

Electron. J. Probab. **27** (2022), article no. 110, 1–66.
 ISSN: 1083-6489 <https://doi.org/10.1214/22-EJP828>

Susceptible–infected epidemics on evolving graphs*

Rick Durrett[†]

Dong Yao[‡]

Abstract

The evoSIR model is a modification of the usual SIR process on a graph G in which $S - I$ connections are broken at rate ρ and the S connects to a randomly chosen vertex. The evoSI model is the same as evoSIR but recovery is impossible. In [14] the critical value for evoSIR was computed and simulations showed that when G is an Erdős-Rényi graph with mean degree 5, the system has a discontinuous phase transition, i.e., as the infection rate λ decreases to λ_c , the fraction of individuals infected during the epidemic does not converge to 0. In this paper we study evoSI dynamics on graphs generated by the configuration model. We show that there is a quantity Δ determined by the first three moments of the degree distribution, so that the phase transition is discontinuous if $\Delta > 0$ and continuous if $\Delta < 0$.

Keywords: susceptible–infected model; configuration model; phase transition.

MSC2020 subject classifications: 60J27.

Submitted to EJP on March 17, 2020, final version accepted on July 9, 2022.

1 Introduction

In the SIR model, individuals are in one of three states: S = susceptible, I = infected, R = removed (cannot be infected). Often this epidemic takes place in a homogeneously mixing population. However, here, we have a graph G that gives the social structure of the population; vertices represent individuals and edges connections between them. $S - I$ edges become $I - I$ at rate λ , i.e., after a time with an exponential(λ) distribution. An individual remains infected for an amount of time T which can be either deterministic or random with a pre-specified distribution. Once individuals leave the infected state, they enter the removed state. In addition, we will allow the graph to evolve: $S - I$ edges are broken at rate ρ and the susceptible individual connects to an individual chosen uniformly at random from the graph. This process is called evoSIR where ‘evo’ stands for ‘evolving’. We will also consider the simpler SI epidemic in which infecteds never recover, and its evolving version *evoSI*, as well as some other variations on this theme.

*Dong Yao is supported by NSF of Jiangsu Province (No. BK20220677).

[†]Duke University, United States of America. E-mail: rtd@math.duke.edu

[‡]Corresponding author. Jiangsu Normal University, China. E-mail: wonderspiritfall@gmail.com

We will use the term *final epidemic size* or *final size of the epidemic* to refer to the number of vertices that are eventually removed in SIR epidemics or eventually infected in SI epidemics. We say a large epidemic (or large outbreak) occurs if the epidemic infects more than ϵn individuals (n is the size of the total population) for some $\epsilon > 0$ independent of n . The critical value is the smallest infection rate such that a large outbreak occurs with probability bounded away from 0 as $n \rightarrow \infty$.

As the main contribution of this paper, we show that, if the underlying graph G is sampled from the configuration model $\text{CM}(n, D)$ (defined in Model 2 in Section 1.2), then there is an explicit quantity Δ given in (1.4.1) (determined by the first three moments of D) such that the following holds for the evoSI model on G :

- If $\Delta > 0$, then the fraction of infected vertices (conditionally on a large outbreak) doesn't converge to 0 as λ approaches the critical value λ_c . In other words, there is a discontinuous phase transition.
- If $\Delta < 0$, then the phase transition is continuous for λ near λ_c .

See Theorem 1.7 for a precise statement.

The introduction is organized into seven subsections. Sections 1.1–1.3 are devoted to a review of previous work. The main results of the paper (Theorem 1.6 and Theorem 1.7) are presented in Section 1.4. Sections 1.5–1.7 sketch the proof of Theorem 1.7.

1.1 DOMath [14]

evoSIR was studied by three Duke undergraduates in the summer and fall of 2018 (Yufeng Jiang, Remy Kassem, and Grayson York) under the direction of Matthew Junge and Rick Durrett. They considered two possibilities for the infection time: the Markovian case in which infections last for an exponential time with mean 1, and the case in which each infection lasts for exactly time 1. Here we restrict our attention to their results for the second case, which is simpler due to its connection with independent bond percolation. In any SIR model each edge will be $S - I$ (or $I - S$) only once. When that happens, in the fixed infection time case without rewiring, the infection will be transferred to the other end with probability

$$\tau^f = \mathbb{P}(T \leq 1) = 1 - e^{-\lambda}, \quad (1.1.1)$$

and the transfers for different edges are independent. Here the ' f ' in the superscript is for "fixed time." Due to the last observation, we can delete edges with probability $e^{-\lambda}$ and the connected components of the resulting graph will give the epidemic sizes when one member of the cluster is infected.

In [14] G was an Erdős-Rényi($n, \mu/n$) random graph in which there are n vertices and each pair is independently connected with probability μ/n . The following result is well-known.

Theorem 1.1. *Consider the SIR process on Erdős-Rényi($n, \mu/n$) with fixed infection time. The reduced graph after deletion of edges as described above is Erdős-Rényi($n, \mu\tau^f/n$). So, if we start with one infected and the rest of the population susceptible, a large outbreak occurs with positive probability if and only if $\mu\tau^f > 1$. If z_0 is the fixed point smaller than 1 of the generating function*

$$G_0(z) = \exp(-\mu\tau^f(1-z)), \quad (1.1.2)$$

then $1 - z_0$ gives both the limiting probability that an infected individual will start a large epidemic, and the fraction of individuals who will become infected when a large epidemic occurs.

Things become more complicated when we introduce rewiring of $S - I$ edges at rate ρ . To be able to use the ideas in the proof of Theorem 1.1, [14] introduced the delSIR model in which edges are deleted instead of rewired. In the evo (or del) version of the model, in order for the infection to be transmitted along an edge, infection must come before any rewiring (or deletion) and before time 1. To compute this probability, note that (i) the probability that infection occurs before rewiring is $\lambda/(\lambda + \rho)$ and (ii) the minimum of two independent exponentials with rates λ and ρ is an exponential with rate $\lambda + \rho$, so the transmission probability is

$$\tau_r^f = \frac{\lambda}{\lambda + \rho}(1 - e^{-(\lambda + \rho)}). \quad (1.1.3)$$

Here the ‘ r ’ subscript is for “rewire.” By a standard coupling argument one can show that evoSI dominates delSI.

Lemma 1.2. *For fixed parameters, there exists a coupling of evoSI and delSI so that there are no fewer infections in the delSI model than in evoSI. Same is true if we replace SI by SIR.*

The next result, Theorem 1 in [14], shows that evoSIR has the same critical value as delSIR, and in the subcritical case the expected cluster sizes are the same.

Theorem 1.3. *The critical value λ_c for a large epidemic in fixed infection time delSIR or evoSIR epidemic with rewiring is given by the solution of $\mu\tau_r^f(\lambda) = 1$. Moreover, if $\lambda < \lambda_c$, then the ratio of the expected epidemic size in delSIR to the size in evoSIR converges to 1 as the number of vertices goes to ∞ .*

The formula for the critical value is easily seen to be correct for the delSIR since, by the reasoning above, there is a large epidemic if and only if the reduced graph in which edges are retained with probability τ_r^f has a giant component. From Lemma 1.2 one can see that the delSIR model has a larger ($\geq$) critical value than evoSIR. Thus, one only has to prove the reverse inequality. Intuitively, the equality of the two critical values holds because a subcritical delSIR epidemic dies out quickly, so it is unlikely that rewirings will influence the outcome.

When n is large, the degree distribution, which is $\text{Binomial}(n - 1, \mu/n)$, is approximately Poisson with mean μ . Due to Poisson thinning, the number of new infections directly caused by one I in delSIR in an otherwise susceptible population is asymptotically Poisson with mean $\mu\tau_r^f$, and hence has limiting generating function of the distribution $\text{Poisson}(\mu\tau_r^f)$

$$G_1(z) = \exp(-\mu\tau_r^f(1 - z)). \quad (1.1.4)$$

The following result, Theorem 2 in [14], identifies the probability of a large outbreak.

Theorem 1.4. *If $z_0 < 1$ is the fixed point of $G_1(z)$, then $1 - z_0$ gives the probability of a large delSIR or evoSIR outbreak.*

In the case of the delSIR model, $1 - z_0$ is the fraction of individuals infected in a large epidemic. It is easy to see that this proportion goes to 0 at the critical value $\mu_c = 1/\tau_r^f = 1$. The next simulation suggests that this is not true in the case of evoSIR.

1.2 Britton et al. [3, 16]

As the authors of [14] were finishing up the writing of their paper, they learned of two papers by Britton and collaborators that study epidemics on evolving graphs with exponential infection times. [3] studies a one parameter family of models (SIR- ω) that interpolates between delSIR and evoSIR. To facilitate later referencing we attach labels to the next two descriptions.

Model 1. SIR- ω epidemic. In this model, an infected individual infects each neighbor at rate λ , and recovers at rate γ . A susceptible individual drops its connection to an infected individual at rate ω . The edge is rewired with probability α and dropped with probability $1 - \alpha$. Since evoSIR ($\alpha = 1$) and delSIR ($\alpha = 0$) have the same critical values and survival probability it follows that this holds for all $0 \leq \alpha \leq 1$ since by a coupling argument the final epidemic size of the case $0 < \alpha < 1$ can be sandwiched between delSIR and evoSIR.

These epidemics take place on graphs generated by

Model 2. Configuration model. Given a nonnegative integer n and a positive integer valued random variable D , take n i.i.d. copies $D_1, \dots, D_n$ of D . If the sum $\sum_{i=1}^n D_i$ is odd, then we replace D_n by $D_n + 1$. We then construct a graph G on n vertices as follows. We attach $D_1, \dots, D_n$ half-edges to vertices $1, 2, \dots, n$, respectively and then pair these half-edges uniformly at random to form a graph. We call this random graph *the configuration model* on n vertices with degree distribution D and denote it by $\text{CM}(n, D)$. We assume D has finite second moment so that the resulting graph will be a simple graph with nonvanishing probability as $n \rightarrow \infty$. See Theorem 3.1.2 in [7]. We refer readers to [22, Chapter 7], [24, Chapters 4 and 7] and [23, Chapter 2] for more details on the configuration model.

Remark 1.5. Throughout the paper, unless otherwise specified, we always consider the annealed probability measure with respect to the configuration model. In other words the randomness is taken over both the degrees $D_1, \dots, D_n$ and the construction of $\text{CM}(n, D)$ based on the degrees.

Britton, Juher, and Saldana [3] studied the initial phase of the epidemic starting with one infected at vertex x (chosen uniformly at random from all vertices) using a branching process approximation. Let Z_m^n be the number of vertices at distance m from x in the graph $\text{CM}(n, D)$. For any fixed $k \in \mathbb{N}$, $\{Z_m^n, 0 \leq m \leq k\}$ converges to the following two-phase branching process $\{Z_m, 0 \leq m \leq k\}$. The number of children in the first generation has the distribution D while subsequent generations have the distribution $D^* - 1$ where D^* is the *size-biased* degree distribution

$$\mathbb{P}(D^* = j) = \frac{j p_j}{m_1}, \quad j \geq 0.$$

Here $p_j = \mathbb{P}(D = j)$ and $m_1 = \mathbb{E}(D)$ is the mean of D . Later we will also use $m_i := \mathbb{E}(D^i)$ to denote the i -th moment of D for $i \geq 1$. This follows from the construction of the configuration model: x connects to other vertices with probability proportional to their degrees so individuals in generations $m \geq 1$ have the $D^* - 1$ children instead of D . The ‘-1’ is because one edge is used in making the connection from x . Before moving on to epidemics on the configuration model, we note that if

$$G(z) = \sum_{k=0}^{\infty} p_k z^k, \tag{1.2.1}$$

then the generating function of $D^* - 1$ is

$$\hat{G}(z) = \sum_{j=1}^{\infty} \frac{j p_j}{m_1} z^{j-1} = \frac{G'(z)}{G'(1)}. \tag{1.2.2}$$

In the SIR- ω model, the probability that an infection will cross an $S - I$ edge is

$$\tau = \frac{\lambda}{\lambda + \gamma + \omega}.$$

Thus we get another two-phase branching process $\bar{Z}_m$ defined as follows. $\bar{Z}_0 = 1$, $\bar{Z}_1 = \text{Binomial}(D, \tau)$ and future generations have offspring distribution $\text{Binomial}(D^* - 1, \tau)$. One can see from this description that the limiting branching process $\bar{Z}_m$ will have positive survival probability if

$$1 < \tau \mathbb{E}(D^* - 1) = \tau \left(\frac{m_2}{m_1} - 1 \right) = \frac{(m_2 - m_1)\tau}{m_1}.$$

Correspondingly, there will be a large epidemic in the SIR- ω model if

$$R_0 = \frac{\lambda}{\lambda + \gamma + \omega} \cdot \frac{m_2 - m_1}{m_1} > 1. \quad (1.2.3)$$

This follows from results on percolation in random graphs (see [9] and [12]). Thus for fixed values of γ and ω

$$\lambda_c = (\gamma + \omega) \frac{m_1}{m_2 - 2m_1}. \quad (1.2.4)$$

When $\gamma = 0$ and $\omega = \rho$ which is the SI- ω model in our notation,

$$\lambda_c = \frac{\rho m_1}{m_2 - 2m_1}. \quad (1.2.5)$$

The critical values of delSI and evoSI only depends on the ratio ρ/λ , so it is natural to define a parameter (this α is different from the α used in the definition of SIR- ω model)

$$\alpha = \rho m_1 / \lambda \quad (1.2.6)$$

that has $\alpha_c = m_2 - 2m_1$. We will only consider the del and evo endpoints of the one parameter family of models SIR- ω , so after the discussion of previous work is completed, there should be no confusion between our α and theirs.

Work of Leung, Ball, Sirl, and Britton [16] demonstrated the paradoxical fact that individual preventative measures may lead to a larger final size of the epidemic. They proved this rigorously for SI epidemics on the configuration model with two degrees and conducted simulation studies for many social networks. Note that in Figure 2 (taken from Figure 1 of [16]) the final size increases with the rewiring rate when the rewiring rate is small. This simulation does not suggest that the phase transition was discontinuous.

On the other hand, Figure 3 from [14] gives a similar simulation that clearly shows the discontinuity.

1.3 Ball and Britton [2]

Ball and Britton [2] analyzed the evoSIR and evoSI epidemics on Erdős-Rényi graphs. Their construction uses properties that are special to that case. See [2, Section 2.3] for details and also [4, 21] for earlier examples of the use of this construction. They solved the case of the SI- ω model on Erdős-Rényi graphs completely, but for the SIR- ω model there is a gap between the necessary and sufficient condition for a discontinuous phase transition. See (1.3.3) and the comment after it.

To prove results about the epidemics on an Erdős-Rényi random graph with mean degree μ they first consider a tree in which each vertex has a $\text{Poisson}(\mu)$ number of descendants and develop a branching process approximation for the SIR- ω epidemic in which infections (births) cross an edge with probability $\lambda/(\lambda + \gamma + \omega)$. Let $I(t)$ be the total number of infected individuals, $I_E(t)$ be the number of infectious edges, and $T(t)$ be the total progeny in the branching process on the tree. Let $I^n(t)$, $I_E^n(t)$ and $T^n(t)$ be the corresponding quantities for an Erdős-Rényi($n, \mu/n$) random graph on n vertices where

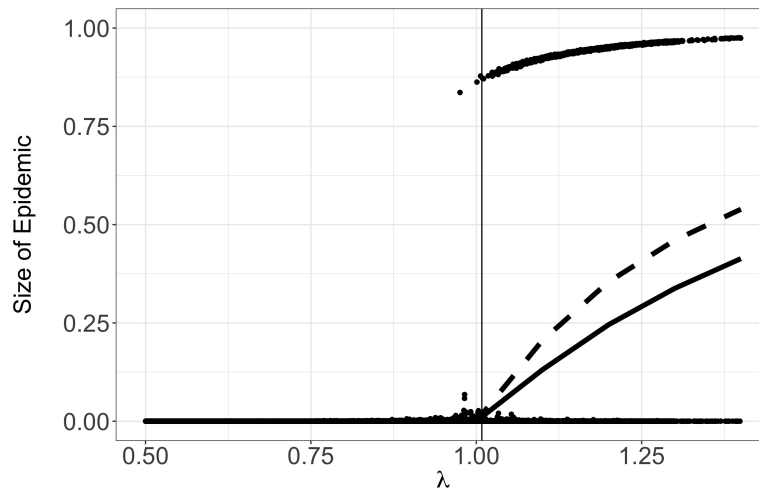


Figure 1: Simulation of the fixed time evoSIR on an Erdős-Rényi graph with $\mu = 5$, $\rho = 4$ and λ varying. $\lambda_c \approx 1.0084$ in agreement with Theorem 1.3. The bottom curve is the final size of the delSIR epidemic with the same parameters. The dashed line above it is an approximation derived in [14] that turned out to be inaccurate. The top curve comes from simulating evoSIR.

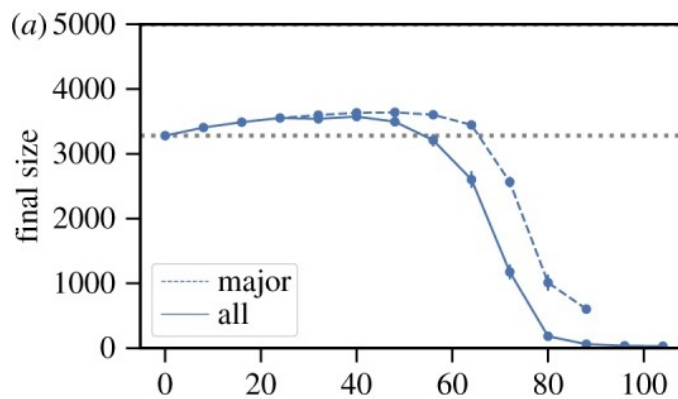


Figure 2: Social distancing can lead to an increase in the final epidemic size in the configuration model. The x -axis indicates the rewiring rate. The horizontal line is the final size when $\omega = 0$.

initially a randomly chosen vertex is infected. They show in their Theorem 2.1 that if $t_n = \inf\{t : T(t) \geq \log n\}$ then the two systems can be defined on the same space so that

$$\sup_{0 \leq t \leq t_n} |(I^n(t), I_E^n(t), T^n(t)) - (I(t), I_E(t), T(t))| \xrightarrow{\mathbb{P}} 0$$

as $n \rightarrow \infty$.

Let $S^n(t)$ be the number of susceptibles at time t and $W^n(t)$ be the number of susceptible-susceptible edges created by rewiring by time t and let

$$X^n(t) = (S^n(t), I^n(t), I_E^n(t), W^n(t)).$$

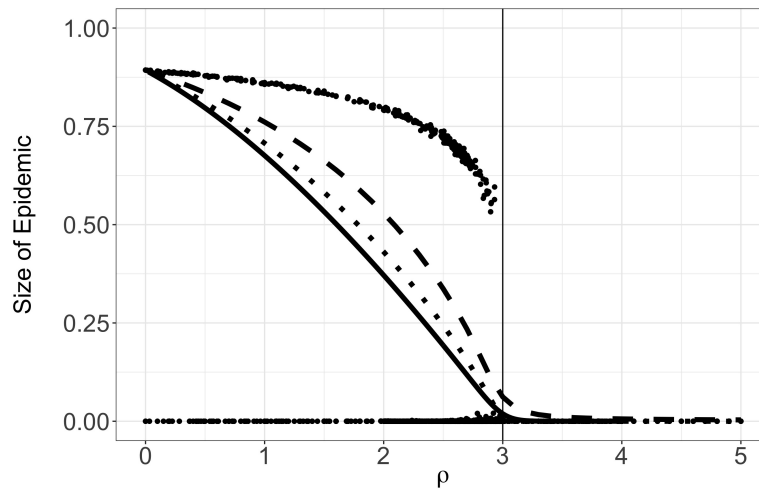


Figure 3: Simulation of a continuous time Erdős-Rényi graph with $\mu = 5$, $\gamma = 1$, $\lambda = 1$ and ρ varying. Note that $\rho_c \approx 3$, the value predicted by (1.2.4). The lowest line is the final size of delSIR which is continuous. The dotted and dashed lines are approximations derived in [14] that turned out to be inaccurate. The top curve comes from simulating evoSIR.

Let $x(t) = (s(t), i(t), i_E(t), w(t))$ be the solution of the ODE

$$\begin{aligned} \frac{ds}{dt} &= -\lambda i_E, \\ \frac{di}{dt} &= -\gamma i + \lambda i_E, \\ \frac{di_E}{dt} &= -\lambda i_E + \lambda \mu i_E s - \lambda \frac{i_E^2}{2} + 2\lambda i_E \frac{w}{s} - \omega i_E (1 - \alpha + \alpha(1 - i)), \\ \frac{dw}{dt} &= w\alpha i_E s - 2\lambda i_E \frac{w}{s}. \end{aligned} \quad (1.3.1)$$

Here α is the probability that an edge is rewired as in the definition of SIR- ω model. Theorem 2.2 in [2], which is proved using results of Darling and Norris [5], shows that for any $t_0 > 0$

$$\sup_{0 \leq t \leq t_0} |X^n(t)/n - x(t)| \xrightarrow{\mathbb{P}} 0$$

as $n \rightarrow \infty$, provided that $I^n(t)/n \rightarrow i(0) > 0$. It is interesting to note, see their Section 3, that the ODE in (1.3.1) is closely related to the “pair approximation” for SIR- ω model.

To explain the phrase in quotes, we note that “mean-field equations” come from pretending that the states of site are independent; the pair approximation from assuming it is a Markov chain. In practice, this approach means that probabilities involving three sites are reduced to probabilities involving 1 and 2 sites using a conditional independence property. For the details of the computation see Chapter 7 in [1].

Letting $T^n = n - S^n(\infty)$ (which is the final size of the epidemic) they make Conjecture 2.1 that one can interchange two limits $n \rightarrow \infty$ and $t \rightarrow \infty$ to conclude

$$T^n/n \rightarrow 1 - s(\infty).$$

To formulate a result that is independent of the validity of the conjecture they let

$$x^\epsilon(t) = (s^\epsilon(t), i^\epsilon(t), i_E^\epsilon(t), w^\epsilon(t))$$

be the solution to the ODE when

$$x^\epsilon(0) = (1 - \epsilon, \epsilon, L^{-1}\epsilon, 0) \quad \text{where} \quad L = \frac{\lambda}{\lambda(\mu - 1) - \omega}.$$

Letting $\tau_{SIR} = 1 - \lim_{\epsilon \downarrow 0} s^\epsilon(\infty)$, their Theorem 2.3 states that

$$\lim_{\lambda \downarrow \lambda_c} \tau_{SIR} \begin{cases} = 0 & \text{if } \gamma > \omega(2\alpha - 1) \text{ or } \mu < 2\omega\alpha/[\omega(2\alpha - 1) - \gamma], \\ > 0 & \text{if } \gamma < \omega(2\alpha - 1) \text{ and } \mu > 2\omega\alpha/[\omega(2\alpha - 1) - \gamma]. \end{cases} \quad (1.3.2)$$

When $\alpha = 1$, $\omega = \rho$ and $\gamma = 1$ then we have a discontinuous phase transition if

$$\rho > 1 \quad \text{and} \quad \mu > \frac{2\rho}{\rho - 1}.$$

For the case of the delSIR model ($\alpha = 0$, $\omega = \rho$ and $\gamma = 1$) studied in [14], the transition is always continuous. This follows from Theorem 2.3 since $1 > -\rho$. From the last calculation we see that the phase transition is always continuous if $\alpha < 1/2$.

In the case of the SI- ω model they show (see Theorem 2.6 of [2]) that if $\mu > 1$ and ω and α are held fixed then the phase transition is discontinuous if and only if $\alpha > 1/3$ and $\mu > 3\alpha/(3\alpha - 1)$. When $\alpha = 1$ this is $\mu > 3/2$ which is the condition in Example 1.10 in Section 1.4.

Theorem 2.4 in [2] gives results for the epidemic starting from a single infected individual. If we let τ_{SIR}^1 be the limiting fraction of final epidemic size conditionally on a large outbreak then

$$\lim_{\lambda \downarrow \lambda_c} \tau_{SIR}^1 \begin{cases} = 0 & \text{if } \gamma > \omega(3\alpha - 1) \text{ or } \mu < 3\omega\alpha/[\omega(3\alpha - 1) - \gamma], \\ > 0 & \text{if } \gamma < \omega(3\alpha - 1) \text{ and } \mu > 3\omega\alpha/[\omega(3\alpha - 1) - \gamma]. \end{cases} \quad (1.3.3)$$

See their paper for a precise statement. Remark 2.4 in [2] states the conjecture that the 3's in the first condition should be 2's.

For SI- ω model they compute the fraction of final epidemic size τ_{SI}^1

$$\lim_{\lambda \downarrow \lambda_c} \tau_{SI}^1 := \tau_0(\mu, \alpha).$$

To give the value of $\tau_0(\mu, \alpha)$, we need some notations. For $\mu > 1$ and $\alpha \in [0, 1]$ let

$$\theta(\mu, \alpha) = \frac{2\alpha(\mu - 1)}{\mu + \alpha(\mu - 1)} \quad \text{and} \quad f_0(x) = \log(1 - x) + \frac{x}{1 - \theta(\mu, \alpha)},$$

then $\tau_0(\mu, \alpha)$ is the largest solution in $[0, 1]$ of $f_0(x) = 0$.

Ball and Britton also made connections of their paper with our paper (as well as an earlier version of this paper) in [2, Section 4]. In particular, they showed in their Figure 5 that in the case $\mu = 2$ their predicted final size $\tau_0(2, 1)$ agrees with simulation results well.

1.4 Statement of our main results

From now on the reader can forget about the meaning of notations used by Ball and Britton. We fix ρ , the rewiring rate, and vary λ . We let $\alpha = \rho m_1/\lambda$. In view of the definition of Δ in (1.4.1), the natural assumption is $\mathbb{E}(D^3) < \infty$. Some of our results can be proved under this assumption, while some need something a little stronger. Specifically, we need finite fifth moment to prove (1.4.2). To simplify things **we assume** $\mathbb{E}(D^5) < \infty$ **throughout**.

Theorem 1.6. Assume $\mathbb{E}(D^5) < \infty$. Consider the delSI model and evoSI model on $\text{CM}(n, D)$. (i) the critical values of delSI and evoSI are the same, i.e.,

$$\alpha_c = m_2 - 2m_1; \text{ equivalently, } \lambda_c = \frac{\rho m_1}{m_2 - 2m_1}.$$

(ii) When $\alpha < \alpha_c$, which is the supercritical case, the probability of a large epidemic is the same in the two models, which is equal to the survival probability $q(\lambda)$ of the two-stage branching process $\bar{Z}_m$ defined in Section 1.2 (with the τ there equal to $\lambda/(\lambda + \rho)$ in our notation).

The proof of Theorem 1.6 that we give in Section 2 is very similar to one for Theorem 1.3 given in [14] for Erdős-Rényi random graphs. Here the fact that we have only $\mathbb{E}(D^5) < \infty$ rather than exponential upper bounds on $\mathbb{P}(D \geq k)$ changes some of the estimates.

Here and in what follows, formulas are sometimes easier to evaluate if we use the “factorial moments” $\mu_k = \mathbb{E}[D(D-1) \cdots (D-k+1)]$, since these can be computed from the k -th derivative of the generating function. To translate between the two notations:

$$\mu_1 = m_1, \quad \mu_2 = m_2 - m_1, \quad \mu_3 = m_3 - 3m_2 + 2m_1.$$

In particular $\alpha_c = m_2 - 2m_1 = \mu_2 - \mu_1$.

Our next result gives an almost sufficient and necessary condition for the discontinuous phase transition of evoSI. We use the word ‘almost’ since the case $\Delta = 0$ is not treated here.

Theorem 1.7. Assume $\mathbb{E}(D^5) < \infty$. Consider the evoSI epidemic on the configuration model $\text{CM}(n, D)$ with one uniformly randomly chosen vertex initially infected. Let

$$\Delta = -\frac{\mu_3}{\mu_1} + 3(\mu_2 - \mu_1). \quad (1.4.1)$$

Let I_∞ be the final epidemic size. If $\Delta > 0$, then there is a discontinuous phase transition. For some $\epsilon_0 > 0$ and some $\delta_0 > 0$,

$$\lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}_1(I_\infty/n > \epsilon_0 | I_\infty/n > \eta) = 1 \quad \text{for all } \alpha_c - \delta_0 < \alpha < \alpha_c. \quad (1.4.2)$$

If $\Delta < 0$, then there a continuous phase transition. For any $\epsilon > 0$, there exists some $\delta > 0$, so that

$$\lim_{n \rightarrow \infty} \mathbb{P}_1(I_\infty/n > \epsilon) = 0 \quad \text{for } \alpha_c - \delta < \alpha < \alpha_c. \quad (1.4.3)$$

To see what this result says we consider some examples.

Example 1.8. Random r -regular graph, $r \geq 3$. Here nr must be even. If we choose the degree distribution $\mathbb{P}(D = r) = 1$, and condition the graph to be simple, i.e., no self-loops or parallel edges then the result is a random regular graph. The case $r = 2$ is excluded because in that case the graph consists of a number of circles. The critical value is $\alpha_c = m_2 - 2m_1 = r^2 - 2r > 0$ when $r > 2$. For $k \leq r$,

$$\mu_k = r(r-1) \cdots (r-k+1),$$

so

$$\begin{aligned} \Delta &= -(r-1)(r-2) + 3(r(r-1) - r) \\ &= -(r-1)(r-2) + 3r(r-2) = (r-2)(2r+1) > 0, \end{aligned}$$

and the phase transition is discontinuous for all $r \geq 3$.

Example 1.9. Geometric(p). The factorial moments are $\mu_1 = 1/p$, $\mu_2 = 2(1-p)/p^2$, and $\mu_3 = 6(1-p)^2/p^3$. $\alpha_c = \mu_2 - \mu_1 = (2-3p)/p^2$, so we need to take $p < 2/3$ to have $\alpha_c > 0$.

$$\begin{aligned}\Delta &= -\frac{6(1-p)^2}{p^2} + 3\left(\frac{2(1-p)}{p^2} - \frac{1}{p}\right) \\ &= -\frac{6}{p^2} + \frac{12}{p} - 6 + \frac{6}{p^2} - \frac{6}{p} - \frac{3}{p} = \frac{3}{p} - 6,\end{aligned}$$

so the phase transition is discontinuous if $p < 1/2$.

Our last example concerns the configuration model generated from Poisson distribution:

Example 1.10. Poisson(μ). The factorial moments $\mu_k = \mu^k$, so the critical value $\alpha_c = \mu_2 - \mu_1 = \mu^2 - \mu$, which is positive if $\mu > 1$. This condition is natural since if $\mu < 1$ then there is no giant component in the graph and a large epidemic is impossible.

$$\Delta = -\mu^2 + 3(\mu^2 - \mu) = 2\mu^2 - 3\mu,$$

so the phase transition for evoSI is discontinuous if $\mu > 3/2$, which is the result given in [2].

Remark 1.11. We believe that the result of Example 1.10 holds for Erdős-Rényi($n, \mu/n$). To prove this rigorously, one first has to prove a quenched version of Theorem 1.7 (i.e., showing that (1.4.2) and (1.4.3) hold with high probability over *any* degree sequence $D_1, \dots, D_n$ (that are *not necessarily i.i.d.*) such that the k -th factorial moment of the empirical distribution converges to μ^k for any $k \geq 0$). We believe that this can be shown using the same ideas in the proof of Theorem 1.7. Then one can transfer results for the configuration model to Erdős-Rényi($n, \mu/n$) using [22, Theorems 7.18 and 7.19], which says that conditionally on having the same degrees, the random graphs generated from these two models have the same distribution.

As a notational note, in this paper we will use $C, C_1, C_2, \dots$ to denote various constants whose specific values might change from line to line. Occasionally when we have an important constant we will number it by the formula it first appeared in, e.g., $C_{2.3.2}$ below.

1.5 Sketch of Proof of Theorem 1.7

The proof of Theorem 1.7 is done by constructing auxiliary models that are upper/lower bounds for evoSI. We introduce a process which we call avoSI (avo is short for avoiding infection) in Section 3.1 and prove that the final set of infected sites in avoSI stochastically dominates evoSI. We also construct a lower bounding process which we call AB-avoSI in Section 4.1, where we prove that the final set of infected sites in evoSI stochastically dominates AB-avoSI. The AB in the name comes from the two counters associated with half-edges that prevent transmission of infections along $S-I$ edges created by $I-I$ rewirings.

The starting point to analyze evoSI via avoSI and AB-avoSI is the following Lemma 1.12. Let $q(\lambda)$ be the survival probability for the two-phase branching process $\{\bar{Z}_m, m \geq 0\}$ introduced in Section 1.2. Recall that the individual in the first generation has offspring distribution Binomial($D, \lambda/(\lambda + \rho)$) while later generations have offspring distribution Binomial($D^* - 1, \lambda/(\lambda + \rho)$) where D^* is the size-biased version of D ,

$$\mathbb{P}(D^* = j) = \frac{j\mathbb{P}(D = j)}{\mathbb{E}(D)}, j \geq 0.$$

Lemma 1.12. *AvoSI, evoSI, AB-avoSI and delSI have the same critical value λ_c and in the supercritical regime $\lambda > \lambda_c$ the probability of a large outbreak is equal to $q(\lambda)$ in all four models.*

Proof. Results in Section 3.1 and 4.1 imply that

$$\text{avoSI} \succeq \text{evoSI} \succeq \text{AB-avoSI} \succeq \text{delSI}$$

where epidemic1 $\succeq$ epidemic2 means that the two epidemics can be constructed on the same space so that the final epidemic size in epidemic1 is greater than or equal to that of epidemic2. In fact, we will prove this chain of comparisons in Lemmas 3.2, 4.1 and 4.2, respectively. It remains to show that avoSI and delSI has the same critical value and probability of a large outbreak. This is proved in Lemma 3.3. $\square$

Below we will use λ_c and $\alpha_c = \rho\lambda_c/m_1$ to denote the critical value. Recall the definition of the generating function G in (1.2.1). Consider a function f defined by

$$f(w) = \log \left(\frac{m_1 w}{G'(w) + \alpha(1-w)G(w)} \right) + \frac{\alpha}{2}(w-1)^2. \quad (1.5.1)$$

Theorem 1.13. *Assume $\mathbb{E}(D^5) < \infty$. Consider the avoSI epidemic on the configuration model $\text{CM}(n, D)$ with one uniformly randomly chosen vertex initially infected. Suppose $\alpha < \alpha_c$ so that we are in the supercritical regime. Let $\tilde{I}_\infty$ be the final epidemic size. Set*

$$\sigma = \sup\{w : 0 < w < 1, f(w) = 0\} \quad \text{with } \sup(\emptyset) = 0, \quad (1.5.2)$$

$$\nu = 1 - \exp \left(-\frac{\alpha}{2}(\sigma-1)^2 \right) G(\sigma). \quad (1.5.3)$$

If we suppose

($\star$) either $\sigma = 0$ or $0 < \sigma < 1$ and there is a $\delta > 0$ so that $f < 0$ on $(\sigma - \delta, \sigma)$, then for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n < \nu + \epsilon) = \lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n > \nu - \epsilon | \tilde{I}_\infty/n > \eta) = 1.$$

Though ν does not give the correct final size of the evoSI epidemic, the formula for $f(w)$ is accurate enough for w near 1 to identify when the phase transition is continuous.

Theorem 1.14. *Consider the avoSI epidemic on the configuration model $\text{CM}(n, D)$ and let $\tilde{I}_\infty$ be the final epidemic size. Set*

$$\Delta = -\frac{\mu_3}{\mu_1} + 3(\mu_2 - \mu_1). \quad (1.5.4)$$

If $\Delta < 0$, then there a continuous phase transition. For any $\epsilon > 0$, there exists some $\delta > 0$, so that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n > \epsilon) = 0 \quad \text{for } \alpha_c - \delta < \alpha < \alpha_c. \quad (1.5.5)$$

We can show that $\Delta > 0$ implies that there is a discontinuous phase transition in avoSI, but that result does not help us prove Theorem 1.7. To get Theorem 1.14 from Theorem 1.13 we compute, see Section 3.6, that

$$\begin{aligned} f'(1) &= -\left(\frac{m_2 - 2m_1}{m_1} - \frac{\rho}{\lambda} \right) \quad \text{which is } < 0 \text{ for } \alpha < \alpha_c, \\ f'(1) &= 0, \quad f''(1) = \Delta \quad \text{when } \alpha = \alpha_c. \end{aligned} \quad (1.5.6)$$

When $\Delta > 0$, as w decreases from 1 the curve of f turns up, and σ stays bounded away from 0. When $\Delta < 0$, the curve of f turns down, and σ converges to 1 as $\alpha \rightarrow \alpha_c$. See Figure 4.

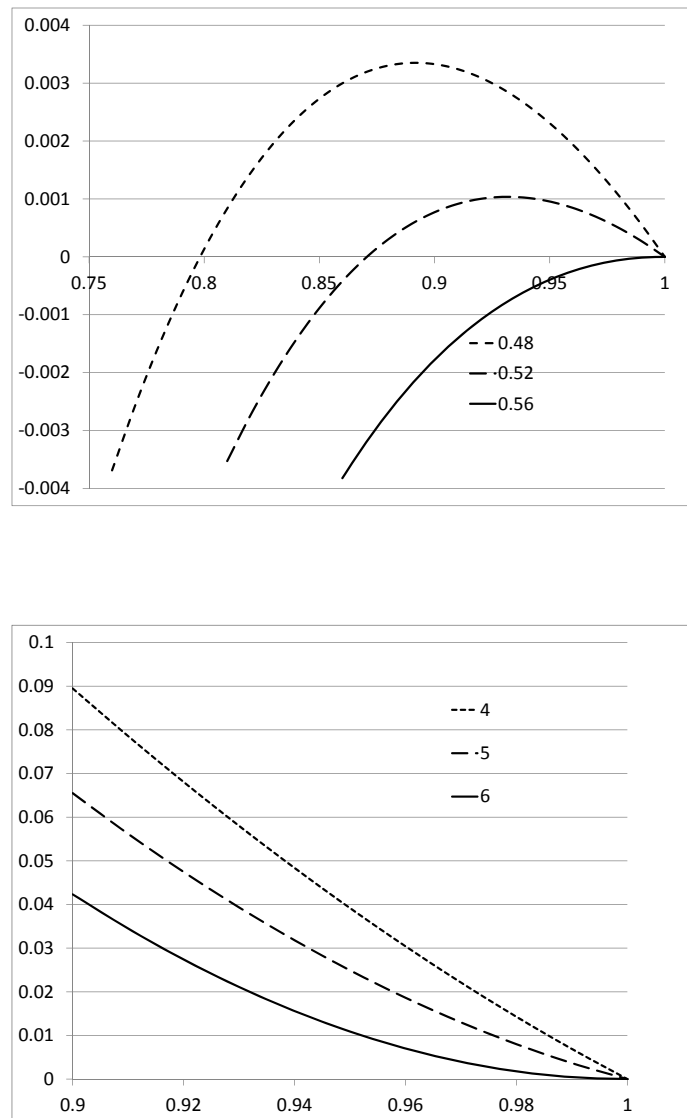


Figure 4: The behavior of $f(w)$ near 1 with respect to different α 's for the Erdős-Rényi graph. In the top graph $\mu = 1.4$, which has $\Delta < 0$. $\alpha_c = \mu^2 - \mu = .56$. Notice that as α increases to 0.56 the intersection with the x axis tends to 1, so the transition is continuous. In the bottom graph $\mu = 3$, which has $\Delta > 0$. $\alpha_c = \mu^2 - \mu = 6$. Notice that when $\alpha \leq \alpha_c$, $f(w) > 0$ for $w \in [0.9, 1)$, so σ is bounded away from 1.

Theorem 1.15. Consider the AB-evoSI epidemic on the configuration model $\text{CM}(n, D)$. Let $\tilde{I}_\infty$ be the final epidemic size. Let Δ be the quantity defined in (1.5.4). If $\Delta > 0$ then there is a discontinuous phase transition. For some $\epsilon_0 > 0$ and some $\delta_0 > 0$,

$$\lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}_1(\tilde{I}_\infty/n > \epsilon_0 | \tilde{I}_\infty/n > \eta) = 1 \quad \text{for all } \alpha_c - \delta_0 < \alpha < \alpha_c. \quad (1.5.7)$$

Theorem 1.7 follows from Theorem 1.14 and Theorem 1.15.

Proof of Theorem 1.7. Lemma 1.12 implies that

$$\lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n > \eta) = \lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n > \eta). \quad (1.5.8)$$

Equations (1.5.8), (1.5.7) and the fact $\text{evoSI} \succeq \text{AB-avoSI}$ (proved in Lemma 4.1) imply that

$$\lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(I_\infty/n > \epsilon_0 | I_\infty/n > \eta) \geq \lim_{\eta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}(\tilde{I}_\infty/n > \epsilon_0 | \tilde{I}_\infty/n > \eta) = 1,$$

for $\alpha_c - \delta_0 < \alpha < \alpha_c$. This is exactly (1.4.2). Equation (1.4.3) follows from (1.5.5) and the fact that $\text{avoSI} \succeq \text{evoSI}$. $\square$

1.6 Sketch of Proof of Theorem 1.13

To begin to explain the ideas behind the analysis of epidemics on evolving graphs we need to recall some history. Volz [27] was the first to derive a limiting ODE system for an SIR epidemic on a (static) graph generated by the configuration model. Miller [18] later simplified the derivation to produce a single ODE. The results of Volz and Miller were based on heuristic computations, but later their conclusion was made rigorous by Decreusefond et al [6] assuming $\mathbb{E}(D^5) < \infty$.

Janson, Lukzak, and Windridge [13] proved the result under more natural assumptions. They studied the epidemic on the graph by revealing its edges dynamically while the epidemic spreads. Recall that the configuration is constructed using half-edges. The authors in [13] call a half-edge free if it has not yet been paired with another half-edge. They call a half-edge susceptible, infected or removed according to the state of its vertex. To modify their construction to include rewiring we add the third bullet below. Hereafter we use “randomly chosen” and “at random” to mean that the distribution of the choice is uniform over the set of possibilities.

- Each free infected half-edge chooses a free half-edge at rate λ . Together the pair forms an edge and is removed from the collection of half-edges. If the pairing is with a susceptible half-edge then its vertex becomes infected and all its edges become infected half-edges.
- Infected vertices recover and enter the removed state at rate 1.
- Each infected half-edge gets removed from the vertex that it is attached to at rate ρ and immediately becomes re-attached to a randomly chosen vertex.

To analyze the avoSI model we follow the approach in Janson, Luczak, and Windridge [13] and construct the graph as we run the infection process. The construction of this process and its coupling to evoSI are described in Section 3.1. Initially the graph consists of half-edges connected to vertices, as in the configuration model construction before the half-edges are paired. Let $\tilde{X}_t$ be the total number of half-edges at time t and let $\tilde{X}_{I,t}$ be the number of half-edges that are attached to infected vertices and let $\tilde{S}_{t,k}$ be the number of susceptible vertices with k half-edges at time t . The evolution of $\tilde{S}_{t,k}$ in avoSI is given by, see (3.2.1),

$$d\tilde{S}_{t,k} = - \left(\lambda \tilde{X}_{I,t} \frac{k \tilde{S}_{t,k}}{\tilde{X}_t - 1} \right) dt + \left(1_{\{k \geq 1\}} \rho \tilde{X}_{I,t} \frac{\tilde{S}_{t,k-1}}{n} \right) dt - \left(\rho \tilde{X}_{I,t} \frac{\tilde{S}_{t,k}}{n} \right) dt + d\tilde{M}_{t,k},$$

where $\tilde{M}_{t,k}$ is a martingale and we have returned to using λ as the infection rate and ρ as the rewiring rate.

Following [13] we time-change the process by multiplying the original transition rates by $(\tilde{X}_t - 1)/(\lambda \tilde{X}_{I,t})$. Let $\bar{X}_t$ be the number of half-edges at time t in the time changed

process, and let $\bar{X}_{S,t}$ be the number of half-edges that are attached to susceptible vertices. Using $\bar{S}_{t,k}$ for the time-changed process the new dynamics are, see (3.2.3),

$$\begin{aligned} d\left(\frac{\bar{S}_{t,k}}{n}\right) &= -\left(k\frac{\bar{S}_{t,k}}{n}\right)dt + \left(1_{(k \geq 1)}\frac{\rho}{\lambda}\frac{\bar{X}_t - 1}{n}\frac{\bar{S}_{t,k-1}}{n}\right)dt \\ &\quad - \left(\frac{\rho}{\lambda}\frac{\bar{X}_t - 1}{n}\frac{\bar{S}_{t,k}}{n}\right)dt + d\left(\frac{\bar{M}_{t,k}}{n}\right). \end{aligned} \quad (1.6.1)$$

Note that, thanks to the time change, the number of infected half-edges $\bar{X}_{I,t}$ no longer appears in the equation. Let γ_n be the first time there are no infected half-edges. Let $w(t) = \exp(-t)$ and $m_1 = \mathbb{E}(D)$. The key to the proof of Theorem 1.13 is to show

$$\begin{aligned} \sup_{0 \leq t \leq \gamma_n} \left| \frac{\bar{X}_t}{n} - m_1 w(t) \right| &\xrightarrow{\mathbb{P}} 0, \\ \sup_{0 \leq t \leq \gamma_n} \left| \frac{\sum_{k=0}^{\infty} \bar{S}_{t,k}}{n} - F_0(w(t)) \right| &\xrightarrow{\mathbb{P}} 0, \\ \sup_{0 \leq t \leq \gamma_n} \left| \frac{\sum_{k=0}^{\infty} k \bar{S}_{t,k}}{n} - F_1(w(t)) \right| &\xrightarrow{\mathbb{P}} 0, \end{aligned} \quad (1.6.2)$$

where

$$\begin{aligned} F_0(w) &= \exp(-(\alpha/2)(w-1)^2)G(w), \\ F_1(w) &= \exp(-(\alpha/2)(w-1)^2)w(G'(w) + \alpha(1-w)G(w)). \end{aligned} \quad (1.6.3)$$

From the results above, we see that

$$\frac{\bar{X}_t}{\bar{X}_{S,t}} \xrightarrow{\mathbb{P}} \frac{m_1 w}{\exp(-(\alpha/2)(w-1)^2) \cdot (G'(w) + \alpha(1-w)G(w))}. \quad (1.6.4)$$

The logarithm of the right-hand side is $f(w)$. Under assumption $(\star)$,

$$\sigma = \sup\{w : 0 < w < 1, f(w) = 0\}$$

gives the time $z = -\log(\sigma)$ at which the infection dies out in the time-changed process and ν defined in (1.5.3) gives the fraction of sites which have been infected.

There are four steps in the proof of (1.6.2):

- In Section 3.2 we show that for each fixed $k \in \mathbb{N}$, $\{\bar{S}_{t,k}/n, t \geq 0\}_{n \geq 1}$ is a tight sequence of processes.
- In Section 3.3 we show that any subsequential limit satisfies a system of differential equations (3.3.4) that has a unique solution $\bar{s}_{t,k}$, so $\bar{S}_{t,k}/n \rightarrow \bar{s}_{t,k}$.
- Section 3.4 we deal with the technicality of showing that the limit of $\sum_{k=0}^{\infty} k \bar{S}_{t,k}/n$ is the sum of the limits $\sum_{k=0}^{\infty} k \bar{s}_{t,k}$.
- In Section 3.5 we complete the proof by establishing the formulas for σ and ν .

1.7 Sketch of Proof of Theorem 1.15

In the AB-avoSI model, each half-edge i has two indices $A(i, t)$ and $B(i, t)$.

- The infection index $A(i, t) = 0$ if i has not been infected by time t . If i first become an infected half-edge at time s , then we set $A(i, t) = s$ for all $t \geq s$.
- The rewiring index $B(i, t) = 0$ if i has not been rewired by time t . If i gets rewired at time s , then we update the value of $B(i, s)$ to be s , regardless of whether i has been rewired before or not. $B(i, \cdot)$ remains constant between consecutive rewirings.

Suppose an infected half-edge i pairs with a susceptible half-edge j at time t , then (in the AB-avoSI model) i will transmit an infection to j if and only if $A(i, t) > B(j, t)$. See Section 4.1 for more details about the AB-avoSI model and its relationship to evoSI. Let $\tilde{S}_{t,k}$ be the number of susceptible vertices with k half-edges at time t and set

$$G_{i,j} = \{I(i, t) = 1, A(i, t) \leq B(j, t)\}.$$

Here $I(i, t)$ is an indicator function such that $I(i, t) = 1$ if half-edge i is an infected half-edge at time t (see the first paragraph of Section 4.2 for the definitions of the notations $I(i, t)$, $S(i, t)$, $S(i, k, t)$, $D(j, t)$ appearing below). As in the avoSI model we make a time-change by multiplying the original transition rates by $(\tilde{X}_t - 1)/(\lambda \tilde{X}_{I,t})$. Using a hat to denote the quantities after the time-change in the AB-avoSI model, we have that, for all $k \geq 0$,

$$\begin{aligned} d\hat{S}_{t,k} = & -k\hat{S}_{t,k} dt + 1_{\{k \geq 1\}} \frac{\rho}{\lambda} \frac{\hat{S}_{t,k-1}}{n} (\hat{X}_t - 1) dt - \frac{\rho}{\lambda} \frac{\hat{S}_{t,k}}{n} (\hat{X}_t - 1) dt \\ & + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} 1_{\{S(j,k+1,t)=1\}} \right) dt + d\hat{M}_{t,k}, \end{aligned}$$

where $\hat{M}_{t,k}$ is a martingale. See equation (4.2.2). This system of equations is not solvable but we can expand in powers of t to study the time-changed system for small t . If we let $\hat{X}_t$, $\hat{X}_{I,t}$, $\hat{S}_t$ be the number of half-edges, the number of infected half-edges and the number of susceptible vertices, respectively, then we have

$$\begin{aligned} \hat{X}_{I,t} = & \hat{X}_{I,0} + \int_0^t \left(-2(\hat{X}_u - 1) + \sum_{k=0}^{\infty} k^2 \hat{S}_{u,k} - \frac{\rho}{\lambda} \frac{\hat{S}_u}{n} (\hat{X}_u - 1) \right) du + \hat{M}_t \\ & - \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} (D(j, t) - 1) 1_{\{S(j,t)=1\}} \right) dt, \end{aligned} \quad (1.7.1)$$

where $\hat{M}_t$ is a martingale. See equation (4.3.12). Define

$$E(t) = \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} D(j, t) 1_{\{S(j,t)=1\}} \right). \quad (1.7.2)$$

By expanding $\hat{S}_{t,k}$ around $t = 0$ up to the second order, we get, for any $\epsilon > 0$, λ close to λ_c and t_0 close to 0,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(\hat{X}_{I,t} \geq \left(\frac{\rho m_1}{2\lambda_c^2} (\lambda - \lambda_c) t + \frac{m_1 \Delta}{4} t^2 - \epsilon \right) n \right. \\ \left. - \int_0^t E(u) du, \forall 0 \leq t \leq \gamma_n \wedge t_0 \right) = 1. \end{aligned} \quad (1.7.3)$$

See (4.3.3) of Lemma 4.4. Here the Δ is the same as the one in (1.4.1), i.e.,

$$\Delta = -\frac{\mu_3}{\mu_1} + 3(\mu_2 - \mu_1).$$

Remark 1.16. In the case of avoSI, we have that, for n large,

$$\bar{X}_{I,t} = \bar{X}_t - \bar{X}_{S,t} = \bar{X}_{S,t} \left(\frac{\bar{X}_t}{\bar{X}_{S,t}} - 1 \right) \sim \bar{X}_{S,t} (f(\exp(-t)) - 1)$$

with f defined in (1.5.1). By expanding f up to the second order and using (1.5.6) along with the fact that $\exp(-t) = 1 - t + t^2/2 + o(t^2)$, we get

$$\begin{aligned}\bar{X}_{I,t} &\sim nm_1(-f'(1)(t - t^2/2) + f''(1)t^2/2 + o(t^2)) \\ &\geq nm_1(\rho(\lambda - \lambda_c)t/(2\lambda_c^2) + (\Delta t^2)/4)\end{aligned}\quad (1.7.4)$$

for small t and λ close to λ_c . In the case of AB-avoSI, see (4.5.12), we have,

$$n \left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2} t + \frac{m_1 \Delta}{8} t^2 - \epsilon_2 - \epsilon_6 \right)$$

as a lower bound when $t > \epsilon$ (the ϵ_2 and ϵ_6 here are some small numbers depending on ϵ). The two expansions do not match but both linear terms vanish at λ_c and the quadratic terms have the same sign so this is good enough.

The proof of Theorem 1.15 is organized into five steps:

- In Section 4.1 we define the AB-avoSI process and prove that evoSI stochastically dominates AB-avoSI.
- In Section 4.2 we derive basic moment estimates for various quantities that will prepare us for later proofs. See Lemma 4.3.
- In Section 4.3 we give rough upper and lower bounds for $\hat{I}_t$ and $\hat{X}_{I,t}$ involving the integral of $E(t)$. See Lemma 4.4. We also give an easy upper bound for $E(t)$ in (4.3.25).
- In Section 4.4 we decompose $E(t)$ into two parts (see (4.4.1)) and give refined bounds for each part. See Lemmas 4.5 and 4.6.
- In Section 4.5 we combine our estimates to complete the proof.

2 Proof of Theorem 1.6

2.1 Coupling of evoSI and delSI

We first prove Lemma 1.2 before proving Theorem 1.6. We define three sequences of random variables, which will serve as the joint randomness to couple evoSI and delSI:

- Let $T_{e,\ell}$, $\ell \geq 1$ be independent exponential random variables with mean $1/\lambda$.
- Let $R_{e,\ell}$, $\ell \geq 1$ be independent exponential random variables with mean $1/\rho$.
- Let $U_{e,\ell}$, $\ell \geq 1$ be independent random variables chosen uniformly at random from all vertices.

Construction of evoSI. We define three sets of edges in evoSI at time t :

- *Active edges*, denoted by $\mathcal{E}_t^a$, are the edges at time t that connect an infected vertex and a susceptible vertex.
- *Uninfected edges*, denoted by $\mathcal{E}_t^0$, connect two susceptible vertices.
- *Inactive edges*, denoted by $\mathcal{E}_t^i$, have both ends infected. Once an edge becomes inactive it remains inactive forever.

The three sets form a partition of all edges.

The three set-valued processes just defined are right-continuous pure jump processes. At time 0 we randomly choose a vertex u_0 to be infected. $\mathcal{E}_0^a$ consists of the edges with one endpoint at u_0 . $\mathcal{E}_0^0$ is the collection of all edges in the graph minus the set $\mathcal{E}_0^a$. $\mathcal{E}_0^i = \emptyset$. We will consider the corresponding sets for delSI, but they will be denoted by $\mathcal{D}_t$ to avoid confusion.

For each undirected edge e , let $\tau_{e,\ell}^e$ be the ℓ -th time the edge becomes active (the superscript ‘ e ’ is short for evo). To make it easier to describe the dynamics, suppose that at time $\tau_{e,\ell}^e$ we have $e = \{x_{e,\ell}, y_{e,\ell}\}$ with $x_{e,\ell}$ infected and $y_{e,\ell}$ susceptible.

- Let $T_{e,\ell}$, $\ell \geq 1$ be the time between $\tau_{e,\ell}$ and the infection of $y_{e,\ell}$ by $x_{e,\ell}$.
- Let $R_{e,\ell}$, $\ell \geq 1$ be the time between $\tau_{e,\ell}$ and the time when $y_{e,\ell}$ breaks its connection to $x_{e,\ell}$ and rewires.
- Let $U_{e,\ell}$ be the vertex that $y_{e,\ell}$ connects to at time $\tau_{e,\ell}^e + R_{e,\ell}$ (if rewiring occurs).

Initial step. To simplify writing formulas, let

$$S_{e,\ell} = \min\{T_{e,\ell}, R_{e,\ell}\}. \quad (2.1.1)$$

At time 0, a randomly chosen vertex u_0 is infected. For a vertex x , let $\mathcal{N}^0(x, t)$, $\mathcal{N}^a(x, t)$ and $\mathcal{N}^i(x, t)$ be the collection of uninfected, active and inactive edges that are connected to x at time t , respectively. At time 0 the edges $\mathcal{N}^0(u_0, 0) = \{e_{0,1}, \dots, e_{0,k}\}$ are added to the list of active edges, where k is the degree of u_0 . We have $\mathcal{E}_0^a = \{e_{0,1}, \dots, e_{0,k}\}$. Suppose $e_{0,j}$ connects u_0 and y_j . At time

$$J_1^e = \min_{1 \leq j \leq k} S_{e_{0,j},1}$$

the first event occurs. (J is for jump and e stands for evo.) Let i be the index that achieves the minimum.

- (i) If $R_{e_{0,i},1} < T_{e_{0,i},1}$, then at time J_1^e vertex y_i breaks its connection with u_0 and rewires to $U_{e_{0,i},1}$. If $U_{e_{0,i},1}$ is susceptible at time J_1^e , we move the edge $e_{0,i}$ to $\mathcal{E}_{J_1^e}^0$. On the initial step this will hold unless $U_{e_{0,i},1} = u_0$ in which case nothing has changed.
- (ii) If $T_{e_{0,i},1} < R_{e_{0,i},1}$ then at time J_1^e vertex y_i becomes infected by u_0 . We move $e_{0,i}$ to $\mathcal{E}_{J_1^e}^i$. We add edges in $\mathcal{N}^0(y_i, J_1^e -)$ to $\mathcal{E}_{J_1^e}^a$.

Induction step. For any active edges e present at time t , let $L^e(e, t) = \sup\{\ell : \tau_{e,\ell}^e \leq t\}$ and let

$$V^e(e, t) = \tau_{e,L^e(e,t)}^e + S_{e,L^e(e,t)}$$

be the time of the next event (infection or rewiring) to affect edge e . Again, the superscripts ' e ' in $L^e(e, t)$ and $V^e(e, t)$ imply that these quantities are for the evoSI model. Suppose we have constructed the process up to time J_m^e for some $m \in \mathbb{N}$. If there are no active edges present at time J_m^e , the construction is done. Otherwise, let

$$J_{m+1}^e = \min_{e \in \mathcal{E}_{J_m^e}^a} V^e(e, J_m^e).$$

Let e_m be the edge that achieves the minimum of $V^e(e, J_m^e)$, let $x(e_m)$ be the infected endpoint of e_m and $y(e_m)$ be the susceptible endpoint of e_m . To simplify notation let $L_m = L^e(e_m, J_m^e)$.

- (i) If $R_{e_m, L_m} < T_{e_m, L_m}$, then at time J_{m+1}^e vertex $y(e_m)$ breaks its connection with $x(e_m)$ and rewires to U_{e_m, L_m} . If U_{e_m, L_m} is susceptible at time J_{m+1}^e , then e_m is moved to $\mathcal{E}_{J_{m+1}^e}^0$. Otherwise it remains active.
- (ii) If $T_{e_m, L_m} < R_{e_m, L_m}$, then at time J_{m+1}^e the vertex $y(e_m)$ is infected by $x(e_m)$ and e_m is moved to $\mathcal{E}_{J_{m+1}^e}^i$. Further,

- all edges e' in $\mathcal{N}^0(y(e_m), J_{m+1}^e -)$ are moved to $\mathcal{E}_{J_{m+1}^e}^a$. Since $y(e_m)$ has just become infected, the other end of e' must be susceptible at time J_{m+1}^e .
- all edges e'' in $\mathcal{N}^a(y(e_m), J_{m+1}^e -)$ are moved to $\mathcal{E}_{J_{m+1}^e}^i$. Since $y(e_m)$ has just become infected, (a) the other end of e'' must be infected at time J_{m+1}^e , and (b) e'' cannot have been inactive earlier.

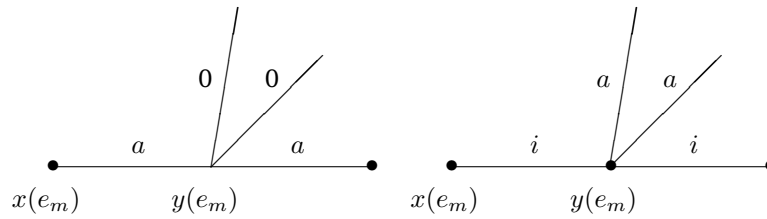


Figure 5: Change of the status of different edges after $y(e_m)$ becomes infected by $x(e_m)$. Black dots mark infected sites.

Construction of delSI. There is no rewiring in delSI, each edge will become active at most once. Thus for each undirected edge e we only need two exponential random variables $T_{e,1}$ and $R_{e,1}$, defined in the beginning of Section 2.1. This allows us to couple evoSI and delSI. Also, we use $\mathcal{D}_t^0, \mathcal{D}_t^a$ and $\mathcal{D}_t^i$ to represent the set of uninfected, active and inactive edges in delSI, respectively.

Initial step. At time 0, a randomly chosen vertex u_0 is infected. The edges $\mathcal{N}^0(u_0, 0) = \{e_{0,1}, \dots, e_{0,k}\}$ are added to the list of active edges. We have $\mathcal{D}_0^a = \{e_{0,1}, \dots, e_{0,k}\}$. Suppose e_j connects u_0 and y_j . At time

$$J_1^d = \min_{1 \leq j \leq k} S_{e_{0,j},1}$$

the first event occurs. The superscript ‘d’ stands for delSI and $S_{e,1} = \min\{T_{e,1}, R_{e,1}\}$, as defined in (2.1.1). Let i be the index that achieves the minimum.

(i) If $R_{e_{0,i},1} < T_{e_{0,i},1}$, then at time J_1^a the edge $e_{0,i}$ is removed from the graph (and hence also from the set $\mathcal{D}_{J_1^d}^a$).

(ii) $T_{e_{0,i},1} < R_{e_{0,i},1}$ then at time J_1^a vertex y_i is infected by x_i . We move $e_{0,i}$ to $\mathcal{D}_{J_1^d}^i$. We add $\mathcal{N}^0(x_i, J_1^d -)$ to $\mathcal{D}_{J_1^d}^a$.

Induction step. For any active edge e at time t , let $\tau_{e,1}^d$ be the first time that e becomes active in the delSI process. We also let

$$V^d(e, t) = \tau_{e,1}^d + S_{e,1},$$

be the time of the next event (infection or rewiring) to affect edge e . Suppose we have constructed the process up to time J_m^d . If there are no active edges present at time J_m^d , the construction is done. Otherwise, we let

$$J_{m+1}^d = \min_{e \in \mathcal{D}_{J_m^d}^a} V(e, J_m^d).$$

Let e_m be the edge that achieves the minimum of $V^d(e, J_m^d)$, let $x(e_m)$ be the infected endpoint of e_m and $y(e_m)$ be the susceptible endpoint of e .

(i) If $R_{e_m,1} < T_{e_m,1}$, then at time J_{m+1}^d the edge e_m is deleted from the graph (and hence also from the set $\mathcal{D}_{J_{m+1}^d}^a$).

(ii) If $T_{e_m,1} < R_{e_m,1}$, then at time J_{m+1}^d the vertex $y(e_m)$ is infected by $x(e_m)$. We move e_m to $\mathcal{D}_{J_{m+1}^d}^i$.

- We move all edges e' in $\mathcal{N}_{y(e_m), J_{m+1}^d}^0$ to $\mathcal{D}_{J_{m+1}^d}^a$. Since $y(e_m)$ has just become infected, the other end of e' must be susceptible at time J_{m+1}^d .
- We move all edges e'' in $\mathcal{N}_{y(e_m), J_{m+1}^d}^a$ to $\mathcal{D}_{J_{m+1}^d}^i$. Since $y(e_m)$ has just become infected, (i) the other end of e'' must be infected at time J_{m+1}^d , and (ii) e'' cannot have been inactive earlier.

We now prove by induction that avoSI dominates evoSI.

Lemma 2.1. *All vertices infected in delSI are also infected in evoSI and are infected earlier in avoSI than evoSI.*

Proof. The induction hypothesis holds for the first vertex since u_0 is infected at time 0 in both evoSI and delSI. Suppose the induction holds up to the k -st infected vertex in delSI. Assume at time t , y becomes the $(k+1)$ -st infected vertex in delSI and y is infected by vertex x through edge e . We see from the construction of delSI that this implies $T_{e,1} < R_{e,1}$. Suppose x was infected at time $s < t$ in delSI. The induction hypothesis implies that x has also been infected in evoSI at a time $s' \leq s$. There are two possible cases for y in evoSI:

- y has already become infected by time $s' + T_{e,1}$.
- y was still susceptible right before $s' + T_{e,1}$. In this case, y will be infected at time $s' + T_{e,1} \leq s + T_{e,1} = t$.

In either case x has been infected by time t in evoSI. This completes the induction step and thus proves Lemma 1.2. $\square$

2.2 The infected sites in delSI

In the introduction we have noted that delSI is equivalent to independent bond percolation. That is, we keep each edge independently with probability $\lambda/(\lambda + \rho)$ and find the component containing the initially infected vertex (say, vertex 1). To compute the size of the delSI epidemic starting from vertex 1, we apply a standard algorithm, see e.g., [17], for computing the size of the component containing 1 in the reduced graph in which edges have independently been deleted with probability $\rho/(\lambda + \rho)$. We call this the *exploration process of delSI*. At step 0 the active set $\mathcal{A}_0 = \{1\}$, the unexplored set $\mathcal{U}_0 = \{2, \dots, n\}$, and the removed set $\mathcal{R}_0 = \emptyset$. Here removed means these sites are no longer needed in the computation. In the SI model sites never enter the removed state, Let $\eta_{i,j} = \eta_{j,i} = 1$ if there is an edge connecting i and j in the reduced graph. If $\eta_{i,j} = 1$ an infection at i is transmitted to j . At step t if $\mathcal{A}_t \neq \emptyset$ we pick an $i_t \in \mathcal{A}_t$ and update the sets as follows.

$$\begin{aligned}\mathcal{R}_{t+1} &= \mathcal{R}_t \cup \{i_t\}, \\ \mathcal{A}_{t+1} &= (\mathcal{A}_t - \{i_t\}) \cup \{y \in \mathcal{U}_t : \eta_{i_t,y} = 1\}, \\ \mathcal{U}_{t+1} &= \mathcal{U}_t - \{y \in \mathcal{U}_t : \eta_{i_t,y} = 1\}.\end{aligned}$$

When $\mathcal{A}_t = \emptyset$ we have found the cluster containing 1 in the reduced graph, which will be the final set of infected sites in the SI model. The *exploration process of the configuration model* can be similarly defined (just without deletion of edges) and we let $\mathcal{J}_t$ be the set of active sites at step t in this exploration. We set $R_t = |\mathcal{R}_{t+1}| = t + 1$, $A_t = |\mathcal{A}_{t+1}|$ and $J_t = |\mathcal{J}_{t+1}|$. We make a time shift so that A_t and J_t can be coupled with two random walks with i.i.d. increments (see Lemma 2.2 below).

We now study the exploration process of the configuration model itself as well as the delSI process on such graph. Let ψ_0 have generating function G defined in (1.2.1) and ζ_0

have a $\text{Binomial}(D, \lambda/(\lambda + \rho))$ distribution whose generating function is denoted by G^ρ . The two generating functions are related by

$$G^\rho(z) = G\left(\frac{\lambda}{\lambda + \rho}z + \frac{\rho}{\lambda + \rho}\right).$$

Recall the definition of the generating function $\hat{G}$ in (1.2.2). We can similarly define $\hat{G}^\rho$. Let $\{\chi_i, i \geq 1\}$ and $\{\xi_i, i \geq 1\}$ be independent random variables with generating functions $\hat{G}$ and $\hat{G}^\rho$, respectively. Define two random walks (for integer-valued t):

$$W_0 = \psi_0 \quad W_t = W_0 + \sum_{r=1}^t (\chi_r - 1); \quad S_0 = \zeta_0 \quad S_t = S_0 + \sum_{r=1}^t (\xi_r - 1). \quad (2.2.1)$$

Let $\tau_0^W = \inf\{t \geq 0 : W_t = 0\}$ and set $\bar{W}_t = W_{t \wedge \tau_0^W}$. We define $\bar{S}_t$ in a similar way.

Lemma 2.2. *We can couple $\{J_t, 0 \leq t \leq n^{1/3} \log n\}$ and $\{\bar{W}_t, 0 \leq t \leq n^{1/3} \log n\}$ so that*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left((J_t)_{t=0}^{n^{1/3} \log n} = (\bar{W}_t)_{t=0}^{n^{1/3} \log n}\right) = 1. \quad (2.2.2)$$

Similarly, there exists a coupling of $\{A_t, 0 \leq t \leq n^{1/3} \log n\}$ and $\{\bar{S}_t, 0 \leq t \leq n^{1/3} \log n\}$ so that

$$\lim_{n \rightarrow \infty} \mathbb{P}\left((A_t)_{t=0}^{n^{1/3} \log n} = (\bar{S}_t)_{t=0}^{n^{1/3} \log n}\right) = 1. \quad (2.2.3)$$

Note that $\bar{W}_t$ and $\bar{S}_t$ can be viewed as the exploration processes of two-phase branching process Z_m and $\bar{Z}_m$ (both are defined in Section 1.2), respectively. The proof of Lemma 2.2 is deferred to the end of this section. For the rest of this section we will always work on the event

$$\{A_t = \bar{S}_t, 0 \leq t \leq n^{1/3} \log n\} \cap \{J_t = \bar{W}_t, 0 \leq t \leq n^{1/3} \log n\}$$

and hence assume that A_t and J_t have independent increments until they hit 0.

2.3 Proof of Theorem 1.6(i)

The formula for the critical value of delSI follows from standard results on percolation in random graphs. Note that in delSI each edge is kept with probability $\lambda/(\lambda + \rho)$. Using [12, Theorem 3.9] we see that

$$\lambda_c(\text{delSI}) = \frac{\rho m_1}{m_2 - 2m_1}.$$

Equivalently, we have

$$\alpha_c(\text{delSI}) = \frac{\rho m_1}{\lambda_c(\text{delSI})} = m_2 - 2m_1.$$

Recall that Lemma 1.2 shows that the final set of infected individuals in delSI is contained in the analogous set for evoSI with the same parameters so

$$\lambda_c(\text{evoSI}) \leq \lambda_c(\text{delSI}).$$

To prove that the two are equal we will show that if $\lambda < \lambda_c(\text{delSI})$ then evoSI dies out, i.e., infects only a vanishing portion of the total population as $n \rightarrow \infty$.

To compare the two evolutions, we will first run the delSI epidemic to completion. Once this is done we will randomly rewire the edges deleted in delSI. If the rewiring creates a new infection in evoSI, then we have to continue to run the process. If not, then the infected sites in the two processes are the same. Let $\mathcal{R}$ be the set of sites that

are eventually infected in delSI, and let $\mathcal{R}'$ be the set of eventually infected sites in evoSI. Let $R' = |\mathcal{R}'|$ and $R = |\mathcal{R}|$.

To get started we use a result of Janson [11] about graphs with specified degree distributions. He works in the set-up introduced by Molloy-Reed [19, 20] where the degree sequence d_i^n , $1 \leq i \leq n$ is specified and one assumes only that limiting moments exist as well some other technical assumptions that are satisfied in our case (for a more recent example see [13]):

$$\frac{1}{n} \sum_{i=1}^n d_i^n \rightarrow \mu, \quad \frac{1}{n} \sum_{i=1}^n d_i^n (d_i^n - 1) \rightarrow \theta. \quad (2.3.1)$$

The next result is Theorem 1.1 in [11]. The μ and θ in this theorem have the same meaning as (2.3.1). The “whp” below is short for with high probability, and means that the probability the inequality holds tends to 1 as $n \rightarrow \infty$. Let $D^n = d_{u_0}^n$ where u_0 is randomly chosen from $\{1, 2, \dots, n\}$.

Theorem 2.3. Suppose $\mu > 0$, $\theta > 1$, and $\mathbb{P}(D^n \geq k) \leq Ck^{1-\gamma}$ for some $\gamma > 3$ and $C < \infty$. Then there is a constant $C_{2.3.2}$ which depends on C so that the largest component $\mathcal{C}_1$ has

$$|\mathcal{C}_1| \leq C_{2.3.2} n^{1/(\gamma-1)} \quad \text{whp.} \quad (2.3.2)$$

Theorem 2.3 can also be applied to the setting where the degrees are random rather than deterministic. We have assumed that in the original graph $\mathbb{E}D^5 < \infty$, so $\mathbb{P}(D \geq k) \leq k^{-3}\mathbb{E}(D^3)$ and we have

$$\mathbb{P}(D^n \geq k) = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(D_i \geq k) = \mathbb{P}(D_1 \geq k) \leq k^{-3}\mathbb{E}(D^3) \leq Ck^{-3}.$$

It follows that

$$|\mathcal{C}_1| \leq C_{2.3.2} n^{1/3} \quad \text{whp.} \quad (2.3.3)$$

Let N_d be the number of deleted edges in the exploration process of delSI. One vertex is removed from the construction on each step, so whp the number of steps is $\leq r_0 := C_{2.3.2} n^{1/3}$. Note that $r_0 \ll n^{1/3} \log n$ so we can assume J_t has independent increments by Lemma 2.2. It follows that

$$N_d \leq r_0 + \bar{W}_{r_0} \leq r_0 + \psi_0 + \sum_{r=1}^{r_0} \chi_r.$$

where we recall that the χ_i are independent with the distribution $D^* - 1$. The inequality comes from the fact that we are counting all the edges even if they are not deleted. Since $\mathbb{E}(D^3) < \infty$, we have

$$\text{Var}(\chi_i) \leq \mathbb{E}(\chi_i^2) = \mathbb{E}[(D^* - 1)^2] \leq \mathbb{E}[(D^*)^2] = \frac{\mathbb{E}(D^3)}{\mathbb{E}(D)} < \infty$$

and $\text{Var}(\psi_0) \leq \mathbb{E}\psi_0^2 = \mathbb{E}D^2 < \infty$. Let $1/3 < a < 1/2$. Using Chebyshev's inequality

$$\mathbb{P}(N_d \geq r_0 \mathbb{E}\chi_1 + \mathbb{E}\psi_0 + r_0 + n^a) \leq (r_0 \mathbb{E}\chi_1^2 + \mathbb{E}\psi_0^2)/n^{2a} \leq Cn^{1/3-2a}. \quad (2.3.4)$$

A similar argument shows that

$$\mathbb{P}(A_{r_0} + R_{r_0} \geq r_0 \mathbb{E}\xi_1 + \mathbb{E}\zeta_0 + r_0 + n^a) \leq (r_0 \mathbb{E}\xi_1^2 + \mathbb{E}\zeta_0^2)/n^{2a} \leq Cn^{1/3-2a}. \quad (2.3.5)$$

Since $a > 1/3$, when n is large we can upper bound $r_0 \mathbb{E}\chi_1 + \mathbb{E}\psi_0 + r_0 + n^a$ and $r_0 \mathbb{E}\xi_1 + \mathbb{E}\zeta_0 + r_0 + n^a$ by $2n^a$.

At step r_0 we use random variables independent of delSI to randomly rewire the deleted edges. Let Y be the number of edges deleted up to time r_0 that rewire to the set $\mathcal{A}_{r_0+1} \cup \mathcal{R}_{r_0+1}$. By construction

$$Y = \text{Binomial}(N_d, (A_{r_0} + R_{r_0})/n).$$

Using (2.3.4), (2.3.5) and $r_0 \leq C_{2.3.2} n^{1/3}$ we see that on a set with probability $\geq 1 - Cn^{-(2a-1/3)}$

$$Y \preceq \text{Binomial}(2n^a, 2n^{a-1}) \equiv \bar{Y},$$

where $\equiv$ indicates that the last equality defines $\bar{Y}$. From this we get

$$\mathbb{P}(Y \geq 1) \leq \mathbb{P}(\bar{Y} \geq 1) \leq \mathbb{E}(\bar{Y}) = \frac{4n^{2a}}{n} \rightarrow 0, \quad (2.3.6)$$

since $a < 1/2$. Since $\{Y = 0\} \subset \{\mathcal{R} = \mathcal{R}'\}$, this shows $\mathbb{P}(\mathcal{R} = \mathcal{R}') \rightarrow 1$ as $n \rightarrow \infty$ and completes the proof of (i). To prepare for the proof of (ii) note that the conclusion (the set of infected sites coincide in delSI and evoSI) holds as long as the number of steps is smaller than $Cn^{1/3}$ even if the epidemic is supercritical.

2.4 Proof of Theorem 1.6(ii)

We need the following ingredient in the proof.

Lemma 2.4. *There is a $\gamma > 0$ so that*

$$\mathbb{P}(0 < A_{\log n} < \gamma \log n | A_0 > 0) \leq \frac{C}{\log n}.$$

Proof of Lemma 2.4. On the event $\{A_{\log n} > 0\}$ we have $A_{\log n} = S_{\log n}$, which is, by definition,

$$\zeta_0 + \sum_{r=1}^{\log n} \xi_r.$$

Hence if we take $\gamma = \mathbb{E}(\xi_1)/2$ then we have

$$\begin{aligned} \mathbb{P}(0 < A_{\log n} < \gamma \log n | A_0 > 0) &\leq \mathbb{P}\left(\sum_{r=1}^{\log n} \xi_r \leq \frac{\mathbb{E}(\xi_1) \log n}{2}\right) \\ &\leq \mathbb{P}\left(\left|\sum_{r=1}^{\log n} \xi_r - \mathbb{E}(\xi_1) \log n\right| \geq \frac{\mathbb{E}(\xi_1) \log n}{2}\right) \\ &\leq \frac{4\text{Var}(\xi_1) \log n}{(\mathbb{E}(\xi_1))^2 \log^2 n} \leq \frac{C}{\log n}. \quad \square \end{aligned}$$

Let B_d and B_e be the events that there is a large epidemic in delSI and evoSI respectively. We now use Lemma 2.4 to show the difference between the probabilities of these two events vanishes asymptotically.

Lemma 2.5. *Suppose $\lambda > \lambda_c(\text{delSI})$. As $n \rightarrow \infty$, $\mathbb{P}(B_e) - \mathbb{P}(B_d) \rightarrow 0$.*

Proof of Lemma 2.5. Clearly $\mathbb{P}(B_d) \leq \mathbb{P}(B_e)$. Let S_t be the random walk defined in (2.2.1). To prove Lemma 2.5 we need the following lemma. Now we return to the proof of Lemma 2.5. Let $F_0 = \{A_{\log n} = 0\}$, $F_1 = \{0 < A_{\log n} < \gamma \log n\}$ and $F_2 = \{A_{\log n} \geq \gamma \log n\}$. Decomposing B_d into three parts and using Lemma 2.4:

$$\mathbb{P}(B_d) = \sum_{i=0}^2 \mathbb{P}(B_d | F_i) \mathbb{P}(F_i).$$

We note that

- $\mathbb{P}(B_d|F_0) = 0$ by the definition of F_0 .
- $\mathbb{P}(F_1)$ converges to 0 by Lemma 2.4.
- $\mathbb{P}(B_d|F_2)$ converges to 1 by the same argument as the proof of Theorem 2.9(c) in [13] (see page 750-752 in [13]).

Therefore we have $\mathbb{P}(B_d) - \mathbb{P}(F_2) \rightarrow 0$. As for $\mathbb{P}(B_e)$, we note that by the remark at the end of the proof of (i) one has $\mathbb{P}(\mathcal{A}_{\log n} = \mathcal{A}'_{\log n}) \rightarrow 1$ where $\mathcal{A}'_{\log n}$ is the number of active sites in evoSI at step $\log n$. Using the decomposition $\mathbb{P}(B_e) = \sum_{i=0}^2 \mathbb{P}(B_e \cap F'_i)$ where the event F'_i is defined in a similar way to F_i with $A_{\log n}$ replaced by $A'_{\log n}$, we see

- $\mathbb{P}(B_e \cap F'_1) \leq \mathbb{P}(F'_1) \leq \mathbb{P}(F_1) + o(1) = o(1)$.
- $\mathbb{P}(B_e \cap F'_0) = o(1)$ by definition of B_e .
- $\mathbb{P}(B_e \cap F'_2) = \mathbb{P}(F'_2) + o(1) = \mathbb{P}(F_2) + o(1)$.

It follows that $\mathbb{P}(B_e) - \mathbb{P}(F_2) \rightarrow 0$. This implies that $\mathbb{P}(B_e) - \mathbb{P}(B_d) \rightarrow 0$. It remains to compute the limit of $\mathbb{P}(F_2)$. Since $\mathbb{P}(F_1) = o(1)$, we have

$$\mathbb{P}(F_2) = \mathbb{P}(F_2 \cup F_1) + o(1) = \mathbb{P}(A_{\log n} > 0) + o(1). \quad (2.4.1)$$

This completes the proof of Lemma 2.5. $\square$

To compute the limit of $\mathbb{P}(A_{\log n} > 0)$, note that due to Lemma 2.2, A_t can be coupled with the exploration process of the two-phase branching process $\bar{Z}_m$ defined in Section 1.2. Therefore we have (recall that $q(\lambda)$ is the survival probability of $\bar{Z}_m$)

$$\mathbb{P}(A_{\log n} > 0) = \mathbb{P}(\bar{Z}_m > 0, \forall m) + o(1) = q(\lambda) + o(1).$$

By (2.4.1) we see that $\mathbb{P}(F_2) \rightarrow q(\lambda)$ as well. This implies that both $\mathbb{P}(B_e)$ and $\mathbb{P}(B_d)$ converge to $q(\lambda)$ as $n \rightarrow \infty$ and completes the proof of Theorem 1.6(ii). Note that using the fact that delSI is equivalent to independent bond percolation, the statement $\mathbb{P}(B_d) \rightarrow q(\lambda)$ also follows from standard results on percolation in random graphs. See, e.g., [12, Theorem 3.9]. It remains to prove Lemma 2.2 to complete the proof of Theorem 1.6.

Proof of Lemma 2.2. We only prove equation (2.2.2) since the other one follows from (2.2.2) and the fact that delSI is equivalent to percolation with edge retaining probability $\lambda/(\lambda + \rho)$. The proof consists of two steps. First, we define an empirical version of W_t . Let $D_1, \dots, D_n$ be i.i.d. random variables sampled from the distribution of D . Given a sample of $D_1, \dots, D_n$, let ψ_0^n be sampled from the (random) distribution

$$\mathbb{P}_n(\psi_0^n = k) = \frac{1}{n} |\{1 \leq i \leq n : D_i = k\}|, \quad \forall k \geq 0,$$

which is the sample empirical distribution. Let $\chi_r^n, r \geq 1$ have the distribution

$$\mathbb{P}_n(\chi_r^n = k) = \frac{1}{D_1 + \dots + D_n} (k+1) |\{1 \leq i \leq n : D_i = k+1\}| \quad \forall k \geq 0.$$

Define

$$W_t^n = \psi_0^n + \sum_{r=1}^t (\chi_r^n - 1), \quad t \in \mathbb{N}.$$

In other words, W_t^n is a random walk in the random environment given by $D_1, \dots, D_n$. We define $\bar{W}_t^n$ to be $W_{t \wedge \tau'}^n$ where τ' is the first time W_t^n hits zero. Note that the condition $\mathbb{E}(D^5) < \infty$ implies that $\max_{1 \leq i \leq n} D_i = o(n^{1/4} \log n)$ whp for any $\epsilon > 0$. Indeed we have

$$\mathbb{P}\left(\max_{1 \leq i \leq n} D_i > n^{1/4} \log n\right) \leq n \mathbb{P}(D_1 > n^{1/4} \log n) \leq n \cdot n^{-1} \log^{-4} n \mathbb{E}(D^4) \leq C \log^{-4} n. \quad (2.4.2)$$

Consequently, for n large enough, whp

$$\sqrt{\frac{n}{\max_{1 \leq i \leq n} D_i}} > n^{3/8} / \log n.$$

Using Lemma 2.12 in [23], we see that (recall that J_t is the exploration process of the configuration model starting from a uniformly randomly chosen vertex)

- J_t can be coupled with $\bar{W}_t^n$ with high probability up to time $n^{3/8} / \log n$.
- Whp the subgraph obtained by exploring the neighborhoods of $n^{3/8} / \log n$ vertices is a tree.

Using $n^{1/3} \log n \ll n^{3/8} / \log n$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(J_t = \bar{W}_t^n, 0 \leq t \leq n^{1/3} \log n \right) = 1. \quad (2.4.3)$$

To prove (2.2.2) it remains to show that one can couple W_t^n and W_t up to step $n^{1/3} \log n$. To this end, we use the characterization of total variation distance in terms of optimal coupling. It is well known that for any two random variables X and Y ,

$$d_{\text{TV}}(X, Y) = \inf_{\text{all couplings of } X, Y} \mathbb{P}(X \neq Y).$$

See [26, Theorem 1.14] for instance. Using $\mathbf{D}_n$ to denote the degree sequence $D_1, \dots, D_n$, it suffices to show that

$$d_{\text{TV}}^{\mathbf{D}_n}((\psi_0^n, \{\chi_r^n, 1 \leq r \leq n^{1/3} \log n\}), (\psi_0, \{\chi_r, 1 \leq r \leq n^{1/3} \log n\})) \xrightarrow{\mathbb{P}} 0. \quad (2.4.4)$$

Here the superscript $\mathbf{D}_n$ indicates that we are considering the quenched law of ψ_0^n and ξ_0^n . Since conditionally on $\mathbf{D}_n$, ψ_0^n and $\chi_r^n, r \geq 1$ are all independent,

$$\begin{aligned} & d_{\text{TV}}^{\mathbf{D}_n}((\psi_0^n, \{\chi_r^n, 1 \leq r \leq n^{1/3} \log n\}), (\psi_0, \{\chi_r, 1 \leq r \leq n^{1/3} \log n\})) \\ & \leq d_{\text{TV}}^{\mathbf{D}_n}(\psi_0^n, \psi_0) + \sum_{r=1}^{n^{1/3} \log n} d_{\text{TV}}^{\mathbf{D}_n}(\chi_r^n, \chi_r) \\ & \leq d_{\text{TV}}^{\mathbf{D}_n}(\psi_0^n, \psi_0) + (\log n) n^{1/3} d_{\text{TV}}^{\mathbf{D}_n}(\chi_1^n, \chi_1). \end{aligned} \quad (2.4.5)$$

We need the following lemma to control $\mathbf{D}_n$. Recall that we let $p_k = \mathbb{P}(D = k)$ and $m_1 = \mathbb{E}(D)$. Also recall that we assume $\mathbb{E}(D^5) < \infty$.

Lemma 2.6. For any $\epsilon > 0$,

$$\mathbb{P} \left(\left| \sum_{i=1}^n D_i - nm_1 \right| > n^{1/2+\epsilon} \right) \leq Cn^{-2\epsilon}. \quad (2.4.6)$$

Let N_k be the cardinality of the set $\{i : D_i = k\}$. We have

$$\mathbb{P} \left(|N_k - np_k| > n^\epsilon (np_k)^{1/2} \right) \leq n^{-2\epsilon}. \quad (2.4.7)$$

Proof of Lemma 2.6. Both (2.4.6) and (2.4.7) follow from Markov's inequality. For (2.4.6) we note that

$$\mathbb{P} \left(\left| \sum_{i=1}^n D_i - nm_1 \right| > n^{1/2+\epsilon} \right) \leq \frac{n \text{Var}(D_1)}{n^{1+2\epsilon}} \leq Cn^{-2\epsilon}.$$

For the second inequality, we note that $\mathbb{E}(1_{\{D_1=k\}} - p_k)^2 = p_k(1 - p_k) \leq p_k$. It follows that

$$\mathbb{E}(N_k - np_k)^2 = \mathbb{E} \left(\sum_{i=1}^n (1_{\{D_i=k\}} - p_k) \right)^2 = n \mathbb{E}(1_{\{D_1=k\}} - p_k)^2 \leq np_k.$$

Therefore

$$\mathbb{P}\left(|N_k - np_k| > n^\epsilon (np_k)^{1/2}\right) \leq \frac{\mathbb{E}(N_k - np_k)^2}{n^{2\epsilon} (np_k)} \leq n^{-2\epsilon}. \quad (2.4.8)$$

□

Let $\epsilon_1 = 1/100, \epsilon_2 = 1/8 + 1/100$. Consider the event

$$\Omega_n = \left\{ \max_{1 \leq i \leq n} D_i \leq n^{1/4} \log n, \left| \sum_{i=1}^n D_i - nm_1 \right| \leq n^{1/2+\epsilon_1} \right\} \quad (2.4.9)$$

$$\cap \left\{ |N_k - np_k| \leq n^{\epsilon_2} (np_k)^{1/2} \text{ for } 1 \leq k \leq n^{1/4} \log n \right\}.$$

Lemma 2.6 and equation (2.4.2) imply that $\mathbb{P}(\Omega_n) \rightarrow 1$ by the union bound. We now control $d_{\text{TV}}^{\text{Dn}}(\chi_1^n, \chi_1)$ on Ω_n . For $k \geq n^{1/4} \log n$, we have $\mathbb{P}_n(\chi_1^n = k) = 0$ on Ω_n almost surely. We also have $\sum_{k \geq n^{1/4} \log n} \mathbb{P}(\chi_1 = k) \leq n^{-1} \mathbb{E}(D^4) \leq Cn^{-1}$. Hence we have

$$\sum_{k \geq n^{1/4} \log n} |\mathbb{P}_n(\chi_1^n = k) - \mathbb{P}(\chi_1 = k)| \leq Cn^{-1}. \quad (2.4.10)$$

For $0 \leq k \leq n^{1/4} \log n$, using the definition of Ω_n we get, for n large,

$$\begin{aligned} |\mathbb{P}_n(\chi_1^n = k) - \mathbb{P}(\chi_1 = k)| &= \left| \frac{(k+1)N_{k+1}}{\sum_{i=1}^n D_i} - \frac{(k+1)p_{k+1}}{m_1} \right| \\ &\leq \left| \frac{(k+1)N_{k+1}}{\sum_{i=1}^n D_i} - \frac{n(k+1)p_{k+1}}{\sum_{i=1}^n D_i} \right| + \left| \frac{n(k+1)p_{k+1}}{\sum_{i=1}^n D_i} - \frac{(k+1)p_{k+1}}{m_1} \right| \\ &\leq \frac{(k+1)n^{\epsilon_2} (np_k)^{1/2}}{m_1 n/2} + \frac{(k+1)p_{k+1} |nm_1 - \sum_{i=1}^n D_i|}{nm_1 \sum_{i=1}^n D_i} \\ &\leq Cn^{\epsilon_2} n^{-1/2} (k+1)^{-1} + C(k+1)p_{k+1} n^{-1/2+\epsilon_1}, \end{aligned} \quad (2.4.11)$$

where we have used the fact that $p_k \leq Ck^{-4}$ since $\mathbb{E}(D^4) < \infty$ in the last step. Summing the last line of (2.4.11) over k from 0 to $n^{1/4} \log n$ and using the facts

$$\sum_{k=1}^n k^{-1} \sim \log n, \quad \sum_{k=0}^{\infty} (k+1)p_{k+1} = \mathbb{E}(D) < \infty,$$

we get, on Ω_n ,

$$\sum_{k=0}^{n^{1/4} \log n} |\mathbb{P}_n(\chi_1^n = k) - \mathbb{P}(\chi_1 = k)| \leq Cn^{\epsilon_2} n^{-1/2} \log n + Cn^{-1/2+\epsilon_1}. \quad (2.4.12)$$

Inserting the values of $\epsilon_1 = 1/100, \epsilon_2 = 1/8 + 1/100$,

$$\sum_{k=0}^{n^{1/4} \log n} |\mathbb{P}_n(\chi_1^n = k) - \mathbb{P}(\chi_1 = k)| \leq Cn^{1/100} \log n (n^{1/8-1/2} + n^{-1/2}). \quad (2.4.13)$$

Combining (2.4.10) and (2.4.13),

$$d_{\text{TV}}^{\text{Dn}}(\chi_1^n, \chi_1) = \frac{1}{2} \sum_{k \geq 0} |\mathbb{P}_n(\chi_1^n = k) - \mathbb{P}(\chi_1 = k)| \leq Cn^{-73/200} \log n. \quad (2.4.14)$$

One can similarly show that $Cn^{-73/200} \log n$ also serves as an upper bound for $d_{\text{TV}}^{\text{Dn}}(\psi_0^n, \psi_0)$. Thus on Ω_n

$$d_{\text{TV}}^{\text{Dn}}(\psi_0^n, \psi_0) + (\log n) n^{1/3} d_{\text{TV}}^{\text{Dn}}(\chi_1^n, \chi_1) \leq Cn^{-73/200+1/3} \log^2 n,$$

which converges to 0 as $n \rightarrow \infty$. This together with (2.4.5) implies (2.4.4) and thus completes the proof of Lemma 2.2.

□

3 Upper bound on evoSI

3.1 avoSI

As mentioned in the introduction, in this we will construct a model that serves as an upper bound for evoSI. We first introduce a model *C-avoSI* where ‘C’ stands for ‘coupled’ which means we couple the structure of the graph with the epidemic. The ‘avo’ stands for avoiding infection.

The C-avoSI process is constructed as follows. First recall that in the construction of the configuration model each vertex is assigned a random number of half-edges initially. In the beginning all half-edges attached to the n vertices are unpaired. The half-edges attached to infected nodes are called infected half-edges and those attached to susceptible nodes are susceptible half-edges. Recall that “randomly chosen” and “at random” mean that the distribution of the choice is uniform over the set of possibilities.

- At rate λ each infected half-edge pairs with a randomly chosen half-edge in the pool of all half-edges excluding itself. If the vertex y associated with that half-edge is susceptible then it becomes infected. Note that if vertex y changes from state S to I then all half-edges attached to y become infected half-edges.
- Each infected half-edge gets removed from the vertex that it is attached to at rate ρ and immediately becomes re-attached to a randomly chosen vertex in the pool of all vertices.

For the purpose of comparisons it is convenient to give a reformulation of C-avoSI where the graph has been constructed before the epidemic. We call this process *avoSI*. To describe this process we define the notion of ‘stable edge’. We say that an edge between two vertices x and y is *stable* if one of the following conditions hold:

- Both x and y are in state S .
- Either x or y has sent an infection to the other one through this edge.

We say that an edge is *unstable* if both conditions fail. Note that an $S - I$ pair is necessarily unstable since the vertex in state I has not sent an infection to the vertex in state S . We define the avoSI process as follows:

- Each infected vertex sends infections to its neighbors at rate λ . If the neighbor has already been infected then nothing changes. Once a vertex receives an infection, it stays infected forever.
- A vertex in state S will break its connection with a vertex in state I at rate ρ and rewire to another randomly chosen vertex. The events for different $S - I$ connections are independent.
- For every unstable $I - I$ edge, each of the two I 's will rewire at rate ρ to another uniformly chosen vertex. (This also explains why we call such edges ‘unstable’, since they may evolve.)

Lemma 3.1. *The C-avoSI process and avoSI process running on a configuration model have the same law in terms of the evolution of the set of infected vertices.*

Proof of Lemma 3.1. It suffices to construct a graph G that has the law of the configuration model such that the set of infected vertices in C-avoSI evolves in the same way as that of the avoSI with initial underlying graph being G . Given an outcome of C-avoSI, the graph G can be constructed as follows. Let H be the collection of vertices and half-edges.

- We assign a unique label to each half-edge in C-avoSI and correspondingly label the half-edges in H .

- Whenever two half-edges combine into one edge in C-evoSI process we pair the two half-edges with the same labels in H . It is clear that the pairings of half-edges are done at random. We pair the remaining half-edges at random after there is no infected half-edge in the system. This forms the graph G . Since the pairings of all half-edges are at random, we deduce that G itself has the law of $\text{CM}(n, D)$.
- Whenever a pairing occurs in C-avoSI, an infected vertex has sent an infection to one of its neighbors(s) in avoSI.
- Whenever an infected half-edge h attached to x rewires to another vertex y in the C-avoSI, the corresponding edge e in avoSI, which contains h as one of its two half-edges, breaks from x and reconnects to vertex y .

The process we construct on G has the same law as avoSI. Indeed, an infected half-edge can be rewired in C-avoSI if and only if it has not been paired. This exactly corresponds to the notion of ‘unstable’ that we used in the construction of avoSI process. Hence Lemma 3.1 follows. $\square$

From now on we will not distinguish between the C-avoSI and the avoSI since they are equivalent. The avoSI process stochastically dominates the evoSI process, as shown in the lemma below.

Lemma 3.2. *The final size of infected vertices in avoSI stochastically dominates that in evoSI.*

We couple the evoSI and avoSI as follows. The evoSI is constructed in the same manner as we did in the comparison of evoSI and delSI. See Section 2.1. We will use the variables $T_{e,\ell}, R_{e,\ell}, U_{e,\ell}, \ell \geq 1$ (defined at the beginning of Section 2.1) in the construction of avoSI to couple evoSI and avoSI.

Construction of avoSI. For avoSI we need four sets to partition the set of all edges. We use $\mathcal{A}$ to avoid confusion with the sets used in the construction of evoSI (not to be confused with the $\mathcal{A}_t$ used in Section 2).

- *Active edges with one end infected* at time t , denoted by $\mathcal{A}_t^{a,1}$, are the edges at time t that connect an infected vertex and a susceptible vertex.
- *Active edges with both ends infected*, denoted by $\mathcal{A}_t^{a,2}$, are the unstable edges at time t that connect two infected vertices.
- *Uninfected edges*, denoted by $\mathcal{A}_t^0$, connect two susceptible vertices.
- *Inactive edges*, denoted by $\mathcal{A}_t^i$, consist of stable infected edges. Once an edge becomes inactive it remains inactive forever.

We set $\mathcal{A}_t = \mathcal{A}_t^{a,1} \cup \mathcal{A}_t^{a,2}$. The four sets form a partition of all edges. They are right-continuous pure jump processes. At time 0 we randomly choose a vertex u_0 to be infected. $\mathcal{A}_0^{a,1}$ consists of the edges with one endpoint at u_0 . $\mathcal{A}_0^0$ is the collection of all edges in the graph minus the set $\mathcal{A}_0^{a,1}$. $\mathcal{A}_0^i = \mathcal{A}_0^{a,2} = \emptyset$.

For each undirected edge e , we let random variables $T_{e,\ell}, R_{e,\ell}, U_{e,\ell}$ be the same as those used in the construction of evoSI. We set $S_{e,\ell} = \min\{T_{e,\ell}, R_{e,\ell}\}$. We also let $V_{e,\ell}'$ be independent uniform random variables that take values in the two endpoints of e . To make it easier to describe the dynamics, suppose that at time $\tau_{e,\ell}^a$ (the ℓ -th time e becomes active in avoSI) we have $e = \{x_{e,\ell}, y_{e,\ell}\}$ with $x_{e,\ell}$ infected and $y_{e,\ell}$ susceptible.

The difference between evoSI and avoSI in terms of these clocks is as follows. For any edge e connecting x and y , once one endpoint $x_{e,\ell}$ becomes infected, the clocks $T_{e,\ell}$ and $R_{e,\ell}$ start running. If the other endpoint $y_{e,\ell}$ also becomes infected through other

edges at (relative) time $w_{e,\ell} < S_{e,\ell}$ (here ‘relative’ means we only count the time after the infection of x), then we replace the clocks $T_{e,\ell}$ and $R_{e,\ell}$ by

$$T'_{e,\ell} = \frac{T_{e,\ell} - w_{e,\ell}}{2} \text{ and } R'_{e,\ell} = \frac{R_{e,\ell} - w_{e,\ell}}{2}$$

since the rates are now twice as fast.

To see that this construction give the correct dynamics of avoSI. We note that conditionally on $T_{e,\ell}, R_{e,\ell} > w_{e,\ell}$, $(T_{e,\ell} - w_{e,\ell})/2$ and $(R_{e,\ell} - w_{e,\ell})/2$ are independent exponential random variables with parameters 2λ and 2ρ , respectively. This corresponds to an unstable $I - I$ pair where each of the two I 's attempt to send infection to the other I at rate λ and rewire from the other I at rate ρ . The variable $V'_{e,\ell}$ corresponds to the vertex with small rewiring time.

If $T_{e,\ell} > R_{e,\ell}$ then $T'_{e,\ell} > R'_{e,\ell}$ and vice versa. Using this and the fact that the uniform variable $U_{e,\ell}$ is the same in evoSI and avoSI we get the following proposition:

Proposition 1. If edge e is rewired in evoSI, then e must be rewired to the same vertex in avoSI (as long as either endpoint of e becomes infected in avoSI). Also, the time it take to e to be rewired in evoSI is not smaller than the time in avoSI.

We now formally construct the avoSI process by induction.

Initial step. At time 0, a randomly chosen vertex u_0 is infected. Let

$$\mathcal{N}^0(x, t), \mathcal{N}^{a,1}(x, t), \mathcal{N}^{a,2}(x, t), \mathcal{N}^i(x, t)$$

be the sets of edges connected to x that belong to the sets $\mathcal{A}_t^0, \mathcal{A}_t^{a,1}, \mathcal{A}_t^{a,2}$ and $\mathcal{A}_t^i$, respectively. At time 0 all edges in $\{e_{0,1}, \dots, e_{0,k}\}$ are added to the list of active edges so that $\mathcal{A}_0^{a,1} = \{e_{0,1}, \dots, e_{0,k}\}$. Suppose e_j connects u_0 and y_j . At time

$$J_1^a = \min_{1 \leq j \leq k} S_{e_{0,j},1}$$

the first event occurs. (The superscript ‘a’ indicates the avoSI model.) Let i be the index that achieves the minimum.

(i) If $R_{e_{0,i},1} < T_{e_{0,i},1}$, then at time J_1^a vertex y_i breaks its connection with u_0 and rewires to $U_{e_{0,i},1}$. If $U_{e_{0,i},1}$ is susceptible at time J_1^a , we move the edge $e_{0,i}$ to $\mathcal{A}_{J_1^a}^0$. On the initial step this will hold unless $U_{e_{0,i},1} = u_0$ in which case nothing has changed.

(ii) If $T_{e_{0,i},1} < R_{e_{0,i},1}$ then at time J_1^a vertex y_i becomes infected by u_0 . We move $e_{0,i}$ to $\mathcal{A}_{J_1^a}^i$. We move edges in $\mathcal{N}^0(y_i, J_1^a -)$ to $\mathcal{A}_{J_1^a}^{a,1}$.

Induction step. For any active edge e at time t , let $L^a(e, t) = \sup\{\ell : \tau_{e,\ell}^a \leq t\}$ and set

$$V^a(e, t) = \begin{cases} \tau_{e,L^a(e,t)} + S_{e,L^a(e,t)} & \text{if } e \in \mathcal{A}_t^{a,1}, \\ \frac{\tau_{e,L^a(e,t)} + I(x^*, y^*) + S_{e,L^a(e,t)}}{2} & \text{if } e \in \mathcal{A}_t^{a,2}, \end{cases}$$

where $I(x^*, y^*)$ is the first time that both $x^* = x(e, L^a(e, t))$ and $y^* = y(e, L^a(e, t))$ become active vertices and the second line comes from the computation

$$\begin{aligned} & \tau_{e,L^a(e,t)} + (I(x^*, y^*) - \tau_{e,L^a(e,t)}^a) + \frac{S_{e,L^a(e,t)} - (I(x^*, y^*) - \tau_{e,L^a(e,t)}^a)}{2} \\ &= \frac{\tau_{e,L^a(e,t)}^a + I(x^*, y^*) + S_{e,L^a(e,t)}}{2}. \end{aligned}$$

Then $V^a(e, t)$ is the time of the next event (infection or rewiring) to affect edge e . Suppose we have constructed the process up to time J_m . If there are no active edges present at time J_m , the construction is done. Otherwise, we let

$$J_{m+1}^a = \min_{e \in \mathcal{A}_{J_m}^a} V^a(e, J_m^a).$$

Let e_m be the edge that achieves the minimum of $V^a(e, J_m^a)$. If e_m only has one endpoint infected at time J_m^a then we let $x(e_m)$ be the infected endpoint of e_m and $y(e_m)$ be the susceptible endpoint of e_m . To simplify notation let $L_m = L(e_m, J_m^a)$.

(i) If $R_{e_m, L_m} < T_{e_m, L_m}$ and $e_m \in \mathcal{A}_{J_m^a}^{a,1}$, then at time J_{m+1}^a vertex $y(e_m)$ breaks its connection with $x(e_m)$ and rewires to U_{e_m, L_m} . If U_{e_m, L_m} is susceptible at time J_{m+1}^a , then e_m is moved to $\mathcal{A}_{J_{m+1}^a}^{a,1}$. Otherwise it remains in $\mathcal{A}_{J_{m+1}^a}^{a,1}$.

(ii) If $R_{e_m, L_m} < T_{e_m, L_m}$ and $e_m \in \mathcal{A}_{J_m^a}^{a,2}$, then at time J_{m+1}^a vertex V'_{e_m, L_m} breaks its connection with the other end of e_m and rewires to U_{e_m, L_m} . If U_{e_m, L_m} is susceptible at time J_{m+1}^a , then e_m is moved to $\mathcal{A}_{J_{m+1}^a}^{a,1}$. Otherwise e_m stays in the set $\mathcal{A}_{J_{m+1}^a}^{a,2}$.

(iii) If $T_{e_m, L_m} < R_{e_m, L_m}$ and $e_m \in \mathcal{A}_{J_m^a}^{a,1}$, then at time J_{m+1}^a the vertex $y(e_m)$ is infected by $x(e_m)$ and e_m is moved to $\mathcal{A}_{J_{m+1}^a}^{i,1}$.

- All edges e' in $\mathcal{N}_{y(e_m), J_{m+1}^a}^0$ are moved to $\mathcal{A}_{J_{m+1}^a}^{a,1}$. Since $y(e_m)$ has just become infected, the other end of e' must be susceptible at time J_{m+1}^a .
- All edges e'' in $\mathcal{N}_{y(e_m), J_{m+1}^a}^{a,1}$ are moved to $\mathcal{A}_{J_{m+1}^a}^{a,2}$. Since $y(e_m)$ has just become infected, (i) the other end of e'' must be infected at time J_{m+1}^a , and (ii) e'' cannot have been inactive earlier.

(iv) If $T_{e_m, L_m} < R_{e_m, L_m}$ and $e_m \in \mathcal{A}_{J_m^a}^{a,2}$, then at time J_{m+1}^a , e_m is moved to $\mathcal{A}_{J_{m+1}^a}^{i,2}$. There are no changes for other edges.

The avoSI process stops when there are no active edges.

Proof of Lemma 3.2. We now prove by induction that all vertices infected in evoSI are also infected in avoSI and actually they are infected earlier in avoSI than evoSI.

The induction hypothesis holds for the first vertex since initially u_0 is infected in both evoSI and avoSI. Suppose the induction holds up to the k -th infected vertex in evoSI. Assume at time t , y becomes the $(k+1)$ -th infected vertex in evoSI. We assume that y is infected by vertex x through edge e . Note that e has possibly gone through a series of rewirings before connecting vertex y . We assume that e connects vertices x_ℓ and y_ℓ after the $(\ell-1)$ -th rewiring. We also assume when x infects y through e , e has been rewired r times. This implies that

$$T_{e, \ell} > R_{e, \ell} \text{ for } 1 \leq \ell \leq r \text{ and } T_{e, r+1} < R_{e, r+1}. \quad (3.1.1)$$

We let $m(x) = \inf\{i : x \in \{x_k, y_k\} \text{ for all } i \leq k \leq r+1\}$. We now divide the analysis into two cases: $m = 1$ and $1 < m \leq r+1$.

Case 1. $m = 1$ so that $y = y_k$ for all $1 \leq k \leq r+1$. If we assume that y has not been infected by time t , then $x_1, \dots, x_r$ must have been infected at the time that the rewiring occurred and $x_{r+1} = x$ infected y at time t in evoSI. By the induction hypothesis we see that $x_1, \dots, x_{r+1}$ are also infected in avoSI. By Proposition 1 and (3.1.1), e breaks its connection with x_1 and reconnects to x_2 , then to x_3 and after r rewirings to x_{r+1} . If y is already infected before x_{r+1} sends an infection to it then we are done. Otherwise since $T_{e, \ell+1} < R_{e, \ell+1}$ we see that x_{r+1} will send an infection to y in avoSI as well. In any case we have proved that x will also be infected in avoSI and is infected earlier. For a picture see Figure 6.

Case 2. If $m > 1$, then again by Proposition 1, the induction hypothesis and (3.1.1) we see that in the avoSI picture, e will be rewired at least r times and y becomes an endpoint of e exactly after $m-1$ rewirings. After this point we can repeat the analysis in the case of $m = 1$ to deduce that x is also infected no later than t in avoSI. For a picture see Figure 7. $\square$

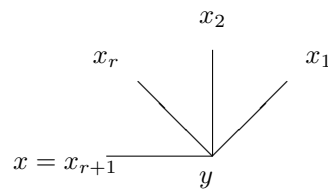


Figure 6: Illustration of the case $m = 1$. If y is infected before time t in avoSI then the induction step holds true. If not, then since $x_1, \dots, x_{r+1}$ are the rewirings in evoSI, they are also infected in avoSI and the rewiring occurs as does the infection of x by x_{r+1} .

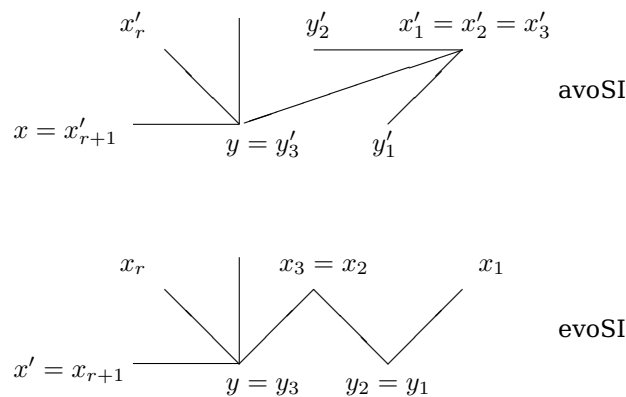


Figure 7: Illustration of the case $m > 1$. We assume that at the time of the rewiring x_1, y_1 and x_2 are infected. Since one flips coins to determine the end that rewires, the sequence of edges (x'_k, y'_k) in avoSI is different from the edges in evoSI. However, thanks to the use of $U_{e,\ell}$ to determine the new endpoint, the second rewiring brings the edge to y , and there is a correspondence between the vertices in the two processes indicated by the drawing.

We now show that avoSI and delSI actually have the same critical value (and thus also have the same critical value as evoSI by Theorem 1.6).

Lemma 3.3. *Theorem 1.6 still holds if we replace evoSI by avoSI.*

Proof. As we will see, Lemma 3.3 follows by repeating the proof of Theorem 1.6. Since avoSI stochastically dominates evoSI (in terms of final epidemic size) and evoSI dominates delSI, avoSI must also dominate delSI. We claim that if we run avoSI on a tree and no edge is rewired to vertices that are infected up to time t , then there are no unstable $I - I$ pairs up to time t . This immediately implies that the evolution of the avoSI process is equal to that of the evoSI process starting from the same initially infected vertex up to time t . To prove the claim, suppose there is an unstable $I - I$ pair connecting vertex x and y , then there must be two infection paths that lead to the infections of x and y . Here an infection path for x is just a sequence of vertices $u_0 \rightarrow u_1 \rightarrow \dots \rightarrow x$ where the

former vertex in the chain infects the latter. Since the edge between x and y is unstable, we see that if we consider the union of the two infection paths together with the edge (x, y) then we get a cycle of infected vertices. This contradicts with the assumption that the original graph is a tree and no edges are rewired to vertices infected by t (so that rewirings will not help create any cycle of infected vertices).

As we mentioned in the proof of Lemma 2.2, by Lemma 2.12 in [23], whp the subgraph obtained by exploring the neighborhoods of $n^{3/8}/\log n$ vertices of any fixed vertex is a tree. The proof of Theorem 1.6(i) implies that with high probability no edge is rewired to infected vertices up to step $O(n^{1/3})$ (i.e., up to the exploration of the neighborhoods of $O(n^{1/3})$ vertices). Hence the condition of the above claim is satisfied. Using the conclusion of the claim we see that whp we can couple the avoSI process and evoSI process such that they coincide up to step $O(n^{1/3})$. Therefore the proof of Theorem 1.6 also applies to the comparison of avoSI and delSI and hence Lemma 3.3 follows. $\square$

3.2 Tightness of $\{\bar{S}_{t,k}/n, t \geq 0\}_{n \geq 1}$

We first consider $\tilde{S}_{t,k}$, the number of susceptible vertices with k half-edges at time t in the original avoSI process (i.e., without the time change). We have the following equation

$$d\tilde{S}_{t,k} = - \left(\lambda \tilde{X}_{I,t} \frac{k \tilde{S}_{t,k}}{\tilde{X}_t - 1} \right) dt + \left(1_{\{k \geq 1\}} \rho \tilde{X}_{I,t} \frac{\tilde{S}_{t,k-1}}{n} \right) dt - \left(\rho \tilde{X}_{I,t} \frac{\tilde{S}_{t,k}}{n} \right) dt + d\tilde{M}_{t,k}. \quad (3.2.1)$$

To explain the terms

1. At rate $\lambda \tilde{X}_{I,t}$ infections occur. The infected half-edge attaches to a susceptible vertex with k half-edges, which we call an S^k , with probability $k \tilde{S}_{t,k}/(\tilde{X}_t - 1)$. The -1 in the denominator is because the half-edge will not connect to itself.
2. At rate $\rho \tilde{X}_{I,t}$ rewirings occur. If $k \geq 1$, the half-edge gets attached to an S^{k-1} with probability $(k-1) \tilde{S}_{t,k-1}/(\tilde{X}_t - 1)$, promoting it to an S^k .
3. If the rewired half-edge gets attached to an S^k , which occurs with probability $k \tilde{S}_{t,k}/(\tilde{X}_t - 1)$, it is promoted to an S^{k+1} and an S^k is lost.
4. If Z_t is a Markov chain with generator L then Dynkin's formula implies

$$f(Z_t) - \int_0^t Lf(Z_s) ds \text{ is a martingale.}$$

See Chapter 4, Proposition 1.7 in [8]. Fortunately, we do not need an explicit formula for the martingale. All that is important is that when $f(Z_t) = \tilde{S}_{t,k}$, $\tilde{M}_{t,k}$ has jumps equal to ± 1 .

The equation for $\bar{S}_{t,k}$ can be then obtained by multiplying the first three terms in the right hand side of (3.2.1) by the time change, leading to

$$\begin{aligned} d\bar{S}_{t,k} = & - \left(\lambda \bar{X}_{I,t} \frac{\bar{X}_t - 1}{\lambda \bar{X}_{I,t}} \frac{k \bar{S}_{t,k}}{\bar{X}_t - 1} \right) dt + \left(1_{\{k \geq 1\}} \rho \bar{X}_{I,t} \frac{\bar{X}_t - 1}{\lambda \bar{X}_{I,t}} \frac{\bar{S}_{t,k-1}}{n} \right) dt \\ & - \left(\rho \bar{X}_{I,t} \frac{\bar{X}_t - 1}{\lambda \bar{X}_{I,t}} \frac{\bar{S}_{t,k}}{n} \right) dt + d\bar{M}_{t,k}. \end{aligned} \quad (3.2.2)$$

Here $\bar{M}_{t,k}$ is a time-changed version of the previous martingale so it is also a martingale with jumps ± 1 .

Canceling common factors and dividing both sides of (3.2.2) by n we get

$$\begin{aligned} d\left(\frac{\bar{S}_{t,k}}{n}\right) &= -\left(k\frac{\bar{S}_{t,k}}{n}\right)dt + \left(1_{(k \geq 1)}\frac{\rho}{\lambda}\frac{\bar{X}_t - 1}{n}\frac{\bar{S}_{t,k-1}}{n}\right)dt \\ &\quad - \left(\frac{\rho}{\lambda}\frac{\bar{X}_t - 1}{n}\frac{\bar{S}_{t,k}}{n}\right)dt + d\left(\frac{\bar{M}_{t,k}}{n}\right). \end{aligned} \quad (3.2.3)$$

We now show that for all fixed $k \geq 0$ and $T > 0$,

$$\sup_{0 \leq t \leq T \wedge \gamma_n} |\bar{M}_{t,k}|/n \xrightarrow{\mathbb{P}} 0. \quad (3.2.4)$$

To do this we note that the expected value of quadratic variation of $\bar{M}_{t \wedge \gamma_n, k}$ evaluated at time T , which is also equal to $\mathbb{E}(\bar{M}_{T \wedge \gamma_n, k}^2)$, is bounded above by the expectation of total number of jumps in the whole avoSI process, which is equal to

$$\mathbb{E}\left(X_0/2 + \sum_{j=1}^{X_0} N_j\right).$$

Here we have a factor of 2 in the denominator because each pairing event takes two half-edges and N_j is the number of times that half-edge j gets transferred to another vertex. Note that an infected half-edge gets rewired before being paired with probability at most $\rho/(\lambda + \rho)$ and susceptible half-edge cannot get rewired unless the vertex it is attached to becomes infected. Thus, N_j is stochastically dominated by a Geometric($\rho/(\lambda + \rho)$) distributed random variable, so that for all j , $\mathbb{E}(N_j) \leq C$ for some constant C . Therefore by L^2 maximal inequality applied to the submartingale $|\bar{M}_{t \wedge \gamma_n, k}|$, we obtain that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T \wedge \gamma_n} \bar{M}_{t,k}^2\right) \leq 4\mathbb{E}(\bar{M}_{T \wedge \gamma_n, k}^2) \leq Cn, \quad (3.2.5)$$

where C is a constant whose value is unimportant.

Since $\bar{S}_{0,k}/n \leq 1$, we know $\{\bar{S}_{0,k}/n\}_{n \geq 1}$ is a tight sequence of random variables. To establish tightness of $\{\bar{S}_{t,k}/n, t \geq 0\}_{n \geq 1}$ we need to show for any fixed $\epsilon, \delta > 0$, there is a $\theta > 0$ and integer n_0 so that for $n \geq n_0$

$$\mathbb{P}\left(\sup_{|t_1 - t_2| \leq \theta, t_1, t_2 \leq T} |\bar{S}_{t_1 \wedge \gamma_n, k} - \bar{S}_{t_2 \wedge \gamma_n, k}|/n \geq \delta\right) \leq \epsilon. \quad (3.2.6)$$

Assuming (3.2.6) for the moment, we see that $\{\bar{S}_{t \wedge \gamma_n, k}/n, t \geq 0\}$, as an element of $\mathbb{D}$, the space of right continuous paths with left limits, satisfies condition (ii) of Proposition 3.26 in [10]. Consequently, $\{\bar{S}_{t \wedge \gamma_n, k}/n, t \geq 0\}_{n \geq 1}$ is a tight sequence. It remains to prove (3.2.6). We note that there exists a constant C so that

$$\mathbb{P}(\bar{X}_0/n > C) \leq \epsilon/3, \quad (3.2.7)$$

since $\mathbb{E}(\bar{X}_0) = \mathbb{E}(\sum_{i=1}^n D_i) = nm_1$. Hence using $\bar{S}_{t,k-1} + \bar{S}_{t,k} \leq n$ and $\bar{X}_t \leq \bar{X}_0$ with (3.2.3)

$$\begin{aligned} &\mathbb{P}\left(\sup_{|t_1 - t_2| \leq \theta, t_1, t_2 \leq T} |\bar{S}_{t_1 \wedge \gamma_n, k} - \bar{S}_{t_2 \wedge \gamma_n, k}|/n \geq \delta\right) \\ &\leq \mathbb{P}\left(k\theta > \frac{\delta}{4}\right) + 2\mathbb{P}\left(\frac{\rho}{\lambda}\frac{X_0}{n}\theta \geq \frac{\delta}{4}\right) + \mathbb{P}\left(2\sup_{0 \leq t \leq T \wedge \gamma_n} |\bar{M}_{t,k}/n| \geq \frac{\delta}{4}\right). \end{aligned}$$

Using (3.2.7) and (3.2.5), we see that if we pick θ small and n large then the last line is $\leq \epsilon$. This proves (3.2.6) and thus completes the proof of tightness of the sequence $\{\bar{S}_{t \wedge \gamma_n, k}/n, t \geq 0\}_{n \geq 1}$.

3.3 Convergence of $\{\bar{S}_{t,k}/n, t \geq 0\}_{n \geq 1}$

Note that the evolution for $\bar{X}_t$ has the same transition rates as the time-changed SIR dynamics defined in [13] so their equation (3.4) also holds true in avoSI, which gives for any fixed T ,

$$\sup_{0 \leq t \leq T \wedge \gamma_n} \left| \frac{\bar{X}_t}{n} - m_1 \exp(-2t) \right| \xrightarrow{\mathbb{P}} 0. \quad (3.3.1)$$

In order to upgrade $\sup_{0 \leq t \leq T \wedge \gamma_n}$ to $\sup_{0 \leq t \leq \gamma_n}$, we note that for any $\epsilon > 0$, we can pick a sufficiently large T so that $m_1 \exp(-2T) < \epsilon$. Then by the monotonically decreasing property of $\bar{X}_t$,

$$\begin{aligned} \mathbb{P} \left(\gamma_n > T, \sup_{T < t \leq \gamma_n} \left| \frac{\bar{X}_t}{n} - m_1 \exp(-2t) \right| > 3\epsilon \right) &\leq \mathbb{P}(\gamma_n > T, \bar{X}_T/n > 2\epsilon) \\ &\leq \mathbb{P} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \left| \frac{\bar{X}_t}{n} - m_1 \exp(-2t) \right| > \epsilon \right) \leq \epsilon, \end{aligned} \quad (3.3.2)$$

for n large enough. Thus we deduce

$$\mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n} \left| \frac{\bar{X}_t}{n} - m_1 \exp(-2t) \right| > \epsilon \right) \leq 2\epsilon$$

for n sufficiently large, which proves the first equation of (1.6.2).

By the tightness of $\{\bar{S}_{t \wedge \gamma_n, k}/n, t \geq 0\}_{n \geq 1}$, we see for any subsequence of $\bar{S}_{t,k}/n$ we can extract a further subsequence that converges in distribution to a process $\bar{s}_{t,k}$ with continuous sample path. By the Skorokhod representation theorem we can assume the convergence is actually in the almost sure sense and we can also assume that $\bar{X}_t/n$ converges a.s. to $m_1 \exp(-2t)$. Having established tightness, a standard argument implies that we can show the convergence of $\bar{S}_{t,k}/n$ by establishing that the limit $\bar{s}_{t,k}$ is independent of the subsequence. First consider the case $k = 0$. The first two terms on the right-hand side of (3.2.3) are 0, so using (3.2.4) and first equation of (1.6.2) we see that any subsequential limit $\bar{s}_{t,0}$ has to satisfy the equation

$$\bar{s}_{t,0} = -\frac{\rho}{\lambda} m_1 \int_0^t \exp(-2z) \bar{s}_{z,0} dz. \quad (3.3.3)$$

Since $z \rightarrow \exp(-2z)$ is Lipschitz continuous this equation has a unique solution.

Repeating this process for $k \geq 1$ we see that any subsequential limit $\bar{s}_{t,k}$ of $\bar{S}_{t,k}/n$ satisfies the differential equation

$$\bar{s}'_{t,k} = -k \bar{s}_{t,k} - \alpha \exp(-2t) \bar{s}_{t,k} + 1_{\{k \geq 1\}} \alpha \exp(-2t) \bar{s}_{t,k-1}, \quad (3.3.4)$$

where $\alpha = \rho m_1 / \lambda$. This system of equations can be solved explicitly. First we rewrite the equations as

$$\bar{s}'_{t,k} + [k + \alpha \exp(-2t)] \bar{s}_{t,k} = 1_{\{k \geq 1\}} \alpha \exp(-2t) \bar{s}_{t,k-1}.$$

Define

$$g_{t,k} = \exp(kt + (\alpha/2)(1 - \exp(-2t))) \bar{s}_{t,k}, \quad (3.3.5)$$

then

$$\begin{aligned} g'_{t,0} &= \alpha \exp(-2t) g_{t,0} + \exp((\alpha/2)(1 - \exp(-2t))) \bar{s}'_{t,0} \\ &= \alpha \exp(-2t) \exp((\alpha/2)(1 - \exp(-2t))) \bar{s}_{t,0} \\ &\quad + \exp(\alpha/2)(1 - \exp(-2t)) [-\alpha \exp(-2t)] \bar{s}_{t,0} = 0. \end{aligned}$$

Let $A_{k,t} = \exp(kt + (\alpha/2)(1 - \exp(-2t)))$. An almost identical calculation for $k \geq 1$ gives

$$\begin{aligned} g'_{t,k} &= [k + \alpha \exp(-2t)]g_{t,k} + A_{k,t}\bar{s}'_{t,k} \\ &= [k + \alpha \exp(-2t)]A_{t,k}\bar{s}_{t,k} \\ &\quad + A_{t,k}\{[-k - \alpha \exp(-2t)]\bar{s}_{t,k} + \alpha \exp(-2t)\bar{s}_{t,k-1}\}, \end{aligned}$$

so we have

$$g'_{t,k} = A_{t,k}\alpha \exp(-2t)\bar{s}_{t,k-1} = \alpha \exp(-t)g_{t,k-1}.$$

Making the change of variable $s = \alpha(1 - \exp(-t))$ and letting $h_{s,k} = g_{t,k}$, we see that $h_{s,0}$ is constant in s and

$$h'_{s,k} = h_{s,k-1}, k \geq 1,$$

from which we see $h_{s,k}$ is a polynomial of degree k in s and for all $\ell \leq k$ the ℓ -th derivative of $h_{s,k}$ at $s = 0$ equals $h_{0,k-\ell}$. From this we obtain that for all k ,

$$h_{s,k} = \sum_{\ell=0}^k \frac{h_{0,k-\ell}}{\ell!} s^\ell.$$

The initial conditions are $g_{0,k} = h_{0,k} = s_{0,k} = p_k = \mathbb{P}(D = k)$. It follows that

$$h_{s,k} = \sum_{\ell=0}^k \frac{p_{k-\ell} s^\ell}{\ell!}, \quad (3.3.6)$$

and hence using definitions of $\bar{s}$, g , and h

$$g_{t,k} = h_{s,k} = \sum_{\ell=0}^k \frac{p_{k-\ell}}{\ell!} (\alpha(1-w))^k, \quad (3.3.7)$$

$$\bar{s}_{t,k} = \exp\left(-\frac{\alpha}{2}(1-w^2)\right) w^k \sum_{\ell=0}^k \frac{p_{k-\ell}}{\ell!} (\alpha(1-w))^\ell, \quad (3.3.8)$$

where $w = w(t) = \exp(-t)$.

3.4 Summing the $\bar{s}_{t,k}$

We pause to record the following fact which we will use later. From the explicit expression for $\bar{s}_{t,k}$ in (3.3.8), dropping the factor $\exp(-\alpha(1-w^2)/2) \leq 1$ and writing $k = (k-\ell) + \ell$,

$$\begin{aligned} \sup_{t \geq 0} \sum_{k \geq K} k \bar{s}_{t,k} &\leq \sup_{0 \leq w \leq 1} \sum_{k \geq K} \sum_{\ell=0}^k \frac{(k-\ell)p_{k-\ell}}{\ell!} (\alpha(1-w))^\ell w^k, \\ &\quad + \sup_{0 \leq w \leq 1} \sum_{k \geq K} \sum_{\ell=1}^k \frac{\ell p_{k-\ell}}{\ell!} (\alpha(1-w))^\ell w^k. \end{aligned} \quad (3.4.1)$$

The $\ell = 0$ term in the first sum is bounded by

$$\sum_{k \geq K} k p_k.$$

The remainder of the two sums is bounded by

$$\sup_w w^K (1-w) \left[\sum_{k \geq \ell, \ell \geq 1} \left(\frac{\alpha^\ell (k-\ell)p_{k-\ell}}{\ell!} + \frac{\alpha^\ell \ell p_{k-\ell}}{\ell!} \right) \right]. \quad (3.4.2)$$

Interchanging the order of summation in the double sum and letting $m = k - \ell$, the double sum in (3.4.2) is bounded by

$$\sum_{\ell=1}^{\infty} \ell \cdot \frac{\alpha^\ell}{\ell!} \left(\sum_{m=0}^{\infty} (mp_m + p_m) \right) \leq C_{\alpha,D}.$$

Combining our calculations leads to

$$\limsup_{K \rightarrow \infty} \sup_{t \geq 0} \sum_{k \geq K} k \bar{s}_{t,k} \leq C \limsup_{K \rightarrow \infty} \left(\sum_{k \geq K} kp_k + \sup_w w^K (1 - w) \right) = 0. \quad (3.4.3)$$

We can use this bound to show that $\sum_{k=0}^{\infty} \bar{s}_{t,k}/n$ converges to $\sum_{k=0}^{\infty} \bar{s}_{t,k}$ as well as $\sum_{k=0}^{\infty} k \bar{s}_{t,k}/n$ converges to $\sum_{k=0}^{\infty} k \bar{s}_{t,k}$. Since the proofs are similar, we only prove the second result. We fix a large number K and observe that $\sum_{k \geq K} k \bar{s}_{t,k}$ satisfies the equation

$$\begin{aligned} d \left(\sum_{k \geq K} k \bar{s}_{t,k} \right) &= - \left(\sum_{k \geq K} k^2 \bar{s}_{t,k} \right) dt + \rho \bar{X}_{I,t} \frac{X_t - 1}{\lambda \bar{X}_{I,t}} \frac{K \bar{s}_{t,K-1}}{n} dt \\ &\quad + \sum_{k \geq K} \rho \bar{X}_{I,t} \frac{\bar{X}_t - 1}{\lambda \bar{X}_{I,t}} \frac{\bar{s}_{t,k}}{n} dt + d\widehat{M}_{t,K}. \end{aligned} \quad (3.4.4)$$

Here $\widehat{M}_{t,K}$ is a martingale that satisfies

$$\sum_{t > 0} (\widehat{M}_{t,k} - \widehat{M}_{t-,k})^2 \leq 3 \sum_{\ell=1}^n Q_\ell^2, \quad (3.4.5)$$

where Q_ℓ is the number of half-edges that vertex ℓ has before it becomes infected. This follows from the observation that there are two sources for the jump of $\sum_{k \geq K} k \bar{s}_{t,k}$:

- A susceptible vertex ℓ with at least K half-edges gets infected. Then $\sum_{k \geq K} k \bar{s}_{t,k}$ drops by the number of half-edges of vertex ℓ , which is Q_ℓ . Each vertex can contribute to this type of jumps at most once.
- A half-edge of an infected vertex gets transferred to a susceptible vertex of degree at least $K - 1$. Then $\sum_{k \geq K} k \bar{s}_{t,k}$ increases by either $\leq K$ (if the vertex gaining a half-edge had $K - 1$ half-edges before) or 1 (if the vertex gaining a half-edge had at least K half-edges before). Vertex ℓ can contribute at most Q_ℓ times to jumps of size 1 and at most once to jumps of size K .

Initially vertex ℓ has D_ℓ half-edges. As time grows the half-edges of other vertices might be transferred to vertex ℓ , the number of which is dominated by

$$V = \text{Binomial}(W, 1/n) \quad \text{where } W = \sum_{m=1}^n D_m. \quad (3.4.6)$$

Thus we have

$$\mathbb{E}[Q_\ell^2] \leq \mathbb{E}[(D_\ell + V)^2] \leq 2\mathbb{E}D^2 + 2\mathbb{E}V^2. \quad (3.4.7)$$

Conditioning on the value of W we have

$$\mathbb{E}V^2 = (1/n)(1 - 1/n)\mathbb{E}W + \mathbb{E}(W/n)^2 \leq C. \quad (3.4.8)$$

It follows from (3.4.5) and (3.4.8) that

$$\mathbb{E} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} |\widehat{M}_{t,k}|^2 \right) \leq 4\mathbb{E}(\widehat{M}_{T \wedge \gamma_n, k}^2) \leq Cn. \quad (3.4.9)$$

Writing (3.4.4) as an integral equation and dropping the negative term $-(\sum_{k \geq K} k^2 \bar{S}_{t,k})$ we see

$$\begin{aligned} \sum_{k \geq K} k \bar{S}_{z \wedge \gamma_n, k} &\leq \sum_{k \geq K} k \bar{S}_{0,k} + \frac{\rho}{\lambda} K \int_0^z \frac{\bar{X}_{u \wedge \gamma_n}}{n} \bar{S}_{u \wedge \gamma_n, K-1} du \\ &\quad + \frac{\rho}{\lambda} \int_0^z \frac{\bar{X}_{u \wedge \gamma_n}}{n} \sum_{k \geq K} \bar{S}_{u \wedge \gamma_n, k} + \widehat{M}_{z \wedge \gamma_n, k}. \end{aligned}$$

Take $\sup_{0 \leq z \leq t}$ on both sides we have

$$\begin{aligned} \sup_{0 \leq z \leq t \wedge \gamma_n} \sum_{k \geq K} k \bar{S}_{z,k} &\leq \sum_{k \geq K} k \bar{S}_{0,k} + \frac{\rho}{\lambda} K \int_0^t \sup_{0 \leq u \leq z \wedge \gamma_n} \frac{\bar{X}_u}{n} \bar{S}_{u, K-1} dz \\ &\quad + \frac{\rho}{\lambda} \int_0^t \sup_{0 \leq u \leq z \wedge \gamma_n} \frac{\bar{X}_u}{n} \sum_{k \geq K} \bar{S}_{u,k} dz + \sup_{0 \leq z \leq t \wedge \gamma_n} \widehat{M}_{z,k}. \end{aligned} \quad (3.4.10)$$

Dividing both sides of (3.4.10) by n , taking the square and using $(a + b + c + d)^2 \leq 4(a^2 + b^2 + c^2 + d^2)$ we have

$$\begin{aligned} \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \frac{\sum_{k \geq K} k \bar{S}_{z,k}}{n} \right)^2 &\leq 4 \left(\sum_{k \geq K} \frac{k \bar{S}_{0,k}}{n} \right)^2 + 4 \left(\frac{\rho}{\lambda} K \right)^2 t \int_0^t \left(\sup_{0 \leq u \leq z \wedge \gamma_n} \frac{\bar{X}_u}{n} \bar{S}_{u, K-1} \right)^2 dz \\ &\quad + 4 \left(\frac{\rho}{\lambda} \right)^2 t \int_0^t \left(\sup_{0 \leq u \leq z \wedge \gamma_n} \frac{\bar{X}_u}{n} \sum_{k \geq K} \bar{S}_{u,k} \right)^2 dz + 4 \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \widehat{M}_{z,k} \right)^2, \end{aligned} \quad (3.4.11)$$

where we have also used the Cauchy-Schwarz inequality to conclude that for any function $g = g(u)$,

$$\left(\int_0^t \left[\sup_{0 \leq u \leq z \wedge \gamma_n} g(u) \right] dz \right)^2 \leq t \int_0^t \left[\sup_{0 \leq u \leq z \wedge \gamma_n} g^2(u) \right] dz.$$

If we use $\widehat{\mathbb{E}}$ to denote the conditional expectation with respect to the σ -algebra generated by $\bar{X}_0$, then for any $\epsilon > 0$, using equation (3.4.9) we can find a constant (depending on ϵ) $C_{3.4.12} > 0$ so that

$$\mathbb{P} \left(\widehat{\mathbb{E}} \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \widehat{M}_{z,k}^2 \right) > C_{3.4.12} n \right) \leq \epsilon. \quad (3.4.12)$$

We then take another constant $C_{3.4.13}$ such that

$$\mathbb{P}(\bar{X}_0/n > C_{3.4.13}) \leq \epsilon. \quad (3.4.13)$$

Taking the conditional expectation of (3.4.11) with respect to $\bar{X}_0$ we see on the event

$$\Omega_0 = \{\bar{X}_0/n \leq C_{3.4.13}\} \cap \left\{ \widehat{\mathbb{E}} \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \widehat{M}_{z,k}^2 \right) \leq C_{3.4.12} n \right\}$$

which has probability $\geq 1 - 2\epsilon$, there exists a constant $C_{3.4.14}$ such that

$$\begin{aligned} \widehat{\mathbb{E}} \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \frac{\sum_{k \geq K} k \bar{S}_{z,k}}{n} \right)^2 &\leq C_{3.4.14} t \int_0^t \widehat{\mathbb{E}} \left[\left(\sup_{0 \leq u \leq z \wedge \gamma_n} \frac{\sum_{k \geq K} k \bar{S}_{u,k}}{n} \right)^2 \right] dz \\ &+ C_{3.4.14} \left(K^2 t \widehat{\mathbb{E}} \left[\left(\sup_{0 \leq z \leq t \wedge \gamma_n} \frac{S_{z,K-1}}{n} \right)^2 \right] + \frac{1}{n} + \widehat{\mathbb{E}} \left[\left(\sum_{k \geq K} \frac{k \bar{S}_{0,k}}{n} \right)^2 \right] \right). \end{aligned} \quad (3.4.14)$$

If we let

$$\phi(t) = \widehat{\mathbb{E}} \left(\sup_{0 \leq z \leq t \wedge \gamma_n} \frac{\sum_{k \geq K} k \bar{S}_{z,k}}{n} \right)^2,$$

and B equals the second line in (3.4.14), then for $0 \leq t \leq T$ we have

$$\phi(t) \leq C_{3.4.14} T \int_0^t \phi(s) ds + B.$$

Gronwall's inequality gives

$$\text{if } \phi(t) \leq \alpha(t) + \int_0^t \beta(s) \phi(s) ds, \text{ then } \phi(t) \leq \alpha(t) \exp \left(\int_0^t \beta(s) ds \right), \quad (3.4.15)$$

provided that $\beta(t) \geq 0$ and $\alpha(t)$ is nondecreasing. So applying (3.4.15) we have

$$\phi(t) \leq B \exp(C_{3.4.14} T^2)$$

on Ω_0 , that is,

$$\begin{aligned} \widehat{\mathbb{E}} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k S_{t,k}}{n} \right)^2 \\ \leq C_{3.4.14} \exp(C_{3.4.14} T^2) \left(K^2 t \widehat{\mathbb{E}} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{S_{t,K-1}}{n} \right)^2 + \frac{1}{n} + \widehat{\mathbb{E}} \left(\sum_{k \geq K} \frac{k \bar{S}_{0,k}}{n} \right)^2 \right). \end{aligned} \quad (3.4.16)$$

To control the first term on the right we use the convergence of $\bar{S}_{t,K-1}/n$ to $\bar{s}_{t,K-1}$ in probability as well as the bounded convergence theorem (since $\bar{S}_{t,K-1}/n \leq 1$) to obtain that

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\bar{S}_{t,K-1}}{n} \right)^2 \leq \left(\sup_{0 \leq t \leq T} \bar{s}_{t,K-1} \right)^2. \quad (3.4.17)$$

Using (3.4.17) and (3.4.3), if we first pick a large K , then for all n sufficiently large all three terms on right hand side of (3.4.16) smaller than $\epsilon^2/3$ with probability at least $1 - \epsilon$, which in turn implies that there is a set Ω_1 with $\mathbb{P}(\Omega_1) \geq 1 - 3\epsilon$, so that on Ω_1

$$\widehat{\mathbb{E}} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k S_{t,k}}{n} \right)^2 \leq \epsilon^2.$$

It follows that

$$\mathbb{E} \left(1_{\Omega_1} \widehat{\mathbb{E}} \left(\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k S_{t,k}}{n} \right)^2 \right) = \mathbb{E} \left(1_{\Omega_1} \sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k S_{t,k}}{n} \right)^2 \leq \epsilon^2.$$

Using $\mathbb{P}(\Omega_1) \geq 1 - 3\epsilon$ and the Chebyshev's inequality, we see that with probability $\geq 1 - 4\epsilon$,

$$\sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k S_{t,k}}{n} \leq \epsilon.$$

Fixing t and ϵ and using the triangle inequality, we get

$$\begin{aligned} \sup_{0 \leq t \leq T \wedge \gamma_n} \left| \frac{\sum_{k \geq 0} k \bar{S}_{t,k}}{n} - \sum_{k \geq 0} k \bar{s}_{t,k} \right| &\leq \sup_{0 \leq t \leq T \wedge \gamma_n} \left| \frac{\sum_{k=0}^K k \bar{S}_{t,k}}{n} - \sum_{k=0}^K k \bar{s}_{t,k} \right| \\ &+ \sup_{0 \leq t \leq T \wedge \gamma_n} \frac{\sum_{k \geq K} k \bar{S}_{t,k}}{n} + \sup_{0 \leq t \leq T} \sum_{k \geq K} k \bar{s}_{t,k}. \end{aligned} \quad (3.4.18)$$

By first choosing K large enough and then n large enough we can make both the first and second terms on the right hand side of (3.4.18) smaller than ϵ with probability at least $1 - 4\epsilon$. The third term can also be made smaller than ϵ using (3.4.3).

Since ϵ is arbitrary we see that

$$\sup_{0 \leq t \leq T \wedge \gamma_n} \left| \frac{\sum_{k=0}^{\infty} k \bar{S}_{t,k}}{n} - \sum_{k=0}^{\infty} k \bar{s}_{t,k} \right| \xrightarrow{\mathbb{P}} 0.$$

To find $\sum_{k=0}^{\infty} \bar{s}_{t,k}$ and $\sum_{k=0}^{\infty} k \bar{s}_{t,k}$, recall that we set $w = w(t) = \exp(-t)$ and $G(w) = \mathbb{E}(w^D)$. The limit of the fraction of susceptible nodes $\bar{s}_t$ satisfies

$$\begin{aligned} \bar{s}_t &= \sum_{k=0}^{\infty} \bar{s}_{t,k} = \exp(-(\alpha/2)(1-w^2)) \sum_{r,\ell \geq 0} \frac{p_r}{\ell!} (\alpha(1-w))^\ell w^{r+\ell} \\ &= \exp(-(\alpha/2)(1-w^2)) G(w). \end{aligned} \quad (3.4.19)$$

The limit of (scaled) number of susceptible half-edges satisfies

$$\bar{x}_{S,t} = \sum_{k=0}^{\infty} k \bar{s}_{t,k} = \exp(-\frac{\alpha}{2}(1-w^2)) \sum_{r,\ell \geq 0} (r+\ell) \frac{p_r}{\ell!} (\alpha(1-w))^\ell w^{r+\ell}.$$

The double sum equals

$$w \sum_{r \geq 0} r p_r w^{r-1} \sum_{\ell \geq 0} \frac{(\alpha(1-w)w)^\ell}{\ell!} + \alpha w(1-w) \sum_{r \geq 0} p_r w^r \sum_{\ell \geq 1} \frac{(\alpha(1-w)w)^{\ell-1}}{(\ell-1)!},$$

so we have

$$\bar{x}_{S,t} = \exp(-(\alpha/2)(1-w^2)) w(G'(w) + \alpha(1-w)G(w)). \quad (3.4.20)$$

Extension to time γ_n . We have proved the second and third statements of (1.6.2) with $0 \leq t \leq \gamma_n$ replaced by $0 \leq t \leq T \wedge \gamma_n$ for any fixed T . To upgrade this to $0 \leq t \leq \gamma_n$, note that $\sum_{k=0}^{\infty} k \bar{S}_{t,k} \leq \bar{X}_t$. Picking a large T satisfying $(m_1 + \alpha) \exp(-T) \leq \epsilon$ and re-using equation (3.3.2) we obtain that

$$\begin{aligned} &\mathbb{P} \left(\gamma_n > T, \sup_{T \leq t \leq \gamma_n} \left| \frac{\sum_{k=0}^{\infty} k \bar{S}_{t,k}}{n} - \exp \left(-\frac{\alpha}{2}(w-1)^2 \right) w(G'(w) + \alpha(1-w)G(w)) \right| \geq 3\epsilon \right) \\ &\leq \mathbb{P} \left(\gamma_n > T, \sup_{T \leq t \leq \gamma_n} \sum_{k=0}^{\infty} k \bar{S}_{t,k} / n > 2\epsilon \right) \leq \mathbb{P}(\gamma_n > T, \bar{X}_T / n > 2\epsilon) \leq \epsilon. \end{aligned} \quad (3.4.21)$$

This proves the third equation of (1.6.2). The proof of the second equation in (1.6.2) is slightly more complicated. Again we fix a large T such that $(m_1 + \alpha) \exp(-T) \leq \epsilon$ and

$$\left| \exp(-\alpha/2)G(0) - \exp\left(-\alpha/2(\exp(-T) - 1)^2\right)G(\exp(-T)) \right| \leq \epsilon. \quad (3.4.22)$$

Equation (3.4.22) and the fact that $\exp\left(-\alpha/2(\exp(-t) - 1)^2\right)G(\exp(-t))$ is decreasing in t imply that

$$\sup_{t, t' > T} \left| \exp\left(-\alpha/2(\exp(-t) - 1)^2\right)G(\exp(-t)) - \exp\left(-\alpha/2(\exp(-t') - 1)^2\right)G(\exp(-t')) \right| \leq \epsilon. \quad (3.4.23)$$

We first estimate the term $\bar{S}_{T,0}$ for large T . Using the weaker version (i.e., with $\sup_{0 \leq t \leq T \wedge \gamma_n}$) of the second equation of (1.6.2) we see that for n large enough,

$$\mathbb{P}\left(\gamma_n > T, \left| \bar{S}_{T,0}/n - \exp\left(-\alpha/2(\exp(-T) - 1)^2\right)G(\exp(-T)) \right| > \epsilon\right) \leq \epsilon. \quad (3.4.24)$$

On the event $\{\gamma_n > T\}$, $\sup_{T \leq t \leq \gamma_n} |\bar{S}_{t,0} - \bar{S}_{T,0}|$ can be bounded by $\bar{X}_T$, since in order to lose a susceptible vertex of degree 0 there must be a half-edge transferred to it. It follows that

$$\mathbb{P}\left(\gamma_n > T, \sup_{T \leq t \leq \gamma_n} |\bar{S}_{t,0} - \bar{S}_{T,0}|/n > 2\epsilon\right) \leq \mathbb{P}\left(\gamma_n > T, \bar{X}_T/n > 2\epsilon\right) \leq \epsilon. \quad (3.4.25)$$

Combining equations (3.4.24), (3.4.25) and (3.4.23) we see that

$$\mathbb{P}\left(\gamma_n > T, \sup_{T < t \leq \gamma_n} \left| \bar{S}_{t,0}/n - \exp\left(-\alpha/2(\exp(-t) - 1)^2\right)G(\exp(-t)) \right| > 4\epsilon\right) \leq 3\epsilon. \quad (3.4.26)$$

Using $\sum_{k \geq 1} \bar{S}_{t,k} \leq \sum_{k=0}^{\infty} k \bar{S}_{t,k}$ and equation (3.4.21) we see that

$$\mathbb{P}\left(\gamma_n > T, \sup_{T \leq t \leq \gamma_n} \sum_{k \geq 1} \bar{S}_{t,k}/n > 2\epsilon\right) \leq \epsilon. \quad (3.4.27)$$

It follows from (3.4.27) and (3.4.26) that

$$\mathbb{P}\left(\gamma_n > T, \sup_{T < t \leq \gamma_n} \left| \sum_{k=0}^{\infty} \bar{S}_{t,k}/n - a_S \exp\left(-\alpha/2(w(t) - 1)^2\right)G(w(t)) \right| > 6\epsilon\right) \leq 4\epsilon,$$

which proves the second equation of (1.6.2) and concludes the proof of (1.6.2).

3.5 Proof of Theorem 1.13

Recall for all $t \leq \gamma_n$ we have $\bar{X}_t \geq \bar{X}_{S,t}$ and at γ_n we have $\bar{X}_t = \bar{X}_{S,t}$ since γ_n is the time that we run out of infected half-edges and the dynamics stop. Note that by the definition of f in (1.5.1) we have

$$\exp(f(w)) = \frac{\bar{x}_t}{\bar{x}_{S,t}}.$$

We can rewrite f as

$$f(w) = \log(m_1) + \log(w) - \log(G'(w) + \alpha(1-w)G(w)) + \frac{\alpha}{2}(w-1)^2.$$

Taking the derivative we have

$$f'(w) = \frac{1}{w} - \frac{G''(w) - \alpha G(w) + \alpha(1-w)G'(w)}{G'(w) + \alpha(1-w)G(w)} + \alpha(w-1). \quad (3.5.1)$$

To evaluate $f'(1)$, recall that $G(1) = 1$, $G'(1) = m_1$ and $G''(1) = \mathbb{E}[D(D-1)] = m_2 - m_1$. The terms with $1-w$ vanish at $w = 1$. Using $\alpha = \rho m_1 / \lambda$ from (1.2.6), it follows that

$$f'(1) = 1 - \frac{m_2 - m_1 - \rho m_1 / \lambda}{m_1} = - \left(\frac{m_2 - 2m_1}{m_1} - \frac{\rho}{\lambda} \right).$$

Theorem 1.6 tells us that in the supercritical case, we have $\lambda > (\rho m_1) / (m_2 - 2m_1)$, and hence $f'(1) < 0$, which implies that f is positive on $(1 - \delta, 1)$ for some $\delta > 0$. Theorem 1.6 also shows that when $\eta > 0$ is small,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\bar{I}_\infty / n > \eta) = q(\lambda) > 0, \quad (3.5.2)$$

where $q(\lambda)$ is the survival probability of the two-phase branching process $\bar{Z}_m$ (defined in Section 1.2). Let $t_\eta < \delta$ be some small number depending on η such that

$$1 - \exp\left(-\frac{\alpha}{2}(\exp(-t_\eta) - 1)^2\right) G(\exp(-t_\eta)) < \eta/4. \quad (3.5.3)$$

Conditionally on $\bar{I}_\infty / n > \eta$ for some small η , the second equation of (1.6.2) implies that with high probability γ_n is also bounded from below by t_η (depending on η), since otherwise we would have

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{P}(\gamma_n < t_\eta, \bar{I}_{\gamma_n} / n > \eta) \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{P}\left(\gamma_n < t_\eta, \bar{S}_{\gamma_n} / n > \exp\left(-\frac{\alpha}{2}(\exp(-\gamma_n) - 1)^2\right) G(\exp(-\gamma_n)) - \eta/4, \bar{I}_{\gamma_n} / n > \eta\right) \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{P}\left(\bar{S}_{\gamma_n} / n > \exp\left(-\frac{\alpha}{2}(\exp(-t_\eta) - 1)^2\right) G(\exp(-t_\eta)) - \eta/4, \bar{I}_{\gamma_n} / n > \eta\right) = 0, \end{aligned}$$

where the last equality is due to (3.5.3) and the fact that $\bar{S}_{\gamma_n} + \bar{I}_{\gamma_n} = n$. We have also used the definition of γ_n so that $\bar{I}_\infty = \bar{I}_{\gamma_n}$ since no more vertices can be infected after γ_n .

Recalling the definition of the σ in statement of Theorem 1.13 we see that for any $\epsilon > 0$,

$$\inf_{t_\eta < t < -\log(\sigma + \epsilon)} \frac{\bar{x}_t}{\bar{x}_{S,t}} > 1,$$

which implies that for some $\epsilon' > 0$ and all $t_\eta < t < -\log(\sigma + \epsilon)$,

$$\bar{x}_t - \bar{x}_{S,t} > \epsilon'. \quad (3.5.4)$$

The first and third equations of (1.6.2) and (3.5.4) imply that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{P}(t_\eta < \gamma_n < -\log(\sigma + \epsilon)) \\ & \leq \limsup_{n \rightarrow \infty} \mathbb{P}(t_\eta < \gamma_n < -\log(\sigma + \epsilon), (\bar{X}_{\gamma_n} - \bar{X}_{S,\gamma_n}) / n > \epsilon' / 2) = 0, \end{aligned}$$

where we have used the fact that $\bar{X}_{\gamma_n} = \bar{X}_{S,\gamma_n}$. Since we have already shown conditionally on $\bar{I}_\infty / n > \eta$ whp $\gamma_n > t_\eta$, we see that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\gamma_n > -\log(\sigma + \epsilon) | \bar{I}_\infty / n > \eta) = 1.$$

Using the monotonically decreasing property of $\bar{S}_t = \sum_{k=0}^{\infty} \bar{S}_{t,k}$ and the second equation of (1.6.2), we see that conditionally on $\bar{I}_{\infty}/n > \eta$, for any $\epsilon > 0$ whp

$$\bar{S}_{\gamma_n}/n \leq \exp\left(-\frac{\alpha}{2}(\sigma + \epsilon - 1)^2\right) G(\sigma + \epsilon) + \epsilon.$$

Hence for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(I_{\infty}/n > \nu - \epsilon | I_{\infty}/n > \eta) = 1,$$

where $\nu = 1 - \exp(-\alpha/2(\sigma - 1)^2)G(\sigma)$.

For the other direction, $(\star)$ implies that we can pick $\delta' > 0$ so that $f(w) < 0$ on $(\sigma - \delta', \sigma)$. (1.6.2) thus implies with high probability γ_n cannot be larger than $-\log(\sigma - \delta')$ since otherwise we would have $\bar{X}_{-\log(\sigma - \delta')} < \bar{X}_{S, -\log(\sigma - \delta')}$, which is impossible. Since δ' can be taken arbitrarily small we conclude that for any $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}(\bar{I}_{\infty}/n < \nu + \epsilon) = 1.$$

3.6 Proof of Theorem 1.14

From (3.5.1) we see that as $\lambda \rightarrow \lambda_c$ we have $f'(1) \rightarrow 0$. The second derivative of f is given by

$$f''(w) = -\frac{1}{w^2} - \frac{G''' + \alpha(1-w)G'' - 2\alpha G'}{G' + \alpha(1-w)G} + \left(\frac{G'' - \alpha G + \alpha(1-w)G'}{G' + \alpha(1-w)G}\right)^2 + \alpha.$$

Recall μ_k denotes the factorial moment $\mathbb{E}[D(D-1)\cdots(D-k+1)]$. We have assumed that $\mathbb{E}(D^5) < \infty$ so

$$G'''(1) = \mathbb{E}[D(D-1)(D-2)] = \mu_3 < \infty.$$

Since the terms with $1-w$ vanish at $w = 1$, inserting the values of $G(1)$, $G'(1)$, $G''(1)$ and $G'''(1)$ we get

$$\begin{aligned} f''(1) &= -1 - \frac{G'''(1) - 2\alpha G'(1)}{G'(1)} + \left(\frac{G''(1) - \alpha G(1)}{G'(1)}\right)^2 + \alpha \\ &= -1 - \frac{\mu_3}{m_1} + 2\alpha + \left(\frac{m_2 - m_1 - \alpha}{m_1}\right)^2 + \alpha. \end{aligned}$$

Theorem 1.6 tells us that $\alpha_c = m_2 - 2m_1$ so

$$-1 + \left(\frac{m_2 - m_1 - \alpha_c}{m_1}\right)^2 = 0.$$

From this, we see that at α_c

$$f''(1) = \frac{-\mu_3 + 3m_1m_2 - 6m_1^2}{m_1}.$$

Using $\mu_1 = m_1$, $\mu_2 = m_2 - m_1$ this can be written as

$$f''(1) = -\frac{\mu_3}{\mu_1} + 3(\mu_2 - \mu_1) \equiv \Delta. \quad (3.6.1)$$

By Theorem 1.13 it suffices to prove that in the case $\Delta < 0$, σ converges to 1 as $\lambda \rightarrow \lambda_c$. Equation (3.6.1) and the assumption $\Delta < 0$ imply that for λ close to λ_c , in a (non-shrinking) neighborhood of 1, $f''(w)$ has to be bounded from above by some negative constant. Since $f'(1)$ converges to 0 as $\lambda \rightarrow \lambda_c$ we conclude that for any fixed $w < 1$ and all λ sufficiently close to λ_c one can find $\hat{w} \in (w, 1)$ so that $f(\hat{w}) < 0$. Using the definition of σ we see $\sigma > \hat{w} > w$. Letting $w \rightarrow 1$, we see that σ has to converge to 0 as $\lambda \rightarrow \lambda_c$ and thus ν converges to 0. Hence we have a continuous phase transition.

4 Lower bound on evoSI

4.1 AB-avoSI

Roughly speaking, the avoSI process serves as an upper bound because certain $I - I$ pairs can rewire, which may leads to additional infections. To get a lower bound, we need to find a way to ensure rewired $I - I$ edges will not transmit infections. This motivates the AB-avoSI process defined as follows. For each half-edge h we give it two indices:

- The infection index $A(h, t) = 0$ if h has not been infected by time t . If h first becomes an infected half-edge at time s , then we set $A(h, t) = s$ for all $t \geq s$.
- The rewiring index $B(h, t) = 0$ if the half-edge h has not rewired by time t . If h gets rewired at time s , then we update the value of $B(h, s)$ to be s , no matter whether h has been rewired before or not. In other words, if we let $\tau_m(h)$ be the time when h is rewired for the m -th time (possibly ∞) with $\tau_0(h) = 0$, then $B(h, t) = \tau_n(h)$ for $\tau_m(h) \leq t < \tau_{m+1}(h)$.

We define the C-AB-avoSI process as follows. As in Section 3.1, C is for coupled.

- At rate λ each infected half-edge h_1 pairs with a randomly chosen half-edge. Suppose h_1 gets paired with half-edge h_2 at time t . If h_2 is susceptible and $B(h_2, t) < A(h_1, t)$ then the vertex associated with half-edge h_2 becomes infected. Otherwise h_1 will not pass infection to the vertex associated with h_2 . The reader will see the reason for this condition in the proof of Lemma 4.1. Note that if vertex y associated with h_2 changes from state S to I then all half-edges attached to y become infected.
- Each infected half-edge gets removed from the vertex that it is attached to at rate ρ and immediately becomes re-attached to a randomly chosen vertex.

Similarly to the relation between C-avoSI and avoSI, one can also define the AB-avoSI such that the C-AB-avoSI has the same law as the AB-avoSI on the configuration model. The construction of the graph G for AB-avoSI follows the same route of avoSI (see the proof of Lemma 3.1). Given the graph G , we view each (full) edge as being composed of two half-edges and assign the two indices $A(\cdot, t)$ and $B(\cdot, t)$ as defined above to every half-edge. The evolution of AB-avoSI is then similar to avoSI, except that each time when infected vertex x tries to infect susceptible vertex y through edge e , we will compare $A(h_1, t)$ and $B(h_2, t)$, where h_1 is the half-edge of e with one end at x and h_2 is the other half-edge of e with one end at y . If $A(h_1, t) > B(h_2, t)$ then the infection will pass through. Otherwise the infection will not pass through, which means x has made an attempt but y remains uninfected. No matter whether the infection passes through or not, we let this $S - I$ edge be deemed stable and it will not get rewired later on. Also from this point on, x will never pass infection to y through e . As a comparison, $S - I$ edges in avoSI are always unstable and are subject to potential rewiring.

One can show that evoSI stochastically dominates AB-avoSI.

Lemma 4.1. *There exists a coupling of avoSI and evoSI such that if a vertex is infected in AB-avoSI then it is also infected in evoSI.*

Proof of Lemma 4.1. Following the proof of Lemma 3.2, we see it suffices to show $I - I$ rewirings will not create additional infections in AB-avoSI. Assume at some time t_1 an edge e between two infected vertices x and y is broken from y and reconnects to z . For a picture see Figure 8.

We would like to show that after time t_1 no infections can pass through e to create additional infected vertices. Denote the half-edge attached to x by h_1 and the other

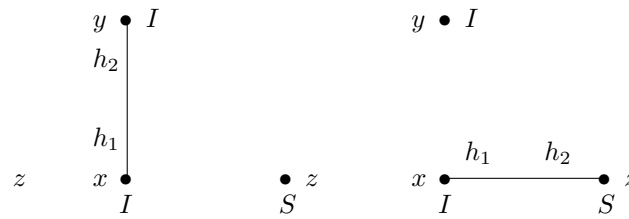


Figure 8: In the transition from the first drawing to the second, x rewires its connection from y to z . Note that h_1 and h_2 are the two half-edges comprising the edge e , which is not drawn in the figure. Vertex x cannot pass an infection to z through h_1 and h_2 since $B(h_2, t) > A(h_1, t)$.

half-edge by h_2 . Since x and y must be infected before time t_1 , we see that, according to the definition of the infection index, for all $t \geq t_1$,

$$A(h_1, t) < t_1, A(h_2, t) < t_1. \quad (4.1.1)$$

We next consider the rewiring index. Since the half-edge h_2 is rewired at time t_1 , using the definition of rewiring index we see that for all $t \geq t_1$,

$$B(h_2, t) \geq t_1. \quad (4.1.2)$$

Combining (4.1.1) and (4.1.2) we see that

$$B(h_2, t) > A(h_1, t).$$

This implies that infections cannot pass from the vertex associated to h_1 to the vertex associated to h_2 . Now we consider the other direction, i.e., from h_2 to h_1 . Note that x has already been infected. There are two possible cases:

- If h_1 stays with x forever then the vertex associated with h_1 is always infected after time t_1 (and hence there is no additional infected vertex).
- If h_1 is rewired to some other vertex x' at time $t_2 > t_1$ then for all $t > t_2$,

$$B(h_1, t) \geq t_2 > t_1 > A(h_2, t),$$

which implies that after time t_2 infections cannot go from the vertex associated with h_2 to the vertex associated with h_1 .

In both cases there will be no additional infected vertices stemming from the rewiring of e . Thus we have completed the proof of Lemma 4.1. $\square$

We end this section by showing AB-avoSI dominates delSI. This is used in the proof Lemma 1.12.

Lemma 4.2. *There exists a coupling of AB-avoSI and delSI such that if a vertex is infected in AB-avoSI then it is also infected in delSI.*

Proof. This can be proved in a similar way to the proof of Lemma 1.2. We construct AB-avoSI using the variables $\{T_{e,\ell}, R_{e,\ell}, V'_{e,\ell}, \ell \geq 1\}$ as we did in the proof of Lemma 3.2 except that for AB-avoSI certain infections may not pass through depending on the relative size of infection index and rewiring index. We also construct delSI using $T_{e,1}$ and

$R_{e,1}$ as we did in the proof of Lemma 1.2. We can then prove Lemma 4.2 by repeating the induction argument used in the proof of Lemma 1.2. Note that if edge e is initially present between x and y and $T_{e,1} < R_{e,1}$ then either end of e can be infected by the other end because the rewiring indices for both half-edges of e are equal to 0 at the time of infection. Therefore we don't need to worry about infections not being transmitted successfully in AB-evoSI. $\square$

4.2 Moment bounds

Let $\tilde{X}_t$ and $\tilde{X}_{I,t}$ be the number of total half-edges and infected half-edges in the AB-avoSI process. As in the analysis of avoSI, we multiply the original transition rates by $(\tilde{X}_t - 1)/(\lambda \tilde{X}_{I,t})$. We will use a hat to denote the quantities after the time change. The evolution equation for $\hat{X}_t$ has the same form as avoSI and hence the first equation of (1.6.2) also holds for AB-avoSI. Now we consider the evolution of the number of susceptible half-edges $\hat{X}_{S,t}$. We need a bit more notation to describe this. For half-edge i , we let $I(i, t) = 1$ if i is an infected half-edge at time t (which also means it hasn't been paired) and $I(i, t) = 0$ otherwise. We can define $S(i, t)$ similarly. We also let $S(i, k, t) = 1$ if i is attached to a susceptible vertex with k half-edges at time t . Finally, we let $v(j, t)$ be the vertex that half-edge j is attached to at time t and $D(j, t)$ be the number of half-edges attached to $v(j, t)$ at time t .

We first write down the equation for $\hat{S}_{t,k}$. To reduce the size of formulas we let

$$G_{i,j} = \{I(i, t) = 1, A(i, t) \leq B(j, t)\}. \quad (4.2.1)$$

Reasoning as in the derivation of (3.2.3) gives

$$\begin{aligned} d\hat{S}_{t,k} = & -k\hat{S}_{t,k} dt + 1_{\{k \geq 1\}} \frac{\rho}{\lambda} \frac{\hat{S}_{t,k-1}}{n} (\hat{X}_t - 1) dt - \frac{\rho}{\lambda} \frac{\hat{S}_{t,k}}{n} (\hat{X}_t - 1) dt \\ & + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} 1_{\{S(j,k+1,t)=1\}} \right) dt + d\hat{M}_{t,k}, \end{aligned} \quad (4.2.2)$$

where $\hat{M}_{t,k}$ is a martingale. Summing (4.2.2) over k from 0 to ∞ and noting that the second and third term cancel, we get

$$d\hat{S}_t = -\hat{X}_{S,t} dt + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} 1_{\{S(j,t)=1\}} \right) dt + dM_{1,t}. \quad (4.2.3)$$

Multiplying both sides of (4.2.2) by k and summing over k , we get

$$\begin{aligned} d\hat{X}_{S,t} = & -\sum_{k=0}^{\infty} k^2 \hat{S}_{t,k} dt + \frac{\rho}{\lambda} \frac{\hat{S}_t}{n} (\hat{X}_t - 1) dt \\ & + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} (D(j, t) - 1) 1_{\{S(j,t)=1\}} \right) dt + dM_{2,t} \end{aligned} \quad (4.2.4)$$

for some martingale term $M_{2,t}$.

Analogously, if we multiply both sides of (4.2.2) by k^2 , k^3 and k^4 , respectively, then we get

$$\begin{aligned} d \left(\sum_{k=0}^{\infty} k^2 \hat{S}_{t,k} \right) = & -\sum_{k=0}^{\infty} k^3 \hat{S}_{t,k} dt + 2 \frac{\rho}{\lambda} \frac{\hat{X}_{S,t}}{n} (\hat{X}_t - 1) + \frac{\rho}{\lambda} \frac{\hat{S}_t}{n} (\hat{X}_t - 1) \\ & + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} (D(j, t) - 1)^2 1_{\{S(j,t)=1\}} \right) dt + dM_{3,t}, \end{aligned} \quad (4.2.5)$$

$$d\left(\sum_{k=0}^{\infty} k^3 \hat{S}_{t,k}\right) = -\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k} dt + 3 \frac{\rho}{\lambda} \frac{\sum_{k=0}^{\infty} k^2 \hat{S}_{t,k}}{n} (\hat{X}_t - 1) + 3 \frac{\rho}{\lambda} \frac{\hat{X}_{S,t}}{n} (\hat{X}_t - 1) + \frac{\rho}{\lambda} \frac{\hat{S}_t}{n} (\hat{X}_t - 1) + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\bar{X}_0} 1_{G_{i,j}} (D(j,t) - 1)^3 1_{\{S(j,t)=1\}} \right) dt + dM_{4,t}, \quad (4.2.6)$$

and

$$d\left(\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}\right) = -\sum_{k=0}^{\infty} k^5 \hat{S}_{t,k} dt + 4 \frac{\rho}{\lambda} \frac{\sum_{k=0}^{\infty} k^3 \hat{S}_{t,k}}{n} (\hat{X}_t - 1) + 6 \frac{\rho}{\lambda} \frac{\sum_{k=0}^{\infty} k^2 \hat{S}_{t,k}}{n} (\hat{X}_t - 1) + 4 \frac{\rho}{\lambda} \frac{\hat{X}_{S,t}}{n} (\hat{X}_t - 1) + \frac{\rho}{\lambda} \frac{\hat{S}_t}{n} (\hat{X}_t - 1) + \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\bar{X}_0} 1_{G_{i,j}} (D(j,t) - 1)^4 1_{\{S(j,t)=1\}} \right) dt + dM_{5,t}. \quad (4.2.7)$$

The main result of this section is the following lemma.

Lemma 4.3. *There exists two constants $C_{4.2.9}, C_{4.2.10} > 0$, such that for every $\epsilon > 0$ whp we have*

$$\sum_{j=1}^5 \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{j,t}| \right) \leq n^{5/6}, \quad (4.2.8)$$

$$\sup_{0 \leq t \leq \gamma_n \wedge 1} \sum_{k=0}^{\infty} (k+1)^4 \hat{S}_{t,k} \leq C_{4.2.9} n, \quad (4.2.9)$$

$$\sum_{j=1}^{\hat{X}_0} D(j,t)^2 1_{\{S(j,t)=1, B(j,t)>0\}} \leq n(C_{4.2.10} t + \epsilon), \quad \text{for all } 0 \leq t \leq \gamma_n \wedge 1 \quad (4.2.10)$$

Proof of Lemma 4.3. We first prove equation (4.2.8). The proof is based on analyzing the quadratic variation of $M_{1,t}, \dots, M_{5,t}$. Since the proofs for the five quantities are similar we only give the details for $M_{5,t}$, the martingale associated with $\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}$. Let Q_x be the number of half-edges that vertex x originally has plus the half-edges that has been rewired to vertex x before x becomes infected. Let $D_x(t)$ be the number of half-edges that x has at time t . We necessarily have $D_x(t) \leq Q_x$ as long as x is susceptible at time t . Note that the jumps of $\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}$ have the following three sources.

- An infected half-edge pairs with a susceptible half-edge attached to a vertex x with $D_x(t)$ half-edges and passes the infection to x . This decreases $\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}$ by $D_x(t)^4$. Such type of jumps can occur at most once for each susceptible vertex.
- An infected half-edge pairs with a susceptible half-edge attached to a vertex x with $D_x(t)$ half-edges but does not pass the infection. This decreases $\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}$ by

$$D_x(t)^4 - (D_x(t) - 1)^4 \leq 15Q_x^3.$$

Such type of jumps can occur at most Q_x times for vertex x .

- An infected half-edge is rewired to a susceptible vertex x with $D_x(t)$ half-edges. This gives an increase of

$$(D_x(t) + 1)^4 - D_x(t)^4 \leq 15Q_x^3$$

to $\sum_{k=0}^{\infty} k^4 \hat{S}_{t,k}$. Such type of jumps can happen at most Q_x times for vertex x .

It follows from the above analysis that the quadratic variation of $M_{5,t}$ is bounded by

$$\sum_{x=1}^n (Q_x^4)^2 + \sum_{x=1}^n 15(Q_x^3)^2 Q_x + \sum_{x=1}^n 15(Q_x^3)^2 Q_x \leq C_{4.2.11} \sum_{x=1}^n Q_x^8. \quad (4.2.11)$$

Using the Burkholder-Davis-Gundy inequality (see, e.g., [15, Theorem 7.34] with $p = 5/4$),

$$\mathbb{E} \left[\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{5,t}|^{5/4} \right] \leq C_{4.2.12} \mathbb{E} \left[\left(\sum_{x=1}^n Q_x^8 \right)^{5/8} \right]. \quad (4.2.12)$$

To bound the right-hand side we need the following well-known fact: for any $p \geq 1$ and positive numbers $a_1, \dots, a_m$,

$$\left(\sum_{i=1}^m a_i^p \right)^{1/p} \leq \sum_{i=1}^m a_i. \quad (4.2.13)$$

In words the L^p norm under the counting measure decreases as p increases. This is easily seen to be true by noting that

$$\sum_{i=1}^m \frac{a_i}{\left(\sum_{j=1}^m a_j^p \right)^{1/p}} = \sum_{i=1}^m \left(\frac{a_i^p}{\sum_{j=1}^m a_j^p} \right)^{1/p} \geq \sum_{i=1}^m \frac{a_i^p}{\sum_{j=1}^m a_j^p} = 1,$$

since

$$\frac{a_i^p}{\sum_{j=1}^m a_j^p} \in [0, 1] \text{ and } 1/p \leq 1.$$

Applying (4.2.13) with $p = 8/5$ and $a_i = Q_i^8$ gives that

$$\left(\sum_{x=1}^n Q_x^8 \right)^{5/8} \leq \sum_{x=1}^n Q_x^5. \quad (4.2.14)$$

As we argued in the proof of (3.4.7), if $V = \text{Binomial}(\hat{X}_0, 1/n)$, then Q_x is dominated by $D_x + V$. To bound the fifth moment of the sum we note that if Y and Z are nonnegative random variables,

$$\mathbb{E}(Y + Z)^5 \leq \mathbb{E}(2 \max\{Y, Z\})^5 \leq 32[\mathbb{E}Y^5 + \mathbb{E}Z^5]. \quad (4.2.15)$$

We claim that the 5-th moment of $\text{Binomial}(m, p)$ is bounded by

$$mp + \dots + (mp)^5.$$

To see this, let $Y_1, \dots, Y_m$ be i.i.d. Bernoulli variable with mean p .

$$\mathbb{E} \left[\left(\sum_{i=1}^m Y_i \right)^5 \right] = \sum_{i_1, \dots, i_5} \mathbb{E}(Y_{i_1} \dots Y_{i_5}) = \sum_{i_1, \dots, i_5} p^{\# \text{ of distinct elements among } i_1, \dots, i_5}. \quad (4.2.16)$$

Note that the number of ordered tuples $(i_1, \dots, i_5)$ such that there are ℓ distinct elements among them is bounded by m^ℓ . We conclude that

$$\mathbb{E} \left[\left(\sum_{i=1}^m Y_i \right)^5 \right] \leq \sum_{\ell=1}^5 m^\ell p^\ell, \quad (4.2.17)$$

which verifies the claim. Using this claim we have that

$$\mathbb{E}(V^5) \leq \sum_{\ell=1}^5 \frac{\mathbb{E}(\widehat{X}_0^\ell)}{n^\ell} \leq C_{4.2.18}, \quad (4.2.18)$$

so that by (4.2.15),

$$\sum_{x=1}^n \mathbb{E}(Q_x^5) \leq 32 \sum_{i=1}^n (\mathbb{E}(D_i^5) + \mathbb{E}(V^5)) \leq C_{4.2.19} n. \quad (4.2.19)$$

Equations (4.2.12), (4.2.14) and (4.2.19) imply that

$$\mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{5,t}| > n^{5/6} \right) = \mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{5,t}|^{5/4} > (n^{5/6})^{5/4} \right) \leq \frac{Cn}{n^{25/24}} \rightarrow 0. \quad (4.2.20)$$

Using the reasoning that led to (4.2.11), we can deduce the same bounds (with different constants) for $M_{1,t}, M_{2,t}, M_{3,t}, M_{4,t}$ and hence equation (4.2.8) follows by using Markov's inequality.

To prove equation (4.2.9), define the event Ω_n to be

$$\left\{ \left| \sum_{k=0}^{\infty} k^i \widehat{S}_{0,k} - n \sum_{k=0}^{\infty} k^i p_k \right| \leq n, \sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{i,t}| \leq n, \text{ for } i = 1, 2, 3, 4 \right\}. \quad (4.2.21)$$

The assumption that $\mathbb{E}(D^5) < \infty$ and equation (4.2.8) imply that $\mathbb{P}(\Omega_n) \rightarrow 1$ as $n \rightarrow \infty$. By the definition of $G_{i,j}$ in (4.2.1),

$$\frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{G_{i,j}}(D(j,t) - 1)^2 1_{\{S(j,t)=1\}} \right) \leq \frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{\{I(i,t)=1\}} (D(j,t) - 1)^2 1_{\{S(j,t)=1\}} \right),$$

which is bounded by

$$\begin{aligned} \sum_{j=1}^{\widehat{X}_0} D(j,t)^2 1_{\{S(j,t)=1\}} &= \sum_{r=1}^n \sum_{j=1}^{\widehat{X}_0} 1_{\{v(j,t)=r\}} 1_{\{r \text{ is susceptible at } t\}} D_r(t)^2 \\ &= \sum_{r=1}^n 1_{\{r \text{ is susceptible at } t\}} D_r(t)^3 = \sum_{k=0}^{\infty} k^3 \widehat{S}_{t,k}. \end{aligned} \quad (4.2.22)$$

Here $D_r(t)$ is the number of half-edges that vertex r has at time t . Using (4.2.5) and (4.2.22) we obtain that

$$\sum_{k=0}^{\infty} k^2 \widehat{S}_{t,k} - \sum_{k=0}^{\infty} k^2 \widehat{S}_{0,k} \leq \int_0^t \left(2 \frac{\rho}{\lambda} \frac{\widehat{X}_{S,u}}{n} (\widehat{X}_u - 1) + \frac{\rho}{\lambda} \frac{\widehat{S}_u}{n} (\widehat{X}_u - 1) \right) dt + M_{3,t}. \quad (4.2.23)$$

On the event Ω_n , we have that $\sum_{k=0}^{\infty} k^2 \widehat{S}_{0,k} \leq (m_2 + 1)n$, $\widehat{X}_{S,u} \leq \widehat{X}_0 \leq (m_1 + 1)n$ and $M_{3,t} \leq n$. Therefore, using (4.2.23) we see that, there exists a constant $C_{4.2.24}$ such that

$$\sum_{k=0}^{\infty} k^2 \widehat{S}_{t,k} \leq (m_2 + 1)n + \int_0^t \left(\frac{2\rho}{\lambda} (m_1 + 1)^2 n + \frac{\rho}{\lambda} (m_1 + 1)n \right) du + n \leq C_{4.2.24} n \quad (4.2.24)$$

for all $0 \leq t \leq \gamma_n \wedge 1$. Analogously to the proof of (4.2.22), we can show

$$\frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{G_{i,j}}(D(j,t) - 1)^3 1_{\{S(j,t)=1\}} \right) \leq \sum_{k=0}^{\infty} k^4 \widehat{S}_{t,k}. \quad (4.2.25)$$

Using (4.2.25) and (4.2.6),

$$\begin{aligned} \sum_{k=0}^{\infty} k^3 \widehat{S}_{t,k} - \sum_{k=0}^{\infty} k^3 \widehat{S}_{0,k} &\leq \int_0^t \left(3 \frac{\rho}{\lambda} \frac{\sum_{k=0}^{\infty} k^2 \widehat{S}_{u,k}}{n} (\widehat{X}_u - 1) \right. \\ &\quad \left. + 3 \frac{\rho}{\lambda} \frac{\widehat{X}_{S,u}}{n} (\widehat{X}_u - 1) + \frac{\rho}{\lambda} \frac{\widehat{S}_u}{n} (\widehat{X}_u - 1) \right) du + M_{4,t}. \end{aligned}$$

On the event Ω_n , using (4.2.24), we see that, for all $0 \leq t \leq \gamma_n \wedge 1$,

$$\begin{aligned} \sum_{k=0}^{\infty} k^3 \widehat{S}_{t,k} &\leq (m_1 + 1)n + n + \\ &\quad \int_0^t \left(\frac{3\rho}{\lambda} C_{4.2.24}(m_1 + 1)n + \frac{3\rho}{\lambda} (m_1 + 1)^2 n + \frac{\rho}{\lambda} (m_1 + 1)n \right) du \leq C_{4.2.26}n, \end{aligned} \quad (4.2.26)$$

for some constant $C_{4.2.26} > 0$. Proceeding in a similar fashion and using (4.2.7) we can show that on Ω_n there exists a constant $C_{4.2.27} > 0$, such that whp for all $0 \leq t \leq \gamma_n \wedge 1$,

$$\sum_{k=0}^{\infty} k^4 \widehat{S}_{t,k} \leq C_{4.2.27}n. \quad (4.2.27)$$

Equation (4.2.9) follows from (4.2.27) and the fact that $\mathbb{P}(\Omega_n) \rightarrow 1$ since

$$\sum_{k=0}^{\infty} (k+1)^4 \widehat{S}_{t,k} \leq 16 \sum_{k=0}^{\infty} (k^4 + 1) \widehat{S}_{t,k} \leq 16 \sum_{k=0}^{\infty} k^4 \widehat{S}_{t,k} + 16 \sum_{k=0}^{\infty} \widehat{S}_{t,k} \leq 16 \sum_{k=0}^{\infty} k^4 \widehat{S}_{t,k} + 16n.$$

We now turn to the proof of equation (4.2.10). Set

$$H(t) = \sum_{j=1}^{\widehat{X}_0} (D(j, t) - 1)^2 1_{\{S(j, t)=1, B(j, t)>0\}}.$$

Note that $H(0) = 0$. Using Dynkin's formula,

$$H(t) = \int_0^t h(s) ds + M_{6,t},$$

where $M_{6,t}$ is a martingale associated with $H(t)$ and $h(t)$ is the rate of change of $H(t)$. We now control $h(t)$ and $M_{6,t}$ by analyzing the jumps of $H(t)$ ($H(t)$ is a pure jump process). Note that there are three types of jumps:

- An infected half-edge pairs with a half-edge attached to susceptible vertex x and makes x infected. This does not increase $H(t)$ and thus makes a non-positive contribution to $h(t)$. The absolute value of the jump size of $H(t)$ is bounded by $D_x(t)^3$ where $D_x(t)$ is the number of half-edges that x has at time t . For each vertex x such jumps can happen at most once.
- An infected half-edge pairs with a half-edge attached to susceptible vertex x but x stays susceptible after the pairing. This does not increase $H(t)$ and thus makes a non-positive contribution to $h(t)$. The absolute value of the jump size of $H(t)$ is bounded above by

$$D_x(t)^2 + |(D_x(t) - 1)^2 - (D_x(x) - 2)^2| D_x(t) \leq 3(D_x(t) + 1)^2.$$

To see this, note that the loss of a half-edge j attached to x makes a twofold contribution to $H(t)$. First, j is no longer a half-edge so $H(t)$ has to decrease

by $(D(j, t) - 1)^2 = (D_x(t) - 1)^2$. Second, for each of the remaining $D_x(t) - 1$ half-edges attached to x , its contribution to $H(t)$ changes from $(D_x(t) - 1)^2$ to $(D_x(t) - 2)^2$.

For each vertex x such type of jumps can occur at most Q_x times.

- An infected half-edge is rewired to a susceptible vertex x . This increases $H(t)$ by at most

$$D_x^2(t) + |(D_x(t)^2 - (D_x(t) - 1)^2)| D_x(t) \leq 3(D_x(t) + 1)^2 \leq 3(Q_x + 1)^2.$$

The rate that x receives a rewired half-edge is equal to

$$\frac{\hat{X}_0}{\lambda \hat{X}_{I,t}} \rho \hat{X}_{I,t} \frac{1}{n} = \frac{\rho \hat{X}_0}{\lambda n} \leq \frac{(m_1 + 1)\rho}{\lambda}$$

on the event Ω_n (defined in (4.2.21)). Here the factor of $1/n$ comes from Poisson thinning since each half-edge is rewired to a uniformly chosen vertex independently. For each vertex x such type of jumps can occur at most Q_x times.

Therefore on Ω_n we have

$$h(t) \leq \frac{(m_1 + 1)\rho}{\lambda} \sum_{x=1}^n 3(D_x(t) + 1)^2 1_{\{x \text{ is susceptible at } t\}} \leq \frac{(m_1 + 1)\rho}{\lambda} \sum_{k=0}^{\infty} 3(k + 1)^4 \hat{S}_{t,k}.$$

Equation (4.2.9) shows that with high probability $\sum_{k=0}^{\infty} (k + 1)^4 \hat{S}_{t,k} \leq C_{4.2.9} n$. Since Ω_n also holds with high probability, we deduce that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\int_0^t h(s) ds \leq C_{4.2.28} n t, \forall 0 \leq t \leq \gamma_n \right) = 1. \quad (4.2.28)$$

The above analysis of the jumps of $H(t)$ also implies that the quadratic variation of $M_{6,t}$ is bounded by

$$(Q_x^3)^2 + (3(Q_x + 1)^2)^2 Q_x + (3(Q_x + 1)^2)^2 Q_x \leq 20(Q_x + 1)^6.$$

We now bound the $2/3$ -th moment of the quadratic variation. Applying (4.2.13) with $p = 3/2$ and $a_i = (Q_i + 1)^6$ gives that

$$\left(\sum_{i=1}^n (Q_i + 1)^6 \right)^{2/3} \leq \sum_{i=1}^n (Q_i + 1)^4.$$

Using this and the Burkholder-Davis-Gundy inequality ([15, Theorem 7.34] with $p = 4/3$),

$$\mathbb{E} \left[\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{6,t}|^{4/3} \right] \leq C' \mathbb{E} \left[\left(\sum_{i=1}^n (Q_i + 1)^6 \right)^{2/3} \right] \leq C' \mathbb{E} \left(\sum_{i=1}^n (Q_i + 1)^4 \right) \leq Cn.$$

This implies that

$$\mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{6,t}| > n^{4/5} \right) = \mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{6,t}|^{4/3} > n^{4/5 \cdot 4/3} \right) \leq \frac{Cn}{n^{16/15}} \rightarrow 0. \quad (4.2.29)$$

Equation (4.2.10) follows from equations (4.2.28) and (4.2.29). $\square$

4.3 Rough upper and lower bounds for $\hat{X}_{I,t}$ and $\hat{I}_t$

Define

$$E(t) = \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{G_{i,j}} D(j,t) 1_{\{S(j,t)=1\}} \right) \quad (4.3.1)$$

where $G_{i,j} = \{I(i,t) = 1, A(i,t) \leq B(j,t)\}$ was defined in (4.2.1).

Lemma 4.4. *There exists two constants $\lambda_0, t_0 \in (0, 1)$, such that for all $\lambda < \lambda_c + \lambda_0$, $t < t_0$ and $\epsilon > 0$, the following four inequalities hold whp for $0 \leq t \leq \gamma_n \wedge t_0$:*

$$\hat{X}_{I,t} \leq \left(\frac{2\rho m_1}{\lambda_c^2} (\lambda - \lambda_c) t + m_1 \Delta t^2 + \epsilon \right) n, \quad (4.3.2)$$

$$\hat{X}_{I,t} \geq \left(\frac{\rho m_1}{2\lambda_c^2} (\lambda - \lambda_c) t + \frac{m_1 \Delta}{4} t^2 - \epsilon \right) n - \int_0^t E(u) du, \quad (4.3.3)$$

$$\hat{I}_t \leq (2m_1 t + \epsilon) n, \quad (4.3.4)$$

$$\hat{I}_t \geq \left(\frac{m_1 t}{2} - \epsilon \right) n - \int_0^t E(u) du. \quad (4.3.5)$$

Proof of Lemma 4.4. We first prove equation (4.3.2). Using Lemma 4.3 we see that there exist a constant C such that whp the absolute value of the right hand sides of (4.2.4)–(4.2.6) are all upper bounded by Cn for $t \leq \gamma_n \wedge 1$. Denote this event by $G_1(n)$. We have $\mathbb{P}(G_1(n)) \rightarrow 1$ as $n \rightarrow \infty$. Denote the event

$$\left\{ \left| \sum_{k=0}^{\infty} k^i \hat{S}_{0,k} - n \sum_{k=0}^{\infty} k^i p_k \right| \leq \epsilon n, \quad \text{for } i = 1, 2, 3 \right\}$$

by $G_2(n)$. We have $\mathbb{P}(G_2(n)) \rightarrow 1$ as $n \rightarrow \infty$ since $\mathbb{E}(D^3) < \infty$. Therefore we get

$$\lim_{n \rightarrow \infty} \mathbb{P}(G_1(n) \cap G_2(n)) = 1,$$

which implies that whp for all $0 \leq t \leq \gamma_n \wedge 1$, we have

$$\left| \sum_{k=0}^{\infty} k^3 \bar{S}_{t,k} - n \sum_{k=0}^{\infty} k^3 p_k \right| + \left| \hat{X}_{S,t} - n \sum_{k=0}^{\infty} k p_k \right| \leq 4n(Ct + \epsilon). \quad (4.3.6)$$

The definition of $E(t)$ in (4.3.1) implies that

$$E(t) \leq \sum_{j=1}^{\hat{X}_0} D(j,t) 1_{\{S(j,t)=1, B(j,t)>0\}}.$$

Therefore by (4.2.10) we see that whp for all $t \leq \gamma_n \wedge 1$,

$$E(t) \leq n(Ct + \epsilon). \quad (4.3.7)$$

Now it follows from the integral form of (4.2.3), (4.3.6) and (4.3.7) that whp

$$\left| \hat{S}_t - n(1 - m_1 t) \right| \leq n(Ct^2 + \epsilon). \quad (4.3.8)$$

Similarly to the proof of (4.3.8), whp for all $0 \leq t \leq \gamma_n \wedge 1$

$$\left| \sum_{k=0}^{\infty} k^2 \bar{S}_{t,k} - n \left(\sum_{k=0}^{\infty} k^2 d_k + \left(- \sum_{k=0}^{\infty} k^3 p_k + 2 \frac{\rho}{\lambda} m_1^2 + \frac{\rho}{\lambda} m_1 \right) t \right) \right| \leq n(Ct^2 + \epsilon). \quad (4.3.9)$$

The evolution equation for $\widehat{X}_t$ in AB-avoSI has the same form as avoSI, i.e.,

$$\widehat{X}_t - \widehat{X}_0 = - \int_0^t 2(\widehat{X}_u - 1) du + M_{0,t}. \quad (4.3.10)$$

Hence, as in the case of avoSI, equation (3.4) in [13] holds for AB-avoSI, which implies

$$\sup_{0 \leq t \leq \gamma_n \wedge 1} \left| \frac{\widehat{X}_t}{n} - m_1 \exp(-2t) \right| \xrightarrow{\mathbb{P}} 0. \quad (4.3.11)$$

Combining (4.2.4) and (4.3.10) and using $\widehat{X}_{I,t} = \widehat{X}_t - \widehat{X}_{S,t}$, we get

$$\begin{aligned} \widehat{X}_{I,t} = & \widehat{X}_{I,0} + \int_0^t \left(-2(\widehat{X}_u - 1) + \sum_{k=0}^{\infty} k^2 \widehat{S}_{u,k} - \frac{\rho}{\lambda} \frac{\widehat{S}_u}{n} (\widehat{X}_u - 1) \right) du + (M_{0,t} - M_{2,t}) \\ & - \frac{1}{\widehat{X}_{I,t}} \int_0^t \left(\sum_{i,j=1}^{\overline{X}_0} 1_{G_{i,j}}(D(j,u) - 1) 1_{\{S(j,u)=1\}} \right) du. \end{aligned} \quad (4.3.12)$$

Dropping the term in the second line of (4.3.12) and bounding

$$\sup_{0 \leq t \leq \gamma_n \wedge 1} \left| \widehat{X}_{I,0} + M_{0,t} - M_{2,t} \right|$$

by ϵn (which holds with high probability by (4.2.8)),

$$\widehat{X}_{I,t} \leq \epsilon n + \int_0^t \left(-2(\widehat{X}_u - 1) + \sum_{k=0}^{\infty} k^2 \widehat{S}_{u,k} - \frac{\rho}{\lambda} \frac{\widehat{S}_u}{n} (\widehat{X}_u - 1) \right) du. \quad (4.3.13)$$

Using (4.3.11), (4.3.8) and (4.3.9) to approximate $\widehat{X}_t$, $\widehat{S}_t$ and $\sum_{k=0}^{\infty} k^2 \widehat{S}_{t,k}$ up to the first order, respectively, we obtain that whp,

$$\begin{aligned} \widehat{X}_t - \widehat{X}_{S,t} \leq & n \int_0^t \left(-2m_1 \exp(-2u) + \sum_{k=0}^{\infty} k^2 p_k + \left(- \sum_{k=0}^{\infty} k^3 p_k + 2 \frac{\rho}{\lambda} m_1^2 + \frac{\rho}{\lambda} m_1 \right) u \right. \\ & \left. - \frac{\rho}{\lambda} (1 - m_1 u) m_1 \exp(-2u) + C u^2 \right) du + 2n\epsilon. \end{aligned} \quad (4.3.14)$$

We would like to expand the integrand of (4.3.14) in powers of u . Since $\rho/\lambda_c = (m_2 - 2m_1)/m_1$, the constant term is

$$-2m_1 + m_2 - \frac{\rho}{\lambda} m_1 = m_1 \left(\frac{\rho}{\lambda_c} - \frac{\rho}{\lambda} \right) = \frac{m_1 \rho}{\lambda_c^2} (\lambda - \lambda_c) + O((\lambda - \lambda_c)^2). \quad (4.3.15)$$

for $\lambda > \lambda_c$. Therefore for λ sufficiently close to λ_c we have

$$-2m_1 + m_2 - \frac{\rho}{\lambda} m_1 \leq \frac{2m_1 \rho}{\lambda_c^2} (\lambda - \lambda_c). \quad (4.3.16)$$

Note that $e^{-2u} = 1 - 2u + 2u^2 + \dots$, so the coefficient in front of u is

$$4m_1 + \left(-m_3 + \frac{\rho}{\lambda} 2m_1^2 + \frac{\rho}{\lambda} m_1 \right) + \frac{\rho}{\lambda} m_1^2 + \frac{\rho}{\lambda} 2m_1 = 4m_1 - m_3 + \frac{\rho}{\lambda} (3m_1^2 + 3m_1). \quad (4.3.17)$$

At $\lambda = \lambda_c$, this coefficient is equal to

$$4m_1 - m_3 + (3 + 3m_1)(m_2 - 2m_1) = -m_3 + 3m_2 - 2m_1 + 3m_2 m_1 - 6m_1^2. \quad (4.3.18)$$

We claim that the quantity in (4.3.18) is exactly equal to $m_1\Delta$. Indeed, from the definition of Δ we see that

$$\Delta = -\frac{\mu_3}{\mu_1} + 3(\mu_2 - \mu_1) = -\frac{m_3 - 3m_2 + 2m_1}{m_1} + 3(m_2 - 2m_1)$$

so that

$$m_1\Delta = -m_3 + 3m_2 - 2m_1 + 3m_2m_1 - 6m_1^2.$$

Using (4.3.17) with the equations that follow, we see that for λ close to λ_c we have

$$4m_1 - m_3 + 2\frac{\rho}{\lambda}m_1^2 + \frac{\rho}{\lambda}m_1 + \frac{\rho}{\lambda}m_1(m_1 + 2) \leq \frac{3m_1\Delta}{2}. \quad (4.3.19)$$

Using (4.3.16) and (4.3.19) in (4.3.14), we see that for some constant $C' > 0$ and all $t \leq m_1\Delta/(2C')$,

$$\begin{aligned} \widehat{X}_{I,t} &\leq n \left(\int_0^t \left(\frac{2\rho}{\lambda_c^2}(\lambda - \lambda_c) + \frac{3m_1\Delta}{2}u + C'u^2 \right) du + 2\epsilon \right) \\ &\leq \left(\frac{2\rho}{\lambda_c^2}(\lambda - \lambda_c)t + \frac{3m_1\Delta}{4}t^2 + \frac{C't^3}{3} + 2\epsilon \right) n \\ &\leq \left(\frac{2\rho}{\lambda_c^2}(\lambda - \lambda_c)t + m_1\Delta t^2 + 2\epsilon \right) n. \end{aligned} \quad (4.3.20)$$

This proves (4.3.2) since ϵ is arbitrary.

The proof of (4.3.3) is parallel to the proof of (4.3.2), except that we now replace the second line of (4.3.12) by $E(u)$ (defined in (4.3.1)). We can do this because

$$\frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{G_{i,j}}(D(j,t) - 1) 1_{\{S(j,t)=1\}} \right) \leq E(t),$$

which is true by the definition of $E(t)$ in (4.3.1).

Since $\widehat{I}_t = n - \widehat{S}_t$, equation (4.2.3) implies

$$d\widehat{I}_t = \widehat{X}_{S,t} dt - \frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{G_{i,j}} 1_{\{S(j,t)=1\}} \right) dt - dM_{1,t}. \quad (4.3.21)$$

Using $\widehat{I}_0 = 1$ and the inequality

$$\frac{1}{\widehat{X}_{I,t}} \left(\sum_{i,j=1}^{\widehat{X}_0} 1_{G_{i,j}} 1_{\{S(j,t)=1\}} \right) \leq E(t),$$

which follows from the fact that $D(j,t) \geq 1$, we see that

$$\left| \widehat{I}_t - \left(1 + \int_0^t \widehat{X}_{S,u} du \right) \right| \leq \int_0^t E(u) du + |M_{1,t}|.$$

The rest of proofs for (4.3.4) and (4.3.5) are parallel to (4.3.2) and (4.3.3). We omit further details. $\square$

Let

$$L(t) = \sum_{j=1}^{\widehat{X}_0} D(j,t) 1_{\{S(j,t)=1, B(j,t)>0\}}. \quad (4.3.22)$$

Since $\sum_{i=1}^{\hat{X}_0} 1_{\{I(i,t)=1\}} = \hat{X}_{I,t}$ we have

$$\begin{aligned} E(t) &= \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{\{I(i,t)=1\}} D(j,t) 1_{\{S(j,t)=1\}} 1_{\{A(i,t) \leq B(j,t)\}} \right) \\ &\leq \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{\{I(i,t)=1\}} D(j,t) 1_{\{S(j,t)=1\}} 1_{\{B(j,t) > 0\}} \right) = L(t). \end{aligned} \quad (4.3.23)$$

The definition of $L(t)$ and the fact $D(j,t) \leq D(j,t)^2$ imply that

$$L(t) \leq \sum_{j=1}^{\hat{X}_0} D(j,t)^2 1_{\{S(j,t)=1, B(j,t) > 0\}}.$$

Using the fact $D(j,t) \leq D(j,t)^2$ again, (4.2.10) implies that

$$\lim_{n \rightarrow \infty} \mathbb{P}(L(t) \leq n(C_{4.2.10}t + \epsilon), \forall 0 \leq t \leq \gamma_n \wedge 1) = 1. \quad (4.3.24)$$

Combining (4.3.23) and (4.3.24) we see that for some constant $C_{4.3.25} > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\int_0^t E(u) du \leq (C_{4.3.25}t^2 + \epsilon)n, \forall 0 \leq t \leq \gamma_n \wedge t_0 \right) = 1. \quad (4.3.25)$$

The bound provided in (4.3.25) is not enough for our purpose (though we will also use it in the proof of Theorem 1.15). We will prove refined bounds in the next section.

4.4 More refined bounds

Equation (4.3.3) implies that we can get a lower bound for $\hat{X}_{I,t}$ if we can upper bound the term $\int_0^t E(u) du$. To this end, we let b be some number in $(0, 1)$ to be determined. We can decompose $E(t)$ into two parts:

$$\begin{aligned} E_1(t) &:= \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{\{I(i,t)=1\}} D(j,t) 1_{\{S(j,t)=1\}} 1_{\{A(i,t) < bt\}} \right), \\ E_2(t) &= \frac{1}{\hat{X}_{I,t}} \left(\sum_{i,j=1}^{\hat{X}_0} 1_{\{I(i,t)=1\}} D(j,t) 1_{\{S(j,t)=1\}} 1_{\{B(j,t) > bt\}} \right). \end{aligned} \quad (4.4.1)$$

Since $G_{i,j} = \{I(i,t) = 1, A(i,t) \leq B(j,t)\}$, we have that $E(t) \leq E_1(t) + E_2(t)$. Now we set

$$\begin{aligned} X(I, b, t) &:= \sum_{i=1}^{\bar{X}_0} 1_{\{A(i,t) \leq bt, I(i,t)=1\}} \leq \hat{X}_{I,t}, \\ L(b, t) &:= \sum_{j=1}^{\bar{X}_0} D(j,t) 1_{\{S(j,t)=1, B(j,t) > bt\}} \leq L(t). \end{aligned} \quad (4.4.2)$$

Recalling the definition of $L(t)$ in (4.3.22), we see that

$$E_1(t) = \frac{X(I, b, t)}{\hat{X}_{I,t}} L(t), \quad E_2(t) = L(b, t). \quad (4.4.3)$$

In the next two lemmas we give bounds on $L(b, t)$ and $X(I, b, t)$.

Lemma 4.5. *There exists a constant $C_{4.4.4}$ so that for any fixed $b \in (0, 1)$ and any $\epsilon > 0$, whp for all $0 \leq t \leq \gamma_n \wedge 1$,*

$$L(b, t) \leq n(C_{4.4.4}(1-b)^{1/2}t + \epsilon). \quad (4.4.4)$$

Proof. Using the Cauchy-Schwartz inequality,

$$\begin{aligned} L(b, t) &= \sum_{j=1}^{\hat{X}_0} D(j, t) 1_{\{S(j, t)=1, B(j, t)>bt\}} \\ &\leq \left(\sum_{j=1}^{\hat{X}_0} D(j, t)^2 1_{\{S(j, t)=1, B(j, t)>0\}} \right)^{1/2} \left(\sum_{j=1}^{\hat{X}_0} 1_{\{bt \leq B(j, t) \leq t\}} \right)^{1/2}. \end{aligned} \quad (4.4.5)$$

The first term in the second line of (4.4.5) has already been controlled by equation (4.2.10), i.e.,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{j=1}^{\hat{X}_0} D(j, t)^2 1_{\{S(j, t)=1, B(j, t)>0\}} \leq n(C_{4.2.10}t + \epsilon), \forall 0 \leq t \leq \gamma_n \wedge 1 \right) = 1. \quad (4.4.6)$$

Let $N(t)$ be the number of rewiring events that occur by time t . Then we have

$$\sum_{j=1}^{\hat{X}_0} 1_{\{bt \leq B(j, t) \leq t\}} \leq N(t) - N(bt). \quad (4.4.7)$$

Now we write down the evolution equation for $N(t)$

$$N(t) = \int_0^t q(u) du + M_{7,t},$$

where

$$q(t) = \rho \hat{X}_{I,t} \frac{\hat{X}_t - 1}{\lambda \hat{X}_{I,t}} \leq \frac{\rho \hat{X}_0}{\lambda} \quad (4.4.8)$$

and $M_{7,t}$ is some martingale. The assumption $\mathbb{E}(D^5) < \infty$ implies that the event $\Omega_n^* = \{\hat{X}_0 \leq 2m_1 n\}$ has probability tending to 1 as $n \rightarrow \infty$. On Ω_n^* , using (4.4.8) we have

$$q(t) \leq 2m_1 \rho n / \lambda.$$

It follows that

$$\lim_{n \rightarrow \infty} \mathbb{P}(q(t) \leq 2m_1 \rho n / \lambda, \forall t \geq 0) = 1. \quad (4.4.9)$$

Note that $N(t)$ is a pure jump process with jump size equal to 1. It follows that the expected value of quadratic variation of $M_{7,t}$ up to time 1 is upper bounded by

$$\mathbb{E} \left(\sup_{t \geq 0} q(t) \right) \leq \frac{\rho}{\lambda} \mathbb{E}(\hat{X}_0) \leq \frac{m_1 \rho}{\lambda}.$$

This implies that for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{7,t}| \leq \epsilon n \right) = 1. \quad (4.4.10)$$

From the definition of $N(t)$, we see that for $t \leq 1$,

$$|N(t) - N(bt)| = \left| \int_{bt}^t q(u) du + M_{7,t} - M_{7,bt} \right| \leq \left| \int_{bt}^t q(u) du \right| + 2 \sup_{0 \leq t \leq \gamma_n \wedge 1} |M_{7,t}|.$$

Thus by (4.4.9) and (4.4.10),

$$\lim_{n \rightarrow \infty} \mathbb{P}(|N(t) - N(bt)| \leq Cn((1-b)t + \epsilon), \forall 0 \leq t \leq \gamma_n) = 1. \quad (4.4.11)$$

Combining (4.4.7) and (4.4.11),

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{j=1}^{\hat{X}_0} 1_{\{bt \leq B(j,t) \leq t\}} \leq Cn((1-b)t + \epsilon) \right) = 1. \quad (4.4.12)$$

Equation (4.4.4) now follows from (4.4.5), (4.4.6) and (4.4.12). $\square$

Let t_0 and λ_0 be the two constants given in the statement of Lemma 4.4. Based on the calculations that led to (4.4.54) we let

$$\begin{aligned} U(t) = & C_{4.4.54} \left[\exp \left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t} \right) (\lambda - \lambda_c)t + (1-b)t \right. \\ & \left. + \exp \left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t} \right) t^2 + (\lambda - \lambda_c)t^2 + t^3 \right] + C_{4.4.52}\sqrt{\epsilon}. \end{aligned} \quad (4.4.13)$$

Lemma 4.6. For any $\lambda < \lambda_c + \lambda_0$, whp for all $0 \leq t \leq t_0 \wedge \gamma_n$ we have

$$X(I, b, t) \leq U(t)n. \quad (4.4.14)$$

Proof. We define the events $H_k(i, t)$, $1 \leq k \leq 4$ for $1 \leq i \leq \hat{X}_0$ (i is any half-edge) as follows.

$$\begin{aligned} H_1(i, t) &= \{I(i, bt) = 1, i \text{ didn't get rewired or paired in } [bt, t]\}, \\ H_2(i, t) &= \{I(i, bt) = 1, i \text{ got rewired to an infected vertex at its first rewiring in } [bt, t]\}, \\ H_3(i, t) &= \{I(i, bt) = 1, i \text{ got rewired to a susceptible vertex at its first rewiring in } [bt, t] \\ &\quad \text{and that vertex later became infected in } [bt, t]\}, \\ H_4(i, t) &= \{0 < A(i, t) \leq bt, S(i, bt) = 1, v(i, bt) \text{ got infected in } [bt, t]\}. \end{aligned}$$

We claim that

$$\{A(i, t) \leq bt, I(i, t) = 1\} \subset \cup_{k=1}^4 H_k(i, t).$$

Indeed, either $I(i, bt) = 1$ or $S(i, bt) = 1$ must hold. The case of $S(i, bt) = 1$ corresponds to $H_4(i, t)$. On the other hand, if $I(i, bt) = 1$ and $I(i, t) = 1$, then there are three possible cases: i didn't get rewired in $[bt, t]$, i was rewired to an infected vertex or i was rewired to a susceptible vertex which later became infected. The first case corresponds to $H_1(i, t)$ while the second and third case are covered in $H_2(i, t)$ and $H_3(i, t)$, respectively.

Let $H_k(t)$ be the number of half-edges i for which $H_k(i, t)$ occurs. It follows from the claim and the definition of $X(I, b, t)$ that

$$X(I, b, t) \leq \sum_{k=1}^4 H_k(t). \quad (4.4.15)$$

We first estimate $H_1(t)$. The rate for a half-edge to be paired at time u is $(\hat{X}_u - 1)/(\lambda \hat{X}_{I,u})$. Hence conditionally on $\hat{X}_{I,s}$, $bt \leq s \leq t$, $H_1(t)$ is stochastically dominated by a

$$\text{Binomial} \left(\hat{X}_{I,bt}, \exp \left(-\int_{bt}^t \frac{\hat{X}_u}{\lambda \hat{X}_{I,u}} du \right) \right).$$

random variable. Here we have a binomial distribution because the Poisson clocks on different infected half-edges are independent of each other. Let $C_{4.4.16} = \max\{2\rho m_1/\lambda_c^2, m_1\Delta, 1\}$. Then the event

$$L_1(n) = \{\widehat{X}_{I,t} \leq C_{4.4.16}n((\lambda - \lambda_c)t + t^2 + \epsilon), \widehat{X}_t \geq nm_1 \exp(-2t)/2, \forall 0 \leq t \leq \gamma_n \wedge t_0\} \quad (4.4.16)$$

has probability tending to 1 as $n \rightarrow \infty$ by (4.3.2) of Lemma 4.4 and (4.3.11). On the event $L_1(n)$, we have that

$$\int_{bt}^t \frac{\widehat{X}_u}{\widehat{X}_{I,u}} du \geq \int_{bt}^t \frac{m_1 \exp(-2u)/2}{C_{4.4.16}((\lambda - \lambda_c)u + u^2 + \epsilon)} du \geq \frac{C_{4.4.17}e^{-2t}(1-b)t}{(\lambda - \lambda_c)t + t^2 + \epsilon}. \quad (4.4.17)$$

Thus on $L_1(n)$, $H_1(t)$ is stochastically dominated by

$$W_0(t) := \text{Binomial}\left(C_{4.4.16}n((\lambda - \lambda_c)t + t^2 + \epsilon), \exp\left(-\frac{C_{4.4.17}e^{-2t}(1-b)t}{(\lambda - \lambda_c)t + t^2 + \epsilon}\right)\right). \quad (4.4.18)$$

For $\sqrt{\epsilon} \leq t \leq t_0$, we have

$$\frac{(1-b)t}{(\lambda - \lambda_c)t + t^2 + \epsilon} \geq \frac{(1-b)t}{2((\lambda - \lambda_c)t + t^2)} = \frac{1-b}{2(\lambda - \lambda_c + t)}.$$

Hence $W_0(t)$ is stochastically dominated by

$$W_1(t) := \text{Binomial}\left(C_{4.4.16}n((\lambda - \lambda_c)t + t^2 + \epsilon), \exp\left(-\frac{C_{4.4.17}e^{-2t}(1-b)}{2(\lambda - \lambda_c + t)}\right)\right) \quad (4.4.19)$$

for $\sqrt{\epsilon} \leq t \leq t_0$. Define

$$U_0(t) = C_{4.4.16}((\lambda - \lambda_c)t + t^2 + \epsilon) \exp\left(-\frac{C_{4.4.17}e^{-2t}(1-b)}{2(\lambda - \lambda_c + t)}\right). \quad (4.4.20)$$

For $t > \sqrt{\epsilon}$, there exists a constant $C_{4.4.21} = C_{4.4.21}(\lambda_0, t_0, \epsilon)$ depending on λ_0 , t_0 and ϵ such that

$$U_0(t) \geq C_{4.4.21}. \quad (4.4.21)$$

We need a large deviations bound for sums of Bernoulli random variables.

Lemma 4.7. Consider n i.i.d. Bernoulli random variables $Y_1, \dots, Y_n$. Let $\mu = \sum_{k=1}^n \mathbb{E}(Y_k)$. Then we have

$$\mathbb{P}\left(\sum_{k=1}^n Y_k \geq 3\mu\right) \leq \exp(-\mu). \quad (4.4.22)$$

Proof of Lemma 4.7. By [25, Theorem 2.3.1], we have

$$\mathbb{P}\left(\sum_{k=1}^n Y_k \geq 3\mu\right) \leq \exp(-\mu) \left(\frac{e\mu}{3\mu}\right)^{3\mu} \leq \exp(-\mu). \quad \square$$

Using Lemma 4.7, we have, for $t \geq \sqrt{\epsilon}$,

$$\mathbb{P}(W_1(t) > 3U_0(t)n) \leq \exp(-U_0(t)n) \leq \exp(-C_{4.4.21}n). \quad (4.4.23)$$

Using the definition of $L_1(n)$ in (4.4.16) and the fact that for $\sqrt{\epsilon} \leq t \leq t_0$, $W_1(t)$ dominates $W_0(t)$ which in turn dominates $H_1(t)$ (see (4.4.18)), we get

$$\mathbb{P}(\{H_1(t) > 3U_0(t)n\} \cap L_1(n)) \leq \exp(-C_{4.4.21}(\lambda_0, t_0, \epsilon)n). \quad (4.4.24)$$

For $0 \leq t \leq \sqrt{\epsilon}$, on the event $L_1(n)$,

$$\widehat{X}_{I,t} \leq C_{4.4.16}n((\lambda - \lambda_c)t + t^2 + \epsilon) \leq C_{4.4.16}n((\lambda - \lambda_c)\sqrt{\epsilon} + \epsilon + \epsilon) \leq C_{4.4.25}n\sqrt{\epsilon}. \quad (4.4.25)$$

for sufficiently small ϵ . Since $H_1(t) \leq \widehat{X}_{I,t}$ (by the definition of $H_1(t)$),

$$H_1(t) \leq C_{4.4.25}n\sqrt{\epsilon} \quad \text{for } t \leq \sqrt{\epsilon}.$$

Thus if we define

$$U_1(t) = U_0(t) + C_{4.4.25}\sqrt{\epsilon}, \quad (4.4.26)$$

then we have

$$\mathbb{P}(\{H_1(t) > 3U_1(t)n\} \cap L_1(n)) \leq \exp(-C_{4.4.21}n) \quad (4.4.27)$$

for all $0 \leq t \leq t_0$. Setting $t_0^\ell = t_0\ell/n^{3/2}$, $0 \leq \ell \leq n^{3/2}$, we get

$$\mathbb{P}\left(\left(\bigcup_{\ell=0}^{n^{3/2}-1}\{H_1(t_0^\ell) > 3U_1(t_0^\ell)n\}\right) \cap L_1(n)\right) \leq n^{3/2}\exp(-C_{4.4.21}n). \quad (4.4.28)$$

Denote the oscillation of $H_1(t)$ in $[t_0^\ell, t_0^{\ell+1}]$ by $\omega(H_1(t), t_0^\ell, t_0^{\ell+1})$. Here, the oscillation of any function (deterministic or random) $g(t)$ in an interval $[a, b]$ is defined to be

$$\sup_{a \leq t_1 \leq t_2 \leq b} |g(t_1) - g(t_2)|.$$

Consider the event that there is at most six pairings (i.e., a half-edge pairs with another half-edge) and rewirings (i.e., a half-edge is rewired to another vertex) occurring in $[bt_0^\ell, bt_0^{\ell+1}] \cup [t_0^\ell, t_0^{\ell+1}]$ and denote it by Ω^ℓ . On Ω^ℓ we have

$$\omega(H_1(t), t_0^\ell, t_0^{\ell+1}) \leq 6 \max_{1 \leq i \leq n} Q_i, \quad (4.4.29)$$

where Q_i is the number of half-edges that vertex i has before it becomes infected. By (4.4.29), Markov's inequality and (4.2.19) (together with $Q_i^5 \geq Q_i^4$) we have

$$\mathbb{P}(\{\omega(H_1(t), t_0^\ell, t_0^{\ell+1}) \geq \epsilon n\} \cap \Omega^\ell) \leq \mathbb{P}\left(6 \max_{1 \leq i \leq n} Q_i \geq \epsilon n\right) \leq \frac{6^4 \mathbb{E}(\sum_{i=1}^n Q_i^4)}{\epsilon^4 n^4} \leq \frac{C}{\epsilon^4 n^3}. \quad (4.4.30)$$

Now we control the probability of $(\Omega^\ell)^c$. Note that the rate for a rewiring or pairing to occur is equal to $(\rho \widehat{X}_{I,t} + \lambda \widehat{X}_{I,t})(\widehat{X}_t - 1)/(\lambda \widehat{X}_{I,t})$ which is bounded by $(\lambda + \rho)\widehat{X}_0/\lambda$. On the event $\Omega_n^* := \{\widehat{X}_0 \leq 2m_1 n\}$ (which holds with high probability) this quantity is upper bounded by $2(\rho + \lambda)m_1 n/\lambda$. Using this we get

$$\mathbb{P}((\Omega^\ell)^c \cap \Omega_n^*) \leq C(n \cdot n^{-3/2})^6 \leq Cn^{-3}. \quad (4.4.31)$$

Combining (4.4.30) and (4.4.31) we get

$$\mathbb{P}(\{\omega(H_1(t), t_0^\ell, t_0^{\ell+1}) > \epsilon n\} \cap \Omega_n^*) \leq \frac{C}{\epsilon^4 n^3}. \quad (4.4.32)$$

By the union bound for probabilities,

$$\mathbb{P}\left(\left(\bigcup_{\ell=0}^{n^{3/2}-1}\{\omega(H_1(t), t_0^\ell, t_0^{\ell+1}) > \epsilon n\}\right) \cap \Omega_n^*\right) \leq \frac{C}{\epsilon^4 n^{3/2}}. \quad (4.4.33)$$

Combining (4.4.28), (4.4.33) and the facts $\mathbb{P}(L_1(n)) \rightarrow 1, \mathbb{P}(\Omega_n^*) \rightarrow 1$, we get

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left(\bigcup_{\ell=0}^{n^{3/2}-1}\{\omega(H_1(t), t_0^\ell, t_0^{\ell+1}) \leq \epsilon n\}\right) \cap \left(\bigcap_{\ell=1}^{n^{3/2}-1}\{H_1(t_0^\ell) \leq 3U_1(t_0^\ell)n\}\right)\right) = 1. \quad (4.4.34)$$

On the event

$$\left(\cap_{\ell=0}^{n^{3/2}-1} \{w(H_1(t), t_0^\ell, t_0^{\ell+1}) \leq \epsilon n\}\right) \cap \left(\cap_{\ell=1}^{n^{3/2}} \{H_1(t_0^\ell) \leq 3U_1(t_0^\ell)n\}\right),$$

we necessarily have $H_1(t) \leq 3U_1(t)n + \epsilon n \leq 4U_1(t)n$ for all $0 \leq t \leq t_0 \wedge \gamma_n$. Hence we get

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_1(t) \leq 4U_1(t)n, \forall 0 \leq t \leq t_0 \wedge \gamma_n) = 1. \quad (4.4.35)$$

Now we turn to the control of $H_2(t)$. We set the event

$$L_2(n) := \{\widehat{I}_t \leq (2m_1t + \epsilon)n, \forall 0 \leq t \leq \gamma_n \wedge t_0\}.$$

By equation (4.3.4) we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(L_2(n)) = 1. \quad (4.4.36)$$

On $L_2(n)$, $H_2(t)$ is stochastically dominated by

$$W_2(t) := \text{Binomial}(C_{4.4.37}n((\lambda - \lambda_c)t + t^2 + \epsilon), C'_{4.4.37}(t + \epsilon)). \quad (4.4.37)$$

for some constants $C_{4.4.37}$ and $C'_{4.4.37}$. Now we define

$$U_2(t) = C_{4.4.37}((\lambda - \lambda_c)t + t^2 + \epsilon)C'_{4.4.37}(t + \epsilon) + C_{4.4.25}\sqrt{\epsilon}. \quad (4.4.38)$$

Following the proof of (4.4.35), one can derive analogous inequalities to (4.4.27) and (4.4.33) for $H_2(t)$. Combining these two inequalities we obtain that

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_2(t) \leq 4U_2(t), \forall 0 \leq t \leq t_0 \wedge \gamma_n) = 1. \quad (4.4.39)$$

We omit further details.

It remains to control $H_3(t)$ and $H_4(t)$. For any vertex x , let $R(x)$ be the indicator function of the event that vertex x has received at least one rewired edge when x first becomes infected and let Q_x be the number of half-edges that x has just before it becomes infected. Let $R(x, t)$ be the indicator of the event that x has received at least one rewired half-edge by time t . Then we have, by the definitions of $H_3(t)$ and $H_4(t)$,

$$H_3(t) + H_4(t) \leq \sum_{x=1}^n R(x, t) Q_x 1_{\{x \text{ was infected in } [bt, t]\}}. \quad (4.4.40)$$

Denote the right hand side of (4.4.40) by $N(bt, t)$, then we can decompose $N(bt, t)$ into a drift part and a martingale part for any fixed t :

$$N(bt, t) = \int_{bt}^t \bar{h}_t(u) du + M_{8,t}. \quad (4.4.41)$$

Let $D_x(u)$ be the number of half-edges of vertex x at time u and $D(j, u)$ the number of half-edges that $v(j, u)$ has at time u (recall that $v(j, u)$ is the vertex that half-edge j is attached to at time u). The process $N(bt, u)$, $bt \leq u \leq t$ has a positive jump whenever a susceptible vertex with at least one rewired half-edge gets infected. The probability that x is infected (given an infection event occurs) is equal to $D_x(u)/(\widehat{X}_u - 1)$ and the contribution to $N(bt, t)$ is equal to $R(x, u)D_x(u)$. Thus $\bar{h}_t(u)$ satisfies

$$\begin{aligned} \bar{h}_t(u) &\leq \lambda \widehat{X}_{I,u} \frac{\widehat{X}_u - 1}{\lambda \widehat{X}_{I,u}} \cdot \frac{\sum_{x=1}^n R(x, u) D_x^2(u) 1_{\{x \text{ is susceptible at time } u\}}}{\widehat{X}_u - 1} \\ &\leq \sum_{j=1}^{\widehat{X}_0} D(j, u)^2 1_{\{S(j,t)=1, B(j,t)>0\}}, \end{aligned} \quad (4.4.42)$$

where the second inequality follows from changing the order of summation:

$$\begin{aligned}
 & \sum_{v=1}^n D_x^2(u) 1_{\{x \text{ is susceptible at time } u\}} R(x, u) \\
 & \leq \sum_{x=1}^n D_x^2(u) 1_{\{x \text{ is susceptible at time } u\}} \sum_{j=1}^{\hat{X}_0} 1_{\{v(j,u)=x, B(j,u)>0\}} \\
 & = \sum_{j=1}^{\hat{X}_0} \sum_{x=1}^n 1_{\{v(j,u)=x\}} D(j, u)^2 1_{\{S(j,u)=1, B(j,u)>0\}} \\
 & = \sum_{j=1}^{\hat{X}_0} D(j, u)^2 1_{\{S(j,u)=1, B(j,u)>0\}}.
 \end{aligned} \tag{4.4.43}$$

In the first step of (4.4.43) we used the definition of $R(x, u)$ so that

$$R(x, u) \leq \sum_{j=1}^{\hat{X}_0} 1_{\{v(j,u)=x, B(j,u)>0\}}.$$

We denote the event

$$L_3(n) = \left\{ \sum_{j=1}^{\bar{X}_0} D(j, t)^2 1_{\{S(j,t)=1, B(j,t)>0\}} \leq n(C_{4.2.10}t + \epsilon), \forall 0 \leq t \leq \gamma_n \wedge 1 \right\}.$$

Then by (4.2.10) we have $\mathbb{P}(L_3(n)) \rightarrow 1$ as $n \rightarrow \infty$. On $L_3(n)$, using (4.4.42) and setting $C_{4.4.44} = C_{4.2.10}$, we see that for all $0 \leq t \leq t_0$,

$$\bar{h}_t(u) \leq n(C_{4.4.44}t + \epsilon), \forall bt \leq u \leq t. \tag{4.4.44}$$

The definition of $N(bt, t)$ as the right hand side of (4.4.40) implies that we can upper bound the quadratic variation of $M_{8,t}$ by $\sum_{i=1}^n Q_i^2$ where Q_i is the number of half-edges that vertex i has before it becomes infected. Using this and the Burkholder-Davis-Gundy inequality we have

$$\mathbb{E}(M_{8,t}^4) \leq C\mathbb{E}\left[\left(\sum_{i=1}^n Q_i^2\right)^2\right] \leq C\mathbb{E}\left(n \sum_{i=1}^n Q_i^4\right) \leq C'n^2. \tag{4.4.45}$$

The second inequality in (4.4.45) is due to the Cauchy-Schwartz inequality

$$\left(\sum_{i=1}^n (Q_i^2)^2\right) \left(\sum_{i=1}^n 1^2\right) \geq \left(\sum_{i=1}^n Q_i^2\right)^2,$$

and the third inequality follows from (4.2.19). Using (4.4.45) we have

$$\mathbb{P}(|M_{8,t}| \geq \epsilon n) \leq \frac{\mathbb{E}(M_{8,t}^4)}{\epsilon^4 n^4} \leq \frac{C}{\epsilon^4 n^2}. \tag{4.4.46}$$

Now using (4.4.40), (4.4.41), (4.4.44) and (4.4.46), we get, for any $0 \leq t \leq t_0$,

$$\mathbb{P}(L_3(n) \cap \{H_3(t) + H_4(t) \geq (1-b)t(C_{4.4.44}t + \epsilon)n + \epsilon n\}) \leq \frac{C}{\epsilon^4 n^2}. \tag{4.4.47}$$

Define

$$U_3(t) = (1-b)t(C_{4.4.44}t + \epsilon) + \epsilon. \tag{4.4.48}$$

Now we can repeat the proof of (4.4.35) (i.e., divide $[0, t_0]$ into $n^{3/2}$ intervals and use a union bound) to get

$$\lim_{n \rightarrow \infty} \mathbb{P}(H_3(t) + H_4(t) \leq 2U_3(t)n, \forall 0 \leq t \leq t_0 \wedge \gamma_n) = 1. \quad (4.4.49)$$

Combining (4.4.35), (4.4.39) and (4.4.49), we have that with high probability

$$X(I, b, t) \leq 5(U_1(t) + U_2(t) + U_3(t))n.$$

Using (4.4.20), (4.4.26), (4.4.38), and (4.4.48), the right-hand side is

$$\begin{aligned} & 5 \left[C_{4.4.16}((\lambda - \lambda_c)t + t^2 + \epsilon) \exp\left(-\frac{C_{4.4.17}e^{-2t}(1-b)}{2(\lambda - \lambda_c + t)}\right) \right. \\ & \quad + (C_{4.4.37}(\lambda - \lambda_c)t + t^2 + \epsilon)C'_{4.4.37}(t + \epsilon) + 2C_{4.4.25}\sqrt{\epsilon} \\ & \quad \left. + (1-b)t(C_{4.4.44}t + \epsilon) + \epsilon \right] n. \end{aligned} \quad (4.4.50)$$

To make the computation easier to write we note that when $t \leq t_0 \leq 1$,

$$\exp\left(-\frac{C_{4.4.17}e^{-2t}(1-b)}{2(\lambda - \lambda_c + t)}\right) \leq \exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) := F. \quad (4.4.51)$$

Expanding the terms in 4.4.50 and putting the terms with ϵ or $\sqrt{\epsilon}$ together, we bound (4.4.50) by

$$\begin{aligned} & 5 \left[C_{4.4.16}((\lambda - \lambda_c)t + t^2)F + (C'_{4.4.37}(\lambda - \lambda_c)t + t^2)C'_{4.4.37}t \right. \\ & \quad \left. + C_{4.4.44}(1-b)t \right] n + C_{4.4.52}\sqrt{\epsilon}n. \end{aligned} \quad (4.4.52)$$

We also used the fact that $\epsilon \leq \sqrt{\epsilon}$ in (4.4.52). Sorting the terms by powers of t we get

$$\begin{aligned} & 5 \left[C_{4.4.16}F(\lambda - \lambda_c)t + C_{4.4.44}(1-b)t \right. \\ & \quad \left. + C_{4.4.16}Ft^2 + C_{4.4.37}C'_{4.4.37}[(\lambda - \lambda_c)t^2 + t^3] \right] n + C_{4.4.52}\sqrt{\epsilon}n. \end{aligned} \quad (4.4.53)$$

Simplifying constants we have

$$\begin{aligned} X(I, b, t) & \leq C_{4.4.54} \left[\exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) (\lambda - \lambda_c)t \right. \\ & \quad \left. + \exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) t^2 + (\lambda - \lambda_c)t^2 + (1-b)t^2 + t^3 \right] n + C_{4.4.52}\sqrt{\epsilon}n, \end{aligned} \quad (4.4.54)$$

which completes the proof of Lemma 4.6. $\square$

4.5 Completing the proof of Theorem 1.15

Proof of Theorem 1.15. We set $\gamma_n = \inf\{t > 0 : \hat{X}_{I,t} = 0\}$. We now condition on a large outbreak so that $\hat{I}_\infty > \eta n$ for some fixed $\eta > 0$. By (4.3.4), we see that, conditionally on $\hat{I}_\infty > \eta n$, with high probability $\gamma_n > \epsilon$ for some $\epsilon > 0$. That is,

$$\lim_{n \rightarrow \infty} \mathbb{P}(\gamma_n > \epsilon | \hat{I}_\infty/n > \eta) = 1. \quad (4.5.1)$$

Let t_0 and λ_0 be given by the statement of Lemma 4.4. We let $\lambda_1 < \lambda_0, t_1 < t_0$ be two constants (independent of ϵ) and $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5, \epsilon_6$ be some small numbers (depending on

ϵ) to be determined. Recall the definition of $U(t)$ in (4.4.13). We set $C_{4.4.52}\sqrt{\epsilon}$ in (4.4.13) to be ϵ_4 . In other words,

$$U(t) = C_{4.4.54} \left(\exp \left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t} \right) (\lambda - \lambda_c)t + \left(\exp \left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t} \right) + (\lambda - \lambda_c) + (1-b) \right) t^2 + t^3 \right) + \epsilon_4, \quad (4.5.2)$$

Previous results imply that the following inequalities hold whp on $0 \leq t \leq \gamma_n \wedge t_0$. The numbers on the left give the formula numbers for these statements.

$$(4.3.3) \quad \hat{X}_{I,t} \geq \left(\frac{\rho m_1(\lambda - \lambda_c)}{2\lambda_c^2} t + \frac{m_1 \Delta}{4} t^2 - \epsilon_2 \right) n - \int_0^t E(u) du.$$

$$(4.3.5) \quad \hat{I}_t \geq \left(\frac{m_1 t}{2} - \epsilon_5 \right) n - \int_0^t E(u) du.$$

$$(4.3.25) \quad \int_0^t E(u) du \leq (C_{4.3.25} t^2 + \epsilon_1) n.$$

$$(4.3.24) \quad L(t) \leq n(C_{4.2.10} t + \epsilon_1).$$

$$(4.4.4) \quad L(b, t) \leq n(C_{4.4.4}(1-b)^{1/2} t + \epsilon_3).$$

$$(4.4.14) \quad X(I, b, t) \leq U(t)n.$$

Let Ω_n be the event that all of the last six formulas together with the event $\{\gamma_n > \epsilon\}$ hold. Combining (4.5.1) and the fact that $\liminf_{n \rightarrow \infty} \mathbb{P}(\hat{I}_\infty/n > \eta) > 0$ (since $\lambda > \lambda_c$),

$$\lim_{n \rightarrow \infty} \mathbb{P}(\Omega_n | \hat{I}_\infty/n > \eta) = 1. \quad (4.5.3)$$

Now we define

$$\tau = \inf \left\{ \epsilon \leq t \leq t_0 \wedge \gamma_n : \int_0^t E(u) du > \left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2} t + \frac{m_1 \Delta}{8} t^2 + \epsilon_6 \right) n \right\}, \quad (4.5.4)$$

where $\inf \emptyset$ here is set to be $t_0 \wedge \gamma_n$. We want to select the parameters λ_1 and t_1 so that whenever $\lambda - \lambda_c < \lambda_1$ and the outcome is in Ω_n , we have

$$t_1 \leq \tau < \gamma_n, \quad (4.5.5)$$

$$\hat{I}_{t_1} \geq \left(\frac{m_1 t_1}{2} - \epsilon_5 \right) n - \int_0^{t_1} E(u) du > \frac{m_1 t_1}{8} n. \quad (4.5.6)$$

This implies

$$\frac{\hat{I}_\infty}{n} > \frac{m_1 t_1}{8},$$

which proves (1.5.7), as desired. We now divide the proof of (4.5.5) and (4.5.6) into five steps:

- In Step 1, we choose appropriate ϵ_1 , ϵ_2 and ϵ_6 to ensure $\gamma_n > \tau > \epsilon$.
- In Step 2, we show that, under conditions (4.5.19) and (4.5.20) below, there exists a constant t_1 such that $\tau \geq t_1$.
- In Step 3, we show that (4.5.19) and (4.5.20) can be satisfied by choosing appropriate values of the parameters involved. The first three steps combined give (4.5.5).
- In Step 4, we prove (4.5.6).
- Finally, we summarize the choices of the parameters.

Step 1. We first show that $\tau > \epsilon$. Using (4.3.25),

$$\int_0^\epsilon E(u) du \leq (C_{4.3.25}\epsilon^2 + \epsilon_1)n, \quad (4.5.7)$$

We set

$$\epsilon_1 = C_{4.3.25}\epsilon^2, \quad (4.5.8)$$

and

$$\epsilon_2 = \epsilon_6 = \frac{\rho m_1(\lambda - \lambda_c)}{16\lambda_c^2}\epsilon. \quad (4.5.9)$$

We now require

$$C_{4.3.25}\epsilon^2 + \epsilon_1 = 2C_{4.3.25}\epsilon^2 \leq \frac{\epsilon_6}{2} = \frac{\rho m_1(\lambda - \lambda_c)}{32\lambda_c^2}\epsilon, \quad (4.5.10)$$

which holds true if $2C_{4.3.25}\epsilon^2 \leq \rho m_1(\lambda - \lambda_c)\epsilon/32\lambda_c^2$ or equivalently,

$$\epsilon < \frac{\rho m_1(\lambda - \lambda_c)}{64C_{4.3.25}\lambda_c^2}. \quad (4.5.11)$$

We now prove $\tau > \epsilon$. Indeed, using (4.5.7)-(4.5.10),

$$\int_0^\epsilon E(u)du \leq \frac{\rho m_1(\lambda - \lambda_c)}{32\lambda_c^2}\epsilon n < \left(\frac{\rho m_1(\lambda - \lambda_c)}{2\lambda_c^2}t + \frac{m_1\Delta}{4}t^2 - \epsilon_2 \right) n.$$

This proves $\tau > \epsilon$ by the definition of τ .

Now we show that $\gamma_n > \tau$. If $\gamma_n > t_0$ then this is trivial. So we assume that $\gamma_n \leq t_0$. Equation (4.3.3) and the definition of τ in (4.5.4) imply that, for $\epsilon < t \leq \tau$,

$$\begin{aligned} \hat{X}_{I,t} &\geq \left(\frac{\rho m_1(\lambda - \lambda_c)}{2\lambda_c^2}t + \frac{m_1\Delta}{4}t^2 - \epsilon_2 \right) n - \int_0^t E(u)du \\ &\geq \left(\frac{\rho m_1(\lambda - \lambda_c)}{2\lambda_c^2}t + \frac{m_1\Delta}{4}t^2 - \epsilon_2 \right) n - \left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}t + \frac{m_1\Delta}{8}t^2 + \epsilon_6 \right) n \\ &\geq \left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}t + \frac{m_1\Delta}{8}t^2 - \epsilon_2 - \epsilon_6 \right) n. \end{aligned} \quad (4.5.12)$$

Note that when $t > \epsilon$,

$$\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}t + \frac{m_1\Delta}{8}t^2 - \epsilon_2 - \epsilon_6 > \frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}\epsilon + \frac{m_1\Delta}{8}\epsilon^2 - \epsilon_2 - \epsilon_6 > 0,$$

due to (4.5.9). This proves $\gamma_n > \tau$.

Step 2. We turn now to the second requirement that $\tau \geq t_1$ for some appropriately chosen constant t_1 . We will show this by contradiction. Assume $\tau < t_1$ which is also smaller than t_0 . Since we have already proved $\gamma_n > \tau > \epsilon$ in **Step 1**, it follows that

$$\int_0^\tau E(u)du = \left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}\tau + \frac{m_1\Delta}{8}\tau^2 + \epsilon_6 \right) n. \quad (4.5.13)$$

We split the integral $\int_0^\tau E(u) du$ into two parts: $\int_0^\epsilon E(u) du$ and $\int_\epsilon^\tau E(u) du$. This first part has already been controlled in (4.5.7) and (4.5.10). For the second part $\int_\epsilon^\tau E(u)du$, (4.4.3) implies that

$$E(t) \leq E_1(t) + E_2(t) = \frac{X(I, b, t)}{\hat{X}_{I,t}}L(t) + L(b, t). \quad (4.5.14)$$

For $\epsilon < t < \tau$, using (4.4.15) and (4.5.12), we get

$$\frac{X(I, b, t)}{\widehat{X}_{I, t}} \leq \frac{U(t)n}{\left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}t + \frac{m_1\Delta}{8}t^2 - \epsilon_2 - \epsilon_6\right)n}. \quad (4.5.15)$$

Using (4.5.9) and $t > \epsilon$, we see

$$\left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}t + \frac{m_1\Delta}{8}t^2 - \epsilon_2 - \epsilon_6\right)n \geq \max\left\{\frac{\rho m_1(\lambda - \lambda_c)}{8\lambda_c^2}tn, \frac{m_1\Delta}{8}t^2n, \epsilon_2n\right\}.$$

Hence using (4.5.2) and (4.5.15) we get

$$\begin{aligned} \frac{X(I, b, t)}{\widehat{X}_{I, t}} &\leq \frac{C_{4.4.54} \exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) (\lambda - \lambda_c)t}{\frac{\rho m_1(\lambda - \lambda_c)}{8\lambda_c^2}t} \\ &\quad + \frac{C_{4.4.54} \left(\exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) + (\lambda - \lambda_c) + (1 - b) + t\right) t^2}{\frac{m_1\Delta}{8}t^2} + \frac{\epsilon_4}{\epsilon_2} \\ &\leq \exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda - \lambda_c + t}\right) \left(\frac{8C_{4.4.54}\lambda_c^2}{\rho m_1} + \frac{8C_{4.4.54}}{m_1\Delta}\right) \\ &\quad + \frac{8C_{4.4.54}}{m_1\Delta} ((\lambda - \lambda_c) + (1 - b) + t) + \frac{\epsilon_4}{\epsilon_2}. \end{aligned} \quad (4.5.16)$$

Let $V(\lambda, t)$ denote the quantity on the last two lines of (4.5.16). $V(\lambda, t)$ is increasing with respect to both λ and t . Therefore for $\lambda \leq \lambda_c + \lambda_1$ and $t \leq t_0$, which we have supposed is < 1 , $V(\lambda, t)$ is bounded above by its value at $(\lambda_c + \lambda_1, t_1)$

$$\begin{aligned} V^* &:= \exp\left(-\frac{C_{4.4.51}(1-b)}{\lambda_1 + t_0}\right) \left(\frac{8C_{4.4.54}\lambda_c^2}{\rho m_1} + \frac{8C_{4.4.54}}{m_1\Delta}\right) \\ &\quad + \frac{8C_{4.4.54}}{m_1\Delta} ((\lambda - \lambda_c) + (1 - b) + t_1) + \frac{\epsilon_4}{\epsilon_2}. \end{aligned} \quad (4.5.17)$$

Using (4.5.14), (4.5.17), (4.3.24), and (4.4.4) we see that

$$\begin{aligned} \int_{\epsilon}^{\tau} E(u) du &\leq \int_{\epsilon}^{\tau} \frac{X(I, b, u)}{\widehat{X}_{I, u}} L(u) du + \int_{\epsilon}^{\tau} L(b, u) du \\ &\leq \int_{\epsilon}^{\tau} V^*(C_{4.2.10}u + \epsilon_1)n du + \int_{\epsilon}^{\tau} n(C_{4.4.4}(1-b)^{1/2}u + \epsilon_3) du \\ &\leq \left(V^*C_{4.2.10}\frac{\tau^2}{2} + V^*\epsilon_1\tau + C_{4.4.4}(1-b)^{1/2}\frac{\tau^2}{2} + \epsilon_3\tau\right)n. \end{aligned} \quad (4.5.18)$$

We want to choose our parameters so that

$$\frac{V^*C_{4.2.10} + C_{4.4.4}(1-b)^{1/2}}{2} \leq \frac{m_1\Delta}{16}, \quad (4.5.19)$$

$$V^*\epsilon_1 + \epsilon_3 \leq \frac{\epsilon_6}{2}. \quad (4.5.20)$$

Indeed, if (4.5.19) and (4.5.20) hold, then by (4.5.7), (4.5.10) and (4.5.18) we have

$$\int_0^{\tau} E(u) du = \int_0^{\epsilon} E(u) du + \int_{\epsilon}^{\tau} E(u) du \leq \left(\frac{\epsilon_6}{2} + \frac{m_1\Delta}{16}\tau^2 + \frac{\epsilon_6}{2}\right)n, \quad (4.5.21)$$

which is smaller than

$$\left(\frac{\rho m_1(\lambda - \lambda_c)}{4\lambda_c^2}\tau + \frac{m_1\Delta}{8}\tau^2 + \epsilon_6\right)n.$$

This contradicts with (4.5.13) and proves $\tau \geq t_1$ on Ω_n .

Step 3. (4.5.19) will hold if

$$V^* \leq \frac{m_1 \Delta}{32C_{4.2.10}} \quad \text{and} \quad (1-b)^{1/2} \leq \frac{m_1 \Delta}{32C_{4.4.4}}.$$

We have $\epsilon_1 \leq \epsilon_6/2$ by (4.5.10), so (4.5.20) will hold if

$$V^* \leq 1/4 \quad \text{and} \quad \epsilon_3 = \frac{\epsilon_6}{4}.$$

To make sure (4.5.19) and (4.5.20) are satisfied, it suffices to have

$$V^* \leq K := \min \left\{ \frac{m_1 \Delta}{32C_{4.2.10}}, \frac{1}{4} \right\}, \quad (1-b)^{1/2} \leq \frac{m_1 \Delta}{32C_{4.4.4}}, \quad \epsilon_3 = \frac{\epsilon_6}{4C_{4.4.4}}. \quad (4.5.22)$$

Using the definition of $V(\lambda, t)$ in (4.5.17), (4.5.22) can be satisfied if we first choose b sufficiently close to 1 such that

$$\frac{8C_{4.4.54}}{m_1 \Delta} (1-b) \leq \frac{K}{8}, \quad (4.5.23)$$

then choose λ_1, t_1 and ϵ_4 such that

$$\exp \left(-\frac{C_{4.4.51}(1-b)}{\lambda_1 + t_1} \right) \left(\frac{bC_{4.4.54}\lambda_c^2}{\rho m_1} + \frac{8C_{4.4.54}}{m_1 \Delta} \right) + \frac{8C_{4.4.54}}{m_1 \Delta} (\lambda - \lambda_c) + \frac{8C_{4.4.54}}{m_1 \Delta} t_1 < \frac{K}{8}, \quad (4.5.24)$$

and finally take

$$\epsilon_4 = \frac{\epsilon_2}{8} K. \quad (4.5.25)$$

Step 4. It remains to take care of (4.5.6). Using (4.3.5) and (4.3.25), we have

$$\hat{I}_{t_1} \geq \left(\frac{m_1 t_1}{2} - \epsilon_5 \right) n - \int_0^{t_1} E(u) du \geq \left(\frac{m_1 t_1}{2} - \epsilon_5 - C_{4.3.25} t_1^2 - \epsilon_1 \right) n. \quad (4.5.26)$$

Recall that we set $\epsilon_1 = C_{4.3.25} \epsilon^2$ in (4.5.8). Thus we have

$$\hat{I}_{t_1} \geq \left(\frac{m_1 t_1}{2} - \epsilon_5 - C_{4.3.25} t_1^2 - C_{4.3.25} \epsilon^2 \right) n, \quad (4.5.27)$$

which is bigger than $nm_1 t_1/8$ if

$$t_1 \leq \frac{m_1}{16C_{4.3.25}}, \quad \epsilon_5 = \frac{m_1 t_1}{16}, \quad \epsilon < \left(\frac{m_1 t_1}{16C_{4.3.25}} \right)^{1/2}. \quad (4.5.28)$$

Summary we can first choose b close to 1, then take λ_1 and t_1 sufficiently small such that (4.5.24) holds true and

$$t_1 \leq \min \left\{ \frac{m_1}{16C_{4.3.25}}, t_0 \right\}.$$

Then we let ϵ be smaller than

$$\min \left\{ \frac{\rho m_1 (\lambda - \lambda_c)}{64C_{4.3.25}\lambda_c^2}, \left(\frac{m_1 t_1}{16C_{4.4.54}} \right)^{1/2} \right\}.$$

Finally we determine $\epsilon_1, \dots, \epsilon_6$ using (4.5.9), (4.5.8), (4.5.22), (4.5.25) and (4.5.28).

$$\begin{aligned} \epsilon_2 = \epsilon_6 &= \frac{\rho m_1 (\lambda - \lambda_c)}{16\lambda_c^2} \epsilon, \quad \epsilon_1 = C_{4.3.25} \epsilon^2, \quad \epsilon_3 = \frac{\epsilon_6}{4}, \\ \epsilon_4 &= \frac{\epsilon_2}{8} \min \left\{ \frac{m_1 \Delta}{32C_{4.2.10}}, \frac{1}{4} \right\}, \quad \epsilon_5 = \frac{m_1 t_0}{16}. \end{aligned} \quad \square$$

References

- [1] Håkan Andersson and Tom Britton, *Stochastic epidemic models and their statistical analysis*, Lecture Notes in Statistics, vol. 151, Springer-Verlag, New York, 2000. MR1784822
- [2] Frank Ball and Tom Britton, *Epidemics on networks with preventive rewiring*, Random Structures Algorithms **61** (2022), no. 2, 250–297. MR4456028
- [3] Tom Britton, David Juher, and Joan Saldaña, *A network epidemic model with preventive rewiring: comparative analysis of the initial phase*, Bull. Math. Biol. **78** (2016), no. 12, 2427–2454. MR3571300
- [4] Tom Britton and Philip D. O'Neill, *Bayesian inference for stochastic epidemics in populations with random social structure*, Scand. J. Statist. **29** (2002), no. 3, 375–390. MR1925565
- [5] R. W. R. Darling and J. R. Norris, *Differential equation approximations for Markov chains*, Probab. Surv. **5** (2008), 37–79. MR2395153
- [6] Laurent Decreusefond, Jean-Stéphane Dhersin, Pascal Moyal, and Viet Chi Tran, *Large graph limit for an SIR process in random network with heterogeneous connectivity*, Ann. Appl. Probab. **22** (2012), no. 2, 541–575. MR2953563
- [7] Rick Durrett, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 20, Cambridge University Press, Cambridge, 2007. MR2271734
- [8] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: characterization and convergence*, John Wiley and Sons, 2009. MR0838085
- [9] N. Fountoulakis, *Percolation on sparse random graphs with given degree sequence*, Internet Math. **4** (2007), no. 4, 329–356. MR2522948
- [10] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR1943877
- [11] Svante Janson, *The largest component in a subcritical random graph with a power law degree distribution*, Ann. Appl. Probab. **18** (2008), no. 4, 1651–1668. MR2434185
- [12] Svante Janson, *On percolation in random graphs with given vertex degrees*, Electron. J. Probab. **14** (2009), no. 5, 87–118. MR2471661
- [13] Svante Janson, Malwina Łuczak, and Peter Windridge, *Law of large numbers for the SIR epidemic on a random graph with given degrees*, Random Structures Algorithms **45** (2014), no. 4, 726–763, [Paging previously given as 724–761]. MR3275704
- [14] Yufeng Jiang, Remy Kassem, Grayson York, Mathew Junge, and Rick Durrett, *SIR epidemics on evolving graphs*, arXiv preprint arXiv:1901.06568 (2019).
- [15] Fima C. Klebaner, *Introduction to stochastic calculus with applications*, second ed., Imperial College Press, London, 2005. MR2160228
- [16] Ka Yin Leung, Frank Ball, Sirl David, and Tom Britton, *Individual preventive social distancing during an epidemic may have negative population-level outcomes*, Journal of the Royal Society Interface **15** (2018), no. 145.
- [17] Anders Martin-Löf, *Symmetric sampling procedures, general epidemic processes and their threshold limit theorems*, J. Appl. Probab. **23** (1986), no. 2, 265–282. MR839984
- [18] Joel C. Miller, *A note on a paper by Erik Volz: SIR dynamics in random networks [mr2358436]*, J. Math. Biol. **62** (2011), no. 3, 349–358. MR2771177
- [19] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Random Structures Algorithms **6** (1995), no. 2-3, 161–179. MR1370952
- [20] Michael Molloy and Bruce Reed, *The size of the giant component of a random graph with a given degree sequence*, Combin. Probab. Comput. **7** (1998), no. 3, 295–305. MR1664335
- [21] Peter Neal, *SIR epidemics on a Bernoulli random graph*, J. Appl. Probab. **40** (2003), no. 3, 779–782. MR1993267
- [22] Remco van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, [43], Cambridge University Press, Cambridge, 2017. MR3617364

- [23] Remco van der Hofstad, *Stochastic processes on random graphs*, Lecture notes for the 47th Summer School in Probability Saint-Flour (2017).
- [24] Remco van der Hofstad, *Random graphs and complex networks*, vol. 2, 2020+, To appear in Cambridge Series in Statistical and Probabilistic Mathematics. Available at <https://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf>. MR3617364
- [25] Roman Vershynin, *High-dimensional probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, 2018, An introduction with applications in data science, With a foreword by Sara van de Geer. MR3837109
- [26] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR1964483
- [27] Erik Volz, *SIR dynamics in random networks with heterogeneous connectivity*, J. Math. Biol. **56** (2008), no. 3, 293–310. MR2358436

Acknowledgments. We are grateful to the referees for constructive comments.