asymptotic $(n = \infty)$ values given by Theorem 2.1. The table indicates that the asymptotic approximations are useful for even moderate n. The approximations are better for the model with $I \equiv 1$ than that with $I = 0$ or ∞, as one might expect since there is less variability in the process with $I \equiv 1$, and improve with increasing σ_D^2. The approximations are noticeably worse when $n = 200$ than for the other values of n. Histograms of the final size of 100,000 simulated epidemics on an NSW network with $n = 200$ and $D \sim \mathrm{Geom}(1/6)$, together with the corresponding density of $\mathrm{N}(n\rho, n\sigma_{\mathrm{NSW}}^2)$ with ρ and σ_{NSW}^2 given by Theorem 2.1, are shown in Figure 1. For the epidemic with $I \equiv 1$, the asymptotic normal distribution gives an excellent approximation, even though n is rather small. The approximation is markedly worse for the model with $I = 0$ or ∞, owing to the increased likelihood of small outbreaks. If ϵ is held fixed, the probability of a small outbreak decreases approximately exponentially with n, as the number of initial infectives is proportional to n, and the approximation improves significantly, particularly in the $I = 0$ or ∞ model.

Turning to Theorem 2.2, Figure 2 shows simulations of the final size of epidemics on an NSW network with one initial infective and I constant. Note that with a single initial infective there is always a non-negligible probability of a minor outbreak, even if n is large. As $n \to \infty$, the probability

D	n	$I \equiv 1$		$I = 0$ or ∞	
		ρ_n	σ_n^2	ρ_n	σ_n^2
$D \equiv 5$	200	0.5097	2.6355	0.4259	9.9771
	500	0.5284	2.3239	0.4957	9.6846
	1000	0.5334	2.2180	0.5200	8.0686
	2000	0.5360	2.1794	0.5295	7.2082
	5000	0.5374	2.1377	0.5351	6.7611
	10000	0.5379	2.1177	0.5366	6.6402
	∞	0.5384	2.1187	0.5384	6.5200
Po(5)	200	0.5659	1.3149	0.4624	10.2610
	500	0.5761	1.0808	0.5504	6.6926
	1000	0.5789	1.0416	0.5708	4.0793
	2000	0.5803	1.0204	0.5766	3.5567
	5000	0.5811	1.0095	0.5798	3.3708
	10000	0.5814	1.0058	0.5807	3.3139
	∞	0.5817	1.0044	0.5817	3.2505
Geom(1/6)	200	0.5160	0.4029	0.4132	8.5412
	500	0.5184	0.3698	0.5055	3.0416
	1000	0.5189	0.3690	0.5163	1.1780
	2000	0.5194	0.3635	0.5180	1.0824
	5000	0.5196	0.3635	0.5190	1.0607
	10000	0.5196	0.3653	0.5194	1.0438
	∞	0.5197	0.3650	0.5197	1.0381
Power(1, 13.796)	200	0.4957	0.4826	0.3900	8.8942
	500	0.4985	0.4260	0.4828	4.1147
	1000	0.4992	0.4237	0.4960	1.8536
	2000	0.4996	0.4177	0.4981	1.6862
	5000	0.4999	0.4164	0.4992	1.6534
	10000	0.4999	0.4182	0.4996	1.6483
	∞	0.5000	0.4180	0.5000	1.6372

TABLE 1
Simulation results against theoretical (asymptotic) calculations for final size of epidemics with $\epsilon = 0.05$ and $p_I = 0.3$ on NSW networks.

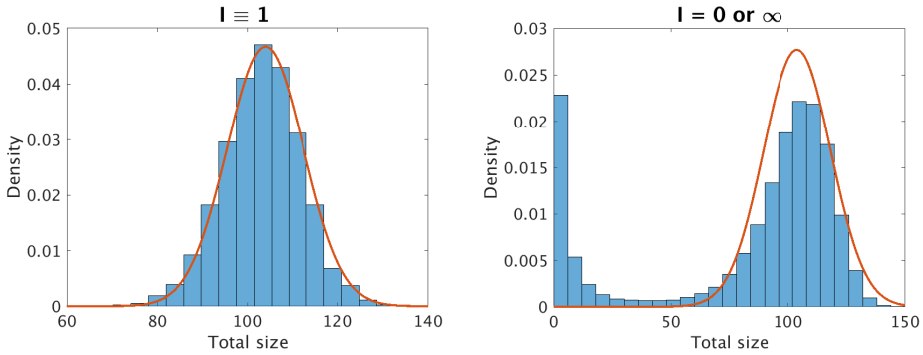

FIG 1. *Histograms of 100,000 simulations of final size for epidemics with* $n = 200, \epsilon = 0.05$ *and* $p_I = 0.3$ *on NSW networks with* $D \sim \text{Geom}(1/6)$*, with asymptotic normal approximation superimposed.*

of a major epidemic converges to the survival probability, p_{maj} say, of a branching process; see, for example, Ball and Sirl (2013), Section 2.1.1, and Theorem 5.1. Moreover, as I is constant, $p_{\text{maj}} = \rho$; see, for example, Kenah and Robins (2007). Superimposed on each histogram in Figure 2 is the density of $\text{N}(n\rho, n\tilde{\sigma}^2_{\text{NSW}})$ multiplied by p_{maj}, which approximates the component of the distribution of $T^{(n)}_{\text{NSW}}$ corresponding to a major outbreak. Even with the very small n, the approximation is very good when $D \sim \text{Geom}(1/6)$ or $D \sim \text{Power}(1, 13.796)$. For both of these degree distributions there is a clear distinction between major and minor outbreaks. That is not the case for the other two degree distributions, though the approximation is still useful. The values of p_{maj} are broadly similar for the four degree distributions. However, the power-law and geometric distributions give greater mass to small values of D than the constant and Poisson distributions; consequently their minor outbreaks are smaller and better separated from major outbreaks.

The upper panels of Figure 3 shows the dependence of ρ (left panel) and $\tilde{\sigma}^2_{\text{MR}}$ and $\tilde{\sigma}^2_{\text{NSW}}$ (right panel) on p_I. The latter two are for the model with I constant. (Recall that, given p_I, the scaled asymptotic mean ρ is independent of the distribution of I.) The asymptotic scaled variances all decrease with p_I and converge to the asymptotic scaled variance of the giant component on the relevant graph as $p_I \uparrow 1$ (cf. Remark 2.6). The asymptotic scaled variances tend to ∞ as $p_I \downarrow p_C$, where $p_C = \mu^{-1}_{\tilde{D}-1}$ is the critical value of p_I so that $R_0 = 1$. Note that $\tilde{\sigma}^2_{\text{NSW}} \geq \tilde{\sigma}^2_{\text{MR}}$ for all choices of D, cf. Remark 2.5. The lower panels of Figure 3 show plots of $(\tilde{\sigma}^2_{\text{NSW}} - \tilde{\sigma}^2_{\text{MR}})/\tilde{\sigma}^2_{\text{MR}}$ against p_I. (The same functions are shown in these two plots but the y-axis is truncated at a smaller value in the right-hand plot to illustrate more clearly the behaviour

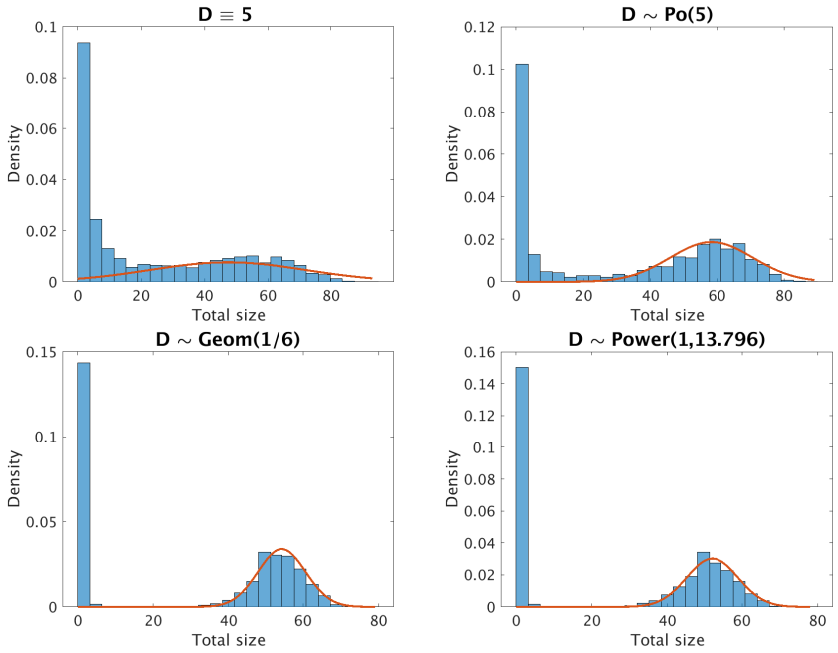

FIG 2. *Histograms of 100,000 simulations of final size for epidemics on NSW networks with* $I \equiv 1, p_I = 0.3, n = 100$ *and one initial infective, with a normal approximation superimposed; see text for details.*

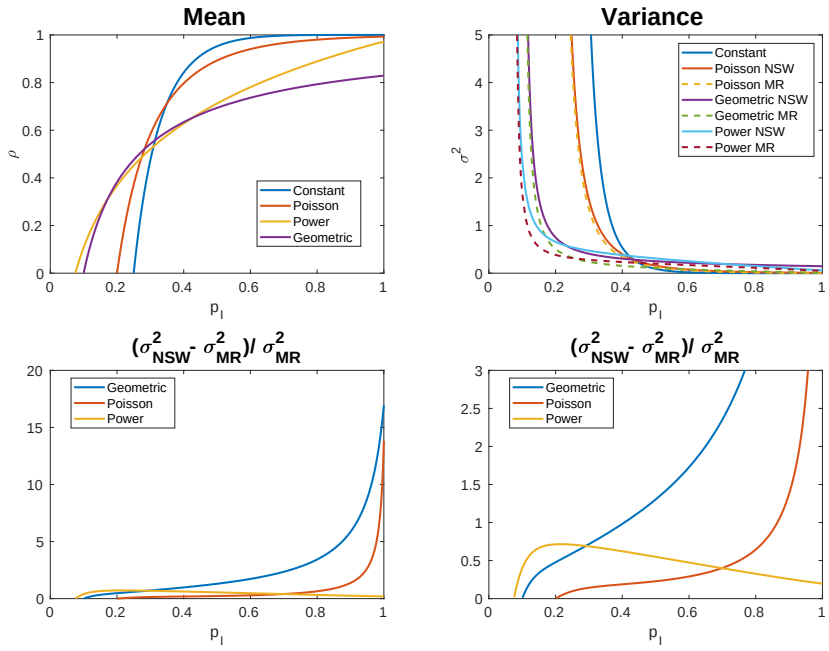

FIG 3. *Dependence of asymptotic means and variances on p_I; see text for details.*

of the functions when p_I is not close to 1.) Note the plots for the Poisson and geometric degree distributions are both increasing with p_I, while that for the Power$(1, 13.796)$ distribution is unimodal.

The final illustrations are concerned with percolation. Figure 4 shows plots of estimates $\hat{\sigma}_n^2$ of the scaled variance $n^{-1}\mathrm{var}(C^{(n)})$ of the size of the largest component, based on $n_{\mathrm{sim}} = 10,000$ simulations for each choice of parameters, together with 95% equal-tailed confidence intervals given by $[(n_{\mathrm{sim}} - 1)/q_2, (n_{\mathrm{sim}} - 1)/q_1]$, where q_1 and q_2 are respectively the 2.5% and 97.5% quantiles of the $\chi^2_{n_{\mathrm{sim}}-1}$ distribution. In all cases $\pi = 0.3$. The filled and unfilled markers correspond to percolation on NSW and MR networks, respectively. The $n \to \infty$ scaled variances, given by Theorem 2.7, are shown by horizontal dashed and solid lines for NSW and MR networks, respectively. The figure suggests that the asymptotic approximations are again generally good, even for moderately-sized networks. For fixed n, the approximation is better for Power$(1, 13.796)$ distribution than for the Po(5) distribution. The plot for site percolation when $D \sim \mathrm{Po}(5)$ is a bit odd, as $\hat{\sigma}_n^2$ is not monotonic in n. This is explored further in Figure 5, which is for percolation on NSW random graphs with $D \sim \mathrm{Po}(5)$. Note that the distribution of $C^{(n)}$ is clearly

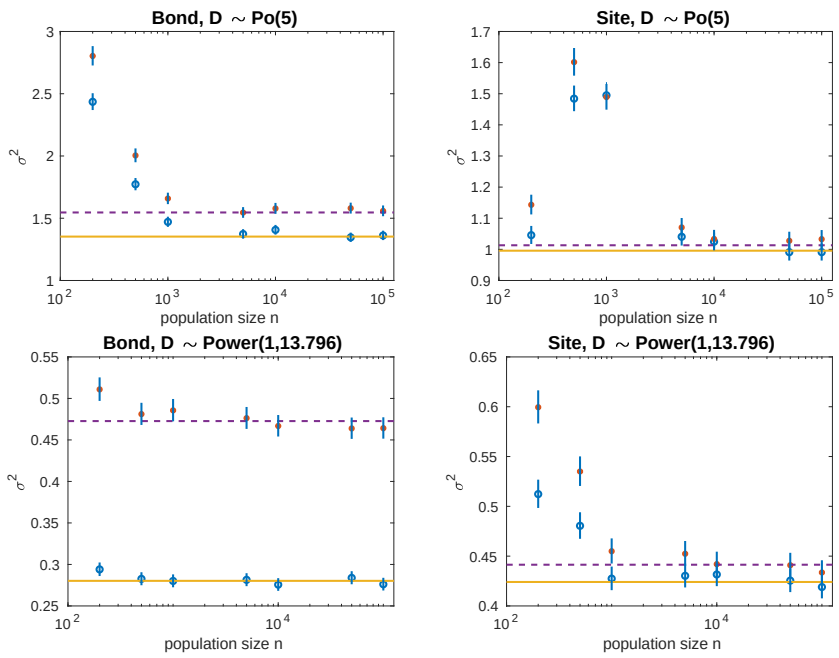

FIG 4. *Plots of scaled variances* $n^{-1}\mathrm{var}(C^{(n)})$ *of the size of the largest component in bond and site percolation on configuration model random graphs; see text for details.*

bimodal for site percolation with $n = 200$ and $\pi = 0.3$. The lower plots in Figure 5 demonstrate that increasing n or π alleviates the issue of small largest components.

Figure 6 shows histograms of the size of the largest component in 100,000 simulated bond and site percolations on an MR random graph with $n = 200, \pi = 0.3$ and $D \sim \mathrm{Power}(1, 13.796)$. Two asymptotic normal approximations are superimposed. The solid lines are the densities of $\mathrm{N}(n\rho, n\sigma^2_{\mathrm{MR,B}})$ (bond percolation) and $\mathrm{N}(n\pi\rho, n\sigma^2_{\mathrm{MR,S}})$ (site percolation), with $\rho, \sigma^2_{\mathrm{MR,B}}$ and $\sigma^2_{\mathrm{MR,S}}$ obtained by setting $D \sim \mathrm{Power}(1, 13.796)$ in Theorem 2.7. An improved approximation (dashed lines) is obtained by instead setting D to be the empirical distribution of $D_1^{(n)}, D_2^{(n)}, \ldots, D_n^{(n)}$. The difference between the approximations is more noticeable for bond percolation. (The support of the histogram has been truncated slightly to make the difference clearer.) The difference decreases with n and is appreciably greater with heavy-tailed degree distributions. Observe from Figures 5 and 6 that the asymptotic normal approximation underestimates the left tail and overestimates the right tail of the distribution of $C^{(n)}$. This phenomenon is present also in the asymptotic

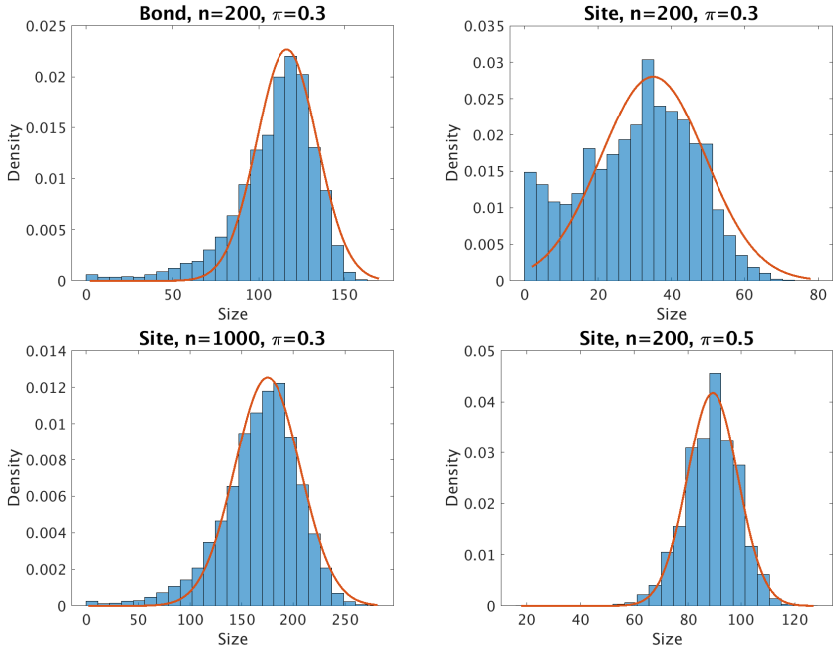

FIG 5. *Histograms 100,000 simulations of size of largest component in percolation on NSW random graphs having* $D \sim \mathrm{Po}(5)$*, with asymptotic normal approximation superimposed.*

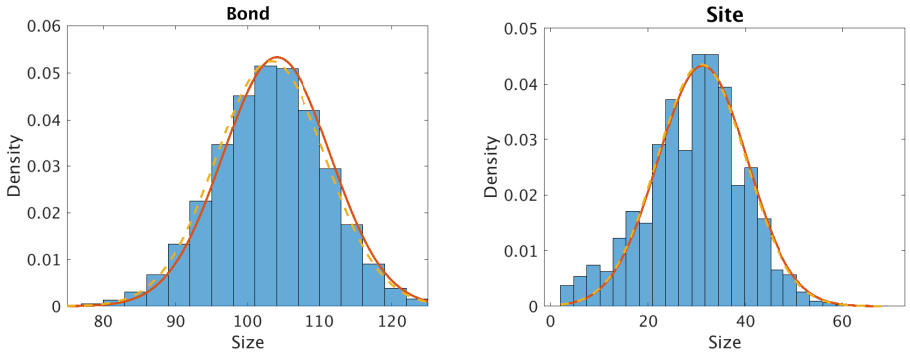

FIG 6. *Histograms 100,000 simulations of size of largest component in percolation on MR random graphs having* $n = 200, \pi = 0.3$ *and* $D \sim \text{Power}(1, 13.796)$, *with two asymptotic normal approximation superimposed; see text for details.*

normal approximation of the epidemic final size $T^{(n)}$.

4. Density dependent population processes. This section collects together some results for density dependent population processes that are required in the paper. It is based on Ethier and Kurtz (1986), Chapter 11, and Pollett (1990), though the statement of the functional central limit theorem is slightly more general than that in those references. The notation is local to this section.

For $n = 1, 2, \ldots$, let $\{\boldsymbol{X}^{(n)}(t) : t \geq 0\} = \{(X_1^{(n)}(t), X_2^{(n)}(t), \ldots, X_p^{(n)}(t)) : t \geq 0\}$ be a continuous-time Markov chain with state space $H^{(n)} \subseteq \mathbb{Z}^p$ and transition intensities of the form

$$q^{(n)}(\boldsymbol{i}, \boldsymbol{i} + \boldsymbol{l}) = n\beta_{\boldsymbol{l}}^{(n)}(n^{-1}\boldsymbol{i}) \qquad (\boldsymbol{i} \in H^{(n)}, \boldsymbol{l} \in \Delta), \tag{4.1}$$

where Δ is the set of possible jumps from a typical state $\boldsymbol{i} = (i_1, i_2, \ldots, i_p)$ and $\beta_{\boldsymbol{l}}^{(n)} : H \to \mathbb{R}$ are functions defined on an open set $H \subseteq \mathbb{R}^p$. Let $H_n = H \cap \{n^{-1}\boldsymbol{i} : \boldsymbol{i} \in \mathbb{Z}^p\}$. (Note that $H_n \subseteq \{n^{-1}\boldsymbol{i} : \boldsymbol{i} \in \mathbb{Z}^p\}$, whereas $H^{(n)} \subseteq \mathbb{Z}^p$.) We require that $\boldsymbol{x} \in H_n$ and $\beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x}) > 0$ imply $\boldsymbol{x} + n^{-1}\boldsymbol{l} \in H_n$. The functions $\beta_{\boldsymbol{l}}^{(n)}$ $(\boldsymbol{l} \in \Delta)$ must satisfy $\beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in H \cap \{n^{-1}\boldsymbol{i} : \boldsymbol{i} \in H^{(n)}\}$. We assume that Δ is finite. The theory in Ethier and Kurtz (1986), Chapter 11, and Pollett (1990) allows Δ to be infinite but only the finite case is required in our application and the results are easier to state in that setting. Suppose that $\beta_{\boldsymbol{l}}(\boldsymbol{x}) = \lim_{n\to\infty} \beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x})$ exists for all $\boldsymbol{l} \in \Delta$ and all $\boldsymbol{x} \in H$; the corresponding family of processes is then called asymptotically density dependent by Pollett (1990). In Ethier and Kurtz (1986), Chapter 11, a

family of processes which satisfies (4.1) with $\beta_{\boldsymbol{l}}^{(n)}$ replaced by $\beta_{\boldsymbol{l}}$ is called a density dependent family, and it is noted that the results usually carry over with little additional effort to the more general form where

$$q^{(n)}(\boldsymbol{i}, \boldsymbol{i}+\boldsymbol{l}) = n\left[\beta_{\boldsymbol{l}}(n^{-1}\boldsymbol{i}) + O(n^{-1})\right] \qquad (\boldsymbol{i} \in H^{(n)}, \boldsymbol{l} \in \Delta).$$

The above framework is slightly different from that in Ethier and Kurtz (1986), which, when extended to asymptotic density dependence, assumes that (i) $H_n = H \cap \{n^{-1}\boldsymbol{i} : \boldsymbol{i} \in H^{(n)}\}$ and (ii) the functions $\beta_{\boldsymbol{l}}^{(n)}$ $(\boldsymbol{l} \in \Delta)$ are nonnegative. The law of large numbers and functional central limit theorem in Ethier and Kurtz (1986) have their origins in Kurtz (1970) and Kurtz (1971), respectively, though the proofs in Ethier and Kurtz (1986) are different. In Pollett (1990), the corresponding theorems do not require (i) or (ii) above but details of the proofs are omitted as they are similar to those in Kurtz (1970) and Kurtz (1971). For completeness, proofs of the law of large numbers and the functional central limit theorem (Theorems 4.1 and 4.2 below) are given in Appendix A. The proofs in Appendix A are based on and follow closely those in Britton and Pardoux (2019) (see also Britton and Pardoux (2020)), which are similar to those in Ethier and Kurtz (1986) but provide appreciably more detail concerning the proof of the functional central limit theorem. The proof of the functional central limit theorem in Appendix A uses Skorohod's theorem, which leads to a simpler proof than that in Britton and Pardoux (2019). Note that the representation of $\{\boldsymbol{X}^{(n)}(t) : t \geq 0\}$ in terms of independent unit-rate Poisson processes (see (A.1) in Appendix A), that is fundamental to the proofs in Ethier and Kurtz (1986), still holds under our conditions, as $\beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in H \cap \{n^{-1}\boldsymbol{i} : \boldsymbol{i} \in H^{(n)}\}$, so the proofs in Ethier and Kurtz (1986) continue to hold after obvious modifications to account for asymptotic density dependence and asymptotically random initial conditions.

Let

$$F(\boldsymbol{x}) = \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l}\beta_{\boldsymbol{l}}(\boldsymbol{x}) \qquad (\boldsymbol{x} \in H). \tag{4.2}$$

THEOREM 4.1. *Suppose that for each compact $K \subset H$,*

$$\sup_{\boldsymbol{x} \in K} \beta_{\boldsymbol{l}}(\boldsymbol{x}) < \infty \qquad \textit{and} \qquad \lim_{n\to\infty} \sup_{\boldsymbol{x} \in K} |\beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x}) - \beta_{\boldsymbol{l}}(\boldsymbol{x})| = 0 \quad (\boldsymbol{l} \in \Delta), \tag{4.3}$$

and there exists $M_K > 0$ such that

$$|F(\boldsymbol{x}) - F(\boldsymbol{y})| < M_K|\boldsymbol{x} - \boldsymbol{y}| \qquad \textit{for all } \boldsymbol{x}, \boldsymbol{y} \in K. \tag{4.4}$$

Suppose also that $n^{-1}\boldsymbol{X}^{(n)}(0) \xrightarrow{p} \boldsymbol{x}_0 \in H$ *as* $n \to \infty$. *Let* $\boldsymbol{x}(t)$ $(t \geq 0)$ *be the unique solution of the integral equation*

$$\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t F(\boldsymbol{x}(u))\,\mathrm{d}u \qquad (t \geq 0) \tag{4.5}$$

and suppose that $\boldsymbol{x}(t)$ *is finite for all* $t \geq 0$. *Then, for all* $t \geq 0$,

$$\sup_{0 \leq u \leq t} |n^{-1}\boldsymbol{X}^{(n)}(u) - \boldsymbol{x}(u)| \xrightarrow{p} 0 \qquad \text{as } n \to \infty. \tag{4.6}$$

Write $F(\boldsymbol{x}) = (F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \ldots, F_p(\boldsymbol{x}))$ and $\boldsymbol{x} = (x_1, x_2, \ldots, x_p)$. Let $\partial F(\boldsymbol{x})$ and $G(\boldsymbol{x})$ be the $p \times p$ matrix functions defined by

$$\partial F(\boldsymbol{x}) = \left[\frac{\partial F_i}{\partial x_j}(\boldsymbol{x})\right] \qquad \text{and} \qquad G(\boldsymbol{x}) = \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l}^\top \boldsymbol{l} \beta_{\boldsymbol{l}}(\boldsymbol{x}).$$

Further, let $\Phi(t,u) = [\phi_{ij}(t,u)]$ $(0 \leq u \leq t < \infty)$ be the solution of the matrix differential equation

$$\frac{\partial}{\partial t}\Phi(t,u) = \partial F(\boldsymbol{x}(t))\Phi(t,u), \quad \Phi(u,u) = I.$$

Note that

$$\phi_{ij}(t,u) = \frac{\partial x_i(t-u)}{\partial x_j(0)} \qquad (i,j = 1,2,\ldots,p).$$

Let $\Rightarrow$ denote weak convergence in the space of right-continuous functions $f : [0,\infty) \to \mathbb{R}^p$ having limits from the left (i.e. càdlàg functions), endowed with the Skorohod metric.

THEOREM 4.2. *Suppose, in addition to the conditions of Theorem 4.1, that* $\beta_{\boldsymbol{l}}$ $(\boldsymbol{l} \in \Delta)$ *and* ∂F *are continuous, and that for each compact* $K \subset H$,

$$\lim_{n \to \infty} \sqrt{n} \sup_{\boldsymbol{x} \in K} |\beta_{\boldsymbol{l}}^{(n)}(\boldsymbol{x}) - \beta_{\boldsymbol{l}}(\boldsymbol{x})| = 0 \qquad (\boldsymbol{l} \in \Delta). \tag{4.7}$$

Suppose also that $\boldsymbol{x}_0 \in H$ *and*

$$\sqrt{n}\left(n^{-1}\boldsymbol{X}^{(n)}(0) - \boldsymbol{x}_0\right) \xrightarrow{D} \boldsymbol{V}(0) \qquad \text{as } n \to \infty, \tag{4.8}$$

where $\boldsymbol{V}(0) \sim N(\boldsymbol{0}, \Sigma_0)$. *Then, as* $n \to \infty$,

$$\left\{\sqrt{n}\left(n^{-1}\boldsymbol{X}^{(n)}(t) - \boldsymbol{x}(t)\right) : t \geq 0\right\} \Rightarrow \{\boldsymbol{V}(t) : t \geq 0\}, \tag{4.9}$$

where $\boldsymbol{x}(t)$ *is given by* (4.5) *and* $\{\boldsymbol{V}(t) : t \geq 0\}$ *is a zero-mean Gaussian process with covariance function given, for* $t_1, t_2 \geq 0$, *by*

$$\begin{aligned}\text{(4.10)}\quad \operatorname{cov}(\boldsymbol{V}(t_1), \boldsymbol{V}(t_2)) =& \Phi(t_1, 0)\Sigma_0\Phi(t_2, 0)^\top \\ &+ \int_0^{\min(t_1,t_2)} \Phi(t_1, u)G(\boldsymbol{x}(u))\Phi(t_2, u)^\top \, \mathrm{d}u.\end{aligned}$$

We also require the following theorem concerned with the first exit of $\{\boldsymbol{X}^{(n)}(t) : t \geq 0\}$ from a region of its state space. The theorem is Ethier and Kurtz (1986), Theorem 11.4.1, expressed now for asymptotically density dependent processes having random initial conditions.

THEOREM 4.3. *Suppose that the conditions of Theorem 4.2 are satisfied. Let* $\varphi : H \to \mathbb{R}$ *be continuously differentiable. Suppose that* $\varphi(\boldsymbol{x}_0) > 0$. *Let* $\tau^{(n)} = \inf\{t \geq 0 : \varphi(n^{-1}\boldsymbol{X}^{(n)}(t)) \leq 0\}$ *and* $\tau = \inf\{t \geq 0 : \varphi(\boldsymbol{x}(t)) \leq 0\}$. *Suppose that* $\tau < \infty$ *and* $\nabla\varphi(\boldsymbol{x}(\tau)) \cdot F(\boldsymbol{x}(\tau)) < 0$, *where* $\cdot$ *denotes inner vector product. Then, as* $n \to \infty$,

$$\text{(4.11)}\qquad \sqrt{n}\left(\tau^{(n)} - \tau\right) \xrightarrow{D} -\frac{\nabla\varphi(\boldsymbol{x}(\tau) \cdot \boldsymbol{V}(\tau)}{\nabla\varphi(\boldsymbol{x}(\tau) \cdot F(\boldsymbol{x}(\tau))}$$

and

$$\sqrt{n}\left(n^{-1}\boldsymbol{X}^{(n)}(\tau^{(n)}) - \boldsymbol{x}(\tau)\right) \xrightarrow{D} \boldsymbol{V}(\tau) - \frac{\nabla\varphi(\boldsymbol{x}(\tau) \cdot \boldsymbol{V}(\tau)}{\nabla\varphi(\boldsymbol{x}(\tau) \cdot F(\boldsymbol{x}(\tau))}F(\boldsymbol{x}(\tau)).$$

Note that (4.11) implies $\tau^{(n)} \xrightarrow{p} \tau$ as $n \to \infty$.

COROLLARY 4.1. *Suppose that the conditions of Theorem 4.3 are satisfied. Let*

$$B = I - \frac{(F(\boldsymbol{x}(\tau)))^\top \nabla\varphi(\boldsymbol{x}(\tau))}{\nabla\varphi(\boldsymbol{x}(\tau)) \cdot F(\boldsymbol{x}(\tau))},$$

Then,

$$\sqrt{n}\left(n^{-1}\boldsymbol{X}^{(n)}(\tau^{(n)}) - \boldsymbol{x}(\tau)\right) \xrightarrow{D} \mathrm{N}(\boldsymbol{0}, B\Sigma(\tau)B^\top) \qquad \text{as } n \to \infty,$$

where

$$\Sigma(\tau) = \Phi(\tau, 0)\Sigma_0\Phi(\tau, 0)^\top + \int_0^\tau \Phi(\tau, u)G(\boldsymbol{x}(u))\Phi(\tau, u)^\top \, \mathrm{d}u.$$

PROOF. Corollary 4.1 follows immediately from Theorems 4.2 and 4.3 on noting that

$$\boldsymbol{V}(\tau) - \frac{\nabla\varphi(\boldsymbol{x}(\tau) \cdot \boldsymbol{V}(\tau)}{\nabla\varphi(\boldsymbol{x}(\tau) \cdot F(\boldsymbol{x}(\tau))}F(\boldsymbol{x}(\tau)) = \boldsymbol{V}(\tau)B^\top.$$

□

5. Proofs.

5.1. *Alternative construction of final size $T^{(n)}$.* We describe first the well-known construction of the final outcome of the epidemic $\mathcal{E}^{(n)}$ using a random directed graph. We then use that construction to show that the distribution of $T^{(n)}$, the final size of $\mathcal{E}^{(n)}$, can be obtained using the location of the first exit of an asymptotically density dependent population process from a given region.

Given a realisation of $\mathcal{G}^{(n)}$, construct a directed random graph $\tilde{\mathcal{G}}^{(n)}$, having vertex set $\mathcal{N}^{(n)} = \{1, 2, \ldots, n\}$, as follows. For each $i = 1, 2, \ldots, n$, by sampling from its infectious period distribution I and then the relevant Poisson processes, draw up a list of individuals i would make contact with if i were to become infected. Then, for each ordered pair of individuals, (i, j) say, a directed edge from i to j is present in $\tilde{\mathcal{G}}^{(n)}$ if and only if j is in i's list. Let $\mathcal{I}^{(n)}$ denote the set of initial infectives in $\mathcal{E}^{(n)}$. For distinct $i, j \in \mathcal{N}^{(n)}$, write $i \rightsquigarrow j$ if and only if there is a chain of directed edges from i to j in $\tilde{\mathcal{G}}^{(n)}$. Let $\mathcal{T}^{(n)} = \{j \in \mathcal{N}^{(n)} \setminus \mathcal{I}^{(n)} : i \rightsquigarrow j \text{ for some i } \in \mathcal{I}^{(n)}\}$ be the set of initial susceptibles that are infected in $\mathcal{E}^{(n)}$, so $T^{(n)} = |\mathcal{T}^{(n)}|$.

Note that $\mathcal{T}^{(n)}$ is determined purely by the presence/absence of directed edges in $\tilde{\mathcal{G}}^{(n)}$ and does not depend on the times of the corresponding infections in $\mathcal{E}^{(n)}$. (This implies that the distribution of $\mathcal{T}^{(n)}$, and hence also $T^{(n)}$ is invariant to the introduction of a latent/exposed period into the model, i.e. the time elapsing after infection of an individual before it is able to infect other individuals.) It follows that the final outcome $\mathcal{T}^{(n)}$ has the same distribution as that of a related epidemic $\tilde{\mathcal{E}}^{(n)}$, with set of initial infectives $\mathcal{I}^{(n)}$, in which for any infective, i say, it is determined upon infection which, if any, of its neighbours i will contact and those contacts take place at the first points of independent Poisson processes, each having rate 1. More precisely, suppose i is infected at time t_0 and i has d neighbours, $i_1, i_2, \ldots, i_d$ say. Let I_i be a realisation of I and, given I_i, let $\chi_{i1}, \chi_{i2}, \ldots, \chi_{id}$ be i.i.d. Bernoulli random variables with success probability $1 - \exp(-\lambda I_i)$. Let $W_{i1}, W_{i2}, \ldots, W_{id}$ be an independent set of i.i.d. unit-mean exponential random random variables. Then i contacts i_j if and only if $\chi_{ij} = 1$, and in that case the contact occurs at time $t_0 + W_{ij}$. Of course, the I and W random variables for any set of distinct infectives are mutually independent.

Given the degrees $D_1^{(n)}, D_2^{(n)}, \ldots, D_n^{(n)}$ and the set of initial infectives $\mathcal{I}^{(n)}$, the random graph $\mathcal{G}^{(n)}$ and the epidemic on it $\tilde{\mathcal{E}}^{(n)}$ can be constructed simultaneously as follows (cf. Ball and Neal (2008)). The process starts with no half-edge paired and all half-edges susceptible. The individuals in $\mathcal{I}^{(n)}$ become infected at time $t = 0$ and all other individuals are susceptible.

When an individual is infected, it determines immediately which, if any, of its unpaired half-edges it will transmit infection along and when it will make those contacts, according to the probabilistic law described above; the half-edges that the individual will infect along and not infect along then become infective and recovered half-edges, respectively. When infection is transmitted along a half-edge that half-edge is paired with a half-edge chosen uniformly at random from all unpaired half-edges, forming an edge in the network, and the two half edges become recovered. If the chosen half-edge is attached to a susceptible individual then that individual is infected and determines immediately along which, if any, of its unpaired half-edges it will transmit infection. If the chosen half-edge is infective or recovered then nothing further happens. The process continues until there is no infective half-edge remaining. (In the epidemic on the NSW random graph, if $D_1^{(n)} + D_2^{(n)} + \cdots + D_n^{(n)}$ is odd then it is possible for the process to end with one unpaired half-edge, which is infective, but we can ignore that possibility because under the conditions of the theorems its probability tends to 0 as $n \to \infty$.)

Note that as we are interested only in the final outcome of the epidemic, it is not necessary to keep track of the degrees of individuals to which infective and unpaired recovered half-edges are attached; we need to know just the total numbers of such half-edges. For $t \geq 0$, let $X_i^{(n)}(t)$ be the number of degree-i susceptible individuals at time t ($i = 0, 1, \ldots, d_{\max}$), and let $Y_E^{(n)}(t)$ and $Z_E^{(n)}(t)$ be the number of infective half-edges (which by definition are unpaired) and the number of unpaired recovered half-edges, respectively, at time t. Let $\{\boldsymbol{W}^{(n)}(t)\} = \{\boldsymbol{W}^{(n)}(t) : t \geq 0\}$, where $\boldsymbol{W}^{(n)}(t) = (\boldsymbol{X}^{(n)}(t), Y_E^{(n)}(t), Z_E^{(n)}(t))$ and $\boldsymbol{X}^{(n)}(t) = (X_0^{(n)}(t), X_1^{(n)}(t), \ldots, X_{d_{\max}}^{(n)}(t))$.

The process $\{\boldsymbol{W}^{(n)}(t)\}$ is a continuous-time Markov chain, whose initial state $\boldsymbol{W}^{(n)}(0)$ is random, even for the epidemic on the MR random graph as the numbers of infective and recovered half-edges created by the initial infectives are random. In the epidemic on the NSW random graph, $X_i^{(n)}(t)$ ($i = 0, 1, \ldots, d_{\max}$) are also random. Before giving the transition intensities of $\{\boldsymbol{W}^{(n)}(t)\}$ some more notation is required.

For simplicity, we take $H^{(n)} = \mathbb{Z}_+^{d_{\max}+3}$ as a state space for $\{\boldsymbol{W}^{(n)}(t)\}$, but note that some states in $H^{(n)}$ are unattainable since clearly $X_i^{(n)}(t) \leq n$ ($i = 0, 1, \ldots, d_{\max}$), $Y_E^{(n)}(t) \leq n d_{\max}$ and $Z_E^{(n)}(t) \leq n d_{\max}$ for all $t \geq 0$. Let $\boldsymbol{n} = (n_0^X, n_1^X, \ldots, n_{d_{\max}}^X, n_E^Y, n_E^Z)$ denote a typical element of $H^{(n)}$. The states with $n_E^Y = 0$ are absorbing, as is the state $(0, 0, \ldots, 0, 1, 0)$. Let $n_E^X = \sum_{i=1}^{d_{\max}} i n_i^X$. For $i = 0, 1, \ldots, d_{\max}$, let $\boldsymbol{e}_i^{\mathrm{S}}$ be the unit vector of length $d_{\max} + 3$ with one in the component corresponding to a degree-i susceptible

so, for example, $\boldsymbol{e}_1^{\mathrm{S}} = (0,1,0,0,\ldots,0)$. Similarly, let $\boldsymbol{e}^{\mathrm{I}} = (0,0,\ldots,0,1,0)$ and $\boldsymbol{e}^{\mathrm{R}} = (0,0,\ldots,0,1)$. For $i = 1,2,\ldots,d_{\max}$ and $k = 0,1,\ldots,i-1$, let $p_{i,k}$ be the probability that if a degree-i susceptible is infected it subsequently transmits infection along k of its remaining $i-1$ half-edges. (Note that a degree-0 susceptible cannot be infected.) Thus $p_{i,k} = \mathrm{P}(X = k)$, where $X \sim \mathrm{Bin}(i-1, 1-\exp(-\lambda I))$. The transition intensities of $\{\boldsymbol{W}^{(n)}(t)\}$ are as follows.

(i) For $i = 1,2,\ldots,d_{\max}$ and $k = 0,1,\ldots,i-1$, an infective half-edge is paired with a degree-i susceptible yielding k infective half-edges and $i-1-k$ recovered half-edges

$$q^{(n)}(\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{e}_i^{\mathrm{S}} + (k-1)\boldsymbol{e}^{\mathrm{I}} + (i-k-1)\boldsymbol{e}^{\mathrm{R}}) = \frac{n_E^Y i n_i^X p_{i,k}}{n_E^X + n_E^Y + n_E^Z - 1};$$

(ii) an infective half-edge is paired with an infective half-edge

$$q^{(n)}(\boldsymbol{n}, \boldsymbol{n} - 2\boldsymbol{e}^{\mathrm{I}}) = \frac{n_E^Y (n_E^Y - 1)}{n_E^X + n_E^Y + n_E^Z - 1};$$

(iii) an infective half-edge is paired with a recovered half-edge

$$q^{(n)}(\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{e}^{\mathrm{I}} - \boldsymbol{e}^{\mathrm{R}}) = \frac{n_E^Y n_E^Z}{n_E^X + n_E^Y + n_E^Z - 1}.$$

Note that these transition intensities are independent of the population size n. We index them by n to connect with theory of density dependent population processes.

Let $\tau^{(n)} = \inf\{t \geq 0 : Y_E^{(n)}(t) = 0\}$. Then $X_i^{(n)}(\tau^{(n)})$ $(i = 0,1,\ldots,d_{\max})$ give the numbers of susceptibles of the different degrees at the end of the epidemic and $T^{(n)} = \sum_{i=1}^{d_{\max}} \left(X_i^{(n)}(0) - X_i^{(n)}(\tau^{(n)}) \right)$.

5.2. *Random time-scale transformation.* We wish to apply Corollary 4.1 to obtain a central limit theorem for $T^{(n)}$ but that corollary cannot be applied directly to $\{\boldsymbol{W}^{(n)}(t)\}$ as $\tau^{(n)} \xrightarrow{p} \infty$ as $n \to \infty$. Thus we consider the following random time-scale transformation of $\{\boldsymbol{W}^{(n)}(t)\}$; cf. Ethier and Kurtz (1986), page 467, Watson (1980) and Janson et al. (2014).

For $t \in [0, \tau^{(n)}]$, let

$$A^{(n)}(t) = \int_0^t \frac{Y_E^{(n)}(u)}{X_E^{(n)}(u) + Y_E^{(n)}(u) + Z_E^{(n)}(u) - 1} \, \mathrm{d}u,$$

where $X_E^{(n)}(u) = \sum_{i=1}^{d_{\max}} iX_i^{(n)}(u)$. Let $\tilde{\tau}^{(n)} = A^{(n)}(\tau^{(n)})$. For $t \in [0, \tilde{\tau}^{(n)}]$, let $U^{(n)}(t) = \inf\{u \geq 0 : A^{(n)}(u) = t\}$ and $\tilde{\boldsymbol{W}}^{(n)}(t) = \boldsymbol{W}^{(n)}\left(U^{(n)}(t)\right)$. Then the process $\{\tilde{\boldsymbol{W}}^{(n)}(t) : 0 \leq t \leq \tilde{\tau}^{(n)}\}$ is Markovian and has transition intensities during $[0, \tilde{\tau}^{(n)})$ given by:

(i) for $i = 1, 2, \ldots, d_{\max}, n_i^X \geq 0$ and $k = 0, 1, \ldots, i-1$,

$$\tilde{q}^{(n)}(\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{e}_i^{\mathrm{S}} + (k-1)\boldsymbol{e}^{\mathrm{I}} + (i-k-1)\boldsymbol{e}^{\mathrm{R}}) = in_i^X p_{i,k};$$

(ii) for $n_E^Y \geq 1$,

$$\tilde{q}^{(n)}(\boldsymbol{n}, \boldsymbol{n} - 2\boldsymbol{e}^{\mathrm{I}}) = n_E^Y - 1;$$

(iii) for $n_E^Z \geq 0$,

$$\tilde{q}^{(n)}(\boldsymbol{n}, \boldsymbol{n} - \boldsymbol{e}^{\mathrm{I}} - \boldsymbol{e}^{\mathrm{R}}) = n_E^Z.$$

In order to define $\tilde{\boldsymbol{W}}^{(n)}(t)$ for $t > \tilde{\tau}^{(n)}$ (so that Corollary 4.1 can be applied) we add the transition intensities:

(ii′) for $n_E^Y \leq 0$,

$$\tilde{q}^{(n)}(\boldsymbol{n}, \boldsymbol{n} + 2\boldsymbol{e}^{\mathrm{I}}) = -n_E^Y.$$

The process $\{\tilde{\boldsymbol{W}}^{(n)}(t)\} = \{\tilde{\boldsymbol{W}}^{(n)}(t) : t \geq 0\}$ is then a continuous-time Markov chain with state space $\tilde{H}^{(n)} = \mathbb{Z}_+^{d_{\max}+1} \times \mathbb{Z} \times \mathbb{Z}_+$, though some states in $\tilde{H}^{(n)}$ are unattainable, and transition intensities given by (i), (ii), (iii) and (ii′) above.

Let $\boldsymbol{l}_{ik}^{(1)} = -\boldsymbol{e}_i^{\mathrm{S}} + (k-1)\boldsymbol{e}^{\mathrm{I}} + (i-k-1)\boldsymbol{e}^{\mathrm{R}}$ $(i = 1, 2, \ldots, d_{\max}; k = 0, 1, \ldots, i-1)$, $\boldsymbol{l}_+^{(2)} = -2\boldsymbol{e}^{\mathrm{I}}$, $\boldsymbol{l}_-^{(2)} = 2\boldsymbol{e}^{\mathrm{I}}$ and $\boldsymbol{l}^{(3)} = -\boldsymbol{e}^{\mathrm{I}} - \boldsymbol{e}^{\mathrm{R}}$. The set of possible jumps of $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$ from a typical state $\boldsymbol{n}$ is $\Delta = \Delta_1 \cup \Delta_2 \cup \Delta_3$, where $\Delta_1 = \{\boldsymbol{l}_{ik}^{(1)} : i = 1, 2, \ldots, d_{\max}; k = 0, 1, \ldots, i-1\}$, $\Delta_2 = \{\boldsymbol{l}_+^{(2)}, \boldsymbol{l}_-^{(2)}\}$ and $\Delta_3 = \{\boldsymbol{l}^{(3)}\}$. Let $\boldsymbol{w} = (\boldsymbol{x}, y_E, z_E)$, where $\boldsymbol{x} = (x_0, x_1, \ldots, x_{d_{\max}})$, and $H = \mathbb{R}^{d_{\max}+3}$. The intensities of the jumps of $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$ admit the form

$$\tilde{q}^{(n)}(\boldsymbol{n}, \boldsymbol{n} + \boldsymbol{l}) = n\tilde{\beta}_{\boldsymbol{l}}^{(n)}(n^{-1}\boldsymbol{n}) \qquad (\boldsymbol{n} \in \tilde{H}^{(n)}, \boldsymbol{l} \in \Delta), \tag{5.1}$$

where the functions $\tilde{\beta}_{\boldsymbol{l}}^{(n)} : H \to \mathbb{R}_+$ $(\boldsymbol{l} \in \Delta)$ are given by

$$\tilde{\beta}_{\boldsymbol{l}}^{(n)}(\boldsymbol{w}) = \begin{cases} \tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w}) & \text{if } \boldsymbol{l} \in \Delta \setminus \{\boldsymbol{l}_+^{(2)}, \boldsymbol{l}_-^{(2)}\}, \\ \tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w}) - n^{-1} & \text{if } \boldsymbol{l} = \boldsymbol{l}_+^{(2)}, \\ \tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w}) & \text{if } \boldsymbol{l} = \boldsymbol{l}_-^{(2)}, \end{cases} \tag{5.2}$$

with

$$\tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w}) = \begin{cases} \tilde{\beta}^{(1)}_{ik}(\boldsymbol{x}, y_E, z_E) = i x_i p_{i,k} & \text{for } \boldsymbol{l} = \boldsymbol{l}^{(1)}_{ik} \in \Delta_1, \\ \tilde{\beta}^{(2)}_{+}(\boldsymbol{x}, y_E, z_E) = y_E 1_{\{y_E > 0\}} & \text{for } \boldsymbol{l} = \boldsymbol{l}^{(2)}_{+}, \\ \tilde{\beta}^{(2)}_{-}(\boldsymbol{x}, y_E, z_E) = -y_E 1_{\{y_E \le 0\}} & \text{for } \boldsymbol{l} = \boldsymbol{l}^{(2)}_{-}, \\ \tilde{\beta}^{(3)}(\boldsymbol{x}, y_E, z_E) = z_E & \text{for } \boldsymbol{l} = \boldsymbol{l}^{(3)}. \end{cases} \tag{5.3}$$

The family of processes $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$ is asymptotically density dependent (see Section 4). The further intensities (ii′) have been chosen so that the resulting drift function satsifies the Lipschitz and differentiability properties required for the application of Theorems 4.1 and 4.2.

5.3. *Proof of Theorem 2.1.* For $t \ge 0$, write $\tilde{\boldsymbol{W}}^{(n)}(t) = (\tilde{\boldsymbol{X}}^{(n)}(t), \tilde{Y}^{(n)}_E(t), \tilde{Z}^{(n)}_E(t))$, where $\tilde{\boldsymbol{X}}^{(n)}(t) = (\tilde{X}^{(n)}_0(t), \tilde{X}^{(n)}_1(t), \ldots, \tilde{X}^{(n)}_{d_{\max}}(t))$. Note that, by construction, $\boldsymbol{W}^{(n)}(\tau^{(n)}) = \tilde{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)})$, so the final size of the epidemic is given by

$$T^{(n)} = \sum_{i=0}^{d_{\max}} \left(\tilde{X}^{(n)}_i(0) - \tilde{X}^{(n)}_i(\tilde{\tau}^{(n)}) \right). \tag{5.4}$$

Note also that $\tilde{\tau}^{(n)} = \inf\{t \ge 0 : \tilde{Y}^{(n)}_E(t) = 0\}$. We use Corollary 4.1 to obtain a central limit theorem for $\tilde{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)})$, and hence for the final size $T^{(n)}$. The asymptotic variance matrix in the central limit theorem for $\tilde{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)})$ is not in closed form. However, we derive a closed-form expression for the asymptotic variance of $T^{(n)}$. The main concepts of the proof are presented here, with some detailed but elementary calculations deferred to Appendix B.

5.3.1. *Deterministic model.* As at (4.2), define the drift function $\tilde{F}(\boldsymbol{w}) = \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l} \tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w})$. Using (5.3),

$$\begin{aligned} \tilde{F}(\boldsymbol{w}) = -\sum_{i=1}^{d_{\max}} \sum_{k=0}^{i-1} i x_i p_{i,k} \left[-\boldsymbol{e}^{\mathrm{S}}_i + (k-1)\boldsymbol{e}^{\mathrm{I}} + (i-k-1)\boldsymbol{e}^{\mathrm{R}} \right] \\ - 2 y_E \boldsymbol{e}^{\mathrm{I}} - z_E(\boldsymbol{e}^{\mathrm{I}} + \boldsymbol{e}^{\mathrm{R}}). \end{aligned} \tag{5.5}$$

Recall that $p_{i,k} = \mathrm{P}(X = k)$, where $X \sim \mathrm{Bin}(i-1, 1-\exp(-\lambda I))$. Thus $\sum_{k=0}^{i-1} p_{i,k} = 1$ and $\sum_{k=0}^{i-1} k p_{i,k} = \mathrm{E}[X] = (i-1)p_I$. Substituting these

into (5.5) and recalling that $q_I = 1 - p_I$ yields

$$\tilde{F}(\boldsymbol{w}) = -\sum_{i=1}^{d_{\max}} i x_i \boldsymbol{e}_i^{\mathrm{S}} + \left\{ \sum_{i=1}^{d_{\max}} i x_i [(i-1)p_I - 1] - 2y_E - z_E \right\} \boldsymbol{e}^{\mathrm{I}} + \left[\sum_{i=1}^{d_{\max}} i(i-1) x_i q_I - z_E \right] \boldsymbol{e}^{\mathrm{R}}. \tag{5.6}$$

For $t \geq 0$, let $\tilde{\boldsymbol{w}}(t) = (\tilde{x}_0(t), \tilde{x}_1(t), \ldots, \tilde{x}_{d_{\max}}(t), \tilde{y}_E(t), \tilde{z}_E(t))$ be defined by

$$\tilde{\boldsymbol{w}}(t) = \tilde{\boldsymbol{w}}(0) + \int_0^t \tilde{F}(\boldsymbol{w}(u))\, \mathrm{d}u. \tag{5.7}$$

Thus $\tilde{\boldsymbol{w}}(t)$ satisfies the differential equations

$$\frac{d\tilde{x}_i}{dt} = -i\tilde{x}_i \quad (i = 0, 1, \ldots, d_{\max}), \tag{5.8}$$

$$\frac{d\tilde{y}_E}{dt} = \sum_{i=2}^{d_{\max}} i(i-1)p_I \tilde{x}_i - \tilde{x}_E - 2\tilde{y}_E - \tilde{z}_E, \tag{5.9}$$

$$\frac{d\tilde{z}_E}{dt} = \sum_{i=2}^{d_{\max}} i(i-1)q_I \tilde{x}_i - \tilde{z}_E, \tag{5.10}$$

where $\tilde{x}_E = \sum_{i=1}^{d_{\max}} i\tilde{x}_i$, having solution (see Appendix B.1)

$$\tilde{x}_i(t) = \tilde{x}_i(0)\mathrm{e}^{-it} \quad (i = 0, 1, \ldots, d_{\max}), \tag{5.11}$$

$$\tilde{y}_E(t) = (\tilde{x}_E(0) + \tilde{y}_E(0) + \tilde{z}_E(0))\mathrm{e}^{-2t} - [\tilde{z}_E(0) + q_I\tilde{x}_E(0)]\mathrm{e}^{-t} - p_I \sum_{i=1}^{d_{\max}} i\tilde{x}_i(0)\mathrm{e}^{-it}, \tag{5.12}$$

$$\tilde{z}_E(t) = [\tilde{z}_E(0) + q_I\tilde{x}_E(0)]\mathrm{e}^{-t} - q_I \sum_{i=1}^{d_{\max}} i\tilde{x}_i(0)\mathrm{e}^{-it}. \tag{5.13}$$

Let $\tilde{\eta}_E(t) = \tilde{x}_E(t) + \tilde{y}_E(t) + \tilde{z}_E(t)$. Note that, for $t \geq 0$,

$$\tilde{\eta}_E(t) = \tilde{\eta}_E(0)\mathrm{e}^{-2t}. \tag{5.14}$$

5.3.2. *Initial conditions.* Consider first the epidemic on the MR random graph. Recall that in the epidemic $\mathcal{E}^{(n)}$, for $i = 0, 1, \ldots, d_{\max}$, there are $v_i^{(n)}$ individuals of degree i, of whom $a_i^{(n)}$ are initially infected. Also,

$\left(\tilde{Y}_E^{(n)}(0), \tilde{Z}_E^{(n)}(0)\right)$ is given by the total numbers of infective and recovered half-edges created by the initial infectives. Thus $\tilde{X}_i^{(n)}(0) = v_i^{(n)} - a_i^{(n)}$ $(i = 0, 1, \ldots, d_{\max})$ and

$$\left(\tilde{Y}_E^{(n)}(0), \tilde{Z}_E^{(n)}(0)\right) = \sum_{i=1}^{d_{\max}} \sum_{j=1}^{a_i^{(n)}} (Y_{ij}, i - Y_{ij}),$$

where Y_{ij} $(i = 1, 2, \ldots, d_{\max}; j = 1, 2, \ldots)$ are independent and $Y_{ij} \sim \mathrm{Bin}(i, 1-\exp(-\lambda I))$. Let $\epsilon_E = \sum_{i=1}^{d_{\max}} i\epsilon_i$. Then, recalling that $\epsilon_i^{(n)} = n^{-1}a_i^{(n)}$,

$$\begin{aligned}
&\sqrt{n}\left[n^{-1}\left(\tilde{Y}_E^{(n)}(0), \tilde{Z}_E^{(n)}(0)\right) - \epsilon_E(p_I, q_I)\right] \\
&\quad = \frac{1}{\sqrt{n}} \sum_{i=1}^{d_{\max}} \left[\sum_{j=1}^{a_i^{(n)}} (Y_{ij}, i - Y_{ij}) - a_i^{(n)} i(p_I, q_I)\right] + \sqrt{n} \sum_{i=1}^{d_{\max}} i(\epsilon_i^{(n)} - \epsilon_i).
\end{aligned}$$

For $i = 1, 2, \ldots, d_{\max}$, let $\sigma_{Y,i}^2 = \mathrm{var}(Y_{i1})$. Now $\lim_{n\to\infty} \sqrt{n}(\epsilon_i^{(n)} - \epsilon_i) = 0$ $(i = 0, 1, \ldots, d_{\max})$, by assumption, so the central limit theorem and Slutsky's theorem imply that

$$\sqrt{n}\left[n^{-1}\left(\tilde{Y}_E^{(n)}(0), \tilde{Z}_E^{(n)}(0)\right) - \epsilon_E(p_I, q_I)\right] \xrightarrow{D} \mathrm{N}\left(\mathbf{0}, \begin{bmatrix} \sigma_Y^2 & -\sigma_Y^2 \\ -\sigma_Y^2 & \sigma_Y^2 \end{bmatrix}\right)$$

as $n \to \infty$, where

$$\sigma_Y^2 = \sum_{i=1}^{d_{\max}} \epsilon_i \sigma_{Y,i}^2. \tag{5.15}$$

(A closed-form expression for σ_Y^2 is given by (B.27) in Appendix B.4.2.) Since $\tilde{\boldsymbol{X}}^{(n)}(0)$ is non-random, it follows using (2.1) that

$$\sqrt{n}\left(n^{-1}\tilde{\boldsymbol{W}}^{(n)}(0) - \tilde{\boldsymbol{w}}(0)\right) \xrightarrow{D} \mathrm{N}\left(\mathbf{0}, \Sigma_0^{\mathrm{MR}}\right) \qquad \text{as } n \to \infty, \tag{5.16}$$

where

$$\tilde{\boldsymbol{w}}(0) = (p_0 - \epsilon_0, p_1 - \epsilon_1, \ldots, p_{d_{\max}} - \epsilon_{d_{\max}}, \epsilon_E p_I, \epsilon_E q_I) \tag{5.17}$$

and

$$\Sigma_0^{\mathrm{MR}} = \begin{bmatrix} 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_Y^2 & -\sigma_Y^2 \\ \mathbf{0} & -\sigma_Y^2 & \sigma_Y^2 \end{bmatrix}. \tag{5.18}$$

Turning to the epidemic on the NSW random graph, recall that in $\mathcal{E}^{(n)}$ the number of initial infectives $a^{(n)}$ is prescribed and the $a^{(n)}$ initial infectives are chosen by sampling uniformly at random without replacement from the population. Thus the network can be constructed using two independent sets of i.i.d. copies of D, viz. $D'_1, D'_2, \ldots, D'_{n-a^{(n)}}$ for the initial susceptibles and $D_1, D_2, \ldots, D_{a^{(n)}}$ for the initial infectives. Let (Y_E, Z_E) be the bivariate random variable obtained by first sampling D and then letting $(Y_E, Z_E) = (Y_E, D - Y_E)$, where $Y_E|D \sim \mathrm{Bin}(D, 1 - \exp(-\lambda I))$. Let $\sigma^2_{Y_E} = \mathrm{var}(Y_E)$, $\sigma_{Y_E,Z_E} = \mathrm{cov}(Y_E, Z_E)$ and $\sigma^2_{Y_E} = \mathrm{var}(Y_E)$. (Closed-form expressions for $\sigma^2_{Y_E}$, $\sigma^2_{Z_E}$ and σ_{Y_E,Z_E} are given in (B.38)-(B.40) in Appendix B.4.3.) Let $\boldsymbol{p} = (p_0, p_1, \ldots, p_{d_{\max}})$ and Σ_{XX} be the $(d_{\max}+1) \times (d_{\max}+1)$ matrix with elements

$$(\Sigma_{XX})_{ij} = \begin{cases} -p_i p_j & \text{if } i \neq j, \\ p_i(1-p_i) & \text{if } i = j. \end{cases} \tag{5.19}$$

Recalling that $\lim_{n\to\infty} \sqrt{n}(\epsilon^{(n)} - \epsilon) = 0$, where $\epsilon^{(n)} = n^{-1}a^{(n)}$, a similar argument to the above shows that

$$\sqrt{n}\left(n^{-1}\tilde{\boldsymbol{W}}^{(n)}(0) - \tilde{\boldsymbol{w}}(0)\right) \xrightarrow{D} \mathrm{N}\left(\boldsymbol{0}, \Sigma_0^{\mathrm{NSW}}\right) \qquad \text{as } n \to \infty, \tag{5.20}$$

where

$$\tilde{\boldsymbol{w}}(0) = ((1-\epsilon)\boldsymbol{p}, \epsilon\mu_D p_I, \epsilon\mu_D q_I) \tag{5.21}$$

and

$$\Sigma_0^{\mathrm{NSW}} = \begin{bmatrix} (1-\epsilon)\Sigma_{XX} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \epsilon\sigma^2_{Y_E} & \epsilon\sigma_{Y_E,Z_E} \\ \boldsymbol{0} & \epsilon\sigma_{Y_E,Z_E} & \epsilon\sigma^2_{Z_E} \end{bmatrix}. \tag{5.22}$$

5.3.3. *Central limit theorem.* We show first that $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$ satisfies the conditions of Theorems 4.1 and 4.2. It is immediate from (5.2) and (5.3) that (4.3) and (4.7) are satisfied. The functions $\tilde{\beta}_{\boldsymbol{l}}(\boldsymbol{w})$ ($\boldsymbol{l} \in \Delta$) are clearly continuous on H. Note from (5.6) that $\tilde{F}(\boldsymbol{w})$ is a linear transformation of $\boldsymbol{w}$, so $\partial\tilde{F}$ is continuous on H and the Lipschitz condition (4.4) holds. Equations (5.11)-(5.13) imply that $\tilde{\boldsymbol{w}}(t)$ is finite for all $t \geq 0$. Finally, (5.16) and (5.20) imply that (4.8) (and hence also the corresponding convergence in probability in Theorem 4.1) holds for the epidemic on an MR and NSW random graph, respectively. Thus all the conditions of Theorem 4.2 are satisfied.

Note that $\tilde{\tau}^{(n)} = \inf\{t \geq 0 : \varphi(n^{-1}\tilde{\boldsymbol{W}}^{(n)}(t)) \leq 0\}$, where $\varphi(\boldsymbol{w}) = \varphi(\boldsymbol{x}, y_E, z_E) = y_E$. Suppose that $\tilde{x}_i(0) = p_i - \epsilon_i$ $(i = 0, 1, \ldots, d_{\max})$, $\tilde{y}_E(0) = p_I \sum_{i=1}^{d_{\max}} i\epsilon_i$ and $\tilde{z}_E(0) = q_I \sum_{i=1}^{d_{\max}} i\epsilon_i$, so $\tilde{\eta}_E(0) = \mu_D$. Let $\tilde{\tau} = \inf\{t \geq 0 : \varphi(\tilde{\boldsymbol{w}}(t)) \leq 0\} = \inf\{t \geq 0 : \tilde{y}_E(t) = 0\}$. Then it follows from (5.12) that $\tilde{\tau}$ satisfies the equation

$$\mathrm{e}^{-\tilde{\tau}} - q_I - \mu_D^{-1} p_I f_{D_\epsilon}^{(1)}(\mathrm{e}^{-\tilde{\tau}}) = 0. \tag{5.23}$$

We show in Appendix B.2 that, under the conditions of Theorem 2.1, the equation (5.23) has a unique solution in $(0, \infty)$. Note that $z = \mathrm{e}^{-\tilde{\tau}}$, where z is defined at (2.2). Also, using (5.11) the deterministic final fraction of the population that is susceptible is given by

$$\sum_{i=0}^{d_{\max}} \tilde{x}_i(\tilde{\tau}) = \sum_{i=0}^{d_{\max}} (p_i - \epsilon_i)\mathrm{e}^{-i\tilde{\tau}} = f_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}).$$

The corresponding deterministic final size is $\rho = 1 - \epsilon - f_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}})$, agreeing with (2.3).

Let $a(\tilde{\tau}) = \nabla\varphi(\boldsymbol{w}(\tilde{\tau})) \cdot \tilde{F}(\boldsymbol{w}(\tilde{\tau}))$. Then, using (5.6), and noting that $\tilde{y}_E(\tilde{\tau}) = 0$,

$$\begin{aligned} a(\tilde{\tau}) &= \sum_{i=1}^{d_{\max}} i\tilde{x}_i(\tilde{\tau})[(i-1)p_I - 1] - 2\tilde{y}_E(\tilde{\tau}) - \tilde{z}_E(\tilde{\tau}) \\ &= p_I \sum_{i=1}^{d_{\max}} i(i-1)\tilde{x}_i(\tilde{\tau}) - \tilde{\eta}_E(\tilde{\tau}). \end{aligned}$$

Thus, using (5.11) and (5.14),

$$a(\tilde{\tau}) = \mathrm{e}^{-2\tilde{\tau}} \left(p_I f_{D_\epsilon}^{(2)}(\mathrm{e}^{-\tilde{\tau}}) - \mu_D \right), \tag{5.24}$$

since $\tilde{\eta}_E(0) = \mu_D$. We show in Appendix B.2 that $a(\tilde{\tau}) < 0$, so we may apply Corollary 4.1.

Writing $\tilde{\boldsymbol{w}} = (\tilde{w}_0, \tilde{w}_1, \ldots, \tilde{w}_p)$, where $p = d_{\max} + 3$, let $\tilde{\Phi}(t, u) = [\tilde{\phi}_{ij}(t, u)]$ $(0 \leq u \leq t < \infty)$, where

$$\tilde{\phi}_{ij}(t, u) = \frac{\partial \tilde{w}_i(t-u)}{\partial \tilde{w}_j(0)} \qquad (i, j = 0, 1, \ldots, p). \tag{5.25}$$

Also, let

$$\tilde{\Sigma}(\tilde{\tau}) = \tilde{\Phi}(\tilde{\tau}, 0)\Sigma_0\tilde{\Phi}(\tilde{\tau}, 0)^\top + \int_0^{\tilde{\tau}} \tilde{\Phi}(\tilde{\tau}, u)\tilde{G}(\tilde{\boldsymbol{w}}(u))\tilde{\Phi}(\tilde{\tau}, u)^\top \, \mathrm{d}u, \tag{5.26}$$

where

$$\tilde{G}(\tilde{\boldsymbol{w}}(u)) = \sum_{\boldsymbol{l}\in\Delta} \boldsymbol{l}^\top \boldsymbol{l} \tilde{\beta}_{\boldsymbol{l}}(\tilde{\boldsymbol{w}}(u)), \tag{5.27}$$

and $\Sigma_0 = \Sigma_0^{\mathrm{MR}}$ or Σ_0^{NSW} depending on whether the epidemic is on an MR or an NSW random graph. Then application of Corollary 4.1 yields

$$\sqrt{n}\left(n^{-1}\tilde{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)}) - \tilde{\boldsymbol{w}}(\tilde{\tau})\right) \xrightarrow{D} \mathrm{N}(\boldsymbol{0}, B\tilde{\Sigma}(\tilde{\tau})B^\top) \qquad \text{as } n\to\infty, \tag{5.28}$$

where

$$B = I - \frac{\left(\tilde{F}(\tilde{\boldsymbol{w}}(\tilde{\tau}))\right)^\top \nabla\varphi(\tilde{\boldsymbol{w}}(\tilde{\tau}))}{\nabla\varphi(\tilde{\boldsymbol{w}}(\tilde{\tau}))\cdot\tilde{F}(\tilde{\boldsymbol{w}}(\tilde{\tau}))}.$$

Thus, recalling (5.4),

$$\sqrt{n}\left(n^{-1}T^{(n)} - \rho\right) \xrightarrow{D} \mathrm{N}(0,\sigma^2) \qquad \text{as } n\to\infty,$$

where

$$\sigma^2 = (\mathbf{1},0,0)B\tilde{\Sigma}(\tilde{\tau})B^\top(\mathbf{1},0,0)^\top. \tag{5.29}$$

Now $\nabla\varphi(\tilde{\boldsymbol{w}}(\tilde{\tau})) = (\boldsymbol{0},1,0)$ and, using (5.6), $(\mathbf{1},0,0)[\tilde{F}(\tilde{\boldsymbol{w}}(\tilde{\tau}))]^\top = -\tilde{x}_E(\tilde{\tau})$, so $(\mathbf{1},0,0)B = (\mathbf{1}, b(\tilde{\tau}), 0)$, where

$$b(\tilde{\tau}) = a(\tilde{\tau})^{-1}\tilde{x}_E(\tilde{\tau}). \tag{5.30}$$

Further, it follows from (5.11)-(5.13) and (5.25) that

$$\left[(\mathbf{1},0,0)\tilde{\Phi}(\tilde{\tau},u)\right]_i = \mathrm{e}^{-i(\tilde{\tau}-u)} \qquad (i=0,1\ldots,d_{\max})$$

and

$$\left[(\boldsymbol{0},1,0)\tilde{\Phi}(\tilde{\tau},u)\right]_i = \begin{cases} i\left[\mathrm{e}^{-2(\tilde{\tau}-u)} - q_I\mathrm{e}^{-(\tilde{\tau}-u)} - p_I\mathrm{e}^{-i(\tilde{\tau}-u)}\right] & \text{if } i=0,1,\ldots,d_{\max}, \\ \mathrm{e}^{-2(\tilde{\tau}-u)} & \text{if } i = d_{\max}+1, \\ -\mathrm{e}^{-(\tilde{\tau}-u)}(1-\mathrm{e}^{-(\tilde{\tau}-u)}) & \text{if } i = d_{\max}+2. \end{cases} \tag{5.31}$$

Further, let

$$\begin{aligned} \boldsymbol{c}(\tilde{\tau},u) &= (\mathbf{1},0,0)B\tilde{\Phi}(\tilde{\tau},u) \\ &= \left(\boldsymbol{c}_S(\tilde{\tau},u), c_I(\tilde{\tau},u), c_R(\tilde{\tau},u)\right), \end{aligned} \tag{5.32}$$

where

$$\boldsymbol{c}_S(\tilde{\tau},u) = (c_0(\tilde{\tau},u), c_1(\tilde{\tau},u), \ldots, c_{d_{\max}}(\tilde{\tau},u)), \tag{5.33}$$

$$c_I(\tilde{\tau},u) = b(\tilde{\tau})\mathrm{e}^{-2(\tilde{\tau}-u)}, \tag{5.34}$$

$$c_R(\tilde{\tau},u) = -b(\tilde{\tau})\mathrm{e}^{-(\tilde{\tau}-u)}(1-\mathrm{e}^{-(\tilde{\tau}-u)}) \tag{5.35}$$

and, for $i = 0, 1, \ldots, d_{\max}$,

$$c_i(\tilde{\tau},u) = \mathrm{e}^{-i(\tilde{\tau}-u)} + b(\tilde{\tau})i\left[\mathrm{e}^{-2(\tilde{\tau}-u)} - p_I\mathrm{e}^{-i(\tilde{\tau}-u)} - q_I\mathrm{e}^{-(\tilde{\tau}-u)}\right]. \tag{5.36}$$

Noting that $\boldsymbol{c}(\tilde{\tau},u)\boldsymbol{l}^\top$ is a scalar, it follows from (5.26) and (5.29) that

$$\sigma^2 = \boldsymbol{c}(\tilde{\tau},0)\Sigma_0\boldsymbol{c}(\tilde{\tau},0)^\top + \sum_{\boldsymbol{l}\in\Delta}\int_0^{\tilde{\tau}}\left(\boldsymbol{c}(\tilde{\tau},u)\boldsymbol{l}^\top\right)^2\tilde{\beta}_{\boldsymbol{l}}(\tilde{\boldsymbol{w}}(u))\,\mathrm{d}u. \tag{5.37}$$

The asymptotic variances, σ^2_{MR} and σ^2_{NSW} in Theorem 2.1 can be obtained by substituting $\Sigma_0 = \Sigma_0^{\mathrm{MR}}$ and $\Sigma_0 = \Sigma_0^{\mathrm{NSW}}$, respectively, in (5.37) and using (5.32)-(5.36), together with (5.3) and (5.11)-(5.13), to evaluate the second term on the right-hand side of (5.37). The details are lengthy and may be found in Appendix B.4.

5.4. *Proof of Theorem 2.2.* We prove Theorem 2.2 for the epidemic on an NSW random graph. The proof for the epidemic on an MR random graph is similar but simpler, as there is no randomness in the degrees of individuals, and is thus omitted. The proof proceeds in two stages. First, in Section 5.4.1, we couple the early stages of the epidemic $\tilde{\mathcal{E}}^{(n)}$, defined in Section 5.1, to a two-type branching process $\tilde{\mathcal{B}}^{(n)}$ which assumes that all infective half-edges in $\tilde{\mathcal{E}}^{(n)}$ are paired with susceptible half-edges. The branching processes $\tilde{\mathcal{B}}^{(n)}$ $(n = 1, 2, \ldots)$ are coupled to a limiting branching process $\tilde{\mathcal{B}}$. The couplings and standard properties of the limiting branching process $\tilde{\mathcal{B}}$ show that, with probability tending to 1 as $n \to \infty$, a major outbreak occurs if and only if the branching process $\tilde{\mathcal{B}}$ does not go extinct, and yield weak convergence results concerning the composition of the population in $\tilde{\mathcal{E}}^{(n)}$ when the number of infective half-edges first reaches $\log n$ in the event of a major outbreak (see Theorem 5.1). Then, in Section 5.4.2, we use the random time-scale transformation introduced in Section 5.2 to determine the asymptotic distribution of the final size of a major outbreak. The argument proceeds as in the proof of Theorem 2.1 but the equation defining $\tilde{\tau}$ now has a solution at 0 and one at $\tilde{\tau} > 0$ (see the discussion following (5.49)) and a lower bounding branching process for the epidemic $\tilde{\mathcal{E}}^{(n)}$ is used to show that $\tilde{\tau} > 0$ is the relevant asymptotic hitting time.

For ease of presentation we assume that, for $n = 1, 2, \ldots$, there is one initial infective in $\tilde{\mathcal{E}}^{(n)}$ (i.e. that $a = 1$), who is chosen by sampling a half-edge uniformly at random from all $D_1^{(n)} + D_2^{(n)} + \cdots + D_n^{(n)}$ half-edges and infecting the individual who owns that half-edge. The proofs are easily extended to $a > 1$ and other ways of choosing the initial infective(s) but the details are more complicated.

5.4.1. *Coupling of epidemic and branching processes.* Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space on which are defined the following independent sets of random variables:

(i) $D_1, D_2, \ldots$ i.i.d. $\sim D$;
(ii) $U_0, U_1, \ldots$ i.i.d. $\sim \mathrm{U}(0,1)$;
(iii) $L_1, L_2, \ldots$ i.i.d. $\sim \mathrm{Exp}(1)$;
(iv) $Y_{01}, Y_{02}, \ldots Y_{0d_{\max}}$, where $Y_{0i} \sim \mathrm{Bin}(i, 1 - \mathrm{e}^{-\lambda I})$;
(v) for $i = 1, 2, \ldots, d_{\max}$, $Y_{i1}, Y_{i2}, \ldots$ i.i.d. $\sim \mathrm{Bin}(i-1, 1 - \mathrm{e}^{-\lambda I})$.

For $n = 1, 2, \ldots$, let $p_i^{(n)} = n^{-1} \sum_{k=1}^n 1_{\{D_k = n\}}$ $(i = 0, 1, \ldots, d_{\max})$ and $\tilde{p}_i^{(n)} = i p_i^{(n)} / \mu_D^{(n)}$ $(i = 1, 2, \ldots, d_{\max})$, where $\mu_D^{(n)} = n^{-1} \sum_{k=1}^n D_k$. Note that by the strong law of large numbers $\mu_D^{(n)} > 0$ (and $\tilde{p}_i^{(n)}$ is well defined) for all sufficiently large n almost surely. Let $\tilde{c}_i^{(n)} = \sum_{j=1}^i \tilde{p}_j^{(n)}$ $(i = 1, 2, \ldots, d_{\max})$. For $x \in (0,1)$, let $\tilde{d}^{(n)}(x) = \min\{i : x \le \tilde{c}_i^{(n)}\}$. Similarly, let $\tilde{p}_i = i p_i / \mu_D$ and $\tilde{c}_i = \sum_{j=1}^i \tilde{p}_j$ $(i = 1, 2, \ldots, d_{\max})$, and let $\tilde{d}(x) = \min\{i : x \le \tilde{c}_i\}$ $(0 < x < 1)$.

For $n = 1, 2, \ldots$, construct on $(\Omega, \mathcal{F}, \mathrm{P})$ a realisation of a two-type continuous-time Markov branching process $\tilde{\mathcal{B}}^{(n)}$, which approximates the process of infected and recovered half-edges in the epidemic $\tilde{\mathcal{E}}^{(n)}$, as follows. The types are denoted I and R depending on whether the individual corresponds to an infective or recovered half-edge. Only type-I individuals have offspring and they do so at their moment of death. Type-R individuals live forever. For $t \ge 0$, let $\hat{Y}_E^{(n)}(t)$ and $\hat{Z}_E^{(n)}(t)$ denote respectively the numbers of type-I and type-R individuals alive in $\tilde{\mathcal{B}}^{(n)}$ at time t. The initial state $(\hat{Y}_E^{(n)}(0), \hat{Z}_E^{(n)}(0))$ is determined as follows. Let $d_0^{(n)} = \tilde{d}^{(n)}(U_0)$, which corresponds to the initial infective in $\tilde{\mathcal{E}}^{(n)}$ having degree $d_0^{(n)}$. If $d_0^{(n)} = 0$ then $(\hat{Y}_E^{(n)}(0), \hat{Z}_E^{(n)}(0)) = (0, 0)$ and $\tilde{\mathcal{B}}^{(n)}$ goes extinct immediately. Alternatively, if $d_0^{(n)} > 0$ then $(\hat{Y}_E^{(n)}(0), \hat{Z}_E^{(n)}(0)) = (Y_{0d_0^{(n)}}, d_0 - Y_{0d_0^{(n)}})$. In that case, for $k = 1, 2, \ldots$, the kth type-I individual born in $\tilde{\mathcal{B}}^{(n)}$ (including the initial individuals) has degree $d_k^{(n)} = \tilde{d}^{(n)}(U_k)$ and lives until age L_k, when it dies. Denote this individual by i^*. Suppose that $d_k^{(n)} = i$ and i^* is the lth degree-i

individual (excluding the initial individuals) born in $\tilde{\mathcal{B}}^{(n)}$. Then, when i^* dies, it leaves Y_{il} type-I and $i-1-Y_{il}$ type-R offspring. Reproduction stops in $\tilde{\mathcal{B}}^{(n)}$ if $\hat{Y}_E^{(n)}(t)=0$. Construct also on $(\Omega,\mathcal{F},\mathrm{P})$ a realisation of a two-type branching process $\tilde{\mathcal{B}}$, defined analogously to $\tilde{\mathcal{B}}^{(n)}$ but using the function $\tilde{d}$ instead of $\tilde{d}^{(n)}$. For $t\geq 0$, let $Y_E(t)$ and $Z_E(t)$ denote respectively the numbers of type-I and type-R individuals alive in $\tilde{\mathcal{B}}$ at time t.

For $n=1,2,\ldots$, construct on $(\Omega,\mathcal{F},\mathrm{P})$ a realisation of the epidemic $\tilde{\mathcal{E}}^{(n)}$, defined in Section 5.1, as follows. Give the n individuals in $\tilde{\mathcal{E}}^{(n)}$ the labels $1,2,\ldots,n$ in increasing order of degree. Now label the $n\mu_D^{(n)}$ half-edges $1,2,\ldots,n\mu_D^{(n)}$, starting with the half-edges (if any) attached to individual 1, then the half-edges (if any) attached to individual 2 and so on. Thus half-edges attached to the same individual have consecutive labels. As in Section 5.1, for $t\geq 0$, let $Y_E^{(n)}(t)$ and $Z_E^{(n)}(t)$ denote respectively the number of infective and recovered half-edges in $\tilde{\mathcal{E}}^{(n)}$ at time t. The initial infective in $\tilde{\mathcal{E}}^{(n)}$ is the individual who owns the half-edge labelled $\left\lfloor n\mu_D^{(n)}U_0\right\rfloor+1$. Note that this individual has degree $d_0^{(n)}=\tilde{d}^{(n)}(U_0)$. If $d_0^{(n)}=0$ then the epidemic stops immediately. Alternatively, if $d_0^{(n)}>0$ then the initial infective infects along $k_0^{(n)}=Y_{0d_0^{(n)}}$ of its half-edges with its remaining $d_0^{(n)}-k_0^{(n)}$ half-edges becoming recovered half-edges. The epidemic stops if $k_0^{(n)}=0$, otherwise the initial infective transmits infection along its infective half-edges at times $L_1,L_2,\ldots,L_{k_0}$.

For $k=1,2,\ldots$, let $l_k^{(n)}=\left\lfloor n\mu_D^{(n)}U_k\right\rfloor+1$ and note that the half-edge having label $l_k^{(n)}$ is attached to an individual having degree $d_k^{(n)}=\tilde{d}^{(n)}(U_k)$. When infection is transmitted along a half-edge that half-edge, l_* say, is attempted to be paired with the half-edge having label $l_k^{(n)}$, where k is the number of the $U_0,U_1,\ldots$ that have been used already in the construction of $\tilde{\mathcal{E}}^{(n)}$. (Thus, for example, the first half-edge emanating from the initial infective is attempted to be paired with the half-edge having label $l_1^{(n)}$.) If the half-edge $l_k^{(n)}$ has already been paired or $l_k^{(n)}=l_*$ then the attempt fails and l_* is attempted to be paired with the half-edge $l_{k+1}^{(n)}$, and so on until a valid pairing is obtained and an edge is formed. Suppose that a valid pairing is made with the half-edge having label l_V. Let i^* be the individual that owns the half-edge l_V and i be the degree of i^*. If i^* is susceptible then it becomes an infective, otherwise nothing happens apart from the formation of the edge. Suppose that i^* is susceptible and is the lth degree-i susceptible to be infected in $\tilde{\mathcal{E}}^{(n)}$, excluding the initial infective. Then i^* infects along Y_{il} of its half-edges and its remaining $i-1-Y_{il}$ half-edges become recovered

half-edges. (When i^* was infected one of its i half-edges was paired to its infector.) Let $k_1 = Y_{il}$. The times of these k_1, infections, relative to the infection time of i^*, are given by $L_{k^*+1}, L_{k^*+2}, \ldots, L_{k^*+k_1}$, where k^* is the number of infective half-edges created in $\tilde{\mathcal{E}}^{(n)}$ prior to the infection of i^*. The epidemic terminates when $Y_E^{(n)}(t) = 0$, or when $Y_E^{(n)}(t) = 1$ and all other half-edges have been paired.

As in Section 5.1, for $t \geq 0$, let $\boldsymbol{X}^{(n)}(t) = (X_0^{(n)}(t), X_1^{(n)}(t), \ldots, X_{d_{\max}}^{(n)}(t))$, where $X_i^{(n)}(t)$ is the number of degree-i susceptible individuals at time t in $\tilde{\mathcal{E}}^{(n)}$. For $n = 1, 2, \ldots$, let $\hat{\tau}^{(n)} = \inf\{t \geq 0 : Y_E^{(n)}(t) \geq \log n\}$, where $\hat{\tau}^{(n)} = \infty$ if $Y_E^{(n)}(t) < \log n$ for all $t \geq 0$. Let $A_{\text{ext}} = \{\omega \in \Omega : \lim_{t\to\infty} Y_E(t) = 0\}$ denote the set on which type-I individuals become extinct in the branching process $\tilde{\mathcal{B}}$ and let α denote the Malthusian parameter of $\tilde{\mathcal{B}}$. Then α is given by the unique real solution of the equation

$$\int_0^\infty \mathrm{e}^{-\alpha t} \mu_{\tilde{D}-1} p_I \mathrm{e}^{-t}\, \mathrm{d}t = 1,$$

so $\alpha = R_0 - 1$, where R_0 is defined at (2.8).

THEOREM 5.1. *(a)* $\lim_{n\to\infty} \mathrm{P}\left(\hat{\tau}^{(n)} = \infty | A_{\text{ext}}\right) = 1$.
(b) Suppose that $R_0 > 1$, *so* $\mathrm{P}(A_{\text{ext}}^C) > 0$. *Then, as* $n \to \infty$,
(i) $\mathrm{P}\left(\hat{\tau}^{(n)} < \infty | A_{\text{ext}}^C\right) \to 1$;
(ii) $\frac{1}{\log n} Y_E^{(n)}(\hat{\tau}^{(n)}) \mid A_{\text{ext}}^C \xrightarrow{p} 1$;
(iii) $\frac{1}{\log n} Z_E^{(n)}(\hat{\tau}^{(n)}) \mid A_{\text{ext}}^C \xrightarrow{p} \alpha^{-1} q_I \mu_{\tilde{D}-1}$;
(iv) $\sqrt{n}\left(n^{-1}\boldsymbol{X}^{(n)}(\hat{\tau}^{(n)}) - \boldsymbol{p}\right) | A_{\text{ext}}^C \xrightarrow{D} \mathrm{N}(\boldsymbol{0}, \Sigma_{XX})$, *where* Σ_{XX} *is defined at* (5.19).

PROOF. The key observations underlying the proof are that (i) the processes $\{(\hat{Y}_E^{(n)}(t), \hat{Z}_E^{(n)}(t)) : t \geq 0\}$ and $\{(Y_E^{(n)}(t), Z_E^{(n)}(t)) : t \geq 0\}$ coincide up until at least the first time that an attempt is made to pair a half-edge with a half-edge belonging to an individual previously used in the construction of $\tilde{\mathcal{E}}^{(n)}$ and (ii) the branching processes $\tilde{\mathcal{B}}^{(n)}$ and $\tilde{\mathcal{B}}$ coincide up until the first time that $\tilde{d}^{(n)}(U_k) \neq \tilde{d}(U_k)$. Thus we show that the probability that the processes $\{(\hat{Y}_E^{(n)}(t), \hat{Z}_E^{(n)}(t)) : 0 \leq t \leq \hat{\tau}^{(n)}\}$ and $\{(Y_E(t), Z_E(t)) : 0 \leq t \leq \hat{\tau}^{(n)}\}$ coincide converges to 1 as $n \to \infty$. The theorem then follows using standard results concerning the asymptotic behaviour of the branching process $\tilde{\mathcal{B}}$.

For $n = 1, 2, \ldots$ and $k = 1, 2, \ldots, n\mu_D^{(n)}$, let $\mathcal{C}^{(n)}(k)$ be the set of half-edges attached to the individual owning half-edge k in $\tilde{\mathcal{E}}^{(n)}$. Thus $k \in \mathcal{C}^{(n)}(k)$ and

$|\mathcal{C}^{(n)}(k)| \leq d_{\max}$. For $n = 1, 2, \ldots$, let

$$M^{(n)} = \min\left\{k \geq 1 : l_k^{(n)} \in \bigcup_{i=0}^{k-1} \mathcal{C}^{(n)}(l_i^{(n)})\right\}$$

be the number of pairings required in $\tilde{\mathcal{E}}^{(n)}$ until an attempt is made to pair a half-edge with a half-edge belonging to a previously used individual. Let $\boldsymbol{D} = (D_1, D_2, \ldots)$. Then, for $m = 1, 2, \ldots$,

$$\begin{aligned} \mathrm{P}(M^{(n)} \leq m|\boldsymbol{D}) &= \mathrm{P}\left(\bigcup_{k=1}^{m}\left\{l_k^{(n)} \in \bigcup_{i=0}^{k-1} \mathcal{C}^{(n)}(l_i^{(n)})\right\}|\boldsymbol{D}\right) \\ &\leq \frac{m(m+1)}{2}\frac{d_{\max}}{n\mu_D^{(n)}}. \end{aligned}$$

Now $\mu_D^{(n)} \xrightarrow{a.s.} \mu_D$ as $n \to \infty$ by the strong law of large numbers. Hence, for any $\gamma \in (0, 1/2)$, $\mathrm{P}(M^{(n)} \leq n^\gamma|\boldsymbol{D}) \xrightarrow{p} 0$ as $n \to \infty$. Thus, for such γ,

$$\lim_{n\to\infty} \mathrm{P}(M^{(n)} \leq n^\gamma) = \lim_{n\to\infty} \mathrm{E}[\mathrm{P}(M^{(n)} \leq n^\gamma|\boldsymbol{D})] = 0, \tag{5.38}$$

as $\mathrm{P}(M^{(n)} \leq n^\gamma|\boldsymbol{D})$ $(n = 1, 2, \ldots)$ is uniformly bounded.

Turning to the coupling of $\tilde{\mathcal{B}}^{(n)}$ and $\tilde{\mathcal{B}}$, for $n = 2, \ldots$,

$$\begin{aligned} \max_{i=1,2,\ldots,d_{\max}-1} |\tilde{c}_i^{(n)} - \tilde{c}_i| &\leq \frac{1}{2}\sum_{i=1}^{d_{\max}} |\tilde{p}_i^{(n)} - \tilde{p}_i| \\ &\leq \frac{1}{\mu_D}\sum_{i=1}^{d_{\max}} i|p_i^{(n)} - p_i|. \end{aligned} \tag{5.39}$$

The first inequality in (5.39) follows from the total variation distance between the distributions $\tilde{p}_i^{(n)}$ $(i = 1, 2, \ldots, d_{\max})$ and $\tilde{p}_i$ $(i = 1, 2, \ldots, d_{\max})$, and the second inequality follows using the triangle inequality and elementary manipulation (Ball and Neal (2017), equation (3.2)). For $n = 1, 2, \ldots$, let $\hat{M}^{(n)} = \min\{k \geq 0 : \tilde{d}^{(n)}(U_k) \neq \tilde{d}(U_k)\}$. Then, for $k = 0, 1, \ldots$,

$$\mathrm{P}\left(\hat{M}_n \leq k\right) \leq \frac{(k+1)(d_{\max}-1)}{\mu_D}\mathrm{E}\left[\sum_{i=1}^{d_{\max}} i|p_i^{(n)} - p_i|\right].$$

Now, for $i = 1, 2, \ldots, d_{\max}$, as $np_i^{(n)} \sim \mathrm{Bin}(n, p_i)$,

$$\mathrm{E}\left[|p_i^{(n)} - p_i|\right] \leq \sqrt{\mathrm{var}(p_i^{(n)})} = \frac{1}{\sqrt{n}}\sqrt{p_i(1-p_i)}.$$

Thus, for any $\gamma \in (0, 1/2)$,

$$\lim_{n\to\infty} \mathrm{P}\left(\hat{M}^{(n)} \le n^{\gamma}\right) = 0. \tag{5.40}$$

For $t \ge 0$, let $T_E(t)$ be the total number of individuals (of either type) alive in the branching process $\tilde{\mathcal{B}}$ during $[0, t]$ and let $T_E(\infty) = \lim_{t\to\infty} T_E(t)$. Thus $T_E(\infty, \omega) < \infty$ if and only if $\omega \in A_{\mathrm{ext}}$. For $n = 1, 2, \ldots$ and $t \ge 0$, let

$$A^{(n)}_{\mathrm{couple}}(t) = \left\{\omega \in \Omega : \left(Y_E^{(n)}(u), Z_E^{(n)}(u)\right) = (Y_E(u), Z_E(u)) \text{ for all } u \in [0, t]\right\}.$$

Let $\tau = \min\{t \ge 0 : Y_E(t) = 0\}$, where $\tau(\omega) = \infty$ if $\omega \in A^C_{\mathrm{ext}}$. Then (5.38) and (5.40) imply that, for $k = 0, 1, \ldots,$

$$\lim_{n\to\infty} \mathrm{P}\left(A^{(n)}_{\mathrm{couple}}(\tau) \cap \{T_E(\infty) = k\}\right) = \mathrm{P}(T_E(\infty) = k).$$

Part (a) of the theorem follows since $A_{\mathrm{ext}} = \bigcup_{k=0}^{\infty}\{T_E(\infty) = k\}$.

Note that although $\tilde{\mathcal{B}}$ is a two-type branching process, since only type-I individuals produce offspring, it can be treated as a single-type branching process consisting only of type-I individuals, in which attached to each (type-I) individual is a random characteristic (see Nerman (1981)) given by its number of type-R offspring. Then it follows from Nerman (1981), Theorem 5.4, that there exists a random variable $W \ge 0$, where $W(\omega) = 0$ if and only if $\omega \in A_{\mathrm{ext}}$ such that, as $t \to \infty$,

$$\mathrm{e}^{-\alpha t} Y_E(t) \xrightarrow{a.s.} W, \tag{5.41}$$

$$\mathrm{e}^{-\alpha t} Z_E(t) \xrightarrow{a.s.} \alpha^{-1} q_I \mu_{\tilde{D}-1} W \tag{5.42}$$

and

$$\mathrm{e}^{-\alpha t} T_E(t) \xrightarrow{a.s.} \alpha^{-1} \mu_{\tilde{D}-1} W. \tag{5.43}$$

(It is easily checked that (5.41)-(5.43) hold by calculating the the appropriate m^{ϕ}_{∞} in Nerman (1981), Theorem 5.4, and using Corollary 3.2 of that paper. For a heuristic argument note, for example, that if $Y_E(u) \approx W\mathrm{e}^{\alpha u}$ $(u \ge 0)$ then $T_E(t) \approx \int_0^t \mu_{\tilde{D}-1}\mathrm{e}^{\alpha u} W \,\mathrm{d}u \approx \alpha^{-1}\mu_{\tilde{D}-1} W\mathrm{e}^{\alpha t}$, since type-$I$ individuals die at rate 1 and have on average $\mu_{\tilde{D}-1}$ offspring.)

For $n = 1, 2, \ldots,$ let $\bar{\tau}^{(n)} = \inf\{t \ge 0 : Y_E(t) \ge \log n\}$, where $\bar{\tau}^{(n)}(\omega) = \infty$ if $\omega \in A_{\mathrm{ext}}$. Then it follows from (5.41) that $\bar{\tau}^{(n)}(\omega) < \infty$ and $\lim_{n\to\infty} \bar{\tau}^{(n)}(\omega) = \infty$, for P-almost all $\omega \in A^C_{\mathrm{ext}}$. Thus, for P-almost all $\omega \in A^C_{\mathrm{ext}}$,

$$\lim_{n\to\infty} \frac{1}{\log n} Y_E(\bar{\tau}^{(n)}) = 1, \tag{5.44}$$

since an individual has at most $d_{\max}$ offspring, and using (5.41)-(5.43),

$$\lim_{n\to\infty} \frac{1}{\log n} Z_E(\bar{\tau}^{(n)}) = \alpha^{-1} q_I \mu_{\tilde{D}-1}$$

and

$$\lim_{n\to\infty} \frac{1}{\log n} T_E(\bar{\tau}^{(n)}) = \alpha^{-1} \mu_{\tilde{D}-1}. \tag{5.45}$$

Now (5.38), (5.40) and (5.45) imply that

$$\lim_{n\to\infty} \mathrm{P}\left(A^{(n)}_{\mathrm{couple}}(\bar{\tau}^{(n)}) | A^C_{\mathrm{ext}} \right) = 1, \tag{5.46}$$

and (i)-(iii) of part (b) of the theorem follow using (5.42) and (5.44).

To prove part (b)(iv), note by the multivariate central limit theorem that

$$\sqrt{n}\left(n^{-1} \boldsymbol{X}^{(n)}(0) - \boldsymbol{p} \right) \xrightarrow{D} \mathrm{N}(\mathbf{0}, \Sigma_{XX}) \qquad \text{as } n \to \infty.$$

Also, (5.45) and (5.46) imply that, for $i = 0, 1, \ldots, d_{\max}$,

$$\lim_{n\to\infty} \mathrm{P}\left(X_i^{(n)}(\hat{\tau}^{(n)}) - X_i^{(n)}(0) \le 2\alpha^{-1} \mu_{\tilde{D}-1} \log n | A^C_{\mathrm{ext}} \right) = 1,$$

so $\frac{1}{\sqrt{n}} \left(X_i^{(n)}(\hat{\tau}^{(n)}) - X_i^{(n)}(0) \right) | A^C_{\mathrm{ext}} \xrightarrow{p} 0$ as $n \to \infty$. Part (b)(iv) now follows using Slutsky's theorem. □

5.4.2. *Epidemic starting with* $\lceil \log n \rceil$ *infectives.* Turning to the proof of Theorem 2.2, note that by Theorem 5.1, $\lim_{n\to\infty} \mathrm{P}\left(G^{(n)} \triangle A_{\mathrm{ext}} \right) = 0$, where $\triangle$ denotes symmetric difference. Hence, using the strong Markov property, we can determine the asymptotic distribution of $T^{(n)}_{\mathrm{NSW}} | G^{(n)}$ by considering the random time-scale transformed process $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$, defined in Section 5.2, with initial state $\tilde{\boldsymbol{W}}^{(n)}(0) = (\boldsymbol{X}^{(n)}(\hat{\tau}^{(n)}), Y_E^{(n)}(\hat{\tau}^{(n)}), Z_E^{(n)}(\hat{\tau}^{(n)}))$. Thus, by Theorem 5.1(b),

$$\sqrt{n}\left(n^{-1} \tilde{\boldsymbol{W}}^{(n)}(0) - (\boldsymbol{p}, 0, 0) \right) \xrightarrow{D} \mathrm{N}(\mathbf{0}, \Sigma_0^{\mathrm{NSW}}) \qquad \text{as } n \to \infty, \tag{5.47}$$

where Σ_0^{NSW} is obtained by setting $\epsilon = 0$ in (5.22).

By Theorem 4.1, for any $t > 0$,

$$\sup_{0 \le u \le t} |n^{-1} \tilde{Y}_E^{(n)}(u) - \tilde{y}_E(u)| \xrightarrow{p} 0 \qquad \text{as } n \to \infty, \tag{5.48}$$

where, using (5.12)

$$\tilde{y}_E(t) = \mu_D \left(\mathrm{e}^{-2t} - q_I \mathrm{e}^{-t}\right) - p_I f_D^{(1)}(\mathrm{e}^{-t}).$$

Note that $\tilde{y}_E(t) = 0$ if and only if

$$\mathrm{e}^{-t} - q_I - \mu_D^{-1} p_I f_D^{(1)}(\mathrm{e}^{-t}) = 0 \tag{5.49}$$

and that $t = 0$ is a solution of (5.49). If $R_0 \leq 1$ then $t = 0$ is the only solution in $[0, \infty)$, but if $R_0 > 1$ then we show in Appendix B.2 that, under the conditions of Theorem 2.2, there is a unique solution in $(0, \infty)$ which we denote by $\tilde{\tau}$.

As in Section 5.3.3, let $\tilde{\tau}^{(n)} = \inf\{t \geq 0 : \varphi(n^{-1}\tilde{\boldsymbol{W}}^{(n)}(t)) \leq 0\}$. Recall that $R_0 > 1$ is assumed in Theorem 2.2. It follows from (5.48) that, $\min\left(\tilde{\tau}^{(n)}, |\tilde{\tau}^{(n)} - \tilde{\tau}|\right) \xrightarrow{p} 0$ as $n \to \infty$. We show below that there exists $\epsilon_0 > 0$ such that

$$\lim_{n\to\infty} \mathrm{P}(\tilde{\tau}^{(n)} < \epsilon_0) = 0. \tag{5.50}$$

Theorem 4.3 and Corollary 4.1, as stated, cannot be applied in the present setting as $\varphi(\tilde{\boldsymbol{w}}(0)) = 0$. However, in the terminology of Theorem 4.3, the proof of Ethier and Kurtz (1986), Theorem 11.4.1, extends easily to the situation when $\tau^{(n)} = \inf\{t \geq \epsilon_0 : \varphi(n^{-1}\boldsymbol{X}^{(n)}(t)) \leq 0\}$ and $\tau = \inf\{t \geq \epsilon_0 : \varphi(\boldsymbol{x}(t)) \leq 0\}$, for fixed $\epsilon_0 > 0$. Hence, so do Theorem 4.3 and Corollary 4.1.

Let

$$\tilde{T}^{(n)} = \sum_{i=0}^{d_{\max}} \left(\tilde{X}_i^{(n)}(0) - \tilde{X}_i^{(n)}(\tilde{\tau}^{(n)})\right). \tag{5.51}$$

Then using (5.50) and the above extension of Corollary 4.1 (we show in Appendix B.2 that $a(\tilde{\tau}) = \nabla\varphi(\boldsymbol{w}(\tilde{\tau})) \cdot \tilde{F}(\boldsymbol{w}(\tilde{\tau})) < 0$), setting $\epsilon = 0$ in the proof of Theorem 2.1 yields

$$\sqrt{n}\left(n^{-1}\tilde{T}^{(n)} - \rho\right) \xrightarrow{D} \mathrm{N}(0, \tilde{\sigma}^2_{\mathrm{NSW}}) \qquad \text{as } n \to \infty.$$

Theorem 2.2 follows using Slutsky's theorem since, by (5.45),

$$\frac{1}{\sqrt{n}}\left(T_{\mathrm{NSW}}^{(n)} - \tilde{T}^{(n)}\right) \xrightarrow{p} 0 \qquad \text{as } n \to \infty.$$

To complete the proof of Theorem 2.2 we use a device, first introduced by Whittle (1955) in the setting of a homogeneously mixing epidemic, to

show that there exists $\epsilon_0 > 0$ such that (5.50) holds. For $n = 1, 2, \ldots,$ construct on $(\Omega, \mathcal{F}, \mathrm{P})$ an epidemic $\check{\mathcal{E}}^{(n)}$ analogously to $\tilde{\mathcal{E}}^{(n)}$ except having initial state given in an obvious notation by

$$\left(\check{\boldsymbol{X}}^{(n)}(0), \check{Y}_E^{(n)}(0), \check{Z}_E^{(n)}(0)\right) = \left(\boldsymbol{X}^{(n)}(\hat{\tau}^{(n)}), Y_E^{(n)}(\hat{\tau}^{(n)}), Z_E^{(n)}(\hat{\tau}^{(n)})\right).$$

Let $\check{\mathcal{B}}^{(n)}$ be the corresponding branching process which assumes all attempted pairing in $\check{\mathcal{E}}^{(n)}$ are valid. Let

$$\check{T}^{(n)} = \sum_{i=0}^{d_{\max}} \left(\check{X}_i^{(n)}(0) - \check{X}_i^{(n)}(\infty)\right)$$

be the final size of $\check{\mathcal{E}}^{(n)}$. Fix $\delta \in (0, 1)$. Then, provided $\check{T}^{(n)} \leq n\delta$, for a given pairing, the probability that an invalid pairing is attempted is at most $p^{(n)}(\delta) = \min(\delta n d_{\max}/(n\mu_D^{(n)}), 1) = \min(\delta d_{\max}/\mu_D^{(n)}, 1)$. It follows that

$$\mathrm{P}\left(\check{T}^{(n)} \geq n\delta\right) \geq \mathrm{P}\left(\check{T}_B^{(n)}(\delta) \geq n\delta\right), \tag{5.52}$$

where $\check{T}_B^{(n)}(\delta)$ is the total number of type-I individuals (excluding the initial individuals) ever alive in the branching process $\check{\mathcal{B}}^{(n)}(\delta)$ that is obtained from $\check{\mathcal{B}}^{(n)}$ by aborting each type-I individual independently with probability $p^{(n)}(\delta)$.

Let $R_0^{(n)}(\delta)$ be the mean number of type-I individuals spawned by a typical type-I individual in $\check{\mathcal{B}}^{(n)}(\delta)$ and let $\pi^{(n)}(\delta)$ be the extinction probability for type-I individuals in $\check{\mathcal{B}}^{(n)}(\delta)$ if the process were to start with one type-I individual. As $n \to \infty$, $p_i^{(n)} \xrightarrow{a.s.} p_i$ $(i = 0, 1, \ldots, d_{\max})$, by the strong law of large numbers, so $\tilde{p}_i^{(n)} \xrightarrow{a.s.} \tilde{p}_i$ $(i = 1, 2, \ldots, d_{\max})$ and $\mu_{\tilde{D}-1}^{(n)} \xrightarrow{a.s.} \mu_{\tilde{D}-1}$, where $\mu_{\tilde{D}-1}^{(n)} = \sum_{i=1}^{d_{\max}} (i-1)\tilde{p}_i^{(n)}$. Thus,

$$R_0^{(n)}(\delta) = (1 - p^{(n)}(\delta))\mu_{\tilde{D}-1}^{(n)} p_I \xrightarrow{a.s.} (1 - p(\delta))R_0 \qquad \text{as } n \to \infty,$$

where $p(\delta) = \min(\delta d_{\max}/\mu_D, 1)$. Further, as $n \to \infty$, the offspring distribution of $\check{\mathcal{B}}^{(n)}(\delta)$ converges almost surely to that of $\tilde{\mathcal{B}}(\delta)$, the branching process obtained from $\tilde{\mathcal{B}}$ by aborting each type-I individual independently with probability $p(\delta)$. Thus, by a simple extension of Britton et al. (2007), Lemma 4.1, $\pi^{(n)}(\delta) \xrightarrow{a.s.} \pi(\delta)$ as $n \to \infty$, where $\pi(\delta)$ is the extinction probability for type-I individuals in $\tilde{\mathcal{B}}(\delta)$ starting from one type-I individual.

Recall that $R_0 > 1$. Thus $\delta \in (0, 1)$ can be chosen sufficiently small such that $(1 - p(\delta))R_0 > 1$, whence $\pi(\delta) < 1$.

Now, $\check{Y}_E^{(n)}(0) \geq \log n$ by construction, so using (5.52), for such δ,

$$\lim_{n\to\infty} \mathrm{P}\left(\check{T}^{(n)} \geq n\delta | \boldsymbol{D}\right) \geq 1 - \lim_{n\to\infty} \left(\pi^{(n)}\right)^{\log n} = 1,$$

for P-almost all $\omega \in A_{\mathrm{ext}}^C$. Hence

$$\lim_{n\to\infty} \mathrm{P}\left(\check{T}^{(n)} \geq n\delta | A_{\mathrm{ext}}^C\right) = 1. \tag{5.53}$$

Finally, for $t \geq 0$, let $\tilde{x}(t) = \sum_{i=0}^{d_{\max}} \tilde{x}_i(t) = f_D(\mathrm{e}^{-t})$, using (5.11) with $\tilde{x}_i(0) = p_i$ $(i = 0, 1, \ldots, d_{\max})$. By construction, $\boldsymbol{W}^{(n)}(\tau^{(n)}) = \tilde{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)})$, so (5.53) and Theorem 4.1 applied to the process $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$, with initial state satisfying (5.47), imply that (5.50) holds with ϵ_0 being given by the unique solution in $(0, \infty)$ of $1 - f_D(\mathrm{e}^{-\epsilon_0}) = \delta$.

5.5. *Proof of Theorem 2.3.* The proof of Theorem 2.3 for the epidemic on the NSW random graph parallels in an obvious fashion the argument in Section 5.4.2 above without the need to condition on A_{ext}^C, so the details are omitted. Again, the proof for the epidemic on the MR random graph is similar but simpler.

5.6. *Proof of Theorem 2.7.* Consider first bond percolation and note that when $I \equiv 1$ and $\lambda = -\log(1 - \pi)$ then in the directed random graph $\tilde{\mathcal{G}}^{(n)}$ defined at the start of Section 5.1, all the possible directed edges in $\tilde{\mathcal{G}}^{(n)}$ (i.e. between pairs of neighbours in $\mathcal{G}^{(n)}$) are present independently, each with probability π. Further, when constructing the final outcome $\mathcal{T}^{(n)}$ of the corresponding epidemic $\mathcal{E}^{(n)}$ using $\tilde{\mathcal{G}}^{(n)}$, for any pair (i, j) of distinct individuals use is made of at most one of the directed edges $i \to j$ and $j \to i$ (if i infects j, whether or not j tries to infect i is immaterial). Thus, in this situation $\tilde{\mathcal{G}}^{(n)}$ can be replaced by $\mathcal{G}_{\mathrm{bond}}^{(n)}$, obtained using bond percolation on $\mathcal{G}^{(n)}$, and in $\mathcal{E}^{(n)}$, an initial susceptible is ultimately infected if and only if in $\mathcal{G}_{\mathrm{bond}}^{(n)}$ there is a chain of edges connecting it to an initial infective.

Suppose that there is one initial infective. Then the final size of the epidemic (including the initial infective) is given by the size of the connected component of $\mathcal{G}_{\mathrm{bond}}^{(n)}$ that contains the initial infective. With probability tending to 1 as $n \to \infty$, a major outbreak occurs in $\mathcal{E}^{(n)}$ if and only if the initial infective belongs to the largest connected component of $\mathcal{G}_{\mathrm{bond}}^{(n)}$, and the final size of a major epidemic is given by the size $C^{(n)}$ of that connected component. Further $C^{(n)}$ has the same asymptotic distribution as $T^{(n)}|G^{(n)}$. Thus (2.11) and (2.12) follow immediately on setting $I \equiv 1$ and

$\lambda = -\log(1-\pi)$ in Theorem 2.2. (Note that $p_I = \pi$ and, since I is constant, $q_I^{(2)} = q_I^2$.)

Turning to site percolation, consider the epidemic $\mathcal{E}^{(n)}$ with $\mathrm{P}(I = \infty) = \pi = 1 - \mathrm{P}(I = 0)$, so each infective infects all of its neighbours with probability π and none of them otherwise. Suppose that there is one initial infective, i^* say, and let $\mathcal{T}^{(n)} = \{j \in \mathcal{N}^{(n)} \setminus \{i^*\} : i^* \rightsquigarrow j\}$ using the directed random graph $\tilde{\mathcal{G}}^{(n)}$. Then $\mathcal{T}^{(n)} \cup \{i^*\}$ differs from the connected component containing i^* in $\mathcal{G}_{\mathrm{site}}^{(n)}$ (site percolation on $\mathcal{G}^{(n)}$) in that $\mathcal{T}^{(n)}$ also includes infected individuals having $I = 0$, which are deleted in $\mathcal{G}_{\mathrm{site}}^{(n)}$. Thus, to obtain a central limit theorem for $C^{(n)}$, we need one for $V^{(n)} = |\{j \in \mathcal{N}^{(n)} \setminus \{i^*\} : I_j = \infty \text{ and } i^* \rightsquigarrow j\}|$, the final size of $\mathcal{E}^{(n)}$ counting only individuals with $I = \infty$. This can be obtained by augmenting the process $\{\boldsymbol{W}^{(n)}(t)\}$ as we now outline.

Let $\{\hat{\boldsymbol{W}}^{(n)}(t)\} = \{\hat{\boldsymbol{W}}^{(n)}(t) : t \geq 0\}$, where

$$\hat{\boldsymbol{W}}^{(n)}(t) = (\boldsymbol{W}^{(n)}(t), V^{(n)}(t)) = (\boldsymbol{X}^{(n)}(t), Y_E^{(n)}(t), Z_E^{(n)}(t), V^{(n)}(t))$$

and $V^{(n)}(t)$ is the total number of initial susceptibles that are infected during $(0, t]$ and have $I = \infty$ (so $V^{(n)}(0) = 0$). A typical element of $\hat{H}^{(n)}$, the state space of $\{\hat{\boldsymbol{W}}^{(n)}(t)\}$, is now $\hat{\boldsymbol{n}} = (n_0^X, n_1^X, \ldots, n_{d_{\max}}^X, n_E^Y, n_E^Z, n^V)$. The transition intensities of $\{\hat{\boldsymbol{W}}^{(n)}(t)\}$ are essentially the same of those of $\{\boldsymbol{W}^{(n)}(t)\}$, given at the end of Section 5.1, except now transitions of type (i) are partitioned according to whether or not the infected susceptible has $I = \infty$. For $i = 1, 2, \ldots, d_{\max}$, we have

$$\hat{q}^{(n)}(\hat{\boldsymbol{n}}, \hat{\boldsymbol{n}} - \boldsymbol{e}_i^{\mathrm{S}} - \boldsymbol{e}^{\mathrm{I}} + (i-1)\boldsymbol{e}^{\mathrm{R}}) = \frac{n_E^Y i n_i^X p_{i,k}(1-\pi)}{n_E^X + n_E^Y + n_E^Z - 1},$$

corresponding to the degree-i infected susceptible having $I = 0$, and

$$\hat{q}^{(n)}(\hat{\boldsymbol{n}}, \hat{\boldsymbol{n}} - \boldsymbol{e}_i^{\mathrm{S}} + (i-1)\boldsymbol{e}^{\mathrm{I}} + \boldsymbol{e}^{\mathrm{V}}) = \frac{n_E^Y i n_i^X p_{i,k}\pi}{n_E^X + n_E^Y + n_E^Z - 1},$$

corresponding to the degree-i infected susceptible having $I = \infty$.

We apply the same random time-scale transformation to $\{\hat{\boldsymbol{W}}^{(n)}(t)\}$ as done to $\{\boldsymbol{W}^{(n)}(t)\}$ in Section 5.2. Denote the time-transformed process by $\{\bar{\boldsymbol{W}}^{(n)}(t)\}$. Let $\hat{\boldsymbol{l}}_{i,0}^{(1)} = \boldsymbol{l}_{i,0}^{(1)}$ and $\hat{\boldsymbol{l}}_{i,i-1}^{(1)} = \boldsymbol{l}_{i,i-1}^{(1)} + \boldsymbol{e}^{\mathrm{V}}$ $(i = 1, 2, \ldots, d_{\max})$, $\hat{\boldsymbol{l}}_{+}^{(2)} = \boldsymbol{l}_{+}^{(2)}$, $\hat{\boldsymbol{l}}_{-}^{(2)} = \boldsymbol{l}_{-}^{(2)}$ and $\hat{\boldsymbol{l}}^{(3)} = \boldsymbol{l}^{(3)}$. The set of possible jumps of $\{\bar{\boldsymbol{W}}^{(n)}(t)\}$ from a typical state $\hat{\boldsymbol{n}}$ is now $\hat{\Delta} = \hat{\Delta}_1 \cup \hat{\Delta}_2 \cup \hat{\Delta}_3$, where $\hat{\Delta}_1 = \bigcup_{i=1}^{d_{\max}} \{\hat{\boldsymbol{l}}_{i,0}^{(1)}\} \cup \{\hat{\boldsymbol{l}}_{i,i-1}^{(1)}\}$, $\hat{\Delta}_2 = \{\hat{\boldsymbol{l}}_{+}^{(2)}, \hat{\boldsymbol{l}}_{-}^{(2)}\}$ and $\hat{\Delta}_3 = \{\hat{\boldsymbol{l}}^{(3)}\}$.

Let $\hat{\boldsymbol{w}} = (\boldsymbol{x}, y_E, z_E, v)$. The family of processes $\{\bar{\boldsymbol{W}}^{(n)}(t)\}$ is again asymptotically density dependent with corresponding functions $\bar{\beta}_{\boldsymbol{l}}(\hat{\boldsymbol{w}})$ ($\boldsymbol{l} \in \hat{\Delta}$) given by (cf. (5.3))

$$\bar{\beta}_{\boldsymbol{l}}(\hat{\boldsymbol{w}}) = \begin{cases} \bar{\beta}^{(1)}_{i,0}(\boldsymbol{x}, y_E, z_E, v) = i x_i (1-\pi) & \text{for } \boldsymbol{l} = \hat{\boldsymbol{l}}^{(1)}_{i,0} \in \hat{\Delta}_1, \\ \bar{\beta}^{(1)}_{i,i-1}(\boldsymbol{x}, y_E, z_E, v) = i x_i \pi & \text{for } \boldsymbol{l} = \hat{\boldsymbol{l}}^{(1)}_{i,i-1} \in \hat{\Delta}_1, \\ \bar{\beta}^{(2)}_{+}(\boldsymbol{x}, y_E, z_E, v) = y_E 1_{\{y_E>0\}} & \text{for } \boldsymbol{l} = \hat{\boldsymbol{l}}^{(2)}_{+}, \\ \bar{\beta}^{(2)}_{-}(\boldsymbol{x}, y_E, z_E, v) = -y_E 1_{\{y_E<0\}} & \text{for } \boldsymbol{l} = \hat{\boldsymbol{l}}^{(2)}_{-}, \\ \bar{\beta}^{(3)}(\boldsymbol{x}, y_E, z_E, v) = z_E 1_{\{z_E>0\}} & \text{for } \boldsymbol{l} = \hat{\boldsymbol{l}}^{(3)}. \end{cases} \tag{5.54}$$

The associated drift function is (cf. (5.6))

$$\begin{aligned} \bar{F}(\hat{\boldsymbol{w}}) = &- \sum_{i=1}^{d_{\max}} i x_i \boldsymbol{e}^{\mathrm{S}}_i + \left\{ \sum_{i=1}^{d_{\max}} i x_i [(i-1)\pi - 1] - 2y_E - z_E \right\} \boldsymbol{e}^{\mathrm{I}} \\ &+ \left[\sum_{i=1}^{d_{\max}} i(i-1)(1-\pi) x_i - z_E \right] \boldsymbol{e}^{\mathrm{R}} + \left[\sum_{i=1}^{d_{\max}} i x_i \pi \right] \boldsymbol{e}^{\mathrm{V}}. \end{aligned}$$

For $t \geq 0$, let $\bar{\boldsymbol{w}}(t) = (\bar{x}_0(t), \bar{x}_1(t), \ldots, \bar{x}_{d_{\max}}(t), \bar{y}_E(t), \bar{z}_E(t), \bar{v}(t))$ be defined analogously to $\tilde{\boldsymbol{w}}(t)$ at (5.7). Noting that $p_I = \pi$ and $q_I = 1 - \pi$, the corresponding deterministic model for $\bar{\boldsymbol{w}}(t)$ is given by (5.8)-(5.10), with $\tilde{x}_i$ replaced by $\bar{x}_i$ etc., augmented with

$$\frac{d\bar{v}}{dt} = \pi \sum_{i=1}^{d_{\max}} i \bar{x}_i. \tag{5.55}$$

Thus, (5.11)-(5.13) still hold, with the above change of notation, and, using (5.11),

$$\bar{v}(t) = \bar{v}(0) + \pi \sum_{i=1}^{d_{\max}} \bar{x}_i(0)(1 - \mathrm{e}^{-it}). \tag{5.56}$$

The stopping time $\tilde{\tau}^{(n)}$ is unchanged, except now $\varphi(\hat{\boldsymbol{w}}) = \varphi(\boldsymbol{x}, y_E, z_E, v) = y_E$. Recall that site percolation corresponds to one initial infective, so under both the MR and NSW models, $\epsilon_i = 0$ and $\tilde{x}_i(0) = p_i$ ($i = 0, 1, \ldots, d_{\max}$). Thus $\tilde{\tau} = \inf\{t \geq 0 : \bar{y}_E(t) = 0\}$ ($= \inf\{t \geq 0 : \tilde{y}_E(t) = 0\}$) is given by the unique solution of (5.49) in $(0, \infty)$. Hence $z = \mathrm{e}^{-\tilde{\tau}}$ is given by the unique solution in $(0, 1)$ of (2.10). Now $\tilde{v}(0) = 0$, since $V^{(n)}(0) = 0$ for all $n = 1, 2, \ldots$, so using (5.56) and recalling that $\rho = 1 - f_D(z)$,

$$\bar{v}(\tilde{\tau}) = \pi \left[1 - f_D(\mathrm{e}^{-\tilde{\tau}}) \right] = \pi\rho.$$

Let $\bar{V}^{(n)} = \bar{V}^{(n)}(\tilde{\tau}^{(n)})$, so in the site percolation setting, $C^{(n)} \stackrel{D}{=} 1 + \bar{V}^{(n)}|G^{(n)}$. Now application of Corollary 4.1, as at (5.28), yields

$$\sqrt{n}\left(n^{-1}\bar{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)}) - \bar{\boldsymbol{w}}(\tilde{\tau})\right) \xrightarrow{D} \mathrm{N}(\mathbf{0}, \bar{B}\bar{\Sigma}(\tilde{\tau})\bar{B}^{\top}) \qquad \text{as } n \to \infty,$$

where

$$\bar{B} = I - \frac{\left(\bar{F}(\bar{\boldsymbol{w}}(\tilde{\tau}))\right)^{\top} \nabla\varphi(\bar{\boldsymbol{w}}(\tilde{\tau}))}{\nabla\varphi(\bar{\boldsymbol{w}}(\tilde{\tau})) \cdot \bar{F}(\bar{\boldsymbol{w}}(\tilde{\tau}))}$$

and $\bar{\Sigma}(\tilde{\tau})$ is obtained by making obvious changes to (5.25)-(5.27) to account for the extra dimension. Further, $\bar{V}^{(n)} = \bar{\boldsymbol{W}}^{(n)}(\tilde{\tau}^{(n)})(\mathbf{0}, 0, 0, 1)^{\top}$, so

$$\sqrt{n}\left(n^{-1}\bar{V}^{(n)} - \pi\rho\right) \xrightarrow{D} \mathrm{N}(0, \bar{\sigma}^2) \qquad \text{as } n \to \infty,$$

where

$$\hat{\sigma}^2 = (\mathbf{0}, 0, 0, 1)\bar{B}\bar{\Sigma}(\tilde{\tau})\bar{B}^{\top}(\mathbf{0}, 0, 0, 1)^{\top}.$$

A simple calculation yields $(\mathbf{0}, 0, 0, 1)\bar{B} = (\mathbf{0}, -\pi b(\tilde{\tau}), 0, 1)$, where $b(\tilde{\tau})$ is given by (5.30) with $a(\tilde{\tau})$ obtained by replacing f_{D_ϵ} by f_D in (5.24). Define $\bar{\Phi}(t, u)$ analogously to $\tilde{\Phi}(t, u)$ at (5.25). Then it follows using (5.11)-(5.13) and (5.56) that

$$\left[(\mathbf{0}, 0, 0, 1)\bar{\Phi}(\tilde{\tau}, u)\right]_i = \begin{cases} \pi\left(1 - re^{-i(\tilde{\tau}-u)}\right) & \text{if } i = 0, 1, \ldots, d_{\max}, \\ 0 & \text{if } i = d_{\max}+1, d_{\max}+2, \\ 1 & \text{if } i = d_{\max}+3, \end{cases}$$

and $\left[(\mathbf{0}, 1, 0, 0)\bar{\Phi}(\tilde{\tau}, u)\right]_i$ is 0, if $i = d_{\max}+3$, and given by the right-hand side of (5.31) (with $p_I = \pi$ and $q_I = 1 - \pi$, if $i = 0, 1, \ldots, d_{\max}+2$.

It follows that $\bar{\boldsymbol{c}}(\tilde{\tau}, u) = (\mathbf{0}, 0, 0, 1)\bar{B}\bar{\Phi}(\tilde{\tau}, u)$ is given by

$$\bar{\boldsymbol{c}}(\tilde{\tau}, u) = -\pi(\boldsymbol{c}(\tilde{\tau}, u), 0) + \pi(\mathbf{1}, 0, 0, 0) + (\mathbf{0}, 0, 0, 1), \tag{5.57}$$

where $\boldsymbol{c}(\tilde{\tau}, u)$ is defined at (5.32), and (cf. (5.37))

$$\bar{\sigma}^2 = \bar{\boldsymbol{c}}(\tilde{\tau}, 0)\bar{\Sigma}_0\bar{\boldsymbol{c}}(\tilde{\tau}, 0)^{\top} + \sum_{\hat{\boldsymbol{l}} \in \hat{\Delta}} \int_0^{\tilde{\tau}} \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u)\hat{\boldsymbol{l}}^{\top}\right)^2 \bar{\beta}_{\boldsymbol{l}}(\bar{\boldsymbol{w}}(u))\, \mathrm{d}u,$$

where $\bar{\Sigma}_0 = \bar{\Sigma}_0^{\mathrm{MR}}$ or $\bar{\Sigma}_0^{\mathrm{NSW}}$, depending on whether the random graph is MR or NSW. Here, $\bar{\Sigma}_0^{\mathrm{MR}}$ and $\bar{\Sigma}_0^{\mathrm{NSW}}$ are the asymptotic variance matrices of $n^{-\frac{1}{2}}\bar{\boldsymbol{W}}^{(n)}(0)$ for the MR and NSW random graphs, respectively; cf. (5.16) and (5.20).

We show in Appendix B.5 that

$$\sum_{\hat{\boldsymbol{l}}\in\hat{\Delta}}\int_0^{\tilde{\tau}}\left(\bar{\boldsymbol{c}}(\tilde{\tau},u)\hat{\boldsymbol{l}}^\top\right)^2\bar{\beta}_{\boldsymbol{l}}(\bar{\boldsymbol{w}}(u))\,\mathrm{d}u=\pi(1-\pi)[\rho-2h(z)(1-z)\mu_D]+\pi^2\sum_{\boldsymbol{l}\in\Delta}\int_0^{\tilde{\tau}}\left(\boldsymbol{c}(\tilde{\tau},u)\boldsymbol{l}^\top\right)^2\tilde{\beta}_{\boldsymbol{l}}(\tilde{\boldsymbol{w}}(u))\,\mathrm{d}u \tag{5.58}$$

and, for both the MR and NSW random graphs, that

$$\bar{\boldsymbol{c}}(\tilde{\tau},0)\bar{\Sigma}_0\bar{\boldsymbol{c}}(\tilde{\tau},0)^\top=\boldsymbol{c}(\tilde{\tau},0)\Sigma_0\boldsymbol{c}(\tilde{\tau},0)^\top. \tag{5.59}$$

Equations (2.13) and (2.14) then follow using Theorem 2.2, noting that $q_I^{(2)}=\pi(1-\pi)$.

5.7. *Simple graphs.* As noted in Section 2.1, the graph $\mathcal{G}^{(n)}$ constructed using the configuration model may contain self-loops and parallel edges. We now apply recent results of Janson (2019) to show that our central limit theorems hold also for the graph $\mathcal{G}_S^{(n)}$, which is distributed as $\mathcal{G}^{(n)}$ conditioned on being simple. We do this for $T_{\mathrm{MR}}^{(n)}$ in Theorem 2.3. Similar arguments hold for the other central limit theorems.

The method of Janson (2019) starts with a realisation of $\mathcal{G}^{(n)}$ and adjusts it by a sequence of switchings, which maintain the degrees of its vertices, until a simple graph, which we denote by $\hat{\mathcal{G}}_S^{(n)}$, is obtained. An edge is called *bad* if it is either a self-loop or a parallel edge. If the realisation of $\mathcal{G}^{(n)}$ contains no bad edge then $\hat{\mathcal{G}}_S^{(n)}=\mathcal{G}^{(n)}$. Otherwise a bad edge is chosen and its endpoints are switched with those of an edge chosen uniformly at random from all edges (see Janson (2019), Section 3.2 for details). The switching process is continued until a simple graph is obtained. Recall that $T_{\mathrm{MR}}^{(n)}$ is the final size of an epidemic on $\mathcal{G}^{(n)}$. Let $\hat{T}_{\mathrm{MR}}^{(n)}$ denote the final size of an epidemic on the graph $\hat{\mathcal{G}}_S^{(n)}$ and $T_{\mathrm{MR,S}}^{(n)}$ denote the final size of an epidemic on the simple graph $\mathcal{G}_S^{(n)}$. Let $S^{(n)}$ be the number of switchings required to obtain a simple graph. Using Janson (2019), Theorem 3.2, $\lim_{K\to\infty}\sup_n \mathrm{P}(S^{(n)}>K)=0$. The distributions of $\hat{\mathcal{G}}_S^{(n)}$ and $\mathcal{G}_S^{(n)}$ differ, though they are equal asymptotically (Janson (2019), Theorem 2.1). Moreover, it follows using Janson (2019), Corollary 2.3 (see also Remark 8.2), that $T_{\mathrm{MR,S}}^{(n)}$ satisfies the central limit theorem in Theorem 2.3 provided

$$\frac{1}{\sqrt{n}}\left(\hat{T}_{\mathrm{MR}}^{(n)}-T_{\mathrm{MR}}^{(n)}\right)\xrightarrow{p}0\qquad\text{as } n\to\infty. \tag{5.60}$$

To prove (5.60), fix $\delta, \epsilon > 0$. There exists integer $K > 0$ such that, for all n, $\mathrm{P}(S^{(n)} > K) < \frac{\delta}{2}$, so we restrict attention to switching processes with $S^{(n)} \leq K$. Let $\mathcal{N}_S^{(n)}$ denote the set of vertices in $\hat{\mathcal{G}}_S^{(n)}$ that have at least one edge that is switched in its construction. Let $\Gamma_S^{(n)}$ be a vector which contains for each vertex in $\mathcal{N}_S^{(n)}$ its degree and the number of edges connected to vertices not in $\mathcal{N}_S^{(n)}$. Note that $\Gamma_S^{(n)}$ remains unchanged during the switching process. Let $\hat{T}_{\mathrm{MR},-}^{(n)}$ denote the final size of an epidemic on $\hat{\mathcal{G}}_S^{(n)}$ in which all the vertices in $\mathcal{N}_S^{(n)}$ are initially recovered; apart from this restriction the initial conditions are the same as those for $T_{\mathrm{MR}}^{(n)}$ and $\hat{T}_{\mathrm{MR}}^{(n)}$. We construct a realisation of $(\hat{\mathcal{G}}_S^{(n)}, \hat{T}_{\mathrm{MR},-}^{(n)})|(\tilde{\mathcal{N}}_S^{(n)}, \Gamma_S^{(n)})$, where $\tilde{\mathcal{N}}_S^{(n)}$ denotes the vertices $\mathcal{N}_S^{(n)}$ together with edges joining them in $\hat{\mathcal{G}}_S^{(n)}$ and edges joining them in $\mathcal{G}^{(n)}$, as follows. Break all the edges except those between individuals in $\mathcal{N}_S^{(n)}$ into half-edges and make all half-edges emanating from individuals in $\mathcal{N}_S^{(n)}$ recovered half-edges. Now use the construction in Section 5.1 to yield a continuous-time Markov chain $\{\check{\boldsymbol{W}}^{(n)}(t)\}$, where $\check{\boldsymbol{W}}^{(n)}(t) = (\check{X}_0^{(n)}(t), \check{X}_1^{(n)}(t), \ldots, \check{X}_{d_{\max}}^{(n)}(t), \check{Y}_E^{(n)}(t), \check{Z}_E^{(n)}(t))$, that is defined analogously to $\{\boldsymbol{W}^{(n)}(t)\}$. The process $\{\check{\boldsymbol{W}}^{(n)}(t)\}$ does not contain the endpoints of edges as they are formed in the corresponding realisation of $\hat{\mathcal{G}}_S^{(n)}$, nor the disease status of individuals, but it can be augmented to carry that information. Let $\check{\tau}^{(n)} = \inf\{t \geq 0 : \check{Y}_E^{(n)}(t) = 0\}$. By applying Theorem 4.1 to the random time-changed version of $\{\check{\boldsymbol{W}}^{(n)}(t)\}$ that is analogous to $\{\tilde{\boldsymbol{W}}^{(n)}(t)\}$, since $S^{(n)} \leq K$ and the degrees are bounded, it is shown easily that, as $n \to \infty$,

(5.61)
$$\frac{1}{n}\check{X}_i^{(n)}(\check{\tau}^{(n)}) \xrightarrow{p} \tilde{x}_i(\tilde{\tau}) \quad (i = 0, 1, \ldots, d_{\max}) \quad \text{and} \quad \frac{1}{n}\check{Z}_E^{(n)}(\check{\tau}^{(n)}) \xrightarrow{p} \tilde{z}_E(\tilde{\tau}),$$

where $\tilde{\tau} > 0$ satisfies (5.23), with $\epsilon = 0$, and $\tilde{x}_i(t)$ and $\tilde{z}_E(t)$ are given by (5.11) and (5.13) with $\tilde{x}_i(0) = p_i$ and $\tilde{z}_E(0) = 0$.

At this stage $\hat{\mathcal{G}}_S^{(n)}$ is partially constructed. To obtain a realisation of $\hat{T}_{\mathrm{MR}}^{(n)}|(\mathcal{N}_S^{(n)}, \Gamma_S^{(n)})$, we now change the disease status of individuals in $\mathcal{N}_S^{(n)}$ to susceptible or infective, according to their initial status in the epidemic on $\mathcal{G}^{(n)}$. Also the disease status of any individual in $\mathcal{N}_S^{(n)}$ that had at least one half-edge paired (and hence received infection) by time $\check{\tau}^{(n)}$ in the augmented version of $\{\boldsymbol{W}^{(n)}(t)\}$ is changed to infective. If no member of $\mathcal{N}_S^{(n)}$ becomes infected then $\hat{T}_{\mathrm{MR}}^{(n)} = \hat{T}_{\mathrm{MR},-}^{(n)}$. Otherwise the epidemic is now spread within $\mathcal{N}_S^{(n)}$, which may lead to infectives within $\mathcal{N}_S^{(n)}$ who have at

least one (unpaired) half-edge, and we carry on the joint construction of $(\hat{\mathcal{G}}_S^{(n)}, \hat{T}_{\mathrm{MR},-}^{(n)})|(\tilde{\mathcal{N}}_S^{(n)}, \Gamma_S^{(n)})$. The construction can also be continued from time $\check{\tau}^{(n)}$ in a similar way to yield a realisation of $(\mathcal{G}^{(n)}, T_{\mathrm{MR}}^{(n)})|(\tilde{\mathcal{N}}_S^{(n)}, \Gamma_S^{(n)})$, the only difference being the spread of infection within $\mathcal{N}_S^{(n)}$ to decide which half-edges emanating from individuals in $\mathcal{N}_S^{(n)}$ become infective half-edges. Let $\hat{T}_{\mathrm{MR},+}^{(n)}$ denote the total size of a similar epidemic, $\mathcal{E}_+^{(n)}$ say, in which at time $\check{\tau}^{(n)}$ *all* individuals in $\mathcal{N}_S^{(n)}$ are made infective. Clearly, realisations of $\hat{T}_{\mathrm{MR}}^{(n)}$ and $T_{\mathrm{MR}}^{(n)}$ can be derived from a realisation of $\mathcal{E}_+^{(n)}$ so that $\hat{T}_{\mathrm{MR},-}^{(n)} \le \hat{T}_{\mathrm{MR}}^{(n)} \le \hat{T}_{\mathrm{MR},+}^{(n)}$ and $\hat{T}_{\mathrm{MR},-}^{(n)} \le T_{\mathrm{MR}}^{(n)} \le \hat{T}_{\mathrm{MR},+}^{(n)}$, whence $|\hat{T}_{\mathrm{MR}}^{(n)} - T_{\mathrm{MR}}^{(n)}| \le \hat{T}_{\mathrm{MR},+}^{(n)} - \hat{T}_{\mathrm{MR},-}^{(n)}$.

To construct $\mathcal{E}_+^{(n)}$, we carry on the exploration process from infective half-edges (if any) emanating from $\mathcal{N}_S^{(n)}$ in turn, only pairing half-edges along which an infective transmits infection After the infection process has finished, we continue by pairing all remaining half-edges to make the graph. If the construction from time $\check{\tau}^{(n)}$ creates a new edge inside $\mathcal{N}_S^{(n)}$, we abort and restart the entire construction from time 0; this happens with probability $o(1)$ as $n \to \infty$, so it may be ignored.

Let i_* denote the first infective half-edge emanating from $\mathcal{N}_S^{(n)}$ that is explored. When infection is transmitted along it, the probability it is paired with a half-edge from a degree-i susceptible is

$$\frac{i\check{X}_i^{(n)}(\check{\tau}^{(n)})}{\check{Z}_E^{(n)}(\check{\tau}^{(n)}) - 1 + \sum_{j=1}^{d_{\max}} i\check{X}_j^{(n)}(\check{\tau}^{(n)})} \xrightarrow{p} \frac{i\tilde{x}_i(\tilde{\tau})}{\tilde{x}_E(\tilde{\tau}) + \tilde{z}_E(\tilde{\tau})} \qquad \text{as } n \to \infty,$$

where $\tilde{x}_E(\tilde{\tau}) = \sum_{i=1}^{d_{\max}} i\tilde{x}_i(\tilde{\tau})$. Let

$$R_0(\tilde{\tau}) = p_I \frac{\sum_{i=1}^{d_{\max}} i(i-1)\tilde{x}_i(\tilde{\tau})}{\tilde{x}_E(\tilde{\tau}) + \tilde{z}_E(\tilde{\tau})} \tag{5.62}$$

be the asymptotic mean number of infections made by the individual owning the half-edge with which i_* is paired. Then $R_0(\tilde{\tau}) < 1$ (see Appendix B.3), as is intuitively plausible since otherwise the epidemic would not have stopped at time $\check{\tau}^{(n)}$. Thus, using (5.61), there exists $\epsilon' > 0$ such that

$$\lim_{n\to\infty} \mathrm{P}\left(\check{Z}_E^{(n)}(\check{\tau}^{(n)}) > n(\tilde{z}_E(\tilde{\tau}) - \epsilon')\right) = 0, \tag{5.63}$$

$$\lim_{n\to\infty} \mathrm{P}\left(\check{X}_i^{(n)}(\check{\tau}^{(n)}) < n(\tilde{x}_i(\tilde{\tau}) + \epsilon')\right) = 0 \quad (i = 1, 2, \ldots, d_{\max}), \tag{5.64}$$

and

$$R_0' = p_I \frac{\sum_{i=1}^{d_{\max}} i(i-1)(\tilde{x}_i(\tilde{\tau}) + \epsilon')}{\tilde{z}_E(\tilde{\tau}) - \epsilon' + \sum_{i=1}^{d_{\max}} i(\tilde{x}_i(\tilde{\tau}) + \epsilon')} < 1.$$

It follows that, with probability tending to 1 as $n \to \infty$, the total number of infectives in the exploration process from the half-edge i_* may be bounded above by the total progeny, Z' say, of a (subcritical) Galton-Watson process having offspring mean R_0'. (Since $\mathrm{P}(Z' < \infty) = 1$, (5.61) ensures that obvious modifications of (5.63) and (5.64) hold throughout the exploration process from i_*.)

Now $\mathrm{E}[Z'] = (1 - R_0')^{-1} < \infty$ and there can be at most $L = 4Kd_{\max}$ infective half-edges emanating from $\mathcal{N}_S^{(n)}$, since $S^{(n)} \leq K$. Thus a simple argument involving exchangeability and Markov's inequality yields

$$\lim_{n\to\infty} \mathrm{P}\left(\hat{T}^{(n)}_{\mathrm{MR},+} - \hat{T}^{(n)}_{\mathrm{MR},-} > \epsilon\sqrt{n} | (\tilde{\mathcal{N}}_S^{(n)}, \Gamma_S^{(n)})\right) \leq \lim_{n\to\infty} \frac{L(1-R_0')^{-1}}{\epsilon\sqrt{n}} = 0,$$

so, since $|\hat{T}^{(n)}_{\mathrm{MR}} - T^{(n)}_{\mathrm{MR}}| \leq \hat{T}^{(n)}_{\mathrm{MR},+} - \hat{T}^{(n)}_{\mathrm{MR},-}$,

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{1}{\sqrt{n}}\left|\hat{T}^{(n)}_{\mathrm{MR}} - T^{(n)}_{\mathrm{MR}}\right| > \epsilon | (\tilde{\mathcal{N}}_S^{(n)}, \Gamma_S^{(n)})\right) = 0. \tag{5.65}$$

The above construction may yield a graph that contains self-loops and/or parallel edges. However, by Janson (2009b), (5.61) and hence also (5.65) still holds if the graph is conditioned on there being no such imperfections. Thus (5.60) follows as $\mathrm{P}(S^{(n)} > K) < \frac{\delta}{2}$ and $\delta, \epsilon > 0$ are arbitrary.

6. Concluding comments. A shortcoming of our results, from a mathematical though not a practical viewpoint, is the requirement of a maximal degree $d_{\max}$. It seems likely that Theorems 2.1-2.3 continue to hold when that requirement is relaxed, subject to appropriate conditions on the degree sequence (MR model) or degree distribution D (NSW model). This conjecture is supported by the numerical illustrations in Section 3, and by the recent work of Barbour and Röllin (2019) and Janson (2018), who inter alia prove central limit theorems for the size of the giant component for configuration model graphs, with asymptotic variances consistent with setting $p_I = 1$ in Theorem 2.2. To extend the present proof to models with unbounded degree would require a functional central limit theorem for density dependent population processes with countable state spaces. Barbour and Luczak (2012) give such a theorem but it is not applicable in our setting as it would require a finite upper bound on the number of neighbours an individual can infect.

Note that (5.28) yields a multivariate central limit theorem for the numbers of susceptibles of different degrees remaining at the end of an epidemic, although we do not derive a closed-form expression for the asymptotic variance matrix. Setting $p_I = 1$, as in Remark 2.6, enables a multivariate central

limit theorem to be obtained for the number of vertices of different degrees in the giant component of an MR random graph, and hence also for the numbers of vertices and edges in the giant component. In particular, the asymptotic variance, σ_E^2 say, of the number of edges in the giant component admits a similar form to (5.37), with $\Sigma_0 = 0$, and it is immediately apparent that σ_E^2 is strictly positive; cf. Janson (2018), Remark 10.6.

The central limit theorems can be extended, at least in principle, to allow for the infection rate λ to depend on the degree of an infective, and also to more general infection processes in which the set of its neighbours that are contacted by a given infective is a symmetric sampling procedure (Martin-Löf (1986)). In both cases it is straightforward to determine the limiting deterministic model in Section 5.3.1 and the equation corresponding to (5.23), which governs $\tilde{\tau}$, but calculation of the asymptotic variances in the central limit theorems is likely to be prohibitive.

The configuration model does not display clustering in the limit as $n \to \infty$ and several authors have considered modifications of the configuration model that introduce clustering. In Trapman (2007) and Coupechoux and Lelarge (2014), in the configuration model construction, for $d = 1, 2, \ldots,$ some individuals having d half-edges are replaced by fully connected cliques, each of size d, with each member of a clique having exactly one half-edge. The half-edges are then paired up in the usual fashion. In Gleeson (2009) and Ball et al. (2010), the network is formed as in the configuration model and the population is also partitioned into fully connected cliques. In both models, the set of edges in the network is the union of those in cliques and the paired half-edges. The methodology in this paper can be extended to this general class of models as follows.

As in Section 5.1, the network and epidemic are constructed simultaneously. The objects counted are now fully susceptible cliques, typed by their size and degree composition, and infective and recovered half-edges. When infection is transmitted down a half-edge that half-edge is paired with a uniformly chosen half-edge as before. If it is paired with a susceptible half-edge, then an epidemic is triggered within the corresponding clique and associated half-edges, leading to the creation of further infective and recovered half-edges and, unless the clique epidemic infects the entire clique, a new susceptible clique having reduced size. Central limit theorems for the final size of epidemics on MR and NSW versions of such random graphs should follow using similar arguments to before but again calculation of the asymptotic variances may be difficult. If the infectious period is constant, so the epidemic model is equivalent to bond percolation on the network, the analysis may perhaps be simplified by first splitting the cliques into com-

ponents determined by bond percolation and then using similar methods to the present paper treating the components as super-individuals.

APPENDIX A: PROOF OF THEOREMS 4.1 AND 4.2

In this appendix we prove Theorems 4.1 and 4.2. As noted in Section 4, the proofs follow closely those of similar theorems in Britton and Pardoux (2019), the main difference being the use of Skorohod's theorem in the proof of Theorem 4.2. As we are primarily interested in Theorem 4.2, we prove Theorem 4.1 under the stronger initial condition (4.8) (repeated in (A.3) below) and indicate how the proof may be extended easily to the corresponding convergence in probability in Theorem 4.1, as stated, and to the corresponding strong law of large numbers in Ethier and Kurtz (1986) and Britton and Pardoux (2019). In this appendix, vectors are column vectors, to aid connection with Ethier and Kurtz (1986) and Britton and Pardoux (2019), and notation is local to it and Section 4.

The process $\{\boldsymbol{X}^{(n)}(t) : t \geq 0\}$ can be expressed as

$$\boldsymbol{X}^{(n)}(t) = \boldsymbol{X}^{(n)}(0) + \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l} Y_{\boldsymbol{l}} \left(\int_0^t n \beta_{\boldsymbol{l}}^{(n)} \left(n^{-1} \boldsymbol{X}^{(n)}(s) \right) \, \mathrm{d}s \right) \quad (t \geq 0), \tag{A.1}$$

where $\{Y_{\boldsymbol{l}}(t) : t \geq 0\}$ $(\boldsymbol{l} \in \Delta)$ are independent unit-rate Poisson processes; see Ethier and Kurtz (1986), Chapter 11, equation (2.1). For $\boldsymbol{l} \in \Delta$, let $\tilde{Y}_{\boldsymbol{l}}(t) = Y_{\boldsymbol{l}}(t) - t$ $(t \geq 0)$. Then

$$\left\{ \frac{1}{\sqrt{n}} \tilde{Y}_{\boldsymbol{l}}(nt) : t \geq 0 \right\} \Rightarrow \{W_{\boldsymbol{l}}(t) : t \geq 0\} \qquad \text{as } n \to \infty, \tag{A.2}$$

where $\{W_{\boldsymbol{l}}(t) : t \geq 0\}$ is a standard Brownian motion (starting at 0); see, for example, Britton and Pardoux (2019), Lemma 2.3.4. Recall (4.8), i. e.

$$\sqrt{n} \left(n^{-1} \boldsymbol{X}^{(n)}(0) - \boldsymbol{x}_0 \right) \xrightarrow{D} \boldsymbol{V}(0) \qquad \text{as } n \to \infty, \tag{A.3}$$

where $\boldsymbol{V}(0) \sim N(\boldsymbol{0}, \Sigma_0)$.

By Skorohod's theorem (see Ethier and Kurtz (1986), page 102), there exists a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, on which are defined the following random quantities:

(i) for each $n = 1, 2, \ldots$, $\{\check{Y}_{\boldsymbol{l}}^{(n)}(t) : t \geq 0\}$ $(\boldsymbol{l} \in \Delta)$ and $\hat{\boldsymbol{X}}^{(n)}(0)$, where $\{\check{Y}_{\boldsymbol{l}}^{(n)}(t) : t \geq 0\}$ $(\boldsymbol{l} \in \Delta) \stackrel{D}{=} \{Y_{\boldsymbol{l}}(t) : t \geq 0\}$ $(\boldsymbol{l} \in \Delta)$ and independently $\hat{\boldsymbol{X}}^{(n)}(0) \stackrel{D}{=} \boldsymbol{X}^{(n)}(0)$;

(ii) $\{\hat{W}_{\boldsymbol{l}}(t) : t \geq 0\}$ ($\boldsymbol{l} \in \Delta$), independent standard Brownian motions, and independently $\hat{\boldsymbol{V}}(0) \sim N(\boldsymbol{0}, \Sigma_0)$;

such that, for $\boldsymbol{l} \in \Delta$,

$$\left\{ \frac{1}{\sqrt{n}} \check{Y}_{\boldsymbol{l}}^{(n)}(nt) : t \geq 0 \right\} \xrightarrow{a.s.} \{\hat{W}_{\boldsymbol{l}}(t) : t \geq 0\} \qquad \text{as } n \to \infty, \tag{A.4}$$

and

$$\sqrt{n}\left(n^{-1} \hat{\boldsymbol{X}}^{(n)}(0) - \boldsymbol{x}_0 \right) \xrightarrow{a.s.} \hat{\boldsymbol{V}}(0) \qquad \text{as } n \to \infty. \tag{A.5}$$

Thus, there exists $A \in \mathcal{F}$ with $\mathrm{P}(A) = 1$, such that the convergences corresponding to (A.4) and (A.5) hold pointwise on A and, for all $\boldsymbol{l} \in \Delta$, the function $\hat{W}_{\boldsymbol{l}}(t, \omega)$ ($t \geq 0$) is continuous for all $\omega \in A$.

For $n = 1, 2, \ldots$, define the process $\{\hat{\boldsymbol{X}}^{(n)}(t) : t \geq 0\}$ by

$$\hat{\boldsymbol{X}}^{(n)}(t) = \hat{\boldsymbol{X}}^{(n)}(0) + \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l} \hat{Y}_{\boldsymbol{l}}^{(n)} \left(\int_0^t n \beta_{\boldsymbol{l}}^{(n)} \left(n^{-1} \hat{\boldsymbol{X}}^{(n)}(s) \right) \mathrm{d}s \right) \quad (t \geq 0), \tag{A.6}$$

where $\hat{Y}_{\boldsymbol{l}}^{(n)}(t) = t + \check{Y}_{\boldsymbol{l}}^{(n)}(t)$ ($t \geq 0, \boldsymbol{l} \in \Delta$). We prove almost sure analogues of Theorems 4.1 and 4.2 for the processes $\{\hat{\boldsymbol{X}}^{(n)}(t) : t \geq 0\}$ ($n = 1, 2, \ldots$) by showing that they satisfy the required convergence pointwise on A. Theorems 4.1 and 4.2 then follow since, for each $n = 1, 2, \ldots$, the process $\{\hat{\boldsymbol{X}}^{(n)}(t) : t \geq 0\}$ has the same law as $\{\boldsymbol{X}^{(n)}(t) : t \geq 0\}$.

We assume that $\omega \in A$ and suppress explicit dependence on ω, so

$$\sqrt{n}\left(n^{-1} \hat{\boldsymbol{X}}^{(n)}(0) - \boldsymbol{x}_0 \right) \to \hat{\boldsymbol{V}}(0) \qquad \text{as } n \to \infty \tag{A.7}$$

and, since convergence in the Skorohod topology is equivalent to locally uniform convergence when the limit process is continuous,

$$\lim_{n \to \infty} \sup_{0 \leq t \leq T} \left| \frac{1}{\sqrt{n}} \check{Y}_{\boldsymbol{l}}^{(n)}(nt) - \hat{W}_{\boldsymbol{l}}(t) \right| = 0 \qquad (\boldsymbol{l} \in \Delta, T \geq 0). \tag{A.8}$$

For $t \geq 0$, let $\bar{\boldsymbol{X}}^{(n)}(t) = n^{-1} \hat{\boldsymbol{X}}^{(n)}(t)$,

$$\bar{\boldsymbol{Y}}^{(n)}(t) = \frac{1}{n} \sum_{\boldsymbol{l} \in \Delta} \boldsymbol{l} \check{Y}_{\boldsymbol{l}}^{(n)} \left(\int_0^t n \beta_{\boldsymbol{l}}^{(n)} \left(\bar{\boldsymbol{X}}^{(n)}(s) \right) \mathrm{d}s \right)$$

and

$$\bar{\boldsymbol{Z}}^{(n)}(t) = \int_0^t \left[F^{(n)}(\bar{\boldsymbol{X}}^{(n)}(s)) - F(\bar{\boldsymbol{X}}^{(n)}(s)) \right] \mathrm{d}s,$$

where $F^{(n)}$ is defined analogously to F at (4.2) but with $\beta_{\boldsymbol{l}}(x)$ replaced by $\beta_{\boldsymbol{l}}^{(n)}(x)$. Then (A.6) yields

$$\bar{\boldsymbol{X}}^{(n)}(t) = \bar{\boldsymbol{X}}^{(n)}(0) + \int_0^t F(\bar{\boldsymbol{X}}^{(n)}(s))\,\mathrm{d}s + \bar{\boldsymbol{Y}}^{(n)}(t) + \bar{\boldsymbol{Z}}^{(n)}(t). \tag{A.9}$$

For $T > 0$, let $K_T = \{\boldsymbol{x}(t) : 0 \le t \le T\}$. Then K_T is a compact subset of the open set H, so there exists $\epsilon = \epsilon(T) > 0$, so that $K_{T,\epsilon} \subset H$, where $K_{T,\epsilon}$ is the (compact) set of consisting of all points in $\mathbb{R}^p$ that are at distance $\le \epsilon$ from K_T. Let $\tau_{T,\epsilon}^{(n)} = \inf\{t > 0 : \bar{\boldsymbol{X}}^{(n)}(t) \notin K_{T,\epsilon}\}$.

Fix $T > 0$. Recalling from (4.5) that $\boldsymbol{x}(t) = \boldsymbol{x}_0 + \int_0^t F(\boldsymbol{x}(s))\,\mathrm{d}s$, it follows from (A.9) and (4.4) that, for $0 \le t \le T$,

$$\begin{aligned}
&\left|\bar{\boldsymbol{X}}^{(n)}\ (t \wedge \tau_{T,\epsilon}^{(n)}) - \boldsymbol{x}(t \wedge \tau_{T,\epsilon}^{(n)})\right| \\
&\qquad \le \left|\bar{\boldsymbol{X}}^{(n)}(0) - \boldsymbol{x}_0\right| + M_{K_{T,\epsilon}} \int_0^{t \wedge \tau_{T,\epsilon}^{(n)}} \left|\bar{\boldsymbol{X}}^{(n)}(s) - \boldsymbol{x}(s)\right| \mathrm{d}s \\
&\qquad\qquad + T \sup_{\boldsymbol{x} \in K_{T,\epsilon}} \left|F^{(n)}(\boldsymbol{x}) - F(\boldsymbol{x})\right| + \left|\bar{\boldsymbol{Y}}^{(n)}(t \wedge \tau_{T,\epsilon}^{(n)})\right| \\
&\qquad \le A_n \exp(-M_{K_{T,\epsilon}} t),
\end{aligned}$$

using Gronwall's inequality (see Ethier and Kurtz (1986), page 498), where

$$A_n = \left|\bar{\boldsymbol{X}}^{(n)}(0) - \boldsymbol{x}_0\right| + T \sup_{\boldsymbol{x} \in K_{T,\epsilon}} \left|F^{(n)}(\boldsymbol{x}) - F(\boldsymbol{x})\right| + \sup_{0 \le t \le T \wedge \tau_{T,\epsilon}^{(n)}} \left|\bar{\boldsymbol{Y}}^{(n)}(t)\right|. \tag{A.10}$$

By (4.3), there exists a constant $C \in (0, \infty)$ such that, for all n,

$$\max_{\boldsymbol{l} \in \Delta} \sup_{0 \le t \le T \wedge \tau_{T,\epsilon}^{(n)}} \left|\beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\right| < C,$$

so it follows from (A.8) and the continuity of $\hat{W}_{\boldsymbol{l}}(t)$ ($\boldsymbol{l} \in \Delta$) that the third term on the right-hand side of (A.10) tends to 0 as $n \to \infty$. The first and second terms tend to 0, using (A.7) and (4.3), respectively. Thus $A_n \to 0$ as $n \to \infty$, so

$$\lim_{n \to \infty} \sup_{0 \le t \le T} \left|\bar{\boldsymbol{X}}^{(n)}(t) - \boldsymbol{x}(t)\right| = 0, \tag{A.11}$$

since $A_n \exp(-M_{K_{T,\epsilon}} T) \le \epsilon$ implies $\tau_{T,\epsilon}^{(n)} > T$, and Theorem 4.1 follows.

Note that if $n^{-1}\boldsymbol{X}^{(n)}(0) \xrightarrow{a.s.} \boldsymbol{x}_0$ as $n \to \infty$, then the above argument is easily modified to show directly from (A.1) (without invoking Skorohod's

theorem) that (4.6) holds with convergence in probability replaced by convergence almost surely, since, for all $\boldsymbol{l} \in \Delta$,

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left| n^{-1} Y_{\boldsymbol{l}}(nt) - nt \right| = 0 \qquad \text{for all } T \ge 0,$$

almost surely. The proof is then essentially that in Ethier and Kurtz (1986) and Britton and Pardoux (2019). If $n^{-1}\boldsymbol{X}^{(n)}(0) \xrightarrow{p} \boldsymbol{x}_0$ as $n \to \infty$, then Skorohod's theorem has to be invoked only for $\boldsymbol{X}^{(n)}(0)$ $(n = 1, 2, \dots)$.

Turning to Theorem 4.2, for $t \ge 0$, let $\boldsymbol{V}^{(n)}(t) = \sqrt{n}[\bar{\boldsymbol{X}}^{(n)}(t) - \boldsymbol{x}(t)]$, $\boldsymbol{R}^{(n)}(t) = \sqrt{n}[F(\bar{\boldsymbol{X}}^{(n)}(t)) - F(\boldsymbol{x}(t))]$, $\tilde{\boldsymbol{Y}}^{(n)}(t) = \sqrt{n}\bar{\boldsymbol{Y}}^{(n)}(t)$ and $\tilde{\boldsymbol{Z}}^{(n)}(t) = \sqrt{n}\bar{\boldsymbol{Z}}^{(n)}(t)$. Then, (A.9) and (4.5) yield

$$\boldsymbol{V}^{(n)}(t) = \boldsymbol{V}^{(n)}(0) + \tilde{\boldsymbol{Z}}^{(n)}(t) + \int_0^t \boldsymbol{R}^{(n)}(s)\,\mathrm{d}s + \tilde{\boldsymbol{Y}}^{(n)}(t) \qquad (t \ge 0). \tag{A.12}$$

Further, for $n = 1, 2, \dots$ and $t \ge 0$, it follows using the mean value theorem that there exists a $p \times p$ matrix $B^{(n)}(t)$ such that

$$\boldsymbol{R}^{(n)}(t) = \partial F(\boldsymbol{x}(t))\boldsymbol{V}^{(n)}(t) + B^{(n)}(t)\boldsymbol{V}^{(n)}(t); \tag{A.13}$$

cf. Britton and Pardoux (2019), Lemma 2.3.2. Moreover, as ∂F is continuous, it follows using (A.11) that

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left\| B^{(n)}(t) \right\| = 0 \qquad (T \ge 0). \tag{A.14}$$

Substituting (A.13) into (A.12) yields

$$\boldsymbol{V}^{(n)}(t) = \boldsymbol{V}^{(n)}(0) + \int_0^t \partial F(\boldsymbol{x}(s))\boldsymbol{V}^{(n)}(s)\,\mathrm{d}s + \boldsymbol{U}(t) + \boldsymbol{\epsilon}^{(n)}(t) \qquad (t \ge 0), \tag{A.15}$$

where $\boldsymbol{U}(t) = \sum_{\boldsymbol{l}\in\Delta} \boldsymbol{l}\hat{W}_{\boldsymbol{l}}\left(\int_0^t \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\,\mathrm{d}s\right)$ and

$$\boldsymbol{\epsilon}^{(n)}(t) = \tilde{\boldsymbol{Z}}^{(n)}(t) + \int_0^t B^{(n)}(s)\boldsymbol{V}^{(n)}(s)\,\mathrm{d}s + \tilde{\boldsymbol{Y}}^{(n)}(t) - \boldsymbol{U}(t). \tag{A.16}$$

We show now that, for all $T > 0$,

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left| \boldsymbol{\epsilon}^{(n)}(t) \right| = 0. \tag{A.17}$$

Fix $T > 0$ and $\boldsymbol{l} \in \Delta$. Then

$$\sup_{0\le t\le T} \left| \frac{1}{\sqrt{n}} \check{Y}_{\boldsymbol{l}}^{(n)} \left(\int_0^t n\beta_{\boldsymbol{l}}^{(n)} \left(\bar{\boldsymbol{X}}^{(n)}(s) \right) \mathrm{d}s \right) - \hat{W}_{\boldsymbol{l}} \left(\int_0^t \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\,\mathrm{d}s \right) \right| \le C_n + D_n, \tag{A.18}$$

where

$$C_n = \sup_{0\le t\le T}\left|\frac{1}{\sqrt{n}}\check{Y}_{\boldsymbol{l}}^{(n)}\left(\int_0^t n\beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\,\mathrm{d}s\right) - \hat{W}_{\boldsymbol{l}}\left(\int_0^t \beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\,\mathrm{d}s\right)\right|$$

and

$$D_n = \sup_{0\le t\le T}\left|\hat{W}_{\boldsymbol{l}}\left(\int_0^t \beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\,\mathrm{d}s\right) - \hat{W}_{\boldsymbol{l}}\left(\int_0^t \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\,\mathrm{d}s\right)\right|.$$

Using the continuity of $\beta_{\boldsymbol{l}}$, the limit in (4.3) and (A.11), there exists a constant $c>0$ such that $\int_0^T \beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\,\mathrm{d}s < c$ for all sufficiently large n, so (A.8) implies that $C_n\to 0$ as $n\to\infty$. Further, $\beta_{\boldsymbol{l}}$ is uniformly continuous on any compact subset of H. Hence, using

$$\begin{aligned}&\left|\beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right) - \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\right| \\ &\qquad\le \left|\beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right) - \beta_{\boldsymbol{l}}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\right| + \left|\beta_{\boldsymbol{l}}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right) - \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\right|,\end{aligned}$$

the second condition in (4.3) and (A.11) imply that

$$\lim_{n\to\infty}\sup_{0\le t\le T}\left|\int_0^t \beta_{\boldsymbol{l}}^{(n)}\left(\bar{\boldsymbol{X}}^{(n)}(s)\right)\,\mathrm{d}s - \int_0^t \beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\,\mathrm{d}s\right| = 0.$$

Thus $D_n\to 0$ as $n\to\infty$, since $\hat{W}_{\boldsymbol{l}}(t)$ is uniformly continuous on any finite interval. It then follows using (A.18) that

$$\lim_{n\to\infty}\sup_{0\le t\le T}\left|\tilde{\boldsymbol{Y}}^{(n)}(t) - \boldsymbol{U}(t)\right| = 0 \qquad (T>0). \tag{A.19}$$

Using Gronwall's inequality, it follows from (A.15) and (A.16) that, for any $T>0$,

$$\begin{aligned}\sup_{0\le t\le T}\left|\boldsymbol{V}^{(n)}(t)\right| \le &\left(\left|\boldsymbol{V}^{(n)}(0)\right| + \sup_{0\le t\le T}\left|\tilde{\boldsymbol{Z}}^{(n)}(t)\right| + \sup_{0\le t\le T}\left|\tilde{\boldsymbol{Y}}^{(n)}(t)\right|\right) \\ &\times \exp\left(\sup_{0\le t\le T}\left\|\partial F(\boldsymbol{x}(t)) + B^{(n)}(t)\right\| T\right).\end{aligned} \tag{A.20}$$

Now $\sup_{n\ge1}\left|\boldsymbol{V}^{(n)}(0)\right| < \infty$, using (A.7); $\sup_{n\ge1}\sup_{0\le t\le T}\left|\tilde{\boldsymbol{Y}}^{(n)}(t)\right| < \infty$, using (A.19) and the continuity of $\boldsymbol{U}$; and

$$\lim_{n\to\infty}\sup_{0\le t\le T}\left|\tilde{\boldsymbol{Z}}^{(n)}(t)\right| = 0 \qquad (T>0), \tag{A.21}$$

using (4.7) and (A.11). Recalling (A.14), it then follows from (A.20) that $\sup_{n\geq 1}\sup_{0\leq t\leq T}\left|\boldsymbol{V}^{(n)}(t)\right| < \infty$, since ∂F is continuous, whence using (A.14),

$$\lim_{n\to\infty}\sup_{0\leq t\leq T}\left|\int_0^t B^{(n)}(s)\boldsymbol{V}^{(n)}(s)\,\mathrm{d}s\right| = 0 \qquad (T>0). \tag{A.22}$$

The limit (A.17) now follows using (A.16), (A.19), (A.21) and (A.22).

The mapping $\boldsymbol{y}\mapsto\eta(\boldsymbol{y})$, which maps $\boldsymbol{y}\in C([0,T];\mathbb{R}^p)$ to $\eta(\boldsymbol{y})=\boldsymbol{x}\in C([0,T];\mathbb{R}^p)$ given by the solution of the integral equation

$$\boldsymbol{x}(t) = \int_0^t \partial F(\boldsymbol{x}(s)))\boldsymbol{x}(s)\,\mathrm{d}s + \boldsymbol{y}(t) \qquad (t\geq 0),$$

is continuous (Britton and Pardoux (2019)). Hence, (A.15) and (A.17) yield

$$\lim_{n\to\infty}\sup_{0\leq t\leq T}\left|\boldsymbol{V}^{(n)}(t)-\boldsymbol{V}(t)\right| = 0 \qquad (T\geq 0), \tag{A.23}$$

where $\boldsymbol{V}(t)$ $(t\geq 0)$ is the solution of the integral equation

$$\boldsymbol{V}(t) = \boldsymbol{V}(0) + \int_0^t \partial F(\boldsymbol{x}(s))\boldsymbol{V}(x)\,\mathrm{d}s + \boldsymbol{U}(t) \qquad (t\geq 0). \tag{A.24}$$

Now $\boldsymbol{V}(t)$ $(t\geq 0)$ is a sample path of the Gaussian process $\{\boldsymbol{V}(t): t\geq 0\}$, defined using (A.24) with $\boldsymbol{V}(0)\sim \mathrm{N}(\boldsymbol{0},\Sigma_0)$ and $\{\boldsymbol{U}(t): t\geq 0\}$ being given by $\boldsymbol{U}(t)=\sum_{\boldsymbol{l}\in\Delta}\boldsymbol{l}W_{\boldsymbol{l}}\left(\int_0^t\beta_{\boldsymbol{l}}(\boldsymbol{x}(s))\,\mathrm{d}s\right)$, where $\{W_{\boldsymbol{l}}(t): t\geq 0\}$ $(\boldsymbol{l}\in\Delta)$ are independent standard Brownian motions that are independent of $\boldsymbol{V}(0)$. The process $\{\boldsymbol{V}(t): t\geq 0\}$ admits the Itô integral representation

$$\boldsymbol{V}(t) = \Phi(t,0)\boldsymbol{V}(0) + \int_0^t \Phi(t,s)\,\mathrm{d}\boldsymbol{U}(s) \qquad (t\geq 0),$$

see Ethier and Kurtz (1986), page 458. Thus $\{\boldsymbol{V}(t): t\geq 0\}$ has zero mean and covariance function given by (4.10), since $\boldsymbol{V}(0)$ and $\{\boldsymbol{U}(t): t\geq 0\}$ are independent. We have thus shown the almost sure analogue of Theorem 4.2 on the probability space $(\Omega,\mathcal{F},\mathrm{P})$, so Theorem 4.2 follows.

APPENDIX B: DETAILED DERIVATIONS

B.1. Deterministic solution (5.11)-(5.13)**.** First note that (5.11) follows immediately from (5.8). Substituting (5.11) into (5.10) yields

$$\frac{d\tilde{z}_E}{dt} = q_I\left[\sum_{i=2}^{d_{\max}} i(i-1)\tilde{x}_i(0)\mathrm{e}^{-it}\right] - \tilde{z}_E$$

and (5.13) follows using the integrating factor e^t.

Multiplying (5.8) by i and adding over $i = 1, 2, \ldots, d_{\max}$ yields

$$\frac{d\tilde{x}_E}{dt} = -\sum_{i=2}^{d_{\max}} i(i-1)\tilde{x}_i - \tilde{x}_E. \tag{B.1}$$

Recall that $\tilde{\eta}_E(t) = \tilde{x}_E(t) + \tilde{y}_E(t) + \tilde{z}_E(t)$. Summing (B.1), (5.9) and (5.10) gives

$$\frac{d\tilde{\eta}_E}{dt} = -2\tilde{\eta}_E,$$

and (5.14) follows. Equation (5.12) then follows since $\tilde{y}_E(t) = \tilde{\eta}_E(t) - \tilde{x}_E(t) - \tilde{z}_E(t)$ and, using (5.11), $\tilde{x}_E(t) = \sum_{i=1}^{d_{\max}} i\tilde{x}_i(0)\mathrm{e}^{-it}$.

B.2. Properties of $\tilde{\tau}$ and $a(\tilde{\tau})$. In this appendix we show that $\tilde{\tau} \in (0, \infty)$ and $a(\tilde{\tau}) < 0$, first in the setting of Theorem 2.1 and then in the setting of Theorems 2.2 and 2.3. Let $G(s) = p_I f_{D_\epsilon}^{(1)}(s) - \mu_D(s - q_I)$ $(0 \le s \le 1)$. Then, from (5.23), $z = \mathrm{e}^{-\tilde{\tau}}$ satisfies $G(z) = 0$. Now $G(0) = p_I f_{D_\epsilon}^{(1)}(0) + \mu_D q_I > 0$, unless $q_I = 0$ and $p_1 - \epsilon_1 = 0$. Also $f_{D_\epsilon}^{(1)}(1) < \mu_D$, since $p_i\epsilon_i > 0$ for at least one $i > 0$, so $G(1) < 0$. Thus, under the conditions of Theorem 2.1, $G(s)$ has at least one zero in $(0, 1)$. Moreover it has precisely one zero, z say, as $G(s)$ is convex on $[0, 1]$, since $G^{(2)}(s) = p_I f_{D_\epsilon}^{(3)}(s) \ge 0$ for all $s \in [0, 1]$. Hence $\tilde{\tau} = -\log z \in (0, \infty)$, as required, and it follows from (5.24) that $a(\tilde{\tau}) < 0$ if and only if $G^{(1)}(z) < 0$. Suppose, for contradiction, that $G^{(1)}(z) \ge 0$. Then, since $G^{(2)}(s) \ge 0$ for all $s \in [z, 1]$,

$$G(1) \ge G(z) + \int_z^1 G^{(1)}(s)\,\mathrm{d}s \ge 0.$$

However, $G(1) < 0$, so $G^{(1)}(z) < 0$, as required.

Turning to the setting of Theorems 2.2 and 2.3, now let $G(s) = p_I f_D^{(1)}(s) - \mu_D(s - q_I)$ $(0 \le s \le 1)$. Then, from (5.49), $z = \mathrm{e}^{-\tilde{\tau}}$ satisfies $G(z) = 0$. Now $G(1) = 0$ and, under the conditions of Theorems 2.2 and 2.3, $G(0) > 0$. Also, using (2.8), $G^{(1)}(1) = p_I f_D^{(2)}(1) - \mu_D > 0$, since $R_0 > 1$. Thus, since $G(s)$ is convex on $[0, 1]$, there exists a unique $z \in (0, 1)$ such that $G(z) = 0$, so $\tilde{\tau} \in (0, \infty)$. Moreover, $G^{(2)}(s) > 0$ for $s \in [z, 1]$ and a similar contradiction argument to before shows that $G^{(1)}(z) < 0$, whence $a(\tilde{\tau}) < 0$.

B.3. Proof of $R_0(\tilde{\tau}) < 1$. Using (5.11) and (5.13), with $\tilde{x}_i(0) = p_i$ and $\tilde{z}_E(0) = 0$, yields $\sum_{i=1}^{d_{\max}} i(i-1)\tilde{x}_i(\tilde{\tau}) = z^2 f_D^{(2)}(z)$, $\tilde{x}_E(\tilde{\tau}) = z f_D^{(1)}(z)$ and $\tilde{z}_E(\tilde{\tau}) = q_I z[\mu_D - f_D^{(1)}(z)]$, where $z = \mathrm{e}^{-\tilde{\tau}}$. As in the previous paragraph, z is

the unique solution of $G(s)=0$ in $(0,1)$, where $G(s)=p_I f_D^{(1)}(s)-\mu_D(s-q_I)$. Thus $p_I f_D^{(1)}(z)=\mu_D(z-q_I)$ and, using (5.62),

$$R_0(\tilde{\tau})<1 \iff p_I z f_D^{(2)}(z) < p_I f_D^{(1)}(z)+q_I\mu_D \iff p_I f_D^{(2)}(z)<\mu_D,$$

which holds as we have shown above that $G^{(1)}(z)<0$.

B.4. Asymptotic variances σ^2_{MR} and σ^2_{NSW}. In this appendix we derive the expressions for σ^2_{MR} and σ^2_{NSW} given in Theorem 2.1. Recalling the partition $\Delta=\Delta_1\cup\Delta_2\cup\Delta_3$ defined in Section 5.2, it follows from (5.37) that

$$\sigma^2_{\mathrm{MR}}=\sigma^2_{0,MR}+\sum_{i=1}^{3}\sigma_i^2 \qquad \text{and} \qquad \sigma^2_{\mathrm{NSW}}=\sigma^2_{0,NSW}+\sum_{i=1}^{3}\sigma_i^2, \tag{B.2}$$

where

$$\sigma^2_{0,MR}=\boldsymbol{c}(\tilde{\tau},0)\Sigma_0^{\mathrm{MR}}\boldsymbol{c}(\tilde{\tau},0)^\top, \qquad \sigma^2_{0,NSW}=\boldsymbol{c}(\tilde{\tau},0)\Sigma_0^{\mathrm{NSW}}\boldsymbol{c}(\tilde{\tau},0)^\top \tag{B.3}$$

and, for $i=1,2,3$,

$$\sigma_i^2=\int_0^{\tilde{\tau}}\sum_{\boldsymbol{l}\in\Delta_i}\left(\boldsymbol{c}(\tilde{\tau},u)\boldsymbol{l}^\top\right)^2\tilde{\beta}_{\boldsymbol{l}}(\tilde{\boldsymbol{w}}(u))\,\mathrm{d}u.$$

We calculate $\sum_{i=1}^3\sigma_i^2$ in Appendix B.4.1, $\sigma^2_{0,MR}$ and then σ^2_{MR} in Appendix B.4.2, and $\sigma^2_{0,NSW}$ and then σ^2_{NSW} in Appendix B.4.3.

B.4.1. *Calculation of* $\sum_{i=1}^3\sigma_i^2$. Noting that $\tilde{y}_E(t)\geq 0$ for $0\leq t\leq\tilde{\tau}$, it follows from (5.3), (5.34) and (5.35) that

$$\sigma_2^2=\int_0^{\tilde{\tau}}4c_I(\tilde{\tau},u)^2\tilde{y}_E(u)\,\mathrm{d}u \tag{B.4}$$

and

$$\begin{aligned}\sigma_3^2 &= \int_0^{\tilde{\tau}}(c_I(\tilde{\tau},u)+c_R(\tilde{\tau},u))^2\tilde{z}_E(u)\,\mathrm{d}u \\ &= \int_0^{\tilde{\tau}}(2c_I(\tilde{\tau},u)-c_J(\tilde{\tau},u))^2\tilde{z}_E(u)\,\mathrm{d}u,\end{aligned} \tag{B.5}$$

where

$$c_J(\tilde{\tau},u)=b(\tilde{\tau})\mathrm{e}^{-(\tilde{\tau}-u)}. \tag{B.6}$$

Now

$$\sigma_1^2 = \int_0^{\tilde{\tau}} \tilde{g}(u)\,\mathrm{d}u, \tag{B.7}$$

where, using (5.3),

$$\begin{aligned}\tilde{g}(u) &= \sum_{\boldsymbol{l}\in\Delta_1} \left(\boldsymbol{c}(\tilde{\tau},u)\boldsymbol{l}^\top\right)^2 \tilde{\beta}_{\boldsymbol{l}}(\tilde{\boldsymbol{w}}(u)) \\ &= \sum_{i=1}^{d_{\max}}\sum_{k=0}^{i-1}\left[\boldsymbol{c}(\tilde{\tau},u)(\boldsymbol{l}_{ik}^{(1)})^\top\right]^2 p_{i,k} i\tilde{x}_i(u). \end{aligned} \tag{B.8}$$

Recalling (5.36), it follows using (5.3) that, for $i = 1, 2, \ldots, d_{\max}$ and $k = 0, 1, \ldots, i-1$,

$$\boldsymbol{c}(\tilde{\tau},u)(\boldsymbol{l}_{ik}^{(1)})^\top = (p_I b(\tilde{\tau}) - 1)\mathrm{e}^{-i(\tilde{\tau}-u)} - 2c_I(\tilde{\tau},u) + [iq_I - (i-k-1)]c_I(\tilde{\tau},u). \tag{B.9}$$

Recall that $p_{i,k} = \mathrm{P}(X = k)$, where $X \sim \mathrm{Bin}(i-1, 1-\exp(-\lambda I))$; see Section 5.1. Elementary calculation yields, for $i = 1, 2, \ldots, d_{\max}$, that

$$\sum_{k=0}^{i-1}(i-k-1)p_{i,k} = (i-1)q_I \tag{B.10}$$

and

$$\sum_{k=0}^{i-1}(i-k-1)^2 p_{i,k} = (i-1)(i-2)q_I^{(2)} + (i-1)q_I. \tag{B.11}$$

Using (B.10) and (B.11), it follows from (B.9) and some algebra that, for $i = 1, 2, \ldots, d_{\max}$

$$\begin{aligned}\sum_{k=0}^{i-1}\left[\boldsymbol{c}(\tilde{\tau},u)(\boldsymbol{l}_{ik}^{(1)})^\top\right]^2 p_{i,k} &= [2c_I(\tilde{\tau},u) - q_I c_J(\tilde{\tau},u)]^2 \\ &\quad + q_I p_I c_J(\tilde{\tau},u)^2(i-1) + (q_I^{(2)} - q_I^2)c_J(\tilde{\tau},u)^2(i-1)(i-2) \\ &\quad + 2[2c_I(\tilde{\tau},u) - q_I c_J(\tilde{\tau},u)]\mathrm{e}^{-i(\tilde{\tau}-u)}[1 - ib(\tilde{\tau})p_I] \\ &\quad + \mathrm{e}^{-2i(\tilde{\tau}-u)}\left[(1-b(\tilde{\tau})p_I)^2 - (i-1)b(\tilde{\tau})p_I(2-3b(\tilde{\tau})p_I)\right. \\ &\qquad \left. +(i-1)(i-2)b(\tilde{\tau})^2 p_I^2\right].\end{aligned} \tag{B.12}$$

Recall from (5.17) and (5.21) that, under both the MR and NSW models, $\tilde{x}_i(0) = p_i - \epsilon_i$ $(i = 0, 1, \ldots, d_{\max})$, so (5.11) yields

$$\tilde{x}_i(u) = (p_i - \epsilon_i)\mathrm{e}^{-iu} \qquad (i = 0, 1, \ldots, d_{\max}). \tag{B.13}$$

For $i, k = 0, 1, \ldots$, let $i_{[k]} = i(i-1)\ldots(i-k+1)$ denote a falling factorial, with the convention that $i_{[0]} = 1$. Then it follows from (B.13) that, for $\theta \in \mathbb{R}$ and $k = 1, 2, \ldots$,

$$\sum_{i=1}^{d_{\max}} i_{[k]} \theta^{i-k} \tilde{x}_i(u) = \mathrm{e}^{-ku} f_{D_\epsilon}^{(k)}(\theta \mathrm{e}^{-u}).$$

Thus, for $k = 1, 2, \ldots$,

$$\sum_{i=1}^{d_{\max}} i_{[k]} \tilde{x}_i(u) = \mathrm{e}^{-ku} f_{D_\epsilon}^{(k)}(\mathrm{e}^{-u}), \tag{B.14}$$

$$\sum_{i=1}^{d_{\max}} i_{[k]} \mathrm{e}^{-i(\tilde{\tau}-u)} \tilde{x}_i(u) = \mathrm{e}^{-k\tilde{\tau}} f_{D_\epsilon}^{(k)}(\mathrm{e}^{-\tilde{\tau}}), \tag{B.15}$$

$$\sum_{i=1}^{d_{\max}} i_{[k]} \mathrm{e}^{-2i(\tilde{\tau}-u)} \tilde{x}_i(u) = \mathrm{e}^{-k(2\tilde{\tau}-u)} f_{D_\epsilon}^{(k)}(\mathrm{e}^{-(2\tilde{\tau}-u)}). \tag{B.16}$$

Substituting (B.12) into (B.8) and using (B.14)-(B.16) yields

$$\begin{aligned}
\tilde{g}(u) = {} & \left[2c_I(\tilde{\tau}, u) - q_I c_J(\tilde{\tau}, u)\right]^2 \tilde{x}_E(u) + c_J(\tilde{\tau}, u)^2 p_I q_I \mathrm{e}^{-2u} f_{D_\epsilon}^{(2)}(\mathrm{e}^{-u}) \\
& + c_J(\tilde{\tau}, u)^2 (q_I^{(2)} - q_I^2) \mathrm{e}^{-3u} f_{D_\epsilon}^{(3)}(\mathrm{e}^{-u}) \\
& + 2\left[2c_I(\tilde{\tau}, u) - q_I c_J(\tilde{\tau}, u)\right]\left[1 - b(\tilde{\tau}) p_I\right] \mathrm{e}^{-\tilde{\tau}} f_{D_\epsilon}^{(1)}(\mathrm{e}^{-\tilde{\tau}}) \\
& - 2b(\tilde{\tau}) p_I \left[2c_I(\tilde{\tau}, u) - q_I c_J(\tilde{\tau}, u)\right] \mathrm{e}^{-2\tilde{\tau}} f_{D_\epsilon}^{(2)}(\mathrm{e}^{-\tilde{\tau}}) \\
& + \left[1 - b(\tilde{\tau}) p_I\right]^2 \mathrm{e}^{-(2\tilde{\tau}-u)} f_{D_\epsilon}^{(1)}(\mathrm{e}^{-(2\tilde{\tau}-u)}) \\
& - b(\tilde{\tau}) p_I \left[2 - 3b(\tilde{\tau}) p_I\right] \mathrm{e}^{-2(2\tilde{\tau}-u)} f_{D_\epsilon}^{(2)}(\mathrm{e}^{-(2\tilde{\tau}-u)}) \\
& + b(\tilde{\tau})^2 p_I^2 \mathrm{e}^{-3(2\tilde{\tau}-u)} f_{D_\epsilon}^{(3)}(\mathrm{e}^{-(2\tilde{\tau}-u)}).
\end{aligned}$$

It then follows using (B.4), (B.5) and (B.7) that

$$\sigma_1^2 + \sigma_2^2 + \sigma_3^2 = \sum_{i=1}^{10} I_i, \tag{B.17}$$

where

$$I_i = \int_0^{\tilde{\tau}} f_i(u)\, \mathrm{d}u,$$

with

$$\begin{aligned}
f_1(u) &= 4c_I(\tilde{\tau},u)^2\left(\tilde{x}_E(u)+\tilde{y}_E(u)+\tilde{z}_E(u)\right),\\
f_2(u) &= \left[c_J(\tilde{\tau},u)^2-4c_I(\tilde{\tau},u)c_J(\tilde{\tau},u)\right]]\tilde{z}_E(u),\\
f_3(u) &= q_I\left[q_Ic_J(\tilde{\tau},u)^2-4c_I(\tilde{\tau},u)c_J(\tilde{\tau},u)\right]]\tilde{x}_E(u),\\
f_4(u) &= c_J(\tilde{\tau},u)^2p_Iq_I\mathrm{e}^{-2u}f^{(2)}_{D_\epsilon}(\mathrm{e}^{-u}),\\
f_5(u) &= c_J(\tilde{\tau},u)^2(q_I^{(2)}-q_I^2)\mathrm{e}^{-3u}f^{(3)}_{D_\epsilon}(\mathrm{e}^{-u}),\\
f_6(u) &= 2\left[2c_I(\tilde{\tau},u)-q_Ic_J(\tilde{\tau},u)\right]\left[1-b(\tilde{\tau})p_I\right]\mathrm{e}^{-\tilde{\tau}}f^{(1)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}),\\
f_7(u) &= -2b(\tilde{\tau})p_I\left[2c_I(\tilde{\tau},u)-q_Ic_J(\tilde{\tau},u)\right]\mathrm{e}^{-2\tilde{\tau}}f^{(2)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}),\\
f_8(u) &= \left[1-b(\tilde{\tau})p_I\right]^2\mathrm{e}^{-(2\tilde{\tau}-u)}f^{(1)}_{D_\epsilon}(\mathrm{e}^{-(2\tilde{\tau}-u)}),\\
f_9(u) &= -b(\tilde{\tau})p_I\left[2-3b(\tilde{\tau})p_I\right]\mathrm{e}^{-2(2\tilde{\tau}-u)}f^{(2)}_{D_\epsilon}(\mathrm{e}^{-(2\tilde{\tau}-u)}),\\
f_{10}(u) &= b(\tilde{\tau})^2p_I^2\mathrm{e}^{-3(2\tilde{\tau}-u)}f^{(3)}_{D_\epsilon}(\mathrm{e}^{-(2\tilde{\tau}-u)}).
\end{aligned}$$

Noting that $\tilde{\eta}_E(0)=\mu_D$, it follows using (5.34) and (5.14) that

$$I_1 = 2b(\tilde{\tau})^2\mu_D\mathrm{e}^{-2\tilde{\tau}}\left(1-\mathrm{e}^{-2\tilde{\tau}}\right). \tag{B.18}$$

Recall from (5.17) and (5.21) that under both the MR and NSW models, $\tilde{x}_E(0)=\sum_{i=1}^{d_{\max}} i(p_i-\epsilon_i)$ and $\tilde{z}_E(0)=q_I\sum_{i=1}^{d_{\max}} i\epsilon_i$. It then follows from (5.13) and (B.13) that

$$\tilde{z}_E(u) = q_I\left(\mu_D\mathrm{e}^{-u}-\tilde{x}_E(u)\right),$$

so

$$f_2(u)+f_3(u) = c_J(\tilde{\tau},u)\mu_Dq_I\left[c_J(\tilde{\tau},u)-4c_I(\tilde{\tau},u)\right]\mathrm{e}^{-u}-c_J(\tilde{\tau},u)^2q_Ip_I\tilde{x}_E(u).$$

Also, setting $k=1$ in (B.14) yields $\tilde{x}_E(u)=\mathrm{e}^{-u}f^{(1)}_{D_\epsilon}(\mathrm{e}^{-u})$. Thus,

$$\begin{aligned}
I_2+I_3+I_4 = &\int_0^{\tilde{\tau}} c_J(\tilde{\tau},u)\mu_Dq_I\left[c_J(\tilde{\tau},u)-4c_I(\tilde{\tau},u)\right]\mathrm{e}^{-u}\,\mathrm{d}u\\
&+q_Ip_I\int_0^{\tilde{\tau}} c_J(\tilde{\tau},u)^2\left[\mathrm{e}^{-2u}f^{(2)}_{D_\epsilon}(\mathrm{e}^{-u})-\mathrm{e}^{-u}f^{(1)}_{D_\epsilon}(\mathrm{e}^{-u})\right]\mathrm{d}u.
\end{aligned} \tag{B.19}$$

The first integral in (B.19) involves only exponential functions and is easily evaluated. Using (B.6), the integrand in the second integral in (B.19) can be expressed as

$$-b(\tilde{\tau})^2\mathrm{e}^{-2\tilde{\tau}}\frac{d}{dt}\left[\mathrm{e}^{u}f^{(1)}_{D_\epsilon}(\mathrm{e}^{-u})\right],$$

so that integral is also easily evaluated. Hence, omitting the details,

(B.20)
$$I_2 + I_3 + I_4 = 2b(\tilde{\tau})^2\mu_D\left(\mathrm{e}^{-2\tilde{\tau}} - \mathrm{e}^{-4\tilde{\tau}}\right) - b(\tilde{\tau})^2\mu_D q_I \mathrm{e}^{-\tilde{\tau}}\left(1 + \mathrm{e}^{-\tilde{\tau}} - 2\mathrm{e}^{-2\tilde{\tau}}\right) + q_I p_I b(\tilde{\tau})^2\left[\mathrm{e}^{-2\tilde{\tau}} f^{(1)}_{D_\epsilon}(1) - \mathrm{e}^{-\tilde{\tau}} f^{(1)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}})\right].$$

Turning to I_5, note that

$$c_J(\tilde{\tau}, u)^2 \mathrm{e}^{-3u} f^{(3)}_{D_\epsilon}(\mathrm{e}^{-u}) = -b(\tilde{\tau})^2 \mathrm{e}^{-2\tilde{\tau}} \frac{d}{dt}\left[f^{(2)}_{D_\epsilon}(\mathrm{e}^{-u})\right],$$

so

$$I_5 = b(\tilde{\tau})^2\left(q_I^{(2)} - q_I^2\right)\mathrm{e}^{-2\tilde{\tau}}\left[f^{(2)}_{D_\epsilon}(1) - f^{(2)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}})\right]. \tag{B.21}$$

The integrals I_6 and I_7 are easily evaluated yielding

$$I_6 + I_7 = 2b(\tilde{\tau})\left\{[1 - p_I b(\tilde{\tau})]\,\mathrm{e}^{-\tilde{\tau}} f^{(1)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - p_I b(\tilde{\tau})\mathrm{e}^{-2\tilde{\tau}} f^{(2)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}})\right\} \times \left(p_I - \mathrm{e}^{-2\tilde{\tau}} + q_I \mathrm{e}^{-\tilde{\tau}}\right). \tag{B.22}$$

To evaluate I_8 to I_{10}, let $J_k = \int_0^{\tilde{\tau}} \mathrm{e}^{-k(2\tilde{\tau}-u)} f^{(k)}_{D_\epsilon}(\mathrm{e}^{-(2\tilde{\tau}-u)})\,\mathrm{d}u$ $(k = 1, 2, \dots)$. Then a simple reduction formula yields

$$J_k = \mathrm{e}^{-(k-1)\tilde{\tau}} f^{(k-1)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - \mathrm{e}^{-2(k-1)\tilde{\tau}} f^{(k-1)}_{D_\epsilon}(\mathrm{e}^{-2\tilde{\tau}}) - (k-1)J_{k-1},$$

for $k = 2, 3, \dots$, and

$$J_1 = f_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - f_{D_\epsilon}(\mathrm{e}^{-2\tilde{\tau}}).$$

Applying these formulae to I_8 to I_{10} yields after some algebra that

(B.23)
$$I_8 + I_9 + I_{10} = f_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - f_{D_\epsilon}(\mathrm{e}^{-2\tilde{\tau}}) + b(\tilde{\tau})p_I(b(\tilde{\tau})p_I - 2)\left[\mathrm{e}^{-\tilde{\tau}} f^{(1)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - \mathrm{e}^{-2\tilde{\tau}} f^{(1)}_{D_\epsilon}(\mathrm{e}^{-2\tilde{\tau}})\right] + b(\tilde{\tau})^2 p_I^2\left[\mathrm{e}^{-2\tilde{\tau}} f^{(2)}_{D_\epsilon}(\mathrm{e}^{-\tilde{\tau}}) - \mathrm{e}^{-4\tilde{\tau}} f^{(2)}_{D_\epsilon}(\mathrm{e}^{-2\tilde{\tau}})\right].$$

Letting $z = \mathrm{e}^{-\tilde{\tau}}$, it follows from (B.17) and the above equations that

(B.24)
$$\begin{aligned}\sum_{i=1}^{3}\sigma_i^2 =& 2\mu_D b(\tilde{\tau})^2(z^2-z^4) - \mu_D q_I b(\tilde{\tau})^2(z+z^2-2z^3)\\ &+ p_I q_I b(\tilde{\tau})^2\left[z^2 f_{D_\epsilon}^{(1)}(1) - z f_{D_\epsilon}^{(1)}(z)\right] + p_I^2 b(\tilde{\tau})^2\left[z^2 f_{D_\epsilon}^{(2)}(z) - z^4 f_{D_\epsilon}^{(2)}(z^2)\right]\\ &+ 2b(\tilde{\tau})\left[(1-p_I b(\tilde{\tau}))z f_{D_\epsilon}^{(1)}(z) - p_I b(\tilde{\tau}) z^2 f_{D_\epsilon}^{(2)}(z)\right]\left(p_I - z^2 + q_I z\right)\\ &+ f_{D_\epsilon}(z) - f_{D_\epsilon}(z^2) + p_I b(\tilde{\tau})\left(p_I b(\tilde{\tau}) - 2\right)\left[z f_{D_\epsilon}^{(1)}(z) - z^2 f_{D_\epsilon}^{(1)}(z^2)\right]\\ &+ (q_I^{(2)} - q_I^2) b(\tilde{\tau})^2 z^2\left[f_{D_\epsilon}^{(2)}(1) - f_{D_\epsilon}^{(2)}(z)\right].\end{aligned}$$

B.4.2. *Calculation of $\sigma^2_{0,MR}$ and σ^2_{MR}.* To determine $\sigma^2_{0,MR}$, note from (B.3), (5.18) and (5.32) that

$$\begin{aligned}\sigma_{0,MR}^2 &= \left[c_I(\tilde{\tau},0) - c_R(\tilde{\tau},0)\right]^2 \sigma_Y^2\\ &= b(\tilde{\tau})^2 z^2 \sigma_Y^2,\end{aligned} \tag{B.25}$$

using (5.34), (5.35) and $z = \mathrm{e}^{-\tilde{\tau}}$. Now σ_Y^2 is given by (5.15), where, for $i = 1, 2, \ldots, d_{\max}$, $\sigma^2_{Y,i} = \mathrm{var}(Y_{i1})$ with $Y_{i1} \sim \mathrm{Bin}(i, 1-\exp(-\lambda I))$. A simple calculation, conditioning on I, yields

$$\sigma^2_{Y,i} = i(i-1)q_I^{(2)} + i q_I - i^2 q_I^2. \tag{B.26}$$

Now $\sum_{i=1}^{d_{\max}} i\epsilon_i = \mu_D - f_{D_\epsilon}^{(1)}(1)$ and $\sum_{i=1}^{d_{\max}} i(i-1)\epsilon_i = f_D^{(2)}(1) - f_{D_\epsilon}^{(2)}(1)$, so

$$\sigma_Y^2 = \left[(q_I^{(2)} - q_I^2)\left(f_D^{(2)}(1) - f_{D_\epsilon}^{(2)}(1)\right) + p_I q_I\left(\mu_D - f_{D_\epsilon}^{(1)}(1)\right)\right]. \tag{B.27}$$

Using (B.2), adding (B.24) and (B.25), after substituting from (B.27), yields

(B.28)
$$\begin{aligned}\sigma^2_{\mathrm{MR}} =& 2\mu_D b(\tilde{\tau})^2(z^2-z^4) - \mu_D q_I b(\tilde{\tau})^2(z+q_I z^2-2z^3) - p_I q_I b(\tilde{\tau})^2 z f_{D_\epsilon}^{(1)}(z)\\ &+ 2b(\tilde{\tau})\left[(1-p_I b(\tilde{\tau}))z f_{D_\epsilon}^{(1)}(z) - p_I b(\tilde{\tau}) z^2 f_{D_\epsilon}^{(2)}(z)\right]\left(p_I - z^2 + q_I z\right)\\ &+ f_{D_\epsilon}(z) - f_{D_\epsilon}(z^2) + p_I b(\tilde{\tau})\left(p_I b(\tilde{\tau}) - 2\right)\left[z f_{D_\epsilon}^{(1)}(z) - z^2 f_{D_\epsilon}^{(1)}(z^2)\right]\\ &- p_I q_I b(\tilde{\tau})^2 z f_{D_\epsilon}^{(1)}(z) + b(\tilde{\tau})^2 p_I^2 z^2\left[f_{D_\epsilon}^{(2)}(z) - z^2 f_{D_\epsilon}^{(2)}(z^2)\right]\\ &+ (q_I^{(2)} - q_I^2) b(\tilde{\tau})^2 z^2\left[f_D^{(2)}(1) - f_{D_\epsilon}^{(2)}(z)\right].\end{aligned}$$

Using (5.23) and recalling that $z = \mathrm{e}^{-\tilde{\tau}}$ yields

$$\mu_D(z - q_I) = p_I f^{(1)}_{D_\epsilon}(z). \tag{B.29}$$

Also, using (5.24), (5.30) and (B.14), with $k = 1$, gives

$$b(\tilde{\tau}) = \frac{f^{(1)}_{D_\epsilon}(z)}{z\left(p_I f^{(2)}_{D_\epsilon}(z) - \mu_D\right)}, \tag{B.30}$$

so

$$b(\tilde{\tau}) z p_I f^{(2)}_{D_\epsilon}(z) = b(\tilde{\tau}) z \mu_D + f^{(1)}_{D_\epsilon}(z). \tag{B.31}$$

Substituting (B.29) and (B.31) into (B.28), recalling (2.3) and noting from (B.29) and (B.30) that $h(z) = b(\tilde{\tau})z$ yields (2.4) after some algebra.

B.4.3. *Calculation of $\sigma^2_{0,NSW}$ and σ^2_{NSW}.* To determine $\sigma^2_{0,NSW}$, first note that (B.3), (5.22) and (5.32) imply that

$$\sigma^2_{0,NSW} = (1 - \epsilon)\sigma^2_A + \epsilon \sigma^2_B, \tag{B.32}$$

where

$$\sigma^2_A = \boldsymbol{c}(\tilde{\tau}, 0) \Sigma_{XX} \boldsymbol{c}(\tilde{\tau}, 0)^\top \tag{B.33}$$

and

$$\sigma^2_B = c_I(\tilde{\tau}, 0)^2 \sigma^2_{Y_E} + 2 c_I(\tilde{\tau}, 0) c_R(\tilde{\tau}, 0) \sigma_{Y_E, Z_E} + c_R(\tilde{\tau}, 0)^2 \sigma^2_{Z_E}. \tag{B.34}$$

Recalling (5.19) and (5.33),

$$\boldsymbol{c}(\tilde{\tau}, 0) \Sigma_{XX} \boldsymbol{c}(\tilde{\tau}, 0)^\top = \sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau}, 0)^2 - \left(\sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau}, 0) \right)^2, \tag{B.35}$$

where from (5.36) and recalling that $z = \mathrm{e}^{-\tilde{\tau}}$,

$$c_i(\tilde{\tau}, 0) = z^i + i b(\tilde{\tau}) z (z - q_I) - p_I b(\tilde{\tau}) i z^i.$$

Thus,

$$\sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau}, 0) = f_D(z) + b(\tilde{\tau}) z (z - q_I) \mu_D - p_I b(\tilde{\tau}) z f^{(1)}_D(z)$$

and

$$\sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau},0)^2 = f_D(z^2) + b(\tilde{\tau})^2 z^2 (z-q_I)^2(\sigma_D^2+\mu_D^2) + p_I^2 b(\tilde{\tau})^2 z^2 \left(z^2 f_D^{(2)}(z^2) + f_D^{(1)}(z^2) \right) + 2b(\tilde{\tau}) z^2 (z-q_I) f_D^{(1)}(z) - 2p_I b(\tilde{\tau}) z^2 f_D^{(1)}(z^2) - 2p_I b(\tilde{\tau})^2 z^2 (z-q_I) \left(z f_D^{(2)}(z) + f_D^{(1)}(z) \right).$$

Note that in the NSW model

$$f_{D_\epsilon}(s) = (1-\epsilon) f_D(s) \qquad (s \in \mathbb{R}). \tag{B.36}$$

Hence, using (B.29),

$$\sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau},0) = f_D(z) - \frac{\epsilon}{1-\epsilon} b(\tilde{\tau}) z (z-q_I) \mu_D$$

and, using (B.29) and (B.31),

$$(1-\epsilon) \sum_{i=0}^{d_{\max}} p_i c_i(\tilde{\tau},0)^2 = f_{D_\epsilon}(z^2) + p_I^2 b(\tilde{\tau})^2 z^2 \left(z^2 f_{D_\epsilon}^{(2)}(z^2) + f_{D_\epsilon}^{(1)}(z^2) \right) - 2p_I b(\tilde{\tau}) z^2 f_{D_\epsilon}^{(1)}(z^2) + (1-\epsilon) b(\tilde{\tau})^2 z^2 (z-q_I)^2 \left(\sigma_D^2 + \mu_D^2 \right) - 2b(\tilde{\tau})^2 z^3 (z-q_I) \mu_D - 2b(\tilde{\tau})^2 z^2 (z-q_I)^2 \mu_D.$$

It then follows using (B.33) and (B.35) that

$$(1-\epsilon)\sigma_A^2 = f_{D_\epsilon}(z^2) - (1-\epsilon) f_D(z)^2 + p_I^2 b(\tilde{\tau})^2 z^2 \left(z^2 f_{D_\epsilon}^{(2)}(z^2) + f_{D_\epsilon}^{(1)}(z^2) \right) - 2p_I b(\tilde{\tau}) z^2 f_{D_\epsilon}^{(1)}(z^2) + 2\epsilon b(\tilde{\tau}) z (z-q_I) \mu_D f_D(z) + b(\tilde{\tau})^2 z^2 (z-q_I)^2 \left[(1-\epsilon)\sigma_D^2 + \frac{1-2\epsilon}{1-\epsilon} \mu_D^2 - 2\mu_D \right] - 2b(\tilde{\tau})^2 z^3 (z-q_I) \mu_D. \tag{B.37}$$

Recall the definition of (Y_E, Z_E) just before (5.19). Now $\mathrm{E}[Y_E|D] = p_I D$ and, using (B.26), $\mathrm{var}(Y_E|D) = D(D-1)q_I^{(2)} + Dq_I - D^2 q_I^2$, so

$$\begin{aligned} \sigma_{Y_E}^2 &= \mathrm{E}\left[\mathrm{var}(Y_E|D)\right] + \mathrm{var}(\mathrm{E}[Y_E|D]) \\ &= (q_I^{(2)} - q_I^2) f_D^{(2)}(1) + p_I q_I \mu_D + p_I^2 \sigma_D^2. \end{aligned} \tag{B.38}$$

Similar arguments show that

$$\sigma_{Z_E}^2 = (q_I^{(2)} - q_I^2) f_D^{(2)}(1) + p_I q_I \mu_D + q_I^2 \sigma_D^2 \tag{B.39}$$

and

$$\sigma_{Y_E,Z_E} = -\left[(q_I^{(2)} - q_I^2) f_D^{(2)}(1) + p_I q_I \mu_D\right] + p_I q_I \sigma_D^2. \tag{B.40}$$

Thus, using (B.34), (5.34), (5.35) and $z = \mathrm{e}^{-\tilde{\tau}}$,

$$\sigma_B^2 = b(\tilde{\tau})^2 z^2 \left[(q_I^{(2)} - q_I^2) f_D^{(2)}(1) + p_I q_I \mu_D + \sigma_D^2 (z - q_I)^2\right].$$

Noting from (B.36) that $f_{D_\epsilon}^{(1)}(1) = (1-\epsilon)\mu_D$ and $f_{D_\epsilon}^{(2)}(1) = (1-\epsilon) f_D^{(2)}(1)$, it follows using (B.25) that

$$\epsilon \sigma_B^2 - \sigma_{0,MR}^2 = \epsilon b(\tilde{\tau})^2 z^2 (z - q_I)^2 \sigma_D^2. \tag{B.41}$$

Note from (B.2) that $\sigma_{\mathrm{NSW}}^2 - \sigma_{\mathrm{MR}}^2 = \sigma_{0,NSW}^2 - \sigma_{0,MR}^2$. Hence, using (B.32), (B.37) and (B.41),

(B.42)

$$\begin{aligned}
\sigma_{\mathrm{NSW}}^2 - \sigma_{\mathrm{MR}}^2 = {} & f_{D_\epsilon}(z^2) - (1-\epsilon) f_D(z)^2 + p_I^2 b(\tilde{\tau})^2 z^2 \left(z^2 f_{D_\epsilon}^{(2)}(z^2) + f_{D_\epsilon}^{(1)}(z^2)\right) \\
& - 2 p_I b(\tilde{\tau}) z^2 f_{D_\epsilon}^{(1)}(z^2) + 2\epsilon b(\tilde{\tau}) z (z - q_I) \mu_D f_D(z) \\
& + b(\tilde{\tau})^2 z^2 (z - q_I)^2 \left[\sigma_D^2 + \frac{1-2\epsilon}{1-\epsilon} \mu_D^2 - 2\mu_D\right] \\
& - 2 b(\tilde{\tau})^2 z^3 (z - q_I) \mu_D.
\end{aligned}$$

Recall that $h(z) = b(\tilde{\tau}) z$. Note from (2.3) and (B.36) that $f_D(z) = (1-\epsilon-\rho)/(1-\epsilon)$ and also that $f_{D_\epsilon}^{(2)}(z) = (1-\epsilon) f_D^{(2)}(z)$. Adding (2.4) and (B.42) then yields (2.6).

B.4.4. *Proof of Remark 2.5.* Suppose that the support of D is not concentrated on a single point. Then (B.33), (B.35) and Jensen's inequality imply $\sigma_A^2 > 0$, since $c_i(\tilde{\tau}, 0)$ $(i = 0, 1, \ldots, d_{\max})$ are not all equal. It then follows from (B.32) and (B.41) that $\sigma_{0,NSW}^2 > \sigma_{0,MR}^2$, whence, using (B.2), $\sigma_{\mathrm{NSW}}^2 > \sigma_{\mathrm{MR}}^2$. Further, setting $\epsilon = 0$ yields $\tilde{\sigma}_{\mathrm{NSW}}^2 > \tilde{\sigma}_{\mathrm{MR}}^2$.

B.5. Asymptotic variances $\sigma^2_{\mathbf{MR,S}}$ and $\sigma^2_{\mathbf{NSW,S}}$. For $i = 1, 2, 3$, let

$$\bar{\sigma}_i^2 = \int_0^{\tilde{\tau}} \sum_{\hat{\boldsymbol{l}} \in \hat{\Delta}_i} \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u) \hat{\boldsymbol{l}}^\top \right)^2 \bar{\beta}_{\boldsymbol{l}}(\bar{\boldsymbol{w}}(u)) \, \mathrm{d}u.$$

Note from (5.57) that $\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_+^{(2)})^\top = -\pi \boldsymbol{c}(\tilde{\tau}, u)(\boldsymbol{l}_+^{(2)})^\top$, so $\bar{\sigma}_2^2 = \pi^2 \sigma_2^2$ since $\bar{\beta}_{\boldsymbol{l}}(\bar{\boldsymbol{w}}(u))$ is independent of $\bar{v}$ for $\boldsymbol{l} \in \hat{\Delta}_2$. Similarly, $\bar{\sigma}_3^2 = \pi^2 \sigma_3^2$.

Turning to $\bar{\sigma}_1^2$, note that for $i = 1, 2, \ldots, d_{\max}$,

$$\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,0}^{(1)})^\top = -\pi \left[\boldsymbol{c}(\boldsymbol{l}_{i,0}^{(1)})^\top \right] \quad \text{and} \quad \bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,i-1}^{(1)})^\top = 1 - \pi \left[\boldsymbol{c}(\boldsymbol{l}_{i,i-1}^{(1)}))^\top \right],$$

whence,

$$\begin{aligned}
&(1-\pi) \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,0}^{(1)})^\top \right)^2 + \pi \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,i-1}^{(1)})^\top \right)^2 \\
&\quad = \pi(1-\pi) + \pi^2 \left[(1-\pi) \left(\boldsymbol{c}(\tilde{\tau}, u) \boldsymbol{l}_{i,0}^{(1)})^\top \right)^2 + \pi \left(\boldsymbol{c}(\tilde{\tau}, u) \boldsymbol{l}_{i,i-1}^{(1)})^\top \right)^2 \right] \\
&\qquad + 2\pi^2 (1-\pi) \boldsymbol{c}(\tilde{\tau}, u) \left(\boldsymbol{l}_{i,0}^{(1)} - \boldsymbol{l}_{i,i-1}^{(1)} \right)^\top.
\end{aligned}$$

Now $\boldsymbol{l}_{i,0}^{(1)} - \boldsymbol{l}_{i,i-1}^{(1)} = (i-1)(\boldsymbol{e}^{\mathrm{R}} - \boldsymbol{e}^{\mathrm{I}})$, so, using (5.32),(5.34) and (5.35),

$$\boldsymbol{c}(\tilde{\tau}, u) \left(\boldsymbol{l}_{i,0}^{(1)} - \boldsymbol{l}_{i,i-1}^{(1)} \right)^\top = -(i-1) c_J(\tilde{\tau}, u),$$

where $c_J(\tilde{\tau}, u) = b(\tilde{\tau}) \mathrm{e}^{-(\tilde{\tau}-u)}$ (see (B.6)). Thus, using (5.54), $\bar{\sigma}_1^2 = \int_0^{\tilde{\tau}} \bar{g}(u) \, \mathrm{d}u$, where

$$\begin{aligned}
\bar{g}(u) &= \sum_{i=1}^{d_{\max}} \left[(1-\pi) \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,0}^{(1)})^\top \right)^2 + \pi \left(\bar{\boldsymbol{c}}(\tilde{\tau}, u)(\hat{\boldsymbol{l}}_{i,i-1}^{(1)})^\top \right)^2 \right] i \tilde{x}_i(u) \\
&= \pi^2 \tilde{g}(u) + \sum_{i=1}^{d_{\max}} \pi(1-\pi) \left[1 - 2(i-1)\pi c_J(\tilde{\tau}, u) \right] i \tilde{x}_i(u),
\end{aligned}$$

and $\tilde{g}(u)$ is given by (B.8) with $p_{i,i-1} = \pi = 1 - p_{i,0}$.

Now $\tilde{x}_i(0) = p_i$ $(i = 0, 1, \ldots, d_{\max})$, so (5.11) yields $\sum_{i=1}^{d_{\max}} i_{[k]} \tilde{x}_i(u) = \mathrm{e}^{-ku} f_D^{(k)}(\mathrm{e}^{-u})$ $(k = 1, 2)$. Hence,

$$\bar{\sigma}_1^2 = \pi^2 \sigma_1^2 + \pi(1-\pi) \int_0^{\tilde{\tau}} \mathrm{e}^{-u} f_D^{(1)}(u) \, \mathrm{d}u - 2\pi^2 (1-\pi) b(\tilde{\tau}) \mathrm{e}^{-\tilde{\tau}} \int_0^{\tilde{\tau}} \mathrm{e}^{-u} f_D^{(2)}(u) \, \mathrm{d}u.$$

Thus, recalling that $z = \mathrm{e}^{-\tilde{\tau}}$ and $\rho = 1 - f_D(z)$,

$$\int_0^{\tilde{\tau}} \mathrm{e}^{-u} f_D^{(1)}(u)\,\mathrm{d}u = f_D(1) - f_D(\mathrm{e}^{-\tilde{\tau}}) = \rho$$

and

$$\int_0^{\tilde{\tau}} \mathrm{e}^{-u} f_D^{(2)}(u)\,\mathrm{d}u = f_D^{(1)}(1) - f_D^{(1)}(\mathrm{e}^{-\tilde{\tau}}) = \mu_D - f_D^{(1)}(z).$$

Further, (2.10) implies $\pi[\mu_D - f_D^{(1)}(z)] = \mu_D(1-z)$, so

$$\sum_{i=1}^{3} \bar{\sigma}_i^2 = \pi^2 \sum_{i=1}^{3} \sigma_i^2 + \pi(1-\pi)[\rho - 2z(1-z)b(\tilde{\tau})\mu_D]$$

and (5.58) follows since $h(z) = b(\tilde{\tau})z$.

For the MR random graph, setting $\epsilon_i = 0$ $(i = 1, 2, \ldots, d_{\max})$ in (5.15), using (5.18) and recalling that $V^{(n)}(0) = 0$ for all n, shows that $\hat{\Sigma}_0 = 0$ and (5.59) follows. For the NSW random graph, a similar argument setting $\epsilon = 0$ in (5.22) and using (5.57) yields

$$\bar{\boldsymbol{c}}(\tilde{\tau}, 0)\bar{\Sigma}_0\bar{\boldsymbol{c}}(\tilde{\tau}, 0)^\top = (\boldsymbol{c}(\tilde{\tau}, 0) + \mathbf{1})\Sigma_{XX}(\boldsymbol{c}(\tilde{\tau}, 0) + \mathbf{1})^\top. \tag{B.43}$$

A simple calculation using (5.19) (or noting that Σ_{XX} is the variance matrix of a single multinomial trial) gives $\mathbf{1}\Sigma_{XX}\mathbf{1}^\top = 0$ and $\boldsymbol{c}(\tilde{\tau}, 0)\Sigma_{XX}\mathbf{1}^\top = 0$, so (5.59) follows from (B.43).

ACKNOWLEDGEMENTS

I thank Svante Janson and Etienne Pardoux for helpful correspondence, concerning simple graphs and density dependent population processes, respectively, and the referees for their careful reading of the paper and constructive comments, which have improved its presentation.

REFERENCES

Andersson, H. (1998) Limit theorems for a random graph epidemic model. *Ann. Appl. Prob.* **8**, 1331–1349.

Ball, F., Britton, T., Leung, K.Y. and Sirl, D. (2019) A stochastic SIR network epidemic model with preventive dropping of edges. *J. Math. Biol.*, **78**, 1875-1951.

Ball, F. and Neal, P. (2008) Network epidemic models with two levels of mixing. *Math. Biosci.* **212**, 69–87.

Ball, F. and Neal, P. (2017) The asymptotic variance of the giant component of configuration model random graphs. *Ann. Appl. Probab.* **27**, 1057–1092.

Ball, F. and Sirl, D. (2013) Acquaintance vaccination in an epidemic on a random graph with specified degree. *J. Appl. Prob.* **50**, 1147–1168.

Ball, F., Sirl, D. and Trapman, P. (2013) Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math. Biosci.* **224**, 53–73.

Barbour, A. D. and Luczak, M. J. (2012) Central limit approximations for Markov population processes with countably many types. *Electron. J. Probab.* **17**(90),1–16.

Barbour, A. and Reinert, G. (2013) Approximating the epidemic curve. *Electron. J. Probab.* **18**(54), 1–30.

Barbour, A.D. and Röllin, A. (2019) Central limit theorems in the configuration model. *Ann. Appl. Probab.* **29**, 1046–1069.

Bollobás, B. (1980) A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1**, 311-316.

Bohman, T. and Picollelli, M. (2012) SIR epidemics on random graphs with a fixed degree sequence. *Random Structures Algorithms* **41**, 179–214.

Britton, T., Janson, S. and Martin-Löf, A. (2007) Graphs with specified degree distributions, simple epidemics and local vaccination strategies. *Adv. Appl. Prob.* **39**, 922–948.

Britton, T. and Pardoux, E. (2019) Stochastic epidemics in a homogeneous community. *Stochastic Epidemic Models with Inference. Lecture Notes in Mathematics* **2255**, 1–120. Springer.

Britton, T. and Pardoux, E. (2020) Stochastic epidemics in a homogeneous community. *arXiv*:1808.05350v3.

Coupechoux, E. and Lelarge, M. (2014) How clustering affects epidemics in random networks. *Adv. Appl. Prob.* **46**, 985–1008.

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

Durrett, R. (2007) *Random Graph Dynamics*. Cambridge University Press, Cambridge.

Ethier, S.N. and Kurtz, T.G. (1986) *Markov Processes: Characterization and Convergence*. Wiley, New York.

Gleeson, J.P. (2009) Bond percolation on a class of clustered random networks. *Phys. Rev. E* **80**, 036107.

van der Hofstad, R. (2016) *Random Graphs and Complex Networks Volume 1*. Cambridge University Press, Cambridge.

Janson, S. (2009a) On percolation in random graphs with given vertex degrees. *Electron. J. Probab.* **14**, 87–118.

Janson, S. (2009b) The probability that a random multigraph is simple. *Combin. Prob. Comput.* **18**, 205–225.

Janson, S. (2010) Asymptotic equivalence and contiguity of some random graphs. *Random Structures Algorithms* **36**, 26–45.

Janson, S. (2014) The probability that a random multigraph is simple. II. *J. Appl. Prob.* **51A**, 123–137.

Janson, S. (2018) Asymptotic normality in random graphs with given vertex degrees. *arXiv*:1812.08063v2.

Janson, S. (2019) Random graphs with given vertex degrees and switchings. *arXiv*:1901.09744v2.

Janson, S., Luczak, M. and Windridge, P. (2014) Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures Algorithms* **45**, 726–763.

Kenah, E. and Robins, J.M. (2007) Second look at the spread of epidemics on networks. *Phys. Rev. E* **76**, 036113.

KhudaBukhsh, W. R., Woroszylo, C., Rempala, G. A. and Koeppl, H. (2017) Functional central limit theorem for susceptible-infected process on configuration model graphs.

arXiv:1703.06328v1.

Kiss, I. Z., Miller, J. C. and Simon, P. (2017) *Mathematics of Epidemics on Networks: From Exact to Approximate Models*. Springer.

Kurtz, T. G. (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* **7**, 49–58.

Kurtz, T. G. (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential equations. *J. Appl. Prob.* **8**, 344–356.

Martin-Löf, A. (1986) Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J. Appl. Prob.* **23**, 265–282.

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

Miller, J.C., Slim, A.C. and Volz, E.M. (2012) Edge-based compartmental modelling for infectious disease spread. *J. R. Soc. Interface* **9**, 890–906.

Molloy, M. and Reed, B. (1995) A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6**, 161–179.

Nerman, O. (1981) On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitsth* **57**, 365–395.

Newman, M.E.J. (2002) The spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128.

Newman, M.E.J., Strogratz, S.H. and Watts, D.J. (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118.

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

Trapman, P. (2007) On analytical approaches to epidemics on networks. *Theor. Popul. Biol.* **71**, 160–173.

Volz, E. M. (2008) SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.* **56**, 293–310.

Watson, R. (1980) A useful random time-scale transformation for the standard epidemic model. *J. Appl. Prob.* **17**, 324–332.

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

Frank Ball
School of Mathematical Sciences,
The University of Nottingham,
University Park,
Nottingham,
NG7 2RD,
UK
E-mail: frank.ball@nottingham.ac.uk