SUBCRITICAL EPIDEMICS ON RANDOM GRAPHS

OANH NGUYEN AND ALLAN SLY

Abstract. We study the contact process on random graphs with low infection rate \( \lambda \). For random \( d \)-regular graphs, it is known that the survival time is \( O(\log n) \) below the critical \( \lambda_c \). By contrast, on the Erdős-Rényi random graphs \( G(n, d/n) \), rare high-degree vertices result in much longer survival times. We show that the survival time is governed by high-density local configurations. In particular, we show that there is a long string of high-degree vertices on which the infection lasts for time \( n^{2+o(1)} \). To establish a matching upper bound, we introduce a modified version of the contact process which ignores infections that do not lead to further infections and allows for a sharper recursive analysis on branching process trees, the local-weak limit of the graph. Our methods, moreover, generalize to random graphs with given degree distributions that have exponential moments.

1. Introduction

Stochastic processes on random networks can be affected by both global properties of the graph like expansion but also local properties like the local-weak limit and the presence of rare local neighbourhoods. Our goal in this paper is to investigate what properties of sparse Erdős-Rényi random graphs determine the survival time for infection processes.

The contact process is a model for the spread of infections a graph \( G = (V, E) \). For a given \( \lambda > 0 \), the contact process \( (X_t)_{t \geq 0} \) with infection rate \( \lambda \) and recovery rate 1 is a continuous-time Markov chain, with state space \( \{0, 1\}^V \), tracking if a vertex \( v \) is either infected \( (X_t(v) = 1) \) or healthy \( (X_t(v) = 0) \). The process evolves according to the following rules:

- Each infected vertex infects each of its neighbors independently at Poisson rate \( \lambda \) and is healed at Poisson rate 1;
- Infection and recovery events in the process happen independently.

On finite graphs, the contact process has a single absorbing state which is no infections and we define the time to reach it as the survival time. The primary question is how quickly this occurs and how it depends on \( \lambda \). We can ask if the behaviour on Erdős-Rényi random graphs is the same as for random \( d \)-regular graphs. In terms of the threshold of \( \lambda \) between fast (polynomial) and slow (exponential) survival times, the answer is yes, both are at approximately \( \frac{1}{d} \) (see [26, 27]) which is determined by the global branching rate of the process. When \( \lambda \) is small, the contact process on random regular graphs survives for time \( O(\log n) \) (see [14, 25]). One can ask if the survival time is still logarithmic for other sparse random graphs, in particular \( G(n, d/n) \)? The answer turns out to be negative for a surprisingly simple reason. Since the largest degree in an Erdős-Rényi random
graph is asymptotically $\frac{\log n}{\log \log n}$, this directly gives a lower bound of $\exp\left(\frac{c \log n}{\log \log n}\right)$ which far exceeds $\log n$.

It is then natural to ask whether $\exp\left(\frac{c \log n}{\log \log n}\right)$ gives the correct order of the survival time. The convergence to equilibrium of the high-temperature Glauber dynamics of the Ising model gives an interesting comparison. Its mixing time is determined by coupling the all plus and minus initial condition and the set of disagreements can be thought of as a kind of infection process where updates can remove disagreements or spread them to neighbours. In [21], it was shown that the mixing time of the Glauber dynamics for the high temperature continuous-time ferromagnetic Ising model on $\mathcal{G}(n, d/n)$ is $\exp\left(O(1) \frac{\log n}{\log \log n}\right)$. It is essentially determined by the time to mix around the maximal degree vertices in the graph. We show, by contrast, that the contact process survives for time polynomial in $n$, much longer than would be predicted by just the high degree vertices. In the following theorem we, moreover, determine the exponent of the polynomial up to the log factors.

**Theorem 1.1.** Let $d > 1$ and $G_n \sim \mathcal{G}(n, d/n)$. There exist positive constants $\lambda_*(d), C(d)$ such that the following holds. The contact process on $G_n$ starting at all vertices infected and infection rate $\lambda < \lambda_*$ has survival time $T_{\lambda,n}$ that satisfies

$$n^{c\lambda^2/\log^2(\lambda^{-1})} \leq T_{\lambda,n} \leq n^{C\lambda^2 \log(\lambda^{-1})}$$

with probability $1 - o(1)$ as $n \to \infty$ where $c = 10^{-9}$.

This result is a special case of our more general results on random graphs with given degree distributions described later in the section. As we will describe, the survival time comes from large connected components of vertices of large constant degree. The discrepancy between the contact process and the Glauber dynamics can be explained by the existence of a phase for the contact process in dimensional 1 but not for the Glauber dynamics. As a result, it is different rare local structures which determine the survival times.

1.1. **Background.** The contact process was first introduced by a work of Harris [10] in which he studied the process on the lattice $\mathbb{Z}^d$. Among other things, he studied the phase diagrams of the contact process which since then has attracted intensive research. For an infinite rooted graph $G$, there are three phases that are of particular interest:

- *(Extinction)* the infection becomes extinct in finite time almost surely;
- *(Weak survival)* the infection survives forever with positive probability, but the root is infected only finitely many times almost surely;
- *(Strong survival)* the infection survives forever and the root gets infected infinitely many times with positive probability.

In that sense, we denote the extinction-survival threshold by

$$\lambda_1(G) = \inf\{\lambda : (X_t) \text{ survives forever with positive probability}\}$$

and the weak-strong survival threshold by

$$\lambda_2(G) = \inf\{\lambda : (X_t) \text{ survives strongly}\}.$$  

For lattices, when the origin is initially infected, it is well known that there is no weak survival phase (see [10], Bezuidenhout-Grimmett [1], the books of Liggett [16], [17] and the references therein).
Beyond lattices, another natural family of graphs to study the contact process is infinite trees. For the infinite $d$-regular tree $\mathbb{T}_d$ with $d \geq 3$, the contact process with the root initially infected has two distinct phase transitions with $0 < \lambda_1(\mathbb{T}_d) < \lambda_2(\mathbb{T}_d) < \infty$, by a series of beautiful works by Pemantle [27] (for $d \geq 4$), Liggett [15] (for $d = 3$), and Stacey [30]. For a Galton-Watson tree $T$ with offspring distribution $\xi$, it is not difficult to see that $\lambda_1(T)$ and $\lambda_2(T)$ are constants depending only on $\xi$, conditioned on $|T| = \infty$. Huang and Durrett [15] proved that on $T \sim \text{GW}(\xi)$ with the root initially infected, $\lambda_2(T) = 0$ if the offspring distribution $\xi$ is subexponential, i.e., $\mathbb{E} e^{c\xi} = \infty$ for all $c > 0$. So in this case, there is only the strong survival phase.

By contrast, if the offspring distribution $\xi$ has an exponential tail, i.e., $\mathbb{E} e^{c\xi} < \infty$ for some $c > 0$, Bhamidi, Nam and the authors [2] showed that there is an extinction phase: $\lambda_1(T) \geq \lambda_0(\xi)$ for some constant $\lambda_0(\xi) > 0$. In [26], Nam et. al. established the first-order asymptotics on $\lambda_1^{GW}(\xi)$ for $\xi$ sufficiently well concentrated around its mean, in particular the case of Poisson. The asymptotics turns out to have the same form as for regular trees.

Aside from infinite graphs, there has been considerable work studying the phase transitions of survival times on large finite graphs. Stacey [29] and Cranston-Mountford-Mourrat-Valesin [4] studied the contact process on the $d$-ary tree $\mathbb{T}_d^h$ of depth $h$ starting from the all-infected state, and their results show that the survival time $T_h$, as $h \to \infty$, satisfies (i) $T_h/h \to \gamma_1$ in probability if $\lambda < \lambda_2(\mathbb{T}_d)$; (ii) $|\mathbb{T}_d^h|^{-1/2} \log T_h \to \gamma_2$ in probability and $T_h/\mathbb{E} T_h \stackrel{d}{\to} \text{Exp}(1)$ if $\lambda > \lambda_2(\mathbb{T}_d)$, where $\gamma_1, \gamma_2$ are constants depending on $d, \lambda$. In [7] [8] [24], similar results were established for the case of the lattice cube $\{1, \ldots, n\}^d$.

Recently in work of Mourrat-Valesin [25] and Lalley-Su [14], it was shown that for any $d \geq 3$, the contact process on the random $d$-regular graph, whose initial configuration is the all-infected state, exhibits the following phase transition:

- **(Short survival)** For $\lambda < \lambda_1(\mathbb{T}_d)$, the process survives for $O(\log n)$-time whp.
- **(Long survival)** For $\lambda > \lambda_1(\mathbb{T}_d)$, the process survives for $e^{\Theta(n)}$-time whp.

For related results, we refer to [3] [4] [14] [18] [22] [23] [25] [26] and the references therein.

Let us move on to finite graphs that are non-regular. We focus on the contact process on random graphs with a given degree distribution. Let $\mu$ be a degree distribution and $G_n \sim \mathcal{G}(n, \mu)$ be a random graph with degree distribution $\mu$, assuming that there exists a unique giant component (for details, see Section 2.2). Consider the contact process on $G_n$ where all vertices are initially infected. In [2], it was shown that if $\mathbb{E}_{D \sim \mu} e^{cD} < \infty$ for some constant $c > 0$, then there exist constants $0 < \underline{\lambda}(\mu) \leq \overline{\lambda}(\mu) < \infty$ such that the survival time $T_{\lambda,n}$ of the process, which is the first time the process reaches the all-healthy state, satisfies the following:

1. For all $\lambda < \underline{\lambda}$, $T_{\lambda,n} \leq n^{1+o(1)}$ whp;
2. For all $\lambda > \overline{\lambda}$, $T_{\lambda,n} \geq e^{\Theta(n)}$ whp.

Based on this result, the short- and long-survival thresholds $\lambda^-_c(\mu), \lambda^+_c(\mu)$ are defined as follows.

$$\lambda^-_c(\mu) := \lim_{\alpha \to \infty} \sup \left\{ \lambda : \lim_{n \to \infty} \mathbb{P}(T_{\lambda,n} \leq n^\alpha) = 1 \right\};$$
$$\lambda^+_c(\mu) := \lim_{\beta \to 0} \inf \left\{ \lambda : \lim_{n \to \infty} \mathbb{P}(T_{\lambda,n} \geq e^{\beta n}) = 1 \right\}.$$
The first-order asymptotics for these thresholds are achieved \cite[Theorem 2]{26}:
\[\lambda_c^- (\mu) \sim \frac{1}{\mathbb{E}_{D \sim \mu} D} \quad \text{and} \quad \lambda_c^+ (\mu) \sim \frac{1}{\mathbb{E}_{D \sim \tilde{\mu}} D}\]
where \(\tilde{\mu}\) is the size-biased distribution of \(\mu\) (see \cite{4} for the precise definition).

On the other hand, if \(\mu\) has a heavier than exponential tail, the story is quite simple \cite{2}: \textit{whp} there is no short survival phase and so, \(\lambda_c^- (\mu) = \lambda_c^+ (\mu) = 0\). Therefore, we shall only restrict to measures with exponential tails in this paper. For these measures, while it is now known that the survival time for \(\lambda > \lambda_c^+ (\mu)\) is exponential \(e^{\Theta(n)}\), little is known about the survival time for small \(\lambda\) except that it is of order \(n^{O(1)}\).

1.2. Main results. As promised, we now present generalizations of Theorem 1.1 for random graphs with given degree distributions. In the following, we state the lower and upper bounds for the survival time separately because the assumptions needed for the lower bound are more relaxed.

**Theorem 1.2** (Lower bound). Let \(\mu\) be any probability measure satisfying \(\mathbb{E}_{D \sim \mu} D (D - 2) > 0\). Let \(G_n \sim \mathcal{G}(n, \mu)\). Let \(\lambda > 0\) be sufficiently small so that \(\lambda < 1/20\) and
\[A = \frac{10^6}{\lambda^2} \log \frac{1}{\lambda} \geq 16 \mathbb{E}_{D \sim \mu} D.\]
Consider the contact process on \(G_n\) where all vertices are initially infected and infection rate \(\lambda\). Then this process still survives at time \(n^{-c/\log(\mu[A, \infty])}\) with probability \(1 - o(1)\) where \(c = 1/120\) and the implicit constant in the \(o(1)\) may depend on \(\lambda\) and \(\mu\).

When \(\mu\) has a truly exponential tail \(\mu[A, \infty) \approx \exp(-A^{1+o(1)})\) for \(A\) sufficiently large, this theorem concludes that the survival time is \(\geq n^{-c\lambda^2 + o(1)}\) with high probability.

Concerning the upper bound, we achieve the same power for the survival time.

**Theorem 1.3** (Upper bound). Let \(\mu\) be a distribution satisfying \(\mathbb{E}_{D \sim \mu} D (D - 2) > 0\) and
\[\mathbb{E}_{D \sim \mu} e^{10D} < \infty.\]
There exists a constant \(C = C(\mu) > 0\) such that for all \(\lambda < 1/C\), the contact process on \(G_n \sim \mathcal{G}(n, \mu)\) starting at all vertices infected and infection rate \(\lambda\) dies before time \(n^{C\lambda^2 \log(\lambda^{-1})}\) with probability \(1 - o(1)\) as \(n \to \infty\) where the implicit constant may depend on \(\lambda\) and \(\mu\).

The constant 10 in \((3)\) may not be optimal.

1.3. Main ideas. In the first part of the paper, we prove the lower bound Theorem 1.2. It is well-known that the contact process survives for a long time on dense sub-graphs. The simplest example is a large star graph, a vertex with \(k\) neighbors. It is easy to verify that the expected survival length behaves like \(e^{ck\lambda^2}\) (see Lemma 3.4). Since the largest degree in an Erdős–Rényi random graph is asymptotically \(\frac{\log n}{\log \log n}\), this directly gives a lower bound of \(\exp(\frac{c \log n}{\log \log n})\) which far exceeds the \(\log n\) survival time in the random regular graph. However, these highest degree vertices are not the optimal local structure. Instead we find that for paths of vertices of degree \(10^6 \lambda^{-2} \log \frac{1}{\lambda}\), the contact process behaves effectively like a supercritical one-dimensional contact process. This has an infection time exponential in the length of the path and we show that with high probability there are paths of logarithmic length which gives the bound in Theorem 1.1.
The second, and more challenging, part is to prove the matching upper bound of Theorem 1.3. As in [2, 26], we first control the contact process on Galton-Watson Branching process trees which are the local weak limits of the random graphs. In [2], we developed a recursive scheme to show that the expected survival time of an infection started from the root is constant, independent on the depth of the tree. This was further refined in [26], giving tail bounds on the conditional expected survival time given the tree as well as for the expected number of infections that reach the leaves. These tail bounds, however, would only translate into an upper bound on the expected survival time on the random graph of $n^{\lambda^{1+o(1)}}$. To explain where this analysis was deficient, it effectively coupled the process with a modified process where the root of the tree must remain infected until all its subtrees became uninfected. The advantage of this coupling is that the process in the subtrees becomes independent. However, if the degree of the root is $D$ then in the modified process, it will take time $\exp(c\lambda D)$ rather than $\exp(c\lambda^2 D)$ until all the children of the root become uninfected. What this analysis misses is that each infected child has only probability $\lambda$ of sending an infection back to its parent.

To give a sharp analysis, we introduce a modified dynamics that stochastically dominates the original where particles are strongly infected if they will send an infection to a neighbour before healing and weakly infected otherwise. Weak infections can then be effectively ignored. A vertex $u$ sends a strong infection to a neighbour $v$ with rate $\frac{\lambda d_v^2}{1+\lambda d_v}$ which is roughly $C\lambda^2$ when the degree of $v$ is constant. With this formulation, we get sharper bounds on the tail of the expected survival time and number of particles to hit the leaves.

To complete the proof, we use this analysis to show by a union bound that for any vertex $v$, starting with $v$ infected, the expected number of infections to exit a ball of radius $C\lambda^2 \log(\lambda^{-1}) \log n$ is $o(1)$. When infections leave the ball, we analyse them as new contact processes started from the exit point and stochastically dominate this process by a subcritical branching process which with high probability terminates in $n^{C\lambda^2 \log(\lambda^{-1})}$ steps. Furthermore, we again use our recursive bounds to show that the maximum expected survival time on balls of radius $C\lambda^2 \log(\lambda^{-1}) \log n$ in the graph is at most $n^{C\lambda^2 \log(\lambda^{-1})}$ with high probability. Combining these results establishes Theorem 1.3.

1.4. Notation. For a tree $T$ and a depth $l$, we denote by $T_l$ the set of vertices of $T$ at depth $l$. We use $T_{\leq l}$ to denote the set of vertices of depth at most $l$. In particular, $T_{\leq l} \sim GW(\xi)_{\leq l}$ denotes the Galton-Watson tree generated up to depth $l$, while the infinite Galton-Watson tree is denoted by $T \sim GW(\xi)$.

Throughout the paper, we often work with the contact process defined on a (fixed) graph generated at random. To distinguish between the two randomness of different nature, we introduce the following notations:

- $\mathbb{P}_{CP}$ and $\mathbb{E}_{CP}$ denote the probability and the expectation, respectively, with respect to the randomness from contact processes.
- $\mathbb{P}_{GW}$ and $\mathbb{P}_{RG}$ denote the probability with respect to the randomness from the underlying graph, when the graph is a Galton-Watson tree and a random graph $G(n, \mu)$, respectively. We write $\mathbb{E}_{GW}$ and $\mathbb{E}_{RG}$ similarly for expectations.
- $\mathbb{P}$ and $\mathbb{E}$ denote the probability and expectation, respectively, with respect to the combined randomness over both the process and the graph. That is, for instance, $\mathbb{E}[\cdot] = \mathbb{E}_{GW}[\mathbb{E}_{CP}[\cdot]]$, if the underlying graph is a Galton-Watson tree.
1.5. **Organization.** In Section 2, we include the preliminaries for the graphical construction of the contact process (Section 2.1), the models of random graphs (Section 6.4). Several useful results on the contact process on star graphs are presented in Section 3. We prove Theorem 1.2 on the lower bound in Section 4. In Section 5, we establish the analog of Theorem 1.3 for trees. And finally, we prove Theorem 1.3 in Section 6.

2. **Preliminaries**

2.1. **Graphical representation of contact processes.** A standard construction of the contact processes uses the *graphical representation*, see for example [17, Chapter 3, section 6]. The idea is to record the infections and recoveries of the contact process on a graph $G$ on the space-time domain $G \times \mathbb{R}_+$. Define i.i.d. Poisson processes $\{N_v(t)\}_{v \in V}$ with rate 1 (that represents the healing clock rings at vertex $v$) and i.i.d. Poisson processes $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \vec{E}}$ with rate $\lambda$ (that represents the clock rings for the infection from $u$ to $v$), where we consider $\vec{E} = \{\vec{uv}, \vec{vu} : (uv) \in E\}$ to be the set of directed edges. Further, we let $\{N_v(t)\}_{v \in V}$ and $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \vec{E}}$ to be mutually independent.

Then the graphical representation is defined as follows:

1. Initially, we have the *empty* domain $V \times \mathbb{R}_+$.
2. For each $v \in V$, mark $\times$ at the point $(v, t)$, at each event time $t$ of $N_v(\cdot)$.
3. For each $\vec{uv} \in \vec{E}$, add an arrow from $(u, t)$ to $(v, t)$, at each event time $t$ of $N_{\vec{uv}}(\cdot)$.

![Figure 1. A realization of the contact process on the interval $V = \{1, \ldots, 5\}$, with initial condition $X_0 = 1_V$. The blue lines describe the spread of infection. We see that $X_s = 1_{\{2,3\}}$.](image)

This gives a geometric picture of the contact process, and further provides a coupling of the processes over all possible initial states. Figure 1 tells us how to interpret the infections at time $t$ based on this graphical representation. The following are two simple yet important observations using this construction.

**Lemma 2.1** ([17]). *Suppose that we have the aforementioned coupling among the contact processes on a graph $G$. Let $T_v$ be the survival time of the contact process on $G$ with $v$ initially infected. Let $T_G$ be the survival time of the contact process on $G$ with all vertices initially infected. Then we have $T_G = \max\{T_v : v \in G\}$.*
Lemma 2.2 ([17]). For a given graph $G = (V,E)$ and any $A \subset V$, let $(X_t)$ be the contact process on $G$ with $A$ initially infected. Consider any (random) subset $I$ of $\mathbb{R}_+$, and define $(X'_t)$ to be the coupled process of $(X_t)$ that has the same initial state, infections and recoveries, except that the recoveries at a fixed vertex $v$ are ignored at times $t \in I$. Then for any $t \geq 0$, we have $X_t \leq X'_t$, i.e., $X_t(v) \leq X'_t(v)$ for all $v$.

2.2. Random graphs and their limiting structure. Let $\mu$ be a probability distribution on $\mathbb{N}$. The random graph $G_n \sim \mathcal{G}(n,\mu)$ with degree distribution $\mu$ is defined as follows:

- Let $d_i \sim \text{i.i.d. } \mu$ for $i = 1, \ldots, n$, conditioned on $\sum_{i=1}^n d_i \equiv 0 \mod 2$. The numbers $d_i$ refer to the number of half-edges attached to vertex $i$.
- Generate the graph $G_n$ by pairing all half-edges uniformly at random.

The resulting graph $G_n$ is also called the configuration model. One may also be interested in the uniform model $G_n^u \sim \mathcal{G}^u(n,\mu)$, which picks a uniformly random simple graph among all simple graphs with degree sequence $\{d_i\}_{i \in [n]} \sim \text{i.i.d. } \mu$. It is well-known that if $\mu$ has a finite second moment, then the two laws $\mathcal{G}(n,\mu)$ and $\mathcal{G}^u(n,\mu)$ are contiguous, in the sense that for any subset $A_n$ of graphs with $n$ vertices,

$$\Pr_{G_n \sim \mathcal{G}(n,\mu)}(G_n \in A_n) \to 0 \quad \text{implies} \quad \Pr_{G_n^u \sim \mathcal{G}^u(n,\mu)}(G_n^u \in A_n) \to 0.$$ 

For details, we refer the reader to Chapter 7 of [31] or to [12]. We remark that when $\mu = \text{Pois}(d)$, the random graph $G_n \sim \mathcal{G}(n,\mu)$ is contiguous to the Erdős-Rényi random graph $G_n^{ER} \sim \mathcal{G}^{ER}(n,d/n)$ as shown in [13], Theorem 1.1.

Furthermore, it is also well-known that the random graph $G_n \sim \mathcal{G}(n,\mu)$ is locally tree-like, and the local neighborhoods converge locally weakly to Galton-Watson trees. To explain this precisely, let us denote the law of Galton-Watson tree with offspring distribution $\mu$ by $\text{GW}(\mu)$, and let $\text{GW}(\mu)_{\leq l}$ be the law of $\text{GW}(\mu)$ truncated at depth $l$, that is, the vertices with distance $> l$ from the root are removed. Further, let $\bar{\mu}$ denote the size-biased distribution of $\mu$, defined by

$$\bar{\mu}(k-1) := \frac{k \mu(k)}{\sum_{k'=1}^{\infty} k' \mu(k')}, \quad k = 1, 2, \ldots.$$ 

Note that if $\mu = \text{Pois}(d)$, then $\bar{\mu} = \mu$. Lastly, define $\text{GW}(\mu, \bar{\mu})_{\leq l}$ to be the Galton-Watson process truncated at depth $l$, such that the root has offspring distribution $\mu$ while all other vertices have offspring distribution $\bar{\mu}$. Then the following lemma shows the convergence of local neighborhoods of $G_n$.

Lemma 2.3 ([6], Section 2.1). Suppose that $\mu$ has a finite mean. Let $l > 0$ and let $v$ denote the vertex in $G_n \sim \mathcal{G}(n,\mu)$ chosen uniformly at random. Then for any rooted tree $(T,x)$ of depth $l$, we have

$$\lim_{n \to \infty} \Pr((N(v,l),v) \cong (T,x)) = \Pr((T,\rho) \sim \text{GW}(\mu,\bar{\mu})_{\leq l})((T,\rho) \cong (T,x)),$$

where $N(v,l)$ is the $l$-neighborhood of $v$ in $G_n$ and $\cong$ denotes the isomorphism of rooted graphs. We say that $G_n$ converges locally weakly to $\text{GW}(\mu,\bar{\mu})$.

We remark that the same holds for a fixed vertex $v \in G_n$. Moreover, we also stress that the condition for $\text{GW}(\mu,\bar{\mu})$ to be supercritical is equivalent to the condition for $\mathcal{G}(n,\mu)$ to have the unique giant component whp (see e.g., [19], or [9], Section 3 for details), which can be addressed as

$$\mathbb{E}_{D \sim \mu} D(D - 2) > 0.$$
3. Contact process on star graphs

In this section, we discuss several results about the contact process on star graphs that will be useful later. For a positive integer \( k \), let \( S_k \) be a star graph with a root, which we usually denote by \( \rho \), and \( k \) leaves. Our goal is to provide a lower bound for the survival time of the contact process on \( S_k \) which we state in Lemma 3.4. Before doing that, we shall need some preparation. Firstly, we show that the root infects many leaves before its recovery.

**Lemma 3.1.** Consider the contact process \((X_t)\) with infection rate \( \lambda \leq 1 \) on the star graph \( S_k \). Let \( \ell = \lambda k \). Assume that the root is infected at time \( 0 \). Let \( C \leq \frac{\ell}{32 \log \ell} \) be a positive number. Then the probability that \( \rho \) infects at least \( C \log \ell \) leaves who remain infected during time \( \left[ \frac{16C \log \ell}{\ell}, 1 \right] \supset [1/2, 1] \) is at least \( 1 - \frac{16}{C \log \ell} - 4\ell^{-C/4} \).

**Proof.** Let \( s = \frac{16C \log \ell}{\ell} \). The probability that the root is still infected until time \( s \) is \( \mathbb{P}(\text{Pois}(s) = 0) = e^{-s} \geq 1 - s = 1 - \frac{16C \log \ell}{\ell} \).

The probability that there are at least \( 8 \log \ell \) infection rings from \( \rho \) to its leaves in time \( [0, s] \) is \( \mathbb{P}(\text{Pois}(\ell s) \geq 8 \log \ell) = \mathbb{P}(\text{Pois}(16C \log \ell) \geq 8 \log \ell) \geq 1 - 2 \exp(-8 \log \ell/6) \geq 1 - 2\ell^{-C} \)

where we use the following standard tail bound for Poisson distribution:

\[
(6) \quad \mathbb{P}(\left| \text{Pois}(\Lambda) - \Lambda \right| \geq x) \leq 2 \exp \left( -\frac{x^2}{2(\Lambda + x)} \right), \quad \forall x > 0.
\]

Under this event, letting \( 8 \log \ell =: 2M \), the probability that these rings correspond to at least \( M \) distinct leaves is at least

\[
1 - \left( \frac{k}{M} \right) M^{2M} k^{2M} \geq 1 - \left( \frac{ekM^2}{Mk^2} \right)^M \geq 1 - \left( \frac{4C \log \ell}{k} \right)^{4C \log \ell} \geq 1 - \exp(-4C \log \ell) = 1 - \ell^{-4C}
\]

where we used \( C \leq k/(32 \log \ell) \).

For each leaf infected before time \( s \), the probability that it remains infected until time \( 1 \) is at least \( \mathbb{P}(\text{Pois}(1) = 0) = e^{-1} \). Thus, given that there are at least \( M \) leaves infected before time \( s \), at least \( M/4 \) of them remain infected until time \( 1 \) with probability at least \( 1 - \exp(-M/16) = 1 - \ell^{-C/4} \) where we used the Chernoff inequality

\[
(7) \quad \mathbb{P}(\text{Bin}(M, q) \leq Mq\delta) \leq \exp \left( -\frac{(1 - \delta)^2 Mq}{2} \right), \quad \forall M \in \mathbb{N}, \ q, \delta \in (0, 1).
\]

Combine all of these events and the union bound, we obtain the desired statement. \( \square \)

The next two lemmas show that once many leaves are infected, they keep the root well-infected which then keeps the number of infected leaves from decreasing.

**Lemma 3.2.** Consider the contact process \((X_t)\) with infection rate \( \lambda \leq 1 \) on the star graph \( S_k \). Let \( M \) by an integer satisfying \( \frac{k}{4} \leq M \leq k \). Assume that at least \( M \) leaves are infected at time \( t \) and remain infected until time \( t + 1/2 \). The probability that \( \rho \) is infected for at least half of the time interval \([t, t + 1/2]\) is at least \( 1 - 2 \exp(-\lambda M/32) \).
Proof. Conditioned on the event that the $M$ leaves remain infected during $I := [t, t + 1/2]$, it holds that $\rho$ is infected with rate at least $\lambda M$. Let $Y \sim \text{Pois}(1/2)$ be the number of healing clocks at $\rho$ in $I$. The probability that $\rho$ is healthy for at least $1/2$ of $I$ is at most

$$
\mathbb{P} \left( \sum_{i=1}^{Y} \eta_i \geq 1/4 \right),
$$

where $\eta_i \sim \exp(\lambda M)$ is the waiting time until the next infection after the $i$-th healing clock. We note that these $\eta_i$ are iid. We have

$$
\mathbb{P} \left( \sum_{i=1}^{Y} \eta_i \geq 1/4 \right) \leq \mathbb{P}(Y \geq \lambda M/8) + \mathbb{P} \left( \sum_{i=1}^{\lambda M/8} \eta'_i \geq \lambda M/4 \right)
$$

where $\eta'_i = \eta_i \lambda M \sim \exp(1)$

$$
\leq \exp(-\lambda M/32) + \mathbb{P} \left( \sum_{i=1}^{\lambda M/8} \eta'_i \geq \lambda M/4 \right) \leq \exp(-\lambda M/32) + \mathbb{P} \left( \sum_{i=1}^{\lambda M/8} \eta'_i \geq \lambda M/4 \right)
$$

by (6)

$$
\leq \exp(-\lambda M/32) + \exp(-\lambda M/8) \mathbb{E} \exp \left( \sum_{i=1}^{\lambda M/8} \eta'_i / 2 \right)
$$

$$
\leq \exp(-\lambda M/32) + 2^{\lambda M/8} \exp(-\lambda M/8) \leq 2 \exp(-\lambda M/32).
$$

This completes the proof. \square

Lemma 3.3. Consider the contact process $(X_t)$ with infection rate $\lambda < 1$ on the star graph $S_k$. Let $M$ be a positive integer satisfying $\frac{32}{\lambda} \leq M \leq \frac{M}{100}$. Assume that at least $M$ leaves are infected at time $t$. Then the probability that at least $\frac{M}{100}$ leaves are infected at time $t + 1$ is at least $1 - 3 \exp(-\lambda M/128)$.

Proof. Let $m$ be the number of leaves infected at time $t$. We have $m \geq M$. Consider two cases.

- Case 1: $m \geq k/2$. Each of these infected leaves has a probability of $\mathbb{P}(\text{Pois}(1) = 0) = e^{-1}$ to remain infected until time $t + 1$. Thus, by Chernoff inequality, at least $m/4 \geq \frac{M}{100}$ of them remain infected until time $t + 1$ with probability at least $1 - \exp(-m/12) \geq 1 - 3 \exp(-\lambda M/128)$.

- Case 2: $M \leq m \leq k/2$. By the same argument as in Case 1, with probability at least $1 - \exp(-m/16)$, at least $M/4$ of these $m$ infected vertices remain infected until time $t + 1/2$. Under this event, by Lemma 3.2, with probability at least $1 - 2 \exp(-\lambda M/128)$, $\rho$ is infected for at least half of the time interval $[t, t + 1/2]$. Thus, each of the remaining $k - m \geq k/2$ leaves receives an infection from $\rho$ during $[t, t + 1/2]$ with probability at least

$$
\mathbb{P}(\text{Pois}(\lambda/2) \geq 1) = 1 - e^{-\lambda/2} \geq \lambda/4.
$$

Once they receive an infection, the probability that they remain infected until time $t + 1$ is at least $e^{-1}$. Thus, the probability that at least $\frac{M}{100}$ vertices are infected at time $t + 1$ is at least $1 - \exp(-\lambda k/100)$.

This gives the desired statement for both cases. \square
Combining these lemmas, we obtain the following lower bound on the survival time of the contact process on star graphs.

**Lemma 3.4.** Consider the contact process \((X_t)\) with infection rate \(\lambda\) on the star graph \(S_k\). Assume that \(\rho\) is infected at time 0. Then the contact process still survives at time \(e^{\lambda^2 k/10^5}\) with probability at least \(1 - \frac{200 \log(\lambda k)}{\lambda^2 k} - 3e^{-\lambda^2 k/10^5}\) for all \(k \geq \lambda^{-2} 10^5\).

**Proof.** Let \(C\) be such that \(\frac{8}{\lambda} \leq C \leq \frac{\lambda k}{32 \log(\lambda k)}\) to be chosen (these inequalities are so that the bounds in the hypotheses of the above lemmas are satisfied). We consider the event \(A\) that consists of the following events

\(A_1\). The root \(\rho\) sends infection to at least \(C \log(\lambda k)\) vertices before time 1/2 who remain infected until time 1,
\(A_2\). At least \(\frac{\lambda k}{100}\) vertices are infected at time 2,
\(A_h\). At least \(\frac{\lambda k}{100}\) vertices are infected at time \(h\) for \(h = 3, 4, \ldots, \lceil e^{\lambda^2 k/10^5} \rceil\).

Under this event, it is clear that the contact process survives until time \(e^{\lambda^2 k/10^5}\). By Lemma 3.1, \(A_1\) happens with probability at least \(1 - \frac{16C \log(\lambda k)}{\lambda k} - 4(\lambda k)^{-C/4}\).

By Lemma 3.3, conditioned on \(A_1\), the event \(A_2\) happens with probability at least

\[1 - 3 \exp(-\lambda C \log(\lambda k)/128) = 1 - 3(\lambda k)^{-\lambda C/128}.\]

Similarly, by Lemma 3.3, each \(h \geq 3\), conditioned on \(A_{h-1}\), the event \(A_h\) happens with probability at least

\[1 - 3 \exp(-\lambda^2 k/(128 \cdot 100)).\]

By the union bound, we obtain

\[1 - \Pr(A) \leq \frac{16C \log(\lambda k)}{\lambda k} + 4(\lambda k)^{-C/4} + 3(\lambda k)^{-\lambda C/128} + 3 \exp(-\lambda^2 k/(128 \cdot 100)) \times \exp(\lambda^2 k/10^5)\]
\[\leq \frac{16C \log(\lambda k)}{\lambda k} + 7(\lambda k)^{-C/4} + 3 \exp(-\lambda^2 k/10^5)\]
\[\leq \frac{23C \log(\lambda k)}{\lambda k} + 3 \exp(-\lambda^2 k/10^5)\] since \(C \geq 4\)
\[\leq \frac{200 \log(\lambda k)}{\lambda^2 k} + 3 \exp(-\lambda^2 k/10^5)\] by choosing \(C = 8/\lambda\).

This completes the proof. \(\square\)

### 4. Proof of the lower bound: Theorem 1.2

Our strategy is to show that in \(G_n\), there exists a long path of high degree vertices which we think about as a long path on which each vertex is attached to a big star (particularly, an \((A - 2)\)-star). For each of these vertices, Lemma 3.4 shows that the contact process on their attached star survives for a long time. So, each infected star has ample time to pass the infection to other stars on the path before being healed. Theorem 1.2 thus follows immediately from the following propositions.
Proposition 4.1. [Existence of long path] Let $\mu$ be any probability measure satisfying $\mathbb{E}_{D \sim \mu} D(D-2) > 0$. Let $G \sim G(n, \mu)$. Let $A > 16 \mathbb{E}_{D \sim \mu} D$ and

$$L = \frac{0.9 \log n}{-\log \mu[A, \infty]}.$$  

With high probability, there exists a path in $G_n$ of length $L$ in which every vertex has at least $A-2$ neighbors outside of the path and these $L(A-2)$ neighbors are all distinct.

Let $P$ be the union of a path as in Proposition 4.1 together with the $(A-2)$-stars attached to vertices on the path. From now on, we simply refer to $P$ as the long path.

Proposition 4.2. [Survival time on the long path] Let $A = \frac{10^6 \lambda}{\lambda^2} \log \lambda$. The contact process on $P$ starting with all vertices infected survives until time $1.01^L \geq n^{-c/\log \mu[A, \infty]}$ with high probability as $L \to \infty$, where $c = 1/120$.

In the rest of this section, we prove these two propositions.

4.1. Existence of a long path. To prove Proposition 4.1, we shall use the cut-off line algorithm introduced in [13] to find a uniform perfect matching of the half-edges in $G_n$.

![Diagram](image-url)

**Figure 2.** The circles ‘o’ represent matched half-edges and the crosses ‘×’ represent unmatched half-edges. The blue half-edge is chosen and matched to the red half-edge which is the highest unmatched half-edge. Then the cut-off line is moved to the new cut-off line (the dashed line).

**Definition 4.3 (Cut-off line algorithm).** A perfect matching of the half-edges of $G_n$ is obtained through the following algorithm (see Figure 2).

- Each half-edge of a vertex $v$ is assigned with a height uniformly chosen in $[0, 1]$ and is placed on the line of vertex $v$.
- The cut-off line is initially set at height 1.
- Pick an unmatched half-edge independent of the heights of all unmatched half-edges and match it to the highest unmatched half-edge. Move the cut-off line to the height of the latter half-edge.
**Proof of Proposition 3.1.** We will use the cut-off line algorithm to explore the edges of $G_n$. Firstly, the degrees of vertices in $G_n$ are sampled and each half-edge of these vertices is assigned a height uniformly chosen in $[0, 1]$ and is placed on the line of the corresponding vertex $v$ as described in the cut-off line algorithm. We shall repeat the following greedy algorithm to find a long path of length $L$. Let $U$ denote the set of visited vertices. We start with $U = \emptyset$.

Let $\varepsilon = 0.05$. We shall perform the algorithm in $n^{1-\varepsilon}$ rounds and show that at least one of them will successfully find a long path. In the $k$-th round where $k = 1, 2, \ldots, n^{1-\varepsilon}$, pick a vertex $v_{k,1} \notin U$ of degree at least $A$ uniformly at random to be the first vertex on the long path for this round. Put $v_{k,1}$ into $U$ and pick $A$ half-edges of $v_{k,1}$ uniformly at random and match them with the highest unmatched half-edges one by one. Put these $A$ newly found neighbors of $v_{k,1}$ into $U$. If any of these visited vertices coincide or have already been in $U$, the round fails and we move to the next round. Otherwise, let $v_{k,2}$ be chosen uniformly at random among these $A$ neighbors of $v_{k,1}$. If $\deg(v_{k,2}) < A$, the round fails and we move to the next round. Otherwise, $v_{k,2}$ will be the second vertex on the long path for this round. We repeat the above process to find $v_{k,3}, \ldots, v_{k,L}$. So, either the round fails during the process or we have found the path $v_{k,1}, \ldots, v_{k,L}$ each of which has $A - 2$ neighbors all distinct.

We now show that each round succeeds with a decent probability. To this end, we first provide some elementary bounds on the number of edges of $G_n$. Let $\mu$ be the set of vertices in $G_n$. For any $v \in \mu$, the number of half-edges is at most $2\mu_v$ and so during the algorithm, the number of matched half-edges is at most $2\mu = 2\mu_v$. During each round, the number of visited vertices is at most $2AL$ and so during the algorithm, the number of matched half-edges is at most $2n^{1-\varepsilon}AL \leq nd/4$ for sufficiently large $n$.

Conditioned on these events, we now provide a lower bound for the probability that a round fails. Let $W$ be the set of vertices in $G_n$ with degree at least $A$. With high probability, $|W| \geq \frac{n\mu(A, \infty)}{2}$ and so its number of half-edges is at least $\frac{n\mu(A, \infty)}{2}$. For any $i \in [L]$ and $k \in [n^{1-\varepsilon}]$, the probability that $v_{k,i} \in W$ is the same as the probability that the highest unmatched half-edge belongs to a vertex in $W$. This happens with probability $\geq \frac{\mu(A, \infty)}{2\mu} = \frac{\mu(A, \infty)}{8d}$ because by [7] the total number of unmatched half-edges is at most $4nd$. 




The total number of half-edges that the vertices visited during the algorithm is \(\leq \Delta n^{1-\epsilon} AL + \sum_{v \in G_n} X_v \leq \Delta n^{1-\epsilon} AL + \delta n/2\). And so, at any stage of the algorithm, the probability that a vertex is revisited is \(\leq \frac{\Delta n^{1-\epsilon} AL + \delta n/2}{nd/4}\) where the denominator bounds from below the number of unmatched half-edges. Thus, for each \(i \in [L]\) and \(k \in [n^{1-\epsilon}]\), the probability that the exploration at \(v_{k,i}\) succeeds is at least \(\frac{\mu(A,\infty) - A \Delta n^{1-\epsilon} AL + \delta n/2}{16d}\) for sufficiently large \(n\). Here we recall that \(\delta\) is a sufficiently small constant depending on \(A\) and \(\mu\).

Thus, the probability that a round \(k \in [n^{1-\epsilon}]\) succeeds is at least \(\frac{A^L(\mu(A,\infty))^{L}}{16d^{1}\epsilon}\) \(\geq (\mu(A,\infty))^{L} = n^{-0.9} = n^{1-2\epsilon}\) where we used \(A \geq 16d\). And so with high probability, that at least one round succeeds. Each successful round gives a desired long path. \(\square\)

4.2. **Long survival on long paths.** To prove Proposition 4.2, we couple the contact process on \(P\) with a percolation process (Lemma 4.4) and then show that this percolation survives for a long time (Lemma 4.5).

Recall the definition of \(A\) in (2) that \(A = \frac{10^6}{\lambda^2} \log \frac{1}{\lambda}\). By Lemma 3.4, the contact process on an \((A-2)\)-star with the root initially infected survives until time \(a^2\) with

\[
\omega = \lfloor e^{\lambda^2/(2d^{10^6})} \rfloor
\]

with probability at least \(1 - 200 \log(\lambda(A-2)/2) - 3 e^{-\lambda^2/(2d^{10^6})} \geq 0.99\) because \(\lambda < 1/20\). This allows us to couple the contact process on \(P\) with a percolation process as follows.

**Lemma 4.4.** [Coupling with Percolation] Consider the following percolation on \([L]\times[0,1,\ldots)\): edges connecting vertices \((i,0)\) and \((i+1,0)\) are open for all \(i\). If \((i,t)\) belongs to an open edge then with probability at least 0.99, the edges connecting \((i,t)\) and \((i,(t+1)), (i+1,(t+1))\) are also open. These events occur independently. There exists a coupling of the contact process on \(P\) with all vertices initially infected and the percolation such that if \((i,t)\) belongs to an open edge in the percolation then there are at least \(\lambda(A-2)/100\) vertices in the \(i\)-th star of \(P\) that are infected at time \(ta\) where \(a\) is as in (10).

**Proof of Lemma 4.4.** For each \(i \in [L]\), let \(A_i^t\) be the event that there are at least \(\lambda(A-2)/100\) vertices in the \(i\)-th star of \(P\) that are infected at time \(ta\). It suffices to show that

\[
\mathbb{P}\left( A_i^{t+1} \cap A_{i-1}^{t+1} \cap A_{i+1}^{t+1} | A_i^t \right) \geq 0.99 =: 1 - 16\epsilon.
\]

Let \(B_i^t\) be the event that there are at least \(\lambda(A-2)/100\) vertices in the \(i\)-th star of \(P\) infected at time \(ta + n\) for all positive integers \(n \leq a\). Applying Lemma 3.3 \(a\) times, we get

\[
\mathbb{P}\left( A_i^{t+1} | A_i^t \right) \geq \mathbb{P}\left( B_i^t | A_i^t \right) \geq 1 - 3a \exp(-\lambda^2(A-2)/12800) \geq 1 - \epsilon.
\]

Note that each vertex infected at certain time \(s\) has a probability of \(\mathbb{P}(\text{Pois}(1/2) = 0) \geq 1/e\) to remain infected until time \(s+1/2\). Thus, under \(B_i^t\), by Chernoff inequality, for all positive integers \(n \leq a\), at least \(M = 0.01\lambda(A-2)/4\) vertices infected at time \(ta + n\) remain infected until time \(ta + n + 1/2\) with probability at least \(1 - a \exp(-\lambda(A-2)/10^4) \geq 1 - \epsilon\). Under this event, applying Lemma 3.2 \(a\) times for each of the time intervals \([ta + n, ta + n + 1/2]\), the root \(v_1\) of the \(i\)-th star is infected for more than \(a/4\) time in the time interval \([ta, ta + a]\) with probability at least

\[
1 - 2a \exp(-\lambda M/32) = 1 - 2a \exp(-\lambda^2(A-2)/12800) \geq 1 - \epsilon.
\]
Thus, the probability that $v_i$ infects $v_{i+1}$ during $[ta, ta + a]$ is at least
\[ P(\text{Pois}(\lambda a/4) \neq 0) = 1 - \exp(-\lambda a/4) \geq 1 - \varepsilon \]
since $a \geq 40/\lambda$.

Once $v_{i+1}$ is infected, in proving Lemma 3.4, we proved that $A_{i+1}^{t+1}$ occurs with probability at least
\[ 1 - \frac{200 \log(\lambda(A-2))}{\lambda^2(A-2)} - 3e^{-\lambda^2(A-2)/10^5} \geq 1 - \varepsilon. \]

Combining all of these events by the union bound, we obtain
\[ P\left( A_{i+1}^{t+1} \Big| A_t^i \right) \geq 1 - 5\varepsilon. \]

Similarly for $A_{i-1}^{t+1}$. By the union bound, we obtain (11). □

Lemma 4.5. [Long survival of percolation] With high probability, the percolation reaches time $1.01L$. In other words, there exists a path of open edges connecting some vertex $(i,0)$ and some vertex $(j,1.01L)$ for some $i,j \in [L]$.

Proof of Lemma 4.5. We use the Peierls argument. We say that a vertex $(i,t) \in [L] \times \mathbb{N}$ is occupied if there exists a path of open edges connecting some vertex $(j,0)$ and $(i,t)$. Observe that the set $C$ of occupied vertices is connected. Consider the subset $C' = C + [-1/2, 1/2]^2$ which is formed by surrounding each vertex in $C$ by the box $[-1/2, 1/2]^2$. Let $B$ be the outer boundary of $C'$ inside $[L] \times \mathbb{N}$.

We observe the following.

1. $B$ is a path in the dual lattice of $\mathbb{Z}^2$. In other words, the edges of $B$ are edges of $(\mathbb{Z} + 1/2)^2$.
2. If the percolation does not reach time $1.01L$ then $B$ lies entirely in $[L] \times [0, 1.01L]$.

Let $\alpha$ be the length of $B$. Each edge in $B$ is dual to a closed edge in the percolation; however, these dual edges may not be distinct. So, to bound the probability of having a path $B$ of length $\alpha$, we need to find a lower bound for the number of distinct closed edges. Walking from the left-most vertex on $B$ gives an orientation to each of the edges in $B$. Let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be the number of edges in $B$ that are in the leftward, rightward, upward, and downward directions, respectively. We have $\alpha = \sum_{i=1}^{4} \alpha_i$.

We further decompose $\alpha_3$ as follows. Let $\alpha_{3,1}, \alpha_{3,2}, \alpha_{3,3}$ be the number of upward edges that follow an edge that is leftward, rightward, upward, respectively. We have
\[ \alpha_3 = \alpha_{3,1} + \alpha_{3,2} + \alpha_{3,3}. \]

Similarly, let $\alpha_{4,1}, \alpha_{4,2}, \alpha_{4,4}$ be the number of up edges that follow a left, right, down edge, respectively. We have
\[ \alpha_4 = \alpha_{4,1} + \alpha_{4,2} + \alpha_{4,4}. \]

Observe that

1. Edges counted in $\alpha_1, \alpha_{3,3}$ and $\alpha_{4,4}$ correspond to distinct closed edges uniquely determined by $B$. 


(2) $\alpha_2 \leq \alpha_1$ because the directed path $B$ goes from its left-most vertex to its right-most one.
(3) $\alpha_3, 1 + \alpha_4, 1 \leq \alpha_1$
(4) $\alpha_3, 2 + \alpha_4, 2 \leq \alpha_2 \leq \alpha_1$.

Thus,

$$\alpha = (\alpha_1 + \alpha_3, 3 + \alpha_4, 4) + \alpha_2 + \alpha_3, 1 + \alpha_3, 2 + \alpha_4, 1 + \alpha_4, 2 \leq (\alpha_1 + \alpha_3, 3 + \alpha_4, 4) + 3\alpha_1 \leq 4(\alpha_1 + \alpha_3, 3 + \alpha_4, 4).$$

Therefore, the probability that a given path $B$ occurs is at most $0.01^{\alpha/4}$. Given $\alpha$, the number of paths $B$ of length $\alpha$ is at most $1.01^L.3^\alpha$ where the first factor is the number of ways to choose the left-most vertex and at any point on the path, there are at most 3 ways to determine the next direction. By the union bound, the probability that the percolation does not reach time $1.01^L$ is at most

$$\sum_{\alpha=0}^{\infty} 1.01^L.3^\alpha.0.01^{\alpha/4} \ll 1.01^L.0.95^L = o(1)$$

completing the proof.

5. Survival time on trees

To prove Theorem 1.3, we first derive an upper bound on the survival time on trees. Following the methods in [2] and [26], we can only get an upper bound of $n^{\lambda_1+o(1)}$ rather than $n^{\lambda_2+o(1)}$. The key new idea here is that as $\lambda$ is small compared to the degree of a typical vertex, a significant portion of infections are weak in the sense the vertex infected recovers before spreading the infection. Therefore, we come up with a modified process that focuses on strong infections and this allows us to get the power $\lambda_2$ in the upper bound.

5.1. The coupling with the starred contact process. In this section, we introduce the starred contact process which takes into account whether a new infection at a vertex $u$ is strong. If we ignore the weak infections, the contact process basically remains the same except that at any time, there could be vertices that are infected and then recover at a later time without passing the infection to any other vertices.

To be specific, let us sample the Poisson clocks for recoveries and infections on $G_n$. Using these clocks, the contact process $(X_t)_{t\geq0}$ with $X_0 = S \subset V$ is read from the graphical representation (See Section 2.1). We now construct the following Markov process $(X_t^*)_t_{t\geq0}$ which we call the starred contact process from the same set of clocks.

At time 0, $X_0^*(u) = 0$ if one of the following hold

- $u \notin S$,
- $u \in S$ and the recovery clock at $u$ rings before any infection clocks from $u$ to its neighbors.

In other words, $u$ recovers before sending any infection out.

Otherwise, we say that $u$ is strongly infected and set $X_0^*(u) = 1^*$. All vertices remain in their states until any of the clocks ring, say at time $t$. Consider the following cases depending on the nature of the clock.
• If the clock is a recovery at a vertex $u$ then set $X^*_t(u) = 0$.
• If the clock is an infection from a healthy vertex $v$ (that is $X^*_t(v) = 0$) to another vertex $u$, nothing changes.
• If the clock is an infection from an infected vertex $v$ (that is $X^*_t(v)$ equals 1 or 1*) to another infected vertex $u$, we set $X^*_t(u) = 1$.
• If the clock is an infection from an infected vertex $v$ (that is $X^*_t(v)$ equals 1 or 1*) to a healthy vertex $u$, then if the first clock after time $t$ at $u$ is a recovery clock rather than an infection clock from $u$ to any of its neighbors, then $u$ remains healthy at time $t$. Otherwise, set $X^*_t(u) = 1$. In either cases, set $X^*_t(v) = 1$.

In any cases, the state at other vertices remain the same. We observe that for any time $t$ and any vertex $u$, if $X_t(u) = 1$ and $X^*_t(u) = 0$ then in the process $(X_s)$, after this time, $u$ will recover before infecting any other vertices. Denote by $R$ and $R^*$ the survival time of the processes $(X_t)$ and $(X^*_t)$, respectively. The above observation implies that

$$R \leq R^* + \max_{i \leq n} \tau_i$$

where $\tau_1, \ldots, \tau_n$ are iid exp(1) random variables. The sum $\sum_{i=1}^n \tau_i$ bounds the total healing time of all vertices $u$ with $X_{R^*}(u) = 1$.

Since whp, $\max_{i \leq n} \tau_i \tau_i = O(\log n)$, we conclude that for all $G_n$, whp,

$$R \leq R^* + O(\log n).$$

Thus, it suffices to bound $R^*$. For the rest of this paper, we shall only work on the starred contact process and so, for simplicity, we write $(X_t)$ for the starred contact process from now on.

5.2. Recursive Survival Time on Trees. Before analyzing the process on graphs, we study it in the simpler case of trees. Then in passing from trees to graphs we will utilizes the fact that Galton-Watson trees are local weak limits of random graphs as discussed in Section 2.2.

Let $T$ be a (deterministic) tree and consider the starred process on $T$ with $\rho$ initially at $1^*$. Let $R(T)$ be the survival time of this process. Let $R(T) = E_{cr} R(T)$.

Let $T^+$ be the tree $T$ together with an added vertex $\rho^+$ which is the parent of $\rho$. Consider the starred contact process on a tree $T^+$ with $\rho^+$ always infected and $\rho$ initially at state $1^*$ which means that $\rho$ will first send an infection to either $\rho^+$ or its children. Let $S(T)$ be the first time that the process returns to the all healthy state on $T$ and $S(T) = E_{cr} S(T)$ be the expected excursion time. The addition of $\rho^+$ makes sure that the contact process on $T^+$ never dies out and so the process no longer has an absorbing state.

Note that if $\rho$ has at least 1 child, then with probability at least $\frac{1}{2}$, the first infection from $\rho$ is sent to its children rather than to $\rho^+$. Thus,

$$R(T) \leq 2S(T).$$

Let $D$ be the number of children of $\rho$ and $d_i$ be the number of children of the $i$-th child of $\rho$. Let $T_i$ be the subtree of $T$ containing $\rho_i$ and all of its successors.
An infection from \(\rho\) to \(\rho_i\) is strong if \(\rho_i\) sends an infection to either \(\rho\) or one of its children before recovering. This happens with probability \(\frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)}\). We call this number the weight of \(T_i\). That is
\[
\text{(14)} \quad w(T_i) = \frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)}.
\]
Similarly,
\[
\text{(15)} \quad w(T) = \frac{\lambda(D + 1)}{1 + \lambda(D + 1)}.
\]

The following recursion allows us to control the growth of \(S(T)\).

**Lemma 5.1** (Recursion for \(S\)). We have
\[
S(T) \leq \frac{1}{1 + \lambda(D + 1)} - \frac{1}{\lambda(D + 1)} + w(T)^{-1} \prod_{i=1}^{D} (1 + \lambda w(T_i)) S(T_i).
\]

**Proof.** Let \((X_t)\) be the starred contact process on \(T^+\) with \(\rho^+\) always infected and \(\rho\) initially at \(1^*\). We consider the process \((X_t^\#)\) on \(T^+\) that also has \(\rho^+\) always infected, also starts at \(\rho\) being \(1^*\) and generated by the same recovery and infection clocks as \((X_t)\) except that we ignore any recoveries of \((X_t)\) at \(\rho\) if at least one of its successors is still infected at the time. Let \(S^\#(T)\) be the expected excursion time of \((X_t^\#)\) that is the expected time difference from the time \(\rho\) is strongly infected to the first time that all vertices in \(T\) recover. Since \(S(T) \leq S^\#(T)\), it suffices to show that
\[
\text{(16)} \quad S^\#(T) = \frac{1}{1 + \lambda(D + 1)} - \frac{1}{\lambda(D + 1)} + w(T)^{-1} \prod_{i=1}^{D} (1 + \lambda w(T_i)) S(T_i).
\]
Let \(t_0\) be the first time that an event at \(\rho\) happens, which could either be a recovery at \(\rho\) or an infection from \(\rho\). We have
\[
\mathbb{E}t_0 = \frac{1}{1 + \lambda(D + 1)}.
\]
Since \(\rho\) is at \(1^*\) initially, this first event has to be an infection from \(\rho\) to a child \(\rho_i\) or to \(\rho^+\). Note that conditioning on the event that \(\rho\) is \(1^*\) at time 0 does not change the expectation of \(t_0\).

After time \(t_0\), if \(\rho\) infects \(\rho_i\), the healing clock at \(\rho\) is deactivated, the infection is spread across \(T\) until all of \(T \setminus \{\rho\}\) recovers. Denote by \(t_1\) the first time that this happens. If at time \(t_0\), \(\rho\) infects \(\rho^+\) then \(t_1 = t_0\).

Let \((X_t^\otimes)\) be the starred contact process on \(T\) with \(\rho\) always infected and one of its children initially at \(1^*\). Let \(S^\otimes\) be the expected excursion time. Let
\[
\text{(17)} \quad \theta = \frac{1}{D} \sum_{i=1}^{D} \mathbb{P}(\text{the infection } \rho \rightarrow t_0 \rho_i \text{ is a strong infection} \mid \rho \rightarrow t_0 \rho_i) = \frac{1}{D} \sum_{i=1}^{D} w(T_i).
\]
be the probability that an infection from \(\rho\) to its children is strong.

We observe that
\[
\text{(18)} \quad \mathbb{E}(t_1 - t_0 \mid \rho \text{ send an infection to a child at time } t_0) = \theta S^\otimes.
\]
Using (18), we shall now prove that

\[ S^\sharp(T) = \mathbb{E}t_0 + \mathbb{E}(t_1 - t_0) + \sum_{k=0}^{\infty} \left( \frac{\lambda D}{1 + \lambda D} \right)^k \left( \frac{1}{1 + \lambda D} \right) \left( \frac{k + 1}{1 + \lambda D} + k\theta S^\otimes \right). \]

To see (19), at time \( t_1 \), only \( \rho \) and \( \rho^+ \) remain infected. The next meaningful event would either be a recovery at \( \rho \) or an infection from \( \rho \) to a child \( \rho_i \). The former corresponds to \( k = 0 \). If the latter happens instead, say at a time \( t_2 \), then the infection will spread across \( T \) again until everything in \( T \setminus \{\rho\} \) is healed. Now, another attempt to heal \( \rho \) happens again with probability of success \( \frac{1}{1 + \lambda D} \). If success, we set \( k = 1 \). Otherwise, the whole process repeats and \( k \) is correspondingly the number of fail attempts to heal \( \rho \). The probability of \( k \) is \( \left( \frac{\lambda D}{1 + \lambda D} \right)^k \left( \frac{1}{1 + \lambda D} \right) \) and the time spent is \( \frac{k + 1}{1 + \lambda D} + k\theta S^\otimes \) according to (18).

By algebraic manipulation, we obtain

\[ S^\sharp(T) = \frac{1}{1 + \lambda(D + 1)} + \frac{D}{D + 1} \theta S^\otimes + \sum_{k=0}^{\infty} \left( \frac{\lambda D}{1 + \lambda D} \right)^k \left( \frac{1}{1 + \lambda D} \right) \left( \frac{k + 1}{1 + \lambda D} + k\theta S^\otimes \right) \]

\[ = \frac{1}{1 + \lambda(D + 1)} + \frac{D}{D + 1} \theta S^\otimes + 1 + \lambda D\theta S^\otimes \]

\[ = \frac{1}{1 + \lambda(D + 1)} - \frac{1}{\lambda(D + 1)} + \frac{1 + \lambda(D + 1)}{\lambda(D + 1)} (1 + \lambda D\theta S^\otimes). \]

We now evaluate \( S^\otimes \) using the observation that \( (X^\otimes_i) \), as \( \rho \) is always infected, is just the product of the root-added contact processes on the \( T_i \)'s, where \( \rho \) plays the role of \( \rho_i^+ \) – the added-root. Let \( \pi_i \) be the stationary distribution of the root-added starred contact process on \( T_i \). Then, \( \pi_i(0) \) is the fraction of time that \( T_i \) is all healthy and so

\[ \pi_i(0) = \frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)} \left( \frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)} \right)^{-1} \]

\[ = \frac{1}{1 + \lambda w(T_i) S(T_i)} \]

where \( \lambda \frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)} \) is the rate at which \( \rho \) sends a strong infection to \( \rho_i \).

Similarly, let \( \pi \) be the stationary distribution of \( (X^\otimes) \). Then, \( \pi(0) \), the fraction of time that \( T \setminus \{\rho\} \) is all healthy, is

\[ \pi(0) = \frac{\lambda D\theta}{\lambda D\theta + S^\otimes} = \frac{1}{1 + \lambda D\theta S^\otimes} \]

where \( \lambda D\theta \) is the rate at which \( \rho \) sends a strong infection to one of the \( \rho_i \).

Since \( \pi = \prod_{i=1}^{D} \pi_i \), we have

\[ \frac{1}{1 + \lambda D\theta S^\otimes} = \prod_{i=1}^{D} \frac{1}{1 + \lambda w(T_i) S(T_i)} \]

and so,

\[ 1 + \lambda D\theta S^\otimes = \prod_{i=1}^{D} (1 + \lambda w(T_i) S(T_i)). \]
Combining this with (20), we get (16). □

5.3. Recursion for the total number of infections at the leaves. To pass from Galton-Watson trees to random graphs, it is crucial to control the effect of cycles as they allow the infection to transfer from one local neighborhood to another. To this end, one needs to control the number of times the infection hits the boundary of a local neighborhood. Therefore, following [26], in this section, we study the total infections at leaves of (finite) trees defined as follows.

**Definition 5.2 (Total infections at leaves).** Let $T$ be a finite tree rooted at $\rho$, set

(22) \[ l = \max \{ \text{dist}(\rho, v) : v \in T \}. \]

be the depth of the tree and

\[ \mathcal{L} := \{ v \in T : \text{dist}(\rho, v) = l \} \]

be the collection of depth-$l$ leaves of $T$. Suppose that $l \geq 1$ and consider the root-added, starred contact process $(X_t)$ on $T^+ = T \cup \{\rho^+\}$ with $\rho^+$ always infected and $\rho$ initially at state $1^*$. For $v \in \mathcal{L}$, define the total infections at $v$ by

\[ M_{l,v}(T) := \text{the number of infections at } v \text{ in } (X_t) \text{ during time } t \in [0, S(T)], \]

where $S(T)$ is the first time that $(X_t)$ returns to the all healthy state on $T$. In other words, we count the number of times $t$ such that $X_t(v) = 1^*$ and $X_{t-}(v) = 0$ for $t \leq S(T)$. Then, we define the total infections at depth-$l$ leaves (during a single excursion) by

\[ M_l(T) = \sum_{v \in \mathcal{L}} M_{l,v}(T). \]

For $l' > l$, we set $M_{l'}(T) \equiv 0$.

We also denote the expected total infections at depth-$l$ leaves by $M_l(T) = \mathbb{E}_{cp} M_l(T)$. Also, as above, we write $M_{l'}(T) = 0$ for $l' > l$. Moreover, if the tree depth is 0 (that is, $T$ consists of a single vertex), we set $M_0(T) \equiv 1$.

Let $(T, \rho)$ be a finite rooted tree of depth $l$. The goal of this section is to establish the following recursive bound.

**Lemma 5.3.** For a finite rooted tree $(T, \rho)$ of depth $l$, let $D = \text{deg}(\rho)$ and $T_1, \ldots, T_D$ be the subtrees rooted at the children $v_1, \ldots, v_D$ of $\rho$. Then, $M_l(T)$, the expected total infections at depth-$l$ leaves on $T$, satisfies the following.

(23) \[ w(T) M_l(T) \leq \lambda \sum_{i=1}^{D} w(T_i) M_{l-1}(T_i) \prod_{1 \leq j \leq D, j \neq i} (1 + \lambda w(T_j) S(T_j)) \]

where we recall that

(24) \[ w(T_i) = \frac{\lambda(d_i + 1)}{1 + \lambda(d_i + 1)} \quad \text{and} \quad w(T) = \frac{\lambda(D + 1)}{1 + \lambda(D + 1)}. \]

**Proof of Lemma 5.3.** We shall use the same notations as in the proof of Lemma 5.1. In particular, we consider the following processes on $T$:

- $(X_t)$ the root-added process with $\rho$ initially at state $1^*$,
Thus, the identity in (26) is also equal to $M$.

Now we control $w$ where (25) $M$ and the end of the excursion of $(X^\rho)$ that this is a strong infection to the $\rho$.

On the other hand, let $L$ during one excursion of $(X^\rho)$, $\rho$ first sends an infection to either $\rho^+$ or one of the $\rho_i$. The probability that this is a strong infection to the $\rho_i$ is $\frac{D}{D+1} \theta$ where $\theta$ is defined in (17). Between this first infection and the end of the excursion of $(X^\rho)$, the number of excursions of $(X^\rho)$ is a geometric random variable with “success” probability $\frac{1}{1+\lambda D \theta}$. This is because $\rho$ recovers at rate 1 and sends a strong infection to the $\rho_i$ with rate $\lambda D \theta$; and the excursion of $(X^\rho)$ ends precisely when $\rho$ recovers. So, the expected number of excursions of $(X^\rho)$ in one excursion of $(X^\rho)$ is $\left(\frac{D}{D+1} + \lambda \right) \theta$. So,

$$(25) \quad M^\rho(T) = \left(\frac{1}{D+1} + \lambda \right) D \theta M^\rho(T).$$

Now we control $M^\rho(T)$ in terms of $\{M_{l-1}(T_i)\}$. We observe that

$$\lim_{t \to \infty} \frac{1}{t} \sum_{v \in L} E_{^CP} \left[ \{ s \in [0,t] : X^\rho_s(v) = 1^* \text{ and } X^\rho_{s-}(v) = 0 \} \right]$$

$$= \lim_{t \to \infty} \frac{\text{the number of excursions of } (X^\rho) \text{ in } [0,t]}{t_0} \frac{M^\rho(T)}{M^\rho(T)} = \frac{M^\rho(T)}{(\lambda D \theta)^{-1} + S^\rho(T)}.$$

On the other hand, let $L_i = \{ v \in T_i : \text{dist}(v, \rho_i) = l - 1 \}$ for each $i \in [D]$ and $(X^{(i)})$ be the restriction of $(X^\rho)$ on $T_i$. Note that $(X^{(i)})$ is simply the root-added process on $T_i$ with $\rho$ playing the role of the added root.

$$\sum_{v \in L_i} \left[ \{ s \in [0,t] : X^\rho_s(v) = 1^* \text{ and } X^\rho_{s-}(v) = 0 \} \right]$$

$$= \sum_{i=1}^D \sum_{v \in L_i} \left[ \{ s \in [0,t] : X^{(i)}_s(v) = 1^* \text{ and } X^{(i)}_{s-}(v) = 0 \} \right].$$

Thus, the identity in (26) is also equal to

$$\sum_{i=1}^D \frac{M_{l-1}(T_i)}{(\lambda w(T_i))^{-1} + S(T_i)}$$

where $w(T_i)$, as defined in (14), is the probability that an infection from $\rho$ to $\rho_i$ is strong. Therefore,

$$M^\rho(T) = \frac{1}{\lambda D \theta} (1 + \lambda D \theta S^\rho(T)) \sum_{i=1}^D \lambda w(T_i) M_{l-1}(T_i) (1 + \lambda w(T_i) S(T_i))^{-1}$$

$$= \frac{1}{D \theta} \sum_{i=1}^D w(T_i) M_{l-1}(T_i) \prod_{1 \leq j \leq D \atop j \neq i} (1 + \lambda w(T_j) S(T_j)).$$
where in the second equality we used equation (21). Finally, combining this with (25) and $M^2(T) \geq M_1(T)$, we deduce the conclusion.

5.4. The tail bound lemma. In the previous sections, we have derived recursive bounds for the survival time $S(T)$ and the number of infections at the leaves $M_1(T)$. These quantities are random variables defined on the probability space of the Galton-Watson tree. In this section, we provide a key lemma that gives tail estimates for $S(T)$ and $M_1(T)$ using the recursions. Let $D \sim \mu$ where $\mu$ is a probability measure on $\mathbb{N}$ satisfying $\mathbb{E}e^{3D} < \infty$. In particular, the measure $\mu$ in Theorem 1.3 satisfies this condition. For the rest of the paper, let $A, B, c$ be any constants that satisfy $A \ll 1, c \ll B$. For definiteness, we set

$$A = \log \mathbb{E}e^{3D} + 2e, \quad c = \frac{1}{A}, \quad B = 24A.$$ 

We assume that the infection rate $\lambda$ is sufficiently small compared to $B^{-1}$. Let

$$q = \frac{c}{16\lambda^2 \log(\lambda^{-1})}$$

and

$$p = (B\lambda)^{-q} \exp(-c\lambda^{-2}/B).$$

**Definition 5.4.** A nonnegative random variable $X$ is said to have a good tail if

$$\mathbb{P}(X \geq t) \leq f(t) := \begin{cases} 
\exp(-ct\lambda^{-1}) & \text{for } A\lambda \leq t \leq \frac{1}{B\lambda} 
\end{cases} \begin{cases} 
pt^{-q}(\log(eBt\lambda))^{-2} & \text{for } \frac{1}{B\lambda} \leq t.
\end{cases}$$

A random variable $X$ is said to have a strong tail if

$$\mathbb{P}(X \geq t) \leq f(3t)/3 \quad \text{for all } t \geq A\lambda/3.$$ 

Since $f$ is deceasing, it is clear that a random variable with a strong tail also has a good tail. Moreover, if $X_1, X_2, X_3$ have strong tails then $X_1 + X_2 + X_3$ has a good tail because

$$\mathbb{P}(X_1 + X_2 + X_3 \geq t) \leq \mathbb{P}(X_1 \geq t/3) + \mathbb{P}(X_2 \geq t/3) + \mathbb{P}(X_3 \geq t/3).$$

A random variable with a good tail starts with an exponential tail and then a power law tail. The extra term $\log(eBt\lambda))^{-2}$ in the latter tail is merely a technical device (used in (43)) for us to prove the following recursive lemma which is essential for our subsequent analysis.

**Lemma 5.5.** [Recursive tail bound] Let $D \sim \mu$ where $\mu$ is a probability measure on $\mathbb{N}$ satisfying $\mathbb{E}e^{3D} < \infty$. There exists $\lambda_0(\mu) > 0$ such that the following holds. Let $X_i$ be independent, not necessarily identically distributed, nonnegative random variables with good tails. Then for $0 < \lambda < \lambda_0(\mu)$,

$$X = \prod_{i=1}^{D}(1 + \lambda X_i) - 1,$$

has a strong tail.

**Remark 5.6.** We remark that this lemma works for any $q$ satisfying

$$20 \log \frac{1}{\lambda} < q < \frac{c}{8\lambda^2 \log(\lambda^{-1})}.$$ 

Here, we set $q$ as in (28) for clarity.
Proof. For notational convenience, we let $\Lambda = \frac{1}{X}$. We split the proof into steps depending on the range of $t$ in $\mathbb{P}(X \geq t)$.

5.4.1. Small $t$: $\frac{4\Lambda}{3} \leq t \leq \frac{\Lambda}{3B}$. We need to show that

$$P(X \geq t) \leq \exp(-3ct\Lambda)/3 = f(3t)/3 =: p_t.$$  

Let $M = t\Lambda \geq A/3$ by (27). Since $\mathbb{E}e^{3D} < \infty$ by (3), we have

$$\mathbb{P}(D \geq M) \leq \mathbb{E}e^{3D} \exp(-3M) \leq \exp(-3ct\Lambda)/24 = p_t/8.$$  

So, it suffices to assume that $D \leq M$. The next lemma allows to use truncation on $X_i$.

**Lemma 5.7.** We have

$$\mathbb{P}
\left(\exists i \leq M : X_i \geq \frac{2\Lambda}{B}\right) \leq \exp(-3ct\Lambda)/24 = p_t/8.$$  

Proof. Since the $X_i$ have good tails, by the union bound, we have

$$\mathbb{P}
\left(\exists i \leq M : X_i \geq \frac{2\Lambda}{B}\right) \leq M2^{-q}\exp(-c\Lambda^2/B)\leq \Lambda^{2}2^{-q}\exp(-c\Lambda^2/B) \leq \frac{1}{24}\exp(-c\Lambda^2/B) \leq \exp(-3ct\Lambda)/24.$$  

where we used (27), (28) and $t \leq \frac{\Lambda}{3B}$.

Lemma 5.7 allows us to get the tail bound for the sum of $X_i$ as follows.

**Lemma 5.8.** We have

$$\mathbb{P}
\left(\sum_{i=1}^{M} X_i \geq (\log 2)\Lambda\right) \leq \exp(-3ct\Lambda)/12 = p_t/4.$$  

Proof. Let $X'_i = X_i1_{X_i \leq \frac{2\Lambda}{B}}$. By Lemma 5.7, it suffices to show that

$$\mathbb{P}
\left(\sum_{i=1}^{M} X'_i \geq (\log 2)\Lambda\right) \leq \exp(-(ct\Lambda)/24.$$  

Let $\xi_i = \Lambda X'_i \leq \frac{2\Lambda^2}{B}$ then

$$\mathbb{E}e^{c\xi_i/4} = 1 + \frac{c}{4} \int_{0}^{\frac{2\Lambda^2}{B}} e^{cx/4} \mathbb{P}(\xi_i \geq x)dx$$

$$\leq 1 + e + \frac{c}{4} \int_{\frac{2\Lambda^2}{B}}^{\Lambda^2} e^{cx/4} \mathbb{P}(\xi_i \geq x)dx \quad \text{since } cA \leq 4 \text{ by (27)}$$

$$\leq 1 + e + \frac{c}{4} \int_{A}^{\Lambda^2} e^{-cx/2}dx + \frac{c}{4} \int_{\frac{2\Lambda^2}{B}}^{\Lambda^2} e^{cx/4} e^{-c\Lambda^2/B}dx$$

$$\leq 1 + e + 2 + 1 \leq e^2.$$  

(36)
And so
\[
\Pr \left( \sum_{i=1}^{M} \xi_i \geq (\log 2)\Lambda^2 \right) \leq e^{-c(\log 2)\Lambda^2/4}e^{2M} = \exp \left( 2t\Lambda - c(\log 2)\Lambda^2/4 \right) \leq \exp(-3ct\Lambda)/24
\]
where we used (27) to see that
\[
2t\Lambda + 3ct\Lambda \leq \Lambda^2 (2 + c)/B \leq c(\log 2)\Lambda^2/8.
\]
\(\square\)

Combining the above lemmas, we get
\[
\Pr(X \geq t) \leq \frac{p_t}{2} + \Pr \left( \prod_{i=1}^{M} (1 + \lambda X_i) - 1 \geq t, \sum_{i=1}^{M} X_i \leq (\log 2)\Lambda, X_i \leq \frac{2\Lambda}{B} \text{ for all } i \right).
\]

When \(\sum_{i=1}^{M} X_i \leq (\log 2)\Lambda\), we have
\[
\prod_{i=1}^{M} (1 + \lambda X_i) \leq e^{\lambda \sum_{i=1}^{M} X_i} \leq 1 + 2\lambda \sum_{i=1}^{M} X_i
\]
where in the second inequality, we used that \(e^s \leq 1 + 2s\) for all \(0 \leq s \leq \log 2\).

Thus,
\[
\Pr(X \geq t) \leq \frac{p_t}{2} + \Pr \left( 2\lambda \sum_{i=1}^{M} X_i 1_{X_i \leq \frac{2\Lambda}{B}} \geq t \right).
\]

Let \(\xi_i = \Lambda X_i 1_{X_i \leq \frac{2\Lambda}{B}}\) then as in (36),
\[
\mathbb{E}e^{c\xi_i/4} \leq e^2.
\]

and so for \(\Lambda\) sufficiently large,
\[
\Pr \left( 2\sum_{i=1}^{M} X_i 1_{X_i \leq \frac{2\Lambda}{B}} \geq t\Lambda \right) \leq e^{-ct\Lambda^2/8}e^{2M} = \exp \left( 2t\Lambda - ct\Lambda^2/8 \right) \leq \exp(-3ct\Lambda)/6 = \frac{p_t}{2}.
\]

This proves (34).

### 5.4.2. Large \(t\): \(t > \frac{\Lambda}{3B}\)

We need to show that
\[
\Pr(X \geq t) \leq f(3t)/3 = p(3t)^{-q(\log(3eBt\lambda))^{-2}/3} =: p_t.
\]

Firstly, we note that for every \(r > 0\), by Markov’s inequality and the assumption (3) on the exponent moment of \(\mu\),
\[
\Pr(D \geq r) \leq \mathbb{E}e^{3D}e^{-3r}.
\]

Let \(Y_i = \log(1 + \lambda X_i)\). Then \(\log X \leq \sum_{i=1}^{D} Y_i\) and by the good tail of \(X_i\),
\[
\Pr(Y_i \geq s) \leq \begin{cases} 
\exp(-c(e^s - 1)\Lambda^2) & \text{if } \log(1 + A\Lambda^2) \leq s \leq \beta, \\
p\lambda^q(e^s - 1)^{-q} \log^{-2}(eB(e^s - 1)) & \text{if } s > \beta,
\end{cases}
\]
where \(\beta = \log \left( 1 + \frac{1}{B} \right)\).
Let $d = ED$, we get

$$P(X \geq t) = P(\log X \geq \log t) \leq P \left( \sum_{i=1}^{D} Y_i \geq \log t \right)$$

(41)

$$\leq P \left( \sum_{i=1}^{2d} Y_i \geq \log t \right) + \sum_{r=2d+1}^{\infty} P \left( \sum_{i=1}^{r} Y_i \geq \log t, D = r \right).$$

To handle $P(\sum_{i=1}^{r} Y_i \geq \log t)$, we decompose the sum $\sum Y_i$ into subsums in which $Y_i \leq \beta$, and $Y > \beta$ and denote the probability that each subsum is large as follows

$$p_{1,h,s} = \max_{1 \leq i_1 < \cdots < i_h \leq r} P \left( \sum_{j=1}^{h} Y_{i_j} \geq s, Y_{i_j} \leq \beta \right)$$

$$p_{2,h,s} = \max_{1 \leq i_1 < \cdots < i_h \leq r} P \left( \sum_{j=1}^{h} Y_{i_j} \geq s, Y_{i_j} > \beta \right).$$

Observe that if $\sum_{i=1}^{r} Y_i \geq \log t$ then by setting $s$ to be the integer part of the first subsum, the second subsum must be at least $\log t - s - 1$. Thus,

$$P \left( \sum_{i=1}^{r} Y_i \geq \log t \right) \leq \sum_{k=0}^{r} \binom{r}{k} \sum_{s \in \mathbb{N}, s \leq \log t} p_{1,k,s} p_{2,r-k,\log t - s - 1}.$$  

(42)

In the next two lemmas, we bound these probabilities.

**Lemma 5.9.** We have for every $k, s \in \mathbb{N}$,

$$p_{1,k,s} \leq 8^k e^{-\frac{1}{2} c A^2 s}.$$

**Proof.** Let $Y_i' = Y_i 1_{Y_i \leq \beta}$. Then

$$P \left( \sum_{i=1}^{k} Y_i \geq s, Y_i \leq \beta \right) \leq P \left( \sum_{i=1}^{k} Y_i' \geq s \right).$$

For a positive number $z$, we have $e^{z Y_i'} \leq 1 + e^{z Y_i'} 1_{Y_i \leq \beta} = 1 + (1 + \lambda X_i)^z 1_{X_i \leq A/B}$ and so

$$E e^{z Y_i'} \leq 1 + (1 + A \lambda^2)^z + \lambda z \int_{A \lambda}^{A/B} (1 + \lambda t)^z - 1 e^{t(X_i \geq t)} dt$$

$$\leq 1 + (1 + A \lambda^2)^z + \lambda z \int_{A \lambda}^{A/B} (1 + \lambda t)^z - 1 e^{-c A t} dt$$

$$\leq 1 + (1 + A \lambda^2)^z + \lambda z \int_{A \lambda}^{A/B} e^{-t(c A - \lambda(z-1))} dt.$$  

By choosing $z = \frac{1}{2} c A^2$, we get

$$E e^{z Y_i'} \leq 1 + (1 + A \lambda^2) c A^2/2 + \frac{1}{2} c A \int_{A \lambda}^{\infty} e^{-t c A^2 / 2} dt \leq 1 + 2 e^{c A / 2} \leq 8.$$  

where we used $c A \leq 1$. Thus, by Markov’s inequality,

$$P \left( \sum_{i=1}^{k} Y_i' \geq s \right) \leq e^{-zs} \prod_{i=1}^{k} E e^{z Y_i'} \leq 8^k e^{-\frac{1}{2} c A^2 s}.$$
giving the desired bound. □

**Lemma 5.10.** For every \( m \geq 1 \) and \( s' > 2\beta m \), we have

\[
P_{2,m,s'} \leq \left( \frac{8p^2q\lambda^q}{\beta^2+q} \right)^m e^{-q(s'-\beta m)}(s' - \beta m)^{-2}.
\]

**Proof.** We have

\[
P \left( \sum_{i=1}^{m} Y_i \geq s', Y_i > \beta \right) \leq \sum_{u \in \mathbb{N}^m : \sum_{i=1}^{m} u_i \geq (s'/\beta - m)} \prod_{i=1}^{m} \mathbb{P}(Y_i \in \beta(u_i, u_i + 1))
\]

\[
\leq \sum_{u} \prod_{i=1}^{m} \mathbb{P}(Y_i > \beta u_i).
\]

where the last sum runs over all \( m \)-tuples \((u_1, \ldots, u_m) \in \mathbb{N}^m \) with \( u_i \geq 1 \) and \( \sum u_i = s'/\beta - m \) (note that the condition \( s' > 2\beta m \) implies that such a tuple exists). By (40) (with \( s = \beta u_i \geq \beta \)), we have

\[
p_{2,m,s'} \leq p^m \lambda^m q \sum_{u} \prod_{i=1}^{m} (e^{\beta u_i} - 1)^{-q} \prod_{i=1}^{m} \log^{-2}(e\beta u_i - 1)).
\]

We observe since \( \beta \leq 1 \) we have that for all \( x \geq \beta \), we have \( e^x - 1 \geq \frac{x}{\beta+1} - \frac{\beta}{\beta+1} e^x \geq \frac{\beta}{\beta+1} e^x \). Thus,

\[
p_{2,m,s'} \leq p^m \lambda^m q \sum_{u} \prod_{i=1}^{m} \left( \frac{\beta}{2} e^{\beta u_i} \right)^{-q} \prod_{i=1}^{m} \log^{-2}(e\beta u_i / 2)
\]

\[
= p^m \lambda^m q \left( \frac{2}{\beta} \right)^{qm} \sum_{u} e^{-q\beta \sum_{i=1}^{m} u_i} \prod_{i=1}^{m} (\beta u_i + \tau)^{-2}
\]

\[
\leq p^m \lambda^m q \left( \frac{2}{\beta} \right)^{qm} e^{-q(s'-\beta m)} \sum_{u} \prod_{i=1}^{m} (\beta u_i + \tau)^{-2}
\]

where \( \tau = \log(e\beta / 2) \geq \log(2e/5) > \beta \) as \( B \geq 32 \).

To handle the sum-product, we use the following observation that holds for all integer \( w \geq 2 \):

\[
\sum_{v=1}^{w-1} (\tau + \beta v)^{-2}(\tau + \beta(w - v))^{-2} \leq \frac{2}{w^2} \sum_{v=1}^{w/2} (\tau + \beta v)^{-2}(\tau + \beta w/2)^{-2}
\]

\[
\leq \frac{8}{\beta^2(2\tau + \beta w)^2} \sum_{v=1}^{\infty} (v + \tau)^{-2} \leq \frac{8}{\beta \tau(\tau + \beta w)^2}.
\]

(43)

By induction over \( m \) we have that

\[
\sum_{u} \prod_{i=1}^{m} (\beta u_i + \tau)^{-2} = \sum_{u_1 + \cdots + u_m = s'/\beta - m} \prod_{i=1}^{m} (\beta u_i + \tau)^{-2} \leq \frac{8^m}{\beta^{m-1}} (\tau + s' - \beta m)^{-2}.
\]

Thus,

\[
p_{2,m,s'} \leq \left( \frac{8p^2q\lambda^q}{\beta^2+q} \right)^m e^{-q(s'-\beta m)}(s' - \beta m)^{-2} \leq \left( \frac{8p^2q\lambda^q}{\beta^2+q} \right)^m e^{-q(s'-\beta m)}(s' - \beta m)^{-2}.
\]

□

25
We now put everything together in (42). We control $\sum Y_i$ by a case analysis writing

$$S_1 = \{ s \in \mathbb{N} : s > (\log t) / 2 \},$$

$$S_2 = S_{2,k,r} = \{ s \in \mathbb{N} : s \leq (\log t) / 2, s' \leq 2(r-k) \beta \},$$

and

$$S_3 = S_{3,k,r} = \{ s \in \mathbb{N} : s \leq (\log t) / 2, s' > 2(r-k) \beta \}$$

where $s' = \log t - s - 1$, and setting

$$p \left( \sum_{i=1}^{r} Y_i \geq \log t \right) \leq \sum_{k=0}^{r} \left( \binom{r}{k} \right) \sum_{s \in S_1} p_{1,k,s} + \sum_{k=0}^{r-1} \left( \binom{r}{k} \right) \sum_{s \in S_2} p_{2,r-k,s'}$$

$$+ \sum_{k=0}^{r-1} \left( \binom{r}{k} \right) \sum_{s \in S_3} p_{1,k,s} p_{2,r-k,s'} =: T_1 + T_2 + T_3.$$

Note that in the sums defining $T_2$ and $T_3$, the index $k$ is at most $r-1$ because otherwise $s$ would have been $[\log t] > \log t / 2$. By Lemma 5.10 and $t \geq \Lambda / (3B)$, we have

$$T_1 \leq \sum_{k=0}^{r} \left( \binom{r}{k} \right) \sum_{s \in S_1} 8^k e^{-\frac{1}{2}c\Lambda^2 s} \leq 2^r 8^r e^{-\frac{1}{4}c\Lambda^2 \log t}$$

$$\leq 16^r p(3t)^{-q} (\log (3eBt\lambda))^{-2} (9\Lambda) = 16^r p_t / (3\Lambda).$$

The contribution from $S_2$ is bounded by

$$T_2 \leq \sum_{k=0}^{r-1} \left( \binom{r}{k} \right) \sum_{s \in S_2} \max_{1 \leq i_1 < \cdots < i_{r-k} \leq r} \prod_{j=1}^{r-k} p(Y_{i_j} > \beta)$$

$$\leq \sum_{k=0}^{r-1} \left( \binom{r}{k} \right) \sum_{s \in S_2} (p\lambda^q B^q)^{r-k} \text{ by (40)}.$$

We note that for $s \in S_2$,

$$2(r-k) \beta \geq s' = \log t - s - 1 \geq \log t / 3.$$

So,

$$r - k \geq \frac{\log t}{6\beta} \geq \frac{\log t}{3}.$$

Hence,

$$T_2 \leq 2^r |S_2| (p\lambda^q B^q)^{\log t / 3} \leq 2^r (\log t) (p\lambda^q B^q)^{\log t / 3} \leq 16^r p_t / (3\Lambda).$$

By Lemma 5.10 we obtain

$$T_3 \leq t^{-q} e^q \sum_{k=0}^{r-1} \left( \binom{r}{k} \right) 8^k \left( \frac{8p2^q\lambda^q e^q}{\beta^q + 2} \right)^{r-k} \sum_{s \in S_3} e^{-\frac{i}{2}c\Lambda^2 (s' - \beta(r-k))^2}$$

where we used $q \leq \frac{1}{2}c\Lambda^2$ to get $e^{-\frac{1}{2}c\Lambda^2 e^{-qs'}} \leq e^{-q(s+s')} = t^{-q} e^q$. By the definition of $S_3$, we have $s \leq \log t / 2$ and so $s' \geq \log t / 3$. Since $s' \geq 2(r-k) \beta$, we have $s' - (r-k) \beta \geq s'/2 \geq \log t / 6$. Thus,
$$\left(s' - \beta(r - k)\right)^{-2} \leq 36(\log(3eBt\lambda))^{-2}$$. So

$$T_3 \leq \frac{36t^{-q}(\log(3eBt\lambda))^{-2}e^q}{\beta} \sum_{k=0}^{r-1} \binom{r}{k} s^k \left(\frac{8p2^{q}\lambda^q e^{q\beta}}{\beta^{q+2}}\right)^{r-k} \sum_{s=0}^{\infty} e^{-\frac{1}{4} s^2 s}$$

$$\leq 72t^{-q}(\log(3eBt\lambda))^{-2}e^q 16^r \left(\frac{8p2^{q}\lambda^q e^{q\beta}}{\beta^{q+2}}\right) \leq 16^r p_t/(3\Lambda)$$.

where in the second inequality, we used that $$\frac{8p2^{q}\lambda^q e^{q\beta}}{\beta^{q+2}} \leq 1$$. So, in total,

$$\text{where in the second inequality, we used that } \frac{8p2^{q}\lambda^q e^{q\beta}}{\beta^{q+2}} \leq 1$$. So, in total,

$$\mathbb{P} \left( \sum_{i=1}^{r} Y_i \geq \log t \right) \leq 16^r p_t/\Lambda.$$}

Combining this with (39) and (41), we obtain

$$\frac{\mathbb{P}(X \geq t)}{p_t} \leq \frac{16^d}{\Lambda} + \sum_{r=2d+1}^{\infty} \mathbb{E}e^{3D}e^{-3r} \frac{16^r}{\Lambda} \leq 1$$

for sufficiently small $$\lambda$$. This completes the proof of Lemma 5.5.

### 5.5. Tail bounds for the excursion time and the total infections at the leaves.

Having proved the recursive Lemma 5.5, we are now ready to deduce tail estimates for $$S(T_{\leq l})$$ and $$M(T_{\leq l})$$ where $$T \sim \text{GW}(\mu)$$.

#### 5.5.1. Tail bounds for the excursion time.

**Lemma 5.11.** Let $$l \geq 0$$ be an integer. Let $$T \sim \text{GW}(\mu)$$ and $$S(T_{\leq l})$$ be the expected survival time on the truncated tree $$T_{\leq l}$$. Then $$w(T_{\leq l})S(T_{\leq l})$$ has a good tail where $$w(T_{\leq l})$$ is the weight defined in (15).

**Proof.** We prove by induction on $$l$$. The case that $$l = 0$$ is trivial because $$w(T_{\leq l})S(T_{\leq l}) = \frac{\lambda}{1+\lambda}$$. Assume that we have the statement for $$l - 1$$. Let $$X = w(T_{\leq l})S(T_{\leq l})$$ and $$X_i = w(T_i)S(T_i)$$ where $$T_i$$ is the subtree of $$T_{\leq l}$$ containing the $$i$$-th child of the root of $$T_{\leq l}$$ and its successors. By Lemma 5.1, we have

$$X \leq \frac{-1}{(1 + \lambda(D + 1))^2} + \prod_{i=1}^{D}(1 + \lambda X_i) \leq X_a + \prod_{i=1}^{D}(1 + \lambda X_i) - 1$$

where $$X_a = 2\lambda(D + 1) + \left[\frac{\lambda(D+1)}{1+\lambda(D+1)}\right]^2$$ contains the constant part.

By the induction hypothesis, $$X_i$$ has a good tail for all $$i$$. Thus, by Lemma 5.5, $$\prod_{i=1}^{D}(1 + \lambda X_i) - 1$$ has a strong tail. Since $$\mathbb{E}e^{3D} < \infty$$ by (3), one can see that both $$2\lambda(D + 1)$$ and $$\left[\frac{\lambda(D+1)}{1+\lambda(D+1)}\right]^2$$ have strong tails. Thus, the sum of those three has a good tail; so does $$X$$. This completes the proof of the lemma. \qed
5.5.2. Tail bounds for the total infections at the leaves.

**Lemma 5.12.** Let \( l \geq 0 \) be an integer. Let \( T \sim GW(\mu) \) and \( M_l(T_{\leq l}) \) be the expected total infections at depth-\( l \) leaves on \( T_{\leq l} \). Then \( (1.5)^l w(T_{\leq l}) M_l(T_{\leq l}) \) has a good tail.

**Proof of Lemma 5.12.** Let

\[
W(T_{\leq l}) = \max \left\{ (1.5)^l w(T_{\leq l}) M_l(T_{\leq l}), w(T_{\leq l}) S(T_{\leq l}) \right\}.
\]

We shall prove by induction that \( W(T) \) has a good tail. This will imply that \( (1.5)^l w(T_{\leq l}) M_l(T_{\leq l}) \) also has a good tail. For \( l = 0 \), we have \( W(T_{\leq 0}) = \frac{\lambda}{1+\lambda} \) which trivially has a good tail. Assume that the statement holds for \( l - 1 \).

Using the same notations as in the proof of Lemma 5.11, we have

\[
w(T_{\leq l}) S(T_{\leq l}) \leq X_a + \prod_{i=1}^D (1 + \lambda w(T_i) S(T_i)) - 1 \leq X_a + \prod_{i=1}^D (1 + 2\lambda W^{l-1}(T_i)) - 1.
\]

By Lemma 5.3

\[
(1.5)^l w(T) M(T) \leq 1.5\lambda \sum_{i=1}^D W^{l-1}(T_i) \prod_{1 \leq j \leq D \atop j \neq i} (1 + \lambda W^{l-1}(T_j)) \leq \frac{3}{2} \left[ \prod_{i=1}^D (1 + 2\lambda W^{l-1}(T_i)) - 1 \right].
\]

Thus,

\[
W(T) \leq X_a + \frac{3}{2} \left[ \prod_{i=1}^D (1 + 2\lambda W^{l-1}(T_i)) - 1 \right].
\]

By induction hypothesis, \( W^{l-1}(T_i) \) has a good tail for all \( i \). By Lemma 5.5 \( \prod_{i=1}^D (1 + 2\lambda W^{l-1}(T_i)) - 1 \) has a strong tail. Moreover, \( X_a \) also has a strong tail. So, for all \( t \geq A\lambda \),

\[
P(W(T) \geq t) \leq P\left( \prod_{i=1}^D (1 + 2\lambda W^{l-1}(T_i)) - 1 \geq \frac{t}{3} \right) + P(X_a \geq \frac{t}{3}) \leq 2f(t)/3 \leq f(t).
\]

Thus, \( W(T) \) also has a good tail. \( \square \)

6. **Proof of Theorem 1.3: upper bound for survival time on graphs**

As mentioned in Lemma 2.3, the local neighborhood of a vertex in \( G_n \sim G(n, \mu) \) is asymptotically a truncated GW tree. In Section 5, we have derived upper bounds for the expected survival time and number of infections at the leaves on truncated GW trees. To pass this result to the actual local neighborhoods, we show in Section 6.1 that all local neighborhoods of \( G_n \) are contained in the so-called edge-added Galton-Watson trees (E\text{GW}) which consist of truncated GW trees and cycles (G\text{WC}). We then derive similar upper bounds for the expected survival time and number of infections at the leaves on G\text{WC} in Section 6.2 and on E\text{GW} in Section 6.3. The derivation of these bounds are similar (though slightly more complicated) to the bounds for GW trees that we have
done in Section 5. We will combine all of these ingredients and finish the proof of Theorem 1.3 in Section 6.4.

6.1. Coupling the local neighborhood of $G_n$. In this section, we follow [2, 26] and show the coupling between random graphs and variants of GW trees. We utilize the notion of augmented distributions, which allows us to stochastically dominate $N(v, l)$ by a larger geometry.

**Definition 6.1** (Definition 4.2, [2]). (Augmented distribution) Let $\varepsilon \in (0, 1)$ and $\mu \equiv \{p_k\}_{k \in \mathbb{N}}$ be a probability measure on $\mathbb{N}$ such that $\mathbb{E}_{\mu} e^{cD} < \infty$ for some $c > 0$. Let

$$k_0 := \max\{k : \sum_{j \geq k} j\sqrt{p_j} \geq \varepsilon/10\} < \infty \quad \text{and} \quad k_{\text{max}} := \sup\{k : p_k > 0\} \in \mathbb{N} \cup \{\infty\}.$$ 

If $k_0 < k_{\text{max}}$, we define the augmented distribution $\mu^\varepsilon$ of $\mu$ by

$$\mu^\varepsilon(k) := \frac{1}{Z} \begin{cases} \frac{p_k}{2} & \text{if } k \leq k_0; \\ \sqrt{p_k} & \text{if } k > k_0, \end{cases}$$

where $Z = \sum_{k \leq k_0} p_k/2 + \sum_{k > k_0} \sqrt{p_k}$ is the normalizing constant. When $k_0 = k_{\text{max}}$, we set

$$\mu^\varepsilon(k) := \frac{1}{Z} \begin{cases} \frac{p_k}{2} & \text{if } k < k_0; \\ \sqrt{p_k} & \text{if } k = k_0, \end{cases}$$

for the normalizing constant $Z = \sum_{k < k_0} p_k/2 + \sqrt{p_{k_0}}$.

Here are some basic properties of augmented distributions.

**Lemma 6.2** (Lemma 4.3, [2]). Let $\mu$ be a probability distribution on $\mathbb{N}$.

1. If $\mu$ has an exponential tail, then so does $\mu^\varepsilon$. More specifically, if $\mathbb{E}_{\mu} e^{cD} < \infty$ then $\mathbb{E}_{\mu^\varepsilon} e^{cD} < \infty$.
2. Let $D_1, \ldots, D_n$ be $n$ independent samples of $\mu$. For a subset $\Delta \subset [n]$, let $\{p_k^\Delta\}_k$ denote the empirical distribution of $\{D_i\}_{i \in [n] \setminus \Delta}$. With high probability over the choice of $D_i$’s, $\{p_k^\Delta\}_k$ is stochastically dominated by $\mu^\varepsilon$, for any $\Delta \in [n]$ with $|\Delta| \leq n/3$.

**Remark 6.3.** The i.i.d $D_i$ in Lemma 6.2 can be viewed as a degree sequence of $G_n \sim G(n, \mu)$. Consider the exploration procedure starting from a single fixed vertex $v$, which, at each step, reveals a vertex adjacent to the current explored neighborhood and the half-edges incident to the new vertex. Then the second statement says that when the exploration process revealed $N \leq n/3$ vertices inside the local neighborhood of $v$, the empirical degree distribution of the $n - N$ unexplored vertices is stochastically dominated by $\mu^\varepsilon$, with high probability.

Using the above properties of augmented distributions, we develop a coupling argument to dominate $N(v, L) \subset G_n$ where $G_n \sim G(n, \mu)$ by a Galton-Watson type branching process. To this end, we first take account of the effect of emerging cycles in $N(v, L)$.

For a constant $\gamma > 0$, let $\mathcal{A}(\gamma)$ be the event that $N(v, \gamma \log n)$ in $G_n$ contains at most one cycle for all $v \in G_n$. The following lemma shows that we typically have $\mathcal{A}(\gamma)$ for some constant $\gamma$.

**Lemma 6.4** (Lemma 4.5, [2]). There exists $\gamma = \gamma(\mu) > 0$ such that for $G_n \sim G(n, \mu)$,

$$\mathbb{P}(G \in \mathcal{A}(\gamma)) = 1 - o(1).$$
We now define the three variants of GW trees that will be used. We use the definitions from [26, Section 6.1].

**Definition 6.5 (GWC).** Let $\xi$ be a positive, integer-valued random variable, and let $m, l \geq 1$ be nonnegative integers. Then, $\mathcal{H}^{m,l} \sim \text{GWC}^1(\xi, m) \leq l$, the Galton-Watson-on-cycle process of type one (in short, GWC$^1$-process), is generated according to the following procedure:

1. Consider a length-$m$ cycle $C = v_1v_2 \ldots v_mv_1$.
2. At each $v_j$ for $j = 1, \ldots, m$, attach $T_j \sim \text{i.i.d.} \ GW(\xi) \leq l$ by setting $v_j$ as its root.

We designate vertex $v_1$ as the root of $\mathcal{H}^{m,l}$ and denote $\rho = v_1$. Note that $m = 1$ corresponds to the usual Galton-Watson trees.

Next, we introduce the Galton-Watson-on-cycle process of type two (in short, GWC$^2$-process), which can be thought as a certain subgraph of GWC$^1$-processes.

**Definition 6.6 (GWC).** Let $\xi$ be a positive, integer-valued random variable, and let $m, l \geq 1$ be integers. Then, $\dot{\mathcal{H}}^{m,l} \sim \text{GWC}^2(\xi, m) \leq l$, the Galton-Watson-on-cycle process of type two (in short, GWC$^2$-process), is generated according to the following procedure:

1. Consider a length-$m$ cycle $C = v_1v_2 \ldots v_mv_1$.
2. At each $v_j$, $j \in \{2, \ldots, m\}$, attach $T_j \sim \text{i.i.d.} \ GW(\xi) \leq l$ by setting $v_j$ as its root. At $v_1$, we do nothing.

We designate vertex $v_1$ as the root of $\dot{\mathcal{H}}^{m,l}$ and denote $\rho = v_1$.

We now introduce the Galton-Watson type process which will be coupled with the local neighborhoods of $G_n$.

**Definition 6.7 (EGW).** Let $l, L, m$ be positive integers with $m \geq 2$ and $l \leq L$, and let $\xi$ be a probability distribution on $\mathbb{N}$. We define $\text{EGW}(\xi; l, m) \leq L$, the edge-added Galton-Watson process (in short, EGW-process) as follows:

1. Generate a $\text{GW}(\xi) \leq L$ tree, conditioned on survival until depth $l$.
2. At each vertex $u$ at depth $l$, add an independent $\text{GWC}^2(\xi, m) \leq L-l$ process (see Definition 6.6) rooted at $u$. Here we preserve the existing subtrees at $u$ which comes from $\text{GW}(\xi) \leq L$ tree from the above step.

Let $\xi'$ be another probability measure on $\mathbb{N}$. Then $\text{EGW}(\xi, \xi'; l, m) \leq L$ denotes the EGW-process where the root has degree distribution $\xi$, and all the descendants have $\xi'$. And in the second step of the definition, we add $\text{GWC}^2(\xi'; m) \leq L-l$ instead.

The next step is to show that the local neighborhood $N(v, L)$ is dominated by a combined law of the above graphs. In what follows, for two probability measures $\nu_1$ and $\nu_2$ on graphs, we say $\nu_1$ stochastically dominates $\nu_2$ and write $\nu_1 \geq_{st} \nu_2$ if there exists a coupling between $S_1 \sim \nu_1$ and $S_2 \sim \nu_2$ such that $S_2 \subset S_1$ in terms of isomorphic embeddings of graphs, i.e., there exists an injective graph homomorphism from $S_2$ into $S_1$. 

Fix a vertex \( v \in G_n \), and consider its local neighborhood \( N(v, \gamma \log n) \) with \( \gamma \) as in Lemma 6.4. Let \( \mathcal{A} = \mathcal{A}(\gamma) \) be the event in the same lemma. For each \( l, s \) with \( s \geq 2 \), we define the event \( \mathcal{B}_{l,s} \) to be the intersection of \( \mathcal{A} \) and the event that \( N(v, \gamma \log n) \) forms a cycle of length \( s \) at distance \( l \) from \( v \).

For the given degree distribution \( \mu \), let \( \tilde{\mu} \) be its size-biased distribution (see [4] for definition), and \( \tilde{\mu}^l := \tilde{\mu}_{(1, \infty)} \) denote the distribution \( \tilde{\mu} \) conditioned on being in the interval \([1, \infty)\). Let \( \mu^s \) and \( \tilde{\mu}^s \) be the augmented distributions of \( \mu \) and \( \tilde{\mu} \), respectively. Further, let \( \eta, \eta_{l,s} \) and \( \eta_0 \) denote the probability measures on rooted graphs describing the laws of \( N(v, \gamma \log n) \), \( \mathcal{E}_{GW}(\mu^s, \tilde{\mu}^s; l, s) \leq \gamma \log n \) and \( \mathcal{G}_W(\mu^s, \tilde{\mu}^s) \leq \gamma \log n \), respectively.

**Lemma 6.8** (Lemma 4.8, [2]). (The coupling) Under the above setting, we have the following stochastic domination:

\[
\eta_1 \mathcal{A} \leq_{st} \sum_{l,s \geq 2} b_{s,l} \eta_{s,l} + b_0 \eta_0,
\]

where \( b_{l,s} = \mathbb{P}(\mathcal{B}_{l,s}) \), \( b_0 = 1 - \sum_{l,s} b_{s,l} \).

### 6.2. Recursive analysis for unicyclic graphs \( \mathcal{G}_W^1 \) and \( \mathcal{G}_W^2 \).

In this section, we derive results similar to Section 5 for \( \mathcal{G}_W^1 \) and \( \mathcal{G}_W^2 \). Let us start with the definitions of survival time and number of infection at the leaves, first for \( \mathcal{G}_W^1 \) in Definition 6.9 and then for \( \mathcal{G}_W^2 \) in Definition 6.10.

**Definition 6.9** (Root-added contact process on \( \mathcal{G}_W^1 \)-processes). Let \( m, l \geq 1 \) be integers, and \( H \) be a graph that consists of a length-\( m \) cycle \( C = v_1 v_2 \ldots v_m v_1 \) and the trees \( T_1, \ldots, T_m \) rooted at \( v_1, \ldots, v_m \), respectively. Consider the root-added starred contact process \((X_t)\) on \( H \) which is the starred contact process on the graph \( H \cup v_1^+ \) with the permanently infected parent \( v_1^+ \) having a single connection with \( v_1 \). The vertex \( v_1 \) is initially at state \( 1^* \).

1. Let \( S(H) \) be the excursion time which is the first time when \( H \) returns to the all-healthy state. \( S(H) = \mathbb{E}_{cp} S(H) \) denotes the expected excursion time.
2. Let \( \mathcal{L}_j = \{v \in T_j : \text{dist}(v, v_j) = l\} \). For \( v \in \mathcal{L}_j \) for some \( j \), we define the total infections at \( v \) by

\[
M_{t,v}(H) := |\{s \in [0, S(H)] : X_s(v) = 1^* \text{ and } X_{s-}(v) = 0\}|.
\]

Then, the total infections at depth-\( l \) leaves and its expectation are given as

\[
M_l(H) := \sum_{j=1}^m \sum_{v \in \mathcal{L}_j} M_{t,v}(H), \quad \text{and} \quad \mathbb{E}_{cp} M_l(H).
\]

**Definition 6.10** (Root-added contact process on \( \mathcal{G}_W^2 \)-processes). For \( \mathcal{H}^m, l \sim \mathcal{G}_W^2(\xi, m) \leq I \), we define the root-added (starred) contact process on \( \mathcal{H}^m, l \) without adding a new parent to the root. Instead, we fix the root \( \rho = v_1 \) to be permanently infected. Assume that \( v_2 \) is initially at state \( 1^* \). Then, we define \( S_2(\mathcal{H}^m, l), M_2(\mathcal{H}^m, l) \) to be the expected excursion time and the expected total number of infections at depth-\( l \) leaves as in Definition 6.9. If \( v_m \) is initially at state \( 1^* \) instead, the corresponding identities are denoted by \( S_m(\mathcal{H}^m, l), M_m(\mathcal{H}^m, l) \).

We also write

\[
S(\mathcal{H}^m, l) = \frac{1}{w(\mathcal{H}^m, l)} \left( w_2(\mathcal{H}^m, l) S_2(\mathcal{H}^m, l) + w_m(\mathcal{H}^m, l) S_m(\mathcal{H}^m, l) \right),
\]
and

\[ M_t(\hat{H}^{m,l}) = \frac{1}{w(\hat{H}^{m,l})} \left( w_2(\hat{H}^{m,l}) M_{t,2}(\hat{H}^{m,l}) + w_m(\hat{H}^{m,l}) M_{t,m}(\hat{H}^{m,l}) \right) \]

where we define the weights to be

\[ w_2(\hat{H}^{m,l}) = \frac{\lambda(D_2 + 2)}{1 + \lambda(D_2 + 2)}, \quad w_m(\hat{H}^{m,l}) = \frac{\lambda(D_m + 2)}{1 + \lambda(D_m + 2)}, \quad w(\hat{H}^{m,l}) = w_2(\hat{H}^{m,l}) + w_m(\hat{H}^{m,l}) \]

with \( D_j \) being the number of children of \( v_j \) in \( T_j \).

The goal of this section is to establish the following analog of Lemmas 5.11 and 5.12.

**Lemma 6.11.** Assume that \( \mathbb{E} e^{3\xi} < \infty \). Let \( m, l \geq 1 \) be integers and \( \mathcal{H}^{m,l} \sim \text{GWC} \left( \xi, \mu \right) \) and \( \hat{H}^{m,l} \sim \text{GWC} \left( \xi, m \right) \). The following random variables have good tails:

\[ w(\mathcal{H}^{m,l}) S(\mathcal{H}^{m,l}), \quad w(\mathcal{H}^{m,l}) S(\hat{H}^{m,l}), \quad (1.5)^l w(\mathcal{H}^{m,l}) M_t(\mathcal{H}^{m,l}) \quad \text{and} \quad (1.5)^{l-1} w(\mathcal{H}^{m,l}) M_t(\hat{H}^{m,l}), \]

where the constants \( A, B, c \) in the definition of goodness (Definition 5.4) are any constants satisfying \( \text{Definition 2.7} \) with \( \xi \) in place of \( D \) and where

\[ w(\mathcal{H}^{m,l}) = \frac{\lambda(D_1 + 3)}{1 + \lambda(D_1 + 3)}. \]

Though the proof of Lemma 6.11 is technically more complicated, the idea is similar to the derivation of Lemmas 5.11 and 5.12. Therefore, we defer its proof to the Appendix 7.

### 6.3. Recursive analysis for \textbf{EGW}.

We now prove similar results for \textbf{EGW}-processes. We start with the definitions of survival time and total number of infections.

**Definition 6.12 (Excursion time and total infections at leaves for \textbf{EGW}).** Let \( h, l, m \) be nonnegative integers with \( m \geq 2 \), \( \nu, \nu' \) be probability measures on \( \mathbb{N} \), and \( \mathcal{H} \sim \text{EGW}(\nu, \nu'; h, m) \). We connect a permanently infected parent \( \rho^+ \) to the root \( \rho \) of \( \mathcal{H} \). Consider the root-added starred contact process \( (X_t) \) on \( \mathcal{H} \) with \( \rho^+ \) always infected and \( \rho \) initially at \( 1^* \).

- The excursion time \( S(\mathcal{H}) \) is the first time when \( (X_t) \) becomes all-healthy on \( \mathcal{H} \), and \( S(\mathcal{H}) = \mathbb{E}_{cv} S(\mathcal{H}) \) denotes the expected excursion time on \( \mathcal{H} \).
- Let \( \mathcal{L} \) be the collection of bottom leaves of \( \mathcal{H} \), that is, denoting \( \{C_j\} \) to be the length-\( m \) cycles in \( \mathcal{H} \),

\[ \mathcal{L} = \{ v \in \mathcal{H} : \text{dist}(v, \rho) \geq l \text{ and dist}(v, C_j) \geq l - h \text{ for all } j \} \]

- Let \( v \in \mathcal{L} \) and denote the total infections at \( v \) by

\[ M_{t,v}(\mathcal{H}) := |\{ s \in [0, S(\mathcal{H})] : X_s(v) = 1^* \text{ and } X_{s-}(v) = 0 \}|. \]

Then, we set the total infections at depth-\( l \) leaves as

\[ M_l(\mathcal{H}) := \sum_{v \in \mathcal{L}} M_{t,v}(\mathcal{H}). \]

- We also let \( M_l(\mathcal{H}) := \mathbb{E}_{cv} M_l(\mathcal{H}) \) be the expected total infections at leaves.

For an \textbf{EGW}-process \( \mathcal{H} \), the tail probabilities of \( S(\mathcal{H}) \) and \( M_l(\mathcal{H}) \) are good, up to appropriate rescales.
Proposition 6.13. Let $\mu$ be a probability measure satisfying the hypothesis of Theorem 1.3. Let $m, l, h$ be nonnegative integers such that $m \geq 2$ and $l \geq h+1$. Let $\mathcal{H} \sim \text{EGW}(\mu^2, \bar{\mu}; h, m)_{\leq l}$. Then, it holds that $w(\mathcal{H})S(\mathcal{H})$ and $1.5w(\mathcal{H})M(\mathcal{H})$ have good tails where $w(\mathcal{H}) = \frac{\lambda(D+1)}{1+\lambda(D+1)}$ and $D$ is the number of children of the root.

Proof. Let $\mathcal{H}' \sim \text{EGW}(\bar{\mu}; h, m)_{\leq l}$. We shall establish the same conclusion for both $\mathcal{H}$ and $\mathcal{H}'$ by an induction on $h$. When $h = 0$, since $\mathcal{H}' \sim \text{GWC}^1(\mu^2, m)_{\leq l}$ and $\mathcal{H} \sim \text{GWC}^1(\mu^2, \bar{\mu}; m)_{\leq l}$, the $\text{GWC}^1$-process in which only the root of the Galton-Watson tree has offspring $\mu^2$ and all others have offspring $\bar{\mu}$. Therefore, Lemma 6.11 gives we have the desired statement for both $\text{EGW}(\bar{\mu}; 0, m)_{\leq l}$ and $\text{EGW}(\mu^2, \bar{\mu}; 0, m)_{\leq l}$ for all $m \geq 2, l \geq 0$.

Suppose that we have the statement for $\mathcal{H}' \sim \text{EGW}(\bar{\mu}; h, m)_{\leq l}$ with $m \geq 2, h \geq 0$ and $l \geq h+1$. Let $\rho_1$ be the root of $\mathcal{H}_1 \sim \text{EGW}(\mu^2; h+1, m)_{\leq l}$ and $D_1 \sim \bar{\mu}^2$ be its degree. Then the subgraphs $H_1^1, \ldots, H_{D_1}^1$, rooted at each child of $\rho_1$ are i.i.d. $\text{EGW}(\bar{\mu}; h, m)_{\leq l-1}$. Since the desired tail statement holds for $H_1^1, \ldots, H_{D_1}^1$ by the induction hypothesis, we use Lemma 5.5 and a recursive bound completely similar to that of Lemma 5.1 to conclude that $\mathcal{H}_1$ satisfies the tail statement as well. The result for $\mathcal{H}_2 \sim \text{EGW}(\mu^2, \bar{\mu}; h+1, m)_{\leq l}$ follows similarly as $\mathcal{H}_2$ has subgraphs $H_1^2, \ldots, H_{D_2}^2 \sim$ i.i.d. $\text{EGW}(\bar{\mu}; h, m)_{\leq l-1}$ with $D_2 \sim \mu^2$. We note that by Condition 3 and the first part of Lemma 6.2, $\mu^2$ and $\bar{\mu}^2$ satisfy the hypothesis of Lemma 5.5 which allows the above arguments to go through. \hfill \Box

6.4. Finishing the proof of Theorem 1.3. We are now ready to finish the proof of Theorem 1.3. It suffices to show that with high probability on the randomness of $G_n$, for all vertices $v \in V(G_n)$, with probability at least $1 - o(\frac{1}{n})$, the starred contact process $(X^{(v)}_t)$ with $v$ being strongly infected initially dies out before time $n^{\lambda^2 \log(\lambda^{-1})}$. We then take the union bound over all $v \in V(G_n)$ and note that by using the graphical representation, the survival time of the contact process starting with all vertices infected is the maximum of the survival times of the individual $(X^{(v)}_t)$ to obtain the statement of the theorem (we also use (13) to go from the starred contact process to the original contact process).

6.4.1. Handling individual local pieces. Recall the definition of $q$ in (28)

$$q = \frac{c}{16\lambda^2 \log(\lambda^{-1})}.$$ 

Let $\delta = \frac{10}{q}, r = \delta \log n$. We note that for $\lambda$ sufficiently small (depending on $\mu$), $\delta$ is also sufficiently small. In particular, we assume $\lambda$ is small enough so that $\delta \leq \gamma/100$, the constant appearing in Lemma 6.4. By this lemma, with high probability, $N(v, r)$ contains at most one cycle for all $v \in V(G_n)$. For each $v$, consider the local neighborhood $B_v$ of $v$ as follows.

- $B_v := N(v, r)$, if $N(v, r)$ is a tree.
- If $N(v, r)$ contains the unique cycle $C_v$ at distance $h$ from $v$, then

$$B_v := N(v, r) \cup \left[ \bigcup_{u \in C_v} N(u, r - h) \right].$$

We define $B_v^+$ the same way except replacing $r$ with $10r$. With this definition, $B_v \subset B_u^+$ for all $u \in B_v$. We first provide a bound on the size of $B_v^+$. 

33
Lemma 6.14. With high probability, for all $v$, it holds that
\[ |B_v^+| \leq n^{C\delta}/2 \]
where $C$ is some constant depending on $\mu$.

For the rest of this section, we assume (wlog) that $C > 2$.

Proof. First, observe that for a fixed $v$, $|B_v^+|$ is bounded by the number of vertices in $T_{\leq 20r}$ where $T \sim \GW(\mu, \bar{\mu})$. To bound the latter, we follow [20] Theorem 2.3] and show the following at the end of this section.

Lemma 6.15. Let $T \sim \GW(\mu, \bar{\mu})$ where $\mu$ has an exponential tail $\mathbb{E}_{N \sim \mu} e^{cN} < \infty$ for some $c > 0$. Then there exists $t > 0$ such that
\[ \sup_s \mathbb{E} \exp(tZ_s \bar{d}^{-s}) < \infty \]
where $\bar{d} = \mathbb{E}_{D \sim \bar{\mu}} D$ and $Z_s$ is the number of vertices at level $s$ of $T$.

Assuming this lemma, by Markov’s inequality, with probability $1 - O(\log n n^{-C})$ for all $0 \leq s \leq 20r$, 
\[ Z_s \leq t^{-1} C \bar{d}^s \log n. \]

On this event,
\[ |B_v^+| \leq \sum_{s=1}^{20r} Z_s \leq t^{-1} C (\log n)^2 \cdot n^{20\delta \log \bar{d}} \leq n^{C\delta}/2. \]

Taking union bound over $v$, we obtain the result. \hfill \Box

For the contact process $(X_t^{(v)})$ on $G_n$ and for $A \subset V(G_n)$, let $(X_t^{(v)}(t))$ be the contact process restricted on $A$, that is we only use the recovery and infection clocks inside $A$. Let $S_v$ be the expected survival time of the contact process $(X_t^{(v)}(t))$. Under $(X_t^{(v)}(t))$, let $M_v$ be the expected total infections at the leaves $\mathcal{L}(B_v)$ which is defined by
\[ \mathcal{L}(B_v) := \begin{cases} \{ u \in B_v : \text{dist}(u,v) = r \} & \text{if } B_v \text{ is a tree}, \\ \{ u \in B_v : \text{dist}(u,v) \geq r \text{ and dist}(u,C_v) \geq r - h \} & \text{if } C_v \text{ exists}, \end{cases} \]
where $C_v$ is the unique cycle (if exists) in $N(v) \cap A$ and $h = \text{dist}(v,C_v)$. Note that this definition is consistent with the definition of leaves of $\EGW$-processes.

Lemma 6.16. We have that
\[ \mathbb{P} \left[ \max_{v \in V(G_n)} S_v \leq n^\delta, \max_{v \in V(G_n)} M_v \leq \frac{1}{\log n} \right] \geq 1 - o(1). \]

Proof. Fix $v \in V(G_n)$. By Lemma 6.15, $B_v$ and $B_v^+$ are dominated by $\EGW(\mu^2, \bar{\mu}^2, \lambda, \delta \log n)$ and $\GW(\mu^2, \bar{\mu}^2, \delta \log n)$. By Proposition 6.13, it holds that $\frac{\lambda(D+1)}{1+\lambda(D+1)} S_v$ and $1.5^{\delta \log n} \frac{\lambda(D+1)}{1+\lambda(D+1)} M_v$ have good tails where $D = \deg_{G_n}(v) \sim \mu$. Thus,
\[ \mathbb{P}_{RG}(S_v > n^\delta) \leq \mathbb{P}_{RG} \left( \frac{\lambda(D+1)}{1+\lambda(D+1)} S_v \geq \lambda n^\delta \right). \]
For $n$ sufficiently large, we have $\lambda n^\delta \geq \frac{1}{\log n}$ and so for sufficiently large $n$,

$$\Pr_{RG}(S_v > n^\delta) \leq C_{\lambda,\mu} n^{-\delta q + o(1)} \leq n^{-3/2},$$

where $C_{\lambda,\mu}$ is some constant depending on $\mu$ and $\lambda$. Similarly,

$$\Pr_{RG}\left(M_v > \frac{1}{\log n}\right) \leq \Pr_{RG}\left(1.5^\delta \log n \frac{\lambda(D + 1)}{1 + \lambda(D + 1)} M_v > \lambda n^\delta/4\right) \leq C_{\lambda,\mu} n^{-q\delta/4} \leq C_{\lambda,\mu} n^{-2}.$$

Taking the union bound over $n$ vertices of $G_n$, we obtain the conclusion. □

Let $E_n$ be the intersection of the good events in Lemmas 6.14 and 6.16. By Markov’s inequality and Lemma 6.16, we obtain the following.

**Lemma 6.17.** Let $G_n \in E_n$. Let $C$ be the constant in Lemma 6.14. For each vertex $v \in V(G_n)$, let $S_v$ be the survival time of $(X_{B_v}(v))(t)$ and $M_v$ be the number of infections at the leaves of $B_v$ of $(X_{B_v}(v))(t)$. We have

$$\Pr_{CP}(S_v \geq jn^{(C+1)\delta}) \leq 2^{-j} \text{ for all } j \geq 1$$

and

$$\Pr_{CP}(M_v \geq jn^{(3C+1)\delta}) \leq 2^{-j+1} \text{ for all } 1 \leq j \leq C \log n.$$

**Proof.** To see (48), at any time $t \leq S_v$, the probability that $(X_{B_v}(v))$ survives until time $t + n^{C\delta}$ is bounded as follows.

$$\Pr_{CP}(S_v > t + n^{(C+1)\delta}|S_v \geq t) \leq \sum_{u \in B_v} \Pr_{CP}(S_v > n^{(C+1)\delta}|X_{B_v}(v) \text{ starts with } u \text{ initially at } 1^*)$$

$$\leq |B_v| n^{-C\delta} \leq 1/2$$

where in the second inequality we use Markov’s inequality, Lemma 6.16 and the fact that $B_v^+$ contains $B_v$. Applying this probability for $t = kn^{(C+1)\delta}$, $k = 0, \ldots, j - 1$, we obtain the statement of the lemma.

Next, for (49), let $C_{v,t_0}$ be the total number of infection clock rings in $B_v$ before time $t_0$. We observe that $M_v \leq C_{v,t_0}$. Thus,

$$\Pr_{CP}(M_v \geq jn^{(3C+1)\delta}) \leq \Pr_{CP}(S_v \geq jn^{(C+1)\delta}) + \Pr_{CP}(C_{v,jn^{(C+1)\delta}} \geq jn^{(3C+1)\delta}).$$

By (48), the first term is bounded by $2^{-j}$. Since $|B_v| \leq n^{C\delta}$, $C_{v,jn^{(C+1)\delta}} \leq \text{Pois}(\lambda jn^{(3C+1)\delta})$. Thus, by the tail bound (6) for Poisson distributions, the last term in (50) is also bounded by $2^{-j}$. □

**Lemma 6.18.** Let $k = n^{(6C+6)\delta}$. Let $M_1, \ldots, M_k$ be independent random variables satisfying (49) and $\mathbb{E}M_i \leq \frac{1}{\log n}$. We have

$$\Pr\left(\sum_{i=1}^{k} M_i \geq k - 1\right) \leq n^{-C/2}.$$

**Proof.** By (49),

$$\Pr(M_i \geq C \log n \cdot n^{(3C+1)\delta}) \leq 2n^{-C}.$$

And so, as $\delta$ is sufficiently small,

$$\Pr(M_i \leq C \log n \cdot n^{(3C+1)\delta}, i = 1, \ldots, k) \geq 1 - 2kn^{-C} \geq 1 - \frac{1}{2n^{C/2}}.$$
Applying the Azuma’s inequality to the random variables \(X_i := M_i 1_{M_i \leq C \log n \cdot n (3C+1)\delta}\) with \(\mathbb{E}X_i \leq \frac{1}{\log n}\), we get
\[
\mathbb{P}(\sum_{i=1}^{k} X_i \geq k - 1) \leq \exp \left( -\frac{k^2}{8 \sum_{i=1}^{k} ||X_i||_2^2} \right) \leq \exp \left( -\frac{k}{8C^2(\log n)^2 \cdot n (6C+2)\delta} \right) \leq \frac{1}{2n^{C/2}}
\]
proving the lemma. \(\square\)

6.4.2. Combining the local pieces. Finally, we show how to combine the local information to bound the survival time on the whole graph. We use the following decomposition of the contact process \((X_t^{(u)})\) on \(G_n\) by stochastically dominating it by a collection of processes running on local neighborhoods \(\{B_u\}_u\).

1. Initially, run \((X_t^{(u)}(t))\).
2. In \(B_u\), when some \(u \in \mathcal{L}(B_u)\) becomes strongly infected in \(B_u\) at time \(t\) (and has been healthy until time \(t-\)), initiate an independent copy of \((X_t^{(u)})\) on \(B_u\).
3. Repeat Step 2 on every running copies of \((X_t^{(u)})\) until all the processes have terminated, that is, when all vertices in every generated copy are healthy.

The survival time of this decomposed process stochastically dominates the process on \(G_n\). Let \(R_v\) denote the termination time of the decomposed process and let \(R_v\) be its expectation. Let \(U_v\) be the enumeration of vertices \(u\) of which a copy of \((X_t^{(u)})\) has been generated during this process, counted with multiplicities. The size of \(U_v\) can be controlled by \(M_v\).

**Lemma 6.19.** Let \(C\) be the constant in Lemma 6.14. For each \(v \in V(G_n)\),
\[
(51) \quad \mathbb{P}(\|U_v\| \geq n^{(6C+6)\delta}) \leq n^{-C/2}.
\]
Thus, with high probability, for all \(v \in V\), we have \(\|U_v\| \leq n^{(6C+6)\delta}\).

**Proof of Lemma 6.19.** We think about \(U_v\) as the total number of vertices in the tree with a root at \(v\) and for each vertex \(u\) of the tree, the number of its children equals the number of infections at the leaves of \(B_u\). In particular, \(U_v\) can be thought of as a Galton-Watson tree in which \(M_v\) is the number of vertices in the first generation, denoted by \(v_1, \ldots, v_k\). Then \(M_{v_1} + \cdots + M_{v_k}\) is the number of vertices in the second generation and so on. Assume that the total number of vertices of the tree is \(\|U_v\| = K\). We then have
\[
1 + M_1 + M_2 + \ldots + M_K = K
\]
where \(M_i\) is the number of children of the \(i\)-th vertex of \(U_v\). These \(M_i\) are independent and satisfies (49) by Lemma 6.17. Note that for every \(l \leq K\), we have \(1 + M_1 + \ldots + M_l \geq l\) as the LHS is at least the number of vertices in the tree formed by vertices \(1, \ldots, l\). Thus, letting \(k = n^{(6C+6)\delta}\), by Lemma 6.18 we obtain
\[
(52) \quad \mathbb{P}(\|U_v\| \geq n^{(6C+6)\delta}) \leq \mathbb{P}(1 + M_1 + \ldots + M_k \geq k) \leq n^{-C/2}
\]
proving the lemma. \(\square\)

Back to the proof of Theorem 1.3 thanks to Lemma 6.19 we conclude that with high probability, for all \(v \in V(G_n)\), at any time \(t\), the number of infected local neighborhoods \(B_u\) is at most \(n^{(6C+6)\delta}\).
By Lemma 6.16, conditioned on this event and on \((X_t^{(v)})\), the expected survival time of \((X^{(v)}_s)_{s \geq t}\) is at most

\[
\sum_{|X_t^{(v)} \cap U_v|} n^\delta \leq n^{(6C+7)\delta}
\]

and so \((X_s)\) survives until time \(t + 2n^{(6C+7)\delta}\) with probability at most 1/2. Repeating this argument for \(t = a \cdot 2n^{(6C+7)\delta}\) with \(a \in [1, 2, \ldots, 2 \log n]\), we conclude that the process \((X_t^{(v)})\) survives until time \(4 \log n \cdot n^{(6C+7)\delta}\) with probability at most \(n^{-2}\).

Taking the union bound over all \(n\) vertices \(v\), with probability at most \(n^{-1}\), the contact process starting at all vertices infected survives until time \(4 \log n \cdot n^{(7C+72)\delta}\), concluding the proof of Theorem 1.3.

**Proof of Lemma 6.15.** Let \(N_i \sim \tilde{\mu}\) be independent. We have for all \(s \geq 1\),

\[
\mathbb{E}\exp(tZ_{s+1}\tilde{d}^{-(s+1)}) = \mathbb{E}\exp\left(\sum_{i=1}^{Z_s} tN_i \tilde{d}^{-(s+1)}\right) = \mathbb{E}\left[\mathbb{E}\left(\sum_{i=1}^{Z_s} tN_i \tilde{d}^{-(s+1)}\right) | \mathcal{Z}_s\right]
\]

\[
= \mathbb{E}\left(\mathbb{E}\left(tN_1 \tilde{d}^{-(s+1)}\right)^Z_s\right) = \mathbb{E}\exp\left(\mathbb{E}\exp(tN_1 \tilde{d}^{-(s+1)})\right).
\]

Since \(\mathbb{E}_{N \sim \mu} e^{CN} < \infty\), we have \(\mathbb{E}_{N \sim \mu} e^{CN} < \infty\) for all \(u < c\). Let \(b = c/2\). For all \(u \in (0, b]\), we have

\[
\mathbb{E}e^{uN_i} = \mathbb{E}\sum_{i=0}^{\infty} \frac{u^i N_i}{i!} \leq 1 + u\tilde{d} + \frac{u^2}{b^2} \sum_{i=2}^{\infty} \frac{b^i N_i}{i!}
\]

\[
\leq e^{u\tilde{d}} + \frac{u^2}{b^2} e^{bN_i} \leq e^{u\tilde{d}} + \exp\left(\frac{u^2}{b^2} e^{bN_i}\right) - 1 \leq \exp\left(u\tilde{d}(1 + u\alpha)\right)
\]

where \(\alpha\) is a large constant. Let \(t > 0\) be such that

\[
\exp\left(2\alpha t \sum_{i=1}^{\infty} \tilde{d}^{-i}\right) < 2 \quad \text{and} \quad 2t \leq b.
\]

We define \(t_n = t \exp\left(2\alpha t \sum_{i=n+1}^{\infty} \tilde{d}^{-i}\right) \leq t \exp\left(2\alpha t \sum_{i=1}^{\infty} \tilde{d}^{-i}\right) < 2t \leq b\). Then, for all \(s \geq 1\),

\[
\mathbb{E}\exp(t_{s+1}Z_{s+1}\tilde{d}^{-(s+1)}) = \mathbb{E}\exp\left(Z_s \log\left(\mathbb{E}\exp(t_{s+1}\tilde{d}^{-(s+1)}N_1)\right)\right)
\]

\[
\leq \mathbb{E}\exp\left(Z_s \tilde{d}^{-s} t_{s+1} (1 + \alpha t_{s+1} \tilde{d}^{-(s+1)})\right) \quad \text{by (53)}
\]

\[
\leq \mathbb{E}\exp\left(Z_s \tilde{d}^{-s} t_{s+1} (1 + 2t \alpha \tilde{d}^{-(s+1)})\right) \quad \text{because} \ t_{s+1} \leq 2t
\]

\[
\leq \mathbb{E}\exp\left(t_s Z_s \tilde{d}^{-s}\right).
\]

Thus,

\[
\sup_{s \geq 2} \mathbb{E}\exp\left(t_s Z_s \tilde{d}^{-s}\right) = \mathbb{E}\exp\left(t_1 Z_1 \tilde{d}^{-1}\right) < \infty
\]

for \(t\) sufficiently small.  

\(\Box\)
In this section, we prove Lemma \ref{lem:main}. Let \( v_1 \ldots v_m v_1 \) be its cycle part in \( \mathcal{H}^{m,l} \), and let \( T_j, j \in [m] \) be i.i.d. \( GW(\xi) \leq l \) rooted at \( v_j \). Let \( D_j \) be the degree of \( v_j \) in \( T_j \), and let \( T_{j,1}, \ldots, T_{j,D_j} \) be the subtrees rooted at \( u_{j,1}, \ldots, u_{j,D_j} \), the children of \( v_j \) in \( T_j \). Then, \( \mathcal{H}^{m,l} \) is simply the graph that consists of the cycle \( v_1 \ldots v_m v_1 \) and the trees \( T_j \) with \( j \in \{2, \ldots, m\} \). Let \( d_{j,1}, \ldots, u_{j,D_j} \) be the number of children of \( u_{j,1}, \ldots, u_{j,D_j} \), respectively.

We recall the weights
\[
\begin{align*}
    w_2(\mathcal{H}^{m,l}) &= \frac{\lambda(D_2 + 2)}{1 + \lambda(D_2 + 2)}, \\
    w_m(\mathcal{H}^{m,l}) &= \frac{\lambda(D_m + 2)}{1 + \lambda(D_m + 2)}, \\
    w(\mathcal{H}^{m,l}) &= w_2(\mathcal{H}^{m,l}) + w_m(\mathcal{H}^{m,l})
\end{align*}
\]

We prove Lemma \ref{lem:main} by induction on \( m + l \). If \( m = l = 1 \), the result follows directly from Lemmas \ref{lem:main} and \ref{lem:main} because \( \mathcal{H}^{m,l} \sim GW(\xi) \leq l \) and \( \mathcal{H}^{m,l} \) consists of a single vertex \( v_1 \). Assume that we have the stated result for all \( m', l' \) such that \( m' + l' < m + l \).

### 7.1. Proof of Lemma \ref{lem:main} for \( w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l}) \)

We first derive a recursive inequality for \( S(\mathcal{H}^{m,l}) \) which is similar to Lemma \ref{lem:main}.

**Lemma 7.1.** We have
\[
\frac{1}{2} w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l}) \leq \frac{2\lambda(D_2 + 2) + \lambda^2(D_2 + 2)^2}{(1 + \lambda(D_2 + 2))^2} + \frac{2\lambda(D_m + 2) + \lambda^2(D_m + 2)^2}{(1 + \lambda(D_m + 2))^2} + (\Pi_2 - 1) + (\Pi_m - 1)
\]

where
\[
\Pi_2 = (1 + \lambda w(\mathcal{H}^{m-1,l})S(\mathcal{H}^{m-1,l})) \prod_{i=1}^{D_2} (1 + \lambda w(T_{2,i})S(T_{2,i}))
\]

and
\[
\Pi_m = (1 + \lambda w(\mathcal{H}^{m-1,l})S(\mathcal{H}^{m-1,l})) \prod_{i=1}^{D_m} (1 + \lambda w(T_{m,i})S(T_{m,i})).
\]

**Proof.** Let us start by analyzing \( S_2(\mathcal{H}^{m,l}) \). The analysis basically follows Lemma \ref{lem:main} except that \( v_1 \) may send an infection to \( v_m \) before \( v_2 \) sends out an infection. In that case, \( S_m(\mathcal{H}^{m,l}) \) also comes into play.

Let \( (X_t) \) be the root-added starred contact process on \( \mathcal{H}^{m,l} \) as in Definition \ref{def:starred}. Let \( (X_{2,t}^\sharp) \) be the process for which the healing clock at \( v_2 \) is disabled until all vertices of \( \mathcal{H}^{m,l} \setminus \{v_1,v_2\} \) are healthy. Let \( S_2(\mathcal{H}^{m,l}) \) be the expected excursion time of \( (X_{2,t}^\sharp) \). We have \( S_2(\mathcal{H}^{m,l}) \leq S_2(\mathcal{H}^{m,l}) \).

Let \( t_0 \) be the first time that an event at \( v_2 \) happens, which must be an infection from \( v_2 \) to \( \{v_1,v_3\} \cup \bigcup_{i=1}^{D_2} v_{2,i} \) rather than a recovery at \( v_2 \) because \( v_2 \) is initially at state 1*'. We have
\[
(54) \quad E t_0 = \frac{1}{1 + \lambda(D_2 + 2)}.
\]
Let \( \bar{t} \) be the first time that \( v_1 \) sends a strong infection to \( v_m \). Let \( t_1 \) be the first time in \( (X_t) \) after time \( t_0 \) that all vertices in \( \mathcal{H}^{m,l} \setminus \{v_1, v_2\} \) are healthy. Let \( S_2^{\otimes} \) be the expected excursion time of the contact process \( (X_{2,t}^\otimes) \) on \( \mathcal{H}^{m,l} \) with \( v_1 \) and \( v_2 \) always infected and one of their neighbors initially at \( 1^* \). For this product process, it is helpful to think about \( v_1 \) and \( v_2 \) as a giant vertex whose neighbors are \( v_3, v_m \) and \( v_2 \), with \( i = 1, \ldots, D_2 \). Let \( \theta_2^{\otimes} \) be the probability that an infection from \( \{v_1, v_2\} \) to \( \{v_3, v_m\} \cup \{v_2, i\} \) is a strong infection. Let \( \theta_2 \) be the probability that an infection from \( v_2 \) to \( \{v_3\} \cup \{v_2, i\} \) is a strong infection. By writing down the formula for \( \theta_2^{\otimes} \) and \( \theta_2 \) in terms of \( D_m, D_3, d_{2,i} \), it is clear that

\[
\theta_2 \leq \frac{D_2 + 2}{D_2 + 1} \theta_2^{\otimes}.
\]

If \( \bar{t} > t_0 \), then

\[
\mathbb{E}(t_1 - t_0 | \bar{t} > t_0) \leq \frac{D_2 + 1}{D_2 + 2} \theta_2 S_2^{\otimes} \leq \theta_2^{\otimes} S_2^{\otimes}
\]

where \( \frac{D_2 + 1}{D_2 + 2} \) is the probability that the infection from \( v_2 \) at time \( t_0 \) is not sent to \( v_1 \).

If \( \bar{t} < t_0 \), we need to take into account the spread of infection from \( v_m \) and so by using the graphical representation, we get

\[
\mathbb{E}(t_1 - t_0 | \bar{t} < t_0) \leq \max\{S_m(\mathcal{H}^{m,l}), \frac{D_2 + 1}{D_2 + 2} \theta_2 S_2^{\otimes}\} \leq S_m(\mathcal{H}^{m,l}) + \theta_2^{\otimes} S_2^{\otimes}.
\]

Combining these equations, we get

\[
\mathbb{E}(t_1 - t_0) \leq \mathbb{P}(\bar{t} < t_0) S_m(\mathcal{H}^{m,l}) + \theta_2^{\otimes} S_2^{\otimes} \leq \frac{\lambda w_m(\mathcal{H}^{m,l})}{1 + \lambda(D_2 + 2)} S_m(\mathcal{H}^{m,l}) + \frac{\lambda w_m(\mathcal{H}^{m,l})}{1 + \lambda(D_2 + 2)} \theta_2^{\otimes} S_2^{\otimes}
\]

where in the last equality, we noted that \( v_1 \) sends a strong infection to \( v_m \) with rate \( \lambda w_m(\mathcal{H}^{m,l}) \).

From time \( t_1 \), we bound the survival time of \( (X_t)_{t \geq t_1} \) by \( (X^\otimes_{2,t})_{t \geq t_1} \) and use the same argument as for \([19]\) to obtain

\[
S_2(\mathcal{H}^{m,l}) \leq \mathbb{E} t_0 + \mathbb{E}(t_1 - t_0) + \sum_{k=0}^{\infty} \left( \frac{\lambda(D_2 + 2)}{1 + \lambda(D_2 + 2)} \right)^k \left( \frac{1}{1 + \lambda(D_2 + 2)} \right)^k \left( \frac{k + 1}{1 + \lambda(D_2 + 2)} + k \theta_2 S_2^{\otimes} \right)
\]

\[
\leq \frac{1}{1 + \lambda(D_2 + 2)} + \frac{\lambda w_m(\mathcal{H}^{m,l})}{1 + \lambda(D_2 + 2)} S_m(\mathcal{H}^{m,l}) + \theta_2^{\otimes} S_2^{\otimes} + 1 + \lambda(D_2 + 2) \theta_2^{\otimes} S_2^{\otimes}
\]

\[
= \frac{1}{1 + \lambda(D_2 + 2)} + \frac{\lambda w_m(\mathcal{H}^{m,l})}{1 + \lambda(D_2 + 2)} S_m(\mathcal{H}^{m,l}) + \frac{1 + \lambda(D_2 + 2)}{1 + \lambda(D_2 + 2)} (1 + \lambda(D_2 + 2)) \theta_2^{\otimes} S_2^{\otimes}
\]

For \( S_2^{\otimes} \), we have the following analog of \([21]\)

\[
1 + \lambda(D_2 + 2) \theta_2^{\otimes} S_2^{\otimes} = \left( 1 + \frac{\lambda^2(D_2 + 2)}{1 + \lambda(D_2 + 2)} S_m(\mathcal{H}^{m-1,l}) \right) + \frac{\lambda^2(D_m + 2)}{1 + \lambda(D_m + 2)} S_m(\mathcal{H}^{m-1,l}) \right)
\]

\[
\prod_{i=1}^{D_1} (1 + \lambda w(T_{2,i}) S(T_{2,i})) = \left( 1 + \lambda w(\mathcal{H}^{m-1,l}) S(\mathcal{H}^{m-1,l}) \right) \prod_{i=1}^{D_2} (1 + \lambda w(T_{2,i}) S(T_{2,i}) = \Pi_2.
\]

where \( \mathcal{H}^{m-1,l} \sim \text{GWC}^2(\xi, m - 1)^l \) contains the cycle \( v_2, v_3, \ldots, v_m \) and the trees \( T_i \) attached to \( v_3, \ldots, v_m \).
part. Thus, by induction hypothesis and Lemma 5.5, Π₂ ≤ \frac{-1}{(1+\lambda(D_m+2))^2} + w_2(\mathcal{H}^{m,l})w_m(\mathcal{H}^{m,l})S_m(\mathcal{H}^{m,l}) + \Pi_2

Similarly, we have

w_m(\mathcal{H}^{m,l})S_m(\mathcal{H}^{m,l}) ≤ \frac{-1}{(1+\lambda(D_m+2))^2} + w_2(\mathcal{H}^{m,l})S_2(\mathcal{H}^{m,l}) + \Pi_m.

Adding these up and arranging, we get

\begin{align*}
w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l}) &\leq \frac{-1}{(1+\lambda(D_m+2))^2} + \frac{\lambda w_2(\mathcal{H}^{m,l})w_m(\mathcal{H}^{m,l})}{1+\lambda(D_m+2)}S_m(\mathcal{H}^{m,l}) + \Pi_2 + \Pi_m \\
eq &\frac{2\lambda(D_m+2) + \lambda^2(D_m+2)^2}{1+\lambda(D_m+2)^2} + \frac{2\lambda(D_m) + \lambda^2(D_m)^2}{1+\lambda(D_m)^2} + (\Pi_2 - 1) + (\Pi_m - 1).
\end{align*}

completing the proof of the recursion. □

By induction hypothesis and Lemma 5.5, Π₂ − 1 and Πₘ − 1 have strong tails. So do the constant part. Thus, \(w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l})\) has a good tail, proving Lemma 6.11 for \(w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l})\).

7.2. Proof of Lemma 6.11 for \(S(\mathcal{H}^{m,l})\). By the same argument as in Lemmas 5.1 and 7.1 we obtain the following recursion.

Lemma 7.2. Under the hypothesis of Lemma 6.11, we have

\begin{align*}
S(\mathcal{H}^{m,l}) &\leq \frac{1}{1+\lambda(D_1+3)} - \frac{1}{\lambda(D_1+3)} + \frac{w(\mathcal{H}^{m,l})}{1+\lambda(D_1+3)} \left(1 + \lambda w(\mathcal{H})S(\mathcal{H})\right) \prod_{i=1}^{D_1} (1 + \lambda w(\mathcal{T}_{1,i})S(\mathcal{T}_{1,i})).
\end{align*}

By the induction hypothesis, each \(w(\mathcal{T}_{1,i})S(\mathcal{T}_{1,i})\) has a good tail. As proved in the previous section, \(w(\mathcal{H})S(\mathcal{H})\) also has a good tail. Combining this with Lemma 6.2 and Lemma 6.3 as in the proof of Lemma 5.11 we conclude that \(w(\mathcal{H}^{m,l})S(\mathcal{H}^{m,l})\) has a good tail.

7.3. Proof of Lemma 6.11 for \(M_t(\mathcal{H})\). Following the proof of Lemma 5.3 we derive a recursive bound for \(M_t(\mathcal{H})\).

Lemma 7.3. Under the hypothesis of Lemma 6.11, we have

\begin{align*}
\frac{1}{2} w(\mathcal{H}^{m,l})M_t(\mathcal{H}^{m,l}) &\leq \lambda(\Sigma_2 + \Sigma_m)
\end{align*}

where

\begin{align*}
\Sigma_2 &= \left[1 + \lambda w(\mathcal{H}^{m-1,l})S(\mathcal{H}^{m-1,l})\right] \left[\sum_{i=1}^{D_2} w_{2,i}M_{t-1}(\mathcal{T}_{2,i}) \prod_{j\neq i,1\leq j \leq D_2} (1 + \lambda w(\mathcal{T}_{2,j})S(\mathcal{T}_{2,j}))\right] \\
&+ w(\mathcal{H}^{m-1,l})M_t(\mathcal{H}^{m-1,l}) \prod_{j=1}^{D_2} (1 + \lambda w(\mathcal{T}_{2,j})S(\mathcal{T}_{2,j}))
\end{align*}

and similarly for \(\Sigma_m\).
Proof of Lemma 7.4. We shall reuse the notations in the proof of Lemma 7.1. We denote the corresponding expected total number of infections at depth-$l$ leaves of the processes $(X_{2,t}^\varphi)$ and $(X_{2,t}^{\varphi}_{\infty})$ by $M^1_2$ and $M^2_2$.

Similar to the derivation of [25], we observe that after the recovery of the first infection from $v_2$ in $(X_{2,t}^\varphi)$, the number of excursions of $(X_{2,t}^{\varphi}_{\infty})$ included in a single excursion of $(X_{2,t}^\varphi)$ is the same as a geometric random variable with success probability $(1 + \lambda(D_2 + 2)\theta_2^\varphi)^{-1}$. Therefore, the expected number of infections at depth-$l$ leaves after time $t_1$ is

$$\lambda(D_2 + 2)\theta_2^\varphi M^2_2(\hat{H}^{m,l}).$$

By the same argument as in [56], we bound from above the expected number of infections at depth-$l$ leaves of $(X_t)$ during $[0, t_1]$ by

$$\frac{\lambda w_m(\hat{H}^{m,l})}{1 + \lambda(D_2 + 2)} M_{l,m}(\hat{H}^{m,l}) + \theta_2^\varphi M^2_2.$$

Now we attempt to control $M^\varphi(\hat{H}^{m,l})$ in terms of $(M(T_{2,i}))$ and $M(\hat{H}^{m-1,l})$ where $\hat{H}^{m-1,l} \sim \text{GWC}^2(\xi, m - 1)^l$ contains the cycle $v_2, v_3, \ldots, v_m$ and the trees $T_i$ attached to $v_3, \ldots, v_m$. Let $L$ be the collection of depth-$l$ leaves of $\hat{H}^{m,l}$. We observe that

$$\lim_{t_0 \to \infty} \frac{1}{t_0} \sum_{v \in L} \mathbb{E}_{CP} \left[ \left\{ s \in [0, t_0] : X_{2,s}^\varphi(v) = 1^* \text{ and } X_{2,s-}^\varphi(v) = 0 \right\} \right] = \frac{M^\varphi(\hat{H}^{m,l})}{(\lambda(D_2 + 2)\theta_2^\varphi)^{-1} + S^\varphi(\hat{H}^{m,l})}.$$

On the other hand, by splitting $L$ into parts corresponding to $T_{2,i}$ and $\hat{H}^{m-1,l}$, the above is also equal to

$$\sum_{i=1}^{D_2} \frac{M_{l-1}(T_{2,i})}{(\lambda w(T_{2,i}))^{-1} + S(T_{2,i})} + \frac{M_{l}(\hat{H}^{m-1,l})}{(\lambda w(\hat{H}^{m-1,l}))^{-1} + S(\hat{H}^{m-1,l})}.$$

Combining these equations and [57], we get

$$\begin{align*}
(D_2 + 2)\theta_2^\varphi M^\varphi(\hat{H}^{m,l}) &= \left[ 1 + \lambda w(\hat{H}^{m-1,l})S(\hat{H}^{m-1,l}) \right] \left[ \sum_{i=1}^{D_2} w(T_{2,i})M_{l-1}(T_{2,i}) \prod_{j \neq i, 1 \leq j \leq D_2} (1 + \lambda w(T_{2,j})S(T_{2,j})) \right] \\
&\quad + w(\hat{H}^{m-1,l})M_{l}(\hat{H}^{m-1,l}) \prod_{j=1}^{D_2} (1 + \lambda w(T_{2,j})S(T_{2,j})) = \Sigma_2.
\end{align*}$$

By this together with [58], [59], and $M^2(\hat{H}^{m,l}) \geq M_{l}(\hat{H}^{m,l})$, we conclude that

$$M_{1,2}(\hat{H}^{m,l}) \leq \frac{\lambda w_m(\hat{H}^{m,l})}{1 + \lambda(D_2 + 2)} M_{l,m}(\hat{H}^{m,l}) + \frac{1 + \lambda(D_2 + 2)}{D_2 + 2} \Sigma_2.$$

Thus,

$$w_2(\hat{H}^{m,l})M_{1,2}(\hat{H}^{m,l}) \leq \frac{1}{41} w_m(\hat{H}^{m,l}) M_{l,m}(\hat{H}^{m,l}) + \lambda \Sigma_2.$$
Adding this with the analog for $M_{i,m}(\hat{H}^{m,l})$ and rearranging, we obtain
\[ w(\hat{H}^{m,l})M_i(\hat{H}^{m,l}) \leq 2\lambda\Sigma_2 + 2\lambda\Sigma_m \]
as claimed. \qed

To conclude that $(1.5)^{l-1}w(\hat{H}^{m,l})M_i(\hat{H}^{m,l})$ has a good tail, we proceed as in the proof of Lemma 5.12. We set
\[ W_i = \max\{(1.5)^{l-1}w(\mathcal{T}_{2,i})M_{l-1}(\mathcal{T}_{2,i}), w(\mathcal{T}_{2,i})S(\mathcal{T}_{2,i})\}, \quad i = 1, \ldots, D_2 \]
and
\[ W_0 = \max\{(1.5)^{l-1}w(\hat{H}^{m-1,l})M_i(\hat{H}^{m-1,l}), w(\hat{H}^{m-1,l})S(\hat{H}^{m-1,l})\} \]
where we used the subscript 0 in place of $\hat{H}^{m-1,l}$ for notational convenience. Then, the quantity $\Sigma_2$ in Lemma 7.3 satisfies
\[ 2(1.5)^{l-1}\lambda\Sigma_2 \leq 3 \sum_{i=0}^{D_2} \lambda W_i \prod_{0 \leq j \leq D_2, j \neq i} (1 + \lambda W_j) \leq 3 \left( \prod_{i=0}^{D_2} (1 + 2\lambda W_i) - 1 \right). \]

We have by the induction hypothesis that $W_i$ has a good tail for all $0 \leq i \leq D_2$. By Lemma 5.5, $\lambda\Sigma_2$ has a strong tail. Similarly, $\lambda\Sigma_m$ also has a strong tail. And so does, $(1.5)^{l-1}w(\hat{H}^{m,l})M_i(\hat{H}^{m,l})$.

7.4. **Proof of Lemma 6.11 for $M_i(\hat{H}^{m,l})$.** For the tail of $M_i(\hat{H}^{m,l})$, we have the following recursive inequality which can be derived analogously as Lemma 5.3.

**Lemma 7.4.** Under the hypothesis of Lemma 6.11, we have
\[ w(\hat{H}^{m,l})M_i(\hat{H}^{m,l}) \leq w(\hat{H}^{m,l})M_i(\hat{H}^{m,l}) \prod_{j=1}^{D_1} (1 + \lambda w(T_{1,j})S(T_{1,j})) \]
\[ + \left[ 1 + \lambda w(\hat{H}^{m,l})S(\hat{H}^{m,l}) \right] \sum_{i=1}^{D_1} w(T_{1,i})M_{l-1}(T_{1,i}) \prod_{j \neq i, 1 \leq j \leq D_1} (1 + \lambda w(T_{1,j})S(T_{1,j})) \].

From here, the conclusion that $(1.5)^{l}w(\hat{H}^{m,l})M_i(\hat{H}^{m,l})$ has a good tail follows from the same argument as above.

**References**

[1] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *The Annals of Probability*, 18(4):1462–1482, 1990.

[2] Shankar Bhamidi, Danny Nam, Oanh Nguyen, and Allan Sly. Survival and extinction of epidemics on random graphs with general degrees. *The Annals of Probability*, 49(01):244–286, 2021.

[3] V. H. Can and B. Schapira. Metastability for the contact process on the configuration model with infinite mean degree. *Electronic Journal of Probability*, 20(26):1–22, 2015.

[4] S. Chatterjee and R. Durrett. Contact process on random graphs with power law degree distribution have critical value 0. *The Annals of Probability*, 37(6):2332–2356, 2009.

[5] M. Cranston, T. Mountford, J.-C. Mourrat, and D. Valesin. The contact process on finite trees revisited. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(2):385–408, 2014.

[6] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Brazilian Journal of Probability and Statistics*, 24(2):137–211, 2010.

[7] R. Durrett and X.-F. Liu. The contact process on a finite set. *The Annals of Probability*, 16(3):1158–1173, 1988.
[8] R. Durrett and R. H. Schonmann. The contact process on finite set. ii. *The Annals of Probability*, 16(3):1570–1583, 1988.
[9] Richard Durrett. *Random graph dynamics*, volume 200. Cambridge university press Cambridge, 2007.
[10] T. E. Harris. Contact interactions on a lattice. *The Annals of Probability*, 2(6):969–988, 1974.
[11] X. Huang and R. Durrett. The contact process on random graphs and galton-watson trees. *arXiv:1810.06040*, 2018.
[12] S. Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18(1-2):205–225, 2009.
[13] J. H. Kim. Poisson cloning model for random graphs, International Congress of Mathematicians. Vol. III. *Eur. Math. Soc.*, pages 873–897, 2006.
[14] S. Lalley and W. Su. Contact processes on random regular graphs. *Ann. Appl. Probab.*, 27(4):2061–2097, 2017.
[15] T. M. Liggett. Multiple transition points for the contact process on the binary tree. *The Annals of Probability*, 24(4):1675–1710, 1996.
[16] T. M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer-Verlag, New York, 1999.
[17] T. M. Liggett. *Interacting Particle Systems*. Springer, Berlin, 2005.
[18] L. Méndard and A. Singh. Percolation by cumulative merging and phase transition for the contact process on some random graphs. *Ann. Sci. École Norm. Supér.*, 49(5):1189–1238, 2016.
[19] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures Algorithms*, 6(2-3):161–179, 1995.
[20] Elchanan Mossel and Allan Sly. Rapid mixing of gibbs sampling on graphs that are sparse on average. *Random Structures & Algorithms*, 35(2):250–270, 2009.
[21] Elchanan Mossel and Allan Sly. Exact thresholds for ising–gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294–328, 2013.
[22] T. Mountford, J.-C. Mourrat, D. Valesin, and Q. Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Process. Appl.*, 126(7):1974–2013, 2016.
[23] T. Mountford, D. Valesin, and Q. Yao. Metastable densities for the contact process on power law random graphs. *Electronic Journal of Probability*, 18(103):1–36, 2013.
[24] T. S. Mountford. A metastable result for the finite multidimensional contact process. *Canad. Math. Bull.*, 36:216–226, 1993.
[25] J.-C. Mourrat and D. Valesin. Phase transition of the contact process on random regular graphs. *Electronic Journal of Probability*, 21(31):1–17, 2016.
[26] Danny Nam, Oanh Nguyen, and Allan Sly. Critical value asymptotics for the contact process on random graphs. *to appear in the Transactions of the American Mathematical Society*.
[27] R. Pemantle. The contact process on trees. *The Annals of Probability*, 20(4):2089–2116, 1992.
[28] B. Schapira and D. Valesin. Extinction time for the contact process on general graphs. *Probability Theory Related Fields*, 169(3-4):871–899, 2017.
[29] A. Stacey. The contact process on finite homogeneous trees. *Probability Theory Related Fields*, 121(4):551–576, 2001.
[30] A. M. Stacey. The existence of an intermediate phase for the contact process on trees. *The Annals of Probability*, 24(4):1711–1726, 1996.
[31] R. van der Hofstad. *Random Graphs and Complex Networks, Volume 1*. Cambridge Series in Statistical and Probabilistic Mathematics. [43]. Cambridge University Press, Cambridge, 2017.

**Division of Applied Mathematics, Brown University, Providence, RI 02906**

*Email address: oanh.nguyen1@brown.edu*

**Department of Mathematics, Princeton University, Princeton, NJ 08544**

*Email address: asly@princeton.edu*