LAW OF LARGE NUMBERS FOR THE SIR EPIDEMIC ON A RANDOM GRAPH WITH GIVEN DEGREES

SVANTE JANSON, MALWINA LUCZAK, PETER WINDRIDGE

ABSTRACT. We study the susceptible-infective-recovered (SIR) epidemic on a random graph chosen uniformly subject to having given vertex degrees. In this model infective vertices infect each of their susceptible neighbours, and recover, at a constant rate.

Suppose that initially there are only a few infective vertices. We prove there is a threshold for a parameter involving the rates and vertex degrees below which only a small number of infections occur. Above the threshold a large outbreak occurs with probability bounded away from zero. Our main result is that, conditional on a large outbreak, the evolutions of certain quantities of interest, such as the fraction of infective vertices, converge to deterministic functions of time.

We also consider more general initial conditions for the epidemic, and derive criteria for a simple vaccination strategy to be successful.

In contrast to earlier results for this model, our approach only requires basic regularity conditions and a uniformly bounded second moment of the degree of a random vertex.

En route, we prove analogous results for the epidemic on the configuration model multigraph under much weaker conditions. Essentially, our main result requires only that the initial values for our processes converge, i.e. it is the best possible.

1. INTRODUCTION

The Markovian SIR process is a simple model for a disease spreading around a finite population in which each individual is either susceptible, infective or recovered. Individuals are represented by vertices in a graph G with edges corresponding to potentially infectious contacts. Infective vertices become recovered at rate $\rho \geqslant 0$ and infect each neighbour at rate $\beta > 0$; those are the only possible transitions, i.e. recovered vertices never become infective.

The applicability, behaviour and tractability of the model depends heavily on how G is chosen. In classical formulations G is the complete graph (see [14] for a historical account) but gradually attention has shifted towards more realistic models where individuals may vary in how many contacts they have.

Particular interest has focused on the case that G itself is random. Several families of random graph have been considered, such as Erdős–Rényi $G(n, p)$ graphs [33], those with local household structure [5] and other forms of clustering [11], see the recent survey [20].

The present paper concerns SIR epidemics on random graphs with a given degree sequence. These random graphs are commonly used to model the internet, scientific collaboration networks and sexual contact networks [35; 34; 4] (and the references therein). Random graphs with given degree sequence are normally constructed via the configuration model,

introduced by Bollobás, see [9]. In recent years, various properties of these graphs have been studied, such as the appearance and size of a giant component [31; 32; 26], as well as the near-critical behaviour [29; 27]. Other quantities investigated include the size of the k-core [24; 25; 40], diameter [18], chromatic number [19] and matching number [7].

There have been a number of studies of SIR epidemics on random graphs with a given degree sequence. A set of non-linear ordinary differential equations summarising the time evolution of the epidemic were obtained heuristically by Volz [41]. Another non-rigorous derivation of these equations is given in [30].

Decreusefond et al. [16] study a measure-valued process describing the degrees of susceptible individuals and the number of edges between different types of vertices. They prove that, as the population size grows to infinity, the measure-valued process converges to a deterministic limit, from which the Volz equations may be derived as a corollary. The results in [16] are proven under the conditions that the fifth moment of the degree of a random vertex is uniformly bounded, and that, asymptotically, the proportion of vertices infective at time zero is positive.

Bohman and Picollelli [8] study the SIR process dynamics on the configuration model with bounded vertex degrees, starting from a single infective. They use a multitype branching process approximation for both the early and final stages of the epidemic. The middle phase of the epidemic, while there are at least a moderate number of infectives, is analysed using Wormald's differential equations method.

Barbour and Reinert [6] use multitype branching process approximations to prove results approximating the entire course of an SIR epidemic within a more general non-Markovian framework, allowing degree dependent infection and recovery time distributions. A result for graphs with a given degree sequence with bounded vertex degrees follows as a corollary.

See also [13] for the SIS epidemic process on a random graph with given degrees, which exhibits very different behaviour compared to the SIR epidemic studied here.

Our contribution. In this paper we analyse the SIR epidemic on graphs with a given degree sequence for an arbitrary number of initially infective vertices, assuming only basic regularity conditions and uniform boundedness of the second moment of the degree distribution. This contrasts with the earlier works mentioned above, which require either uniform boundedness of the fifth moment [16], or uniformly bounded degrees [8; 6]. In our proof we study the configuration model epidemic under the weaker condition that the degree of a randomly chosen susceptible vertex is uniformly integrable. This is the best possible condition for a 'law of large numbers' result in the spirit of [41; 30; 16; 8], since it amounts to convergence of the average number of susceptible contacts at the epidemic's epoch; see Remark 2.2. Our approach extends techniques of [24; 26] and leads to fairly simple proofs.

The rest of the paper is laid out as follows. In Section 2, we define the model and notation, and state our assumptions and results. In Section 3 we consider a time-changed version of the epidemic, as a tool to be used in our proofs. In Section 4, we prove our results for multigraphs with a given degree sequence defined by the configuration model. In Section 5, we study more carefully the probability of a large outbreak and the size of a small outbreak, obtaining more detailed forms of statements in Theorem 2.9(i) and (ii)(c). In Section 6, we transfer the results from multigraphs to simple graphs with a given degree sequence. In Section 7 we discuss briefly what happens when the second moment of the degree of a random vertex is not uniformly bounded. Section 8 contains a few remarks on the random

time shift used in our proof. Appendix A contains a technical lemma on the time change in Section 3. Appendix B is a summary of the main notation used in the paper.

2. Model, notation, assumptions and results

For $n \in \mathbb{N}$ and a sequence $(d_i)_1^n$ of non-negative integers, let $G = G(n, (d_i)_1^n)$ be a simple graph (i.e. with no loops or double edges) on n vertices, chosen uniformly at random from among all graphs with degree sequence $(d_i)_1^n$. (We tacitly assume that there is some such graph, so $\sum_{i=1}^n d_i$ must be even, at least.)

Given the graph G, the SIR epidemic evolves as a continuous-time Markov chain. At any time, each vertex is either susceptible, infected or recovered. Each infective vertex recovers at rate $\rho \geqslant 0$ and also infects each susceptible neighbour at rate $\beta > 0$.

We assume that there are initially n_S, n_I, and n_R susceptible, infective and recovered vertices, respectively. Further, we assume that, for each $k \geqslant 0$, there are respectively $n_{\text{S},k}$, $n_{\text{I},k}$ and $n_{\text{R},k}$ of these vertices with degree k. Thus, $n_\text{S} + n_\text{I} + n_\text{R} = n$ and $n_\text{S} = \sum_{k=0}^\infty n_{\text{S},k}$, $n_\text{I} = \sum_{k=0}^\infty n_{\text{I},k}$, $n_\text{R} = \sum_{k=0}^\infty n_{\text{R},k}$. We write n_k to denote the total number of vertices with degree k; thus, for each k, $n_k = n_{\text{S},k} + n_{\text{I},k} + n_{\text{R},k}$. Note that all these parameters, as well as the sequence $(d_i)_1^n$, depend on the number n of vertices, although we omit explicit mention of this in the notation. For technical reasons, note that they do not have to be defined for all integers n; a subsequence is enough.

We consider asymptotics as $n \to \infty$, and all unspecified limits below are as $n \to \infty$. Throughout the paper we use the notation o_p in a standard way. That is, for a sequence of random variables $(Y^{(n)})_1^\infty$ and real numbers $(a_n)_1^\infty$, '$Y^{(n)} = o_p(a_n)$' means $Y^{(n)}/a_n \xrightarrow{\text{P}} 0$. Similarly, $Y^{(n)} = O_\text{p}(1)$ means that for every $\varepsilon > 0$ there exists K_ε such that $\mathbb{P}(|Y^{(n)}| > K_\varepsilon) < \varepsilon$ for all n. For a sequence $(Y_t^{(n)})_1^\infty$ of real-valued stochastic processes defined on a subset E of $\mathbb{R}$ and a real-valued function y on E, '$Y_t^{(n)} \xrightarrow{\text{P}} y(t)$ uniformly on $E' \subseteq E$' means $\sup_{t \in E'} |Y_t^{(n)} - y(t)| \xrightarrow{\text{P}} 0$. Given a sequence of events $(\mathcal{E}_n)_1^\infty$, event $\mathcal{E}_n$ is said to hold w.h.p. (with high probability) if the probability of $\mathcal{E}_n$ converges to 1.

We assume the following regularity conditions for the degree sequence asymptotics.

(D1) The fractions of initially susceptible, infective and recovered vertices converge to some $\alpha_\text{S}, \alpha_\text{I}, \alpha_\text{R} \in [0, 1]$, i.e.

$$n_\text{S}/n \to \alpha_\text{S}, \qquad n_\text{I}/n \to \alpha_\text{I}, \qquad n_\text{R}/n \to \alpha_\text{R}. \qquad (2.1)$$

Further, $\alpha_\text{S} > 0$.

(D2) The degree of a randomly chosen susceptible vertex converges to a probability distribution $(p_k)_0^\infty$, i.e.

$$n_{\text{S},k}/n_\text{S} \to p_k, \qquad k \geqslant 0. \qquad (2.2)$$

Further, this limiting distribution has a finite and positive mean

$$\lambda := \sum_{k=0}^{\infty} k p_k \in (0, \infty). \tag{2.3}$$

(D3) The average degree of a randomly chosen susceptible vertex converges to λ, i.e.

$$\sum_{k=0}^{\infty} k n_{\mathrm{S},k}/n_{\mathrm{S}} \to \lambda. \tag{2.4}$$

(D4) The average degree over all vertices converges to $\mu > 0$, i.e.

$$\sum_{k=0}^{\infty} k n_k/n = \sum_{i=1}^{n} d_i/n \to \mu, \tag{2.5}$$

and, in more detail, for some $\mu_{\mathrm{S}}, \mu_{\mathrm{I}}, \mu_{\mathrm{R}}$,

$$\sum_{k=0}^{\infty} k n_{\mathrm{S},k}/n \to \mu_{\mathrm{S}}, \tag{2.6}$$

$$\sum_{k=0}^{\infty} k n_{\mathrm{I},k}/n \to \mu_{\mathrm{I}}, \qquad \sum_{k=0}^{\infty} k n_{\mathrm{R},k}/n \to \mu_{\mathrm{R}}. \tag{2.7}$$

(D5) The maximum degree of the initially infective vertices is not too large:

$$\max\{k : n_{\mathrm{I},k} > 0\} = o(n). \tag{2.8}$$

(D6) Either $p_1 > 0$ or $\rho > 0$ or $\mu_{\mathrm{R}} > 0$.

Remark 2.1. Obviously, $\alpha_{\mathrm{S}} + \alpha_{\mathrm{I}} + \alpha_{\mathrm{R}} = 1$ and $\mu_{\mathrm{S}} + \mu_{\mathrm{I}} + \mu_{\mathrm{R}} = \mu$. Further, assumptions (D1)–(D3) imply $\sum_{k=0}^{\infty} k n_{\mathrm{S},k}/n \to \alpha_{\mathrm{S}} \lambda$. Thus, $\mu_{\mathrm{S}} = \alpha_{\mathrm{S}} \lambda$ and (2.6) is redundant.

The assumptions $\alpha_{\mathrm{S}} > 0$ in (D1) and $\lambda > 0$ in (D2) mean that there are initially a significant number of susceptibles with non-zero degree. They are included to avoid trivialities, and, in particular, imply that $n_{\mathrm{S}} \geqslant 1$ for large enough n.

Remark 2.2. Assumptions (D1)–(D3) together imply that $\sum_{k=0}^{\infty} k n_{\mathrm{S},k}/n$ is uniformly summable, i.e. for any $\varepsilon > 0$ there exists K such that $\sum_{k=K+1}^{\infty} k n_{\mathrm{S},k}/n < \varepsilon$ for n large enough. Conversely, (D1), (D2) and uniform summability of $\sum_{k=0}^{\infty} k n_{\mathrm{S},k}/n$ imply (D3).

Remark 2.3. In particular, the uniform summability in Remark 2.2 implies that $\max\{k : n_{\mathrm{S},k} > 0\} = o(n)$. This and assumption (D5) imply, using (2.5),

$$\sum_{k=0}^{\infty} k^2 (n_{\mathrm{S},k} + n_{\mathrm{I},k}) \leqslant o(n) \sum_{k=0}^{\infty} k n_k = o(n^2). \tag{2.9}$$

Conversely, (2.9) implies (2.8). We do not need the corresponding condition for initially recovered vertices, but since these only play a passive role, it would be essentially no loss of generality to assume that the maximum degree of all vertices $\max_i d_i = \max\{k : n_k > 0\} = o(n)$.

It will be convenient for us to work with multigraphs, that is to allow loops and multiple edges. Let $G^*(n, (d_i)_1^n)$ be the random multigraph with given degree sequence $(d_i)_1^n$ defined by the configuration model: we take a set of d_i half-edges for each vertex i and combine half-edges into edges by a uniformly random matching (see e.g. [9]). Conditioned on the multigraph being simple, we obtain $G = G(n, (d_i)_1^n)$, the uniformly distributed random graph with degree sequence $(d_i)_1^n$.

The configuration model has been used in the study of epidemics in a number of earlier works, see, for example, [1; 4; 12; 16; 8].

We prove our results for the SIR epidemic on G^*, and, by conditioning on G^* being simple, we deduce that these results also hold for the SIR epidemic on G . Our argument relies on the probability that G^* is simple being bounded away from zero as $n \to \infty$. By the main theorem of [21] this occurs provided the following condition holds.

(G1) The degree of a randomly chosen vertex has a bounded second moment, i.e.

$$\sum_{k=0}^{\infty} k^2 n_k = O(n). \tag{2.10}$$

Remark 2.4. Assumption (G1) implies that the distribution $(p_k)_0^\infty$ has a finite second moment, i.e. $\sum_{k=0}^{\infty} k^2 p_k < \infty$. Note also that (G1) implies (2.9) and thus (D5).

Remark 2.5. Although we use (G1) in order to draw conclusions for the simple graph G, we suspect that the results hold even without it. Bollobás and Riordan [10] have recently shown results for a related problem (the size of the giant component in G) from the multigraph case without using (G1); they show that even if the probability that the multigraph is simple is almost exponentially small, the error probabilities in their case are even smaller. We have not attempted doing anything similar here.

We study the SIR epidemic on the multigraph G^*, revealing its edges dynamically while the epidemic spreads. To be precise, we call a half-edge free if it is not yet paired to another half-edge. We start with d_i half-edges attached to vertex i and all half-edges free. We call a half-edge susceptible, infective or recovered according to the type of vertex it belongs to.

Now, each free infective half-edge chooses a free half-edge at rate β, uniformly at random from among all the other free half-edges. Together the pair form an edge, and are removed from the pool of free half-edges. If the chosen half-edge belongs to a susceptible vertex then that vertex becomes infective. Infective vertices also recover at rate ρ.

We stop the above process when there are no free infective half-edges, at which point the epidemic stops spreading. Some infective vertices may remain but they will recover at i.i.d. exponential times without affecting any other vertex. In any case, they turn out to be irrelevant for our purposes. Some susceptible and recovered half-edges may also remain, and these are paired off uniformly at time ∞ to reveal the remaining edges in G^*. This step is unimportant for the spread of the epidemic, but we perform it for the purpose of transferring our results to the simple graph G.

Clearly, if all the pairings are completed then the resulting graph is the multigraph G^*. Moreover, the quantities of interest (numbers of susceptible, infective and recovered vertices at each time t) have the same distribution as if we were to reveal the multigraph G^* first and run the SIR epidemic on G^* afterwards.

For $t \geqslant 0$, let S_t, I_t and R_t denote the numbers of susceptible, infective and recovered vertices, respectively, at time t. Thus S_t is decreasing and R_t is increasing. Also $S_0 = n_{\mathrm{S}}$, $I_0 = n_{\mathrm{I}}$ and $R_0 = n_{\mathrm{R}}$.

For the dynamics described above (with half-edges paired off dynamically, as needed), for $t \geqslant 0$, let $X_{\mathrm{S},t}$, $X_{\mathrm{I},t}$ and $X_{\mathrm{R},t}$ be the number of free susceptible, infective and recovered half-edges at time t, respectively. Thus $X_{\mathrm{S},t}$ is decreasing, $X_{\mathrm{S},0} = \sum_{k=0}^{\infty} k n_{\mathrm{S},k}$, $X_{\mathrm{I},0} = \sum_{k=0}^{\infty} k n_{\mathrm{I},k}$ and $X_{\mathrm{R},0} = \sum_{k=0}^{\infty} k n_{\mathrm{R},k}$. The variables $X_{\mathrm{S},t}$, $X_{\mathrm{I},t}$ and $X_{\mathrm{R},t}$ are convenient tools for the analysis of S_t, I_t and R_t, but not 'observable' quantities. (They have no interpretation for the version of the SIR process on G^* where the multigraph is constructed upfront.) For the uniformly random graph G with degree sequence $(d_i)_1^n$, the variables $X_{\mathrm{S},t}$, $X_{\mathrm{I},t}$ and $X_{\mathrm{R},t}$, for $t \geqslant 0$, are defined as above conditioned on the final multigraph G^* being a simple graph.

2.1. Results.
We will show that, upon suitable scaling, the processes $S_t, I_t, R_t, X_{\mathrm{S},t}, X_{\mathrm{I},t}, X_{\mathrm{R},t}$ converge to deterministic functions. The limiting functions will be written in terms of a parameterisation $\theta_t \in [0,1]$ of time solving an ordinary differential equation given below. The function θ_t can be interpreted as the limiting probability that a given initially susceptible half-edge has not been paired with a (necessarily infective) half-edge by time t. This means that the probability that a given degree k initially susceptible vertex is still susceptible at time t is asymptotically close to θ_t^k. With this in mind, we define the function v_{S} by

$$v_{\mathrm{S}}(\theta) := \alpha_{\mathrm{S}} \sum_{k=0}^{\infty} p_k \theta^k, \qquad \theta \in [0,1], \tag{2.11}$$

so the limiting fraction of susceptible vertices is $v_{\mathrm{S}}(\theta_t)$ at time t. Similarly, for the number of susceptible half-edges we define

$$h_{\mathrm{S}}(\theta) := \alpha_{\mathrm{S}} \sum_{k=0}^{\infty} k \theta^k p_k = \theta v_{\mathrm{S}}'(\theta), \qquad \theta \in [0,1]. \tag{2.12}$$

For the total number of free half-edges, we let

$$h_X(\theta) := \mu \theta^2, \qquad \theta \in [0,1]. \tag{2.13}$$

The intuition here is that two free half-edges disappear each time an edge is formed by pairing, so a random free half-edge is paired with intensity twice the intensity of a susceptible free half-edge, and so the probability that a given half-edge is still free at time t is asymptotically close to θ_t^2. For the numbers of half-edges of the remaining types, we define (with justification in the proof below), for $\theta \in [0,1]$,

$$h_{\mathrm{R}}(\theta) := \mu_{\mathrm{R}} \theta + \frac{\mu \rho}{\beta} \theta (1 - \theta), \tag{2.14}$$

$$h_{\mathrm{I}}(\theta) := h_X(\theta) - h_{\mathrm{S}}(\theta) - h_{\mathrm{R}}(\theta). \tag{2.15}$$

Thus $h_X(\theta) = h_{\mathrm{S}}(\theta) + h_{\mathrm{I}}(\theta) + h_{\mathrm{R}}(\theta)$. The corresponding limit functions for infective and recovered vertices are more easily described by differential equations, which will be introduced in (2.20) and (2.26). Note that

$$v_{\mathrm{S}}(1) = \alpha_{\mathrm{S}}, \tag{2.16}$$

$$h_{\mathrm{S}}(1) = \alpha_{\mathrm{S}} \lambda = \mu_{\mathrm{S}}, \qquad h_{\mathrm{R}}(1) = \mu_{\mathrm{R}}, \qquad h_{\mathrm{I}}(1) = \mu - \mu_{\mathrm{S}} - \mu_{\mathrm{R}} = \mu_{\mathrm{I}}. \tag{2.17}$$

We also introduce the 'infective pressure'

$$p_I(\theta) := \frac{h_I(\theta)}{h_X(\theta)}, \qquad (2.18)$$

which appears in the differential equations (2.19) and (2.25) below.

Our first two theorems concern the case where the initially infective population is macroscopic, so that the course of the epidemic is approximately deterministic for a long time, until shortly before extinction.

Theorem 2.6. *Let us consider the SIR epidemic on the multigraph G^* with degree sequence $(d_i)_1^n$. Suppose that (D1)–(D6) are satisfied. Let $\mu_I > 0$.*

(a) *There is a unique $\theta_\infty \in (0,1)$ with $h_I(\theta_\infty) = 0$. Further, h_I is strictly positive on $(\theta_\infty, 1]$ and strictly negative on $(0, \theta_\infty)$.*

(b) *There is a unique continuously differentiable function $\theta_t : [0, \infty) \to (\theta_\infty, 1]$ such that*

$$\frac{d}{dt}\theta_t = -\beta\theta_t p_I(\theta_t), \qquad \theta_0 = 1. \qquad (2.19)$$

Furthermore, $\theta_t \searrow \theta_\infty$ as $t \to \infty$.

(c) *Let $\hat{I}_t$ be the unique solution to*

$$\frac{d}{dt}\hat{I}_t = \frac{\beta h_I(\theta_t) h_S(\theta_t)}{h_X(\theta_t)} - \rho\hat{I}_t, \ t \geqslant 0, \qquad \hat{I}_0 = \alpha_I, \qquad (2.20)$$

and $\hat{R}_t := 1 - v_S(\theta_t) - \hat{I}_t$. Then, uniformly on $[0, \infty)$,

$$S_t/n \xrightarrow{\text{P}} v_S(\theta_t), \qquad I_t/n \xrightarrow{\text{P}} \hat{I}_t, \qquad R_t/n \xrightarrow{\text{P}} \hat{R}_t, \qquad (2.21)$$

$$X_{S,t}/n \xrightarrow{\text{P}} h_S(\theta_t), \qquad X_{I,t}/n \xrightarrow{\text{P}} h_I(\theta_t), \qquad X_{R,t}/n \xrightarrow{\text{P}} h_R(\theta_t), \qquad (2.22)$$

and, consequently, $X_t/n \xrightarrow{\text{P}} h_X(\theta_t)$.

(d) *Hence, the number $S_\infty := \lim_{t\to\infty} S_t$ of susceptibles that escape infection satisfies*

$$S_\infty/n \xrightarrow{\text{P}} v_S(\theta_\infty).$$

Theorem 2.7. *Let us consider the SIR epidemic on the uniform simple graph G with degree sequence $(d_i)_1^n$. Suppose that (D1)–(D6) and (G1) are satisfied. Let $\mu_I > 0$. Then the conclusions of Theorem 2.6 hold.*

Decreusefond et al. [16] obtain a related result, assuming that the fifth moment of the degree of a random vertex is uniformly bounded as $n \to \infty$.

Remark 2.8. We can give examples in which S_t/n, I_t/n and R_t/n fail to converge to deterministic limits when assumption (D5) does not hold. Generally, we believe (D5) is necessary for the convergence to deterministic limits in Theorems 2.6 and 2.7, but have not attempted to prove it.

Our third and fourth theorems concern the case where there are initially a small number of infectives. Let

$$\Re_0 := \left(\frac{\beta}{\rho + \beta}\right)\left(\frac{\alpha_S}{\mu}\right)\sum_{k=0}^{\infty}(k-1)kp_k; \qquad (2.23)$$

this quantity can be interpreted as the basic reproductive ratio of the epidemic. When $\Re_0 > 1$, then there is a positive probability that a large epidemic develops in the population,

as previously identified in the literature on epidemic models, such as [2; 34; 41; 8]. In that event, once established, the evolution of the epidemic is approximately deterministic, as in Theorems 2.6 and 2.7.

We will prove below that there is a unique $\theta_\infty \in (0,1)$ with $h_I(\theta_\infty) = 0$, and that if there is a large epidemic, the number of susceptibles that never get infected is approximately $nv_S(\theta_\infty)$. Fix a number $s_0 \in (v_S(\theta_\infty), \alpha_S)$, i.e., between the (approximate) fractions of susceptibles at the beginning and at the end of the epidemic in the case that a large epidemic develops, and let

$$T_0 := \inf\{t \geqslant 0 : S_t \leqslant ns_0\}. \tag{2.24}$$

(This means that $T_0 = \infty$ if S_t never falls below ns_0. We will see that this corresponds to the case of a small outbreak.) We shift the initial condition of the limiting differential equation, now defined on $(-\infty, \infty)$, so that $t = 0$ corresponds to the time T_0 in the random process, by which the fraction of susceptible individuals has fallen from about $\alpha_S = v_S(1)$ to some fixed smaller s_0. By time T_0, a positive fraction of the population has been infected, and from that point onwards the quantities of interest follow a law of large numbers. The exact choice of s_0 is unimportant.

We extend the processes to be defined on $(-\infty, \infty)$ by taking $S_t = S_0$ for $t < 0$, and similarly for the other processes.

Theorem 2.9. *Let us consider the SIR epidemic on the multigraph G^* with degree sequence $(d_i)_1^n$. Suppose that (D1)–(D6) and (G1) are satisfied. Suppose also that $\alpha_I = \mu_I = 0$ but there is initially at least one infective vertex with non-zero degree.*

(i) *If $\mathfrak{R}_0 \leqslant 1$ then the number $n_S - S_\infty$ of initially susceptible vertices that ever get infected is $o_p(n)$.*

(ii) *Suppose $\mathfrak{R}_0 > 1$.*

    (a) *There is a unique $\theta_\infty \in (0,1)$ with $h_I(\theta_\infty) = 0$. Further, h_I is strictly positive on $(\theta_\infty, 1)$ and strictly negative on $(0, \theta_\infty)$.*

    (b) *Let $s_0 \in (v_S(\theta_\infty), v_S(1))$. Then there is a unique continuously differentiable $\theta_t : \mathbb{R} \to (\theta_\infty, 1)$ such that*

$$\frac{d}{dt}\theta_t = -\beta\theta_t p_I(\theta_t), \qquad \theta_0 = v_S^{-1}(s_0). \tag{2.25}$$

    Furthermore, $\theta_t \searrow \theta_\infty$ as $t \to \infty$ and $\theta_t \nearrow 1$ as $t \to -\infty$.

    (c) *Let T_0 be defined by (2.24). Then $\liminf_{n\to\infty} \mathbb{P}(T_0 < \infty) > 0$. Furthermore, if the initial number of infective half-edges $X_{I,0} \to \infty$, then $\mathbb{P}(T_0 < \infty) \to 1$.*

    (d) *Let $\hat{I}_t$ be the unique solution to*

$$\frac{d}{dt}\hat{I}_t = \frac{\beta h_I(\theta_t) h_S(\theta_t)}{h_X(\theta_t)} - \rho\hat{I}_t, \qquad \lim_{t\to-\infty} \hat{I}_t = 0, \tag{2.26}$$

    and $\hat{R}_t := 1 - v_S(\theta_t) - \hat{I}_t$.
    Conditional on $T_0 < \infty$, then, uniformly on $(-\infty, \infty)$,

$$S_{T_0+t}/n \xrightarrow{\mathrm{p}} v_S(\theta_t), \qquad I_{T_0+t}/n \xrightarrow{\mathrm{p}} \hat{I}_t, \qquad R_{T_0+t}/n \xrightarrow{\mathrm{p}} \hat{R}_t, \tag{2.27}$$

$$X_{S,T_0+t}/n \xrightarrow{\mathrm{p}} h_S(\theta_t), \qquad X_{I,T_0+t}/n \xrightarrow{\mathrm{p}} h_I(\theta_t), \qquad X_{R,T_0+t}/n \xrightarrow{\mathrm{p}} h_R(\theta_t), \tag{2.28}$$

and, consequently, also $X_{T_0+t}/n \xrightarrow{\mathrm{p}} h_X(\theta_t)$.

8

(e) *Conditional on $T_0 < \infty$, the number of susceptibles that escape infection satisfies*

$$S_\infty / n \xrightarrow{\ \mathrm{p}\ } v_{\mathrm{S}}(\theta_\infty).$$

(f) *The number of susceptibles that ever get infected $S_0 - S_\infty$ satisfies $S_0 - S_\infty = o_{\mathrm{p}}(n)$ on the event $T_0 = \infty$, in the sense that, for all $\varepsilon > 0$, $\mathbb{P}(T_0 = \infty, S_0 - S_\infty > \varepsilon n) = o(1)$ as $n \to \infty$.*
Similarly, $X_{S,0} - X_{S,\infty} = o_{\mathrm{p}}(n)$, $\sup_{t \geqslant 0} X_{\mathrm{I},t} = o_{\mathrm{p}}(n)$, $\sup_{t \geqslant 0}(X_0 - X_t) = o_{\mathrm{p}}(n)$ on $T_0 = \infty$.

The same result holds even without assumption (G1), except that, in this case, it is possible to have $\theta_t : \mathbb{R} \to (\theta_\infty, 1]$ with $\theta_t = 1$ for $t \leqslant \hat{A}_0$, for some $\hat{A}_0 < 0$.

Theorem 2.10. *Let us consider the SIR epidemic on the uniform simple graph G with degree sequence $(d_i)_1^n$. Suppose that (D1)–(D6) and (G1) are satisfied. Let $\mu_{\mathrm{I}} = 0$. Then the conclusions of Theorem 2.10 hold.*

It is possible that $\mathbb{P}(T_0 = \infty) \to 0$, and then the statements in Theorem 2.9(ii)(f) are trivial. In order to prove that quantities of interest are small conditional on $T_0 = \infty$ in this case, we would need to study the speeed at which $\mathbb{P}(T_0 = \infty) \to 0$. Nevertheless, the theorem shows a dichotomy when $\mathfrak{R}_0 > 1$: w.h.p. either $T_0 = \infty$ and the outbreak is small, with only a few individuals infected; or $T_0 < \infty$ and the outbreak is large, with $(\alpha_S - v_{\mathrm{S}}(\theta_\infty))n + o(n)$ individuals infected (a more detailed description of the evolution is given in (ii)(d)).

In fact, we will take $T_0 < \infty$ as the definition of a large outbreak. (Formally this depends on the choice of s_0, but the theorem shows that any two choices w.h.p. yield the same result.) Thus the probability $\mathbb{P}(T_0 < \infty)$ in (ii)(c) is, by definition, the probability of a large outbreak. We give a formula for this probability in Theorem 5.3, using a branching process approximation to the early stage of the epidemic (or an equivalent approximation using a random walk), see further Sections 4.3.1 and 5. The condition $\mathfrak{R}_0 > 1$ can be interpreted as supercriticality of this branching process approximation.

Furthermore, it turns out that a small outbreak is really small. In Theorem 5.4, we sharpen Theorem (ii)(f) by showing that, for a small outbreak, only $O_{\mathrm{p}}(1)$ susceptibles are infected, both in the (sub)critical case ($\mathfrak{R}_0 \leqslant 1$) provided $X_{\mathrm{I},0} = O(1)$, and in the supercritical case ($\mathfrak{R}_0 > 1$).

For bounded degree sequences and $\mathfrak{R}_0 \neq 1$, a result similar to Theorems 2.9 and 2.10 is proven in [8] by Bohman and Picollelli. The threshold in $\mathfrak{R}_0$ for a possible large outbreak and the final size of a large outbreak are derived heuristically in [2; 34; 41]. All the above papers assume that initially there are no recovered vertices. Our motivation for allowing the presence of initially recovered individuals is to be able to analyse simple vaccination strategies, see Section 2.2. Before doing that, we give some connections to related results.

Theorems 2.6 and 2.7 imply $X_{\mathrm{I},t}/X_t \xrightarrow{\ \mathrm{p}\ } \mathrm{p}_{\mathrm{I}}(\theta_t)$ and $X_{\mathrm{S},t}/X_t \xrightarrow{\ \mathrm{p}\ } \mathrm{p}_{\mathrm{S}}(\theta_t)$ uniformly when $\mu_{\mathrm{I}} > 0$, where $\mathrm{p}_{\mathrm{S}}(\theta) := h_{\mathrm{S}}(\theta)/h_X(\theta)$ is defined analogously to p_{I} in (2.18). Theorems 2.9 and 2.10 yield the same result for the time shifted process when $\mu_{\mathrm{I}} = 0$ and $\mathfrak{R}_0 > 1$, conditional on a large outbreak. To explain the connection with [41], let

$$g(\theta) := \sum_{k=0}^{\infty} p_k \theta^k, \qquad \theta \in [0, 1], \tag{2.29}$$

the probability generating function for the asymptotic degree distribution of initially susceptible vertices. Note that $v_S(\theta) = \alpha_S g(\theta)$ and $h_S(\theta) = \alpha_S \theta g'(\theta)$. Differentiating $p_I(\theta_t)$ and $p_S(\theta_t)$ yields, using (2.19) and (2.12)–(2.15),

$$\frac{dp_I(\theta_t)}{dt} = p_I(\theta_t)\left(-(\rho + \beta) + \beta p_I(\theta_t) + \beta p_S(\theta_t)\theta_t\frac{g''(\theta_t)}{g'(\theta_t)}\right), \tag{2.30}$$

$$\frac{dp_S(\theta_t)}{dt} = \beta p_I(\theta_t)p_S(\theta_t)\left(1 - \theta_t\frac{g''(\theta_t)}{g'(\theta_t)}\right). \tag{2.31}$$

These are the 'Volz equations' [41, Table 3] mentioned in the introduction. Volz [41] derived them heuristically, assuming that the number of edges from a newly infective vertex to susceptible, infective and recovered vertices has multinomial distribution with parameters p_I, p_S and $1 - p_S - p_I$.

Theorem 2.10 relates to the existence of a giant component in G as follows. The epidemic spreads only within connected components of G. Further, if there are no recoveries, then all vertices connected to an initially infective vertex eventually get infected. Indeed, when $\rho = \mu_I = \mu_R = 0$, the threshold $\mathfrak{R}_0 > 1$ is equivalent to $\sum_{k=0}^{\infty} k(k-2)p_k > 0$; this is the well known condition of Molloy and Reed [31] for existence of a giant component. Also, in part (ii)(a), the equation defining $\theta_\infty \in (0, 1)$ becomes $\lambda\theta_\infty^2 - \sum_{k=0}^{\infty} kp_k\theta_\infty^k = 0$ in this case. With this value of θ_∞, and assuming $\alpha_I = \alpha_R = 0$ so $\alpha_S = 1$, it is known that $1 - v_S(\theta_\infty) = 1 - \sum_{k=0}^{\infty} p_k\theta_\infty^k$ is the fraction of vertices in the giant component [32] (see also [26]).

The connection to the giant component explains why (D6) is needed, at least when $\mathfrak{R}_0 = 1$. Suppose that (D6) is not satisfied, i.e. both $\rho = \mu_R = 0$ and $p_1 = 0$. If also $\mu_I = 0$ then $\mathfrak{R}_0 = 1$ is equivalent to $\sum_{k=0}^{\infty} k(k-2)p_k = 0$, and so only p_0 and p_2 can be non-zero. At least three different types of behaviour of component sizes in G are possible in this case. We will demonstrate them with the following examples from [26, Remark 2.7], see also Remark 4.1. We assume that $n_R = 0$ in each example.

The first example is a random 2-regular graph, that is $n_2 = n$ for all n. In this case, all the components are cycles. Let $V_1 \geqslant V_2 \geqslant$ denote the ordered component sizes. Then V_1/n converges weakly to a non-degenerate distribution on $[0, 1]$, and the same holds for V_2/n, V_3/n, and so on [3, Lemma 5.7]. Let us suppose that there is initially a lone infective vertex. The number of vertices in the component it occupies (and hence eventually infects) is a size biased sample from $(V_1, V_2, \ldots)$, and, divided by n, also converges to a non-degenerate distribution on $[0, 1]$.

For the second example, we suppose that $n = n_1 + n_2$, where $n_2/n \to 1$, n_1 is even and $n_1 \to \infty$. The desired graph can be obtained from a random 2-regular graph with $n - n_1/2$ vertices by selecting $n_1/2$ degree 2 vertices at random, one after another, and creating two degree 1 vertices out of each one. During this procedure, components are chosen in a size-biased fashion and split uniformly (except on the first attempt, since they were all cycles to begin with). The largest component in the resulting graph contains only $o_p(n)$ vertices, and so only $o_p(n)$ susceptibles are infected if n_I is bounded.

For our third example, we take $n = n_2 + n_4$, where $n_2/n \to 1$ and $n_4 \to \infty$. Each vertex of degree 4 can be obtained by merging a pair of vertices of degree 2. Analogously to the previous example, we see there is a unique giant component with $n - o_p(n)$ vertices. Hence,

even if only a single given vertex is initially infective, then all $n - o_p(n)$ susceptibles in the giant component succumb to infection w.h.p.

2.2. Vaccination.

Let us suppose that we vaccinate some susceptible vertices before the epidemic process starts. The vaccine is assumed perfect, so that a vaccinated vertex never becomes infective. In particular, vaccinated vertices behave like recovered vertices in the SIR dynamics. Let us use this fact to analyse degree dependent vaccination strategies by applying Theorems 2.9 and 2.10 to a suitably modified degree sequence.

We assume that each initially susceptible vertex of degree $k \geqslant 0$ is vaccinated with probability $\pi_k \in [0, 1)$, independently of all the others. Here are two examples of such strategies.

Uniform vaccination. We vaccinate every susceptible vertex with the same probability $\pi_k = v$ for all k and some $v \in [0, 1)$, independently of all the others. The total number V of vaccinations thus satisfies $V/n_S \xrightarrow{\text{p}} v$, using the law of large numbers.

Edgewise vaccination. We vaccinate the end point of each susceptible half edge with probability $v \in [0, 1)$, independently of all the other half-edges. Thus the probability that a degree k susceptible is vaccinated is $\pi_k := 1 - (1 - v)^k$, and, under our assumptions, the total number V of vaccinations satisfies $V/n_S \xrightarrow{\text{p}} \sum_{k=0}^{\infty} p_k \pi_k$.

These strategies are considered in [12], where their efficacy in a related epidemic model (equivalent to constant recovery times) is compared, along with two other strategies (uniform acquaintance vaccination strategy and another edgewise strategy, neither of which can be studied with the present argument).

As noted above, vaccinating a vertex amounts to changing its type from susceptible to recovered. Let us calculate the (random) number of vertices of each type post-vaccination, and show that assumptions (D1)–(D6) hold for the modified degree sequence.

We add a tilde to our notation for the post-vaccinated epidemic. Thus $\tilde{n}_{S,k}$ denotes the number of degree $k \geqslant 0$ initially susceptible vertices that remain unvaccinated. We have $\tilde{n}_{S,k} \sim \text{Binomial}(n_{S,k}, 1 - \pi_k)$, and, by the law of large numbers,

$$\tilde{n}_{S,k} = n_{S,k}(1 - \pi_k) + o_p(n) = n\alpha_S p_k(1 - \pi_k) + o_p(n).$$

Using the uniform summability of $\sum_{k=0}^{\infty} k n_{S,k}/n$ (see Remark 2.2)

$$\tilde{n}_S := \sum_{k=0}^{\infty} \tilde{n}_{S,k} = n\alpha_S \sum_{k=0}^{\infty} p_k(1 - \pi_k) + o_p(n),$$

whence

$$\frac{\tilde{n}_S}{n} \xrightarrow{\text{p}} \alpha_S \sum_{k=0}^{\infty} p_k(1 - \pi_k) =: \tilde{\alpha}_S > 0.$$

Furthermore,

$$\frac{\tilde{n}_{S,k}}{\tilde{n}_S} \xrightarrow{\text{p}} \frac{p_k(1 - \pi_k)}{\sum_{k=0}^{\infty} p_k(1 - \pi_k)} =: \tilde{p}_k,$$

and $\tilde{p}_1 > 0$ if $p_1 > 0$. The mean of $(\tilde{p}_k)_0^{\infty}$ is

$$\tilde{\lambda} := \sum_{k=0}^{\infty} k\tilde{p}_k \leqslant \alpha_S \lambda / \tilde{\alpha}_S < \infty.$$

11

By the uniform summability of $\sum_{k=0}^{\infty} k n_{S,k}/n$, and the inequality $\tilde{n}_{S,k} \leqslant n_{S,k}$, we also have

$$\frac{1}{\tilde{n}_S} \sum_{k=0}^{\infty} k \tilde{n}_{S,k} \xrightarrow{\mathrm{p}} \sum_{k=0}^{\infty} k \tilde{p}_k.$$

The initial number $\tilde{n}_{R,k} = n_{R,k} + (n_{S,k} - \tilde{n}_{S,k})$ of degree k vertices that are either recovered or have been vaccinated satisfies

$$\tilde{n}_{R,k} = n_{R,k} + n \alpha_S \pi_k p_k + o_p(n),$$

and so, again using the uniform summability of $\sum_{k=0}^{\infty} k n_{S,k}/n$,

$$\sum_{k=0}^{\infty} k \tilde{n}_{R,k}/n \to \tilde{\mu}_R := \mu_R + \alpha_S \sum_{k=0}^{\infty} k \pi_k p_k.$$

The number $\tilde{n}_{I,k}$ of infective vertices of degree k after the vaccinations is unchanged, i.e. $\tilde{n}_{I,k} = n_{I,k}$. It follows that

$$\tilde{\alpha}_I = \lim_{n \to \infty} \sum_{k=0}^{\infty} \tilde{n}_{I,k}/n = \alpha_I, \qquad \tilde{\mu}_I = \lim_{n \to \infty} \sum_{k=0}^{\infty} k \tilde{n}_{I,k}/n = \mu_I.$$

Similarly, $\tilde{n}_k = n_k$, and so $\tilde{\mu} = \mu$. Furthermore,

$$\sum_{k=0}^{\infty} k \tilde{n}_{S,k}/n \to \tilde{\mu}_S = \tilde{\alpha}_S \tilde{\lambda} = \alpha_S \sum_{k=0}^{\infty} k p_k (1 - \pi_k) = \mu_S - \alpha_S \sum_{k=0}^{\infty} k \pi_k p_k.$$

These limits may be assumed almost sure by the Skorokhod coupling lemma. Hence, we can apply Theorems 2.9 and 2.10 to the post-vaccination epidemic with the modified values $\tilde{\alpha}_S, \tilde{\alpha}_I, \tilde{\alpha}_R, \tilde{\lambda}, \tilde{\mu}_I, \tilde{\mu}_R$ and $(\tilde{p}_k)_0^{\infty}$, with the understanding that R_t and $X_{R,t}$ now include vaccinated vertices and half-edges (though these could easily be subtracted off). The limiting deterministic evolution and final size follow immediately. Rather than restating the results in full, we simply give criteria for the vaccination programme to be successful.

Corollary 2.11. *Let us consider the SIR epidemic on the uniform simple graph G. Suppose that (D1)–(D6) and (G1) hold. Suppose further that each initially susceptible vertex of degree k is vaccinated with probability $\pi_k \in [0, 1)$ independently of the others, and $\mu_I = 0$. Let*

$$\widetilde{\mathfrak{R}}_0 := \left(\frac{\beta}{\rho + \beta} \right) \left(\frac{\alpha_S}{\mu} \right) \sum_{k=0}^{\infty} k(k-1) p_k (1 - \pi_k). \tag{2.32}$$

If $\widetilde{\mathfrak{R}}_0 \leqslant 1$, then the total number of susceptible vertices that get infected is $o_p(n)$. If $\widetilde{\mathfrak{R}}_0 > 1$, then there exists $\delta > 0$ such that at least δn susceptibles get infected with probability bounded away zero.

The same result holds for the SIR epidemic on the multigraph G^*, even without assumption (G1).

3. The time-changed epidemic

In this and the next two sections, we consider the SIR epidemic on the random multigraph G^*; in Section 6 we transfer the results to the simple random graph G.

A key step in the proof is to alter the speed of the process by multiplying each transition rate by a constant depending on the current state. The constant is chosen so that each free susceptible half-edge gets infected at unit rate (or, equivalently, so that the infection pressure on the population is 1). Specifically, if there are $x_I \geqslant 1$ free infective half-edges, and a total of x free half-edges of any type, we multiply all transition rates out of such a state by $(x-1)/\beta x_I > 0$. Thus each infective vertex recovers at rate $\rho(x-1)/\beta x_I$, and each free infective half-edge pairs off at rate $(x-1)/x_I$. This change of rates accelerates the epidemic in its 'slow' phases, when the number of free infective half-edges is $o(n)$ (beginning and end of the epidemic). Later, we will invert the time change to obtain the original process.

We use Greek letters (τ and σ) for the time variable of the altered process as a reminder of the rate modification. The notation for the numbers of half-edges and vertices of each type in the modified process follows that for the original process, except that we superscript each variable with a prime. For example, $X'_{I,\tau}$ denotes the number of infective half-edges at time $\tau \geqslant 0$ in the modified process.

Let

$$\tau^* := \inf\{\tau \geqslant 0 : X'_{I,\tau} = 0\} \tag{3.1}$$

be the time at which the modified process stops, when there are no free infective half-edges.

Lemma 3.1. *Suppose that (D1)–(D5) hold. Fix $\tau_1 > 0$. Then, uniformly over $[0, \tau_1 \wedge \tau^*]$,*

$$S'_\tau/n \xrightarrow{\mathrm{P}} v_S(e^{-\tau}), \tag{3.2}$$

$$X'_{S,\tau}/n \xrightarrow{\mathrm{P}} h_S(e^{-\tau}), \tag{3.3}$$

$$X'_\tau/n \xrightarrow{\mathrm{P}} h_X(e^{-\tau}), \tag{3.4}$$

$$X'_{R,\tau}/n \xrightarrow{\mathrm{P}} h_R(e^{-\tau}), \tag{3.5}$$

and, consequently,

$$X'_{I,\tau}/n \xrightarrow{\mathrm{P}} h_I(e^{-\tau}). \tag{3.6}$$

Proof. For each $k \in \mathbb{Z}^+$, let $S'_\tau(k)$ denote the number of susceptible vertices with k half-edges at time $\tau \geqslant 0$ (we omit the qualifier $\tau \leqslant \tau^*$ throughout the proof; any occurrence of 'τ' is understood to mean '$\tau \wedge \tau^*$'). Thus $S'_\tau = \sum_{k=0}^\infty S'_\tau(k)$ and $X'_{S,\tau} = \sum_{k=0}^\infty k S'_\tau(k)$ for each τ. Also, $S'_0(k) = n_{S,k}$, for each k.

For each k, the only jumps in $S'_\tau(k)$ are decrements by 1, and these occur when an infective half-edge pairs off and chooses one of the half-edges belonging to a susceptible vertex of degree k. Hence, with the modified transition rates,

$$dS'_\tau(k) = -\beta X'_{I,\tau} \left(\frac{X'_\tau - 1}{\beta X'_{I,\tau}}\right) \left(\frac{k S'_\tau(k)}{X'_\tau - 1}\right) d\tau + dM_{S,\tau}(k)$$

$$= -k S'_\tau(k) \, d\tau + dM_{S,\tau}(k), \tag{3.7}$$

where $(M_{S,\tau}(k))_{\tau \geqslant 0}$ is a martingale starting from $M_{S,0}(k) = 0$ [17, Proposition 1.7].

13

The differential notation in (3.7) means that

$$S'_\tau(k) = S'_0(k) - k \int_0^\tau S'_\sigma(k)\, d\sigma + M_{S,\tau}(k). \tag{3.8}$$

Since $S'_0(k) = n_{S,k}$, it follows that

$$
\begin{aligned}
\left| S'_\tau(k) - n_{S,k} e^{-k\tau} \right| &= \left| S'_\tau(k) - n_{S,k} \left(1 - k \int_0^\tau e^{-k\sigma}\, d\sigma \right) \right| \\
&= \left| \int_0^\tau k \left(-S'_\sigma(k) + n_{S,k} e^{-k\sigma} \right) d\sigma + M_{S,\tau}(k) \right| \\
&\leqslant k \int_0^\tau \left| S'_\sigma(k) - n_{S,k} e^{-k\sigma} \right| d\sigma + |M_{S,\tau}(k)|.
\end{aligned} \tag{3.9}
$$

Consequently, using Gronwall's inequality (for positive bounded functions) [38, Appendix §1],

$$
\begin{aligned}
\sup_{\tau \leqslant \tau_1} |S'_\tau(k) - n_{S,k} e^{-k\tau}| &\leqslant k \int_0^{\tau_1} \sup_{\tau \leqslant \sigma} |S'_\tau(k) - n_{S,k} e^{-k\tau}|\, d\sigma + \sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)| \\
&\leqslant e^{k\tau_1} \sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)|,
\end{aligned}
$$

and it follows that

$$\sup_{\tau \leqslant \tau_1} \left| S'_\tau(k)/n - \alpha_S p_k e^{-k\tau} \right| \leqslant |n_{S,k}/n - \alpha_S p_k| + e^{k\tau_1} \sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)|/n. \tag{3.10}$$

The first term on the right goes to zero as $n \to \infty$ by (D1) and (D2). Let us show that $\sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)|/n \xrightarrow{\text{p}} 0$.

The martingale $M_{S,\tau}(k)$ is right continuous and has left limits (càdlàg), and it is also of finite variation. The quadratic variation process of such a martingale is the running sum of its (countably many) squared jumps [28, Theorem 26.6]. The jumps in $M_{S,\tau}(k)$ are by (3.8) the same as the jumps in $S'_\tau(k)$. Each jump in $S'_\tau(k)$ is a decrement by one, and there are at most $S'_0(k)$ such jumps. Hence the quadratic variation $[M_S(k)]_\tau$ of $M_{S,\tau}(k)$ satisfies

$$[M_S(k)]_\tau = \sum_{0 \leqslant \sigma \leqslant \tau} (\Delta M_{S,\sigma}(k))^2 \leqslant S'_0(k) = n_{S,k} \leqslant n,$$

for any $\tau \geqslant 0$. In particular, $\mathbb{E}[M_S(k)]_\tau < \infty$ for every $\tau \geqslant 0$, so $M_{S,\tau}(k)$ is square integrable and $\mathbb{E}\, M_{S,\tau}(k)^2 = \mathbb{E}[M_S(k)]_\tau$ [37, Corollary 3 after Theorem II.6.27, p. 73]. Hence, Doob's L^2-inequality [28, Proposition 7.16] yields

$$\mathbb{E} \sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)|^2 \leqslant 4\, \mathbb{E}\, M_{S,\tau_1}(k)^2 = 4\, \mathbb{E}[M_S(k)]_{\tau_1} = O(n). \tag{3.11}$$

It follows that $\sup_{\tau \leqslant \tau_1} |M_{S,\tau}(k)| = o_p(n)$, and so, by (3.10), for each k,

$$\sup_{\tau \leqslant \tau_1} \left| S'_\tau(k)/n - \alpha_S p_k e^{-k\tau} \right| \xrightarrow{\text{p}} 0. \tag{3.12}$$

Let $\varepsilon > 0$. By assumptions (D1)–(D3), there exists $K > 0$ such that $\sum_{k=K+1}^\infty k n_{S,k}/n < \epsilon$ for any n, see Remark 2.2. Further, K can be chosen large enough that $\sum_{k=K+1}^\infty k p_k < \varepsilon$.

14

Consequently,

$$\sup_{\tau \leqslant \tau_1} \left| X'_{S,\tau}/n - h_S(e^{-\tau}) \right| = \sup_{\tau \leqslant \tau_1} \left| \sum_{k=0}^{\infty} k S'_\tau(k)/n - \alpha_S \sum_{k=0}^{\infty} k p_k e^{-\tau k} \right|$$

$$\leqslant \sum_{k=0}^{K} k \sup_{\tau \leqslant \tau_1} \left| S'_\tau(k)/n - \alpha_S p_k e^{-k\tau} \right| + \sum_{k=K+1}^{\infty} k \left(n_{S,k}/n + p_k \right)$$

$$\leqslant \sum_{k=0}^{K} k \sup_{\tau \leqslant \tau_1} \left| S'_\tau(k)/n - \alpha_S p_k e^{-k\tau} \right| + 2\epsilon, \tag{3.13}$$

and the same bound applies to $\sup_{\tau \leqslant \tau_1} |S'_\tau/n - v_S(e^{-\tau})|$. The finite sum in the last line of (3.13) tends to zero in probability, by (3.12). This completes our proof of (3.2) and (3.3).

We prove (3.4) and (3.5) similarly. The total number of free half-edges decrements by 2 whenever an infective half-edge pairs off. Then

$$dX'_\tau = -2\beta X'_{I,\tau} \left(\frac{X'_\tau - 1}{\beta X'_{I,\tau}} \right) d\tau + dM_{X,\tau} = -2(X'_\tau - 1) \, d\tau + dM_{X,\tau}, \tag{3.14}$$

where $(M_{X,\tau})_{\tau \geqslant 0}$ is a càdlàg martingale satisfying $M_{X,0} = 0$. Writing the equation in integrated form and proceeding as in (3.9),

$$\left| X'_\tau - X'_0 e^{-2\tau} \right| \leqslant 2 \int_0^\tau \left| X'_\sigma - X'_0 e^{-2\sigma} \right| d\sigma + 2\tau + |M_{X,\tau}|.$$

Gronwall's inequality yields

$$\sup_{\tau \leqslant \tau_1} \left| X'_\tau - X'_0 e^{-2\tau} \right| \leqslant e^{2\tau_1} \left(2\tau_1 + \sup_{\tau \leqslant \tau_1} |M_{X,\tau}| \right),$$

and so,

$$\sup_{\tau \leqslant \tau_1} \left| X'_\tau/n - h_X(e^{-\tau}) \right| \leqslant \left| X'_0/n - \mu \right| + e^{2\tau_1} \left(2\tau_1 + \sup_{\tau \leqslant \tau_1} |M_{X,\tau}| \right)/n.$$

By (D4) $X'_0/n = \sum_{k=0}^{\infty} k n_k/n \to \mu$, and so the first term above converges to 0. To estimate the martingale term, note that the jumps in $M_{X,\tau}$ are the same as the jumps in X'_τ. At each jump, X'_τ decreases by 2, and there is at most one jump in X'_τ for each of the X'_0 initial half-edges. Hence, the quadratic variation $[M_X]_{\tau_1}$ satisfies

$$[M_X]_{\tau_1} \leqslant \sum_{\sigma \geqslant 0} (\Delta M_{X,\sigma})^2 \leqslant 4X'_0 = O(n).$$

Proceeding as in (3.11), we see that $\sup_{\tau \leqslant \tau_1} |M_{X,\tau}| = o_p(n)$, and (3.4) follows.

Now, the number of free recovered half-edges $X'_{R,\tau}$ decreases by 1 when an infective half-edge is paired with a recovered half-edge, and increases by $k \geqslant 0$ when an infective vertex with k free half-edges recovers. With the modified rates, we have

$$dX'_{R,\tau} = \left(-\beta X'_{I,\tau} \left(\frac{X'_{R,\tau}}{X'_\tau - 1} \right) + \rho X'_{I,\tau} \right) \left(\frac{X'_\tau - 1}{\beta X'_{I,\tau}} \right) d\tau + dM_{R,\tau}$$

$$= -X'_{R,\tau} \, d\tau + \rho \beta^{-1}(X'_\tau - 1) \, d\tau + dM_{R,\tau}, \tag{3.15}$$

where $(M_{R,\tau})_{\tau \geqslant 0}$ is a càdlàg martingale with $M_{R,0} = 0$.

Differentiating the expression for h_{R} in (2.14) shows that

$$\frac{d}{d\sigma}h_{\mathrm{R}}(e^{-\sigma}) = -h_{\mathrm{R}}(e^{-\sigma}) + \rho\beta^{-1}h_X(e^{-\sigma}).$$

Since $h_{\mathrm{R}}(1) = \mu_{\mathrm{R}}$, integrating, subtracting the integrated form of (3.15) divided by n, and taking the absolute value yields

$$\left|X'_{\mathrm{R},\tau}/n - h_{\mathrm{R}}(e^{-\tau})\right| \leqslant \int_0^\tau \left|X'_{\mathrm{R},\sigma}/n - h_{\mathrm{R}}(e^{-\sigma})\right| d\sigma + E_\tau, \tag{3.16}$$

where the absolute error term E_τ is given by

$$E_\tau := \left|X'_{\mathrm{R},0}/n - \mu_{\mathrm{R}}\right| + \frac{\rho}{\beta}\int_0^\tau \left|X'_\sigma/n - h_X(e^{-\sigma})\right| d\sigma + \frac{\rho\tau}{n\beta} + \frac{|M_{\mathrm{R},\tau}|}{n}.$$

Let us show that $\sup_{\tau\leqslant\tau_1}|E_\tau| \xrightarrow{\mathrm{P}} 0$; then (3.5) follows from (3.16) by applying Gronwall's inequality.

First of all, $X'_{\mathrm{R},0}/n = \sum_{k=0}^\infty k n_{\mathrm{R},k}/n \to \mu_{\mathrm{R}}$ by (D4). The integrand in the second term tends to zero uniformly by the convergence (3.4) of X'_σ already proven. Finally, the jumps in $M_{\mathrm{R},\tau}$ are due to either an infective vertex recovering (which happens at most once for each vertex, and only for vertices that were initially infective or initially susceptible) or an infective half-edge pairing to a recovered half-edge (which happens at most once for each half-edge). It follows that

$$[M_{\mathrm{R}}]_{\tau_1} \leqslant \sum_{\sigma\geqslant 0}(\Delta M_{\mathrm{R},\sigma})^2 \leqslant X_0 + \sum_{k=0}^\infty k^2(n_{\mathrm{S},k} + n_{\mathrm{I},k}) = o(n^2),$$

by (D5) and (2.9), and so $\sup_{\tau\leqslant\tau_1}|M_{\mathrm{R},\tau}| = o_{\mathrm{p}}(n)$.

Finally, the convergence (3.6) of $X'_{\mathrm{I},\tau}/n = X'_\tau/n - X'_{\mathrm{S},\tau}/n - X'_{\mathrm{R},\tau}/n$ follows by applying the triangle inequality. $\qquad\square$

Remark 3.2. Any given susceptible half-edge gets infected at unit rate in the modified process, and so each initial degree k susceptible gets infected after an $\mathrm{Exp}(k)$ time, independently of all the other susceptibles. This observation can be used to give an alternative proof of (3.12) using the Glivenko–Cantelli lemma for convergence of empirical distributions, as in [26]. Using the Glivenko–Cantelli lemma would allow us to take $\tau_1 = \infty$ in (3.2)–(3.4). However, this strengthening would give no extra benefit, since τ^* is bounded w.h.p., as shown in Section 4 below.

Further, it is possible to obtain quantitative statements of convergence for $\sup_{\tau\leqslant\tau_1}|M_{\mathrm{S},\tau}|/n$, $\sup_{\tau\leqslant\tau_1}|M_{X,\tau}|/n$ and $\sup_{\tau\leqslant\tau_1}|M_{\mathrm{R},\tau}|/n$ using techniques such as those in [15].

3.1. **Inverting the rate change to recover the original process.** To close this section, we explain how to rescale time in order to obtain the original process. To this end, we define

$$A_\tau = \int_0^\tau \frac{1}{\beta}\left(\frac{X'_\sigma - 1}{X'_{\mathrm{I},\sigma}}\right) d\sigma, \qquad \tau \geqslant 0, \tag{3.17}$$

where we regard the bracketed term in the integrand as being equal to $1/2$ if $X'_{\mathrm{I},\sigma} = 0$, i.e. if $\sigma \geqslant \tau^*$ (till then, the bracketed term is at least $1/2$, since $X'_\sigma \geqslant X'_{\mathrm{I},\sigma}$, and, if $X'_{\mathrm{I},\sigma} = 1$, then $X'_\sigma \geqslant 2$). Thus A is strictly increasing and continuous. We let $\tau(t) : [0,\infty) \to [0,\infty)$ be its strictly increasing continuous inverse, so $A_{\tau(t)} = t$ for any $t \geqslant 0$.

16

The processes in their original time scaling can then be realised as

$$S_t = S'_{\tau(t)}, \qquad I_t = I'_{\tau(t)}, \qquad R_t = R'_{\tau(t)}, \qquad \ldots, \qquad t \geqslant 0, \qquad (3.18)$$

by applying Lemma A.1 to the underlying Markov evolution of the epidemic and graph dynamics. Since the epidemic stops when we run out of free infective half-edges, it makes no difference to (3.18) if we replace $\tau(t)$ by $\bar{\tau}(t) := \tau(t) \wedge \tau^*$ above; this will be convenient for the proofs below.

4. Proofs for multigraphs: Theorems 2.6 and 2.9

We continue to study the SIR epidemic on the random multigraph G^*. We assume (D1)–(D6), unless otherwise stated. For simplicity, we also assume (G1), and leave the minor modifications in the general case to Section 7. (We are mainly interested in the simple random graph G, where we have to assume (G1).)

We begin with the (sub)critical regime (part (i) of Theorem 2.9).

4.1. **Subcriticality: proof of Theorem 2.9 part (i).** Suppose that the hypotheses $\mathfrak{R}_0 \leqslant 1$ and $\mu_{\mathrm{I}} = 0$ are satisfied. We must show only $o_{\mathrm{p}}(n)$ susceptibles ever get infected, and it is sufficient to prove this for the modified epidemic studied in Section 3. The key step is proving that the modified epidemic dies almost instantly, i.e. $\tau^* \xrightarrow{\mathrm{P}} 0$.

For this purpose, we first show that $h_{\mathrm{I}}(\theta) < 0$ for $\theta \in (0, 1)$. It is enough to consider

$$h(\theta) := h_{\mathrm{I}}(\theta)/\theta = \mu\theta - \alpha_{\mathrm{S}} \sum_{k=0}^{\infty} k\theta^{k-1} p_k - \mu_{\mathrm{R}} - \rho\mu(1-\theta)/\beta. \qquad (4.1)$$

By assumption, $h(1) = h_{\mathrm{I}}(1) = \mu_{\mathrm{I}} = 0$, see (2.17). Furthermore,

$$h(0) = -\alpha_{\mathrm{S}} p_1 - \mu_{\mathrm{R}} - \rho\mu/\beta < 0, \qquad (4.2)$$

by (D6).

Differentiating h and substituting the identity $\beta\alpha_{\mathrm{S}} \sum_{k=0}^{\infty} k(k-1)p_k = \mu(\rho + \beta)\mathfrak{R}_0$ yields

$$\beta h'(\theta) = (\rho + \beta)\mu - \beta\alpha_{\mathrm{S}} \sum_{k=0}^{\infty} k(k-1)\theta^{k-2} p_k$$
$$\geqslant (\rho + \beta)\mu(1 - \mathfrak{R}_0)$$
$$\geqslant 0, \qquad (4.3)$$

for $\theta \in (0, 1)$. If $\sum_{k=3}^{\infty} p_k > 0$ and $\theta < 1$, then there is strict inequality going from the first to second line in (4.3), and thus $h'(\theta) > 0$. If $\sum_{k=3}^{\infty} p_k = 0$, then h is linear and $h(0) < 0$ by (4.2). In any case, recalling that $h(1) = 0$, we obtain $h(\theta) < 0$ for $\theta \in (0, 1)$.

We now take $\varepsilon > 0$ and apply Lemma 3.1 with $\tau_1 = \varepsilon$. It follows from (3.6) that

$$\sup_{\tau \leqslant \tau^* \wedge \varepsilon} \left| X'_{\mathrm{I},\tau}/n - h_{\mathrm{I}}(e^{-\tau}) \right| < |h_{\mathrm{I}}(e^{-\varepsilon})|/2 \qquad (4.4)$$

w.h.p. On the event that inequality (4.4) holds, we have $\tau^* < \varepsilon$; otherwise the left hand side of (4.4) is at least $|h_{\mathrm{I}}(e^{-\varepsilon})|$, since $X'_{\mathrm{I},\varepsilon} \geqslant 0$ and $h_{\mathrm{I}}(e^{-\varepsilon}) < 0$. Hence w.h.p. $\tau^* < \varepsilon$.

It follows that $\tau^* \xrightarrow{\mathrm{P}} 0$, and, by (3.2), the number $n_{\mathrm{S}} - S'_{\tau^*}$ of susceptibles that ever get infected satisfies

$$(n_{\mathrm{S}} - S'_{\tau^*})/n \xrightarrow{\mathrm{P}} \alpha_{\mathrm{S}} - v_{\mathrm{S}}(e^{-0}) = 0. \qquad (4.5)$$

17

Remark 4.1. If (D6) does not hold, then $h(0) = 0 = h(1)$. By (4.3), then $h(\theta) = 0$ for every θ. This happens only in the case $p_0 + p_2 = 1$. (In this case (2.23) yields $\mathfrak{R}_0 = 1$.) Lemma 3.1 is still valid, but it is no longer true in general that $\tau^* \xrightarrow{\text{P}} 0$. For example, if all vertices have degree 2 and there is initially a single infective vertex, then there will be two free infective half-edges until one pairs off with the other, and it is easy to see that τ^* has an Exp(1) distribution. Furthermore, $v_S(\theta) = \theta^2$ and it follows from (3.3) that $S'_{\tau^*}/n \xrightarrow{\text{d}} v_S(e^{-\tau^*}) = e^{-2\tau^*} \sim B(\frac{1}{2}, 1)$ (a beta distribution with density $\frac{1}{2}x^{-1/2}$ on $[0, 1]$), so in this case there is a non-degenerate limiting distribution of the size of the epidemic.

Recall that this example was considered at the end of Section 2.1. It is easily seen that for the two modifications considered there, with some vertices of degree 1 or 4, we have $\tau^* \xrightarrow{\text{P}} 0$ and $\tau^* \xrightarrow{\text{P}} \infty$, respectively, and thus $S'_{\tau^*}/n \xrightarrow{\text{P}} 1$ and $S'_{\tau^*}/n \xrightarrow{\text{P}} 0$.

4.2. Many initially infective half-edges: proof of Theorem 2.6. We now consider the case where there is initially a large number of infective half-edges, i.e. $\mu_I > 0$.

(a) It suffices to prove that h defined in (4.1) has a unique root $\theta_\infty \in (0, 1)$, since $h(1) = \mu_I > 0$ and $h(0) = -\alpha_S p_1 - \mu_R - \rho\mu/\beta < 0$ by (4.2) and (D6). Calculating $h''(\theta) = -\alpha_S \sum_{k=0}^\infty k(k-1)(k-2)\theta^{k-3}p_k \leqslant 0$ shows that h is concave on $(0, 1)$. This, together with the inequality $h(1) > 0$, implies uniqueness of θ_∞.

(b) By (G1), $\sum_{k=0}^\infty k^2 p_k < \infty$, see Remark 2.4, and by (2.12)–(2.15), the derivative of h_I is bounded on $[0, 1]$. Hence, p_I is Lipschitz continuous on $[\theta_\infty, 1]$. Consequently, both existence and uniqueness of the solution θ_t, $t \geqslant 0$, to (2.19) follow from standard theory. Note that the constant function θ_∞ is another solution to the differential equation, so $\theta_t > \theta_\infty$ for all t by uniqueness of solutions. Thus θ_t is strictly decreasing and bounded below, so the limit $\lim_{t\to\infty} \theta_t$ exists, and must be a zero of p_I, i.e., by part (a), the limit equals θ_∞.

In the proof below we will use a more explicit form of the solution. Let $\hat{\tau}_\infty := -\ln\theta_\infty$, and define $\hat{A} : [0, \hat{\tau}_\infty) \to [0, \infty)$ by

$$\hat{A}_\tau := \int_0^\tau \frac{d\sigma}{\beta p_I(e^{-\sigma})}, \qquad 0 \leqslant \tau < \hat{\tau}_\infty. \tag{4.6}$$

The integrand is strictly positive on $[0, \hat{\tau}_\infty)$ and so $\hat{A}$ is strictly increasing. Furthermore, $p_I(e^{-\hat{\tau}_\infty}) = 0$ and $p'_I(e^{-\sigma})$ is bounded for σ in a neighbourhood of $\hat{\tau}_\infty$. Hence $p_I(e^{-\sigma}) = O(\hat{\tau}_\infty - \sigma)$ for $\sigma \in [0, \hat{\tau}_\infty]$. It follows that $\hat{A}_\tau \nearrow \infty$ as $\tau \nearrow \hat{\tau}_\infty$. The inverse $\hat{\tau} : [0, \infty) \to [0, \hat{\tau}_\infty)$ of $\hat{A}$ is strictly increasing and continuously differentiable, and satisfies $\hat{\tau}'(t) = \beta p_I(\exp(-\hat{\tau}(t)))$ by the Inverse Function Theorem, and $\hat{\tau}(0) = 0$. So $\theta_t = \exp(-\hat{\tau}(t))$ solves (2.19).

(c) We first show that $\tau^* \xrightarrow{\text{P}} \hat{\tau}_\infty$. Let us take a sufficiently small $\varepsilon > 0$ and define

$$\delta := \inf_{\tau \leqslant \hat{\tau}_\infty - \varepsilon} h_I(e^{-\tau}) \wedge |h_I(e^{-(\hat{\tau}_\infty + \varepsilon)})|. \tag{4.7}$$

By part (a), $\delta > 0$. Then Lemma 3.1 (with $\tau_1 = \hat{\tau}_\infty + \varepsilon$) shows that w.h.p.

$$\sup_{\tau \leqslant \tau^* \wedge (\hat{\tau}_\infty + \varepsilon)} \left| X'_{I,\tau}/n - h_I(e^{-\tau}) \right| < \delta/2. \tag{4.8}$$

We claim that, on that event, $\hat{\tau}_\infty - \varepsilon < \tau^* < \hat{\tau}_\infty + \varepsilon$. Indeed, if $\tau^* \leqslant \hat{\tau}_\infty - \varepsilon$, then the left hand side of (4.8) it is at least $|X'_{I,\tau^*}/n - h_I(e^{-\tau^*})| = h_I(e^{-\tau^*}) \geqslant \delta$, by definition of

δ and the fact that $X'_{I,\tau^*} = 0$. If $\tau^* \geqslant \hat{\tau}_\infty + \varepsilon$, then the left hand side of (4.8) is at least $|X'_{I,\hat{\tau}_\infty+\varepsilon}/n - h_I(e^{-(\hat{\tau}_\infty+\varepsilon)})| \geqslant |h_I(e^{-(\hat{\tau}_\infty+\varepsilon)})| \geqslant \delta$, by definition of δ and the fact that $X'_{I,\hat{\tau}_\infty+\varepsilon} \geqslant 0$ and $h_I(e^{-(\hat{\tau}_\infty+\varepsilon)}) < 0$. Thus, w.h.p.

$$\hat{\tau}_\infty - \varepsilon < \tau^* < \hat{\tau}_\infty + \varepsilon. \tag{4.9}$$

Our next task is to return to the original process, and, to this end, we need to study the process $(A_\tau)_{\tau \geqslant 0}$ defined in (3.17) and its inverse $\tau(t)$. The integrand in (3.17) converges to $1/(\beta p_I(e^{-\sigma}))$ uniformly in probability on $[0, \hat{\tau}_\infty - \varepsilon]$ by (3.4), (3.6), and the fact that $h_I(e^{-\sigma}) \geqslant \delta > 0$ for $0 \leqslant \sigma \leqslant \hat{\tau}_\infty - \varepsilon$. Therefore,

$$A_\tau \xrightarrow{P} \hat{A}_\tau \tag{4.10}$$

uniformly over $0 \leqslant \tau \leqslant \hat{\tau}_\infty - \varepsilon$, where $\hat{A}_\tau$ is as in (4.6).

Let $t_1 := \hat{A}_{\hat{\tau}_\infty - 2\varepsilon}$. (We assume $\varepsilon < \hat{\tau}_\infty/2$.) The uniform convergence (4.10) and strict monotonicity of $\hat{A}$ imply that $A_{\hat{\tau}_\infty - \varepsilon} > t_1$ w.h.p. On this event, $\tau(t) \leqslant \tau(t_1) < \hat{\tau}_\infty - \varepsilon$ for $t \leqslant t_1$, and so (4.10) shows that

$$t - \hat{A}_{\tau(t)} = A_{\tau(t)} - \hat{A}_{\tau(t)} \xrightarrow{P} 0, \tag{4.11}$$

uniformly on $t \leqslant t_1$. Recall from the proof of (b) that $\hat{\tau}(t)$ is the inverse of the function $\hat{A}_\tau$; then $0 \leqslant \hat{\tau}'(t) = \beta p_I(\exp(-\hat{\tau}(t)) \leqslant \beta$, and so $\hat{\tau}$ is uniformly continuous. Hence, (4.11) implies

$$\sup_{t \leqslant t_1} |\hat{\tau}(t) - \tau(t)| \xrightarrow{P} 0. \tag{4.12}$$

Recall the definition $\bar{\tau}(t) := \tau(t) \wedge \tau^*$. If $t \geqslant t_1$, then w.h.p., using (4.12) and (4.9),

$$\hat{\tau}_\infty - 3\varepsilon = \hat{\tau}(t_1) - \varepsilon < \bar{\tau}(t_1) \leqslant \bar{\tau}(t) \leqslant \tau^* < \hat{\tau}_\infty + \varepsilon,$$

from which it follows that w.h.p.

$$\sup_{t \geqslant t_1} |\hat{\tau}(t) - \bar{\tau}(t)| < 3\varepsilon. \tag{4.13}$$

We conclude from (4.12)–(4.13) that $\bar{\tau}(t) \xrightarrow{P} \hat{\tau}(t)$, and hence that $\exp(-\bar{\tau}(t)) \xrightarrow{P} \theta_t = \exp(-\hat{\tau}(t))$, uniformly over all $t \geqslant 0$. It now follows that

$$\sup_{t \geqslant 0} |S_t/n - v_S(\theta_t)| = \sup_{t \geqslant 0} |S'_{\bar{\tau}(t)}/n - v_S(\theta_t)|$$

$$\leqslant \sup_{t \geqslant 0} |S'_{\bar{\tau}(t)}/n - v_S(e^{-\bar{\tau}(t)})| + \sup_{t \geqslant 0} |v_S(e^{-\bar{\tau}(t)}) - v_S(\theta_t)|$$

converges to zero in probability, by (3.2) and the uniform continuity of v_S on $[0, 1]$. The convergence statements in (2.22) follow similarly from Lemma 3.1 and uniform continuity of the functions h_S, h_I, h_R. The convergence of X_t/n also follows, as $X_t = X_{S,t} + X_{I,t} + X_{R,t}$.

Since $R_t/n = 1 - S_t/n - I_t/n$, it remains to prove convergence for the fraction I_t/n of infective vertices in (2.21). We will work directly with the original process, using a compactness argument. The number of infective vertices I_t increases by 1 when a free infective half-edge is paired with a free suceptible half-edge, and decreases by 1 when an infective vertex recovers. Therefore,

$$dI_t = \beta X_{I,t} \left(\frac{X_{S,t}}{X_t - 1} \right) dt - \rho I_t dt + dM_{I,t}, \tag{4.14}$$

where $(M_{\mathrm{I},t})_{t\geqslant 0}$ is a càdlàg martingale with $M_{\mathrm{I},0} = 0$ and quadratic variation

$$[M_{\mathrm{I}}]_t \leqslant \sum_{s\geqslant 0}(\Delta I_s)^2 \leqslant 2n,$$

since each vertex can only get infected or recover at most once. As in the proof of Lemma 3.1, Doob's inequality then gives $\sup_{t\geqslant 0}|M_{\mathrm{I},t}| = o_{\mathrm{p}}(n)$.

Writing (4.14) in integral form and dividing by n, we obtain

$$(I_t - I_0 - M_{\mathrm{I},t})/n = \int_0^t \left(\left(\frac{\beta X_{\mathrm{I},s}X_{\mathrm{S},s}}{n(X_s - 1)}\right) - \frac{\rho I_s}{n}\right) ds, \tag{4.15}$$

and the integrand is bounded in absolute value (by $2\beta\mu + \rho$ for n large). Hence, $(I_t - I_0 - M_{\mathrm{I},t})/n$, $n \geqslant 0$, is a uniformly Lipschitz family, and it is also uniformly bounded on each finite interval $[0, t_1]$. Thus, the Arzela–Ascoli theorem implies that it is tight in $C[0, t_1]$ for any $t_1 > 0$ [28, Theorems A2.1 and 16.5], and so also in $C[0, \infty)$. This then implies that there exists a subsequence along which the process converges in distribution in $C[0, \infty) \subset D[0, \infty)$, and the same holds for $(I_t - I_0)/n$ in $D[0, \infty)$ since $\sup_{t\geqslant 0}|M_{\mathrm{I},t}| = o_{\mathrm{p}}(n)$. The convergence may be assumed almost sure by the Skorokhod coupling lemma.

Hence, along the subsequence, I_t/n a.s. converges, uniformly on compact sets, to a continuous limit $\hat{I}_t$. Clearly, $\hat{I}_0 = \alpha_{\mathrm{I}}$. Let us show that $\hat{I}_t$ is deterministic. Since $M_{\mathrm{I},t} = o_{\mathrm{p}}(n)$, (4.15), the dominated convergence theorem, and the uniform convergence (2.22) of X_t/n, $X_{\mathrm{I},t}/n$ and $X_{\mathrm{S},t}/n$ (which may also be assumed a.s.), together imply that

$$\hat{I}_t = \lim_{n\to\infty} I_t/n = \alpha_{\mathrm{I}} + \int_0^t \left(\left(\frac{\beta h_{\mathrm{I}}(\theta_s)h_{\mathrm{S}}(\theta_s)}{h_X(\theta_s)}\right) - \rho\hat{I}_s\right) ds.$$

Consequently, $\hat{I}_t$ is continuously differentiable, and, differentiating, we obtain the differential equation (2.20) with the unique solution

$$\hat{I}_t = \alpha_{\mathrm{I}}e^{-\rho t} + \int_0^t e^{-\rho(t-s)}\left(\frac{\beta h_{\mathrm{I}}(\theta_s)h_{\mathrm{S}}(\theta_s)}{h_X(\theta_s)}\right) ds. \tag{4.16}$$

Hence, the subsequential limit $\hat{I}_t$ is uniquely determined, which implies convergence along the original sequence, and the convergence may be assumed almost sure. The limit is continuous, so the convergence is uniform on $[0, t_1]$, for any $t_1 > 0$.

Let $t_1 > 0$ and let $t \geqslant t_1$. The number of recovered vertices R_t is increasing in t, so

$$0 \leqslant I_t/n = 1 - S_t/n - R_t/n \leqslant 1 - S_t/n - R_{t_1}/n.$$

But $R_{t_1}/n \xrightarrow{\mathrm{p}} 1 - \hat{I}_{t_1} - v_{\mathrm{S}}(\theta_{t_1})$ and $S_t/n \geqslant \inf_{t\geqslant 0}S_t/n \xrightarrow{\mathrm{p}} v_{\mathrm{S}}(\theta_\infty)$. So for any $t_1 > 0$ and $\varepsilon > 0$, w.h.p. for all $t \geqslant t_1$,

$$0 \leqslant I_t/n \leqslant v_{\mathrm{S}}(\theta_{t_1}) - v_{\mathrm{S}}(\theta_\infty) + \hat{I}_{t_1} + \varepsilon, \tag{4.17}$$

and $v_{\mathrm{S}}(\theta_{t_1}) \to v_{\mathrm{S}}(\theta_\infty)$ as $t_1 \to \infty$, by continuity of v_{S}.

In the case $\rho > 0$, we make a change of variable $s = t - u$ in (4.16). Then $h_{\mathrm{I}}(\theta_{t-u}) \to h_{\mathrm{I}}(\theta_\infty) = 0$ as $t \to \infty$, and, together with the dominated convergence theorem, this shows that $\hat{I}_t \to 0$ as $t \to \infty$. Since $|I_t/n - \hat{I}_t| \leqslant \max\{I_t/n, \hat{I}_t\}$, it follows from (4.17) that w.h.p.

$$\sup_{t\geqslant t_1}|I_t/n - \hat{I}_t| \leqslant v_{\mathrm{S}}(\theta_{t_1}) - v_{\mathrm{S}}(\theta_\infty) + \sup_{t\geqslant t_1}\hat{I}_t + \varepsilon, \tag{4.18}$$

which is at most 2ε if t_1 is large enough.

In the case $\rho = 0$ (no recoveries), we instead note that I_t and $\hat{I}_t$ are non-decreasing. Then, since $\hat{I}_t \leqslant 1$, $\hat{I}_\infty := \lim_{t \to \infty} \hat{I}_t$ exists, and w.h.p.

$$\sup_{t \geqslant t_1}(\hat{I}_t - I_t/n) \leqslant \hat{I}_\infty - I_{t_1}/n \leqslant \hat{I}_\infty - \hat{I}_{t_1} + \varepsilon. \tag{4.19}$$

Together with (4.17), this shows that $\sup_{t \geqslant t_1} |I_t/n - \hat{I}_t| \leqslant 2\varepsilon$ w.h.p. if t_1 is large enough.

In summary, in both cases, for any $\varepsilon > 0$ we can choose t_1 such that $\sup_{t \geqslant t_1} |I_t/n - \hat{I}_t| \leqslant 2\varepsilon$ w.h.p. Since also $I_t/n \xrightarrow{\text{P}} \hat{I}_t$ uniformly on $[0, t_1]$, for any $t_1 > 0$, it follows that $I_t/n \xrightarrow{\text{P}} \hat{I}_t$ uniformly over the whole of $[0, \infty)$. This completes the proof of (2.21).

(d) This statement is an immediate consequence of (2.21).

Remark 4.2. If (D6) does not hold then $h(0) = 0$ and $h(1) = \mu_{\text{I}} > 0$, so $h(\theta) \geqslant \mu_{\text{I}}\theta > 0$ for $\theta \in (0, 1]$, by concavity of h. So the only root of $h_{\text{I}}(\theta)$ in $[0, 1]$ is at zero, and the argument in (c) above shows that, for any fixed $\tau_1 > 0$, $\tau^* > \tau_1$ w.h.p. Thus $\tau^* \xrightarrow{\text{P}} \infty$, and it follows from (3.2) that $S'_{\tau^*}/n \xrightarrow{\text{P}} v_{\text{S}}(0) = \alpha_{\text{S}} p_0$. Hence, apart from the isolated vertices, all but $o_{\text{p}}(n)$ of the susceptible vertices succumb to infection.

4.3. Supercriticality: proof of Theorem 2.9 part (ii).

We now consider the 'supercritical' regime, where the basic reproductive ratio $\mathfrak{R}_0 > 1$ and there are few initially infective vertices ($\mu_{\text{I}} = 0$). Let us start by noting that $h_{\text{I}}(1) = \mu_{\text{I}} = 0$, see (2.17).

We will say that an event $\mathcal{E}_n$ holds w.h.p. conditional on event $\mathcal{D}_n$ if $\mathbb{P}(\mathcal{E}_n \mid \mathcal{D}_n) = 1 - o(1)$ as $n \to \infty$, or, equivalently, if $\mathbb{P}(\mathcal{E}_n^c \mid \mathcal{D}_n) = o(1)$ as $n \to \infty$. If $\mathbb{P}(\mathcal{D}_n)$ is bounded away from 0 as $n \to \infty$, then $\mathcal{E}_n$ holds w.h.p. conditional on $\mathcal{D}_n$ if and only if $\mathbb{P}(\mathcal{D}_n, \mathcal{E}_n^c) = o(1)$.

(a) It is sufficient to show that h defined in (4.1) has a unique root in $(0, 1)$. Differentiating h and then substituting $\beta\alpha_{\text{S}} \sum_{k=0}^\infty k(k-1)p_k = \mu(\rho + \beta)\mathfrak{R}_0$ yields, cf. (4.3),

$$\beta h'(1) = \mu(\rho + \beta)(1 - \mathfrak{R}_0) < 0.$$

Together with the fact that $h(1) = h_{\text{I}}(1) = \mu_{\text{I}} = 0$, this shows that $h(\theta) > 0$ for all $\theta < 1$ close enough to 1. Furthermore, $h(0) = -\alpha_{\text{S}} p_1 - \mu_{\text{R}} - \rho\mu/\beta < 0$ by (D6), and so there is at least one root in $(0, 1)$. Uniqueness follows from concavity of h, as in the proof of Theorem 2.6(a).

(b) As in the proof of Theorem 2.6, existence and uniqueness of the solution to (2.25) follow from Lipschitz continuity of p_{I} and standard theory, but in the proof below we will use a more explicit form of the solution.

Recall that there is an extra parameter $s_0 \in (v_{\text{S}}(\theta_\infty), v_{\text{S}}(1))$ used in the initial condition of (2.25) to calibrate the time shift. Let $\hat{\tau}_0 := -\ln v_{\text{S}}^{-1}(s_0)$ and $\hat{\tau}_\infty := -\ln\theta_\infty$; thus $0 < \hat{\tau}_0 < \hat{\tau}_\infty$. We define $\hat{A} : (0, \hat{\tau}_\infty) \to \mathbb{R}$ by

$$\hat{A}_\tau := \int_{\hat{\tau}_0}^\tau \frac{d\sigma}{\beta p_{\text{I}}(e^{-\sigma})}, \qquad 0 < \tau < \hat{\tau}_\infty. \tag{4.20}$$

Then $\hat{A}$ is strictly increasing, $\hat{A}_{\hat{\tau}_0} = 0$, and $\hat{A}_\tau \nearrow \infty$ as $\tau \nearrow \hat{\tau}_\infty$. By (G1), p_{I} is Lipschitz continuous on $[\theta_\infty, 1]$ and $p_{\text{I}}(e^{-\sigma}) = O(\sigma)$, because $p_{\text{I}}(1) = 0$; hence $\hat{A}_\tau \searrow -\infty$ as $\tau \searrow 0$. The inverse $\hat{\tau} : \mathbb{R} \to (0, \hat{\tau}_\infty)$ of $\hat{A}$ is strictly increasing and continuously differentiable. Also,

$$\hat{\tau}'(t) = \beta p_{\text{I}}(\exp(-\hat{\tau}(t))), \tag{4.21}$$

21

and $\hat{\tau}(0) = \hat{\tau}_0$. Hence, $\theta_t = \exp(-\hat{\tau}(t))$ satisfies (2.25). The final statements follow from $\hat{\tau}(t) \nearrow \hat{\tau}_\infty = -\ln \theta_\infty$ as $t \to \infty$ and $\hat{\tau}(t) \searrow 0$ as $t \to -\infty$.

(c) The proof of this part is deferred till Section 4.3.1.

(d) Proceeding as in the proof of Theorem 2.6(c), but now defining

$$\delta := \inf_{\varepsilon \leqslant \tau \leqslant \hat{\tau}_\infty - \varepsilon} h_{\mathrm{I}}(e^{-\tau}) \wedge |h_{\mathrm{I}}(e^{-(\hat{\tau}_\infty + \varepsilon)})| \tag{4.22}$$

instead of (4.7), we deduce that, for any $\varepsilon > 0$, either $0 \leqslant \tau^* < \varepsilon$ or $\hat{\tau}_\infty - \varepsilon < \tau^* < \hat{\tau}_\infty + \varepsilon$, w.h.p.; in other words, the time-changed epidemic dies out either almost instantaneously or around time $\hat{\tau}_\infty := -\ln \theta_\infty$.

Recall that $T_0 := \inf\{t \geqslant 0 : S_t \leqslant ns_0\}$, and assume that $\varepsilon > 0$ is small enough that $v_{\mathrm{S}}(e^{-\varepsilon}) > s_0$. By (3.2),

$$\inf_{t \geqslant 0} S_t/n = S'_{\tau^*}/n = v_{\mathrm{S}}(e^{-\tau^*}) + o_{\mathrm{p}}(1),$$

and so $\mathbb{P}(\tau^* < \varepsilon, T_0 < \infty) = o(1)$.

For the rest of this proof, we will condition on the event $T_0 < \infty$. Since by part (c), $\mathbb{P}(T_0 < \infty)$ is bounded away from 0 as $n \to \infty$, the previous two paragraphs show that, for sufficiently small $\varepsilon > 0$, $\hat{\tau}_\infty - \varepsilon < \tau^* < \hat{\tau}_\infty + \varepsilon$ holds w.h.p. conditional on $T_0 < \infty$.

As in the proof of Theorem 2.6, we need to study the inverse $\tau(t)$ of the process A_τ defined in (3.17), now with the added complication of the time shift T_0. Here, we extend the definition of τ to $(-\infty, 0)$ by letting $\tau(t) := 0$ for $t < 0$; note that this agrees with our convention to take $S_t = S_0$ for $t < 0$, etc., and that (3.18) holds for all $t \in (-\infty, \infty)$.

The calibration time $T_0 = \inf\{t \geqslant 0 : S'_{\tau(t)} \leqslant s_0 n\}$ satisfies $\tau(T_0) \xrightarrow{\mathrm{P}} \hat{\tau}_0 := -\ln v_{\mathrm{S}}^{-1}(s_0)$, by (3.2) and the fact that $v_{\mathrm{S}}(e^{-\tau})$ is strictly decreasing. (We use part (c) to guarantee that the convergence in probability survives the conditioning. This is done without further comment below.)

By (3.4), (3.6), and the fact that $\inf_{\varepsilon \leqslant \sigma \leqslant \hat{\tau}_\infty - \varepsilon} h_{\mathrm{I}}(e^{-\sigma}) > 0$,

$$A_\tau - A_{\tau'} \xrightarrow{\mathrm{P}} \int_{\tau'}^{\tau} \frac{d\sigma}{\beta p_{\mathrm{I}}(e^{-\sigma})} = \hat{A}_\tau - \hat{A}_{\tau'}, \tag{4.23}$$

uniformly over $\tau, \tau' \in [\varepsilon, \hat{\tau}_\infty - \varepsilon]$. Let us assume that ε is small enough that $\hat{\tau}_0 \in [2\varepsilon, \hat{\tau}_\infty - 2\varepsilon]$. Taking $\tau' = \tau(T_0)$ in (4.23) and using the fact that $\tau(T_0) \xrightarrow{\mathrm{P}} \hat{\tau}_0$ gives

$$A_\tau - T_0 = A_\tau - A_{\tau(T_0)} \xrightarrow{\mathrm{P}} \hat{A}_\tau - \hat{A}_{\hat{\tau}_0} = \hat{A}_\tau, \tag{4.24}$$

uniformly over $\tau \in [\varepsilon, \hat{\tau}_\infty - \varepsilon]$.

Now suppose that $t_1 > 0$ is given. We may assume that ε is small enough that $t_1 < (-\hat{A}_\varepsilon) \wedge \hat{A}_{\hat{\tau}_\infty - \varepsilon}$ (recall from part (b) that the right-hand side diverges as $\varepsilon \to 0$). Then $A_\varepsilon - T_0 \xrightarrow{\mathrm{P}} \hat{A}_\varepsilon < -t_1$ by (4.24), so, w.h.p. conditional on $T_0 < \infty$, $A_\varepsilon < T_0 - t_1$ and thus $\tau(T_0 - t_1) > \varepsilon$. (In particular, $T_0 - t_1 > 0$, w.h.p. conditional on $T_0 < \infty$. Since t_1 is arbitrary, this shows $T_0 \xrightarrow{\mathrm{P}} \infty$.)

Similarly, w.h.p. conditional on $T_0 < \infty$, $A_{\hat{\tau}_\infty - \varepsilon} - T_0 > t_1$ and $\tau(T_0 + t_1) < \hat{\tau}_\infty - \varepsilon$. Thus, w.h.p. conditional on $T_0 < \infty$, $\tau(T_0 + t) \in [\varepsilon, \hat{\tau}_\infty - \varepsilon]$ for all $t \in [-t_1, t_1]$. So, taking $\tau = \tau(T_0 + t)$ in (4.24) shows that $\hat{A}_{\tau(T_0 + t)} \xrightarrow{\mathrm{P}} t$ uniformly on $t \in [-t_1, t_1]$. Thus $\tau(T_0 + t) \xrightarrow{\mathrm{P}} \hat{\tau}(t)$ uniformly over $t \in [-t_1, t_1]$ by uniform continuity of $\hat{\tau}(t)$, and this holds for any $t_1 > 0$.

Furthermore, as shown above, w.h.p. conditional on $T_0 < \infty$, $\tau(T_0 + t) < \hat{\tau}_\infty - \varepsilon < \tau^*$ for all $t \in [-t_1, t_1]$, and thus $\bar{\tau}(T_0 + t) := \tau(T_0 + t) \wedge \tau^* = \tau(T_0 + t)$ for $t \leqslant t_1$. Hence,

$$\bar{\tau}(T_0 + t) \xrightarrow{\text{P}} \hat{\tau}(t), \tag{4.25}$$

uniformly over $t \in [-t_1, t_1]$. As before, uniform convergence for $|t| \geqslant t_1$ follows by monotonicity: if $t \geqslant t_1$ then w.h.p. conditional on $T_0 < \infty$, by (4.25), $0 \leqslant \bar{\tau}(T_0 - t) \leqslant \bar{\tau}(T_0 - t_1) < \hat{\tau}(-t_1) + \varepsilon$ and $\hat{\tau}(t_1) - \varepsilon < \bar{\tau}(T_0 + t_1) \leqslant \bar{\tau}(T_0 + t) \leqslant \tau^* < \hat{\tau}_\infty + \varepsilon$. Hence, w.h.p. conditional on $T_0 < \infty$, $|\bar{\tau}(T_0 - t) - \hat{\tau}(-t)| \leqslant \max\{\bar{\tau}(T_0 - t), \hat{\tau}(-t)\} \leqslant \hat{\tau}(-t_1) + \varepsilon$ and $|\bar{\tau}(T_0 + t) - \hat{\tau}(t)| \leqslant \hat{\tau}_\infty - \hat{\tau}(t_1) + \varepsilon$ for all $t \geqslant t_1$, and the right hand sides can be made arbitrarily small by choosing t_1 large and ε small. Hence, (4.25) holds uniformly on $\mathbb{R}$.

Consequently, $\exp(-\bar{\tau}(T_0 + t)) \xrightarrow{\text{P}} \exp(-\hat{\tau}(t)) = \theta_t$ uniformly on $\mathbb{R}$. The convergence $S_{T_0+t}/n \xrightarrow{\text{P}} v_{\text{S}}(\theta_t)$ in (2.27) and the limits in (2.28) then follow from Lemma 3.1 and uniform continuity of the limit functions, just as in the proof of Theorem 2.6.

The compactness argument from Section 4.2 shows that I_{T_0+t}/n converges in distribution in $D(-\infty, \infty)$ along a subsequence to a differentiable process $\hat{I}_t$ that satisfies

$$\frac{d}{dt}\hat{I}_t = \beta \frac{h_{\text{I}}(\theta_t) h_{\text{S}}(\theta_t)}{h_X(\theta_t)} - \rho \hat{I}_t, \qquad t \in \mathbb{R}. \tag{4.26}$$

Recovered vertices never become infective, and so, by (2.27) and the fact that $n_{\text{I}}/n \to \alpha_{\text{I}} = 0$,

$$0 \leqslant I_{T_0+t}/n \leqslant n_{\text{I}}/n + (n_{\text{S}} - S_{T_0+t})/n \xrightarrow{\text{P}} \alpha_{\text{S}} - v_{\text{S}}(\theta_t). \tag{4.27}$$

Hence, $\hat{I}_t \leqslant \alpha_{\text{S}} - v_{\text{S}}(\theta_t)$, which tends to zero as $t \to -\infty$ and hence $\hat{I}_t \to 0$ as well. It follows that the subsequential limit $\hat{I}_t$ is unique and deterministic, given by

$$\hat{I}_t = \int_{-\infty}^t e^{-\rho(t-s)} \frac{\beta h_{\text{I}}(\theta_s) h_{\text{S}}(\theta_s)}{h_X(\theta_s)} \, ds, \tag{4.28}$$

which implies that $I_{T_0+t}/n \to \hat{I}_t$ in distribution along the original sequence. The convergence can be assumed almost sure by the Skorokhod coupling lemma, and so, since $\hat{I}_t$ is continuous, $I_{T_0+t}/n \to \hat{I}_t$ uniformly on $[-t_1, t_1]$, a.s. for any $t_1 > 0$.

Take any $\varepsilon > 0$. Then, w.h.p. conditional on $T_0 < \infty$, $I_{T_0+t}/n < \alpha_{\text{S}} - v_{\text{S}}(\theta_{-t_1}) + \varepsilon$ for any $t \leqslant -t_1$, by (4.27) and monotonicity of S_t. Thus, w.h.p. conditional on $T_0 < \infty$, $\sup_{t \leqslant -t_1} |I_{T_0+t}/n - \hat{I}_t| < 2\varepsilon$ if t_1 is large enough. Furthermore, the argument leading to (4.18) and (4.19) gives $\sup_{t \geqslant t_1} |I_{T_0+t}/n - \hat{I}_t| < 2\varepsilon$ w.h.p. conditional on $T_0 < \infty$, for t_1 large enough.

Hence $I_{T_0+t}/n \xrightarrow{\text{P}} \hat{I}_t$ uniformly on $\mathbb{R}$ and the remaining parts of (2.27) follow.

(e) This is an immediate consequence of (d).

(f) Given $s \in (v_{\text{S}}(\theta_\infty), v_{\text{S}}(1))$, let $T(s) := \inf\{t \geqslant 0 : S_t \leqslant ns\}$, so that $T_0 = T(s_0)$. We will show that, for any $s_1, s_2 \in (v_{\text{S}}(\theta_\infty), v_{\text{S}}(1))$ with $s_1 < s_2$, $\mathbb{P}(T(s_2) < \infty, T(s_1) = \infty) = o(1)$. (Clearly, $\mathbb{P}(T(s_2) = \infty, T(s_1) < \infty) = 0$.) By part (e), conditional on the event $T(s_2) < \infty$, $S_\infty/n \xrightarrow{\text{P}} v_{\text{S}}(\theta_\infty)$, and so $S_\infty/n < s_1$ w.h.p. conditional on $T(s_2) < \infty$, which implies that, w.h.p. conditional on $T(s_2) < \infty$, $S_t \leqslant ns_1$ for some $t < \infty$ and thus $T(s_1) < \infty$. So $\mathbb{P}(T(s_2) < \infty, T(s_1) = \infty) = o(1)$, as required.

Now, recall that $s_0 \in (v_{\text{S}}(\theta_\infty), v_{\text{S}}(1))$. Let $\varepsilon > 0$; by the above, $\mathbb{P}(T_0 = \infty, T(v_{\text{S}}(1) - \epsilon/2) < \infty) = o(1)$, and so $\mathbb{P}(T_0 = \infty, S_t \leqslant n(v_{\text{S}}(1) - \epsilon/2) \text{ for some } t) = o(1)$. It then follows that

23

$\mathbb{P}(T_0 = \infty, S_\infty < n(v_S(1) - \epsilon/2)) = o(1)$, and so $\mathbb{P}(T_0 = \infty, S_0 - S_\infty > \varepsilon n) = o(1)$. Since ε can be taken arbitrarily small, we have that $S_0 - S_\infty = o_p(n)$ on the event $T_0 = \infty$.

Since $S_0 - S_\infty = o_p(n)$ on $T_0 = \infty$, also $\sup_{t \geqslant 0} I_t \leqslant n_I + S_0 - S_\infty = o_p(n)$ on $T_0 = \infty$. Further, the uniform summability of $\sum_{k=0}^{\infty} k n_{S,k}/n$ (see Remark 2.2) implies that the number $X_{S,0} - X_{S,\infty}$ of half-edges that ever get infected is $o_p(n)$ on $T_0 = \infty$ also. Hence $\sup_{t \geqslant 0} X_{I,t} = o_p(n)$ on $T_0 = \infty$, and only $o_p(n)$ half-edges pair off before the end of the epidemic, so $\sup_{t \geqslant 0}(X_0 - X_t) = o_p(n)$ on $T_0 = \infty$.

4.3.1. *Proof of part (c): there is a large epidemic with probability bounded away from zero.* To study the beginning phase of the epidemic, we concentrate on the number $X_{I,t}$ of free infective half-edges. It is convenient to colour the half-edges of a newly infected vertex according to whether the vertex recovers before the half-edge can pair off. More specifically, as soon as a new vertex is infected, we give it a random recovery time with distribution $\mathrm{Exp}(\rho)$ and each of its remaining half-edges an infection time with distribution $\mathrm{Exp}(\beta)$; these times are independent of each other and everything else. We then colour each of these half-edges red if its infection time is smaller than the recovery time of the vertex, and black otherwise. The black half-edges thus will recover without pairing off, so we can ignore them, while the red half-edges will pair off at some time. Let Z_t be the number of free red half-edges at time t (for this proof, we need only consider the process in its original time scale).

We fix a small $\varepsilon > 0$ and stop if and when either $X_{I,t}$ becomes at least εn or at least εn infective half-edges have paired off. Of course, we also stop if $X_{I,t}$ becomes 0. Before stopping, the number X_t of free half-edges thus satisfies

$$\sum_{k=0}^{\infty} k n_k = X_0 \geqslant X_t \geqslant X_0 - 2\varepsilon n. \tag{4.29}$$

Furthermore, before stopping, there are at least $n_{S,k} - \varepsilon n$ remaining susceptible vertices of degree k, for each k. Hence, when an infective half-edge pairs off, it will infect a vertex of degree k with probability at least

$$\frac{k(n_{S,k} - \varepsilon n)}{\sum_{k=0}^{\infty} k n_k} = \frac{k(\alpha_S p_k + o(1) - \varepsilon)n}{(\mu + o(1))n} = \frac{k(\alpha_S p_k - \varepsilon)}{\mu} + o(1) \geqslant \frac{k(\alpha_S p_k - 2\varepsilon)}{\mu}, \tag{4.30}$$

if n is large enough. (Uniformly in k because we only have to consider the finite set of k for which the final expresssion is nonnegative.) If a vertex of degree k is infected, this produces $k - 1$ new free infective half-edges. Of these, some random number Y_{k-1} will be red; the distribution of Y_{k-1} can easily be given explicitly, see Lemma 5.1 below, but here we only need the simple fact that its mean is

$$\mathbb{E}[Y_{k-1}] = \frac{\beta}{\beta + \rho}(k - 1). \tag{4.31}$$

Since the infecting half-edge is removed from among the free infective half-edges, the net change in Z_t is $Y_{k-1} - 1$.

When a half-edge is paired off, there is also a possibility that it gets joined to another infective half-edge, in which case that half-edge is removed from the free infective half-edges and Z_t is reduced by 1 or 2 (depending on whether the second half-edge was black or red);

this happens with probability at most

$$\frac{X_{I,t}}{\sum_{k=0}^{\infty} kn_k - 2\varepsilon n} \leqslant \frac{\varepsilon n}{(\mu + o(1) - 2\varepsilon)n} < 2\varepsilon/\mu \tag{4.32}$$

if n is large and ε small.

Finally, it may happen that the infective half-edge connects to a recovered half-edge. In this case Z_t is reduced by 1.

Consequently, if t_i is the i:th time an infective half-edge is paired off, then

$$\Delta Z_{t_i} \geqslant -1 + U_i, \tag{4.33}$$

where U_i is independent of the previous history and and has the mixture distribution

$$U_i = \begin{cases} Y_{k-1} & \text{with probability } k(\alpha_S p_k - 2\varepsilon)_+/\mu, \text{ for each } k \geqslant 1, \\ -1 & \text{with probability } 2\varepsilon/\mu \\ 0 & \text{otherwise.} \end{cases} \tag{4.34}$$

We may assume that $\varepsilon \leqslant \mu/2$, and then U_i is well-defined, because either $(\alpha_S p_k - 2\varepsilon)_+ = 0$ for every $k \geqslant 1$, or there exists a k with $\alpha_S p_k > 2\varepsilon$, and then

$$\sum_{k=0}^{\infty} k(\alpha_S p_k - 2\varepsilon)_+ \leqslant \sum_{k=0}^{\infty} k\alpha_S p_k - 2\varepsilon = \alpha_S \lambda - 2\varepsilon \leqslant \mu - 2\varepsilon.$$

The random variable U_i has expectation

$$\mathbb{E}\, U_i = \sum_{k=1}^{\infty} \frac{k(\alpha_S p_k - 2\varepsilon)_+}{\mu} \mathbb{E}\, Y_{k-1} - \frac{2\varepsilon}{\mu} = \frac{\beta}{(\beta + \rho)\mu} \sum_{k=1}^{\infty} (k-1)k(\alpha_S p_k - 2\varepsilon)_+ - \frac{2\varepsilon}{\mu} \tag{4.35}$$

which, as $\varepsilon \searrow 0$, converges by monotone convergence to

$$\frac{\beta}{(\beta + \rho)\mu} \sum_{k=1}^{\infty} (k-1)k\alpha_S p_k = \mathfrak{R}_0 > 1. \tag{4.36}$$

Thus we may assume that ε is chosen so small that $\mathbb{E}\, U_i > 1$, and thus $\mathbb{E}[U_i - 1] > 0$.

By (4.33), the sequence $Z_{t_i} - Z_0$ $(i \geqslant 1)$, until we stop, dominates a random walk starting at 0, with i.i.d. increments $U_i - 1$ such that $\mathbb{E}[U_i - 1] > 0$. It is well-known that such a random walk is transient and diverges to $+\infty$, so its minimum M_- is a.s. finite. If the process stops at time t_i by running out of infective half-edges, we must have $Z_{t_i} = 0$, and thus $M_- \leqslant -Z_0^{(n)}$ (the n dependency is added to Z here as a reminder that M does *not* depend on n).

Suppose that the initial number of infective half-edges $X_{I,0} = \sum_{k=0}^{\infty} kn_{I,k} \to \infty$ (however slowly). Then also the initial number of red infective half-edges $Z_0^{(n)} \to \infty$ in probability; to see this, observe that either there are at least $\sqrt{X_{I,0}}$ infective vertices with at least one half-edge (and the half-edges for different vertices are coloured independently), or there is at least one infective vertex with at least $\sqrt{X_{I,0}}$ half-edges (each of which was coloured independently, given the recovery time). Hence $\mathbb{P}(M_- \leqslant -Z_0^{(n)}) \to 0$, and w.h.p. we do not stop by running out of red half-edges.

Since the process must stop at some point, w.h.p. either the process stops by producing at least εn infective half-edges or because at least εn half-edges have paired off. In either case, at least one of $X_{I,t}/n$ and X_t/n differs by at least ε from its initial value, so we cannot

25

have convergence to the trivial constant solution. Thus w.h.p. $T_0 < \infty$, see (f). We deduce that $X_{I,0} \to \infty$ implies $T_0 < \infty$ w.h.p.

Finally, provided there is initially at least one infective half-edge, we have $\mathbb{P}(Z_0^{(n)} \geqslant 1) \geqslant \beta/(\rho + \beta)$. Further, $Z_0^{(n)} \geqslant 1$ occurs independently of the steps in the random walk above and it is straightforward to show $\mathbb{P}(M_- = 0) > 0$. Hence, the probability of stopping by running out of red half-edges is at most

$$\mathbb{P}(M_- \leqslant -Z_0^{(n)}) \leqslant 1 - \frac{\beta}{\rho + \beta}\mathbb{P}(M_- = 0) < 1. \tag{4.37}$$

It follows that, with probability bounded away from zero, we stop by producing at least εn infective half-edges or because at least εn half-edges have been paired off. As noted above, this implies that $X_{I,t}/n$ and X_t/n cannot converge to constants, and thus w.h.p. $T_0 < \infty$.

5. The probability of a large outbreak and the size of a small outbreak

In this section, we obtain an expression for the probability $\mathbb{P}(T_0 < \infty)$ in Theorem 2.9(ii)(c) using a branching process approximation of the initial steps. We thus assume (D1)–(D6), $\alpha_I = \mu_I = 0$ and $\mathfrak{R}_0 > 1$. We begin by finding the distribution of the variables Y_{k-1} defined in Section 4.3.1. We denote falling factorials by $(n)_j := n \cdots (n - j + 1)$, and let $_2F_1$ denote the hypergeometric function, as defined in e.g. [36].

Lemma 5.1. *Consider an infective vertex with $\ell \geqslant 0$ free half-edges and let Y_ℓ be the number of them that pair off before the vertex recovers (assuming that none of them is chosen by another infective half-edge in a pairing event). If $\rho > 0$, then*

$$\mathbb{P}(Y_\ell = j) = \frac{\rho\,\ell!\,\Gamma(\ell + \rho/\beta - j)}{\beta\,(\ell - j)!\,\Gamma(\ell + \rho/\beta + 1)} = \frac{\rho}{\ell\beta + \rho} \cdot \frac{(\ell)_j}{(\ell + \rho/\beta - 1)_j}, \tag{5.1}$$

and the probability generating function $g_\ell(x) := \mathbb{E}\,x^{Y_\ell}$ is given by

$$g_\ell(x) = \sum_{j=0}^{\ell} \mathbb{P}(Y_\ell = j)x^j = \frac{\rho}{\ell\beta + \rho} \cdot {}_2F_1\left(-\ell, 1; -\ell - \frac{\rho}{\beta} + 1; x\right). \tag{5.2}$$

If $\rho = 0$, then $Y_\ell = \ell$ and $g_\ell(x) = x^\ell$.

Proof. The case $\rho = 0$ is trivial, so assume $\rho > 0$. By conditioning on the time of recovery, and changing variables $x = e^{-\beta t}$ to obtain a beta integral,

$$\mathbb{P}(Y_\ell = j) = \int_0^\infty \binom{\ell}{j}\left(1 - e^{-\beta t}\right)^j e^{-(\ell - j)\beta t}e^{-\rho t}\rho\,dt$$

$$= \int_0^1 \binom{\ell}{j}(1 - x)^j x^{\ell - j + \rho/\beta - 1}\frac{\rho}{\beta}\,dx$$

$$= \frac{\rho}{\beta}\binom{\ell}{j}\frac{\Gamma(\ell - j + \rho/\beta)\Gamma(j + 1)}{\Gamma(\ell + \rho/\beta + 1)},$$

which can be written as in (5.1), and (5.2) follows from the definition of hypergeometric functions. $\square$

Remark 5.2. The standard hypergeometric distribution obtained by drawing n balls from an urn with N balls of which m are white has probability generating function $c \cdot {}_2F_1(-n, -m; N - n - m + 1; x)$, where $c = (N - n)!\,(N - m)!/\big((N - m - n)!\,N!\big)$ is a normalization constant. Hence the distribution of Y_ℓ can formally be regarded as a 'negative hypergeometric distribution', with parameters $(n, m, N) = (\ell, -1, -1 - \rho/\beta)$.

Next, we define ξ to be the random variable obtained by mixing Y_{k-1} with probabilities $\alpha_{\mathrm{S}} k p_k/\mu$; to be precise,

$$\xi = \begin{cases} Y_{k-1} & \text{with probability } \alpha_{\mathrm{S}} k p_k/\mu, \text{ for each } k \geqslant 1, \\ 0 & \text{with probability } \mu_{\mathrm{R}}/\mu, \end{cases} \tag{5.3}$$

where the probabilities sum to 1, since we assume $\mu_{\mathrm{I}} = 0$. (If $\mu_{\mathrm{R}} = 0$, we thus mix using the size-biased distribution corresponding to $(p_k)_0^\infty$.) Note that this is the limiting case $\varepsilon = 0$ of (4.34). We have, by (4.31) and (2.23), cf. (4.35),

$$\mathbb{E}\,\xi = \sum_{k=1}^\infty \frac{\alpha_{\mathrm{S}} k p_k}{\mu}\,\mathbb{E}\,Y_{k-1} = \frac{\alpha_{\mathrm{S}}}{\mu} \cdot \frac{\beta}{\beta + \rho} \sum_{k=1}^\infty (k - 1)k p_k = \mathfrak{R}_0 > 1. \tag{5.4}$$

Theorem 5.3. *Suppose that the assumptions of Theorem 2.9(ii) (Theorem 2.10(ii)) hold. Let $q \in (0, 1)$ be the extinction probability for a Galton–Watson process with offspring distribution ξ, starting with a single individual. Then for the SIR epidemic on the multigraph G^* (simple graph G)*

$$\mathbb{P}(T_0 = \infty) = \prod_{k=1}^\infty g_k(q)^{n_{\mathrm{I},k}} + o(1). \tag{5.5}$$

Hence the probability of a large outbreak is $1 - \prod_{k=1}^\infty g_k(q)^{n_{\mathrm{I},k}} + o(1)$.

Note that $g_k(q)$ is the extinction probability if we start the Galton–Watson process with Y_k individuals instead of one. Hence the product in (5.5) is the extinction probability if we start with a number of individuals distributed as Y_k for each initially infected vertex of degree k, with these numbers independent. (This is just the number of initially red infective half-edges in Section 4.3.1.)

Before proving Theorem 5.3, we state another theorem, saying that a small outbreak infects only $O_{\mathrm{p}}(1)$ vertices. This is valid in the supercritical case without further assumption, and also in the (sub)critical case (when the outbreak is small w.h.p.) provided the initial number of infected half-edges $X_{\mathrm{I},0} = \sum_{k=0}^\infty k n_{\mathrm{I},k}$, is bounded. (This is equivalent to the initial number of infected individuals and their degrees being bounded.)

Theorem 5.4. *Suppose that the assumptions of Theorem 2.9 (Theorem 2.10) are satisfied. Then the following holds for the SIR epidemic on the multigraph G^* (simple graph G).*

(i) *If $\mathfrak{R}_0 \leqslant 1$ and further the initial number of infective half-edges $X_{\mathrm{I},0} = O(1)$, then the number $n_{\mathrm{S}} - S_\infty$ of susceptible vertices that ever get infected is $O_{\mathrm{p}}(1)$.*

(ii) *If $\mathfrak{R}_0 > 1$, then a small outbreak infects only $O_{\mathrm{p}}(1)$ vertices. In other words, for every $\varepsilon > 0$ there exists $K_\varepsilon < \infty$ such that*

$$\mathbb{P}(T_0 = \infty \text{ and } n_{\mathrm{S}} - S_\infty > K_\varepsilon) < \varepsilon. \tag{5.6}$$

Remark 5.5. The assumption $X_{I,0} = O(1)$ is easily seen to be necessary in part (i) of this result. Indeed, if $X_{I,0} \to \infty$, then the number of initially infective half-edges that pair off before recovering tends to infinity in probability. While there are at least δn susceptible vertices, then the probability that at a pairing event an infective half-edge is joined to a half-edge at a susceptible vertex is asymptotically bounded away from 0. Hence, the number of susceptibles infected tends to infinity in probability.

In the supercritical case this can be assumed because of Theorem 2.9(ii)(c) (Theorem 2.10(ii)(c)).

We leave it to the reader to study the precise size of the epidemic in the subcritical case when $X_{I,0} \to \infty$. The critical case ($\mathfrak{R}_0 = 1$) is examined in a forthcoming work [23].

Theorems 5.3 and 5.4 are valid both in the simple graph case and the multigraph case of Theorem 2.9. We prove the multigraph case here and defer the simple graph case to Section 6.

Proof of Theorems 5.3 and 5.4 in the multigraph case. We begin with the supercritical case $\mathfrak{R}_0 > 1$. Since $\mathbb{E}\,\xi = \mathfrak{R}_0 > 1$ by (5.4), the Galton–Watson process is supercritical, and thus $0 < q < 1$ as asserted.

Fix a small $\varepsilon > 0$. We have shown in Section 4.3.1 that, if $X_{I,0} \to \infty$, then $\mathbb{P}(T_0 < \infty) \to 1$. It follows that there exists a finite N (depending on the parameters β, ρ, $(p_k)_0^\infty$, etc., but not on n) such that, if $X_{I,0} \geqslant N$, then $\mathbb{P}(T_0 < \infty) > 1 - \varepsilon$. (If not, then there would exist a sequence of initial conditions satisfying (D1)–(D6) with $X_{I,0} \to \infty$, and thus $n \to \infty$, but $\mathbb{P}(T_0) \leqslant 1 - \varepsilon$.)

Now consider the Galton–Watson process $(W_i)_{i=0}^\infty$ with offspring distribution ξ and some initial W_0. Since the process is supercritical, a.s. either $W_i = 0$ for some i, and thus also for all larger i, or $W_i \to \infty$ as $i \to \infty$. We consider also the corresponding random walk stopped at 0, defined by $\hat{Z}_0 = W_0$, $\hat{Z}_{i+1} = 0$ if $\hat{Z}_i = 0$ and $\hat{Z}_{i+1} = \hat{Z}_i - 1 + \xi_{i+1}$ if $\hat{Z}_i > 0$, where (ξ_i) are i.i.d. copies of ξ. This can be regarded as the Galton–Watson process modified so that different individuals give birth at different times, and thus $W_i \to \infty$ (0) $\iff \hat{Z}_i \to \infty$ (0).

A.s., $\hat{Z}_i \to \infty$ unless the random walk hits 0 and thus is absorbed there. Thus, $\mathbb{P}(0 < \hat{Z}_i < N) \to 0$ as $i \to \infty$, and it follows that there exists some $K < \infty$ such that, for any initial number $\hat{Z}_0$, $\mathbb{P}(0 < \hat{Z}_K < N) < \varepsilon$ and furthermore $\mathbb{P}(\lim_{i \to \infty} \hat{Z}_i = \infty) > \mathbb{P}(\hat{Z}_K \geqslant N) - \varepsilon$. (It suffices to consider a finite number of $\hat{Z}_0$, since, if $\hat{Z}_0$ is large, then $\mathbb{P}(\inf_i \hat{Z}_i < N) < \varepsilon$.) From now on, we let $\hat{Z}_0 = W_0$ be the initial number of red infective half-edges.

Let us now return to the epidemic, and consider only the K first half-edges that pair off. Let us colour the half-edges as in Section 4.3.1. Since we consider only a fixed number of steps, each half-edge that pairs off infects a susceptible vertex of degree k with probability, cf. (4.30),

$$\frac{k(n_{S,k} + O(1))}{\sum_{k=0}^\infty kn_k + o(n)} = \frac{\alpha_S k p_k}{\mu} + o(1). \tag{5.7}$$

Hence we can w.h.p. couple the first K steps of the epidemic and the random walk $\hat{Z}_i$ such that $Z_{t_i} = \hat{Z}_i$, for $i \leqslant K$, where, as above Z_{t_i} is the number of red infective half-edges when i half-edges have paired off.

If we have at least N red infective half-edges after K steps, let us restart the epidemic there, and regard the situation after K steps as a new initial configuration, omitting the $2K$ half-edges that have been paired. Since only $o(n)$ half-edges have changed status, the

28

assumptions (D1)–(D6) still hold; however, we now start with at least N infective half-edges. Thus, by our choice of N, $\mathbb{P}(T_0 < \infty) > 1 - \varepsilon$.

Combining these facts, we see that, with probability $1 - O(\varepsilon) + o(1)$, either:

(i) the Galton–Watson process (W_i) dies out, $\hat{Z}_K = 0$, the epidemic infects at most K vertices, and $T_0 = \infty$; or

(ii) the Galton–Watson process (W_i) survives, $\hat{Z}_K \geqslant N$, there are at least N free infective half-edges after K steps, and $T_0 < \infty$.

Since ε is arbitrary, (5.5) follows, using the comment above that the extinction probability of (W_i) is the product in (5.5), and so does (5.6).

This proves the theorems in the supercritical case. In the (sub)critical case $\mathfrak{R}_0 \leqslant 1$, we only have to prove Theorem 5.4(i). This follows by comparison with a random walk as in the supercritical case; the details are simpler and are omitted. $\qquad\square$

Remark 5.6. By standard branching process theory, the extinction probability q is the unique root in $(0, 1)$ of

$$q = \mathbb{E}\, q^\xi = \frac{\mu_{\mathrm{R}}}{\mu} + \frac{\alpha_{\mathrm{S}}}{\mu} \sum_{k=1}^{\infty} k p_k g_{k-1}(q), \tag{5.8}$$

where g_{k-1} is given by (5.2). Hence q and the (asymptotic) probability $1 - \prod_{k=1}^{\infty} g_k(q)^{n_{\mathrm{I},k}}$ of a large outbreak can be determined to arbitrary precision given the parameters of the model.

Remark 5.7. We have in Theorem 5.3 used a single-type Galton–Watson process, for simplicity. It is perhaps more natural to use a multi-type Galton–Watson process, with types $0, 1, 2, \ldots$, where an individual of type k has a number of children distributed as Y_k and each child is randomly assigned type ℓ with probability $\alpha_{\mathrm{S}}(\ell + 1)p_{\ell+1}/\mu + \delta_{\ell,0}\mu_{\mathrm{R}}/\mu$, $\ell \geqslant 0$; we start with $n_{\mathrm{I},k}$ individuals of type k for each $k \geqslant 0$. The total number of individuals in each generation, ignoring their types, will form the single-type Galton–Watson process above, with offspring distribution ξ (starting as explained above). Hence the two different branching processes have the same extinction probability, which is the product in (5.5).

6. Transfer to the simple random graph: proofs of Theorems 2.7 and 2.10

In this section we consider the simple random graph G. We assume (G1), in addition to (D1)–(D6).

All results in Theorems 2.6 and 2.9 about convergence in probability, or results holding w.h.p., immediately transfer from the random multigraph G^* to the simple graph G by conditioning on G^* being simple, since $\liminf_{n \to \infty} \mathbb{P}(G^* \text{ is simple}) > 0$ by assumption (G1) and [21].

It remains to show Theorem 2.10(ii)(c), and the more precise Theorem 5.3, for G. Again, the case $X_{\mathrm{I},0} \to \infty$, when $\mathbb{P}(T_0 < \infty) \to 1$, is clear. By considering subsequences, it thus suffices to consider the case $X_{\mathrm{I},0} = O(1)$, i.e. a bounded number of initially infected half-edges. Our assumptions allow a number $n_{\mathrm{I},0}$ of isolated infected vertices, but they do not affect the edges in the graph or the infections at all, so we may simply delete them and assume $n_{\mathrm{I},0} = 0$. (Since we assume $n_{\mathrm{I}}/n \to \alpha_{\mathrm{I}} = 0$, the number of isolated infected vertices $o(n)$ and their deletion will not affect the assumtions.) We may thus assume that there is a bounded number of initially infected vertices, with uniformly bounded degrees. By again considering subsequences, we may assume that the numbers $n_{\mathrm{I},k}$ of initially infected vertices

are constant (i.e., do not depend on n), with only a finite number different from 0. This assumption implies that the number $\varkappa := \prod_{k=1}^{\infty} g_k(q)^{n_{I,k}}$ in Theorem 5.3 is constant, and the conclusion (5.5) may be written

$$\mathbb{P}(T_0 = \infty) \to \varkappa. \tag{6.1}$$

We have shown this for the random multigraph G^*, and we want to show that it holds also if we condition on G^* being simple, i.e., that the events $\{T_0 = \infty\}$ and $\{G^* \text{ is simple}\}$ are asymptotically independent.

We thus continue to work with G^* and the configuration model. Let W be the number of loops and pairs of parallel edges in G^*; thus G^* is simple if and only if $W = 0$. We write $W = \sum_{\alpha \in \mathcal{A}} I_\alpha$, where the index set $\mathcal{A} = \mathcal{A}' \cup \mathcal{A}''$ with $\mathcal{A}'$ the set of all pairs of two half-edges at the same vertex and $\mathcal{A}''$ the set of all pairs of two pairs $\{a_i, a_j\}$ and $\{b_i, b_j\}$ with a_i and b_i distinct half-edges at some vertex i and a_j and b_j distinct half-edges at some other vertex j; I_α is the indicator that these half-edges form a loop or a pair of parallel edges, respectively.

We let $\mathcal{L}$ be the event that at most $\log n$ vertices will be infected. By the definition of T_0, $\mathcal{L}$ implies $T_0 = \infty$ (for large n), and by Theorem 5.4(ii), $\mathbb{P}(T_0 = \infty \text{ and not } \mathcal{L}) = o(1)$. Hence $\mathcal{L}$ and $\{T_0 = \infty\}$ coincide w.h.p., and it suffices to show that

$$\mathbb{P}(\mathcal{L} \mid W = 0) \to \varkappa. \tag{6.2}$$

We do this by inverting the conditioning and using the method of moments, in the same way as in the proof of a similar result in [25]. (See in particular the general formulation in [25, Proposition 7.1]. However, we prefer to give a self-contained proof here.)

First, since we already know $\mathbb{P}(\mathcal{L}) \to \varkappa > 0$ and $\liminf \mathbb{P}(W = 0) > 0$, (6.2) is equivalent to

$$\mathbb{P}(\mathcal{L} \text{ and } W = 0) = \mathbb{P}(\mathcal{L}) \, \mathbb{P}(W = 0) + o(1) \tag{6.3}$$

and thus to

$$\mathbb{P}(W = 0 \mid \mathcal{L}) = \mathbb{P}(W = 0) + o(1). \tag{6.4}$$

By again considering a subsequence, we may assume that $W \overset{d}{\longrightarrow} \widehat{W}$ for some random variable $\widehat{W}$, with convergence of all moments, where furthermore the distribution of $\widehat{W}$ is determined by its moments (at least among non-negative distributions), see [21] if the maximum degree $\max d_i = o(n^{1/2})$ and [22] in general. ($\widehat{W}$ has a Poisson distribution if $\max d_i = o(n^{1/2})$, but not in general, see [22].)

We write $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$, where $\mathcal{A}_2$ is the set of all $\alpha \in \mathcal{A}$ that include a half-edge at an initially infected vertex, and $\mathcal{A}_1$ are the others. Correspondingly, we have $W = W_1 + W_2$, where $W_1 := \sum_{\mathcal{A}_1} I_\alpha$ and $W_2 := \sum_{\mathcal{A}_2} I_\alpha$.

Since the number of initially infected half-edges is $O(1)$, we have, using (2.10) and denoting the total number of half-edges by $N := \sum_k k n_k \sim \mu n$,

$$\mathbb{E}\, W_2 \leqslant \frac{O(1)}{N-1} + \frac{O(1) \sum_i d_i^2}{(N-1)(N-3)}, \tag{6.5}$$

where the first term on the right-hand side is an upper bound on the number of loops at infective vertices, and the second term is an upper bound on the number of pairs of parallel edges coming out of infective vertices. (The probability that a particular half-edge at an infective vertex joins to vertex i is $d_i/(N-1)$, and the probability that another half-edge at the same infective vertex joins to vertex i is $(d_i - 1)/(N-3)$.) Thus $W_2 \overset{P}{\longrightarrow} 0$, and

$W_1 = W - W_2 \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$. For each $m \geqslant 1$, $\mathbb{E}\, W_1^m \leqslant \mathbb{E}\, W^m \to \mathbb{E}\, \widehat{W}^m$, so each moment of W_1 is bounded as $n \to \infty$, and thus the convergence $W_1 \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$ implies moment convergence

$$\mathbb{E}\, W_1^m \to \mathbb{E}\, \widehat{W}^m \tag{6.6}$$

for every $m \geqslant 1$.

We next consider the conditioned variable $(W_1 \mid \mathcal{L})$. We want to show that $(W_1 \mid \mathcal{L}) \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$ by the method of moments, so we fix $m \geqslant 1$ and write

$$\mathbb{E}(W_1^m \mid \mathcal{L}) = \sum_{\alpha_1,\ldots,\alpha_m \in \mathcal{A}_1} \mathbb{E}(I_{\alpha_1} \cdots I_{\alpha_m} \mid \mathcal{L}). \tag{6.7}$$

Consider some $\alpha_1, \ldots, \alpha_m \in \mathcal{A}_1$. If we condition on $I_{\alpha_1} \cdots I_{\alpha_m} = 1$, we have fixed the pairing of $O(1)$ half-edges, but the remaining half-edges are paired uniformly, so if we remove the edges given by $\alpha_1, \ldots, \alpha_m$, we have another instance of the configuration model, say $\bar{G}^*$. Obviously, our assumptions (D1)–(D6) and (G1) hold for $\bar{G}^*$ too. Furthermore, the probability that the infection will spread to any vertex involved in $\alpha_1, \ldots, \alpha_m$ before $\log n$ vertices have been infected is $O\big(\log n \max_i d_i / n\big) = o(1)$, since $\max_i d_i = O(n^{1/2})$ by (G1). Hence, w.h.p., the extra edges given by $\alpha_1, \ldots, \alpha_m$ will not affect whether $\mathcal{L}$ occurs or not. Consequently, Theorem 5.3 applied to $\bar{G}^*$ shows that

$$\mathbb{P}(\mathcal{L} \mid I_{\alpha_1} \cdots I_{\alpha_m} = 1) = \mathbb{P}(\mathcal{L} \text{ in } \bar{G}^*) + o(1) = \varkappa + o(1), \tag{6.8}$$

uniformly in all $\alpha_1, \ldots, \alpha_m \in \mathcal{A}_1$ (for a fixed m). (The estimate (6.8) fails if some $\alpha_i \in \mathcal{A}_2$, for example if α_1 is a loop at an initially infected vertex. This is the reason for considering W_2 separately.) We invert the conditioning again and obtain, uniformly in all $\alpha_1, \ldots, \alpha_m \in \mathcal{A}_1$,

$$\begin{aligned}
\mathbb{E}(I_{\alpha_1} \cdots I_{\alpha_m} \mid \mathcal{L}) &= \frac{\mathbb{P}(I_{\alpha_1} \cdots I_{\alpha_m} = 1 \text{ and } \mathcal{L})}{\mathbb{P}(\mathcal{L})} \\
&= \frac{(\varkappa + o(1))\, \mathbb{P}(I_{\alpha_1} \cdots I_{\alpha_m} = 1)}{\varkappa + o(1)} \\
&= \big(1 + o(1)\big)\, \mathbb{E}(I_{\alpha_1} \cdots I_{\alpha_m}).
\end{aligned} \tag{6.9}$$

Summing over all $\alpha_1, \ldots, \alpha_m \in \mathcal{A}_1$ yields, using (6.6),

$$\mathbb{E}(W_1^m \mid \mathcal{L}) = \big(1 + o(1)\big)\, \mathbb{E}(W_1^m) \to \mathbb{E}\, \widehat{W}^m. \tag{6.10}$$

This holds for each $m \geqslant 1$, and thus $(W_1 \mid \mathcal{L}) \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$ by the method of moments. Furthermore, $W_2 \overset{\mathrm{p}}{\longrightarrow} 0$, and thus $(W_2 \mid \mathcal{L}) \overset{\mathrm{p}}{\longrightarrow} 0$. Consequently, $(W \mid \mathcal{L}) = (W_1 + W_2 \mid \mathcal{L}) \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$. In particular, $\mathbb{P}(W = 0 \mid \mathcal{L}) \to \mathbb{P}(\widehat{W} = 0)$. Since also $W \overset{\mathrm{d}}{\longrightarrow} \widehat{W}$ and thus $\mathbb{P}(W = 0) \to \mathbb{P}(\widehat{W} = 0)$, the equation (6.4) follows, which, as stated above, implies (6.2) and completes the proof.

7. No second moment assumption

The main reason for making the assumption (G1) on a bounded second moment of the degree distribution is that it allows us to study the epidemic on the configuration model multigraph instead of the uniform simple graph, see Section 6. However, we have used (G1) also in some parts of the proofs for the multigraph case. We will now show that this is not necessary, so that, in the multigraph case, the only assumption on the degree distribution

required is the uniform summability of $\sum_k k n_{\mathrm{S},k}/n$ inherent in assumptions (D1)–(D3), see Remark 2.2. The uniform summability assumption seems indispensible.

In the proof of Theorem 2.6(b), using an identical argument but without assuming (G1), h_I and p_I are Lipschitz continuous on $[\theta_\infty, \theta_1]$ for any $\theta_1 < 1$, so that p_I is locally Lipschitz continuous on $[\theta_\infty, 1)$. Hence, if we choose $b \in (\theta_\infty, 1)$, then there is a unique solution $\tilde{\theta}_t$, $t \geqslant 0$, to the differential equation with initial condition $\tilde{\theta}_0 = b$, and we may extend $\tilde{\theta}_t$ to a solution $\tilde{\theta}_t : (a, \infty) \to (\theta_\infty, 1)$ for some $a < 0$, where, by choosing $|a|$ maximal, $\tilde{\theta}_t \to 1$ as $t \searrow a$. Note that $a > -\infty$, since $\mathrm{p}_\mathrm{I}(1) > 0$ and p_I is continuous. The translate $\theta_t := \tilde{\theta}_{t+a}$ then is a solution to (2.19). Uniqueness is proved similarly: if θ_t and $\bar{\theta}_t$ are two solutions to (2.19), then we may translate them so that the translates θ_{t-a} and $\bar{\theta}_{t-\bar{a}}$ (with $a, \bar{a} < 0$) have the same value b at 0, with $b \in (\theta_\infty, 1)$. By the uniqueness, as long as $\theta \in [\theta_\infty, 1)$, the two translates θ_{t-a} and $\bar{\theta}_{t-\bar{a}}$ coincide on the set where they are both less than 1, and, by continuity, they coincide completely and thus $a = \bar{a}$ and $\theta_t = \bar{\theta}_t$. The rest of the proof is the same.

In the proof of Theorem 2.9(ii)(b), we note first that, as in the proof of Theorem 2.6(b) just given, p_I is locally Lipschitz continuous on $[\theta_\infty, 1)$, which implies uniqueness of the solution to (2.25). Furthermore, as before, we define $\hat{A}$ by (4.20) and note that

$$\lim_{\tau \to 0} \hat{A}_\tau = \hat{A}_0 := -\int_0^{\hat{\tau}_0} \frac{d\sigma}{\beta \mathrm{p}_\mathrm{I}(e^{-\sigma})}. \tag{7.1}$$

If $\hat{A}_0 = -\infty$, everything is as before. However, if we do not assume (G1), it is possible that the integral converges and thus $\hat{A}_0 > -\infty$. In this case, the inverse $\hat{\tau} : [\hat{A}_0, \infty) \to [0, \hat{\tau}_\infty)$ of $\hat{A}$ is continuously differentiable on $(\hat{A}_0, \infty)$, and, if we extend it by defining $\hat{\tau}(t) = 0$ for $t < \hat{A}_0$, then it is continuous on $\mathbb{R}$ and satisfies (4.21) for all $t \neq \hat{A}_0$. (Recall that $\mathrm{p}_\mathrm{I}(1) = 0$.) Since the right-hand side of (4.21) is continuous on $\mathbb{R}$, it follows that $\hat{\tau}(t)$ is continuously differentiable everywhere and that (4.21) is satisfied also at $\hat{A}_0$. Thus, again, $\theta_t = \exp(-\hat{\tau}(t))$ satisfies (2.25), although now $\theta_t = 1$ for $t \leqslant \hat{A}_0$. (In this case, there is a unique solution to (2.25) for $\theta_0 < 1$ but infinitely many solutions for $\theta_0 = 1$.) The rest of the proof is the same.

The proof of Theorem 2.9(ii)(d) remains the same if $\hat{A}_0 = -\infty$. If $\hat{A}_0 > -\infty$, then the interval $[-t_1, t_1]$ should be replaced by $[t_2, t_1]$ with $-\hat{A}_0 < t_2 < t_1 < \infty$. The convergence (4.25) then follows as before, uniformly on each such $[t_2, t_1]$. Furthermore, the same monotonicity argument as before, now using $\hat{\tau}(t_2) \to 0$ as $t_2 \to \hat{A}_0$, shows that (4.25) holds uniformly on $\mathbb{R}$. The rest is the same as before.

Let us confirm it is indeed possible that $\hat{A}_0 > -\infty$. By (7.1) and (2.18),

$$\hat{A}_0 > -\infty \iff \int_0^{\hat{\tau}_0} \frac{d\sigma}{\mathrm{p}_\mathrm{I}(e^{-\sigma})} < \infty \iff \int_0^{\hat{\tau}_0} \frac{d\sigma}{h_\mathrm{I}(e^{-\sigma})} < \infty \iff \int_0^{x_0} \frac{dx}{h_\mathrm{I}(1-x)} < \infty, \tag{7.2}$$

for any small positive x_0 ($x_0 < 1 - \theta_\infty$). Recall that $h_\mathrm{I}(1) = \mu_\mathrm{I} = 0$. Thus, (2.13)–(2.15) imply that, as $x \to 0$, $h_\mathrm{I}(1-x) = h_\mathrm{S}(1) - h_\mathrm{S}(1-x) + O(x)$. It follows easily from (7.2) that

$$\hat{A}_0 > -\infty \iff \int_0^{x_0} \frac{dx}{h_\mathrm{S}(1) - h_\mathrm{S}(1-x)} < \infty. \tag{7.3}$$

Example 7.1. Suppose that $p_k \sim k^{-\alpha-1}$ as $k \to \infty$, with $1 < \alpha \leqslant 2$. (This means that the asymptotic degree distribution has moments of order $< \alpha$, but not of order α.) Then (2.12)

32

implies that if $1 < \alpha < 2$, then $h'_{\mathrm{S}}(1-x) \asymp x^{\alpha-2}$ as $x \to 0$ and thus $h_{\mathrm{S}}(1) - h_{\mathrm{S}}(1-x) \asymp x^{\alpha-1}$ so the integral in (7.3) converges and $\hat{A}_0 > -\infty$. However, if $\alpha = 2$, then $h'_{\mathrm{S}}(1-x) \asymp |\log x|$, so $h_{\mathrm{S}}(1) - h_{\mathrm{S}}(1-x) \asymp x|\log x|$ and $\hat{A}_0 = -\infty$.

We end this section by noting that, when there is no second moment, then $\mathfrak{R}_0$ is infinite (and thus the process is supercritical, for any β and ρ). On the other hand, if (G1) holds, then $\mathfrak{R}_0$ is finite, cf. Remark 2.4.

8. The random time shift

In Theorems 2.9(ii) and 2.10(ii), we use a (random) time-shift T_0, which can be interpreted as the time it takes for the epidemic to grow from a small number of initially infected to a large outbreak.

Suppose for simplicity that $\sum_{k=1}^{\infty} k n_{\mathrm{I},k} \to \infty$, so $T_0 < \infty$ w.h.p. by Theorem 2.9(ii) for G^*, and by Theorem 2.10(ii) for G. Then, for any $\delta > 0$, (4.24) shows that

$$T_0 \geqslant T_0 - A_\delta \xrightarrow{\mathrm{P}} -\hat{A}_\delta. \tag{8.1}$$

Suppose first that (G1) holds. Letting $\delta \to 0$, we have $-\hat{A}_\delta \to \infty$, as noted after (4.20); thus $T_0 \xrightarrow{\mathrm{P}} \infty$.

On the other hand, if we assume that (G1) does not hold (and we thus consider the multigraph case), the situation is more complicated. It is possible that $\hat{A}_0 = -\infty$, as seen in Example 7.1, and then $T_0 \xrightarrow{\mathrm{P}} \infty$ as above. However, if $\hat{A}_0 > -\infty$, then we obtain only $T_0 \geqslant |\hat{A}_0| - \varepsilon$ w.h.p., for every $\varepsilon > 0$. It is possible to give examples showing that in this case, both $T_0 \xrightarrow{\mathrm{P}} \infty$ and $T_0 \xrightarrow{\mathrm{P}} |\hat{A}_0| < \infty$ are possible (as well as intermediate cases). We omit the details.

Appendix A. Time-changed Markov chains

The following lemma justifies the time change in Section 3. A more general result appears in [39, III. (21.7)] but we include a proof in the simpler setting of Markov chains.

Lemma A.1. *Suppose $(Y_\tau)_{\tau \geqslant 0}$ is a continuous time Markov chain with finite state space E, and infinitesimal transition rates $(q(i,j))_{i,j \in E}$.*

Let $f : E \to (0, \infty)$ be strictly positive and define the strictly increasing process

$$A_\tau = \int_0^\tau f(Y_\sigma)\, d\sigma, \qquad \tau \geqslant 0, \tag{A.1}$$

and its inverse $\tau(t)$, $t \geqslant 0$.

Then the time-changed process $(Y_{\tau(t)})_{t \geqslant 0}$ is again Markov and has infinitesimal rates $(q(i,j)/f(i))_{i,j \in E}$.

Proof. The paths of Y are piecewise constant and right continuous. Let $J_0 = 0$ and $J_{k+1} = \inf\{\tau \geqslant J_k : Y_\tau \neq Y_{J_k}\}$, $k \geqslant 0$, be the jump times of $(Y_\tau)_{\tau \geqslant 0}$. The state space E is finite, so the process is non-explosive and a.s. $J_k \to \infty$. Also, without loss of generality we may assume no state is absorbing, i.e. the rate of leaving $q(i) := -q(i,i) = \sum_{j \neq i} q(i,j) > 0$ is strictly positive for each $i \in E$. Thus $J_k < \infty$ a.s. for each k.

For notational ease, let $\tilde{Y}_t := Y_{\tau(t)}$, $t \geqslant 0$. Since $\tau(A_{J_k}) = J_k$, the jump times of $\tilde{Y}$ are given by $\tilde{J}_k := A_{J_k}$, $k \geqslant 0$. It follows that the holding times for $\tilde{Y}$ are

$$
\begin{aligned}
\tilde{J}_{k+1} - \tilde{J}_k &= A_{J_{k+1}} - A_{J_k} \\
&= (J_{k+1} - J_k) f(Y_{J_k}) \\
&= ((J_{k+1} - J_k) q(Y_{J_k})) \left(\frac{f(Y_{J_k})}{q(Y_{J_k})} \right) \\
&= T_k \, \tilde{q}(\tilde{Y}_{\tilde{J}_k})^{-1}
\end{aligned}
$$

$k \geqslant 0$, where $T_k = (J_{k+1} - J_k) q(Y_{J_k})$ and $\tilde{q}(i) = q(i)/f(i)$, $i \in E$. The Markov property of $(Y_\tau)_{\tau \geqslant 0}$ implies the T_k are independent $\mathrm{Exp}(1)$ random variables that are also independent of the embedded jump chain $(Y_{J_k})_{k \geqslant 0} = (\tilde{Y}_{\tilde{J}_k})_{k \geqslant 0}$. The latter has transition kernel $(q(i,j)/q(i))_{i,j \in E} = (\tilde{q}(i,j)/\tilde{q}(i))_{i,j \in E}$, where $\tilde{q}(i,j) = q(i,j)/f(i)$, $i,j \in E$. It follows that $\tilde{Y}_t$ is a Markov chain with transition rates $\tilde{q}(i,j)$. $\qquad \square$

Appendix B. Summary of notation

For ease of reference we summarise the main notation used to describe the epidemic. A subscript of S, I or R always refers to susceptible, infective or recovered, respectively. The initial conditions are as follows.

$n_{\mathrm{S}}, n_{\mathrm{I}}, n_{\mathrm{R}}$	Number of vertices (of given type).
$n_{\mathrm{S},k}, n_{\mathrm{I},k}, n_{\mathrm{R},k}$	Number of degree $k \geqslant 0$ vertices.
$\alpha_{\mathrm{S}}, \alpha_{\mathrm{I}}, \alpha_{\mathrm{R}}$	Limiting fractions of vertices.
$(p_k)_0^\infty, \lambda$	Limiting degree distribution for a randomly chosen susceptible and its mean.
$\mu = \mu_{\mathrm{S}} + \mu_{\mathrm{I}} + \mu_{\mathrm{R}}$	Limiting mean degree (for any vertex).

The stochastic processes are denoted as follows.

S_t, I_t, R_t	Number of vertices (of given type) at time $t \geqslant 0$.
$S_t(k), I_t(k), R_t(k)$	Number of degree $k \geqslant 0$ vertices.
$X_{\mathrm{S},t}, X_{\mathrm{I},t}, X_{\mathrm{R},t}$	Number of free half-edges.
$X_t = X_{\mathrm{S},t} + X_{\mathrm{I},t} + X_{\mathrm{R},t}$	Total number of half-edges.

Time-changed versions of the above processes are superscripted with a prime and use greek time indices. In addition we have $\tau(t)$ and A_τ, the time change and its inverse. The limiting trajectories for these processes are denoted as follows.

$\hat{\tau}(t), \hat{A}_\tau$	Time change and its inverse.
$\theta_t = e^{-\hat{\tau}(t)}$	Parameterisation of time.
$v_{\mathrm{S}}(\theta_t), \hat{I}_t, \hat{R}_t$	Fraction of vertices (of given type) at time $t \geqslant 0$.
$h_{\mathrm{S}}(\theta_t), h_{\mathrm{I}}(\theta_t), h_{\mathrm{R}}(\theta_t), h_X(\theta_t)$	Number of free half-edges divided by n.
$\mathrm{p_S}(\theta_t) = h_{\mathrm{S}}(\theta_t)/h_X(\theta_t),$	Susceptible proportion of free half-edges.
$\mathrm{p_I}(\theta_t) = h_{\mathrm{I}}(\theta_t)/h_X(\theta_t)$	Infective proportion of free half-edges.

The following quantities are also important in our analysis.

$\theta_\infty = \lim_{t\to\infty} \theta_t$	Root of h_I corresponding to the end of the epidemic.
τ^*	Duration of the time-changed epidemic.
$\hat\tau_\infty = -\ln\theta_\infty$	Limiting duration of the time-changed epidemic, assuming a large outbreak.
$s_0 \in (v_\mathrm{S}(\theta_\infty), v_\mathrm{S}(1))$	Fraction of susceptibles used in calibration time T_0.

References

[1] H. Andersson. Limit theorems for a random graph epidemic model. *Ann. Appl. Probab.*, 8(4):1331–1349, 1998.

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

[3] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach.* EMS, Zürich, 2003.

[4] F. Ball and P. Neal. Network epidemic models with two levels of mixing. *Math. Biosci.*, 212(1):69–87, 2008.

[5] F. Ball, D. Sirl, and P. Trapman. Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math. Biosci.*, 224(2):53–73, 2010.

[6] A. D. Barbour and G. Reinert. Approximating the epidemic curve. *Electron. J. Probab.*, 18:no. 54, 2013.

[7] T. Bohman and A. Frieze. Karp–Sipser on random graphs with a fixed degree sequence. *Combin. Probab. Comput.*, 20(5):721–741, 2011.

[8] T. Bohman and M. Picollelli. SIR epidemics on random graphs with a fixed degree sequence. *Random Struct. Alg.*, 41(2):179–214, 2012.

[9] B. Bollobás. *Random graphs.* Cambridge Univ. Press, Cambridge, second edition, 2001.

[10] B. Bollobás and O. Riordan. An old approach to the giant component problem. *arXiv:1209.3691*, September 2012.

[11] T. Britton, M. Deijfen, A. N. Lagerås, and M. Lindholm. Epidemics on random graphs with tunable clustering. *J. Appl. Probab.*, 45(3):743–756, 2008.

[12] T. Britton, S. Janson, and A. Martin-Löf. Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv. Appl. Probab.*, 39(4):922–948, 2007.

[13] S. Chatterjee and R. Durrett. Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability*, pages 2332–2356, 2009.

[14] D. J. Daley and J. Gani. *Epidemic modelling: an introduction.* Cambridge Univ. Press, Cambridge, 1999.

[15] R. W. R. Darling and J. R. Norris. Differential equation approximations for Markov chains. *Probab. Surv.*, 5:37–79, 2008.

[16] L. Decreusefond, J.-S. Dhersin, P. Moyal, and V. C. Tran. Large graph limit for an SIR process in random network with heterogeneous connectivity. *Ann. Appl. Probab.*, 22(2):541–575, 2012.

[17] S. N. Ethier and T. G. Kurtz. *Markov processes: Characterization and convergence.* Wiley, New York, 1986.

[18] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Struct. Alg.*, 31(4):482–516, 2007.

[19] A. Frieze, M. Krivelevich, and C. Smyth. On the chromatic number of random graphs with a fixed degree sequence. *Combin. Probab. Comput.*, 16(5):733–746, 2007.

[20] T. House. Modelling epidemics on networks. *Contemp. Phys.*, 53(3):213–225, 2012.

[21] S. Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18(1-2):205–225, 2009.

[22] S. Janson. The probability that a random multigraph is simple, II. *J. Appl. Probab.*, to appear.

[23] S. Janson, M. Luczak, and P. Windridge. The critical SIR epidemic on a random graph with given degrees. *(In preparation)*.

[24] S. Janson and M. J. Luczak. A simple solution to the *k*-core problem. *Random Struct. Alg.*, 30(1-2):50–62, 2007.

[25] S. Janson and M. J. Luczak. Asymptotic normality of the *k*-core in random graphs. *Ann. Appl. Probab.*, 18(3):1085–1137, 2008.

[26] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Struct. Alg.*, 34(2):197–216, 2009.

[27] A. Joseph. The component sizes of a critical random graph with given degree sequence. *arXiv:1012.2352*, December 2010.

[28] O. Kallenberg. *Foundations of modern probability*. Springer-Verlag, New York, second edition, 2002.

[29] M. Kang and T. G. Seierstad. The critical phase for random graphs with a given degree sequence. *Combin. Probab. Comput.*, 17(1):67–86, 2008.

[30] J. C. Miller. A note on a paper by Erik Volz: SIR dynamics in random networks. *J. Math. Biol.*, 62(3):349–358, 2011.

[31] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Struct. Alg.*, 6(2-3):161–179, 1995.

[32] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, 7(3):295–305, 1998.

[33] P. Neal. SIR epidemics on a Bernoulli random graph. *J. Appl. Probab.*, 40(3):779–782, 2003.

[34] M. E. J. Newman. Spread of epidemic disease on networks. *Phys. Rev. E*, 66(1):016128, 11, 2002.

[35] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.

[36] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST handbook of mathematical functions*. Cambridge Univ. Press, Cambridge, 2010. Also available as *NIST Digital Library of Mathematical Functions*. `http://dlmf.nist.gov/`.

[37] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[38] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[39] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1. Foundations.* Cambridge Univ. Press, Cambridge, 2000. Reprint of the second (1994) edition.

[40] C. M. Sato. On the robustness of random k-cores. *European J. Combin.*, 41:163–182, 2014.

[41] E. Volz. SIR dynamics in random networks with heterogeneous connectivity. *J. Math. Biol.*, 56(3):293–310, 2008.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se
URL: http://www2.math.uu.se/~svante/

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.
E-mail address: m.luczak@qmul.ac.uk
URL: http://www.maths.qmul.ac.uk/~luczak/

School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.
E-mail address: p.windridge@qmul.ac.uk
URL: http://www.maths.qmul.ac.uk/~pwindridge/