SIR epidemics on random graphs with a fixed degree sequence

Tom Bohman* Michael Picollelli†

September 12, 2011

Abstract

Let $\Delta > 1$ be a fixed positive integer. For $\mathbf{z} \in \mathbb{R}_+^\Delta$ let $G_{\mathbf{z}}$ be chosen uniformly at random from the collection of graphs on $\|\mathbf{z}\|_1 n$ vertices that have $z_i n$ vertices of degree i for $i = 1, \ldots, \Delta$. We determine the likely evolution in continuous time of the SIR model for the spread of an infectious disease on $G_{\mathbf{z}}$, starting from a single infected node. Either the disease halts after infecting only a small number of nodes, or an epidemic spreads to infect a linear number of nodes. Conditioning on the event that more than a small number of nodes are infected, the epidemic is likely to follow a trajectory given by the solution of an associated system of ordinary differential equations. These results also give the likely number of nodes infected during the course of the epidemic and the likely length in time of the epidemic.

1 Introduction

The Susceptible-Infectious-Recovered (SIR) model is a classical epidemiological model for the spread of a disease on a network. Let $G = (V, E)$ be a graph; this is the network on which the epidemic spreads. Let $\lambda > 0$ and $\rho > 0$ be fixed constants. These are the rate of infection and the rate of recovery for the disease in question. The vertex set V is partitioned into three parts: a set S of nodes that are susceptible to infection, a set I of infectious nodes and a set R of recovered nodes. Each infectious node sends infections to each of its neighbors (as determined by the edge set E) at times determined by independent exponential random variables with parameter λ. An infectious node recovers at a random time given by an independent exponential random variable with parameter ρ. A recovered node sends no more infections and is immune to further infection. Such a node will remain in the set R even if it receives an infection. Susceptible nodes move to the set I when they receive infections. Throughout this note we assume that the epidemic begins with a single (and arbitrary) infected node. The methods introduced here extend to other initial conditions, but we do not discuss these extensions for the sake of simplicity.

We note that classical epidemiological theory is not based on a fixed network but rather an assumption that each individual in a population has a probability of being infected that is proportional to the number of infected individuals in the population. There has

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, USA. Supported in part by NSF grants DMS-0701183 and DMS-1001638. E-mail: tbohman@math.cmu.edu

†Department of Electrical & Computer Engineering, University of Delaware, Newark, DE, USA; E-mail: mpicolle@udel.edu

been a growing interest in SIR and related models on specific networks in recent years as such models are much closer to the reality of the spread of disease in human (and other) populations. For example, real world populations are heterogeneous in the sense that different individuals may play very different roles with respect to the spread of a certain disease. (E.g. in the case of the spread of sexually transmitted diseases some individuals will have far more contacts than others and therefore will play a far greater role in an epidemic.) However, there are a number of difficulties in arriving at specific networks that are useful in modeling real world populations. Thus there has been an interest in considering graphs drawn at random from collections of networks with properties that mimic those of real world networks [24] [33]. In this paper we analyze the propagation of an SIR epidemic on graphs chosen at random from the collection of graphs with some fixed degree sequence and a finite maximum degree. For a treatment of the SIR model on the Erdős-Renyi random graph $G_{n,p}$ see Neal [31].

Let $\Delta > 0$ be a fixed positive integer. For $\mathbf{z} \in \mathbb{R}_+^\Delta$ let $G_\mathbf{z}$ be chosen uniformly at random from the collection of graphs on $\|\mathbf{z}\|_1 n$ vertices that have $z_i n$ vertices of degree i for $i = 1, \dots, \Delta$. This model has been the subject of much recent study. There is considerable interest in the analysis of the phase transition [29] [23] [16] and the size of the largest component in the sub-critical phase [30] [18] [34]. Results have also been obtained on the diameter [14], the k-core [20] [21], the matching number [10] and the chromatic number [15]. Here we study the likely evolution of the SIR model for the spread of an infectious disease on $G_\mathbf{z}$, starting from a single infected node. A discrete time analog of the SIR epidemic on $G_\mathbf{z}$ was considered by Andersson [2],[3] as well as Britton, Janson, and Martin-Löf [13]. In this paper we work with the continuous time SIR model and obtain a detailed description of the continuous time evolution of the epidemic.

We work with the configuration model for generating a random graph with a specified degree sequence, which we now briefly describe. We refer the reader to [22] for a more thorough discussion. Every vertex v is associated with a set of $d(v)$ *configuration points* (or configuration *positions*, or *half-edges*), where $d(v)$ is the degree of v. We choose a perfect matching on the set of all configuration positions uniformly at random. Such a perfect matching is called a *configuration*. Given a configuration, we generate a multigraph (i.e. a graph in which there may be loops or multiple edges) on the original vertex set by placing edges between the vertices associated with matched pairs of configuration positions. For the degree sequences we consider here, it is known (see [19]) that the probability a random configuration yields a simple graph (i.e. neither loops nor multiple edges) is bounded away from 0 as the number of vertices grows. Additionally, conditioned on the multigraph being simple, the configuration model provides a uniform distribution on the set of labeled simple graphs with the given degrees. Finally, we note that if we condition on the event that some fixed collection of pairs appear in a random configuration, then the partner of any unmatched configuration point is uniform on the set of (other) unmatched points.

Our probability space is the product of a randomly chosen configuration and sequences of exponential random variables that determine the evolution of an SIR epidemic. Our results are stated relative to this product space. Given this assumed probability space, we can view the epidemic and the underlying graph as being revealed together; that is, we only observe a match between configuration points when an infection is being sent along the edge in question. To be precise, when a node v becomes infected we sample an independent exponential random variable with mean λ^{-1} for each of the remaining (i.e. unmatched) configuration points associated with v and an independent exponential random variable with mean ρ^{-1}. The mean λ^{-1} exponentials are the transmission times for infection along the associated edges while the the mean ρ^{-1} exponential is the recovery time of v. When an infection corresponding to one of these configuration points occurs (assuming this happens before v recovers) we choose the matched configuration point uniformly at random from the collection of remaining configuration points and proceed.

Throughout this note, we assume that n is fixed and sufficiently large, and probability asymptotics are stated in the limit as $n \to \infty$. All random variables we introduce are defined on the product space outlined above and hence depend implicitly on n. In the course of our discussion we will introduce some limit objects. These are fixed functions on the reals and absolute constants (that, of course, do not depend on n). We emphasize that all constants hidden in $O(\cdot)$ or $\Theta(\cdot)$ notation depend only on the parameters $\lambda, \rho, \mathbf{z}$ and are independent of n.

In order to summarize our results, we introduce some random variables. Let $S(t), Y(t)$ and $R(t)$ be the sets of susceptible, infectious and recovered vertices, respectively, at time t. Let the random variable t_{end} be the smallest time $t > 0$ such that $Y(t) = \emptyset$; this time marks the end of the spread of the disease. Note further that $R(t_{\text{end}})$ is the set of nodes that were infected in the course of spread of the disease. Of course, our results depend on the relationships between the parameter $\mathbf{z}, \lambda$ and ρ. Define

$$\|\mathbf{z}\| = \sum_{i=1}^{\Delta} iz_i$$

and

$$\hat{d} = \hat{d}_{\mathbf{z}} = \sum_{i=2}^{\Delta} \frac{(i-1)iz_i}{\|\mathbf{z}\|}.$$

Note that $\hat{d}$ is the expected degree of the vertex corresponding to a randomly chosen configuration position after we remove the chosen configuration position. (This parameter plays a central role in the appearance of a giant component in $G_{\mathbf{z}}$, see [29] [16].) We begin by noting that if the rate of change in the number of infectious configuration position is not positive then the disease quickly dies out. Let $\mathcal{S}$ be the event that $R(t_{\text{end}})$ contains at most $\log^{4/3} n$ vertices.

Theorem 1. *If $\lambda(\hat{d}-1) - \rho < 0$ then $Pr(\mathcal{S}) \to 1$.*

Now we turn to the more interesting case $\lambda(\hat{d}-1) - \rho > 0$, where there is a nonzero probability of a full-scale epidemic. It turns out that the full-scale phase of the epidemic is very likely to closely follow a fixed trajectory. The lengths of time before and after the full-scale phase (these are time periods during which the number of infectious vertices is small) are not tightly concentrated. To describe this phenomena we introduce some more notation. Let $Q_Y(t)$ be the number of remaining configuration points that correspond to vertices in $Y(t)$ (these are the infectious configuration positions). For $\epsilon > 0$ set

$$t_{\epsilon\uparrow} = \inf\{t : Q_Y(t) > \epsilon n\} \qquad t_{\epsilon\downarrow} = \inf\{t > t_{\epsilon\uparrow} : Q_Y(t) < \epsilon n/2\}.$$

(If $Q_Y(t) < \epsilon n$ for all t then we set $t_{\epsilon\uparrow} = t_{\epsilon\downarrow} = \infty$.)

Theorem 2. *Let $\lambda(\hat{d}-1)-\rho>0$. There exists a constant $p<1$ such that $Pr(\mathcal{S})\to p$. Furthermore, there is a constant M and a continuous function $g:\mathbb{R}\to\mathbb{R}_+$ with $\lim_{x\to-\infty}g(x)=\lim_{x\to\infty}g(x)=0$ such that for $\epsilon>0$ sufficiently small we have the following when we condition on $\overline{\mathcal{S}}$:*

(i) With high probability $t_{\epsilon\uparrow}=\Theta(\log n)$ and $t_{\text{end}}-t_{\epsilon\downarrow}=\Theta(\log n)$.

(ii) With high probability

$$Q_Y(t)=g(a+t-t_{\epsilon\uparrow})n\pm n^{1/2+o(1)}\qquad \text{for all }\ t\in[t_{\epsilon\uparrow},t_{\epsilon\uparrow}+b]$$

where a is the smallest solution of $g(a)=\epsilon$ and $b-a$ is the largest solution of $g(a+b)=\epsilon/2$. In particular, the difference $t_{\epsilon\downarrow}-t_{\epsilon\uparrow}$ tends to the fixed real b as n tends to infinity.

(iii) With high probability $|R(t_{\text{end}})|=Mn\pm n^{1/2+o(1)}$.

We emphasize that, while the limiting probability p that the disease infects only a small number of nodes depends on the degree of the initially infectious vertex, the function g and constant M that describe a full-scale epidemic are independent of the degree of the first infected vertex. While Theorem 2 gives a considerable amount of information about the evolution of an epidemic when one occurs, we obtain more detailed results below. For example, we explicitly determine the function g and constant M given in Theorem 2. We also determine the trajectories of the degree sequences of susceptible and infectious vertices.

Our results both extend previous work and provide rigorous verification of results found heuristically. Andersson [3], for example, considered a discrete-time SIR epidemic on $G_{\mathbf{z}}$, establishing that the initial growth over time is essentially given by that of a branching process. An analysis of the evolution over time for a continuous-time SIR model was later obtained by Volz [37] (cf. Miller [27]) through a heuristic analysis of a system of differential equations and supported by simulation results. Britton, Janson and Martin-Löf [13] also considered a discrete-time SIR epidemic on $G_{\mathbf{z}}$, initiated with a single infected individual, and established conditions for a full-scale epidemic (i.e. $|R(t_{\text{end}})|=\Theta(n)$) as well as the final asymptotic size within a $o(n)$ error. Theorems 1 and 2, then, can be viewed as a natural extension of these results to the continuous-time SIR epidemic. And Theorem 2 also offers an additional conclusion: while the entire epidemic is likely to run its course in $\Theta(\log(n))$ time, the duration in which a significant proportion of individuals actively infect is essentially constant. Barbour [8] (cf. Britton [12]) proved that the full-scale epidemic phase of the classical continuous time SIR model (in which there is no underlying graph) occurs in constant time, and Theorem 2 shows that SIR epidemics on sparse networks exhibit similar behavior.

While our account of the evolving epidemic is mostly complete, there is a major deficiency in these results: the fact that our results are restricted to initial degree sequences with a bounded maximum degree. An extension of our analysis to unbounded degree sequences remains an open issue.

The remainder of the paper is organized as follows. In the next Section we give a more detailed description of the epidemic, including a discussion of the function g and the constant M in Theorem 2. Theorems 1 and 2 follow more or less directly from the detailed description of the evolution contained in Theorems 3, 4, 5 and 6 (given in Section 2). A comparison of our results and methods with the literature follows in Section 2.1. Section 3 contains some technical preliminaries. In Section 4 we analyze the beginning of the spread of the disease and prove Theorem 3. The epidemic phase is studied in Section 5, in which we prove Theorem 4 and 5. Finally, in Section 6 we consider the end of the spread of the disease and prove Theorem 6.

2 Evolution in Detail

To achieve a more detailed description of the evolution (and prove Theorems 1 and 2) we discretize and re-parameterize time. We introduce a sequence of random times. Set $t_0 = 0$, $t_1 = 1$ and

$$t_j = \begin{cases} t_{j-1} + \frac{1}{Q_Y(t_{j-1})} & \text{if } Q_Y(t_{j-1}) \neq 0 \\ \infty & \text{otherwise} \end{cases}$$

for $j \geq 2$. We emphasize that t_j is a random variable. The main virtue of this parametrization is that the expected changes in all random variables in one time step are bounded by an absolute constant.

Now we are ready to discuss the initial spread of the disease. Define

$$\begin{aligned} X_{\text{end}} &= \inf\{i \geq 0 | Q_Y(t_i) = 0\} \\ X_{\uparrow} &= \inf\{i \geq 0 | Q_Y(t_i) \geq n^{\alpha}\} \end{aligned}$$

where $\alpha > 0$ is small and fixed. (We could set $\alpha = 1/4$, but in an attempt to make the paper easier to read we do not assign a fixed value to α.) Further define

$$\begin{aligned} F_I(i) = &\text{ the number of infections sent in the interval } [0, t_i] \\ F_R(i) = &\text{ the number of recoveries in the interval } [0, t_i] \end{aligned}$$

Theorem 3. *(i) If $\lambda(\hat{d} - 1) < \rho$ then with high probability we have*

$$X_{\text{end}} \leq \log n \qquad \textit{and} \qquad F_I(X_{\text{end}}) + F_R(X_{\text{end}}) \leq \log^{4/3} n.$$

(ii) If $\lambda(\hat{d} - 1) > \rho$ then with high probability we have either

$$X_{\text{end}} \leq \log n \qquad \textit{and} \qquad F_I(X_{\text{end}}) + F_R(X_{\text{end}}) \leq \log^{4/3} n \tag{1}$$

or

$$t_{X_\uparrow} = O(\log n) \qquad \textit{and} \qquad F_I(X_\uparrow) + F_R(X_\uparrow) \leq O\left(n^{4\alpha/3}\right). \tag{2}$$

Furthermore, the probability of the event given in (1) tends to a fixed constant $p < 1$ (where this constant p depends on the degree of the initial infected vertex).

(N.b. The lower bound on $t_{\epsilon\uparrow}$ in part (i) of Theorem 2 follows from Theorem 4.) We remark in passing that the $\log n$ that appears in Part (i) of Theorem 3 and (1) can be replaced by any function that tends to infinity as $n \to \infty$.

We describe the evolution of the full-scale phase of the epidemic by showing that the evolution, properly scaled, closely follows its expected trajectory. Before describing the scaling and trajectory, we list the random variables we track. Set

$$\begin{aligned} S_i(t) &= \text{ the number of susceptible vertices of degree } i \\ Y_i(t) &= \text{ the number of infectious vertices of degree } i \end{aligned}$$

Furthermore, let $Q(t)$ be the total number of remaining configuration positions and let $Q_R(t)$ be the number of remaining configuration points that correspond to vertices in R. (Note that we have $Q_R(t) = Q(t) - Q_Y(t) - \sum_{i=1}^{\Delta} iS_i(t)$.) All of these variables have linear scaling in the sense that the random variables divided by n are concentrated around some continuous function. The scaling of time is more delicate. We scale time relative to the random times steps t_j by introducing a 'time' variable x which is related to time steps t_j by the linear scaling that associates t_j with $x \approx j/n$. Note that the 'time' variable x has a useful interpretation with respect to the evolution of the epidemic: Since the expected number of infections sent in the time interval $[t_j, t_{j+1})$ is roughly λ, at a particular value of x roughly $\lambda x n$ infections have been sent. Now we introduce our trajectories. Define

$$q_Y(x) = -\frac{\rho\|\mathbf{z}\|}{\lambda}\sqrt{\frac{\|\mathbf{z}\| - 2\lambda x}{\|\mathbf{z}\|}} + \left(1 + \frac{\rho}{\lambda}\right)(\|\mathbf{z}\| - 2\lambda x) - \sum_{i=1}^{\Delta} i z_i \left(\frac{\|\mathbf{z}\| - 2\lambda x}{\|\mathbf{z}\|}\right)^{i/2}.$$

Furthermore, for $0 \le x \le \|\mathbf{z}\|/(2\lambda)$ define

$$\begin{aligned} q(x) &= \|\mathbf{z}\| - 2\lambda x, \\ s_i(x) &= z_i \left(\frac{\|\mathbf{z}\| - 2\lambda x}{\|\mathbf{z}\|}\right)^{i/2}, \\ q_R(x) &= \frac{\rho\|\mathbf{z}\|}{\lambda}\sqrt{\frac{\|\mathbf{z}\| - 2\lambda x}{\|\mathbf{z}\|}} - \frac{\rho}{\lambda}(\|\mathbf{z}\| - 2\lambda x). \end{aligned} \tag{3}$$

We stress that $q_Y(x), q(x), s_i(x)$ and $q_R(x)$ are *not* function of n; rather, they are the 'limit' objects that describe the evolution of the epidemic when n is large. Let x^* be the least positive root of $q_Y(x)$. (The properties of $q_Y(x)$ are discussed in Section 5.1 below.) Define

$$\sigma = \sigma(n) = \sup\{x < x^* : q_Y(x) \ge \log^4(n) n^{-1/2}\}.$$

It is useful to 're-start' time at $X_\uparrow$. Set $u_0 = t_{X\uparrow}$ and

$$u_{i+1} = t_{X_\uparrow + i + 1} = u_i + \frac{1}{Q_Y(u_i)}$$

for $i \ge 0$. Now we are ready to state our first dynamic concentration result.

Theorem 4. *Let $\lambda(\hat{d} - 1) > \rho$. If we condition on the event $\overline{\mathcal{S}}$ then with high probability we have*

$$\begin{aligned} Q(u_k) &= nq(k/n) + O\left(\log^2(n) n^{1/2}\right) \\ Q_R(u_k) &= nq_R(k/n) + O\left(\log^3(n) n^{1/2}\right) \\ S_i(u_k) &= ns_i(k/n) + O\left(\log^3(n) n^{1/2}\right) \quad \text{for } i = 1, \ldots, \Delta \end{aligned}$$

for $k = 0, \ldots, \sigma n$.

Note that the evolution of $Q_Y(u_k)$ is also determined by Theorem 4. In particular, if we condition on the event $\overline{\mathcal{S}}$ then with high probability we have

$$Q_Y(u_k) = nq_Y(k/n) + O\left(\log^3(n)n^{1/2}\right) \tag{4}$$

for $k = 0, \ldots, \sigma n$. However the values of $Y_1(u_k), \ldots, Y_{\Delta-1}(u_k)$ are not determined by Theorem 4. They are excluded because the system of differential equations that governs their evolution is not Lipschitz when $Q_Y(t)$ is small. So the standard methods do not apply. However, we do manage to overcome this to establish the trajectory of $Y_1(u_k), \ldots, Y_{\Delta-1}(u_k)$ over most of the epidemic. For $\eta > 0$ a constant define

$$\sigma_\eta = \max\{x < x^* : q_Y(x) \geq \eta\}.$$

Theorem 5. *Let $\lambda(\hat{d}-1) > \rho$. There exist functions $y_1(x), \ldots, y_{\Delta-1}(x)$ (these are given by the solution of (19) below) such that for any $\eta > 0$ constant we have the following: If we condition on the event $\overline{\mathcal{S}}$ then with high probability*

$$Y_i(u_k) = ny_i(k/n) + O\left(n^{3/4}\right) \quad \textit{for } i = 1, \ldots, \Delta - 1$$

and $k = 0, \ldots, \sigma_\eta n$.

A more precise account of the evolution of $Q_Y(t)$ and $Y_i(t)$ for $t_{X_\uparrow} < t < t_{\epsilon\uparrow}$ is given in Section 5 below. (More precise than given by (4) and Theorem 5.) Indeed, such an account is necessary for our proof of Theorem 4 because of the central role $Q_Y(t)$ plays in the process. Furthermore, as we have already noted the usual methods do not apply to the early evolution of $Y_i(t)$ as the expected one-step changes in these variables do not satisfy a useful Lipschitz condition when $Q_Y(t)$ is small. The methods developed to overcome this technical obstacle are a central contribution of this paper and are discussed in more detail in the proof of Theorem 5 in Section 5 below. This completes (for the moment) our description of the full-scale phase of the epidemic.

The following Theorem describes the evolution beyond time $u_{\sigma n}$.

Theorem 6. *Let $\lambda(\hat{d}-1) > \rho$. If we condition on the event $\overline{\mathcal{S}}$ then with high probability, $X_{end} - \sigma n = \Theta(\log^4(n)\sqrt{n})$ and $t_{\text{end}} - u_{\sigma n} = \Theta(\log n)$.*

As the step by step changes in all the relevant random variables are bounded (see Lemma 17) we arrive at the following Corollary of Theorems 4 and 6.

Corollary 7. *Let $\lambda(\hat{d}-1) > \rho$. If we condition on the event $\overline{\mathcal{S}}$ then with high probability for $u_{\sigma n} \leq t \leq t_{\text{end}}$ we have*

$$\begin{aligned}
Q(t) &= nq(x^*) + O\left(\log^5(n)\sqrt{n}\right) \\
Q_R(t) &= nq_R(x^*) + O\left(\log^5(n)\sqrt{n}\right) \\
S_i(t) &= ns_i(x^*) + O\left(\log^5(n)\sqrt{n}\right) \quad \textit{for } i = 1, \ldots, \Delta
\end{aligned}$$

We close this Section with a discussion of the constant M and the function g of Theorem 2. From Corollary 7 we can infer that

$$M = \sum_{i=1}^{\Delta} (z_i - s_i(x^*)),$$

where the functions s_i are explicitly defined in (3).

To prove part (ii) of Theorem 2 we use the results above along with the facts that, for some $\delta > 0$, we have $q'_Y > 0$ on an interval containing $[0, \delta]$, $q'_Y < 0$ on an interval containing $[x^* - \delta, x^*]$, q'_Y is bounded on $[0, x^*]$, and q_Y is bounded below on an open interval containing $[\delta, x^* - \delta]$. (These properties follow from the observations in Section 5.1.) Consider the function $f : (0, x^*) \to \mathbb{R}$ defined by

$$f(x) = \int_{x^*/2}^{x} \frac{1}{q_Y(w)} dw.$$

Note that f is a bijection, and define $g : \mathbb{R} \to \mathbb{R}_+$ by

$$g(s) = q_Y(f^{-1}(s)).$$

Next, let $a = \inf\{s : g(s) \geq \epsilon\}$ and $b = \sup\{s : g(a + s) = \epsilon/2\}$. Note further that we have $f^{-1}(a) = \inf\{x > 0 : q_Y(x) = \epsilon\}$. For $t_{\epsilon\uparrow} \leq t \leq t_{\epsilon\uparrow} + b$ define x_t by $f(x_t) = a + t - t_{\epsilon\uparrow}$, that is,

$$\int_{f^{-1}(a)}^{x_t} \frac{1}{q_Y(w)} \, dw = t - t_{\epsilon\uparrow}.$$

As

$$g(a + t - t_{\epsilon\uparrow}) = g(f(x_t)) = q_Y(x_t),$$

it suffices to show that if the estimate (4) holds then $Q_Y(t) = q_Y(x_t)n + O(\log^3(n)\sqrt{n})$ for $t_{\epsilon\uparrow} \leq t \leq t_{\epsilon\uparrow} + b$.

Let $t_{\epsilon\uparrow} \leq t \leq t_{\epsilon\uparrow} + b$. There exist (random) steps ℓ_0 and ℓ_1 such that $t_{\epsilon\uparrow} \in [u_{\ell_0}, u_{\ell_0+1})$ and $t \in [u_{\ell_1}, u_{\ell_1+1})$. It follows from (4) that we have $\ell_0 = f^{-1}(a)n + O(\log^3(n)\sqrt{n})$. Furthermore,

$$\begin{aligned}
t - t_{\epsilon\uparrow} &= \sum_{j=\ell_0}^{\ell_1} u_{j+1} - u_j \pm O\left(\frac{1}{n}\right) \\
&= \sum_{j=\ell_0}^{\ell_1} \frac{1}{Q_Y(u_j)} \pm O\left(\frac{1}{n}\right) \\
&= \sum_{j=\ell_0}^{\ell_1} \frac{1}{nq_Y(j/n) \pm O\left(\log^3(n)n^{1/2}\right)} \pm O\left(\frac{1}{n}\right) \\
&= \left[\sum_{j=\ell_0}^{\ell_1} \frac{1}{nq_Y(j/n)}\right] \pm O\left(\log^3(n)n^{-1/2}\right) \\
&= \left[\int_{\ell_0/n}^{\ell_1/n} \frac{1}{q_Y(w)} \, dw\right] \pm O\left(\log^3(n)n^{-1/2}\right).
\end{aligned}$$

It follows that $\ell_1 = x_t n + O\left(\log^3(n)n^{1/2}\right)$. Again appealing to (4) we have

$$
\begin{aligned}
Q_Y(t) &= Q_Y(u_{\ell_1}) + O(\log(n)) \\
&= nq_Y(\ell_1/n) + O\left(\log^3(n)n^{1/2}\right) \\
&= nq_Y(x_t) + O\left(\log^3(n)n^{1/2}\right).
\end{aligned}
$$

This is the desired estimate.

2.1 Comparison with the literature

The more common approach (see, for example, [2], [3], [7], [13], [32], [37]) to modeling SIR epidemics on random networks with a specified degree sequence constructs the underlying graph by first selecting the degree of each vertex independently according to a common distribution. In this subsection, we assume that $\mathbf{z} \in \mathbb{R}_+^{\boldsymbol{\Delta}}$ satisfies $\sum_{i=1}^{\Delta} z_i = 1$ and that D_v is a random variable with distribution $\mathbf{z}$ (i.e. $Pr[D_v = i] = z_i$). (We choose not, in general, to set $\sum_i z_i = 1$, as our expressions are not simplified by such an assumption.) We begin this subsection by showing that our results agree with the known results for the model with degrees given by D_v, with exponential infection and recovery rates with parameters λ and ρ, respectively. We note that $E[D_v] = \|\mathbf{z}\|$ and $E[D_v(D_v - 1)] = \|\mathbf{z}\|\hat{d}$, defined in the Introduction. We follow Newman [32], but see also [2], [7].

One parameter of interest in this analysis is the transmissibility, p_T, defined as the probability that an infected individual infects a susceptible neighbor. As recovery rates and infection rates are independent and exponentially distributed, this is simply the probability that $\tau_I < \tau_R$, where $\tau_I \in Exp(\lambda)$ and $\tau_R \in Exp(\rho)$ are independent: direct integration yields

$$
p_T = 1 - \frac{\rho}{\lambda + \rho} = \frac{\lambda}{\lambda + \rho}.
$$

Let

$$
R_0 = p_T \cdot \frac{E[D_v(D_v - 1)]}{E[D_v]}.
$$

It is known that if $R_0 > 1$ then with non-zero probability an epidemic infects a positive fraction of the vertices, while if $R_0 \leq 1$ then with high probability only a vanishing proportion of the population is infected. As $R_0 = \frac{\lambda}{\lambda+\rho}\hat{d}$ (with $\hat{d}$ defined above), Theorems 1 and 2 yield essentially the same conclusions. (R_0 is typically referred to as the *basic reproductive number*.)

Suppose now that $R_0 > 1$ and a full-scale epidemic occurs. We turn to the the final proportion of the population infected. Let $G(x) = \sum_{i=1}^{\Delta} z_i x^i$, the probability generating function for D_v. Define q to be the probability that a vertex v at the end of a randomly chosen edge (or containing a randomly chosen configuration point) does *not* receive infection from any of its remaining $d(v) - 1$ edges (or configuration points) in the epidemic. Fix a vertex v of degree k and one of its configuration points, the probability that v receives an infection from the neighbor of that point is $p_T(1 - q)$ (since its neighbor must both become infected and transmit along that edge prior to recovery), and hence with probability $\zeta = 1 - p_T + p_T q$, v does not receive an infection from that configuration point. As these events are effectively (or, at least, asymptotically) independent for distinct configuration points within the same vertex, the probability a given vertex of degree k remains uninfected is ζ^k, and therefore the fraction of the population infected is

$$1 - G(\zeta).$$

It remains to compute ζ (equivalently, q). The probability that a random half-edge falls on a vertex of degree k is $kz_k/(\sum_k kz_k) = kz_k/E[D_v]$. And in this event, the probability that v does not receive an infection along one of the other edges incident with v is ζ^{k-1}. It follows that that ζ is the unique solution in $(0,1)$ to

$$1 - \frac{1}{p_T} + \frac{\zeta}{p_T} = q = \frac{\sum_{k=1}^{\Delta} kz_k\zeta^{k-1}}{\sum_k kz_k} = \frac{G'(\zeta)}{E[D_v]}. \tag{5}$$

This analysis is made rigorous in Andersson [3].

To arrive at the same conclusion from our analysis, we begin by claiming that

$$\zeta = \left(\frac{\|\mathbf{z}\| - 2\lambda x^*}{\|\mathbf{z}\|}\right)^{1/2} \tag{6}$$

is the unique solution of (5) in $(0,1)$. To see that this is the case, recall that we defined x^* to be the least positive real such that $q_Y(x^*) = 0$; that is, x^* is defined to be the least positive real such that

$$-\frac{\rho\|\mathbf{z}\|}{\lambda}\zeta + \left(\frac{\lambda+\rho}{\lambda}\right)\|\mathbf{z}\| \cdot \zeta^2 - \sum_{i=1}^{\Delta} iz_i\zeta^i = 0.$$

Note that $\zeta \in (0,1)$ follows from $x^* \in (0, \frac{\|\mathbf{z}\|}{2\lambda})$ (see Lemma 15 in Section 5.1). Dividing this expression this by $\|\mathbf{z}\|\zeta$ gives exactly (5). Our expression for the likely number of nodes that are infected in the course of an epidemic is

$$M = \sum_{i=1}^{\Delta}(z_i - s_i(x^*))n = \left[1 - \sum_{i=1}^{\Delta} z_i\zeta^i\right] n = [1 - G(\zeta)]\, n$$

as predicted above.

We note in passing that the fact that ζ as defined in (6) is the limit as $n \to \infty$ of the probability that a fixed vertex v gets an infection through a particular configuration point follows naturally from our analysis. To see this, we return to our random time-steps. Provided there are still infected vertices of positive degree at time t_j, we have $t_{j+1} - t_j = 1/Q_Y(t_j)$. As each configuration point in a vertex in $Y(t_j)$ has a probability of roughly $\lambda/Q_Y(t_j)$ of sending an infection in $[t_j, t_{j+1}]$, and as there are (by definition) $Q_Y(t_j)$ such points, it follows that we expect $\approx \lambda$ infections to be sent per time step. As the limiting continuous variable x is associated with these steps via the linear scaling $x = j/n$, our interpretation of x^* is that approximately $\lambda x^* n$ infections are sent throughout the entire epidemic. As we only expose edges when an infection is sent, it then follows that ζ (interpreted now as the probability that v is not infected through a particular configuration point) is also be the probability that a given configuration point is not paired during through $\lambda x^* n$ pairings. Therefore,

$$\begin{aligned}\zeta = \prod_{i=1}^{\lambda x^* n} \frac{\|\mathbf{z}\|n - 2i}{\|\mathbf{z}\|n - 2i + 1} &= \prod_{i=1}^{\lambda x^* n} \left(1 - \frac{1}{\|\mathbf{z}\|n - 2i + 1}\right) \\ &\approx \exp\left(-\sum_{i=1}^{\lambda x^* n} \frac{1}{\|\mathbf{z}\| - 2i + 1}\, d\tau\right) \\ &\approx \exp\left(-\int_0^{\lambda x^*} \frac{1}{\|\mathbf{z}\| - 2\tau}\, d\tau\right) \\ &= \left(\frac{\|\mathbf{z}\| - 2\lambda x^*}{\|\mathbf{z}\|}\right)^{1/2}.\end{aligned}$$

We close this Section by commenting on some of the similarities between our analysis and methods that are commonplace in epidemiology. (We refer the reader to the survey of Britton [12] for a more thorough discussion of the literature.)

One observation we make use of is that when the number of infecting configuration points is small relative to the size of the entire remaining graph, new infections are very unlikely to hit previously-infected vertices - this will allow us to couple the infection with a continuous time multitype branching process (see Sections 4 and 6), and the coupling remains valid until the first time an infected vertex is re-infected (also known as a *ghost contact*, see Mollison [28]). See [26], [5], [6], for example, for similar applications of branching processes in the analysis of epidemiological models.

As mentioned above, our choice of time-scaling will allow us to assume that, on average, λ infections are sent per unit time, and it will turn out that, on average, ρ configuration points move from the infected to the recovered set during that time as well. This allows us to determine the functions q, q_Y, q_R, s_i, y_i from the previous section as the solution to a system of ordinary differential equations (see Section 5). The use of differential equations to model SIR epidemic dynamics dates back to the work of Kermack and McKendrick [25]. A common approach to proving such approximations is to assume, say, an ε fraction of the population is initially infected, apply limit theorems for stochastic differential equations, and then taking the resulting limits as $\varepsilon \to 0$. In the context of a network model like $G_{\mathbf{z}}$, such an approach is highly suspect as there is some *distribution* of vertex degrees of infectious vertices when ϵn vertices are infected. A proper rigorous treatment must prove that the assumed degree sequence at this point is the one that is likely to be produced by the epidemic dynamics. In this paper, we determine this distribution in 2 steps. First, we utilize the branching process coupling to initially grow the infection to a large but explicitly sublinear size, roughly n^α infections, for some small $\alpha > 0$. From that point forward, we apply the differential equations method to show that the degree sequences are dynamically concentrated around their expected trajectory into and through the epidemic phase of the spread of the disease. Establishing dynamic concentration just after n^α infections have occurred is the main technical difficulty here (as the usual Lipschitz condition does not hold).

3 Preliminaries

3.1 Concentration Inequalities

Our arguments rely on martingale large-deviation inequalities. In most cases we use the standard Azuma-Hoeffding inequality (see [1] or [38]). We also make use of the following variation.

Lemma 8. *Let $0 \equiv X_0, X_1, \dots$ be a sequence of random variables measurable with respect to a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots$ such that $|X_{i+1} - X_i| < c$. Suppose there exists a nondecreasing function $h : \{0, \dots, t\} \to \mathbb{R}_+$ such that*

$$X_i > h(i) \quad \Rightarrow \quad E[X_{i+1} - X_i | \mathcal{F}_i] \leq 0.$$

Then for all $\alpha \geq 2c$,

$$Pr\left(\exists i, 0 \leq i \leq t : X_i \geq h(i) + \alpha\right) \leq (t+1) \exp\left(-\frac{\alpha^2}{8tc^2}\right).$$

An analogous result holds if $X_i < -h(i)$ implies $E[X_{i+1} - X_i | \mathcal{F}_i] \geq 0$.

Proof. For each r, $r = 0, \dots, t$, define a stopping time $U(r)$ to be the least $k \in \{r, \dots, t\}$ such that $X_k < h(k)$ if such a k exists, t otherwise. By the condition given, the sequence $X_r, \dots, X_{U(r)}$ forms a supermartingale. Therefore, by Azuma-Hoeffding and the union bound, the probability that there exists a pair (r, a) such that $r \leq a \leq U(r)$ and $X_a \geq X_r + \alpha/2$ is at most $(t+1)\exp\left(-\frac{\alpha^2}{8tc^2}\right)$. It therefore suffices to show the conclusion provided no such pair exists.

We inductively define the following sequence: $r_0 = 0$, and, if $U(r_{i-1}) < t$, $r_i = U(r_{i-1}) + 1$ for $i \geq 1$. We note that $X_r \leq X_0 + \alpha$ for $0 \leq r \leq U(0)$. Furthermore, for each r_i, $i > 1$, and each a such that $r_i \leq a \leq U(r_i)$ we have

$$X_a \leq X_{r_i} + \frac{\alpha}{2} \leq X_{U(r_{i-1})} + c + \frac{\alpha}{2} \leq h(r_i - 1) + c + \frac{\alpha}{2} \leq h(a) + \alpha.$$

□

3.2 Linear Algebra

Define

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 & \cdots & 0 \\ 2 & -2 & 0 & 0 & \cdots & 0 \\ 0 & 3 & -3 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & \Delta-2 & -(\Delta-2) & 0 \\ 0 & 0 & \cdots & 0 & \Delta-1 & -(\Delta-1) \end{pmatrix}. \tag{7}$$

This matrix plays a central role in our analysis of this process. Here we gather some elementary observations regarding A that will be used in the proofs of Theorems 3 and 6. We begin by noting that A is diagonalizable.

Lemma 9. *The matrix A has eigenvalues $-1, -2, \ldots, -(\Delta - 1)$.*

Proof. Let $\mathbf{u}$ be the vector

$$\mathbf{u} = (1 \;\; 2 \;\; \cdots \;\; (\Delta - 1)). \tag{8}$$

The vector $\mathbf{u}^T$ is a right eigenvector corresponding to the eigenvalue -1. For $i = 2, \ldots, \Delta - 1$, define $\mathbf{v}_i \in \mathbb{R}^{\Delta-1}$ by

$$(\mathbf{v}_i)_j = (-1)^j \binom{i}{j}. \tag{9}$$

Note that $\mathbf{v}_i$ is a left eigenvector of A corresponding to eigenvalue $-i$.

□

Recall that we have an initial degree distribution $\mathbf{z} \in \mathbb{R}_+^\Delta$, and as the process evolves an approximation for the distribution of the susceptibles, $\mathbf{s}(t) \in \mathbb{R}_+^\Delta$. (These correspond to actual degree sequences $\mathbf{Z}, \mathbf{S}(t) \in \mathbb{Z}_+^\Delta$. For example, the initial degree sequence is $S_i(0) = Z_i \in \{\lfloor z_i n \rfloor, \lceil z_i n \rceil\}$ for $i = 1, \ldots, \Delta$.) Given a triple of approximate distributions $\mathbf{s}, \mathbf{y}, q$ or a triple of degree sequences $\mathbf{S}(t), \mathbf{Y}(t), Q(t)$ define

$$\begin{aligned}
\mathbf{d}_s &= \begin{pmatrix} \frac{2s_2}{q} & \frac{3s_3}{q} & \cdots & \frac{\Delta s_\Delta}{q} \end{pmatrix} \\
\mathbf{D}_S &= \begin{pmatrix} \frac{2S_2}{Q} & \frac{3S_3}{Q} & \cdots & \frac{\Delta S_\Delta}{Q} \end{pmatrix}.
\end{aligned}$$

Note $||\mathbf{z}|| = q(0)$, and let

$$\mathbf{d} = \mathbf{d}_s(0) \qquad\qquad \mathbf{D} = \mathbf{D}_S(0).$$

We additionally define

$$\hat{d}_s = \mathbf{u}\mathbf{d}_s^T = \sum_{i=2}^{\Delta} \frac{(i-1)is_i}{q} \qquad\qquad \hat{D}_S = \mathbf{u}(\mathbf{D}_S)^T = \sum_{i=2}^{\Delta} \frac{(i-1)iS_i}{Q}.$$

Recall that $\hat{d} = \mathbf{u}\mathbf{d}^T$, and similarly define $\hat{D} = \hat{D}_S(0) = \mathbf{u}\mathbf{D}^T$.

We will often work with the matrix $A + \mathbf{u}^T\mathbf{d}$. The following observation follows from Lemma 9.

Corollary 10. *The matrix $A + \mathbf{u}^T\mathbf{d}$ is diagonalizable with right eigenvector $\mathbf{u}^T$ corresponding to eigenvalue $\hat{d} - 1$ and left eigenvectors $\mathbf{v}_i$ corresponding to $-i$, $i = 2, \ldots, \Delta - 1$.*

We will want to work with a full basis of left eigenvectors of $A + \mathbf{u}^T\mathbf{d}$. To this end we let $\mathbf{v}_1$ be the left eigenvector of $A + \mathbf{u}^T\mathbf{d}$ corresponding to eigenvalue $\hat{d} - 1$ with the property that $\mathbf{v}_1\mathbf{u}^T = 1$. We note an additional property of $\mathbf{v}_1$.

Claim 11. *If $\hat{d} > 1$ then $(\mathbf{v}_1)_j > 0$ for $j = 1, 2, \ldots, \Delta - 1$.*

Proof. We note that

$$(\hat{d} - 1)(\mathbf{v}_1)_{\Delta-1} = \left(\mathbf{v}_1(A + \mathbf{u}^T\mathbf{d})\right)_{\Delta-1} = -(\Delta - 1)(\mathbf{v}_1)_{\Delta-1} + \frac{\Delta z_\Delta}{q},$$

and, as $z_\Delta > 0$, we have $(\mathbf{v}_1)_{\Delta-1} > 0$. Now, assume for the sake of contradiction that there exists an index j such that $(\mathbf{v}_1)_j < 0$. Let $k = \max\{j : (\mathbf{v}_1)_j < 0\}$. Then, as $k < \Delta - 1$, we have

$$(\hat{d} - 1)(\mathbf{v}_1)_k = -k(\mathbf{v}_1)_k + (k+1)(\mathbf{v}_1)_{k+1} + \frac{(k+1)z_{k+1}}{q},$$

and consequently $(\mathbf{v}_1)_k > 0$, a contradiction. $\square$

We conclude this section by introducing notation for the change of basis between the basis $\mathbf{v}_1, \ldots, \mathbf{v}_{\Delta-1}$ of left eigenvectors of $A + \mathbf{u}^T\mathbf{d}$ and the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_{\Delta-1}$ of $\mathbb{R}^{\Delta-1}$. Define P to be the matrix with the following property: If $\mathbf{y} \in \mathbb{R}^{\Delta-1}$ and $\mathbf{y} = \sum_{i=1}^{\Delta-1} c_i \mathbf{v}_i$, then $c_i = (\mathbf{y}P)_i = \mathbf{y}P_i$, where P_i denotes the ith column of P.

3.3 Branching Processes

We couple both the initial and final phases of the epidemic with continuous-time multi-type branching processes, which are covered in Chapter V of Athreya and Ney [4]. We reproduce the necessary facts here for reference, attempting to maintain the notation of [4] whenever possible. In what follows, let $\mathbb{Z}_+$ denote the nonnegative integers.

For $i = 1, \ldots, \Delta - 1$ and $\mathbf{j} \in \mathbb{Z}_+^{\Delta-1}$ let $p^{(i)}(\mathbf{j}) \in [0, 1]$ such that

$$\sum_{\mathbf{j} \in (\mathbb{Z}_+)^{\Delta-1}} p^{(i)}(\mathbf{j}) = 1.$$

The branching processes we consider have the following form: particles have type i, $i = 1, 2, \ldots, \Delta - 1$. Each particle has an exponentially-distributed lifetime with parameter a_i, at which time it dies, producing $\mathbf{j}$ offspring ($\mathbf{j}_k$ offspring of type k) according to $p^{(i)}(\mathbf{j})$. For $\mathbf{x} \in [0, 1]^{\Delta-1}$, $i = 1, 2, \ldots, \Delta - 1$, we define the probability generating function

$$f^{(i)}(\mathbf{x}) = \sum_{\mathbf{j} \in (\mathbb{Z}_+)^{\Delta-1}} p^{(i)}(\mathbf{j})\mathbf{x}^{\mathbf{j}},$$

where we take $\mathbf{x}^{\mathbf{j}} = \prod_{k=1}^{\Delta-1} x_k^{j_k}$. We note the following fact regarding these probability generating functions: If

$$\left.\frac{\partial f^{(i)}(\mathbf{x})}{\partial x_j}\right|_{\mathbf{x}=\mathbf{1}} < \infty, \text{ for all } i, j, \tag{10}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$, then with probability 1 the process does not produce infinitely many particles over any finite time interval.

Now, let $m_{ij}(t)$ denote the expected number of descendants of type j of a particle of type i that after a period of time of length t. Then, letting $M(t) = (m_{ij}(t))$, provided (10) holds, $m_{ij}(t) < \infty$, and there exists $B \in \mathbb{R}^{(\Delta-1)\times(\Delta-1)}$ such that

$$M(t) = \exp(Bt) = \sum_{k=0}^{\infty} \frac{t^k B^k}{k!}. \tag{11}$$

Moreover, the matrix $B = (b_{ij})$ has entries

$$b_{ij} = a_i \left(\left. \frac{\partial f^{(i)}(\mathbf{x})}{\partial x_j} \right|_{\mathbf{x}=\mathbf{1}} - \delta_{ij} \right), \tag{12}$$

where δ_{ij} is the Kronecker delta function.

Our primary interest in these branching processes is the extinction probabilities. Specifically, let $\pi^{(i)}$ denote the probability that the branching process, initiated with a single particle of type i, eventually dies out. Then the extinction probability vector $\boldsymbol{\pi}$ is the unique solution of the system of equations

$$f^{(i)}(\mathbf{x}) = x_i, \; i = 1, \dots, \Delta - 1, \tag{13}$$

for $\mathbf{x} \in [0,1]^{\Delta-1}$. We do not explicitly calculate extinction probabilities here. Our main interest is in determining whether or not $\boldsymbol{\pi} = \mathbf{1}$. To this end, suppose there is a time t_0 such that $m_{ij}(t_0) > 0$ for all i, j; we call such a process **positive regular**. Let $\eta_1, \dots, \eta_{\Delta-1}$ be the (complex) eigenvalues of B. As the eigenvectors of B are eigenvectors of $M(t)$, this implies that the eigenvalues of $M(t)$ are

$$e^{\eta_1 t}, e^{\eta_2 t}, \cdots, e^{\eta_{\Delta-1} t}.$$

The Perron-Frobenius Theorem (applied to $M(t)$) yields that we can arrange the eigenvalues so that η_1 is real, and

$$\eta_1 > Re(\eta_2) \geq \cdots \geq Re(\eta_{\Delta-1}).$$

If the branching process is positive regular then

$$\eta_1 \leq 0 \Rightarrow \boldsymbol{\pi} = \mathbf{1} \qquad \text{and} \qquad \eta_1 > 0 \Rightarrow \boldsymbol{\pi} < \mathbf{1}.$$

4 Initial growth

In this section we prove Theorem 3 which describes the initial growth of the set of infected vertices.

We begin by coupling our process with a continuous time branching process. Our branching process has individuals of type i for $i = 1, 2, \dots, \Delta - 1$, and the lifetime of an individual of type i is distributed as $Exp(\rho + i\lambda)$. Upon an individual's death it produces no children with probability $\frac{\rho}{\rho+i\lambda}$. Otherwise (i.e. with probability $\frac{i\lambda}{\rho+i\lambda}$) the individual produces 2 offspring, one of type $i-1$ (we treat a child of type 0 as no child whatsoever) and another child of type T where $T \in \{1, \dots, \Delta - 1\}$ is chosen at random with the probability $Pr(T = i) = i \cdot S_i(0)/Q(0)$ where $\mathbf{S}(0) = \mathbf{z}n$. Note that this continuous time branching process has probability generating functions

$$f^{(i)}(\mathbf{x}) = \frac{\rho}{\rho + i\lambda} + \frac{i\lambda}{\rho + i\lambda} \sum_{j=0}^{\Delta-1} \frac{(j+1)S_{j+1}(0)}{Q(0)} x_{i-1} x_j \qquad \text{for } i = 1, \dots, \Delta - 1$$

where we set $x_0 = 1$.

The coupling is formally defined as follows. We generate sequences of i.i.d. exponential random variables with the parameters λ and ρ and i.i.d. uniform random choices of configuration positions with replacement (i.e. from the full initial collection of configuration positions). We run the SIR model with the appropriate exponential random variables and the given list of randomly chosen configuration positions, simply repeating any random choice that results in a configuration point that has already been seen. These same sequences are used to generate the branching process. We determine the time of death of each individual as the minimum of appropriate exponential random variables, and whether that death is an actual death or the release of a new infection is determined by which of the variables achieves the minimum. In the case of a new infection, the type of the second offspring is then determined by one of the i.i.d. uniform random choices of a configuration position.

Let $\mathcal{C}(t)$ be the event that the set of infected nodes in the SIR process is the same as the set of living individuals in the branching process (the degrees of the infected nodes should agree with the types of the living individuals) up to time t. Note that the two processes do not diverge until there is a choice of configuration position that corresponds to the same vertex as some previously chosen configuration position. Suppose that at time s_k in the SIR model the k^{th} infection is sent. The probability that this infection hits a vertex for which some configuration position has already been sampled is at most $\frac{\Delta k}{Q}$. So the probability that any of the first k infections hits a previously-infected vertex is at most $\Delta k^2/(Q-2k)$. In other words,

$$Pr(\mathcal{C}(s_k)) \geq 1 - \frac{\Delta k^2}{Q-2k}.$$

Note that the events treated in Theorem 3 involve far fewer than $\sqrt{n}$ infections. Thus, it suffices to prove Theorem 3 for the branching process.

We use the branching process notation and machinery as introduced in Section 3.3. As the probability generating functions are quadratic, the first partials exist and are finite at $\mathbf{x} = \mathbf{1}$, implying that with probability 1 the process remains finite for all finite t. Let $\boldsymbol{\pi}$ be the vector of extinction probabilities; that is, $\boldsymbol{\pi}_i$ is the probability of extinction when starting the process with a single individual of type i. Note that, as $\lambda > 0$ and $z_\Delta > 0$, over any finite time interval t, $m_{ij}(t) > 0$, as a particle of type i can produce a particle of type $\Delta - 1$, which then produces one of type $\Delta - 2$, etc. In other words the process is positive regular. Note further that we have

$$B = \lambda(A + \mathbf{u}^T\mathbf{D}) - \rho I. \tag{14}$$

In particular, by Corollary 10, we can conclude that

$$\lambda(\hat{D} - 1) - \rho \leq 0 \Rightarrow \boldsymbol{\pi} = \mathbf{1} \qquad \text{and} \qquad \lambda(\hat{D} - 1) - \rho > 0 \Rightarrow \boldsymbol{\pi} < \mathbf{1}.$$

Of course, we are interested in the behavior of the SIR model for large n. We therefore assume that n is large enough so that $\lambda(\hat{d} - 1) - \rho$ and $\lambda(\hat{D} - 1) - \rho$ are on the same side of 0. Of course, the case $\lambda(\hat{d} - 1) - \rho = 0$ is more delicate. We do not treat this case here.

It remains to show that the branching processes either quickly dies out or quickly achieves a large size. We treat the two cases separately with the following notational assumptions. Let $X_i(t)$ denote the number of particles of type i alive at time t, and let $\mathbf{X}(t)$ denote the vector $(X_i(t))$. We define

$$Z(t) = \sum_{i=1}^{\Delta-1} iX_i(t) = \mathbf{X}(t)\mathbf{u}^T.$$

Note that non-extinction at time t is equivalent to $Z(t) \geq 1$.

4.1 $\lambda(\hat{d}-1) - \rho > 0$

Our first observation is that we have

$$E[\mathbf{X}(t+1/m) \mid \mathbf{X}(t)] = \mathbf{X}(t)M(1/m).$$

By Corollary 10, (11) and (14) we therefore have

$$E[Z(t+1/m) \mid \mathbf{X}(t)] = \exp\left\{\frac{\lambda(\hat{D}-1)-\rho}{m}\right\} Z(t) \geq \left(1 + \frac{\lambda(\hat{D}-1)-\rho}{m}\right) Z(t). \quad (15)$$

For n sufficiently large, there exists a $c > 0$ such that $\lambda(\hat{D}-1) - \rho > c$. Conditioning on non-extinction this implies that

$$E[Z(t+1/m) \mid \mathbf{X}(t)] \geq Z(t)\left(1 + \frac{c}{m}\right). \quad (16)$$

We define a sequence of random times as follows. Set $t_0 = 0, t_1 = 1$ and for $j = 2, \ldots$ set

$$t_j = \begin{cases} t_{j-1} + 1/Z(t_{j-1}) & \text{if } Z(t_{j-1}) > 0 \\ t_{j-1} & \text{if } Z(t_{j-1}) = 0. \end{cases}$$

For ease of notation set $Z_j = Z(t_j)$. Note that (16) implies that, conditioning on $Z_j > 0$, we have

$$E[Z_{j+1} \mid Z_j] \geq Z_j + c.$$

Define the sequence Z'_j of random variables as follows. Set $Z'_0 = 1$. Set $Z'_j = Z'_{j-1}$ if for any $a \in \{1, \ldots, j\}$ we have $Z_{a-1} = 0$ or the number of births and deaths in $(t_{a-1}, t_a]$ is at least $\max\{\log^{1/3} n, a^{1/3}\}$. Otherwise, let

$$Z'_j = Z_j - jc.$$

Then the Z'_j form a submartingale, and $|Z'_{j-1} - Z'_j| \leq \max\{\log^{1/3} n, j^{1/3}\}$ for all j. (Note that our 'truncation' of Z_j slightly reduces the one-step expected change. However, it is easy to see that we still have the supermartingale condition: There is some 'room' in (16) and we can bound the contribution to the expectation from large one-step changes using binomial approximations as in the proof of Claim 12 below.) By the Azuma-Hoeffding inequality we have

$$j > \log n \quad \Rightarrow \quad Pr(Z'_j < -cj/2) \leq \exp\left\{\frac{-c^2 j^2}{8 \cdot j \cdot j^{2/3}}\right\} = \exp\left\{-\frac{c^2 j^{1/3}}{8}\right\}.$$

It follows that with high probability we have $Z'_j > -cj/2$ for $j = \log n, \ldots, 2n^\alpha/c$. In particular, we have one of the following events with high probability:

(i) $Z_j = 0$ for some $j \leq \log n$,

(ii) There exists an index j such that the number of births and deaths in $[t_j, t_{j+1})$ is greater then $\max\{\log^{1/3} n, (j+1)^{1/3}\}$, or

(iii) event (ii) does not hold and $Z_j > cj/2$ for $j = \log n, \ldots, 2n^\alpha/c$.

Note that in event (iii) we have $Z_{2n^\alpha/c} \geq n^\alpha$,

$$t_{2n^\alpha/c} < \log n + \sum_{j=\log n}^{2n^\alpha/c} \frac{2}{cj} = O(\log n),$$

and the total number of births and deaths up to time $t_{2n^\alpha/c}$ is at most $(2n^\alpha/c)^{4/3}$.

It remains to show that the probability of event (ii) tends to 0 as n goes to infinity. This follows from the following Claims and the union bound (note that we may assume $Z_j > cj/2$ for $j \geq \log n$).

Claim 12. *Let $Z(t) = m$. If $5e(\rho + \Delta\lambda) < m \leq \log n$, the probability that at least $\log^{1/3} n$ births or deaths occur in $(t, t+1/m]$ is $O((1/2)^{\log^{1/3} n})$. If $m > \log n$ then the probability at least $(2m/c)^{1/3}$ births or deaths occur in $(t, t+1/m]$ is $O((1/m^{1/3})^{m^{1/3}})$.*

Proof. To determine whether or not a particle of type i alive or born in the process dies or gives birth, we sample from a sequence of independent exponential random variables with parameter $\rho + i\lambda$. As a result, the probability that a particle of any type dies or gives birth is at most $\frac{\rho+\Delta\lambda}{m}$, and a birth or death produces at most 2 new particles. Therefore, the probability that there are at least ℓ births or deaths is bounded above by the probability that $W \geq \ell$, where $W \in Bin(m + 2\ell, \frac{\rho+\Delta\lambda}{m})$. First suppose $5e(\rho + \Delta\lambda) < m \leq \log n$.

$$\begin{aligned} Pr[W \geq \log^{1/3} n] &\leq \binom{m + 2\log^{1/3} n}{\log^{1/3} n} \left(\frac{\rho + \Delta\lambda}{m}\right)^{\log^{1/3} n} \\ &\leq \left(\frac{e(\rho + \Delta\lambda)(m + 2\log^{1/3} n)}{m \log^{1/3} n}\right)^{\log^{1/3} n} \\ &\leq \left(\frac{1}{2}\right)^{\log^{1/3} n}. \end{aligned}$$

Now suppose $m > \log n$.

$$Pr[W \geq (2m/c)^{1/3}] \leq \binom{m + 2(2m/c)^{1/3}}{(2m/c)^{1/3}} \left(\frac{\rho + \Delta\lambda}{m}\right)^{(2m/c)^{1/3}} \leq \left(\frac{3e(\rho + \Delta\lambda)}{m^{1/3}}\right)^{m^{1/3}}.$$

∎

Claim 13. *Let $Z(t) = m$. If $m \leq 5e(\rho + \Delta\lambda)$, the probability that at least $\log^{1/3} n$ births or deaths occur in $(t, t+1/m]$ is $e^{-\omega(\log\log n)}$.*

Proof. Consider the forest of descendants of the individuals present in $Z(t)$. Note that this is a subtree of a binary tree. If more than $\log^{1/3} n$ births or deaths occur in this interval then there is some descendant at depth $\log_2(\log^{1/3} n) = \frac{1}{3}\log_2 \log n$ from the root. Consider the ancestors of this individual. At least half of them must have life spans of less than $6/(\log_2 \log n)$. So the probability of this event is at most

$$5e(\rho + \Delta\lambda) \cdot 2^{\frac{1}{3}\log_2 \log n} \cdot 2^{\frac{1}{3}\log_2 \log n} \cdot \left[\frac{6(\rho+\Delta\lambda)}{\log_2 \log n}\right]^{\frac{1}{6}\log_2 \log n}.$$

This first moment calculation is as follows: first we choose one of the individuals present in $Z(t)$ and choose a path in the full binary tree of the appropriate depth starting at the chosen individual, we then choose half of the vertices on this path to have short life-spans and finally we multiply by the probability that all of these life-spans are short. □

4.2 $\lambda(\hat{d}-1) - \rho < 0$

Our approach is fundamentally the same as the argument in Section 4.1. As above, set $t_0 = 0, t_1 = 1$ and for $j = 2, \dots$ set

$$t_j = \begin{cases} t_{j-1} + 1/Z(t_{j-1}) & \text{if } Z(t_{j-1}) > 0 \\ t_{j-1} & \text{if } Z(t_{j-1}) = 0 \end{cases}$$

and $Z_j = Z(t_j)$. We begin with the observation

$$E[Z_{j+1}|\mathbf{X}(t_j)] = \exp\left\{\frac{\lambda(\hat{D}) - \rho}{Z_j}\right\} Z_j \le Z_j - c \tag{17}$$

for some constant $c > 0$. Define the sequence Z'_j of random variables as follows. Set $Z'_0 = 1$. Set $Z'_j = Z'_{j-1}$ if for any $a \in \{1, \dots, j\}$ we have $Z_{a-1} = 0$ or $Z_a > \log n$ or the number of births and deaths in $(t_{a-1}, t_a]$ is at least $\log^{1/3} n$. Otherwise, let $Z'_j = Z_j + cj$. Then the Z'_j form a supermartingale, $|Z'_{j+1} - Z'_j| \le \log^{1/3} n$. By the Azuma-Hoeffding inequality we have

$$Pr\left(Z'_{\log n} \ge c \log n\right) \le \exp\left\{-\frac{c^2 \log^2 n}{3\log^{2/3} n \cdot \log n}\right\} = \exp\left\{-\frac{\log^{1/3} n}{3}\right\}. \tag{18}$$

Noting that Claim 13 holds here, Theorem 3 follows.

5 Epidemic

We begin by deriving our system of differential equations (which is (19) below). We then discuss the solution of this system in Section 5.1. In Section 5.2 we establish some technical preliminaries for the proofs of Theorems 4 and 5. These proofs follow in Sections 5.3 and 5.4, respectively.

Recall that we set

$$t_0 = 0 \qquad t_1 = 1 \qquad \text{and} \qquad t_{i+1} = t_i + 1/Q_Y(t_i) \ \text{ for } i = 1, \dots.$$

and assume

$$
\begin{aligned}
S_i(t_k) &\approx ns_i(k/n) \text{ for } i = 1, \dots, \Delta \\
Y_i(t_k) &\approx ny_i(k/n) \text{ for } i = 1, \dots, \Delta - 1 \\
Q_R(t_k) &\approx nq_R(k/n)
\end{aligned}
$$

As is usual in application of the differential equations method for establishing dynamic concentration of random graph process, the derivatives of our continuous variables are determined by the one-step expected changes in the associated random variables. Note that the expected number of infections sent from vertices infected at time t_k in the interval $[t_k, t_k + 1/Q_Y(t_k))$ is approximately λ. Furthermore, the probability that any one vertex that is infectious at time t_k recovers is $\rho/Q_Y(t_k)$. Defining $x = x(k) = k/n$, we should have

$$
\begin{aligned}
\frac{dy_i}{dx} &\approx E\left[Y_i(t_{k+1}) - Y_i(t_k) | \mathcal{F}_k\right] \\
&\approx \lambda \left[-\frac{iY_i(t_k)}{Q_Y(t_k)} + \frac{(i+1)Y_{i+1}(t_k)}{Q_Y(t_k)} + \frac{(i+1)S_{i+1}(t_k)}{Q(t_k)} - \frac{iY_i(t_k)}{Q(t_k)} + \frac{(i+1)Y_{i+1}(t_k)}{Q(t_k)} \right] \\
&\qquad - \rho \frac{Y_i(t_k)}{Q_Y(t_k)} \\
&\approx (i+1)\lambda \frac{s_{i+1}(x)}{q(x)} - \rho \frac{y_i(x)}{q_Y(x)} + \lambda \left(1 + \frac{q_Y(x)}{q(x)}\right) \left((i+1)\frac{y_{i+1}(x)}{q_Y(x)} - i\frac{y_i(x)}{q_Y(x)}\right).
\end{aligned}
$$

where

$$
q_Y(x) = \sum_{i=1}^{\Delta-1} iy_i(x) \qquad \text{and} \qquad q(x) = q_Y(x) + q_R(x) + \sum_{i=1}^{\Delta} is_i(x).
$$

We emphasize that in this calculation we pass from the expected one-step change, which is a function of the n, to the limiting functions, y_i, s_i, q etc, that are *not* functions of n. Making analogous calculations for S_i and Q_R we arrive at the system of differential equations

$$
\begin{aligned}
\frac{ds_i}{dx} &= -\lambda i \frac{s_i}{q} \\
\frac{dy_i}{dx} &= (i+1)\lambda \frac{s_{i+1}}{q} - \rho \frac{y_i}{q_Y} + \lambda \left(1 + \frac{q_Y}{q}\right) \left((i+1)\frac{y_{i+1}}{q_Y} - i\frac{y_i}{q_Y}\right) \\
\frac{dq_R}{dx} &= \rho - \lambda \frac{q_R}{q}.
\end{aligned}
\tag{19}
$$

with initial conditions

$$
y_i(0) = 0 \qquad s_i(0) = z_i \qquad q_R(0) = 0.
$$

We view this as a differential equation on the bounded region $\mathcal{D}$ given by the set of vectors $\mathbf{w} = (\mathbf{s}, \mathbf{y}, q_R) \in [0, \|\mathbf{z}\|]^{2\Delta}$. For $\mathbf{w} \in \mathcal{D}$ set

$$
q_Y(\mathbf{w}) = q_Y(\mathbf{y}) = \sum_{i=1}^{\Delta-1} iy_i \qquad\qquad q(\mathbf{w}) = q_Y(\mathbf{w}) + \sum_{i=1}^{\Delta} is_i + q_R.
$$

Note that for every point $\mathbf{w}$ in the interior of $\mathcal{D}$ we have $q_Y > 0$ and $q > 0$, and the system (19) satisfies a local Lipschitz condition on the interior of $\mathcal{D}$. Consequently, for every point $\mathbf{w}$ in the interior of $\mathcal{D}$ a solution exists passing through $\mathbf{w}$ and extending to points arbitrarily close to the boundary (see Theorem 11 in Chapter 2 of [17]). However, we are interested in the solution through the point in $\mathcal{D}$ corresponding to the initial condition $(\mathbf{s}, \mathbf{y}, q_R) = (\mathbf{z}, 0, 0)$. This is a point on the boundary of $\mathcal{D}$ and the system is *not* locally Lipschitz here. So the existence of a unique solution to the system (19) does not follow immediately from the standard theory. We establish the existence of this solution in the next section, in which we prove the following theorem. Recall that x^* is the least positive root of the following expression:

$$(\|\mathbf{z}\| - 2\lambda x) - \sum_{i=1}^{\Delta} \frac{i z_i}{\|\mathbf{z}\|^{i/2}} (\|\mathbf{z}\| - 2\lambda x)^{i/2} - \left(\frac{\rho\sqrt{\|\mathbf{z}\|}}{\lambda} (\|\mathbf{z}\| - 2\lambda x)^{1/2} - \frac{\rho}{\lambda} (\|\mathbf{z}\| - 2\lambda x) \right).$$

Define

$$s_i(x) = \frac{z_i}{\|\mathbf{z}\|^{i/2}} (\|\mathbf{z}\| - 2\lambda x)^{i/2} \qquad q_R(x) = \frac{\rho\sqrt{\|\mathbf{z}\|}}{\lambda} (\|\mathbf{z}\| - 2\lambda x)^{1/2} - \frac{\rho}{\lambda} (\|\mathbf{z}\| - 2\lambda x).$$

Theorem 14. *There exists a function $\boldsymbol{y}(x) : [0, x^*] \to [0, \|\mathbf{z}|]^{\Delta - 1}$ such that $\boldsymbol{s}(x), \boldsymbol{y}(x), q_R(x)$ is the unique solution to the system of differential equations (19). Furthermore, for $0 < x < x^*$ the vector $\boldsymbol{y}(x)$ is on the interior of $\mathcal{D}$, $\boldsymbol{y}(x^*) = 0$ and $q'_Y(x^*) < 0$.*

We prove Theorem 14 in Section 5.1. With the appropriate trajectory in hand, we return to the SIR process and prove Theorems 4 and 5 in Sections 5.2–5.4.

As noted above, there are two significant complications for our application of the differential equations method: our trend hypothesis and Lipschitz constant for dy_i/dx deteriorate dramatically for Q_Y small. We deal with the Lipschitz constant by removing y_i from the system, working with a smaller system ((20) below) before moving on to the full system (19). We deal with the trend hypothesis by treating the start of the spread of the disease separately and employing a kind of 'self-correcting' or 'mean-reversion' property that certain random variables in this system exhibit. Consider, for example, the random variable S_i. (We do not need the 'mean-reversion' property to handle this random variable, but it is a good example of the phenomenon.) As the expected change in S_i is roughly $-\lambda i \frac{S_i}{Q}$ and the derivative of s_i is $-\lambda i \frac{s_i}{q}$, if $S_i(u_r)$ is significantly greater than $s_i(r/n)n$ we expect S_i to decrease faster than its continuous approximation $s_i(r/n)n$, and if $S_i(u_r)$ is significantly smaller than $s_i(r/n)n$ we expect its decrease to be slower than its continuous approximation. This balancing effect should push $S_i(u_r)$ towards its expected trajectory $ns_i(r/n)$ and ensure that this trajectory is indeed a good approximation. The use of this affect is the key to our proof of Theorem 5. We also use it in the proof of Theorem 4. (Even though other methods would suffice.) See [36] for another application of this 'self-correcting' phenomenon.

5.1 Solution of the ODE

We begin by noting that while the function that determined dy_i/dx is not well-behaved at $x = 0$ the other derivatives are well-behaved here (as they have no q_Y in the denominator).

In particular, note that summing over the relevant variables in (19) we have $dq/dx = -2\lambda$. We consider the following restricted version of (19).

$$\frac{dq}{dx} = -2\lambda \qquad \frac{ds_i}{dx} = -\lambda\frac{is_i}{q} \qquad \frac{dq_R}{dx} = \rho - \lambda\frac{q_R}{q} \tag{20}$$

with initial conditions $q(0) = \|\mathbf{z}\|, \mathbf{s}(0) = \mathbf{z}(0), q_R(0) = 0$. Note that (20) has solution

$$\begin{aligned} q(x) &= \|\mathbf{z}\| - 2\lambda x, \\ s_i(x) &= \frac{z_i}{\|\mathbf{z}\|^{i/2}}(\|\mathbf{z}\| - 2\lambda x)^{i/2}, \\ q_R(x) &= \frac{\rho\sqrt{\|\mathbf{z}\|}}{\lambda}(\|\mathbf{z}\| - 2\lambda x)^{1/2} - \frac{\rho}{\lambda}(\|\mathbf{z}\| - 2\lambda x). \end{aligned}$$

Note further that it follows that we also have

$$\begin{aligned} q_Y(x) &= q(x) - \sum_{i=1}^{\Delta} is_i(x) - q_R(x) \\ &= (\|\mathbf{z}\| - 2\lambda x) - \sum_{i=1}^{\Delta} \frac{iz_i}{\|\mathbf{z}\|^{i/2}}(\|\mathbf{z}\| - 2\lambda x)^{i/2} \\ &\quad - \left(\frac{\rho\sqrt{\|\mathbf{z}\|}}{\lambda}(\|\mathbf{z}\| - 2\lambda x)^{1/2} - \frac{\rho}{\lambda}(\|\mathbf{z}\| - 2\lambda x)\right). \end{aligned}$$

We are now in the position to note some useful properties of the function $q_Y(x)$. (Recall that x^* is the least positive root of $q_Y(x)$.)

Lemma 15. *We have* $x^* < \frac{\|z\|}{2\lambda}$, $q_Y'(0) > 0$ *and* $q_Y'(x^*) < 0$.

Proof. Consideration of $q_Y(x)/(\|\mathbf{z}\| - 2\lambda x)$ and the fact that $\rho > 0$ yields $x^* < \frac{\|\mathbf{z}\|}{2\lambda}$.

Observe that

$$q_Y'(x) = \lambda(\hat{d}_s(x) - 1) - \rho - \lambda\frac{q_Y(x)}{q(x)}. \tag{21}$$

Noting that $\hat{d}_s(0) = \hat{d}$ and $q_Y(0) = 0$, we have $q_Y'(0) > 0$.

Since $q_Y(x)$ is positive on $(0, x^*)$, we have $q_Y'(x^*) \leq 0$. Suppose for a contradiction that $q_Y'(x^*) = 0$. Then

$$q_Y''(x^*) = \lambda\hat{d}_s'(x^*) - \lambda\frac{q(x^*)q_Y'(x^*) - q_Y(x^*)q'(x^*)}{q^2(x^*)} = \lambda\hat{d}_s'(x^*) < 0,$$

as

$$\hat{d}_s'(x^*) = -\lambda\left(\sum_{i=3}^{\Delta}\frac{(i-2)(i-1)iz_i}{\|\mathbf{z}\|^{i/2}}(\|\mathbf{z}\| - 2\lambda x^*)^{(i-4)/2}\right)$$

and $x^* < \frac{\|\mathbf{z}\|}{2\lambda}$. This implies $q_Y(x^* - \delta) < 0$ for some $\delta > 0$, a contradiction. ∎

Now, let us return to the system (19): it remains to determine the existence and uniqueness of a solution for $y_i(x)$. For $\eta > 0$ let

$$\mathcal{D}_\eta = \{\mathbf{w} \in \mathcal{D} : q_Y(\mathbf{w}) > \eta\}.$$

Note that a solution for $y_i(x)$ exists on $\mathcal{D}_\eta$ for any $\eta > 0$. The main issue is establishing the existence of a solution at $x = 0$. The rest of this section is dedicated to doing so.

We have

$$\begin{aligned} \frac{d\mathbf{y}}{dx} &= \frac{\mathbf{y}}{q_Y}\left(\lambda\left(A + \mathbf{u}^T\mathbf{d}_s\right) - \rho I\right) + \lambda\frac{\mathbf{y}}{q}A \\ &= \frac{\mathbf{y}}{q_Y}\left(\lambda\left(A + \mathbf{u}^T\mathbf{d}\right) - \rho I\right) + \lambda\frac{\mathbf{y}}{q}A + \lambda\frac{\mathbf{y}}{q_Y}\mathbf{u}^T\left(\mathbf{d}_s - \mathbf{d}\right). \end{aligned} \tag{22}$$

Our motivation for (22) is the following: to show that (19) has a solution we change basis from the standard basis to the basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{\Delta-1}$ constructed in Section 3.2. Specifically, let $\mathbf{y} = \sum_{i=1}^{\Delta-1} c_i\mathbf{v}_i$. We observe that, as $\mathbf{v}_1\mathbf{u}^T = 1$ and $\mathbf{v}_i\mathbf{u}^T = 0$ for $i > 1$, we have $c_1 = q_Y$. Thus, applying P to both sides of (22) will give us a system of differential equations in the c_i, with initial conditions $c_i(0) = 0$. Noting that

$$\frac{\mathbf{y}}{q_Y}(\lambda(A + \mathbf{u}^T\mathbf{d}) - \rho I) = (\lambda(\hat{d} - 1) - \rho)\mathbf{v}_1 + \sum_{i=2}^{\Delta-1}\frac{-i\lambda - \rho}{q_Y}c_i\mathbf{v}_i,$$

$$\frac{\mathbf{y}}{q}\lambda A = \frac{q_Y}{q}\mathbf{v}_1\lambda A + \sum_{i=2}^{\Delta-1}\frac{-\lambda i}{q}c_i\mathbf{v}_i \qquad \text{and} \qquad \lambda\frac{\mathbf{y}}{q_Y}\mathbf{u}^T\left(\mathbf{d}_s - \mathbf{d}\right) = \lambda(\mathbf{d}_s - \mathbf{d}),$$

we have the following system for the c_i, $i > 1$:

$$\frac{dc_i}{dx} = c_i\left(\frac{-i\lambda - \rho}{q_Y} - \frac{i\lambda}{q}\right) + \left(\frac{\lambda q_Y}{q}\mathbf{v}_1 A + \lambda(\mathbf{d}_s - \mathbf{d})\right)P_i, \ i = 2, \ldots, \Delta - 1. \tag{23}$$

Lemma 16. *For each i, $i = 2, 3, \ldots, \Delta - 1$, there is a $\delta = \delta_i > 0$ and a $M = M_i > 0$ such that there exists a unique solution $c_i(x)$ to the differential equation (23) on the region $(0, \delta] \times \mathbb{R}$ such that $\lim_{x\to 0^+} c_i(x) = 0$. Moreover, that solution lies entirely in the region*

$$\mathcal{D} = \mathcal{D}_i := \{(x, c) \in \mathbb{R}^2 \mid x \in [0, \delta], -Mx^2 \le c \le Mx^2\}.$$

Note that Lemma 16 suffices to complete the proof of Theorem 14. We simply take η less than $q_Y(\delta_i)$ for all i and pass to the region $\mathcal{D}_\eta$ where a unique solution to the odes with the desired properties is known to exist.

Proof of Lemma 16. Consider a fixed $i \ge 2$. Define

$$f(x) = f_i(x) = \frac{i\lambda + \rho}{q_Y} + \frac{i\lambda}{q} \tag{24}$$

$$g(x) = g_i(x) = \left(\frac{\lambda q_Y}{q}\mathbf{v}_1 A + \lambda(\mathbf{d}_s - \mathbf{d})\right)P_i. \tag{25}$$

Note that

$$c(x) = c_i(x) = e^{\int_x^{\frac{x^*}{2}} f(\eta)d\eta} \left[\int_0^x g(\psi) e^{-\int_\psi^{\frac{x^*}{2}} f(\eta)d\eta} d\psi \right]$$

satisfies (23). Appealing to Lemma 15 (and the fact that ds_i/dx is bounded) we see that there exist constants K, δ such that

$$0 \leq x < \delta \quad \Rightarrow \quad |h(x)| < Kx.$$

In particular, it follows that if $0 < x < \delta$ then

$$|c(x)| \leq \left| \int_0^x g(\psi) e^{-\int_\psi^x f(\eta)d\eta} \, d\psi \right| \leq \int_0^x K\psi \, d\psi = \frac{Kx^2}{2}.$$

Finally, we establish the uniqueness of the solution $c_i(x)$. Suppose there exists a solution $c^*(x)$ to (23) on $(0, \delta_i]$ such that $c^*(x_0) \neq c_i(x_0)$ for some $x_0 \in (0, \delta]$, and $\lim_{x \to 0^+} c^*(x) = 0$. Define $p(x) := c^*(x) - c_i(x)$ and assume without loss of generality $p(x_0) > 0$. Then $\lim_{x \to 0^+} p(x) = 0$, but

$$\frac{dp}{dx} = (c^*(x) - c_i(x)) \left(-\frac{i\lambda + \rho}{q_Y(x)} - \frac{i\lambda}{q(x)} \right) < 0,$$

for $0 < x \leq x_0$, a contradiction.

□

5.2 Boundedness and Trend Hypotheses

In Sections 5.3 and 5.4 we apply the differential equations method for random graph processes to prove Theorems 4 and 5, respectively. Two key ingredients in any application of this method are the 'trend hypothesis' which establishes one step expected changes and the 'boundedness hypothesis' which bounds the one step change in any of the variables.

Lemma 17 (Boundedness Hypothesis)**.** *If $Q_Y(t) \geq \log^3(n)/2$, then the probability that more than $\log(n)$ vertices either recover or send or receive infections in the time interval $(t, t + \frac{1}{Q_Y(t)}]$ is $o(n^{-2})$.*

We refer to at least $\log n$ vertices recovering or sending or receiving infections over the time interval $[u_r, u_{r+1})$ as a **large jump** at time u_r. We prove Theorem 4 and 5 working with the version of the process where we stop as soon as a large jump occurs. Note that, by Lemma 17, the probability of stopping the process for this reason is $o(1)$.

Lemma 18 (Trend Hypothesis)**.** *For $Q_Y(t) \geq \log^3(n)/2$ we have*

$$
\begin{aligned}
E\left[Y_i\left(t+\frac{1}{Q_Y(t)}\right)-Y_i(t)\Big|\mathcal{F}_t\right] &= \frac{\lambda}{Q_Y(t)}\left(1+\frac{Q_Y(t)}{Q(t)}\right)((i+1)Y_{i+1}(t)-iY_i(t)) \quad (26)\\
&\quad -\frac{\rho}{Q_Y(t)}Y_i(t)+\frac{\lambda(i+1)S_{i+1}(t)}{Q(t)}+O\left(\frac{\log^2(n)}{Q_Y(t)}\right),\\
E\left[S_i\left(t+\frac{1}{Q_Y(t)}\right)-S_i(t)\Big|\mathcal{F}_t\right] &= -\lambda\frac{iS_i(t)}{Q(t)}+O\left(\frac{\log^2(n)}{Q_Y(t)}\right), \quad (27)\\
E\left[Q_R\left(t+\frac{1}{Q_Y(t)}\right)-Q_R(t)\Big|\mathcal{F}_t\right] &= \rho-\lambda\frac{Q_R(t)}{Q(t)}+O\left(\frac{\log^2(n)}{Q_Y(t)}\right), \quad (28)\\
E\left[Q\left(t+\frac{1}{Q_Y(t)}\right)-Q(t)\Big|\mathcal{F}_t\right] &= -2\lambda+O\left(\frac{\log(n)}{Q_Y(t)}\right). \quad (29)
\end{aligned}
$$

Furthermore, these estimates hold if we stop the process as soon as $\log n$ vertices recover or send or receive infections during the time interval $[t, t+1/Q_Y(t))$.

Proof of Lemma 17. Let $t_0 = t, t_1 = t+\frac{1}{Q_Y(t)}$. Let C be the number of configuration points sending infections and D be the number of vertices that recover in $(t_0, t_1]$. We bound the number of configuration points which send infections by sampling from an infinite list of i.i.d. exponential random variables with parameter λ. We associate the first $Q_Y(t_0)$ with configuration points in $Y(t_0)$ and for each vertex outside $Y(t_0)$ which becomes infected, we sample its infecting times from the same list, starting with the $(Q_Y(t_0)+1)$st entry. As each infection yields at most Δ new infectious configuration points, we have

$$
\begin{aligned}
Pr[C \geq \log(n)/3] &\leq \binom{Q_Y(t)+\Delta\log(n)}{\log(n)/3}\left(\frac{\lambda}{Q_Y(t)}\right)^{\log(n)/3}\\
&\leq \binom{2Q_Y(t)}{\log(n)/3}\left(\frac{\lambda}{Q_Y(t)}\right)^{\log(n)/3}\\
&\leq \left(\frac{6\lambda e}{\log(n)}\right)^{\log(n)/3} = o\left(\frac{1}{n^2}\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
Pr[D \geq \log(n)/3 \mid C \leq \log(n)/3] &\leq \binom{|Y(t)|+\log(n)}{\log(n)/3}\left(\frac{\rho}{Q_Y(t)}\right)^{\log(n)/3}\\
&\leq \left(\frac{6\rho e}{\log(n)}\right)^{\log(n)/3} = o\left(\frac{1}{n^2}\right).
\end{aligned}
$$

□

Proof of Lemma 18. Let $t_0 = t$, $t_1 = t+\frac{1}{Q_Y(t)}$. As in the proof of Lemma 17, let C denote the number of infections sent in $(t_0, t_1]$ and let D be the number of recoveries. By Lemma 17 the probability that at least $\log n$ vertices send or receive infections or recover is $o(1/n^2)$. As $C, D = O(n)$ and $1/n = o(\log n/Q_Y(t))$, we may henceforth consider the process in which we stop if and when $\log n$ vertices send or receive infections or recover. Given this assumption, we stochastically bound C as $W_1 \leq C \leq W_2$, where

$$W_1 \sim Bin\left(Q_Y - \lfloor \Delta \log(n) \rfloor, \frac{\lambda}{Q_Y}\right) \quad \text{and}$$
$$W_2 \sim Bin\left(Q_Y + \lfloor \Delta \log(n) \rfloor, \frac{\lambda}{Q_Y}\right).$$

Similar bounds apply to D. Therefore we have

$$E[C] = \lambda + O\left(\frac{\log(n)}{Q_Y}\right) \quad \text{and}$$
$$E[D] = \frac{\rho|Y|}{Q_Y} + O\left(\frac{\log(n)}{Q_Y}\right),$$

yielding (29).

Furthermore, our stopping rule yields

$$\begin{aligned} S_i(s) &= S_i(t) + O\left(\log(n)\right), \\ Y_i(s) &= Y_i(t) + O\left(\log(n)\right), \text{ and} \\ Q_R(s) &= Q_R(t) + O\left(\log(n)\right), \end{aligned}$$

for all $s \in [t_0, t_1)$. Consequently, $Q_Y(s) = Q_Y(t) + O\left(\log n\right)$ and $Q(s) = Q(t) + O\left(\log n\right)$ for all $s \in [t_0, t_1)$.

Now we establish (26). The remaining statements follow similarly. Suppose $0 \leq C, D \leq \log n$. For $0 \leq j \leq C$ the j^{th} infection can be viewed as being selected uniformly at random from the configuration points in $Y(s)$ where s is the time that this infection is sent. So the probability that it lies in an infected vertex of degree i is

$$\frac{iY_i(s)}{Q_Y(s)} = \frac{iY_i(t)}{Q_Y(t)} + O\left(\frac{\log(n)}{Q_Y(t)}\right).$$

In this event we decrease Y_i by one and increase Y_{i-1} by one. The probability this infection hits an infected vertex of degree $k \neq i$ is

$$\frac{kY_k(s)}{Q(s) - 1} = \frac{kY_k(t)}{Q(t)} + O\left(\frac{\log(n)}{Q(t)}\right)$$

(note that we must account for the infecting configuration point when $k = i$, but we arrive at the same result). We get a similar formula for the probability of the event that the infection hits a susceptible vertex of degree i, and in this event we increase Y_{i-1} by one. Similarly, when the j^{th} vertex recovers it is chosen uniformly at random from available infectious vertices, and therefore selects a vertex of degree i with probability $\frac{Y_i(t)}{|Y(t)|} + O\left(\frac{\log(n)}{Q_Y(t)}\right)$. Considering now the change to Y_i we have

$$\begin{aligned} E\left[Y_i(t_1) - Y_i(t_0) \mid C, D\right] &= C\left(\frac{(i+1)S_{i+1}(t)}{Q(t)} + \frac{Y_{i+1}(t)}{Q_Y(t)} + \frac{Y_{i+1}(t)}{Q(t)} - \frac{Y_i(t)}{Q_Y(t)} - \frac{Y_i(t)}{Q(t)}\right) \\ &\quad -D\frac{Y_i(t)}{|Y(t)|} + O\left(\frac{\log^2(n)}{Q_Y(t)}\right). \end{aligned}$$

$\square$

5.3 Proof of Theorem 4

As discussed in Section 5.2, we work with the version of the process where we stop as soon as there is a large jump. As the error term in the trend hypothesis depends on $Q_Y(t)$, we begin with the following Lemma that ensures that $Q_Y(t)$ grows to a sufficient size. The proof of Lemma 19 appears at the end of this section.

Lemma 19 (Initial Steps)**.** *Let $K = n^{3/4}$. Then*

$$\begin{aligned} Q(u_r) &= nq\left(\frac{r}{n}\right) + O(n^{1/2}), \\ Q_Y(u_r) &= nq_Y\left(\frac{r}{n}\right) + O(n^{1/2}), \\ Q_R(u_r) &= nq_R\left(\frac{r}{n}\right) + O(n^{1/2}), \text{ and} \\ S_i(u_r) &= ns_i\left(\frac{r}{n}\right) + O(n^{1/2}), \quad \text{for } i = 1, 2, \dots, \Delta, \end{aligned}$$

for $r = 0, 1, 2, \dots, K$ with probability at least $1 - o(1)$.

Set $K = n^{3/4}$ and let M be the minimum of $\lfloor \sigma n \rfloor$ and the least index $j > K$ such that either a large jump occurs in (u_{j-1}, u_j) or $Q_Y(u_j) < \sqrt{n}$. We argue that, with high probability, the following inequalities hold for $j = K, \dots, M$:

$$\left| Q(u_j) - q\left(\frac{j}{n}\right)n \right| \le \log^2(n)\sqrt{n} \tag{30}$$

$$\left| S_i(u_j) - s_i\left(\frac{j}{n}\right)n \right| \le \log^3(n)\sqrt{n}, \quad i = 1, \dots, \Delta \tag{31}$$

$$\left| Q_R(u_j) - q_R\left(\frac{j}{n}\right)n \right| \le \log^3(n)\sqrt{n}. \tag{32}$$

Note that provided these inequalities hold, we have $Q_Y(u_j) = q_Y(\frac{j}{n})n + O(\log^3(n)\sqrt{n})$ for $j = K, \dots, M$. As $q_Y(x) \ge \log^4(n)n^{-1/2}$ for $x \in (\frac{K}{n}, \sigma)$, if (30)–(32) hold and there is no large jump then $M = \sigma n$. This completes the proof.

We first prove (30). Let $X = Q(u_K) - q(K/n)n$, and note $X = O(\sqrt{n})$ by Lemma 19. By Lemma 18, as $Q_Y(u_j) \ge n^{1/2}/2$ for $j = K, \dots, M$, there is a nonnegative function $f(j) = O(\frac{(j-K)\log(n)}{\sqrt{n}})$, $f(K) = 0$, so that the sequences

$$A_j = Q(u_j) - q\left(\frac{j}{n}\right)n - X + f(j)$$

and

$$B_j = Q(u_j) - q\left(\frac{j}{n}\right)n - X - f(j)$$

form sub- and super-martingales, respectively, with differences bounded by $3\log(n)$. Note that (30) failing is contained in the event that $A_j \le -\log^2(n)\sqrt{n}/2$ or $B_j \ge \log^2(n)\sqrt{n}/2$ for some j, $K \le j \le M$. By Azuma-Hoeffding, the probability of this event is at most

$$2\exp\left\{-\Omega\left(\frac{\log^4(n)}{\log^2(n)}\right)\right\} = o(1).$$

We now prove (31) using Lemma 8. Let M' be the minimum of M and the smallest $j > K$ for which (30) fails to hold. Fix an i, $1 \le i \le \Delta$, and let $W = S_i(u_K) - s_i(K/n)n$. Consider the sequence

$$A_j = S_i(u_j) - s_i(j/n)n - W,$$

for $j = K, \ldots, M'$. As there are no large jumps, $|A_{j+1} - A_j| \le \log(n)$. By (30) and Lemma 18, as $Q_Y(u_j) \ge \sqrt{n}$ for $j = K, \ldots, M' - 1$, we have

$$\begin{aligned}
E[A_{j+1} - A_j | \mathcal{F}_{u_j}] &= -\frac{\lambda i S_i(u_j)}{Q(u_j)} - \left(s_i\left(\frac{j+1}{n}\right) - s_i\left(\frac{j}{n}\right)\right)n + O\left(\frac{\log^2(n)}{\sqrt{n}}\right) \\
&= -\frac{\lambda i S_i(u_j)}{Q(u_j)} - s_i'\left(\frac{j}{n}\right) + O\left(\frac{1}{n} + \frac{\log^2(n)}{\sqrt{n}}\right) \\
&= -\frac{\lambda i S_i(u_j)}{Q(u_j)} - \left(-\frac{\lambda i s_i(j/n)}{q(j/n)}\right) + O\left(\frac{\log^2(n)}{\sqrt{n}}\right) \\
&= -\frac{\lambda i S_i(u_j)}{q(j/n)n + O(\log^2(n)\sqrt{n})} - \left(-\frac{\lambda i s_i(j/n)}{q(j/n)}\right) \\
&\qquad + O\left(\frac{\log^2(n)}{\sqrt{n}}\right) \\
&= -\frac{\lambda i}{q(j/n)n}\left(S_i(u_j) - s_i(j/n)n\right) + O\left(\frac{\log^2(n)}{\sqrt{n}}\right) \\
&= -\frac{\lambda i}{q(j/n)n} A_j + O\left(\frac{\log^2(n)}{\sqrt{n}}\right).
\end{aligned}$$

In particular, there is a nonnegative function $h(j) = O(\log^2(n)\sqrt{n})$ so that the conditions of Lemma 8 apply. Consequently the probability that there exists a step j such that $A_j > h(j) + \log^3(n)\sqrt{n}/2$ is at most

$$\sigma n \exp\left\{-\Omega\left((\log^4(n)\right)\right\} = o(1).$$

As W and $h(j)$ are $o(\log^3(n)\sqrt{n})$, the event that the upper bound in (31) fails is contained in the event that such a step j exists. Repeating this argument for the sequence $-A_j$ gives the lower bound in (31).

Similar arguments yield (32). Theorem 4 follows from the union bound.

Proof of Lemma 19. We first observe that if f is a differentiable function defined on $[0, \delta]$ for some $\delta > 0$ whose derivative satisfies a Lipschitz condition, then for n sufficiently large and $x \in [0, K/n]$ we have

$$|f(x) - (f(0) + xf'(0))| = O\left(x^2\right) = O\left(\frac{K^2}{n^2}\right) = O\left(n^{-1/2}\right).$$

So, it suffices to argue that, for $r = 0, \ldots, K$, with high probability we have the following inequalities:

$$\begin{aligned}
|Q(u_r) - (Q(0) - 2\lambda r)| &\leq 3n^{4\alpha/3} + 2\lambda r^{2/3}, & (33)\\
|Q_Y(u_r) - (n^\alpha + (\lambda(\hat{d}-1) - \rho)r)| &\leq \frac{n^\alpha}{2} + (\lambda(\hat{d}-1) - \rho)r^{2/3}, & (34)\\
|Q_R(u_r) - \rho r| &\leq 2\Delta n^{4\alpha/3} + \rho r^{2/3}, \text{ and} & (35)\\
\left| S_i(u_r) - \left(z_i n - \frac{iz_i}{\|\mathbf{z}\|}\lambda r \right) \right| &\leq 2n^{4\alpha/3} + \left(\frac{iz_i}{\|\mathbf{z}\|} \right) r^{2/3}. & (36)
\end{aligned}$$

Define the time J to be the least $r \leq K$ such that one of these conditions fails or a large jump takes place between u_{r-1} and u_r. Recall that we stop the process as soon as a large jump occurs, and note that conditions (33)-(36) trivially hold for $r \ll n^\alpha / \log n$.

We bound the probability that there exists a step $r < K$ such that the lower bound in (34) fails. The arguments for the remaining conditions are analogous. Define $\varphi = \lambda(\hat{d}-1) - \rho$. Observe that if (34) holds at time u_r, then $Q_Y(u_r) \geq \frac{n^\alpha}{2}$, so Lemmas 17 and 18 apply. We have

$$E\left[Q_Y(u_{r+1}) - Q_Y(u_r) \middle| \mathcal{F}_{u_r}\right] = \varphi + O\left(\frac{r + n^{4\alpha/3}}{n} + \frac{\log^2(n)}{n^\alpha/2 + (r - r^{2/3})\varphi} \right).$$

There is a function

$$\begin{aligned}
H(r) = O\left(\sum_{j=1}^{r} \left(\frac{j + n^{4\alpha/3}}{n} + \frac{\log^2(n)}{n^\alpha/2 + \varphi(j - j^{2/3})} \right) \right) \\
= O\left(\frac{r^2 + rn^{4\alpha/3}}{n} + \log^2(n)\log(r) \right)
\end{aligned}$$

such that the random variables $Z_0 = 0$,

$$Z_r = Q_Y(u_r) - Q_Y(u_0) - r\varphi + H(r)$$

for $1 \leq r \leq K$ form a submartingale with differences bounded by $O(\log(n))$. Suppose the lower bound in (34) fails at time u_r; that is, suppose

$$Q_Y(u_r) \leq \frac{n^\alpha}{2} + \varphi(r - r^{2/3}).$$

We have

$$\begin{aligned}
Z_r &= Q_Y(u_r) - \varphi r - Q_Y(u_0) + H(r)\\
&\leq \frac{n^\alpha}{2} - \varphi r^{2/3} - Q_Y(u_0) + O\left(r^{1/2}\right)\\
&\leq -\left(\frac{\varphi}{2}\right) r^{2/3},
\end{aligned}$$

for n sufficiently large. Therefore, by Azuma-Hoeffding and the assumption $r = \Omega(n^\alpha / \log n)$, we note that this happens with probability at most

$$\exp\left\{ -\Omega\left(\frac{r^{4/3}}{\log^2(n) r} \right) \right\} \leq \exp\left\{ -\Omega\left(\frac{n^{\alpha/3}}{\log^{5/2}(n)} \right) \right\}.$$

□

5.4 Proof of Theorem 5

We apply the general theorem of Wormald (Theorem 5.2 of [38]). However, we run into the difficulty that this system does not satisfy a Lipschitz condition at the origin. To circumvent this we prove an initial Lemma which gets the process into the region on which we can apply the Theorem of Wormald. Recall that for $\eta > 0$ a constant,

$$\mathcal{D}_\eta = \{\mathbf{w} \in \mathcal{D} : q_Y(\mathbf{w}) > \eta\}.$$

In order to reach $\mathcal{D}_\eta$ we do not analyze the early evolution of $Y_1, Y_2, \dots, Y_{\Delta-1}$. Instead, we change variables to take advantage of the understanding of the function $c_i(x)$ for x small established in Section 5.1. Define

$$\mathbf{C}(t) = \mathbf{Y}(t)P$$

where P is the change of basis matrix introduced in Section 3.2. We note that $C_1(t) = Q_Y(t)$, which we already know is well-behaved. It remains to study the early evolution of $C_2(t), \dots, C_{\Delta-1}(t)$.

Lemma 20. *Set $L = \delta n$. With high probability we have*

$$C_i(u_r) = nc_i(r/n) + O\left(\sqrt{n} + rn^{-1/4}\right)$$

for $r = 0, \dots, L$ and $i = 2, \dots, \Delta - 1$.

With this Lemma in hand we, turn to applying Wormald's Theorem 5.1 with 2 modifications. Our starting point corresponds to $x = \delta$. This change does not otherwise alter the proof. The more significant change is that we must include an error at the initial point. How does this initial error influence the subsequent errors in the induction that proves Theorem 5.1 of Wormald? Recall that in said proof we set parameters λ, β and γ and the increment in our inductive step is $w = n\lambda/\beta$. Then B_j is the allowed error after step j, which is determined by the recurrence

$$B_{j+1} = \left(1 + \frac{Bw}{n}\right) B_j + \left(\frac{Bw^2}{n} + Bw\lambda\right)$$

where B is an absolute constant. In [38] we simply set $B_0 = 0$. If instead we allow $B_0 > 0$ then the solution of the recurrence is

$$B_j = B_0 \left(1 + \frac{Bw}{n}\right)^j + \left(\frac{Bw^2}{n} + Bw\lambda\right) \sum_{k=0}^{j-1} \left(1 + \frac{Bw}{n}\right)^k = O\left(B_0\right) + O\left(\lambda n\right).$$

Therefore, we apply Wormald's Theorem 5.1 with parameters

$$\beta = \log^2 n \qquad \gamma = \frac{1}{n^2} \qquad \lambda = n^{-1/4} \qquad B_0 = n^{3/4},$$

(taking δ sufficiently small) to complete the proof of Theorem 5.

Proof of Lemma 20. We note at the outset that we may assume that inequalities (33)-(36) hold for $0 \le r \le K$ and that the estimates given by Theorem 4 hold for $K < r \le L$.

We determine the expected one-step changes in $C_2, \dots C_{\Delta-1}$ by following the linear algebra that gave (23). To this end, let $i \in \{2, \dots, \Delta-1\}$ be fixed and define

$$G_i(t) = \left(\frac{\lambda Q_Y(t)}{Q(t)} \mathbf{v}_1 A + \lambda(\mathbf{D}_S(t) - \mathbf{d}) \right) P_i.$$

Applying the trend hypothesis we have

$$E[C_i(u_{r+1}) - C_i(u_r) | \mathcal{F}_{u_r}] = C_i(u_r) \left(-\frac{\lambda i + \rho}{Q_Y(u_r)} - \frac{\lambda i}{Q(u_r)} \right) + G_i(u_r) + O\left(\frac{\log^2 n}{Q_Y(u_r)} \right). \quad (37)$$

Note that we have a good set-up for 'self-correction' in the expression for the expected change in C_i. We also note that by our stopping rule we have

$$|C_i(u_{r+1}) - C_i(u_r)| \le 2\Delta \left(\max_{i,j} |P_{ij}| \right) \log n = O(\log n)$$

for $r = 0, \dots, L$. Note that by (33)-(36) and Theorem 4 we have

$$G_i(u_r) = O\left(\frac{r + n^{4\alpha/3}}{n} \right) \qquad \text{for } r = 0, 1, \dots, K = n^{3/4} \quad (38)$$

$$G_i(u_r) = g_i(r/n) + O(n^{-1/4}) \qquad \text{for } r = 0, 1, \dots, L = \delta n. \quad (39)$$

(Recall that the function $g_i(x)$ is defined in (25).)

First we establish an upper bound on $C_i(u_r)$ for r small.

Claim 21. *With high probability we have*

$$|C_i(u_r)| = O\left(\frac{r^2}{n} + \frac{r n^{4\alpha/3}}{n} + \sqrt{n} \right)$$

for $r = 0, \dots, K$.

Proof. By (37) there is a function $h(r) = O(G_i(r) Q_Y(u_r) + \log^2(n))$ such that

$$\begin{aligned} C_i(u_r) > h(r) \qquad &\Rightarrow \qquad E[C_i(u_{r+1}) - C_i(u_r) | \mathcal{F}_{u_r}] \le 0 \\ C_i(u_r) < -h(r) \qquad &\Rightarrow \qquad E[C_i(u_{r+1}) - C_i(u_r) | \mathcal{F}_{u_r}] \ge 0. \end{aligned}$$

Applying Lemma 8 to the sequence $C_i(u_r)$ (and using (38)) gives that with high probability we have

$$|C_i(u_r)| \le h(r) + \sqrt{n} = O\left(\frac{r^2}{n} + \frac{r n^{4\alpha/3}}{n} + \sqrt{n} \right)$$

for $r = 0, \dots, K$. ☐

Note that Claim 21 establishes Lemma 20 for $r = 0, \dots, n^{3/4}$ as we have $c_i(r/n) = O(r^2/n^2)$ and this implies that $n c_i(r/n) = O(\sqrt{n})$ for $r \le n^{3/4}$.

To deal with $r = K+1, \ldots, L$, we define the stopping time L' to be the smallest $K < r \le L$ such that there is a big jump at time u_{r-1} or $C_i(u_r) \ge (M+2)r^2/n$ (recall that M is the constant from Lemma 16). Define

$$Z_r = C_i(u_r) - nc_i(r/n).$$

Note that the event $C_i(u_j) \ge (M+2)j^2/n$ is contained in the event $|C_i(u_j) - nc_i(j/n)| > \sqrt{n} + jn^{-1/4}$ as $|c_i(x)| \le Mx^2$ and $j \ge n^{3/4}$ implies $jn^{-1/4} \le j^2/n$. As the second derivative of $c_i(x)$ is bounded we have

$$\begin{aligned} n\left(c_i\left(\frac{r+1}{n}\right) - c_i\left(\frac{r}{n}\right)\right) &= c_i'\left(\frac{r}{n}\right) + O\left(\frac{1}{n}\right) \\ &= c_i(r/n)\left(\frac{-i\lambda-\rho}{q_Y(r/n)} - \frac{i\lambda}{q(r/n)}\right) + g_i(r/n) + O\left(\frac{1}{n}\right) \end{aligned}$$

Using this fact and (39) we have

$$\begin{aligned} E[Z_{r+1} - Z_r|\mathcal{F}_{u_r}] &= C_i(u_r)\left(-\frac{\lambda i+\rho}{Q_Y(u_r)} - \frac{\lambda i}{Q(u_r)}\right) - C_i(u_r)\left(-\frac{\lambda i+\rho}{nq_Y(r/n)} - \frac{\lambda i}{nq(r/n)}\right) \\ &\quad + G_i(u_r) - g_i\left(\frac{r}{n}\right) \\ &\quad + C_i(u_r)\left(-\frac{\lambda i+\rho}{nq_Y(r/n)} - \frac{\lambda i}{nq(r/n)}\right) \\ &\quad -nc_i\left(\frac{r}{n}\right)\left(-\frac{\lambda i+\rho}{nq_Y(r/n)} - \frac{\lambda i}{nq(r/n)}\right) + O\left(\frac{\log^2(n)}{Q_Y(u_r)}\right) + O\left(\frac{1}{n}\right) \\ &= C_i(u_r)\left(-(\lambda i+\rho)\frac{nq_Y(r/n) - Q_Y(u_r)}{nq_Y(r/n)Q_Y(u_r)} - \lambda i\frac{nq(r/n) - Q(u_r)}{nq(r/n)Q(u_r)}\right) \\ &\quad + Z_r\left(-\frac{\lambda i+\rho}{nq_Y(r/n)} - \frac{\lambda i}{nq(r/n)}\right) + O\left(\frac{\log^2(n)}{Q_Y(u_r)} + n^{-1/4}\right) \\ &= -\Omega\left(\frac{Z_r}{r}\right) + O\left(\frac{r^2}{n}\cdot\frac{n^{3/4}}{r^2}\right) + O\left(\frac{\log^2(n)}{r} + n^{-1/4}\right) \\ &= -\Omega\left(\frac{Z_r}{r}\right) + O\left(n^{-1/4}\right). \end{aligned}$$

It follows that there exists a nonnegative, nondecreasing function $h(r) = O\left(rn^{-1/4}\right)$ such that

$$Z_r > h(r) \;\Rightarrow\; E[Z_{r+1} - Z_r|\mathcal{F}_{u_r}] \le 0 \quad \text{and} \quad Z_r < -h(r) \;\Rightarrow\; E[Z_{r+1} - Z_r|\mathcal{F}_{u_r}] \ge 0.$$

Appealing to Lemma 8 once more, noting that $r \ge n^{3/4}$ implies $rn^{-1/4} \ge r^{2/3}$, the claim follows. $\square$

6 The End

In this Section we prove Theorem 6 which describes the end of an epidemic. As the reader might anticipate, we view the end of the epidemic as a mirror image of the emergence of the epidemic. Given Theorem 4, we expect the change in Q_Y over additional steps to be $\approx q'_Y(x^*)$, so after $\Theta(\log^4(n)\sqrt{n})$ additional steps, we expect the infection to die out. To be precise, we will show that after $\Theta(\log^4(n)\sqrt{n})$ additional steps, with high probability Q_Y drops below $\log^4(n)$ and this occurs in $\Theta(\log(n))$ (actual) time. At that point we couple the remainder with a branching process to establish that the epidemic indeed comes to an end.

We begin by noting that since $q'_Y(x^*) < 0$ we have $|\sigma - x^*| = O\left(\log^4(n)n^{-1/2}\right)$. Thus, Theorem 4 implies

$$\begin{aligned} S_i(u_{\sigma n}) &= ns_i(x^*) + O\left(\log^4(n)\sqrt{n}\right) &(40)\\ Q(u_{\sigma n}) &= nq(x^*) + O\left(\log^4(n)\sqrt{n}\right), \text{ and} &(41)\\ Q_R(u_{\sigma n}) &= nq_R(x^*) + O\left(\log^4(n)\sqrt{n}\right). &(42)\end{aligned}$$

The following Lemma establishes the first part of our claim.

Lemma 22. *Let $K = 4\log^4(n)\sqrt{n}/|q'_Y(x^*)|$. With high probability there exists an index j, $\sigma n \le j \le \sigma n + K$, such that:*

$$Q_Y(u_j) \quad < \quad \log^4(n),$$

$u_j - u_{\sigma n} = \Theta(\log(n))$, and $O(\log^4(n)\sqrt{n})$ infections and $O(\log^5(n)\sqrt{n})$ recoveries occur in $(u_{\sigma n}, u_j)$.

Proof of Lemma 22. Let the stopping time J be the minimum of $\sigma n + K$ and the smallest index $\sigma n \le j$ such that $Q_Y(u_j) < \log^3(n)$, $Q_Y(u_j) > n^{2/3}$ or at least $n^{2/3}$ infections have been sent in $(u_{\sigma n}, u_j)$ or a large jump occurs at step u_{j-1}. For $\sigma n \le j < J$, we have $S_i(u_j) = s_i(x^*)n + O(n^{2/3})$ and $Q(u_j) = q(x^*)n + O(n^{2/3})$.

Applying Lemma 18 with (40)-(42) and (21) we have

$$\begin{aligned} E[Q_Y(u_{j+1}) - Q_Y(u_j) \mid \mathcal{F}_{u_j}] &= \lambda\left(\hat{D}_S(u_j) - 1 - \frac{Q_Y(u_j)}{Q(u_j)}\right) - \rho + O\left(\frac{\log^2(n)}{Q_Y(u_j)}\right)\\ &= \lambda\left(\hat{d}_s(x^*) - 1\right) - \rho + O\left(\frac{1}{n^{1/3}} + \frac{\log^2(n)}{\log^3(n)}\right)\\ &= q'_Y(x^*) + O\left(\frac{1}{\log(n)}\right)\end{aligned}$$

for $j < J$.

Set $\eta = 1/(2|q'_Y(x^*)|)$. Define $T_0 = \sigma n$, $m_0 = Q_Y(u_{\sigma n})$, and for $r > 0$ let

$$T_r = T_{r-1} + \eta m_{r-1} \quad \text{and} \quad m_r = Q_Y(u_{T_r}).$$

Let W be the least r such that $m_r < \log^4(n)$. We argue that with high probability, for $0 \le r < W$, the following hold:

(i) $m_{r+1} \le 3m_r/4$,

(ii) $Q_Y(u_j) \ge m_r/4$ for $j = T_r, \ldots, T_{r+1}$, and

(iii) at most $2\lambda\eta m_r$ infections are sent and $\log(n)m_r$ recoveries occur in $(u_{T_r}, u_{T_{r+1}})$.

Note that (i) and (ii) imply $W = \Theta(\log(n))$ and $T_W \leq \sigma n + K$ (as $m_0 \leq (3/2)\log^4(n)\sqrt{n}$), (ii) implies $u_{T_{r+1}} - u_{T_r} \leq 4\eta$ and hence $u_{T_W} - u_{\sigma n} = O(\log(n))$, and (i) and (iii) together imply the remainder of the Lemma.

Now consider a fixed r and suppose $\log^4(n) < m_r < n^{3/5}$, and at most $n^{3/5}$ infections have occurred in (u_{T_0}, u_{T_r}) (i.e. $T_r < J$). Define the following sequences of random variables: $A_0 = B_0 = D_0 = 0$, and for $1 \leq \ell \leq \eta m_r$, let

$$A_\ell = Q_Y(u_{T_r+\ell}) - (m_r + \ell q'_Y(x^*)) - \frac{\ell}{\sqrt{\log(n)}}$$
$$B_\ell = Q_Y(u_{T_r+\ell}) - (m_r + \ell q'_Y(x^*)) + \frac{\ell}{\sqrt{\log(n)}}$$
$$D_\ell = Q(u_{T_r+\ell}) - (Q(u_{T_r}) - 2\lambda\ell) + \frac{\ell}{\sqrt{\log(n)}},$$

where we stop the process at the minimum of J and the first step $j > T_r$ such that any of the following conditions hold:

$$A_j > \frac{m_r}{6}, \quad B_j < -\frac{m_r}{6} \quad \text{or} \quad D_j < -\lambda\eta m_r.$$

Note that A_j is a supermartingale with differences bounded by $\Delta\log(n)$, while B_j is a submartingale with differences bounded by $\Delta\log(n)$ and D_j is a submartingale with differences bounded by $3\log(n)$. The event $T_r < J \leq T_{r+1}$ with no large jumps is contained in the union of events $A_{\eta m_r} \geq m_r/6$, $B_{\eta m_r} \leq -m_r/6$ and $D_{\eta m_r} \leq -\lambda\eta m_r$. Applying Azuma-Hoeffding and the union bound, the probability of any of these events is $\exp(-\Omega(\log^2(n))) = o(1/\log(n))$, as $m_r \geq \log^4(n)$, and the lemma follows by induction. □

Now, let $T > \sigma n$ and suppose that $0 < Q_Y(u_T) < \log^4(n)$. We couple again with a branching process as in Section 4 initiating it at time u_T with $|Y_i(u_T)|$ particles of type i. This time, however we sample from the fixed set $Q(u_T)$ of configuration positions with fixed sets of susceptible $S_1(u_T), \dots, S_\Delta(u_T)$. If a sampled configuration position does not lie in this set then no offspring is produced.

As above, let $X_i(t)$ be the number of particles of type i alive at time t, and let

$$W(t) = \sum_{i=1}^{\Delta-1} iX_i(t).$$

Note that the coupling fails if our random choice of configuration position hits the same vertex twice *or* hits a vertex from $Y(u_T)$. Provided Q_Y remains small enough, this occurs with probability $o(1)$. Set $v_0 = u_T$ and for $i \geq 1$ set $v_i = u_{T+i} = v_{i-1} + 1/Q_Y(v_{i-1})$.

Proceeding as above, we obtain that the expected value matrix $M(t) = \exp(B_n t)$, where

$$B_n = \lambda(A + \mathbf{u}\mathbf{D}_S(u_T)^T) - \rho I.$$

Furthermore

$$E\left[W_{j+1}|\mathcal{F}_{v_j}\right] = \exp\left\{\lambda(\hat{d}_s(u_T) - 1) - \rho\right\} W_j.$$

Note that

$$\hat{D}_S(u_T) = \hat{d}_s(x^*) + O\left(\frac{\log^4(n)}{n^{1/2}}\right)$$

by Lemma 22. For n sufficiently large we have

$$\lambda(\hat{D}_S(u_T) - 1) - \rho < 0$$

and the W_i form a supermartingale. Finally, we apply the analysis of Section 4.2 directly by viewing the branching process as $|Y(u_T)|$ independent branching processes initiated with individual particles, and applying the union bound to each we have with probability $1-o(1)$, $Q_Y(u_T + \log(n)) = 0$ and at most $O(\log^{16/3}(n))$ additional infections are sent, completing the proof.

Acknowledgment. We thank the anonymous referees for many helpful suggestions.

References

[1] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley 2000.

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

[3] H. Andersson, Limit theorems for a random graph epidemic model, *Annals of Applied Probability* **8** (1998), 1331-1349.

[4] K.B. Athreya and P.E. Ney, *Branching Processes*, Springer-Verlag, 1970.

[5] F. Ball, The threshold behaviour of epidemic models, *J. Appl. Probab.* **20** (1983), 227-241.

[6] F. Ball and P. Donnelly, Strong approximations for epidemic models, *Stochastic Proesses and their Applications* **55** (1995), 1-21.

[7] F. Ball and P. Neal, Network epidemic models with two levels of mixing, *Mathematical Biosciences* **212** (2008), 69-87.

[8] A.D. Barbour, The duration of the closed stochastic epidemic, *Biometrika* **62** (1975b) 477-482.

[9] A. Beveridge, T. Bohman, A. Frieze and O. Pikhurko, Product rule wins a competitive game. *Proceedings of the American Mathematical Society* **135** (2007) 3061-3071.

[10] T. Bohman and A. Frieze, Karp-Sipser on random graphs with a fixed degree sequence, submitted.

[11] B. Bollobás, *Random Graphs*, Cambridge University Press, 2001.

[12] T. Britton, Stochastic epidemic models: A survey, *Mathematical Biosciences* **225** (2010), 24-35.

[13] T. Britton, S. Janson, and A. Martin-Löf, Graphs with specificied degree distributions, simple epidemics, and local vaccination strategies, *Advances in Applied Probility* **39** (2007), 922-948.

[14] D. Fernholz and V. Ramachandran, The diameter of sparse random graphs, *Random Structures and Algorithms* **31** (2007), 482-516.

[15] A. Frieze, M. Krivelevich, C. Smyth, On the chromatic number of random graphs with a fixed degree sequence. *Combinatorics Probability and Computing*, **16** (2007) 733–746.

[16] H. Hatami, M. Molloy, The scaling window for a random graph with a given degree sequence, arXiv:0907.4211.

[17] W. Hurewicz, *Lectures on Ordinary Differential Equations*, Dover Publications, 2002.

[18] S. Janson, The largest component in a subcritical random graph with a power law degree distribution, *Annals of Applied Probability* **18** (2008), 1651-1668.

[19] S. Janson, The probability that a random multigraph is simple. *Combinatorics, Probability and Computing* **18** (2009), no. 1-2, 205-225.

[20] S. Janson and M. Luczak, A simple solution to the k-core problem, *Random Structures and Algorithms* **30** (2007), 50 - 62.

[21] S. Janson and M. Luczak, Asymptotic normality of the k-core in random graphs, *Annals of Applied Probability* **18** (2008), 1085-1137.

[22] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, John Wiley and Sons, 2000.

[23] M. Kang, T. Seierstad, The critical phase for random graphs with a given degree sequence, *Combinatorics, Probability and Computing* **17** (2008), 67-86.

[24] M. Keeling and K. Eames, Networks and epidemic models. *Journal of the Royal Society Interface* **2** (2005) 295-307.

[25] W. O. Kermack and A. G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, *Proc. Roy. Soc. Lond. A* **115** (1927), 700-721.

[26] J. Metz, The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small iniital infections, *Acta Biotheoretica* **27** (1978), 75-123.

[27] J. Miller, A note on a paper by Erik Volz [arXiv:0705.2092]: SIR dynamics in random networks, arXiv:0909.4485.

[28] D. Mollison, Spatial contact models for ecological and epidemic spread, *J. Roy. Stat. Soc. B* **39** (1977), 283-326.

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

[30] M. Molloy and B. Reed, The Size of the Largest Component of a Random Graph on a Fixed Degree Sequence. *Combinatorics, Probability and Computing* **7** (1998) 295-306.

[31] P. Neal, SIR epidemics on a Bernoulli random graph, *Journal of Applied Probility* **40** (2003), 779-782.

[32] M. Newman, Spread of epidemic disease on networks, *Phys. Rev. E* **66** (2002), 016128.

[33] M. Newman, S. Strogatz, and D. Watts, Random graphs with arbitrary degree distributions and their applications, *Physical Review E* **64** (2001) 026118.

[34] B. Pittel, The largest component in a subcritical random graph with a power law degree distribution, *Annals Applied Probability* **18** (2008), 1636-1650.

[35] J. Spencer and N. Wormald, Birth control for giants. *Combinatorica*, **27** (2007) 587–628.

[36] A. Telcs, N. Wormald and S. Zhou, Hamiltonicity of random graphs produced by 2-processes, *Random Structures and Algorithms* **31** (2007), 450-481.

[37] E. Volz, SIR Dynamics in Random Networks with Heterogeneous Connectivity, *Journal of Mathematical Biology*, DOI: 10.1007/s00285-007-0116-4 (Published online: 1 August 2007)

[38] N. Wormald, The differential equations method for random graph processes and greedy algorithms. pp. 73-155 in *Lectures on Approximation and Randomized Algorithms*, Karonski and Prömel eds. PWN, Warsaw 1999.