



Mathematical Statistics
Stockholm University

The duration of an *SIR* epidemic on a configuration model

Abid Ali Lashari
Ana Serafimović
Pieter Trapman

Research Report 2018:13

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

<http://www.math.su.se>



Mathematical Statistics
Stockholm University
Research Report **2018:13**,
<http://www.math.su.se>

The duration of an *SIR* epidemic on a configuration model

Abid Ali Lashari, Ana Serafimović and Pieter Trapman

May 2018

Abstract

We consider the spread of a stochastic *SIR* (Susceptible, Infectious, Recovered) epidemic on a configuration model random graph. We focus especially on the final stages of the outbreak and provide limit results for the duration of the entire epidemic, while we allow for non-exponential distributions of the infectious period and for both finite and infinite variance of the asymptotic degree distribution in the graph.

Our analysis relies on the analysis of some subcritical continuous time branching processes and on ideas from first-passage percolation.

Keywords: *SIR* epidemics; Time to extinction; Branching process approximation; First passage percolation.

1 Introduction

Mathematical models have been widely used to study the spread of infectious diseases and to design control strategies for reducing the impact of those diseases [11]. In many models, a fundamental assumption for the spread of so-called Susceptible-Infectious-Recovered (*SIR*) epidemics is that the individuals are uniformly mixing, i.e. all pairs of individuals in the population contact each other at the same rate, independently of each other. However, in real life, individuals do not mix homogeneously within populations. In order to gain some realism, a (social) network structure may be introduced to the models where contacts are only possible between “neighbours” (pairs of individuals that share a connection in the network; see e.g. [2, 21]). In this set-up, each vertex represents an individual and an edge represents that two individuals have a relationship that makes it possible for the disease to transmit from one to the other.

Much work has already been done for (variants) of epidemics on random graphs, e.g. by calculating the final size of the epidemic (the fraction of the population infected during the epidemic) and the probability of a large or major outbreak [10, 4]. In this paper we focus on the (random) duration of an epidemic on a configuration model graph. This duration is especially relevant for animal diseases. When outbreaks of those diseases occur, often trade bans are imposed on import from affected counties. So, longer durations of epidemics might lead to severe economic costs.

For homogeneously mixing populations Barbour [6] provides rigorous results on the duration of (Markov) *SIR* epidemics and Britton [10] also sketched some results about the duration of epidemic in a homogeneously mixing population. A corollary of their results is that the time until the epidemic goes extinct is $\Theta(\log n)$, where n is the population size and the “order notation” is discussed in Section 2.1.

We consider *SIR* epidemics (see Section 2.3) on configuration model graphs in the large population limit. Configuration model graphs are random graphs with specified vertex degrees (see Section 2.2, or for a detailed description see [12, 14]). In this graph each individual/vertex has his or her fixed degree (number of neighbours) with whom he/she can interact. The neighbouring vertices are chosen in such a way that the graph is uniform among all possible multigraphs with the given degree sequence.

We only consider major outbreaks of the epidemic, i.e. we assume that the number of ultimately infected individuals is of the same order as the number of individuals in the population. The beginning (until a small but non-negligible fraction of the population is infected) and the middle part (until a small but non-negligible fraction of the ultimately infected individuals

still has to be infected) of a major outbreak on a configuration model have been studied before (e.g. in [7, 25, 17]). Volz [25] studied a deterministic model for the spread of an *SIR* epidemic through a network using a set of differential equations, keeping track of the probability that a vertex of given degree avoids infection as a function of time. Barbour and Reinert [7] study (among other things) a stochastic model for the spread of *SIR* epidemics on a configuration model with bounded degrees and minor conditions on the infectious period distribution. The approach of the paper is tailored for finding the distribution of the time a typical individual in the population gets infected, but is not directly suitable for finding the time of the last infected individual recovering (the end of the epidemic). Janson et al. [17] study the spread of Markov *SIR* epidemics on quite general configuration models and their analysis heavily relies on the memoryless infectious period. As in [25] and [7], Janson et al. do not study the time until the end of the epidemic.

The spread of epidemics on random graphs can also be studied using first-passage percolation [9, 8, 1]. In first passage percolation i.i.d. weights (lengths) are assigned to edges in the graph and questions regarding distances between typical vertices in the graph can be answered. In epidemiological terms the distance between a uniformly chosen vertex and the initially infected vertex may be interpreted as the time of infection of that uniformly chosen vertex in an *SI* epidemic (i.e. an *SIR* epidemic with infinite infectious period). In this setting the question regarding the duration of an *SI* epidemic corresponds to the flooding time of the giant component of the random graph [1].

In the analysis of first passage percolation on random graphs in [9, 8] growing “balls” around vertices are explored and the time at which the balls touch provides precise results on the distance between the center vertices of those balls. These methods are very well suited for obtaining the asymptotic distribution of the distance between two vertices, but are less fit for finding flooding times and diameters (however, see [1]).

As written above, we focus on the duration of the whole epidemic, and in particular on the end of the epidemic. We allow for quite general infectious period distributions (see Theorem 2.1 below), and do not have to restrict ourselves to infinite infectious periods as is the case in the first passage percolation literature. Furthermore, we pose milder conditions on the degree distribution of the configuration model than Barbour and Reinert [7], who also allow for relatively general infectious period distributions. Our approach is to use the results of [7] (which are obtained through methods similar to those used in first passage percolation) to obtain the time until a typical vertex gets infected and then use subcritical branching processes to approximate the time between the infection of a typical vertex and the end of the

whole epidemic. We show that the duration of the epidemic divided by $\log n$ converges in probability to a (specified) constant. We note that our result is weaker in nature than the results of [7, 9, 8], where asymptotic distributions of infection times/distances of uniformly chosen vertices minus their typical distances are provided. However, as stated, we allow for more general distributions of the infectious period and degree distributions.

1.1 Outline of paper

The paper is structured as follows. In Section 2 we give a formal definition of the model and we provide the main theorem of this paper, which deals with the duration of the epidemic.

In Section 3 we discuss some techniques for analysing epidemics on graphs. In particular, we introduce so called “epidemic generated graphs”. Furthermore, we summarise results on continuous time branching processes that we need in the paper.

In Section 4 heuristics are given for the main theorem, while in Sections 5 and 6 this theorem is proved rigorously. In this proof the durations of the initial and final phase of the epidemic are analysed separately.

2 Definitions, notation and main results

2.1 Basic notation

The following basic notation and definitions are used throughout this paper (see also e.g. [18, Section 1.2]). For real-valued functions, f and g and $x \rightarrow \infty$ we say,

$$\begin{aligned} f(x) &= O(g(x)) && \text{if } \limsup f(x)/g(x) < \infty, \\ f(x) &= o(g(x)) && \text{if } \lim f(x)/g(x) = 0, \\ f(x) &= \Theta(g(x)) && \text{if } 0 < \liminf |f(x)/g(x)| \leq \limsup |f(x)/g(x)| < \infty. \end{aligned}$$

All random processes and random variables that we consider are defined on a rich enough probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which we do not further specify. The population size is always denoted by n . In this paper, asymptotic results and limits are for $n \rightarrow \infty$, unless explicitly stated otherwise. We say that an event occurs with high probability (w.h.p.) if the probability of the event converges to 1. Furthermore, $\xrightarrow{a.s.}$ denotes almost sure convergence, $\xrightarrow{\mathbb{P}}$ denotes convergence in probability, and \xrightarrow{d} denotes convergence in distribution.

Throughout, the cardinalities of a set \mathcal{X} is denoted by $|\mathcal{X}|$.

2.2 Construction of the random graph and assumptions on the degree distribution

The epidemic spreads on a random graph $G^{(n)} = (V^{(n)}, E^{(n)})$. The set $V^{(n)}$ consists of n vertices that represent the individuals, and the edge set $E^{(n)}$ represent connections/relationships of individuals through which infection might transmit. For $v \in V^{(n)}$, the degree of vertex v (i.e. the number of edges adjacent to vertex v) is denoted by $d_v^{(n)}$. $G^{(n)}$ is generated through a configuration model with given degree sequence $\{d_v^{(n)}\}_{v \in V^{(n)}}$.

The graph is constructed by assigning $d_v^{(n)}$ half-edges (edges with only one endpoint assigned to a vertex) to the vertex v for $v \in V^{(n)}$ and pairing those half-edges uniformly at random. By this construction every vertex has the right degree, although it is possible that there is more than one edge between a pair of vertices (parallel edges) or that an edge connects a vertex to itself (a self-loop). In the graph, parallel edges are counted separately in the degree and a self loop adds two to the degree of a vertex.

Define

$$\ell(n) = \sum_{v \in V^{(n)}} d_v^{(n)} \quad \text{and} \quad \ell_2(n) = \sum_{v \in V^{(n)}} (d_v^{(n)})^2. \quad (1)$$

We assume that there exists a random variable D such that as $n \rightarrow \infty$,

$$(A1) \quad n^{-1} \sum_{v \in V^{(n)}} \mathbf{1}(d_v^{(n)} = k) \rightarrow p_k = \mathbb{P}(D = k),$$

$$(A2) \quad n^{-1} \ell(n) \rightarrow \mathbb{E}[D] < \infty,$$

$$(A3) \quad n^{-1} \ell_2(n) \rightarrow \mathbb{E}[D^2] \in (0, \infty].$$

Observe that $\ell(n)$ is even, since every edge in $E^{(n)}$ adds 2 to the total degree of the graph. We note that if the degrees of vertices are i.i.d. and distributed as the random variable D , then the above assumptions are satisfied w.h.p. The ‘‘size biased’’ random variable \tilde{D} is defined through

$$\mathbb{P}(\tilde{D} = k) = \tilde{p}_k = \frac{k p_k}{\mathbb{E}[D]}.$$

Let $D^{(n)}$ be a random variable with the same distribution as the degree of a vertex chosen uniformly at random from the graph. That is

$$\mathbb{P}(D^{(n)} = k) = n^{-1} \sum_{v \in V^{(n)}} \mathbf{1}(d_v^{(n)} = k) \quad \text{for } k \in \mathbb{N}_{\geq 0}.$$

By (A1) and (A2), $D^{(n)} \xrightarrow{d} D$ and $\mathbb{E}[D^{(n)}] \xrightarrow{d} \mathbb{E}[D]$. Similarly, $\tilde{D}^{(n)}$ is the size biased variant of $D^{(n)}$, i.e.

$$\mathbb{P}(\tilde{D}^{(n)} = k) = \frac{k \mathbb{P}(D^{(n)} = k)}{\mathbb{E}[D^{(n)}]} = \frac{k \sum_{v \in V^{(n)}} \mathbf{1}(d_v^{(n)} = k)}{\ell(n)} \quad \text{for } k \in \mathbb{N}_{\geq 0}.$$

Note that $\tilde{D}^{(n)}$ is distributed as the degree of an end vertex of a uniformly chosen edge from the graph. By (A1) and (A2), $\tilde{D}^{(n)} \xrightarrow{d} \tilde{D}$, while by (A3), $\mathbb{E}[\tilde{D}^{(n)}] = \frac{\ell_2(n)}{\ell(n)} \xrightarrow{d} \mathbb{E}[D^2]/\mathbb{E}[D] = \mathbb{E}[\tilde{D}]$. Note that $\mathbb{E}[\tilde{D}]$ may be infinite.

For the epidemic process on the graph, we merge parallel edges and ignore self-loops. Because $\mathbb{E}(D) < \infty$, this assumption has no impact on the asymptotic degree distribution although the number of self-loops and parallel edges diverges if $\text{Var}(D) = \infty$ [14, p. 219].

2.3 The *SIR* epidemic

We consider an *SIR* (Susceptible \rightarrow Infectious \rightarrow Recovered) epidemic on $G^{(n)}$. We say that a vertex is susceptible, infectious or recovered if the individual it represents is in this “infection state”. Neighbours in the population contact each other according to independent homogeneous Poisson processes with rate β , and if the contact is between a susceptible and an infectious vertex, then the susceptible one becomes immediately infectious itself. Infectious vertices stay so for a random period distributed as the random variable L . All infectious periods and Poisson processes are independent of each other. A contact by an infectious vertex is called an *infectious contact*, whether or not the “contactee” is susceptible. Throughout we assume that at time 0, there is one infectious individual chosen uniformly at random from the population and all other individuals are susceptible. It is straightforward to extend the model to other initial conditions.

Let $\phi_L(\beta) = \int_0^\infty e^{-\beta t} L(dt)$ be the Laplace transform of L . The probability, ψ say, that an infected vertex makes a contact with a given neighbour during its infectious period (and infect it if it is still susceptible) is given by

$$\psi = \int_0^\infty \beta e^{-\beta t} \mathbb{P}(L > t) dt = 1 - \int_0^\infty \beta e^{-\beta t} \mathbb{P}(L \leq t) dt = 1 - \phi_L(\beta), \quad (2)$$

where the last identity is obtained through partial integration.

We denote the sets of susceptible, infectious and recovered individuals at time t by $S(t) = S^{(n)}(t)$, $I(t) = I^{(n)}(t)$ and $R(t) = R^{(n)}(t)$ respectively. If no confusion is possible, we sometimes suppress the (n) superscript. We say that the epidemic goes extinct or ends before time t if $|I^{(n)}(t)| = 0$.

Throughout we use continuous time branching processes [16] to approximate the epidemic processes. We consider those processes for which there exists a number α (called Malthusian parameter, or real-time growth rate) which satisfies

$$\int_0^\infty e^{-\alpha t} \mu(dt) = 1, \quad (3)$$

where $\mu(dt)$ denotes the mean measure for births of children of a particle. Below we define and justify a branching process approximation of the *SIR*

epidemic model. The approximating branching process has mean offspring measure

$$\mu'(dt) = \mathbb{E}[\tilde{D} - 1] \beta e^{-\beta t} \mathbb{P}(L > t) dt. \quad (4)$$

Note that (using (2) for the last equality),

$$R_0 := \mu'(\infty) = \int_0^\infty \mu'(dt) = \psi \mathbb{E}(\tilde{D} - 1) \quad (5)$$

corresponds to the expected number of offspring of a particle during all of its life. If $R_0 > 1$ the epidemic is supercritical and α exists and is strictly positive. If on the other hand $R_0 < 1$, the process is subcritical and α might exist and if it does, α is strictly negative. If $R_0 = 1$ the epidemic is critical and the corresponding α trivially equals 0.

In epidemic literature R_0 is arguably the most studied quantity (e.g. [11]). It is usually defined as the average number of secondary infections caused by a typical infected individual in the early stages of an epidemic in a further susceptible population. This definition is consistent with equation (5).

2.4 The main results

In this subsection we state the main results of the paper. The proofs will be provided in Sections 5 and 6. We consider an *SIR* epidemic on the configuration model graph $G^{(n)} = (V^{(n)}, E^{(n)})$ with degrees satisfying assumptions (A1)-(A3). The infectious periods are distributed as L , and neighbours contact each other according to independent Poisson processes with intensity β .

Define the time until extinction of an epidemic in a population of size n by

$$T^* = T^*(n) = \inf\{t \geq 0; |I^{(n)}(t)| = 0\}. \quad (6)$$

For $\mathbb{E}[\tilde{D} - 1] < \infty$ and $R_0 > 1$, let α' be the solution of

$$1 = \int_0^\infty e^{-\alpha' t} \mathbb{E}[\tilde{D} - 1] \beta e^{-\beta t} \mathbb{P}(L > t) dt = \int_0^\infty e^{-\alpha' t} \mu'(dt), \quad (7)$$

where $\mu'(dt)$ is defined in (4). If $\mathbb{E}[\tilde{D} - 1] = \infty$, we set $\alpha' = \infty$.

If the following equation has a solution, then α^* is the solution of

$$1 = \int_0^\infty e^{-\alpha^* t} \mu^*(dt), \quad (8)$$

where

$$\mu^*(dt) = \mathbb{E} \left[(\tilde{D} - 1) (1 - \psi + \psi \tilde{q}^*)^{\tilde{D}-2} \right] \beta e^{-\beta t} \mathbb{P}(L > t) dt, \quad (9)$$

and \tilde{q}^* is the smallest positive solution of

$$s = \mathbb{E} \left[(1 - \psi + \psi s)^{\tilde{D}-1} \right]. \quad (10)$$

By standard theory on supercritical branching processes [16], we obtain $\tilde{q}^* \in (0, 1)$, (because it can be interpreted as the extinction probability of a supercritical branching process [4]). By Lemma 2.1 below the branching process defined through $\mu^*(dt)$ is subcritical. A sufficient (but not necessary) condition for $\alpha^* > -\infty$ to exist is $\mathbb{P}(L > t_0) = 0$ for some $t_0 > 0$ and $\mathbb{E}[\tilde{D} - 1] < \infty$.

Before stating the main theorem, we provide the following Lemma, the proof of which is provided in Section 6. Recall R_0 is defined in (5).

Lemma 2.1. *Assume $R_0 > 1$. If equation (8) has a solution, then the solution α^* is strictly negative.*

The main theorem then reads.

Theorem 2.1. *Conditioned on a large outbreak, and assuming that*

- (i) *there exist $c > 1$ and $\alpha_c < 0$ such that $c = \int_0^\infty e^{-\alpha_c t} \mu^*(dt)$,*
 - (ii) *for α^* as in (8), $\int_0^\infty e^{(|\alpha^*|+\eta)t} L(dt) < \infty$ for some $\eta > 0$ and*
 - (iii) *$\mathbb{P}(L > t + s) < e^{-|\alpha^*|(s-s_0)} \mathbb{P}(L > t)$ for all $s, t > 0$ and some $s_0 > 0$,*
- we have that $\frac{T^*(n)}{\log n} \xrightarrow{\mathbb{P}} \frac{1}{\alpha'} + \frac{1}{|\alpha^*|}$.*

Remark. *We believe that the assumptions of Theorem 2.1 can be replaced by assuming that α^* exists and $\int_0^\infty t e^{|\alpha^*|t} L(dt) < \infty$ (see Lemma 3.2 below). However, in our proofs we use the conditions (i)-(iii), and we believe that the current assumptions are a small price to pay for the availability of the proof techniques we provide in the paper.*

Remark. *Intuition from first passage percolation (e.g. [9, 8]) and research on the epidemic curve [17, 7] suggests that (possibly with some extra conditions on the distributions of the infectious period and degrees) $T^*(n) - \left(\frac{1}{\alpha'} + \frac{1}{|\alpha^*|}\right) \log n$ might converge in distribution to an a.s. finite random variable. We did not try to prove this or identify which extra conditions would be necessary for such a proof.*

In the theorem the summand $\frac{1}{\alpha'}$ is related to the duration of the early stage, i.e. the exponentially growing phase, of the epidemic, while the summand $\frac{1}{|\alpha^*|}$ is related to the duration of the final phase, i.e. the exponentially declining phase, of the epidemic. The condition $\int_0^\infty e^{(|\alpha^*|+\eta)t} L(dt) < \infty$ for some $\eta > 0$ guarantees that none of the individuals infected during the epidemic will stay infectious for a time longer than $\log[n]/|\alpha^*|$. This condition

is satisfied if $\mathbb{P}(L > t)$ decays faster than exponential, but is not satisfied if $\mathbb{P}(L > t)$ decays slower than exponential.

In order to prove Theorem 2.1 we use some lemmas. Let

$$q^* = \mathbb{E}[(1 - \psi + \psi \tilde{q}^*)^D], \quad (11)$$

where \tilde{q}^* is defined through (10). Copying the steps of the corresponding result for random intersection graphs as provided in [5], we obtain

Lemma 2.2. *Conditioned on a major outbreak, $n^{-1}|S^{(n)}(\infty)| \xrightarrow{\mathbb{P}} q^*$.*

In order to formulate the main lemmas, define for $\gamma \in (0, 1 - q^*)$

$$T'_\gamma = T'_\gamma(n) := \inf\{t > 0; n^{-1}|S^{(n)}(t)| < (1 - \gamma)\}. \quad (12)$$

Theorem 2.1 now follows trivially from the following lemmas, where the first is about the duration of the initial phase of the epidemic and the second about the duration of the final phase.

Lemma 2.3. *Conditioned on a major outbreak and for all $\gamma \in (0, 1 - q^*)$, we have $\frac{T'_\gamma(n)}{\log n} \xrightarrow{\mathbb{P}} \frac{1}{\alpha'}$.*

Lemma 2.4. *Assume that the assumptions of Theorem 2.1 are met. Conditioned on a major outbreak, there exists $\gamma \in (0, 1 - q^*)$, such that*

$$\frac{T^*(n) - T'_\gamma(n)}{\log n} \xrightarrow{\mathbb{P}} \frac{1}{|\alpha^*|}.$$

Note that Lemma 2.3 implies that Lemma 2.4 actually holds for all $\gamma \in (0, 1 - q^*)$.

3 The epidemic on the graph

3.1 The epidemic generated graph

3.1.1 Susceptibility sets and the epidemic generated graph

In this section, we construct G^{epi} , the so-called epidemic generated graph (see e.g. [19, 22, 5]). In the random graph G , we replace all undirected edges by two directed ones, pointing in opposite directions. For $u \in V^{(n)}$, we assign to u a value L_u , where L_u is distributed as L . L_u may be interpreted as the infectious period of u . Furthermore, we assign an exponentially distributed random variables $\{\tau_{uv}; (u, v) \in E^{(n)}\}$ with expectation $1/\beta$ to the directed edges in $E^{(n)}$. τ_{uv} may be interpreted as the time between the infection of u

(if ever) and the first contact after the infection of u between u and v . All random variables defined in this paragraph are independent.

In order to obtain G^{epi} , we thin G by keeping all directed edges $(u, v) \in E^{(n)}$ satisfying $\tau_{uv} < L_u$ and delete all others. This implies that we delete an edge emanating from vertex u with probability $e^{-\beta L_u}$. Deletion of the edge from u to v means that if u gets infected, it will not succeed in passing the infection to vertex v during the time it is infectious. G^{epi} is a directed graph and the vertices that can be reached starting from the initially infectious vertex correspond to the ultimately recovered vertices.

Recall that the initial infective vertex in the epidemic is chosen uniformly at random from the population. So, we pick a vertex v_0 uniformly at random as a starting vertex for exploring G^{epi} . The vertices towards which v_0 has a directed edge in G^{epi} are the first generation vertices in the exploration process. The vertices that can be reached by an edge from the first generation vertex are the second generation vertices (if they are not in an earlier generation) and so on. It can be shown (see e.g. [4]) that the set of vertices that can be reached from v_0 is with high probability either $\Theta(1)$ or $\Theta(n)$. Because it is unlikely that there are short circuits in this exploration process, the exploration process is well approximated by a branching process, which is described in Section 3.2.

For future reference we define the susceptibility set $\mathcal{S}(u)$ of vertex u . The susceptibility set of u is the set of all vertices such that if they were initially infected then u would be ultimately recovered [3] i.e. $v \in \mathcal{S}(u)$ if and only if there is a path from v to u in G^{epi} . In the large population limit, the initial infected individual v_0 is in the susceptibility set of u w.h.p. if both the susceptibility set of u and the forward exploration set of v_0 are large. If one of those sets is small then v_0 is w.h.p not in $\mathcal{S}(u)$.

The susceptibility set can be described by a discrete time process (for detail see [5]), which is constructed by using the epidemic generated graph in the following way. Let u be a vertex chosen uniformly at random from the n vertices in the population. The degree of u is then distributed as $D^{(n)}$. The probability that a given neighbour of u has an edge to u in the epidemic generated graph is ψ (defined as in (2)) and the events that neighbours of u have an edge to u in the epidemic generated graph are independent for different neighbours. Those neighbours with an edge to u are the first generation of the backward process. The vertices in the first generation of the backward process have (by the construction of the configuration model) degrees asymptotically distributed as $\tilde{D}^{(n)}$, where one of his or her neighbours is u . The other neighbours are with high probability all different and not in the first generation of the backward process. We continue in this fashion: vertices with an edge towards k -th generation vertices in the epidemic generated

graph, which are not in any of the generations up to and including k are in generation $k + 1$. This generation view of the susceptibility set of u allows us to approximate the backward process by a Galton-Watson process of which the matriarch (the first ancestor) has a different offspring distribution than all other individuals in the process [16]. Below we use that $D^{(n)} \xrightarrow{d} D$ and $\tilde{D}^{(n)} \xrightarrow{d} \tilde{D}$. For the approximating Galton-Watson process the probability of extinction is therefore given by equation (11).

This result can be obtained in a similar fashion as the results on the backward branching process in [4, 5].

3.1.2 Construction of the epidemic generated graph

For the proof of the main theorem we rely on the following explicit step-by-step construction of the epidemic generated graph, G^{epi} , or more precisely, of the cluster of vertices in G^{epi} reachable from the initially infectious vertex. Label the vertices in $V^{(n)}$ by $1, 2, \dots, n$, such that

$$\mathbf{d}^{(n)} = d_1^{(n)}, \dots, d_n^{(n)} = d_1, \dots, d_n$$

is a non-decreasing degree sequence satisfying assumptions (A1)-(A3). Let

$$\mathbf{s} = \mathbf{s}^{(n)} = \{(1, 1), (1, 2), \dots, (1, d_1), (2, 1), \dots, (2, d_2), \dots, (n, 1), \dots, (n, d_n)\}$$

be the set containing $\ell(n)$ elements, corresponding to the half-edges and let

$$\mathbf{x} = \mathbf{x}^{(n)} = (x_1, x'_1), (x_2, x'_2), \dots$$

be an infinite sequence of (2 dimensional) elements of $\mathbf{s}^{(n)}$, where the elements are chosen independently and uniformly at random. Let x_0 be the initially infected vertex, which is chosen uniformly at random from the population. Furthermore, let L_v ($v \in \{1, 2, \dots, n\}$) be the infectious period of vertex v , if this vertex becomes infective during the epidemic.

We use the following process of partitions of the set of half edges.

$$\{\mathcal{E}^{(n)}(t); t \geq 0\} = \{\mathcal{E}(t); t \geq 0\} = \{(\mathcal{E}_1^{(n)}(t), \mathcal{E}_2^{(n)}(t), \mathcal{E}_3^{(n)}(t), \mathcal{E}_4^{(n)}(t); t \geq 0\}.$$

In this process at time t , $\mathcal{E}_1^{(n)}(t)$ is the set of unpaired half-edges belonging to susceptible vertices, $\mathcal{E}_2^{(n)}(t)$ is the set of unpaired half-edges belonging to infectious vertices, $\mathcal{E}_3^{(n)}(t)$ is the set of unpaired half-edges belonging to recovered vertices and $\mathcal{E}_4^{(n)}(t)$ is the set of paired half edges. Let $\sigma(v)$ be the time the first half-edge belonging to vertex v is added to $\{\mathcal{E}_4^{(n)}(t); t \geq 0\}$. For $t \geq 0$, let $\mathcal{V}^{(n)}(t)$ be the set of vertices to which at least one of the half-edges in $\mathcal{E}_2^{(n)}(t) \cup \mathcal{E}_3^{(n)}(t) \cup \mathcal{E}_4^{(n)}(t)$ belongs. So, $\mathcal{V}^{(n)}(t)$ corresponds to the set of

vertices which is no longer susceptible at time t , i.e. $\mathcal{V}^{(n)}(t) = V^{(n)}(t) \setminus S^{(n)}(t)$. Throughout the process the sequence $\mathbf{x}^{(n)}$ is also explored element by element and $\mathbf{x}^{(n)}(t)$ is the set of elements of $\mathbf{x}^{(n)}$ explored before or at time t .

The construction of $\{\mathcal{E}^{(n)}(t); t \geq 0\}$ is as follows.

- Start of construction: Choose a vertex (say x_0) uniformly at random. This vertex is the initial infected vertex. Set $\sigma(x_0) = 0$. Note that x_0 has degree d_{x_0} in G .

At time $t = 0$, $\mathcal{E}_2^{(n)}(0) = \{(v, v') \in \mathbf{s}^{(n)}; v = x_0\}$ consists of all half edges attached to x_0 , while all other half edges are in $\mathcal{E}_1^{(n)}(0) = \{(v, v') \in \mathbf{s}^{(n)}; v \neq x_0\}$. None of the elements of $\mathbf{x}^{(n)}$ are explored yet at time 0, i.e. $\mathbf{x}^{(n)}(0) = \emptyset$.

- Assume that at time t , $\mathbf{x}^{(n)}$ is explored up to and including (x_k, x'_k) , i.e. $\mathbf{x}^{(n)}(t) = \{(x_1, x'_1), \dots, (x_k, x'_k)\}$. Define

$$t_+(t) = \min(\{\sigma(v) + L(v); (v, v') \in \mathcal{E}_2^{(n)}(t)\} \cup \{\sigma(v) + \tau_{v, v'}; (v, v') \in \mathcal{E}_2^{(n)}(t)\}),$$

which can be interpreted as the first time after time t something changes in the process. In the interval $[t, t_+(t))$ the process $\mathcal{E}^{(n)}(t)$ is constant, while if $t_+(t) = \sigma(u) + L(u)$ then all $\{(v, v') \in \mathcal{E}_2^{(n)}(t); v = u\}$ are in $\mathcal{E}_3^{(n)}(t_+(t))$. If $t_+(t) = \sigma(u) + \tau_{u, u'}$, then $(u, u') \in \mathcal{E}_4^{(n)}(t_+(t))$. In addition, consider (x_{k+1}, x'_{k+1}) , which is the half-edge (u, u') “wants to” be paired with if it is still possible. The half edge (x_{k+1}, x'_{k+1}) is considered explored from $t_+(t)$ on, i.e. $(x_{k+1}, x'_{k+1}) \in \mathbf{x}^{(n)}(t_+(t))$. We distinguish between the following cases for further changes in $\mathcal{E}^{(n)}(t)$ at time $t_+(t) = \sigma(u) + \tau_{u, u'}$.

- If $(x_{k+1}, x'_{k+1}) \in \mathcal{E}_1^{(n)}(t)$, then $(x_{k+1}, x'_{k+1}) \in \mathcal{E}_4^{(n)}(t_+(t))$, while all $d_{x_{k+1}} - 1$ other half edges belonging to x_{k+1} (which necessarily belong to $\mathcal{E}_1^{(n)}(t)$) move to $\mathcal{E}_2^{(n)}(t_+(t))$. Furthermore, $\sigma(x_{k+1}) = t_+(t)$.
 - If $(x_{k+1}, x'_{k+1}) \in \mathcal{E}_2^{(n)}(t) \cup \mathcal{E}_3^{(n)}(t)$, then $(x_{k+1}, x'_{k+1}) \in \mathcal{E}_4^{(n)}(t_+(t))$, while none of the other half edges change group.
 - If $(x_{k+1}, x'_{k+1}) \in \mathcal{E}_4^{(n)}(t)$, then take the above steps with (x_{k+1}, x'_{k+1}) replaced by (x_{k+2}, x'_{k+2}) and so on (while treating all considered half-edges as explored).
- Continue the above construction until $\mathcal{E}_2^{(n)}(t) = \emptyset$. That is, until there are no unpaired half-edges left which belong to infectious vertices.

3.2 Branching processes theory background

Throughout the manuscript we use several continuous time branching processes. In this section we summarise some of the results we use in the analysis of the duration of the epidemic. The branching processes are two stage branching processes in the sense that the reproduction law for the matriarch (the first ancestor) is different from that of the other particles in the process. In the exposition below we use a single stage branching process, but extending the results to two stage branching processes is straightforward. For further theory we refer to [16, Chapter 6] and [14, Chapter 3].

Assume that particles give birth to other particles according to a random point process distributed as $\{\xi(t); t \geq 0\}$. Define $\mu(t) = \mathbb{E}[\xi(t)]$. If $\mu(\infty) > 1$ then equation (3) has a strictly positive solution α , which is called the Malthusian parameter of the process. We call a process supercritical if $\mu(\infty) > 1$, critical if $\mu(\infty) = 1$ and subcritical if $\mu(\infty) < 1$.

Let $Z(t)$ be the number of particles in the branching process at time t and $Z^{tot}(t)$ the number of particles born in the branching process up to time t . Furthermore, let $Z(t; a)$ be the number of particles alive at time t , born after time $t - a$. The following Lemma follows immediately from Theorems 2.1 and 2.4 of [15] and Theorem 5.4 of [20].

Lemma 3.1. *Assume $\mu(\infty) > 1$ and let α be the Malthusian parameter defined in (3). Furthermore, for $\log^+ t := \max(0, \log t)$, if there exist $\epsilon > 0$ such that $\int_0^\infty t(\log^+ t)^{1+\epsilon} e^{-\alpha t} \mu(dt) < \infty$, then almost surely and in expectation*

$$e^{-\alpha t} Z(t) \rightarrow W \quad \text{and} \quad e^{-\alpha t} Z^{tot}(t) \rightarrow W' \quad \text{as } t \rightarrow \infty, \quad (13)$$

where W and W' are a.s. finite random variables satisfying

$$\mathbb{P}(W > 0) = \mathbb{P}(W' > 0) = \mathbb{P}(Z(t) \not\rightarrow 0, \text{ for } t \rightarrow \infty).$$

If in addition

$$\mathbb{E} \left[\int_0^\infty e^{-\alpha t} \xi(dt) \log^+ \left(\int_0^\infty e^{-\alpha t} \xi(dt) \right) \right] < \infty,$$

then a.s. on $\{Z(t) \rightarrow \infty\}$ we have

$$\frac{Z(t; a)}{Z(t)} \rightarrow \frac{\int_0^a \mathbb{P}(L > u) e^{-\alpha u} du}{\int_0^\infty \mathbb{P}(L > u) e^{-\alpha u} du}, \quad \text{as } t \rightarrow \infty. \quad (14)$$

We need the following immediate Corollary of this Lemma.

Corollary 3.1. *Assume that the conditions of Lemma 3.1 hold. For $k \in \mathbb{N}$, define $\hat{T}_k = \inf\{t \geq 0; Z(t) \geq k\}$ and $\hat{T}'_k = \inf\{t \geq 0; Z^{\text{tot}}(t) \geq k\}$. Then a.s. on $\{Z(t) \rightarrow \infty \text{ as } t \rightarrow \infty\}$ we have that*

$$\frac{\hat{T}_k}{\log k} \rightarrow \frac{1}{\alpha} \quad \text{and} \quad \frac{\hat{T}'_k}{\log k} \rightarrow \frac{1}{\alpha} \quad \text{as } k \rightarrow \infty. \quad (15)$$

To approximate the end of an epidemic we use a subcritical branching process. For these branching processes equation (3) does not necessarily have a solution. However if it has, then we may obtain some useful results. First note that $\alpha < 0$. Let the life-length of particles be distributed as L . From Theorem 6.2 of [16], we immediately obtain

Lemma 3.2. *Let $\mu(\infty) < 1$ and $Z(0) = 1$. Assume (i) equation (3) has a solution, (ii) $\int_0^\infty te^{|\alpha|t}L(dt) < \infty$, (iii) $\int_0^\infty te^{|\alpha|t}\mu(dt) < \infty$ and (iv) $\mathbb{E}\left[\int_0^\infty e^{|\alpha|t}\xi(dt)\log^+(\xi(\infty))\right] < \infty$, then $e^{|\alpha|t}\mathbb{P}[Z(t) > 0]$ converges to a strictly positive and finite limit.*

Below we use the following Corollary of this Lemma.

Corollary 3.2. *Assume that the conditions of Lemma 3.2 hold. For $k \in \mathbb{N}$, define $\hat{T}_k^* = \inf\{t \geq 0; Z(t) = 0\}$ conditioned on $Z(0) = k$. Then, $\frac{\hat{T}_k^*}{\log k} \xrightarrow{\mathbb{P}} \frac{1}{|\alpha|}$ for $k \rightarrow \infty$.*

Proof. It is enough to prove that for every $\delta > 0$,

$$\mathbb{P}\left(\hat{T}_k^* \leq \frac{1+\delta}{|\alpha|} \log k\right) = \mathbb{P}\left(Z\left(\frac{1+\delta}{|\alpha|} \log k\right) = 0 \mid Z(0) = k\right) \rightarrow 1$$

and

$$\mathbb{P}\left(\hat{T}_k^* \leq \frac{1-\delta}{|\alpha|} \log k\right) = \mathbb{P}\left(Z\left(\frac{1-\delta}{|\alpha|} \log k\right) = 0 \mid Z(0) = k\right) \rightarrow 0$$

as $k \rightarrow \infty$. Note that $\{Z(t); t \geq 0\}$ is distributed as $\{\sum_{j=1}^k Z_j(t); t \geq 0\}$, where the $\{Z_j(t); t \geq 0\}$ are independent branching processes distributed as the subcritical branching process conditioned on $k = 1$. Therefore,

$$\{Z(t) = 0\} = \cap_{j=1}^k \{Z_j(t) = 0\}$$

and we obtain that

$$\mathbb{P}(Z(t) = 0 \mid Z(0) = k) = (\mathbb{P}(Z(t) = 0 \mid Z(0) = 1))^k.$$

So,

$$\begin{aligned} \mathbb{P}\left(Z\left(\frac{1+\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = k\right) \\ = \left(\mathbb{P}\left(Z\left(\frac{1+\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = 1\right)\right)^k. \end{aligned}$$

By Lemma 3.2 we know that there exists $t_0 > 0$ such that for all $t > t_0$ we have both

$$\mathbb{P}(Z(t) > 0 \mid Z(0) = 1) < e^{-|\alpha|(1-\delta/2)t} \quad \text{and} \quad \mathbb{P}(Z(t) > 0 \mid Z(0) = 1) > e^{-|\alpha|(1+\delta)t}.$$

So, we obtain

$$\begin{aligned} \mathbb{P}\left(Z\left(\frac{1+\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = k\right) &= \left(\mathbb{P}\left(Z\left(\frac{1+\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = 1\right)\right)^k \\ &> \left(1 - e^{-(1-\delta/2)(1+\delta)\log k}\right)^k = \left(1 - k^{-(1+\delta/2-\delta^2/2)}\right)^k = \left(1 - \frac{k^{-(\delta-\delta^2)/2}}{k}\right)^k, \end{aligned}$$

which converges to 1 for $\delta < 1$, by $(1 - ck^{-1})^k \rightarrow e^{-c}$ as $k \rightarrow \infty$. Similarly,

$$\begin{aligned} \mathbb{P}\left(Z\left(\frac{1-\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = k\right) &= \left(\mathbb{P}\left(Z\left(\frac{1-\delta}{|\alpha|}\log k\right) = 0 \mid Z(0) = 1\right)\right)^k \\ &< \left(1 - e^{-(1+\delta)(1-\delta)\log k}\right)^k = \left(1 - k^{-(1-\delta^2)}\right)^k = \left(1 - \frac{k^{\delta^2}}{k}\right)^k, \end{aligned}$$

which converges to 0, by $(1 - ck^{-1})^k \rightarrow e^{-c}$ as $k \rightarrow \infty$ and the proof of the Corollary is complete. \square

4 Heuristics

In this subsection we provide some heuristic arguments for Theorem 2.1. If a large outbreak occurs, the epidemic can be subdivided into three phases, which can be roughly described as follows. Let $\epsilon > 0$ be small. In the initial phase the number of susceptible vertices decreases from $n - 1$ to $(1 - \epsilon)n$. In the intermediate phase the number of susceptible vertices decreases from $(1 - \epsilon)n$ to $(q^* + \epsilon)n$. While the final stage of the epidemic last from the moment that the number of susceptible vertices is $(q^* + \epsilon)n$ until there are no more infectious vertices in the population.

4.1 The initial and intermediate phase of the epidemic

The primary intuition for the initial phase is that the number of infectious vertices at time t (i.e. $|I(t)|$) and the number of vertices infected before time t (i.e. $|I(t)| + |R(t)|$) are “well approximated” by a branching process with mean measure given by (4) as long as $n^{-1}|S(t)| > 1 - \epsilon$ for $\epsilon > 0$ but small. The result of Lemma 2.3 then follows by applying Corollary 3.1 to $k = \epsilon n$.

To justify the use of (4), assume that the degree of a vertex uniformly taken from the population of size n has exactly the same distribution function as D , then a newly infected vertex has degree distribution \tilde{D} , because of size biasing (see e.g. [12]). Apart from one (the infector) all of the neighbours of this newly infected vertex are susceptible with high probability. A newly infected vertex stays infectious for a random time L . Neighbours contacts each other with intensity β , and if the contact is between a susceptible and an infectious vertex then the susceptible one becomes infected, which can be interpreted as being a child of his or her infector in the approximating branching process. So in an approximating branching process we obtain expression (4):

$$\mu'(dt) = \mathbb{E}[\tilde{D} - 1] \beta e^{-\beta t} \mathbb{P}(L > t) dt,$$

where $\mathbb{E}[\tilde{D} - 1]$ is the expected number of susceptible neighbours of a newly infected vertex, $\beta e^{-\beta t}$ is the density of the time since infection of the first contact with a given neighbour, while $\mathbb{P}(L > t)$ is the probability that the vertex is still infectious at this time of first contact. The Malthusian parameter of this approximating branching process is therefore given by (7).

In the intermediate phase of the epidemic, $|S(t)|$, $|I(t)|$, and the number of infectious-susceptible neighbour pairs are all $\Theta(n)$. This implies that changes in $n^{-1}|S(t)|$, occur at an $\Theta(1)$ rate and the intermediate phase has duration $\Theta(1)$.

Our proof of Lemma 2.3, however makes use of the fact that the initial and intermediate phase of the epidemic are, with some extra conditions on D and L , studied in [7]. In [7] the (random) evolution of $|S(T_0 + (\alpha')^{-1}(\frac{1}{2} \log[n] + t))|$ is studied, where $T_0 = \inf\{t \geq 0; |S(t)| \leq n - \sqrt{n}\}$ is the time when \sqrt{n} vertices are infected or recovered. As a corollary of the results of [7] it follows that for $T'_\gamma(n)$ defined as in Lemma 2.3, $T'_\gamma(n) - (\alpha')^{-1} \log n$ converges in distribution as $n \rightarrow \infty$. We avoid the extra conditions of [7] at the cost of only being able to study the convergence (in probability) of $T'_\gamma(n)/(\log n)$.

4.2 The final phase of the epidemic

In order to describe the end of the epidemic more work is required. We use that for $1 - q^* - \gamma > 0$ but small, the time interval between $T'_\gamma(n)$ and $T^*(n)$,

none of the quantities $n^{-1}|S^{(n)}(t)|$ and $n^{-1}|\mathcal{E}_1^{(n)}(t)|$, $n^{-1}|\mathcal{E}_2^{(n)}(t)|$, $n^{-1}|\mathcal{E}_3^{(n)}(t)|$ and $n^{-1}|\mathcal{E}_4^{(n)}(t)|$ (as defined in Section 3.1.2) change much. So, we assume that during the final stages of the epidemic, the environment of newly infected vertices is more or less constant. That is, we assume that the degree distribution and the fraction of the neighbours which are still susceptible of newly infected vertices are constant during this final phase. In particular, the degree distribution of a vertex infected during the final phase of the epidemic should be well approximated by the size biased degree distribution of ultimately susceptible vertices, while the fraction of susceptible neighbours of a newly infected vertex in this phase of the epidemic should be well approximated by the fraction of susceptible neighbours of ultimately susceptible vertices. We now find those quantities.

Let D^* be a random variable, such that the degree of a uniformly chosen ultimately susceptible vertex converges in distribution to D^* as $n \rightarrow \infty$. And let p_{ss}^* be the probability that a given neighbour of an ultimately susceptible vertex is ultimately susceptible itself. Below we show that p_{ss}^* is indeed well defined, and whether a given neighbour of an ultimately susceptible vertex is susceptible is independent of the degree of that vertex.

The end of the epidemic is then described by offspring measure

$$\mu^*(dt) = \mathbb{E}[\tilde{D}^* - 1] p_{ss}^* \beta e^{-\beta t} \mathbb{P}(L > t) dt, \quad (16)$$

which is derived in the same way as equation (4) and where \tilde{D}^* is the size-biased variant of D^* . Below we then derive that

$$\mathbb{E}[\tilde{D}^* - 1] = \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^*)^{\tilde{D}-1}]}{\tilde{q}} \quad \text{and} \quad p_{ss}^* = \frac{\tilde{q}^*}{1 - \psi + \psi \tilde{q}^*}.$$

Combining the above with (16) and Corollary 3.2 then gives Lemma 2.4.

4.2.1 Degree distribution of ultimately susceptible individuals

It is important to note that in the epidemic process the probability of a vertex being ultimately recovered or susceptible does not depend on its infectious period, even when infectious periods are random. This fact help us to derive the probability of a vertex being ultimately susceptible and of degree k (as in [2]), which then yields the degree distribution of the ultimately susceptible.

Assume the epidemic takes off, which occurs with the same probability as the survival of an approximating forward branching process, see e.g. [4]. Recall that there is only one initially infectious individual. So, as $n \rightarrow \infty$, the probability that a uniformly chosen vertex is the initial infectious vertex

goes to zero. Therefore, the probability that a uniformly chosen vertex v is ultimately susceptible (i.e. it escapes the epidemic) is given by

$$\xi = \sum_{k=0}^{\infty} \xi_k p_k, \quad (17)$$

where ξ_k is probability that a vertex of degree k does not acquire the infection by any of its neighbours until the end of the epidemic. We denote a neighbour of vertex v by u . Recall that $1 - \psi$ is the probability that u does not have an infectious contact to v . Let \tilde{q}^* denote the probability that u escapes the epidemic (we determine \tilde{q}^* later). Then, ξ_k is given by

$$\xi_k = \sum_{l=0}^k (1 - \psi)^l \binom{k}{l} (\tilde{q}^*)^{k-l} (1 - \tilde{q}^*)^l = (1 - \psi + \psi \tilde{q}^*)^k. \quad (18)$$

Using (18) in (17), the probability for a uniformly chosen vertex x to escape the epidemic is then given by

$$\xi = \sum_{k=0}^{\infty} p_k (1 - \psi + \psi \tilde{q}^*)^k. \quad (19)$$

Similarly, the probability \tilde{q}^* that the neighbour of a vertex escaping infection also escapes infection is given by

$$\tilde{q}^* = \sum_{k=0}^{\infty} \tilde{\xi}_k \tilde{p}_k, \quad (20)$$

where $\tilde{\xi}_k$ is the probability that a degree k vertex does not acquire the infection from $k - 1$ given neighbouring vertices and is defined as

$$\tilde{\xi}_k = \sum_{l=0}^{k-1} (1 - \psi)^l \binom{k-1}{l} (\tilde{q}^*)^{k-l-1} (1 - \tilde{q}^*)^l = (1 - \psi + \psi \tilde{q}^*)^{k-1}. \quad (21)$$

Here we consider only $k - 1$ of the k neighbours of u because we are exploring the second generation of the susceptibility set of v and we assume that u does not acquire infection from v . Equations (20) and (21) give that \tilde{q}^* is the smallest solution of

$$\tilde{q}^* = \sum_{k=0}^{\infty} \tilde{p}_k (1 - \psi + \psi \tilde{q}^*)^{k-1}. \quad (22)$$

This result gives the implicit expression for the probability \tilde{q}^* that neighbour u escapes the epidemic. Moreover, from (18) we obtain the probability that

a vertex of degree k escapes the epidemic. From this we deduce that the probability that an ultimately susceptible individual has degree k (say p_k^*) is given by

$$p_k^* = \frac{\xi_k p_k}{\xi} = \frac{(1 - \psi + \psi \tilde{q}^*)^k p_k}{\sum_{j=1}^{\infty} p_j (1 - \psi + \psi \tilde{q}^*)^j}, \quad (23)$$

where ξ is a normalising constant and is defined in (19). The size biased distribution of the ultimately susceptible individuals is given through

$$\begin{aligned} \tilde{p}_k^* &= \frac{k p_k^*}{\sum_{j=1}^{\infty} j p_j^*} = \frac{k p_k (1 - \psi + \psi \tilde{q}^*)^k}{\sum_{j=1}^{\infty} j p_j (1 - \psi + \psi \tilde{q}^*)^j}, \\ &= \frac{\tilde{p}_k (1 - \psi + \psi \tilde{q}^*)^{k-1}}{\sum_{j=1}^{\infty} \tilde{p}_j (1 - \psi + \psi \tilde{q}^*)^{j-1}} = \frac{\tilde{p}_k (1 - \psi + \psi \tilde{q}^*)^{k-1}}{\tilde{q}^*}. \end{aligned} \quad (24)$$

4.2.2 Fraction of ultimately susceptible neighbours of an ultimately susceptible vertex

Let v be an arbitrary vertex of degree k and u one of its neighbours. We compute the fraction of neighbours of the ultimately susceptible which are also ultimately susceptible as the following conditional probability:

$$\begin{aligned} p_{ss}^* &= \mathbb{P}(u \text{ is ultimately susceptible} \mid v \text{ is ultimately susceptible}), \\ &= \frac{\mathbb{P}(v \text{ and } u \text{ are ultimately susceptible})}{\mathbb{P}(v \text{ is ultimately susceptible})} = \frac{\tilde{q}^* \tilde{\xi}_k}{\xi_k} = \frac{\tilde{q}^*}{1 - \psi + \psi \tilde{q}^*}. \end{aligned} \quad (25)$$

Note that this probability is independent of the degree k of the vertex v .

In computing the above probability, \tilde{q}^* is the probability that the initially susceptible neighbour u escapes the infection from all its neighbouring vertices, apart from possibly v , $\tilde{\xi}_k$ is the probability that v escapes infection from all of its neighbours, apart from possibly u , and ξ_k is the unconditional probability that vertex v does not acquire the infection until the end of the epidemic.

5 Proof of Lemma 2.3

We split up the proof in the following two lemmas which trivially lead to the proof of Lemma 2.3.

Lemma 5.1. $\frac{T'_\gamma(n)}{\log n} \leq \frac{1}{\alpha'} + \delta$ *w.h.p. for every $\delta > 0$.*

Lemma 5.2. *Assume $\mathbb{E}[D^2] < \infty$, then $\frac{T'_\gamma(n)}{\log n} \geq \frac{1}{\alpha' + \delta}$ w.h.p. for every $\delta > 0$.*

Note that for $\mathbb{E}[D^2] = \infty$, also $\alpha' = \infty$ and the equivalent of Lemma 5.2 is meaningless. Lemma 5.1 still holds in that case.

Proof of Lemma 5.1. Assume first that $D^{(n)}$ has uniformly bounded support, that is, there exist $K > 0$ such that $\mathbb{P}(D^{(n)} > K) = 0$ for all $n \in \mathbb{N}$. Furthermore, we assume that there exists $L_{\max} > 0$ such that $\mathbb{P}(L > L_{\max}) = 0$, i.e. we assume that L has bounded support. Under those assumptions the conditions of [7, Thm. 3.3] are satisfied. Note that in the notation of [7], λ is the Malthusian parameter (α' in our notation) and N is the population size (n in our notation). Furthermore, it is easily deduced from equation (3.11) and the definition of τ_N on page 27 of [7] that $\tau_N / [\log N] \xrightarrow{\mathbb{P}} 1/(2\lambda)$ on the event of a large outbreak. Finally, the expression $\hat{s}_l(u)$ in [7] is independent of N for all $l \in \{1, 2, \dots, K\}$. Translating the notation of [7, Thm. 3.3] to our notation we obtain as an immediate corollary that for every $\gamma \in (0, 1 - q^*)$ and every $\delta > 0$,

$$n^{-1} |S^{(n)}(((\alpha')^{-1} + \delta) \log n)| < q^* + \gamma \quad \text{w.h.p.}$$

To obtain the results without the extra conditions, let $K = K(\delta)$ be a large constant satisfying some properties specified later. Mark (before the pairing) all half-edges belonging to vertices with degree strictly larger than K . By assumptions (A1) and (A2) one can make the fraction of half-edges that are marked to be arbitrary small by choosing K and n large enough. The next step is to pair all half-edges (ignoring whether they are marked and unmarked) uniformly at random as before. Then delete all edges which contain at least one marked half-edge. If a fraction $\delta_1 = \delta_1(K)$ of the half-edges is marked then the remaining degree distribution of the graph is dominated by a Mixed Binomial distribution with random “number of trials parameter” $D^{(n)} \mathbf{1}(D^{(n)} \leq K)$ and “probability parameter” $1 - \delta_1$ (The domination is because we ignore that some marked half-edges will be paired with other marked half-edges). Let $D_K^{(n)}$ be distributed as the Mixed Binomial random variable, and $\tilde{D}_K^{(n)}$ be the size-biased variant of $D_K^{(n)}$. It follows immediately from assumptions (A1) and (A2) that

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}[D_K^{(n)}] = \mathbb{E}[D].$$

Furthermore, both for $\mathbb{E}[D^2] = \infty$ and $\mathbb{E}[D^2] < \infty$, we obtain from (A1) and (A3) that

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}[(D_K^{(n)})^2] = \mathbb{E}[D^2].$$

In particular, we obtain that

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}[\tilde{D}_K^{(n)} - 1] = \mathbb{E}[\tilde{D} - 1].$$

In addition we consider an epidemic on the newly created (thinned) graph with infectious period distribution

$$L' = L\mathbf{1}(L < L_{\max}) + L_{\max}\mathbf{1}(L \geq L_{\max}).$$

So, in the new model we have deleted some edges and shortened some infectious periods, which make that the epidemic spreads faster in the original model.

For this new epidemic we deduce from (7) that the Malthusian parameter is the x satisfying

$$\frac{1}{\mathbb{E}[\tilde{D}_K^{(n)} - 1]} = \int_0^{L_{\max}} e^{-xt}\beta e^{-\beta t}\mathbb{P}(L > t)dt =: f(x, L_{\max}). \quad (26)$$

Note that $f(x, L_{\max})$ is continuous and decreasing in x and continuous and increasing in L_{\max} . Furthermore, if $\psi\mathbb{E}[\tilde{D} - 1] > 1$, then

$$f(0, \infty) = \psi > \frac{1}{\mathbb{E}[\tilde{D} - 1]}.$$

While $\lim_{x \rightarrow \infty} f(x, L_{\max}) = 0$ for all $L_{\max} > 0$. It follows that the solution of (26) converges to α' as $K \rightarrow \infty$ and $L_{\max} \rightarrow \infty$. In particular, for every $\delta > 0$, there exists K_0 and L_0 such that for all $K > K_0$ and $L_{\max} > L_0$, the x solving (26) satisfies $1/x < 1/\alpha' + \delta/2$.

So, by choosing L_{\max} and K large enough (but finite), we are in the realm of [7, Thm. 3.3] and for the corresponding model we obtain that for every $\gamma \in (0, 1 - q^*)$ and $\delta > 0$ with high probability it holds that,

$$n^{-1}|S^{(n)}((\alpha')^{-1} + \delta/2 + \delta/2) \log n| < q^* + \gamma,$$

which finishes the proof of Lemma 5.1. \square

Proof of Lemma 5.2. In order to prove the lemma we prove the following stronger statement: The number of vertices affected by the epidemic up to time $\frac{\log n}{\alpha' + \delta}$ satisfies $|n - S(\frac{\log n}{\alpha' + \delta})| = o(n)$ with high probability for all $\delta > 0$.

Now, for $\delta_1 > 0$, let $\alpha_1 = \alpha_1(\delta_1)$ satisfy

$$\frac{1}{\mathbb{E}[\tilde{D}] + \delta_1} = \int_0^\infty e^{-\alpha_1 t}\beta e^{-\beta t}\mathbb{P}(L > t)dt. \quad (27)$$

As before, because $R_0 > 1$, we know that α_1 exists and is positive for all $\delta_1 \geq 0$ and is continuous increasing in δ_1 on $[0, \infty)$. In particular, for every $\delta > 0$, we can and do choose $\delta_1 > 0$ such that $\alpha_1(\delta_1) < \alpha' + \delta/2$.

We use the notation of Section 3.1.2, where the vertices in $V^{(n)}$ are labelled such that the degree sequence d_1, d_2, \dots, d_n is non-decreasing. Recall that $\ell(n) = \sum_{v=1}^n d_v$ and $\ell_2(n) = \sum_{v=1}^n (d_v)^2$, which by assumption (A3) and the assumption $\mathbb{E}[D^2] < \infty$ (or equivalently $\mathbb{E}[\tilde{D}] < \infty$) is $O(n)$.

Let $\epsilon_1 \in (0, 1)$ be a number to be specified later. For $i \leq \epsilon_1 n$ define the random variable $D^{(n)}(\mathbf{x}; i)$ through

$$\mathbb{P}(D^{(n)}(\mathbf{x}; i) = k) = \frac{\sum_{v=1}^n \mathbf{1}(d_v = k) \mathbf{1}(v \notin \{x_0, x_1, \dots, x_i\})}{\sum_{v=1}^n \mathbf{1}(v \notin \{x_0, x_1, \dots, x_i\})}.$$

That is, $D^{(n)}(\mathbf{x}; i)$ is the degree distribution of the vertices not chosen in the first i elements of \mathbf{x} .

Note that for all $i \leq i_0 = \lfloor \epsilon_1 n \rfloor$, the random variable $D^{(n)}(\mathbf{x}; i)$ is stochastically dominated by $D''^{(n)}(\epsilon_1)$, which is defined through

$$\mathbb{P}(D''^{(n)}(\epsilon_1) = k) = \frac{\sum_{v=i_0+1}^n \mathbf{1}(d_v = k)}{n - i_0}.$$

Let $\tilde{D}''^{(n)}(\epsilon_1)$ be the size biased variant of $D''^{(n)}(\epsilon_1)$. It follows that

$$\mathbb{E}[\tilde{D}''^{(n)}(\epsilon_1)] = \frac{\sum_{v=i_0+1}^n (d_v)^2}{\sum_{v=i_0+1}^n d_v}.$$

Observe that $\sum_{v=i_0+1}^n (d_v)^2 \leq \sum_{v=1}^n (d_v)^2 = \ell_2(n)$, while

$$\sum_{v=i_0+1}^n d_v = \sum_{v=1}^n d_v - \sum_{v=1}^{i_0} d_v \geq \ell(n) - i_0 \mathbb{E}[D^{(n)}] \geq \ell(n)(1 - \epsilon_1).$$

So,

$$\mathbb{E}[\tilde{D}''^{(n)}(\epsilon_1)] \leq \frac{\ell_2(n)}{\ell(n)(1 - \epsilon_1)} = \frac{1}{1 - \epsilon_1} \mathbb{E}[\tilde{D}^{(n)}].$$

Note that (apart from possibly x_0), in the construction of $\{\mathcal{E}(t); t \geq 0\}$ the degree of a vertex added to the exploration is stochastically smaller than $\tilde{D}''^{(n)}(\epsilon_1)$, as long as $t < t_0$, where $t_0 = t_0(\epsilon_1) = \max\{t > 0; |\mathbf{x}(t)| \leq i_0\}$. That is up to we add the i_0 -th vertex to $\{\mathcal{E}(t); t \geq 0\}$ the number of vertices in $\mathcal{E}(t)$ is less than the number of particles in a branching process with offspring measure $\mu''^{(n)}(dt; i_0) = \mathbb{E}[\tilde{D}''^{(n)}(\epsilon_1)] \beta e^{-\beta t} \mathbb{P}(L > t) dt$. Denote the number of particles in this branching process at time t ($t \geq 0$) by $Z''^{(n)}(t)$.

Because $\mathbb{E}[\tilde{D}^{(n)}] \rightarrow \mathbb{E}[\tilde{D}]$ as $n \rightarrow \infty$, we have that for every $\delta_1 > 0$, we can choose $\epsilon_1 = \epsilon_1(\delta_1) > 0$ and $n_0 = n_0(\delta_1) \in \mathbb{N}$, such that $\mathbb{E}[\tilde{D}''^{(n)}(\epsilon_1)] < \mathbb{E}[\tilde{D}] + \delta_1$ for all $n > n_0$.

So for $\epsilon_1 = \epsilon_1(\delta_1)$ and $n_0 = n_0(\delta_1)$ as above, $\{Z''^{(n)}(t); t \in (0, t_0)\}$ is dominated by a branching process $\{Z''(t); t \in (0, t_0)\}$ with offspring measure

$$\mu''(dt) = (\mathbb{E}[\tilde{D}] + \delta_1)\beta e^{-\beta t}\mathbb{P}(L > t)dt.$$

This branching process has a Malthusian parameter $\alpha_1(\delta_1)$, which satisfies equation (27) and is less than $\alpha' + \delta/2$.

Now observe that by Lemma 3.1, with high probability

$$|V^{(n)}\left(\frac{\log n}{\alpha' + \delta}\right) \setminus S^{(n)}\left(\frac{\log n}{\alpha' + \delta}\right)| = O\left(e^{(\alpha' + \frac{\delta}{2})\frac{\log n}{\alpha' + \delta}}\right) = O\left(n^{\frac{\alpha' + \delta/2}{\alpha' + \delta}}\right) = o(n).$$

Since $i_0 = \theta(n)$ and the number of individuals infected before time t is stochastically less than $Z''(\frac{\log n}{\alpha' + \delta})$, we obtain that $n - |S(\frac{\log n}{\alpha' + \delta})| = o(n)$ with high probability for all $\delta > 0$. \square

6 The final stage of the epidemic and proof of Lemma 2.4

6.1 Proof of Lemma 2.1

Before we prove Lemma 2.4, we first provide the proof of Lemma 2.1.

Proof of Lemma 2.1. Because (8) has a solution and $\int_0^\infty e^{-xt}\mu^*(dt)$ is decreasing in x , we obtain that $\alpha^* < 0$ if and only if $\int_0^\infty \mu^*(dt) = \int_0^\infty e^{0 \cdot t}\mu^*(dt) < 1$. We obtain from (2) that

$$\begin{aligned} \int_0^\infty \mu^*(dt) &= \int_0^\infty \mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi\tilde{q}^*)^{\tilde{D}-2}]\beta e^{-\beta t}\mathbb{P}(L > t)dt \\ &= \psi\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi\tilde{q}^*)^{\tilde{D}-2}]. \end{aligned}$$

The function

$$g(s) = \sum_{k=1}^{\infty} \tilde{p}_k(1 - \psi + \psi s)^{k-1}$$

is convex and analytic on $s \in [0, 1]$ and has derivative

$$g'(s) = \psi \sum_{k=1}^{\infty} (k-1)\tilde{p}_k(1 - \psi + \psi s)^{k-2} = \psi\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi s)^{\tilde{D}-2}].$$

Furthermore, by the definition of \tilde{q}^* (see (10)) and the convexity of $g(s)$, \tilde{q}^* and 1 are the only two solutions of the equation $s = g(s)$ in $[0, 1]$. This, together with $\int_0^\infty \mu^*(dt) = g'(\tilde{q}^*)$, the convexity of g and $\tilde{q}^* < 1$ implies that $\int_0^\infty \mu^*(dt) < 1$, which finishes the proof. \square

6.2 Time until the end of the epidemic

In this section we use the construction of the epidemic generated graph as presented in Section 3.1.2. We restrict ourselves to major outbreaks. Our approach is to define a random time $t_1 = t_1^{(n)}$, when the fraction of susceptible vertices among all vertices is larger than, but close to, its asymptotic value and sandwich (w.h.p.) the process $\{|I^{(n)}(t)|; t \geq t_1^{(n)}\}$ between two branching processes and then find the time until those branching processes go extinct.

We start with an almost trivial observation, which we need in the proof of Lemma 2.4. Let D_x^* be a random variable with distribution defined through

$$p_k^*(x) = \mathbb{P}(D_x^* = k) = \frac{p_k x^k}{\sum_{l=0}^{\infty} p_l x^l} \quad (28)$$

for $k \geq 0$. Here $p_k = \mathbb{P}(D = k)$ as in Section 2.2.

Claim 6.1. *For $x \in (0, 1)$ all moments of the random variable D_x^* are finite, regardless of the distribution of D .*

Proof. We use the D'Alembert ratio test for convergence of series [24, p. 65]. Consider the j th moment of D_x^* :

$$\mathbb{E}[(D_x^*)^j] = \sum_{k=1}^{\infty} k^j p_k^*(x) = \sum_{k=1}^{\infty} k^j \frac{p_k x^k}{\sum_{l=0}^{\infty} p_l x^l} = \frac{\mathbb{E}[D^j x^D]}{\mathbb{E}[x^D]}.$$

To apply the ratio test we consider the limit of two subsequent terms

$$\lim_{k \rightarrow \infty} \frac{(k+1)^j x^{k+1}}{k^j x^k} = x \in (0, 1),$$

by assumption. This implies that the sum $\sum_{k=1}^{\infty} k^j x^k$ converges. It then follows that

$$\sum_{k=1}^{\infty} k^j p_k^*(x) = \sum_{k=1}^{\infty} k^j \frac{p_k x^k}{\sum_{l=0}^{\infty} p_l x^l} \leq \sum_{k=1}^{\infty} k^j \frac{x^k}{\sum_{l=0}^{\infty} p_l x^l} < \infty.$$

□

Let $\{\mathcal{E}^{(n)}(t); t \geq 0\} = \{(\mathcal{E}_1^{(n)}(t), \mathcal{E}_2^{(n)}(t), \mathcal{E}_3^{(n)}(t), \mathcal{E}_4^{(n)}(t); t \geq 0\}$ be as in Section 3.1.2. In our analysis below we consider $|\mathcal{E}_1^{(n)}(t)|$, $|\mathcal{E}_4^{(n)}(t)|$ and

$$\sum_{v \in S^{(n)}(t)} d_v \mathbf{1}(d_v \geq k) = \sum_{v \in V^{(n)}} d_v \mathbf{1}(d_v \geq k) \mathbf{1}(v \in S^{(n)}(t)).$$

Note that $|\mathcal{E}_1^{(n)}(t)|$ and $\sum_{v \in S^{(n)}(t)} d_v \mathbf{1}(d_v \geq k)$ are decreasing in t , while $|\mathcal{E}_4^{(n)}(t)|$ is increasing in t . For $\epsilon \in (0, \psi(1 - \tilde{q}^*))$ define

$$\begin{aligned} t_a^{(n)}(\epsilon) &:= \inf\{t > 0; |\mathcal{E}_1^{(n)}(t)| \leq \mathbb{E}[(1 - \psi + \psi\tilde{q}^* + \epsilon)^{\tilde{D}}] \ell(n)\}, \\ t_b^{(n)}(\epsilon) &:= \inf\{t > 0; |\mathcal{E}_4^{(n)}(t)| \geq \ell_n - 1 - (1 - \psi + \psi\tilde{q}^* + \epsilon)^2 \ell(n)\}, \\ t_c^{(n)}(\epsilon) &:= \inf\{t > 0; \sum_{v \in S^{(n)}(t)} d_v \mathbf{1}(d_v \geq k) \leq \mathbb{E}[\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi\tilde{q}^* + \epsilon)^{\tilde{D}}] \ell(n)\}, \end{aligned}$$

where the infimum of an empty set is ∞ . Let

$$t_1^{(n)}(\epsilon) := \max(t_a^{(n)}(\epsilon), t_b^{(n)}(\epsilon), t_c^{(n)}(\epsilon))$$

and define the event $\mathcal{A}_1^{(n)}(\epsilon) = \{t_1^{(n)}(\epsilon) < \infty\}$. Let $\mathcal{A}_2^{(n)}(\epsilon)$ be the event that all of the following events hold.

$$\begin{aligned} |\mathcal{E}_1^{(n)}(\infty)| &> \mathbb{E}[(1 - \psi + \psi\tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n) \\ |\mathcal{E}_4^{(n)}(\infty)| &< \ell_n - 1 - (1 - \psi + \psi\tilde{q}^* - \epsilon)^2 \ell(n) \\ \sum_{v \in S^{(n)}(\infty)} d_v \mathbf{1}(d_v \geq k) &> \mathbb{E}[\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi\tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n). \end{aligned}$$

Finally define $\mathcal{A}^{(n)}(\epsilon) = \mathcal{A}_1^{(n)}(\epsilon) \cap \mathcal{A}_2^{(n)}(\epsilon)$.

Lemma 6.1. *For all $\epsilon \in (0, \psi(1 - \tilde{q}^*))$ and conditioned on a large outbreak, $\mathbb{P}(\mathcal{A}^{(n)}(\epsilon)) \xrightarrow{\mathbb{P}} 1$ and there exists $c_1 > 0$, such that*

$$|S^{(n)}(t_1^{(n)}(\epsilon))| - |S^{(n)}(\infty)| = c_1 n \quad w.h.p.$$

Proof. We start with some definitions. Let $K_1 = K_1^{(n)}(\epsilon)$ be a Poisson distributed random variable with expectation $\ell(n) |\log(1 - \psi + \psi\tilde{q}^* + \epsilon/2)|$ and let $K_2 = K_2^{(n)}(\epsilon)$ be a Poisson distributed random variable with expectation $\ell(n) |\log(1 - \psi + \psi\tilde{q}^* - \epsilon/2)|$. Both K_1 and K_2 are independent of the epidemic process. Let $\{|\mathbf{x}(t)|; t \geq 0\}$ be as in Section 3.1.2 and define the random time $t'(K_i) := \inf\{t \geq 0; |\mathbf{x}(t)| \geq K_i\}$ for $i \in \{1, 2\}$.

Because the elements of \mathbf{x} are i.i.d. and uniform among all $\ell(n)$ half-edges, we have by well-known properties of the Poisson distribution (see e.g. [23, p. 317]) that the number of times a given half-edge is among the first K_1 (resp. K_2) elements of \mathbf{x} is Poisson distributed with expectation $|\log(1 - \psi + \psi\tilde{q}^* + \epsilon/2)|$ (resp. Poisson distributed with expectation $|\log(1 - \psi + \psi\tilde{q}^* - \epsilon/2)|$) and independent for different half-edges. This implies that the events that different half-edges are not among the first K_1 elements of \mathbf{x} are independent and have probability $e^{-|\log(1 - \psi + \psi\tilde{q}^* + \epsilon/2)|} = 1 - \psi + \psi\tilde{q}^* + \epsilon/2$. Similarly, the events that different half-edges are not among the first K_2 elements of \mathbf{x} are independent and have probability $1 - \psi + \psi\tilde{q}^* - \epsilon/2$.

It is easy to see that the probability that none of the half-edges belonging to a uniformly chosen vertex is part of the first K_1 elements of \mathbf{x} is given by

$$\sum_{k=0}^{\infty} \mathbb{P}(D^{(n)} = k)(1 - \psi + \psi\tilde{q}^* + \epsilon/2)^k.$$

So, using a variant of the (weak) law of large numbers (e.g. [13, Problem 7.11.20]) and $D^{(n)} \xrightarrow{d} D$, we obtain

$$\frac{1}{n} |S^{(n)}(t'(K_1))| \xrightarrow{\mathbb{P}} \sum_{k=0}^{\infty} \mathbb{P}(D = k)(1 - \psi + \psi\tilde{q}^* + \epsilon/2)^k. \quad (29)$$

By Lemma 2.2 and equation (11) we know that conditioned on a large outbreak

$$\frac{1}{n} |S^{(n)}(\infty)| \xrightarrow{\mathbb{P}} q^* = \sum_{k=0}^{\infty} \mathbb{P}(D = k)(1 - \psi + \psi\tilde{q}^*)^k. \quad (30)$$

Because $\sum_{k=0}^{\infty} \mathbb{P}(D^{(n)} = k)x^k$ is strictly increasing on $x \in [0, 1)$, and $D^{(n)} \xrightarrow{d} D$, (29) and (30) imply that there exists $c_1 > 0$, such that

$$|S^{(n)}(t'(K_1))| - |S^{(n)}(\infty)| > c_1 n \quad \text{w.h.p.} \quad (31)$$

Equation (30) also immediately gives that

$$\frac{1}{n} |S^{(n)}(\infty)| > \sum_{k=0}^{\infty} \mathbb{P}(D = k)(1 - \psi + \psi\tilde{q}^* - \epsilon/2)^k \quad \text{w.h.p.}$$

and therefore,

$$\mathbf{x}(\infty) < K_2(\epsilon) \quad \text{w.h.p.} \quad (32)$$

Similarly, the probability that a uniformly chosen half-edge belongs to a vertex of which none of the half-edges is part of the first K_1 elements of \mathbf{x} is given by

$$\frac{1}{\ell(n)} \mathbb{E}[|\mathcal{E}_1^{(n)}(t'(K_1))|] = \sum_{k=0}^{\infty} \mathbb{P}(\tilde{D}^{(n)} = k)(1 - \psi + \psi\tilde{q}^* + \epsilon/2)^k.$$

Using the same arguments as above this implies that $t'(K_1) > t_a^{(n)}(\epsilon)$ w.h.p. Combined with (31) this implies $|S^{(n)}(t_a^{(n)}(\epsilon))| - |S^{(n)}(\infty)| > c_1 n$ w.h.p. Furthermore, by (32)

$$\frac{1}{\ell(n)} \mathbb{E}[|\mathcal{E}_1^{(n)}(\infty)|] > \sum_{k=0}^{\infty} \mathbb{P}(\tilde{D}^{(n)} = k)(1 - \psi + \psi\tilde{q}^* - \epsilon/2)^k$$

and again using the same weak law of large numbers arguments we obtain that

$$\frac{1}{\ell(n)} |\mathcal{E}_1^{(n)}(\infty)| > \sum_{k=0}^{\infty} \mathbb{P}(\tilde{D}^{(n)} = k) (1 - \psi + \psi \tilde{q}^* - \epsilon/2)^k \quad \text{w.h.p.}$$

Now we turn our attention to $|\mathcal{E}_4^{(n)}(t)|$. In G , all half-edges are paired uniformly at random. For a half-edge not to be part of $\mathcal{E}_4^{(n)}(t)$, neither the half-edge itself nor its partner should be part of $\mathbf{x}(t)$. So,

$$\frac{\ell(n) - |\mathcal{E}_4^{(n)}(t'(K_1))| - 1}{\ell(n)} < \frac{\ell(n) - |\mathcal{E}_4^{(n)}(t'(K_1))|}{\ell(n)} \xrightarrow{\mathbb{P}} (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^2. \quad (33)$$

So, $t'(K_1) > t_b^{(n)}(\epsilon)$ w.h.p. Combined with (31) this gives $|S^{(n)}(t_b^{(n)}(\epsilon))| - |S^{(n)}(\infty)| > c_1 n$. From (32) we obtain in a similar fashion that

$$\frac{\ell(n) - |\mathcal{E}_4^{(n)}(\infty)| - 1}{\ell(n)} > (1 - \psi + \psi \tilde{q}^* - \epsilon/2)^2 \quad \text{w.h.p.}$$

Finally, we consider $\sum_{v \in S^{(n)}(t)} d_v(d_v - 1)$. We know,

$$\begin{aligned} & \mathbb{E}\left[\frac{1}{\ell(n)} \sum_{v \in V^{(n)}} d_v(d_v - 1) \mathbf{1}(v \in S(t'(K_a)))\right] \\ &= \frac{n}{\ell(n)} \frac{1}{n} \sum_{v \in V^{(n)}} d_v(d_v - 1) (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^{d_v} \\ &= \frac{1}{\mathbb{E}[D^{(n)}]} \mathbb{E}[D^{(n)}(D^{(n)} - 1) (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^{D^{(n)}}] \\ &= \mathbb{E}[(\tilde{D}^{(n)} - 1) (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^{\tilde{D}^{(n)}}], \end{aligned}$$

which by $D^{(n)} \xrightarrow{\mathbb{P}} D$, the bounded convergence theorem and Claim 6.1 converges in probability to $\mathbb{E}[(\tilde{D} - 1) (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^{\tilde{D}}]$. Using a law of large number argument as before we then obtain that,

$$\frac{1}{\ell(n)} \sum_{v \in V^{(n)}} d_v(d_v - 1) \mathbf{1}(v \in S^{(n)}(t)) \xrightarrow{\mathbb{P}} \mathbb{E}[(\tilde{D} - 1) (1 - \psi + \psi \tilde{q}^* + \epsilon/2)^{\tilde{D}}],$$

which implies that $t'(K_a) > t_c(\epsilon)$ w.h.p. Combined with (31) this gives $|S^{(n)}(t_c^{(n)}(\epsilon))| - |S^{(n)}(\infty)| > c_1 n$.

The proof of

$$\sum_{v \in S^{(n)}(\infty)} d_v(d_v - 1) > \mathbb{E}[(\tilde{D} - 1) (1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n) \quad \text{w.h.p.}$$

is completely analogous to the proof of $|\mathcal{E}_1^{(n)}(\infty)| > \mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n)$ w.h.p. Combining all the above results finishes the proof of the lemma. \square

Now we are almost ready to prove Lemma 2.4. In the proof we consider who infected whom, and since individuals can be infected only once, this leads to a tree representation of the infection process: the infection tree. We say that vertex u is an ancestor of v if the path from the initial infectious vertex to v in the infection tree contains u . To be complete we say that v is an ancestor of itself. Let v be a vertex infected at time $\sigma(v)$. Then define $\{J_v^{(n)}(t); t \geq 0\}$, through

$$J_v^{(n)}(t) = I^{(n)}(\sigma(v) + t) \cap \{u \in V^{(n)}; v \text{ is an ancestor of } u\}.$$

Furthermore, let $V_*^{(n)}(t) \subset V^{(n)}$, be the set of vertices infected after time t .

In the proofs below we use the standard notation that for any function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(t-) = \lim_{s \nearrow t} f(s)$.

6.2.1 Proof of Lemma 2.4

Proof of Lemma 2.4. Throughout the proof we restrict ourselves to the event $\mathcal{A}^{(n)}(\epsilon)$ (defined as in Lemma 6.1) for some $\epsilon > 0$ conveniently chosen. Because there exists $c_1 > 0$ such that $|S^{(n)}(t_1^{(n)}(\epsilon))| - |S^{(n)}(\infty)| > c_1 n$ w.h.p., we immediately obtain that there exists $\gamma \in (0, 1 - q^*)$ such that $T'_\gamma(n) \in (t_1^{(n)}(\epsilon), T^*(n))$ w.h.p. The proof consists of the following steps:

1. There exists with high probability $\gamma > 0$ such that for $v \in V_*^{(n)}(T'_\gamma(n))$ and for $\delta \in (0, |\alpha^*|)$ small enough, we can construct a branching process which is dominated by $\{J_v^{(n)}(t); t \geq 0\}$ and has Malthusian parameter larger than $-(|\alpha^*| + \delta/2)$ (i.e. the absolute value of the Malthusian is less than $|\alpha^*| + \delta/2$) and a branching process which dominates $\{J_v^{(n)}(t); t \geq 0\}$ and has Malthusian parameter less than $-(|\alpha^*| - \delta/2)$ (i.e. the absolute value of the Malthusian is larger than $|\alpha^*| - \delta/2$).
2. Show that there exists $\gamma > 0$ and $\delta > 0$ such that the dominating and dominated branching process satisfy the conditions of Lemma 3.2.
3. Show that the number of vertices infected after time $T'_\gamma(n)$, which are infected by vertices infected before time $T'_\gamma(n)$ is $\theta(n)$.
4. Show that for every $\delta \in (0, |\alpha^*|)$, there exist $\gamma > 0$, such that $T^*(n) - T'_\gamma(n) < \frac{\log n}{|\alpha^*| - \delta}$ w.h.p.
5. Show that for every $\delta \in (0, 1)$, there exist $\gamma > 0$, such that $T^*(n) - T'_\gamma(n) > (1 - \delta) \frac{\log n}{|\alpha^*|}$ w.h.p.

Combining the statements of step 4 and 5 immediately imply Lemma 2.4.

Step 1:

Let $\epsilon > 0$ be small and chosen appropriately later. If at time t a half-edge from $\mathcal{E}_2^{(n)}(t-)$ is paired with another half-edge, this other half-edge belongs to $\mathcal{E}_1^{(n)}(t-)$ with probability

$$\kappa^{(n)}(t) = \frac{|\mathcal{E}_1^{(n)}(t-)|}{\ell(n) - |\mathcal{E}_4^{(n)}(t-)| - 1}.$$

Here $\ell(n) - |\mathcal{E}_4^{(n)}(t-)|$ is the number of not-yet paired vertices just before time t and the -1 appear in the denominator because the half-edge from $\mathcal{E}_2^{(n)}(t-)$ cannot be paired with itself. Furthermore, the probability that if the half-edge is paired with a half-edge from $\mathcal{E}_1^{(n)}(t-)$, it belongs to a vertex of degree at least k is given by

$$\pi_{\geq k}^{(n)}(t) = \frac{\sum_{v \in S^{(n)}(t-)} d_v \mathbf{1}(d_v \geq k)}{|\mathcal{E}_1^{(n)}(t-)|}.$$

The quantities $|S^{(n)}(t)|$ and $\{|\mathcal{E}_1^{(n)}(t)|; t \geq 0\}$ are decreasing in t , while $\{|\mathcal{E}_4^{(n)}(t)|; t \geq 0\}$ is increasing in t . So, for $t_1 := t_1^{(n)}(\epsilon)$ as in Lemma 6.1, and $t > t_1$ and on $\mathcal{A}^{(n)}(\epsilon)$,

$$\kappa^{(n)}(t) \leq \frac{|\mathcal{E}_1^{(n)}(t_1)|}{\ell(n) - |\mathcal{E}_4^{(n)}(\infty)| - 1} \leq \frac{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] \ell(n)}{(1 - \psi + \psi \tilde{q}^* - \epsilon)^2 \ell(n)} =: \kappa_+(\epsilon) \quad (34)$$

and

$$\kappa^{(n)}(t) \geq \frac{|\mathcal{E}_1^{(n)}(\infty)|}{\ell(n) - |\mathcal{E}_4^{(n)}(t_1)| - 1} \geq \frac{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n)}{(1 - \psi + \psi \tilde{q}^* + \epsilon)^2 \ell(n)} =: \kappa_-(\epsilon). \quad (35)$$

Similarly, on $\mathcal{A}^{(n)}(\epsilon)$ and for $k \geq 1$

$$\begin{aligned} \pi_{\geq k}^{(n)}(t) &= \frac{\sum_{v \in S^{(n)}(t-)} d_v \mathbf{1}(d_v \geq k)}{|\mathcal{E}_1^{(n)}(t-)|} \leq \min \left(1, \frac{\sum_{v \in S^{(n)}(t_1)} d_v \mathbf{1}(d_v \geq k)}{|\mathcal{E}_1^{(n)}(\infty)|} \right) \\ &\leq \min \left(1, \frac{\mathbb{E}[(\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] \ell(n)}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \ell(n)} \right) =: \mathbb{P}(\tilde{D}^+(\epsilon) \geq k). \quad (36) \end{aligned}$$

That is, $\tilde{D}^+(\epsilon)$ stochastically dominates the random variable defined through

$\pi_{\geq k}^{(n)}(t)$ for $t > t_1^{(n)}(\epsilon)$. Furthermore,

$$\begin{aligned} \pi_{\geq k}^{(n)}(t) &= \frac{\sum_{v \in \mathcal{S}^{(n)}(t-)} d_v \mathbf{1}(d_v \geq k)}{|\mathcal{E}_1^{(n)}(t-)|} \geq \frac{\sum_{v \in \mathcal{S}^{(n)}(\infty)} d_v \mathbf{1}(d_v \geq k)}{|\mathcal{E}_1^{(n)}(t_1)|} \\ &\geq \frac{\mathbb{E}[(\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}})] \ell(n)}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] \ell(n)} =: \mathbb{P}(\tilde{D}^- (\epsilon) \geq k). \end{aligned} \quad (37)$$

That is, $\tilde{D}^- (\epsilon)$ is stochastically dominated by the random variable defined through $\pi_{\geq k}^{(n)}(t)$ for $t > t_1^{(n)}(\epsilon)$.

Let v be a vertex infected at time t . Then v has degree distribution defined through $\pi_{\geq k}^{(n)}(t)$. One of the d_v half-edges attached to v is paired at time t , while the other $d_v - 1$ are still unpaired at time t . Let L_v be the infectious period of v and let $\tau_{v,1}, \tau_{v,2}, \dots, \tau_{v,d_v-1}$ be independent exponentially distributed random variables with expectation $1/\beta$ assigned to the different un-paired half-edges of v . If $\tau_{v,i} \leq L_v$, then $t + \tau_{v,i}$ is the time at which a contact is made along the half-edge (and the half-edge is paired), while if $\tau_{v,i} > L_v$, then $\tau_{v,i}$ does not have an epidemiological interpretation. If $\tau_{v,i} \leq L_v$ then the contact made at time $t + \tau_{v,i}$ is with a susceptible with probability $\kappa^{(n)}(t)$. By (34) and (36) we thus obtain that for all $v \in V_*^{(n)}(t_1^{(n)}(\epsilon))$, $\{|J_v^{(n)}(t)|; t \geq 0\}$ is stochastically dominated by a branching process in which particles give birth at ages given by the point process

$$\{\hat{\xi}_\epsilon^+(t); t \geq 0\} = \left\{ \sum_{k=1}^{\tilde{D}^+(\epsilon)-1} \mathbf{1}(\tau_{v,k} < \min(L, t)) Y_k^+(\epsilon); t \geq 0 \right\},$$

where $Y_k^+(\epsilon)$ is a Bernoulli random variable with success probability $\kappa_+(\epsilon)$ and all defined random variables are independent. The mean offspring measure of this branching process is then given by

$$\{\mu_\epsilon^+(t); t \geq 0\} = \mathbb{E}[\tilde{D}^+(\epsilon) - 1] \kappa_+(\epsilon) \mathbb{P}(\tau_{v,k} < \min(L, t)),$$

where

$$\begin{aligned} \mathbb{E}[\tilde{D}^+(\epsilon) - 1] &\leq \sum_{k=1}^{\infty} \frac{\mathbb{E}[(\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}})]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]} - 1 \\ &= \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}, \end{aligned}$$

while (34) gives

$$\kappa_+(\epsilon) = \frac{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]}{(1 - \psi + \psi \tilde{q}^* - \epsilon)^2}.$$

Finally,

$$\mathbb{P}(\tau_{v,k} < \min(L, t)) = \int_0^t (1 - e^{-\beta s}) L(ds) = \int_0^t \beta e^{-\beta s} \mathbb{P}(L > s) ds. \quad (38)$$

Combining the above terms we obtain that

$$\begin{aligned} \mu_\epsilon^+(dt) &\leq \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] (1 - \psi + \psi \tilde{q}^* - \epsilon)^2} \beta e^{-\beta t} \mathbb{P}(L > t) dt \\ &= K_+(\epsilon) \mu^*(dt), \end{aligned} \quad (39)$$

where

$$K_+(\epsilon) = \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]}{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^*)^{\tilde{D}-2}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}+2}]} \quad (40)$$

and $\mu^*(dt)$ is defined in (9).

Similarly, by (35) and (37) we obtain that for all $v \in V_*^{(n)}(t_1)$, the process $\{|J_v^{(n)}(t)|; t \geq 0\}$ stochastically dominates a branching process in which particles give birth at ages given by the point process

$$\{\hat{\xi}_\epsilon^-(t); t \geq 0\} = \left\{ \sum_{k=1}^{\tilde{D}^-(\epsilon)-1} \mathbf{1}(\tau_{v,k} < \min(L, t)) Y_k^-(\epsilon); t \geq 0 \right\},$$

where $Y_k^-(\epsilon)$ is a Bernoulli random variable with success probability $\kappa_-(\epsilon)$ and all defined random variables are independent. The mean offspring measure of this branching process is given by

$$\{\mu_\epsilon^-(t); t \geq 0\} = \mathbb{E}[\tilde{D}^-(\epsilon) - 1] \kappa_-(\epsilon) \mathbb{P}(\tau_{v,k} < \min(L, t)),$$

where

$$\begin{aligned} \mathbb{E}[\tilde{D}^-(\epsilon) - 1] &= \sum_{k=1}^{\infty} \frac{\mathbb{E}[(\mathbf{1}(\tilde{D} \geq k)(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]} - 1 \\ &= \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}]}, \end{aligned}$$

while (34) gives

$$\kappa_-(\epsilon) = \frac{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}{(1 - \psi + \psi \tilde{q}^* + \epsilon)^2}.$$

As for the upperbound

$$\mathbb{P}(\tau_{v,k} < \min(L, t)) = \int_0^t (1 - e^{-\beta s})L(ds) = \int_0^t \beta e^{-\beta s} \mathbb{P}(L > s) ds.$$

Combining this with (38), we obtain that

$$\begin{aligned} \mu_\epsilon^-(dt) &= \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}{\mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}}] (1 - \psi + \psi \tilde{q}^* + \epsilon)^2} \beta e^{-\beta t} \mathbb{P}(L > t) dt \\ &= K_-(\epsilon) \mu^*(dt), \end{aligned} \quad (41)$$

where

$$K_-(\epsilon) = \frac{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* - \epsilon)^{\tilde{D}}]}{\mathbb{E}[(\tilde{D} - 1)(1 - \psi + \psi \tilde{q}^*)^{\tilde{D}-2}] \mathbb{E}[(1 - \psi + \psi \tilde{q}^* + \epsilon)^{\tilde{D}+2}]} \quad (42)$$

and $\mu^*(dt)$ is defined in (9).

Because there exists with high probability $\gamma \in (0, 1 - q^*)$ such that $T'_\gamma(n) \in (t_1^{(n)}(\epsilon), T^*(n))$, we obtain that with high probability and for all $v \in V_*^{(n)}(T'_\gamma(n))$, we have constructed a branching process with reproduction process $\{\hat{\xi}_\epsilon^+(t); t \geq 0\}$ and mean offspring measure $\{\mu_\epsilon^+(t); t \geq 0\}$, which stochastically dominates the process $\{|J_v^{(n)}(t)|; t \geq 0\}$ and an independent branching process with reproduction process $\{\hat{\xi}_\epsilon^-(t); t \geq 0\}$ and mean offspring measure $\{\mu_\epsilon^-(t); t \geq 0\}$, which is stochastically dominated by the process $\{|J_v^{(n)}(t)|; t \geq 0\}$.

Step 2:

In this step we wish to show that there exists $\epsilon > 0$ such that

- (i) there exists $\alpha_\epsilon^- < 0$ such that $1 = \int_0^\infty e^{-\alpha_\epsilon^- t} \mu_\epsilon^-(dt)$,
- (ii) $\int_0^\infty t e^{|\alpha_\epsilon^-| t} L(dt) < \infty$,
- (iii) $\int_0^\infty t e^{|\alpha_\epsilon^-| t} \mu_\epsilon^-(dt) < \infty$ and
- (iv) $\mathbb{E} \left[\int_0^\infty e^{|\alpha_\epsilon^-| t} \hat{\xi}_\epsilon^-(dt) \log^+(\xi(\infty)) \right] < \infty$.

After that we show that the corresponding results with the $-$ superscript replaced by $+$ easily follow.

All expectations in $K_+(\epsilon)$ and $K_-(\epsilon)$ as defined in (40) and (42), are finite by Claim 6.1 and continuous in ϵ . Furthermore, $K_+(\epsilon)$ is clearly increasing in ϵ , while $K_-(\epsilon)$ is decreasing. So, $\mu_\epsilon^+(dt)$ decreases for every $t > 0$ to $\mu^*(dt)$ as $\epsilon \searrow 0$ and $\mu_\epsilon^-(dt)$ increases for every $t > 0$ to $\mu^*(dt)$ as $\epsilon \searrow 0$.

By the conditions in Theorem 2.1, we know that there exist $c > 1$ and $\alpha_c < 0$ such that $c = \int_0^\infty e^{-\alpha_c t} \mu^*(dt)$. Assume that ϵ is small enough for

$K_-(\epsilon) \in (1/c, 1)$. It follows that

$$\int_0^\infty e^{-\alpha_c t} \mu_\epsilon^-(dt) = \int_0^\infty e^{-\alpha_c t} K_-(\epsilon) \mu^*(dt) = cK_-(\epsilon) > 1.$$

From $\int_0^\infty e^{-\alpha_c t} \mu^*(dt) = c < \infty$, we deduce that $g(x) := \int_0^\infty e^{-xt} \mu_\epsilon^-(dt)$ is continuous and decreasing in x on $[\alpha_c, \infty)$, with $g(\alpha_c) = cK_-(\epsilon) > 1$ and $g(0) = \mu_\epsilon^-(\infty) < 1$. So there exists $\alpha_\epsilon^- < 0$ such that $1 = \int_0^\infty e^{-\alpha_\epsilon^- t} \mu_\epsilon^-(dt)$.

The justification that there exists $\alpha_\epsilon^+ < 0$ such that $1 = \int_0^\infty e^{-\alpha_\epsilon^+ t} \mu_\epsilon^+(dt)$, runs entirely parallel.

To show that $\int_0^\infty t e^{|\alpha_\epsilon^-|t} L(dt) < \infty$, we observe that because $K_-(\epsilon)$ is decreasing and continuous in ϵ , α_ϵ^- is negative and continuous and decreasing in ϵ as long as $\mu_\epsilon^-(\infty) < 1$. Therefore, $|\alpha_\epsilon^-|$ is increasing in ϵ and $|\alpha_\epsilon^-| \searrow \alpha^*$ as $\epsilon \searrow 0$. In particular, for $\epsilon > 0$ but small enough we have that $|\alpha_\epsilon^-| < |\alpha^*| + \eta$, where η is as in the assumption of Theorem 2.1. The desired result now follows immediately from this assumption.

Observe that if $\int_0^\infty e^{(|\alpha_\epsilon^-| + \eta)t} \mu_\epsilon^-(dt) < \infty$ for some $\eta > 0$, then we also have that $\int_0^\infty t e^{|\alpha_\epsilon^-|t} \mu_\epsilon^-(dt) < \infty$. Then setting $\eta = \beta > 0$ gives

$$\begin{aligned} \int_0^\infty e^{(|\alpha_\epsilon^-| + \eta)t} \mu_\epsilon^-(dt) &= K_-(\epsilon) \int_0^\infty e^{(|\alpha_\epsilon^-| + \beta)t} \beta e^{-\beta t} \mathbb{P}(L > t) dt \\ &= \frac{K_-(\epsilon)}{|\alpha_\epsilon^-|} \int_0^\infty t e^{|\alpha_\epsilon^-|t} \beta L(dt). \end{aligned}$$

From $\int_0^\infty t e^{|\alpha_\epsilon^-|t} L(dt) < \infty$ it then follows that $\int_0^\infty t e^{|\alpha_\epsilon^-|t} \mu_\epsilon^-(dt)$, so the third condition of Lemma 3.2 is satisfied.

Finally,

$$\begin{aligned} &\mathbb{E} \left[\int_0^\infty e^{|\alpha_\epsilon^-|t} \hat{\xi}_\epsilon^-(dt) \log^+(\hat{\xi}_\epsilon^-(\infty)) \right] \\ &\leq \mathbb{E} \left[\int_0^\infty e^{|\alpha_\epsilon^-|t} \hat{\xi}_\epsilon^-(dt) \log^+(\tilde{D}^-(\epsilon) - 1) \right] \\ &= \mathbb{E}[(\tilde{D}^-(\epsilon) - 1) \log^+(\tilde{D}^-(\epsilon) - 1)] \kappa_-(\epsilon) \int_0^\infty e^{|\alpha_\epsilon^-|t} \beta e^{-\beta t} P(L > t) dt \\ &\leq \mathbb{E}[(\tilde{D}^-(\epsilon) - 1)^2] \kappa_-(\epsilon) \int_0^\infty e^{|\alpha_\epsilon^-|t} \beta e^{-\beta t} P(L > t) dt \\ &= \frac{\mathbb{E}[(\tilde{D}^-(\epsilon) - 1)^2]}{\mathbb{E}[\tilde{D}^-(\epsilon) - 1]} \int_0^\infty e^{|\alpha_\epsilon^-|t} \mu_\epsilon^-(dt). \end{aligned}$$

It follows from Claim 6.1 that the quotient of expectations is finite, while the integral is finite by step (iii). So assumption (iv) is met. Conditions (ii)-(iv) are also satisfied for $\{\hat{\xi}_\epsilon^+(t); t \geq 0\}$ together with $\{\mu_\epsilon^+(t); t \geq 0\}$, because $|\alpha_\epsilon^+| < |\alpha_\epsilon^-|$ and $\mu_\epsilon^+(dt)/\mu_\epsilon^-(dt) \leq K_+(\epsilon)/K_-(\epsilon) < \infty$.

Step 3:

Let $\gamma \in (0, 1 - q^*)$ and $\gamma' \in (\gamma, 1 - q^*)$. By the definition

$$T'_\gamma(n) = \inf\{t > 0; n^{-1}|S^{(n)}(t)| < 1 - \gamma\},$$

we obtain that if $T'_{\gamma'}(n) < \infty$, then $n^{-1}|S^{(n)}(t)| \geq 1 - \gamma'$ for $t < T'_{\gamma'}(n)$ and in particular, $n^{-1}|S^{(n)}(T'_{\gamma'}(n))| \geq 1 - \gamma'$. Combined with Lemma 2.2 this gives that $|S^{(n)}(T'_{\gamma'}(n))| - |S^{(n)}(\infty)| = \theta(n)$ w.h.p.

For $t > 0$, let $V_{\dagger}^{(n)}(t) \subset V_*^{(n)}(t)$ be the set of vertices infected after time t , which are infected by vertices infected before time t (in the language of [9], $V_{\dagger}^{(n)}(t)$ is the coming generation). So, $V_{\dagger}^{(n)}(t)$ is the subset of $V_*^{(n)}(t)$ of which the infecting vertex is not in $V_*^{(n)}(t)$. Assume that $|V_{\dagger}^{(n)}(t)| = o(n)$. From Step 1 we know that $|V_*^{(n)}(t)|$ is stochastically smaller than the total progeny of $|V_{\dagger}^{(n)}(t)|$ sub-critical branching processes with mean offspring measure μ_{ϵ}^+ and thus expected total number of children per particle $\mu_{\epsilon}^+(\infty)$. However the total size of such a branching process has expected size $(1 - \mu_{\epsilon}^+(\infty))^{-1} = \theta(1)$. This implies that if $|V_{\dagger}^{(n)}(T'_{\gamma'}(n))| = o(n)$, then $\mathbb{E}[|V_*^{(n)}(T'_{\gamma'}(n))|] = o(n)$, which implies that $|V_*^{(n)}(T'_{\gamma'}(n))| = |S^{(n)}(T'_{\gamma'}(n))| - |S^{(n)}(\infty)| = o(n)$ w.h.p., which is a contradiction. This finishes step 3.

Step 4:

Let $\delta \in (0, |\alpha^*|)$. We can choose $\epsilon > 0$, such that α_{ϵ}^+ exists and $|\alpha_{\epsilon}^+| \in (|\alpha^*| - \delta, |\alpha^*|)$. Furthermore, by Lemma 6.1 we know that $t_1^{(n)}(\epsilon) < T^*(n)$ w.h.p. and we may w.h.p. choose γ such that $T'_{\gamma}(n) \in (t_1^{(n)}(\epsilon), T^*(n))$. For $v \in V_{\dagger}^{(n)}(T'_{\gamma}(n))$ let u be the ‘‘infector’’ of v . We consider the difference in infection time

$$\sigma(v) - \sigma(u) = (T'_{\gamma}(n) - \sigma(u)) + (\sigma(v) - T'_{\gamma}(n)).$$

Let τ be an exponentially distributed random variable with parameter β and let L_v be the infectious period of v , which is distributed as L and independent of τ . Then,

$$\begin{aligned} & \mathbb{P}(\sigma(v) - T'_{\gamma}(n) > x | v \in V_{\dagger}^{(n)}(T'_{\gamma}(n))) \\ &= \mathbb{P}(\sigma(v) - T'_{\gamma}(n) > x | \sigma(v) - \sigma(u) > T'_{\gamma}(n) - \sigma(u), \sigma(v) - \sigma(u) < \infty) \\ &= \mathbb{P}(\sigma(v) - \sigma(u) > x + T'_{\gamma}(n) - \sigma(u) | \sigma(v) - \sigma(u) > T'_{\gamma}(n) - \sigma(u), \sigma(v) - \sigma(u) < \infty) \\ &= \mathbb{P}(\tau > x + T'_{\gamma}(n) - \sigma(u) | \tau > T'_{\gamma}(n) - \sigma(u), L > \tau) \\ &= \mathbb{P}(\tau > x + T'_{\gamma}(n) - \sigma(u) | \tau > T'_{\gamma}(n) - \sigma(u)) \frac{\mathbb{P}(L > \tau | \tau > x + T'_{\gamma}(n) - \sigma(u), \tau > T'_{\gamma}(n) - \sigma(u))}{\mathbb{P}(L > \tau | \tau > T'_{\gamma}(n) - \sigma(u))} \\ &= \mathbb{P}(\tau > x) \frac{\mathbb{P}(L > \tau | \tau > x + T'_{\gamma}(n) - \sigma(u))}{\mathbb{P}(L > \tau | \tau > T'_{\gamma}(n) - \sigma(u))} \\ &= \mathbb{P}(\tau > x) \frac{\mathbb{P}(L > \tau + x + T'_{\gamma}(n) - \sigma(u))}{\mathbb{P}(L > \tau + T'_{\gamma}(n) - \sigma(u))} \\ &= \mathbb{P}(\tau > x) \mathbb{P}(L > \tau + x + T'_{\gamma}(n) - \sigma(u) | L > \tau + T'_{\gamma}(n) - \sigma(u)), \end{aligned} \tag{43}$$

where we used the memoryless property of the exponential distribution in the last two steps. Now let s_0 be such that $\mathbb{P}(L > t + s) < e^{-|\alpha^*|(s-s_0)}\mathbb{P}(L > t)$

for all t and s (which is possible by assumption (iii) of Lemma 2.1). Then (43) together with τ being exponentially distributed with expectation $1/\beta$ gives,

$$\begin{aligned} \mathbb{P}(\sigma(v) - T'_\gamma(n) > x | v \in V_\dagger^{(n)}(T'_\gamma(n))) \\ < \mathbb{P}(\tau > x) e^{-|\alpha^*|(x-s_0)} = e^{-(|\alpha^*|+\beta)x} e^{|\alpha^*|s_0}. \end{aligned} \quad (44)$$

Next observe that for $t > T'_\gamma(n)$,

$$|I^{(n)}(t)| = \sum_{v \in V_\dagger^{(n)}(T'_\gamma(n))} |J_v(t - \sigma(v))| + |I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))|. \quad (45)$$

This implies that

$$\begin{aligned} \mathbb{P}(|I^{(n)}(t)| > 0) \\ \leq \sum_{v \in V_\dagger^{(n)}(T'_\gamma(n))} \mathbb{P}(|J_v(t - \sigma(v))| > 0) + \mathbb{P}(|I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))| > 0). \end{aligned} \quad (46)$$

Recall that $\{J_v(t); t \geq 0\}$ is dominated by a branching process with mean offspring measure $\{\mu_\epsilon^+(t); t \geq 0\}$. Consider a sequence of i.i.d. copies of this process and let $Z_{\epsilon,k}^+(t)$ be the number of alive particles in the k -th copy of this branching process at time t . So, (46) implies

$$\begin{aligned} \mathbb{P}(|I^{(n)}(t)| > 0) \\ \leq \sum_{k=1}^{|V_\dagger^{(n)}(T'_\gamma(n))|} \mathbb{P}(|Z_{\epsilon,k}^+(t - \sigma(v))| > 0) + \mathbb{P}(|I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))| > 0). \end{aligned} \quad (47)$$

Lemma 3.2 and Step 2 above now give that for all $k \in \mathbb{N}$ there exists $c_1 \in (0, \infty)$ such that $e^{|\alpha_\epsilon^+|t} \mathbb{P}(Z_{\epsilon,k}^+(t) > 0) \rightarrow c_1$, which implies that there exists $c_2 > c_1$ such that $\mathbb{P}(Z_{\epsilon,k}^+(t) > 0) < c_2 e^{-|\alpha_\epsilon^+|t}$ for all $t > 0$. So, the probability that $v \in V_\dagger^{(n)}(T'_\gamma(n))$ still has any offspring $t_* = \frac{\log n}{|\alpha^*| - \delta}$ time units after $T'_\gamma(n)$ is bounded from above by

$$\begin{aligned} \int_0^{t_*} \mathbb{P}(|Z_{\epsilon,k}^+(t-s)| > 0) \mathbb{P}(\sigma(v) - T'_\gamma(n) \in ds | v \in V_\dagger^{(n)}(T'_\gamma(n))) \\ + \mathbb{P}(\sigma(v) - T'_\gamma(n) > t_* | v \in V_\dagger^{(n)}(T'_\gamma(n))). \end{aligned} \quad (48)$$

By (44), Lemma 3.2 and $t^* \rightarrow \infty$ as $n \rightarrow \infty$ we obtain that (48) is bounded

from above by

$$\begin{aligned}
& e^{|\alpha^*|s_0} \left(\int_0^{t_*} c_2 e^{-|\alpha_\epsilon^+|(t^*-s)} (|\alpha^*| + \beta) e^{-(|\alpha^*|+\beta)s} ds + e^{-(|\alpha^*|+\beta)t_*} \right) \\
&= e^{|\alpha^*|s_0} \left(c_2 e^{-|\alpha_\epsilon^+|t^*} \int_0^{t^*} (|\alpha^*| + \beta) e^{-(|\alpha^*| - |\alpha_\epsilon^+| + \beta)s} ds + e^{-(|\alpha^*|+\beta)t_*} \right) \\
&= e^{|\alpha^*|s_0} \left(c_2 \frac{|\alpha^*| + \beta}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} (e^{-|\alpha_\epsilon^+|t^*} - e^{-(|\alpha^*|+\beta)t_*}) + e^{-(|\alpha^*|+\beta)t_*} \right) \\
&= e^{|\alpha^*|s_0} \left(c_2 \frac{|\alpha^*| + \beta}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} \left(n^{-\frac{|\alpha_\epsilon^+|}{|\alpha^*|-\delta}} - n^{-\frac{|\alpha^*|+\beta}{|\alpha^*|-\delta}} \right) + n^{-\frac{|\alpha^*|+\beta}{|\alpha^*|-\delta}} \right) \\
&= e^{|\alpha^*|s_0} \left(c_2 \frac{|\alpha^*| + \beta}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} n^{-\frac{|\alpha_\epsilon^+|}{|\alpha^*|-\delta}} - \frac{(c_2 - 1)(|\alpha^*| + \beta) + |\alpha_\epsilon^+|}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} n^{-\frac{|\alpha^*|+\beta}{|\alpha^*|-\delta}} \right) \\
&= \frac{1}{n} e^{|\alpha^*|s_0} \left(c_2 \frac{|\alpha^*| + \beta}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} n^{-\frac{|\alpha_\epsilon^+| - (|\alpha^*| - \delta)}{|\alpha^*|-\delta}} - \frac{(c_2 - 1)(|\alpha^*| + \beta) + |\alpha_\epsilon^+|}{|\alpha^*| - |\alpha_\epsilon^+| + \beta} n^{-\frac{\delta+\beta}{|\alpha^*|-\delta}} \right)
\end{aligned}$$

Because $|\alpha_\epsilon^+| > |\alpha^*| - \delta$ this expression is $o(1/n)$. Together with $|V_{\dagger}^{(n)}(T'_\gamma(n))| = \theta(n)$, this in turn implies that the expression

$$\begin{aligned}
& |V_{\dagger}^{(n)}(T'_\gamma(n))| \\
& \sum_{k=1} \mathbb{P}(|Z_{\epsilon,k}^+(t^* - \sigma(v))| > 0)
\end{aligned}$$

as given in equation (47) is $o(1)$.

We also know that

$$\mathbb{P}(|I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))| > 0) \leq |I^{(n)}(T'_\gamma(n))| \max_{s>0} \mathbb{P}(L_u > t + s | L_u > s).$$

By assumption (iii) of Theorem 2.1 and $|I^{(n)}(T'_\gamma(n))| \leq n$ we obtain that

$$\mathbb{P}(|I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))| > 0) \leq n e^{|\alpha^*|s_0} e^{-|\alpha^*|t}.$$

So with $t_* = \frac{\log n}{|\alpha^*| - \delta}$ as above we obtain that

$$\begin{aligned}
& \mathbb{P}(|I^{(n)}(t_*) \cap I^{(n)}(T'_\gamma(n))| > 0) \\
& \leq n e^{|\alpha^*|s_0} n^{-\frac{|\alpha^*|}{|\alpha^*|-\delta}} = e^{|\alpha^*|s_0} n^{-\frac{\delta}{|\alpha^*|-\delta}} = o(1).
\end{aligned}$$

Therefore, if $t_1^{(n)}(\epsilon) < T^*(n)$ and for all $\delta \in (0, \alpha^*)$,

$$\begin{aligned}
& \mathbb{P} \left(T^*(n) - T'_\gamma(n) > \frac{\log n}{|\alpha^*| - \delta} \right) \\
& = \mathbb{P} \left(|I^{(n)} \left(T'_\gamma(n) + \frac{\log n}{|\alpha^*| - \delta} \right)| > 0 \right) = o(1)
\end{aligned}$$

as desired and Step 4 is completed.

Step 5:

Let $\delta \in (0, |\alpha^*|)$. We can choose $\epsilon > 0$, such that α_ϵ^- exists and $|\alpha_\epsilon^-| \in (|\alpha^*|, |\alpha^*|(1 + \delta/2))$. Furthermore, by Lemma 6.1 we know that $t_1^{(n)}(\epsilon) < T^*(n)$ w.h.p. and we choose γ such that $T'_\gamma(n) \in (t_1^{(n)}(\epsilon), T^*(n))$.

Observe that for $t > T'_\gamma(n)$,

$$\begin{aligned} |I^{(n)}(t)| &= \sum_{v \in V_\dagger^{(n)}(T'_\gamma(n))} |J_v(t - \sigma(v))| + |I^{(n)}(t) \cap I^{(n)}(T'_\gamma(n))| \\ &\geq \sum_{v \in V_\dagger^{(n)}(T'_\gamma(n))} |J_v(t - \sigma(v))|, \end{aligned} \quad (49)$$

where $J_v(s) = 0$ for $s < 0$.

Recall that $\{J_v(t); t \geq 0\}$ dominates a branching process with mean offspring measure $\{\mu_\epsilon^-(t); t \geq 0\}$. Consider a sequence of i.i.d. copies of this branching process and let $Z_{\epsilon,k}^-(t)$ be the number of alive particles in the k -th copy of this process at time t . So, $|I^{(n)}(t)|$ is stochastically larger than $\sum_{k=1}^{|V_\dagger^{(n)}(T'_\gamma(n))|} Z_{\epsilon,k}^-(t - \sigma(v))$. By the independence of the branching processes we then obtain that

$$\begin{aligned} \mathbb{P}(|I^{(n)}(T'_\gamma(n) + t)| = 0) &\leq \prod_{k=1}^{|V_\dagger^{(n)}(T'_\gamma(n))|} \mathbb{P}(Z_{\epsilon,k}^-(T'_\gamma(n) + t - \sigma(v)) = 0) \\ &\leq \prod_{k=1}^{|V_\dagger^{(n)}(T'_\gamma(n))|} \mathbb{P}(Z_{\epsilon,k}^-(t) = 0) = \left(\mathbb{P}(Z_{\epsilon,1}^-(t) = 0)\right)^{|V_\dagger^{(n)}(T'_\gamma(n))|} \\ &= \mathbb{P}\left(Z_{\epsilon,1}^-(t) = 0 \mid Z_{\epsilon,1}^-(0) = |V_\dagger^{(n)}(T'_\gamma(n))|\right). \end{aligned} \quad (50)$$

For the second inequality we used that $\{Z_{\epsilon,k}^-(t) = 0\}$ is increasing in t . Now we can apply Corollary 3.2, which gives that for all $\delta > 0$

$$\mathbb{P}\left(Z_{\epsilon,1}^- \left((1 - \delta/3) \frac{\log |V_\dagger^{(n)}(T'_\gamma(n))|}{|\alpha_\epsilon^-|} \right) = 0 \mid Z_{\epsilon,1}^-(0) = |V_\dagger^{(n)}(T'_\gamma(n))| \right) \rightarrow 0.$$

By step 3 we obtain that $\log |V_\dagger^{(n)}(T'_\gamma(n))| = \theta(n)$, which implies that

$$\mathbb{P}\left(Z_{\epsilon,1}^- \left((1 - \delta/2) \frac{\log n}{|\alpha_\epsilon^-|} \right) = 0 \mid Z_{\epsilon,1}^-(t) = |V_\dagger^{(n)}(T'_\gamma(n))| \right) \rightarrow 0.$$

and thus by (50) we obtain

$$\mathbb{P} \left(|I^{(n)} \left(T'_\gamma(n) + (1 - \delta/2) \frac{\log n}{|\alpha_\epsilon^-|} \right) | = 0 \right) \rightarrow 0.$$

By $|\alpha_\epsilon^-| < |\alpha^*|(1 + \delta/2)$ we then obtain

$$\mathbb{P} \left(|I^{(n)} \left(T'_\gamma(n) + \frac{1 - \delta/2 \log n}{1 + \delta/2 |\alpha_\epsilon^-|} \right) | = 0 \right) \rightarrow 0.$$

Because $1 - \delta < \frac{1 - \delta/2}{1 + \delta/2}$, this implies that

$$\mathbb{P} \left(|I^{(n)} \left(T'_\gamma(n) + (1 - \delta) \frac{\log n}{|\alpha_\epsilon^-|} \right) | = 0 \right) \rightarrow 0,$$

which in turn leads to

$$\begin{aligned} \mathbb{P} \left(T^*(n) - T'_\gamma(n) < (1 - \delta) \frac{\log n}{|\alpha^*|} \right) \\ = \mathbb{P} \left(|I^{(n)} \left(T'_\gamma(n) + (1 - \delta) \frac{\log n}{|\alpha^*|} \right) | = 0 \right) \rightarrow 0. \end{aligned}$$

□

References

- [1] Amini, H. and Lelarge, M. (2015). The diameter of weighted random graphs. *The Annals of Applied Probability*, 25(3):1686–1727.
- [2] Andersson, H. (1999). Epidemic models and social networks. *Mathematical Scientist*, 24(2):128–147.
- [3] Ball, F. and Lyne, O. D. (2001). Stochastic multi-type SIR epidemics among a population partitioned into households. *Advances in Applied Probability*, 33(01):99–123.
- [4] Ball, F., Sirl, D., and Trapman, P. (2009). Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Advances in Applied Probability*, 41(03):765–796.
- [5] Ball, F., Sirl, D., and Trapman, P. (2014). Epidemics on random intersection graphs. *The Annals of Applied Probability*, 24(3):1081–1128.

- [6] Barbour, A. D. (1975). The duration of the closed stochastic epidemic. *Biometrika*, 62(2):477–482.
- [7] Barbour, A. D. and Reinert, G. (2013). Approximating the epidemic curve. *Electron. J. Probab*, 18(54):1–30.
- [8] Bhamidi, S., van der Hofstad, R., and Hooghiemstra, G. (2017). Universality for first passage percolation on sparse random graphs. *The Annals of Probability*, 45(4):2568–2630.
- [9] Bhamidi, S., Van Der Hofstad, R., and Komjáthy, J. (2014). The front of the epidemic spread and first passage percolation. *Journal of Applied Probability*, 51(A):101–121.
- [10] Britton, T. (2010). Stochastic epidemic models: a survey. *Mathematical biosciences*, 225(1):24–35.
- [11] Diekmann, O., Heesterbeek, H., and Britton, T. (2012). *Mathematical tools for understanding infectious disease dynamics*. Princeton University Press.
- [12] Durrett, R. (2007). *Random graph dynamics*. Cambridge university press Cambridge.
- [13] Grimmett, G. and Stirzaker, D. (2001). *Probability and random processes*. Oxford university press, third edition.
- [14] Van der Hofstad, R. (2016). *Random graphs and complex networks*, volume 1. Cambridge University Press.
- [15] Iksanov, A. and Meiners, M. (2015). Rate of convergence in the law of large numbers for supercritical general multi-type branching processes. *Stochastic Processes and their Applications*, 125(2):708–738.
- [16] Jagers, P. (1975). *Branching processes with biological applications*. Wiley.
- [17] 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(4):726–763.
- [18] Janson, S., Luczak, T., and Rucinski, A. (2011). *Random graphs*, volume 45. John Wiley & Sons.

- [19] Ludwig, D. (1975). Final size distribution for epidemics. *Mathematical Biosciences*, 23(1-2):33–46.
- [20] Nerman, O. (1981). On the convergence of supercritical general (cmj) branching processes. *Probability Theory and Related Fields*, 57(3):365–395.
- [21] Newman, M. E. (2002). Spread of epidemic disease on networks. *Physical review E*, 66(1):016128.
- [22] Pellis, L., Ball, F., and Trapman, P. (2012). Reproduction numbers for epidemic models with households and other social structures. I. definition and calculation of R_0 . *Mathematical biosciences*, 235(1):85–97.
- [23] Resnick, S. I. (2013). *Adventures in stochastic processes*. Springer Science & Business Media.
- [24] Rudin, W. (1964). *Principles of mathematical analysis*, volume 3. McGraw-hill New York.
- [25] Volz, E. (2008). SIR dynamics in random networks with heterogeneous connectivity. *Journal of mathematical biology*, 56(3):293–310.