CRITICAL VALUE ASYMPTOTICS FOR THE CONTACT PROCESS ON RANDOM GRAPHS

DANNY NAM, OANH NGUYEN, AND ALLAN SLY

ABSTRACT. Recent progress in the study of the contact process (see Shankar Bhamidi, Danny Nam, Oanh Nguyen, and Allan Sly [Ann. Probab. 49 (2021), pp. 244–286]) has verified that the extinction-survival threshold λ_1 on a Galton-Watson tree is strictly positive if and only if the offspring distribution ξ has an exponential tail. In this paper, we derive the first-order asymptotics of λ_1 for the contact process on Galton-Watson trees and its corresponding analog for random graphs. In particular, if ξ is appropriately concentrated around its mean, we demonstrate that $\lambda_1(\xi) \sim 1/\mathbb{E}\xi$ as $\mathbb{E}\xi \to \infty$, which matches with the known asymptotics on d-regular trees. The same results for the short-long survival threshold on the Erdős-Rényi and other random graphs are shown as well.

CONTENTS

1. INTRODUCTION

The contact process is a model of the spread of disease. For a graph $G = (V, E)$, the contact process on G with infection rate λ and recovery rate 1 is a continuous-time Markov chain, in which 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 rate λ and is healed at rate 1.

• Infection and recovery events in the process happen independently.

The contact process was first introduced by a work of Harris [7] 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.

Denote the extinction-survival threshold by

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

and the weak-strong survival threshold by

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

For the lattice, when the origin is initially infected, it is well known that there is no weak survival phase, that is, $0 < \lambda_1(\mathbb{Z}^d) = \lambda_2(\mathbb{Z}^d) < \infty$ (see [7], Bezuidenhout-Grimmett [2], and also the books of Liggett [13], [14] and the references therein). On the other hand, for the infinite d-regular tree $\mathbb{T}_d$ with $d \geq 3$, we have that 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 work by Pemantle [18] (for $d \geq 4$), Liggett [12] (for $d = 3$), and Stacey [20] (for a shorter proof that works for all $d \geq 3$). Moreover, we have from [18] that

$$(1.1) \qquad \frac{1}{d-1} \leq \lambda_1(\mathbb{T}_d) < \frac{1}{d-2} \quad \text{and} \quad \frac{1}{2\sqrt{d-1}} \leq \lambda_2(\mathbb{T}_d).$$

In particular, the first-order asymptotics of $\lambda_1(\mathbb{T}_d)$ is $1/d$ as d becomes large.

Much less is known about the contact process on random trees. First of all, for a Galton-Watson tree $\mathcal{T}$ with offspring distribution ξ, it is not difficult to see that $\lambda_1(\mathcal{T})$ and $\lambda_2(\mathcal{T})$ are constants which are the same for a.e. $\mathcal{T} \sim \mathsf{GW}(\xi)$ conditioned on $|\mathcal{T}| = \infty$, and hence the constants $\lambda_1^{\mathrm{GW}}(\xi)$ and $\lambda_2^{\mathrm{GW}}(\xi)$ are well-defined. Huang and Durrett [8] proved that on $\mathcal{T} \sim \mathsf{GW}(\xi)$ with the root initially infected, $\lambda_2^{\mathrm{GW}}(\xi) = 0$ if the offspring distribution ξ 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 ξ has an exponential tail, i.e., $\mathbb{E}e^{c\xi} < \infty$ for some $c > 0$, Bhamidi and the authors [3] showed that there is an extinction phase: $\lambda_1^{\mathrm{GW}}(\xi) \geq \lambda_0(\xi)$ for some constant $\lambda_0(\xi) > 0$. Our first main result derives the first-order asymptotics on $\lambda_1^{\mathrm{GW}}(\xi)$ for ξ *concentrated* around its mean, which turns out to have the same form as (1.1).

Theorem 1. *Let $\{\xi_k\}$ be a sequence of nonnegative integer-valued random variables with $\mathbb{E}\xi_k =: d_k \to \infty$ as $k \to \infty$. Assume that there exists a collection of positive constants $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ for which*

$$(1.2) \qquad \begin{aligned} &\mathbb{P}(\xi_k \geq (1+\delta)d_k) \leq \exp(-c_\delta d_k) \text{ for all } \delta \in (0,1); \\ &\mathbb{P}(\xi_k \geq (1+a)d_k) \leq \exp(-c_1 a d_k) \text{ for all } a \geq 1, \end{aligned}$$

for all large enough k. Consider the contact process on the Galton-Watson tree $\mathcal{T}_k \sim \mathsf{GW}(\xi_k)$ with the root initially infected. Then, the extinction-survival threshold

$\lambda_1^{\mathrm{GW}}(\xi_k)$ *satisfies*

$$\lim_{k\to\infty} \lambda_1^{\mathrm{GW}}(\xi_k)d_k = 1. \tag{1.3}$$

Remark 1.1. The concentration condition (1.2) we impose resembles the form of large deviation estimates. Notice that the family of Poisson distributions $\{\mathrm{Pois}(d)\}_{d>0}$ satisfies (1.2). This fact allows us to deduce an analogous result for Erdős-Rényi random graphs in Corollary 3.

Remark 1.2. The conclusion (1.3) is not always true without assuming (1.2). In fact, for any given $d > 0$, we can construct ξ with $\mathbb{E}\xi = d$ and arbitrarily small $\lambda_1^{\mathrm{GW}}(\xi) > 0$ (compared to $1/d$), by truncating a heavy-tailed distribution at a large enough degree. Understanding the size of $\lambda_1^{\mathrm{GW}}(\xi)$ in full generality seems to be a challenging problem.

Remark 1.3. In [23], Xue showed (1.3) for $\xi_k = \mathbf{Bin}(k,p)$ with a fixed constant $p > 0$. Their result works for the case where the recovery and infection rates are given by i.i.d. random variables for the contact process on a regular tree. See, e.g., [19, 22, 24] for more works from this perspective.

In [3], the proof of $\lambda_1^{\mathrm{GW}}(\xi) > 0$ introduced a new method recursive analysis on Galton-Watson trees that controlled the expected survival times, but the quantitative lower bounds on λ_1 they deduced were far from being sharp. Our main contribution is to introduce an alternative tree recursion, and develop techniques to control the tail probabilities of the survival time over the level of Galton-Watson trees in addition to its expectation. A detailed outline is given in Section 3.1.

A naturally related object is the contact process on a random graph with a given degree distribution. Let μ be a degree distribution and $G_n \sim \mathcal{G}(n,\mu)$ be a random graph with degree distribution μ, assuming the giant component condition (2.2) (for details, see Section 2.2). Consider the contact process on $G_n \sim \mathcal{G}(n,\mu)$ where all vertices are initially infected. In [3], it was shown that if $\mathbb{E}_{D\sim\mu}e^{cD} < \infty$ for some constant $c > 0$, then there exist constants $0 < \underline{\lambda}(\mu) \le \overline{\lambda}(\mu) < \infty$ such that the survival time $T_{\lambda,n}$ of the process satisfies the following:

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

On the other hand if μ has a subexponential tail, they proved that whp there is no short survival phase. Based on this result, we formally define the short- and long-survival thresholds $\lambda_c^-(\mu)$, $\lambda_c^+(\mu)$ as follows.

$$\lambda_c^-(\mu) := \sup\left\{\lambda : \lim_{n\to\infty}\mathbb{P}(T_{\lambda,n} \le n^{C(\lambda)}) = 1 \text{ for some } C(\lambda) < \infty\right\}; \tag{1.4}$$
$$\lambda_c^+(\mu) := \inf\left\{\lambda : \lim_{n\to\infty}\mathbb{P}(T_{\lambda,n} \ge e^{c(\lambda)n}) = 1 \text{ for some } c(\lambda) > 0\right\}.$$

The second result of the paper verifies the first-order asymptotics for $\lambda_c^-(\mu)$ and $\lambda_c^+(\mu)$ which have the same form as Theorem 1.

Theorem 2. *Let (μ_k) be a sequence of degree distributions with the size-biased distribution $\widetilde{\mu}_k$ (see (2.1) for its precise definition). Suppose that the mean d_k of $\widetilde{\mu}_k$ tends to infinity as $k \to \infty$. Moreover, assume that $(\widetilde{\mu}_k)$ satisfies the concentration condition (1.2) for fixed positive constants $\mathfrak{c} = \{c_\delta\}_{\delta\in(0,1]}$. Then, the short- and*

long-survival thresholds $\lambda_c^-(\mu_k)$, $\lambda_c^+(\mu_k)$ of the contact process on $G_n \sim \mathcal{G}(n, \mu_k)$ satisfy

$$\lim_{k \to \infty} \lambda_c^-(\mu_k) d_k = \lim_{k \to \infty} \lambda_c^+(\mu_k) d_k = 1.$$

A proof of $\lambda_c^-(\mu) > 0$ for μ with an exponential tail was given in [3], which relied on estimating the probability of having an infection deep inside Galton-Watson trees which are local weak limits of the random graphs—see Section 2.2 for details. However, we will see in Section 3.2 that controlling such an event is insufficient for Theorem 2 since λ is not small enough. To overcome this issue, we first take the expectation of the event over the level of the contact process, and then study its tail probability over the level of Galton-Watson trees. This new approach turns out to provide a substantial improvement from [3]. On the other hand, we will soon see the generalized result on λ_c^+ in Theorem 5, and it requires a different approach in spreading infections and an improved structural analysis of random graphs than in [3], which we overview in Section 3.3.

We can deduce an analog of Theorem 2 for the Erdős-Rényi random graph, arguably one of the most well-known models of random graphs, based on the contiguity between $\mathcal{G}_{\mathrm{ER}}(n, d/n)$ and $\mathcal{G}(n, \mathrm{Pois}(d))$ (see Section 2.2 for details).

Corollary 3. *Consider the contact process on the random graph $G_n \sim \mathcal{G}_{\mathrm{ER}}(n, d/n)$ with all vertices initially infected. Then, the short- and long- survival thresholds of the process, defined analogously as (1.4), satisfy*

$$\lim_{d \to \infty} \lambda_{\mathrm{ER}}^-(d) d = \lim_{d \to \infty} \lambda_{\mathrm{ER}}^+(d) d = 1.$$

As mentioned in [3], we actually expect the transition from polynomial- to exponential-time survival is sharp and happens at the extinction-survival threshold of the corresponding Galton-Watson tree. Namely,

Conjecture 4. *Let μ be a degree distribution satisfying the giant component condition (2.2), and let $\widetilde{\mu}$ be its size-biased distribution (definition given in (2.1)). Recalling (1.4), we have*

$$\lambda_c^-(\mu) = \lambda_c^+(\mu) = \lambda_1^{\mathrm{GW}}(\widetilde{\mu}).$$

The special case of random regular graphs ($\mu \equiv d$) was established by Mourrat-Valesin [17] and Lalley-Su [11] who showed that the short-long survival threshold for random d-regular graph occurs exactly at $\lambda_1(\mathbb{T}_d)$.

The next result establishes one inequality of the conjecture, by showing that the intensity which gives a supercritical contact process on a Galton-Watson tree implies an exponential time survival on the corresponding random graph.

Theorem 5. *Let μ be a degree distribution satisfying the giant component condition (2.2), and let $\widetilde{\mu}$ be its size-biased distribution with $d := \mathbb{E}_{D \sim \widetilde{\mu}} D$. Recalling (1.4), we have*

$$(1.5) \qquad \lambda_c^+(\mu) \leq \lambda_1^{\mathrm{GW}}(\widetilde{\mu}) \leq \frac{1}{d-1}.$$

In the case of $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$, one may ask if the survival time of the contact process on $G_n \sim \mathcal{G}(n, \mu)$ is still exponentially long, even with a single initial infection. The following theorem gives an affirmative answer to this question.

Theorem 6. *For all* $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$, *whp in* $G_n \sim \mathcal{G}(n, \mu)$, *the contact process on* G_n *starting from a single infection at a site chosen uniformly at random survives until time* $e^{\Theta(n)}$ *with positive probability. Moreover, the same holds for the Erdős-Rényi random graph* $G_n \sim \mathcal{G}_{\mathrm{ER}}(n, d/n)$, *when* $\lambda > \lambda_1^{\mathrm{GW}}(Pois(d))$.

2. Preliminaries

In this section, we set up the notations and review preliminary concepts on the contact process and random graphs.

2.1. The contact process and its graphical representation.

Let $G = (V, E)$ be a graph (finite or infinite) and $A \subset V$. The configuration space of infections is $\{0, 1\}^V$, and $\mathbf{1}_A \in \{0, 1\}^V$ is the configuration of which the vertices in A are infected and the others are healthy. We denote the contact process with initial infections at A by

$$(X_t) \sim \mathsf{CP}^\lambda(G; \mathbf{1}_A),$$

where λ is the intensity of infection and $\mathbf{1}_A \in \{0, 1\}^V$ describes the initial condition. We sometimes write $\mathsf{CP}^\lambda(G)$ when the initial condition is irrelevant in the context. Also, $\mathbf{0}, \mathbf{1}$ denote all-healthy, all-infected configurations, respectively, and we write $\mathbf{1}_v$ for $\mathbf{1}_{\{v\}}$. The transition rule of the continuous-time Markov chain $(X_t) \sim \mathsf{CP}^\lambda(G; \mathbf{1}_A)$ can be defined as follows:

- X_t becomes $X_t - \mathbf{1}_v$ with rate 1 for each v such that $X_t(v) = 1$.
- X_t becomes $X_t + \mathbf{1}_u$ with rate $\lambda N_t(u)$ for each u with $X_t(u) = 0$, where $N_t(u) := \sum_{v \sim u} X_t(v)$ denotes the number of infected neighbors of u at time t.

The dynamics of the contact process can be interpreted by the *graphical representation* which provides a convenient coupling of the process. We briefly discuss this notion following Chapter 3, Section 6 of [14]. Let $\{N_v(t)\}_{v \in V}$ (resp., $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \vec{E}}$) be the family of independent Poisson processes with rate 1 (resp., rate λ), where $\vec{E} = 2E := \{\vec{uv}, \vec{vu} : (uv) \in E\}$ is the set of directed edges. We set $\{N_v(t)\}_{v \in V}$ to be independent of $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \vec{E}}$ as well. Note that all event times of the Poisson processes are distinct almost surely. We generate the graphical representation of $\mathsf{CP}^\lambda(G; \mathbf{1}_A)$ as follows:

1. Initially, we have the empty space-time domain $V \times \mathbb{R}_+$.
2. For each $v \in V$, place symbol $\times$ at (v, t), for each event time t of N_v. The symbol $\times$ describes the time when v gets recovered.
3. For each $\vec{uv} \in \vec{E}$, place an arrow from (u, t) to (v, t), for each event time t of $N_{\vec{uv}}$. The arrow indicates that the infection is passed from u to v at time t if $X_t(u) = 1$.

Therefore, as described in Figure 1, we can read off the diagram starting from the bottom horizontal line and obtain (X_t). Construction of the graphical representation will play an important role in Sections A.2 and 6.3 when we introduce a *decomposition* of the contact process.
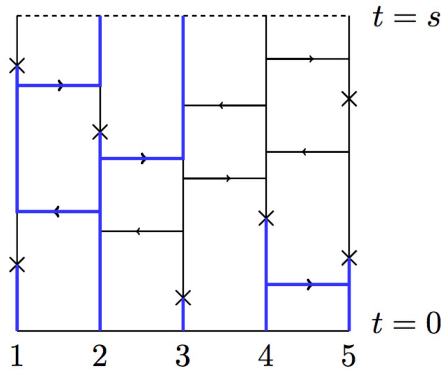
FIGURE 1. An instance of the graphical representation of the contact process on $V = \{1, \ldots, 5\}$ with initial configuration $\mathbf{1}_V$. Note that $X_s = \mathbf{1}_{\{2,3\}}$.

2.2. Random graphs and their limiting structure.

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

- Let $d_i \sim$ i.i.d. μ 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* $G_n^{\mathsf{u}} \sim \mathcal{G}^{\mathsf{u}}(n, \mu)$, which picks a uniformly random simple graph among all simple graphs with degree sequence $\{d_i\}_{i \in [n]} \sim$ i.i.d. μ. It is well-known that if μ has a finite second moment, then the two laws $\mathcal{G}(n, \mu)$ and $\mathcal{G}^{\mathsf{u}}(n, \mu)$ are *contiguous*, in the sense that for any subset A_n of graphs with n vertices,

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

For details, we refer the reader to Chapter 7 of [21] or to [9]. 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 $G_n^{\text{ER}} \sim \mathcal{G}_{\text{ER}}(n, d/n)$ as shown in [10, 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 μ by $\mathsf{GW}(\mu)$, and let $\mathsf{GW}(\mu)^l$ be the law of $\mathsf{GW}(\mu)$ truncated at depth l, that is, the vertices with distance $> l$ from the root are removed. Further, let $\widetilde{\mu}$ denote the *size-biased* distribution of μ, defined by

$$(2.1) \qquad \widetilde{\mu}(k-1) := \frac{k\mu(k)}{\sum_{k'=1}^{\infty} k'\mu(k')}, \quad k = 1, 2, \ldots.$$

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

Lemma 2.1 ([4, Section 2.1]). *Suppose that μ 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*

$$\lim_{n \to \infty} \mathbb{P}((N(v, l), v) \cong (T, x)) = \mathbb{P}_{(\mathcal{T}, \rho) \sim \mathbf{GW}(\mu, \widetilde{\mu})^l}((\mathcal{T}, \rho) \cong (T, x)),$$

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 $\mathbf{GW}(\mu, \widetilde{\mu})$.

We remark that the same holds for a fixed vertex $v \in G_n$. Moreover, we also stress that the condition for $\mathbf{GW}(\mu, \widetilde{\mu})$ to be supercritical is equivalent to the condition for $\mathcal{G}(n, \mu)$ to have the unique giant component whp (see e.g., [16], or [5, Section 3] for details), which can be addressed as

$$(2.2) \qquad \qquad \mathbb{E}_{D \sim \mu} D(D - 2) > 0.$$

2.3. **Notations.** 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^l and $T_{\leq l}$ to denote the set of vertices of depth at most l. In particular, $\mathcal{T}^l \sim \mathbf{GW}(\xi)^l$ denotes the Galton-Watson tree generated up to depth l, while the infinite Galton-Watson tree is denoted by $\mathcal{T} \sim \mathbf{GW}(\xi)$.

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:

- $\mathbb{P}_{\mathrm{CP}}$ and $\mathbb{E}_{\mathrm{CP}}$ denote the probability and the expectation, respectively, with respect to the randomness from contact processes.
- $\mathbb{P}_{\mathrm{GW}}$ and $\mathbb{P}_{\mathrm{RG}}$ denote the probability with respect to the randomness from the underlying graph, when the graph is a Galton-Watson tree and a random graph $\mathcal{G}(n, \mu)$, respectively. We write $\mathbb{E}_{\mathrm{GW}}$ and $\mathbb{E}_{\mathrm{RG}}$ 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, $\mathbb{E}[\cdot] = \mathbb{E}_{\mathrm{GW}}[\mathbb{E}_{\mathrm{CP}}[\cdot]]$, if the underlying graph is a Galton-Watson tree.

3. MAIN CONCEPTS AND IDEAS

Let us start by emphasizing that even though we borrow some ideas and notations from [3], this manuscript is self-contained and so the reader does not have to be familiar with [3] to read this manuscript. In this section, we briefly introduce the primary notions and discuss the main ideas in the paper. We also address the organization of the rest of the article in Section 3.4.

3.1. **The root-added process and the lower bound of Theorem 1.** In [3], Bhamidi and the authors studied the *root-added contact process* to prove $\lambda_1 > 0$ on $\mathbf{GW}(\xi)$ with ξ having an exponential tail. This notion continues to play a huge role in the current work as well, and hence we begin with explaining its definition and the concept of *excursion time*.

Definition 3.1 (Root-added contact process, [3]). Let T be a (finite or infinite) tree rooted at ρ. Let T^+ be the tree that has a parent vertex ρ^+ of ρ which is connected only with ρ. The *root-added contact process* on T is the continuous-time Markov chain on the state space $\{0, 1\}^T$, defined as the contact process on T^+ with ρ^+ set to be infected permanently (hence we exclude ρ^+ from the state space). That

is, ρ^+ is infected initially, and it does not have a recovery clock attached to itself. Let $\mathbf{CP}_{\rho^+}^{\lambda}(T;x_0)$ denote the root-added contact process on T with initial condition $x_0 \in \{0,1\}^T$. Note that the root-added contact process no longer has an absorbing state.

Definition 3.2 (Survival and excursion times). Let T be a (finite or infinite) tree rooted at ρ. The *survival time* and *excursion time* on T, denoted by $\mathbf{R}(T)$ and $\mathbf{S}(T)$, respectively, are defined as follows:

- $\mathbf{R}(T)$ is the first time when the contact process $\mathbf{CP}^{\lambda}(T;\mathbf{1}_\rho)$ is all-healthy (i.e., when the process terminates). We also denote the *expected survival time* by $R(T) = \mathbb{E}_{\mathrm{CP}}\mathbf{R}(T)$.
- $\mathbf{S}(T)$ is the first time when the root-added contact process $\mathbf{CP}_{\rho^+}^{\lambda}(T;\mathbf{1}_\rho)$ becomes all-healthy on T. We also denote the *expected excursion time* by $S(T) = \mathbb{E}_{\mathrm{CP}}\mathbf{S}(T)$.

Note that the quantities $R(T)$ and $S(T)$ are fixed numbers for each tree (T,ρ) and satisfy $R(T) \leq S(T)$, which can be seen through the coupling via the graphical representation.

The previous work [3] established $\lambda_1 > 0$ by a recursive inequality on the depth of the tree that showed $\mathbb{E}\mathbf{S}(\mathcal{T}) < \infty$ for $\mathcal{T} \sim \mathbf{GW}(\xi)$ with ξ having an exponential tail decay. However, the argument had limitations since it could only deal with a small enough λ. In order to push its applicability to near-criticality, in Section 4 we introduce another recursive inequality based on fundamental properties of the contact process. Using the two different recursions, we can bound the tail probabilities of the *expected* excursion time, namely, $\mathbb{P}_{\mathrm{GW}}(S(\mathcal{T}) \geq t)$, and then the bound easily implies $\mathbb{E}_{\mathrm{GW}}S(\mathcal{T}) < \infty$. This is a substantial improvement from [3] where we could only control its expectation $\mathbb{E}_{\mathrm{GW}}S(\mathcal{T})$ for small enough λ.

3.2. **Deep infections, unicyclic neighbors and the lower bound of Theorem 2.** To establish Theorem 2, we attempt to generalize Theorem 1 based on the fact that the local neighborhoods of $\mathcal{G}(n,\mu)$ look like Galton-Watson trees. There are two major obstacles on carrying out this idea.

1. For a vertex v in $G \sim \mathcal{G}(n,\mu)$, its local neighborhood $N(v,l)$ contains a lot of cycles if $l \geq c_\mu \log n$ for some constant c_μ.
2. Even for small l, there are $o(n)$ vertices that contain a cycle in $N(v,l)$.

3.2.1. *Deep infections.* To overcome the first issue, we show that the probability of having a deep infection of depth $\geq c_\mu \log n$ inside $\mathbf{CP}^{\lambda}(\mathcal{T}, \mathbf{1}_\rho)$ is very small for a Galton-Watson tree $\mathcal{T}$. This leads to the consideration of the *total infections at leaves* of (finite) trees defined as follows.

Definition 3.3 (Total infections at leaves). Let T be a finite tree rooted at ρ, set $l := \max\{\mathrm{dist}(\rho,v) : v \in T\}$ be the depth of the tree and

$$\mathcal{L} := \{v \in T : \mathrm{dist}(\rho,v) = l\}$$

be the collection of depth-l leaves of T. Suppose that $l \geq 1$ and consider the root-added contact process $(X_t) \sim \mathbf{CP}_{\rho^+}^{\lambda}(T;\mathbf{1}_\rho)$. For $v \in \mathcal{L}$, define the *total infections at v*, by

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

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

$$\mathbf{M}^l(T) = \sum_{v \in \mathcal{L}} \mathbf{M}_v^l(T).$$

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

We also denote the *expected total infections at depth-l leaves* by $M^l(T) = \mathbb{E}_{\mathrm{CP}}\mathbf{M}^l(T)$. Also, as above, we write $M^{l'}(T) = 0$ for $l' > l$. Moreover, if the tree depth is 0 (that is, T is a single vertex), we set $\mathbf{M}^0(T) \equiv 1$.

The previous work [3] derived an exponential decay of $\mathbf{M}^l(\mathcal{T}^l)$ in l for $\mathcal{T}^l \sim \mathsf{GW}(\widetilde{\mu})^l$ to deal with the same issue, but as before it had to require λ to be small enough. However, unfortunately, the decay of $\mathbf{M}^l(\mathcal{T}^l)$ is insufficient for our purpose if λ is close to λ_1, due to the reason we explain below.

If $\lambda = (1-\varepsilon)d^{-1}$ with $d = \mathbb{E}_{D \sim \widetilde{\mu}}D$, then for an infected vertex v in $\mathcal{T}^l \sim \mathsf{GW}(\widetilde{\mu})^l$, the expected number of offsprings of v that get infected before v becomes healthy is

$$\frac{\lambda d}{\lambda + 1} \approx \frac{(1-\varepsilon)d}{d+1}, \quad \text{hence} \quad \mathbb{E}[\mathbf{M}^l(\mathcal{T}^l)] \geq \left(\frac{(1-\varepsilon)d}{d+1} \right)^l,$$

and intuitively, the latter quantity essentially corresponds to $\mathbb{P}(\mathbf{M}^l(\mathcal{T}^l) \geq 1)$. To apply a union bound over all vertices in $G_n \sim \mathcal{G}(n, \mu)$, we need this probability to be of order $o(n^{-1})$. That is, we roughly require

$$l \gg \frac{\log n}{\log\{(1+\varepsilon)(1+d^{-1})\}}.$$

This is much larger than our budget $c_\mu \log n$ which is approximately $c \log_d n$. Thus, investigating the tail probability of $\mathbf{M}^l(\mathcal{T}^l)$ is not enough for our purpose. However, in [3], this approach was sufficient since we could set λ as small as we wanted.

In Section 5, we instead focus on studying $\mathbb{P}_{\mathrm{GW}}(M^l(\mathcal{T}^l) \geq t)$, which turns out to have a much better bound than the tail of $\mathbf{M}^l(\mathcal{T}^l)$. Similarly as explained in Section 3.1, we derive two different recursive inequalities on $M^l(T)$ for a deterministic tree T, and prove its tail bound for the case of Galton-Watson trees.

3.2.2. *Unicyclic neighbors.* Another major issue is to deal with the neighborhoods $N(v, l)$ in $G_n \sim \mathcal{G}(n, \mu)$ containing a cycle. We rely on idea as [3], by observing that if μ satisfies (1.2), then there exists $\gamma(\mathfrak{c}) > 0$ such that whp, $N(v, \gamma \log_d n)$ contains at most one cycle for all v with d as in the previous subsection (see Lemma 6.11). Therefore, we study $S(T')$ and $M^l(T')$ as above (precise definitions are given in Section 6.1), for certain unicyclic graphs T' which are closely related to Galton-Watson trees. To this end, we appropriately cover T' by trees and deduce information on $S(T')$ and $M^l(T')$ from the results we obtained on trees. However, formalizing this idea requires a heavy technical work and it is presented in Appendix A. Similar ideas are applied to studying the contact process on G_n. Roughly speaking, we decompose G_n by its local neighborhoods $\{N(v, l)\}_{v \in G_n}$, and derive results on $\mathbf{CP}^\lambda(G_n)$ by using what we know on $N(v, l)$.

3.3. **The proof of Theorems 5 and 6.** The previous work [3] settled that $\lambda_c^+(\mu) < \infty$, which was based on a challenging structural analysis on the configuration model. Roughly speaking, they showed the existence of an *embedded expander*, a subset of large degree vertices in the random graph, on which it is easy to send infections from one vertex to another. Upon establishing its existence, spreading infections on the embedded expander could then be done by a relatively straight-forward way, which was to infect a site at distance l with probability $(c(1 - e^{-\lambda}))^l$, since we could choose λ to be large. One of the main difficulties in establishing the much improved bound $\lambda_c^+(\mu) \leq \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$ is that we need to develop a more efficient way of sending infections from a vertex to another.

The key observation for such improvement is that if $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$, the expected number of infections on $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu})$ grows exponentially in time (Lemmas 7.5 and 7.6). We use this property as our driving force of passing infections on the random graph, which is possible since the local neighborhoods look like Galton-Watson trees. This new method turns out to be substantially better than the aforementioned approach.

However, since we now need to reveal the neighborhoods to check if the infections spread well, the structural analysis on the random graph becomes even more involved than the previous proof in [3]. We carry out by introducing an appropriate notion of *good vertices*, which roughly refer to the sites that are capable of propagating enough infections around them, and showing that any set of δn infected *good vertices* causes $\geq 2\delta n$ *good vertices* to be infected at a later time with high probability except for an exponentially small error.

3.4. **Organization of the article.** Sections 4–6 are devoted to the derivation of the lower bounds of Theorems 1 and 2. In Section 4, we introduce the basic form of the recursion argument on Galton-Watson trees and establish the lower bound of Theorem 1. In Section 5, we extend the recursion criterion to the study of deep infections. Section 6 then concludes the proof of the lower bound of Theorem 2, while the technical works needed to study the unicyclic graphs are deferred to Appendix A. Finally, we finish the proof of Theorems 1 and 2 by settling their upper bounds in Section 7.

4. Survival and excursion times on trees

In this section, we introduce primary recursive argument on the expected excursion time which are used throughout the paper. In Section 4.1, we review some ideas developed in [3], and derive another recursive inequality on excursion times. In Sections 4.2 and 4.3, we prove a tail probability estimate and establish Theorem 1 as its application.

4.1. **Deterministic recursions on trees.** Let T be a finite tree rooted at ρ and recall the definition of the root-added contact process $\mathsf{CP}_{\rho^+}^\lambda(T^+; x_0)$ (Definition 3.1). Let $R(T)$ and $S(T)$ be expected survival and recursion time as in Definition 3.2.

In (T, ρ), let $D = \deg(\rho)$ and $v_1, \ldots, v_D$ be the children of ρ. Further, let $T_1, \ldots, T_D$ be the subtrees from each child of ρ, rooted at $v_1, \ldots, v_D$, respectively. In [3], we proved the following recursion on the excursion times.

Proposition 4.1 ([3, Lemma 3.3]). *Let T and $T_1, \ldots, T_D$ be as above. Then, the expected excursion times $S(T)$ and $S(T_1), \ldots, S(T_D)$ satisfy*

$$(4.1) \qquad S(T) \leq \prod_{i=1}^{D} (1 + \lambda S(T_i)).$$

Even though the proof can be found in [3], we briefly explain it again, mainly because the ideas will be revisited in Proposition 5.2. For a detailed proof, we refer to [3].

Proof. Consider $\mathbf{CP}_\rho^\lambda(T_i; \mathbf{1}_{v_i})$ (the subscript ρ indicates that it serves as the added parent above v_i), the root-added contact process on each T_i. We define $\mathbf{CP}^\otimes(T; \mathbf{1}_{v_i})$ as the contact process on T such that

- The root ρ is infected permanently, i.e., all the recoveries at ρ are ignored.
- Initially, the vertex v_i is infected, and it is the only infected vertex other than ρ.

Note that $\mathbf{CP}^\otimes(T; \mathbf{1}_{v_i})$ can be written as the following form:

$$(4.2) \qquad \mathbf{CP}^\otimes(T; \mathbf{1}_{v_i}) := \left(\otimes_{\substack{j=1 \\ j \neq i}}^{D} \mathbf{CP}_\rho^\lambda(T_j; \mathbf{0}) \right) \otimes \mathbf{CP}_\rho^\lambda(T_i; \mathbf{1}_{v_i}),$$

for each $i \in [D]$. Let $\mathbf{S}_i^\otimes$ denote the excursion time of this process, that is, the first return time to the all-healthy state $\otimes_{j=1}^{D} \mathbf{0}_{T_j}$, and let $S_i^\otimes = \mathbb{E}_{\mathrm{CP}} \mathbf{S}_i^\otimes$. Further, define the average of $S_i^\otimes$ by

$$(4.3) \qquad S^\otimes = \frac{1}{D} \sum_{i=1}^{D} S_i^\otimes.$$

Then, we can control $S(T)$ by $S^\otimes$, based on the following modification of the process $\mathbf{CP}_{\rho+}^\lambda(T; \mathbf{1}_\rho)$ introduced in [3, Lemma 3.3].

- Consider the process $(X_t^\sharp)$ on T that follows the same transition rule as $\mathbf{CP}_{\rho+}^\lambda(T; \mathbf{1}_\rho)$, except for the recoveries at root ρ.
- An independent rate-1 Poisson clock is associated with ρ, and the recovery at ρ is only valid if $X_t^\sharp = \mathbf{1}_\rho$ when the clock rings at time t.

In other words, $(X_t^\sharp)$ is generated by ignoring the recoveries of $(X_t) \sim \mathbf{CP}_{\rho+}^\lambda(T; \mathbf{1}_\rho)$ at ρ if there is another infected vertex at the time of recovery. Let us denote the expected excursion time of this process by $S^\sharp(T)$. Recalling the coupling argument using the graphical representation (Section 2.1), we know that $S^\sharp(T) \geq S(T)$. Moreover, an excursion of $(X_t^\sharp)$ can be described as follows.

1. Initially $X_0^\sharp = \mathbf{1}_\rho$, and we terminate if ρ gets healed before infecting any of its children. Otherwise, suppose that the first child to receive an infection from ρ is v_i.
2. Since ρ stays infected until everyone else is healthy, it is the same as running an excursion of $\mathbf{CP}^\otimes(T; \mathbf{1}_{v_i})$. When the excursion is finished, we go back to Step 1.

The probability that we terminate at Step 1 is $(1 + \lambda D)^{-1}$. So an excursion of $(X_t^\sharp)$ is a series of excursions of $\{\mathbf{CP}^\otimes(T; \mathbf{1}_{v_i})\}_i$, until we stop when having a successful coin toss of probability $(1 + \lambda D)^{-1}$ after each of the excursions.

Furthermore, note that the expected waiting time to see either a recovery at ρ or an infection at a child is $(1 + \lambda D)^{-1}$. Therefore, we obtain that

$$(4.4) \quad S(T) \leq S^\sharp(T) = \sum_{k=0}^{\infty} \left(\frac{1}{1 + \lambda D} \right) \left(\frac{\lambda D}{1 + \lambda D} \right)^k \left[\frac{k+1}{1 + \lambda D} + k S^\otimes \right] = 1 + \lambda D S^\otimes.$$

The final step is to consider the stationary distributions of $\mathsf{CP}_\rho^\lambda(T_i)$ and their product. Let π_i be the stationary distribution of $\mathsf{CP}_\rho^\lambda(T_i)$. Note that the process should always have a nontrivial stationary distribution as the added parent ρ stays infected forever, and it is independent of the initial state (and hence we have omitted it from the notation). By the definition of stationarity, $\pi_i(\mathbf{0})$ corresponds to the fraction of expected time that $\mathsf{CP}_\rho^\lambda(T_i)$ spends at state $\mathbf{0}$ (see [1, Section 2.2.2] for a formal discussion). Therefore, the law of large numbers implies that

$$(4.5) \qquad\qquad \pi_i(\mathbf{0}) = \frac{\lambda^{-1}}{S(T_i) + \lambda^{-1}} = \frac{1}{1 + \lambda S(T_i)}.$$

Similarly, the stationary distribution $\pi^\otimes$ of $\mathsf{CP}^\otimes(T)$ satisfies

$$(4.6) \qquad\qquad \pi^\otimes(\mathbf{0}) = \frac{(\lambda D)^{-1}}{(\lambda D)^{-1} + S^\otimes} = \frac{1}{1 + \lambda D S^\otimes}.$$

Since $\pi^\otimes = \otimes_{i=1}^D \pi_i$, this implies

$$(4.7) \qquad\qquad 1 + \lambda D S^\otimes = \prod_{i=1}^{D}(1 + \lambda S(T_i)).$$

Therefore, we plug this into (4.4) and obtain the conclusion. $\square$

Unfortunately, having (4.1) is insufficient for our purpose. One can see this by taking expectation on each side of (4.1) over $T \sim \mathsf{GW}(\xi)$. In order to yield a meaningful recursion, λ should be small in terms of the exponential moment of ξ, which has nothing to do with $1/\mathbb{E}\xi$ in general (for details, see the proof of [3, Lemma 3.3]). Therefore, we develop another recursion which redeems (4.1). Our first step is to build up a recursion regarding $R(T)$, the expected survival time.

Proposition 4.2. *Let* $T_1, \ldots, T_D$ *be as Proposition* 4.1, *and assume that* $\lambda^2 \sum_{i=1}^D R(T_i) < 1$. *Then,*

$$(4.8) \qquad\qquad R(T) \leq \frac{1 + \lambda \sum_{i=1}^D R(T_i)}{1 - \lambda^2 \sum_{i=1}^D R(T_i)}.$$

Proof. Suppose that we run $\mathsf{CP}^\lambda(T; \mathbf{1}_\rho)$. In the beginning, which we call the *first round*, the infection at the root stays there for a while, then it may infect some children. If a child v_i gets infected, then we can think of it as running a new contact process $\mathsf{CP}^\lambda(T_i; \mathbf{1}_{v_i})$. Here, we should also consider the effect of v_i infecting ρ again, and if this happens, the reinfected ρ starts the *second round* of the dynamics.

The expected survival time of the first round is bounded by

$$1 + \lambda \sum_{i=1}^{D} R(T_i),$$

since the expected survival time of the root is 1 and it sends infections λ times in expectation to each child before dying out. Let Q be the number of infections sent from the children $\{v_i\}$ to ρ in the first round. Then, similarly as before, we have

$$\mathbb{E}Q \le \lambda^2 \sum_{i=1}^{D} R(T_i).$$

Let $\tau_1, \ldots, \tau_Q$ denote the times when an infection is sent from the children $\{v_i\}$ to ρ during the first round. Moreover, let $\mathfrak{G}_0$ refer to the space-time diagram (such as Figure 1) that encodes the graphical representation of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$. In other words, the dynamics of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ can be seen on the diagram $\mathfrak{G}_0$, starting from the single initial infection at ρ at time 0. Let $\mathfrak{G}_1, \ldots, \mathfrak{G}_Q$ be the identical copies of $\mathfrak{G}_0$, and consider the following modification of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ on $\mathfrak{G}_0$:

- All the infections from $\{v_i\}$ to ρ during the first round are ignored on $\mathfrak{G}_0$. In other words, this modification on $\mathfrak{G}_0$ illustrates the first round of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$. Let $\mathcal{R}'$ be the survival time of this process.
- For each $j = 1, \ldots, Q$, consider the contact process described by $\mathfrak{G}_j$, with the single initial infection at ρ at time τ_j. Let $\mathcal{R}_j$ be the survival time after the initial infection is introduced.

Note that the above is a decomposition of the original $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ on $\mathfrak{G}_0$. Therefore, the survival time $\mathcal{R}$ of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ can be written as follows:

$$\mathcal{R} = \mathcal{R}' \vee \max\{\tau_j + \mathcal{R}_j : j = 1, \ldots, Q\}.$$

Since $\tau_j \le \mathcal{R}'$, we see that

$$\mathcal{R} \le \mathcal{R}' + \sum_{i=1}^{Q} \mathcal{R}_j.$$

Observing that $\mathbb{E}\mathcal{R}_j = \mathbb{E}\mathcal{R} = R(T)$ for all j, we take an expectation on both sides and deduce that

$$R(T) \le 1 + \lambda \sum_{i=1}^{D} R(T_i) + \left\{ \lambda^2 \sum_{i=1}^{D} R(T_i) \right\} R(T).$$

Then, the conclusion follows since we assumed $\lambda^2 \sum_{i=1}^{D} R(T_i) < 1$. $\qquad\square$

We are now interested in the relation between $R(T)$ and $S(T)$.

Proposition 4.3. *On a finite rooted tree (T, ρ), let $R(T)$ and $S(T)$ be the expected survival time of $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ and the expected excursion time of $\mathbf{CP}^\lambda_{\rho+}(T; \mathbf{1}_\rho)$, respectively. Then, we have*

(4.9) $$\frac{S(T)}{1 + \lambda S(T)} \le R(T).$$

Proof. Suppose that we are running a root-added contact process $(X_t) \sim \mathbf{CP}^\lambda_{\rho+}(T; \mathbf{1}_\rho)$ until time t_0. Further, let

$$A_{t_0} = \{t \le t_0 : X_t \ne \mathbf{0}\},$$

and let $|A_{t_0}|$ be the Lebesgue measure of the set A_{t_0}. In the root-added contact process, after one excursion we wait $(1/\lambda)$-time in expectation until we start the

next excursion. Therefore, we have

$$(4.10) \qquad \lim_{t_0 \to \infty} \frac{\mathbb{E}_{\mathrm{CP}} |A_{t_0}|}{t_0} = \frac{S(T)}{\lambda^{-1} + S(T)} = \frac{\lambda S(T)}{1 + \lambda S(T)}.$$

On the other hand, $\mathbf{CP}^\lambda_{\rho^+}(T; \mathbf{1}_\rho)$ can be considered as the contact process $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ of which the root ρ receives new infections at every ring of an independent Poisson process with rate λ. Since the rate-λ Poisson process rings λt_0 times in expectation until time t_0, we see that

$$\mathbb{E}_{\mathrm{CP}} |A_{t_0}| \leq \lambda t_0 R(T).$$

Comparing this to (4.10), we obtain that

$$\frac{\lambda S(T)}{1 + \lambda S(T)} \leq \lambda R(T),$$

implying the conclusion. $\hfill\square$

Combining Propositions 4.2 and 4.3, we obtain that

$$(4.11) \qquad \begin{aligned} S(T) \leq \frac{R(T)}{1 - \lambda R(T)} &\leq \frac{1 + \lambda \sum_{i=1}^D R(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^D R(T_i)} \\ &\leq \frac{1 + \lambda \sum_{i=1}^D S(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^D S(T_i)}, \end{aligned}$$

provided that $\lambda + 2\lambda^2 \sum_{i=1}^D S(T_i) < 1$. In the rest of the paper, (4.1) and (4.11) serve as two major recursive inequalities for $S(T)$.

4.2. Recursive tail estimate for Galton-Watson trees.

In this subsection, we establish the primary tail probability estimate on $S(\mathcal{T})$ for a Galton-Watson tree $\mathcal{T}$, and prove Theorem 1 as its application.

Let ξ be an integer-valued random variable that satisfies the concentration condition (1.2) for $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$. For the Galton-Watson tree $\mathcal{T} \sim \mathbf{GW}(\xi)$, the expected excursion time $S(\mathcal{T})$ is now a random variable driven by the randomness from $\mathbf{GW}(\xi)$. The goal of this subsection is to show that $S(\mathcal{T})$ is finite almost surely if $\lambda \leq (1-\varepsilon)d^{-1}$, where $\varepsilon > 0$ is an arbitrarily fixed constant and $d = \mathbb{E}\xi$ is large enough depending on ε. We establish this by proving that the upper tail of $S(\mathcal{T})$ is very light. In what follows, we denote the law of Galton-Watson trees of depth l by $\mathbf{GW}(\xi)^l$.

Theorem 4.4. *Let $l \geq 0$ be an integer, $\varepsilon \in (0,1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any ξ that satisfies $d := \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{T}^l \sim \mathbf{GW}(\xi)^l$ that*

$$(4.12) \qquad \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^l) \geq t\right) \leq t^{-\sqrt{d}} (\log t)^{-2} \text{ for all } t \geq \frac{2}{\varepsilon},$$

where $S(\mathcal{T}^l)$ is the expected excursion time on $\mathcal{T}^l$.

Remark 4.5. The term $(\log t)^{-2}$ in the r.h.s. of (4.12) may look useless, but this plays a key role in carrying out an inductive argument (see Lemma 4.8). We remark that the exponent -2 of $(\log t)^{-2}$ can be replaced by any number smaller than -1. Moreover, the exponent $\sqrt{d}$ can be replaced by any term of order $O(d^{1-\eta})$ for $\eta > 0$.

The rest of the subsection is devoted to the proof of Theorem 4.4. We do this by an induction on l, the tree depth. If $l = 0$, $\mathcal{T}^0$ is just a single vertex ρ and hence $S(\mathcal{T}^0) \equiv 1$, implying (4.12).

Suppose that we have (4.12) for l. Let $(\mathcal{T}^{l+1}, \rho) \sim \mathbf{GW}(\xi)^{l+1}$, and set $D = \deg(\rho) \overset{\mathrm{d}}{=} \xi$. As before, let T_i, $i \in [D]$ denote the subtrees of $\mathcal{T}^{l+1}$, rooted at the child v_i. Let c_1 be the constant in $\mathfrak{c}$ with $\delta = 1$, and we divide t into three regimes as follows.

1. (small) $2/\varepsilon \leq t \leq d^{\frac{1}{10}}$;
2. (intermediate) $d^{\frac{1}{10}} \leq t \leq \exp(\frac{1}{2}c_1\sqrt{d})$;
3. (large) $\exp(\frac{1}{2}c_1\sqrt{d}) \leq t$.

Then, we establish (4.12) on each regime separately. As the proof goes on, we will figure out the conditions for $d_0(\varepsilon, \mathfrak{c})$ as well.

Remark 4.6. There is much freedom in choosing $d^{\frac{1}{10}}$, the threshold between the small and intermediate regime. Indeed, any d^η with $\eta \in (0, 1 - d^{-1/2})$ would work for our purpose. However, the specific choice of $d^{1/10}$ will turn out to be useful later, in the proof of Proposition 6.4 and Lemma A.5. The choice of $\exp(\frac{1}{2}c_1\sqrt{d})$ will be clear in (4.22).

4.2.1. *Proof of Theorem* 4.4 *for small* t. To show (4.12) for small t, we rely on (4.11). For $\mathcal{T}^{l+1} \sim \mathbf{GW}(\xi)^{l+1}$ and T_i, $i \in [D]$ as above, suppose that

$$\sum_{i=1}^{D} S(T_i) \leq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right).$$

Then, from (4.11) and a little bit of algebra we see that

$$S(\mathcal{T}^{l+1}) \leq \frac{1 + \lambda\frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)}{1 - \lambda - 2\lambda^2\frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)} \leq \frac{\frac{2}{\varepsilon} - \frac{1}{3}}{1 - \frac{7}{d\varepsilon}} < \frac{2}{\varepsilon},$$

where the last inequality holds if

$$(4.13) \qquad\qquad d \geq \frac{42}{\varepsilon^2}.$$

Therefore, for d with (4.13), we have
(4.14)
$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq \frac{2}{\varepsilon}\right) \leq \mathbb{P}\left(D \geq \left(1 + \frac{\varepsilon}{6}\right)d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} S(T_i) \geq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right),$$

where T_i, $i \in \mathbb{N}$ are i.i.d. $\mathbf{GW}(\xi)^l$. By the assumption on ξ, we can bound the first term in the r.h.s. by

$$(4.15) \qquad\qquad \mathbb{P}\left(D \geq \left(1 + \frac{\varepsilon}{6}\right)d\right) \leq \exp\left(-c_{\varepsilon'}d\right),$$

for $\varepsilon' = \varepsilon/6$. To deal with the second term, the induction hypothesis tells us that the c.d.f. of $S(T_i)$ has an upper bound

$$\mathbb{P}_{\mathrm{GW}}(S(T_i) \geq s) \leq s^{-\sqrt{d}}$$

for all $s \geq 2/\varepsilon$, and hence we can apply the following lemma.

Lemma 4.7. *Let $\varepsilon \in (0,1)$ be a given constant. Then there exists $d_1 > 0$ independent of ε such that the following holds for all $d \geq d_1$. Let Z_i, $i \in \mathbb{N}$ be i.i.d. positive random variables that satisfies*

$$(4.16) \qquad \mathbb{P}\left(Z_i \geq t\right) \leq 5t^{-\sqrt{d}}, \quad \text{for all } t \geq \frac{2}{\varepsilon}.$$

Then, we have

$$(4.17) \qquad \mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i \geq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right) \leq \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}.$$

We defer the proof of this lemma until Section 4.3, since it is essentially a special case of Lemma 4.8.

Thanks to Lemma 4.7, we combine (4.14), (4.15) and (4.17) to obtain that
$$(4.18)$$
$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq \frac{2}{\varepsilon}\right) \leq \exp(-c_{\varepsilon'}d) + \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\log d^{\frac{1}{10}}\right)^{-2} \leq d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2},$$

for d satisfying

$$(4.19) \qquad \exp(-c_{\varepsilon'}d) \leq \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}.$$

Finally, we clearly see that (4.18) settles (4.12) for all small t, namely, $2/\varepsilon \leq t \leq d^{\frac{1}{10}}$.

4.2.2. *Proof of Theorem 4.4 for intermediate t.* For $t \geq d^{\frac{1}{10}}$, we use (4.1). That is, we attempt to control

$$(4.20) \qquad \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D} \log(1 + \lambda S(T_i)) \geq \log t\right).$$

Let c_1 be the constant given from the assumption of Theorem 4.4 (with $\delta = 1$). Note that we can split the event in the r.h.s. as follows.

$$(4.21) \qquad \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq 2d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1 + \lambda S(T_i)) \geq \log t\right).$$

The concentration assumption (1.2) on ξ tells us that the first term is bounded by

$$(4.22) \qquad \mathbb{P}(D \geq 2d) \leq \exp(-c_1 d) \leq \frac{1}{2}t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for } t \leq e^{\frac{1}{2}c_1\sqrt{d}}.$$

To control the second term in the r.h.s. of (4.21), we use the following lemma.

Lemma 4.8. *Let $\varepsilon \in (0,1)$ be a given constant. Then there exists $d_1(\varepsilon) > 0$ such that the following holds true for all $d \geq d_1$. Let Z_i, $i \in \mathbb{N}$ be i.i.d. positive random variables that satisfies*

$$(4.23) \qquad \mathbb{P}(Z_i \geq t) \leq 5t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq \frac{3}{\varepsilon}.$$

Then, for all $t \geq d^{\frac{1}{10}}$, we have

$$(4.24) \qquad \mathbb{P}\left(\sum_{i=1}^{2d} \log\left(1 + \frac{4}{d}Z_i\right) \geq \log\left(\frac{t}{2}\right)\right) \leq \frac{1}{2}t^{-\sqrt{d}}(\log t)^{-2}.$$

The proof of Lemma 4.8 is postponed until Section 4.3, since it requires a substantial technical work. Note that (4.24) contains a slightly more generalized form than the r.h.s. of (4.21), which will be useful in Sections 5 and 6.

To conclude the proof for intermediate t, (4.21), (4.22) and (4.24) together deduce (4.12) for $t \geq d^{\frac{1}{10}}$. Here, the constant d_0 in the statement of the theorem should satisfy $d_0 \geq d_1$ for d_1 in Lemma 4.8.

4.2.3. *Proof of Theorem 4.4 for large t.* For $t \geq \exp(\frac{1}{2}c_1\sqrt{d})$, we again attempt to control (4.20). Let

$$\Delta_{d,t} = \frac{4\sqrt{d}}{c_1}\log t,$$

and note that we can split the event in the r.h.s. of (4.20) as follows.

(4.25)

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \Delta_{d,t}\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d}\log(1+\lambda S(T_i)) \geq \log t\right)$$

$$+ \sum_{2d \leq r \leq \Delta_{d,t}}\mathbb{P}(D \geq r) \times \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{r}\log(1+\lambda S(T_i)) \geq \log t\right).$$

The concentration assumption (1.2) on ξ tells us that for $t \geq \exp(\frac{1}{2}c_1\sqrt{d})$,

$$(4.26) \quad \mathbb{P}\left(D \geq \Delta_{d,t}\right) \leq \mathbb{P}\left(D \geq d + \frac{\Delta_{d,t}}{2}\right) \leq \exp\left(-2\sqrt{d}\log t\right) \leq \frac{1}{4}t^{-\sqrt{d}}(\log t)^{-2},$$

where the last inequality is satisfied by large d. This controls the first term in the r.h.s. of (4.25). The second term is then estimated by Lemma 4.8, which gives

$$(4.27) \quad \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d}\log(1+\lambda S(T_i)) \geq \log t\right) \leq \frac{1}{2}t^{-\sqrt{d}}(\log t)^{-2}.$$

To bound the last term, we use the following corollary.

Corollary 4.9. *Let $\varepsilon \in (0,1)$ be a given constant. Then there exist $d_1(\varepsilon) > 0$ and an absolute constant $C_1 > 0$ such that the following holds true for all $d \geq d_1$. Let Z_i, $i \in \mathbb{N}$ be i.i.d. positive random variables that satisfies*

$$\mathbb{P}(Z_i \geq t) \leq 5t^{-\sqrt{d}}(\log t)^{-2}, \text{ for all } t \geq \frac{3}{\varepsilon}.$$

Then, for all $t \geq \exp(\frac{1}{2}c_1\sqrt{d})$ and $2d \leq r \leq \Delta_{d,t}$ with $\Delta_{d,t}$ as above, we have

$$(4.28) \quad \mathbb{P}\left(\sum_{i=1}^{r}\log\left(1+\frac{4}{d}Z_i\right) \geq \log\left(\frac{t}{2}\right)\right) \leq t^{-\sqrt{d}}(\log t)^{-2}\exp\left(\frac{C_1 r}{\varepsilon\sqrt{d}}\right).$$

The proof of Corollary 4.9 is postponed until Section 4.3, which will be proven together with Lemma 4.8.

To conclude the proof of Theorem 4.4 for large t, observe that for $2d \leq r \leq \Delta_{d,t}$

$$\mathbb{P}(D \geq r) \leq \mathbb{P}\left(D \geq d + \frac{r}{2}\right) \leq \exp\left(-\frac{c_1 r}{2}\right),$$

by the concentration condition (1.2) of D. Therefore, for sufficiently large d, we have by (4.28) that

$$\mathbb{P}(D \geq r) \times \mathbb{P}_{\text{GW}}\left(\sum_{i=1}^{r} \log(1 + \lambda S(T_i)) \geq \log t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \exp\left(-\frac{c_1 r}{4}\right).$$

Along with (4.25), (4.26) and (4.27), this proves Theorem 4.4 for $t \geq \exp(\frac{1}{2}c_1\sqrt{d})$. $\qquad\square$

As the first application of Theorem 4.4, we establish Theorem 1.

Proof of Theorem 1: The lower bound. Let $\varepsilon \in (0,1)$ be an arbitrary constant and suppose that the collection of distributions $\{\xi_k\}_k$ satisfies (1.2) for $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$. Let $d_0 = d_0(\varepsilon, \mathfrak{c})$ as in Theorem 4.4, and assume that $\mathbb{E}\xi_k \geq d_0$. Then, Theorem 4.4 implies that for $\mathsf{CP}^\lambda_{\rho^+}(\mathcal{T}_k^l; \mathbf{1}_\rho)$ on $\mathcal{T}_k \sim \mathsf{GW}(\xi_k)^l$ with intensity $\lambda = (1-\varepsilon)(\mathbb{E}\xi_k)^{-1}$, we have

$$\mathbb{E}_{\text{GW}}\left[S(\mathcal{T}_k^l)\right] < \frac{3}{\varepsilon}.$$

Thus, we deduce that the expected survival time of $\mathsf{CP}^\lambda(\mathcal{T}_k^l; \mathbf{1}_\rho)$ is also

$$\mathbb{E}_{\text{GW}}\left[R(\mathcal{T}_k^l)\right] < \frac{3}{\varepsilon}.$$

For $\mathcal{T}_k \sim \mathsf{GW}(\xi_k)$, consider the natural coupling of the contact processes on $\mathcal{T}_k$ and $\{\mathcal{T}_k^l\}_{l \in \mathbb{N}}$ using the graphical representation, defined on the same probability space. Then, by the monotone convergence theorem, we have

$$\mathbb{E}_{\text{GW}}[R(\mathcal{T}_k)] = \lim_{l \to \infty} \mathbb{E}_{\text{GW}}[R(\mathcal{T}_k^l)] \leq \frac{3}{\varepsilon},$$

and hence the survival time $\mathbf{R}(\mathcal{T}_k)$ is finite almost surely, which implies that

$$\lambda = \frac{1-\varepsilon}{\mathbb{E}\xi_k} \leq \lambda_1^{\text{GW}}(\xi_k),$$

for all large enough k. $\qquad\square$

4.3. Proof of the induction lemmas.

To conclude Section 4, we discuss the proof of Lemmas 4.7, 4.8 and Corollary 4.9. We first establish Lemma 4.8 and Corollary 4.9 together, and then the proof of Lemma 4.7 will follow based on similar ideas.

Proof of Lemma 4.8 and Corollary 4.9. We first set

$$(4.29) \quad Y_i = \log\left(1 + \frac{4Z_i}{d}\right), \quad \Delta_{d,t} = \frac{4\sqrt{d}}{c_1}\log t, \quad \text{and let} \quad 2d \leq r \leq 2d \vee \Delta_{d,t}.$$

We will attempt to find an upper bound of

$$(4.30) \qquad\qquad \mathbb{P}\left(\sum_{i=1}^{r} \log\left(1 + \frac{4}{d}Z_i\right) \geq \log\left(\frac{t}{2}\right)\right),$$

to cover the case of Corollary 4.9 as well.

Note that the assumption (4.23) on Z_i gives us $\mathbb{E}Z_i \leq \frac{4}{\varepsilon}$. Based on (4.23), we can find an upper bound for $\mathbb{E}Y_i$ by

$$(4.31) \qquad \mathbb{E}Y_i \leq \log \mathbb{E}e^{Y_i} = \log\left(1 + \frac{4\mathbb{E}Z_i}{d}\right) \leq \log\left(1 + \frac{16}{\varepsilon d}\right) \leq \frac{16}{\varepsilon d}.$$

Further, (4.23) tells us that the tail of Y_i is bounded by

(4.32)
$$\mathbb{P}(Y_i \geq s) = \mathbb{P}\left(Z_i \geq \frac{d}{4}(e^s - 1)\right) \leq 5\left(\frac{d}{4}(e^s - 1)\right)^{-\sqrt{d}}\left\{\log\left(\frac{d}{4}(e^s - 1)\right)\right\}^{-2}.$$

In order to have $\sum_{i=1}^r Y_i \geq \log(t/2)$ for $t \geq \sqrt{d}$ and $r \leq 2d \vee \Delta_{d,t}$, we need to have some of Y_i being larger than $\frac{\log d}{80d}$. In the following, k denotes the number of Y_i that are at least $\frac{\log d}{80d}$. Note that in this case, the sum over all Y_i that are smaller than $\frac{\log d}{80d}$ is at most

(4.33)
$$(2d \vee \Delta_{d,t}) \cdot \frac{\log d}{80d} = \frac{1}{40}\log d \vee \frac{\log d}{20c_1\sqrt{d}}\log t =: m(d,t).$$

Then, we can split the probability in (4.30) as follows.

(4.34)
$$\mathbb{P}\left(\sum_{i=1}^r Y_i \geq \log\left(\frac{t}{2}\right)\right)$$
$$\leq \sum_{k=1}^r \binom{r}{k} \sum_{j=1}^{m(d,t)+1} \mathbb{P}\left(\sum_{i=1}^k Y_i \geq \log\left(\frac{t}{2}\right) - j \; ; \; Y_i \geq \frac{\log d}{80d}, \; \forall i \leq k\right)$$
$$\times \mathbb{P}\left(\sum_{i=k+1}^r Y_i \geq j - 1 \; ; \; Y_i \leq \frac{\log d}{80d}, \; \forall i > k\right).$$

Our focus is to control the two terms in the inner sum of the r.h.s. We begin with the first one.

Lemma 4.10. *Under the setting of Lemma 4.8 and (4.29), there exist absolute constants $K, d_0 > 0$ such that for all $d \geq d_0$, $1 \leq k \leq r$ and $1 \leq j \leq m(d,t)+1$, we have*
$$\mathbb{P}\left(\sum_{i=1}^k Y_i \geq \log\left(\frac{t}{2}\right) - j \; ; \; Y_i \geq \frac{\log d}{80d}, \; \forall i \leq k\right)$$
$$\leq \left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k e^{-\sqrt{d}(\log t - j)}(\log t)^{-2}.$$

Proof. Since $e^x - 1 \geq \frac{x}{x+1}e^x \geq \frac{y}{2}e^x$ for all $x \wedge 1 \geq y > 0$, we have that for $s \geq \frac{\log d}{80d}$,
$$\frac{d}{4}(e^s - 1) \geq \frac{d}{4} \cdot \frac{\log d}{160d}e^s = \frac{\log d}{640}e^s.$$

Plugging this into (4.32), we obtain that

(4.35)
$$\mathbb{P}(Y_i \geq s) \leq 5\left(\frac{640}{\log d}\right)^{\sqrt{d}} e^{-s\sqrt{d}}\left(s + \frac{1}{2}\log\log d\right)^{-2},$$

for all $s \geq \frac{\log d}{80d}$, if d is such that

(4.36)
$$\frac{1}{2}\log\log d \leq \log\left(\frac{\log d}{640}\right), \quad \text{i.e.,} \quad d \geq \exp\left(640^2\right).$$

Keeping this in mind, we can rewrite the probability for each k and j as

$$\mathbb{P}\left(\sum_{i=1}^{k} Y_i \geq \log\left(\frac{t}{2}\right) - j \; ; \; Y_i \geq \frac{\log d}{80d}, \; \forall i \leq k\right)$$

$$(4.37) \qquad \leq \sum_{\substack{u_1,\ldots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i \geq \frac{80d}{\log d}\{\log(t/2)-j\}-k}} \left[\prod_{i=1}^{k} \mathbb{P}\left(Y_i \in \frac{\log d}{80d}[u_i, \, u_i+1]\right)\right]$$

$$\leq \sum_{\substack{u_1,\ldots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k \rfloor}} \left[\prod_{i=1}^{k} \mathbb{P}\left(Y_i \geq \frac{\log d}{80d}u_i\right)\right].$$

Then, the r.h.s. of (4.37) is upper bounded by

$$5^k \left(\frac{640}{\log d}\right)^{k\sqrt{d}} e^{-\sqrt{d}\left(\log(t/2)-j-\frac{k\log d}{80d}\right)}$$

$$(4.38) \qquad \times \sum_{\substack{u_1,\ldots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k \rfloor}} \prod_{i=1}^{k}\left(\frac{\log d}{80d}u_i + \frac{1}{2}\log\log d\right)^{-2}.$$

The sum in the right can be controlled inductively in k according to the following lemma.

Lemma 4.11. *Let* $a > b > 0$ *and* $m \in \mathbb{N}$. *Then,*

$$\sum_{u=0}^{m}(a+bu)^{-2}(a+b(m-u))^{-2} \leq \frac{8}{b(a-b)(a+bm)^2}.$$

Proof of Lemma 4.11. The l.h.s. of above can be upper bounded by

$$2\sum_{u=0}^{\lceil m/2 \rceil}(a+bu)^{-2}\left(a+\frac{bm}{2}\right)^{-2} \leq \frac{8}{b^2(2a+bm)^2}\sum_{m=0}^{\infty}\left(u+\frac{a}{b}\right)^{-2}$$

$$\leq \frac{8}{b(a-b)(a+bm)^2}.$$

$\square$

From now on, $K > 0$ is an absolute constant that may vary from line by line. We apply Lemma 4.11 $(k-1)$-times to (4.38) with $a = \frac{1}{2}\log\log d$, $b = \frac{\log d}{80d}$ and see that

$$(4.39) \qquad \sum_{\substack{u_1,\ldots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k \rfloor}} \prod_{i=1}^{k}\left(\frac{\log d}{80d}u_i + \frac{1}{2}\log\log d\right)^{-2}$$

$$\leq \left(\frac{Kd}{\log d}\right)^{k-1}\left(\log\left(\frac{t}{2}\right) - j - \frac{k\log d}{80d}\right)^{-2}.$$

Therefore, we see from (4.37), (4.38) and (4.39) that

$$
\begin{aligned}
(4.40) \quad & \mathbb{P}\left(\sum_{i=1}^{k} Y_i \geq \log\left(\frac{t}{2}\right) - j \; ; \; Y_i \geq \frac{\log d}{80d}, \; \forall i \leq k\right) \\
& \leq \left(\frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^{k} e^{-\sqrt{d}\left(\log(t/2) - j - \frac{k \log d}{80d}\right)} \left(\log\left(\frac{t}{2}\right) - j - \frac{k \log d}{80d}\right)^{-2}.
\end{aligned}
$$

Note that the definition of r (4.29) and $m(d,t)$ (4.33) implies that for $t \geq d^{\frac{1}{10}}$,

$$
\log\left(\frac{t}{2}\right) - m(d,t) - \frac{r \log d}{80d} \geq \log\left(\frac{t}{2}\right) - \frac{1}{4}\log t - \frac{1}{4}\log t \geq \frac{1}{3}\log t.
$$

Therefore, the r.h.s. of (4.40) is at most

$$
\left(\frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^{k} e^{-\sqrt{d}\left(\log(t/2) - j - \frac{k \log d}{80d}\right)} 9(\log t)^{-2}.
$$

Absorbing the constants 9, $2^{\sqrt{d}}$ and the term $\exp(\frac{k \log d}{80\sqrt{d}})$ into K, we obtain the conclusion. $\qquad\square$

Now we turn our attention to bounding the second term in the inner sum of (4.34) in the r.h.s.

Lemma 4.12. *Under the setting of Lemma 4.8 and (4.29), there exist absolute constants $K, d_0 > 0$ such that for all $d \geq d_0$, $1 \leq k \leq r$ and $1 \leq j \leq m(d,t)+1$, we have*

$$
\mathbb{P}\left(\sum_{i=k+1}^{r} Y_i \geq j - 1 \; ; \; Y_i \leq \frac{\log d}{80d}, \; \forall i > k\right) \leq 1 \wedge \left(\frac{\exp\left(19r\, d^{-1/3}\right)}{\exp\left(\varepsilon d^{2/3} j\right)}\right).
$$

Proof. We first observe that by (4.31) and (4.32),

$$
\begin{aligned}
& \mathbb{E}\left[\exp\left(\varepsilon d^{2/3} Y_i\right) \; ; \; Y_i \leq \frac{\log d}{80d}\right] \\
& \qquad \leq \exp\left(\varepsilon d^{2/3}\left(\frac{16}{\varepsilon d}\right)\right) + \mathbb{E}\left[\exp\left(\varepsilon d^{2/3} Y_i\right) \mathbf{1}_{\left\{\frac{16}{\varepsilon d} \leq Y_i \leq \frac{\log d}{80d}\right\}}\right] \\
& \qquad \leq \exp\left(16 d^{-1/3}\right) + \int_{\frac{16}{\varepsilon d}}^{\frac{\log d}{80d}} e^{\varepsilon d^{2/3} y}\left(\frac{d}{4}(e^y - 1)\right)^{-\sqrt{d}} \mathrm{d}y \\
& \qquad \leq 1 + 17 d^{-\frac{1}{3}} + \frac{\log d}{80d} \times e^{\frac{\varepsilon \log d}{80 d^{1/3}}} \times \left(\frac{d}{4} \cdot \frac{16}{\varepsilon d}\right)^{-\sqrt{d}} \leq 1 + 18 d^{-\frac{1}{3}},
\end{aligned}
$$

where the last two inequalities hold for d larger than some absolute constant d_2. Therefore, we obtain that

$$\mathbb{P}\left(\sum_{i=k+1}^{r} Y_i \geq j-1 \,;\, Y_i \leq \frac{\log d}{80d},\, \forall i > k\right)$$

$$\leq \mathbb{P}\left(\sum_{i=1}^{r} Y_i \geq j-1 \,;\, Y_i \leq \frac{\log d}{80d},\, \forall i > k\right)$$

$$\leq 1 \,\wedge\, \exp\left(-\varepsilon d^{2/3}(j-1)\right) \times \left\{\mathbb{E}\left[\exp\left(\varepsilon d^{2/3} Y_i\right) \,;\, Y_i \leq \frac{\log d}{80d}\right]\right\}^r$$

$$\leq 1 \,\wedge\, \frac{\left(1 + 18 d^{-1/3}\right)^r}{\exp\left(\varepsilon d^{2/3}(j-1)\right)} \leq 1 \wedge \left(\frac{\exp\left(19 r\, d^{-1/3}\right)}{\exp\left(\varepsilon d^{2/3} j\right)}\right),$$

concluding the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can now combine (4.34) with Lemmas 4.10 and 4.12 to see that

$$\mathbb{P}\left(\sum_{i=1}^{r} Y_i \geq \log\left(\frac{t}{2}\right)\right)$$

$$(4.41) \qquad \leq \frac{t^{-\sqrt{d}}}{(\log t)^2} \sum_{k=1}^{r} \sum_{j=1}^{m(d,t)+1} \binom{r}{k} \left(\frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^k e^{j\sqrt{d}} \left(\frac{e^{19r\, d^{-1/3}}}{e^{\varepsilon d^{2/3} j}} \wedge 1\right),$$

$$\leq \frac{t^{-\sqrt{d}}}{(\log t)^2} \sum_{k=1}^{r} \binom{r}{k} \left(\frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^k \left[\sum_{j=1}^{m(d,t)+1} e^{j\sqrt{d}} \left(\frac{e^{19r\, d^{-1/3}}}{e^{\varepsilon d^{2/3} j}} \wedge 1\right)\right].$$

The last inner sum over j can be split into two parts, $\varepsilon d^{2/3} j \leq 19 r\, d^{-1/3}$ and $\varepsilon d^{2/3} j > 19 r\, d^{-1/3}$. That is, we divide into two regimes based on $j_0 = 19 r (\varepsilon d)^{-1}$. If, for instance, $2\sqrt{d} \leq \varepsilon d^{2/3}$, then this sum is bounded by $2\exp(j_0\sqrt{d}) = 2\exp(19 r\, \varepsilon^{-1} d^{-1/2})$, since the exponent $\varepsilon d^{2/3} j$ in the denominator becomes at least twice as large as $j\sqrt{d}$ in the numerator. This gives another condition for d_1, namely,

$$(4.42) \qquad\qquad\qquad\qquad d \geq \left(\frac{2}{\varepsilon}\right)^6.$$

Further, the outer sum over k in the r.h.s. of (4.41) is at most

$$(4.43) \quad 2\exp\left(\frac{19r}{\varepsilon\sqrt{d}}\right) \sum_{k=1}^{r} \binom{r}{k} \left(\frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^k \leq 2\exp\left(\frac{19r}{\varepsilon\sqrt{d}}\right) \times \left(1 + \frac{K^{\sqrt{d}} d}{(\log d)^{\sqrt{d}}}\right)^r$$

$$\leq \exp\left(\frac{20r}{\varepsilon\sqrt{d}}\right),$$

where the last inequality holds for large d such that

$$(4.44) \qquad\qquad\qquad\qquad \frac{1}{\varepsilon\sqrt{d}} > \left(\frac{K}{\log d}\right)^{\sqrt{d}} d.$$

Therefore, we establish Corollary 4.9 from (4.41) and (4.43) by setting $C_1 = 20$ and d_0 to satisfy (4.36), (4.42) and (4.44). For Lemma 4.8, we plug in $r = 2d$ in the l.h.s. of (4.43), and see that if d is so large that

$$(4.45) \qquad\qquad\qquad\qquad \log d \geq 3 \cdot K \cdot e^{19/\varepsilon},$$

then by using $\binom{2d}{k} \leq (2d)^k$ and the fact that $\sum_{k \geq 1} u^k \leq 2u$ for $u \leq \frac{1}{2}$,
(4.46)
$$2 \exp\left(\frac{19\sqrt{d}}{\varepsilon}\right) \sum_{k=1}^{2d} \binom{2d}{k} \left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k \leq 4e^{\frac{19\sqrt{d}}{\varepsilon}} \left(\frac{K^{\sqrt{d}}d^2}{(\log d)^{\sqrt{d}}}\right) \leq \frac{4d^2}{2^{\sqrt{d}}} \leq \frac{1}{2}.$$

Therefore, we obtain the conclusion of Lemma 4.8, by setting $d_1(\varepsilon)$ to satisfy (4.42) and (4.45). (Condition (4.36) is absorbed into (4.45).) □

We conclude this section by proving Lemma 4.7. The idea of splitting the probability as (4.34) is used, but here the computation is simpler than the previous one.

Proof of Lemma 4.7. Define $Z_i' = Z_i - \frac{2}{\varepsilon}$, then Z_i' satisfies

$$\mathbb{P}(Z_i' \geq s) \leq 5\left(s + \frac{2}{\varepsilon}\right)^{-\sqrt{d}}, \quad \text{for all } s \geq 0.$$

We claim that

(4.47)
$$\mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i \geq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right) \leq \mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i' \geq \frac{d}{3}\right) \leq \frac{1}{2} d^{-\frac{1}{10}\sqrt{d}} \left(\log\left(d^{\frac{1}{10}}\right)\right)^{-2}.$$

In order to have $\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i' \geq \frac{d}{3}$, some Z_i' must be at least $\frac{1}{10}$. In what follows, k denotes the number of such Z_i'. Note that on the other hand, the sum over all Z_i' that are at most $\frac{1}{10}$ is bounded by $(1 + \frac{\varepsilon}{6})\frac{d}{10} \leq \frac{d}{6}$ from above. Therefore, we can bound the probability in the middle in (4.47) by

(4.48)
$$\mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i' \geq \frac{d}{3}\right) \leq \sum_{k=1}^{(1+\frac{\varepsilon}{6})d} \binom{(1+\frac{\varepsilon}{6})d}{k} \mathbb{P}\left(\sum_{i=1}^{k} Z_i' \geq \frac{d}{6} \; ; \; Z_i' \geq \frac{1}{10}\right).$$

We can go through similar steps as in (4.37) to see that

$$\mathbb{P}\left(\sum_{i=1}^{k} Z_i' \geq \frac{d}{6} \; ; \; Z_i' \geq \frac{1}{10}\right) \leq \sum_{\substack{u_1,\ldots,u_k \geq 0 \\ u_1+\ldots+u_k=\lfloor\frac{5}{3}d-k\rfloor}} \prod_{i=1}^{k} \mathbb{P}\left(Z_i' \geq \frac{u_i}{10}\right)$$

$$\leq \sum_{\substack{u_1,\ldots,u_k \geq 0 \\ u_1+\ldots+u_k=\lfloor\frac{5}{3}d-k\rfloor}} 5^k \prod_{i=1}^{k} \left(\frac{2}{\varepsilon} + \frac{u_i}{10}\right)^{-\sqrt{d}}.$$

The choice of $\{u_i\}_{i=1}^k$ that maximizes the product in the r.h.s. is such that all but exactly one u_i are equal to zero. Hence, its maximum is at most

$$\left(\frac{\varepsilon}{2}\right)^{(k-1)\sqrt{d}} \left[\frac{\varepsilon}{2} + \frac{1}{10}\left(\frac{5d}{3} - k\right)\right]^{-\sqrt{d}} \leq \left(\frac{\varepsilon}{2}\right)^{(k-1)\sqrt{d}} \left(\frac{18}{d}\right)^{\sqrt{d}},$$

using $\frac{k}{10} \leq \frac{d}{9}$. Since the number of $\{u_i\}_{i=1}^k$ satisfying $\sum_{i=1}^k u_i = \lfloor \frac{5}{3}d - k \rfloor$ is at most $\binom{2d}{k}$, the r.h.s. of (4.48) is at most

$$\sum_{k=1}^{(1+\frac{\varepsilon}{6})d} \binom{2d}{k}^2 5^k \left(\frac{\varepsilon}{2}\right)^{(k-1)\sqrt{d}} \left(\frac{18}{d}\right)^{\sqrt{d}}$$

$$\leq 2d \left(\frac{36}{\varepsilon d}\right)^{\sqrt{d}} \sum_{k \geq 1} \left[20d^2 \left(\frac{\varepsilon}{2}\right)^{\sqrt{d}}\right]^k \leq 80d^3 \left(\frac{18}{d}\right)^{\sqrt{d}},$$

using $\sum_{k \geq 1} u^k \leq 2u$ for small enough $u > 0$. Thus, there exists $d_1 > 0$ such that the r.h.s. is smaller than that of (4.47) for $d \geq d_1(\varepsilon)$. $\square$

5. TOTAL INFECTIONS AT LEAVES ON TREES

Let (T, ρ) be a finite rooted tree of *depth l*. That is,

$$(5.1) \qquad\qquad\qquad l = \max\{\text{dist}(\rho, v) : v \in T\}.$$

As discussed in Section 3.2.1, we desire to control the infection going deep inside the tree for subcritical λ. To this end, we investigate $M^l(T)$, the *expected total infections at depth-l leaves* defined in Definition 3.3. In particular, the goal of this section is to establish the following theorem:

Theorem 5.1. *Let $l \geq 0$ be an integer, $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any ξ that satisfies $d = \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, we have for $\lambda = (1 - \varepsilon)d^{-1}$ and $\mathcal{T}^l \sim \mathsf{GW}(\xi)^l$ that*

$$(5.2) \qquad \mathbb{P}_{\text{GW}}\left(M^l(\mathcal{T}^l) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2,$$

where $M^l(\mathcal{T}^l)$ is the expected total infections at depth-l leaves on $\mathcal{T}^l$.

To prove this theorem, we first derive two different recursive inequalities for $M^l(T)$ for a deterministic tree T in Sections 5.1 and 5.2. Then, in Section 5.3, we verify the theorem, which is along the lines of proving Theorem 4.4.

5.1. **The first recursive inequality.** We begin with deriving a recursive inequality on $M^l(T)$ described in the following proposition, which is an analogue of (4.1).

Proposition 5.2. *For a finite rooted tree (T, ρ) of depth l, let $D = \deg(\rho)$ and $T_1, \ldots, T_D$ be the subtrees rooted at each child v_i of ρ. Then, $M^l(T)$, the expected total infections at depth-l leaves on T, satisfies the following.*

$$(5.3) \qquad\qquad M^l(T) \leq \lambda \sum_{i=1}^D M^{l-1}(T_i) \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda S(T_j)).$$

Proof. Recall the processes $\mathsf{CP}^{\otimes}(T; \mathbf{1}_{v_i})$ and $(X_t^{\sharp})$ defined in the proof of Proposition 4.1. Let $\mathbf{S}^{\sharp}(T)$ and $\mathbf{S}_i^{\otimes}(T)$ be the excursion time of $(X_t^{\sharp})$ and $(X_{i,t}^{\otimes}) \sim \mathsf{CP}^{\otimes}(T; \mathbf{1}_{v_i})$, respectively, and set $\mathcal{L} = \{v \in T : \text{dist}(v, \rho) = l\}$. We define the total infections at depth-l leaves $\mathbf{M}^{\sharp}(T)$ (resp., $\mathbf{M}_i^{\otimes}(T)$) of $(X_t^{\sharp})$ (resp., $(X_{i,t}^{\otimes})$) analogously as Definition 3.3.

- For $v \in \mathcal{L}$, let

$$\mathbf{M}_v^\sharp(H) := \left| \left\{ s \in [0, \mathbf{S}^\sharp(H)] : \ X_s^\sharp(v) = 1 \text{ and } X_{s-}^\sharp(v) = 0 \right\} \right|;$$

$$\mathbf{M}_{i,v}^\otimes(H) := \left| \left\{ s \in [0, \mathbf{S}_i^\otimes(H)] : \ X_{i,s}^\otimes(v) = 1 \text{ and } X_{i,s-}^\otimes(v) = 0 \right\} \right|.$$

- We further set the total infections at leaves to be

$$\mathbf{M}^\sharp(T) := \sum_{v \in \mathcal{L}} \mathbf{M}_v^\sharp(T) \quad \text{and} \quad M^\sharp(T) = \mathbb{E}_{\mathrm{CP}} \mathbf{M}^\sharp(T);$$

$$\mathbf{M}_i^\otimes(T) := \sum_{v \in \mathcal{L}} \mathbf{M}_{i,v}^\otimes(T) \quad \text{and} \quad M_i^\otimes(T) = \mathbb{E}_{\mathrm{CP}} \mathbf{M}_i^\otimes(T).$$

By a standard coupling between $(X_t^\sharp)$ and $(X_t) \sim \mathbf{CP}_{\rho^+}^\lambda(T; \mathbf{1}_\rho)$ based on their graphical representations, we have $M^\sharp(T) \geq M^l(T)$.

Moreover, in the perspectives of (4.4), the number of excursions of $\{(X_{i,t}^\otimes)\}$ included in a single excursion of $(X_t^\sharp)$ is the same as a geometric random variable with success probability $(1 + \lambda D)^{-1}$. Therefore, λD excursions of $\{(X_{i,t}^\otimes)\}$ happen in expectation during an excursion of $(X_t^\sharp)$. Since the initial condition $\mathbf{1}_{v_i}$ of the product chain is selected uniformly at random at each excursion, we see that

$$(5.4) \qquad M^\sharp(T) = \lambda D \times \frac{\sum_{i=1}^D M_i^\otimes(T)}{D} = \lambda D M^\otimes(T),$$

where we define $M^\otimes(T)$ as the arithmetic mean of $\{M_i^\otimes(T)\}$.

Now we attempt to control $M^\otimes(T)$ in terms of $\{M(T_i)\}$. To this end, let $(X_t^\otimes) \sim \mathbf{CP}^\otimes(T; \mathbf{0})$, and we observe that due to the same argument as (5.8) and (5.9),

$$(5.5) \qquad \begin{aligned} \lim_{t_0 \to \infty} \frac{1}{t_0} &\mathbb{E}_{\mathrm{CP}} \left[\left| \left\{ s \in [0, t_0] : \ X_s^\otimes(v) = 1 \text{ and } X_{s-}^\otimes(v) = 0 \right\} \right| \right] \\ &= \frac{M^\otimes(T)}{(\lambda D)^{-1} + S^\otimes(T)}. \end{aligned}$$

On the other hand, let $\mathcal{L}_i = \{v \in T_i : \mathrm{dist}(v, v_i) = l - 1\}$ for each $i \in [D]$ and $X_t^{(i)} \sim \mathbf{CP}_\rho^\lambda(T_i; \mathbf{0})$ be the restriction of $(X_t^\otimes)$ on T_i. Note that

$$\sum_{v \in \mathcal{L}} \left| \left\{ s \in [0, t_0] : \ X_s^\otimes(v) = 1 \text{ and } X_{s-}^\otimes(v) = 0 \right\} \right|$$

$$= \sum_{i=1}^D \sum_{v \in \mathcal{L}_i} \left| \left\{ s \in [0, t_0] : \ X_s^{(i)}(v) = 1 \text{ and } X_{s-}^{(i)}(v) = 0 \right\} \right|.$$

Thus, by the same reasoning as above, the l.h.s. of (5.5) is also equal to

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

Therefore, by (4.7), (5.5) implies that

$$M^\otimes(T) = \frac{1}{D} \sum_{i=1}^D M^{l-1}(T_i) \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda S(T_j)).$$

Finally, combining this with (5.4) and $M^\sharp(T) \geq M^l(T)$, we deduce the conclusion. $\qquad\square$

5.2. **The second recursive inequality.** The goal of this section is to obtain the recursive inequality on $M^l(T)$ that is in parallel to (4.11). In Definition 3.3, $\mathbf{M}^l(T)$ and $M^l(T)$ were defined with respect to the root-added contact process $\mathbf{CP}^\lambda_{\rho+}(T; \mathbf{1}_\rho)$. We begin with defining $\bar{\mathbf{M}}^l(T)$ and $\bar{M}^l(T)$ for the (usual) contact process $(\bar{X}_t) \sim \mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$ similarly as Definition 3.3.

- Let (T, ρ) be a rooted tree with depth l and $\mathcal{L} := \{v \in T : \mathrm{dist}(\rho, v) = l\}$ be the collection of depth-l leaves. For each $v \in \mathcal{L}$, we define $\bar{\mathbf{M}}^l_v(T)$ to be the *total infections at* v, that is,

$$\bar{\mathbf{M}}^l_v(T) := \left| \left\{ s \in [0, \mathbf{R}(T)] : \ \bar{X}_s(v) = 1 \text{ and } \bar{X}_{s-}(v) = 0 \right\} \right|,$$

  where $\mathbf{R}(T)$ is the survival time of $(\bar{X}_t)$.
- $\bar{\mathbf{M}}^l(T)$ is the *total infections at depth-l leaves* with respect to $(\bar{X}_t) \sim \mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$, given by

$$\bar{\mathbf{M}}^l(T) := \sum_{v \in \mathcal{L}} \bar{\mathbf{M}}^l_v(T).$$

  As before, we set $\bar{\mathbf{M}}^{l'}(T) \equiv 0$ for $l' > l$.
- $\bar{M}^l(T) := \mathbb{E}_{\mathrm{CP}} \bar{\mathbf{M}}^l(T)$ is the *expected total infections at depth-l leaves* with respect to $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$. We also set $\bar{M}^{l'}(T) = 0$ for $l' > l$.

For (T, ρ) as above, let $D = \deg(\rho)$ and $v_1, \ldots, v_D$ be the children of ρ. We denote the subtrees at each child by $T_1, \ldots, T_D$ as before. Here, each subtree has depth at most $l - 1$. We begin with obtaining the following recursive inequality on $\bar{M}^l(T)$ which is parallel to Proposition 4.2 on $R(T)$.

Corollary 5.3. *Under the above setting, let* $R(\cdot) = \mathbb{E}_{\mathrm{CP}} \mathbf{R}(\cdot)$ *and assume that* $\lambda^2 \sum_{i=1}^D R(T_i) < 1$. *Then, we have*

$$(5.6) \qquad\qquad \bar{M}^l(T) \leq \frac{\lambda \sum_{i=1}^D \bar{M}^{l-1}(T_i)}{1 - \lambda^2 \sum_{i=1}^D R(T_i)}.$$

Proof. When running $\mathbf{CP}^\lambda(T; \mathbf{1}_\rho)$, recall the notion of the *first and second round* discussed in the proof of Proposition 4.2. In the first round of infection, the expected total infections at leaves on T is bounded by

$$\lambda \sum_{i=1}^D \bar{M}^{l-1}(T_i),$$

due to the same reason as before. Also, the expected number of infections sent from the children $\{v_i\}$ to ρ in the first round is at most

$$\lambda^2 \sum_{i=1}^D R(T_i).$$

Thus, we get

$$\bar{M}^l(T) \leq \lambda \sum_{i=1}^D \bar{M}^{l-1}(T_i) + \left\{ \lambda^2 \sum_{i=1}^D R(T_i) \right\} \bar{M}^l(T),$$

which leads to our conclusion. $\qquad\square$

We would like to translate this result to a recursion on $M^l(T)$, and hence we discuss the relation between $M^l(T)$ and $\bar{M}^l(T)$, which is analogous to Proposition 4.3. Recall the definition of $S(T)$ from Definition 3.2.

Corollary 5.4. *On a finite rooted tree (T, ρ) of depth l, let $M^l(T)$ and $\bar{M}^l(T)$ be defined as above and let $S(T)$ denote the expected excursion time on T. Then, we have*

$$(5.7) \qquad M^l(T) \leq (1 + \lambda S(T))\bar{M}^l(T).$$

Proof. We rely on a similar idea as in Proposition 4.3. Let $\mathcal{L} = \{v \in T : \text{dist}(v, \rho) = l\}$ and $(X_t) \sim \mathbf{CP}_{\rho+}^{\lambda}(T; \mathbf{1}_{\rho})$ be the root-added contact process. For $v \in \mathcal{L}$, consider the following quantity:

$$(5.8) \qquad \lim_{t_0 \to \infty} \frac{1}{t_0} \sum_{v \in \mathcal{L}} \mathbb{E}_{\mathrm{CP}}\left[|\{s \in [0, t_0] : X_s(v) = 1 \text{ and } X_{s-}(v) = 0\}|\right].$$

First, by the renewal theorem (see, e.g., [6, Theorem 2.4.6]), observe that the above is the same as
$$(5.9)$$
$$\lim_{t_0 \to \infty} \frac{\text{the number of excursions in } (X_t) \text{ during } [0, t_0]}{t_0} \times M^l(T) = \frac{M^l(T)}{\lambda^{-1} + S(T)},$$

since the sum of the expected excursion time and the expected waiting time until the next excursion is $S(T) + \lambda^{-1}$. On the other hand, since the rate-λ Poisson process rings λt_0 times in expectation during $[0, t_0]$, the l.h.s. of (5.9) is at most $\lambda \bar{M}^l(T)$, by following the discussion in Proposition 4.3. $\qquad \square$

Let us combine (5.6) and (5.7) to deduce a recursion on $M^l(T)$ as follows.

$$M^l(T) \leq \frac{1 + \lambda S(T)}{1 - \lambda^2 \sum_{i=1}^{D} S(T_i)} \lambda \sum_{i=1}^{D} M^{l-1}(T_i).$$

By plugging in (4.11) in the r.h.s., we obtain the first recursion for $M^l(T)$, namely,

$$(5.10) \qquad M^l(T) \leq \frac{\lambda \sum_{i=1}^{D} M^{l-1}(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^{D} S(T_i)},$$

which is valid given that $\lambda + 2\lambda^2 \sum_{i=1}^{D} S(T_i) < 1$.

5.3. Recursive tail estimate for Galton-Watson trees. The goal of this subsection is to establish Theorem 5.1. As one may expect, we go through similar steps as Theorem 4.4 with some appropriate adjustments.

Let ξ be a positive, integer-valued random variable with mean $d = \mathbb{E}\xi$, and let $(\mathcal{T}^l, \rho) \sim \mathbf{GW}(\xi)^l$. Further, we denote the children of the root by $v_1, \ldots, v_D$ where $D = \deg(\rho)$. The subtree of $\mathcal{T}^l$ rooted at v_i is denoted by T_i. Set $\lambda = (1 - \varepsilon)d^{-1}$ as in the assumption.

We establish the theorem by an induction on l. The initial case $l = 0$ is obvious, since $\mathcal{T}^0$ is a single vertex and $M^0(\mathcal{T}^0) = 1$ by its definition (Definition 3.3). From now on, suppose that the conclusion holds for l, and we attempt to prove it for $l + 1$. As before, we split the inequality into three cases, namely,

1. (small) $2/\varepsilon \leq t \leq d^{\frac{1}{10}}$;
2. (intermediate) $d^{\frac{1}{10}} \leq t \leq \exp(\frac{1}{2}c_1\sqrt{d})$;
3. (large) $\exp(\frac{1}{2}c_1\sqrt{d}) \leq t$.

(The reason for the choice of $d^{\frac{1}{10}}$ is explained in Remark 4.6.) For convenience, we define
$$(5.11)$$
$$\widetilde{M}(\mathcal{T}^{l+1}) = \left(1 - \frac{\varepsilon}{10}\right)^{-(l+1)} M^{l+1}(\mathcal{T}^{l+1}), \quad \text{and} \quad \widetilde{M}(T_i) = \left(1 - \frac{\varepsilon}{10}\right)^{-l} M^l(T_i).$$

5.3.1. *Proof of Theorem* 5.1 *for small* t. Suppose that
$$\sum_{i=1}^{D} \widetilde{M}(T_i) \leq (2+\varepsilon)d, \quad \text{and} \quad \sum_{i=1}^{D} S(T_i) \leq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right).$$

Then, the second recursive inequality (5.10) tells us that
$$\widetilde{M}(\mathcal{T}^{l+1}) \leq \left(1 - \frac{\varepsilon}{10}\right)^{-1} \frac{(1-\varepsilon)(2+\varepsilon)}{1 - \frac{1}{d} - \frac{4}{\varepsilon d}(1 + \frac{\varepsilon}{3})} \leq 2,$$

where the last inequality holds for all $d \geq d_0'(\varepsilon)$ for appropriate $d_0'(\varepsilon)$. Therefore, for $t \geq 2$, we have

$$(5.12)$$
$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \left(1 + \frac{\varepsilon}{6}\right)d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} \widetilde{M}(T_i) \geq (2+\varepsilon)d\right)$$
$$+ \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} S(T_i) \geq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right),$$

where T_i, $i \in \mathbb{N}$ are i.i.d. $\mathsf{GW}(\xi)^l$. The first term in the r.h.s. is estimated by the concentration condition (1.2), namely,

$$(5.13) \qquad \mathbb{P}\left(D \geq \left(1 + \frac{\varepsilon}{6}\right)d\right) \leq \exp(-c_{\varepsilon'}d) \leq \frac{1}{4}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2},$$

for $\varepsilon' = \frac{\varepsilon}{6}$, where the last inequality holds if $d \geq d_0'(\varepsilon)$ for appropriate constant $d_0'(\varepsilon) > 0$. Then, Lemma 4.7 controls the last term in the r.h.s. of (5.12), implying the bound (4.16). For the second term, we claim that there exists $d_1(\varepsilon) > 0$ such that if $d \geq d_1(\varepsilon)$, then

$$(5.14) \qquad \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} \widetilde{M}(T_i) \geq (2+\varepsilon)d\right) \leq \frac{1}{4}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}.$$

Indeed, almost the same argument from the proof of Lemma 4.7 can be applied to deduce (5.14). The only two changes we need are to set
$$\widehat{M}(T_i) = \widetilde{M}(T_i) - 2,$$

(in the lemma, it was $Z_i' = Z_i - \frac{2}{\varepsilon}$) and to split $\widehat{M}$ based on $\frac{\varepsilon}{10}$, not $\frac{1}{10}$. Also, Lemma 4.7 holds if $d \geq d_1$ for an absolute constant d_1, but here d_1 depends on ε. We omit the remaining details.

Finally, we obtain the conclusion (5.2) for $l+1$ and $2 \leq t \leq d^{\frac{1}{10}}$ by combining (5.13), Lemma 4.7 and (5.14).

5.3.2. *Proof of Theorem* 5.1 *for intermediate t.* For $t \geq d^{\frac{1}{10}}$, we rely on the second recursive inequality (5.3). However, one issue here is that the quantity in the r.h.s. of (5.3) is no longer a single product of i.i.d. random variables as (4.1), (4.20). To overcome this difficulty, we define

$$(5.15) \qquad W_i := \max\left\{\widetilde{M}(T_i),\, S(T_i)\right\}, \quad \text{for each } i \in \mathbb{N},$$

for $\widetilde{M}(T_i)$ as in (5.11). Then, from a little bit of algebra we see that

$$\lambda \sum_{i=1}^{D} \widetilde{M}(T_i) \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda S(T_j)) \leq \sum_{i=1}^{D} \lambda W_i \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda W_j)$$

$$\leq \prod_{i=1}^{D} (1 + 2\lambda W_i).$$

Based on this observation, we attempt to control

$$(5.16) \quad \begin{aligned} \mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) &\leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D} \log(1 + 2\lambda W_i) \geq \log\left\{\left(1 - \frac{\varepsilon}{10}\right)t\right\}\right) \\ &\leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D} \log(1 + 2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right). \end{aligned}$$

We are also aware of the tail estimate on W_i by the induction hypothesis and Theorem 4.4, namely,
(5.17)
$$\mathbb{P}(W_i \geq s) \leq \mathbb{P}\left(\widetilde{M}(T_i) \geq s\right) + \mathbb{P}(S(T_i) \geq s) \leq 2s^{-\sqrt{d}}(\log s)^{-2}, \quad \text{for all } s \geq \frac{2}{\varepsilon}.$$

This falls into the assumption of Lemma 4.8. Therefore, we deduce the tail estimate on $\widetilde{M}(\mathcal{T}^{l+1})$ for intermediate t by

$$(5.18)$$

$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}(D \geq 2d) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1 + 2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right)$$

$$\leq \exp(-c_1 d) + \frac{1}{2} t^{-\sqrt{d}}(\log t)^{-2} \leq t^{-\sqrt{d}}(\log t)^{-2},$$

where we used the concentration condition (1.2) to bound the tail of D, and the last inequality is true if $d \geq d_0'(\varepsilon)$ for appropriate constant $d_0'(\varepsilon) > 0$.

5.3.3. *Proof of Theorem* 5.1 *for large t.* We continue to rely on W_i defined from (5.15). Set

$$\Delta_{d,t} = \frac{4\sqrt{d}}{c_1} \log t.$$

For large t, we modify (5.16) as in (4.25), which is

(5.19)

$$
\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \Delta_{d,t}\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1 + 2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right)
$$

$$
+ \sum_{2d \leq r \leq \Delta_{d,t}} \mathbb{P}(D \geq r) \times \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{r} \log(1 + 2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right).
$$

Note that $t/2$ in the log in the r.h.s. is a lower bound of $(1 - \frac{\varepsilon}{10})t$ (cf. (5.18)). The first term in the r.h.s. can be controlled by the concentration condition (1.2). By (5.17), we can use Lemma 4.8 to obtain the estimate for the second, and the third term is bounded by the same reasoning as in Section 4.2.3. This concludes the proof of Theorem 5.1. $\qquad\square$

6. Short survival on random graphs: Proof of Theorem 2

Let μ be a given degree distribution satisfying (1.2) for some positive constants $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$. Define $\widetilde{\mu}$ to be its size-biased distribution (2.1) and $d = \mathbb{E}_{D \sim \widetilde{\mu}} D$. In this section, we are interested in the contact process on $G_n \sim \mathcal{G}(n, \mu)$, particularly on its short survival. Our goal is to establish the following theorem, which implies the lower bound of Theorem 2, that is,

$$
\limsup_{k \to \infty} \limsup_{n \to \infty} \lambda_c^-\left(\mu_k\right) d_k \geq 1.
$$

Theorem 6.1. *Let $\varepsilon \in (0,1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true: Let μ be a probability measure on $\mathbb{N}$ whose sized-biased distribution $\widetilde{\mu}$ satisfies $d := \mathbb{E}_{D \sim \widetilde{\mu}} D \geq d_0$ and the concentration condition (1.2) with $\mathfrak{c}$. Further, let $\lambda = (1 - \varepsilon)d^{-1}$ and $G_n \sim \mathcal{G}(n, \mu)$. Then, there exists an event $\mathcal{E}_n$ over graphs such that $\mathbb{P}(G_n \in \mathcal{E}_n) = 1 - o(1)$ and*

(6.1)
$$
\mathbb{P}_{\mathrm{CP}}\left(\mathbf{R}(G_n) \leq n^2 \mid G_n \in \mathcal{E}_n\right) = 1 - o(1),
$$

where $\mathbf{R}(G_n)$ is the survival time of $\mathsf{CP}^\lambda(G_n; \mathbf{1}_{G_n})$.

To establish the theorem, we would like to study the structure of local neighborhoods $N(v, l)$ of $G_n \sim \mathcal{G}(n, \mu)$. Although a neighborhood selected uniformly at random converges weakly to a Galton-Watson tree, there are some neighborhoods who contain a cycle. In Section 6.1, we extend the properties from Sections 4 and 5 to certain Galton-Watson-type random graphs with a cycle that are relevant to the local neighborhoods of G_n. Then in Section 6.2, we develop a coupling between the local neighborhoods and the aforementioned graphs in Section 6.1, following the ideas from [3, Section 4.1]. Finally, we conclude the proof of Theorem 6.1 and establish the lower bound of Theorem 2 in Section 6.3, based on all the properties we obtained in the previous sections.

6.1. **Recursive analysis for unicyclic graphs.** In this subsection, we do the final preliminary work before delving into the proof of Theorem 2. Although we need to consider the neighborhoods $N(v, l)$ inside G_n that contain a cycle, fortunately, it turns out that it is enough to look at the case with exactly one cycle (see the discussion in Section 3.2.2 for a sketchy review, or Section 6.2 for a detailed explanation). Therefore, we are interested in the Galton-Watson type
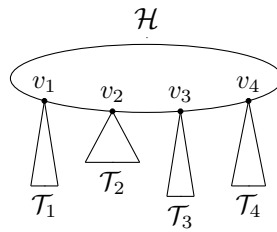
FIGURE 2. An example of $\mathcal{H} \sim \mathsf{GWC}^1(\xi, m)^l$ with $m = 4$. Note that the depth of a subtree $\mathcal{T}_j$ is at most l (not necessarily equal to l).

processes with a single cycle, particularly the ones which are introduced in this subsection.

Definition 6.2 (Galton-Watson-on-cycle process of type one)**.** Let ξ be a positive, integer-valued random variable, and let $m, l \geq 1$ be nonnegative integers. Then, $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi, m)^l$, 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 $C = v_1 v_2 \ldots v_m v_1$.
2. At each v_j for $j = 1, \ldots, m$, attach $\mathcal{T}_j^l \sim$ i.i.d. $\mathsf{GW}(\xi)^l$ by setting v_j as its root.

The resulting graph is called $\mathcal{H}^{m,l}$. 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. Figure 2 illustrates a diagram of a GWC^1-process.

We are again interested in the excursion time and the total infections at leaves on the GWC^1-processes. For concreteness, we present the following definitions.

Definition 6.3. 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 depth $\leq l$ trees $T_1, \ldots, T_m$ rooted at $v_1, \ldots, v_m$, respectively (recall the definition of *tree depth* in (5.1)).

1. The *root-added contact process* $\mathsf{CP}_{v_1^+}^\lambda(H; x)$ is the contact process on the graph $H \cup v_1^+$ with the permanently infected parent ρ^+ having a single connection with v_1, and with the initial condition $x \in \{0, 1\}^H$.

2. $\mathbf{S}(H)$ (resp., $\mathbf{R}(H)$) is the *excursion* (resp. *survival*) *time*, which is the first time when $\mathsf{CP}_{v_1^+}^\lambda(H; \mathbf{1}_{v_1})$ (resp., $\mathsf{CP}^\lambda(H; \mathbf{1}_{v_1})$) returns to the all-healthy state. $S(H) = \mathbb{E}_{\mathrm{CP}} \mathbf{S}(H)$ (resp., $R(H) = \mathbb{E}_{\mathrm{CP}} \mathbf{R}(H)$) denotes the *expected excursion* (resp., *survival*) *time*.

3. Let $\mathcal{L}_j = \{v \in T_j : \mathrm{dist}(v, v_j) = l\}$, $(X_t) \sim \mathsf{CP}_{v_1^+}^\lambda(H; \mathbf{1}_{v_1})$ and $(\bar{X}_t) \sim \mathsf{CP}^\lambda(H; \mathbf{1}_{v_1})$. For $v \in \mathcal{L}_j$ for some j, we define *the total infections at* v by

$$\mathbf{M}_v^l(H) := |\{s \in [0, \mathbf{S}(H)] : X_s(v) = 1 \text{ and } X_{s-}(v) = 0\}|;$$
$$\bar{\mathbf{M}}_v^l(H) := \left|\{s \in [0, \mathbf{R}(H)] : \bar{X}_s(v) = 1 \text{ and } \bar{X}_{s-}(v) = 0\}\right|.$$

Then, the *total infections at leaves* and its expectation are given as

$$\mathbf{M}^l(H) := \sum_{j=1}^{m} \sum_{v \in \mathcal{L}_j} \mathbf{M}_v^l(H), \quad \text{and} \quad M^l(H) := \mathbb{E}_{\mathrm{CP}} \mathbf{M}^l(H);$$

$$\bar{\mathbf{M}}^l(H) := \sum_{j=1}^{m} \sum_{v \in \mathcal{L}_j} \bar{\mathbf{M}}_v^l(H), \quad \text{and} \quad \bar{M}^l(H) := \mathbb{E}_{\mathrm{CP}} \bar{\mathbf{M}}^l(H).$$

Note that $\mathbf{M}^l(H) \equiv \bar{\mathbf{M}}^l(H) \equiv 0$, $M^l(H) = \bar{M}^l(H) = 0$ if all the depths of $T_1, \ldots, T_m$ are smaller than l.

The goal of this section is to establish the following theorem.

Proposition 6.4. *Let $m, l \geq 1$ be integers, $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any ξ that satisfies $d := \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, we have for $\lambda = (1 - \varepsilon)d^{-1}$ and $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi, m)^l$ that*

(6.2)
$$\mathbb{P}_{\mathrm{GW}} \left(S(\mathcal{H}^{m,l}) \geq t \right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq \frac{2}{\varepsilon};$$

$$\mathbb{P}_{\mathrm{GW}} \left(M^l(\mathcal{H}^{m,l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t \right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2,$$

where $S(\mathcal{H}^{m,l})$, $M^l(\mathcal{H}^{m,l})$ are given as in Definition 6.3.

The main idea of the proof is the same as what we saw in Theorems 4.4 and 5.1. However, as one can easily expect, the analysis becomes much more complicated due to the existence of a cycle. In particular, we cannot apply the tree recursion techniques directly. Due to its technicality, the proof of Proposition 6.4 is presented in Appendix A.

Next, we introduce the *Galton-Watson-on-cycle process of type two* (in short, GWC^2-process), which can be thought of as a certain subgraph of GWC^1-processes. Although it is a very similar object to GWC^1-process, the *root-added process* on GWC^2 is defined in a different way, as presented in the following definitions.

Definition 6.5 (Galton-Watson-on-cycle process of type two)**.** Let ξ be a positive, integer-valued random variable, and let $m \geq 2$, $l \geq 1$ be integers. Then, $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, m)^l$, 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 $C = v_1 v_2 \ldots v_m v_1$.
2. At each v_j, $j \in \{2, \ldots, m\}$, attach $\mathcal{T}_j^l \sim$ i.i.d. $\mathsf{GW}(\xi)^l$ by setting v_j as its root. At v_1, we do nothing.

The resulting graph is called $\dot{\mathcal{H}}$. We designate vertex v_1 as the root of $\dot{\mathcal{H}}$ and denote $\rho = v_1$. Figure 3 illustrates an example of GWC^2-process.

Remark 6.6. The GWC^2-process is the same object as the GWC-process defined in [3, Section 4].

Definition 6.7 (Root-added contact process on GWC^2-processes)**.** For $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, m)^l$, we define the *root-added contact process* on $\dot{\mathcal{H}}$ without adding a new parent to the root. Instead, we fix the root $\rho = v_1$ to be permanently infected by itself, and we denote this process by $\mathbf{CP}_{v_1}^\lambda(\dot{\mathcal{H}}^{m,l}; x)$, for an initial configuration
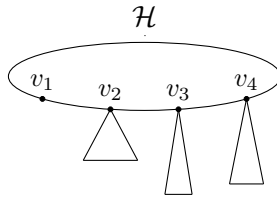
FIGURE 3. An example of $\mathcal{H} \sim \mathsf{GWC}^2(\xi, m)^l$ with $m = 4$

$x \in \{0,1\}^{\dot{\mathcal{H}} \backslash \{v_1\}}$. Then, we define $S_2(\dot{\mathcal{H}})$, $M_2^l(\dot{\mathcal{H}})$ (resp., $S_m(\dot{\mathcal{H}})$, $M_m^l(\dot{\mathcal{H}})$) analogously as Definition 6.3, with respect to $\mathsf{CP}_{v_1}^\lambda(\dot{\mathcal{H}}; \mathbf{1}_{v_2})$ (resp., $\mathsf{CP}_{v_1}^\lambda(\dot{\mathcal{H}}; \mathbf{1}_{v_m})$). We also write

$$S(\dot{\mathcal{H}}) = \frac{1}{2}\left(S_2(\dot{\mathcal{H}}) + S_m(\dot{\mathcal{H}})\right), \quad \text{and} \quad M^l(\dot{\mathcal{H}}) = \frac{1}{2}\left(M_2^l(\dot{\mathcal{H}}) + M_m^l(\dot{\mathcal{H}})\right).$$

Then, the analog of Proposition 6.4 can be derived on GWC^2-process as follows.

Corollary 6.8. *Let* $m \geq 2$, $l \geq 1$ *be integers,* $\varepsilon \in (0,1)$ *and* $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ *be a collection of positive constants. Then there exists* $d_0(\varepsilon, \mathfrak{c}) > 0$ *such that the following holds true. For any* ξ *that satisfies* $d := \mathbb{E}\xi \geq d_0$ *and* (1.2) *with* $\mathfrak{c}$, *we have for* $\lambda = (1 - \varepsilon)d^{-1}$ *and* $\dot{\mathcal{H}}^{m,l} \sim \mathsf{GWC}^2(\xi, m)^l$ *that*

(6.3)
$$\mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}^{m,l}) \geq t\right) \leq 4t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq \frac{3}{\varepsilon};$$
$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}^{m,l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq 4t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 3,$$

where $S(\dot{\mathcal{H}}^{m,l})$, $M^l(\dot{\mathcal{H}}^{m,l})$ *are given as in Definition* 6.7.

As before, we explain the proof of the corollary in Appendix A.3 due to its technicality.

6.2. **Coupling the local neighborhood of random graphs.** Let $G_n \sim \mathcal{G}(n, \mu)$ and $\widetilde{\mu}$ be the size-biased distribution of μ. Define $\mathsf{GW}(\mu, \widetilde{\mu})^l$ to be the Galton-Watson tree of depth l such that the offspring distribution of the root is μ while that of all other descendants is $\widetilde{\mu}$.

For a fixed vertex v in G_n, $N(v, l)$ converges locally weakly to $\mathsf{GW}(\mu, \widetilde{\mu})^l$ as $n \to \infty$, as we briefly saw in Lemma 2.1 and the explanation below it. However, the standard coupling between $N(v, l)$ and $\mathsf{GW}(\mu, \widetilde{\mu})^l$ always has an error at least $\Theta(n^{-1})$. To diminish this error, we introduce the notion of *augmented distribution* (Definition 4.2, [3]), which allows us to stochastically dominate $N(v, l)$ by a larger geometry. For our purpose, the definition will be slightly different from [3] to yield a smaller distortion of the mean of μ. Note that if $\{p_k\}_{k \in \mathbb{N}}$ satisfies $\sum_k p_k e^{ck} < \infty$ for some $c > 0$, then it also has $\sum_k k\sqrt{p_k} < \infty$ by Cauchy-Schwarz inequality (see, for instance, Lemma 4.1 of [3]).

Definition 6.9 (Augmented distribution, [3]). Let $\varepsilon \in (0,1)$ and $\mu \equiv \{p_k\}_{k \in \mathbb{N}}$ be a probability measure on $\mathbb{N}$ such that $\mathbb{E}_{D \sim \mu} e^{cD} < \infty$ for some $c > 0$. Let $k_0 := \max\{k : \sum_{j \geq k} j\sqrt{p_j} \geq \varepsilon/10\}$ (which is finite by the above discussion), and $k_{\max} := \max\{k : p_k > 0\}$, with $k_{\max} = +\infty$ if the maximum does not exist. If

$k_0 < k_{\max}$, we define the *augmented distribution* $\mu^\sharp = \mu^\sharp_\varepsilon$ of μ by

$$\mu^\sharp(k) := \frac{1}{Z} \begin{cases} \left(1 - \frac{\varepsilon}{10}\right) p_k & \text{if } k \leq k_0; \\ \sqrt{p_k} & \text{if } k > k_0, \end{cases}$$

where $Z = \sum_{k \leq k_0} (1 - \frac{\varepsilon}{10}) p_k + \sum_{k > k_0} \sqrt{p_k}$ is the normalizing constant. When $k_0 = k_{\max}$, we set

$$\mu^\sharp(k) := \frac{1}{Z} \begin{cases} \left(1 - \frac{\varepsilon}{10}\right) p_k & \text{if } k < k_0; \\ \sqrt{p_k} & \text{if } k = k_0, \end{cases}$$

for the normalizing constant $Z = \sum_{j < k_0} (1 - \frac{\varepsilon}{10}) p_k + \sqrt{p_{k_0}}$.

Suppose that we generated the i.i.d. degrees $\{D_i\}_{i \in [n]}$ of G_n. Consider the exploration procedure starting from a fixed vertex v and its half-edges, which picks an unmatched half-edge in the *explored* neighborhood of v, reveals its pair half-edge uniformly at random among all the *unmatched* half-edges and absorbs the matched vertex. If the matched half-edge is from the *unexplored* half-edges, then we include the half-edges adjacent to the matched vertex in our explored neighborhood and its (unmatched) half-edges into the explored neighborhood. Then, during the early steps of exploration, the number of newly added (unmatched) half-edges is roughly distributed as $\widetilde{\mu}$, as long as we discover a new half-edge out of the previous explored neighborhood.

However, as we mentioned at the beginning of this subsection, their exact distributions are not precisely $\widetilde{\mu}$. The role of the augmented distribution is to provide a unified law that stochastically dominates the number of newly explored half-edges in all the early steps. The following lemma describes this property.

Lemma 6.10 (Lemma 4.3, [3])**.** *Let $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then, there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds. For any probability measure ν on $\mathbb{N}$ having $d := \mathbb{E}_{D \sim \nu} D \geq d_0$ and the concentration condition* (1.2) *with $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$,*

1. *Let $\nu^\sharp := \nu^\sharp_\varepsilon$ and $d^\sharp := \mathbb{E}_{D^\sharp \sim \nu^\sharp} D^\sharp$ be its mean. Then,*

$$d^\sharp \leq \left(1 + \frac{\varepsilon}{9}\right) d.$$

   Further, let $\widetilde{d}$ (resp. $\widetilde{d^\sharp}$) be the mean of $\widetilde{\nu}$ (resp. $\widetilde{\nu^\sharp}$), the size distribution of ν (resp. $\nu^\sharp$). Then, $\widetilde{d^\sharp} \leq (1 + \frac{\varepsilon}{9}) \widetilde{d}$.

2. *There is a collection of positive constants $\{c'_{\varepsilon'}\}_{\varepsilon' \in [\frac{\varepsilon}{10}, 1]}$ such that $\nu^\sharp$ satisfies*

$$(6.4) \qquad \begin{aligned} \mathbb{P}_{D^\sharp \sim \nu^\sharp} \left(D^\sharp \geq (1 + \varepsilon') d\right) &\leq \exp\left(-c'_{\varepsilon'} d\right), \quad \text{for all } \varepsilon' \in \left[\frac{\varepsilon}{10}, 1\right]; \\ \mathbb{P}_{D^\sharp \sim \nu^\sharp} \left(D^\sharp \geq (1 + a) d\right) &\leq \exp\left(-c'_1 a d\right), \quad \text{for all } a \geq 1. \end{aligned}$$

3. *Let $\{D_i\}_{i \in [n]}$ be a collection of i.i.d. samples of ν. For a subset $\Delta \subset [n]$, let $\{p^\Delta_k\}_{k \in \mathbb{N}}$ denote the empirical distribution of $\{D_i\}_{i \in [n] \setminus \Delta}$. Whp over the choice of $\{D_i\}_{i \in [n]}$, $\{p^\Delta_k\}_{k \in \mathbb{N}}$ is stochastically dominated by $\nu^\sharp$, for any $\Delta \in [n]$ with $|\Delta| \leq n^{\frac{1}{2}}$.*

The third property in the lemma is almost analogous to that of Lemma 4.3, [3]. In Section B.4, we discuss the proof of the lemma, focusing on the aspects which are different from [3].

Now we want to derive a coupling between the local neighborhoods of $G_n \sim \mathcal{G}(n, \mu)$ and the Galton-Watson-type processes. The first thing we should handle is to control the number of cycles in a local neighborhood $N(v, l)$. For a fixed constant $\gamma > 0$, we define the event $\mathcal{E}_n^C(\gamma)$ over the graphs with n vertices to be

$$\mathcal{E}_n^C(\gamma) := \{G_n : \forall v \in G_n, \ N(v, \gamma \log n) \text{ contains at most one cycle}\}.$$

Then, we adopt the following lemma that shows the event $\mathcal{E}_n^C(\gamma)$ is indeed typical for some $\gamma > 0$. The proof of the lemma is presented in Section B.5 in the Appendix, due to its similarity to [3, Lemma 4.5] and [15, Lemma 2.1].

Lemma 6.11 ([3,15])**.** *Let $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then, there exists a constant $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds. For any probability measure μ on $\mathbb{N}$ having $d := \mathbb{E}_{D \sim \mu} D \geq d_0$ and the concentration condition (1.2) with $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$, we have for $G_n \sim \mathcal{G}(n, \mu)$ that*

(6.5) $$\mathbb{P}\left(G_n \in \mathcal{E}_n^C\left(\frac{1}{10 \log d}\right)\right) = 1 - o(1).$$

Remark 6.12. In [3, Lemma 4.5], it was proven that $\mathbb{P}(G_n \in \mathcal{E}_n^C(\gamma)) = 1 - o(1)$ for some constant γ depending on d in a rather implicit sense. Here, we have an additional assumption of (1.2) and it makes it possible to deduce Lemma 6.11, a stronger result. The improved information on the constant γ turns out to be crucial in Section 6.3 when proving Theorem 6.1.

Set $\gamma_1 = \frac{1}{10}$ and $l_n := \frac{\gamma_1}{2} \log_d n$. The remaining task of this subsection is to define the Galton-Watson type process which will be coupled with the local neighborhoods $N(v, l_n)$. The notion was already introduced in [3, Definition 4.7].

Definition 6.13 (Edge-added Galton-Watson process, [3])**.** Let h, m, l be nonnegative integers such that $m \geq 2$ and $l \geq h + 1$, and let ν be a probability distribution on $\mathbb{N}$. We define the *Edge-added Galton-Watson process* (in short, **EGW**-*process*), denoted by **EGW**$(\nu; h, m)^l$ as follows.

1. Generate a **GW**$(\nu)^l$ tree, conditioned on survival until depth l. The root ρ of this tree is also the root of **EGW**$(\nu; h, m)^l$.
2. At each vertex u at depth h, add an independent **GWC**$^2(\nu, m)^{l-h}$ process (see Definition 6.5) rooted at u. Here we preserve the existing subtrees at u which comes from **GW**$(\nu)^l$ tree from Step 1.

Let ν' be another probability measure on $\mathbb{N}$. Then, **EGW**$(\nu, \nu'; h, m)^l$ denotes the **EGW**-process whose root has offspring distribution ν while all other descendants have ν'. Here we also add **GWC**$^2(\nu', m)^{l-h}$ in Step 2 of the definition. We also remark that **EGW**$(\nu; 0, m)^l = $ **GWC**$^1(\nu, m)^l$. An example of **EGW**-process is illustrated in Figure 4.

We develop a coupling between $N(v, \gamma_1 \log_d n)$ and the **EGW**-processes. To this end, we define the notion of stochastic domination between two probability measures η, η' over rooted graphs: we write $\eta \leq_{\text{ST}} \eta'$ if there exists a coupling $((\mathcal{G}, \rho), (\mathcal{G}', \rho'))$ such that $(\mathcal{G}, \rho) \sim \eta$, $(\mathcal{G}', \rho') \sim \eta'$ and $\mathcal{G} \subset \mathcal{G}'$, that is, there is a graph isomorphism that maps ρ to ρ' and embeds $\mathcal{G}$ into $\mathcal{G}'$.

Let v be a fixed vertex in $G_n \sim \mathcal{G}(n, \mu)$, and recall that $\gamma_1 = \frac{1}{10}$ and $l_n = \frac{\gamma_1}{2} \log_d n$. We also abbreviate $\mathcal{E}_n^C = \mathcal{E}_n^C(\gamma_1 / \log d)$. In addition to $\mathcal{E}_n^C$, we define $\mathcal{B}_{h,m}(v)$ to be the subevent of $\mathcal{E}_n^C$ such that $N(v, l_n)$ contains a cycle of length m that is at distance h from v.
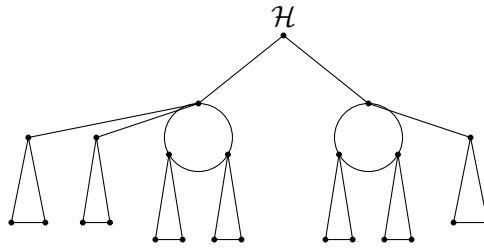
FIGURE 4. An example of $\mathcal{H} \sim \mathbf{EGW}(\nu; h, m)^l$ with $m = 3$ and $h = 1$

Moreover, for a given $\varepsilon \in (0, 1)$ we set $\mu^\sharp = \mu^\sharp_\varepsilon$ to be the augmented distribution of μ, and $\widetilde{\mu}^\sharp$ to be the size-biased distribution of $\mu^\sharp$. Further, let η, η_0 and $\eta_{h,m}$ denote the law of $N(v, l_n)$, $\mathbf{GW}(\mu^\sharp, \widetilde{\mu}^\sharp)^{l_n}$ and $\mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^{l_n}$, respectively. Then, Lemma 4.8 of [3] provides us the following coupling lemma, which plays a crucial role in settling the lower bound of Theorem 2 in the following subsection.

Lemma 6.14 (Lemma 4.8, [3]). *Under the above setting, let* $b_{h,m} := \mathbb{P}(\mathcal{B}_{h,m}(v))$ *and* $b_0 := 1 - \sum_{h,m} b_{h,m}$. *Then we have*

$$\eta \mathbf{1}_{\mathcal{E}_n^C} \leq_{\mathrm{ST}} b_0 \eta_0 + \sum_{h \geq 0, m \geq 2} \eta_{h,m}.$$

So far, we have collected almost all the elements we need in the proof of Theorem 2. The last thing we want is the tail probability estimates for $\mathbf{EGW}$-processes. To be concrete, we begin with defining the quantities of interest.

Definition 6.15 (Excursion time and total infections at leaves for $\mathbf{EGW}$). Let h, m, l be nonnegative integers with $m \geq 2$, ν, ν' be probability measures on $\mathbb{N}$, and $\mathcal{H} \sim \mathbf{EGW}(\nu, \nu'; h, m)^l$. We connect a permanently infected parent ρ^+ to the root ρ of $\mathcal{H}$, and the contact process on the resulting graph is the *root-added contact process* $\mathbf{CP}^\lambda_{\rho^+}(\mathcal{H}; x)$ on $\mathcal{H}$ with initial condition $x \in \{0, 1\}^\mathcal{H}$.

- The *excursion time* $\mathbf{S}(\mathcal{H})$ is the first time when $\mathbf{CP}^\lambda_{\rho^+}(\mathcal{H}; \mathbf{1}_\rho)$ becomes all-healthy on $\mathcal{H}$, and $S(\mathcal{H}) = \mathbb{E}_{\mathrm{CP}} \mathbf{S}(\mathcal{H})$ 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 $\mathcal{H}$,
$$\mathcal{L} = \{v \in \mathcal{H} : \mathrm{dist}(v, \rho) \geq l \text{ and } \mathrm{dist}(v, C_j) \geq l - h \text{ for all } j\}.$$

- Let $v \in \mathcal{L}$. Denoting $(X_t) \sim \mathbf{CP}^\lambda_{\rho^+}(\mathcal{H}; \mathbf{1}_\rho)$, we define the *total infections at* v by
$$\mathbf{M}^l_v(\mathcal{H}) := |\{s \in [0, \mathbf{S}(\mathcal{H})] : X_s(v) = 1 \text{ and } X_{s-}(v) = 0\}|.$$
Then, we set the *total infections at depth-l leaves* as
$$\mathbf{M}^l(\mathcal{H}) := \sum_{v \in \mathcal{L}} \mathbf{M}^l_v(\mathcal{H}).$$

- We also let $M^l(\mathcal{H}) := \mathbb{E}_{\mathrm{CP}} \mathbf{M}^l(\mathcal{H})$ to be the *expected total infections at leaves*.

For an $\mathbf{EGW}$-process $\mathcal{H}$, the tail probabilities of $S(\mathcal{H})$ and $M^l(\mathcal{H})$ can be estimated using Theorems 4.4, 5.1, and Proposition 6.4, which can be described as follows.

Proposition 6.16. *Let m, l, h be nonnegative integers such that $m \geq 2$ and $l \geq h + 1$. Also, let $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then, there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any μ that satisfies $d := \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, let $\mu^\sharp := \mu_\varepsilon^\sharp$, and set $\widetilde{\mu}^\sharp$ to be its size-biased distribution (Definition 6.9). Then, we have for $\lambda = (1 - \varepsilon)d^{-1}$ and $\mathcal{H} \sim \mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^l$ that*

(6.6)
$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}) \geq t\right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq \frac{2}{\varepsilon};$$

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2,$$

where $S(\mathcal{H}^{m,l})$, $M^l(\mathcal{H}^{m,l})$ are given as in Definition 6.15.

Proof. Along with (6.6), we establish the same inequalities for $\mathcal{H} \sim \mathsf{EGW}(\widetilde{\mu}^\sharp; h, m)^l$ in tandem, by an induction on h. When $h = 0$, we have (6.6) for both $\mathsf{EGW}(\widetilde{\mu}^\sharp; 0, m)^l$ and $\mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^l$, since the former one is the same as $\mathsf{GWC}^1(\widetilde{\mu}^\sharp, m)^l$ and the latter is $\mathcal{H} \sim \mathsf{GWC}^1(\mu^\sharp, \widetilde{\mu}^\sharp, m)^l$, the GWC^1-process in which only the root of the Galton-Watson tree rooted at v_1 has offspring $\mu^\sharp$ and all others have offspring $\widetilde{\mu}^\sharp$. By Lemma 6.10, the mean $\widetilde{d} := \mathbb{E}_{D^\sharp \sim \widetilde{\mu}^\sharp} D^\sharp$ of $\widetilde{\mu}^\sharp$ is such that

$$\widetilde{d} \leq \left(1 + \frac{\varepsilon}{9}\right) d,$$

and $\widetilde{\mu}^\sharp$ also satisfies (6.4) with $\mathfrak{c}'$ as in the lemma. Note that this implies (1.2) for $\widetilde{\mu}^\sharp$. Therefore, Proposition 6.4 gives (6.6) for $h = 0$ and all $m \geq 2$, $l \geq 0$.

Suppose that we have (6.6) for $\mathcal{H}' \sim \mathsf{EGW}(\widetilde{\mu}^\sharp; h, m)^l$ with $m \geq 2$, $h \geq 0$ and $l \geq h + 1$. Let ρ_1 be the root of $\mathcal{H}_1 \sim \mathsf{EGW}(\widetilde{\mu}; h + 1, m)^l$ and $D_1 \sim \widetilde{\mu}^\sharp$ be its degree. Then the subgraphs $H_1^1, \ldots, H_{D_1}^1$ rooted at each child of v_1 are i.i.d. $\mathsf{EGW}(\widetilde{\mu}^\sharp; h, m)^{l-1}$. Since we have the tail probability estimates for $H_1^1, \ldots, H_{D_1}^1$ by the induction hypothesis, we follow the proof of Theorems 4.4 (Sections 4.2.1—4.2.3) and 5.1 (Sections 5.3.1—5.3.3) to deduce (6.6) for $\mathcal{H}_1$. The result for $\mathcal{H}_2 \sim \mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h + 1, m)^l$ follows similarly, where the only difference is that $\mathcal{H}_2$ has subgraphs $H_1^2, \ldots, H_{D_2}^2 \sim$ i.i.d. $\mathsf{EGW}(\widetilde{\mu}^\sharp; h, m)^{l-1}$ with $D_2 \sim \mu^\sharp$. $\square$

6.3. **Proof of Theorem 2.** In this subsection, we conclude the proof of the lower bound of Theorem 2, by showing Theorem 6.1.

Let $\varepsilon \in (0, 1)$ and a collection of positive constants $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be given, and let $d_0 = d_0(\varepsilon, \mathfrak{c})$ be the maximal among those from Theorems 4.4, 5.1 and Proposition 6.16. For μ and its size-biased distribution with mean $d := \mathbb{E}_{D \sim \mu'}$, let $\gamma_1 \in (0, \frac{1}{10})$ be a constant satisfying Lemma 6.11, that is,
(6.7)
$$\mathbb{P}_{G_n \sim \mathcal{G}(n,\mu)}\left(G_n \in \mathcal{E}_n^C\right) = 1 - o(1), \quad \text{with } \mathcal{E}_n^C = \mathcal{E}_n^C\left(\frac{\gamma_1}{\log d}\right) \text{ as in the lemma,}$$

and set $l_n = \lfloor \frac{\gamma_1}{2} \log_d n \rfloor$.

In the remaining of this section, we assume that $G_n \sim \mathcal{G}(n, \mu)$ satisfies $\mathcal{E}_n^C$. For each vertex $v \in G_n$, we define the *block* B_v at v as follows.

- $B_v := N(v, l_n)$, if $N(v, l_n)$ is a tree.

- If $N(v, l_n)$ contains the unique cycle C at distance h from v, then

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

Note that we always have $B_v \subset N(v, \gamma_1 \log_d n)$, so that B_v can only have at most one cycle.

Let $\mu^\sharp = \mu_\varepsilon^\sharp$ be the augmented distribution (Definition 6.9) of μ and $\widetilde{\mu}^\sharp$ be its size-biased distribution. Also, set $\eta_0, \eta_{h,m}$ to be the laws of $\mathsf{GW}(\mu^\sharp, \widetilde{\mu}^\sharp)^{l_n}$, $\mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^{l_n}$, respectively, and let $\mathcal{H}$ be a sample from

$$\mathcal{H} \sim b_0 \eta_0 + \sum_{h \geq 0, m \geq 2} \eta_{h,m},$$

where the constants b_0 and $\{b_{h,m}\}_{h \geq 0, m \geq 2}$ are given as in Lemma 6.14. Then, Lemma 6.14 tells us that for fixed v, we can couple $N(v, l_n)$ and $\mathcal{H}$ so that $N(v, l_n) \subset \mathcal{H}$ (in terms of isomorphic embeddings). Furthermore, the same proof as in the lemma gives the coupling between B_v and $\mathcal{H}$ as follows. We refer [3, Lemma 4.8] for its proof.

Corollary 6.17. *Under the above setting, let v be a fixed vertex in G_n and η' denote the law of B_v. Then, we have*

$$(6.8) \qquad \eta' \mathbf{1}_{\mathcal{E}_n^C} \leq_{\mathrm{ST}} b_0 \eta_0 + \sum_{h \geq 0, m \geq 2} \eta_{h,m}.$$

Further, let us consider the collections of the *bottom leaves* of B_v. To this end, let C denote the cycle in B_v if exists, and let
(6.9)
$$\mathcal{L}(B_v) := \begin{cases} \{u \in B_v : \mathrm{dist}(u, v) \geq l_n\} & \text{if } C \text{ does not exist;} \\ \{u \in B_v : \mathrm{dist}(u, v) \geq l_n \text{ and } \mathrm{dist}(u, C) \geq l_n - h\} & \text{if } C \text{ exists,} \end{cases}$$

where $h = \mathrm{dist}(v, C)$. Also, set $\mathcal{L}(\mathcal{H})$ to be the bottom leaves of $\mathcal{H}$, where we follow Definition 6.15 if $\mathcal{H}$ is an EGW-process. An important thing to note is that if $B_v \subset \mathcal{H}$, then we have

$$\mathcal{L}(B_v) \subset \mathcal{L}(\mathcal{H}),$$

in terms of isomorphic embedding. That is, there exists an isomorphism ϕ mapping B_v into $\mathcal{H}$ such that $\phi(v) = \rho$, $\phi(\mathcal{L}(B_v)) \subset \mathcal{L}(\mathcal{H})$.

Since $\mathcal{H}$ is a mixture of Galton-Watson and EGW-processes, Theorems 4.4, 5.1 and Proposition 6.16 imply the tail probabilities of $S(\mathcal{H})$ and $M^{l_n}(\mathcal{H})$. Thus, $R(B_v)$ and $\bar{M}^{l_n}(B_v)$, the expected survival time and the expected total infections at leaves with respect to $\mathsf{CP}^\lambda(B_v; \mathbf{1}_v)$ (Definition 6.15), satisfy

$$\mathbb{P}(R(B_v) \geq t) \leq \mathbb{P}(S(\mathcal{H}) \geq t) \leq 3t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for all } t \geq \frac{2}{\varepsilon};$$

$$(6.10) \qquad \mathbb{P}\left(\bar{M}^{l_n}(B_v) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} t \right)$$

$$\leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} t \right) \leq 3t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for all } t \geq 2.$$

Based on the above discussions, we can deduce the following lemma.

Lemma 6.18. *Under the above setting, define the events $\mathcal{E}_n^R$ and $\mathcal{E}_n^M$ over the graphs with n vertices by*

$$\mathcal{E}_n^R := \{R(B_v) \leq n^{2/\sqrt{d}}, \text{ for all } v \in G_n\}$$
$$\mathcal{E}_n^M := \{\bar{M}^{l_n}(B_v) \leq (\log n)^{-1}, \text{ for all } v \in G_n\}.$$

Then, $\mathbb{P}(G_n \in \mathcal{E}_n^R \cap \mathcal{E}_n^M) = 1 - o(1)$.

Proof. Note that (6.10) implies

$$\mathbb{P}\left(R(B_v) \geq n^{2/\sqrt{d}}\right) \leq n^{-2},$$

and also

$$\mathbb{P}\left(\bar{M}^{l_n}(B_v) \geq \frac{1}{\log n}\right) \leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} \frac{\exp(\frac{\varepsilon\gamma_1}{20}\log_d n)}{\log n}\right)$$

$$\leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} \exp\left(\frac{\varepsilon\gamma_1}{30}\log_d n\right)\right)$$

$$\leq \exp\left(-\frac{\varepsilon\gamma_1\sqrt{d}}{30\log d}\log n\right) \leq n^{-2},$$

where the second and the last inequalities are true for d larger than some constant depending on ε (note that γ_1 was an absolute constant). Therefore, the conclusion follows from a union bound over all vertices. □

We conclude this subsection by establishing Theorem 6.1.

Proof of Theorem 6.1. Let $\mathcal{E}_n^C, \mathcal{E}_n^R$ and $\mathcal{E}_n^M$ be the events over graphs of n vertices defined in Lemmas 6.11 and 6.18. Further, set $\gamma_1 \in (0, \frac{1}{10})$ as in (6.7) and define

$$\mathcal{E}_n := \mathcal{E}_n^C \cap \mathcal{E}_n^R \cap \mathcal{E}_n^M.$$

Then, Lemmas 6.11 and 6.18 tell us that

$$\mathbb{P}(G_n \in \mathcal{E}_n) = 1 - o(1).$$

In the remaining proof, we settle (6.1) for $\mathcal{E}_n$ as above.

Let $\mathbf{R}_v(G_n)$ be the survival time of $\mathsf{CP}^\lambda(G_n; \mathbf{1}_v)$. Then, the standard coupling argument (also called graphical representation, see, for instance, Section 2 of [3] for a brief review, and [14] for a detailed introduction) tells us that

$$(6.11) \qquad \mathbf{R}(G_n) = \max\{\mathbf{R}_v(G_n) : v \in G_n\} \leq \sum_{v \in G_n} \mathbf{R}_v(G_n).$$

To investigate $\mathbf{R}_v(G_n)$, we introduce a *decomposition* of $\mathsf{CP}^\lambda(G_n; \mathbf{1}_v)$ using the *blocks* $\{B_u\}_u$. Recall the graphical representation from Section 2.1 and let $\{N_x(t)\}_{x \in V(G_n)}$ and $\{N_{\vec{x}\vec{y}}(t)\}_{\vec{x}\vec{y} \in \vec{E}(G_n)}$ be the Poisson processes defining the recoveries and infections of $\mathsf{CP}^\lambda(G_n; \mathbf{1}_v)$, respectively. Also, recall the definition of the bottom leaves $\mathcal{L}(B_u)$ defined in (6.9).

Definition 6.19 (Decomposition). Let G_n and $\{B_u\}_{u \in G_n}$ be as above. The *decomposition* of $\mathsf{CP}^\lambda(G_n; \mathbf{1}_v)$ by $\{B_u\}_u$ is the coupled process generated as follows.

   1. Initially, run $\mathsf{CP}^\lambda(B_v; \mathbf{1}_v)$ whose recoveries and infections are given by
      $\{(N_x(s))_{s \geq 0}\}_{x \in V(B_v)}$ and $\{(N_{\vec{x}\vec{y}}(s))_{s \geq 0}\}_{\vec{x}\vec{y} \in \vec{E}(B_v)}$.

2. In B_v, when some $u \in \mathcal{L}(B_v)$ becomes infected in B_v at time t (and has been healthy until time $t-$), initiate a copy of $\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)$ with event times given by $\{(N_x(s))_{s \geq t}\}_{x \in V(B_u)}$ and $\{(N_{\vec{xy}}(s))_{s \geq t}\}_{\vec{xy} \in \vec{E}(B_u)}$.

3. Repeat Step 2 to every running copies of $\{\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)\}_{u \in G_n}$ until the process terminates, that is, when all vertices in every generated copy are healthy.

Since the copies generated in this definition share the event times with the original process, it is just another way of interpreting $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$. In particular, both $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$ and its decomposition terminate at the same time. Also, even though the generated copies can be highly correlated with each other, the law of each copy is the same as the contact process with intensity λ. In this approach, we can see that controlling the number of infections at $\mathcal{L}(B_u)$ is crucial, which is done by the second and third lines of (6.10).

Let $\widetilde{\mathbf{R}}_v(G_n)$ denote the termination time of the decomposed process. Since the copies in the above procedure share the event times with the original process, we see that $\widetilde{\mathbf{R}}_v(G_n) = \mathbf{R}_v(G_n)$ and hence

$$(6.12) \qquad R_v(G_n) := \mathbb{E}_{\mathrm{CP}} \mathbf{R}_v(G_n) = \widetilde{R}_v(G_n) := \mathbb{E}_{\mathrm{CP}} \widetilde{\mathbf{R}}_v(G_n).$$

Further, let $\mathbf{U}_v(G_n)$ be the enumeration of vertices u of which a copy of $\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)$ has been generated during this process. If the copies of $\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)$ for u appeared multiple times, we include u multiple times in $\mathbf{U}_v(G_n)$, respecting the multiplicities. Then, we observe that

$$(6.13) \qquad \widetilde{\mathbf{R}}_v(G_n) \leq \sum_{u \in \mathbf{U}_v(G_n)} \mathbf{R}_u(B_u),$$

where we have a slight abuse of notation in the r.h.s. since we omitted the initiation times of the copies $\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)$, $u \in \mathbf{U}_v(G_n)$.

On the event $\mathcal{E}_n^M$, the expected number of new copies generated in Step 2 and 3 from a single $\mathbf{CP}^\lambda(B_u; \mathbf{1}_u)$ is at most $(\log n)^{-1}$, uniformly in u. This gives that

$$\mathbb{E}_{\mathrm{CP}} \mathbf{U}_v(G_n) \leq \sum_{k=0}^\infty (\log n)^{-k} \leq 2,$$

for large enough n. Moreover, on $\mathcal{E}_n^R$, each generated copy survives at most $n^{2/\sqrt{d}}$-time in expectation, so we have from (6.12) and (6.13) that for all $v \in G_n$,

$$R_v(G_n) \leq \mathbb{E}_{\mathrm{CP}} \widetilde{\mathbf{R}}_v(G_n) \leq 2n^{2/\sqrt{d}}.$$

Thus, (6.11) gives $R(G_n) \leq 2n^{1+2d^{-1/2}}$, and hence by Markov's inequality we get

$$\mathbb{P}_{\mathrm{CP}}\left(\mathbf{R}(G_n) \geq n^2 \,\big|\, G_n \in \mathcal{E}_n \right) \leq \frac{1}{\sqrt{n}} = o(1),$$

for any d larger than some absolute constant. This finishes the proof of Theorem 6.1. $\qquad \square$

7. Supercritical phase for general distribution: Proof of Theorems 5 and 6

In this section, we establish Theorems 5 and 6. We first observe that the phase transitions for $\mathbf{GW}(\widetilde{\mu})$ and $\mathbf{GW}(\mu, \widetilde{\mu})$ occur at the same intensity (Lemma 7.1) and then show the second inequality in (1.5) (Lemma 7.2). After verifying the two

lemmas, we prove the first inequality in (1.5) and Theorem 6, which is the main goals of this section. Finally, in Section 7.6, we combine the main results obtained throughout the paper and conclude the proof of Theorems 1 and 2.

Lemma 7.1. *We have*

$$\lambda_1^{\mathrm{GW}}(\widetilde{\mu}) = \lambda_1^{\mathrm{GW}}(\mu, \widetilde{\mu}).$$

Proof. Let $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$. The contact process on $\mathsf{GW}(\mu, \widetilde{\mu})$ with rate λ and the root initially infected has a positive probability of reaching a vertex in the first generation. Once such a vertex is infected, the contact process on its subtree has a positive probability to survive forever. Thus, $\lambda > \lambda_1^{\mathrm{GW}}(\mu, \widetilde{\mu})$.

On the other hand, let $\lambda < \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$. Let τ be the (random) survival time of the contact process on $\mathsf{GW}(\widetilde{\mu})$ with the root initially infected. We have

$$\mathbb{P}(\tau < \infty) = 1.$$

Let $m \sim \mu$ be the number of vertices in the first generation of $\mathcal{T} \sim \mathsf{GW}(\mu, \widetilde{\mu})$. Let (X_t) be the contact process with rate λ on $\mathcal{T}$ and with the root initially infected. Conditioned on m, if $\mathcal{T}$ remains infected at time t, the probability that $\mathcal{T}$ is healed in a finite time before any infection between the root and its children happens is at least

$$\mathbb{P}_{\tau'}\left(\mathbf{Ber}(e^{-2m\lambda\tau'})\right) > 0,$$

where $\tau' = \max\{\tau_0, \tau_1, \ldots, \tau_m\} < \infty$ a.s., $\tau_0 \sim \exp(1)$ upper bounds the survival time of the root and $\tau_1, \ldots, \tau_m \overset{iid}{\sim} \tau$ upper bound the survival time of the contact process on the m subtrees rooted at the vertices in the first generation after time t. Thus, (X_t) dies out with probability 1, proving $\lambda < \lambda_1^{\mathrm{GW}}(\mu, \widetilde{\mu})$ and completing the proof. $\qquad\square$

So, for the rest of the proof, we simply write λ_1 for $\lambda_1^{\mathrm{GW}}(\widetilde{\mu})$ and $\lambda_1^{\mathrm{GW}}(\mu, \widetilde{\mu})$. Next, we prove the second inequality in (1.5).

Lemma 7.2. *Let μ be a degree distribution satisfying the giant component condition (2.2), and let $\widetilde{\mu}$ be its size-biased distribution with $d := \mathbb{E}_{D \sim \widetilde{\mu}} D$. We have*

$$\lambda_1^{\mathrm{GW}}(\widetilde{\mu}) \leq \frac{1}{d-1}.$$

Proof. Let $\lambda > \frac{1}{d-1}$. Let $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu})$ and (X_t) be the contact process on $\mathcal{T}$ with the root ρ initially infected and infection rate λ. Construct a subtree $\mathcal{T}'$ rooted at ρ level by level as follows. Let v be a vertex of $\mathcal{T}$ and let u be its parent. The vertex v belongs to $\mathcal{T}'$ if and only if $u \in X_t \cap \mathcal{T}'$ for some t and after the first time that u is infected, it passes its infection to v before being healed. Observe that $\mathcal{T}'$ is a Galton-Watson tree with branching rate $\frac{\lambda d}{1+\lambda}$ and so, as this branching rate is greater than 1, $\mathcal{T}'$ is infinite with positive probability which means that (X_t) lasts forever with positive probability. In other words, $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$. Hence, we obtain the conclusion. $\qquad\square$

The rest of this section is devoted to the proof of the first inequality in (1.5) and the proof of Theorem 6. Let $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$. We want to show that whp in $G_n \sim \mathcal{G}(n, \mu)$, the contact process on G_n starting with all vertices infected survives until time $e^{\Theta(n)}$ whp and the contact process starting with one random vertex infected survives until time $e^{\Theta(n)}$ with positive probability. Our strategy is to show that whp, if there are many "good" vertices (Definition 7.7) infected at certain

time t, then there are at least twice as many good vertices infected at a later time. Roughly speaking, a good vertex v has a *well-expanding* neighborhood that looks like a Galton-Watson tree. Since λ is in the supercritical regime of the tree, with positive probability, the contact process starting from v infects many vertices on the boundary of that neighborhood.

7.1. Preprocessing the graph.

We first preprocess the graph G_n to get rid of high degree vertices. This will be useful to control the number of explored vertices during our exploration of G_n (we refer to the paragraph following Definition 6.9 for a formal discussion of the term *exploration*). If μ has an infinite support, we let η_0, j_0, and j_1 be positive constants such that $\mu(j_0) > 0$, $j_1 \in [j_0, e^{j_0/10})$, and

$$(7.1) \qquad \mu(j_0, \infty) = \mathbb{P}_{D\sim\mu}(D > j_0) \geq 10\mathbb{E}_{D\sim\mu}D\mathbf{1}_{\{D>j_1\}} =: 10\eta_0.$$

Note that the requirement $j_1 < e^{j_0/10}$ is guaranteed by the assumption that μ has an exponential tail. If μ has a finite support, we let $j_0 = j_1$ be the largest number in the support of μ and η_0 be a sufficiently small constant satisfying

$$(7.2) \qquad \mu(j_0) \geq 10\eta_0.$$

Sample the degrees $D_1, D_2, \ldots, D_n$ of G_n independently according to the measure μ. Let ν be the empirical measure of this sequence (D_i). We have whp,

$$(7.3) \qquad |\nu[0,l] - \mu[0,l]| \leq \eta_0, \quad \forall l \leq j_1 \qquad \text{by Chernoff inequality}$$

and

$$(7.4) \qquad \sum_{i=1}^{n} D_i\mathbf{1}_{\{D_i>j_1\}} \leq 2\eta_0 n \qquad \text{by Chebyshev's inequality.}$$

Consider the subgraph G'_n obtained from G_n by deleting all vertices with degree greater than j_1 together with their matched half-edges. The remaining randomness is the random perfect matching of the remaining half-edges. Let n' be the number of vertices of G'_n. By (7.4), the number of vertices that are affected by this removal is at most $3\eta_0 n$ and in particular, $n' \geq n - 3\eta_0 n$. Let ν' be the empirical measure of the degree sequence of G'_n. We have

$$|\nu'[0,l] - \nu[0,l]| \leq 3\eta_0, \quad \forall l.$$

Combining this with (7.3) gives

$$\text{whp,} \quad |\nu'[0,l] - \mu[0,l]| \leq 4\eta_0, \quad \forall l \leq j_1.$$

Let μ_{η_0,j_0} be the probability measure given by
$$(7.5)$$
$$\mu_{\eta_0,j_0}(A) := \begin{cases} (1 - 4\eta_0)\frac{\mu(A\cap[0,j_0])}{\mu[0,j_0]} + 4\eta_0\mathbf{1}_{\{0\in A\}} & \text{if } \mu \text{ has an infinite support,} \\ \mu(A \cap [0,j_0)) + (\mu(j_0) - 4\eta_0)1_{j_0\in A} + 4\eta_0\mathbf{1}_{\{0\in A\}} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \mu \text{ has a finite support.} \end{cases}$$

By (7.1) and (7.2), we have for every $l \leq j_0$,

$$\nu'[0,l] \leq \max\{\mu[0,l] + 4\eta_0, 1\} \leq \mu_{\eta_0,j_0}[0,l]$$

and so

$$(7.6) \qquad\qquad \text{whp,} \quad \nu' \geq_{st} \mu_{\eta_0,j_0}.$$

Similarly, we have

Lemma 7.3. *With high probability, for any subset Δ of vertices in G'_n with at most $\eta_0 n$ vertices, the empirical measure ν'' of the degree sequence in $G'_n \setminus \Delta$ also satisfies $\nu'' \geq_{st} \mu_{\eta_0, j_0}$.*

Therefore, as long as we have explored at most $\eta_0 n$ vertices, we can use μ_{η_0, j_0} to bound from below the neighborhood of an unexplored vertex.

Let $\widetilde{\mu}_{\eta_0, j_0}$ be the size-biased distribution of μ_{η_0, j_0}. As $\eta_0 \to 0$ (and $j_0 \to \infty$ when μ has an infinite support), $\widetilde{\mu}_{\eta_0, j_0}$ approaches $\widetilde{\mu}$. Let $\tilde{d}$ be the mean of $\widetilde{\mu}_{\eta_0, j_0}$.

7.2. **Good trees.** We define a notion of "good" trees to describe those trees on which the contact process spreads well: the number of infected vertices grows exponentially. As λ is in the supercritical regime of the Galton-Watson tree $\mathsf{GW}(\widetilde{\mu})$, we then show that with positive probability, a random tree is good.

Definition 7.4 (Good trees). A tree T with root ρ is said to be $(s_0, L_0, k, \theta, \varepsilon)$-good if

$$(7.7) \qquad \mathbb{P}_{\mathrm{CP}}\left(|X^{kL_0}_{ks_0} \cap T_{kL_0}| \geq \theta^k\right) > \varepsilon,$$

where $\left(X^{kL_0}_t\right) \sim \mathsf{CP}^\lambda(T_{\leq kL_0}; \mathbf{1}_\rho)$ and T_{kL_0} is the set of vertices at the kL_0-generation.

The following lemma asserts that there is a positive chance for a random tree to be good.

Lemma 7.5. *Let $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu}_{\eta_0, j_0})$. There exist constants $s_0, L_0, \varepsilon > 0$ and $\theta > 1$ such that*

$$(7.8) \qquad \mathbb{P}_{\mathrm{GW}}(\mathcal{T} \text{ is } (s_0, L_0, k, \theta, \varepsilon)\text{-good}) > \varepsilon$$

for all k sufficiently large where the probability is taken over the random graph and the contact process.

For the proof of Lemma 7.5, we first show that

Lemma 7.6. *Let $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu})$. There exist constants $s_0, L_0, \varepsilon > 0$ and $\theta > 1$ such that*

$$(7.9) \qquad \mathbb{E}(|X^{L_0}_{s_0} \cap \mathcal{T}_{L_0}|) \geq \theta,$$

where $(X^{L_0}_t) \sim \mathsf{CP}^\lambda(\mathcal{T}_{\leq L_0}; \mathbf{1}_\rho)$ and the expectation is taken over the random graph and the contact process.

Assuming Lemma 7.6, we prove Lemma 7.5.

Proof of Lemma 7.5. As $\widetilde{\mu}_{\eta_0, j_0} \to \widetilde{\mu}$, it suffices to prove (7.8) for $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu})$.

Let $W_k = |X^{kL_0}_{ks} \cap \mathcal{T}_{kL_0}|$. Then (W_k) dominates a Galton-Watson tree with branching rate $\theta > 1$. Recall that for a Galton-Watson tree (Z_k) with branching rate $\theta > 1$, there exists a nonzero random variable Z such that

$$\frac{|Z_k|}{\theta^k} \to Z \quad a.e.$$

Thus, there exists an $\varepsilon > 0$ such that for sufficiently large k,

$$(7.10) \qquad \mathbb{P}\left(W_k > \varepsilon \theta^k\right) > \varepsilon.$$

Let $F(\mathcal{T}) = \mathbb{P}_{\mathrm{CP}}\left(|X_{kt}^{kL_0} \cap \mathcal{T}_{kL_0}| \geq \varepsilon\theta^k\right)$. We have

$$\mathbb{E}_{\mathrm{GW}}F(\mathcal{T}) > \varepsilon.$$

Since $F(\mathcal{T})$ is a random variable taking values in $[0,1]$, the above inequality implies (7.8) which is

$$\mathbb{P}_{\mathrm{GW}}(F(\mathcal{T}) > \varepsilon/2) > \varepsilon/2$$

because otherwise,

$$\mathbb{E}_{\mathrm{GW}}F(\mathcal{T}) \leq \mathbb{E}_{\mathrm{GW}}F(\mathcal{T})\mathbf{1}_{F(\mathcal{T})\geq\varepsilon/2} + \varepsilon/2 \leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

a contradiction. $\square$

Proof of Lemma 7.6. Let $\lambda' \in (\lambda_1, \lambda)$. Note that (X_t) has the same distribution, up to rescaling the time axis, as a contact process (Y_t) with infection rate λ' and recovery rate $\frac{\lambda'}{\lambda} =: 1-\varepsilon_1$. Let $(Y_t^{L_0})$ be the contact process (Y_t) restricted to $\mathcal{T}_{\leq L_0}$, namely, it only uses the infection and recovery events inside $\mathcal{T}_{\leq L_0}$. It suffices to show that for some s_0, L_0,

$$(7.11) \qquad\qquad \mathbb{E}_{\mathcal{T}}(|Y_{\leq s_0}^{L_0} \cap \mathcal{T}_L|) \geq \theta,$$

where $Y_{\leq s_0}^{L_0} := \bigcup_{t \leq s_0} Y_t^{L_0}$. Indeed, assuming (7.11), we have

$$\mathbb{E}(|X_{\leq s_0}^{L_0} \cap \mathcal{T}_L|) \geq \theta.$$

Let $W_k = |X_{\leq ks}^{kL_0} \cap \mathcal{T}_{kL_0}|$ then by the same argument that derives (7.10), we have

$$(7.12) \qquad\qquad \mathbb{P}\left(W_k > \varepsilon\theta^k\right) > \varepsilon$$

for all sufficiently large k. As a consequence,

$$\mathbb{E}|W_k| \geq \varepsilon^2\theta^k.$$

Thus, there exists $t \in [0, ks]$ such that

$$\mathbb{E}|X_{[t,t+1]}^{kL_0} \cap \mathcal{T}_{kL_0}| \geq \varepsilon^2\theta^k/ks.$$

Hence, by choosing k large, we have

$$\mathbb{E}|X_{t+1}^{kL_0} \cap \mathcal{T}_{kL_0}| \geq e^{-1}\varepsilon^2\theta^k/ks > 1,$$

which proves (7.9) by letting the new s_0 to be $t+1$ and the new L_0 to be kL_0.

Let (Y_t') be the contact process on $\mathcal{T}$ with the root initially infected and with infection rate λ' and recovery rate 1.

Since $\lambda' > \lambda_1$, the contact process (Y_t') survives forever with positive probability and in such event, there is a path of infection from the root to infinity with positive probability. Let $\mathcal{A}_{s_0,L_0}$ be the event that there is a path of infection of $(Y_t')_{t\leq s_0}$ from the root to a leaf of the L_0-generation that lies entirely in $\mathcal{T}_{\leq L_0}$. We have

$$\mathbb{P}_{\mathcal{T},Y'}\left(\bigcup_{s_0\in(0,\infty)} \mathcal{A}_{s_0,L_0} \neq \emptyset \quad \forall L_0\right) > \varepsilon$$

for some constant $\varepsilon > 0$. Thus, for all L_0,

$$\mathbb{P}_{\mathcal{T},Y'}\left(\bigcup_{s_0\in(0,\infty)} \mathcal{A}_{s_0,L_0} \neq \emptyset\right) > \varepsilon.$$

We fix L_0 to be a sufficiently large constant. Since $\mathcal{A}_{s_0,L_0}$ is an increasing event in s_0, there exists $s_0 \geq 1$ such that

$$\mathbb{P}_{\mathcal{T},Y'}\left(\mathcal{A}_{s_0,L_0} \neq \emptyset\right) > \varepsilon.$$

Note that $\mathcal{A}_{s_0,L_0} \subset \mathcal{A}_{s_0,L'_0}$ for any $L'_0 \leq L_0$ and hence

$$(7.13) \qquad \mathbb{P}_{\mathcal{T},Y'}\left(\mathcal{A}_{s_0,L'_0} \neq \emptyset\right) > \varepsilon, \quad \forall L'_0 \leq L_0.$$

Conditioned on the event $\mathcal{A}_{s_0,L_0}$, let W_{s_0,L_0} be the set of vertices in $\mathcal{T}_{\leq L_0}$ that are infected by time s_0 by an infection path that lies entirely in $\mathcal{T}_{\leq L_0}$ then W_{s_0,L_0} is a subtree of $\mathcal{T}$ that contains ρ and at least one vertex of depth L_0. Let $\partial W_{s_0,L_0}$ be the set of vertices of W_{s_0,L_0} that either belong to $\mathcal{T}_{L_0}$ or has at least one child not in W_{s_0,L_0}.

Let C be such that

$$(7.14) \qquad C\varepsilon^2\varepsilon_1\frac{\lambda}{1+\lambda} = 10,$$

where we recall that $\varepsilon_1 = 1 - \frac{\lambda'}{\lambda}$. The reason for this choice of C will be clear in (7.16).

Let $\mathcal{B}_{L_0}$ be the event that there exists a subtree W of $\mathcal{T}_{\leq L_0}$ containing the root, intersecting $\mathcal{T}_{L_0}$ and satisfying $|\partial W| \leq C$. Note that this event $\mathcal{B}_{L_0}$ depends only on the randomness of the tree, but not on the contact process. We shall show later that for any fixed constant C,

$$(7.15) \qquad \lim_{L_0\to\infty}\mathbb{P}\left(\mathcal{B}_{L_0}\right) = 0.$$

Assuming (7.15), we let L_0 be sufficiently large so that $\mathbb{P}(\mathcal{B}_{L_0}) \leq \varepsilon/2$.

We now consider the event $\mathcal{A}_{s_0,L_0} \setminus \mathcal{B}_{L_0}$ (which happens with probability at least $\varepsilon/2$). The process (Y_t) can be obtained from (Y'_t) by censoring each recovery clock with probability $1 - \varepsilon_1$; in other words, with probability $1 - \varepsilon_1$, a recovery clock of (Y'_t) is kept and with probability ε_1, it is ignored. For each vertex $v \in \partial W_{s_0,L_0} \cap \mathcal{T}_{<L_0}$ with a child $u \notin W_{s_0,L_0}$, if $v \in Y'_{s_0}$ then with probability at least $(1 - e^{-s_0})\frac{\lambda}{1+\lambda}$, $u \in Y_{\leq 2s_0}$. If $v \notin Y'_{s_0}$, then there is at least one recovery clock at v in time $[0,s_0]$. By censoring these recovery clocks, with probability at least $\varepsilon_1(1 - e^{-s_0})\frac{\lambda}{1+\lambda}$, we have $u \in Y_{\leq 2s_0}$. Consider the subtree $T_u \subset \mathcal{T}$ rooted at u. We observe that the trees T_u are disjoint and are independent of W_{s_0,L_0} and $(Y'_t)_{t\leq s_0}$. By (7.13), if u is infected at time t, with probability at least ε, there is a path of infection from u to a leaf of the L_0-generation of $\mathcal{T}$ by time $s_0 + t$. In particular, the expected number of infected descendants of u that belong to the L_0-generation of $\mathcal{T}$ by time $s_0 + t$ is at least ε.

Hence, by the choice of C in (7.14),

$$(7.16) \qquad \begin{aligned} &\mathbb{E}(|Y^{L_0}_{\leq 3s_0} \cap \mathcal{T}_{L_0}|) \\ &\geq \mathbb{P}\left(\mathcal{A}_{s_0,L_0} \setminus \mathcal{B}_{L_0}\right)\mathbb{E}\left(|\partial W_{s_0,L_0}|\,\Big|\,\mathcal{A}_{s_0,L_0} \setminus \mathcal{B}_{L_0}\right)\varepsilon_1(1 - e^{-s_0})\frac{\lambda}{1+\lambda}\varepsilon \\ &\geq \frac{\varepsilon}{2}C\varepsilon_1\frac{1}{2}\frac{\lambda}{1+\lambda}\varepsilon > 1, \end{aligned}$$

proving (7.11) by setting $3s_0$ to be the new s_0.

It remains to prove (7.15). Assume that such a W exists. Let P be a path of length L_0 in W connecting ρ and a vertex in the L_0-generation of $\mathcal{T}$. Let v be a vertex in P followed by a child $u \in P$ and consider the subtree $\mathcal{T}_P(v)$ consisting of

v and its descendants who are not descendants of u. If this subtree intersects $\mathcal{T}_{L_0}$ then it must contain a vertex in ∂W. Observe also that these subtrees $\mathcal{T}_P(v)$ are disjoint for all $v \in P$. Therefore, there are at most C vertices v in P whose $\mathcal{T}_P(v)$ intersects $\mathcal{T}_{L_0}$. Assume that the depth of these vertices is $p_1 < \cdots < p_k$ ($k \leq C$).

There are at most $C(L_0 + 1)^C$ ways to choose the sequence $p_1 < \cdots < p_k$. For each such choice, it suffices to show that the probability that there exists a W corresponding to this sequence is exponentially small in L_0, namely there exists a positive constant c such that

$$(7.17) \qquad \mathbb{P}\left(\mathcal{B}_{L_0,p_1,\ldots,p_k}\right) \leq e^{-cL_0},$$

where $\mathcal{B}_{L_0,p_1,\ldots,p_k}$ is the event that there exists a path $P = (\rho, v_1, \ldots, v_{L_0})$ of length L_0 in $\mathcal{T}_{\leq L_0}$ connecting the root and a vertex of depth L_0 such that for all vertices $v \in P$ except those at depths $p_1, \ldots, p_k$, $\mathcal{T}_P(v)$ does not intersect $\mathcal{T}_{L_0}$. Since

$$\mathcal{B}_{L_0} \subset \bigcup_{p_1,\ldots,p_k} \mathcal{B}_{L_0,p_1,\ldots,p_k},$$

by the union bound, we get (7.15).

It is left to prove (7.17). We say that a vertex v is *thin* if it has exactly one child with at least one descendant in the L_0-generation of the tree. Let $\varepsilon_2 > 0$ be the probability that the root of $\mathsf{GW}(\widetilde{\mu})$ has at least 2 children with infinitely many descendants then for a given vertex v, the probability that v is thin is at most $1 - \varepsilon_2$ (with or without conditioning on the event that v has a descendant in the L_0-generation). We have

$$(7.18) \qquad \mathbb{P}\left(\mathcal{B}_{L_0,p_1,\ldots,p_k}\right) \leq (1-\varepsilon_2)^{L_0-k} d^k \varepsilon_2^{-k} \leq (1-\varepsilon_2)^{L_0} \left((1-\varepsilon_2)^{-1} d\varepsilon_2^{-1}\right)^C,$$

which gives (7.17) for sufficiently large L_0.

To see (7.18), take for an example that the sequence $p_1, \ldots, p_k$ is $2, 3$. The probability that ρ is thin is at most $1 - \varepsilon_2$. Conditioned on this event, there is only one choice for v_1, which is the only child of ρ with a descendant in the L_0-generation. The probability that v_1 is thin is again at most $1 - \varepsilon_2$ and conditioned on this event, there is only one choice for v_2. Now, v_2 is not thin and there can be many choices for v_3 among the children of v_2. Likewise, there can be many choices for v_4 among the children of v_3. Given v_2, the expected number of choices for v_4 is the same as

$$(7.19) \qquad \mathbb{E}_{\mathcal{T}\sim\mathsf{GW}(\widetilde{\mu})}\left(|\mathcal{T}_2| \big| \mathcal{T}_{L_0-2} \neq \emptyset\right) \leq \frac{d^2}{\mathbb{P}_{\mathcal{T}\sim\mathsf{GW}(\widetilde{\mu})}\left(\mathcal{T}_{L_0-2} \neq \emptyset\right)} \leq d^2\varepsilon_2^{-1}.$$

Repeating this argument for the chance that $v_4, \ldots, v_{L_0-1}$ are thin, we get

$$\mathbb{P}\left(\mathcal{B}_{L_0,2,3}\right) \leq (1-\varepsilon_2)^{L_0-2} d^2 \varepsilon_2^{-1} \leq (1-\varepsilon_2)^{L_0-2} d^2 \varepsilon_2^{-2}.$$

The proof of (7.18) for a general sequence $p_1, \ldots, p_k$ follows by the same reasoning. We leave the details to the interested reader. $\qquad \square$

7.3. Good vertices.

We are now ready to define the notion of "good" vertices as mentioned earlier. Let s_0, L_0, θ be the constants in Lemma 7.5.

Let c_0 be a small constant such that for $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu}_{\eta_0,j_0})$ and for all $h \geq c_0^{-1}$,

$$(7.20) \qquad \mathbb{P}_{\mathrm{GW}}\left(|\mathcal{T}_h| \geq c_0 \widetilde{d}^h\right) \geq 2c_0.$$

The existence of c_0 follows from the fact that $\frac{|\mathcal{T}_h|}{d^h}$ converges to a nonzero random variable almost surely and that $\widetilde{\mu}_{\eta_0,j_0}$ converges to $\widetilde{\mu}$ and $\widetilde{d}$ converges to d.

For a set of κ "good" vertices, we shall show that there are a lot more vertices infected at a later time (Lemma 7.11). Generally, a good vertex is a vertex with an expanding neighborhood, namely a distance-h neighborhood contains at least $c_0 \tilde{d}^h$ vertices (similar to $\mathcal{T}$ as in (7.20)).

For the proof of Theorem 5, we only need to take $\kappa = \delta n$ while for Theorem 6, we shall need the range of κ to be in $[(\log n)^C, \delta n]$ where C is a large constant and δ is a small constant to be chosen. Thus, for the reader who is only interested in Theorem 5, it suffices to assume that $\kappa = \delta n$ for the rest of the proof. Here, we write for any $\kappa \in [(\log n)^C, \delta n]$.

Let

$$(7.21) \qquad l_1(\kappa) = 2 \log_{\tilde{d}} \log \frac{n}{\kappa}$$

be the depth (up to an additive constant) of the neighborhood for the definition of good vertices. Roughly speaking, the reason for this choice is to have

$$(7.22) \qquad \tilde{d}^{l_1(\kappa)} \gg \log \frac{n}{\kappa} \quad \text{and} \quad j_1^{Cl_1(\kappa)} \kappa \leq \eta_0 n,$$

where the second inequality ensures the hypothesis of Lemma 7.3 and the first inequality, whose purpose will be clear later in Lemma 7.8, is to facilitate the union bound over choices of κ vertices.

Definition 7.7 (*Good* vertices)**.** A vertex $v \in G'_n$ is said to be κ-*good* if

- $N(v,3)$ has $j_0(j_0 - 1)^3$ out-going half-edges,
- Each out-going half-edge of $N(v,3)$ expands well to depth $l_1(\kappa)$, namely, it is matched to a half-edge of a vertex u and the number of out-going half-edges of $N(u, l_1(\kappa)) \setminus N(v, 3 + l_1(\kappa))$ is at least $c_0 \tilde{d}^{l_1(\kappa)}$.

By (7.20), we have for a uniformly chosen vertex v,

$$(7.23) \qquad \mathbb{P}\left(v \text{ is good}\right) \geq (\mu_{\eta_0, j_0}(j_0))^{j_0^4} c_0^{j_0(j_0-1)^3} =: p_0,$$

where the term $(\mu_{\eta_0, j_0}(j_0))^{j_0^4}$ bounds the probability that the first condition in Definition 7.7 happens with j_0^4 bounding the total number of vertices in $N(v,3)$ and the term $c_0^{j_0(j_0-1)^3}$ bounds the probability of the second condition for which we use Lemma 7.3 and (7.20).

7.4. **A graph property.** Let

$$(7.24) \qquad l_2(\kappa) = Al_1(\kappa),$$

where A is defined by the equation

$$(\lambda \tilde{d}/2)\theta^{A/2} = 2 > 1.$$

Intuitively, if there are κ infected vertices that are κ-good, at distance $4 + l_1(\kappa)$ from these vertices, there are about $\kappa \tilde{d}^{l_1(\kappa)}$ vertices by the Definition 7.7. The number of these vertices that are infected at some point in the future is at least $\kappa \tilde{d}^{l_1(\kappa)}(\lambda/2)^{l_1(\kappa)}$. A decent portion of these vertices expands good trees of depth $l_2(\kappa)L_0$ by Lemma 7.5 and so the number of infected vertices at depth $4 + l_1(\kappa) + l_2(\kappa)L_0$ is about $\kappa \tilde{d}^{l_1(\kappa)}(\lambda/2)^{l_1(\kappa)}\theta^{l_2(\kappa)} \geq \kappa 2^{l_1(\kappa)} \gg \kappa$. The rest of this section is to make this heuristic rigorous.

Note that A is independent of C and δ. Let

$$(7.25) \qquad\qquad l(\kappa) = 4 + l_1(\kappa) + l_2(\kappa)L_0.$$

Following the above intuition, we consider a sequence of events on the random graphs G'_n.

- $\mathcal{B}$ is the event that for every $\kappa \in [(\log n)^C, \delta n]$, for every set D of κ vertices of G'_n that are κ-*good* and whose 3-neighborhoods $N(\cdot, 3)$ are all disjoint, it holds that
  (1) there are at least $c_0 \tilde{d}^{l_1(\kappa)} \kappa$ out-going half-edges from $\partial N(D, 4 + l_1(\kappa))$,
  (2) at   least
      $c_0 \varepsilon \tilde{d}^{l_1(\kappa)} \kappa / 4$ of these half-edges expands to an $(s_0, L_0, l_2(\kappa), \theta, \varepsilon)$-good tree of depth $l_2(\kappa)L_0$. Moreover, these good trees are disjoint.
- Let C_1 be a large constant that does not depend on δ and C (in other words, we shall choose $C \gg C_1$ and $\delta^{-1} \gg C_1$). Let $\mathcal{C}$ be the event that the following holds for every $\kappa \in [(\log n)^C, \delta n]$, every set D of at most κ vertices in G'_n, every $a \geq p_0^{-1} \kappa \log \frac{n}{\kappa} + (\log n)^{C_1}$ and $L_1 \leq C \log \log \frac{n}{\kappa}$. Let (X_t) be the contact process on G'_n with D initially infected and let s be any positive number. Let D_s be the set of vertices in $\partial N(D, L_1)$ that are infected at time X_s using only the infection and recovery events within $N(D, L_1)$ and let $\mathcal{F}_\kappa$ be the set of κ-good vertices in G'_n. Then

$$(7.26) \qquad \mathbb{P}_{\mathrm{CP}}\left(|D_s \cap \mathcal{F}_\kappa| \leq a p_0 / 2 \,\Big|\, |D_s| \geq a \right) \leq \exp\left(-3\kappa \log \frac{n}{\kappa} \right).$$

Under the randomness of the perfect matching of half-edges in G'_n, we show in Lemmas 7.8 and 7.10 that $\mathcal{B}$ and $\mathcal{C}$ happen whp, respectively.

Lemma 7.8. *We have*

$$(7.27) \qquad\qquad \mathbb{P}(G'_n \in \mathcal{B}) = 1 - o(1).$$

For the proof of this lemma, we shall use the cut-off line algorithm introduced in [10] to find a uniform perfect matching of the half-edges in G'_n.

Definition 7.9 (Cut-off line algorithm). A perfect matching of the half-edges of G'_n is obtained through the following algorithm.

- Each half-edge of a vertex v is assigned 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.
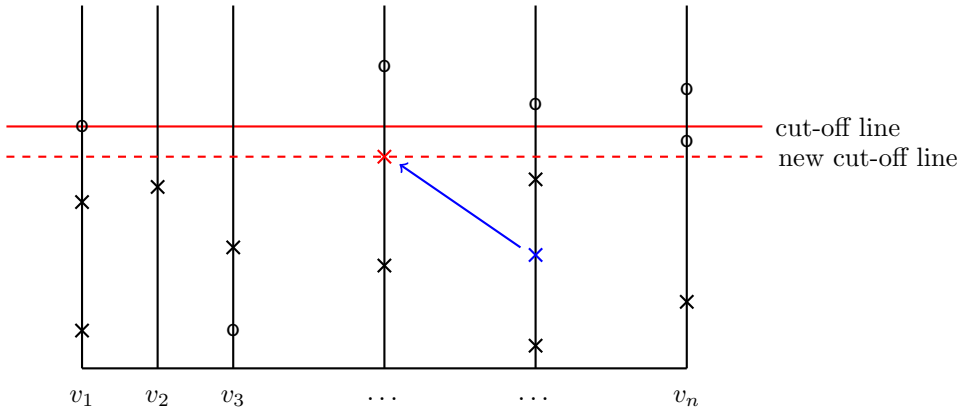
Figure 5 illustrates the algorithm.

FIGURE 5. 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).

Proof of Lemma 7.8. Let D be a set of κ vertices.

Perform the cut-off line algorithm to explore $N(D, 4 + l_1(\kappa))$: in other words, we run the cut-off line algorithm to match the half-edges of D and then $\partial N(D, 1), \ldots, \partial N(D, 3 + l_1(\kappa))$. Starting with the cut-off line at 1, let $1 - a$ be the position of the cut-off line after this exploration.

Let

$$\gamma = \frac{\kappa}{n} \leq \delta.$$

There are at most $2 j_1^{4+l_1(\kappa)} \kappa$ half-edges of G_n' that lie above this cut-off line while whp, there are more than n half-edges in G_n'. Hence, by Chernoff inequality,

$$\mathbb{P}(a \geq \gamma^{0.99}) \leq \mathbb{P}\left(\mathbf{Bin}(n, \gamma^{0.99}) \leq 2 j_1^{4+l_1(\kappa)} \kappa\right)$$
$$\leq \exp\left(-\gamma^{0.99} n/12\right) \leq \exp(-1.01\kappa \log \gamma).$$

For each $j \geq 2$, let n_j be the number of vertices in G_n' with j half-edges above the cutoff line. Under the event that $a \leq \gamma^{0.99}$, if $\sum_{j \geq i} n_j \geq 1.09\kappa/(i-1)$ for some i then there are $\geq 1.09\kappa/(i-1)$ vertices in G_n' with $\geq i$ half-edges above the line $1 - \gamma^{0.99}$. This happens with probability at most

$$\mathbb{P}(\mathbf{Bin}(n, j_1^i \gamma^{0.99i}) \geq 1.09\kappa/(i-1)),$$

which is at most

$$\exp\left(-0.99 \times 1.09 \frac{\kappa}{i-1} \log \frac{1.09 \gamma^{-0.99i+1}}{j_1^i(i-1)}\right)$$
$$\leq \exp\left(-0.99 \times 1.09 \frac{\kappa}{i-1} \log \frac{1.09 \gamma^{-0.99i+1}}{j_1^{j_1+1}}\right)$$

by the following Chernoff inequality for binomial distribution

$$(7.28) \qquad \mathbb{P}\left(\mathbf{Bin}(N, p) \geq xNp\right) \leq \exp\left(-xNp \log x + xNp\right), \quad \forall N, p > 0, x > 1.$$

Since $\gamma \leq \delta$ is sufficiently small compared to j_1, the above exponential is at most $\exp(-1.01\kappa \log \gamma)$ for $i \geq 2$.

Taking union bound over i and D and noting that $\binom{n}{\kappa} = o(\exp(1.01\kappa \log \gamma))$, we conclude that whp, for all D and all $2 \leq i \leq j_1$, we have $\sum_{j \geq i} n_j \leq 1.09\kappa/(i-1)$. Thus

$$(7.29) \qquad \sum_{i=2}^{j_1} in_i \leq 1.09\kappa \left(\frac{j_1}{j_1-1} + \sum_{i=2}^{j_1-1} \left(\frac{1}{i-1} - \frac{1}{i} \right) i \right) \leq 1.1\kappa(2 + \log j_1).$$

Now, assume that D consists of κ-good vertices with disjoint 3-neighborhoods. By definition, the number of out-going half-edges of $N(D,3)$ is $j_0(j_0-1)^3\kappa$. By (7.29), all but at most $1.1\kappa(2 + \log j_1)$ of these half-edges expand disjoint $l_1(\kappa)$-neighborhoods. By the definition of κ-good vertices, each of these neighborhoods has at least $c_0 \tilde{d}^{l_1(\kappa)}$ out-going half-edges. Therefore, the number of out-going half-edges from $\partial N(D, 4 + l_1)$ is at least

$$(7.30) \qquad \left(j_0(j_0-1)^3\kappa - 1.1\kappa(2 + \log j_1) \right) c_0 \tilde{d}^{l_1(\kappa)} \geq c_0 \tilde{d}^{l_1(\kappa)}\kappa.$$

This establishes the first part in the definition of $\mathcal{B}$. For the second part, we enumerate these half-edges by $e_1, e_2, \ldots, e_m$ and explore their $l_2(\kappa)L_0$ neighborhoods one by one. Note that since the maximal degree is j_1, at any point during the exploration, the number of explored vertices is at most $\kappa j_1^{5+l_1(\kappa)+l_2(\kappa)L_0} = \kappa \left(\log \frac{n}{\kappa} \right)^{A_1}$ where A_1 is a constant independent of δ and C. This number is at most $\eta_0 n$ for sufficiently small δ. Thus, by Lemma 7.3 and Lemma 7.5, for each $k \in [1, m]$, the probability that e_k expands to an $(s_0, L_0, l_2(\kappa), \theta, \varepsilon)$-good tree that does not intersect the previously explored neighborhoods (i.e., $N(D, 4 + l_1(\kappa)) \cup \bigcup_{i=1}^{k-1} N(e_i, l_2(\kappa)L_0)$) is at least $\varepsilon/2$, conditioned on these previously explored neighborhoods. Hence, the number of half-edges that expand to disjoint $(s_0, L_0, l_2(\kappa), \theta, \varepsilon)$-good trees is at least $\mathbf{Bin}(c_0 \tilde{d}^{l_1(\kappa)}\kappa, \varepsilon/2)$. By Chernoff inequality,
(7.31)
$$\mathbb{P}\left(\mathbf{Bin}(c_0 \tilde{d}^{l_1(\kappa)}\kappa, \varepsilon/2) \leq c_0 \tilde{d}^{l_1(\kappa)}\kappa\varepsilon/4 \right) \leq \exp\left(-\Omega\left(\tilde{d}^{l_1(\kappa)}\kappa \right) \right) \leq \exp\left(-1.1\kappa \log \frac{n}{\kappa} \right)$$

by (7.22). By taking the union bound over the choices of D, this completes the proof. $\qquad \square$

Lemma 7.10. *We have*

$$\mathbb{P}\left(G_n' \in \mathcal{C} \right) = 1 - o(1).$$

Proof of Lemma 7.10. The event $|D_s| \geq a$ depends only on the L_1-neighborhood of D. Thus, conditioning on this event, we expose the complement of this neighborhood to find the number of κ-good vertices in D_s. Since the probability for a vertex to be κ-good is at least p_0,

$$\mathbb{E}_{\mathrm{RG,CP}}\left(|D_s \cap \mathcal{F}_\kappa| \,\Big|\, |D_s| \geq a \right) \geq ap_0.$$

By Azuma's inequality and noting that $|D_s \cap \mathcal{F}_\kappa|$ depends only on $N(D_s, 4 + l_1(\kappa))$ whose size is at most $j_1^{5+l_1(\kappa)}$,

(7.32)
$$\mathbb{P}_{\text{RG,CP}}\left(|D_s \cap \mathcal{F}_\kappa| \leq a p_0/2 \Big| |D_s| \geq a\right) \leq \exp\left(-\frac{(a p_0/2)^2}{2a j_1^{10+2l_1(\kappa)}}\right) \leq \exp\left(-5\kappa \log \frac{n}{\kappa}\right),$$

where we chose C_1 much larger than A in (7.24).

Writing $\mathbb{P}_{\text{RG,CP}}(\cdot)$ as $\mathbb{E}_{\text{RG}}\left(\mathbb{P}_{\text{CP}}(\cdot)\Big| G_n'\right)$ and applying Markov's inequality, we yield from the above inequality that

$$\mathbb{P}_{\text{RG}}\left(\mathbb{P}_{(X_t)}\left(|D_s \cap \mathcal{F}_\kappa| \leq a p_0/2\right) \geq \exp\left(-3\kappa \log \frac{n}{\kappa}\right) \Big| |D_s| \geq a\right)$$

$$\leq \exp\left(-2\kappa \log \frac{n}{\kappa}\right) = o\left(\binom{n}{\kappa}^{-1}\right).$$

Taking union bound over at most $\kappa\binom{n}{\kappa}$ choices of D, at most n choices of a and over κ and L_1, we obtain

$$\mathbb{P}_{\text{RG}}\left(\mathcal{C}^c\right) \leq \sum_{\kappa=(\log n)^C}^{\delta n} \exp\left(-\kappa \log \frac{n}{\kappa}\right) = o(1).$$

$\square$

Let $\mathcal{E}$ be the event that the following holds for all $\kappa \in [\log^C n, \delta n]$. Let $\mathcal{F}_\kappa$ be the set of κ-good vertices in G_n'. For any set D of κ vertices in G_n' that are κ-good with disjoint 3-neighborhoods, there exists an $s \geq \varepsilon$ such that

(7.33)
$$\mathbb{P}_{\text{CP}}\left(|D_s \cap \mathcal{F}_{2\kappa}| \leq \kappa \log \frac{n}{\kappa}\right) \leq e^{-\kappa},$$

where (X_t) is the contact process on G_n' with D initially infected and D_s is the set of vertices in $\partial N(D, l(\kappa))$ that are infected at time X_s using only the infection and recovery events within $N(D, l(\kappa))$.

Lemma 7.11. *We have for a sufficiently large constant C,*
$$\mathbb{P}\left(G_n' \in \mathcal{E}\right) = 1 - o(1).$$

Proof. Under the event $\mathcal{B}$ in Lemma 7.8, there exists a constant $t > 0$ such that

(7.34)
$$\mathbb{E}_{\text{CP}}\left(|X_{\leq tl_2(\kappa)} \cap \partial N(D, l(\kappa))|\right) \geq \frac{\varepsilon}{4}\kappa c_0 \tilde{d}^{l_1(\kappa)}\left(\frac{(1 - e^{-1})\lambda}{1 + \lambda}\right)^{4+l_1(\kappa)} \theta^{l_2(\kappa)} \geq \theta^{l_2(\kappa)/2}\kappa,$$

where the last inequality follows from (7.25). Dividing the time interval $[0, tl_2(\kappa)]$ into intervals of length 1, there exists an interval $[r, r+1]$ such that

$$\mathbb{E}_{\text{CP}}\left(|X_{[r,r+1]} \cap \partial N(D, l(\kappa))|\right) \geq \theta^{l_2(\kappa)/2}\kappa/(tl_2(\kappa)),$$

where $X_{[r,r+1]} := \bigcup_{x \in [r,r+1]} X_x$. Since

$$\mathbb{E}\mathbf{1}_{v \in X_{r,r+1}} \leq e\mathbb{E}\mathbf{1}_{v \in X_{r+1}},$$

we obtain

$$\mathbb{E}_{\text{CP}}\left(|X_{r+1} \cap \partial N(D, l(\kappa))|\right) \geq e\theta^{l_2(\kappa)/2}\kappa/(tl_2(\kappa)) \geq 4p_0^{-1}\kappa \log \frac{n}{\kappa}.$$

By Azuma's inequality,

$$\text{(7.35)} \quad \mathbb{P}_{\text{CP}}\left(|X_{r+1} \cap \partial N(D, l(\kappa))| \leq 2p_0^{-1}\kappa \log \frac{n}{\kappa}\right) \leq \exp\left(-\frac{\kappa^2 \log^2 \frac{n}{\kappa}}{\kappa j_1^{3l(\kappa)}}\right)$$

$$\leq \exp\left(-3\kappa \log \frac{n}{\kappa}\right).$$

In the notations of Lemma 7.10, $X_{r+1} \cap \partial N(D, l(\kappa))$ is D_r. Thus, (7.35) together with (7.26) give

$$\mathbb{P}_{\text{CP}}\left(|D_s \cap \mathcal{F}_\kappa| \leq \kappa \log \frac{n}{\kappa}\right) \leq 2\exp\left(-3\kappa \log \frac{n}{\kappa}\right) \leq e^{-\kappa},$$

completing the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7.5. **Finishing the proof of Theorems 5 and 6.** We are now ready to finish the proof of the first inequality in (1.5) and Theorem 6. Let G'_n be a graph belonging to both events $\mathcal{B}$ and $\mathcal{C}$ in Lemmas 7.8 and 7.10.

Let v be a uniformly chosen random vertex in G'_n. By Lemma 7.3, whp over G'_n, for any constant C_2, $N(v, C_2 \log \log n) \geq_{st} \mathsf{GW}(\mu_{\eta_0, j_0}, \widetilde{\mu}_{\eta_0, j_0})$. And so, by Lemma 7.5, there exist constants $s_0, L_0, \varepsilon > 0$ and $\theta > 1$ such that

$$\mathbb{P}(N(v, C_2 \log \log n) \text{ contains an } (s_0, L_0, C_2 \log \log n/L_0, \theta, \varepsilon)\text{-good tree}) > \varepsilon,$$

where the probability is over the randomness of v and the randomness of the perfect matching of half-edges in G'_n. This implies that the set, denoted by $\mathcal{E}_1$, of vertices v such that $N(v, C_2 \log \log n)$ contains an $(s_0, L_0, C_2 \log \log n/L_0, \theta, \varepsilon)$-good tree has expectation at least $\varepsilon n'$ where $n' = |G'_n| \geq n/2$. By Azuma's inequality,

$$\mathbb{P}\left(|\mathcal{E}_1| \leq \varepsilon n/4\right) \leq \exp\left(-\frac{\varepsilon^2 n}{32 j_1^{2C_2 \log \log n}}\right) = o(1).$$

Thus, with high probability over G'_n, $|\mathcal{E}_1| \geq \varepsilon n/4$ which means that with positive probability over the choice of v, $N(v, C_2 \log \log n)$ contains an $(s_0, L_0, C_2 \log \log n/L_0, \theta, \varepsilon)$-good tree.

Let (X_t) be the contact process on G'_n starting with v infected. By Definition 7.4 of good trees, with probability at least ε over the contact process, there exists a time at which the number of infected vertices is at least

$$\theta^{C_2 \log \log n/L_0} = (\log n)^{C_2 \log \theta/L_0}.$$

Let

$$\kappa := (\log n)^{C_2 \log \theta/(2L_0)} p_0/(2j_1^4).$$

As we are under the event $\mathcal{C}$ in Lemma 7.10, among these infected vertices, there are at least $(\log n)^{C_2 \log \theta/L_0} p_0/2 \geq \kappa j_1^4$ vertices that are κ-good.

Once there are at least κj_1^4 infected vertices that are κ-good, we observe that a 3-neighborhood contains at most j_1^4 vertices and so at least κ of these vertices have disjoint 3-neighborhoods. Applying Lemma 7.11, we find that with probability at least $e^{-\kappa}$, at a later time, there are at least $2\kappa j_1^4$ infected vertices that are 2κ-good. Repeating this argument for $2\kappa, 4\kappa, \ldots$, we see that with probability at least $\varepsilon/2$, there exists a time at which there are $\geq \delta n$ infected vertices that are δn-good. Now, applying Lemma 7.11 for $e^{\delta n/2}$ times with $\kappa = \delta n$ gives that with probability at least $\varepsilon/4$, the contact process survives up to time $\varepsilon e^{\delta n/2}$, concluding the proof of Theorem 6. If all vertices are initially infected, then the first inequality in (1.5) of

Theorem 5 follows directed by applying Lemma 7.11 for $e^{\delta n/2}$ times with $\kappa = \delta n$ as above.

7.6. Finishing the proof of Theorems 1 and 2.

Finally, we combine the results obtained in this article and establish Theorems 1 and 2. Let us begin with Theorem 1. Its lower bound

$$\lim_{k \to \infty} \lambda_1^{\mathrm{GW}}(\xi_k) d_k \geq 1$$

was proven at the end of Section 4.2. Also, the other inequality was obtained by Theorem 5. For Theorem 2, Theorem 6.1 implies the lower bound

$$\lim_{k \to \infty} \lambda_c^-(\mu_k) d_k \geq 1.$$

Then the upper bound is again given by Theorem 5. Thus, we conclude the proofs of Theorems 1 and 2. □

APPENDIX A. RECURSIVE ANALYSIS FOR UNICYCLIC GRAPHS

Appendix A is devoted to the proof of Proposition 6.4 and Corollary 6.8. The main technical work needed here is to carry out the tree recursions as in Sections 4 and 5 despite the presence of a cycle. As before, our approach is based on an induction argument, which differs depending on whether t in (6.2) is small or large.

In Section A.1, we introduce an appropriate way of covering a unicyclic graph by trees. Based on this notion, in Section A.2 we can establish (6.2) for *small* values of t. Then, we conclude the proof of Proposition 6.4 and Corollary 6.8 in tandem inductively in Section A.3.

A.1. Reduction of unicyclic graphs to trees.

Let H be a graph that consists of a length-m cycle $C = v_1 v_2 \ldots v_m v_1$ and depth $\leq l$ trees $T_1, \ldots, T_m$ rooted at $v_1, \ldots, v_m$, respectively. The goal of this subsection is to introduce an approach that interprets the contact process on H by decomposing it into processes on trees. To this end, we begin with defining the *cover* of H.

Definition A.1 (Cover of H). Let $m' = \lceil \frac{m+1}{2} \rceil$ and H be a unicyclic graph defined as above. The *cover of H* is a pair of graphs A_1 and A_2 defined as follows (an illustration of the graphs can be found in Figure 6):

- A_1 consists of a length-m line $v_{m'} \ldots v_m v_1 v_2 \ldots v_{m'-1} \tilde{v}_{m'}$ and the trees T_j rooted at v_j except at the two endpoints $v_{m'}, \tilde{v}_{m'}$. The root of A_1 is v_1.
- A_2 consists of a length-m line $v_1 \ldots v_m \tilde{v}_1$ and the trees T_j rooted at v_j except at the two endpoints $v_1, \tilde{v}_1$. The root of A_2 is $v_{m'}$.

Using the cover A_1 and A_2 of H, we define the *decomposition* of $\mathbf{CP}^\lambda(H; \mathbf{1}_{v_1})$, similarly as Definition 6.19. Recall the graphical representation of the contact process discussed in Section 2.1, and let $\{N_v(t)\}_{v \in V(H)}$ and $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \vec{E}(H)}$ be the Poisson processes that define the recoveries and infections of $\mathbf{CP}^\lambda(H; \mathbf{1}_{v_1})$.

1. Initially, run $\mathbf{CP}^\lambda(A_1; \mathbf{1}_{v_1})$, whose recoveries and infections are given by $\{(N_v(s))_{s \geq 0}\}_{v \in V(A_1)}$ and $\{(N_{\vec{uv}}(s))_{s \geq 0}\}_{\vec{uv} \in \vec{E}(A_1)}$, respectively.
2. In A_1, if either $v_{m'}$ or $\tilde{v}_{m'}$ becomes infected at time t (and has been healthy until time $t-$), we start running a copy of $\mathbf{CP}^\lambda(A_2; \mathbf{1}_{v_{m'}})$ whose recoveries and infections are defined by $\{(N_v(s))_{s \geq t}\}_{v \in V(A_2)}$ and $\{(N_{\vec{uv}}(s))_{s \geq t}\}_{\vec{uv} \in \vec{E}(A_2)}$, respectively.
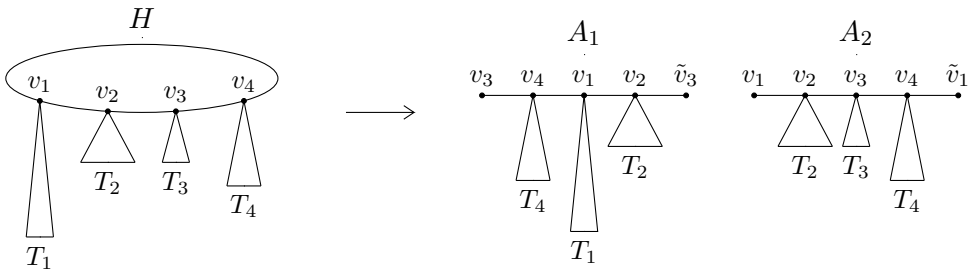
FIGURE 6. An example of the cover (A_1, A_2) of the unicyclic graph H with $m = 4$

3. When either v_1 or $\tilde{v}_1$ becomes infected at time t' (while it has been healthy until $t'-$) in a copy of $\mathbf{CP}^\lambda(A_2; \mathbf{1}_{v_{m'}})$, initiate a copy of $\mathbf{CP}^\lambda(A_1; \mathbf{1}_{v_1})$ that has $\{(N_v(s))_{s \geq t'}\}_{v \in V(A_1)}$ and $\{(N_{\vec{uv}}(s))_{s \geq t'}\}_{\vec{uv} \in \vec{E}(A_1)}$ as its event times, and return to Step 2.
4. The process terminates when all vertices in all copies of two processes are healthy.

Recalling the explanation below Definition 6.19, we need to control the number of infections at the endpoints $v_{m'}, \tilde{v}_{m'}$ and $v_1, \tilde{v}_1$ of A_1 and A_2, respectively. Therefore, we first consider the following graphs $F^{m,l}$ and $\mathcal{F}^{m,l}$ which are basically a one-sided version of A_1 and A_2.

- $F^{m,l}$ consists of a length-$(m-1)$ line $v_1 v_2 \dots v_m$ and depth $\leq l$ trees T_j rooted at v_j except at v_m. $\rho = v_1$ is designated as its root.
- $\mathcal{F}^{m,l}$ denotes the above graph when T_j, $j \in [m-1]$ are i.i.d. $\mathbf{GW}(\xi)^l$. In this case, we write $\mathcal{T}_j$ instead of T_j for each tree rooted at v_i.
- If $m = 1$, we set $\mathcal{F}^{1,l} \sim \mathbf{GW}(\xi)^l$ to be a single vertex v_1.

One can also define the root-added contact process $(X_t) \sim \mathbf{CP}^\lambda_{v_1^+}(F^{m,l}; \mathbf{1}_{v_1})$ by adding a permanently infected parent v_1^+ to v_1, and set the quantities $S(F^{m,l})$ and $M^l(F^{m,l})$ as Definition 6.3. In addition, we consider $B(F^{m,l})$, the *total infections at the end* and its expectation as follows.

$$(A.1) \qquad \mathbf{B}(F^{m,l}) := \left| \left\{ s \in [0, \mathbf{S}(F^{m,l})] : X_s(v_m) = 1 \text{ and } X_{s-}(v_m) = 0 \right\} \right|,$$

and $B(F^{m,l}) := \mathbb{E}_{\mathrm{CP}} \mathbf{B}(F^{m,l})$. Then, we have the following lemma.

Lemma A.2. *Let $m, l \geq 1$ be integers, $\varepsilon \in (0, 1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any ξ that satisfies $d := \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, we have for $\lambda = (1 - \varepsilon)d^{-1}$ and $\mathcal{F}^{m,l}$ as above that*

$$(A.2) \qquad\qquad \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{F}^{m,l}) \geq t\right) \quad \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq \tfrac{2}{\varepsilon};$$

$$(A.3) \quad \mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{F}^{m,l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \quad \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2;$$

$$(A.4) \qquad \mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m,l}) \geq d^{-\frac{3}{4}(m-1)}t\right) \quad \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2,$$

where $S(\mathcal{F}^{m,l})$, $M^l(\mathcal{F}^{m,l})$ and $B(\mathcal{F}^{m,l})$ are given as above.

Proof. We prove this lemma by an induction on m, for each fixed l. Establishing (A.2) and (A.3) is along the same lines as (4.4) and (5.1), so we postpone those parts of the proof to Section B.1 in the Appendix.

In the remaining proof, we show (A.4). When $m = 1$, $\mathcal{F}^{1,l}$ is a single vertex v_1 and we have the conclusion. Suppose that (A.4) holds for m and we want to prove the same thing for $\mathcal{F}^{m+1,l}$. Let $\{\mathcal{T}_j\}_{j=1}^{m-1}$ be i.i.d. $\mathsf{GW}(\xi)^l$ trees in $\mathcal{F}^{m,l}$, each rooted at v_i. Dividing $\mathcal{F}^{m+1,l}$ into subgraphs rooted at each child of v_1, we have the following description:

- Let $u_1, \ldots, u_D$ be the children of v_1 inside $\mathcal{T}_1$ and $T_1, \ldots, T_D$ be the subtrees in $\mathcal{T}_1$ rooted at each of them, respectively. Hence, $T_1, \ldots, T_D \sim$ i.i.d. $\mathsf{GW}(\xi)^{l-1}$ given D.
- Let F be the subtree of $\mathcal{F}^{m+1,l}$ rooted at v_2. That is, F is the graph consists with the length-$(m-1)$ line $v_2 \ldots v_{m+1}$ with each v_i, $2 \leq i \leq m$ being a root of the tree $\mathcal{T}_i$. Hence, $F \sim \mathcal{F}^{m,l}$.

Based on this decomposition of $\mathcal{F} = \mathcal{F}^{m+1,l}$, we can establish the two recursive inequalities as before. To derive the first one, let $(\bar{X}_t) \sim \mathsf{CP}^\lambda(\mathcal{F}, \mathbf{1}_{v_1})$ and

$$\bar{\mathbf{B}}(\mathcal{F}) := \left| \left\{ s \in [0, \mathbf{R}(\mathcal{F})] : \bar{X}_s(v_{m+1}) = 1 \text{ and } \bar{X}_{s-}(v_{m+1}) = 0 \right\} \right|,$$

where $\mathbf{R}(\mathcal{F})$ is the survival time of $(\bar{X}_t)$. We also set $\bar{B}(\mathcal{F}) := \mathbb{E}_{\mathrm{CP}} \bar{\mathbf{B}}(\mathcal{F})$, and let $\bar{\mathbf{B}}(F)$ and $\bar{B}(F)$ be the analogue for F. Then, based on the same idea as Proposition 4.2, we have

$$\bar{B}(\mathcal{F}^{m+1,l}) \leq \frac{\lambda \bar{B}(F)}{1 - \lambda^2 \left(R(F) + \sum_{i=1}^{D} R(T_i) \right)}, \quad \text{if } \lambda^2 \left(R(F) + \sum_{i=1}^{D} R(T_i) \right) < 1.$$

Thus, Proposition 4.3 and (4.11) tell us that

$$(A.5) \qquad B(\mathcal{F}^{m+1,l}) \leq \frac{\lambda B(F)}{1 - \lambda - 2\lambda^2 \left(S(F) + \sum_{i=1}^{D} S(T_i) \right)},$$

$$\text{if } \lambda + 2\lambda^2 \left(S(F) + \sum_{i=1}^{D} S(T_i) \right) < 1,$$

and hence we obtain the first recursive inequality.

The second inequality comes from the same idea as in Proposition 5.2:

$$(A.6) \qquad B(\mathcal{F}^{m+1,l}) \leq \lambda B(F) \prod_{i=1}^{D} (1 + \lambda S(T_i)).$$

As before, we show (A.4) separately on different regimes of t. Details in the proof are also similar to the proof of Theorem 4.4.

Case 1. For $t \leq d^{\frac{1}{10}}$.

Suppose that

$$B(F) \leq d^{-\frac{3}{4}(m-1)+\frac{1}{2}}, \quad \text{and} \quad S(F) + \sum_{i=1}^{D} S(T_i) \leq \frac{2d}{\varepsilon} \left(1 + \frac{\varepsilon}{3} \right).$$

Then, by (A.5), $B(\mathcal{F}^{m+1,l}) \le d^{-\frac{3m}{4}}$. Therefore, for any $t \ge 1$,

(A.7)
$$\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m+1,l}) \ge d^{-\frac{3m}{4}}t\right) \le \mathbb{P}\left(D \ge \left(1 + \frac{\varepsilon}{6}\right)d\right) + \mathbb{P}\left(B(F) \ge d^{-\frac{3}{4}(m-1)+\frac{1}{2}}\right)$$
$$+ \mathbb{P}_{\mathrm{GW}}\left(S(F) + \sum_{i=1}^{D} S(T_i) \ge \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right).$$

By following the logics from Section 4.2.1 and relying on the induction hypothesis that tells us the tail probability of $B(F)$, we see that the r.h.s. of (A.7) is at most $t^{-\sqrt{d}}(\log t)^{-2}$ for all $2 \le t \le d^{\frac{1}{10}}$.

Case 2. For $t \ge d^{\frac{1}{10}}$.

For convenience, set
$$\widehat{S} = \prod_{i=1}^{D}(1 + \lambda S(T_i)), \quad \text{and} \quad \widetilde{B} = d^{\frac{3}{4}(m-1)}B(F).$$

We use (A.6) and hence prove the following.

Claim A.3. Under the above setting, we have

(A.8) $$\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m+1,l}) \ge d^{-\frac{3m}{4}}t\right) \le \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \cdot \widehat{S} \ge \frac{1}{2}d^{\frac{1}{4}}t\right) \le t^{-\sqrt{d}}(\log t)^{-2},$$
$$\text{for all } t \ge d^{\frac{1}{10}}.$$

Note that the first inequality is just a rewriting of (A.6).

The idea is to split the probability as follows: let $\alpha = d^{\frac{1}{10}}$, and we see that
$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \cdot \widehat{S} \ge t\right) \le \mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \ge t\alpha\right) + \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \ge \frac{t}{2\alpha^2}d^{\frac{1}{4}}\right)$$
(A.9)
$$+ \sum_{k=0}^{k_0} \mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \ge t\alpha^{-k}\right) \cdot \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \ge \frac{\alpha^{k-1}}{2}d^{\frac{1}{4}}\right),$$

where we set $k_0 = \lfloor \frac{\log t}{\log \alpha}\rfloor$-1, the point where $\alpha^{k+1} \le t$ is the closest to t. Similar ideas as Lemma 4.8 yields the conclusion (A.8), but the computation is much simpler. We postpone the remaining details of Claim A.3 to Section B.2 in the Appendix. $\qquad\square$

A.2. **Back to unicyclic graphs and the proof of Proposition 6.4 for small t.** In the previous subsection, the graph $\mathcal{F}^{m,l}$ is introduced to describe a way of decomposing GWC^1-processes. Now we combine the pieces together and derive a recursive tail probability estimates for GWC^1-processes.

For integers $m_1, m_2 \ge 1$ and $l \ge 0$, let $\underline{\mathrm{m}} = (m_1, m_2)$ and define the random graph $\mathcal{A}^{\underline{\mathrm{m}},l}$ as

- $\mathcal{A}^{\underline{\mathrm{m}},l}$ consists of a length-$(m_1 + m_2)$ line $v_{-m_1} \ldots v_{-1}v_0v_1 \ldots v_{m_2}$ and i.i.d. trees $\mathcal{T}_j \sim \mathsf{GW}(\xi)^l$ rooted at v_j for $j \in \{-m_1 + 1, \ldots, m_2 - 1\}$. We designate $\rho = v_0$ as the root of $\mathcal{A}^{\underline{\mathrm{m}},l}$.

We define the *root-added contact process* on $\mathcal{A}^{\underline{\mathrm{m}},l}$ by adding a permanently infected parent ρ^+ whose only connection is to $\rho = v_0$. Further, we define $S(\mathcal{A}^{\underline{\mathrm{m}},l})$ and $M^l(\mathcal{A}^{\underline{\mathrm{m}},l})$ as Definition 6.3, and $B(\mathcal{A}^{\underline{\mathrm{m}},l})$ as (A.1) but considering the infections

at both endpoints v_{-m_1}, v_{m_2} (not just a single endpoint as in (A.1)). Under this setting, we first establish the tail probability estimates for $\mathcal{A}^{\underline{m},l}$.

Lemma A.4. *Let $m_1, m_2, l \geq 1$ be integers, $\varepsilon \in (0,1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be a collection of positive constants. Then there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that the following holds true. For any ξ that satisfies $d := \mathbb{E}\xi \geq d_0$ and (1.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{A}^{\underline{m},l}$ as above that*

$$(A.10) \qquad\qquad \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{A}^{\underline{m},l}) \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq \tfrac{2}{\varepsilon};$$

$$(A.11) \quad \mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{A}^{\underline{m},l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2;$$

$$(A.12) \quad \mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{A}^{\underline{m},l}) \geq d^{-\frac{3}{4}(m_1 \wedge m_2)}t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } t \geq 2,$$

where $S(\mathcal{A}^{\underline{m},l})$, $M^l(\mathcal{A}^{\underline{m},l})$ and $B(\mathcal{A}^{\underline{m},l})$ are given as above.

We can prove this lemma based on Lemma A.2, by a straight-forward application of the recursive argument from Theorems 4.4 and 5.1. To prevent repeating the same argument many times, we postpone the proof until Section B.3 in the Appendix.

We conclude this subsection by settling the following lemma, which establishes Proposition 6.4 for $t \leq d^{\frac{1}{10}}$.

Lemma A.5. *Let $m \geq 2$ and $l \geq 1$ be integers, $\varepsilon \in (0,1)$ and $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$ be positive constants. Then, under the setting of Proposition 6.4, we have*

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}^{m,l}) \geq t\right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } \frac{2}{\varepsilon} \leq t \leq d^{\frac{1}{10}};$$

$$(A.13)$$

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{H}^{m,l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq 3t^{-\sqrt{d}}(\log t)^{-2} \text{ for all } 2 \leq t \leq d^{\frac{1}{10}}.$$

Proof. Let $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi, m)^l$, set $m' = \lceil \frac{m+1}{2} \rceil$ and consider the cover A_1 and A_2 of $\mathcal{H}^{m,l}$ defined in Definition A.1. Set $m_1 = m + 1 - m'$ and $m_2 = m' - 1$. Then, we can see that $m_1, m_2 \geq 1$ for $m \geq 2$ and $A_1, A_2 \sim \mathcal{A}^{\underline{m},l}$ for $\underline{m} = (m_1, m_2)$, with $\mathcal{A}^{\underline{m},l}$ defined as in the beginning of Section A.2. Thus, we have (A.10), (A.11) and (A.12), where the last one in particular implies that

$$\mathbb{P}_{\mathrm{GW}}\left(B(A_1) \geq d^{-\frac{3}{4}}t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq 2,$$

and the same thing for A_2 as well.

Recall the decomposition of the contact process on $\mathcal{H}^{m,l}$ by its cover (Definition 6.19). According to its formulation, the quantities $R(\mathcal{H}^{m,l})$ and $\bar{M}^l(\mathcal{H}^{m,l})$ (see Definition 6.3 for their definitions) can be estimated as follows.

$$R(\mathcal{H}^{m,l}) \leq R(A_1) + B(A_1)R(A_2) + B(A_1)B(A_2)R(\mathcal{H}^{m,l}), \text{ and hence}$$
$$R(\mathcal{H}^{m,l}) \leq \frac{R(A_1) + B(A_1)R(A_2)}{1 - B(A_1)B(A_2)}.$$

We have the same thing for $\bar{M}^l(\mathcal{H}^{m,l})$, namely,

$$\bar{M}^l(\mathcal{H}^{m,l}) \leq \frac{\bar{M}^l(A_1) + B(A_1)\bar{M}^l(A_2)}{1 - B(A_1)B(A_2)}.$$
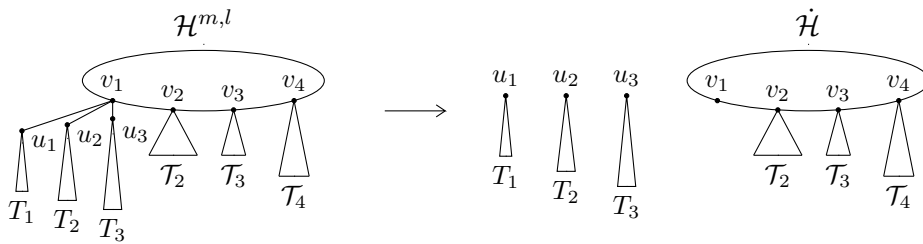
FIGURE 7. An illustration of dividing $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi,m)^l$ into $T_1, T_2, T_3 \sim \mathsf{GW}(\xi)^{l-1}$ and $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi,m)^l$, given that $m = 4$ and $D = 3$

Observe that Proposition 4.3 and Corollary 5.4 can be applied to $\mathcal{H}^{m,l}$ so that

$$\frac{S(\mathcal{H}^{m,l})}{1 + \lambda S(\mathcal{H}^{m,l})} \leq R(\mathcal{H}^{m,l}), \quad \text{and} \quad M^l(\mathcal{H}^{m,l}) \leq (1 + \lambda S(\mathcal{H}^{m,l}))\bar{M}^l(\mathcal{H}^{m,l}),$$

and hence we have

$$S(\mathcal{H}^{m,l}) \leq \frac{S(A_1) + B(A_1)S(A_2)}{1 - B(A_1)B(A_2) - \lambda S(A_1) - \lambda B(A_1)S(A_2)}$$

$$M^l(\mathcal{H}^{m,l}) \leq \frac{M^l(A_1) + B(A_1)M^l(A_2)}{1 - B(A_1)B(A_2) - \lambda S(A_1) - \lambda B(A_1)S(A_2)}.$$

To establish the first inequality of (A.13), we first observe that if

$$t \leq d^{\frac{1}{10}}, \quad S(A_1),\ S(A_2) \leq \left(1 - d^{-\frac{5}{9}}\right)t, \quad \text{and} \quad B(A_1),\ B(A_2) \leq d^{-\frac{5}{8}},$$

then $S(\mathcal{H}^{m,l}) \leq t$. Therefore, for $t \leq d^{\frac{1}{2}}$, we have

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}^{m,l}) \geq t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(S(A_1) \geq \left(1 - d^{-\frac{5}{9}}\right)t\right) + \mathbb{P}_{\mathrm{GW}}\left(S(A_2) \geq \left(1 - d^{-\frac{5}{9}}\right)t\right)$$

$$+ \mathbb{P}_{\mathrm{GW}}\left(B(A_1) \geq d^{-\frac{5}{8}}\right) + \mathbb{P}_{\mathrm{GW}}\left(B(A_2) \geq d^{-\frac{5}{8}}\right).$$

Since $(1 - d^{-\frac{5}{9}})^{-\sqrt{d}} \leq 1 + 2d^{-\frac{1}{18}}$, applying Lemma A.4 gives us that the r.h.s. of the above is bounded by $3t^{-\sqrt{d}}(\log t)^{-2}$ for $2/\varepsilon \leq t \leq d^{\frac{1}{10}}$. We can also do the same thing for $M^l(\mathcal{H}^{m,l})$ to deduce the conclusion. $\qquad\square$

A.3. **Proof of Proposition 6.4 for large t.** In this section, we conclude the proof of Proposition 6.4. Consider $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi,m)^l$.

For a GWC^1-process $\mathcal{H}^{m,l} \sim \mathsf{GWC}^1(\xi,m)^l$, we can derive a recursive inequality on $\mathcal{H}^{m,l}$ similar to (4.1) using GWC^2-process. As described in Figure 7 let $v_1 \ldots v_m v_1$ be its cycle part, and let $\mathcal{T}_j$, $j \in [m]$ be i.i.d. $\mathsf{GW}(\xi)^l$ rooted at each v_i. Let D be the degree of v_1 in $\mathcal{T}_1$, and let $T_1, \ldots, T_D$ be the subtrees rooted at $u_1, \ldots u_D$, the children of v_1 in $\mathcal{T}_1$. Further, let $\dot{\mathcal{H}}$ be the graph that consists of the cycle $v_1 \ldots v_m v_1$ and the trees $\mathcal{T}_j$ rooted at v_j for $j \in \{2, \ldots, m\}$, which is $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi,m)^l$. Also, note that $T_1, \ldots, T_D \sim$ i.i.d. $\mathsf{GW}(\xi)^{l-1}$. Then, Lemma A.6 shows the connection between GWC^1- and GWC^2-processes. Its proof is based on the same ideas as Proposition 4.1, and it will be presented in Section A.3.2.

Lemma A.6. *Under the above setting, we have*

$$(A.14) \qquad S(\mathcal{H}^{m,l}) \leq (1 + 2\lambda S(\dot{\mathcal{H}})) \prod_{i=1}^{D} (1 + \lambda S(T_i)).$$

This shows that understanding the tail probabilities on GWC^2-processes is helpful in estimating those on GWC^1-processes.

We prove Proposition 6.4 and Corollary 6.8 in tandem. This is done by an inductive argument which we detail below. Let $l \geq 1$ be a fixed integer. We establish both (6.2) and (6.3) for all m via the following steps.

  Step 1. Show that (6.2) (resp., (6.3)) are true for $m = 1$ (resp., $m = 2$).
  Step 2. For each integer $m \geq 2$, prove that if (6.3) is true for m, then we have (6.2) for m.
  Step 3. For each integer $m \geq 2$, prove that if (6.2) and (6.3) are true for m, then we have (6.3) for $m + 1$.
    3-1. Show that (6.3) holds when $t \leq d^{\frac{1}{10}}$ (for $m + 1$).
    3-2. Show that (6.3) holds when $t \geq d^{\frac{1}{10}}$ (for $m + 1$).

A.3.1. *Proof of Proposition* 6.4 *and Corollary* 6.8, *Step* 1. Proving (6.2) for $m = 1$ is straight-forward, since $\mathcal{H}^{1,l} \sim \mathsf{GW}(\xi)^l$ and the result follows from Theorems 4.4 and 5.1. To establish (6.3) for $m = 2$, we divide (6.3) into two cases, $t \leq d^{\frac{1}{10}}$ and $t \geq d^{\frac{1}{10}}$. Observe that $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, 2)^l$ is the same as $(\mathcal{T}_2, v_2) \sim \mathsf{GW}(\xi)^l$ with a parent $\rho = v_1$ connected to v_2 via a pair of edges.

Case 1. $t \leq d^{\frac{1}{10}}$.

Letting $S(\dot{\mathcal{H}}) = S_2(\dot{\mathcal{H}})$, $M^l(\dot{\mathcal{H}}) = M_2^l(\dot{\mathcal{H}})$ as Definition 6.7, an analog of Proposition 4.3 tells us that

$$(A.15) \qquad \frac{S(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \leq R(\mathcal{T}_2) \leq S(\mathcal{T}_2), \quad \text{and} \quad \frac{M^l(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \leq \bar{M}^l(\mathcal{T}_2) \leq M^l(\mathcal{T}_2).$$

The left inequality can be rewritten as

$$S(\dot{\mathcal{H}}) \leq \frac{S(\mathcal{T}_2)}{1 - 2\lambda S(\mathcal{T}_2)}.$$

Based on this information and on Theorems 4.4 and 5.1, we deduce that for $3/\varepsilon \leq t \leq d^{\frac{1}{10}}$,

$$(A.16) \qquad \mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}) \geq t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}_2) \geq \left(1 - d^{-\frac{3}{4}}\right) t\right) \leq 2t^{-\sqrt{d}}(\log t)^{-2},$$

since $(1 - d^{-\frac{3}{4}})^{-\sqrt{d}} < \frac{3}{2}$ and $|\log(1 - d^{-\frac{3}{4}})| \leq \frac{1}{100}\log t$ for large d. For the same reason, we obtain that

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{T}_2) \geq \left(1 - d^{-\frac{3}{4}}\right)\left(1 - \frac{\varepsilon}{10}\right)^l t\right)$$
$$+ \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}_2) \geq d^{\frac{1}{5}}\right),$$

and the r.h.s. is bounded by $2t^{-\sqrt{d}}(\log t)^{-2}$.

Case 2. $t \geq d^{\frac{1}{10}}$.

We can deduce the conclusion by exactly the same argument as Theorem 4.4 (Sections 4.2.2, 4.2.3) and Theorem 5.1 (Sections 5.3.2, 5.3.3). Indeed, the recursive

inequalities (4.1), (5.3) do not change for $\dot{\mathcal{H}}$ since they are independent of the number of connections between v_1 and v_2.

A.3.2. *Proof of Proposition* 6.4 *and Corollary* 6.8, *Step 2.* We have already established (6.2) for $t \leq d^{\frac{1}{10}}$ in Lemma A.5. Hence, we focus on settling the other part, $t \geq d^{\frac{1}{10}}$. We begin with presenting the proof of Lemma A.6, which shows the relation between GWC^1- and GWC^2-processes.

Proof of Lemma A.6. Let $\mathcal{H}^{m,l}$, $\dot{\mathcal{H}}$, D, $\{T_i\}_{i=1}^D$ and $\{u_i\}_{i=1}^D$ be as in the statement of the lemma. As we run $\mathsf{CP}_{v_1^+}^{\lambda}(\mathcal{H}^{m,l}; \mathbf{1}_{v_1})$, we adopt the analogous idea as Proposition 4.1 by ignoring the recoveries at v_1 except when all the other vertices are healthy. This translates the original process $\mathsf{CP}_{v_1^+}^{\lambda}(\mathcal{H}^{m,l}; \mathbf{1}_{v_1})$ to the product chain of $\mathsf{CP}_{v_1}^{\lambda}(T_i)$, $i \in [D]$ and $\mathsf{CP}_{v_2}^{\lambda}(\dot{\mathcal{H}})$ in the perspectives of Proposition 4.1. In $\mathsf{CP}_{v_2}^{\lambda}(\dot{\mathcal{H}}; \mathbf{0})$, the first infection from v_2 happens with rate 2λ which is passed to either v_3 or v_{m+1}, each with probability $\frac{1}{2}$. Therefore, the expected excursion time for this chain is $S(\dot{\mathcal{H}}) = \frac{1}{2}(S_3(\dot{\mathcal{H}}) + S_{m+1}(\dot{\mathcal{H}}))$. Based on these observations, we follow the logic (4.4), (4.5) and (4.6) from Proposition 4.1 to obtain (A.14). $\square$

Having Lemma A.6 and (6.3) for m, we can establish the first inequality of (6.2) by the same reasoning as Sections 4.2.2 and 4.2.3 (Lemma 4.8 and Corollary 4.9).

For the second line of (6.2), we have the following recursive inequality for $M^l(\mathcal{H}^{m,l})$ which can be derived analogously as Proposition 5.2.
(A.17)
$$M^l(\mathcal{H}^{m,l}) \leq 2\lambda M^l(\dot{\mathcal{H}}) \prod_{i=1}^D (1+\lambda S(T_i)) + \lambda(1+2\lambda S(\dot{\mathcal{H}})) \sum_{i=1}^D M^{l-1}(T_i) \prod_{\substack{j=1 \\ j \neq i}}^D (1+\lambda S(T_j)).$$

As done in Section 5.3.2, we set

$$T_0 := \dot{\mathcal{H}}, \quad \widetilde{M}_i := \left(1 - \frac{\varepsilon}{10}\right)^{-(l-1)} M^{l-1}(T_i), \quad \text{and} \quad W_i := \widetilde{M}_i \vee S(T_i).$$

Moreover, let $\widetilde{M} = (1 - \frac{\varepsilon}{10})^{-l} M^l(\mathcal{H}^{m,l})$. Then, we can derive the following from the above inequality.

$$\left(1 - \frac{\varepsilon}{10}\right) \widetilde{M} \leq \prod_{i=0}^D (1 + 4\lambda W_i).$$

Also, the assumption (6.3) and Theorem 5.1 tell us the tail bound on W_i, namely,

$$\mathbb{P}_{\mathrm{GW}}(W_i \geq t) \leq 4t^{-\sqrt{d}}(\log t)^{-2} + t^{-\sqrt{d}}(\log t)^{-2} \leq 5t^{-\sqrt{d}}(\log t)^{-2},$$

which falls into the assumption of Lemma 4.8 and Corollary 4.9. Therefore, we conclude the proof of (6.2) by obtaining that for all $t \geq d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(2 \prod_{i=0}^D (1 + 4\lambda W_i) \geq t\right) \leq 4t^{-\sqrt{d}}(\log t)^{-2}.$$
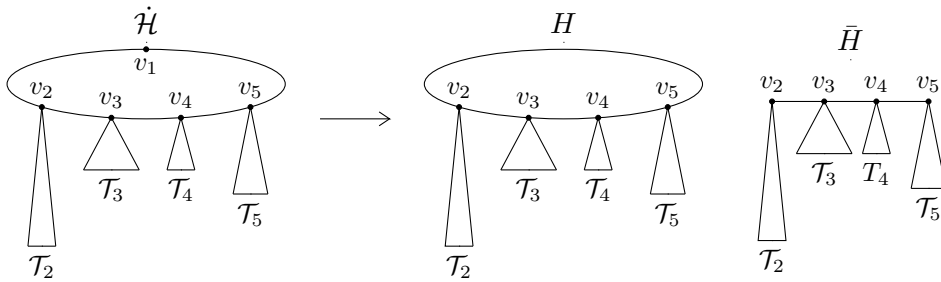
FIGURE 8. A description of $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, m+1)^l$, H and $\bar{H}$ in Section A.3.3 with $m = 4$

A.3.3. *Proof of Proposition* 6.4 *and Corollary* 6.8, *Step 3-1.* Let $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, m+1)^l$, $v_1 v_2 \ldots v_{m+1} v_1$ be its cycle part with root $\rho = v_1$ and $\mathcal{T}_2, \ldots, \mathcal{T}_{m+1}$ be the i.i.d. $\mathsf{GW}(\xi)^l$ trees rooted at $v_2, \ldots, v_{m+1}$, respectively. As described in Figure 8, let H and $\bar{H}$ be the graph defined as follows:

- H consists of a cycle $v_2 \ldots v_{m+1} v_2$ and the trees $\mathcal{T}_j$ rooted at v_j, $j \in \{2, \ldots, m+1\}$.
- $\bar{H}$ is obtained by removing the edge $(v_2 v_{m+1})$ from H. In other words, it consists of a length-$(m-1)$ path $v_2 \ldots v_{m+1}$ and the trees $\mathcal{T}_j$ rooted at v_j, $j \in \{2, \ldots, m+1\}$.

Also, let $R_j(\bar{H})$ (resp., $\bar{M}_j^l(\bar{H})$) denote the expected survival time (resp., expected total infections at leaves) of the (usual) contact process $\mathsf{CP}^\lambda(\bar{H}; \mathbf{1}_{v_j})$.

In the process $\mathsf{CP}_{v_1}^\lambda(\dot{\mathcal{H}}; \mathbf{0})$, the permanently infected root v_1 infects v_2 and v_{m+1} with rate λ each. Hence, the process can be interpreted similarly as in Definition 6.19:

- At rate 2λ, initiate a copy of either $\mathsf{CP}^\lambda(\bar{H}; \mathbf{1}_{v_2})$ or $\mathsf{CP}^\lambda(\bar{H}; \mathbf{1}_{v_{m+1}})$, each chosen with probability $\frac{1}{2}$ and whose event times are coupled with $\mathsf{CP}_{v_1}^\lambda(\dot{\mathcal{H}}, \mathbf{0})$.

Therefore, for $S(\dot{\mathcal{H}}), M^l(\dot{\mathcal{H}})$ defined as Definition 6.7, the argument from Proposition 4.3 and Corollary 5.4 implies
(A.18)
$$\frac{S(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \le \frac{R_2(\bar{H}) + R_{m+1}(\bar{H})}{2}, \quad \text{and} \quad \frac{M^l(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \le \frac{\bar{M}_2^l(\bar{H}) + \bar{M}_{m+1}^l(\bar{H})}{2}.$$

Note that $H \sim \mathsf{GWC}^1(\xi, m)^l$, and let $S_i(H)$ and $M_i^l(H)$ be defined as Definition 6.3, with respect to the root-added process $\mathsf{CP}_{v_i^+}^\lambda(H; \mathbf{1}_{v_i})$. Also, we set

$$S(H) = \frac{S_2(H) + S_{m+1}(H)}{2}, \quad \text{and} \quad M^l(H) = \frac{M_2^l(H) + M_{m+1}^l(H)}{2}.$$

Then, since $H' \subset H$, we use $R_i(H') \le S_i(H)$ and $M_i'(H') \le M_i(H)$ in (A.18) to see that

$$S(\dot{\mathcal{H}}) \le \frac{S(H)}{1 - 2\lambda S(H)}, \quad \text{and} \quad \frac{M^l(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \le M^l(H).$$

Now, we go through the same computations as (A.15) and (A.16). Namely, we deduce that for $3/\varepsilon \le t \le d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}) \ge t\right) \le \mathbb{P}_{\mathrm{GW}}\left(S(H) \ge \left(1 - d^{-\frac{3}{4}}\right)t\right) \le 4t^{-\sqrt{d}}(\log t)^{-2},$$

since $(1 - d^{-\frac{3}{4}})^{-\sqrt{d}} < \frac{5}{4}$ and $|\log(1 - d^{-\frac{3}{4}})| \le \frac{1}{100}\log t$. For the same reason, we obtain that

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \ge \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \le \mathbb{P}_{\mathrm{GW}}\left(M^l(H) \ge \left(1 - d^{-\frac{3}{4}}\right)\left(1 - \frac{\varepsilon}{10}\right)^l t\right)$$
$$+ \mathbb{P}_{\mathrm{GW}}\left(S(H) \ge d^{\frac{1}{5}}\right),$$

and the r.h.s. is bounded by $4t^{-\sqrt{d}}(\log t)^{-2}$. Hence, we finish the proof for Step 3-1.

A.3.4. *Proof of Proposition* 6.4 *and Corollary* 6.8, *Step 3-2.* Let $\dot{\mathcal{H}} \sim \mathsf{GWC}^2(\xi, m+1)^l$ and let $\mathcal{T}_2, \ldots, \mathcal{T}_{m+1}$ as above. Consider the graphs H_2 and H_{m+1} defined as follows.

- H_2 consists of a cycle $v_2 \ldots v_{m+1}$ and the trees $\mathcal{T}_3, \ldots, \mathcal{T}_{m+1}$ rooted at $v_3, \ldots, v_{m+1}$. The vertex $\rho = v_2$ is designated as the root.
- H_{m+1} consists of a cycle $v_2 \ldots v_{m+1}$ and the trees $\mathcal{T}_2, \ldots, \mathcal{T}_m$ rooted at $v_2, \ldots, v_m$. The vertex $\rho = v_{m+1}$ is designated as the root.

Note that $H_2 \sim \mathsf{GWC}^2(\xi, m)^l$, and we let $S_3(H_2), S_{m+1}(H_2)$ as Definition 6.7 with $S(H_2) := \frac{1}{2}(S_3(H_2) + S_{m+1}(H_2))$. Similarly, define $M^l(H_2)$, $S(H_{m+1})$ and $M^l(H_{m+1})$. Further, let D_2 (resp. D_{m+1}) be the degree of v_2 in the tree $\mathcal{T}_2$, and denote the subtrees at each child of v_2 in $\mathcal{T}_2$ by T_i^2, $i \in [D_2]$. Define D_{m+1} and T_i^{m+1}, $i \in [D_{m+1}]$ analogously with respect to $\mathcal{T}_{m+1}$.

Then, an analogous idea as Lemma A.6 shows that

$$S_2(\dot{\mathcal{H}}) \le (1 + 2\lambda S(H_2))\prod_{i=1}^{D_2}(1 + \lambda S(T_i^2));$$

(A.19)

$$S_{m+1}(\dot{\mathcal{H}}) \le (1 + 2\lambda S(H_{m+1}))\prod_{i=1}^{D_{m+1}}(1 + \lambda S(T_i^{m+1})).$$

(Indeed, this was also proven in [3, Lemma 4.11], and its verification is based on the same idea as Proposition 4.1.) Note that we are assuming the tail probabilities for $S(H_2)$ and $S(H_{m+1})$ satisfy (6.3), while those for $S(T_i)$ are given by Theorem 4.4. Based on this observation, we can follow the same analysis done in Lemma 4.8, Corollary 4.9 and Sections 4.2.2, 4.2.3 to see that for $t \ge d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}) \ge t\right) \le \mathbb{P}_{\mathrm{GW}}\left((1 + 2\lambda S(H_2))\prod_{i=1}^{D_2}(1 + \lambda S(T_i^2)) \ge t\right)$$
$$+ \mathbb{P}_{\mathrm{GW}}\left((1 + 2\lambda S(H_{m+1}))\prod_{i=1}^{D_{m+1}}(1 + \lambda S(T_i^{m+1})) \ge t\right)$$
$$\le 2t^{-\sqrt{d}}(\log t)^{-2} + 2t^{-\sqrt{d}}(\log t)^{-2} = 4t^{-\sqrt{d}}(\log t)^{-2}.$$

The logic to derive (6.3) for $M^l(\dot{\mathcal{H}})$ is similar. We first derive the recursive inequalities, based on (A.19) and the ideas from Proposition 5.2, We have

$$M_a^l(\dot{\mathcal{H}}) \leq 2\lambda M^l(H_a) \prod_{i=1}^{D_a}(1 + \lambda S(T_i^a))$$

$$+ \lambda(1 + 2\lambda S(H_a)) \sum_{i=1}^{D_a} M^{l-1}(T_i^a) \prod_{\substack{j=1 \\ j \neq i}}^{D_a}(1 + \lambda S(T_j^a)),$$

for $a \in \{2, m+1\}$. As done in Section 5.3.2, we set

$$T_0^a := H_a, \quad \widetilde{M}_i^a := \left(1 - \frac{\varepsilon}{10}\right)^{-(l-1)} M^{l-1}(T_i^a), \quad \text{and} \quad W_i^a := \widetilde{M}_i^a \vee S(T_i^a),$$

and derive from the above inequality that with $\widetilde{M}_a = (1 - \frac{\varepsilon}{10})^{-l} M_a(\dot{\mathcal{H}})$,

$$\left(1 - \frac{\varepsilon}{10}\right) \widetilde{M}_a \leq \prod_{i=0}^{D_a}(1 + 4\lambda W_i^a), \quad a \in \{2, \, m+1\}.$$

Then, the rest of the proof goes the same as Section A.3.2, yielding that

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right)$$

$$\leq \sum_{a \in \{2, m+1\}} \mathbb{P}_{\mathrm{GW}}\left(2 \prod_{i=0}^{D_a}(1 + 4\lambda W_i^a) \geq t\right) \leq 4t^{-\sqrt{d}}(\log t)^{-2},$$

for all $t \geq d^{\frac{1}{10}}$, hence establishing Step 3-2. Combining the argument in Sections A.3.1—A.3.4, we conclude the proof of Proposition 6.4 and Corollary 6.8. $\qquad\square$

Appendix B. Proof of technical lemmas

B.1. **Remaining proof of Lemma A.2.** We finish the proof of Lemma A.2 by establishing (A.2) and (A.3). The ideas are very similar as in the proofs of Theorems 4.4 and 5.1.

Remaining proof of Lemma A.2. Let $\mathcal{F}^{m,l}$, $\{v_j\}_{j=1}^m$ and $\{\mathcal{T}_j\}_{j=1}^m$ be as in the definition of $\mathcal{F}^{m,l}$. Further, recall the definitions of D, $\{u_i\}_{i=1}^D$, $\{T_i\}_{i=1}^D$ and F from the proof of Lemma A.2.

For any fixed l, assume that (A.2) and (A.3) are true for m. We show that under this assumption, they are true for $m+1$. (Note that the case $m = 1$ is trivial, since $\mathcal{F}^{1,l}$ is a single vertex.) We begin with presenting the recursion inequalities which follow from straight-forward generalizations of previous propositions.

We first have

$$S(\mathcal{F}^{m,l}) \leq \frac{1 + \lambda\left(S(F) + \sum_{i=1}^D S(T_i)\right)}{1 - \lambda - 2\lambda^2\left(S(F) + \sum_{i=1}^D S(T_i)\right)};$$

$$M(\mathcal{F}^{m,l}) \leq \frac{\lambda\left(M(F) + \sum_{i=1}^D M(T_i)\right)}{1 - \lambda - 2\lambda^2\left(S(F) + \sum_{i=1}^D S(T_i)\right)},$$

which are basically rephrasings of (4.11) and (5.10), respectively. Based on these inequalities, we can repeat the argument from Sections 4.2.1 and 5.3.1 to deduce (A.2) and (A.3) for $t \leq d^{\frac{1}{10}}$ and $m + 1$.

For the other case, $t \geq d^{\frac{1}{10}}$, we observe that

$$S(\mathcal{F}^{m,l}) \leq (1 + \lambda S(F)) \prod_{i=1}^{D} (1 + \lambda S(T_i));$$

$$M(\mathcal{F}^{m,l}) \leq \lambda M(F) \prod_{i=1}^{D} (1 + \lambda S(T_i)) + \lambda (1 + \lambda S(F)) \sum_{i=1}^{D} M(T_i) \prod_{\substack{j=1 \\ j \neq i}}^{D} (1 + \lambda S(T_j)),$$

which come from (4.1) and (5.3). Note that the inequality for $M(\mathcal{F}^{m,l})$ is also reminiscent of (A.17). Then, we follow the logic from Sections 4.2.2, 4.2.3, 5.3.2 and 5.3.3 to obtain (A.2) and (A.3) for $t \geq d^{\frac{1}{10}}$ and $m + 1$.   □

B.2. **Proof of Claim A.3.** In this subsection, we finish the remaining proof of Claim A.3

Remaining proof of Claim A.3. Recall the definitions of $\widetilde{B}$ and $\widehat{S}$ in (A.8). In the setting of Claim A.3, we are assuming (A.4) for m, which is

(B.1) $$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq 2.$$

The bound on $\widehat{S}$ was derived in Sections 4.2.2 and 4.2.3 which yielded that

(B.2) $$\mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq d^{\frac{1}{10}}.$$

Set $\alpha = d^{\frac{1}{10}}$ and $k_0 = \lfloor \frac{\log t}{\log \alpha} \rfloor - 1$. The rest of the proof is analyzing (A.9). The assumptions (B.1) and (B.2) give that each of the first two terms in its r.h.s. is smaller than $\frac{1}{4} t^{-\sqrt{d}} (\log t)^{-2}$. Then, the third term can also be controlled based on observing that for $k \in [k_0]$ and $t \geq d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\alpha^{-k}\right) \cdot \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq \frac{\alpha^{k-1}}{2} d^{\frac{1}{4}}\right)$$

(B.3)
$$\leq \left(\frac{\alpha^k}{t} \cdot \frac{2}{d^{\frac{1}{4}}\alpha^{k-1}}\right)^{\sqrt{d}} \left(\frac{1}{\log(t\alpha^{-k})}\right)^2 \left(\frac{1}{\log\left(\frac{1}{2}\alpha^{k-1}d^{\frac{1}{4}}\right)}\right)^2$$

$$\leq \left(d^{\frac{1}{8}}t\right)^{-\sqrt{d}} \frac{1}{(\log t - k\log\alpha)^2} \frac{1}{\left(\frac{1}{8}\log d + k\log\alpha\right)^2}.$$

Observing that $\log t - k_0 \log \alpha \geq \log \alpha = \frac{1}{10} \log d$, we can apply Lemma 4.11 in order to control the sum of (B.3) over k. All in all, we deduce that (A.8) holds true for $t \geq d^{\frac{1}{10}}$.   □

B.3. **Proof of Lemma A.4.** The goal of this subsection is to present the proof of Lemma A.4.

Proof of Lemma A.4. We begin with establishing the recursive inequalities for $S(\mathcal{A}^{\underline{m},l})$, $M(\mathcal{A}^{\underline{m},l})$ and $B(\mathcal{A}^{\underline{m},l})$. To this end, we first define some subgraphs of $\mathcal{A}^{\underline{m},l}$ to generate the recursions. Recall the definitions of $\{v_j\}_{j=-m_1}^{m_2}$ and $\{\mathcal{T}_j\}_{j=-m_1+1}^{m_2-1}$ from the beginning of Section A.2.

- Let D be the degree of v_0 inside $\mathcal{T}_0$, and let $T_1, \dots, T_D$ be the subtrees of $\mathcal{T}_0$ rooted at each child of v_0 in $\mathcal{T}_0$.
- Let F_1 be the graph consists of the length-$(m_1 - 1)$ line $v_{-1} \dots v_{-m_1}$ and the trees $\mathcal{T}_{-1}, \dots, \mathcal{T}_{-m_1+1}$ rooted at $v_{-1}, \dots, v_{-m_1+1}$, respectively.
- Let F_2 be defined similarly in the other branch of $\mathcal{A}^{\underline{m},l}$. That is, it consists of the length-$(m_2 - 1)$ line $v_1 \dots v_{m_2}$ and the trees $\mathcal{T}_1, \dots, \mathcal{T}_{m_2-1}$ rooted at $v_1, \dots, v_{m_2-1}$, respectively.
- For convenience, we set $T_{D+1} := F_1$ and $T_{D+2} := F_2$. Note that the graphs $T_1, \dots, T_{D+2}$ are all trees.

Based on the same ideas as the derivation of (4.11), (5.10) and (A.5), we have the first recursive inequalities as follows.

(B.4)
$$S(\mathcal{A}^{\underline{m},l}) \leq \frac{1 + \sum_{i=1}^{D+2} S(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^{D+2} S(T_i)}; \quad M(\mathcal{A}^{\underline{m},l}) \leq \frac{\lambda \sum_{i=1}^{D+2} M(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^{D+2} S(T_i)};$$
$$B(\mathcal{A}^{\underline{m},l}) \leq \frac{\lambda(B(T_{D+1}) + B(T_{D+2}))}{1 - \lambda - 2\lambda^2 \sum_{i=1}^{D+2} S(T_i)}.$$

Moreover, the second recursive inequalities are obtained from Propositions 4.1, 5.2 and from (A.6), which read

$$S(\mathcal{A}^{\underline{m},l}) \leq \prod_{i=1}^{D+2} (1 + \lambda S(T_i)); \quad M(\mathcal{A}^{\underline{m},l}) \leq \lambda \sum_{i=1}^{D+2} M(T_i) \prod_{\substack{j=1 \\ j \neq i}}^{D+2} (1 + \lambda S(T_i));$$

(B.5)
$$B(\mathcal{A}^{\underline{m},l}) \leq \lambda \sum_{i=D+1}^{D+2} B(T_i) \prod_{\substack{j=1 \\ j \neq i}}^{D+2} (1 + \lambda S(T_i)).$$

Note that we know the tail probabilities for $S(T_i)$, $i \in [D + 2]$ by Theorems 4.4, 5.1 and Lemma A.2. Therefore, we can deduce (A.10) and (A.11) by following Sections 4.2.1—4.2.3 and 5.3.1—5.3.3.

To see (A.12), we rely on Lemma A.2. For $t \leq d^{\frac{1}{10}}$, we have the following inequality reminiscent of (A.7).

$$\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{A}^{\underline{m},l}) \geq d^{\frac{3m_1 \wedge m_2}{4}} t\right) \leq \mathbb{P}\left(D \geq \left(1 + \frac{\varepsilon}{6}\right) d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D+2} S(T_i) \geq \frac{2d}{\varepsilon}\left(1 + \frac{\varepsilon}{3}\right)\right)$$
$$+ P\left(B(F_1) \geq d^{-\frac{3}{4}(m_1-1)+\frac{1}{2}}\right) + P\left(B(F_2) \geq d^{-\frac{3}{4}(m_2-1)+\frac{1}{2}}\right).$$

By estimating the terms in the r.h.s. as in Case 1 of the proof of Lemma A.2 and Section 4.2.1, we see that the above is smaller than $t^{-\sqrt{d}}(\log t)^{-2}$ for $t \leq d^{\frac{1}{10}}$.

For $t \geq d^{\frac{1}{10}}$, we set

$$\widetilde{B}(\mathcal{A}^{\underline{m},l}) := d^{\frac{3}{4}(m_1 \wedge m_2)} B(\mathcal{A}^{\underline{m},l}), \quad \text{and} \quad \widetilde{B}_a := d^{\frac{3}{4}(m_a-1)} B(F_a), \quad a \in \{1, 2\},$$

and observe that by (B.5),

$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}(\mathcal{A}^{\underline{m},l}) \geq t\right) \leq \sum_{a \in \{1,2\}} \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}_a \prod_{\substack{j=1 \\ j \neq D+a}}^{D+2} (1 + \lambda S(T_i)) \geq \frac{d^{\frac{1}{4}} t}{3}\right).$$

Then, the summand in the r.h.s. can be bounded by applying Claim A.3 and implies (A.12) for $t \geq d^{\frac{1}{10}}$. $\square$

B.4. Proof of Lemma 6.10. Here, we establish Lemma 6.10. Indeed, the third statement in the lemma was proven in Lemma 4.3 of [3]. Therefore, we focus on the first two statements of the lemma.

Proof of Lemma 6.10. We begin with establishing the first statement. We only show
$$d^\sharp \leq \left(1 + \frac{\varepsilon}{9}\right) d,$$
since $\widetilde{d^\sharp} \leq (1 + \frac{\varepsilon}{9})\widetilde{d}$ follows analogously. Let Z be as in Definition 6.9. We first note that
$$Z \geq \sum_k \left(1 - \frac{\varepsilon}{10}\right) p_k = 1 - \frac{\varepsilon}{10}.$$
Moreover, when $k_0 < k_{\max}$,
$$\sum_{k \leq k_0} k\left(1 - \frac{\varepsilon}{10}\right) p_k + \sum_{k > k_0} k\sqrt{p_k} \leq \sum_k k\left(1 - \frac{\varepsilon}{10}\right) p_k + \frac{\varepsilon}{10} \leq \left(1 - \frac{\varepsilon}{10}\right) d + \frac{\varepsilon}{10}.$$
This gives us $d^\sharp \leq \left(1 + \frac{\varepsilon}{9}\right) d$, for all d larger than some absolute constant. In the other case when $k_0 = k_{\max}$, we are done if $k_0 \leq (1 + \frac{\varepsilon}{10})d$. If not, the concentration condition gives that
$$p_{k_0} \leq \exp(-c_{\varepsilon'} d)$$
with $\varepsilon' = \frac{\varepsilon}{10}$, and hence for d larger than some constant depending on ε and $\mathfrak{c}$, we obtain
$$\sum_{k < k_0} k p_k + k_0 \sqrt{p_{k_0}} \leq \left(1 - \frac{\varepsilon}{10}\right) d + \varepsilon,$$
which again implies $d^\sharp \leq \left(1 + \frac{\varepsilon}{9}\right) d$.

For the second, we only establish the first of (6.4), since the second line follows analogously. Let $\varepsilon' \in [\frac{\varepsilon}{10}, 1]$. For any integer $k \in [(1 + \varepsilon')d, 2d]$, we have
$$\mu^\sharp(k) \leq \frac{\sqrt{p_k}}{Z} \leq 2\exp\left(-\frac{c_{\varepsilon'}}{2}d\right).$$
Therefore, we obtain that
$$\mathbb{P}_{D^\sharp \sim \mu^\sharp}\left(D^\sharp \geq (1 + \varepsilon')d\right) \leq \sum_{k = \lceil (1+\varepsilon')d \rceil}^{\lfloor 2d \rfloor} 2\exp\left(-\frac{c_{\varepsilon'}}{2}d\right) + \sum_{k \geq \lceil 2d \rceil} 2\exp\left(-\frac{c_1 k}{2}\right)$$
$$\leq 2d\exp\left(-\frac{c_{\varepsilon'}}{2}d\right) + \frac{2e^{-c_1 d}}{1 - e^{-\frac{c_1}{2}}} \leq \exp\left(-\frac{c_{\varepsilon'} \wedge c_1}{3}d\right),$$
where the last inequality holds for any d larger than some constant $d_0(\mathfrak{c})$. Therefore, we can set $c_{\varepsilon'}' = \frac{1}{3}(c_{\varepsilon'} \wedge c_1)$ and obtain the first line of (6.4). $\square$

B.5. Proof of Lemma 6.11. Lemma 4.5 of [3] had a very similar statement as Lemma 6.11, but it was weaker in terms of determining the scope of γ. It turns out that in order to prove the stronger version, Lemma 6.11, the following bound on the size of $N(v, l)$ is needed.

Lemma B.1. *Under the setting of Lemma 6.11, let $\mu^{\sharp} := \mu_{\varepsilon}^{\sharp}$ be the augmented distribution of μ and $\widetilde{\mu}^{\sharp}$ be its size-biased distribution. Further, let $L_n := \frac{1}{10} \log_d n$ and $\mathcal{T} \sim \mathsf{GW}(\mu^{\sharp}, \widetilde{\mu}^{\sharp})^{L_n}$, where d is the mean of $\widetilde{\mu}$, the size-biased distribution of ν. Then, there exists $d_0(\varepsilon, \mathfrak{c}) > 0$ such that if $d \geq d_0$, then we have*

$$\mathbb{P}\left(|\mathcal{T}| \geq n^{\frac{1}{5}}\right) \leq \exp\left(-n^{\frac{1}{20}}\right).$$

Proof. Let $c_1' \in \mathfrak{c}'$ be the constant in Lemma 6.10. Then, we first note that the second statement of Lemma 6.10 tells us the following:

$$(\text{B.6}) \qquad \mathbb{E}_{D \sim \widetilde{\mu}^{\sharp}}\left[\exp\left(\frac{c_1'}{2} D\right)\right] \leq \exp(2d) + \sum_{k \geq 2d} e^{\frac{c_1' k}{2}} e^{-c_1' k} \leq e^{3d},$$

where the last inequality holds if d is larger than some constant depending on c_1'. Moreover, it is straight-forward to see that the same bound holds for $\mu^{\sharp}$ as well. We also note that

$$(\text{B.7}) \qquad \left(\frac{6d}{c_0'}\right)^{L_n} \leq n^{\frac{1}{9}},$$

if d is larger than some constant depending on c_0'.

Let Y_h be the number of vertices at depth h of $\mathcal{T}$. Further, let $D \sim \widetilde{\mu}^{\sharp}$ and set $c'' = \frac{c_0'}{2}$. Note that Hölder's inequality implies $\mathbb{E}e^{aX} \leq (\mathbb{E}e^X)^a$ for $0 \leq a \leq 1$. This gives that

$$\mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^h Y_h\right)\right] = \mathbb{E}\left[\mathbb{E}_D\left[\exp\left(\left(\frac{c''}{3d}\right)^h D\right)\right]^{Y_{h-1}}\right]$$

$$\leq \mathbb{E}\left[\exp\left\{3d \cdot \frac{(c'')^{h-1}}{(3d)^h} Y_{h-1}\right\}\right] = \mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{h-1} Y_h\right)\right].$$

Hence, we obtain that

$$(\text{B.8}) \qquad \mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{L_n} \sum_{h=0}^{L_n} Y_h\right)\right] \leq \sum_{h=0}^{L_n} \left(\frac{c''}{3d}\right)^h \leq 2,$$

if d is larger than some constant depending on $\mathfrak{c}$. (Note that (B.6) holds for $\mu^{\sharp}$ as well.) Therefore, combining (B.8) with (B.7), we deduce that

$$\mathbb{P}\left(\sum_{h=0}^{L_n} Y_h \geq n^{\frac{1}{5}}\right) \leq \exp\left(-\left(\frac{c''}{3d}\right)^{L_n} n^{\frac{1}{5}}\right) \times \mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{L_n} \sum_{h=0}^{L_n} Y_h\right)\right]$$

$$\leq \exp\left(-n^{\frac{1}{20}}\right).$$

$\square$

Proof of Lemma 6.14. In Lemma 4.5, [3], its proof reduces to obtaining tail probability bounds of $\sum_{h=0}^{L_n} Y_h$ for Y_h as above. Having Lemma B.1 in hand, following the proof of Lemma 4.5, [3] deduces Lemma 6.14. $\square$

References

[1] D. Aldous and J. A. Fill, *Reversible Markov chains and random walks on graphs*, Unfinished monograph, recompiled 2014, available at `http://www.stat.berkeley.edu/$\sim$aldous/RWG/book.html`, 2002.

[2] Carol Bezuidenhout and Geoffrey Grimmett, *The critical contact process dies out*, Ann. Probab. **18** (1990), no. 4, 1462–1482. MR1071804

[3] Shankar Bhamidi, Danny Nam, Oanh Nguyen, and Allan Sly, *Survival and extinction of epidemics on random graphs with general degree*, Ann. Probab. **49** (2021), no. 1, 244–286, DOI 10.1214/20-AOP1451. MR4203338

[4] Amir Dembo and Andrea Montanari, *Gibbs measures and phase transitions on sparse random graphs*, Braz. J. Probab. Stat. **24** (2010), no. 2, 137–211, DOI 10.1214/09-BJPS027. MR2643563

[5] Rick Durrett, *Random graph dynamics*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 20, Cambridge University Press, Cambridge, 2007. MR2271734

[6] Rick Durrett, *Probability: theory and examples*, 4th ed., Cambridge Series in Statistical and Probabilistic Mathematics, vol. 31, Cambridge University Press, Cambridge, 2010, DOI 10.1017/CBO9780511779398. MR2722836

[7] T. E. Harris, *Contact interactions on a lattice*, Ann. Probability **2** (1974), 969–988, DOI 10.1214/aop/1176996493. MR356292

[8] Xiangying Huang and Rick Durrett, *The contact process on random graphs and Galton Watson trees*, ALEA Lat. Am. J. Probab. Math. Stat. **17** (2020), no. 1, 159–182, DOI 10.30757/alea.v17-07. MR4057187

[9] Svante Janson, *The probability that a random multigraph is simple*, Combin. Probab. Comput. **18** (2009), no. 1-2, 205–225, DOI 10.1017/S0963548308009644. MR2497380

[10] Jeong Han Kim, *Poisson cloning model for random graphs*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 873–897. MR2275710

[11] Steven Lalley and Wei Su, *Contact processes on random regular graphs*, Ann. Appl. Probab. **27** (2017), no. 4, 2061–2097, DOI 10.1214/16-AAP1249. MR3693520

[12] Thomas M. Liggett, *Multiple transition points for the contact process on the binary tree*, Ann. Probab. **24** (1996), no. 4, 1675–1710, DOI 10.1214/aop/1041903202. MR1415225

[13] Thomas M. Liggett, *Stochastic interacting systems: contact, voter and exclusion processes*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 324, Springer-Verlag, Berlin, 1999, DOI 10.1007/978-3-662-03990-8. MR1717346

[14] Thomas M. Liggett, *Interacting particle systems*, Classics in Mathematics, Springer-Verlag, Berlin, 2005. Reprint of the 1985 original, DOI 10.1007/b138374. MR2108619

[15] Eyal Lubetzky and Allan Sly, *Cutoff phenomena for random walks on random regular graphs*, Duke Math. J. **153** (2010), no. 3, 475–510, DOI 10.1215/00127094-2010-029. MR2667423

[16] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993), Random Structures Algorithms **6** (1995), no. 2-3, 161–179, DOI 10.1002/rsa.3240060204. MR1370952

[17] Jean-Christophe Mourrat and Daniel Valesin, *Phase transition of the contact process on random regular graphs*, Electron. J. Probab. **21** (2016), Paper No. 31, 17, DOI 10.1214/16-EJP4476. MR3492935

[18] Robin Pemantle, *The contact process on trees*, Ann. Probab. **20** (1992), no. 4, 2089–2116. MR1188054

[19] Jonathon Peterson, *The contact process on the complete graph with random vertex-dependent infection rates*, Stochastic Process. Appl. **121** (2011), no. 3, 609–629, DOI 10.1016/j.spa.2010.11.003. MR2763098

[20] A. M. Stacey, *The existence of an intermediate phase for the contact process on trees*, Ann. Probab. **24** (1996), no. 4, 1711–1726, DOI 10.1214/aop/1041903203. MR1415226

[21] Remco van der Hofstad, *Random graphs and complex networks. Vol. 1*, Cambridge Series in Statistical and Probabilistic Mathematics, [43], Cambridge University Press, Cambridge, 2017, DOI 10.1017/9781316779422. MR3617364

[22] Xiaofeng Xue, *Contact processes with random vertex weights on oriented lattices*, ALEA Lat. Am. J. Probab. Math. Stat. **12** (2015), no. 1, 245–259. MR3343484

[23] Xiaofeng Xue, *Critical value for the contact process with random recovery rates and edge weights on regular tree*, Phys. A **462** (2016), 793–806, DOI 10.1016/j.physa.2016.06.001. MR3531601

[24] Xiaofeng Xue and Yu Pan, *Contact processes with random recovery rates and edge weights on complete graphs*, J. Stat. Phys. **169** (2017), no. 5, 951–971, DOI 10.1007/s10955-017-1898-4. MR3719634

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544
Email address: `dhnam@princeton.edu`

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544
Current address: Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912
Email address: `oanh_nguyen1@brown.edu`

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544
Email address: `asly@princeton.edu`