From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Mathematics

2021

Phase transitions of epidemics on random graphs

Nam, Dong Hoon

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01ws859j73h

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Phase transitions of epidemics on random graphs

Dong Hoon Nam

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Mathematics
Adviser: Allan Murray Sly

May 2021

© Copyright by Dong Hoon Nam, 2021.

All Rights Reserved

Abstract

This thesis focuses on the study of two mathematical models of epidemics, namely the *contact process* and the *threshold-θ contact process*. Until recently, rigorous understanding of their phase transition was mostly limited to the cases of elementary underlying networks such as lattices, homogeneous trees, and random regular graphs. The purpose of this work is to present recent advances in analyzing their behavior on general random graphs, including Galton-Watson trees, Erdős-Rényi random graphs, and the configuration model.

In the contact process, each individual is either infected or healthy. An infected individual can infect its neighbors or become healthy, with both events happening independently at certain rates. Its behavior can intuitively be understood as follows: when the infection is weak, the epidemic dies out quickly (extinction phase); on the other hand, when the infection is strong, the epidemic survives for a long time (survival phase). We propose new techniques to study the survival time of the process and establish the necessary and sufficient criteria for the existence of the extinction phase on Galton-Watson trees. Moreover, the corresponding analogs on random graphs are developed. Additionally, we extend our methods to derive the asymptotics for the value of the infection rate at the threshold.

We then consider the discrete-time threshold-θ contact process, where an individual becomes infected in the next time step with probability p, if it had at least θ neighboring infections in the current step. On (sparse) Erdős-Rényi graphs with large enough average degree, we establish a discontinuous phase transition in the emergence of metastability: (i) if p is large enough, the process starting from a high density of infections shows a long survival, keeping its infection density above a certain level; (ii) if either p is small or the initial state has a low density of infections, the process exhibits a short survival. We further present the analogs of the result on other random graphs. Finally, we discuss the contrasting result for the low-degree random regular graph on which the process only displays a short survival.

Acknowledgements

First of all, I would like to sincerely thank Prof. Allan Sly, my PhD advisor, for all the guidance and support during my years at Princeton. In the beginning of graduate school, I was truly lucky to attend his wonderful lectures, and they led me into the world of probability. In addition to his endless enthusiasm and deep insight into the subject, he has been extremely generous with his time and also in sharing interesting problems and ideas. Talking about challenging math problems with him is one of the things that I am going to miss the most from my Princeton life. He has been much more than just an amazing mentor and a great collaborator, and I will never have enough words to thank him.

I am deeply indebted to Prof. Rick Durrett, whose constant encouragement and helpful comments broadened my perspectives. I was fortunate to participate in the 2019 AMS MRC he organized, where I learned lots about stochastic spatial models, including the threshold contact process discussed in this thesis. I appreciate him taking the time and energy to talk with me about many interesting topics and to review some of my drafts very carefully.

My deep appreciation extends to Prof. Michael Aizenman, who taught me the basics of probability and statistical mechanics in the early periods of my PhD. I was able to establish a firm understanding on the fundamentals thanks to his guidance and insights. I also want to thank Prof. Shankar Bhamidi, Prof. Eyal Lubetzky, and Prof. Miklos Racz for their interest, encouragement, and support in my research.

I take this opportunity to express my sincere gratitude to Prof. Evita Nestoridi and Prof. Oanh Nguyen. They have been wonderful mentors, amazing co-workers, and reliable friends who have made my research and life at Princeton much easier and more enjoyable. They were always there to help me with everything and inspired me in many different ways.

My thanks also goes to the many bright colleagues working in probability whom I have tremendously benefited from working with. In particular, I would like to thank Lingfu Zhang, Dor Elboim, Dylan Airey, Shuangping Li, Dr. Duncan Dauvergne, and Youngtak Sohn for motivating discussions and enthusiastic collaborations.

I am wholeheartedly grateful to Prof. Donghun Lee, Prof. Junehyuk Jung, Dr. In-Jee Jeong, and Dr. Joonhyun La for their help and friendship. Their support was essential in

settling at Princeton when I first came, and our conversations on math, life, and our careers have been an invaluable source of inspiration.

At Princeton, I have been truly fortunate to have wonderful classmates and friends. Thanks to their support, I have been able to not lose myself during a long journey of completing my PhD. In particular, I would like to thank Seok Hyeong, Noah, Gyujin, Sanghak, Sanghoon, Seungwook, Sang Woo, Woon Sang, Sangmin, Dahan, Nick, Sam, Alex, Wonjae, Yura, and Yeongjun.

The students who took my MAT 201 class in Fall 2018 deserve special thanks for making my life vivid and exciting. Interacting with them was always pleasurable, and they made teaching and office hours a source of refreshment.

I sincerely thank Jill Leclair and Marianne Ostendorf for their hospitality and for helping me with any issues related to my Ph.D studies and beyond. I am also grateful to the Samsung Scholarship for its financial support, which enabled me to focus more on my research.

I am further indebted to Prof. Nam-Gyu Kang, Prof. Sanghyuk Lee, and Prof. Panki Kim for their encouragement and support. I learned a lot from them during my undergraduate years and was able to build my passion in mathematics.

Last but not least, I deeply thank my family, my dearest friends Junho and Kyeongsik, and Erin, my significant other, for their companionship and moral support over the years, especially during the painful period of COVID-19. They have always been a great source of motivation and added meaning to my life.

Dedicated to my family.

Contents

Abstract iii

Acknowledgements iv

1 Introduction **1**

1.1 The contact process 2

1.1.1 Graphical representation 3

1.2 The threshold-θ contact process 4

1.3 Notations 5

1.4 Random graphs and their limiting structure 6

1.4.1 Galton-Watson trees 6

1.4.2 The Erdős-Rényi random graph 7

1.4.3 The configuration model 8

1.4.4 Local weak convergence 9

1.5 Results and related works: the contact process 10

1.5.1 On Galton-Watson trees 11

1.5.2 On the configuration model and the Erdős-Rényi random graphs . . 14

1.5.3 Related works 17

1.6 Results and related works: the threshold contact process 19

1.6.1 Results for more general random graphs 23

1.6.2 Metastability and the discontinuous phase transition on $\mathcal{G}(n, \mu)$ . . . 26

1.6.3 Discontinuous phase transition on infinite Galton-Watson trees . . . 27

1.6.4 Extinction on the infinite $(\theta + 1)$-regular tree 28

1.7 Organization 29

2 Phase transition of the contact process on random graphs **30**

2.1 Overview of the chapter 30

2.1.1 Organization 32

2.2 Extinction in Galton-Watson trees 32

2.2.1 Expected survival time in finite trees 33

2.2.2 Exponential decay of the infection depth 37

2.3 Short survival in random graphs 43

2.3.1 Coupling the local neighborhood 44

2.3.2 Estimating the survival time 49

2.3.3 Bounding the survival time: Proof of Theorem 2.3.1 53

2.4 Long survival in random graphs: Proof of Theorem 4, part 2 55

2.4.1 A structural lemma 55

2.4.2 Proof of Theorem 4, part 2 56

2.5 Long survival in random graphs: Proof of Theorem 5 59

2.6 Proof of the structural lemma of random graphs 63

2.6.1 The embedded expander property in the subexponential case 63

2.6.2 The embedded expander property in the exponential case 75

2.7 Proof of technical lemmas 76

2.7.1 Proof of Lemma 2.3.3 76

2.7.2 Local neighborhoods contain at most one cycle 78

2.7.3 Survival time on cyclic graphs 81

2.7.4 Survival time of the delayed process on cyclic graphs 82

2.7.5 Eliminating high-degree vertices from random graphs 85

3 First-order asymptotics of the critical value **89**

3.1 Overview of the chapter 89

3.1.1 The lower bound of Theorem 3 89

3.1.2 Deep infections, unicyclic neighbors and the lower bound of Theorem 7 90

3.1.3 Proof of Theorems 10 and 11 93

3.1.4 Organization 93

3.2 Survival and excursion times on trees 94
3.2.1 Deterministic recursions on trees 94
3.2.2 Recursive tail estimate for Galton-Watson trees 97
3.2.3 Proof of the recursive tail estimate: small t 99
3.2.4 Proof of the recursive tail estimate: intermediate t 100
3.2.5 Proof of the recursive tail estimate: large t 101
3.2.6 Proof of the induction lemmas 104
3.3 Total infections at leaves on trees 111
3.3.1 The first recursive inequality 111
3.3.2 The second recursive inequality 113
3.3.3 Recursive tail estimate for Galton-Watson trees 116
3.3.4 Tail estimate for total infections: small t 116
3.3.5 Tail estimate for total infections: intermediate t 118
3.3.6 Tail estimate for total infections: large t 119
3.4 Short survival on random graphs: proof of Theorem 7 119
3.4.1 Recursive analysis for unicyclic graphs 120
3.4.2 Coupling the local neighborhood of random graphs 123
3.4.3 Short survival on random graphs: Proof of Theorem 7 129
3.5 Recursive analysis for unicyclic graphs 133
3.5.1 Reduction of unicyclic graphs to trees 134
3.5.2 Back to unicyclic graphs and tail estimates for small t 138
3.5.3 Tail estimates on unicyclic graphs for large t: Overview 141
3.5.4 Tail estimates on unicyclic graphs for large t: Step 1 142
3.5.5 Tail estimates on unicyclic graphs for large t: Step 2 143
3.5.6 Tail estimates on unicyclic graphs for large t: Step 3-1 145
3.5.7 Tail estimates on unicyclic graphs for large t: Step 3-2 146
3.6 Supercritical phase for general distribution: Proof of Theorems 10 and 11 . 148
3.6.1 Preprocessing the graph 150
3.6.2 Good trees 152
3.6.3 Good vertices 158

3.6.4 A graph property of neighborhood expansion 159
3.6.5 Finishing the proof of Theorems 10 and 11 164
3.6.6 Finishing the proof of Theorems 3 and 7 166
3.7 Proof of technical lemmas 166
3.7.1 Remaining proof of Lemma 3.5.2 166
3.7.2 Proof of Claim 3.5.3 167
3.7.3 Proof of Lemma 3.5.4 168
3.7.4 Proof of Lemma 3.4.8 170
3.7.5 Refined analysis of cycle existence in local neighborhoods 171

4 The threshold-θ contact process on random graphs 174
4.1 Overview of the chapter 174
4.1.1 Organization 176
4.2 Structural analysis of random graphs 176
4.2.1 The cut-off line algorithm 177
4.2.2 Expansion of small subsets 183
4.2.3 Expansion of large subsets 186
4.3 Discontinuous phase transitions on random graphs 188
4.3.1 Proof of Theorem 12-1 188
4.3.2 Proof of Theorem 12-2: overview 189
4.3.3 Reducing the infection to a small size 190
4.3.4 Elemination of small infections 192
4.3.5 Proof of Theorem 14 193
4.4 Short survival on low-degree random regular graphs:
Proof of Theorem 13 197

Chapter 1

Introduction

Understanding the spread of epidemics on networks is a fundamental problem in science. Mathematicians and physicists have designed various mathematical models of infectious diseases and there have been extensive studies to investigate them using numerous techniques. In this thesis, we consider two models of epidemics, the *contact process* and the *threshold-θ contact process*. Even though their definitions are relatively simple and fundamental, rigorous understanding of their behavior has been mostly limited to the cases of elementary underlying networks such as lattices, homogeneous trees, and random regular graphs. The goal of this thesis is to present recent advances in analyzing their dynamics on general random graphs, including Galton-Watson trees, Erdős-Rényi random graphs, and the configuration model.

Our research interest in the mathematical study of epidemics arises from the following simple thoughts:

1. If the strength of infection is weak, then the disease will die out quickly;

2. If the disease is strongly infectious, it will survive in the network for a long time.

In the models of our interest, there will be a parameter that determines the infection strength. Somewhat surprisingly, these models often display *phase transitions*, which can roughly be described as follows:

- When the infection strength gradually increases (or decreases), there is a threshold (or thresholds) where the macroscopic dynamics of the disease changes discontinuously.

One example of such a transition is the following: the disease survives for time polynomial in the number of individuals in the network when the strength of infection is below the threshold, but it persists for an exponentially long time on the network beyond the threshold. The purpose of this work is to understand such phase transitions rigorously, for our two disease models defined on random graphs.

In the following sections of this chapter, we explain how the models are defined and discuss their basic properties. We then give an introduction to the background materials which are necessary in comprehending the rest of this thesis. Finally, we briefly survey the main results of this work which are based on the works of the author and his coworkers [BNNS21, NNS21, Nam19][1].

1.1 The contact process

The contact process, which is the first model of our interest, is introduced by Harris in 1974 [Har74]. As a stochastic process defined on a graph, its transitions are given as follows:

- Each vertex is either *infected* or *healthy*.
- Each infected vertex infects each of its neighbors independently at rate λ, and it becomes healthy at rate 1 independently of all the infections.
- Infection and recovery events in the process happen independently from vertex to vertex.

For further discussions, we first set up some notations and give a formal definition of the model. For a graph $G = (V, E)$ (finite or infinite), the contact process on G with infection rate $\lambda \in (0, \infty)$ is a continuous-time Markov chain on the state space $\{0,1\}^V$, where 0 (resp. 1) corresponds to the healthy (resp. infected) individual. When the vertices in $A \subset V$ are infected, we denote the state by the indicator $\mathbf{1}_A$:

$$\mathbf{1}_A(v) := \begin{cases} 1 & \text{if } v \in A; \\ 0 & \text{if } v \notin A. \end{cases} \tag{1.1.1}$$

[1]The works [BNNS21, NNS21] were presented in scholary conferences [BNNS18, BNNS19, BNNS20].

Note that any state in $\{0,1\}^V$ can be written as an indicator $\mathbf{1}_A$ for some $A \subset V$. If $\mathbf{1}_A$ is the initial state of the process, we denote the process by

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

We will frequently use the notation $\mathbf{0}$ for the all-healthy state $\mathbf{0} = \mathbf{1}_\emptyset$, and also write $\mathbf{1}_v = \mathbf{1}_{\{v\}}$ if the state has a single infected vertex v. The transition rule of the process can be described 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 $N_t(u)$ is the number of neighbors v of u with $X_t(v) = 1$.

We sometimes write $\mathbf{CP}^\lambda(G)$ when knowing the initial condition is unnecessary. For convenience, we usually denote the state space by $\{0,1\}^G$, that is, identifying G as the set of vertices.

1.1.1 Graphical representation

We briefly discuss a coupling method of the contact processes using a *graphical representation* based on [Lig05, Chapter 3, Section 6]. The idea is to record the infections and recoveries happening in $\mathbf{CP}^\lambda(G; \mathbf{1}_A)$ on the space-time domain $G \times \mathbb{R}_+$. Define i.i.d. Poisson processes $\{N_v(t)\}_{v \in V}$ with rate 1 and i.i.d. Poisson processes $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \overrightarrow{E}}$ with rate λ, where $\overrightarrow{E} = \{\vec{uv},\ \vec{vu} : (uv) \in E\}$ is the set of directed edges. Further, we let $\{N_v(t)\}_{v \in V}$ and $\{N_{\vec{uv}}(t)\}_{\vec{uv} \in \overrightarrow{E}}$ to be mutually independent. Then the graphical representation is defined as follows:

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

This gives a geometric picture of $\mathbf{CP}^\lambda(G; \mathbf{1}_A)$, and further provides a coupling of the processes over all possible initial states. Figure 1.1 tells us how to interpret the infections at

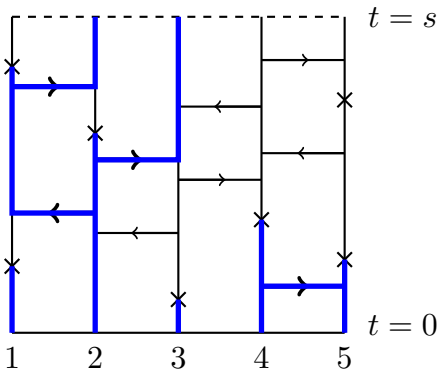


Figure 1.1: A realization of the contact process on the interval $V = \{1, \dots, 5\}$, with initial condition $X_0 = \mathbf{1}_V$. The blue lines describe the spread of infection. We see that $X_s = \mathbf{1}_{\{2,3\}}$

time t based on this graphical representation. We point out two lemmas which are easy consequences of the above construction. For proofs, see, e.g., [Lig05].

Lemma 1.1.1. *Suppose that we have the aforementioned coupling among the contact processes (with any initial states) on a graph G. Let T_v and T_G be the first time when $\mathbf{CP}^\lambda(G; \mathbf{1}_v)$ and $\mathbf{CP}^\lambda(G; \mathbf{1}_G)$ reach the all-healthy state $\mathbf{0}$, respectively. Then we have $T_G = \max\{T_v : v \in G\}$.*

Lemma 1.1.2. *For a given graph $G = (V, E)$ and any $A \subset V$, let $(X_t) \sim \mathbf{CP}^\lambda(G; \mathbf{1}_A)$. Consider any (random) subset $\mathcal{I}$ of $\mathbb{R}_+$, and define (X'_t) to be the coupled process of (X_t) that has the same initial state, infections and recoveries, except that the recoveries at a fixed vertex u are ignored at times $t \in \mathcal{I}$. Then for any $t \geq 0$, we have $X_t \leq X'_t$, i.e., $X_t(v) \leq X'_t(v)$ for all v.*

1.2 The threshold-θ contact process

The next model of interest is the discrete-time threshold-θ contact process with $\theta \geq 2$. As before, it is defined on a graph $G = (V, E)$, and the configuration of the process at time $t \in \mathbb{N}$ is $X_t \in \{0, 1\}^V$, where 0 and 1 denote 'healthy' and 'infected' individuals, respectively. Suppose that the state X_t at time t is given. Then its transitions are defined as follows.

- If a vertex v has at least θ neighbors u with $X_t(u) = 1$, then $X_{t+1}(v) = 1$ with probability p and $X_{t+1}(v) = 0$ with probability $1 - p$.

- If v has less than θ infected neighbors at time t, then $X_{t+1}(v) = 0$ with probability 1.
- At each time step, transitions happen independently and simultaneously at all sites.

We write

$$(X_t) \sim \mathsf{TCP}^{\theta,p}(G; \mathbf{1}_A),$$

to denote the law of the discrete-time threshold-θ contact process with infection probability p defined on the graph G with initial state $\mathbf{1}_A$.

Regarding 0 and 1 as 'vacant' and 'occupied by a particle,' respectively, the process can be thought of a population growth model with a nonlinear reproduction rule [Too74, BG91, Che92, Che94, DG85, MS09], where we need multiple neighboring parents to generate a new particle. It can also be seen as a dynamical version of the bootstrap percolation process (see, e.g., [AL03] for a review on this topic), which allows occupied sites to become vacant.

One interesting aspect of the threshold contact process is that on $\mathbb{Z}^d$ (with edges between every nearest neighbors), it is predicted to display a discontinuous phase transition [Too74]. To be specific, on the infinite graph $\mathbb{Z}^d$, we expect the density of infection in the upper invariant measure (i.e., the law of the system at $t \to \infty$ limit started from all-infected initial state) to be discontinuous at p_c, the point where a nonzero invariant measure emerges. However, its rigorous verification remains unanswered. Further discussions will be presented in Section 1.6.

1.3 Notations

In the rest of this thesis, we rely on the following notation: For two positive sequences (a_n) and (b_n), we say that $a_n = O(b_n)$ or $b_n = \Omega(a_n)$ if there exists a constant C independent of n such that $a_n \le Cb_n$ for all n. If $a_n = O(b_n)$ and $b_n = O(a_n)$, we write $a_n = \Theta(b_n)$.

Furthermore, we say that an event happens *with high probability* (in short, whp) if the probability that it happens converges to 1 as n tends to infinity. Here, n refers to the number of vertices in the underlying graph we work with.

For a tree T and an integer l, we denote by T_l the set of vertices of T at depth l. We use T^l and $T_{\le l}$ to denote the set of vertices of depth at most l. In particular, $\mathcal{T}^l \sim \mathsf{GW}(\xi)^l$ denotes

the Galton-Watson tree generated up to depth l, while the (possibly infinite) Galton-Watson tree is denoted by $\mathcal{T} \sim \mathbf{GW}(\xi)$. (See the following section for a formal definition of $\mathbf{GW}(\xi)$)

Throughout the article, we often work with $\mathbf{CP}^\lambda$ and $\mathbf{TCP}^{\theta,p}$ defined on a *fixed* graph G, but generated randomly. That is, we first produce a graph randomly, fix the resulting graph G, and run the process $\mathbf{CP}^\lambda$ or $\mathbf{TCP}^{\theta,p}$ on G. To eliminate ambituity between the two randomness of different nature, we introduce the following notations:

- $\mathbb{P}_{\text{CP}}$ and $\mathbb{E}_{\text{CP}}$ denote the probability and the expectation, respectively, with respect to the randomness from the contact processes $\mathbf{CP}^\lambda$ or $\mathbf{TCP}^{\theta,p}$;
- $\mathbb{P}_{\text{GW}}$ and $\mathbb{P}_{\text{RG}}$ 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. ($\mathcal{G}(n,\mu)$ is defined in the following section.) We write $\mathbb{E}_{\text{GW}}$ and $\mathbb{E}_{\text{RG}}$ 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, $\mathbb{E}[\cdot] = \mathbb{E}_{\text{GW}}[\mathbb{E}_{\text{CP}}[\cdot]]$, if the underlying graph is a Galton-Watson tree.

$\mathbb{P}$ and $\mathbb{E}$ are used as the symbols of probability and expectation of other general random variables as well, when their meaning is clear in the context.

1.4 Random graphs and their limiting structure

In this section, we briefly review fundamental models of random graphs and their basic properties. They will serve as the underlying networks of the disease in the rest of this thesis.

1.4.1 Galton-Watson trees

Let ξ be a probability distribution on $\mathbb{N} = \{0, 1, 2, \ldots\}$. The *Galton-Watson tree with offspring distribution* ξ is a random tree generated by the following procedure: We start with one particle, called the *root* of the tree. It has k children where k is a random number sampled from ξ. Then, each of these children (if there is any) also has children with the same

offspring distribution ξ, independently of each other and of their parent. This continues forever or until there are no more children.

Througout this article, the Galton-Watson tree with offspring distribution ξ obtained as above is denoted by $\mathcal{T} \sim \mathsf{GW}(\xi)$, where $\mathsf{GW}(\xi)$ is the symbol that refers to its law. The following Proposition describes a fundamental property of $\mathsf{GW}(\xi)$. For a proof, see, e.g., [LP16, Proposition 5.4].

Proposition 1.4.1. *Let $\mathcal{T} \sim \mathsf{GW}(\xi)$, and let $|\mathcal{T}|$ denote the number vertices in $\mathcal{T}$. Moreover, let m denote the mean of ξ. Then, the following hold true.*

- *If $m \leq 1$ and ξ is not identically 1, then $\mathbb{P}(|\mathcal{T}| < \infty) = 1$;*
- *Otherwise, $\mathbb{P}(|\mathcal{T}| = \infty) > 0$.*

Furthermore, it is not difficult to compute $\mathbb{P}(|\mathcal{T}| = \infty)$ using the moment generating function of ξ. We refer to [LP16, Chapter 5] for details.

Our interest in this article is the study of infinite Galton-Watson trees, and hence we always assume that $m := \mathbb{E}\xi > 1$. (The case $\xi \equiv 1$ is less interesting, since the resulting graph is the same as $\mathbb{Z}$.) The main reason why we want $\mathcal{T}$ to be infinite is to have the possibility of the contact process running forever on $\mathcal{T}$, which is impossible to happen on a finite graph.

1.4.2 The Erdős-Rényi random graph

The Erdős-Rényi random graph graph is arguably the most fundamental random graph model in the mathematics literature. The model, proposed by Erdős and Rényi [ER59, Bol01], is defined as follows: We start with an empty graph with n vertices. For each pair of vertices (so there are $\frac{n(n-1)}{2}$ of them), connect them with an edge with probability p independently of others. We denote the resulting graph as $G \sim \mathcal{G}^{\mathrm{ER}}(n, p)$.

The structure of the Erdős-Rényi graph largely depends on the scale of the parameter p. Our interest is the case where $p = d/n$ for a fixed constant $d > 0$, and it is sometimes called the *sparse* Erdős-Rényi graph. For such p, $\mathcal{G}^{\mathrm{ER}}(n, p)$ maintains a $O(1)$ average degree as n tends to infinity, and as we will see later in Section 1.4.4, it is closely related with Galton-Watson trees and the configuration model.

The sparse Erdős-Rényi graph $\mathcal{G}^{\text{ER}}(n, \frac{d}{n})$ exhibits a phase transition at $d = 1$. The following theorem summarizes the known result, and we refer to [Bol01, JLR00] for details. A more detailed description of the phase transition is discussed in [FK15, Chapter 2].

Theorem 1.4.2. *$G \sim \mathcal{G}^{\text{ER}}(n, \frac{d}{n})$ satisfies the following* whp.

- *If $d < 1$, the size of the largest connected component of G is $O(\log n)$;*
- *If $d > 1$, there exists a single connected component of G of size $\Theta(n)$, and the rest of the components are of size $O(\log n)$.*

Remark 1.4.3. When $d = 1$, the largest components are of size $\Theta(n^{2/3})$. More precisely, the profile of the component sizes rescaled by $n^{2/3}$ converges in distribution to the excursion lengths of the process $(Y_t - \min_{s \le t} Y_s)_{t \ge 0}$ where $Y_t := B_t - \frac{t^2}{2}$ for the standard Brownian motion (B_t). For details and its proof, see, e.g., [Dur07, Chapter 2].

The $\Theta(n)$-sized component in the second case of the theorem is called the *giant component*. We are only interested in this case, $d > 1$, since the contact process on G cannot exhibit an exponentially long survival unless there is such a giant. Thus, we assume $d > 1$ for all Erdős-Rényi graphs $\mathcal{G}^{\text{ER}}(n, \frac{d}{n})$ in the rest of this work.

1.4.3 The configuration model

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 $i = 1, \dots, n$, conditioned on $\sum_{i=1}^{n} d_i \equiv 0 \bmod 2$. 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 graph uniformly at random among all *simple* graphs with degree sequence $\{d_i\}_{i \in [n]} \sim$ i.i.d. μ. The difference here is that the configuration model is not necessarily a simple graph. However, if the second moment of

μ is finite, we have the following contiguity between the two models. For details, see, e.g., [vdH17], Chapter 7.

Lemma 1.4.4 ([Jan09, vdH17]). *Suppose that $\mathbb{E}_{D\sim\mu}D^2 < \infty$. Then, uniformly in n, we have*

$$\mathbb{P}_{G_n\sim\mathcal{G}(n,\mu)}(G_n \text{ is simple}) \in (0,1).$$

In particular, for any subset A_n of graphs with n vertices,

$$\mathbb{P}_{G_n\sim\mathcal{G}(n,\mu)}(G_n \in A_n) \to 0 \quad \text{implies} \quad \mathbb{P}_{G_n^u\sim\mathcal{G}^u(n,\mu)}(G_n^{\mathsf{u}} \in A_n) \to 0. \tag{1.4.1}$$

We remark that when $\mu = \mathrm{Pois}(d)$, the random graph $G_n \sim \mathcal{G}(n,\mu)$ is contiguous to the Erdős-Rényi random graph $G_n^{\mathrm{ER}} \sim \mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$ in the sense of (1.4.1). See [Kim06, Theorem 1.1] for details.

Throughout the article, we study the configuration model $\mathcal{G}(n,\mu)$ instead of $\mathcal{G}^{\mathsf{u}}(n,\mu)$, under the assumption $\mathbb{E}_{D\sim\mu}D^2 < \infty$. Further, it is well known that $G \sim \mathcal{G}(n,\mu)$ (and hence, $G \sim \mathcal{G}^{\mathsf{u}}(n,\mu)$) whp contains the unique connected component of size linear in n, if and only if $\mathbb{E}_{D\sim\mu}D(D-2) > 0$ (the meaning of this condition is explained in the following subsection; for details, see [MR95]), which is an analog of Theorem 1.4.2. Hence, we always assume $\mathbb{E}_{D\sim\mu}D(D-2) > 0$, which is the most interesting case for us. Otherwise, the graph decomposes into many small components and the contact process would not exhibit $e^{\Theta(n)}$-survival time for any $\lambda > 0$.

1.4.4 Local weak convergence

Given a sequence of random graphs G_n, let $N(v,r)$ be an induced subgraph of G_n consisting of vertices of distance at most r from v. Let $\mathbb{P}_n$ be the distribution of the neighborhood $N(v,r)$ where v is a uniformly chosen vertex of G_n. We say that a random rooted tree $\mathcal{T}$ is the local weak limit of G_n if for any finite r and any rooted tree T of depth at most r,

$$\lim_{n\to\infty} \mathbb{P}_n(N(v,r) \cong T) = \mathbb{P}(\mathcal{T}^r \cong T)$$

where $\cong$ denotes the graph isomorphsim between rooted graphs, and $\mathcal{T}^r$ is the rooted subtree of the first r generations of $\mathcal{T}$ with the same root.

We shall use the following known convergence of $\mathcal{G}(n,\mu)$ to its corresponding Galton-Watson tree. Define the *size-biased distribution* $\widetilde{\mu}$ to be

$$\widetilde{\mu}(k-1) = \frac{k\mu(k)}{\sum_{i=1}^{\infty} i\mu(i)}, \quad k = 1, 2, \dots. \tag{1.4.2}$$

Note that if $\mu = \mathrm{Pois}(d)$ then $\widetilde{\mu} = \mu$. Let $\mathcal{T}(\mu) \sim \mathbf{GW}(\mu, \widetilde{\mu})$ be the *size-biased* Galton-Watson tree in which the root has the offspring distribution μ while all other vertices have the offspring distribution $\widetilde{\mu}$. We also stress that the Galton-Watson tree is supercritical if and only if $\mathbb{E}_{\widetilde{D}\sim\widetilde{\mu}}\widetilde{D} > 1$, equivalent to $\mathbb{E}_{D\sim\mu}D(D-2) > 0$ which we saw above.

Lemma 1.4.5 ([DM10], 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 $\mathbf{GW}(\mu,\widetilde{\mu})$.

Similarly, the Galton-Watson tree with offspring distribution Pois(d) is the local weak limit of the Erdős-Rényi random graph $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$.

1.5 Results and related works: the contact process

We move back to the discussion of the epidemics models, and overview the main results developed in this work. We begin with the contact process, based on [BNNS21, NNS21].

For an infinite graph G with its root ρ, consider the contact process $(X_t) \sim \mathbf{CP}^{\lambda}(G; \mathbf{1}_\rho)$ that starts from a single infection at the root. The process can show the following three types of behaviors 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

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

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 [Har74, BG90], and also the books of Liggett [Lig99, Lig05] 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 works by Pemantle [Pem92] (for $d \geq 4$), Liggett [Lig96] (for $d = 3$), and Stacey [Sta96] (for a shorter proof that works for all $d \geq 3$). Moreover, we have from [Pem92] that

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

In particular, the first-order asymptotics of $\lambda_1(\mathbb{T}_d)$ is

$$\lambda_1(\mathbb{T}_d) \sim \frac{1}{d}, \tag{1.5.1}$$

as d becomes large.

1.5.1 On Galton-Watson trees

A natural interest is then to study the phase diagram of the contact process on Galton-Watson trees, although much less is known about this case. The first goal of this thesis is

to establish the necessary and sufficient criterion for $\lambda_1 > 0$. In particular, we deal with the extinction in Galton-Watson trees with unbounded offspring distribution, which is the result that has been first verified in [BNNS21]. Since the contact process would never run forever on a finite graph, our interest is in supercritical Galton-Watson trees in order to consider infinite trees.

For a Galton-Watson tree $\mathcal{T}$ with offspring distribution ξ such that $\mathbb{E}\xi > 1$, 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 $|\mathcal{T}| = \infty$. For details, see [Pem92, Proposition 3.1] or [Hua20, Proposition 1]. Thus, the constants $\lambda_1^{\mathrm{GW}}(\xi)$ and $\lambda_2^{\mathrm{GW}}(\xi)$ are well-defined.

Theorem 1 ([BNNS21])**.** *Consider the contact process on the Galton-Watson tree with offspring distribution ξ, and suppose that only the root ρ of the tree is initially infected. If ξ has an exponential tail, i.e., $\mathbb{E}e^{c\xi} < \infty$ for some $c > 0$, then there exists a constant $\lambda_0 = \lambda_0(\xi) > 0$ such that for all $\lambda < \lambda_0$ and for a.e. $\mathcal{T} \sim \mathsf{GW}(\xi)$, the process $\mathsf{CP}^\lambda(\mathcal{T}; \mathbf{1}_\rho)$ dies out almost surely. Therefore, in such a case, we have $\lambda_1^{\mathrm{GW}}(\xi) > 0$.*

Recently, Huang and Durrett [HD20a] proved that on Galton-Watson trees, $\lambda_2^{\mathrm{GW}}(\xi) = 0$ if the offspring distribution ξ is subexponential, i.e., $\mathbb{E}e^{c\xi} = \infty$ for all $c > 0$. Combining Theorem 1 with their result, we have the complete characterization on the existence of extinction on Galton-Watson trees. Moreover, Theorem 1 establishes a stronger version of the following conjecture by Huang and Durrett:

Conjecture 2 ([HD20a])**.** *Suppose that $\mathbb{P}(\xi \geq k) = (1-p)^k$ for all k larger than some constant K, and consider the contact process on the Galton-Watson tree with offspring distribution ξ. Then λ_2, the weak-strong survival threshold, is strictly positive.*

Remark 1.5.1. Durrett offered $1,000 for answering Conjecture 2, in his presentation at Northeast Probability Seminar on November 15th, 2018 (Figure 1.2[2]).

The challenge in understanding the infection time on trees with unbounded degree distributions is that the infection persists for a long time around high degree vertices as there are many neighbors from which it can be reinfected. Indeed, it was shown in [BBCS05]

[2]Presentation slides available at Durrett's website: https://services.math.duke.edu/∼rtd/Talks/zoe.pdf

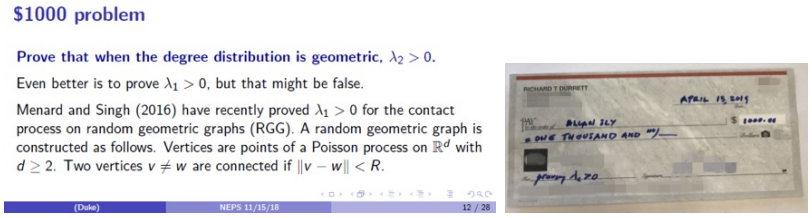


Figure 1.2: A slide from Durrett's presentation at Northeast Probability Seminar in November 2018, announcing Conjecture 2 (left), and the prize for solving it (right).

that the infection will last time $e^{c_\lambda k}$ in a neighborhood of a vertex of degree k with positive probability for some $c_\lambda > 0$. Thus exponential tails on the degree distribution are needed for there to be few enough high degree vertices in the tree for extinction to be certain.

We can further derive 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 (1.5.1) [NNS21].

Theorem 3 ([NNS21])**.** *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*

$$\begin{aligned} &\mathbb{P}(\xi_k \geq (1+\delta)d_k) \leq \exp(-c_\delta d_k) \quad \text{for all } \delta \in (0,1); \\ &\mathbb{P}(\xi_k \geq (1+a)d_k) \leq \exp(-c_1 a d_k) \quad \text{for all } a \geq 1. \end{aligned} \tag{1.5.2}$$

for all large enough k. Consider the contact process on the Galton-Watson tree $\mathcal{T}_k \sim \mathbf{GW}(\xi_k)$ with the root initially infected. Then, the extinction-survival threshold $\lambda_1^{\mathrm{GW}}(\xi_k)$ satisfies

$$\lim_{k\to\infty} \lambda_1^{\mathrm{GW}}(\xi_k) d_k = 1. \tag{1.5.3}$$

Remark 1.5.2. The concentration condition (1.5.2) we impose resembles the form of large deviation estimates. Notice that the family of Poisson distributions $\{\mathrm{Pois}(d)\}_{d>0}$ satisfies (1.5.2). This fact allows us to deduce an analogous result for Erdős-Rényi random graphs in Corollary 8.

Remark 1.5.3. The conclusion (1.5.3) is not always true without assuming (1.5.2). 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 very challenging problem.

Remark 1.5.4. In [Xue16], Xue showed (1.5.3) 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., [Pet11, Xue15, XP17] for more works from this perspective.

1.5.2 On the configuration model and the Erdős-Rényi random graphs

The next object of interest is the contact process on random graphs. For the contact process on the Erdős-Rényi random graph $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$, no rigorous results were known regarding its phase diagram—whether it shows short or long survival, or both. In this work, we prove that on $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$, the contact process exhibits two different phases depending on λ, as a consequence of an analogous criterion on more general random graphs.

We focus on studying the contact process on the configuration model with degree distribution μ, which we denote by $G \sim \mathcal{G}(n,\mu)$ (definitions given in the previous section). For the contact process on $G \sim \mathcal{G}(n,\mu)$, our main goal is to study how long the process survives in terms of the size of the graph. The next result of this thesis establishes the necessary and sufficient criterion for the contact process on $G \sim \mathcal{G}(n,\mu)$ to display the short survival phase. As mentioned in the previous section, we assume throughout that μ satisfies

$$\mathbb{E}_{D\sim\mu} D(D-2) > 0 \quad \text{and} \quad \mathbb{E}_{D\sim\mu} D^2 < \infty. \tag{1.5.4}$$

Theorem 4 ([BNNS21])**.** *Suppose that μ satisfies* (1.5.4) *and there exists some constant $c > 0$ such that $\mathbb{E}_{D\sim\mu} e^{cD} < \infty$. Consider $\mathbf{CP}^{\lambda}(G; \mathbf{1}_G)$, the contact process on $G \sim \mathcal{G}(n,\mu)$ where all vertices are initially infected. Then, there exist constants $0 < \underline{\lambda}(\mu) \le \overline{\lambda}(\mu) < \infty$ such that the following hold, whp over the choice of G:*

(1) For all $\lambda < \underline{\lambda}$, the survival time of the process is at most $n^{1+o(1)}$-time whp.

(2) For all $\lambda > \overline{\lambda}$, the survival time of the process is $e^{\Theta(n)}$-time whp.

Theorem 5 ([BNNS21])**.** *Suppose that μ satisfies (1.5.4) and $\mathbb{E}_{D\sim\mu} e^{cD} = \infty$ for all $c > 0$. Consider the contact process on $G \sim \mathcal{G}(n,\mu)$ where all vertices are initially infected. Then,*

for any fixed $\lambda > 0$ and whp *over the choice of G, the survival time of the process is $e^{\Theta(n)}$-time* whp.

For the case of Erdős-Rényi random graphs, which are contiguous to $\mathcal{G}(n,\mu)$ with $\mu = \mathrm{Pois}(d)$ (as we saw in Section 1.4.3), we can show that the contact process on $\mathcal{G}^{\mathrm{ER}}(n,\frac{d}{n})$ exhibits two different phases, as a consequence of Theorem 4.

Corollary 6 ([BNNS21])**.** *For any fixed $d > 1$, consider the contact process on $G \sim \mathcal{G}^{\mathrm{ER}}(n,\frac{d}{n})$ where all vertices are initially infected. Then there exist constants $0 < \underline{\lambda}(d) \le \overline{\lambda}(d) < \infty$ such that the following holds,* whp *over the choice of G:*

(1) For all $\lambda < \underline{\lambda}$, the survival time of the process is at most $n^{1+o(1)}$-time whp.

(2) For all $\lambda > \overline{\lambda}$, the survival time of the process is $e^{\Theta(n)}$-time whp.

Remark 1.5.5. One might be interested in studying the contact process on $G \sim \mathcal{G}(n,\mu)$, with the initial condition such that only a single vertex is infected. When a uniformly random vertex in G is infected initially while all the other ones are healthy, we will later see that Theorems 4 and 5 continue to hold, if we change "whp" to "with positive probability" at the end of the statements of Theorems 4-(2) and 5. To be precise, we mean that whp over the choice of $G \sim \mathcal{G}(n,\mu)$, with positive probability over the choice of the initially infected vertex v and with positive probability over the contact process. Proofs are given in Remark 2.4.5 for exponential distributions and Remark 2.5.4 for subexponential distributions.

Based on these results, we can formally define the short- and long-survival thresholds $\lambda_c^-(\mu)$, $\lambda_c^+(\mu)$ as follows.

$$\begin{aligned} \lambda_c^-(\mu) &:= \sup\left\{\lambda : \lim_{n\to\infty} \mathbb{P}(T_{\lambda,n} \le n^{C(\lambda)}) = 1 \text{ for some } C(\lambda) < \infty\right\}; \\ \lambda_c^+(\mu) &:= \inf\left\{\lambda : \lim_{n\to\infty} \mathbb{P}(T_{\lambda,n} \ge e^{c(\lambda)n}) = 1 \text{ for some } c(\lambda) > 0\right\}. \end{aligned} \tag{1.5.5}$$

In particular, Theorems 4 and 5 tell us that $0 < \lambda_c^-(\mu) \le \lambda_c^+(\mu) < \infty$ if μ has an exponential tail, and that $\lambda_c^-(\mu) = \lambda_c^+(\mu) = 0$ if μ is subexponential. The next result provides the first-order asymptotics for $\lambda_c^-(\mu)$ and $\lambda_c^+(\mu)$ for the former case, which is an analog of Theorem 3.

Theorem 7 ([NNS21]). *Let (μ_k) be a sequence of degree distributions with the size-biased distribution $\widetilde{\mu}_k$ (definition given in* (1.4.2)*). 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* (1.5.2) *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*

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

We can deduce an analog of Theorem 7 for the Erdős-Rényi random graph, based on the contiguity between $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$ and $\mathcal{G}(n, \mathrm{Pois}(d))$.

Corollary 8 ([NNS21]). *Consider the contact process on the random graph $G_n \sim \mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$ with all vertices initially infected. Then, the short- and long- survival thresholds of the process, defined analogously as* (1.5.5)*, satisfy*

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

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 9 ([BNNS21, NNS21]). *Let μ be a degree distribution satisfying the giant component condition* (1.5.4)*, and let $\widetilde{\mu}$ be its size-biased distribution. Recalling* (1.5.5)*, 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 [MV16] and Lalley-Su [LS17] who showed that the short-long survival threshold for random d-regular graph occurs exactly at $\lambda_1(\mathbb{T}_d)$.

The following 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 10 ([NNS21]). *Let μ be a degree distribution satisfying the giant component condition* (1.5.4)*, and let $\widetilde{\mu}$ be its size-biased distribution with $d := \mathbb{E}_{D\sim\widetilde{\mu}}D$. Recalling* (1.5.5)*, we have*

$$\lambda_c^+(\mu) \le \lambda_1^{\mathrm{GW}}(\widetilde{\mu}) \le \frac{1}{d-1}. \tag{1.5.6}$$

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 11 ([NNS21]). *For all $\lambda > \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$ and* whp *over the choice of $G_n \sim \mathcal{G}(n,\mu)$, 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 $G_n \sim \mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$, when $\lambda > \lambda_1^{\mathrm{GW}}(\mathrm{Pois}(d))$.*

1.5.3 Related works

In [Har74], Harris first introduced the contact process on $\mathbb{Z}^d$ and showed that the extinction-survival threshold $\lambda_c(\mathbb{Z}^d)$ satisfies $0 < \lambda_c(\mathbb{Z}^d) < \infty$ for any d. Building upon this work, the model on $\mathbb{Z}^d$ has been studied intensively and we refer to Liggett [Lig99] for a survey of results. Pemantle [Pem92] studied the contact process on the infinite d-ary tree $\mathbb{T}_d$ and showed that it exhibits three different phases—extinction, weak survival and strong survival—for $d \ge 3$. This result was later supplemented by [Lig96, Sta96] for the case $d = 2$.

Less is known for the contact process on general Galton-Watson trees. Recently, Huang and Durrett [HD20a] proved that on Galton-Watson trees, $\lambda_2 = 0$ if the offspring distribution ξ is subexponential. Along with Theorem 1, we now have the complete characterization of the existence of extinction in the contact process on Galton-Watson trees.

There has been considerable amount of work studying the phase transitions of survival times on large finite graphs. Stacey [Sta01] and Cranston et. al. [CMMV14] studied the contact process on the d-ary tree $\mathbb{T}_d^h$ of depth h starting from the all-infected state, and their results show that the survival time T_h, as $h \to \infty$, satisfies (i) $T_h/h \to \gamma_1$ in probability if $\lambda < \lambda_2(\mathbb{T}_d)$; (ii) $|\mathbb{T}_d^h|^{-1}\log \mathbb{E}T_h \to \gamma_2$ in probability and $T_h/\mathbb{E}T_h \xrightarrow{d} \mathrm{Exp}(1)$ if $\lambda > \lambda_2(\mathbb{T}_d)$,

where γ_1, γ_2 are constants depending on d, λ. In [DL88, DS88, Mou93], similar results were established for the case of the lattice cube $\{1, \dots, n\}^d$.

Recently in work of Mourrat and Valesin [MV16] and Lalley and Su [LS17], it was shown that for any $d \geq 3$, the contact process on the random d-regular graph, whose initial configuration is the all-infected state, exhibits the following phase transition, whp over the choice of $G \sim \mathcal{G}(n, d)$:

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

Moreover in [LS17], a "cutoff phenomenon" of the fraction of infected vertices was established. In [MV16], the same result as above is proven for $G \sim \mathcal{G}(n, \mu)$ with bounded μ (Theorems 1.3 and 1.4 of [MV16]). For an unbounded degree distribution μ, Chatterjee and Durrett [CD09] proved that if μ obeys a power law, then the contact process always displays long survival for any $\lambda > 0$, though their survival time was slightly weaker than exponential ($e^{n^{1-\delta}}$ for any $\delta > 0$). This result was later generalized in [MMVY16] to an exponential survival. Our Theorems 4 and 5 extend the aforementioned results to any general μ. In [SV17], a long survival on general graphs for $\lambda > \lambda_c(\mathbb{Z})$ was settled, with survival time at least $\exp(|G|/\{\log |G|\}^\kappa)$ for any $\kappa > 1$.

On random graphs with power-law degree distributions, metastability properties on the size of infected vertices were studied in [MVY13, CS15]. For other types of random graphs, recently in [MS16] it was shown that the contact process on random geometric graphs exhibits both short and long survival.

It is worth mentioning the works [AS10, SS14, MMVY16, DR17, Swa18, SV18, Xue18, JLM19, SS19, BdSV20, HP20, LMSV20] which study various different aspects of the contact process.

Furthermore, there are a plenty of variants of the contact process that have been of extensive interest in the literature, with the threshold-θ contact process being one of them. We address some of them [CSW18, FMMV19, FMV20, MPV19, DY20, HD20b, YD20, FMUV21] and references therein for the readers who are interested in this direction.

1.6 Results and related works: the threshold contact process

We move on to the discussion of the discrete-time threshold-θ contact process. Regarding this process, we outline the main results developed in this thesis based on [Nam19] and review the relevant background and literature.

Previously, Chatterjee and Durrett [CD13] studied this model on random d-regular graphs and infinite d-regular trees, and showed that for any $d \geq \theta + 2$, there is such a discontinuous phase transition. In particular, on random d-regular graph, they proved that

- If p is close to 1, the process starting from all-infected initial condition maintains high density of infection for exponentially long time;
- On the other hand, if the initial fraction of infection is very small, the process becomes extinct in $O(\log n)$ time,

where both events happen with probability tending to 1 as $n \to \infty$. Further, they asked if this is true for the configuration model with a general degree distribution μ. Our main result answers this question for μ with a light enough tail. Recall the definition of $G_n \sim \mathcal{G}(n, \mu)$.

Theorem 12. *Let $\theta \geq 2$ be an integer and μ be a distribution on $\mathbb{N}$ such that $D \sim \mu$ satisfies $\mathbb{P}(D \geq \theta + 2) = 1$ and $\mathbb{E}D^k < \infty$ for all $k > 0$. Let $(X_t) \sim \mathsf{TCP}^{\theta,p}(G_n; X_0)$ denote the discrete-time threshold-θ contact process on G_n with parameter p, starting from the state $X_0 \in \{0,1\}^{G_n}$. Then, whp over the choice of G_n, the following hold:*

1. *There exist $p_1 \in (0,1)$ and $\varepsilon_1, \gamma_1 > 0$ such that if $p \geq p_1$ and $|X_0| \geq (1-\varepsilon_1)n$, then*

$$\lim_{n\to\infty} \inf_{|X_0| \geq (1-\varepsilon_1)n} \mathbb{P}_{\mathrm{CP}}\Big(|X_t| \geq (1-\varepsilon_1)n \text{ for all } t \in [0, e^{\gamma_1 n}] \,\Big|\, X_0,\ G_n \Big) = 1.$$

   The constants $p_1, \varepsilon_1, \gamma_1$ depend on μ, θ.

2. *For any $p \in (0,1)$, there exist constants $\varepsilon_2, C_2 > 0$ depending on p, μ and θ, such that*

$$\lim_{n\to\infty} \inf_{|X_0| \leq \varepsilon_2 n} \mathbb{P}_{\mathrm{CP}}\Big(X_{C_2 \log n} = \mathbf{0} \,\Big|\, X_0,\ G_n \Big) = 1.$$

3. *There exist* $p_2 \in (0,1)$ *and* $C > 0$ *depending on* μ *and* θ*, such that for any* $p \le p_2$*,*

$$\lim_{n\to\infty} \mathbb{P}_{\mathrm{CP}}\Big(X_{C\log n} = \mathbf{0} \,\Big|\, X_0 = \mathbf{1}_{G_n},\, G_n\Big) = 1.$$

The second item directly implies the third item: for instance by setting $p_2 < \varepsilon_2$, we get $|X_1| \le \varepsilon_2 n$ and then a logarithmic time survival, both with probability tending to one. Nevertheless, we include item 3 separately for the clarity of the exposition.

For a given graph G_n, let $\tau_0(G_n) = \tau_0(G_n, \theta, p)$ be the expected time the the process on G_n starting from the all-infected state reaches at the all-healthy condition. Formally,

$$\tau_0(G_n, \theta, p) := \mathbb{E}_{\mathrm{CP}}\big[\inf\{t : X_t = \mathbf{0}\} \,\big|\, X_0 = \mathbf{1}_{G_n},\, G_n\big]. \tag{1.6.1}$$

We define the critical probability p_c as

$$p_c = p_c(\mu) := \lim_{\beta\to 0} \inf\Big\{p : \lim_{n\to\infty} \mathbb{P}_{\mathcal{G}(n,\mu)}\big(\tau_0(G_n, p) \ge e^{\beta n}\big) = 1\Big\}. \tag{1.6.2}$$

In words, p_c is the smallest p that admits an exponentially long running time on a typical graph $G_n \sim \mathcal{G}(n,\mu)$. We see that the first and the third conclusions of Theorem 12 imply $0 < p_c < 1$.

Theorem 12 further tells us that the threshold-θ contact process on $G_n \sim \mathcal{G}(n,\mu)$ exhibits a discontinuous phase transition, described in the following sense. If $p > p_c$ and $X_0 = \mathbf{1}_{G_n}$, then the process runs for exponentially long time by the definition of p_c. During this time period, $|X_t|/n$ should stay above ϵ_2, except for the final $O(\log n)$ time, because any low density state dies out rapidly due to Theorem 12-2. Thus, the process lingers at a *metastable* state for exponentially long time, until it manages to get through an exponentially small bottlenect to become extinct rapidly. Furthermore, the density of infection of the metastable state is larger than $\varepsilon_2 > 0$ for any $p > p_c$, which does not approach to zero as $p \searrow p_c$. On the other hand, for $p < p_c$, the process does not possess a nontrivial metastable state since the order of the survival time is strictly smaller than an exponential in n. See Section 1.6.2 for a formal description of the discontinuous phase transition, and Section 1.6.3 for its connection to the infinite Galton-Watson trees.

The following provide heuristic descriptions why we would expect such a discontinuous phase transition.

► *Mean-field calculation and metastability.* Consider the threshold-θ contact process on a $(\theta+2)$-regular tree, and suppose that we are interested in the flow of infection towards the root. Let v be a vertex other than the root and assume that at time t each child $u \sim v$ is infected with probability q independently of each other. Then, by only looking at the infections among its children, the probability that v is infected at time $t+1$ is at least

$$(\theta+1)pq^{\theta}(1-q) + pq^{\theta+1} = pq^{\theta}((\theta+1) - \theta q) =: f(q).$$

One can see that $f(q)$ is increasing in q, and that for p close enough to 1, there is a fixed point of $f(q_0) = q_0$ such that $q_0 < 1$. Thus, if we had $q > q_0$, the density of occupied sites is at least $f(q) > q_0$. This implies that the threshold-θ contact process on infinite $(\theta+2)$-regular tree should survive for large enough p and hence have a nontrivial stationary distribution. Since we assumed that μ is lower bounded by $\theta+2$, the local neighborhoods of $G_n \sim \mathcal{G}(n,\mu)$ roughly dominate $(\theta+2)$-regular trees, inferring the existence of metastability on G_n for large p.

► *Extinction of a low-density state.* Suppose that the local neighborhood $N(v, R+1)$ of $v \in G_n$ is a tree for some constant R. In such a case, if only the sites in $N(v,R)$ are occupied initially, then $X_t(u) = 0$ for all t and $u \notin N(v,R)$. This shows that the infections starting from a 'bounded tree' in G_n cannot percolate, and hence they may become extinct rapidly.

Recall that G_n is *locally treelike*, and the limiting local structure around a randomly selected vertex is a Galton-Watson tree. Moreover, if we choose $R = c\log n$ for a small constant $c > 0$, then the neighborhoods $N(v,R)$ for $v \in G_n$ typically look like trees with size at most $\bar{d}^R$, for some large constant $\bar{d}$ (greater than the mean of the size-biased distribution of μ). Indeed, we can show that for all $v \in G_n$, $N(v,R)$ can contain at most one cycle whp: see Lemma 2.3.5.

Suppose that a low density state is generated by i.i.d. Ber(ε) at each site. If ε is small

enough so that $(\varepsilon\bar{d})^R = o(n^{-1})$, the occupied sites at time 1 in $N(v, R)$ are likely to be surrounded by vacant sites for all v, by a union bound. Since $N(v, R)$ is a tree or contains just a single cycle, the infections inside it would die out rapidly from the aforementioned intuition.

This heuristics gives an idea why a "well-spread" infections of a low density must go through a rapid extinction. It infers Theorem 12-3, since if the probability parameter satisfies $p \le \varepsilon$, then the state X_1 at time 1 is stochastically dominated by i.i.d. $\mathrm{Ber}(\varepsilon)$ regardless of the initial state X_0. The role of Theorem 12-2 is to show that such a phenomenon happens for *any* choice of εn initial infections in G_n.

Fontes and Schonmann [FS08] studied the continuous-time threshold-$\theta \ge 2$ contact process on infinite d-regular trees. They showed that for large d, the process exhibits a discontinuous phase transition by a mean-field analysis. Later, Chatterjee and Durrett [CD13] extended the result to all $d \ge \theta + 2$ for the discrete-time model, as a result of establishing the same type of conclusion on random d-regular graphs. The main contribution of this work is to demonstrate an analogous phenomenon on general random graphs. For a more detailed review on related subjects, we refer the reader to [CD13] which has a nice summary on the literature.

On the other hand, [CD13] conjectured that the threshold-θ contact process should die out on random $(\theta+1)$-regular graphs, implying that the condition $D \ge \theta+2$ in Theorem 12 is not just a technical device. We verify this conjecture by establishing a polynomial upper bound on the survival time.

Theorem 13. *Let $p \in (0, 1)$ be a given number and let $\theta \ge 2$ be an integer. There exists a constant C_p such that the following hold true, whp over the choice of $G_n \sim \mathcal{G}(n, \theta + 1)$:*

- *Let (X_t) denote the threshold-θ contact process on G_n with parameter p. We have*

$$\lim_{n\to\infty} \mathbb{P}_{\mathrm{CP}}\big(X_{n^{C_p}} = \mathbf{0} \,\big|\, X_0 = \mathbf{1}_{G_n}, G_n\big) = 1.$$

In order to prove Theorem 13, consider a cycle C inside $G_n \sim \mathcal{G}(n, \theta + 1)$. Note that

$$\text{if } X_t(C) = \mathbf{0}, \text{ then } X_s(C) = \mathbf{0} \text{ for all } s \ge t,$$

due to the definition of the threshold-θ contact process. Further, at each time step, the probability of the entire cycle C becoming healthy is at least $(1-p)^{|C|}$. Therefore, we show that the graph G_n can be covered with cycles of length $O(\log n)$ and then convert each of them to be all-healthy in time $(1-p)^{-O(\log n)}$, thus obtaining that the process can survive at most polynomially long time. Details are discussed in Section 4.4.

1.6.1 Results for more general random graphs

It turns out that the condition $D \geq \theta + 2$ can easily be removed if we know the existence of the $(\theta+2)$-*core* of a macroscopic size inside G_n, which is defined to be the largest induced subgraph with minimal degree $\theta + 2$. Kim [Kim06] introduced a robust technique that finds cores inside Erdős-Rényi random graphs, which was later generalized to the case of $G_n \sim \mathcal{G}(n, \mu)$ in [JL07].

Let $D \sim \mu$, $h \in (0,1)$ and $D_h = \mathbf{Bin}(D, h)$, where **Bin** is a shorthand for a binomial random variable. That is, for each $r \in \mathbb{N}$,

$$\mathbb{P}(D_h = l) = \sum_{j \geq l} \mathbb{P}(\mathbf{Bin}(j,h) = l)\, \mathbb{P}(D = j).$$

Proposition 1.6.1 ([Kim06, JL07])**.** *Let r be a positive integer and μ be a degree distribution with $d = \mathbb{E}_{D\sim\mu} D$. Define the functions $F_r = F_r(h)$ and $\rho_r = \rho_r(h)$ by*

$$\begin{aligned} F_r(h) &= \mathbb{E}\left[D_h \mathbf{1}_{\{D_h \geq r\}}\right] = \sum_{l \geq r} \sum_{j \geq l} l\, \mathbb{P}(\mathbf{Bin}(j,h) = l)\, \mathbb{P}(D = j); \\ \rho_r(h) &= \mathbb{P}(D_h \geq r) = \sum_{l \geq r} \sum_{j \geq l} \mathbb{P}(\mathbf{Bin}(j,h) = l)\, \mathbb{P}(D = j). \end{aligned} \tag{1.6.3}$$

If there exists $h \in (0,1)$ such that

$$dh^2 < F_r(h), \tag{1.6.4}$$

then whp *over the choice of G_n, there exists the r-core K_n inside $G_n \sim \mathcal{G}(n,\mu)$. The size of K_n tends to $|K_n|/n \to \rho_r(\hat{h})$ in probability, where $\hat{h}$ is the largest $h \leq 1$ such that $dh^2 = F_r(h)$.*

Remark 1.6.2. For Erdős-Rényi random graphs $\mathcal{G}^{\text{ER}}(n, \frac{d}{n})$, the condition (1.6.4) is equiv-

alent to

$$d > d_r := \left\{ \frac{\alpha}{\mathbb{P}(\mathrm{Pois}(\alpha) \geq r-1)} \,:\, \alpha > 0 \right\}. \tag{1.6.5}$$

Furthermore, Theorem 12 assumed μ to have finite k-th moments for all k. Our method generalizes to μ with a finite K-th moment for some large K. To introduce this extension, we first define $\mathcal{M}(k, M)$ to be the collection of probability distributions μ on $\mathbb{N}$ that satisfies $\mathbb{E}_{D\sim\mu} D^k \leq M$.

Theorem 14. *Let $\theta \geq 2$ be an integer, set $k_1 = 6$ and let $M > 0$ be a fixed constant. There exist a collection of graphs $\mathcal{A}_n$ for each n, and a constant $K = K(k_1, M)$ such that for all $\mu \in \mathcal{M}(k_1, M)$ satisfying $\mathbb{E}_{D\sim\mu} D^K < \infty$ and (1.6.4) with $r = \theta + 2$, the following hold true:*

1. $\lim_{n\to\infty} \mathbb{P}_{G_n\sim\mathcal{G}(n,\mu)}(G_n \in \mathcal{A}_n) = 1$. *In the following items, let $(X_t) \sim \mathsf{TCP}^{\theta,p}(G_n; X_0)$.*

2. *There exist $p_1 \in (0,1)$ and $\varepsilon_1, \gamma_1 > 0$ such that for any $p \in [p_1, 1]$ and $G_n \in \mathcal{A}_n$,*

$$\lim_{n\to\infty} \mathbb{P}_{\mathrm{CP}}\Big(|X_t| \geq (1-\varepsilon_1)\rho_{\theta+2}(\hat{h})n,\ \forall t \in [0, e^{\gamma_1 n}] \,\Big|\, X_0 \equiv \mathbf{1},\, G_n \Big) = 1,$$

   where $\rho_{\theta+2}$ and $\hat{h}$ are as in Proposition 1.6.1. The constants $p_1, \varepsilon_1, \gamma_1$ depend on μ and θ.

3. *For any $p > 0$, there exist $\varepsilon_2, C_2 > 0$ such that for any $G_n \in \mathcal{A}_n$,*

$$\lim_{n\to\infty} \inf_{|X_0|\leq\varepsilon_2 n} \mathbb{P}_{\mathrm{CP}}\Big(X_{C_2 \log n} = \mathbf{0} \,\Big|\, X_0,\, G_n \Big) = 1.$$

   The constants ε_2, C_2 depend on p, μ and θ.

Remark 1.6.3. The notation $\mathcal{A}_n$ substitutes "whp over the choice of G_n," and it is introduced to give a clear description of a complicated statement.

Remark 1.6.4. Simulations suggest that the phase transition of the threshold-θ contact process could be continuous when μ is heavy-tailed. According to [VD13], $\mathsf{TCP}^{2,p}(G_n)$ on $G_n \sim \mathcal{G}(n, \mu)$ with

$$\mathbb{P}_{D\sim\mu}(D = k) \sim Ck^{-\alpha}, \quad 2 < \alpha \leq 3, \tag{1.6.6}$$

is expected to show a continuous phase transition. The belief is that the existence of many high-degree vertices prevents Theorem 14-3 from being true. However, if α in (1.6.6) is greater than 3, their simulation predicts a discontinuous phase transition. Unfortunately, as stated in Theorem 14, our methods can cover μ with a lighter tail satisfying the assumption in the statement.

As we have seen before, the Erdős-Rényi random graph $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$ is contiguous to $\mathcal{G}(n, \mu)$ with $\mu = \mathrm{Pois}(d)$ in the sense that for an event $\mathcal{E}_n$,

$$\lim_{n\to\infty} \mathbb{P}_{G_n \sim \mathcal{G}(n,\mathrm{Pois}(d))}(G_n \in \mathcal{E}_n) = 0 \quad \text{implies} \quad \lim_{n\to\infty} \mathbb{P}_{G_n \sim \mathcal{G}^{\mathrm{ER}}(n,\frac{d}{n})}(G_n \in \mathcal{E}_n) = 0.$$

Furthermore, Poisson distribution satisfies the assumption of Theorem 14, since all its moments are finite. Thus, for $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$, we have an analogue of Theorem 14 for $d > d_{\theta+2}$.

Corollary 15. *Let $\theta \geq 2$ be an integer and $d > d_{\theta+2}$ for $d_{\theta+2}$ defined as (1.6.5). There exists a collection $\mathcal{A}_n$ of graphs for each n, such that*

1. $\lim_{n\to\infty} \mathbb{P}_{G_n \sim \mathcal{G}^{\mathrm{ER}}(n,d/n)}(G_n \in \mathcal{A}_n) = 1$. *In the following items, let (X_t) denote the threshold-θ contact process on G_n with parameter p.*

2. *There exist $p_1 \in (0,1)$ and $\varepsilon_1, \gamma_1 > 0$ such that if $p \geq p_1$ then for every $G_n \in \mathcal{A}_n$,*

$$\lim_{n\to\infty} \mathbb{P}_{\mathrm{CP}}\Big(|X_t| \geq (1-\varepsilon_1)\rho_{\theta+2}(\hat{h})n,\ \forall t \in [0, e^{\gamma_1 n}] \,\Big|\, X_0 = \mathbf{1}_{G_n},\, G_n\Big) = 1,$$

   where $\rho_{\theta+2}$ and $\hat{h}$ are as in Proposition 1.6.1. The constants $p_1, \varepsilon_1, \gamma_1$ depend on d and θ.

3. *For any $p \in (0,1)$, there exist $\varepsilon_2, C_2 > 0$ such that for every $G_n \in \mathcal{A}_n$,*

$$\lim_{n\to\infty} \inf_{|X_0| \leq \varepsilon_2 n} \mathbb{P}_{\mathrm{CP}}\Big(X_{C_2 \log n} = \mathbf{0} \,\Big|\, X_0,\, G_n\Big) = 1.$$

   The constants ε_2, C_2 depend on p, d and θ.

1.6.2 Metastability and the discontinuous phase transition on $\mathcal{G}(n,\mu)$

In this subsection, we give a formal description on the metastable behavior and the discontinuous phase transition discussed in Theorem 12. Similarly as τ_0 in (1.6.1), we can define τ_ρ for $\rho \in [0,1)$ as

$$\tau_\rho(G_n,p) := \mathbb{E}_{\mathrm{CP}}\left[\inf\left\{t : \frac{|X_t|}{n} \le \rho\right\} \,\middle|\, X_0 = \mathbf{1}_{G_n},\, G_n\right].$$

Then, for each $p \in (p_c(\mu),1)$, the *lower* and *upper* metastabe densities can be defined as

$$\rho_-(\mu,p) := \lim_{\beta\to 0}\inf\left\{\rho : \lim_{n\to\infty}\mathbb{P}_{\mathcal{G}(n,\mu)}\big(\tau_0(G_n,p) - \tau_\rho(G_n,p) \ge e^{\beta n}\big) = 1\right\};$$
$$\rho_+(\mu,p) := \lim_{\beta\to 0}\sup\left\{\rho : \lim_{n\to\infty}\mathbb{P}_{\mathcal{G}(n,\mu)}\big(\tau_\rho(G_n,p) \ge e^{\beta n}\big) = 1\right\}.$$

In words, ρ_- is the smallest possible density of infection that can survive for an exponentially long time, and ρ_+ is the largest possible density that can be observed after an exponentially long time starting from $X_0 = \mathbf{1}_{G_n}$. Due to the definition of $p_c(\mu)$ in (1.6.2), note that we have $\rho_-(\mu,p) \le \rho_+(\mu,p)$. We also define $\rho_-(\mu,p) = \rho_+(\mu,p) = 0$ when $p \in [0,p_c(\mu))$. Recalling Theorem 12 and the discussions below it, we see that the functions $\rho_-(\mu,p), \rho_+(\mu,p)$ are both discontinuous at $p_c(\mu)$, when μ satisfies the assumption of Theorem 12 or 14. Thus, the phase transition from a short-surival (when $p < p_c(\mu)$) to a long-survival (when $p > p_c(\mu)$) is discontinuous, in terms of the infection density of metastable states.

We believe that $\rho_+(\mu,p) = \rho_-(\mu,p) =: \rho(\mu,p)$, and expect the dynamics of the supercritical threshold-$\theta \ge 2$ contact process on a typical $G_n \sim \mathcal{G}(n,\mu)$ to be described as follows:

1. Starting from an all-infected state, it reaches at the density $n^{-1}|X_t| = \rho(\mu,p) + o(1)$ in time $t \le n^{O(1)}$.

2. It stays at the same density level of $\rho(\mu,p) + o(1)$ for an expnentially long time until it reaches at density $\rho(\mu,p) + \Theta(1)$.

3. At density $\rho(\mu,p)+\Theta(1)$, it can either go back to the previous step of density $\rho(\mu,p)+o(1)$, or becomes extinct. Both transitions happen rapidly, in time $n^{O(1)}$.

The definition of $\rho_{\pm}(\mu, p)$ and the predicted dynamics illustrated here are motivated by the theory of metastable Markov processes. We refer to the survey [Bov06] and references therein for a further background on this topic.

1.6.3 Discontinuous phase transition on infinite Galton-Watson trees

A natural interest is to study the threshold-θ contact process on infinite Galton-Watson trees, which are local weak limits of random graphs (for details, see, e.g., [DM10, Section 2.1]). In fact, it is more straight-forward to consider the critical probability and the discontinuous phase transition on infinite graphs, in the following sense. Let $T \sim \mathsf{GW}(\mu')$ be the Galton-Watson tree with offspring distribution μ', conditioned on $|T| = \infty$. Consider the threshold-θ contact process starting from the all-infected state $\mathbf{1}_T$, and denote its law at time t by $\xi_t^{p,T}$. Since the process is a Markov chain, we can consider its equilibrium $\xi_\infty^{p,T} \stackrel{d}{=} \lim_{t\to\infty} \xi_t^{p,T}$, and define

$$p_c^{\mathrm{GW}}(T) := \sup\left\{p : \xi_\infty^{p,T} \equiv \mathbf{0}\right\}.$$

Several fundamental questions can arise from this formulation.

Question 1. Is there a global critical probability $p_c^{\mathrm{GW}}(\mu')$ such that $p_c^{\mathrm{GW}}(T) = p_c^{\mathrm{GW}}(\mu')$ for almost every $T \sim \mathsf{GW}(\mu')$ conditioned on $|T| = \infty$?

Question 2. Let μ' be a size-biased distribution of μ. Assuming that the previous question is true, is $p_c^{\mathrm{GW}}(\mu') = p_c(\mu)$ for $p_c(\mu)$ given in (1.6.2)?

(Note that in Question 2, the Galton-Watson tree with the size-biased distribution μ' is the limiting local structure of $\mathcal{G}(n, \mu)$: see Section 1.4.4.) We believe that the answer for both questions is yes. For the linear contact process, it is not difficult to see that the answer for the first question is affirmative, and Theorem 10 gives a partial answer to Quesion 2.

Furthermore, denoting the root of T by ρ, we can conceptually define the equilibrium density $\rho(p, T)$ as

$$\rho(p, T) := \lim_{R\to\infty} \frac{1}{|N(\rho, R)|} \sum_{v\in N(\rho,R)} \mathbb{P}_{X\sim\xi_\infty^{p,T}}(X(v) = 1).$$

Note that the definition cannot be as simple as [CD13, Section 1.2.2], since the tree is not homogeneous. Also, it is an increasing function in p due to monotonicity. We believe that the above limit is well-defined, and expect that there exists $\rho(p, \mu')$ such that $\rho(p, \mu') = \rho(p, T)$ for almost every $T \sim \mathbf{GW}(\mu')$ conditioned on $|T| = \infty$, similarly as in Question 1. Moreover, an interesting problem is to establish a discontinuous phase transition, extending Theorem 12 to the case of infinite Galton-Watson trees:

Question 3. Let μ' be an unbounded degree distribution that has finite k-th moments for all $k > 0$ and is bounded by $\theta + 1$ from below. Show that for almost every $T \sim \mathbf{GW}(\mu')$ conditioned on $|T| = \infty$, the equilibrium density $\rho(p, \mu') = \rho(p, T)$ is well-defined, and is discontinuous at $p_c^{\mathrm{GW}}(\mu')$.

However, for the threshold-θ contact process, there does not seem to be an obvious way to translate the results on random graphs (Theorems 12, 14) to those on infinite Galton-Watson trees, which was possible for random d-regular graphs and the infinite d-regular tree [CD13]. In [CD13], they used the fact that the depth-$\log\log n$ local neighborhoods of $\mathcal{G}(n, d)$ look the same as a d-regular tree for $n - o(n)$ vertices. Nevertheless, in $\mathcal{G}(n, \mu)$, only a very small fraction of vertices has a local neighborhood structure that looks like a *fixed* Galton-Watson tree.

Note that if μ is bounded, say, by M, we know the answer by comparing the Galton-Watson tree with the infinite M-regular tree [CD13].

1.6.4 Extinction on the infinite $(\theta + 1)$-regular tree

Although we proved in Theorem 13 that the threshold-θ contact process on $G_n \sim \mathcal{G}(n, \theta+1)$ does not have a metastable state, it is not clear if the process on the infinite $(\theta + 1)$-regular tree always dies out. The proof for Theorem 13 cannot be applied, since there are no cycles in the infinite tree, which played a huge role in $\mathcal{G}(n, \theta + 1)$.

Question 4. Does the threshold-θ contact process on the infinite $(\theta + 1)$-regular tree with all-infected initial condition die out in finite time?

1.7 Organization

In Chapter 2, we study the phase transition of the contact process on random graphs, establishing Theorems 1, 4, 5, and Corollary 6.

Our next goal in Chapter 3 is to understand the asymptotics of the critical value. Here, we demonstrate Theorems 3, 7, 10, 11, and Corollary 8.

The final chapter, Chapter 4, is devoted to the threshold contact process, and we verify Theorems 12, 13, 14, and Corollary 15.

Chapter 2

Phase transition of the contact process on random graphs

We begin our study of the contact process, and prove some of the results discussed in Section 1.5; namely, Theorems 1, 4, 5, and Corollary 6. The logic and technical details developed in this chapter are all based on the work [BNNS21] by Bhamidi, Nguyen, Sly, and the author. Before delving into the actual proof, we first outline the main ideas and techniques used in this chapter.

2.1 Overview of the chapter

The analysis of the subcritical contact process (i.e., extinction and short survival) relies on three main ideas which we now describe. Here, we assume that the offspring distribution of the Galton-Watson trees and the degree distribution of the configuration model have exponential tails.

► *Modified process: preventing recoveries at the root.* One main difficulty in studying the contact process on Galton-Watson trees comes from complicated dependencies inside the given tree. To overcome this obstacle, we consider the following modification of the process:

- A vertex is added above the root that is always infected. As such the chain no longer has an absorbing state.

- Recoveries at the root only occur when none of its descendants are infected at the time of recovery. All the other infections and recoveries are the same as the original process.

In the modified process, when the root is infected, the processes inside each subtree from a child of the root behave independently. By relating the stationary probability of the root being healthy to the extinction time we develop a recursive relationship over the tree height. As a result, we show that the expected survival time of the contact process with small enough λ is bounded above by a fixed constant, for any finite-depth Galton-Watson trees.

▶ *Exponential decay of infection depth: the delayed process.* To relate the finite Galton-Watson trees to the infinite tree, we prove that the probability that the infection goes deeper than depth h decays exponentially in h. To this end, we introduce the *delayed process*, which spends exponentially longer time at states containing deeper infections. Based on a similar argument introduced above, we show that the expected survival time of the delayed process on the Galton-Watson tree is bounded by a constant if λ is small enough. This will imply that in the original process, the infection can go deeper than h at most with an exponentially small probability in h. Thus, the contact process on the infinite Galton-Watson tree can be regarded as that on a large-depth finite tree, which will be useful in studying the contact process on $\mathcal{G}(n, \mu)$.

▶ *Coupling the local neighborhoods of $\mathcal{G}(n, \mu)$.* To study the contact process on $G \sim \mathcal{G}(n, \mu)$ exhibiting short survival, we attempt to dominate the local neighborhoods of the graph by Galton-Watson trees, in terms of isomorphic inclusions of graphs. However, some of the local neighborhoods $N(v, r)$ will contain a cycle, and hence we introduce modified Galton-Watson type processes that contain a cycle and behave similarly as the Galton-Watson trees. After dominating the local neighborhoods of G by the new branching processes, we study the contact process on the latter graphs and bound its survival time based on the aforementioned ideas.

On the other hand, when studying the long survival for $G \sim \mathcal{G}(n, \mu)$, we rely on the existence of what we call *embedded expanders* inside the graph. Roughly speaking, we call

a subset W of vertices in G an *embedded expander*, if (i) all vertices in W have high degree, say, at least M, and (ii) distance R-neighborhood of every subset $W' \subset W$ of at most a certain size intersects with W at more than $2|W'|$ vertices (See a precise definition in Lemma 2.4.2). As noted above, we expect an infection at a degree M vertex to last for at least time exponential in M.

Intuitively, if a subset W' of an embedded expander W is infected, then the infections inside W' would happen repeatedly for a reasonably long time due to its large degrees, and hence it will not die out whp before infecting its neighbors within distance R. Thus, if W is an embedded expander, then the infection originally on W' is likely to spread over $2|W'|$ vertices after some time. In Sections 2.4 and 2.5 we make this intuition rigorous and prove the existence of an embedded expander inside G. For the latter argument, we partially rely on the *Cut-off line algorithm* (Definition 2.6.3) which was introduced in [Kim06] to find the cores of random graphs.

For Theorem 5, we show that if μ is subexponential, then we can find an embedded expander in G such that R is arbitrarily smaller than M. Therefore, even if λ is very small, it will be possible for infections in the embedded expander to travel the distance of R before dying out.

2.1.1 Organization

The rest of the chapter is organized as follows. We prove Theorem 1, Theorem 4-(1), Theorem 4-(2) in Sections 2.2, 2.3, and 2.4 respectively. In Section 2.5, we prove Theorems 5. In Section 2.6, we prove a structural lemma of random graphs which plays a crucial role in establishing Theorems 4-(2) and 5. Finally, Section 2.7 is devoted to demonstrating some of the technical lemmas developed in the chapter.

2.2 Extinction in Galton-Watson trees

Let ξ be a random variable on $\mathbb{N}$ having an exponential tail, namely, $\mathbb{E}\exp(c\xi) = M < \infty$ for some constants $c, M > 0$. Throughout this section we assume $\mathbb{E}\xi > 1$, which makes $\mathcal{T} \sim \mathsf{GW}(\xi)$, the Galton-Watson tree with offspring distribution ξ, survive forever with

positive probability. We also denote the depth-L Galton-Watson tree by $\mathcal{T}_{\leq L} \sim \mathsf{GW}(\xi)^L$, and its root denoted by ρ.

The goal of this section is to establish Theorem 1. To this end, we prove the following in the next two sections:

- We first show that for small enough λ, the expected survival time of $\mathsf{CP}^\lambda(\mathcal{T}_{\leq L}; \mathbf{1}_\rho)$ is bounded by a constant uniform in L.
- Then, we prove that for small enough λ, the probability that the infection in the process goes deeper than h decays exponentially in h.

At the end, we will combine the two to see that the death-survival threshold λ_1 of the infinite Galton-Watson tree is strictly positive. Moreover, both properties will be essential in Section 2.3.

2.2.1 Expected survival time in finite trees

In this section, we prove the following theorem:

Theorem 2.2.1. *Let L be an arbitrary integer and $\xi, \mathcal{T}_{\leq L}$ be defined as above. Let $\boldsymbol{R}_L$ be the first time when $\mathsf{CP}^\lambda(\mathcal{T}_{\leq L}; \mathbf{1}_\rho)$ reaches state $\mathbf{0}$. Then there exist constants $C, \lambda_0 > 0$ depending only on ξ such that for any $\lambda \leq \lambda_0$ and L, we have $\mathbb{E}\boldsymbol{R}_L \leq C$.*

As mentioned in Section 1.3, $\mathbb{E}\mathbf{R}_L = \mathbb{E}_{\mathrm{GW}}[\mathbb{E}_{\mathrm{CP}}[\mathbf{R}_L]]$.

Let $D \sim \xi$ denote the degree of the root ρ, and $v_1, \ldots, v_D$ be the children of ρ. Further, let $\mathcal{T}_{v_i}$ denote the subtree of $\mathcal{T}$ rooted at v_i. To establish Theorem 2.2.1, our attempt is to study the effect of joining the subtrees $\mathcal{T}_{v_i}$ together at ρ, and hence expressing $\mathbf{R}_L$ in terms of $\mathbf{R}_{L-1}$. The main difficulty of this approach comes from the fact that the contact process on $\mathcal{T}_{\leq L}$ does not behave independently on each subtree $\mathcal{T}_{v_i}$. To overcome this obstacle, we study the contact process in a slightly different setting, by adding a parent ρ^+ above the root ρ which is infected permanently.

Definition 2.2.2 (Root-added contact process)**.** Let T be a finite 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 $\mathbf{CP}^{\lambda}_{\rho^+}(T^+; x_0)$ 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.

By adding a permanently infected parent, we can take advantage of independence between different subtrees as well as the stationary distribution of the process, as briefly discussed in Section 2.1. In the following lemma, we formally introduce the "modified process" explained in Section 2.1 and construct a quantitative recursion argument in terms of the tree depth.

Lemma 2.2.3. *Let L be an arbitrary integer and $\xi, \mathcal{T}_{\leq L}$ be defined as above. Define $\boldsymbol{S}_L$ to be the first time when $\mathbf{CP}^{\lambda}_{\rho^+}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$ reaches state $\mathbf{0}$. Then there exists a constant $\lambda_0 > 0$ depending only on ξ such that for any $\lambda \leq \lambda_0$ and L, $\mathbb{E}\boldsymbol{S}_L \leq e$.*

Proof. We build an inductive argument in terms of L, by considering the modified contact process $(\widetilde{X}_t) \sim \widetilde{\mathbf{CP}}^{\lambda}_{\rho^+;\rho}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$ defined as follows:

- $(\widetilde{X}_t)$ is coupled with $(X_t) \sim \mathbf{CP}^{\lambda}_{\rho^+}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$ in the sense that they share the same locations of recovery and infection clocks. In particular, ρ^+ is permanently infected in $(\widetilde{X}_t)$.
- In $(\widetilde{X}_t)$, the recovery at ρ at time s is valid if and only if $\widetilde{X}_s = \mathbf{1}_\rho$. Otherwise, we ignore the recovery at ρ. In other words, when there exists an infected vertex other than ρ and ρ^+, the recovery at ρ is neglected.

Let $\widetilde{\mathbf{S}}_L$ be the first time when $\widetilde{X}_t$ reaches the all-healthy state $\mathbf{0}$. Then Lemma 1.1.2 tells us that $\mathbf{S}_L \leq \widetilde{\mathbf{S}}_L$. Assume that we started running $(\widetilde{X}_t)$ from $t = 0$. Then there are two possibilities for the transition to the second state from the initial state $\mathbf{1}_\rho$:

A. ρ becomes healthy;

B. ρ infects one of its children, say, v_i.

When $\mathbf{A}$ happens, then $\widetilde{\mathbf{S}}_L$ is just the time elapsed until encountering $\mathbf{A}$. If we let D denote the number of children of ρ, then probability of the event $\mathbf{A}$ is $\frac{1}{1+\lambda D}$ and also

$$\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}_L|\mathbf{A}] = \frac{1}{1+\lambda D}.$$

On the other hand, when $\mathbf{B}$ happens, then the recoveries at ρ are neglected until all of its descendants are healthy. Therefore, the infection and recovery occurring inside the subtrees $\mathcal{T}_{v_i} \cup \{\rho\}$ become independent of each other until all of them become completely healthy at the same time. Hence, after the occurrence of $\mathbf{B}$, where we have $\mathbf{1}_{\{\rho,v_i\}}$ as its new initial state, $(\widetilde{X}_t)$ can be viewed as the product chain $\left(X_t^{\otimes}\right)$ of root-added contact processes defined as follows.

$$\left(X_t^{\otimes}\right) \;\sim\; \mathsf{CP}_{\rho}^{\otimes}(\mathcal{T}_{\le L}; \mathbf{1}_{v_i}) := \left(\otimes_{\substack{j=1 \\ j\neq i}}^{D} \mathsf{CP}_{\rho}^{\lambda}(\mathcal{T}_{v_j}^{+}; \mathbf{0}) \right) \otimes \mathsf{CP}_{\rho}^{\lambda}(\mathcal{T}_{v_i}^{+}; \mathbf{1}_{v_i}).$$

(Here for each $\mathcal{T}_{v_j}$, we view ρ as its permanently infected "added" parent of the root v_i.) Note that this perspective is valid until $\widetilde{X}_t$ returns back to $\mathbf{1}_{\rho}$.

Let $\widetilde{\mathbf{S}}_i^{\otimes}$ denote the time that the above product chain on $\cup_{j=1}^{D}\mathcal{T}_{v_j}$ started from the state $\mathbf{1}_{v_i}$ reaches the all-healthy state $\mathbf{0}$. At time $s = \widetilde{\mathbf{S}}_i^{\otimes}$, $\widetilde{X}_s$ is again in the state $\mathbf{1}_{\rho}$, hence it again meets with either $\mathbf{A}$ or $\mathbf{B}$ in the next step. Note that in this situation the expected waiting time to encounter either event is $\frac{1}{1+\lambda D}$. Also, define $\widetilde{\mathbf{S}}^{\otimes}$ to be the average of $\widetilde{\mathbf{S}}_i^{\otimes}$ over all i, recalling that when event $\mathbf{B}$ occurs, each child v_i is infected with equal probability. Then, if we continue this procedure until $\widetilde{X}_t$ reaches $\mathbf{0}$, we get

$$\begin{aligned} \mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}_L \mid \mathcal{T}_{\le L}] = \sum_{k=0}^{\infty} \left(\frac{\lambda D}{1+\lambda D}\right)^k \frac{1}{1+\lambda D} \\ \times \left[(k+1)\frac{1}{1+\lambda D} + k\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\otimes} \;\middle|\; \{\mathcal{T}_{v_i} : i \in [D]\} \right] \right]. \end{aligned} \tag{2.2.1}$$

Simplifying the sum then gives

$$\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}_L \mid \mathcal{T}_{\le L}] = 1 + \lambda D \mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\otimes} \;\middle|\; \{\mathcal{T}_{v_i} : i \in [D]\} \right],$$

which implies

$$\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}_L \,|\, D] = 1 + \lambda D \,\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\otimes} | D]. \tag{2.2.2}$$

The next step of the proof is to estimate $\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\otimes}|D]$ by relating it to the stationary distributions of the root-added contact processes. Let $\pi^{(D)}$ be the stationary distribution of the product chain $\mathsf{CP}_{\rho}^{\otimes}(\mathcal{T}_{\le L})$ (when defining $\pi^{(D)}$, note that the initial state of the process is irrelevant). We also let π_i be the stationary distribution of $\mathsf{CP}_{\rho}^{\lambda}(\mathcal{T}_i^{+})$. Then we have

$$\pi^{(D)} = \otimes_{i=1}^{D} \pi_i.$$

For any state x on $\mathcal{T}_{\le L} \setminus \{\rho\}$, $\pi^{(D)}(x)$ is proportional to the expected time that the chain $(X_t^{\otimes}) \sim \mathsf{CP}_{\rho}^{\otimes}(\mathcal{T}_{\le L})$ stays at state x. Moreover, the expected time for the chain to stay at $\mathbf{0}$ is $(\lambda D)^{-1}$, and after escaping from $\mathbf{0}$, it spends time $\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\otimes}|\mathcal{T}_{\le L}]$ in expectation before returning back to $\mathbf{0}$. Also note that the ratio between them is precisely described by the stationary mass $\pi^{(D)}(\mathbf{0})$ (see Section 2.2.2 (Lemma 2.5 in particular) of [AF02] for a formal discussion). Therefore,

$$\pi^{(D)}(\mathbf{0}) = \frac{(\lambda D)^{-1}}{(\lambda D)^{-1} + \mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\otimes} \,|\, \mathcal{T}_{\le L}]} = \frac{1}{1 + \lambda D \,\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\otimes} \,|\, \mathcal{T}_{\le L}]}. \tag{2.2.3}$$

Similarly, we have

$$\pi_i(\mathbf{0}) = \frac{1}{1 + \lambda \mathbb{E}_{\mathrm{CP}}[\mathbf{S}_{L-1} \,|\, \mathcal{T}_{v_i}]}, \tag{2.2.4}$$

where $\mathbf{S}_{L-1}$ is the first time when $(X_t^i) \sim \mathsf{CP}_{\rho}^{\lambda}(\mathcal{T}_{v_i}^{+}; \mathbf{1}_{v_i})$ reaches state $\mathbf{0}$. Here, note that $\mathbf{S}_{L-1}$ matches with the definition from the statement of this lemma since $\mathcal{T}_{v_i} \sim \mathsf{GW}(\xi)^{L-1}$. Therefore, we obtain that

$$1 + \lambda D \,\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\otimes} \,\middle|\, \mathcal{T}_{\le L}\right] = \prod_{i=1}^{D} (1 + \lambda \mathbb{E}_{\mathrm{CP}}[\mathbf{S}_{L-1} | \mathcal{T}_{v_i}]). \tag{2.2.5}$$

Since $\{\mathcal{T}_{v_i}\}_{i \ge 1}$ are i.i.d. $\mathsf{GW}(\xi)^{L-1}$ for all i, integrating (2.2.5) over the randomness of $\{\mathcal{T}_{v_i} : i \in [D]\}$ tells us that

$$1 + \lambda D \mathbb{E}\left[\widetilde{\mathbf{S}}^{\otimes} \,\middle|\, D\right] = (1 + \lambda \mathbb{E}[\mathbf{S}_{L-1}])^{D} \le \exp\{(\lambda \mathbb{E}[\mathbf{S}_{L-1}]) D\}.$$

Combining this with (2.2.2), we get

$$\mathbb{E}\left[\widetilde{\mathbf{S}}_L \middle| D\right] \leq \exp\left\{(\lambda \mathbb{E}[\mathbf{S}_{L-1}])D\right\}. \tag{2.2.6}$$

In the last step of the proof, we complete the inductive argument using the fact that $D \sim \xi$ has an exponential tail. Let us set $c, M > 0$ to be the constants satisfying $\mathbb{E}\exp(cD) = M$. When $L = 0$, we trivially have that $\mathbb{E}S_0 = 1$. Define K and λ_0 as

$$K = e \cdot \max\{\log M,\, 1\}, \quad \lambda_0 = \frac{c}{K}.$$

Suppose that $\mathbb{E}[\mathbf{S}_{L-1}] \leq e$. Then for any $\lambda \leq \lambda_0$, we have

$$\begin{aligned}\mathbb{E}\mathbf{S}_L \leq \mathbb{E}\widetilde{\mathbf{S}}_L &\leq \mathbb{E}_{D\sim\xi}\left[\exp(\lambda \mathbb{E}[\mathbf{S}_{L-1}]D)\right] = \mathbb{E}_{D\sim\xi}\left[\exp\left(\frac{\lambda\mathbb{E}[\mathbf{S}_{L-1}]}{c}\cdot cD\right)\right] \\ &\leq \exp\left\{\log M \frac{\lambda\mathbb{E}[\mathbf{S}_{L-1}]}{c}\right\} \leq e,\end{aligned}$$

where we used Jensen's inequality to deduce the first inequality in the second line. Finally, an elementary induction argument implies the desired result. □

Proof of Theorem 2.2.1. For $\mathbf{R}_L, \mathbf{S}_L$ defined as in the statement of Theorem 2.2.1 and Lemma 2.2.3, respectively, we have $\mathbb{E}\mathbf{R}_L \leq \mathbb{E}\mathbf{S}_L$ due to Lemma 1.1.2. Therefore, setting λ_0 as in the proof of Lemma 2.2.3 and $C = e$, we obtain $\mathbb{E}\mathbf{R}_L \leq C$ for all $\lambda \leq \lambda_0$ and L. □

2.2.2 Exponential decay of the infection depth

In this section, we show that the maximal depth that the infection can reach before dying out decays exponentially.

For any integer L, let $\mathcal{T}_{\leq L} \sim \mathsf{GW}(\xi)^L$ and $\mathcal{T}^+_{\leq L}$ be the graph obtained by adding a new parent root ρ^+ above ρ in $\mathcal{T}_{\leq L}$ as before. For each state $x \in \{0,1\}^{\mathcal{T}_{\leq L}}$, define the depth of x in $\mathcal{T}^+_{\leq L}$ to be

$$r(x) = r(x; \mathcal{T}^+_{\leq L}) = \max\{d(\rho^+, v) : x(v) = 1\}.$$

For $x = \mathbf{0}$, we set $r(\mathbf{0}) = 0$. Consider the root-added process $(X_t) \sim \mathsf{CP}^\lambda_{\rho^+}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$

(Definition 2.2.2), and let $\mathbf{S}_L$ be the first time then the process reaches the state $\mathbf{0}$. Let $H = \max\{r(X_t) : t \in [0, \mathbf{S}_L]\}$ be the maximal depth that the process reaches during an excursion from $\mathbf{0}$. Our goal in this section is to establish the following theorem and conclude the proof of Theorem 1.

Theorem 2.2.4. *Let $L > 0$ be any integer and let $\mathcal{T}_{\leq L}$, $\boldsymbol{S}_L$ and H be as above. There exist constants $K, \lambda_0 > 0$ depending only on ξ such that for all $\lambda \leq \lambda_0$, $h > 0$ and $m > 0$, we have*

$$\mathbb{P}_{\mathrm{CP}}(H > h \,|\, \mathcal{T}_{\leq L}) \leq 2m(K\lambda)^h,$$

with probability at least $1 - m^{-1}$ over the choice of $\mathcal{T}_{\leq L}$.

In order to control the deepest depth of infection, we introduce the *delayed* contact process.

Definition 2.2.5 (Delayed contact process)**.** Let $\mathcal{S}^+$ be a graph rooted at ρ^+ and $\mathcal{S} = \mathcal{S}^+ \setminus \{\rho^+\}$. For any two states $x, y \in \{0,1\}^{\mathcal{S}}$, let Q_{xy} be the rate of transition from x to y in the contact process $\mathsf{CP}^{\lambda}_{\rho^+}(\mathcal{S}^+)$. For a fixed constant $\theta \in (0,1)$, the *delayed contact process*, denoted by $\mathsf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{S}^+; x_0)$, is the continuous-time Markov chain on $\{0,1\}^{\mathcal{S}}$ with initial state x_0 and transition rate

$$Q^{(\theta)}_{xy} = \theta^{r(x;\mathcal{S}^+)} Q_{xy} = \theta^{r(x)} Q_{xy}.$$

According to the definition, in the delayed contact process, we spend exponentially longer time in the states with deeper depths. Let $\pi_{\mathcal{S}}, \nu^{\theta}_{\mathcal{S}}$ denote the stationary distributions of $\mathsf{CP}^{\lambda}_{\rho^+}(\mathcal{S}^+)$ and $\mathsf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{S}^+)$, respectively. Then,

$$\nu^{\theta}_{\mathcal{S}}(x) = \frac{\theta^{-r(x)}\pi_{\mathcal{S}}(x)}{\sum_y \theta^{-r(y)}\pi_{\mathcal{S}}(y)}, \tag{2.2.7}$$

where the summation is over all possible states $y \in \{0,1\}^{\mathcal{S}}$.

Suppose we have a lower bound on $\nu^{\theta}_{\mathcal{S}}(\mathbf{0})$. Then this implies an upper bound on $\pi_{\mathcal{S}}(x)$, by

$$\pi_{\mathcal{S}}(x) = \frac{\theta^{r(x)}\nu^{\theta}_{\mathcal{S}}(x)}{\sum_y \theta^{r(y)}\nu^{\theta}_{\mathcal{S}}(y)} \leq \frac{\theta^{r(x)}\nu^{\theta}_{\mathcal{S}}(x)}{\nu^{\theta}_{\mathcal{S}}(\mathbf{0})},$$

which intuitively infers that it is (exponentially) unlikely to see states of having very deep infections until the process comes back to $\mathbf{0}$. Based on this intuition, we establish the following proposition.

Proposition 2.2.6. *Let $L > 0$ be any integer and $\mathcal{T}_{\leq L} \sim \mathsf{GW}(\xi)^L$. Set $\nu^{\theta}_{\mathcal{T}_{\leq L}}$ to denote the stationary distribution of $\mathsf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{T}^+_{\leq L})$ on the space $\{0,1\}^{\mathcal{T}_{\leq L}}$. Then there exist constants $K, \lambda_0 > 0$ depending only on ξ such that for all $\lambda \leq \lambda_0$ and L,*

$$\mathbb{E}\left[\nu^{\theta}_{\mathcal{T}_{\leq L}}(\mathbf{0})^{-1}\right] \leq 2,$$

where θ is given by $\theta = K\lambda$.

Proof. Let $\mathbf{S}^{\theta}_L$ be the first time when $(X_t) \sim \mathsf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$ reaches state $\mathbf{0}$. We first derive an analog of (2.2.2) based on the methods from Lemma 2.2.3. To this end, define $(\widetilde{X}_t) \sim \widetilde{\mathsf{DP}}^{\lambda,\theta}_{\rho^+;\rho}(\mathcal{T}^+_{\leq L}; \mathbf{1}_\rho)$ to be the modification of (X_t) in such a way that

1. $(\widetilde{X}_t)$ shares the same infection and recovery clocks as (X_t).
2. In $(\widetilde{X}_t)$, recovery attempt at ρ is ignored if there exists an infected vertex other than ρ^+ and ρ at that moment.

Let $D \sim \xi$ denote the number of children of ρ, and let $\{\mathcal{T}_{u_i} : i = 1, \ldots, D\}$ be the subtrees from the children $u_1, \ldots, u_D$ of ρ. Then, when there is an infected vertex other than ρ^+ and ρ in $\widetilde{X}_t$, it can be regarded as the slowed-down version of process $\mathsf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}_{\leq L})$, where it spends longer time by the factor of θ^{-1} at each state, since the tree $\mathcal{T}_{\leq L}$ is one depth lower than $\mathcal{T}^+_{\leq L}$. Note that in $\mathsf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}_{\leq L})$, ρ is the permanently infected parent that has D children.

Let $\widetilde{\mathbf{S}}^{\theta}_L$ be the first time when $\widetilde{X}_t$ becomes $\mathbf{0}$. Also, let $\widetilde{\mathbf{S}}^{\theta}_i$ be the first time when $\mathsf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}_{\leq L}; \mathbf{1}_{u_i})$ is $\mathbf{0}$, and set $\widetilde{\mathbf{S}}^{\theta}$ to be the average of $\widetilde{\mathbf{S}}^{\theta}_i$ over $i = 1, \ldots, D$. Then, we can apply the same argument as Lemma 2.2.3 to this setting and deduce that

$$\begin{aligned}\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta}_L \,\middle|\, \mathcal{T}_{\leq L}\right] &= \sum_{k=0}^{\infty}\left(\frac{\lambda D}{1+\lambda D}\right)^k \frac{1}{1+\lambda D}\left[\frac{k+1}{\theta(1+\lambda D)} + \frac{k}{\theta}\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta} \,\middle|\, \mathcal{T}_{\leq L}\right]\right] \\ &= \frac{1}{\theta}\left(1 + \lambda D \mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta} \,\middle|\, \mathcal{T}_{\leq L}\right]\right).\end{aligned} \tag{2.2.8}$$

Now we relate these equations with the stationary distributions. Let $\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}, \nu^{\theta}_{\mathcal{T}_{u_i}}$ be the stationary distributions of $\mathbf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}_{\le L})$, $\mathbf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}^{+}_{u_i})$, respectively. Further, define $\nu^{\otimes}_{\mathcal{T}_{\le L}} = \otimes_{i=1}^{D}\nu^{\theta}_{\mathcal{T}_{u_i}}$. In contrast to what we had in Lemma 2.2.3, we do not necessarily have $\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}} = \nu^{\otimes}_{\mathcal{T}_{\le L}}$.

For each state $x \in \Omega_L := \{0,1\}^{\cup_{i=1}^{D}\mathcal{T}_{u_i}}$ of $\mathbf{DP}^{\lambda,\theta}_{\rho}(\mathcal{T}_{\le L})$, we decompose it into $x = (x_i)_{i=1}^{D}$, where $x_i \in \Omega_i := \{0,1\}^{\mathcal{T}_{u_i}}$. Setting $\pi_{\mathcal{T}_{u_i}}$ to be the stationary distribution of $\mathbf{CP}^{\lambda}_{\rho}(\mathcal{T}^{+}_{u_i})$ and $\pi^{\otimes}_{\mathcal{T}_{\le L}} := \otimes_{i=1}^{D}\pi_{\mathcal{T}_{u_i}}$, the equation (2.2.7) implies that

$$\begin{aligned}
\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(x) &= \frac{\theta^{-r(x;\mathcal{T}_{\le L})}\pi^{\otimes}_{\mathcal{T}_{\le L}}(x)}{\sum_{y\in\Omega_L}\theta^{-r(y;\mathcal{T}_{\le L})}\pi^{\otimes}_{\mathcal{T}_{\le L}}(y)} = \frac{\theta^{-r(x;\mathcal{T}_{\le L})}\prod_{i=1}^{D}\pi_{\mathcal{T}_{u_i}}(x_i)}{\sum_{y\in\Omega_L}\theta^{-r(y;\mathcal{T}_{\le L})}\prod_{i=1}^{D}\pi_{\mathcal{T}_{u_i}}(y_i)};\\
\nu^{\otimes}_{\mathcal{T}_{\le L}}(x) &= \prod_{i=1}^{D}\left[\frac{\theta^{-r(x_i;\mathcal{T}_{u_i})}\pi_{\mathcal{T}_{u_i}}(x_i)}{\sum_{y_i\in\Omega_i}\theta^{-r(y_i;\mathcal{T}_{u_i})}\pi_{\mathcal{T}_{u_i}}(y_i)}\right]\\
&= \frac{\theta^{-\sum_{i=1}^{D}r(x_i;\mathcal{T}_{u_i})}\prod_{i=1}^{D}\pi_{\mathcal{T}_{u_i}}(x_i)}{\sum_{y\in\Omega_L}\theta^{-\sum_{i=1}^{D}r(y_i;\mathcal{T}_{u_i})}\prod_{i=1}^{D}\pi_{\mathcal{T}_{u_i}}(y_i)}.
\end{aligned} \tag{2.2.9}$$

Notice that

$$r(x;\mathcal{T}_{\le L}) = \max\{r(x_i;\mathcal{T}_{u_i}) : i = 1,\dots,D\} \le \sum_{i=1}^{D} r(x_i;\mathcal{T}_{u_i}).$$

Therefore, deeper states tend to have larger weight in $\nu^{\otimes}_{\mathcal{T}_{\le L}}$ than in $\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}$, which implies that

$$\nu^{\otimes}_{\mathcal{T}_{\le L}}(\mathbf{0}) \le \widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(\mathbf{0}).$$

Moreover, we have the following equations as an analog of (2.2.3), (2.2.4):

$$\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(\mathbf{0}) = \frac{1}{1+\lambda D\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\theta}|\mathcal{T}_{\le L}]}, \quad \nu^{\otimes}_{\mathcal{T}_{\le L}}(\mathbf{0}) = \prod_{i=1}^{D}\left[\frac{1}{1+\lambda\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\theta}_{L-1}|\mathcal{T}_{u_i}\right]}\right]. \tag{2.2.10}$$

We combine our discussion with (2.2.8) to deduce that

$$\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\theta}_{L}\middle|\,\mathcal{T}_{\le L}\right] \le \mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta}_{L}\middle|\,\mathcal{T}_{\le L}\right] \le \frac{1}{\theta}\prod_{i=1}^{D}\left(1+\lambda\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\theta}_{L-1}\middle|\,\mathcal{T}_{u_i}\right]\right),$$

and hence

$$\mathbb{E}\left[\mathbf{S}_L^\theta\right] \le \frac{1}{\theta}\mathbb{E}_{D\sim\xi}\left[\exp\left(\lambda\mathbb{E}\left[\mathbf{S}_{L-1}^\theta\right]D\right)\right]. \tag{2.2.11}$$

The final step is to adjust the constants and deduce the conclusion. Let $c, M > 0$ be constants satisfying $\mathbb{E}_{D\sim\xi}\exp(cD) = M$. We set $K, \lambda_0 > 0$ and θ as

$$K = \max\left\{\frac{2\log M}{c\log 2}, 2\right\}, \quad \lambda_0 = \frac{1}{2K}, \quad \theta = K\lambda,$$

where $\lambda \in (0, \lambda_0]$ is arbitrary. For $L = 0$, we have $\mathbb{E}\mathbf{S}_0^\theta = \theta^{-1}$. Suppose that $\mathbb{E}\mathbf{S}_{L-1}^\theta \le 2/\theta$. Then, the right hand side of (2.2.11) can be bounded by

$$\begin{aligned}\frac{1}{\theta}\mathbb{E}_{D\sim\xi}\left[\exp\left(\lambda\mathbb{E}\left[\mathbf{S}_{L-1}^\theta\right]D\right)\right] &= \frac{1}{\theta}\mathbb{E}_{D\sim\xi}\left[\exp\left(\frac{\lambda\mathbb{E}\left[\mathbf{S}_{L-1}^\theta\right]}{c}\cdot cD\right)\right] \\ &\le \frac{1}{\theta}\exp\left\{\log M\left(\frac{\lambda}{c}\mathbb{E}\left[\mathbf{S}_{L-1}^\theta\right]\right)\right\} \le \frac{1}{\theta}M^{\frac{2}{cK}} \le \frac{2}{\theta},\end{aligned}$$

where the first inequality is due to Jensen's inequality. Therefore, for K, λ_0 as above, we have $\mathbb{E}\mathbf{S}_L^\theta \le 2/\theta$ for all $\lambda \le \lambda_0$ with $\theta = K\lambda$. Finally, note that $\nu^\theta_{\mathcal{T}_{\le L}}(\mathbf{0})$ is given by

$$\nu^\theta_{\mathcal{T}_{\le L}}(\mathbf{0}) = \frac{1}{1 + \lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}_L^\theta|\mathcal{T}_{\le L}]}.$$

Thus, we obtain the desired conclusion by taking expectation over its reciprocal and plugging in the estimate $\mathbb{E}\mathbf{S}_L^\theta \le 2/\theta$. □

Proof of Theorem 2.2.4. Let $L > 0$ be an arbitrary integer and let $\mathcal{T}_{\le L} \sim \mathsf{GW}(\xi)^L$. Also, let K, λ_0 be the constants given by Lemma 2.2.6, and let $\pi_{\mathcal{T}_{\le L}}$, $\nu^\theta_{\mathcal{T}_{\le L}}$ be the stationary distributions of $\mathsf{CP}^\lambda_{\rho^+}(\mathcal{T}^+_{\le L})$ and $\mathsf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{T}^+_{\le L})$, respectively, with $\theta = K\lambda$.

Set $\Omega = \{0,1\}^{\mathcal{T}_{\le L}}$, and define

$$A := \{x \in \Omega : r(x; \mathcal{T}^+_{\le L}) \ge h\}.$$

We first observe that

$$\frac{\pi_{\mathcal{T}_{\le L}}(A)}{\pi_{\mathcal{T}_{\le L}}(\mathbf{0})} = \frac{\sum_{x\in A}\theta^{r(x;\mathcal{T}^+_{\le L})}\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(x)}{\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(\mathbf{0})} \le \frac{\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(A)}{\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(\mathbf{0})}\theta^h.$$

Proposition 2.2.6 and Markov's inequality imply that with probability $1 - m^{-1}$ over the choice of $\mathcal{T}_{\le L}$, we have $\widetilde{\nu}^{\theta}_{\mathcal{T}_{\le L}}(\mathbf{0})^{-1} \le 2m$, and hence for such choices

$$\frac{\pi_{\mathcal{T}_{\le L}}(A)}{\pi_{\mathcal{T}_{\le L}}(\mathbf{0})} \le 2m(K\lambda)^h. \tag{2.2.12}$$

Moreover, if $(X_t) \sim \mathsf{CP}^{\lambda}_{\rho^+}(\mathcal{T}^+_{\le L})$ hits A, then the expected time needed for X_t to escape from A is at least 1. Indeed, it takes a unit expected time just to heal one infected site of depth at least h. In other words, if we set $\mathbf{S}_L, H$ as in the statement and define $\gamma_L(h) := |\{t \in [0, \mathbf{S}_L] : X_t \in A\}|$ where $|\cdot|$ denotes the Lebesgue measure, then

$$\mathbb{E}_{\mathrm{CP}}[\gamma_L(h) \,|\, H \ge h,\, \mathcal{T}_{\le L}] \ge 1.$$

Combining this with (2.2.12) tells us that

$$\begin{aligned}\mathbb{P}_{\mathrm{CP}}(H \ge h \,|\, \mathcal{T}_{\le L}) &\le \mathbb{E}_{\mathrm{CP}}[\gamma_L(h) \,|\, H \ge h, \mathcal{T}_{\le L}]\,\mathbb{P}_{\mathrm{CP}}(H \ge h \,|\, \mathcal{T}_{\le L}) \le \mathbb{E}_{\mathrm{CP}}[\gamma_L(h) \,|\, \mathcal{T}_{\le L}] \\ &\le \frac{\pi_{\mathcal{T}_{\le L}}(A)}{\pi_{\mathcal{T}_{\le L}}(\mathbf{0})} \le 2m(K\lambda)^h,\end{aligned}$$

with probability $1 - m^{-1}$ over the choice of $\mathcal{T}_{\le L}$. □

We conclude this section by completing the proof of Theorem 1.

Proof of Theorem 1. Let K, λ_0 be given as Theorem 2.2.4, and set $\lambda \le \lambda_0$ to be a constant such that $K\lambda < 1$. Let $\delta > 0$ be any given small number, and set h to be the constant satisfying $(K\lambda)^h = \frac{\delta^2}{8}$. Further, let $\mathcal{T} \sim \mathsf{GW}(\xi)$ and ρ be its root.

Define $E(h)$ to be the event that the infection inside $\mathsf{CP}^{\lambda}(\mathcal{T}; \mathbf{1}_\rho)$ does not go deeper than depth h until dying out. Then, Theorem 2.2.4 implies that

$$\mathbb{P}(E(h)) \ge 1 - \delta,$$

which can be seen by setting $m = \frac{2}{\delta}$.

Let $\mathcal{T}_h$ be the truncated tree of $\mathcal{T}$ at depth h, and couple the processes $\mathsf{CP}^\lambda(\mathcal{T}; \mathbf{1}_\rho)$ and $\mathsf{CP}^\lambda(\mathcal{T}_h; \mathbf{1}_\rho)$ by identifying the recoveries and infections inside $\mathcal{T}_h$. Then, on $E(h)$, $\mathsf{CP}^\lambda(\mathcal{T}; \mathbf{1}_\rho)$ can be regarded as $\mathsf{CP}^\lambda(\mathcal{T}_h; \mathbf{1}_\rho)$. Let $\mathbf{R}$ and $\mathbf{R}_h$ be the times when $\mathsf{CP}^\lambda(\mathcal{T}; \mathbf{1}_\rho)$ and $\mathsf{CP}^\lambda(\mathcal{T}_h; \mathbf{1}_\rho)$ reaches $\mathbf{0}$. Then, Theorem 2.2.1 tells us that

$$\mathbb{E}[\mathbf{R}|E(h)] = \mathbb{E}[\mathbf{R}_h|E(h)] \le \frac{\mathbb{E}\mathbf{R}_h}{\mathbb{P}(E(h))} < \infty.$$

Thus, for $(X_t) \sim \mathsf{CP}^\lambda(\mathcal{T}; \mathbf{1}_\rho)$, we have

$$\mathbb{P}(X_t \neq \mathbf{0} \text{ for all } t \ge 0) \le \delta.$$

Since this holds true for all $\delta > 0$, we conclude that $\lambda_1^{\mathrm{GW}}(\xi) \ge \lambda > 0$. □

2.3 Short survival in random graphs

We turn our attention to the contact process on random graphs $G \sim \mathcal{G}(n, \mu)$. Throughout the rest of the chapter, μ is a probability distribution on $\mathbb{N}$ that satisfies for $D \sim \mu$,

$$\sigma^2 := \mathbb{E}D^2 < \infty \quad \text{and} \quad b := \frac{\mathbb{E}D(D-1)}{\mathbb{E}D} > 1,$$

as discussed in Section 1.4. In this section, in particular, we assume that μ has an exponential tail, i.e., $\mathbb{E}_{D\sim\xi} \exp(cD) < \infty$ for some $c > 0$. Our goal is to establish Theorem 4-(1), by proving the following:

Theorem 2.3.1. *Let μ be as above and $G \sim \mathcal{G}(n, \mu)$. For a vertex $v \in G$, let $\boldsymbol{T}_v$ denote the time when $\mathsf{CP}^\lambda(G; \mathbf{1}_v)$ reaches the state $\mathbf{0}$. Then there exist events $\mathcal{E}_1$, $\mathcal{E}_2(G)$, and constants $B, \lambda_0 > 0$ depending on μ such that the following hold:*

- *$\mathcal{E}_1$ is an event over the random graphs such that $\mathbb{P}_{\mathrm{RG}}(G \in \mathcal{E}_1) = 1 - o(1)$.*
- *$\mathcal{E}_2(G)$ is an event over the contact process $\mathsf{CP}^\lambda(G)$ such that*

$$\mathbb{P}_{\mathrm{CP}}(\mathcal{E}_2 \,|\, G) = 1 - o(1),$$

for all $G \in \mathcal{E}_1$.

- *For all* $\lambda \in (0, \lambda_0)$ *we have*
$$\frac{1}{n} \sum_{v \in G} \mathbb{E}_{\mathrm{CP}}[\boldsymbol{T}_v | G] \le B,$$
for all large enough n *and all* $G \in \mathcal{E}_1$.

Then, our main theorem follows simply by applying Markov's inequality.

Proof of Theorem 4-(1). Let T be the time when $\mathsf{CP}^\lambda(G; \mathbf{1}_G)$ reaches the state $\mathbf{0}$. For a graph $G \in \mathcal{E}_1$, the event $\mathcal{E}_2 = \mathcal{E}_2(G)$ given in Theorem 2.3.1, and any constant $C > 0$, we have

$$\begin{aligned}\mathbb{P}_{\mathrm{CP}}\left(\mathbf{T} > Cn \,|\, G, \mathcal{E}_2\right) &= \mathbb{P}_{\mathrm{CP}}\left(\max_{v \in G} \mathbf{T}_v > Cn \,\middle|\, G, \mathcal{E}_2\right) \\ &\le \sum_{v \in G} \mathbb{P}_{\mathrm{CP}}\left(\mathbf{T}_v > Cn \,|\, G, \mathcal{E}_2\right) \le \frac{B}{C},\end{aligned}$$

where the first equality is due to Lemma 1.1.1, and the second is from Markov's inequality. Since the events $\mathcal{E}_1$ and $\mathcal{E}_2$ given $G \in \mathcal{E}_1$ both hold whp, we obtain the conclusion. □

In the rest of the section we focus on proving Theorem 2.3.1. Our proof relies much on the fact that the local neighborhood $N(v, L) := \{u \in G : \mathrm{dist}(u, v) \le L\}$ of a fixed vertex v roughly looks like a Galton-Watson branching process. Hence the results from Section 2.2 will play a huge role in this section as well.

However, since $\mathcal{G}(n, \mu)$ contains cycles with nontrivial probability, we introduce a variant of Galton-Watson trees that can cover the effect of cycles in $\mathcal{G}(n, \mu)$, and develop a delicate coupling argument with the local neighborhood $N(v, L)$. This new branching process will stochastically dominate $N(v, L)$ in terms of isomorphic embeddings of graphs, and hence the contact process will survive for a longer time. The result will then follow by showing Theorem 2.3.1 for this new graph.

2.3.1 Coupling the local neighborhood

Let $G \sim \mathcal{G}(n, \mu)$, where μ has an exponential tail, and let μ' denote the size-biased distribution of μ. As discussed in Section 1.4.4, it is well known that the local neighborhood

$N(v, L)$ around $v \in G$ behaves roughly as the Galton-Watson process $\mathsf{GW}(\mu, \mu')^L$. However, the standard coupling between the two object produces an error at least $\Theta(n^{-1})$. Therefore, we consider augmented versions of μ, μ' to stochastically dominate $N(v, L)$ by a larger geometry.

Definition 2.3.2 (Augmented distribution)**.** Let μ be a probability distribution on $\mathbb{N}$ with an exponential tail. Let $k_0 = \max\{k : \sum_{j \geq k} \sqrt{p_j} \geq 1/2\}$, and $k_{\max} := \max\{k : p_k > 0\}$, with $k_{\max} = +\infty$ if the maximum does not exist. When $k_0 < k_{\max}$, we define the augmented distribution $\mu^\sharp$ of μ by

$$\mu^\sharp(j) = \frac{1}{Z} \begin{cases} p_j/2 & \text{if } j \leq k_0; \\ \sqrt{p_j} & \text{if } j > k_0, \end{cases}$$

where $Z = \sum_{j \leq k_0} p_j/2 + \sum_{j > k_0} \sqrt{p_j}$. If $k_0 = k_{\max}$, then we let

$$\mu^\sharp(j) = \frac{1}{Z} \begin{cases} p_j/2 & \text{if } j < k_0; \\ \sqrt{p_j} & \text{if } j = k_0, \end{cases}$$

where $Z = \sum_{j < k_0} p_j/2 + \sqrt{p_{k_0}}$.

We observe some of the basic properties of augmented distributions in the following lemma. The proof is based on elementary applications of estimating large deviation events, and is postponed to Section 2.7.1 since it is a bit technical and less related with the main theme of the work.

Lemma 2.3.3. *Let μ be a probability distribution on $\mathbb{N}$.*

(1) If μ has an exponential tail, then so does $\mu^\sharp$.

(2) Let $D_1, \ldots, D_n$ be n independent samples of μ. For a subset $\Delta \subset [n]$, let $\{p_k^\Delta\}_k$ denote the empirical distribution of $\{D_i\}_{i \in [n] \setminus \Delta}$. With high probability over the choice of D_i's, $\{p_k^\Delta\}_k$ is stochastically dominated by $\mu^\sharp$, for any $\Delta \in [n]$ with $|\Delta| \leq n/3$

Remark 2.3.4. The i.i.d D_i in the second condition of Lemma 2.3.3 can be viewed as a degree sequence of $G \sim \mathcal{G}(n, \mu)$. Consider the exploration procedure starting from a single fixed vertex v, which, at each step, reveals a vertex adjacent to the current explored

neighborhood and the half-edges incident to the new vertex. Then the second statement says that when the exploration process revealed $N \leq n/3$ vertices inside the local neighborhood of v, the empirical degree distribution of the $n - N$ unexplored vertices is stochastically dominated by $\mu^\sharp$, with high probability.

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

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

Lemma 2.3.5. *There exists $\gamma = \gamma(\mu) > 0$ such that for $G \sim \mathcal{G}(n, \mu)$, $\mathbb{P}(G \in \mathcal{A}(\gamma)) = 1 - o(1)$.*

This is a well-known property that holds true in general for various types of random graphs. Our proof of this lemma will be very similar to that of Lemma 2.1 in [LS10]. However, it is more technical due to generality of the model and hence we postpone the proof to Section 2.7.2.

Fix a constant $\gamma_1 > 0$ satisfying the condition in Lemma 2.3.5, and let $\mathcal{A} = \mathcal{A}(\gamma_1)$ for convenience. In the following, we define two Galton-Watson type branching processes, which are used to stochastically dominate $N(v, \gamma_1 \log n)$.

Definition 2.3.6 (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, $\mathcal{H} \sim \mathsf{GWC}(\xi, m)^l$, the *Galton-Watson-on-cycle process* (in short, **GWC**-*process*), is generated according to the following procedure:

1. Consider a length-m cycle $C = v_1 v_2 \dots v_m v_1$.
2. At each v_j, $j \in \{2, \dots, m\}$, 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 $\mathcal{H}$. We designate vertex v_1 as the root of $\mathcal{H}$ and denote $\rho = v_1$. Figure 2.1 illustrates an example of **GWC**-process.

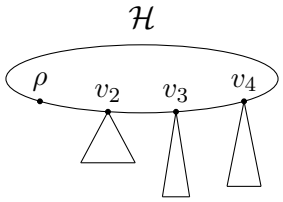


Figure 2.1: An example of $\mathcal{H} \sim \mathsf{GWC}(\xi, m)^L$ with $m = 4$.

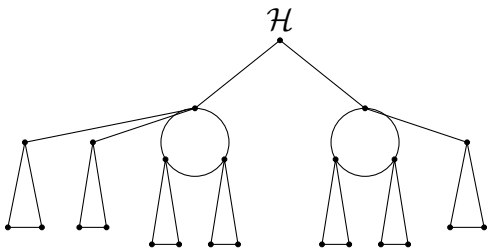


Figure 2.2: An example of $\mathcal{H} \sim \mathsf{EGW}(\nu; h, m)^L$ with $m = 3$ and $h = 1$.

Definition 2.3.7 (Edge-added Galton-Watson process)**.** Let h, m, L be positive integers with $s \geq 2$ and $h \leq L$, and let ξ be a probability distribution on $\mathbb{N}$. We define $\mathsf{EGW}(\xi; h, m)^L$, the edge-added Galton-Watson process (in short, **EGW**-process) as follows:

1. Generate a $\mathsf{GW}(\xi)^L$ tree, conditioned on survival until depth h.

2. At each vertex v at depth h, add an independent $\mathsf{GWC}(\xi; m)^{L-h}$ process rooted at v. Here we preserve the existing subtrees from v.

Let ξ' be another probability measure on $\mathbb{N}$. Then $\mathsf{EGW}(\xi, \xi'; h, m)^L$ denotes the **EGW**-process where the root has degree distribution ξ, and all the descendants have ξ'. Here we also add $\mathsf{GWC}(\xi'; m)^{L-h}$ in the second step of the definition. Figure 2.2 illustrates an example of **EGW**-process.

We now develop an argument showing that the local neighborhood $N(v, L)$ is dominated by a combined law of **EGW**-processes. In what follows, for two probability measures ν_1 and ν_2 on rooted graphs, we say ν_1 *stochastically dominates* ν_2 and write $\nu_1 \geq_{st} \nu_2$ if there exists a coupling between $(S_1, \rho_1) \sim \nu_1$ and $(S_2, \rho_2) \sim \nu_2$ such that $S_2 \subset S_1$ in terms of isomorphic embeddings of graphs, i.e., there exists an injective graph homomorphism from S_2 into S_1 that maps ρ_2 to ρ_1.

Fix a vertex $v \in G$, and consider its local neighborhood $N(v, L_n)$ where $L_n = \gamma_1 \log n$

with γ_1 as in Lemma 2.3.5. For each h, s with $s \geq 2$, we define the event $\mathcal{B}_{h,s}(v)$ to be the subevent of $\mathcal{A}$ such that in addition to $\mathcal{A}$, $N(v, L_n)$ forms a cycle of length s at distance h from v.

For the given degree distribution μ, let μ' be its size-biased distribution, and $\widetilde{\mu} := \mu'_{[1,\infty)}$ denote the distribution μ' conditioned on being in the interval $[1, \infty)$. Let $\mu^\sharp$ and $\widetilde{\mu}^\sharp$ be the augmented distributions of μ and $\widetilde{\mu}$, respectively. Further, let η, $\eta_{h,s}$ and η_0 denote the probability measures on rooted graphs describing the laws of $N(v, L_n)$, $\mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, s)^{L_n}$ and $\mathsf{GW}(\mu^\sharp, \widetilde{\mu}^\sharp)^{L_n}$, respectively.

Lemma 2.3.8. *Under the above setting, for a fixed vertex $v \in G$ we have the following stochastic domination:*

$$\eta \mathbf{1}_{\mathcal{A}} \leq_{st} \sum_{h,s:s\geq 2} b_{h,s}\eta_{h,s} + b_0\eta_0,$$

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

Proof. We study $N(v, L_n)$ from an exploration procedure point of view, in terms of the breadth-first search algorithm. Initially before exploring anything, we have n vertices with each of them having i.i.d. μ half-edges. The term "explore" means that we match a pair of half-edges and form an edge between their endpoint vertices. For convenience, we initially impose an arbitrary ordering on all half-edges before exploring anything. We consider the following exploration procedure:

- We start from the single vertex v and the half-edges adjacent to it.
- Suppose that we explored up to depth-t neighborhood of v. Let $\partial N(v, t)$ denote the unmatched half-edges on the boundary of $N(v, t)$, and we explore the half-edges in $\partial N(v, t)$ one by one, respecting the aforementioned ordering. During the (t, i)-th exploration step for $1 \leq i \leq |\partial N(v, t)|$, the i-th half-edge in $\partial N(v, t)$ is paired with a uniformly random unexplored half-edge.

Let $N(v, t; i)$ denote the explored neighborhood until (t, i)-th exploration step. Also, let $H_t = |\partial N(v, t)|$. During the (t, i)-th exploration step, the i-th half-edge of $\partial N(v, t)$ seeks for its uniformly random pair from the unexplored half-edges. Therefore, if we have yet

explored fewer than $\frac{n}{3}$ vertices, then after pairing a half-edge, the number of newly added half-edges to $N(v,t;i)$ from $N(v,t;i-1)$ is stochastically dominated by $\widetilde{\mu}^\sharp$, due to Lemma 2.3.3. This implies that conditioned on the event that $N(v,t;i)$ does not contain any cycles, $N(v,t;i)$ is stochastically dominated by $\mathcal{T}_{t,i}$, where $\mathcal{T}_{t,i}$ is generated by adding new offsprings according to $\widetilde{\mu}^\sharp$ to i vertices of depth t, to the Galton-Watson tree $\mathcal{T}_t \sim \mathsf{GW}(\mu^\sharp, \widetilde{\mu}^\sharp)^t$.

Define (T, I) to be the index of the exploration step when a cycle is formed. In other words, the I-th half-edge in $\partial N(v,T)$ is either paired to a j-th half-edge in $\partial N(v,T)$ for some $j > I$ or to one of the newly explored half-edges during the (T,k)-th exploration step for some $k < I$. Note that on the event $\mathcal{A}$, either the unique (T,I) exists or it does not exist up to exploring $N(v,L_n)$.

Suppose that there exists unique valid (T,I). Let C be the cycle formed at this step and $v(C)$ be the vertex in C that is closest to v. Up to the $(T,I-1)$-th exploration step, we can stochastically dominate $N(v,T;I-1)$ by $\mathcal{T}_{T,I-1}$ as mentioned above. Let $w(C)$ be the vertex in $\mathcal{T}_{T,I-1}$ corresponding to $v(C)$ via an isomorphic embedding of $N(v,T;I-1)$ into $\mathcal{T}_{T,I-1}$. At (T,I)-th exploration step, we add $\mathcal{S} \sim \mathsf{GWC}(\widetilde{\mu}^\sharp; |C|)^{L_n}$ at $w(C)$. Note that this GWC-process $\mathcal{S}$ can be coupled with C and its descendants in $N(v,L_n)$ in the sense that each Galton-Watson subtree hanging to the cycle of $\mathcal{S}$ stochastically dominates the corresponding subtree in $N(v,L_n)$ hanging to C.

Let h denote the distance from v to $v(C)$. Completing the rest of the exploration as discussed above, $N(v,L_n)$ is stochastically dominated by $\mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, |C|)^{L_n}$, given that there exists the unique valid (T,I). This implies that on the event $\mathcal{A}$, $N(v,L_n)$ is stochastically dominated by a combined law of EGW-processes, and in this combination, the probability mass of appearance of $\mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, s)_{L_n}$ should be $b_{h,s} = \mathbb{P}(\mathcal{B}_{h,s})$. This concludes the proof of the claimed result. □

2.3.2 Estimating the survival time

Thanks to Lemma 2.3.8, we now study the contact process on edge-added Galton-Watson processes. On such graphs, we first show that the expected survival time of the contact process is bounded by a constant when the infection rate is small enough, as an analog of Theorem 2.2.1.

Proposition 2.3.9. *Let l, s, L be any integers such that $s \geq 2$ and $L \geq h$. Let $\boldsymbol{R}_{h,s,L}$ denote the first time when $\mathsf{CP}^{\lambda}(\mathcal{S}; \mathbf{1}_{\rho})$ reaches at state $\mathbf{0}$, where $\mathcal{S} \sim \mathsf{EGW}(\mu^{\sharp}, \widetilde{\mu}^{\sharp}; h, s)^{L}$ rooted at ρ. Then there exist constants $C, \lambda_0 > 0$ depending only on μ such that for all $\lambda \leq \lambda_0$, s, h and L, we have $\mathbb{E}\boldsymbol{R}_{h,s,L} \leq C$.*

Remark 2.3.10. Since the coupling given in Lemma 2.3.8 only works until depth $L_n = \gamma_1 \log n$, we need to show that the contact process on edge-added Galton-Watson process does not go deeper than $\gamma_1 \log n$ with probability $1 - o(n^{-1})$. Note that the $o(n^{-1})$ error is needed when applying a union bound over all vertices in order to translate our results to $G \sim \mathcal{G}(n, \mu)$. This will be done in the next section based on Theorem 2.2.4.

To establish Proposition 2.3.9, we develop a recursive argument on both s and L to deduce an analog of Lemma 2.2.3 for **GWC**- and **EGW**-processes. The idea will be similar to that of Lemma 2.2.3, which is to utilize the notion of *root-added contact process* (Definition 2.2.2). We first extend the result of Lemma 2.2.3 to the case of **GWC**-processes: in the following lemma, we estimate the time that the *root-added* contact process $\mathsf{CP}^{\lambda}_{\rho}(\mathcal{S})$ reaches $\mathbf{0}$, where $\mathcal{S} \sim \mathsf{GWC}(\widetilde{\mu}^{\sharp}; s)^{L}$ (note that the state space is now $\{0,1\}^{\mathcal{S}\setminus\{\rho\}}$). Here, we fix the root ρ of $\mathcal{S}$ to be the permanently infected parent. There is a slight difference from the previous root-added contact processes considered in Lemma 2.2.3, since now the permanently infected parent has two children rather than one. We pick a child v of ρ and study $\mathsf{CP}^{\lambda}_{\rho}(\mathcal{S}; \mathbf{1}_{v})$.

Lemma 2.3.11. *Let s, L be any integers with $s \geq 2$ and let $\mathcal{S} \sim \mathsf{GWC}(\widetilde{\mu}^{\sharp}; s)^{L}$ be a **GWC**-process rooted at ρ, with $\widetilde{\mu}^{\sharp}$ as before. Let v be any neighbor of ρ, and let $\boldsymbol{S}_{s,L}$ denote the first time when $\mathsf{CP}^{\lambda}_{\rho}(\mathcal{S}; \mathbf{1}_{v})$ reaches at state $\mathbf{0}$. Then there exists a constant $\lambda_0 > 0$ depending only on μ such that for any $\lambda \leq \lambda_0$ and s, L, $\mathbb{E}\boldsymbol{S}_{s,L} \leq 2e$.*

Proof. Let us first study the case of $s \geq 3$. We will develop an inductive argument on s, similarly as in Lemma 2.2.3. Let $(X_t) \sim \mathsf{CP}^{\lambda}_{\rho}(\mathcal{S}; \mathbf{1}_{v})$, and consider the modified version $(\widetilde{X}_t) \sim \widetilde{\mathsf{CP}}^{\lambda}_{\rho;v}(\mathcal{S}; \mathbf{1}_{v})$ of (X_t), defined as

1. $(\widetilde{X}_t)$ is coupled with (X_t) in the sense that they share the same locations of recoveries and infections. In particular, ρ is infected in $(\widetilde{X}_t)$.

2. In $(\widetilde{X}_t)$, the recovery at v at time s is valid if and only if $\widetilde{X}_s = \mathbf{1}_v$. Otherwise, we ignore the recovery at v. In other words, when there exists an infected vertex other than ρ and v, the recovery at v is neglected.

The modified process $\widetilde{\mathsf{CP}}^{\lambda}_{\rho;v}(\mathcal{S};\mathbf{1}_v)$ plays the same role as the $\widetilde{\mathsf{CP}}$-process introduced in the proof of Lemma 2.2.3, which we now detail. Let $D \sim \widetilde{\mu}^{\sharp}$ to satisfy $\deg(v) = D+2$, and let $u_1,\dots,u_D$ be the neighbors of v which are not on the cycle of $\mathcal{S}$. Let $\mathcal{T}_{u_i}$ denote the subtrees branching from u_i, which has the law of i.i.d. $\mathsf{GW}(\widetilde{\mu}^{\sharp})^{L-1}$, and regard $v \in \mathcal{T}^{+}_{u_i}$ as the permanently infected parent of u_i. Further, call $\mathcal{S}' = \mathcal{S} \setminus \cup_{i=1}^{D}\mathcal{T}_{u_i}$, and define $\mathsf{CP}^{\lambda}_{\rho,v}(\mathcal{S}')$ to be the contact process on $\mathcal{S}'$ in which ρ and v are set to be infected permanently. As we run the process $(\widetilde{X}_t)$ from $t=0$,

A. The second state of $\widetilde{X}_t$ is $\mathbf{0}$ with probability $\frac{1}{1+(D+2)\lambda}$. Here, $D+2$ comes from $D+1$ possible new infections from v, and one possible infection from ρ to its child other than v. When this happens, the expected waiting time until the transition to $\mathbf{0}$ is $\frac{1}{1+(D+2)\lambda}$.

B. Otherwise, $\{\rho,v\}$ infects a uniformly random neighbor U before v is healed, and $(\widetilde{X}_t)$ then can be regarded as a product chain of $\mathsf{CP}^{\lambda}_{\rho,v}(\mathcal{S}')$ and $\{\mathsf{CP}^{\lambda}_{v}(\mathcal{T}^{+}_{u_i}) : i \in [D]\}$ with initial state $\mathbf{1}_U$, until $\widetilde{X}_t$ returns back to $\mathbf{1}_v$. Denote this product chain by $\mathsf{CP}^{\otimes}_{\rho;v}(\mathcal{S})$ (whose state space is $\{0,1\}^{\mathcal{S}\setminus\{\rho,v\}}$). Here, U can be thought of the first infected vertex besides $\{\rho,v\}$ in $\mathsf{CP}^{\otimes}_{\rho;v}(\mathcal{S};\mathbf{0})$.

Let $\widetilde{\mathbf{S}}_{s,L}$ be the first time that $\widetilde{\mathsf{CP}}^{\lambda}_{\rho;v}(\mathcal{S};\mathbf{1}_v)$ becomes $\mathbf{0}$, and let $\mathbf{S}^{\otimes}$ denote the first time that $\mathsf{CP}^{\otimes}_{\rho;v}(\mathcal{S};\mathbf{1}_U)$ reaches all-healthy state except ρ, v. Then, similarly as in Lemma 2.2.3, the above reasoning implies that

$$\begin{aligned}\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}_{s,L}\,\middle|\,\mathcal{S}\right] &= \sum_{k=0}^{\infty}\left(\frac{(D+2)\lambda}{1+(D+2)\lambda}\right)^{k}\frac{1}{1+(D+2)\lambda} \\ &\qquad\times\left[\frac{k+1}{1+(D+2)\lambda} + k\,\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\otimes}\,\middle|\,\mathcal{S}\right]\right] \\ &= 1+(D+2)\lambda\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\otimes}\,\middle|\,\mathcal{S}\right].\end{aligned} \tag{2.3.1}$$

Therefore, we have

$$\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}_{s,L}\,\middle|\, D\right] = 1 + (D+2)\lambda\mathbb{E}_{\mathrm{CP}}\left[\mathbf{S}^{\otimes}\middle|\, D\right]. \tag{2.3.2}$$

Now we take account of the stationary distributions of the above processes to obtain the conclusion. Let $\pi^{\otimes}$, π' and π_i be the stationary distributions of $\mathbf{CP}^{\otimes}_{\rho;v}(\mathcal{S})$, $\mathbf{CP}^{\lambda}_{\rho,v}(\mathcal{S}')$ and $\mathbf{CP}^{\lambda}_{v}(\mathcal{T}^{+}_{u_i})$, respectively. Then clearly, $\pi^{\otimes} = (\otimes_{i=1}^{D}\pi_i)\otimes\pi'$. We can relate these objects with the running times similarly as (2.2.3, 2.2.4), by

$$\begin{aligned}
\pi^{\otimes}(\mathbf{0}) &= \frac{1}{1+(D+2)\lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}^{\otimes}|\mathcal{S}]};\\
\pi_i(\mathbf{0}) &= \frac{1}{1+\lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}_{L-1}|\mathcal{T}_{u_i}]};\\
\pi'(\mathbf{0}) &= \frac{1}{1+2\lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}_{s-1,L}|\mathcal{S}']},
\end{aligned} \tag{2.3.3}$$

where $\mathbf{S}_{L-1}$ denotes the first time when $\mathbf{CP}^{\lambda}_{v}(\mathcal{T}^{+}_{u_i};\mathbf{1}_{u_i})$ becomes $\mathbf{0}$, and $\mathbf{S}_{s-1,L}$ is the time it takes for $\mathbf{CP}^{\lambda}_{\rho,v}(\mathcal{S}';\mathbf{0})$ to return to $\mathbf{0}$ after the first infection besides $\{\rho,v\}$ occurs. Note that the existence of the infection other than $\{\rho,v\}$ in $\mathbf{CP}^{\lambda}_{\rho,v}(\mathcal{S}';\mathbf{0})$ is guaranteed by the condition $s\geq 3$. Also, notice that $\mathcal{S}'$ can be regarded as $\mathcal{S}''\sim\mathbf{GWC}(\widetilde{\mu}^{\sharp};s-1)^{L}$, since the processes $\mathbf{CP}^{\lambda}_{\rho,v}(\mathcal{S}';\mathbf{1}_w)$ (with w being a neighbor of $\{\rho,v\}$ in $\mathcal{S}'$) and $\mathbf{CP}^{\lambda}_{\rho'}(\mathcal{S}'';\mathbf{1}_{v'})$ for the root ρ' and one of its neighbor v' of $\mathcal{S}''$ share the same law. This implies that the notation $\mathbf{S}_{s-1,L}$ in (2.3.3) matches with the definition of it given in the statement of the lemma.

Therefore, combining (2.3.2) and (2.3.3) gives that

$$\mathbb{E}\left[\mathbf{S}_{s,L}|D\right] \leq \mathbb{E}\left[\widetilde{\mathbf{S}}_{s,L}\middle|\, D\right] \leq \left(1+2\lambda\mathbb{E}[\mathbf{S}_{s-1,L}]\right)\left(1+\lambda\mathbb{E}[\mathbf{S}_{L-1}]\right)^{D}. \tag{2.3.4}$$

Since we already have a bound for $\mathbb{E}[\mathbf{S}_{L-1}]$ due to Lemma 2.2.3, we deduce the desired result by manipulating (2.3.4), as in the final step of the proof of Lemma 2.2.3. Namely, for any $\lambda\leq\min\{\lambda_0,(4e)^{-1}\}$ with λ_0 given as in Lemma 2.2.3, (2.3.4) gives us that

$$\mathbb{E}[\mathbf{S}_{s,L}] \leq e(1+2\lambda\mathbb{E}[\mathbf{S}_{s-1,L}]),$$

and hence $\mathbb{E}[\mathbf{S}_{s-1,L}]\leq 2e$ implies $\mathbb{E}[\mathbf{S}_{s,L}]\leq 2e$.

The case $s=2$ is simpler, since $\mathbf{GWC}(\widetilde{\mu}^{\sharp};2)^{L}$ is the same as the law of $\mathcal{T}^{+}_{\leq L}$ for $\mathcal{T}_{\leq L}\sim$

$\mathbf{GW}(\widetilde{\mu}^\sharp)^L$, except that the parent ρ^+ of ρ in $\mathcal{T}^+_{\le L}$ is now connected with ρ by two edges. From the contact process point of view, this means that the intensity of infection from ρ^+ to ρ is 2λ, and everything else is identical to the case of $\mathbf{GW}(\widetilde{\mu}^\sharp)^L$. Hence, the same proof as Lemma 2.2.3 can be replicated, and we obtain that there exists a constant $\lambda_0 > 0$ such that $\mathbb{E}[\mathbf{S}_{2,L}] \le 2e$ for all $\lambda \le \lambda_0$. We leave the details of the proof to the reader. □

Proof of Proposition 2.3.9. It can be proven by the same way as Lemma 2.3.11. For completeness, we present the proof in Section 2.7.3. □

2.3.3 Bounding the survival time: Proof of Theorem 2.3.1

Let h, s, L be arbitrary integers with $s \ge 2$ and $L \ge h$, and consider an edge-added Galton-Watson process $\mathcal{S} \sim \mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, s)^L$.

We can extend the result of Theorem 2.2.4 and Proposition 2.2.6 to the case of edge-added Galton-Watson processes. The method will be the same as Proposition 2.3.9 and Lemma 2.3.11, appropriately adjusted to the current setting of *delayed* contact process (Definition 2.2.5). We state the result in the following lemma, whose proof is deferred to Section 2.7.4, since it is similar to the previous proofs but more technical.

Lemma 2.3.12. *Let $\mathcal{S} \sim \mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, s)^L$, and $\nu^\theta_{\mathcal{S}}$ be the stationary distribution of $\mathbf{DP}^{\lambda,\theta}_{\rho^+}(\mathcal{S}^+)$ on the space $\{0,1\}^{\mathcal{S}}$. Then there exist constants $C, \lambda_0 > 0$ depending only on μ such that for all $\lambda \le \lambda_0$, we have $\mathbb{E}[\nu^\theta_{\mathcal{S}}(\mathbf{0})^{-1}] \le 2$ for $\theta = C\lambda$.*

Based on Proposition 2.2.6 and Lemma 2.3.12, we have an analog of Theorem 2.2.4 for **EGW**-processes. Thus, we can complete the proof of Theorem 4-(1), by combining the previous results to build up a coupling between the contact processes on local neighborhood $N(v, L)$ and on **EGW**-processes.

Proof of Theorem 2.3.1. Let $G \sim \mathcal{G}(n, \mu)$ for any large enough n and let $\gamma_1 > 0$ be the constant satisfying Lemma 2.3.5. For each $v \in G$, let $\mathcal{A}_v$ be the event that $N_v := N(v, \gamma_1 \log n)$ in G contains at most one cycle. Then, the proof of Lemma 2.3.5 tells us $\mathbb{P}(\mathcal{A}_v) \ge 1 - o(n^{-1})$.

Let λ'_0 be the minimum between the λ_0's given by Proposition 2.3.9 and Lemma 2.3.12, C be as in 2.3.12, and $\theta = C\lambda$ for $\lambda \le \lambda'_0$. Further, let $(X_t) \sim \mathbf{CP}^\lambda(N_v; \mathbf{1}_v)$ and $H_v :=$

$\max\{r(X_t) : t \geq 0\}$, where $r(X_t)$ denotes the maximal depth among the infected sites in X_t. Note that $r(X_t)$ stays 0 after X_t becomes $\mathbf{0}$.

Define the event $\mathcal{B}_v$ as

$$\mathcal{B}_v := \{\mathbb{P}(H_v \geq h \,|\, N_v) \leq 2n^2(C\lambda)^h, \text{ for all } h\}.$$

Following the same proof as Theorem 2.2.4 based on Lemma 2.3.12, we have $\mathbb{P}(\mathcal{B}_v) \geq 1-n^{-2}$, by dominating N_v by the **EGW**-processes as Lemma 2.3.8. Assume that $(C\lambda_0')^{\gamma_1 \log n} \leq n^{-4}$ by making λ_0' smaller if needed. Then, the event $\mathcal{C}_v$ given by

$$\mathcal{C}_v = \mathcal{C}_v(G) := \{H_v < \gamma_1 \log n\},$$

satisfies $\mathbb{P}(\mathcal{C}_v|\mathcal{A}_v \cap \mathcal{B}_v) = O(n^{-2})$. Note that $\mathcal{A}_v \cap \mathcal{B}_v$ is an event over the random graph G, while $\mathcal{C}_v$ is an event over the contact process (X_t) given the graph G. By the aforementioned coupling of N_v and the **EGW**-process, Proposition 2.3.9 gives that

$$\mathbb{E}_{\mathrm{CP}}[\mathbf{T}_v \,|\, G, \mathcal{C}_v] \leq B, \tag{2.3.5}$$

for all $G \in \mathcal{A}_v \cap \mathcal{B}_v$, with some constant $B = B(\mu) > 0$.

Define the events

$$\mathcal{E}_1 := \cap_{v \in G}(\mathcal{A}_v \cap \mathcal{B}_v), \quad \text{and} \quad \mathcal{E}_2(G) := \cap_{v \in G}\mathcal{C}_v.$$

Then, the above discussion shows that $\mathbb{P}_{\mathrm{RG}}(G \in \mathcal{E}_1) = 1 - o(1)$ and $\mathbb{P}_{\mathrm{CP}}((X_t) \in \mathcal{E}_2(G)|G \in \mathcal{E}_1) = 1 - o(1)$, and hence (2.3.5) holds the same given $\mathcal{E}_1$ and $\mathcal{E}_2(G)$, namely,

$$\mathbb{E}_{\mathrm{CP}}[\mathbf{T}_v \,|\, G \in \mathcal{E}_1, \mathcal{E}_2(G)] \leq B,$$

under a possible modification of B if needed. Therefore, by linearity of expectation, summing the above over all $v \in G$ gives the conclusion. □

Proof of Corollary 6-(1). The statement follows immediately from the contiguity between

$\mathcal{G}(n, \mathrm{Pois}(d))$ and $\mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$. To be precise, for any subset A_n of graphs with n vertices,

$$\mathbb{P}_{G\sim\mathcal{G}(n,\mu)}(G \in A_n) \to 0 \quad \text{implies} \quad \mathbb{P}_{G\sim\mathcal{G}^{\mathrm{ER}}(n,\frac{d}{n})}(G \in A_n) \to 0,$$

where $\mu = \mathrm{Pois}(d)$. Since the statement is true whp for $G \sim \mathcal{G}(n, \mu)$, the configuration model, it is also true whp for $G \sim \mathcal{G}^{\mathrm{ER}}(n, \frac{d}{n})$. □

2.4 Long survival in random graphs: Proof of Theorem 4, part 2

2.4.1 A structural lemma

Our main tool to prove long survival time is the structural lemma below whose proof is deferred to Section 2.6. As mentioned in Section 2.1, we show that the random graph G_n contains a large (α, R)-*embedded expander*. Once some subset of this expander is infected, it is likely to spread the infection over its R-neighborhood whose size more than doubles the original subset. We define an embedded expander as follows.

Definition 2.4.1 (Embedded expander)**.** For two positive numbers α and R, we say that a subset of vertices W_0 is an (α, R)-*embedded expander* of G_n if for every subset $A \subset W_0$ with $|A| \le \alpha|W_0|$, we have

$$|N(A, R) \cap W_0| \ge 2|A|. \tag{2.4.1}$$

where $N(A, R)$ is the collection of all vertices in G_n of distance at most R from A.

The following lemma concerns the existence of such an (α, R)-*embedded expander* in the random graph G_n.

Lemma 2.4.2. *Suppose that μ satisfies (1.5.4) and there exists some constant $c > 0$ such that $\mathbb{E}_{D\sim\mu}e^{cD} < \infty$. Let $G \sim \mathcal{G}(n, \mu)$. There exist positive constants α, β, R, j such that the following holds* whp*. There exist a subgraph $\bar{G}_n$ of G_n whose maximal degree is at most $2j$ and an (α, R)-embedded expander W_0 of $\bar{G}_n$ with $|W_0| \ge \beta n$.*

2.4.2 Proof of Theorem 4, part 2

We first make a simple observation.

Lemma 2.4.3. *Let R be a positive integer constant. Consider the contact process Y_t with infection rate λ on a path of length R connecting two vertices v and u. Then there exist positive constants C and λ_0 depending only on R such that for all $\lambda \geq \lambda_0$, we have*

$$\mathbb{P}_{\text{CP}}\left(u \in Y_{t+C} \middle| v \in Y_t\right) \geq \frac{3}{4}.$$

Proof. Let C be a sufficiently large constant compared to R. Let $\mathcal{A}$ be the event that the infection on $N(v,1)$ survives in the entire time interval $[t, t+C]$.

By [CD09, Lemma 1.1], for sufficiently large λ compared to C,

$$\mathbb{P}_{\text{CP}}\left(\mathcal{A} \middle| v \in Y_t\right) \geq \frac{99}{100}.$$

Since there is a path of length R from v to u, by [CD09, Lemma 2.4],

$$\mathbb{P}_{\text{CP}}\left(u \in Y_{t''} \text{ for some } t'' \in [t', t'+R] \middle| N(v,1) \cap Y_{t'} \neq \emptyset\right) \geq \frac{1}{e^{6R}},$$

and so as C is large compared to R, we have

$$\begin{aligned}\mathbb{P}_{\text{CP}}&\left(u \in Y_{t'} \text{ for some } t' \in [t, t+C+R] \middle| v \in Y_t\right) \\ &\geq \mathbb{P}\left(\mathbf{Bin}\left(\frac{C}{R}, \frac{1}{e^{6R}}\right) \geq 1\right) - \frac{1}{100} \geq \frac{98}{100}.\end{aligned}$$

Assume that $u \in Y_{t'}$ for some $t' \in [t, t+C+R]$. Fix a neighbor u' of u. By [CD09, Lemma 1.1] again, the contact process on the edge (u, u') survives in the entire interval $[t', t'+3C]$ with probability at least $\frac{99}{100}$. Since $t+C+R \in [t', t'+2C]$, there is at least one clock ring in $[t+C+R, t'+3C]$ and the last clock ring before time $t+3C$ is an infection clock from u' to u rather than the recovery clock at u, with probability at least $\frac{99}{100}$. If u' is already infected at that time, u will be infected. Otherwise, u has already been infected and remains infected. In either case, u is infected at time $t+3C$ with probability at least $\frac{96}{100} \geq \frac{3}{4}$. By replacing

C by $C/3$, we complete the proof. □

Let $\alpha, \beta, j, R, \bar{G}_n$ and W_0 as in Lemma 2.4.2. It suffices to show that the contact process (X_t) on $\bar{G}_n$ with all vertices infected initially survives for $e^{\Theta(n)}$-time with probability at least $1 - e^{-\Omega(n)}$ over the contact process. For the rest of this proof, all the vertices, edges, paths, and balls are of $\bar{G}_n$ unless otherwise noted.

Let $X_t^0 = X_t \cap W_0$ be the collection of infected vertices of W_0 at time t. We show that, thanks to the expander property of W_0, with very high probability, after some time C, the number of infected vertices in W_0 increases.

Lemma 2.4.4. *Let C and λ_0 be the constants in Lemma 2.4.3. There exists a positive constant C' depending only on j and R such that for all $\lambda \geq \lambda_0$ and for every integer $a \in (0, \alpha\beta n]$,*

$$\mathbb{P}_{\mathrm{CP}}\left(|X_{t+C}^0| \leq \frac{5}{4}a \middle| |X_t^0| = a\right) \leq 2\exp\left(-\frac{a}{C'}\right). \tag{2.4.2}$$

Proof. We will use Azuma's inequality. Let $G_{n,t}$ be the induced subgraph of $\bar{G}_n$ on the set $\bigcup_{v \in X_t^0} N(v, R)$. Let $(\hat{X}_{t'})_{t' \in [t, t+C]}$ be the contact process on $G_{n,t}$ with $\hat{X}_t := X_t^0$ (so $(\hat{X}_{t'})$ only uses the infection and recovery clocks of vertices and edges inside $G_{n,t}$). Let $\mathcal{X}$ be the infected vertices u of $\hat{X}_{t+C}$ such that there exists $v \in X_t^0$ and a directed path of infection on the graphical representation of $(\hat{X}_{t'})$ from (v, t) to $(u, t+C)$ and the vertices of the path lie entirely in $B(v, R)$. We have

$$\mathcal{X} \subset \hat{X}_{t+C} \subset X_{t+C}^0.$$

Since the maximal degree in $\bar{G}_n$ is at most $2j$, the number of vertices, denoted by a', in $G_{n,t}$ is at most $a(2j)^{R+1}$. Enumerate the vertices in $G_{n,t}$ by $v_1, \ldots, v_{a'}$. For each $i = 0, 1, 2, \ldots, a'$, let $\mathcal{F}_i$ be the σ-algebra generated by the randomness of the recovery clocks and infection clocks during time $(t, t+C]$ on the vertices $v_1, \ldots, v_i$ and edges connecting them. Let

$$\mathcal{X}_i := \mathbb{E}\left(|\mathcal{X}| \middle| \mathcal{F}_i, X_t^0, |X_t^0| = a\right).$$

We have $\mathcal{X}_{a'} = \mathcal{X}$. By Azuma's inequality, we have for every $s > 0$,

$$\mathbb{P}\left(|\mathcal{X}_{a'} - \mathcal{X}_0| \geq s\right) \leq 2\exp\left(-\frac{s^2}{2a'K^2}\right) \tag{2.4.3}$$

where

$$K := \max_j ||\mathcal{X}_j - \mathcal{X}_{j-1}||_\infty \leq |B\left(v_{i+1}, R\right)| \leq (2j)^{R+1}.$$

From Lemma 2.4.3 and the expander properties of W_0 as in Lemma 2.4.2, we obtain

$$\mathcal{X}_0 = \mathbb{E}\left(|\mathcal{X}| \,\middle|\, |X_t^0| = a\right) \geq \frac{3}{4}\left|N(X_t^0, R) \cap W_0\right| \geq \frac{3a}{2}. \tag{2.4.4}$$

Thus, by (2.4.3) for $s = \frac{a}{4}$ and the fact that $\mathcal{X}_{a'} \leq \left|X_{t+C}^0\right|$, we obtain (2.4.2). □

Proof of Theorem 4-(2). For the lower bound on survival time, initially, all vertices in W_0 are infected so $|X_0^0| \geq \beta n$. Let t_1 be the first time that $|X_{t_1}^0| = \alpha\beta n$. At time $t_2 = t_1 + C$, we have $|X_{t_2}^0| \geq \frac{5}{4}\alpha\beta n$ with probability at least $1 - \frac{2}{m^2}$ where $m := \exp\left(\frac{\alpha\beta n}{2C'}\right)$ by Lemma 2.4.4. Let t_3 be the first time after t_2 that $|X_{t_3}^0| = \alpha\beta n$ again. Repeating this process m times, we get that the contact process survives until time mC with probability at least $1 - 2/m$ by the union bound, proving the lower bound for Theorem 4-(2).

As for the upper bound, observe that for any time t, the probability that the contact process dies out during the time interval $[t, t+1]$ is at least the probability that for each vertex v in G_n, at least one of the infection clocks from a neighbor u of v to v or the recovery clock at v rings in $[t, t+1]$ and the last clock rings before time $t+1$ is the recovery clock at v. Thus, the probability that the process dies out during $[t, t+1]$ is at least $\prod_v \frac{c}{\deg(v)}$. By Cauchy-Schwartz inequality and the fact that whp, the total degrees in G_n is $O(n)$, we have $\prod_v \frac{c}{\deg(v)} \geq e^{-c'n}$ for some small constants c, c'. Therefore, whp, the contact process dies out before time $e^{2c'n}$. □

Remark 2.4.5. To prove the corresponding result (Remark 1.5.5) for μ having an exponential tail, when initially, there is only one uniformly chosen vertex v infected in G_n, observe that with positive probability over the choice of v, v belongs to W_0. Thus, it suffices to condition on this event and show that with positive probability over the contact process, the

process survives until time e^{cn} for some constant c. Let $\lambda_{0,R}$, C_R and C'_R be the constants λ_0, C and C' corresponding to R in Lemmas 2.4.3 and 2.4.4, respectively. By (2.4.1), for any $k \geq 1$,

$$|N(v, kR) \cap W_0| \geq 2^k.$$

Since the number of vertices in $N(v, kR)$ is at most $(2j)^{kR+1}$, there are at most $(2j)^{2kR+2}$ edges in $N(v, kR)$. Thus, there exists a constant $\lambda_{j,k,R}$ such that for all $\lambda \geq \lambda_{j,k,R}$, the probability that all the infection clocks in $N(v, kR)$ ring before time $t_0 = \frac{\Theta(1)}{(2j)^{kR+1}}$ and all recovery clocks ring after time t_0 is at least $3/4$. Thus, with probability at least $3/4$, there exists $t_1 \leq t_0$ at which $|X^0_{t_1}| \geq 2^k \geq (5/4)^k$.

Conditioning on this event and applying Lemma 2.4.4, we get that with probability at least

$$1 - \sum_{i=k}^{\infty} 2 \exp\left(-\frac{5^i}{C'_R 4^i}\right)$$

there exists a time $t \geq t_1$ at which $|X^0_t| \geq \alpha\beta n$. Let k be a sufficiently large constant, this probability is at least $1/2$. Finally, conditioned on this event, the same argument as in the proof of Theorem 4-(2) shows that starting from this t, the contact process survives until time $e^{-\Omega(n)}$ whp. Altogether, the contact process (X_t) starting from v survives until time $e^{-\Omega(n)}$ with probability at least $1/4 - o(1)$ for all $\lambda \geq \max\{\lambda_{0,R}, \lambda_{j,k,R}\}$ as desired.

Remark 2.4.6. Corollary 6-(2) follows from Theorem 4-(2) in the exact same way as we deduced Corollary 6-(1) from Theorem 4-(1).

2.5 Long survival in random graphs: Proof of Theorem 5

In this section, we prove Theorem 5 following the same strategy as in the proof of Theorem 4-(2). The following structural lemma is an analog of Lemma 2.4.2 for subexponential distributions.

Lemma 2.5.1. *Suppose that μ satisfies (1.5.4) and $\mathbb{E}_{D\sim\mu} e^{cD} = \infty$ for all $c > 0$. Let $G \sim \mathcal{G}(n, \mu)$. For any $\delta > 0$, there exist $\alpha, \beta, j, R > 0$ with $R \leq \delta j$ such that the following holds whp. There exist a subgraph $\bar{G}_n$ of G_n whose maximal degree is at most $2j$ and an (α, R)-embedded expander W_0 of $\bar{G}_n$ with $|W_0| \geq \beta n$ and $deg_{\bar{G}_n} w \geq j/2$ for all $w \in W_0$.*

Fix $\lambda > 0$. Let α, β, j, R, $\bar{G}_n$, W_0 be as in Lemma 2.5.1 where $\delta > 0$ is a sufficiently small constant depending on λ. It suffices to show that the contact process (X_t) on $\bar{G}_n$ with all vertices infected initially survives for $e^{\Omega(n)}$-time with probability at least $1 - e^{-\Omega(n)}$ over the contact process. For the rest of the proof, all the vertices, edges, paths, and balls are of $\bar{G}_n$.

Let X_t^0 be the collection of infected vertices of W_0 at time t. We show the following analog of Lemma 2.4.3.

Lemma 2.5.2. *There exists a constant c depending only on λ such that for sufficiently large n and for every $u \in N(v, R) \cap W_0$, we have*

$$\mathbb{P}_{\mathrm{CP}}\left(u \in X^0_{t+2e^{cj}} \middle| v \in X_t^0\right) \geq \frac{3}{4}. \tag{2.5.1}$$

Assuming Lemma 2.5.2, we prove the following analog of Lemma 2.4.4.

Lemma 2.5.3. *Let c be the constant in Lemma 2.5.2. There exists a positive constant C' depending only on j and R such that for every integer $a \in (0, \alpha\beta n]$,*

$$\mathbb{P}_{\mathrm{CP}}\left(\left|X^0_{t+e^{cj}}\right| \leq \frac{5}{4}a \middle| |X_t^0| = a\right) \leq 2\exp\left(-\frac{a}{C'}\right). \tag{2.5.2}$$

The proof of this lemma is identical to the proof of Lemma 2.4.4. Using this lemma, the proof of Theorem 5 is identical to that of Theorem 4-(2). It remains to prove Lemma 2.5.2.

Proof of Lemma 2.5.2. Let Γ be a path of length at most R connecting v and u. Since $v \in W_0$, $N_{v,u} := \{v\} \cup (N(v, 1) \setminus \Gamma)$ contains a star with $j/4$ leaves. Let $\mathcal{A}$ be the event that the infection on $N_{v,u}$ survives in the entire time interval $\left[t, t + e^{cj}\right]$ for some constant c depending only on λ.

By [BBCS05, Lemma 5.3], by choosing δ sufficiently small in Lemma 2.5.1 and using the inequality $j \geq R/\delta \geq 1/\delta$, we have

$$\mathbb{P}_{\mathrm{CP}}\left(\mathcal{A} \middle| v \in X_t^0\right) \geq \frac{99}{100}.$$

Since there is a path of length at most R from v to u, by [CD09, Lemma 2.4], there exists

a constant $c' > 0$ depending only on λ such that for any time t',

$$\mathbb{P}_{\mathrm{CP}}\left(u \in X^0_{t''} \text{ for some } t'' \in [t', t'+R+1] \middle| N_{v,u} \cap X^0_{t'} \neq \emptyset\right) \geq c'^{R+1}.$$

Thus,

$$\begin{aligned}
&\mathbb{P}_{\mathrm{CP}}\left(u \in X^0_{t'} \text{ for some } t' \in [t, t+e^{cj}+R+1] \middle| v \in X^0_t\right) \\
&\quad \geq \mathbb{P}\left(\mathbf{Bin}\left(\frac{e^{cj}}{R}, c'^{R+1}\right) \geq 1\right) - \frac{1}{100} \\
&\quad \geq \mathbb{P}\left(\mathbf{Bin}\left(\frac{e^{cR/\delta}}{R}, c'^{R+1}\right) \geq 1\right) - \frac{1}{100} \geq \frac{98}{100}
\end{aligned}$$

by choosing δ sufficiently small (for example, $\delta = \frac{c}{100 \log c'^{-1}}$) in Lemma 2.5.1. Since $R \leq \delta j$, we can replace the interval $[t, t+e^{cj}+R+1]$ in the above inequality by the bigger interval $[t, t+2e^{cj}]$.

Assume that $u \in X^0_{t'}$ for some $t' \in [t, t+2e^{cj}]$. By the third inequality in [CD09, Lemma 2.3] and Markov's inequality, with probability at least $\frac{99}{100}$, there exists a time $t'' \in [t', t'+e^{cj}]$ such that there are at least $\lambda j/4$ neighbors of u infected at time t''. By [CD09, Lemma 2.2], with probability at least $\frac{99}{100}$, there are at least $\lambda j/10$ neighbors of u infected at any time in the time interval $[t'', t''+3e^{cj}]$. Since $t+2e^{cj} \in [t'', t''+3e^{cj}]$, the probability that u is infected at time $t+2e^{cj}$ is at least the probability that the last clock rings before time $t+2e^{cj}$ is an infection clock rather than the recovery clock at u. Since there are at least $\lambda j/10$ neighbors of u infected at any time in the interval $[t'', t+2e^{cj}]$, the probability of the above event is

$$\frac{\lambda^2 j/10}{\lambda^2 j/10 + 1} \geq \frac{99}{100}.$$

That completes the proof of Lemma 2.5.2. □

Remark 2.5.4. To prove the corresponding result (Remark 1.5.5) for subexponential μ when initially, there is only one uniformly chosen vertex v infected in G_n, observe that with positive probability over the choice of v, v belongs to W_0. Thus, it suffices to condition on this event and show that with positive probability over the contact process, the process survives until time $e^{\Omega(n)}$.

Since W_0 is an (α, R)-embedded expander, observe that for every $A \subset W_0$ with $|A| \leq \alpha|W_0|/4$, we have

$$|N(A, 2R)| \geq 4|A|.$$

Thus, using the same proof as for Lemmas and 2.5.2 and 2.5.3, one can see that there exist constants c' and C'' such that for all integer $a \in (0, \alpha\beta n/4]$,

$$\mathbb{P}_{\text{CP}}\left(\left|X^0_{t+e^{c'j}}\right| \leq 2a \middle| |X^0_t| = a\right) \leq 2\exp\left(-\frac{a}{C''}\right). \tag{2.5.3}$$

Let k be the largest number such that $2^k \leq \alpha\beta n/4$. Let $\mathcal{A}_0$ be the event that $X^0_0 = \{v\} \subset W_0$. For each $i = 1, \ldots, k$, let $\mathcal{A}_i$ be the event that

$$\left|X^0_{ie^{c'j}}\right| \geq 2^i.$$

Let $\mathcal{A}^*$ be the event that the contact process survives up to time $e^{\Omega(n)}$. We want to show that

$$\mathbb{P}_{\text{CP}}\left(\mathcal{A}^* \middle| X^0_0 = \{v\}\right) = \Omega(1). \tag{2.5.4}$$

In fact,

$$\begin{aligned}
\mathbb{P}_{\text{CP}}\left(\mathcal{A}^* \middle| X^0_0 = \{v\}\right) &\geq \mathbb{P}_{\text{CP}}\left(\bigcap_{i=1}^{k} \mathcal{A}_i \cap \mathcal{A}^* \middle| X^0_0 = \{v\}\right) \\
&\geq \prod_{i=1}^{k} \mathbb{P}_{\text{CP}}\left(\mathcal{A}_i \middle| \mathcal{A}_{i-1}\right) \mathbb{P}_{\text{CP}}\left(\mathcal{A}^* \middle| \mathcal{A}_k\right).
\end{aligned}$$

By (2.5.3), for each $i = 1, \ldots, k$, $\mathbb{P}_{\text{CP}}\left(\mathcal{A}_i \middle| \mathcal{A}_{i-1}\right) \geq 1 - 2\exp\left(-\frac{2^{i-1}}{C''}\right)$. Finally, by the same argument as in the proof of Theorem 4-(2), once there are about $\Theta(n)$ vertices in W_0 infected, the contact process survives for an exponentially long time whp. In other words, $\mathbb{P}_{\text{CP}}\left(\mathcal{A}^* \middle| \mathcal{A}_k\right) = 1 - o(1)$. Let k' be the smallest number such that $2^{k'-1} \geq 10C''$.

Combining all of these inequalities, we obtain

$$\begin{aligned}\mathbb{P}_{\mathrm{CP}}\left(\mathcal{A}^*\middle|X_0^0=\{v\}\right) &\geq \frac{98}{100}\prod_{i=2}^{k'}\left(1-2\exp\left(-\frac{2^{i-1}}{C''}\right)\right)\\ &\quad\times\prod_{i=k'+1}^{k}\left(1-2\exp\left(-\frac{2^{i-1}}{C''}\right)\right)\\ &\geq \frac{98}{100}\left(1-4e^{-10}\right)\prod_{i=2}^{k'}\left(1-2\exp\left(-\frac{2^{i-1}}{C''}\right)\right)=\Omega(1)\end{aligned}$$

as desired. That completes the proof of Remark 1.5.5.

2.6 Proof of the structural lemma of random graphs

In this section, we prove the structural Lemmas 2.4.2 and 2.5.1. We start by proving Lemma 2.5.1 for subexponential μ in Section 2.6.1. The proof of Lemma 2.4.2 is very similar and is presented in Section 2.6.2.

2.6.1 The embedded expander property in the subexponential case

In this subsection, we establish Proof of Lemma 2.5.1.

Step 1. Preprocessing. In this step, we eliminate high-degree vertices in G_n so that the degrees become bounded. This will allow us to control the size of the neighborhoods that we explore in the next steps. We prove in Lemma 2.6.1 that the elimination does not significantly affect relevant parameters of G_n.

Let $b=\frac{\sum_{l=1}^{\infty}l(l-1)\mu(l)}{\sum_{l=1}^{\infty}l\mu(l)}>1$ be the branching rate of μ. Let d be the mean $d=\mathbb{E}_{D\sim\mu}D>1$. For a constant j, consider the graph $\bar{G}_n$ obtained from G_n by deleting all vertices with degree at least $2j+1$ together with their half-edges and their matches. Let $\bar{n}$ be the number of vertices of $\bar{G}_n$ and $0\leq d_1\leq\cdots\leq d_{\bar{n}}\leq 2j$ be the degree sequence of vertices in $\bar{G}_n$.

The branching rate $\bar{b}$ of a (deterministic) degree sequence (d_i) is defined to be the branching rate of the empirical measure generated by (d_i), namely,

$$\bar{b}=\frac{\sum_l l(l-1)\#\{i:d_i=l\}}{\sum_l l\#\{i:d_i=l\}}.$$

Throughout the proof, ε can be any small constant (for example, $\varepsilon = 1/2$).

Lemma 2.6.1 (Eliminating high-degree vertices)**.** *Let j_0, ε be any positive constants with $\varepsilon < 1$. There exists a positive constant $j \geq j_0$ such that the following hold* whp.

1. *Conditioned on the degree sequence $(d_1, \ldots, d_{\bar{n}})$, the edges of $\bar{G}_n$ form a uniformly chosen perfect matching of its half-edges.*

2. *The number of vertices and the total degree in $\bar{G}_n$ (which is twice the number of edges of $\bar{G}_n$) satisfy*

$$\bar{n} \geq (1-\varepsilon)n \quad \textit{and} \quad d_1 + \cdots + d_{\bar{n}} \in (1-\varepsilon, 1+\varepsilon)\, nd.$$

3. *The branching rate $\bar{b}$ of the degree sequence of $\bar{G}_n$ satisfies $\bar{b} \in (1-\varepsilon, 1+\varepsilon)b$.*

4. *For all i, $0 \leq d_i \leq 2j$. The number of vertices with large degree is as expected*

$$\# \left\{ i \in \{1, \ldots, \bar{n}\} : d_i \in \left[\frac{j}{2}, 2j\right] \right\} \geq \varepsilon n \mu[j, 2j].$$

To simplify the notation, for the rest of this section 2.6, we define

$$\mathfrak{u}_j := \mu[j, 2j].$$

Proof. Since the proof of Items 1-3 is rather standard, we defer it to Section 2.7.5. Here, we only prove Item 4. Choose j large enough such that

$$\mathbb{E}_{D\sim\mu} D \mathbf{1}_{D \geq 2j+1} \leq \varepsilon/4.$$

For each vertex $v \in G_n$, consider the random variable

$$X_v := \deg_{G_n}(v) \mathbf{1}_{\deg_{G_n}(v) \geq 2j+1}.$$

These random variables are independent with mean at most $\varepsilon/4$ and variance bounded by

the second moment of μ. By Chebyshev's inequality, whp

$$\sum_{v \in G_n} X_v \leq \varepsilon n/2.$$

Thus, whp, the total number of removed half-edges from vertices of degree in $[0, 2j]$ in G_n is at most $\sum_{v \in G_n} X_v \leq \varepsilon n/2$.

By Chernoff inequality, whp, the number of half-edges of G_n of vertices of degree in $[0, j)$ and $[j, 2j]$ are $(1-\varepsilon, 1+\varepsilon)nd$ and $(1-\varepsilon, 2+2\varepsilon)jn\mathfrak{u}_j$, respectively. Since j is sufficiently large, $j\mathfrak{u}_j$ is very small compared to d. The first deleted half-edge has probability roughly $\frac{j\mathfrak{u}_j}{d}$ to be from vertices of degree in $[j, 2j]$. Ideally, one expects to delete at most $\frac{j\mathfrak{u}_j}{d}\varepsilon nd$ half-edges from these vertices. We will show that it is the case, namely,

Claim 2.6.2. *The number of half-edges deleted from vertices of degree in $[j, 2j]$ is at most*

$$10\varepsilon jn\mathfrak{u}_j \quad \textit{whp}.$$

Assuming the claim, the number of vertices originally with degree in G_n in $[j, 2j]$ and with degree less than $j/2$ in $\bar{G}_n$ is at most

$$\frac{10\varepsilon jn\mathfrak{u}_j}{j/2} = 20\varepsilon n\mathfrak{u}_j \quad \text{whp}.$$

By Chernoff inequality, in G_n, the number of vertices of degree in $[j/2, 2j]$ is in $(1-\varepsilon, 2+2\varepsilon)n\mathfrak{u}_j$ (where we choose j so that $\mu[j/2, j) \leq (1+\varepsilon)\mathfrak{u}_j$). Hence, whp, the number of vertices in $\bar{G}_n$ with degree in $[j/2, 2j]$ is in $(1-21\varepsilon, 2+2\varepsilon)n\mathfrak{u}_j$, completing the proof of Item 4. □

To prove Claim 2.6.2, we will use the following cut-off line algorithm to find the random matches of the deleted high-degree vertices in G_n.

Definition 2.6.3 (Cut-off line algorithm)**.** Given a graph G_n in which each vertex v has degree $d_{G_n}(v)$. 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.

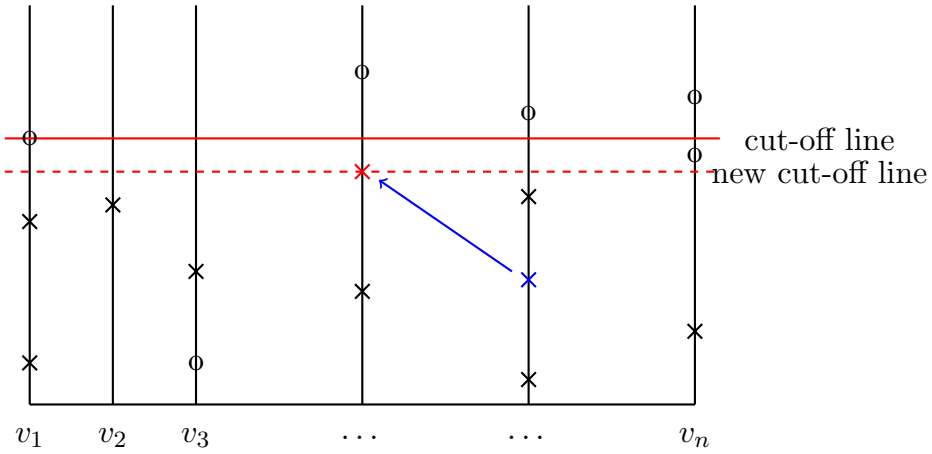


Figure 2.3: The circles 'o' represent matched half-edges and the crosses '×' represent unmatched half-edges. The blue half-edge is chosen and matched to the red half-edge which is the highest unmatched half-edge. Then the cut-off line is moved to the new cut-off line (dashed).

- Set the cut-off line 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 2.3 illustrates the algorithm.

Proof of Claim 2.6.2. For each half-edge of a vertex in G_n with degree at least $2j+1$, we choose their matches according to the above algorithm. Assume that at the end of this process of deleting half-edges of vertices with degree at least $2j+1$, the cut-off line is at height $h \in (0,1)$. Note that a half-edge of a vertex whose degree in G_n lies in $[j, 2j]$ is deleted if and only if it is above the cut-off line. We show that whp, $h \geq 1-4\varepsilon$. Indeed, Item 2 implies that the number of half-edges of G_n is at least $(1-\varepsilon)nd/2$ and in choosing the heights of these half-edges, the number of half-edges with heights above $1-4\varepsilon$ is at least $2\varepsilon(1-2\varepsilon)nd$ whp, by Chernoff inequality, which contradicts the event that we only delete at most εnd half-edges altogether. Hence the cut-off line is above $1-4\varepsilon$ whp. Since the number of half-edges with degree in G_n belonging to $[j, 2j]$ is at most $(1+\varepsilon)2jn\mathfrak{u}_j$ whp, the number of half-edges above $1-4\varepsilon$ is at most $8\varepsilon(1+\varepsilon)jn\mathfrak{u}_j \leq 10\varepsilon jn\mathfrak{u}_j$ as claimed. That proves Claim 2.6.2. □

Step 2. Exploration. Let W be the set of vertices in $\bar{G}_n$ whose degrees belong to $[j/2, 2j]$. We shall find the desired (α, R)-embedded expander W_0 inside W. In this step, we explore the R-neighborhoods of these high-degree vertices in W.

After having preprocessed the graph G_n to obtain $\bar{G}_n$, the remaining randomness is the perfect matching of the half-edges in $\bar{G}_n$. In this step, we condition on the preprocessing step and write the probability in terms of the randomness of the perfect matching in $\bar{G}_n$. We run the following exploration process to perform some matchings of the half-edges of $\bar{G}_n$. Let R and $\mathfrak{r}$ be some large (bounded) numbers to be chosen (they are chosen in (2.6.14)).

1. For each vertex $v \in W$, set $r_v = 1$.

2. If $r_v < R$ for all $v \in W$, explore the neighborhood $B_{\bar{G}_n}(v, r_v)$ simultaneously for all $v \in W$; noting that if $u \in B_{\bar{G}_n}(v, r_v) \cap B_{\bar{G}_n}(v', r_{v'})$, we stop exploring the branch starting at u. If $r_v \geq R$ for some $v \in W$, the process terminates. Otherwise, go to (3).

3. For each vertex $v \in W$, if $N(v, r_v)$ intersects at most $100\mathfrak{r}$ other balls $B_{\bar{G}_n}(v', r_{v'})$ $(v' \in W)$, set $r_v := r_v + 1$. Otherwise, keep r_v intact. If none of the r_v is increases in this step, the process terminates. Otherwise, go back to (2).

We show that when the exploration process terminates, the number of vertices $v \in W$ at which the process stops before reaching radius R is insignificant. For that, we choose R and $\mathfrak{r}$ so that the expected number of vertices of W that lie in a neighborhood $N(v, 2R)$ is small compared to $\mathfrak{r}$ (see (2.6.1)). And so, it is unlikely that the different neighborhoods $N(v, R)$ intersect frequently.

Lemma 2.6.4. *Let R and $\mathfrak{r}$ be positive numbers bounded by some constants and satisfying*

$$\frac{\bar{b}^{2R-1} j^2 \mathfrak{u}_j}{d} \leq \frac{\mathfrak{r}}{10}. \tag{2.6.1}$$

The number of $v \in W$ with $r_v = R$ is at least $0.98|W|$ whp.

Note that $\bar{b}$ is a random variable and so are R and $\mathfrak{r}$. Nevertheless, we will choose R and $\mathfrak{r}$ so that they are bounded.

Proof. For this proof, we write $N(v, r_v)$ for $B_{\bar{G}_n}(v, r_v)$ for simplicity. We first show that for each $v \in W$, it is likely that $r_v = R$; more specifically,

$$\mathbb{P}\left(r_v = R\right) \geq 0.99. \tag{2.6.2}$$

Indeed, if $r_v < R$ then $N(v, R)$ intersects more than $100\mathfrak{r}$ other balls $N(u, R)$ $(u \in W)$, which implies that $N(v, 2R)$ contains more than $100\mathfrak{r}$ elements of W. By Markov's inequality,

$$\mathbb{P}\left(r_v < R\right) \leq \frac{\mathbb{E}\left|N(v, 2R) \cap W\right|}{100\mathfrak{r}}.$$

By Items 2 and 4 of Lemma 2.6.1, we have

$$\mathbb{E}\left|N(v, 2R) \cap W\right| \leq 2j\bar{b}^{2R-1}\frac{2j(2+\varepsilon)n\mathfrak{u}_j}{(1-\varepsilon)nd} \leq \mathfrak{r}$$

where we used (2.6.1). Thus, we get (2.6.2).

It remains to show that whp, at least $0.98|W|$ vertices v satisfy $r_v = R$. We derive this from (2.6.2) and Azuma's inequality. Let X be the number of vertices $v \in W$ with $r_v = R$. By (2.6.2) and Item 4 of Lemma 2.6.1,

$$\mathbb{E}X \geq 0.99|W| \geq 0.99(1-\varepsilon)n\mathfrak{u}_j.$$

Enumerate the vertices of $\bar{G}_n$ by $u_1, \ldots, u_{\bar{n}}$. Let $\mathcal{F}_i$ be the σ-algebra generated by the matchings of the half-edges of vertices $u_1, \ldots, u_i$. We will apply the Azuma's inequality to the martingale $\mathbb{E}\left(X - \mathbb{E}X\middle|\mathcal{F}_i\right), i = 0, \ldots, \bar{n}$. Since the maximal degree in $\bar{G}_n$ is $2j$, we have that for every i,

$$\left|\mathbb{E}\left(X - \mathbb{E}X\middle|\mathcal{F}_{i+1}\right) - \mathbb{E}\left(X - \mathbb{E}X\middle|\mathcal{F}_i\right)\right| \leq 4j \max_{u \in \bar{G}_n}\left|N(u, 2R)\right| \leq 4j(2j)^{2R} \tag{2.6.3}$$

where in the first inequality, we observed that for any fixed matching of the half-edges of vertices $u_1, \ldots, u_i$ and any two different matchings of the half-edges of u_{i+1}, there exists a bijection between the extensions of these matchings into perfect matchings of $\bar{G}_n$ such that

the number of different matchings are at most $4j$. Using Azuma's inequality and the fact that j and R are constants, we obtain

$$\mathbb{P}\left(|X - \mathbb{E}X| \geq \varepsilon \mathbb{E}X\right) \leq \exp\left(-\frac{\varepsilon^2(\mathbb{E}X)^2}{2\bar{n}(4j(2j)^{2R})^2}\right) = \exp\left(-\Omega(n)\right). \tag{2.6.4}$$

Thus, whp, $X \geq (1-\varepsilon)\mathbb{E}X \geq 0.98|W|$ as stated. □

Step 3. Finding W_0. Consider a new graph G_n' with vertex set $V' \cup V''$ where V'' are the vertices of $\bar{G}_n$ that have not been touched in the exploration step 2 and each element of V' is a ball $N(v, r_v)$ whose half-edges are the unmatched half-edges of $N(v, r_v)$. If there is an unmatched half-edge that belongs to at least 2 balls, we choose one such ball at random and associate this half-edge to that ball. The remaining randomness is the uniform perfect matching of the half-edges in G_n'. We show that many vertices in V' have high degree. Note that $|V'| = |W|$.

Lemma 2.6.5 (V' has high degree)**.** *There exist positive constants $\varepsilon', \varepsilon''$ and R_0 depending only on μ such that for all bounded positive numbers $R_1, R, \mathfrak{r}$ satisfying*

$$\begin{aligned} R_0 \leq \min\{R_1, R - R_1\}, \quad & 800\mathfrak{r} \leq \varepsilon'^2(\bar{b}(1-\varepsilon''))^{R_1-1}j, \\ \frac{\bar{b}^{2R_1-1}j^2\mathfrak{u}_j}{d} \leq \frac{1}{10^4}, \quad & \frac{\bar{b}^{2R-1}j^2\mathfrak{u}_j}{d} \leq \frac{\mathfrak{r}}{10}, \end{aligned} \tag{2.6.5}$$

the number of vertices in V' with degree at least M is at least $\frac{\varepsilon'}{2}|V'|$ whp where

$$M = \frac{\varepsilon'^3(\bar{b}(1-\varepsilon''))^{R-1}j}{8}. \tag{2.6.6}$$

We note that in (2.6.5), the last two inequalities are for Lemma 2.6.4 to hold. The constant R_1 is mainly for technical reasons. The condition that $800\mathfrak{r} \leq \varepsilon'(\bar{b}(1-\varepsilon''))^{R_1-1}j$ is there so that when we ignore at most $100\mathfrak{r}$ possible common branches, the number of remaining branches is still significant.

Proof. First, we will show that

Claim 2.6.6. *For every $v \in W$, with probability at least ε', the number of half-edges on the*

boundary of $N(v,R)$ that do not belong to any of the balls $N(v',R), v' \in W \setminus \{v\}$ is at least M.

Indeed, let $u_1, \ldots, u_l$ be the vertices of distance R_1 from v. Since $\frac{\bar{b}^{2R_1-1}j^2\mathfrak{u}_j}{d} < \frac{1}{10^4}$, by Lemma 2.6.4, with probability at least 0.99, $N(v,R_1)$ does not intersect any other balls $N(v',R_1)$.

Since $\frac{\bar{b}^{2R-1}j^2\mathfrak{u}_j}{d} < \frac{\mathfrak{r}}{10}$, by Lemma 2.6.4, with probability at least 0.99, $N(v,R)$ intersects at most $100\mathfrak{r}$ other balls $N(v',R), v' \in W$. Consider the branches $B(u_h, R-R_1)$ consisting of vertices at distance r from v and distance $r-R_1$ from u_h for $R_1 \le r \le R$. Conditioned on $u_1, \ldots, u_l$, it suffices to show that with probability at least 0.9, for any choices of sets $A \subset [1,\ldots,l]$ with $|A| \le 100\mathfrak{r}$,

$$\left| \bigcup_{h \in [1,\ldots,l]\setminus A} \partial B(u_h, R-R_1) \right| \ge M. \tag{2.6.7}$$

Letting $Z_{R'}$ be the number of children in the R'-th generation of the corresponding size-biased Galton-Watson process, we have for any $\varepsilon'' \in (0,1)$, $\frac{Z_{R'}}{(\bar{b}(1-\varepsilon''))^{R'}}$ converges almost surely to some random variable Z (with Z not identically 0 and taking values in $[0,\infty]$) as $R' \to \infty$ (see for example, [AN72, pages 24-29]). For some sufficiently small constant ε' (that only depends on μ), we have

$$\mathbb{P}(Z \ge 4\varepsilon') \ge 4\varepsilon'.$$

Thus, for a sufficiently large R_0 and $R_0 \le R_1, R_0 \le R - R_1$, we have

$$\mathbb{P}\left(l \ge \varepsilon'(\bar{b}(1-\varepsilon''))^{R_1-1}j\right) \ge \varepsilon'$$

and

$$\mathbb{P}\left(|\partial B(u_h, R-R_1)| \ge \varepsilon'(\bar{b}(1-\varepsilon''))^{R-R_1}\right) \ge \varepsilon'.$$

We can choose ε'' so small that $\bar{b}(1-\varepsilon'') > 1$. Under the event that $l \ge \varepsilon'(\bar{b}(1-\varepsilon''))^{R_1-1}j$, let X_h $(h = 1,\ldots,l)$ be the indicator of the event that $|\partial B(u_h, R-R_1)| \ge \varepsilon'(\bar{b}(1-\varepsilon''))^{R-R_1}$. When R_0 is sufficiently large, l is also large. Thus, with probability at least 0.9, at least

$\varepsilon' l/4$ indices $h \in [1, \dots, l]$ have $X_h = 1$. Under this event, since $100\mathfrak{r} \le \varepsilon' l/8$ by (2.6.5), for any choices of sets $A \subset [1, \dots, l]$ with $|A| \le 100\mathfrak{r}$, there are at least $\varepsilon' l/8$ indices $h \notin A$ with $X_h = 1$, which implies

$$\sum_{h \in [1,\dots,l] \setminus A} |\partial B(u_h, R - R_1)| \ge \varepsilon'^2 (\bar{b}(1-\varepsilon''))^{R-R_1} l/8.$$

Since $j, \bar{b}$ and R are bounded by some constant, for each v, the boundaries $\partial B(u_h, R - R_1)$ are disjoint with probability at least 0.9, proving (2.6.7) and Claim 2.6.6.

Next, by using the Azuma's inequality the same way that we used in proving Lemma 2.6.4, we obtain that whp, there are at least $(1-\varepsilon)\varepsilon'|W|$ vertices v satisfying the event in Claim 2.6.6 simultaneously. Combining this with Lemma 2.6.4 completes the proof of Lemma 2.6.5. □

Note that the (random) edges of G_n' form a uniformly chosen perfect matching of its half-edges. On the half-edges of G_n', consider the following coloring scheme on the half-edges:

- For each vertex v with at least M half-edges in G_n', choose exactly M half-edges among them uniformly at random and color them blue;
- Perform the uniform random matching among all half-edges in G_n', and let W_1' be the induced subgraph on vertices with degree at least M;
- Let $\mathfrak{K}$ be the number of blue edges, formed by two blue half-edges, and for each $v \in W_1'$, let $\deg_b(v)$ be the number of blue edges adjacent to v.

Conditioned on $\mathfrak{K}$, we see that the distribution of $\{\deg_b(v)\}_{v \in W_1'}$ is given by

$$\{B_v\}_{v \in W_1'}, \quad B_v \sim \text{i.i.d. } \mathbf{Bin}(M, \theta) \text{ conditioned on } \sum_{v \in W_1'} B_v = 2\mathfrak{K}.$$

Note that due to the conditioning on the sum of $\{B_v\}$, their distribution is well-defined

regardless of the specific value of θ. However, for explicitness, we let

$$\theta := \frac{2\mathfrak{K}}{M|W_1'|}.$$

Lemma 2.6.7. *Let ε' be as in Lemma 2.6.5. With high probability,*

$$\theta \geq \frac{\varepsilon' \mathfrak{u}_j M}{60d}. \tag{2.6.8}$$

Proof. By Item 4 of Lemma 2.6.1, Lemmas 2.6.4 and 2.6.5, the number of vertices in W_1' is at least $\frac{\varepsilon'|W|}{2} \geq \frac{\varepsilon' n \mathfrak{u}_j}{2}$.

Thus, the total number of blue half-edges in W_1' is at least $\frac{\varepsilon' n \mathfrak{u}_j M}{2}$ whp while the total degree in G_n' is at most $2nd$ by Item 2 of Lemma 2.6.1. Let

$$\theta_0 = \frac{\#\{\text{number of blue half-edges in } W_1'\}}{\#\{\text{number of half-edges in } G_n'\}} =: \frac{d_{W_1'}}{d_{G_n'}} \geq \frac{\varepsilon' \mathfrak{u}_j M}{4d}.$$

Since $d_{W_1'} = M|W_1'|$, it suffices to show that whp, $\mathfrak{K} \geq \theta_0 d_{W_1'}/12$. Indeed, splitting the set of blue half-edges of W_1' into two parts of equal size A and B (independent of their heights). We perform the cut-off line algorithm 2.6.3 to match the half-edges of A first. Since at least $d_{W_1'}/4 = \theta_0 d_{G_n'}/4$ highest half-edges in G_n' have been matched during this step, the cut-off line is below $1 - \frac{\theta_0}{5}$ whp (otherwise, by Chernoff inequality, the number of half-edges above the cut-off line is at most $\theta_0 d_{G_n'}/4$). By Chernoff inequality again, the number of half-edges of B that lie above $1 - \theta_0/5$ is at least $\theta_0 d_{W_1'}/12$ whp. Since all edges between A and B are inside W_1', $\mathfrak{K} \geq \theta_0 d_{W_1'}/12$ whp. □

Lemma 2.6.7 allows to find a high-degree core of W_1'.

Lemma 2.6.8. *Let M and θ be as in Lemmas 2.6.5 and 2.6.7. Assume that $\theta M \geq 100$. Then whp, W_1' contains a subgraph W_0' with the following properties.*

- *The number of vertices in W_0' is at least $|W_1'|/2$.*
- *Each vertex in W_0' has degree at least $\frac{\theta M}{20}$ inside W_0'; in other words, W_0' is an $\frac{\theta M}{20}$-core.*
- *Each vertex in W_0' has degree at most $(2j)^R$.*

Proof. Let $s = \frac{\theta M}{20}$. The last property follows from the fact that the maximum degree of vertices in V' is $(2j)^R$. For the rest of this proof, we only look at the blue half-edges in W_1' that are matched to another blue half-edge in W_1'. To find W_0', we use the cut-off line algorithm 2.6.3 to find a uniform perfect matching of these half-edges of W_1' as follows. Each of these half-edges is re-assigned a height uniformly chosen in $[0, 1]$. If there is a vertex in W_1' with less than s unmatched half-edges (equivalently, less than s half-edges below the cut-off line), match its half-edges to the highest unmatched half-edges and move the cut-off line accordingly. Remove this vertex. Repeat this step until there are no such vertices left.

Let W_0' be the set of remaining vertices. It remains to show that $|W_0'| \geq \frac{|W_1'|}{2}$. We show that after the removal, the cut-off line is above $2/3$ whp. Assuming this, we have

$$|W_0'| \geq |W_1'| - N'$$

where N' is the number of vertices in W_1' having less than s half-edges below the line $2/3$. By Lemma 2.6.7, the number of internal half-edges of each vertex $v \in W_1'$ has degree distribution $\mathbf{Bin}(M, \theta)$. Thus, the distribution of the number of its half-edges that lie below the line $2/3$ is $\mathbf{Bin}\left(M, \frac{2\theta}{3}\right)$. So, the probability that v has less than s half-edges below the line $2/3$ is at most

$$\mathbb{P}\left(\mathbf{Bin}\left(M, \frac{2\theta}{3}\right) \leq \frac{\theta M}{20}\right) \leq \exp\left(-\frac{\theta M}{12}\right) \leq \frac{1}{200}.$$

where we used the Chernoff inequality and the assumption $\theta M \geq 100$. By Chernoff inequality, we have whp, $N' \leq \frac{|W_1'|}{100}$. And so, $|W_0'| \geq \frac{|W_1'|}{2}$ as desired.

Now, we prove that after the removal process above, the cut-off line is above $2/3$ whp. Let a be the number of removed vertices ($0 \leq a \leq |W_1'|$). The total number of matched half-edges is at most $2as$ because each time we remove a vertex, at most $2s$ half-edges are matched. Thus, the total number of half-edges above the cut-off line is at most $2as \leq 2s|W_1'|$.

On the other hand, given a vertex v with degree distribution $\mathbf{Bin}(M, \theta)$, the distribution of the number of its half-edges that lie above the line $2/3$ is $\mathbf{Bin}\left(M, \frac{\theta}{3}\right)$. By Chernoff inequality, the number of half-edges of W_1' above the line $2/3$ is at least $M\theta|W_1'|/4 =$

$5s|W_1'| > 2s|W_1'|$. Thus, whp, the cut-off line is above $2/3$, completing the proof of Lemma 2.6.8. □

Next, we show that W_0' is an embedded expander.

Lemma 2.6.9. *Let M and θ be as in Lemmas 2.6.5 and 2.6.7. Assume that $\theta M \geq 100$. There exists a constant $\alpha > 0$ such that whp, W_0' is an $(\alpha, 1)$-embedded expander.*

Proof. Let $N = |W_0'|$, $s = \frac{\theta M}{20}$. Note that $N = \Theta(n)$ and $s \geq 5$ is a constant. It suffices to show that for any subset A of the vertex set of W_0' of size αN, the size of $N(A, 1)$ is at least twice that of A. In other words, whp, for every $m \leq \alpha N$ and subsets of vertices A, B with $|A| = m, |B| = 2m$, the neighbors of A are not contained fully in B. Fix two sets A and B with $|A| = m, |B| = 2m$. By Lemma 2.6.8, the number of half-edges in A is at least ms, the number of half-edges in B is $b \leq (2j)^R 2m$ and the total number of half-edges in W_0' is $c \geq Ns$. The probability that all the neighbors of A belong to B is at most

$$\frac{b}{c-1}\frac{b-1}{c-3}\cdots\frac{b-ms+1}{c-2ms+1} \leq \left(\frac{b}{Ns-2ms+1}\right)^{ms} \leq \left(\frac{4(2j)^R m}{Ns}\right)^{ms}.$$

Taking the union bound over m and choices of A, B, we get that the probability that W_0' is not an $(\alpha, 1)$-embedded expander is at most

$$\sum_{m=1}^{\alpha N} \binom{N}{2m}^2 \left(\frac{4(2j)^R m}{Ns}\right)^{ms} \leq \sum_{m=1}^{\log n} \frac{C\log^{s-4} n}{N^{s-4}} + \sum_{m=\log n}^{\alpha N} \left(C\alpha^{s-4}\right)^m \tag{2.6.9}$$

for some constant C depending only on j, R and s. Choosing $\alpha \leq \frac{1}{2C}$ makes the r.h.s. of (2.6.9) of order $o(1)$. This completes the proof of Lemma 2.6.9. □

Proof of Lemma 2.5.1. Since W_0' is a subset of V', each of its vertices is a ball $N(v, r_v)$ in $\bar{G}_n$, a subgraph of G_n. Let W_0 be the collection of all such centers v. Clearly, $|W_0| = |W_0'| = \Theta(n)$ and W_0 is an $(\alpha, 2R+1)$-embedded expander of G_n. To finish the proof of Lemma 2.5.1 for subexponential degree distributions, it remains to show that there exists a choice of $j, R, \mathfrak{r}$ satisfying the assumptions of the previous Lemmas (in particular, Lemmas 2.6.4, 2.6.5, 2.6.8 and 2.6.9). In other words, for any given positive constants R_0, δ and ε'

(δ and ε' can be arbitrarily small and R_0 can be arbitrarily large), we show that there exist a constant j and random variables $R, R_1, \mathfrak{r}$ such that the following conditions holds:

$$R \le \delta j, \tag{2.6.10}$$

$$R_0 \le \min\{R_1, R - R_1\}, \tag{2.6.11}$$

$$\frac{\bar{b}^{2R_1-1} j^2 \mathfrak{u}_j}{d} \le \frac{1}{10^4}, \quad 800\mathfrak{r} \le \varepsilon'^2 (\bar{b}(1-\varepsilon''))^{R_1-1} j, \quad \frac{\bar{b}^{2R-1} j^2 \mathfrak{u}_j}{d} \le \frac{\mathfrak{r}}{10}, \tag{2.6.12}$$

$$\frac{\varepsilon'^7 (\bar{b}(1-\varepsilon''))^{2R-2} j^2 \mathfrak{u}_j}{64 \cdot 60 d} \ge 100. \tag{2.6.13}$$

Note that (2.6.10) comes merely from the statement of Lemma 2.4.2. For given j and $\bar{b}$, we define $\mathfrak{r}$, R and R_1 by the following equations so that (2.6.12) holds automatically

$$\frac{\bar{b}^{2R_1-1} j^2 \mathfrak{u}_j}{d} = \frac{1}{10^4}, \quad 800\mathfrak{r} = \varepsilon'^2 (\bar{b}(1-\varepsilon''))^{R_1-1} j, \quad \frac{\bar{b}^{2R-1} j^2 \mathfrak{u}_j}{d} = \frac{\mathfrak{r}}{10}. \tag{2.6.14}$$

Since μ has finite second moment, $\lim_{j\to\infty} j^2 \mathfrak{u}_j = 0$ and so (2.6.11) holds when j is sufficiently large. Since μ is subexponential, $\mathfrak{u}_j \ge e^{-\varepsilon_0 j}$ for any constant $\varepsilon_0 > 0$ and for large j. By choosing ε_0 to be small compared to δ, (2.6.10) holds. The inequality (2.6.13) holds automatically. This completes the proof of Lemma 2.5.1 for subexponential distributions μ. □

Remark 2.6.10. Observe that the above subgraph W_0 is also an $(\alpha, 2R+1)$-embedded expander of the graph $\bar{G}_n$ with $\deg_{\bar{G}_n} w \ge j/2$ for all $w \in W_0$.

2.6.2 The embedded expander property in the exponential case

We continue our discussion and verify Lemma 2.4.2.

In Section 2.6.1, we used the assumption that μ is subexponential only in the last step of choosing the parameters R, R_1 and $\mathfrak{r}$. In particular, we used this assumption to obtain that j can be an arbitrarily large constant while $\mathfrak{u}_j := \mu[j, 2j]$ remains positive and $\mathfrak{u}_j \ge e^{-\varepsilon_0 j}$. The inequality $\mathfrak{u}_j \ge e^{-\varepsilon_0 j}$ is only used to show that with the choice of parameters $R, R_1, \mathfrak{r}$ as in (2.6.14), (2.6.10) holds as stated in Lemma 2.5.1. Here, when μ has an exponential

tail, Lemma 2.4.2 does not assert that $R \leq \delta j$ and so there is no need for $\mathfrak{u}_j \geq e^{-\varepsilon_0 j}$.

Thus, if μ has an infinite support, one can still find an arbitrarily large constant j for which $\mathfrak{u}_j > 0$. This is therefore enough for the rest of the proof of Section 2.6.1 to follow, proving Lemma 2.4.2 for such μ.

If the support of μ is finite, let j be the largest integer in the support of μ. Let $\mathfrak{u}_j$ be an arbitrarily small constant with $\mathfrak{u}_j \leq \mu[j, 2j] = \mu(j)$ ($\mathfrak{u}_j$ could be much smaller than $\mu(j)$). Since the degrees in G_n are already bounded, there is no need to run the Step 1 of preprocessing the graph as in Section 2.6.1. So, for this case, $\bar{G}_n = G_n, \bar{n} = n, \bar{b} = b$ and so on. For the exploration, Step 2, we choose W by assigning each vertex in G_n of degree j to W independently with probability $\frac{\mathfrak{u}_j}{\mu(j)}$. The rest of the proof follows without any changes. In (2.6.14), we choose $\mathfrak{u}_j$ to be sufficiently small so that (2.6.11) holds and so does (2.6.13). □

2.7 Proof of technical lemmas

To conclude Chapter 2, we present the remaining proofs of some technical lemmas develop through this chapter.

2.7.1 Proof of Lemma 2.3.3

Here we prove Lemma 2.3.3, which is based on an elementary analysis of large deviation events.

Proof of Lemma 2.3.3. The first statement follows directly from the Cauchy-Schwarz inequality: If we have $\mathbb{E}_{D\sim\mu} \exp(3\varepsilon D) < \infty$ for some $\varepsilon > 0$, then

$$\left(\sum_k e^{\varepsilon k}\sqrt{p_k}\right)^2 \leq \left(\sum_k e^{3\varepsilon k} p_k\right)\left(\sum_k e^{-\varepsilon k}\right) < \infty.$$

For the second statement, let n be a given large enough integer, and define

$$k_n = \min\left\{k : \sum_{j\geq k} p_k \leq \frac{1}{n\log\log n}\right\}.$$

Let D_i for $i = 1, \dots n$ be i.i.d samples from μ. We start by studying the empirical distribution of the D_i. First, by a simple union bound, the definition of k_n implies that

$$\mathbb{P}\left(\exists i \in [n] : D_i \geq k_n\right) \leq \frac{1}{\log\log n} = o(1).$$

Moreover, since μ has an exponential tail, there exists a constant $C > 0$ depending on μ such that $k_n \leq C \log n$. Recall the definition of $k_0 = \max\{k : \sum_{j \geq k} \sqrt{p_k} \geq 1/2\}$. Our next goal is to show that with high probability, the number of i such that $D_i = k$ is at most $\frac{1}{2} n \sqrt{p_k}$ for all $k_0 \leq k \leq k_n$. We consider two possible cases of k as follows:

1. For k such that $p_k \leq (n \log^2 n)^{-1}$, Markov's inequality implies that

$$\mathbb{P}(|\{i : D_i = k\}| \geq 1) \leq (\log^2 n)^{-1}. \tag{2.7.1}$$

2. For k such that $p_k \geq (n \log^2 n)^{-1}$, we use the following large deviation estimate for binomials (Corollary 22.9 of [FK15]): for $c > 1$,

$$\mathbb{P}(\mathbf{Bin}(n, p) \geq cnp) \leq \exp\{-np(c \log c + 1 - c)\}.$$

This gives that

$$\begin{aligned} &\mathbb{P}\left(|\{i : D_i = k\}| \geq \frac{1}{2} n \sqrt{p_k}\right) \\ &\leq \exp\left(\frac{1}{2} n \sqrt{p_k}(1 - 2\sqrt{p_k} + \log(2\sqrt{p_k}))\right) \leq \exp\left(-n^{1/3}\right). \end{aligned} \tag{2.7.2}$$

Since $k_n \leq C \log n$, applying a union bound on (2.7.1, 2.7.2) tells us that

$$\mathbb{P}\left(\exists\, k_0 \leq k \leq k_n : |\{i : D_i = k\}| \geq \frac{1}{2} n \sqrt{p_k}\right) = o(1). \tag{2.7.3}$$

When μ satisfies $k_0 < k_{\max}$, (2.7.3) implies that the empirical distribution of $\{D_i : i \in [n]\}$ is stochastically dominated by $\mu^\sharp$, since $\mu^\sharp(k) \geq \sqrt{p_k}$ by the definition of k_0. On the other hand, if $k_0 = k_{\max}$, the stochastic domination becomes trivial because we only augment the weight of $k_{\max}$ in μ. Since taking out any $n/3$ entries from $[n]$ can only increase each

probability mass of the empirical distribution of $\{D_i : i \in [n]\}$ by a factor of $3/2$, with high probability we have for each $k_0 \le k$,

$$|\{i : D_i = k\}| \le \frac{3}{4} n \sqrt{p_k},$$

and hence we conclude the second statement of Lemma 2.3.3. □

2.7.2 Local neighborhoods contain at most one cycle

In this subsection, we prove Lemma 2.3.5. We use Lemma 2.3.3 to bound the probability of $N(v, c \log n)$ having at least two cycles.

Proof of Lemma 2.3.5. Let v be an arbitrary vertex in $G \sim \mathcal{G}(n, \mu)$, fixed before we explore the matchings of half-edges. We again study the local neighborhood $N(v, L)$ by exploration process, particularly in terms of the breadth-first search perspective. We start exploring from the single vertex v, and at time s we explore all the vertices of distance s from v, based on what we explored until time $s-1$. Let V_s be the collection of vertices explored at time s, and set $X_s = |V_s|$.

We will bound the probability of discovering at least two cycles during the exploration process until depth $L = c \log n$ (c will be determined later). Let $\mu^\sharp$ be the augmented distribution (Definition 2.3.2) and let $\widetilde{\mu}^\sharp := (\mu^\sharp)'_{[1,\infty)}$ denote its size-biased distribution conditioned on being inside the interval $[1, \infty)$. Also, let $\mathcal{T} \sim \mathsf{GW}(\mu^\sharp, \widetilde{\mu}^\sharp)^L$ and let Y_s be the number of vertices in $\mathcal{T}$ at depth s. Then, Lemma 2.3.3 implies that there exists a coupling between $(X_s)_{s \le L}$ and $(Y_s)_{s \le L}$ in such a way that $X_s \le Y_s$ for all $s \le L$, as long as $\sum_{s=0}^{L} X_s \le n/3$. Define B to be the event that $\sum_{s=0}^{L} X_s \le n/3$. On B, we clearly have

$$\sum_{v \notin \cup_{s \le L} V_s} \deg(v) \ge \frac{2n}{3}.$$

Assume that when moving from V_s to V_{s+1}, we pair the half-edges adjacent to V_s one by one. Let H_s be the number of unpaired half-edges adjacent to the vertices in V_s. For $s \le L$ and $1 \le i \le H_s$, let $\mathcal{F}_{s,i}$ be the σ-algebra generated by the exploration process until pairing the $(i-1)$-th half-edge. Set $W_{s,i}$ to be the collection of unpaired half-edges at that

moment. Further, let $A_{s,i}$ be the event that the i-th half-edge adjacent to V_s is paired with a half-edge in $W_{s,i}$. Then clearly,

$$\mathbb{P}(A_{s,i}|\mathcal{F}_{s,i}, B) \leq \frac{3|W_{s,i}|}{2n}.$$

We bound the size of $W_{s,i}$ based on the following observations:

1. Since the exploration of half-edges in H_{s-1} for $s \geq 1$ is done independently step by step, we can stochastically dominate $|H_s|$ by i.i.d $\zeta_j \sim \widetilde{\mu}^{\sharp}$ as

$$|H_s| \ \leq_{st} \ \sum_{j=1}^{Y_s} \zeta_j \ \stackrel{d}{=} \ Y_{s+1}.$$

2. Since $\zeta \sim \widetilde{\mu}^{\sharp}$ satisfies $\zeta \geq 1$, we can bound $|W_{s,i}|$ similarly by

$$|W_{s,i}| \ \leq_{st} \ \sum_{j=1}^{H_s} \zeta_j \ \leq_{st} \ \sum_{j=1}^{Y_{s+1}} \zeta_j \ \stackrel{d}{=} \ Y_{s+2}.$$

Therefore, combining above argument gives that

$$\sum_{j=1}^{H_s} \mathbf{1}_{A_{s,j}} \ \leq_{st} \ \mathbf{Bin}\left(Y_{s+1}, \frac{3Y_{s+2}}{2n}\right),$$

where the l.h.s is conditioned on the event B. Notice that the l.h.s of the above inequality stochastically dominates the number of cycles formulated during the exploration of depth from s to $s+1$. Hence the number of cycles in $N(v, L)$ conditioned on B is stochastically dominated by

$$\mathbf{Bin}\left(\sum_{s=1}^{L+1} Y_s, \frac{3Y_{L+2}}{2n}\right). \tag{2.7.4}$$

Now we bound the size of Y_s to conclude our argument. Let $\zeta \sim \mu^{\sharp}$ and $\zeta' \sim \widetilde{\mu}^{\sharp}$, and let ε, M be the constants that satisfy $\max\{\mathbb{E}e^{\varepsilon\zeta}, \mathbb{E}e^{\varepsilon\zeta'}\} \leq M$. Set K to be a large constant such that $M^{1/K} = e^{\varepsilon}$. Then we observe that Y_s/K^s has an exponential tail for all s, since

$$\begin{aligned}\mathbb{E}\left[\exp\left(\varepsilon\frac{Y_s}{K^s}\right)\right] &= \mathbb{E}\left[\mathbb{E}_{\zeta'}\left[\exp\left(\varepsilon\frac{\zeta'}{K^s}\right)\right]^{Y_{s-1}}\right] \\ &\leq \mathbb{E}\left[M^{Y_{s-1}/K^s}\right] = \mathbb{E}\left[\exp\left(\varepsilon\frac{Y_{s-1}}{K^{s-1}}\right)\right],\end{aligned} \tag{2.7.5}$$

where the inequality is due to Jensen's inequality. Iterating this $(s-1)$-times gives that the l.h.s is bounded by e^{ε}. Set $c > 0$ to be the constant satisfying $K^{c\log n+3} = n^{1/6}$. Based on the above observation, we bound the quantity (2.7.4) as follows.

$$\begin{aligned}\mathbb{P}\left(\mathbf{Bin}\left(\sum_{s=1}^{L+1} Y_s, \frac{3Y_{L+2}}{2n}\right) \geq 2\right) \;\leq\; &\mathbb{P}\left(\mathbf{Bin}\left(n^{1/5}, \frac{3}{2}n^{-4/5}\right) \geq 2\right) \\ &+\; \mathbb{P}\left(\sum_{s=1}^{L+2} Y_s \geq n^{1/5}\right).\end{aligned} \tag{2.7.6}$$

It is easy to see that the first term in the r.h.s is bounded by $o(n^{-1})$. The second term can be bounded using (2.7.5). Namely,

$$\begin{aligned}\mathbb{P}\left(\sum_{s=1}^{L+2} Y_s \geq n^{1/5}\right) &\leq e^{-\varepsilon n^{1/30}}\mathbb{E}\left[\exp\left(\varepsilon K^{-(L+3)}\sum_{s=1}^{L+2} Y_s\right)\right] \\ &\leq e^{-\varepsilon n^{1/30}}\mathbb{E}\left[\exp\left(\varepsilon K^{-(L+3)}\sum_{s=1}^{L+1} Y_s\right)\cdot\exp\left(\varepsilon K^{-(L+2)}Y_{L+1}\right)\right].\end{aligned}$$

Iterating this $(L+1)$ more times, we obtain

$$\begin{aligned}\mathbb{P}\left(\sum_{s=1}^{L+2} Y_s \geq n^{1/5}\right) &\leq e^{-\varepsilon n^{1/30}}\mathbb{E}\left[\exp\left(\varepsilon\left(\sum_{s=1}^{L+2} K^{-s}\right)\right)\right] \\ &\leq \exp\left(\varepsilon(1-n^{1/30})\right) = o(n^{-1}),\end{aligned}$$

as long as $K \geq 2$. Applying our estimates to (2.7.6), we conclude the desired result. □

2.7.3 Survival time on cyclic graphs

In this subsection, we present the proof of Proposition 2.3.9. It follows from the same technique as Lemma 2.3.11.

Proof of Proposition 2.3.9. Let $s \geq 2$ and L be any integers, and we build up an inductive argument starting from $h = 0$.

Let $\mathcal{S}_0 \sim \mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; 0, s)^L$, and $\rho^+ \in \mathcal{S}_0^+$ be the parent of ρ as before. Define $\mathbf{S}_{0,s,L}$ to be the first time when $(X_t) \sim \mathsf{CP}^\lambda_{\rho^+}(\mathcal{S}_0^+; \mathbf{1}_\rho)$ reaches state $\mathbf{0}$. Similarly as in Lemmas 2.2.3 and 2.3.11, we consider $(\widetilde{X}_t) \sim \widetilde{\mathsf{CP}}^\lambda_{\rho^+;\rho}(\mathcal{S}_0^+; \mathbf{1}_\rho)$, which is coupled with (X_t) in such a way that they share the same infection and recovery clocks, except that in $(\widetilde{X}_t)$, the recovery at ρ is ignored if at that time there exists an infected vertex other than ρ and ρ^+. Letting $D \sim \mu^\sharp$ be $D + 3 = \deg(\rho; \mathcal{S}_0^+)$, $\mathcal{T}_{u_1}, \ldots, \mathcal{T}_{u_D}$ be the i.i.d. $\mathsf{GW}(\widetilde{\mu}^\sharp)^{L-1}$ subtrees from the children of ρ and $\mathcal{S}'$ be the $\mathsf{GWC}(\widetilde{\mu}^\sharp; s)^L$ process that also hangs at ρ, we obtain the following by repeating the same argument in Lemma 2.3.11.

$$\mathbb{E}[\mathbf{S}_{0,s,L}|D] \leq (1 + \lambda\mathbb{E}[\mathbf{S}_{L-1}])^D(1 + 2\lambda\mathbb{E}[\mathbf{S}_{s,L}]), \tag{2.7.7}$$

where $\mathbf{S}_{L-1}$, $\mathbf{S}_{s,L}$ are as in the statements of Lemmas 2.2.3 and 2.3.11, respectively.

For general $h \neq 0$, we first develop the same argument in terms of $\mathcal{S}'_h \sim \mathsf{EGW}(\widetilde{\mu}^\sharp; h, s)^L$. Let $\mathbf{S}'_{h,s,L}$ be the first time when $\mathsf{CP}^\lambda_{\rho^+}((\mathcal{S}'_h)^+; \mathbf{1}_\rho)$ reaches at $\mathbf{0}$. For $D' \sim \widetilde{\mu}^\sharp$ denoting $D' + 1 = \deg(\rho; (\mathcal{S}'_h)^+)$, the subgraphs of descendents from the children of ρ consist of D' i.i.d $\mathsf{EGW}(\widetilde{\mu}^\sharp; h - 1, s)^L$ processes. Therefore, repeating the previous reasoning gives that

$$\mathbb{E}[\mathbf{S}'_{h,s,L}|D'] \leq (1 + \lambda\mathbb{E}[\mathbf{S}'_{h-1,s,L}])^{D'}. \tag{2.7.8}$$

Finally, the subgraphs of descendents (from children of ρ) of $\mathcal{S}_h \sim \mathsf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, s)^L$ for $h \geq 1$ with $\deg(\rho) = D$ consist of D i.i.d. $\mathsf{EGW}(\widetilde{\mu}^\sharp; h - 1, s)^L$ processes. Therefore, we deduce that $\mathbf{S}_{h,s,L}$, the first time when $\mathsf{CP}^\lambda_{\rho^+}(\mathcal{S}_h; \mathbf{1}_\rho)$ reaches $\mathbf{0}$, satisfies

$$\mathbb{E}[\mathbf{S}_{h,s,L}|D] \leq (1 + \lambda\mathbb{E}[\mathbf{S}'_{h-1,s,L}])^D. \tag{2.7.9}$$

Here, the law of D follows the conditional distribution of $\mu^\sharp$ being inside the interval $[1,\infty)$.

Combining the three equations (2.7.7), (2.7.8) and (2.7.9), we obtain that there exists a constant λ_0 such that for all $\lambda \le \lambda_0$, h, s and L,

$$\mathbb{E}[\mathbf{S}_{h,s,L}] \le 2e,$$

by manipulating the constants in the same way as Lemma 2.3.11. Then, the standard coupling between contact processes tells us that $\mathbf{R}_{h,s,L} \le_{st} \mathbf{S}_{h,s,L}$, concluding the proof. □

2.7.4 Survival time of the delayed process on cyclic graphs

Here, we discuss the proof of Lemma 2.3.12. We first prove the corresponding result for **GWC**-processes. Let $\widetilde{\mu}^\sharp$ be as in Section 2.3. Namely, $\widetilde{\mu}^\sharp$ is the augmented distribution of $\mu'_{[1,\infty)}$, where $\mu'_{[1,\infty)}$ is the size-biased distribution of μ conditioned on being in $[1,\infty)$.

Lemma 2.7.1. *Let $\mathcal{S} \sim \mathbf{GWC}(\widetilde{\mu}^\sharp; s)^L$, and $\nu^\theta_{\mathcal{S}}$ be the stationary distribution of $\mathbf{DP}^{\lambda,\theta}_\rho(\mathcal{S})$ on the space $\{0,1\}^{\mathcal{S}\setminus\{\rho\}}$, which is the delayed contact process on $\mathcal{S}$ with ρ set to be infected permanently. Then there exist constants $C, \lambda_0 > 0$ depending only on μ such that for all $\lambda \le \lambda_0$ and s, L with $s \ge 2$, we have $\mathbb{E}[\nu^\theta_{\mathcal{S}}(\mathbf{0})^{-1}] \le 2$ for $\theta = C\lambda$.*

Proof. Let v, v' be the two neighbors of ρ in $\mathcal{S}$. Let $\mathbf{S}^\theta_v$ denote the first time when $\mathbf{DP}^{\lambda,\theta}_\rho(\mathcal{S}; \mathbf{1}_v)$ reaches $\mathbf{0}$, and define $\mathbf{S}^\theta_{v'}$ analogously. Then, we set $\mathbf{S}^\theta_L = \frac{1}{2}(\mathbf{S}^\theta_v + \mathbf{S}^\theta_{v'})$.

As before, we build up an inductive argument on s. The case $s = 2$ is essentially the same as Proposition 2.2.6, since $\mathcal{S} \sim \mathbf{GWC}(\widetilde{\mu}^\sharp; 2)^L$ can be thought of as $\mathcal{T}^+_{\le L}$ with $\mathcal{T}_{\le L} \sim \mathbf{GW}(\widetilde{\mu}^\sharp)^L$, where ρ^+ in $\mathcal{T}^+_{\le L}$ is connected with ρ by a double-edge. Thus, the same proof of Proposition 2.2.6 can be applied, and we leave the details to the reader.

The general case $s \ge 3$ is also similar to the previous arguments of Propositions 2.3.9 and 2.2.6, but there is a subtle difference in comparing the stationary distributions, which makes the current case more technical. As before, we start with introducing a modified process as follows.

Let v be a neighbor of ρ in $\mathcal{S}$, and $(X_t) \sim \mathbf{DP}^{\lambda,\theta}_\rho(\mathcal{S}; \mathbf{1}_v)$. Define $(\widetilde{X}_t) \sim \widetilde{\mathbf{DP}}^{\lambda,\theta}_{\rho;v}(\mathcal{S}; \mathbf{1}_v)$ as

1. $(\widetilde{X}_t)$ has the same infection and recovery clocks at (X_t).

2. In $(\widetilde{X}_t)$, any recovery attempt at v is ignored if there exists an infected vertex other than ρ and v at that moment.

If $\{\rho, v\}$ infects a (random) neighbor U before v becomes healthy, then $(\widetilde{X}_t)$ behaves as $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\mathcal{S}; \mathbf{1}_U)$ (meaning that we fix both ρ, v to be infected forever), until $\widetilde{X}_t$ comes back to $\mathbf{1}_v$. Let $\widetilde{\mathbf{S}}^{\theta}_{s,L}$ denote the first time when $\widetilde{\mathbf{DP}}^{\lambda,\theta}_{\rho;v}(\mathcal{S}; \mathbf{1}_v)$ becomes $\mathbf{0}$, and $\widetilde{\mathbf{S}}^{\theta}$ be the first time it takes for $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\mathcal{S}; \mathbf{0})$ to return to $\mathbf{0}$ after infecting a (random) vertex U other than ρ and v. Setting $D \sim \widetilde{\mu}^{\sharp}$ to be $\deg(v) = D + 2$, the same reasoning as (2.2.1, 2.3.1) implies that

$$\begin{aligned}\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta}_{s,L} \,\middle|\, \mathcal{S}\right] &= \sum_{k=0}^{\infty} \left(\frac{(D+2)\lambda}{1+(D+2)\lambda}\right)^k \frac{1}{1+(D+2)\lambda} \\ &\quad \times \left[\frac{k+1}{\theta(1+(D+2)\lambda)} + k\,\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta} \,\middle|\, \mathcal{S}\right]\right] \\ &= \frac{1}{\theta}\left(1 + (D+2)\lambda\theta\mathbb{E}_{\mathrm{CP}}\left[\widetilde{\mathbf{S}}^{\theta} \,\middle|\, \mathcal{S}\right]\right).\end{aligned} \tag{2.7.10}$$

Now we take account of the stationary measures to compare the running times. We first set up some notations as follows.

- $\nu'_{\mathcal{S}}$ and $\pi'_{\mathcal{S}}$ are the stationary distribution of $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\mathcal{S})$ and $\mathbf{CP}^{\lambda}_{\rho,v}(\mathcal{S})$, respectively.
- $\mathcal{T}_{u_1}, \ldots, \mathcal{T}_{u_D}$ denote the subtrees from the children $u_1, \ldots, u_D$ of v outside the cycle. Note that these subtrees are i.i.d $\mathbf{GW}(\widetilde{\mu}^{\sharp})^{L-1}$.
- Set $\widetilde{\mathcal{S}} = \mathcal{S} \setminus \cup_{i=1}^{D} \mathcal{T}_{u_i}$. $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\widetilde{\mathcal{S}})$ denotes the delayed contact process on $\widetilde{\mathcal{S}}$ that fixes both ρ, v to be infected permanently, which has the depth $r(x; \widetilde{\mathcal{S}})$ computed with respect to ρ. In particular, all possible states of $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\widetilde{\mathcal{S}})$ have depth $r(x)$ at least one, since v is always infected.
- $\nu_{\mathcal{T}_{u_i}}$ and $\nu_{\widetilde{\mathcal{S}}}$ are the stationary distributions of $\mathbf{DP}^{\lambda,\theta}_{v}(\mathcal{T}^{+}_{u_i})$ and $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\widetilde{\mathcal{S}})$, respectively. Moreover, $\pi_{\mathcal{T}_{u_i}}$ and $\pi_{\widetilde{\mathcal{S}}}$ denote the stationary distributions of $\mathbf{CP}^{\lambda}_{v}(\mathcal{T}^{+}_{u_i})$ and $\mathbf{CP}^{\lambda}_{\rho,v}(\widetilde{\mathcal{S}})$, respectively. Also, set
$$\nu^{\otimes}_{\mathcal{S}} = \left(\otimes_{i=1}^{D} \nu_{\mathcal{T}_{u_i}}\right) \otimes \nu_{\widetilde{\mathcal{S}}}.$$

Note that $\pi'_{\mathcal{S}} = \left(\otimes_{i=1}^{D} \pi_{\mathcal{T}_{u_i}}\right) \otimes \pi_{\widetilde{\mathcal{S}}}$. Keeping in mind that $r(\mathbf{0}; \mathcal{S}) = 1$ in $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\mathcal{S})$, we obtain

by using (2.2.9) that

$$\nu'_{\mathcal{S}}(\mathbf{0}) \geq \nu_{S}^{\otimes}(\mathbf{0}). \tag{2.7.11}$$

Moreover, observe that if we merge ρ and v in $\widetilde{\mathcal{S}}$ into a single vertex ρ', then the resulting graph $\widetilde{\mathcal{S}}'$ satisfies $\widetilde{\mathcal{S}}' \sim \mathsf{GWC}(\widetilde{\mu}^{\sharp}; s-1)^{L}$, and we can consider the natural one-to-one correspondence between the two state spaces $\{0,1\}^{\widetilde{\mathcal{S}}\setminus\{\rho,v\}}$ and $\{0,1\}^{\widetilde{\mathcal{S}}'\setminus\{\rho'\}}$. Thus, we can regard them as

$$\Omega = \{0,1\}^{\widetilde{\mathcal{S}}\setminus\{\rho,v\}} = \{0,1\}^{\widetilde{\mathcal{S}}'\setminus\{\rho'\}}.$$

For any $x \in \Omega \setminus \{\mathbf{0}\}$, note that

$$r(x;\widetilde{\mathcal{S}}) \in \left\{ r(x;\widetilde{\mathcal{S}}'),\ r(x;\widetilde{\mathcal{S}}')+1 \right\}.$$

In particular, $r(x;\widetilde{\mathcal{S}}) - 1 \leq r(x;\widetilde{\mathcal{S}}')$. Further, we have $r(\mathbf{0};\widetilde{\mathcal{S}}) = 1$ and $r(\mathbf{0};\widetilde{\mathcal{S}}') = 0$. This implies that if $\nu_{\widetilde{\mathcal{S}}'}$ denotes the stationary distribution of $\mathsf{DP}_{\rho'}^{\lambda,\theta}(\widetilde{\mathcal{S}}')$, then

$$\nu_{\widetilde{\mathcal{S}}'}(\mathbf{0}) \leq \nu_{\widetilde{\mathcal{S}}}(\mathbf{0}).$$

Therefore, combining with (2.7.11), we have

$$\nu'_{\mathcal{S}}(\mathbf{0}) \geq \left(\prod_{i=1}^{D} \nu_{\mathcal{T}_{u_i}}(\mathbf{0}) \right) \cdot \nu_{\widetilde{\mathcal{S}}'}(\mathbf{0}). \tag{2.7.12}$$

We can relate the quantities in (2.7.12) with the running times of the delayed processes. Let $\mathbf{S}_{L-1}^{\theta}$ be the first time when $\mathsf{DP}_{\rho}^{\lambda,\theta}(\mathcal{T}_{u_i}; \mathbf{1}_{u_i})$ returns to $\mathbf{0}$. Similarly, let $v_1, v_2 \notin \{\rho, v\}$ be the two neighbors of $\{\rho, v\}$ in $\mathcal{S}'$, let $\mathbf{S}_{v_i}^{\theta}$ be the first time when $\mathsf{DP}_{\rho,v}^{\lambda,\theta}(\mathcal{S}'; \mathbf{1}_{v_i})$ reaches $\mathbf{0}$, and observe that $\mathbf{S}_{s-1,L}^{\theta} \overset{d}{=} \frac{1}{2}(\mathbf{S}_{v_1}^{\theta} + \mathbf{S}_{v_2}^{\theta})$, where the definition of $\mathbf{S}_{s-1,L}^{\theta}$ is given in the

beginning of the proof. Continuing similarly as (2.3.3, 2.2.10), we get that

$$\begin{aligned}\nu'_{\mathcal{S}}(\mathbf{0}) &= \frac{1}{1+(D+2)\lambda\theta\mathbb{E}_{\mathrm{CP}}[\widetilde{\mathbf{S}}^{\theta}|\mathcal{S}]};\\ \nu_{\mathcal{T}_{u_i}}(\mathbf{0}) &= \frac{1}{1+\lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}^{\theta}_{L-1}|\mathcal{T}_{u_i}]};\\ \nu_{\widetilde{\mathcal{S}}'}(\mathbf{0}) &= \frac{1}{1+2\lambda\mathbb{E}_{\mathrm{CP}}[\mathbf{S}^{\theta}_{s-1,L}|\widetilde{\mathcal{S}}]},\end{aligned} \tag{2.7.13}$$

where the additional factor of θ in the first identity comes from the fact that $r(\mathbf{0};\mathcal{S}) = 1$ in $\mathbf{DP}^{\lambda,\theta}_{\rho,v}(\mathcal{S})$. Plugging these into (2.7.12) and using (2.7.10), we obtain that

$$\mathbb{E}[\mathbf{S}^{\theta}_{s,L}|D] \le \mathbb{E}\left[\widetilde{\mathbf{S}}^{\theta}_{s,L}\middle| D\right] \le \frac{1}{\theta}(1+\lambda\mathbb{E}[\mathbf{S}^{\theta}_{L-1}])^{D}(1+2\lambda\mathbb{E}[\mathbf{S}^{\theta}_{s-1,L}]).$$

Arguing similarly as Lemma 2.3.11 and Proposition 2.2.6, we deduce that there exist constants $C, \lambda_0 > 0$ depending on μ such that for all $\lambda \le \lambda_0$ and s, L with $s \ge 2$, $\mathbb{E}[\mathbf{S}^{\theta}_{s,L}] \le 3/\theta$ for $\theta = C\lambda$. Setting C to satisfy $C \ge 3$, and applying this to the right-hand side of the above equation (which is written in terms of $(s-1, L)$) gives the desired conclusion. □

Proof of Lemma 2.3.12. To finish the proof of Lemma 2.3.12, we argue similarly as Proposition 2.3.9. Namely, we establish the result for $\mathbf{EGW}(\widetilde{\mu}^{\sharp}; h, s)^{L}$ and then extend it to the general case $\mathbf{EGW}(\mu^{\sharp}, \widetilde{\mu}^{\sharp}; h, s)^{L}$. In both steps, we appeal to the same technique as Proposition 2.2.6, which is simpler than what is done here for the **GWC**-processes. We omit the details due to similarity. □

2.7.5 Eliminating high-degree vertices from random graphs

In this subsection, we provide the details of the proof of Lemma 2.6.1, Items 1-3.

Item 1 follows from the definition of G_n that its edges are obtained from a uniformly chosen perfect matching of the half-edges.

For Item 2, choose j large enough such that

$$\delta := \mathbb{E}_{D\sim\mu} D\mathbf{1}_{D\ge 2j+1} \le \varepsilon/4.$$

For each vertex $v \in G_n$, consider the random variable

$$X_v := \deg_{G_n}(v)\mathbf{1}_{\deg_{G_n}(v)\geq 2j+1}.$$

These random variables are independent with mean δ and variance bounded by the second moment of μ. By Chebyshev's inequality, whp

$$\sum_{v\in G_n} X_v \leq \varepsilon n/2.$$

Thus, whp, the total number of removed half-edges is at most $2\sum_{v\in G_n} X_v \leq \varepsilon n$. So is the number of removed vertices. Thus, $\bar{n} \geq (1-\varepsilon)n$. Applying Chernoff inequality to the random variables $\bar{X}_v := \deg_{G_n}(v)\mathbf{1}_{\deg_{G_n}(v)\leq 2j}$ we obtain that whp,

$$\sum_{v\in G_n} \bar{X}_v \in (1-\varepsilon, 1+\varepsilon)nd.$$

Combining this with the fact that the total number of deleted half-edges is at most εnd whp, we get

$$(1-2\varepsilon)nd \leq \sum_{v\in G_n} \bar{X}_v - \varepsilon n \leq d_1 + \cdots + d_{\bar{n}} \leq \sum_{v\in G_n} \bar{X}_v \leq (1+\varepsilon)nd$$

completing the proof of Item 2.

To prove Item 3, let $\varepsilon' = \min\{\varepsilon, \varepsilon(b-1)\}$. Let k_0 be a large constant such that for all $h \geq k_0$, the branching rate of $\mu_{[0,h]} \in (1-\varepsilon', 1+\varepsilon')b$, namely

$$\frac{\mathbb{E}_{D\sim\mu} D(D-1)\mathbf{1}_{0\leq D\leq h}}{\mathbb{E}_{D\sim\mu} D\mathbf{1}_{0\leq D\leq h}} \in (1-\varepsilon', 1+\varepsilon')b, \tag{2.7.14}$$

$$\begin{aligned} &\mathbb{E}_{D\sim\mu} D\mathbf{1}_{0\leq D\leq k_0} \in (1-\varepsilon', 1+\varepsilon')d, \\ \text{and} \quad &\mathbb{E}_{D\sim\mu} D^2\mathbf{1}_{0\leq D\leq k_0} \geq \frac{2}{\varepsilon'}\mathbb{E}_{D\sim\mu} D^2\mathbf{1}_{k_0<D}. \end{aligned} \tag{2.7.15}$$

Let $k \geq k_0$ be such that

$$\mathbb{E}_{D\sim\mu} D^2\mathbf{1}_{k_0<D\leq k} \geq \frac{2}{\varepsilon'}\mathbb{E}_{D\sim\mu} D^2\mathbf{1}_{k<D}. \tag{2.7.16}$$

which k exists because of the boundedness of $\mathbb{E}_{D\sim\mu}D^2$. Note that (2.7.16) implies that

$$\mathbb{E}_{D\sim\mu}\left(D\mathbf{1}_{k_0<D\le k}\right) \ge \frac{2}{\varepsilon'}\mathbb{E}_{D\sim\mu}\left(D\mathbf{1}_{k<D}\right). \tag{2.7.17}$$

We now show that for all constant $j \ge k$, $\bar{b} \in (1-\varepsilon', 1+\varepsilon')b$ whp. Let E_l and $\bar{E}_l$ be the number of half-edges attached to vertices of degree l in G_n and $\bar{G}_n$ respectively. We need to show that whp,

$$\bar{b} = \frac{\sum_{l=0}^{2j}(l-1)\bar{E}_l}{\sum_{l=0}^{2j}\bar{E}_l} \in (1-\varepsilon, 1+\varepsilon)b. \tag{2.7.18}$$

Since j is a constant and $E_l = l\sum_{v\in G_n}\mathbf{1}_{\deg_{G_n}(v)=l}$, by Chernoff inequality, whp we have,

$$\begin{aligned}
&\sum_{l=0}^{2j}E_l \in (1-\varepsilon', 1+\varepsilon')\sum_{l=0}^{2j}l\mu(l)n \quad \text{and} \\
&\sum_{l=0}^{2j}(l-1)E_l \in (1-\varepsilon', 1+\varepsilon')\sum_{l=0}^{2j}l(l-1)\mu(l)n.
\end{aligned} \tag{2.7.19}$$

This together with (2.7.14) and (2.7.15) give

$$\frac{\sum_{l=0}^{2j}(l-1)E_l}{\sum_{l=0}^{2j}E_l} \in (1-3\varepsilon', 1+3\varepsilon')b \quad \text{and} \quad \sum_{l=0}^{2j}l\mu(l) \in (1-\varepsilon', 1+\varepsilon')d.$$

Since the total number of removed half-edges is at most $\varepsilon' nd$ whp,

$$\sum_{l=0}^{2j}\bar{E}_l \in \sum_{l=0}^{2j}E_l + (-\varepsilon', 0)nd \subset (1-2\varepsilon', 1+2\varepsilon')\sum_{l=0}^{2j}E_l.$$

From this and (2.7.19), (2.7.18) reduces to proving that

$$\sum_{l=0}^{2j}l\bar{E}_l \in (1-\varepsilon', 1+\varepsilon')\sum_{l=0}^{2j}lE_l.$$

The upper bound is straightforward. To prove the lower bound, let N_l be the number of vertices of degree l in G_n. Since the number of deleted half-edges is at most $\sum_{l=2j+1}^{\infty}lN_l$,

we have by Markov's inequality, (2.7.17), and Chernoff inequality, whp

$$\begin{aligned}\sum_{l=0}^{2j}(E_l-\bar{E}_l) &\le \sum_{l=2j+1}^{\infty} lN_l \le \frac{1}{\varepsilon'}\mathbb{E}\sum_{l=2j+1}^{\infty} lN_l \\ &\le \frac{1}{2}\mathbb{E}\sum_{l=k_0+1}^{2j} lN_l \le \sum_{l=k_0+1}^{2j} lN_l.\end{aligned} \tag{2.7.20}$$

Thus, we have

$$\sum_{l=0}^{2j} l(E_l-\bar{E}_l) \le \sum_{l=k_0+1}^{2j} l^2N_l \le \varepsilon'\sum_{l=0}^{2j} l^2N_l = \varepsilon'\sum_{l=0}^{2j} lE_l \tag{2.7.21}$$

where the first inequality follows from (2.7.20) and the fact that the left-hand side of (2.7.21) is largest when the deleted half-edges counted in $\sum_{l=0}^{2j}(E_l-\bar{E}_l)$ are drawn from vertices of highest degrees possible and the second inequality follows from the Chernoff inequality and (2.7.15). That completes the proof of (2.7.18) and hence Item 3. □

Chapter 3

First-order asymptotics of the critical value

We continue our study of the contact process on random graphs. In Chapter 2, we saw the existence of contrasting phases when the offspring distribution (or the degree distribution) has an exponential tail decay. In this chapter, we develop stronger methods for a more precise investigation of the critical value, the threshold where the phase transition happens. Our main goal is to prove the remaining results introduced in Section 1.5, namely, Theorems 3, 7, 10, 11, and Corollary 8. The logic and technical details developed in this chapter are all based on the work [NNS21] by Nguyen, Sly, and the author.

3.1 Overview of the chapter

In this section, we briefly introduce the primary notions and discuss the main ideas used in the chapter. We also address the organization of the rest of the article in Section 3.1.4. The notations introduced here are mostly consistent with those from the previous chapter, however, some will be different as we have to work with a more complicated objects.

3.1.1 The lower bound of Theorem 3

In Chapter 2, we worked with the root-added contact process defined in Definition 2.2.2. Our argument will heavily rely on this modified process again. In the following, we formally

define the *survival and excursion times*, which are the notions that already have appeared in the previous chapter.

Definition 3.1.1 (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 $\mathsf{CP}^{\lambda}(T; \mathbf{1}_{\rho})$ 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 $\mathsf{CP}^{\lambda}_{\rho^+}(T; \mathbf{1}_{\rho})$ 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 $R(T) \le S(T)$, which can be seen through the coupling via the graphical representation.

We established $\lambda_1^{\mathrm{GW}} > 0$ in Chapter 2 by a recursive inequality on the depth of the tree that showed $\mathbb{E}\mathbf{S}(\mathcal{T}) < \infty$ for $\mathcal{T} \sim \mathsf{GW}(\xi)$ 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 3.2 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, $\mathbb{P}_{\mathrm{GW}}(S(\mathcal{T}) \ge t)$, and then the bound easily implies $\mathbb{E}_{\mathrm{GW}}S(\mathcal{T}) < \infty$. This is a substantial improvement from the methods in Chapter 2 where we could only control its expectation $\mathbb{E}_{\mathrm{GW}}S(\mathcal{T})$ for small enough λ.

3.1.2 Deep infections, unicyclic neighbors and the lower bound of Theorem 7

To establish Theorem 7, we attempt to generalize Theorem 3 based on the fact that the local neighborhoods of $\mathcal{G}(n, \mu)$ look like Galton-Watson trees. There are two major obstacles when 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 $l \ge c_{\mu} \log n$ for some constant c_{μ}.

2. Even for small l, there are $o(n)$ vertices that contain a cycle in $N(v, l)$.

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 $\mathsf{CP}^\lambda(\mathcal{T}, \mathbf{1}_\rho)$ 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 3.1.2 (Total infections at leaves)**.** Let T be a finite tree rooted at ρ, set

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

to 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 $(X_t) \sim \mathsf{CP}^\lambda_{\rho^+}(T; \mathbf{1}_\rho)$. For $v \in \mathcal{L}$, define the *total infections at* v, by

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

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

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

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 $\mathbf{M}^0(T) \equiv 1$.

In the previous chapter, we 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 (Theorem 2.2.4), but as before it had to require λ to be small

enough. However, unfortunately, this 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

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

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 Chapter 2, this approach was sufficient since we could set λ as small as we wanted.

In Section 3.3, 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 3.1.1, 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.

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 the same idea as before, by observing that if μ satisfies (1.5.2), then there exists $\gamma(\mathsf{c}) > 0$ 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 3.4.9). Therefore, we study $S(T')$ and $M^l(T')$ as above, for certain unicyclic graphs T' which are closely related to Galton-Watson trees (precise definitions are given in Section 3.4.1). 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 Section

3.5. 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 $N(v,l)$.

3.1.3 Proof of Theorems 10 and 11

Previously in Chapter 2, we 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 straightforward 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 $\lambda_c^+(\mu) \le \lambda_1^{\mathrm{GW}}(\widetilde{\mu})$ 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 $\mathcal{T} \sim \mathbf{GW}(\widetilde{\mu})$ grows exponentially in time (Lemmas 3.6.5 and 3.6.6). 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 Section 2.6. 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 $\ge 2\delta n$ *good vertices* to be infected at a later time with high probability except for an exponentially small error.

3.1.4 Organization

Sections 3.2–3.4 are devoted to the derivation of the lower bounds of Theorems 3 and 7. In Section 3.2, we introduce the basic form of the recursion argument on Galton-Watson

trees and establish the lower bound of Theorem 3. In Section 3.3, we extend the recursion criterion to the study of deep infections. Section 3.4 then concludes the proof of the lower bound of Theorem 7, while the technical works needed to study the unicyclic graphs are deferred to Section 3.5. Finally, we finish the proof of Theorems 3 and 7 by settling their upper bounds in Section 3.6. Finally, Section 3.7 is devoted to demonstrating some of the technical lemmas developed in the chapter.

3.2 Survival and excursion times on trees

In this section, we introduce primary recursive argument on the expected excursion time which are used throughout the chapter. In Section 3.2.1, we derive another recursive inequality on excursion times, in addition to the previous one developed in Chapter 2. In Sections 3.2.2 and 3.2.6, we prove a tail probability estimate and establish Theorem 3 as its application.

3.2.1 Deterministic recursions on trees

Let T be a finite tree rooted at ρ and recall the definition of the root-added contact process $\mathbf{CP}^{\lambda}_{\rho^+}(T^+; x_0)$ (Definition 2.2.2). Let $R(T)$ and $S(T)$ be expected survival and recursion time as in Definition 3.1.1.

In (T, ρ), let $D = \deg(\rho)$ and $v_1, \dots, v_D$ be the children of ρ. Further, let $T_1, \dots, T_D$ be the subtrees from each child of ρ, rooted at $v_1, \dots, v_D$, respectively. In the proof of Lemma 2.2.3, we proved the following recursion on the excursion times:

Proposition 3.2.1. *Let T and $T_1, \dots, T_D$ be as above. Then, the expected excursion times $S(T)$ and $S(T_1), \dots, S(T_D)$ satisfy*

$$S(T) \le \prod_{i=1}^{D} (1 + \lambda S(T_i)). \tag{3.2.1}$$

It may sound unusual that a step in the proof of a lemma is written as a proposition, but this tells how important this recursion is going to be in this chapter.

Unfortunately, having (3.2.1) is insufficient for our purpose. One can see this by taking

expectation on each side of (3.2.1) 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 (recall the proof Lemma 2.2.3). Therefore, we develop another recursion which redeems (3.2.1). Our first step is to build up a recursion regarding $R(T)$, the expected survival time.

Proposition 3.2.2. *Let $T_1, \dots, T_D$ be as Proposition 3.2.1, and assume that*

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

Then, we have

$$R(T) \le \frac{1 + \lambda \sum_{i=1}^{D} R(T_i)}{1 - \lambda^2 \sum_{i=1}^{D} R(T_i)}. \tag{3.2.2}$$

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 children 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 $\tau_1, \dots, \tau_Q$ 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 1.1) that encodes the graphical representation of $\mathsf{CP}^{\lambda}(T; \mathbf{1}_{\rho})$. In other words, the dynamics of $\mathsf{CP}^{\lambda}(T; \mathbf{1}_{\rho})$ can be seen on the diagram $\mathfrak{G}_0$, starting from the single initial infection

at ρ at time 0. Let $\mathfrak{G}_1, \dots, \mathfrak{G}_Q$ 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 is 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 $j = 1, \dots, Q$, 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:

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

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

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

Proposition 3.2.3. *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*

$$\frac{S(T)}{1 + \lambda S(T)} \le R(T). \tag{3.2.3}$$

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

$$\lim_{t_0\to\infty} \frac{\mathbb{E}_{\mathrm{CP}}|A_{t_0}|}{t_0} = \frac{S(T)}{\lambda^{-1}+S(T)} = \frac{\lambda S(T)}{1+\lambda S(T)}. \tag{3.2.4}$$

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

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

Comparing this to (3.2.4), we obtain that

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

implying the conclusion. □

Combining Propositions 3.2.2 and 3.2.3, we obtain that

$$\begin{aligned} S(T) \le \frac{R(T)}{1-\lambda R(T)} &\le \frac{1+\lambda\sum_{i=1}^{D} R(T_i)}{1-\lambda-2\lambda^2\sum_{i=1}^{D} R(T_i)} \\ &\le \frac{1+\lambda\sum_{i=1}^{D} S(T_i)}{1-\lambda-2\lambda^2\sum_{i=1}^{D} S(T_i)}, \end{aligned} \tag{3.2.5}$$

provided that $\lambda + 2\lambda^2\sum_{i=1}^{D} S(T_i) < 1$. In the rest of the chapter, (3.2.1) and (3.2.5) serve as two major recursive inequalities for $S(T)$.

3.2.2 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 3 as its application.

Let ξ be an integer-valued random variable that satisfy the concentration condition (1.5.2) 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 $\lambda \le (1-\varepsilon)d^{-1}$, 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 $\mathsf{GW}(\xi)^l$.

Theorem 3.2.4. *Let $l \ge 0$ 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 $d := \mathbb{E}\xi \ge d_0$ and (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{T}^l \sim \mathsf{GW}(\xi)^l$ that*

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^l) \ge t\right) \le t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \ge \frac{2}{\varepsilon}, \tag{3.2.6}$$

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

Remark 3.2.5. The term $(\log t)^{-2}$ in the r.h.s. of (3.2.6) may look useless, but this plays a key role in carrying out an inductive argument (see Lemma 3.2.8). 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 3.2.4. 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 (3.2.6).

Suppose that we have (3.2.6) for l. Let $(\mathcal{T}^{l+1}, \rho) \sim \mathsf{GW}(\xi)^{l+1}$, 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) $2/\varepsilon \le t \le d^{\frac{1}{10}}$;
2. (intermediate) $d^{\frac{1}{10}} \le t \le \exp(\frac{1}{2}c_1\sqrt{d})$;
3. (large) $\exp(\frac{1}{2}c_1\sqrt{d}) \le t$.

Then, we establish (3.2.6) on each regime separately. As the proof goes on, we will figure out the conditions for $d_0(\varepsilon, \mathfrak{c})$ as well.

Remark 3.2.6. 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 3.4.4 and Lemma 3.5.5. The choice of $\exp(\frac{1}{2}c_1\sqrt{d})$ will be clear in (3.2.16).

3.2.3 Proof of the recursive tail estimate: small t

We first prove (3.2.6) of Theorem 3.2.4 for small t. To this end, we rely on (3.2.5). For $\mathcal{T}^{l+1} \sim \mathsf{GW}(\xi)^{l+1}$ and $T_i,\ i \in [D]$ as above, suppose that

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

Then, from (3.2.5) and a little bit of algebra we see that

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

where the last inequality holds if

$$d \ge \frac{42}{\varepsilon^2}. \tag{3.2.7}$$

Therefore, for d with (3.2.7), we have

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \ge \frac{2}{\varepsilon}\right) \le \mathbb{P}\left(D \ge \left(1+\frac{\varepsilon}{6}\right)d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} S(T_i) \ge \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right), \tag{3.2.8}$$

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

$$\mathbb{P}\left(D \ge \left(1+\frac{\varepsilon}{6}\right)d\right) \le \exp\left(-c_{\varepsilon'}d\right), \tag{3.2.9}$$

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

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

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

Lemma 3.2.7. *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

$$\mathbb{P}\left(Z_i \geq t\right) \leq 5t^{-\sqrt{d}}, \quad \text{for all } t \geq \frac{2}{\varepsilon}. \tag{3.2.10}$$

Then, we have

$$\mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i \geq \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right) \leq \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}. \tag{3.2.11}$$

We defer the proof of this lemma until Section 3.2.6, since it is essentially a special case of Lemma 3.2.8 below.

Thanks to Lemma 3.2.7, we combine (3.2.8), (3.2.9) and (3.2.11) to obtain that

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq \frac{2}{\varepsilon}\right) \leq \exp(-c_{\varepsilon'}d) + \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\log d^{\frac{1}{10}}\right)^{-2} \leq d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}, \tag{3.2.12}$$

for d satisfying

$$\exp(-c_{\varepsilon'}d) \leq \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}. \tag{3.2.13}$$

Finally, we clearly see that (3.2.12) settles (3.2.6) for all small t, namely, $2/\varepsilon \leq t \leq d^{\frac{1}{10}}$.

3.2.4 Proof of the recursive tail estimate: intermediate t

We continue our discussion and establish (3.2.6) for intermediate t. For $t \geq d^{\frac{1}{10}}$, we use (3.2.1). That is, we attempt to control

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D}\log(1+\lambda S(T_i)) \geq \log t\right). \tag{3.2.14}$$

Let c_1 be the constant given from the assumption of Theorem 3.2.4 (with $\delta = 1$). Note that we can split the event in the r.h.s. as follows.

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq 2d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d}\log(1+\lambda S(T_i)) \geq \log t\right). \tag{3.2.15}$$

The concentration assumption (1.5.2) on ξ tells us that the first term is bounded by

$$\mathbb{P}(D \geq 2d) \leq \exp(-c_1 d) \leq \frac{1}{2} t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for } t \leq e^{\frac{1}{2} c_1 \sqrt{d}}. \tag{3.2.16}$$

To control the second term in the r.h.s. of (3.2.15), we use the following lemma.

Lemma 3.2.8. *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*

$$\mathbb{P}(Z_i \geq t) \leq 5 t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for all } t \geq \frac{3}{\varepsilon}. \tag{3.2.17}$$

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

$$\mathbb{P}\left(\sum_{i=1}^{2d} \log\left(1 + \frac{4}{d} Z_i\right) \geq \log\left(\frac{t}{2}\right) \right) \leq \frac{1}{2} t^{-\sqrt{d}} (\log t)^{-2}. \tag{3.2.18}$$

The proof of Lemma 3.2.8 is postponed until Section 3.2.6, since it requires a substantial technical work. Note that (3.2.18) contains a slightly more generalized form than the r.h.s. of (3.2.15), which will be useful in Sections 3.3 and 3.4.

To conclude the proof for intermediate t, (3.2.15), (3.2.16) and (3.2.18) together deduce (3.2.6) 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 3.2.8.

3.2.5 Proof of the recursive tail estimate: large t

We finish the proof of theorem 3.2.4 by settling (3.2.6) for large t. For $t \geq \exp(\frac{1}{2} c_1 \sqrt{d})$, we again attempt to control (3.2.14). 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 (3.2.14) as follows.

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \Delta_{d,t}\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1+\lambda S(T_i)) \geq \log t\right) + \sum_{2d \leq r \leq \Delta_{d,t}} \mathbb{P}(D \geq r) \times \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{r} \log(1+\lambda S(T_i)) \geq \log t\right). \tag{3.2.19}$$

The concentration assumption (1.5.2) on ξ tells us that for $t \geq \exp(\frac{1}{2}c_1\sqrt{d})$,

$$\mathbb{P}\left(D \geq \Delta_{d,t}\right) \leq \mathbb{P}\left(D \geq d + \frac{\Delta_{d,t}}{2}\right) \leq \exp\left(-2\sqrt{d}\log t\right) \leq \frac{1}{4}t^{-\sqrt{d}}(\log t)^{-2}, \tag{3.2.20}$$

where the last inequality is satisfied by large d. This controls the first term in the r.h.s. of (3.2.19). The second term is then estimated by Lemma 3.2.8, which gives

$$\mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1+\lambda S(T_i)) \geq \log t\right) \leq \frac{1}{2}t^{-\sqrt{d}}(\log t)^{-2}. \tag{3.2.21}$$

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

Corollary 3.2.9. *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*

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

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

$$\mathbb{P}\left(\sum_{i=1}^{r} \log\left(1+\frac{4}{d}Z_i\right) \geq \log\left(\frac{t}{2}\right)\right) \leq t^{-\sqrt{d}}(\log t)^{-2}\exp\left(\frac{C_1 r}{\varepsilon\sqrt{d}}\right). \tag{3.2.22}$$

The proof of Corollary 3.2.9 is postponed until Section 3.2.6, which will be proven together with Lemma 3.2.8.

To conclude the proof of Theorem 3.2.4 for large t, observe that for $2d \le r \le \Delta_{d,t}$

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

by the concentration condition (1.5.2) of D. Therefore, for sufficiently large d, we have by (3.2.22) that

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

Along with (3.2.19), (3.2.20) and (3.2.21), this proves Theorem 3.2.4 for $t \ge \exp(\frac{1}{2} c_1 \sqrt{d})$. □

As the first application of Theorem 3.2.4, we establish Theorem 3.

Proof of Theorem 3: the lower bound. Let $\varepsilon \in (0,1)$ be an arbitrary constant and suppose that the collection of distributions $\{\xi_k\}_k$ satisfies (1.5.2) for $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$. Let $d_0 = d_0(\varepsilon, \mathfrak{c})$ as in Theorem 3.2.4, and assume that $\mathbb{E}\xi_k \ge d_0$. Then, Theorem 3.2.4 implies that for $\mathbf{CP}^{\lambda}_{\rho^+}(\mathcal{T}_k^l; \mathbf{1}_\rho)$ on $\mathcal{T}_k \sim \mathbf{GW}(\xi_k)^l$ with intensity $\lambda = (1-\varepsilon)(\mathbb{E}\xi_k)^{-1}$, we have

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

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

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

For $\mathcal{T}_k \sim \mathbf{GW}(\xi_k)$, 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

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

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

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

for all large enough k. □

3.2.6 Proof of the induction lemmas

To conclude Section 3.2, we discuss the proof of Lemmas 3.2.7, 3.2.8 and Corollary 3.2.9. We first establish Lemma 3.2.8 and Corollary 3.2.9 together, and then the proof of Lemma 3.2.7 will follow based on similar ideas.

Proof of Lemma 3.2.8 and Corollary 3.2.9. We first set

$$Y_i = \log\left(1 + \frac{4Z_i}{d}\right), \quad \Delta_{d,t} = \frac{4\sqrt{d}}{c_1}\log t, \quad \text{and let} \quad 2d \leq r \leq 2d \vee \Delta_{d,t}. \tag{3.2.23}$$

We will attempt to find an upper bound of

$$\mathbb{P}\left(\sum_{i=1}^{r} \log\left(1 + \frac{4}{d}Z_i\right) \geq \log\left(\frac{t}{2}\right)\right), \tag{3.2.24}$$

to cover the case of Corollary 3.2.9 as well.

Note that the assumption (3.2.17) on Z_i gives us $\mathbb{E}Z_i \leq \frac{4}{\varepsilon}$. Based on (3.2.17), we can find an upper bound for $\mathbb{E}Y_i$ by

$$\mathbb{E}Y_i \leq \log \mathbb{E}e^{Y_i} = \log\left(1 + \frac{4\mathbb{E}Z_i}{d}\right) \leq \log\left(1 + \frac{16}{\varepsilon d}\right) \leq \frac{16}{\varepsilon d}. \tag{3.2.25}$$

Further, (3.2.17) tells us that the tail of Y_i is bounded by

$$\mathbb{P}(Y_i \geq s) = \mathbb{P}\left(Z_i \geq \frac{d}{4}(e^s - 1)\right) \leq 5\left(\frac{d}{4}(e^s-1)\right)^{-\sqrt{d}}\left\{\log\left(\frac{d}{4}(e^s-1)\right)\right\}^{-2}. \tag{3.2.26}$$

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

$$(2d \vee \Delta_{d,t}) \cdot \frac{\log d}{80d} = \frac{1}{40} \log d \vee \frac{\log d}{20c_1\sqrt{d}} \log t =: m(d,t). \tag{3.2.27}$$

Then, we can split the probability in (3.2.24) as follows.

$$\begin{aligned} &\mathbb{P}\left(\sum_{i=1}^{r} Y_i \geq \log\left(\frac{t}{2}\right)\right) \\ &\leq \sum_{k=1}^{r} \binom{r}{k} \sum_{j=1}^{m(d,t)+1} \mathbb{P}\left(\sum_{i=1}^{k} Y_i \geq \log\left(\frac{t}{2}\right) - j \ ; \ Y_i \geq \frac{\log d}{80d}, \ \forall i \leq k\right) \\ &\qquad\qquad \times \mathbb{P}\left(\sum_{i=k+1}^{r} Y_i \geq j-1 \ ; \ Y_i \leq \frac{\log d}{80d}, \ \forall i > k\right). \end{aligned} \tag{3.2.28}$$

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

Lemma 3.2.10. *Under the setting of Lemma 3.2.8 and (3.2.23), 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 (3.2.26), we obtain that

$$\mathbb{P}(Y_i \geq s) \leq 5 \left(\frac{640}{\log d}\right)^{\sqrt{d}} e^{-s\sqrt{d}} \left(s + \frac{1}{2}\log\log d\right)^{-2}, \tag{3.2.29}$$

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

$$\frac{1}{2}\log\log d \leq \log\left(\frac{\log d}{640}\right), \quad \text{i.e.,} \quad d \geq \exp\left(640^2\right). \tag{3.2.30}$$

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

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{k} Y_i \geq \log\left(\frac{t}{2}\right) - j \ ; \ Y_i \geq \frac{\log d}{80d}, \ \forall i \leq k\right) \\
&\leq \sum_{\substack{u_1,\dots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i \geq \frac{80d}{\log d}\{\log(t/2)-j\}-k}} \left[\prod_{i=1}^{k} \mathbb{P}\left(Y_i \in \frac{\log d}{80d}[u_i, \ u_i+1]\right)\right] \\
&\leq \sum_{\substack{u_1,\dots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k \rfloor}} \left[\prod_{i=1}^{k} \mathbb{P}\left(Y_i \geq \frac{\log d}{80d} u_i\right)\right].
\end{aligned} \tag{3.2.31}$$

Then, the r.h.s. of (3.2.31) is upper bounded by

$$\begin{aligned}
&5^k \left(\frac{640}{\log d}\right)^{k\sqrt{d}} e^{-\sqrt{d}\left(\log(t/2)-j-\frac{k\log d}{80d}\right)} \\
&\times \sum_{\substack{u_1,\dots,u_k \geq 1: \\ \sum_{i=1}^{k} u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k \rfloor}} \prod_{i=1}^{k} \left(\frac{\log d}{80d} u_i + \frac{1}{2}\log\log d\right)^{-2}.
\end{aligned} \tag{3.2.32}$$

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

Lemma 3.2.11. *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 3.2.11. The l.h.s. of above can be upper bounded by

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

□

From now on, $K > 0$ is an absolute constant that may vary from line by line. We apply

Lemma 3.2.11 $(k-1)$-times to (3.2.32) with $a = \frac{1}{2}\log\log d$, $b = \frac{\log d}{80d}$ and see that

$$\begin{aligned}\sum_{\substack{u_1,\dots,u_k\geq 1:\\ \sum_{i=1}^k u_i = \lfloor \frac{80d}{\log d}\{\log(t/2)-j\}-k\rfloor}} \prod_{i=1}^k \left(\frac{\log d}{80d}u_i + \frac{1}{2}\log\log d\right)^{-2} \\ \leq \left(\frac{Kd}{\log d}\right)^{k-1}\left(\log\left(\frac{t}{2}\right) - j - \frac{k\log d}{80d}\right)^{-2}.\end{aligned} \tag{3.2.33}$$

Therefore, we see from (3.2.31), (3.2.32) and (3.2.33) that

$$\begin{aligned}&\mathbb{P}\left(\sum_{i=1}^k Y_i \geq \log\left(\frac{t}{2}\right) - j \ ;\ Y_i \geq \frac{\log d}{80d},\ \forall i \leq k\right) \\ &\leq \left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k e^{-\sqrt{d}\left(\log(t/2)-j-\frac{k\log d}{80d}\right)}\left(\log\left(\frac{t}{2}\right) - j - \frac{k\log d}{80d}\right)^{-2}.\end{aligned} \tag{3.2.34}$$

Note that the definition of r (3.2.23) and $m(d,t)$ (3.2.27) 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 (3.2.34) 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.

□

Now we turn our attention to bounding the second term in the inner sum of (3.2.28) in the r.h.s..

Lemma 3.2.12. *Under the setting of Lemma 3.2.8 and (3.2.23), 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 (3.2.25) and (3.2.26),

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

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

$$\begin{aligned}\mathbb{P}\left(\sum_{i=k+1}^{r} Y_i \ge j-1\ ;\ Y_i \le \frac{\log d}{80d},\ \forall i>k\right) &\le \mathbb{P}\left(\sum_{i=1}^{r} Y_i \ge j-1\ ;\ Y_i \le \frac{\log d}{80d},\ \forall i>k\right) \\ &\le 1\ \wedge\ \exp\left(-\varepsilon d^{2/3}(j-1)\right)\times\left\{\mathbb{E}\left[\exp\left(\varepsilon d^{2/3}Y_i\right)\ ;\ Y_i\le\frac{\log d}{80d}\right]\right\}^r \\ &\le 1\ \wedge\ \frac{\left(1+18d^{-1/3}\right)^r}{\exp\left(\varepsilon d^{2/3}(j-1)\right)} \le 1\wedge\left(\frac{\exp\left(19r\,d^{-1/3}\right)}{\exp\left(\varepsilon d^{2/3}j\right)}\right),\end{aligned}$$

concluding the proof. □

We can now combine (3.2.28) with Lemmas 3.2.10 and 3.2.12 to see that

$$\begin{aligned}&\mathbb{P}\left(\sum_{i=1}^{r} Y_i \ge \log\left(\frac{t}{2}\right)\right) \\ &\le \frac{t^{-\sqrt{d}}}{(\log t)^2}\sum_{k=1}^{r}\sum_{j=1}^{m(d,t)+1}\binom{r}{k}\left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k e^{j\sqrt{d}}\left(\frac{e^{19r\,d^{-1/3}}}{e^{\varepsilon d^{2/3}j}}\ \wedge\ 1\right), \qquad (3.2.35) \\ &\le \frac{t^{-\sqrt{d}}}{(\log t)^2}\sum_{k=1}^{r}\binom{r}{k}\left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k\left[\sum_{j=1}^{m(d,t)+1} e^{j\sqrt{d}}\left(\frac{e^{19r\,d^{-1/3}}}{e^{\varepsilon d^{2/3}j}}\ \wedge\ 1\right)\right].\end{aligned}$$

The last inner sum over j can be split into two parts, $\varepsilon d^{2/3}j \le 19r\,d^{-1/3}$ and $\varepsilon d^{2/3}j > 19r\,d^{-1/3}$. That is, we divide into two regimes based on $j_0 = 19r(\varepsilon d)^{-1}$. If, for instance, $2\sqrt{d} \le \varepsilon d^{2/3}$, then this sum is bounded by $2\exp(j_0\sqrt{d}) = 2\exp(19r\,\varepsilon^{-1}d^{-1/2})$, 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,

$$d \geq \left(\frac{2}{\varepsilon}\right)^6. \tag{3.2.36}$$

Further, the outer sum over k in the r.h.s. of (3.2.35) is at most

$$\begin{aligned} 2\exp\left(\frac{19r}{\varepsilon\sqrt{d}}\right)\sum_{k=1}^{r}\binom{r}{k}\left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k &\leq 2\exp\left(\frac{19r}{\varepsilon\sqrt{d}}\right)\times\left(1+\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^r \\ &\leq \exp\left(\frac{20r}{\varepsilon\sqrt{d}}\right), \end{aligned} \tag{3.2.37}$$

where the last inequality holds for large d such that

$$\frac{1}{\varepsilon\sqrt{d}} > \left(\frac{K}{\log d}\right)^{\sqrt{d}} d. \tag{3.2.38}$$

Therefore, we establish Corollary 3.2.9 from (3.2.35) and (3.2.37) by setting $C_1 = 20$ and d_0 to satisfy (3.2.30), (3.2.36) and (3.2.38). For Lemma 3.2.8, we plug in $r = 2d$ in the l.h.s. of (3.2.37), and see that if d is so large that

$$\log d \geq 3 \cdot K \cdot e^{19/\varepsilon}, \tag{3.2.39}$$

then by using $\binom{2d}{k} \leq (2d)^k$ and the fact that $\sum_{k\geq 1} u^k \leq 2u$ for $u \leq \frac{1}{2}$,

$$2\exp\left(\frac{19\sqrt{d}}{\varepsilon}\right)\sum_{k=1}^{2d}\binom{2d}{k}\left(\frac{K^{\sqrt{d}}d}{(\log d)^{\sqrt{d}}}\right)^k \leq 4e^{\frac{19\sqrt{d}}{\varepsilon}}\left(\frac{K^{\sqrt{d}}d^2}{(\log d)^{\sqrt{d}}}\right) \leq \frac{4d^2}{2^{\sqrt{d}}} \leq \frac{1}{2}. \tag{3.2.40}$$

Therefore, we obtain the conclusion of Lemma 3.2.8, by setting $d_1(\varepsilon)$ to satisfy (3.2.36) and (3.2.39). (Condition (3.2.30) is absorbed into (3.2.39).) □

We conclude this section by proving Lemma 3.2.7. The idea of splitting the probability as (3.2.28) is used, but here the computation is simpler than the previous one.

Proof of Lemma 3.2.7. 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

$$\mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i \geq \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right) \leq \mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i' \geq \frac{d}{3}\right) \leq \frac{1}{2}d^{-\frac{1}{10}\sqrt{d}}\left(\log\left(d^{\frac{1}{10}}\right)\right)^{-2}. \quad (3.2.41)$$

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 (3.2.41) by

$$\mathbb{P}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} Z_i' \geq \frac{d}{3}\right) \leq \sum_{k=1}^{(1+\frac{\varepsilon}{6})d} \binom{(1+\frac{\varepsilon}{6})d}{k} \mathbb{P}\left(\sum_{i=1}^{k} Z_i' \geq \frac{d}{6}\ ;\ Z_i' \geq \frac{1}{10}\right). \quad (3.2.42)$$

We can go through similar steps as in (3.2.31) to see that

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{k} Z_i' \geq \frac{d}{6}\ ;\ Z_i' \geq \frac{1}{10}\right) &\leq \sum_{\substack{u_1,\ldots,u_k\geq 0\\ u_1+\ldots+u_k=\lfloor\frac{5}{3}d-k\rfloor}} \prod_{i=1}^{k} \mathbb{P}\left(Z_i' \geq \frac{u_i}{10}\right) \\ &\leq \sum_{\substack{u_1,\ldots,u_k\geq 0\\ u_1+\ldots+u_k=\lfloor\frac{5}{3}d-k\rfloor}} 5^k \prod_{i=1}^{k}\left(\frac{2}{\varepsilon}+\frac{u_i}{10}\right)^{-\sqrt{d}}.\end{aligned}$$

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 (3.2.42) 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 (3.2.41) for $d \geq d_1(\varepsilon)$. □

3.3 Total infections at leaves on trees

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

$$l = \max\{\mathrm{dist}(\rho, v) : v \in T\}. \tag{3.3.1}$$

As discussed in Section 3.1.2, 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 3.1.2. In particular, the goal of this section is to establish the following theorem:

Theorem 3.3.1. *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 (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{T}^l \sim \mathsf{GW}(\xi)^l$ that*

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{T}^l) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2, \tag{3.3.2}$$

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 3.3.1 and 3.3.2. Then, in Section 3.3.3, we verify the theorem, which is along the lines of proving Theorem 3.2.4.

3.3.1 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 (3.2.1).

Proposition 3.3.2. *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.*

$$M^l(T) \leq \lambda \sum_{i=1}^{D} M^{l-1}(T_i) \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda S(T_j)). \tag{3.3.3}$$

Proof. Recall the processes $\mathbf{CP}^{\otimes}(T;\mathbf{1}_{v_i})$ and $(X_t^\sharp)$ defined in the proof of Proposition 3.2.1. Let $\mathbf{S}^\sharp(T)$ and $\mathbf{S}_i^\otimes(T)$ be the excursion time of $(X_t^\sharp)$ and $(X_{i,t}^\otimes) \sim \mathbf{CP}^{\otimes}(T;\mathbf{1}_{v_i})$, respectively, and set $\mathcal{L} = \{v \in T : \mathrm{dist}(v,\rho) = l\}$. 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 3.1.2.

- 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, from the viewpoint of (2.2.1), 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

$$M^\sharp(T) = \lambda D \times \frac{\sum_{i=1}^D M_i^\otimes(T)}{D} = \lambda D M^\otimes(T), \tag{3.3.4}$$

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 (3.3.8) and (3.3.9),

$$\begin{aligned}\lim_{t_0\to\infty} \frac{1}{t_0}\mathbb{E}_{\mathrm{CP}}\left[\,\left|\left\{s \in [0,t_0] :\ X_s^\otimes(v) = 1 \text{ and } X_{s-}^\otimes(v) = 0\right\}\right|\,\right] \\ = \frac{M^\otimes(T)}{(\lambda D)^{-1} + S^\otimes(T)}.\end{aligned} \tag{3.3.5}$$

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

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

Thus, by the same reasoning as above, the l.h.s. of (3.3.5) is also equal to

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

Therefore, by (2.2.5), (3.3.5) implies that

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

Finally, combining this with (3.3.4) and $M^\sharp(T) \ge M^l(T)$, we deduce the conclusion. □

3.3.2 The second recursive inequality

The goal of this section is to obtain the recursive inequality on $M^l(T)$ that is in parallel to (3.2.5). In Definition 3.1.2, $\mathbf{M}^l(T)$ and $M^l(T)$ were defined with respect to the root-added contact process $\mathbf{CP}_{\rho^+}^\lambda(T; \mathbf{1}_\rho)$. 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 3.1.2.

- 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 $\bar{\mathbf{M}}_v^l(T)$ to be the *total infections at* v, that is,

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

  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 $\mathsf{CP}^\lambda(T; \mathbf{1}_\rho)$. 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 3.2.2 on $R(T)$.

Corollary 3.3.3. *Under the above setting, let $R(\cdot) = \mathbb{E}_{\mathrm{CP}} \boldsymbol{R}(\cdot)$ and assume that*

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

Then, we have

$$\bar{M}^l(T) \le \frac{\lambda \sum_{i=1}^{D} \bar{M}^{l-1}(T_i)}{1 - \lambda^2 \sum_{i=1}^{D} R(T_i)}. \tag{3.3.6}$$

Proof. When running $\mathsf{CP}^\lambda(T; \mathbf{1}_\rho)$, recall the notion of the *first and second round* discussed in the proof of Proposition 3.2.2. 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

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

which leads to our conclusion. □

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 3.2.3. Recall the definition of $S(T)$ from Definition 3.1.1.

Corollary 3.3.4. *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*

$$M^l(T) \le (1 + \lambda S(T))\bar{M}^l(T). \tag{3.3.7}$$

Proof. We rely on a similar idea as in Proposition 3.2.3. Let $\mathcal{L} = \{v \in T : \mathrm{dist}(v, \rho) = l\}$ and $(X_t) \sim \mathbf{CP}^{\lambda}_{\rho^+}(T; \mathbf{1}_\rho)$ be the root-added contact process. For $v \in \mathcal{L}$, consider the following quantity:

$$\lim_{t_0 \to \infty} \frac{1}{t_0} \sum_{v \in \mathcal{L}} \mathbb{E}_{\mathrm{CP}} \left[\left| \{s \in [0, t_0] : X_s(v) = 1 \text{ and } X_{s-}(v) = 0\} \right| \right]. \tag{3.3.8}$$

First, by the renewal theorem (see, e.g., [Dur10, Theorem 2.4.6]), observe that the above is the same as

$$\lim_{t_0 \to \infty} \frac{\text{the number of excursions in } (X_t) \text{ during } [0, t_0]}{t_0} \times M^l(T) = \frac{M^l(T)}{\lambda^{-1} + S(T)}, \tag{3.3.9}$$

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 (3.3.9) is at most $\lambda \bar{M}^l(T)$, by following the discussion in Proposition 3.2.3. □

Let us combine (3.3.6) and (3.3.7) to deduce a recursion on $M^l(T)$ as follows.

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

By plugging in (3.2.5) in the r.h.s., we obtain the first recursion for $M^l(T)$, namely,

$$M^l(T) \le \frac{\lambda \sum_{i=1}^D M^{l-1}(T_i)}{1 - \lambda - 2\lambda^2 \sum_{i=1}^D S(T_i)}, \tag{3.3.10}$$

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

3.3.3 Recursive tail estimate for Galton-Watson trees

The goal of this subsection is to establish Theorem 3.3.1. As one may expect, we go through similar steps as Theorem 3.2.4 with some appropriate adjustments.

Let ξ be a positive, integer-valued random variable with mean $d = \mathbb{E}\xi$, and let $(\mathcal{T}^l, \rho) \sim \mathsf{GW}(\xi)^l$. Further, we denote the children of the root by $v_1, \dots, v_D$ 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 3.1.2). 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) $2/\varepsilon \le t \le d^{\frac{1}{10}}$;
2. (intermediate) $d^{\frac{1}{10}} \le t \le \exp(\frac{1}{2}c_1\sqrt{d})$;
3. (large) $\exp(\frac{1}{2}c_1\sqrt{d}) \le t$.

(The reason for the choice of $d^{\frac{1}{10}}$ is explained in Remark 3.2.6.) For convenience, we define

$$\widetilde{M}(\mathcal{T}^{l+1}) = \Big(1 - \frac{\varepsilon}{10}\Big)^{-(l+1)} M^{l+1}(\mathcal{T}^{l+1}), \quad \text{and} \quad \widetilde{M}(T_i) = \Big(1 - \frac{\varepsilon}{10}\Big)^{-l} M^l(T_i). \tag{3.3.11}$$

3.3.4 Tail estimate for total infections: small t

To establish Theorem 3.3.1, we begin with case 1: small t. Suppose that

$$\sum_{i=1}^{D} \widetilde{M}(T_i) \le (2+\varepsilon)d, \quad \text{and} \quad \sum_{i=1}^{D} S(T_i) \le \frac{2d}{\varepsilon}\Big(1 + \frac{\varepsilon}{3}\Big).$$

Then, the second recursive inequality (3.3.10) tells us that

$$\widetilde{M}(\mathcal{T}^{l+1}) \le \Big(1 - \frac{\varepsilon}{10}\Big)^{-1} \frac{(1-\varepsilon)(2+\varepsilon)}{1 - \frac{1}{d} - \frac{4}{\varepsilon d}(1 + \frac{\varepsilon}{3})} \le 2,$$

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

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \left(1+\frac{\varepsilon}{6}\right)d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} \widetilde{M}(T_i) \geq (2+\varepsilon)d\right) \\ + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} S(T_i) \geq \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right),\end{aligned} \tag{3.3.12}$$

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 (1.5.2), namely,

$$\mathbb{P}\left(D \geq \left(1+\frac{\varepsilon}{6}\right)d\right) \leq \exp(-c_{\varepsilon'}d) \leq \frac{1}{4}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}, \tag{3.3.13}$$

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 3.2.7 controls the last term in the r.h.s. of (3.3.12), implying the bound (3.2.10). For the second term, we claim that there exists $d_1(\varepsilon) > 0$ such that if $d \geq d_1(\varepsilon)$, then

$$\mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{(1+\frac{\varepsilon}{6})d} \widetilde{M}(T_i) \geq (2+\varepsilon)d\right) \leq \frac{1}{4}d^{-\frac{1}{10}\sqrt{d}}\left(\frac{1}{10}\log d\right)^{-2}. \tag{3.3.14}$$

Indeed, almost the same argument from the proof of Lemma 3.2.7 can be applied to deduce (3.3.14). 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 3.2.7 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 (3.3.2) for $l+1$ and $2 \leq t \leq d^{\frac{1}{10}}$ by combining (3.3.13), Lemma 3.2.7 and (3.3.14).

3.3.5 Tail estimate for total infections: intermediate t

For $t \geq d^{\frac{1}{10}}$, we rely on the second recursive inequality (3.3.3). However, one issue here is that the the quantity in the r.h.s. of (3.3.3) is no longer a single product of i.i.d. random variables as (3.2.1), (3.2.14). To overcome this difficulty, we define

$$W_i := \max\left\{\widetilde{M}(T_i),\ S(T_i)\right\}, \quad \text{for each } i \in \mathbb{N}, \tag{3.3.15}$$

for $\widetilde{M}(T_i)$ as in (3.3.11). Then, from a little bit of algebra we see that

$$\begin{aligned}\lambda \sum_{i=1}^{D} \widetilde{M}(T_i) \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda S(T_j)) \ &\leq \ \sum_{i=1}^{D} \lambda W_i \prod_{\substack{1 \leq j \leq D \\ j \neq i}} (1 + \lambda W_j) \\ &\leq \ \prod_{i=1}^{D} (1 + 2\lambda W_i).\end{aligned}$$

Based on this observation, we attempt to control

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) &\leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D} \log(1 + 2\lambda W_i) \ \geq \ \log\left\{\left(1 - \frac{\varepsilon}{10}\right) t\right\}\right) \\ &\leq \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D} \log(1 + 2\lambda W_i) \ \geq \ \log\left(\frac{t}{2}\right)\right).\end{aligned} \tag{3.3.16}$$

We are also aware of the tail estimate on W_i by the induction hypothesis and Theorem 3.2.4, namely,

$$\mathbb{P}(W_i \geq s) \leq \mathbb{P}\left(\widetilde{M}(T_i) \geq s\right) + \mathbb{P}(S(T_i) \geq s) \leq 2s^{-\sqrt{d}} (\log s)^{-2}, \quad \text{for all } s \geq \frac{2}{\varepsilon}. \tag{3.3.17}$$

This falls into the assumption of Lemma 3.2.8. Therefore, we deduce the tail estimate on $\widetilde{M}(\mathcal{T}^{l+1})$ for intermediate t by

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) &\leq \mathbb{P}(D \geq 2d) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1 + 2\lambda W_i) \ \geq \ \log\left(\frac{t}{2}\right)\right) \\ &\leq \exp(-c_1 d) + \frac{1}{2} t^{-\sqrt{d}} (\log t)^{-2} \leq t^{-\sqrt{d}} (\log t)^{-2},\end{aligned} \tag{3.3.18}$$

where we used the concentration condition (1.5.2) to bound the tail of D, and the last

inequality is true if $d \geq d_0'(\varepsilon)$ for appropriate constant $d_0'(\varepsilon) > 0$.

3.3.6 Tail estimate for total infections: large t

We continue to rely on W_i defined from (3.3.15). Set

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

For large t, we modify (3.3.16) as in (3.2.19), which is

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(\widetilde{M}(\mathcal{T}^{l+1}) \geq t\right) \leq \mathbb{P}\left(D \geq \Delta_{d,t}\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{2d} \log(1+2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right) \\ + \sum_{2d \leq r \leq \Delta_{d,t}} \mathbb{P}(D \geq r) \times \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{r} \log(1+2\lambda W_i) \geq \log\left(\frac{t}{2}\right)\right).\end{aligned} \tag{3.3.19}$$

Note that $t/2$ in the log in the r.h.s. is a lower bound of $(1 - \frac{\varepsilon}{10})t$ (cf. (3.3.18)). The first term in the r.h.s. can be controlled by the concentration condition (1.5.2). By (3.3.17), we can use Lemma 3.2.8 to obtain the estimate for the second, and the third term is bounded by the same reasoning as in Section 3.2.5. This concludes the proof of Theorem 3.3.1. □

3.4 Short survival on random graphs: proof of Theorem 7

Let μ be a given degree distribution satisfying (1.5.2) for some positive constants $\mathfrak{c} = \{c_\delta\}_{\delta \in (0,1]}$. Define $\widetilde{\mu}$ to be its size-biased distribution (1.4.2) 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 7, that is,

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

Theorem 3.4.1. *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 (1.5.2) 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*

$$\mathbb{P}_{\text{CP}}\left(\boldsymbol{R}(G_n) \le n^2 \;\middle|\; G_n \in \mathcal{E}_n\right) = 1 - o(1), \tag{3.4.1}$$

where $\boldsymbol{R}(G_n)$ *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 3.4.1, we extend the properties from Sections 3.2 and 3.3 to certain Galton-Watson-type random graphs with a cycle that are relevant to the local neighborhoods of G_n. Then in Section 3.4.2, we develop a coupling between the local neighborhoods and the aforementioned graphs in Section 3.4.1, following the ideas from Section 2.3.1. Finally, we conclude the proof of Theorem 3.4.1 and establish the lower bound of Theorem 7 in Section 3.4.3, based on all the properties we obtained in the previous sections.

3.4.1 Recursive analysis for unicyclic graphs

In this subsection, we do the final preliminary work before delving into the proof of Theorem 7. 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 3.1.2 for a brief review, or Section 3.4.2 for a detailed explanation). Therefore, we are interested in the Galton-Watson type processes with a single cycle, particularly the ones which are introduced in this subsection. Recall the **GWC**-process defined in Definition 2.3.6. In the following, we introduce the **GWC′**-process that is very similar to **GWC**-processes, which however is still necessary for building our argument.

Definition 3.4.2 (Galton-Watson-on-cycle process of the second type)**.** Let ξ be a positive, integer-valued random variable, and let $m, l \ge 1$ be nonnegative integers. Then, $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi,m)^l$, the *Galton-Watson-on-cycle process of the second type* (in short, **GWC′**-*process*), is generated according to the following procedure:

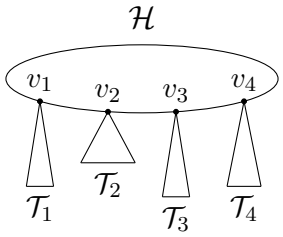


Figure 3.1: An example of $\mathcal{H} \sim \mathsf{GWC}'(\xi, m)^l$ with $m = 4$. Note that the depth of a subtree $\mathcal{T}_j$ is at most l (not necessarily equal to l).

1. Consider a length-m cycle $C = v_1 v_2 \dots v_m v_1$.

2. At each v_j for $j = 1, \dots, m$, 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 3.1 illustrates a diagram of a GWC'-process.

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

Definition 3.4.3. Let $m, l \geq 1$ be integers, and H be a graph that consists of a length-m cycle $C = v_1 v_2 \dots v_m v_1$ and depth $\leq l$ trees $T_1, \dots, T_m$ rooted at $v_1, \dots, v_m$, respectively (recall the definition of *tree depth* in (3.3.1)).

1. The *root-added contact process* $\mathsf{CP}^{\lambda}_{v_1^+}(H; x)$ 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 $\mathsf{CP}^{\lambda}_{v_1^+}(H; \mathbf{1}_{v_1})$ (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 $\mathcal{L}_j = \{v \in T_j : \mathrm{dist}(v, v_j) = l\}$, $(X_t) \sim \mathsf{CP}^{\lambda}_{v_1^+}(H; \mathbf{1}_{v_1})$ 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

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

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

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

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 3.4.4. *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 (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi, m)^l$ that*

$$\begin{aligned} \mathbb{P}_{\text{GW}}\left(S(\mathcal{H}^{m,l}) \geq t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq \frac{2}{\varepsilon}; \\ \mathbb{P}_{\text{GW}}\left(M^l(\mathcal{H}^{m,l}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2, \end{aligned} \tag{3.4.2}$$

where $S(\mathcal{H}^{m,l})$, $M^l(\mathcal{H}^{m,l})$ are given as in Definition 3.4.3.

The main idea of the proof is the same as what we saw in Theorems 3.2.4 and 3.3.1. 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 3.4.4 is presented in Section 3.5.

Recall the original **GWC**-process defined in Definition 2.3.6. In the following, we define the root-added contact process for **GWC**-processes, and introduce its excursion time and total infections.

Definition 3.4.5 (Root-added contact process on **GWC**-processes)**.** For $\dot{\mathcal{H}} \sim \mathbf{GWC}(\xi, m)^l$, 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 $\mathbf{CP}^{\lambda}_{v_1}(\dot{\mathcal{H}}^{m,l}; x)$, for an initial configuration $x \in \{0,1\}^{\dot{\mathcal{H}}\setminus\{v_1\}}$. 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 3.4.3, with respect to $\mathbf{CP}^{\lambda}_{v_1}(\dot{\mathcal{H}}; \mathbf{1}_{v_2})$ (resp., $\mathbf{CP}^{\lambda}_{v_1}(\dot{\mathcal{H}}; \mathbf{1}_{v_m})$). 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 3.4.4 can be derived on **GWC**-process as follows.

Corollary 3.4.6. *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 (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\dot{\mathcal{H}}^{m,l} \sim \mathbf{GWC}(\xi, m)^l$ that*

$$\begin{aligned} \mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}^{m,l}) \geq t\right) &\leq 4t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq \frac{3}{\varepsilon}; \\ \mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}^{m,l}) \geq \left(1-\frac{\varepsilon}{10}\right)^l t\right) &\leq 4t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 3, \end{aligned} \tag{3.4.3}$$

where $S(\dot{\mathcal{H}}^{m,l})$, $M^l(\dot{\mathcal{H}}^{m,l})$ are given as in Definition 3.4.5.

As before, we explain the proof of the corollary in Section 3.5.3 due to its technicality.

3.4.2 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 $\mathbf{GW}(\mu, \widetilde{\mu})^l$ 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 $\mathbf{GW}(\mu, \widetilde{\mu})^l$ as $n \to \infty$, as we briefly saw in Lemma 1.4.5 and the explanation below it. However, the standard coupling between $N(v,l)$ and $\mathbf{GW}(\mu, \widetilde{\mu})^l$ always has an error at least $\Theta(n^{-1})$. To diminish this error, we recall the notion of *augmented distribution* (Definition 2.3.2), which has allowed us to stochastically dominate $N(v,l)$ by a larger geometry. For our purpose, we introduce a

refined version of it which will be slightly different in order 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 the proof of Lemma 2.3.3 in Section 2.7.1).

Definition 3.4.7 (Augmented distribution, an enhanced version)**.** 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, which is an analogue of Lemma 2.3.3

Lemma 3.4.8. *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 (1.5.2) 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*

$$\begin{aligned} &\mathbb{P}_{D^\sharp\sim\nu^\sharp}\left(D^\sharp \geq (1+\varepsilon')d\right) \leq \exp\left(-c'_{\varepsilon'} d\right), \quad \text{for all } \varepsilon' \in \left[\frac{\varepsilon}{10}, 1\right]; \\ &\mathbb{P}_{D^\sharp\sim\nu^\sharp}\left(D^\sharp \geq (1+a)d\right) \leq \exp\left(-c'_1 a d\right), \quad \text{for all } a \geq 1. \end{aligned} \tag{3.4.4}$$

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 2.3.3. In Section 3.7.4, we discuss the proof of the lemma, focusing on the aspects which are different from Lemma 2.3.3.

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 $\mathcal{E}^C_n(\gamma)$ over the graphs with n vertices to be

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

Then, we adopt the following lemma that shows the event $\mathcal{E}^C_n(\gamma)$ is indeed typical for some

$\gamma > 0$. The proof of the lemma is presented in Section 3.7.5, due to its similarity to Lemma 2.3.5 and [LS10, Lemma 2.1].

Lemma 3.4.9. *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 (1.5.2) with $\mathfrak{c} = \{c_\delta\}_{\delta\in(0,1]}$, we have for $G_n \sim \mathcal{G}(n,\mu)$ that*

$$\mathbb{P}\left(G_n \in \mathcal{E}_n^C\left(\frac{1}{10\log d}\right)\right) = 1 - o(1). \tag{3.4.5}$$

Remark 3.4.10. In Lemma 2.3.5, we proved 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 (1.5.2) and it makes it possible to deduce Lemma 3.4.9, a stronger result. The improved information on the constant γ turns out to be crucial in Section 3.4.3 when proving Theorem 3.4.1.

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 $N(v, l_n)$. As we have seen in the previous chapter, this is done by the **EGW**-processes in Definition 2.3.7.

We develop a coupling between $N(v, \gamma_1 \log_d n)$ and the **EGW**-processes. Recall the notion of stochastic domination between probability measures (on rooted graphs) introduced in Section 2.3.1. Although the coupling scheme is developed very similarly as before, we describe the details for completeness as it involves more refined setting and different notations.

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.

Moreover, for a given $\varepsilon \in (0,1)$ we set $\mu^\sharp = \mu_\varepsilon^\sharp$ 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, we have the following

analogue of Lemma 2.3.8, which plays a crucial role in settling the lower bound of Theorem 7 in the following subsection.

Lemma 3.4.11. *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 7. The last thing we want is the tail probability estimates for **EGW**-processes. To be concrete, we begin with defining the quantities of interest.

Definition 3.4.12 (Excursion time and total infections at leaves for **EGW**-processes)**.** 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

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

  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 **EGW**-process $\mathcal{H}$, the tail probabilities of $S(\mathcal{H})$ and $M^l(\mathcal{H})$ can be estimated using Theorems 3.2.4, 3.3.1, and Proposition 3.4.4, which can be described as follows.

Proposition 3.4.13. *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 (1.5.2) with $\mathfrak{c}$, let $\mu^\sharp := \mu^\sharp_\varepsilon$, and set $\widetilde{\mu}^\sharp$ to be its size-biased distribution (Definition 3.4.7). Then, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{H} \sim \mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^l$ that*

$$\begin{aligned} \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}) \geq t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq \frac{2}{\varepsilon}; \\ \mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^l t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2, \end{aligned} \tag{3.4.6}$$

where $S(\mathcal{H})$, $M^l(\mathcal{H})$ are given as in Definition 3.4.12.

Proof. Along with (3.4.6), we establish the same inequalities for $\mathcal{H} \sim \mathbf{EGW}(\widetilde{\mu}^\sharp; h, m)^l$ in tandem, by an induction on h. When $h = 0$, we have (3.4.6) for both $\mathbf{EGW}(\widetilde{\mu}^\sharp; 0, m)^l$ and $\mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h, m)^l$, since the former one is the same as $\mathbf{GWC}'(\widetilde{\mu}^\sharp, m)^l$ and the latter is $\mathcal{H} \sim \mathbf{GWC}'(\mu^\sharp, \widetilde{\mu}^\sharp, m)^l$, the $\mathbf{GWC}'$-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 3.4.8, the mean $\tilde{d} := \mathbb{E}_{D^\sharp \sim \widetilde{\mu}^\sharp} D^\sharp$ of $\widetilde{\mu}^\sharp$ is such that

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

and $\widetilde{\mu}^\sharp$ also satisfies (3.4.4) with $\mathfrak{c}'$ as in the lemma. Note that this implies (1.5.2) for $\widetilde{\mu}^\sharp$. Therefore, Proposition 3.4.4 gives (3.4.6) for $h = 0$ and all $m \geq 2$, $l \geq 0$.

Suppose that we have (3.4.6) for $\mathcal{H}' \sim \mathbf{EGW}(\widetilde{\mu}^\sharp; h, m)^l$ with $m \geq 2$, $h \geq 0$ and $l \geq h+1$. Let ρ_1 be the root of $\mathcal{H}_1 \sim \mathbf{EGW}(\widetilde{\mu}; h+1, m)^l$ and $D_1 \sim \widetilde{\mu}^\sharp$ be its degree. Then the subgraphs $H^1_1, \ldots, H^1_{D_1}$ rooted at each child of v_1 are i.i.d. $\mathbf{EGW}(\widetilde{\mu}^\sharp; h, m)^{l-1}$. Since we have the tail probability estimates for $H^1_1, \ldots, H^1_{D_1}$ by the induction hypothesis, we follow the proof of Theorems 3.2.4 (Sections 3.2.3—3.2.5) and 3.3.1 (Sections 3.3.4—3.3.6) to deduce (3.4.6) for $\mathcal{H}_1$. The result for $\mathcal{H}_2 \sim \mathbf{EGW}(\mu^\sharp, \widetilde{\mu}^\sharp; h+1, m)^l$ follows similarly, where the only difference is that $\mathcal{H}_2$ has subgraphs $H^2_1, \ldots, H^2_{D_2} \sim$ i.i.d. $\mathbf{EGW}(\widetilde{\mu}^\sharp; h, m)^{l-1}$ with $D_2 \sim \mu^\sharp$. □

3.4.3 Short survival on random graphs: Proof of Theorem 7

In this subsection, we conclude the proof of the lower bound of Theorem 7, by showing Theorem 3.4.1.

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 3.2.4, 3.3.1 and Proposition 3.4.13. 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 3.4.9, that is,

$$\mathbb{P}_{G_n\sim\mathcal{G}(n,\mu)}\left(G_n \in \mathcal{E}_n^C\right) = 1 - o(1), \quad \text{with } \mathcal{E}_n^C = \mathcal{E}_n^C\left(\frac{\gamma_1}{\log d}\right) \text{ as in the lemma,} \tag{3.4.7}$$

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 3.4.7) 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 3.4.11. Then, Lemma 3.4.11 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. See Lemma 2.3.8 for its proof.

Corollary 3.4.14. *Under the above setting, let v be a fixed vertex in G_n and η' denote the law of B_v. Then, we have*

$$\eta' \mathbf{1}_{\mathcal{E}_n^C} \leq_{\mathrm{ST}} b_0 \eta_0 + \sum_{h \geq 0, m \geq 2} \eta_{h,m}. \tag{3.4.8}$$

Further, let us consider the the collections of the *bottom leaves* of B_v. To this end, let C denote the cycle in B_v if exists, and let

$$\mathcal{L}(B_v) := \begin{cases} \{u \in B_v : \mathrm{dist}(u, v) \geq l_n\} & \text{if } C \text{ does not exist;} \\ \{u \in B_v : \mathrm{dist}(u, v) \geq l_n \text{ and } \mathrm{dist}(u, C) \geq l_n - h\} & \text{if } C \text{ exists,} \end{cases} \tag{3.4.9}$$

where $h = \mathrm{dist}(v, C)$. Also, set $\mathcal{L}(\mathcal{H})$ to be the bottom leaves of $\mathcal{H}$, where we follow Definition 3.4.12 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 3.2.4, 3.3.1 and Proposition 3.4.13 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 $\mathbf{CP}^{\lambda}(B_v; \mathbf{1}_v)$ (Definition 3.4.12), satisfy

$$\begin{aligned} &\mathbb{P}(R(B_v) \geq t) \leq \mathbb{P}\left(S(\mathcal{H}) \geq t\right) \leq 3t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for all } t \geq \frac{2}{\varepsilon}; \\ &\mathbb{P}\left(\bar{M}^{l_n}(B_v) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} t\right) \\ &\leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} t\right) \leq 3t^{-\sqrt{d}} (\log t)^{-2}, \quad \text{for all } t \geq 2. \end{aligned} \tag{3.4.10}$$

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

Lemma 3.4.15. *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 (3.4.10) implies

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

and also

$$\begin{aligned}
\mathbb{P}\left(\bar{M}^{l_n}(B_v) \geq \frac{1}{\log n}\right) &\leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} \frac{\exp(\frac{\varepsilon\gamma_1}{20}\log_d n)}{\log n}\right) \\
&\leq \mathbb{P}\left(M^{l_n}(\mathcal{H}) \geq \left(1 - \frac{\varepsilon}{10}\right)^{l_n} \exp\left(\frac{\varepsilon\gamma_1}{30}\log_d n\right)\right) \\
&\leq \exp\left(-\frac{\varepsilon\gamma_1\sqrt{d}}{30\log d}\log n\right) \leq n^{-2},
\end{aligned}$$

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 3.4.1.

Proof of Theorem 3.4.1. 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 3.4.9 and 3.4.15. Further, set $\gamma_1 \in (0, \frac{1}{10})$ as in (3.4.7) and define

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

Then, Lemmas 3.4.9 and 3.4.15 tell us that

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

In the remaining proof, we settle (3.4.1) for $\mathcal{E}_n$ as above.

Let $\mathbf{R}_v(G_n)$ be the survival time of $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$. Then, the standard coupling argument (graphical representation) tells us that

$$\mathbf{R}(G_n) = \max\{\mathbf{R}_v(G_n) : v \in G_n\} \le \sum_{v \in G_n} \mathbf{R}_v(G_n). \tag{3.4.11}$$

To investigate $\mathbf{R}_v(G_n)$, we introduce a *decomposition* of $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$ using the *blocks* $\{B_u\}_u$. Recall the graphical representation from Section 1.1.1 and let $\{N_x(t)\}_{x \in V(G_n)}$ and $\{N_{\vec{xy}}(t)\}_{\vec{xy} \in \vec{E}(G_n)}$ be the Poisson processes defining the recoveries and infections of $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$, respectively. Also, recall the definition of the bottom leaves $\mathcal{L}(B_u)$ defined in (3.4.9).

Definition 3.4.16 (Decomposition)**.** Let G_n and $\{B_u\}_{u \in G_n}$ be as above. The *decomposition* of $\mathbf{CP}^\lambda(G_n; \mathbf{1}_v)$ by $\{B_u\}_u$ is the coupled process generated as follows.

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

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 $\{(N_x(s))_{s \ge t}\}_{x \in V(B_u)}$ and $\{(N_{\vec{xy}}(s))_{s \ge t}\}_{\vec{xy} \in \vec{E}(B_u)}$.

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 shares 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 (3.4.10).

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

$\mathbf{R}_v(G_n)$ and hence

$$R_v(G_n) := \mathbb{E}_{\mathrm{CP}}\mathbf{R}_v(G_n) = \widetilde{R}_v(G_n) := \mathbb{E}_{\mathrm{CP}}\widetilde{\mathbf{R}}_v(G_n). \tag{3.4.12}$$

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

$$\widetilde{\mathbf{R}}_v(G_n) \le \sum_{u \in \mathbf{U}_v(G_n)} \mathbf{R}_u(B_u), \tag{3.4.13}$$

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

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

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

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 (3.4.12) and (3.4.13) that for all $v \in G_n$,

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

Thus, (3.4.11) gives $R(G_n) \le 2n^{1+2d^{-1/2}}$, and hence by Markov's inequality we get

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

for any d larger than some absolute constant. This finishes the proof of Theorem 3.4.1. □

3.5 Recursive analysis for unicyclic graphs

This section is devoted to the proof of Proposition 3.4.4 and Corollary 3.4.6. The main technical work needed here is to carry out the tree recursions as in Sections 3.2 and 3.3

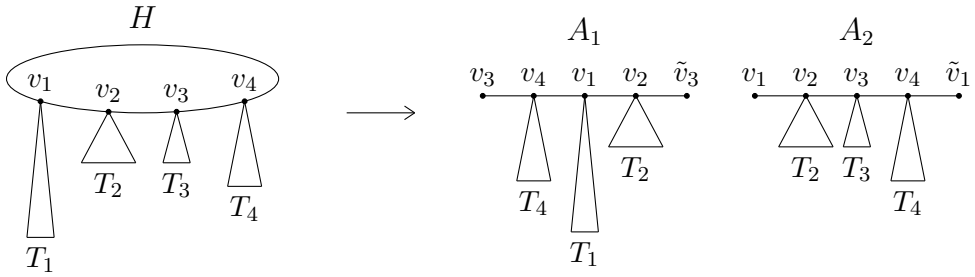


Figure 3.2: An example of the cover (A_1, A_2) of the unicyclic graph H with $m = 4$.

despite the presence of a cycle. As before, our approach is based on an induction argument, which differs depending on whether t in (3.4.2) is small or large.

In Section 3.5.1, we introduce an appropriate way of covering a unicyclic graph by trees. Based on this notion, in Section 3.5.2 we can establish (3.4.2) for *small* values of t. Then, we conclude the proof of Proposition 3.4.4 and Corollary 3.4.6 in tandem inductively in Section 3.5.3.

3.5.1 Reduction of unicyclic graphs to trees

Let H be a graph that consists of a length-m cycle $C = v_1 v_2 \dots v_m v_1$ and depth $\le l$ trees $T_1, \dots, T_m$ rooted at $v_1, \dots, v_m$, 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 3.5.1 (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 3.2):

- A_1 consists of a length-m line $v_{m'} \dots v_m v_1 v_2 \dots v_{m'-1} \tilde{v}_{m'}$ 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 $v_1 \dots v_m \tilde{v}_1$ 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 3.4.16. Recall the graphical representation of the contact process discussed

in Section 1.1.1, 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 the process $\mathbf{CP}^\lambda(A_1;\mathbf{1}_{v_1})$, 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.

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 3.4.16, 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 $v_1v_2\dots v_m$ 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 3.4.3. In addition, we consider $B(F^{m,l})$, the *total infections at the end* and its expectation as follows.

$$\mathbf{B}(F^{m,l}):=\left|\left\{s\in[0,\mathbf{S}(F^{m,l})]:\ X_s(v_m)=1 \text{ and } X_{s-}(v_m)=0\right\}\right|, \tag{3.5.1}$$

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

Lemma 3.5.2. *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 (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{F}^{m,l}$ as above that*

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{F}^{m,l}) \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq \tfrac{2}{\varepsilon}; \tag{3.5.2}$$

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{F}^{m,l}) \geq \left(1-\frac{\varepsilon}{10}\right)^l t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2; \tag{3.5.3}$$

$$\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m,l}) \geq d^{-\frac{3}{4}(m-1)}t\right) \leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2, \tag{3.5.4}$$

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 (3.5.2) and (3.5.3) is along the same lines as (3.2.4) and (3.3.1), so we postpone those parts of the proof to Section 3.7.1.

In the remaining proof, we show (3.5.4). When $m = 1$, $\mathcal{F}^{1,l}$ is a single vertex v_1 and we have the conclusion. Suppose that (3.5.4) 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 Propositions 3.2.2, we have

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

Thus, Proposition 3.2.3 and (3.2.5) tell us that

$$B(\mathcal{F}^{m+1,l}) \le \frac{\lambda B(F)}{1-\lambda-2\lambda^2\left(S(F)+\sum_{i=1}^{D} S(T_i)\right)}, \quad \text{if } \lambda+2\lambda^2\left(S(F)+\sum_{i=1}^{D} S(T_i)\right) < 1, \tag{3.5.5}$$

and hence we obtain the first recursive inequality.

The second inequality comes from the same idea as in Proposition 3.3.2:

$$B(\mathcal{F}^{m+1,l}) \le \lambda B(F)\prod_{i=1}^{D}(1+\lambda S(T_i)). \tag{3.5.6}$$

As before, we show (3.5.4) separately on different regimes of t. Details in the proof are also similar to the proof of Theorem 3.2.4.

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

Suppose that

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

Then, by (3.5.5), $B(\mathcal{F}^{m+1,l}) \le d^{-\frac{3m}{4}}$. Therefore, for any $t \ge 1$,

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m+1,l}) \ge d^{-\frac{3m}{4}}t\right) \le{}& \mathbb{P}\left(D \ge \left(1+\frac{\varepsilon}{6}\right)d\right) + \mathbb{P}\left(B(F) \ge d^{-\frac{3}{4}(m-1)+\frac{1}{2}}\right) \\ &+ \mathbb{P}_{\mathrm{GW}}\left(S(F)+\sum_{i=1}^{D} S(T_i) \ge \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right).\end{aligned} \tag{3.5.7}$$

By following the logics from Section 3.2.3 and relying on the induction hypothesis that tells us the tail probability of $B(F)$, we see that the r.h.s. of (3.5.7) is at most $t^{-\sqrt{d}}(\log t)^{-2}$ for all $2 \le t \le d^{\frac{1}{10}}$.

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

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 (3.5.6) and hence prove the following.

Claim 3.5.3. *Under the above setting, we have*

$$\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{F}^{m+1,l}) \geq d^{-\frac{3m}{4}}t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}\cdot\widehat{S} \geq \frac{1}{2}d^{\frac{1}{4}}t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq d^{\frac{1}{10}}. \tag{3.5.8}$$

Note that the first inequality is just a rewriting of (3.5.6).

The idea is to split the probability as follows: let $\alpha = d^{\frac{1}{10}}$, and we see that

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}\cdot\widehat{S} \geq t\right) \leq{}& \mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\alpha\right) + \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq \frac{t}{2\alpha^2}d^{\frac{1}{4}}\right) \\ &+ \sum_{k=0}^{k_0}\mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\alpha^{-k}\right)\cdot\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq \frac{\alpha^{k-1}}{2}d^{\frac{1}{4}}\right),\end{aligned} \tag{3.5.9}$$

where we set $k_0 = \lfloor\frac{\log t}{\log\alpha}\rfloor$-1, the point where $\alpha^{k+1} \leq t$ is the closest to t. Similar ideas as Lemma 3.2.8 yields the conclusion (3.5.8), but the computation is much simpler. We postpone the remaining details of Claim 3.5.3 to Section 3.7.2. □

3.5.2 Back to unicyclic graphs and tail estimates for small t

Here, we establish Proposition 3.4.4 for small t. In the previous subsection, the graph $\mathcal{F}^{m,l}$ is introduced to describe a way of decomposing **GWC′**-processes. Now we combine the pieces together and derive a recursive tail probability estimates for **GWC′**-processes.

For integers $m_1, m_2 \geq 1$ and $l \geq 0$, 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 $v_{-m_1}\dots v_{-1}v_0v_1\dots v_{m_2}$ and i.i.d. trees $\mathcal{T}_j \sim \mathbf{GW}(\xi)^l$ rooted at v_j for $j \in \{-m_1+1,\dots,m_2-1\}$. 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 3.4.3, and $B(\mathcal{A}^{\underline{\mathrm{m}},l})$ as (3.5.1) but considering the infections at both endpoints v_{-m_1}, v_{m_2} (not just a single endpoint as in (3.5.1)). Under this setting, we first establish the tail probability estimates for $\mathcal{A}^{\underline{\mathrm{m}},l}$.

Lemma 3.5.4. *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 (1.5.2) with $\mathfrak{c}$, we have for $\lambda = (1-\varepsilon)d^{-1}$ and $\mathcal{A}^{\underline{m},l}$ as above that*

$$\begin{aligned}
\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{A}^{\underline{m},l}) \geq t\right) &\leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq \tfrac{2}{\varepsilon}; && (3.5.10)\\
\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{A}^{\underline{m},l}) \geq \left(1-\frac{\varepsilon}{10}\right)^l t\right) &\leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2; && (3.5.11)\\
\mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{A}^{\underline{m},l}) \geq d^{-\frac{3}{4}(m_1\wedge m_2)}t\right) &\leq t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } t \geq 2, && (3.5.12)
\end{aligned}$$

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 3.5.2, by a straight-forward applications of the recursive argument from Theorems 3.2.4 and 3.3.1. To prevent repeating the same argument many times, we postpone the proof until Section 3.7.3.

We conclude this subsection by settling the following lemma, which establishes Proposition 3.4.4 for $t \leq d^{\frac{1}{10}}$.

Lemma 3.5.5. *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 3.4.4, we have*

$$\begin{aligned}
\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}^{m,l}) \geq t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } \frac{2}{\varepsilon} \leq t \leq d^{\frac{1}{10}};\\
\mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{H}^{m,l}) \geq \left(1-\frac{\varepsilon}{10}\right)^l t\right) &\leq 3t^{-\sqrt{d}}(\log t)^{-2} \quad \text{for all } 2 \leq t \leq d^{\frac{1}{10}}.
\end{aligned} \tag{3.5.13}$$

Proof. Let $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi, m)^l$, 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 3.5.1. 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 $A_1, A_2 \sim \mathcal{A}^{\underline{\mathrm{m}},l}$ for $\underline{\mathrm{m}} = (m_1, m_2)$, with $\mathcal{A}^{\underline{\mathrm{m}},l}$ defined as in the

beginning of Section 3.5.2. Thus, we have (3.5.10), (3.5.11) and (3.5.12), 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 3.4.16). According to its formulation, the quantities $R(\mathcal{H}^{m,l})$ and $\bar{M}^l(\mathcal{H}^{m,l})$ (see Definition 3.4.3 for their definitions) can be estimated as follows.

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

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

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

Observe that Proposition 3.2.3 and Corollary 3.3.4 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

$$\begin{aligned} S(\mathcal{H}^{m,l}) &\leq \frac{S(A_1) + B(A_1)S(A_2)}{1 - B(A_1)B(A_2) - \lambda S(A_1) - \lambda B(A_1)S(A_2)}; \\ M^l(\mathcal{H}^{m,l}) &\leq \frac{M^l(A_1) + B(A_1)M^l(A_2)}{1 - B(A_1)B(A_2) - \lambda S(A_1) - \lambda B(A_1)S(A_2)}, \end{aligned}$$

To establish the first inequality of (3.5.13), we first observe that if

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

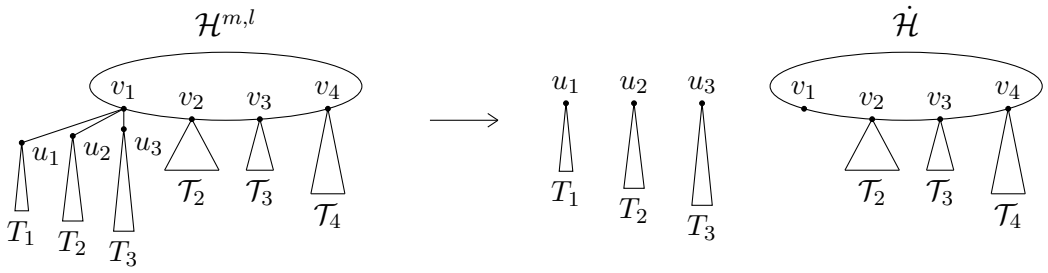


Figure 3.3: An illustration of dividing $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi, m)^l$ into $T_1, T_2, T_3 \sim \mathsf{GW}(\xi)^{l-1}$ and $\dot{\mathcal{H}} \sim \mathsf{GWC}(\xi, m)^l$, given that $m = 4$ and $D = 3$.

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

$$\mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{H}^{m,l}) \ge t\right) \le \mathbb{P}_{\mathrm{GW}}\left(S(A_1) \ge \left(1 - d^{-\frac{5}{9}}\right) t\right) + \mathbb{P}_{\mathrm{GW}}\left(S(A_2) \ge \left(1 - d^{-\frac{5}{9}}\right) t\right) + \mathbb{P}_{\mathrm{GW}}\left(B(A_1) \ge d^{-\frac{5}{8}}\right) + \mathbb{P}_{\mathrm{GW}}\left(B(A_2) \ge d^{-\frac{5}{8}}\right).$$

Since $(1 - d^{-\frac{5}{9}})^{-\sqrt{d}} \le 1 + 2d^{-\frac{1}{18}}$, applying Lemma 3.5.4 gives us that the r.h.s. of the above is bounded by $3t^{-\sqrt{d}}(\log t)^{-2}$ for $2/\varepsilon \le t \le d^{\frac{1}{10}}$. We can also do the same thing for $M^l(\mathcal{H}^{m,l})$ to deduce the conclusion. □

3.5.3 Tail estimates on unicyclic graphs for large t: Overview

In this section, we conclude the proof of Proposition 3.4.4 and also verify Corollary 3.4.6. Consider $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi, m)^l$.

For a GWC'-process $\mathcal{H}^{m,l} \sim \mathsf{GWC}'(\xi, m)^l$, we can derive a recursive inequality on $\mathcal{H}^{m,l}$ similar to (3.2.1) using GWC-process. As described in Figure 3.3 let $v_1 \dots v_m v_1$ 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 $T_1, \dots, T_D$ be the subtrees rooted at $u_1, \dots u_D$, the children of v_1 in $\mathcal{T}_1$. Further, let $\dot{\mathcal{H}}$ be the graph that consists of the cycle $v_1 \dots v_m v_1$ and the trees $\mathcal{T}_j$ rooted at v_j for $j \in \{2, \dots, m\}$, which is $\dot{\mathcal{H}} \sim \mathsf{GWC}(\xi, m)^l$. Also, note that $T_1, \dots, T_D \sim$ i.i.d. $\mathsf{GW}(\xi)^{l-1}$. Then, the following lemma shows the connection between GWC'- and GWC-processes. Its proof is based on the same ideas as Proposition 3.2.1, and it will be presented in Section 3.5.5.

Lemma 3.5.6. *Under the above setting, we have*

$$S(\mathcal{H}^{m,l}) \leq (1+2\lambda S(\dot{\mathcal{H}}))\prod_{i=1}^{D}(1+\lambda S(T_i)). \tag{3.5.14}$$

This shows that understanding the tail probabilities on **GWC**-processes is helpful in estimating those on **GWC′**-processes. The following proposition gives necessary estimates for **GWC**-processes.

We prove Proposition 3.4.4 and Corollary 3.4.6 in tandem. This is done by an inductive argument which we detail below. Let $l \geq 1$ be a fixed integer. We establish both (3.4.2) and (3.4.3) for all m via the following steps.

1. Show that (3.4.2) (resp., (3.4.3)) are true for $m = 1$ (resp., $m = 2$).
2. For each integer $m \geq 2$, prove that if (3.4.3) is true for m, then we have (3.4.2) for m.
3. For each integer $m \geq 2$, prove that if (3.4.2) and (3.4.3) are true for m, then we have (3.4.3) for $m+1$.
   - 3-1. Show that (3.4.3) holds when $t \leq d^{\frac{1}{10}}$ (for $m+1$).
   - 3-2. Show that (3.4.3) holds when $t \geq d^{\frac{1}{10}}$ (for $m+1$).

3.5.4 Tail estimates on unicyclic graphs for large t: Step 1

Proving (3.4.2) for $m = 1$ is straight-forward, since $\mathcal{H}^{1,l} \sim \mathbf{GW}(\xi)^l$ and the result follows from Theorems 3.2.4 and 3.3.1. To establish (3.4.3) for $m = 2$, we divide (3.4.3) into two cases, $t \leq d^{\frac{1}{10}}$ and $t \geq d^{\frac{1}{10}}$. Observe that $\dot{\mathcal{H}} \sim \mathbf{GWC}(\xi, 2)^l$ is the same as $(\mathcal{T}_2, v_2) \sim \mathbf{GW}(\xi)^l$ 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 3.4.5, an analog of Proposition 3.2.3 tells us that

$$\frac{S(\dot{\mathcal{H}})}{1+2\lambda S(\dot{\mathcal{H}})} \leq R(\mathcal{T}_2) \leq S(\mathcal{T}_2), \quad \text{and} \quad \frac{M^l(\dot{\mathcal{H}})}{1+2\lambda S(\dot{\mathcal{H}})} \leq \bar{M}^l(\mathcal{T}_2) \leq M^l(\mathcal{T}_2). \tag{3.5.15}$$

The left inequality can be rewritten as

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

Based on this information and on Theorems 3.2.4 and 3.3.1, we deduce that for $3/\varepsilon \le t \le d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}) \ge t\right) \le \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}_2) \ge \left(1-d^{-\frac{3}{4}}\right)t\right) \le 2t^{-\sqrt{d}}(\log t)^{-2}, \tag{3.5.16}$$

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

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \ge \left(1-\frac{\varepsilon}{10}\right)^l t\right) \le{}& \mathbb{P}_{\mathrm{GW}}\left(M^l(\mathcal{T}_2) \ge \left(1-d^{-\frac{3}{4}}\right)\left(1-\frac{\varepsilon}{10}\right)^l t\right)\\ &+ \mathbb{P}_{\mathrm{GW}}\left(S(\mathcal{T}_2) \ge d^{\frac{1}{5}}\right),\end{aligned}$$

and the r.h.s. is bounded by $2t^{-\sqrt{d}}(\log t)^{-2}$.

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

We can deduce the conclusion by exactly the same argument as Theorem 3.2.4 (Sections 3.2.4, 3.2.5) and Theorem 3.3.1 (Sections 3.3.5, 3.3.6). Indeed, the recursive inequalities (3.2.1), (3.3.3) do not change for $\dot{\mathcal{H}}$ since they are independent of the number of connections between v_1 and v_2.

3.5.5 Tail estimates on unicyclic graphs for large t: Step 2

We have already established (3.4.2) for $t \le d^{\frac{1}{10}}$ in Lemma 3.5.5. Hence, we focus on settling the other part, $t \ge d^{\frac{1}{10}}$. We begin with presenting the proof of Lemma 3.5.6, which shows the relation between **GWC′**- and **GWC**-processes.

Proof of Lemma 3.5.6. 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 $\mathbf{CP}^{\lambda}_{v_1^+}(\mathcal{H}^{m,l};\mathbf{1}_{v_1})$, we adopt the analogous idea as Proposition 3.2.1 by ignoring the recoveries at v_1 except when all the other vertices are healthy. This translates the original process $\mathbf{CP}^{\lambda}_{v_1^+}(\mathcal{H}^{m,l};\mathbf{1}_{v_1})$ to the product chain of $\mathbf{CP}^{\lambda}_{v_1}(T_i)$, $i \in [D]$ and $\mathbf{CP}^{\lambda}_{v_2}(\dot{\mathcal{H}})$ in the perspectives of Proposition 3.2.1. In $\mathbf{CP}^{\lambda}_{v_2}(\dot{\mathcal{H}};\mathbf{0})$, 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 observation, we follow the logic from Proposition 3.2.1 (and Lemma 2.2.3) to obtain (3.5.14). □

Having Lemma 3.5.6 and (3.4.3) for m, we can establish the first inequality of (3.4.2) by the same reasoning as Sections 3.2.4 and 3.2.5 (Lemma 3.2.8 and Corollary 3.2.9).

For the second line of (3.4.2), we have the following recursive inequality for $M^l(\mathcal{H}^{m,l})$ which can be derived analogously as Proposition 3.3.2.

$$M^l(\mathcal{H}^{m,l}) \leq 2\lambda M^l(\dot{\mathcal{H}}) \prod_{i=1}^{D}(1+\lambda S(T_i)) + \lambda(1+2\lambda S(\dot{\mathcal{H}})) \sum_{i=1}^{D} M^{l-1}(T_i) \prod_{\substack{j=1 \\ j\neq i}}^{D}(1+\lambda S(T_j)). \tag{3.5.17}$$

As done in Section 3.3.5, 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 (3.4.3) and Theorem 3.3.1 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 3.2.8 and Corollary 3.2.9. Therefore, we conclude the proof of (3.4.2) by obtaining that for all $t \geq d^{\frac{1}{10}}$,

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \geq \left(1-\frac{\varepsilon}{10}\right)^l t\right) \leq \mathbb{P}_{\mathrm{GW}}\left(2\prod_{i=0}^{D}(1+4\lambda W_i) \geq t\right) \leq 4t^{-\sqrt{d}}(\log t)^{-2}.$$

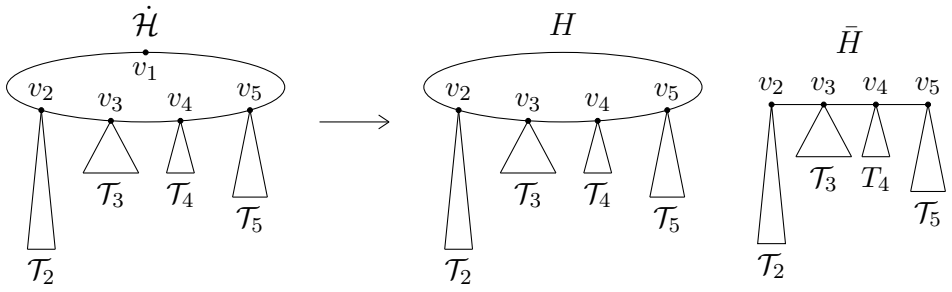


Figure 3.4: A description of $\dot{\mathcal{H}} \sim \mathsf{GWC}(\xi, m+1)^l$, H and $\bar{H}$ in Section 3.5.6 with $m = 4$.

3.5.6 Tail estimates on unicyclic graphs for large t: Step 3-1

Let $\dot{\mathcal{H}} \sim \mathsf{GWC}(\xi, m+1)^l$, $v_1 v_2 \dots v_{m+1} v_1$ be its cycle part with root $\rho = v_1$ and $\mathcal{T}_2, \dots, \mathcal{T}_{m+1}$ be the i.i.d. $\mathsf{GW}(\xi)^l$ trees rooted at $v_2, \dots, v_{m+1}$, respectively. As described in Figure 3.4, let H and $\bar{H}$ be the graph defined as follows:

- H consists of a cycle $v_2 \dots v_{m+1} v_2$ and the trees $\mathcal{T}_j$ rooted at v_j, $j \in \{2, \dots, m+1\}$.
- $\bar{H}$ is obtained by removing the edge $(v_2 v_{m+1})$ from H. In other words, it consists of a length-$(m-1)$ path $v_2 \dots v_{m+1}$ and the trees $\mathcal{T}_j$ rooted at v_j, $j \in \{2, \dots, m+1\}$.

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 $\mathsf{CP}^\lambda_{v_1}(\dot{\mathcal{H}}; \mathbf{0})$, 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 3.4.16:

- 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 $\mathsf{CP}^\lambda_{v_1}(\dot{\mathcal{H}}, \mathbf{0})$.

Therefore, for $S(\dot{\mathcal{H}}), M^l(\dot{\mathcal{H}})$ defined as Definition 3.4.5, the argument from Proposition 3.2.3 and Corollary 3.3.4 implies

$$\frac{S(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \le \frac{R_2(\bar{H}) + R_{m+1}(\bar{H})}{2}, \quad \text{and} \quad \frac{M^l(\dot{\mathcal{H}})}{1 + 2\lambda S(\dot{\mathcal{H}})} \le \frac{\bar{M}_2^l(\bar{H}) + \bar{M}_{m+1}^l(\bar{H})}{2}. \tag{3.5.18}$$

Note that $H \sim \mathsf{GWC}'(\xi, m)^l$, and let $S_i(H)$ and $M_i^l(H)$ be defined as Definition 3.4.3,

with respect to the root-added process $\mathsf{CP}^{\lambda}_{v_i^+}(H;\mathbf{1}_{v_i})$. 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 (3.5.18) 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 (3.5.15) and (3.5.16). 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

$$\begin{aligned}\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}})\ge\left(1-\frac{\varepsilon}{10}\right)^l t\right)&\le\mathbb{P}_{\mathrm{GW}}\left(M^l(H)\ge\left(1-d^{-\frac{3}{4}}\right)\left(1-\frac{\varepsilon}{10}\right)^l t\right)\\&\quad+\mathbb{P}_{\mathrm{GW}}\left(S(H)\ge d^{\frac{1}{5}}\right),\end{aligned}$$

and the r.h.s. is bounded by $4t^{-\sqrt{d}}(\log t)^{-2}$. Hence, we finish the proof for Step 3-1.

3.5.7 Tail estimates on unicyclic graphs for large t: Step 3-2

Let $\dot{\mathcal{H}}\sim\mathsf{GWC}(\xi,m+1)^l$ and let $\mathcal{T}_2,\dots,\mathcal{T}_{m+1}$ as above. Consider the graphs H_2 and H_{m+1} defined as follows.

- H_2 consists of a cycle $v_2\dots v_{m+1}$ and the trees $\mathcal{T}_3,\dots,\mathcal{T}_{m+1}$ rooted at $v_3,\dots,v_{m+1}$. The vertex $\rho=v_2$ is designated as the root.
- H_{m+1} consists of a cycle $v_2\dots v_{m+1}$ and the trees $\mathcal{T}_2,\dots,\mathcal{T}_m$ rooted at $v_2,\dots,v_m$. The vertex $\rho=v_{m+1}$ is designated as the root.

Note that $H_2\sim\mathsf{GWC}(\xi,m)^l$, and we let $S_3(H_2),S_{m+1}(H_2)$ as Definition 3.4.5 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 3.5.6 shows that

$$\begin{aligned} S_2(\dot{\mathcal{H}}) &\le (1+2\lambda S(H_2)) \prod_{i=1}^{D_2} (1+\lambda S(T_i^2)); \\ S_{m+1}(\dot{\mathcal{H}}) &\le (1+2\lambda S(H_{m+1})) \prod_{i=1}^{D_{m+1}} (1+\lambda S(T_i^{m+1})). \end{aligned} \tag{3.5.19}$$

(Indeed, this was also proven in Lemma 2.3.11 and its verification was based on the same idea as Proposition 3.2.1.) Note that we are assuming the tail probabilities for $S(H_2)$ and $S(H_{m+1})$ satisfy (3.4.3), while those for $S(T_i)$ are given by Theorem 3.2.4. Based on this observation, we can follow the same analysis done in Lemma 3.2.8, Corollary 3.2.9 and Sections 3.2.4, 3.2.5 to see that for $t \ge d^{\frac{1}{10}}$,

$$\begin{aligned} \mathbb{P}_{\mathrm{GW}}\left(S(\dot{\mathcal{H}}) \ge t\right) &\le \mathbb{P}_{\mathrm{GW}}\left((1+2\lambda S(H_2)) \prod_{i=1}^{D_2} (1+\lambda S(T_i^2)) \ge t\right) \\ &\quad + \mathbb{P}_{\mathrm{GW}}\left((1+2\lambda S(H_{m+1})) \prod_{i=1}^{D_{m+1}} (1+\lambda S(T_i^{m+1})) \ge t\right) \\ &\le 2t^{-\sqrt{d}}(\log t)^{-2} + 2t^{-\sqrt{d}}(\log t)^{-2} = 4t^{-\sqrt{d}}(\log t)^{-2}. \end{aligned}$$

The logic to derive (3.4.3) for $M^l(\dot{\mathcal{H}})$ is similar. We first derive the recursive inequalities, based on (3.5.19) and the ideas from Proposition 3.3.2, We have

$$M_a^l(\dot{\mathcal{H}}) \le 2\lambda M^l(H_a) \prod_{i=1}^{D_a} (1+\lambda S(T_i^a)) + \lambda(1+2\lambda S(H_a)) \sum_{i=1}^{D_a} M^{l-1}(T_i^a) \prod_{\substack{j=1 \\ j\neq i}}^{D_a} (1+\lambda S(T_j^a)),$$

for $a \in \{2, m+1\}$. As done in Section 3.3.5, 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 $\widetilde{M}_a = (1-\frac{\varepsilon}{10})^{-l} M_a(\dot{\mathcal{H}})$,

$$\Big(1-\frac{\varepsilon}{10}\Big)\,\widetilde{M}_a \le \prod_{i=0}^{D_a}(1+4\lambda W_i^a), \quad a \in \{2,\ m+1\}.$$

Then, the rest of the proof goes the same as Section 3.5.5, yielding that

$$\mathbb{P}_{\mathrm{GW}}\left(M^l(\dot{\mathcal{H}}) \ge \Big(1-\frac{\varepsilon}{10}\Big)^l t\right) \le \sum_{a\in\{2,m+1\}} \mathbb{P}_{\mathrm{GW}}\left(2\prod_{i=0}^{D_a}(1+4\lambda W_i^a) \ge t\right) \le 4t^{-\sqrt{d}}(\log t)^{-2},$$

for all $t \ge d^{\frac{1}{10}}$, hence establishing Step 3-2. Combining the argument in Sections 3.5.4—3.5.7, we conclude the proof of Proposition 3.4.4 and Corollary 3.4.6. □

3.6 Supercritical phase for general distribution: Proof of Theorems 10 and 11

In this section, we establish Theorems 10 and 11. We first observe that the phase transitions for $\mathsf{GW}(\widetilde{\mu})$ and $\mathsf{GW}(\mu,\widetilde{\mu})$ occur at the same intensity (Lemma 3.6.1) and then show the second inequality in (1.5.6) (Lemma 3.6.2). After verifying the two lemmas, we prove the first inequality in (1.5.6) and Theorem 11, which is the main goals of this section. Finally, in Section 3.6.6, we combine the main results obtained throughout the chapter and conclude the proof of Theorems 3 and 7.

Lemma 3.6.1. *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. □

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 (1.5.6).

Lemma 3.6.2. *Let μ be a degree distribution satisfying the giant component condition (1.5.4), 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. □

The rest of this section is devoted to the proof of the first inequality in (1.5.6) and the proof of Theorem 11. 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 3.6.7) 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 an *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.

3.6.1 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 3.4.7 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

$$\mu(j_0, \infty) = \mathbb{P}_{D\sim\mu}(D > j_0) \geq 10\mathbb{E}_{D\sim\mu} D\mathbf{1}_{\{D>j_1\}} =: 10\eta_0. \tag{3.6.1}$$

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

$$\mu(j_0) \geq 10\eta_0. \tag{3.6.2}$$

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,

$$|\nu[0, l] - \mu[0, l]| \leq \eta_0, \quad \forall l \leq j_1 \quad \text{by Chernoff inequality} \tag{3.6.3}$$

and

$$\sum_{i=1}^{n} D_i \mathbf{1}_{\{D_i > j_1\}} \leq 2\eta_0 n \quad \text{by Chebyshev's inequality.} \tag{3.6.4}$$

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 (3.6.4), 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 (3.6.3) gives

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

Let μ_{η_0,j_0} be the probability measure given by

$$\mu_{\eta_0,j_0}(A) := \begin{cases} (1-4\eta_0)\frac{\mu(A\cap[0,j_0])}{\mu[0,j_0]} + 4\eta_0 \mathbf{1}_{\{0\in A\}} & \text{if } \mu \text{ has an infinite support;} \\ \mu(A\cap[0,j_0)) + (\mu(j_0) - 4\eta_0)1_{j_0\in A} + 4\eta_0 \mathbf{1}_{\{0\in A\}} & \text{otherwise.} \end{cases} \tag{3.6.5}$$

By (3.6.1) and (3.6.2), 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

$$\mathsf{whp}, \quad \nu' \geq_{st} \mu_{\eta_0,j_0}. \tag{3.6.6}$$

Similarly, we have

Lemma 3.6.3. *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 $G_n' \backslash \Delta$ 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 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 $\widetilde{\mu}_{\eta_0,j_0}$.

3.6.2 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 3.6.4 (Good trees)**.** A tree T with root ρ is said to be $(s_0, L_0, k, \theta, \varepsilon)$-good if

$$\mathbb{P}_{\mathrm{CP}}\left(|X_{ks_0}^{kL_0} \cap T_{kL_0}| \geq \theta^k\right) > \varepsilon \tag{3.6.7}$$

where $\left(X_t^{kL_0}\right) \sim \mathsf{CP}^{\lambda}(T_{\leq kL_0}; \mathbf{1}_\rho)$ 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 3.6.5. *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*

$$\mathbb{P}_{\mathrm{GW}}(\mathcal{T} \text{ is } (s_0, L_0, k, \theta, \varepsilon)\text{-good}) > \varepsilon \tag{3.6.8}$$

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

For the proof of Lemma 3.6.5, we first show that

Lemma 3.6.6. *Let $\mathcal{T} \sim \mathsf{GW}(\widetilde{\mu})$. There exist constants $s_0, L_0, \varepsilon > 0$ and $\theta > 1$ such that*

$$\mathbb{E}(|X_{s_0}^{L_0} \cap \mathcal{T}_{L_0}|) \geq \theta \tag{3.6.9}$$

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

Assuming Lemma 3.6.6, we prove Lemma 3.6.5.

Proof of Lemma 3.6.5. As $\widetilde{\mu}_{\eta_0,j_0} \to \widetilde{\mu}$, it suffices to prove (3.6.8) for $\mathcal{T} \sim \mathbf{GW}(\widetilde{\mu})$.

Let $W_k = |X_{ks}^{kL_0} \cap \mathcal{T}_{kL_0}|$. 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,

$$\mathbb{P}\left(W_k > \varepsilon\theta^k\right) > \varepsilon. \tag{3.6.10}$$

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 (3.6.8) 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. □

Proof of Lemma 3.6.6. 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,

$$\mathbb{E}_{\mathcal{T}}(|Y_{\leq s_0}^{L_0} \cap \mathcal{T}_L|) \geq \theta \tag{3.6.11}$$

where $Y^{L_0}_{\le s_0} := \bigcup_{t\le s_0} Y^{L_0}_t$. Indeed, assuming (3.6.11), we have

$$\mathbb{E}(|X^{L_0}_{\le s_0} \cap \mathcal{T}_L|) \ge \theta.$$

Let $W_k = |X^{kL_0}_{\le ks} \cap \mathcal{T}_{kL_0}|$ then by the same argument that derives (3.6.10), we have

$$\mathbb{P}\left(W_k > \varepsilon\theta^k\right) > \varepsilon \tag{3.6.12}$$

for all sufficiently large k. As a consequence,

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

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

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

Hence, by choosing k large, we have

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

which proves (3.6.9) 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 $(Y'_t)_{t\le s_0}$ from the root to a leaf of the L_0-generation that lies entirely in $\mathcal{T}_{\le L_0}$. We have

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

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

$$\mathbb{P}_{\mathcal{T},Y'}\left(\mathcal{A}_{s_0,L_0'} \neq \emptyset\right) > \varepsilon, \quad \forall L_0' \leq L_0. \tag{3.6.13}$$

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

$$C\varepsilon^2\varepsilon_1\frac{\lambda}{1+\lambda} = 10 \tag{3.6.14}$$

where we recall that $\varepsilon_1 = 1 - \frac{\lambda'}{\lambda}$. The reason for this choice of C will be clear in (3.6.16).

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,

$$\lim_{L_0\to\infty} \mathbb{P}\left(\mathcal{B}_{L_0}\right) = 0. \tag{3.6.15}$$

Assuming (3.6.15), 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 (3.6.13), 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 (3.6.14),

$$\begin{aligned}\mathbb{E}(|Y^{L_0}_{\leq 3s_0} \cap \mathcal{T}_{L_0}|) &\geq \mathbb{P}\left(\mathcal{A}_{s_0,L_0} \setminus \mathcal{B}_{L_0}\right) \mathbb{E}\left(|\partial W_{s_0,L_0}| \middle| \mathcal{A}_{s_0,L_0} \setminus \mathcal{B}_{L_0}\right) \varepsilon_1(1-e^{-s_0})\frac{\lambda}{1+\lambda}\varepsilon \\ &\geq \frac{\varepsilon}{2} C\varepsilon_1 \frac{1}{2}\frac{\lambda}{1+\lambda}\varepsilon > 1 \end{aligned} \tag{3.6.16}$$

proving (3.6.11) by setting $3s_0$ to be the new s_0.

It remains to prove (3.6.15). 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 contains 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

$$\mathbb{P}\left(\mathcal{B}_{L_0,p_1,\ldots,p_k}\right) \leq e^{-cL_0} \tag{3.6.17}$$

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 (3.6.15).

It is left to prove (3.6.17). 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 $\mathbf{GW}(\widetilde{\mu})$ 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

$$\mathbb{P}\left(\mathcal{B}_{L_0,p_1,\ldots,p_k}\right) \leq (1-\varepsilon_2)^{L_0-k} d^k \varepsilon_2^{-k} \leq (1-\varepsilon_2)^{L_0} \left((1-\varepsilon_2)^{-1} d \varepsilon_2^{-1}\right)^C \tag{3.6.18}$$

which gives (3.6.17) for sufficiently large L_0.

To see (3.6.18), 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

$$\mathbb{E}_{\mathcal{T}\sim\mathbf{GW}(\widetilde{\mu})}\left(|\mathcal{T}_2| \middle| \mathcal{T}_{L_0-2} \neq \emptyset\right) \leq \frac{d^2}{\mathbb{P}_{\mathcal{T}\sim\mathbf{GW}(\widetilde{\mu})}\left(\mathcal{T}_{L_0-2} \neq \emptyset\right)} \leq d^2 \varepsilon_2^{-1}. \tag{3.6.19}$$

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 (3.6.18) for a general sequence $p_1, \ldots, p_k$ follows by the same reasoning. We leave the details to the interested reader. □

3.6.3 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 3.6.5.

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

$$\mathbb{P}_{\mathrm{GW}}\left(|\mathcal{T}_h| \geq c_0 \tilde{d}^h\right) \geq 2c_0. \tag{3.6.20}$$

The existence of c_0 follows from the fact that $\frac{|\mathcal{T}_h|}{\tilde{d}^h}$ converges to a nonzero random variable almost surely and that $\widetilde{\mu}_{\eta_0,j_0}$ converges $\widetilde{\mu}$ and $\tilde{d}$ 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 3.6.10). 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 (3.6.20)).

For the proof of Theorem 10, we only need to take $\kappa = \delta n$ while for Theorem 11, 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 10, 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

$$l_1(\kappa) = 2\log_{\tilde{d}} \log \frac{n}{\kappa} \tag{3.6.21}$$

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

$$\tilde{d}^{l_1(\kappa)} \gg \log \frac{n}{\kappa} \quad \text{and} \quad j_1^{Cl_1(\kappa)}\kappa \leq \eta_0 n \tag{3.6.22}$$

where the second inequality ensures the hypothesis of Lemma 3.6.3 and the first inequality, whose purpose will be clear later in Lemma 3.6.8, is to facilitate the union bound over choices of κ vertices.

Definition 3.6.7 (*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 (3.6.20), we have for a uniformly chosen vertex v,

$$\mathbb{P}\left(v \text{ is good}\right) \geq (\mu_{\eta_0,j_0}(j_0))^{j_0^4} c_0^{j_0(j_0-1)^3} =: p_0 \tag{3.6.23}$$

where the term $(\mu_{\eta_0,j_0}(j_0))^{j_0^4}$ bounds the probability that the first condition in Definition 3.6.7 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 3.6.3 and (3.6.20).

3.6.4 A graph property of neighborhood expansion

Let

$$l_2(\kappa) = Al_1(\kappa) \tag{3.6.24}$$

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 3.6.7. 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 3.6.5 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

$$l(\kappa) = 4 + l_1(\kappa) + l_2(\kappa)L_0. \tag{3.6.25}$$

From 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

$$\mathbb{P}_{\mathrm{CP}} \left(|D_s \cap \mathcal{F}_\kappa| \leq a p_0 / 2 \middle| |D_s| \geq a \right) \leq \exp\left(-3\kappa \log \frac{n}{\kappa} \right). \tag{3.6.26}$$

Under the randomness of the perfect matching of half-edges in G'_n, we show in Lemmas 3.6.8 and 3.6.9 that $\mathcal{B}$ and $\mathcal{C}$ happen whp, respectively.

Lemma 3.6.8. *We have*

$$\mathbb{P}(G'_n \in \mathcal{B}) = 1 - o(1). \tag{3.6.27}$$

For the proof of this lemma, we shall use the cut-off line algorithm introduced in Definition 2.6.3 to find a uniform perfect matching of the half-edges in G'_n. See also Figure 2.3 for reference.

Proof of Lemma 3.6.8. 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 $2j_1^{4+l_1(\kappa)}\kappa$ 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,

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

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

$$\mathbb{P}\left(\mathbf{Bin}(N,p) \geq xNp\right) \leq \exp\left(-xNp\log x + xNp\right), \quad \forall N, p > 0, x > 1. \tag{3.6.28}$$

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

Taking union bound over i and D and noting that $\binom{n}{\kappa} = o\left(\exp\left(1.01\kappa\log\gamma\right)\right)$, 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

$$\sum_{i=2}^{j_1} in_i \leq 1.09\kappa\left(\frac{j_1}{j_1-1} + \sum_{i=2}^{j_1-1}\left(\frac{1}{i-1}-\frac{1}{i}\right)i\right) \leq 1.1\kappa(2+\log j_1). \tag{3.6.29}$$

Now, assume that D consists of κ-good vertices with disjoint 3-neighborhoods. By definition, the number of out-doing half-edges of $N(D,3)$ is $j_0(j_0-1)^3\kappa$. By (3.6.29), 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

$$\left(j_0(j_0-1)^3\kappa - 1.1\kappa(2+\log j_1)\right) c_0\tilde{d}^{l_1(\kappa)} \geq c_0\tilde{d}^{l_1(\kappa)}\kappa. \tag{3.6.30}$$

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 3.6.3 and Lemma 3.6.5, 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,

$$\mathbb{P}\left(\mathbf{Bin}(c_0\tilde{d}^{l_1(\kappa)}\kappa, \varepsilon/2) \leq c_0\tilde{d}^{l_1(\kappa)}\kappa\varepsilon/4\right) \leq \exp\left(-\Omega\left(\tilde{d}^{l_1(\kappa)}\kappa\right)\right) \leq \exp\left(-1.1\kappa\log\frac{n}{\kappa}\right) \tag{3.6.31}$$

by (3.6.22). By taking the union bound over the choices of D, this completes the proof. □

Lemma 3.6.9. *We have*

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

Proof of Lemma 3.6.9. 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,

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

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 at most $j_1^{5+l_1(\kappa)}$,

$$\mathbb{P}_{\mathrm{RG,CP}}\left(|D_s \cap \mathcal{F}_\kappa| \le ap_0/2 \middle| |D_s| \ge a\right) \le \exp\left(-\frac{(ap_0/2)^2}{2aj_1^{10+2l_1(\kappa)}}\right) \le \exp\left(-5\kappa \log \frac{n}{\kappa}\right). \tag{3.6.32}$$

where we chose C_1 much larger than A in (3.6.24).

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

$$\begin{aligned}\mathbb{P}_{\mathrm{RG}}&\left(\mathbb{P}_{(X_t)}\left(|D_s \cap \mathcal{F}_\kappa| \le ap_0/2\right) \ge \exp\left(-3\kappa \log \frac{n}{\kappa}\right) \middle| |D_s| \ge a\right) \\ &\le \exp\left(-2\kappa \log \frac{n}{\kappa}\right) = o\left(\binom{n}{\kappa}^{-1}\right).\end{aligned}$$

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

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

□

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 $s \ge \varepsilon$ such that

$$\mathbb{P}_{\mathrm{CP}}\left(|D_s \cap \mathcal{F}_{2\kappa}| \le \kappa \log \frac{n}{\kappa}\right) \le e^{-\kappa} \tag{3.6.33}$$

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 3.6.10. *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 3.6.8, there exists a constant $t > 0$ such that

$$\mathbb{E}_{\mathrm{CP}}\left(|X_{\leq tl_2(\kappa)} \cap \partial N(D, l(\kappa))|\right) \geq \frac{\varepsilon}{4}\kappa c_0 \tilde{d}^{l_1(\kappa)} \left(\frac{(1-e^{-1})\lambda}{1+\lambda}\right)^{4+l_1(\kappa)} \theta^{l_2(\kappa)} \geq \theta^{l_2(\kappa)/2}\kappa \quad (3.6.34)$$

where the last inequality follows from (3.6.25). Dividing the time interval $[0, tl_2(\kappa)]$ into intervals of length 1, there exists an interval $[r, r+1]$ such that

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

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

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

By Azuma's inequality,

$$\mathbb{P}_{\mathrm{CP}}\left(|X_{r+1} \cap \partial N(D, l(\kappa))| \leq 2p_0^{-1}\kappa\log\frac{n}{\kappa}\right) \leq \exp\left(-\frac{\kappa^2 \log^2\frac{n}{\kappa}}{\kappa j_1^{3l(\kappa)}}\right) \leq \exp\left(-3\kappa\log\frac{n}{\kappa}\right). \quad (3.6.35)$$

In the notations of Lemma 3.6.9, $X_{r+1} \cap \partial N(D, l(\kappa))$ is D_r. Thus, (3.6.35) together with (3.6.26) give

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

completing the proof. □

3.6.5 Finishing the proof of Theorems 10 and 11

We are now ready to finish the proof of the first inequality in (1.5.6) and Theorem 11. Let G'_n be a graph belonging to both events $\mathcal{B}$ and $\mathcal{C}$ in Lemmas 3.6.8 and 3.6.9.

Let v be a uniformly chosen random vertex in G'_n. By Lemma 3.6.3, whp over G'_n, for any constant C_2, $N(v, C_2 \log\log n) \geq_{st} \mathbf{GW}(\mu_{\eta_0,j_0}, \widetilde{\mu}_{\eta_0,j_0})$. And so, by Lemma 3.6.5, 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 the definition 3.6.4 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 3.6.9, 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 3.6.10, 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 3.6.10 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 11. If all vertices are initially infected, then the first inequality in (1.5.6) of Theorem 10 follows directed by applying Lemma 3.6.10 for $e^{\delta n/2}$ times with $\kappa = \delta n$ as above.

3.6.6 Finishing the proof of Theorems 3 and 7

Finally, we combine the results obtained in this article and establish Theorems 3 and 7. Let us begin with Theorem 3. Its lower bound

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

was proven at the end of Section 3.2.2. Also, the other inequality was obtained by Theorem 10. For Theorem 7, Theorem 3.4.1 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 10. Thus, we conclude the proofs of Theorems 3 and 7. □

3.7 Proof of technical lemmas

To conclude Chapter 3, we present the remaining proofs of some technical lemmas develop through this chapter.

3.7.1 Remaining proof of Lemma 3.5.2

We finish the proof of Lemma 3.5.2 by establishing (3.5.2) and (3.5.3). The ideas are very similar as in the proofs of Theorems 3.2.4 and 3.3.1.

Remaining proof of Lemma 3.5.2. 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 3.5.2.

For any fixed l, assume that (3.5.2) and (3.5.3) 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 (3.2.5) and (3.3.10), respectively. Based on these inequalities, we can repeat the argument from Sections 3.2.3 and 3.3.4 to deduce (3.5.2) and (3.5.3) 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 (3.2.1) and (3.3.3). Note that the inequality for $M(\mathcal{F}^{m,l})$ is also reminiscent of (3.5.17). Then, we follow the logic from Sections 3.2.4, 3.2.5, 3.3.5 and 3.3.6 to obtain (3.5.2) and (3.5.3) for $t \geq d^{\frac{1}{10}}$ and $m+1$. □

3.7.2 Proof of Claim 3.5.3

In this subsection, we finish the remaining proof of Claim 3.5.3

Remaining proof of Claim 3.5.3. Recall the definitions of $\widetilde{B}$ and $\widehat{S}$ in (3.5.8). In the setting of Claim 3.5.3, we are assuming (3.5.4) for m, which is

$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq 2. \tag{3.7.1}$$

The bound on $\widehat{S}$ was derived in Sections 3.2.4 and 3.2.5 which yielded that

$$\mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\right) \leq t^{-\sqrt{d}}(\log t)^{-2}, \quad \text{for all } t \geq d^{\frac{1}{10}}. \tag{3.7.2}$$

Set $\alpha = d^{\frac{1}{10}}$ and $k_0 = \lfloor \frac{\log t}{\log \alpha} \rfloor - 1$. The rest of the proof is analyzing (3.5.9). The assumptions (3.7.1) and (3.7.2) 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 $t \geq d^{\frac{1}{10}}$,

$$\begin{aligned}
&\mathbb{P}_{\mathrm{GW}}\left(\widehat{S} \geq t\alpha^{-k}\right) \cdot \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B} \geq \frac{\alpha^{k-1}}{2} d^{\frac{1}{4}}\right) \\
&\leq \left(\frac{\alpha^k}{t} \cdot \frac{2}{d^{\frac{1}{4}}\alpha^{k-1}}\right)^{\sqrt{d}} \left(\frac{1}{\log\left(t\alpha^{-k}\right)}\right)^2 \left(\frac{1}{\log\left(\frac{1}{2}\alpha^{k-1}d^{\frac{1}{4}}\right)}\right)^2 \\
&\leq \left(d^{\frac{1}{8}}t\right)^{-\sqrt{d}} \frac{1}{(\log t - k\log\alpha)^2} \frac{1}{\left(\frac{1}{8}\log d + k\log\alpha\right)^2}.
\end{aligned} \tag{3.7.3}$$

Observing that $\log t - k_0 \log \alpha \geq \log \alpha = \frac{1}{10}\log d$, we can apply Lemma 3.2.11 in order to control the sum of (3.7.3) over k. All in all, we deduce that (3.5.8) holds true for $t \geq d^{\frac{1}{10}}$. $\square$

3.7.3 Proof of Lemma 3.5.4

The goal of this subsection is to present the proof of Lemma 3.5.4.

Proof of Lemma 3.5.4. We begin with establishing the recursive inequalities for $S(\mathcal{A}^{\underline{\mathrm{m}},l})$, $M(\mathcal{A}^{\underline{\mathrm{m}},l})$ and $B(\mathcal{A}^{\underline{\mathrm{m}},l})$. To this end, we first define some subgraphs of $\mathcal{A}^{\underline{\mathrm{m}},l}$ 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 3.5.2.

- Let D be the degree of v_0 inside $\mathcal{T}_0$, and let $T_1, \ldots, T_D$ be the subtrees of $\mathcal{T}_0$ rooted at each child of v_0 in $\mathcal{T}_0$.
- Let F_1 be the graph that consists of the line $v_{-1} \ldots v_{-m_1}$ of length $(m_1 - 1)$ and the trees $\mathcal{T}_{-1}, \ldots, \mathcal{T}_{-m_1+1}$ rooted at $v_{-1}, \ldots, v_{-m_1+1}$, respectively.
- Let F_2 be defined similarly in the other branch of $\mathcal{A}^{\underline{\mathrm{m}},l}$. 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 $T_1, \dots, T_{D+2}$ are all trees.

Based on the same ideas as the derivation of (3.2.5), (3.3.10) and (3.5.5), we have the first recursive inequalities as follows.

$$\begin{aligned} &S(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \frac{1+\sum_{i=1}^{D+2} S(T_i)}{1-\lambda-2\lambda^2\sum_{i=1}^{D+2} S(T_i)}; \quad M(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \frac{\lambda\sum_{i=1}^{D+2} M(T_i)}{1-\lambda-2\lambda^2\sum_{i=1}^{D+2} S(T_i)}; \\ &B(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \frac{\lambda(B(T_{D+1})+B(T_{D+2}))}{1-\lambda-2\lambda^2\sum_{i=1}^{D+2} S(T_i)}. \end{aligned} \tag{3.7.4}$$

Moreover, the second recursive inequalities are obtained from Propositions 3.2.1, 3.3.2 and from (3.5.6), which read

$$\begin{aligned} &S(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \prod_{i=1}^{D+2}(1+\lambda S(T_i)); \quad M(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \lambda\sum_{i=1}^{D+2} M(T_i)\prod_{\substack{j=1\\ j\neq i}}^{D+2}(1+\lambda S(T_i)); \\ &B(\mathcal{A}^{\underline{\mathrm{m}},l}) \le \lambda\sum_{i=D+1}^{D+2} B(T_i)\prod_{\substack{j=1\\ j\neq i}}^{D+2}(1+\lambda S(T_i)). \end{aligned} \tag{3.7.5}$$

Note that we know the tail probabilities for $S(T_i),\ i \in [D+2]$ by Theorems 3.2.4, 3.3.1 and Lemma 3.5.2. Therefore, we can deduce (3.5.10) and (3.5.11) by following Sections 3.2.3—3.2.5 and 3.3.4—3.3.6.

To see (3.5.12), we rely on Lemma 3.5.2. For $t \le d^{\frac{1}{10}}$, we have the following inequality reminiscent of (3.5.7).

$$\begin{aligned} \mathbb{P}_{\mathrm{GW}}\left(B(\mathcal{A}^{\underline{\mathrm{m}},l}) \ge d^{\frac{3m_1\wedge m_2}{4}}t\right) \le \mathbb{P}\left(D \ge \left(1+\frac{\varepsilon}{6}\right)d\right) + \mathbb{P}_{\mathrm{GW}}\left(\sum_{i=1}^{D+2} S(T_i) \ge \frac{2d}{\varepsilon}\left(1+\frac{\varepsilon}{3}\right)\right) \\ + P\left(B(F_1) \ge d^{-\frac{3}{4}(m_1-1)+\frac{1}{2}}\right) + P\left(B(F_2) \ge d^{-\frac{3}{4}(m_2-1)+\frac{1}{2}}.\right) \end{aligned}$$

By estimating the terms in the r.h.s. as in Case 1 of the proof of Lemma 3.5.2 and Section 3.2.3, we see that the above is smaller than $t^{-\sqrt{d}}(\log t)^{-2}$ for $t \le d^{\frac{1}{10}}$.

For $t \geq d^{\frac{1}{10}}$, we set

$$\widetilde{B}(\mathcal{A}^{\underline{\mathrm{m}},l}) := d^{\frac{3}{4}(m_1 \wedge m_2)} B(\mathcal{A}^{\underline{\mathrm{m}},l}), \quad \text{and} \quad \widetilde{B}_a := d^{\frac{3}{4}(m_a - 1)} B(F_a), \;\; a \in \{1,2\},$$

and observe that by (3.7.5),

$$\mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}(\mathcal{A}^{\underline{\mathrm{m}},l}) \geq t\right) \leq \sum_{a \in \{1,2\}} \mathbb{P}_{\mathrm{GW}}\left(\widetilde{B}_a \prod_{\substack{j=1 \\ j \neq D+a}}^{D+2} (1 + \lambda S(T_i)) \geq \frac{d^{\frac{1}{4}} t}{3}\right).$$

Then, the summand in the r.h.s. can be bounded by applying Claim 3.5.3 and implies (3.5.12) for $t \geq d^{\frac{1}{10}}$. □

3.7.4 Proof of Lemma 3.4.8

Here, we establish Lemma 3.4.8. Indeed, the third statement in the lemma was proven in Lemma 2.3.3. Therefore, we focus on the first two statements of the lemma.

Proof of Lemma 3.4.8. 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 3.4.7. 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

$$\sum_{k<k_0} kp_k + k_0\sqrt{p_{k_0}} \le \left(1 - \frac{\varepsilon}{10}\right) d + \varepsilon,$$

which again implies $d^\sharp \le \left(1 + \frac{\varepsilon}{9}\right) d$.

For the second, we only establish the first of (3.4.4), since the second line follows analogously. Let $\varepsilon' \in [\frac{\varepsilon}{10}, 1]$. For any integer $k \in [(1+\varepsilon')d, 2d]$, we have

$$\mu^\sharp(k) \le \frac{\sqrt{p_k}}{Z} \le 2\exp\left(-\frac{c_{\varepsilon'}}{2}d\right).$$

Therefore, we obtain that

$$\begin{aligned}\mathbb{P}_{D^\sharp\sim\mu^\sharp}\left(D^\sharp \ge (1+\varepsilon')d\right) &\le \sum_{k=\lceil(1+\varepsilon')d\rceil}^{\lfloor 2d\rfloor} 2\exp\left(-\frac{c_{\varepsilon'}}{2}d\right) + \sum_{k\ge\lceil 2d\rceil} 2\exp\left(-\frac{c_1 k}{2}\right) \\ &\le 2d\exp\left(-\frac{c_{\varepsilon'}}{2}d\right) + \frac{2e^{-c_1 d}}{1-e^{-\frac{c_1}{2}}} \le \exp\left(-\frac{c_{\varepsilon'}\wedge c_1}{3}d\right),\end{aligned}$$

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 (3.4.4). □

3.7.5 Refined analysis of cycle existence in local neighborhoods

Lemma 2.3.5 had a very similar statement as Lemma 3.4.9, but it was weaker in terms of determining the scope of γ. It turns out that in order to prove the stronger version, Lemma 3.4.9, the following bound on the size of $N(v, l)$ is needed.

Lemma 3.7.1. *Under the setting of Lemma 3.4.9, 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 $d \ge d_0$, then we have*

$$\mathbb{P}\left(|\mathcal{T}| \ge n^{\frac{1}{5}}\right) \le \exp\left(-n^{\frac{1}{20}}\right)$$

Proof. Let $c'_1 \in \mathfrak{c}'$ be the constant in Lemma 3.4.8. Then, we first note that the second

statement of Lemma 3.4.8 tells us the following:

$$\mathbb{E}_{D\sim\widetilde{\mu}^{\sharp}}\left[\exp\left(\frac{c_1'}{2}D\right)\right] \leq \exp(2d) + \sum_{k\geq 2d} e^{\frac{c_1' k}{2}} e^{-c_1' k} \leq e^{3d}, \tag{3.7.6}$$

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

$$\left(\frac{6d}{c_0'}\right)^{L_n} \leq n^{\frac{1}{9}}, \tag{3.7.7}$$

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

$$\begin{aligned}\mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{h}Y_h\right)\right] &= \mathbb{E}\left[\mathbb{E}_D\left[\exp\left(\left(\frac{c''}{3d}\right)^{h}D\right)\right]^{Y_{h-1}}\right] \\ &\leq \mathbb{E}\left[\exp\left\{3d\cdot\frac{(c'')^{h-1}}{(3d)^{h}}Y_{h-1}\right\}\right] = \mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{h-1}Y_h\right)\right].\end{aligned}$$

Hence, we obtain that

$$\mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{L_n}\sum_{h=0}^{L_n}Y_h\right)\right] \leq \sum_{h=0}^{L_n}\left(\frac{c''}{3d}\right)^{h} \leq 2, \tag{3.7.8}$$

if d is larger than some constant depending on $\mathfrak{c}$. (Note that (3.7.6) holds for $\mu^{\sharp}$ as well.) Therefore, combining (3.7.8) with (3.7.7), we deduce that

$$\begin{aligned}\mathbb{P}\left(\sum_{h=0}^{L_n}Y_h\geq n^{\frac{1}{5}}\right) &\leq \exp\left(-\left(\frac{c''}{3d}\right)^{L_n}n^{\frac{1}{5}}\right)\times\mathbb{E}\left[\exp\left(\left(\frac{c''}{3d}\right)^{L_n}\sum_{h=0}^{L_n}Y_h\right)\right] \\ &\leq \exp\left(-n^{\frac{1}{20}}\right).\end{aligned}$$

□

Proof of Lemma 3.4.11. In Lemma 2.3.5, its proof reduces to obtaining tail probability bounds of $\sum_{h=0}^{L_n}Y_h$ for Y_h as above. Having Lemma 3.7.1 in hand, following the proof of

Lemma 2.3.5 deduces Lemma 3.4.11. □

Chapter 4

The threshold-θ contact process on random graphs

In this final chapter, our interest moves on to the study of the *discrete-time threshold-θ contact process* (Section 1.2) on random graphs, with $\theta \geq 2$. This chapter is devoted to verifying the results on its discontinuous phase transition discussed in Section 1.6. The logic and technical details developed in this chapter are all based on the work [Nam19] by the author.

4.1 Overview of the chapter

We begin with explaining the main ideas in the proof of Theorem 12. Working with the generalized setting (Theorem 14) is based on the same strategy, but requires more technicalities.

Let W be the set of infected sites in $G_n \sim \mathcal{G}(n, \mu)$ at some time t. Then, to understand the infections at time $t+1$, the information on the following set is crucial:

$$W^{*\theta} := \big\{u \in V : u \text{ has at least } \theta \text{ neighbors in } W\big\}, \tag{4.1.1}$$

where V denotes the set of vertices of G_n. If we know the size of $W^{*\theta}$, then $|X_{t+1}| \approx p|W^{*\theta}|$.

In order to establish long survival results, we introduce the event $\mathcal{F}^{*\theta}(m_1, m_2)$ such that

$$\mathcal{F}^{*\theta}(m_1, m_2) := \big\{\text{for all } W \subset V \text{ of size } m_1,\ |W^{*\theta}| \geq m_2\big\}, \tag{4.1.2}$$

One can find a similar definition in [CD13]. Then the key property in our argument can be stated as follows.

Proposition 4.1.1 (Expansion of large subsets)**.** *Let μ be a distribution on $\mathbb{N}$ that satisfies $\mathbb{E}_{D\sim\mu}D^6 =: M_6 < \infty$ and is lower bounded by $\theta + 2$. Then, there exists a constant $\varepsilon_1 = \varepsilon_1(\theta, M_6) > 0$ such that for $G_n \sim \mathcal{G}(n, \mu)$,*

$$\mathbb{P}\left(G_n \in \mathcal{F}^{*\theta}\left(\lceil(1-\varepsilon_1)n\rceil, \lceil(1-0.9\varepsilon_1)n\rceil\right)\right) \geq 1 - \frac{2\varepsilon_1^{1/5}}{(\varepsilon_1 n)^5}, \tag{4.1.3}$$

for all large enough n.

For the purpose of proving short survival, it suffices to consider the case $\theta = 2$. Define $\mathcal{E}^{*2}(m_1, m_2)$ to be an event that

$$\mathcal{E}^{*2}(m_1, m_2) := \big\{\text{for all } W \subset V \text{ of size } m_1,\ |W^{*2}| \leq m_2\big\}. \tag{4.1.4}$$

Proposition 4.1.2 (Expansion of small subsets)**.** *Let μ be a distribution on $\mathbb{N}$ that has finite k-th moments for all $k > 0$. Then, for all $\delta > 0$, there exists $\varepsilon_2(\delta, \mu) > 0$ such that for $G_n \sim \mathcal{G}(n, \mu)$ and for any $\varepsilon = \varepsilon(n) \in (0, \varepsilon_2)$ such that εn is an integer and $\lim_{n\to\infty} \varepsilon n = \infty$,*

$$\mathbb{P}\left(G_n \in \mathcal{E}^{*2}(\varepsilon n, (1+\delta)\varepsilon n)\right) \geq 1 - \frac{2\varepsilon^{1/4}}{(\varepsilon n)^{20/\delta}} - e^{-c\delta\varepsilon n}, \tag{4.1.5}$$

for all large enough n, where $c > 0$ is an absolute constant.

Chatterjee and Durrett [CD13] also derived similar estimates for random regular graphs as the above propositions. Thanks to regularity, they had better error probabilies than above, which was roughly $\exp(-c\varepsilon n \log(1/\varepsilon))$. In particular, the errors were $o(1)$ even for $\varepsilon n = O(1)$. In our case, existence of high degree vertices prevents us from achieving the error as small as in [CD13], especially when ε is very small.

It turns out that Theorem 12-1 follows in a straight-forward way from Proposition 4.1.1, based on a similar argument as in Sections 2 of [CD13]. In showing Theorem 12-2, one needs to be more careful after obtaining a small size of infections, since the bound (4.1.5) works for ε such that $\lim_{n\to\infty} \varepsilon n = \infty$. To deal with such difficulty, we bound the changes of infection size by a certain biased random walk which is easier to deal with. For the proof of Theorem 14, we require a generalized version of Proposition 4.1.2 which we introduce in Section 4.2.

In [CD13], they had the conclusions (4.1.3), (4.1.5) as byproducts of an analysis on W^{*1}, but this approach worked only for random regular graphs. To overcome this difficulty, we rather study W^{*2} and $W^{*\theta}$ directly, via describing the random matching among half-edges by certain binomial type random variables. We carry out this step using the "cut-off line algorithm" (Definition 2.6.3, developed by [Kim06]), appropriately modified to fit with our setting. In particular, since high-degree vertices cause problems in applying the algorithm in a way we want, we introduce a method of proving the propositions after getting rid of those problematic vertices.

4.1.1 Organization

The rest of the chapter is organized as follows. In Section 4.2, we discuss the structural properties of $\mathcal{G}(n,\mu)$ and prove Propositions 4.1.1 and 4.1.2. Then, we settle Theorems 12, 14 and Corollary 15 in Section 4.3. Finally, the last section is devoted to the proof of Theorem 13.

4.2 Structural analysis of random graphs

The purpose of this section is to establish Propositions 4.1.1 and a generalized version of Proposition 4.1.2, which will be used in Section 4.3 to prove Theorems 12 and 14. Recalling the definitions of event $\mathcal{E}^{*2}(m_1,m_2)$ in (4.1.4), the generalization of Proposition 4.1.2 can be stated in the following way.

Proposition 4.2.1 (Expansion of small subsets: extended version)**.** *Let $\delta \in (0,1)$ be any given number. For any degree distribution μ such that $\mathbb{E}_{D\sim\mu} D^{k_0} < \infty$ with $k_0 = 20/\delta$, there*

exists a constant $\varepsilon_2(\delta,\mu)$ *such that for* $G_n \sim \mathcal{G}(n,\mu)$ *and for any* $\varepsilon(n) \in (0,\varepsilon_2)$ *such that* εn *is an integer and* $\lim_{n\to\infty} \varepsilon n = \infty$*,*

$$\mathbb{P}\left(G_n \in \mathcal{E}^{*2}(\varepsilon n, (1+\delta)\varepsilon n)\right) \geq 1 - \frac{C\varepsilon^{1/4}}{(\varepsilon n)^{k_0-1}} - e^{-c\delta\varepsilon n},$$

for all large enough n*, where* $c, C > 0$ *are some absolute constants.*

In the following subsection, we introduce the *cut-off line algorithm*. Although this was already used in previous chapters (see Definition 2.6.3), we are going to discuss the details of its formulation again, since we need more involved notations for further purpose. This plays a crucial role in the proof of the above propositions, which is done in Sections 4.2.2 and 4.2.3.

4.2.1 The cut-off line algorithm

In this subsection, we review the cut-off algorithm, set up formal notations and derive some properties that are used in Sections 4.2.2 and 4.2.3. Notations and details we use to describe the algorithm are going to be slightly different from Definition 2.6.3.

Suppose that the degree sequence $\{D_i\}_{i=1}^n$ of the graph $G_n \sim \mathcal{G}(n,\mu)$ is given by i.i.d. samples from μ. Let $V = [n] := \{1,\dots,n\}$ be the vertex set of G_n so that $\deg(i) = D_i$. Let $d = \frac{1}{n}\sum_{i=1}^n D_i$ and $d_0 = \mathbb{E}_{D\sim\mu} D$.

Define $\{a_j\}_{j=1}^{dn}$ be the collection of half-edges. For each half-edge a_j, let v_j be the vertex that a_j is attached to. In other words, the half-edges are associated with the vertices by a map $v : [dn] \to [n]$ satisfying $|\{j \in [dn] : v_j = i\}| = D_i$ for each $i \in [n]$. To generate the full graph G_n, we have to perform a uniform random perfect matching among all half-edges $\{a_j\}_{j=1}^{dn}$.

In the cut-off line algorithm, we consider the rectangle $R = [n] \times [0,1]$. To be specific, on the xy-plane, vertices are located at points $1,\dots,n$ on the x-axis, and each vertex is assigned with a vertical interval of length 1 from $y = 0$ to $y = 1$. Then, we assign $B_j \sim$ i.i.d. Unif$[0,1]$ to each half-edge a_j, and mark the points

$$(v_j, B_j), \quad j = 1,\dots,dn$$

on the rectangle R. We call each point (v_j, B_j) *a clone of* v_j. Note that since $\{B_j\}$ is generated after the map $v : [dn] \to [n]$ is chosen, $\{B_j\}$ is independent of v.

Suppose that we are attempting to match a_1, the first half-edge, to its counterpart chosen uniformly at random. One way to do this is to find l_1 such that

$$B_{l_1} = \max\{B_j : j \in [dn] \setminus \{1\}\},$$

and pair a_1 with a_{l_1}. Since we found the matching pair of a_1, we can both remove a_1 and a_{l_1} among the collection of half-edges, and repeat the process for a_2 (unless $l_1 = 2$) to find its pair. This procedure defines the *cut-off line algorithm*, which we rigorously state as follows.

Definition 4.2.2 (The cut-off line algorithm, revisited)**.** Let $R = [n] \times [0, 1]$, a_j, v_j, and B_j as above and set $H_0 = 1$. At the beginning, dn points (v_j, B_j), $j \in [dn]$ are present on R. At t-th step of the algorithm, we perform the following.

1. Let J_t be the collection of indices $j \in [dn]$ such that the point (v_j, B_j) is present in R.
2. Let j_t be the minimal element in J_t, and let $l_t = \arg\max\{B_j : j \in J_t \setminus \{j_t\}\}$.
3. Set $H_t = B_{l_t}$, remove the points (v_{j_t}, B_{j_t}), (v_{l_t}, B_{l_t}) from R, and place an edge between v_{j_t} and v_{l_t} by pairing a_{j_t} and a_{l_t}.

The horizontal line $y = H_t$ is called the *cut-off line* at step t.

Suppose that we want to match all the half-edges attached at $W \subset V$ of size m. To this end, we first choose the map v in such a way that $v_j \in W$ for all $j \leq \sum_{i \in W} D_i$, so that the first $\sum_{i \in W} D_i$ half-edges are all at W. Then, we perform the cut-off line algorithm until all the points (v_j, B_j), $j \leq \sum_{i \in W} D_i$ are removed from R. Suppose that after T-th step of the algorithm, all half-edges at W are removed for the first time.

Our first goal is to derive estimates on H_T. For our purpose, it is enough to have some cheap estimates as follows, nevertheless sharper ones can be found in [Kim06]. We first make the following simple observation.

Lemma 4.2.3. *Under the above setting, we have that*

$$|\{j \in [nd] : B_j \geq H_T\}| \leq 2 \sum_{i \in W} D_i. \tag{4.2.1}$$

Proof. At each step of the algorithm, two points are removed from R. Note that $T \leq \sum_{i \in W} D_i$ by the definition of T. Therefore, at most $2 \sum_{i \in W} D_i$ points are removed from R after T-th step, which upper bound the l.h.s. of (4.2.1). □

From now on, let $m = \varepsilon n$. Our focus is on deriving a lower bound on H_T, and then we will later see its connection with the size of W^{*l}. Before getting started, we remark two possible cases when H_T can be low.

1. Too many half-edges are attached to W, i.e., many vertices in W have very high degree. If so, we need a number of steps of the algorithm to match all of them, resulting in a low H_T.

2. The number of half-edges with large B_j is small.

We begin with controlling the first possibility, bounding $\sum_{i \in W} D_i$. For later purpose, we prove the following lemma that covers all choices of εn-sized subsets of $[n]$.

Lemma 4.2.4. *Let $\eta > 0$ be an arbitrary number. Suppose that $\mathbb{E}_{D \sim \mu} D^k = M < \infty$ for some $k > 1 + \eta$ and let $\varepsilon = \varepsilon(n) \in (0, 1)$ be such that $\varepsilon n > 0$ is an integer. Then, we have*

$$\mathbb{P}\left(\forall\, W \subset V \text{ with } |W| = \varepsilon n,\ \sum_{i \in W} D_i \leq 3 M^{\frac{1}{k}} \varepsilon^{1 - \frac{1+\eta}{k}} n \right) \leq \frac{2\varepsilon^{\eta}}{(\varepsilon n)^{k-1}}, \tag{4.2.2}$$

for all large enough n.

We give the proof here to explain the odd looking upper bound. The reader who is not curious about the details can skip the proof. The logic is fairly straight-forward based on Hölder's inequality and large deviation estimates, but it requires some care on technical details.

Proof of Lemma 4.2.4. First, note that it suffices to show the inequality for the largest εn values among $\{D_i\}_{i=1}^n$. Let Δ denote the εn-th largest value among $\{D_i\}_{i=1}^n$, and let s be

the maximal integer that satisfies

$$\mathbb{P}_{D\sim\mu}(D \geq s) \geq \varepsilon^{1-\frac{\eta}{k}}. \tag{4.2.3}$$

Then, we have that

$$\mathbb{P}(\Delta < s) \leq \mathbb{P}\left(\mathbf{Bin}(n, \varepsilon^{1-\frac{\eta}{k}}) \leq \varepsilon n\right) = \exp\left(-c_0 n \varepsilon^{1-\frac{\eta}{k}}\right),$$

for some absolute constant $c_0 > 0$, where we obtain the last equality from the fact $\varepsilon n \geq 1$. Let W be the set of εn indices of the εn lagest values among $\{D_i\}_{i=1}^n$. Then, we claim that

$$\mathbb{P}\left(\sum_{i\in W} D_i \leq 2M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}} n + s\varepsilon n\right) \leq \frac{2\varepsilon^\eta}{(\varepsilon n)^{k-1}}. \tag{4.2.4}$$

To establish the above inequality, note that given $\{\Delta \geq s\}$, we have

$$\sum_{i\in W} D_i \leq \sum_{i=1}^n D_i \mathbf{1}_{\{D_i \geq s+1\}} + s\varepsilon n. \tag{4.2.5}$$

Moreover, Hölder's inequality and (4.2.3) implies that

$$\mathbb{E}_{D\sim\mu} D\mathbf{1}_{\{D\geq s+1\}} \leq M^{\frac{1}{k}}\varepsilon^{(1-\frac{\eta}{k})(1-\frac{1}{k})} \leq M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}. \tag{4.2.6}$$

If $Z_i,\ i \in [n]$ are i.i.d. random variables whose c.d.f. satisfies $P(Z_i \geq t) \leq t^{-\alpha}$ for $\alpha > 2$, then the large deviation inequalities for heavy-tailed random variables (see e.g., [Nag79, Theorem 1.9]) tell us that

$$\mathbb{P}\left(\sum_{i=1}^n (Z_i - \mathbb{E}Z_i) \geq bn\right) \leq 2n\mathbb{P}(Z_1 \geq bn),$$

for any constant $b > 0$ and for all large enough n. Applying this to the random variables

$D_i \mathbf{1}_{\{D_i \geq s+1\}}$, we obtain by (4.2.6) that

$$\mathbb{P}\left(\sum_{i=1}^{n} D_i \mathbf{1}_{\{D_i \geq s+1\}} \geq 2M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}n\right) \leq 2n\mathbb{P}\left(D_1 \geq M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}n\right) \leq \frac{2\varepsilon^{\eta}}{(\varepsilon n)^{k-1}}.$$

Thus, we establish (4.2.4).

Moreover, by Markov's inequality,

$$\mathbb{P}_{D\sim\mu}(D \geq s) \leq \frac{M}{s^k}, \quad \text{hence} \quad s \leq M^{\frac{1}{k}}\varepsilon^{-\frac{1}{k}+\frac{\eta}{k^2}} \leq M^{\frac{1}{k}}\varepsilon^{-\frac{1}{k}}. \tag{4.2.7}$$

Combining (4.2.4), (4.2.6) and (4.2.7), we deduce the conclusion. □

Our next objective is to study $e(i, W)$, the number of edges between a vertex i and the set W. Suppose that $y = H_T$ is the cut-off line when all half-edges from W are matched for the first time. Then, observe that

$$e(i, W) = \text{the number of points removed from } \{i\} \times [H_T, 1] \text{ until } T\text{-th step.} \tag{4.2.8}$$

Let $h \in [0, 1]$ and consider $Y_i \sim \mathbf{Bin}(D_i, 1-h),\ i \in [n]$, where the Y_i's are independent. Since the height of the points in R are set to be i.i.d. Unif$[0, 1]$, we can see that given the event $H_T \geq h$, there is a coupling between $\{e(i, W)\}_{i=1}^n$ and $\{Y_i\}_{i=1}^n$ such that

$$e(i, W) \leq Y_i \text{ for all } i \in [n], \text{ conditioned on } \sum_{i=1}^{n} Y_i \geq 2\sum_{i' \in W} D_{i'}, \tag{4.2.9}$$

where the latter conditioning is needed to set the cut-off line above h, based on Lemma 4.2.3. Define the graph property $\mathcal{B}_n$ (i.e., subsets of graphs of n vertices) by

$$\mathcal{B}_n := \left\{\text{for all } W \subset V \text{ with } |W| = \varepsilon n,\ \sum_{i \in W} D_i \leq 3M^{\frac{\delta}{20}}\varepsilon^{1-\frac{\delta}{16}}n\right\}, \tag{4.2.10}$$

the event given in (4.2.2). Now we choose appropriate h, which can be done based on the following estimate on H_T.

Lemma 4.2.5. *Let $\eta > 0$ be arbitrary. Suppose that $\mathbb{E}_{D\sim\mu} D^k = M < \infty$ for some $k > 1+\eta$, $\mathbb{E}_{D\sim\mu} D = d_0$ and let $\varepsilon \in (0,1)$ be a number such that εn is an integer. Let $W \subset V$ be any subset satisfying $|W| = \varepsilon n$, and let H_T be the height of the cut-off line when all clones at W are removed from the rectangle R for the first time. Then, we have that for all large enough n,*

$$\mathbb{P}\left(1 - H_T \geq \frac{7M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}}{d_0} \;\middle|\; \mathcal{B}_n\right) \leq \exp\left\{-c\varepsilon^{1-\frac{1}{k}}n\right\}, \tag{4.2.11}$$

where $\mathcal{B}_n$ is given as (4.2.10) and $c > 0$ is an absolute constant.

Proof. Recall that $d = \frac{1}{n}\sum_{i=1}^{n} D_i$. We first note that if $H_T \leq h$, then

$$|\{j \in [dn] : B_j \geq h\}| \leq 2\sum_{i\in W} D_i,$$

by Lemma 4.2.3. Moreover, observe that given $\{D_i\}_{i=1}^n$,

$$|\{j \in [dn] : B_j \geq h\}| \stackrel{\mathrm{d}}{=} \mathbf{Bin}(dn, 1-h).$$

We also know that whp over the choice of $\{D_i\}_{i=1}^n$, $\sum_{i\in W} D_i \leq 3M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}n$ for all W. Further, we have $d \geq 0.9d_0$ whp. Therefore, we obtain that the l.h.s. of (4.2.11) is at most

$$\mathbb{P}\left(\mathbf{Bin}\left(0.9d_0 n, \frac{7M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}}{d_0}\right) \leq 6M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}n\right). \tag{4.2.12}$$

To control this term, we rely on the following large deviation estimate for binomials (see, e.g., [FK15, Chapter 21]): for all $\gamma > 0$, there exists $c_\gamma > 0$ such that

$$\mathbb{P}\left(\mathbf{Bin}(n,p) \leq (1-\gamma)np\right) \leq \exp(-c_\gamma np). \tag{4.2.13}$$

Applying this to (4.2.12), we see that it is bounded by $\exp\left\{-c\varepsilon^{1-\frac{1}{k}}n\right\}$, for some absolute constant $c > 0$, and hence we obtain the conclusion. □

Therefore, the lemma tells us that we can choose h to be

$$h = 1 - \frac{7M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}}{d_0},$$

so that conditioned on $\mathcal{B}_n$,

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i \geq 6M^{\frac{1}{k}}\varepsilon^{1-\frac{1+\eta}{k}}n \;\middle|\; \mathcal{B}_n\right) \geq 1-\exp\left\{-c\varepsilon^{1-\frac{1}{k}}n\right\}. \tag{4.2.14}$$

4.2.2 Expansion of small subsets

In this subsection, we consider the expansion property of small subsets and prove Proposition 4.2.1.

Let $\delta \in (0,1)$ be given and suppose that μ satisfies $\mathbb{E}_{D\sim\mu}D^{k_0} = M < \infty$ with $k_0 = \frac{20}{\delta}$, as in the statement of Proposition 4.2.1. Let the degree sequence $\{D_i\}_{i=1}^n$ be given by i.i.d. samples of μ and let $\varepsilon = \varepsilon(n) \in (0,1)$ be a small parameter which will be specified later.

Before considering the matching between half-edges, we first exclude vertices with too large D_i from our consideration by a simple concentration argument.

Lemma 4.2.6. *Suppose that μ satisfies $\mathbb{E}_{D\sim\mu}D^{k_0} = M < \infty$ with $k_0 = \frac{20}{\delta}$ and $\varepsilon = \varepsilon(n) \in (0,1)$ obeys $\lim_{n\to\infty}\varepsilon n = \infty$. Then, there exists an absolute constant $c > 0$ such that for all large enough n and $D_i \sim$ i.i.d. μ,*

$$\mathbb{P}_\mu\left(\left|\left\{i\in[n] : D_i \geq \left(\frac{3M}{\delta\varepsilon}\right)^{\delta/20}\right\}\right| \geq \frac{1}{2}\delta\varepsilon n\right) \leq \exp\{-c\delta\varepsilon n\}. \tag{4.2.15}$$

Proof. Let $M_0 = (3M/\delta\varepsilon)^{\delta/20}$. Then, note that

$$|\{i\in[n] : D_i \geq M_0\}| \sim \mathbf{Bin}\left(n, \mathbb{P}_\mu(D\geq M_0)\right).$$

We can easily see that

$$\mathbb{P}_\mu(D\geq M_0) \leq \frac{M}{M_0^{20/\delta}} = \frac{1}{3}\delta\varepsilon,$$

and hence (4.2.15) follows from (4.2.13). □

Let $\mathcal{A}_n$ be the graph property defined as

$$\mathcal{A}_n := \left\{\left|\left\{i\in[n] : D_i \geq \left(\frac{3M}{\delta\varepsilon}\right)^{\delta/20}\right\}\right| < \frac{1}{2}\delta\varepsilon n\right\}, \tag{4.2.16}$$

where $\{D_i\}_{i=1}^n$ is the degree sequence of a graph, and let $\mathcal{B}_n$ be as (4.2.10). Then for $G_n \sim \mathcal{G}(n,\mu)$, we obtain the following by applying Lemmas 4.2.4 and 4.2.6 with $\eta = \frac{1}{4}$.

$$\mathbb{P}\left(G_n \in \mathcal{A}_n \cap \mathcal{B}_n\right) \geq 1 - \frac{2\varepsilon^{1/4}}{(\varepsilon n)^{k_0-1}} - \exp(-c\delta\varepsilon n), \tag{4.2.17}$$

for some absolute constants $c, C > 0$. Also, note that the event $\mathcal{A}_n \cap \mathcal{B}_n$ is measureable with respect to $\{D_i\}_{i=1}^n$, i.e., independent of the matching between half-edges.

Let $W \in V = [n]$, be a fixed subset of vertices with size $|W| = \varepsilon n$. As discussed in the previous subsection, suppose that we generate G_n, by first pairing the half-edges at W following the cut-off line algorithm. Let H_T be the cut-off line when all half-edges at W are paired for the first time.

Let $d_0 = \mathbb{E}_{D\sim\mu} D$. By applying Lemma 4.2.5 with $\eta = \frac{1}{4}$, we have

$$\mathbb{P}\left(1 - H_T \geq \frac{7M^{\frac{\delta}{20}}\varepsilon^{1-\frac{\delta}{16}}}{d_0}\right) \leq \exp\left\{-c\varepsilon^{1-\frac{\delta}{20}}n\right\}, \tag{4.2.18}$$

for some absolute constant $c > 0$. Let

$$h = 1 - \frac{7M^{\frac{\delta}{20}}\varepsilon^{1-\frac{\delta}{16}}}{d_0},$$

and for each $i \in [n]$ define $Y_i \sim \mathbf{Bin}(D_i, 1-h)$ to be mutually independent given $\{D_i\}_{i=1}^n$. Further, let $\mathcal{C}_n$ denote the event that

$$\mathcal{C}_n := \left\{\sum_{i=1}^n Y_i \geq 6M^{\frac{\delta}{20}}\varepsilon^{1-\frac{\delta}{16}}n\right\}. \tag{4.2.19}$$

Then, observe that

$$\begin{aligned}
&\mathbb{P}\left(|W^{*2}| \geq (1+\delta)\varepsilon n \,\middle|\, \mathcal{A}_n \cap \mathcal{B}_n\right) \\
&\quad\leq \mathbb{P}\left(\mathcal{C}_n^c \mid \mathcal{A}_n \cap \mathcal{B}_n\right) \\
&\qquad + \mathbb{P}\Big(\,|\{i \in [n] : Y_i \geq 2\}| \geq (1+\delta)\varepsilon n \,\Big|\, \mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n\Big).
\end{aligned} \tag{4.2.20}$$

The first term in the r.h.s. can be handled by (4.2.14). For the second, we carry out by bounding Y_i by $Z_i \sim \mathbf{Bin}(M_0, 1-h)$ with $M_0 = (3M/\delta\varepsilon)^{\delta/20}$, where Z_i being independent

of each other. Namely,

$$\begin{aligned}
&\mathbb{P}\left(|\{i \in [n] : Y_i \geq 2\}| \geq (1+\delta)\varepsilon n \mid \mathcal{A}_n \cap \mathcal{B}_n \cap \mathcal{C}_n\right) \\
&\quad \leq (1+o(1))\mathbb{P}\Big(|\{i \in [n] : Y_i \geq 2\}| \geq (1+\delta)\varepsilon n \,\Big|\, \mathcal{A}_n\Big) \\
&\quad \leq (1+o(1))\mathbb{P}\left(|\{i \in [n] : Z_i \geq 2\}| \geq \left(1+\frac{\delta}{2}\right)\varepsilon n\right),
\end{aligned} \tag{4.2.21}$$

where the last inequality follows from the definition of $\mathcal{A}_n$ and Lemma 4.2.6. We can then bound

$$\mathbb{P}(Z_i \geq 2) \leq M_0^2(1-h)^2,$$

and hence obtain that $|\{i \in [n] : Z_i \geq 2\}| \leq_{\text{ST}} \mathbf{Bin}(n, M_0^2(1-h)^2)$, where $\leq_{\text{ST}}$ denotes stochastic domination of random variables. Note that the binomials satisfy the following large deviation estimate (see, e.g., [FK15, Chapter 21])

$$\mathbb{P}(\mathbf{Bin}(n,p) \geq \mu np) \leq \exp(-\mu np \log(\mu/e)). \tag{4.2.22}$$

We thus obtain

$$\text{r.h.s. of (4.2.21)} \;\leq (1+o(1))\exp\left\{-\left(1+\frac{\delta}{2}\right)\log\left(\varepsilon^{-1+\frac{\delta}{3}}\right)\right\}, \tag{4.2.23}$$

if ε satisfies

$$\varepsilon^{\frac{\delta}{12}} \leq \frac{d_0^2\delta^{\frac{\delta}{10}}}{150eM^{\frac{\delta}{5}}}. \tag{4.2.24}$$

Therefore, combining (4.2.14), (4.2.20) and (4.2.23) implies that

$$\mathbb{P}\Big(|W^{*2}| \geq (1+\delta)\varepsilon n \,\Big|\, \mathcal{A}_n \cap \mathcal{B}_n\Big) \leq (1+o(1))e^{-\left(1+\frac{\delta}{6}\right)\varepsilon n\log\left(\frac{1}{\varepsilon}\right)}.$$

Then, we can apply a union bound to obtain that

$$\begin{aligned}
&\mathbb{P}\Big(\exists\, W \subset V \text{ such that } |W| = \varepsilon n \text{ and } |W^{*2}| \geq (1+\delta)\varepsilon n \,\Big|\, \mathcal{A}_n \cap \mathcal{B}_n\Big) \\
&\quad \leq (1+o(1))\binom{n}{\varepsilon n}\exp\left\{-\left(1+\frac{\delta}{6}\right)\varepsilon n\log\left(\frac{1}{\varepsilon}\right)\right\} \\
&\quad \leq (1+o(1))\exp\left\{-\frac{\delta}{6}\varepsilon n\log\left(\frac{1}{\varepsilon}\right)\right\}.
\end{aligned} \tag{4.2.25}$$

Finally, we deduce the conclusion by simply observing

$$\mathbb{P}\left(G_n \in \mathcal{E}^{*2}(\varepsilon n, (1+\delta)\varepsilon n)\right) \geq \mathbb{P}(\mathcal{A}_n \cap \mathcal{B}_n) \cdot \mathbb{P}\Big(G_n \in \mathcal{E}^{*2}(\varepsilon n, (1+\delta)\varepsilon n) \,\Big|\, \mathcal{A}_n \cap \mathcal{B}_n\Big),$$

and then using (4.2.17), for any ε satisfying (4.2.24). □

4.2.3 Expansion of large subsets

Now we deal with the expansion property of large subsets and establish Proposition 4.1.1.

Let $\theta \geq 2$ be an integer, and suppose that $D \sim \mu$ satisfies $\mathbb{P}(D \geq \theta + 2) = 1$ and $\mathbb{E}D^6 = M < \infty$. Also, let $\varepsilon > 0$ be such that εn is an integer, and as (4.2.10), define $\mathcal{B}_n$ to be the event that

$$\mathcal{B}_n := \left\{ \text{for all } W \subset V \text{ with } |W| = \varepsilon n, \ \sum_{i \in W} D_i \leq 3M^{\frac{1}{k_0}} \varepsilon^{1-\frac{1+\eta}{k_0}} n \right\},$$

where $k_0, \eta > 0$ can be any numbers such that $1 + \eta < k_0 \leq 6$. Then, Lemma 4.2.4 implies that

$$\mathbb{P}\left(G_n \in \mathcal{B}_n\right) \geq 1 - \frac{2\varepsilon^{\eta}}{(\varepsilon n)^{k_0 - 1}}.$$

Now, let $W \subset V$ be a fixed subset of vertices with size $|W| = \varepsilon n$, and let $U = V \setminus W$. We will study the set $V \setminus U^{*\theta}$, which can be written as

$$V \setminus U^{*\theta} = \{i \in [n] : e(i, W) \geq D_i - \theta + 1\},$$

where $e(i, W)$ denotes the number of edges between vertex i and the set W. As in the previous subsection, consider the independent binomials $Y_i \sim \mathbf{Bin}(D_i, 1-h)$, where

$$h = 1 - \frac{7M^{\frac{1}{k_0}} \varepsilon^{1-\frac{1+\eta}{k_0}}}{d_0}, \tag{4.2.26}$$

with $d_0 = \mathbb{E}_{D \sim \mu} D$. Suppose that we generate G_n starting from matching the half-edges at W, and let H_T be the height of the cut-off line when all half-edges at W are paired for the first time. Recall the definition of $\mathcal{C}_n$ in (4.2.19). Then, by (4.2.9), similar argument as

(4.2.20) tells us that

$$\begin{aligned}&\mathbb{P}\left(\left|V\setminus U^{*\theta}\right|\geq 0.9\varepsilon n\;\middle|\;\mathcal{B}_n\right)\\&\leq \mathbb{P}\left(\mathcal{C}_n^c\mid\mathcal{B}_n\right)+\mathbb{P}\left(|\{i\in[n]:Y_i\geq D_i-\theta+1\}|\geq 0.9\varepsilon n\mid\mathcal{B}_n\cap\mathcal{C}_n\right).\end{aligned}\tag{4.2.27}$$

Observe that

$$\mathbb{P}(\mathbf{Bin}(d,1-h)\geq d-\theta+1)\leq\binom{d}{\theta-1}(1-h)^{d-\theta+1}=:\mathfrak{p}_{\theta,h}(d).$$

Further, we can find the maximum of $\mathfrak{p}_{\theta,h}(d)$ in the regime $d\geq\theta+2$ by noting that

$$\frac{\mathfrak{p}_{\theta,h}(d+1)}{\mathfrak{p}_{\theta,h}(d)}=\frac{d+1}{d-\theta+2}(1-h)\leq\frac{\theta+3}{4}(1-h).$$

Hence, if $1-h\leq\frac{4}{\theta+3}$, then $\mathfrak{p}_{\theta,h}(d)$ is maximal when $d=\theta+2$. Therefore, for such h, we have

$$\mathbb{P}(Y_i\geq D_i-\theta+1)\leq\binom{\theta+2}{\theta-1}(1-h)^3\leq\{(\theta+2)(1-h)\}^3.\tag{4.2.28}$$

Therefore, having (4.2.22) and (4.2.26) in mind, we see that the last term in (4.2.27) is bounded by

$$\begin{aligned}&(1+o(1))\mathbb{P}\left(|\{i\in[n]:Y_i\geq D_i-\theta+1\}|\geq 0.9\varepsilon n\right)\\&\qquad\leq(1+o(1))\mathbb{P}\left(\mathbf{Bin}(n,\{(\theta+2)(1-h)\}^3)\geq 0.9\varepsilon n\right)\\&\qquad\leq\exp\left\{-0.9\varepsilon n\log\left(\varepsilon^{-2+\frac{4+3\eta}{k_0}}\right)\right\},\end{aligned}$$

if ε satisfies

$$\varepsilon\leq e^{-k_0}\left(\frac{d_0}{7(\theta+2)M}\right)^{3k_0}.\tag{4.2.29}$$

Note that we used (4.2.22) for the last inequality. Thus, if we take $k_0=6$, $\eta=\frac{1}{5}$, we get

$$\begin{aligned}\mathbb{P}\left(\left|V\setminus U^{*\theta}\right|\geq 0.9\varepsilon n\;\middle|\;\mathcal{B}_n\right)&\leq e^{-c\varepsilon^{1-\frac{1}{6}}n}+e^{-1.1\varepsilon n\log\left(\frac{1}{\varepsilon}\right)}\\&\leq 2e^{-1.1\varepsilon n\log\left(\frac{1}{\varepsilon}\right)},\end{aligned}\tag{4.2.30}$$

where $c>0$ is an absolute constant from (4.2.14), and the last inequality holds true if we

take ε small so that $c\varepsilon^{1/6} \geq \log(1/\varepsilon)$. Therefore, the conclusion (4.1.3) follows after we take a union bound as in (4.2.25), if we set ε_1 to be small so that it satisfies

$$\varepsilon_1^{1-\frac{1+\eta}{6}} \leq \frac{4d_0}{7(\theta+2)M^{\frac{1}{6}}}, \quad \varepsilon_1 \leq \left(\frac{e^{-\frac{1}{3}}d_0}{7(\theta+2)M}\right)^{3k_0}, \quad c\varepsilon_1^{\frac{1}{6}} \geq \log\left(\frac{1}{\varepsilon_1}\right), \tag{4.2.31}$$

where each condition follows from (4.2.28), (4.2.29) and (4.2.30), respectively. □

4.3 Discontinuous phase transitions on random graphs

In this section, we establish Theorems 12 and 14. Along with the results obtained from above, Theorem 12-1 follows by a simple argument, similarly as [CD13, Section 2]. Theorem 13 requires additional care as mentioned in Section 4.1, where we discuss a biased random walk argument to control the expansion of infections when their size is small. In order to settle Theorem 14, we introduce some background on the k-cores of random graphs in Section 4.3.5.

4.3.1 Proof of Theorem 12-1

Let $\delta_1 = 0.1$ and $\varepsilon_1 > 0$ be the constant as in Proposition 4.1.1. Set

$$p_1 := \frac{1-\varepsilon_1}{1-(1-\frac{\delta_1}{2})\varepsilon_1}.$$

Suppose that the threshold-θ contact process (X_t) with $p \geq p_1$ satisfies $|X_t| \geq (1-\varepsilon_1)n$ at some time t. Recalling the definiton of $\mathcal{F}^{*\theta}(m_1, m_2)$ in (4.1.4), we have

$$\begin{aligned} G_n \in \mathcal{F}^{*\theta}(\lceil(1-\varepsilon_1)n\rceil, \lceil(1-(1-\delta_1)\varepsilon_1)n\rceil) \quad &\text{implies} \\ |X_{t+1}| \ \geq_{\mathrm{ST}} \ \mathbf{Bin}\left(\lceil(1-(1-\delta_1)\varepsilon_1)n\rceil, p\right), & \end{aligned}$$

where $\geq_{\mathrm{ST}}$ denotes stochastic domination between random variables. Thus, on the event $\mathcal{F}^{*\theta}((1-\varepsilon_1)n, (1-(1-\delta_1)\varepsilon_1)n)$, a standard large deviation estimate for binomials (e.g.,

[FK15, Section 22.4]) implies that

$$\mathbb{P}_{\mathrm{CP}}\left(|X_{t+1}| \le (1-\varepsilon_1)n \,|\, |X_t| \ge (1-\varepsilon_1)n\right) \le \exp\left(-\frac{\delta_1^2\varepsilon_1^2}{12}n\right),$$

where $\mathbb{P}_{\mathrm{CP}}$ denotes the probability coming from the randomness of (X_t).

Set $\tau := \exp\left(\frac{\delta_1^2\varepsilon_1^2}{13}n\right)$, and let $X_0 \equiv \mathbf{1}$. Then, the probability that $|X_t| \ge (1-\varepsilon_1)n$ fails for some $t \in [0,\tau]$ is exponentially small in n. Finally, Proposition 4.1.1 tells us that $G_n \in \mathcal{F}^{*\theta}((1-\varepsilon_1)n, (1-(1-\delta_1)\varepsilon_1)n)$ whp. □

4.3.2 Proof of Theorem 12-2: overview

As discussed in Section 4.1, the argument in the previous section or in Section 3 of [CD13] does not work the same for Theorem 12-2. Instead, the proof will proceed in two steps as follows.

1. Starting from $\varepsilon_2 n$ infections where $\varepsilon_2 > 0$ is a small fixed constant, we first reduce the size of infection into $\Delta \log\log n$ in time $C_1 \log n$, where $\Delta > 0$ is another small fixed constant.

2. We eliminate the remaining infections *by luck*, meaning that we rely on the event where all the infections are recovered in a single time step. If this fails and the infection expands into a bigger size than $\Delta \log\log n$, then we show that it quickly returns to size $\Delta \log\log n$ after a short time, where we try another update to kill every infection at once.

We prove the short survival result for the threshold-2 contact process, which clearly implies the result for general $\theta \ge 2$. Let $p \in (0,1)$ be the given probability parameter of the threshold-θ contact process (X_t), and set

$$\delta_2 := \frac{1}{3}(1-p), \qquad \Delta = (3\log(\frac{1}{1-p}))^{-1}.$$

Further, let $\varepsilon_2 = \varepsilon_2(\delta_2)$ as in Proposition 4.1.2. Assume that $G_n \sim \mathcal{G}(n,\mu)$ satisfies

$\mathcal{E}^{*2}(m,(1+\delta_2)m)$ for all $m \in [\Delta \log\log n,\, \varepsilon_2 n]$, which happens with probability

$$\begin{aligned}
&\mathbb{P}\left(G_n \in \bigcap_{m=\Delta\log\log n}^{\varepsilon_2 n} \mathcal{E}^{*2}(m,(1+\delta_2)m)\right) \\
&\quad \geq 1 - \sum_{m=\Delta\log\log n}^{\varepsilon_2 n} \left\{ \frac{C}{m^{k_2-1}} \left(\frac{m}{n}\right)^{\frac{1}{4}} + e^{-c\delta m} \right\} \qquad (4.3.1) \\
&\quad \geq 1 - 2\,(\log n)^{-c\delta\Delta} = 1 - o(1),
\end{aligned}$$

where we set $k_2 = 20/\delta_2$.

4.3.3 Reducing the infection to a small size

Define the event $\mathcal{D}_n$ by

$$\mathcal{D}_n := \bigcap_{m=\Delta\log\log n}^{\varepsilon_2 n} \mathcal{E}^{*2}(m,(1+\delta_2)m),$$

the event in the l.h.s. of (4.3.1).

To begin with, we show that (X_t) started from $|X_0| \leq \varepsilon_2 n$ reaches size $\Delta\log\log n$ in time $O(\log n)$ whp. Therefore, let us assume that $|X_0| \geq \Delta\log\log n$, otherwise there is nothing to prove.

On $G_n \in \mathcal{D}_n$, if X_t at some time t is $\Delta\log\log n \leq |X_t| \leq \varepsilon_2 n$, we have

$$|X_{t+1}| \leq_{\mathrm{ST}} \mathbf{Bin}((1+\delta_2)|X_t|,\, p).$$

Noting that $p(1+\delta_2) \leq 1 - 2\delta_2$, we obtain $\mathbb{E}_{\mathrm{CP}}[|X_{t+1}| \,|\, |X_t|] \leq (1-2\delta_2)|X_t|$, and

$$\mathbb{P}_{\mathrm{CP}}\left(|X_{t+1}| \geq (1-\delta_2)m \,|\, |X_t| = m,\, \mathcal{D}_n\right) \leq \exp\left(-\frac{\delta_2^2 m}{3}\right), \qquad (4.3.2)$$

Let τ_2' be the first time when $|X_t| \leq \sqrt{n}$. Then, there exists some constant C_2' such that

$$\mathbb{P}_{\mathrm{CP}}\left(\tau_2' \geq C_2' \log n \,\middle|\, |X_0| \leq \varepsilon_2 n,\, \mathcal{D}_n\right) \leq \exp\left(-\frac{\delta_2^2 n^{1/2}}{4}\right), \qquad (4.3.3)$$

where the r.h.s. indicates the probability that the events $\{|X_{t+1}| \leq (1-\delta_2)|X_t|\}$ hold for

all t until the infection size reaches $\sqrt{n}$.

To make (X_t) reach $\Delta \log\log n$ from $|X_{\tau_2'}| \le n^{1/2}$, we consider a coupling between (X_t) and a biased random walk which we define as follows. Let Q_s, $s \in [n]$ be i.i.d. random variables given by

$$\mathbb{P}(Q_s = 1+\delta_2) = (\log n)^{-\gamma}, \quad \mathbb{P}(Q_s = 1-\delta_2) = 1-(\log n)^{-\gamma},$$

where $\gamma := \delta_2^2\Delta/3$. Here, $\mathbb{P}(Q_s = 1+\delta_2)$ corresponds to the minimum of the r.h.s. of (4.3.2) over $m \in [\Delta \log\log n,\, \varepsilon_2 n]$. Define $(\widehat{X}_t)_{t\ge\tau_2'}$ by $\widehat{X}_{\tau_2'} = \sqrt{n}$, and

$$\widehat{X}_{\tau_2'+t} = \widehat{X}_{\tau_2'} \prod_{s=1}^{t} Q_s.$$

Let $T > 0$. Then, conditioned on the events

$$G_n \in \mathcal{D}_n \ \text{ and } \ |X_{\tau_2'+t}| \in [\Delta \log\log n,\, \varepsilon_2 n] \ \text{ for all } t \in [0,T], \tag{4.3.4}$$

we have a natural coupling between (X_t) and $(\widehat{X}_t)$ such that

$$\left|X_{\tau_2'+t}\right| \le \widehat{X}_{\tau_2'+t}, \ \text{ for all } t \in [0,T].$$

Let t_2 set to be

$$t_2 = \left\lceil \frac{3\log n}{2\log(1+\delta_2)} \right\rceil.$$

(Note that t_2 is deterministic whereas τ_2' is not.) Then, we can control the number of Q_s that equals $1+\delta_2$ as follows.

$$\begin{aligned} &\mathbb{P}\left(|\{s \in [1,t_2] : Q_s = 1+\delta_2\}| \ge \frac{t_2}{4} \right) \\ &\quad = \mathbb{P}\left(\mathbf{Bin}(t_2,\, (\log n)^{-\gamma}) \ge \frac{t_2}{4} \right) = o(1). \end{aligned} \tag{4.3.5}$$

Since $(1-\delta)(1+\delta) < 1$, this means that $(\widehat{X}_{\tau_2'+t})$ reaches $\Delta \log\log n$ before $\varepsilon_2 n$ in time $t \le t_2$ whp. This implies that if $G_n \in \mathcal{D}_n$, the second event in (4.3.4) holds true whp until $|X_t|$ reaches $\Delta \log\log n$ at some $t \le \tau_2' + t_2$. Thus, we obtain that on $G_n \in \mathcal{D}_n$, $|X_t|$ becomes

at most $\Delta \log\log n$ in time $C_2' \log n + t_2 = O(\log n)$ whp for C_2' as in (4.3.3), if started from $|X_0| \le \varepsilon_2 n$.

4.3.4 Elemination of small infections

Under the same setting as the previous subsection, define $\tau_2(0)$ to be the first time when $|X_t|$ is at most $\Delta \log\log n$. We showed that there exists a constant C_2 such that

$$\mathbb{P}(\tau_2(0) \ge C_2 \log n \,|\, |X_0| \le \varepsilon_2 n,\, \mathcal{D}_n) = o(1).$$

For each $i \in \mathbb{N}$, define $\tau_2(i)$ inductively so that

$$\tau_2(i+1) := \min\{t \ge \tau_2(i) + 1 : |X_t| \le \Delta \log\log n\}.$$

At time $\tau_2(0)+1$, we obtain $X_{\tau_2(0)+1} = \mathbf{0}$ with probability at least

$$(1-p)^{\Delta \log\log n} = (\log n)^{-1/3}. \tag{4.3.6}$$

If we fail to achieve such a lucky event, we claim that

$$\mathbb{P}\left(\tau_2(1) - \tau_2(0) \ge \sqrt{\log n} \;\middle|\; \tau_2(0),\, \mathcal{D}_n\right) \le \exp\left(-\frac{1}{4}\sqrt{\log n}\right), \tag{4.3.7}$$

for some $\alpha > 0$. This can be shown using a similar estimate as (4.3.5). That is,

$$\begin{aligned} &\mathbb{P}\left(\left|\left\{s \in \left[1, \sqrt{\log n}\right] : Q_s = 1 + \delta_2\right\}\right| \ge \frac{\sqrt{\log n}}{4}\right) \\ &\quad = \mathbb{P}\left(\mathbf{Bin}(\sqrt{\log n},\, (\log n)^{-\gamma}) \ge \frac{\sqrt{\log n}}{4}\right) \le \exp\left(-\frac{1}{4}\sqrt{\log n}\right), \end{aligned} \tag{4.3.8}$$

where the last inequality follows from (4.2.22). Noting that $|X_{\tau_2(0)+1}| \le 2\Delta \log\log n$, the event inside the l.h.s. of (4.3.8) implies that $(X_{\tau_2(0)+t})$ reaches $\Delta \log\log n$ before εn. Of course, this happens the same for $\tau_2(i+1) - \tau_2(i)$ with any $i \in \mathbb{N}$.

Now, let $I = \min\{i \in \mathbb{N} : \tau_2(i+1) - \tau_2(i) \geq \sqrt{\log n}\}$. Then, (4.3.7) tells us that

$$\mathbb{P}\left(I \leq \sqrt{\log n} \,\middle|\, |X_0| \leq \varepsilon_2 n, \mathcal{D}_n \right) \leq \sqrt{\log n} \cdot e^{-\frac{1}{4}\sqrt{\log n}} \leq e^{-\frac{1}{5}\sqrt{\log n}}. \tag{4.3.9}$$

This implies that $\tau_2(\sqrt{\log n}) \leq \sqrt{\log n} \cdot \sqrt{\log n} = \log n$ whp. Moreover, since $X_{\tau_2(i)+1} = \mathbf{0}$ with probability at least $(\log n)^{-1/3}$ by (4.3.6),

$$\begin{aligned}
&\mathbb{P}(X_{\tau_2(i)+1} \neq \mathbf{0} \text{ for all } i \leq \sqrt{\log n}) \\
&\quad \leq \mathbb{P}\left(\mathbf{Bin}\left(\sqrt{\log n}, (\log n)^{-\frac{1}{3}} \right) = 0 \right) \leq \exp\left(-(\log n)^{\frac{1}{6}} \right),
\end{aligned}$$

which tells us that $X_{\tau_2(\sqrt{\log n})} = \mathbf{0}$ whp. Along with the discussion in Section 4.3.3, we deduce that on $G_n \in \mathcal{D}_n$, the process started from $|X_0| \leq \varepsilon_2 n$ dies out in time $(C_2+1)\log n$ whp. Since $\mathbb{P}(G_n \in \mathcal{D}_n) = 1 - o(1)$ by (4.3.1), the conclusion of Theorem 12-2 follows. □

4.3.5 Proof of Theorem 14

The goal of this subsection is to establish Theorem 14 and Corollary 15. To this end, we first show that Proposition 4.1.1 continues to hold when we replace the assumption $D \geq \theta + 2$ by the $(\theta+2)$-*core condition*. Recall that the r-core of G is the largest induced subgraph of G whose minimal degree is r.

Let μ be a given degree distribution with $d = \mathbb{E}_{D\sim\mu} D$. Recall Proposition 1.6.1 where we defined $F_r(h)$ and $\rho_r(h)$, and saw that the existence of $h \in (0,1)$ satisfying $dh^2 < F_r(h)$ implies the existence of the r-core of size $\approx \rho_r(\hat{h})n$ in $G_n \sim \mathcal{G}(n,\mu)$ whp, as mentioned in the proposition.

Let K_n be the r-core of G_n, and consider the event $\mathcal{F}^{*\theta}(m_1, m_2)$ for K_n. That is,

$$K_n \in \mathcal{F}^{*\theta}(m_1, m_2) \text{ if and only if } \forall\, W \subset K_n \text{ of size } m_1,\ |W^{*\theta}| \geq m_2.$$

Here, note that $W^{*\theta}$ is defined in terms of K_n, so that it contains vertices in K_n that has at least θ neighbors in W. Under this setting, we have the following generalization of Proposition 4.1.1.

Proposition 4.3.1. *Let $\theta \geq 2$ be an integer and μ be a degree distribution that satisfies*

$\mathbb{E}_{D\sim\mu}D^6 = M < \infty$ *and (1.6.4) with* $r = \theta + 2$. *Let* $\rho = \rho_{\theta+2}(\hat{h})$ *with* $\rho_{\theta+2}$ *and* $\hat{h}$ *as in (1.6.3) and the discussion below (1.6.4). Then, there exists a constant* $\varepsilon_1(\theta, M) > 0$ *such that for* $G_n \sim \mathcal{G}(n, \mu)$, *its* $(\theta+2)$*-core* K_n *is of size* $1 - \frac{1}{100}\varepsilon_1 \le |K_n|/\rho n \le 1 + \frac{1}{100}\varepsilon_1$ *whp, and satisfies*

$$\mathbb{P}\left(K_n \in \mathcal{F}^{*\theta}\left(\lceil(1-\varepsilon_1)\rho n\rceil, \lceil(1-0.9\varepsilon_1)\rho n\rceil\right)\right) \ge 1 - o(1), \tag{4.3.10}$$

for all large enough n, *where* $C > 0$ *is an absolute constant.*

Proof. Let $\varepsilon > 0$ be a small constant that will be chosen later. Let $\{D_i\}_{i=1}^n$ be the degree sequence of G_n generated by i.i.d. samples of μ, and define the event $\mathcal{B}_n$ by

$$\mathcal{B}_n := \left\{\text{for all } W \subset V \text{ with } |W| = \varepsilon n,\ \sum_{i\in W} D_i \le 3M^{\frac{1}{6}}\varepsilon^{\frac{4}{5}}n\right\}.$$

Then, by Lemma 4.2.4, we have

$$\mathbb{P}\left(G_n \in \mathcal{B}_n\right) \ge 1 - \frac{C\varepsilon^{\frac{1}{5}}}{(\varepsilon n)^5}.$$

Now, given the degree sequence $\{D_i\}_{i=1}^n$ of G_n, we give an ordering to all half-edge clones as in Definition 4.2.2. Then we generate K_n, the $(\theta+2)$-core, using the cut-off line algorithm as follows:

0. Set the initial height of the cut-off line $H_0 = 1$. Let L_0 be the set of vertices i with $D_i \le \theta + 1$, and R_0 be the set of half-edge clones attached to L_0. We call a vertex is *light* if it has less than $\theta + 2$ clones under the cut-off line.

1. At step t, we match the half-edge a_t in R_{t-1} that is chosen with respect to the prescribed order with the highest unmatched clone a_t' (i.e., having the largest B_j in the language of Definition 4.2.2) and set H_t to be its height.

2. Let v_t be the vertex containing a_t'. Update R_t by

$$R_t = R_{t-1} \setminus \{a_t, a_t'\}.$$

Moreover, add all the half-edges at v_t below H_t to R_t, if the number of them is less than $\theta+2$, i.e., v_t becomes *light* at step t.

Let $\mathcal{K}_n$ be the event that K_n, the $(\theta+2)$-core of G_n, is of size

$$1-\frac{\varepsilon}{100}\le\frac{|K_n|}{\rho n}\le 1+\frac{\varepsilon}{100}.$$

If μ satisfies (1.6.4) with $r=\theta+2$, then we have $\mathcal{K}_n$ whp, so that the process terminates in the sense that at least ρn vertices remain *heavy* when R_t becomes empty. Let H_{CORE} be the height of the cut-off line when R_t is empty for the first time. Then the vertices V_{CORE} remaining at that time forms K_n, and each $i\in V_{\mathrm{CORE}}$ has $D_i'\ge\theta+2$ clones below H_{CORE}.

Let $U\subset V_{\mathrm{CORE}}$ be any fixed subset of size $\lceil\varepsilon|V_{\mathrm{CORE}}|\rceil$. Lemma 4.2.4 implies that if $G_n\in\mathcal{B}_n$, regardless of the choice of U

$$\sum_{i\in U}D_i'\le\sum_{i\in U}D_i\le 3M^{\frac{1}{6}}\varepsilon^{\frac{4}{5}}n, \tag{4.3.11}$$

by choosing $k=6$ and $\eta=\frac{1}{5}$ in (4.2.2). Suppose that we start running the cut-off line algorithm again from the initial height H_{CORE} to match all the remaining half-edges at U. Then, thanks to (4.3.11), we can repeat the same argument as proof of Proposition 4.1.1 and obtain

$$\begin{aligned}&\mathbb{P}\left(\left|V_{\mathrm{CORE}}\setminus U^{*\theta}\right|\ge 0.9\varepsilon|V_{\mathrm{CORE}}|\,\middle|\,|V_{\mathrm{CORE}}|,\,\mathcal{B}_n\cap\mathcal{K}_n\right)\\&\quad\le 2\exp\left(-1.1\varepsilon\log\left(\frac{1}{\varepsilon}\right)|V_{\mathrm{CORE}}|\right).\end{aligned}$$

The only difference is to replace d_0 in Lemma 4.2.5 and (4.2.18) by $\hat{d}_0$, where

$$\hat{d}_0=F_{\theta+2}(\hat{h}),$$

with $F_{\theta+2}$ and $\hat{h}$ as in (1.6.3) and the discussion below (1.6.4). Here, $\hat{d}_0n/2$ corresponds to the expected total number of edges in K_n (see [JL07, Theorem 2.3]).

Therefore, applying the union bound over all choices of U deduces the conclusion for ε_1 satisfying (4.2.31), since the event $\mathcal{B}_n\cap\mathcal{K}_n$ happens whp for $G_n\sim\mathcal{G}(n,\mu)$. □

We conclude the section by proving Theorem 14 and Corollary 15.

Proof of Theorem 14. Let $\rho = \rho_{\theta+2}(\hat{h})$, where $\rho_{\theta+2}$ and $\hat{h}$ are as in (1.6.3) and the discussion below (1.6.4). To establish the first part of the theorem, we only focus on the infections inside K_n, the $(\theta+2)$-core of G_n, and repeat the argument in Section 4.3.1 with

$$p := \frac{1 - \frac{99}{100}\varepsilon_1}{1 - \frac{95}{100}\varepsilon_1},$$

for ε_1 as in Proposition 4.3.1. Then, starting from $X_0 \equiv \mathbf{1}$, $(1-\varepsilon_1)$ fraction of K_n will remain infected for exponentially long time, and hence we deduce Theorem 14-1.

Suppose that we want to establish the second part of the theorem with the above p. Recalling Propositin 4.2.1 and the discussion in Section 4.3.2, we required μ to satisfy

$$\mathbb{E}_{D\sim\mu} D^{20/\delta} < \infty,$$

for $\delta = \frac{1}{3}(1-p)$. Here, we can see that $\frac{20}{\delta} = 60(1-p)^{-1} \le 1500\varepsilon_1^{-1}$. Based on the conditions for ε_1 given in (4.2.31), we can set ε_1 to be

$$\varepsilon_1^{-1} = C_1'(\theta M)^{C_2},$$

for some absolute constants $C_1', C_2 > 0$. Setting $C_1 = 1500C_1'$, μ falls into the regime where Proposition 4.2.1 works for $\delta = \frac{1}{3}(1-p)$, and hence we can repeat the argument in Section 4.3.2 to deduce Theorem 14-2. □

Proof of Corollary 15. We can instead work with $G_n \sim \mathcal{G}(n, \mathrm{Pois}(d))$ (Section 1.4.3). Then the first part of the Corollary can be proven analogously as above. Moreover, since the Poisson distribution has all polynomial moments, we can establish the second part for all p as in Section 4.3.2. □

4.4 Short survival on low-degree random regular graphs: Proof of Theorem 13

In this section, we establish Theorem 13. Denoting the threhold-θ contact process by (X_t) and random $(\theta+1)$-regular graph by $G_n \sim \mathcal{G}(n, \theta+1)$, the proof consists of the following observations.

O1. For a cycle C inside G_n, if $X_t(C) \equiv \mathbf{0}$ at some time t, then $X_s(C) \equiv \mathbf{0}$ for all $s \geq t$.

O2. There exists a constant $\kappa(\theta)$ such that whp, $G_n \sim \mathcal{G}(n, \theta+1)$ can be covered by cycles of length at most $\kappa \log n$.

The first observation is based on the definition of (X_t). That is, if the entire cycle C is healthy, then at each $v \in C$ we can only find at most $\theta - 1$ infected neighbors of v, and hence v stays healthy in the next time step. The second one is obtained by the following proposition.

Proposition 4.4.1. *Let $d \geq 3$ be an integer. There exists a constant $\kappa(d) > 0$ such that whp over $G_n \sim \mathcal{G}(n, d)$, we have that for all $v \in V(G_n)$, there is a cycle $C_v \ni v$ of length at most $\kappa \log n$.*

Assuming Proposition 4.4.1, the observations above easily imply Theorem 13.

Proof of Theorem 13. Suppose that $G_n \sim \mathcal{G}(n, \theta+1)$ can be covered by cycles of length at most $\kappa \log n$ for some constant $\kappa > 0$. Set $a = 1 + \kappa \log(\frac{1}{1-p})$. For a cycle C of length at most $\kappa \log n$, we have

$$\mathbb{P}_{\mathrm{CP}}\left(X_{n^a}(C) \neq \mathbf{0}\right) \leq \mathbb{P}\left(\mathbf{Bin}\left(n^a, (1-p)^{\kappa \log n}\right) = 0\right) \leq e^{-n},$$

where $(1-p)^{\kappa \log n}$ corresponds to the probability that the entire cycle C gets recovered at a single time step.

Let $C_1, \ldots, C_m$ be the cycles of length at most $\kappa \log n$ that cover G_n. Clearly $m \leq n$,

and the above estimate tells us that

$$\mathbb{P}_{\mathrm{CP}}\left(\bigcap_{l=1}^{m}\{X_{ln^a}(C_l)\equiv\mathbf{0}\}\right)\geq 1-ne^{-n},$$

where we devote each time interval $[(l-1)n^a+1,\, ln^a]$ to eliminating infections inside C_l. Further, the observation O1 tells us that if $X_{ln^a}(C_l)\equiv\mathbf{0}$ for each l, then $X_{mn^a}\equiv\mathbf{0}$. Therefore, the infection can survive at most n^{a+1}-time whp, given that G_n can be covered with cycles of length at most $\kappa\log n$. Since the latter holds whp over the choice of $G_n\sim\mathcal{G}(n,\theta+1)$, the conclusion of Theorem 13 follows. □

We conclude this section by showing Proposition 4.4.1, which can be done similarly as [BFdlV82].

Proof of Proposition 4.4.1. Let v be a fixed vertex in $G_n\sim\mathcal{G}(n,d)$ and set $L=2\log_{d-1}n$. We will show that the probability that the neighborhood $N(v,L)$ contains a cycle crossing v is at least $1-o(n^{-1})$. Then, the error $o(n^{-1})$ allows us to take a union bound over all vertices, hence implying Proposition 4.4.1.

Let $u_1,\dots,u_d$ be the neighbors of v (if v has a self-loop, then we are done, so we assume that it has d neighbors). Define the *branch* of u_i with respect to v by

$$Br(u_i,v;l):=\big\{u\in V(G_n):\exists\text{ a path of length}\leq l\text{ between }u\text{ and }u_i\text{ that does not cross }v\big\}.$$

If there exist $i\neq j$ such that $Br(u_i,v;L_0+1)\,\cap\,Br(u_j,v;L_0+1)\neq\emptyset$ with L_0 given by $L_0=\lfloor\log_{d-1}(n^{1/2}\log n)\rfloor$, then we are done. Indeed, we claim that

$$\mathbb{P}\left(Br(u_1,v;L_0+1)\cap Br(u_2,v;L_0+1)\neq\emptyset\right)\geq 1-o(n^{-1}). \tag{4.4.1}$$

To see this, we first see that the branches have enough expansion with large probability. Suppose that we have explored $N(v,h)$, the vertices at distance h. At this point, the vertices at distance exactly h from v have unmatched half-edges which will be matched in the next step when we explore $N(v,h+1)$. Let $\partial N(v,h)$ be the collection of those unmatched half-edges, and let $\partial Br(i,h-1)$ be the ones who are attached to vertices in $Br(u_i,v;h-1)$.

Then, as shown in equation (1) of [BFdlV82],

$$\mathbb{P}\left(|\partial Br(i, L_0)| \geq \frac{1}{2}(d-1)^{L_0+1} \text{ for all } i \in [d]\right) \geq 1 - o(n^{-1}). \tag{4.4.2}$$

Further, given that $|\partial Br(1, L_0)|, |\partial Br(2, L_0)| \geq \frac{1}{2} n^{1/2} \log n$, the probability that there exist two half-edges, each from $Br(1, L_0)$ and $Br(2, L_0)$, which are paired in the next exploration step is at least

$$\mathbb{P}\left(\mathbf{Bin}\left(\frac{n^{1/2}\log n}{2}, \frac{\log n}{2dn^{1/2}}\right) \geq 1\right) \geq 1 - \exp\left(-\frac{(\log n)^2}{4d}\right). \tag{4.4.3}$$

Therefore, combining (4.4.2) and (4.4.3) implies (4.4.1), and hence we obtain the conclusion.

□

Bibliography

[AF02] D. Aldous and J. A. Fill. *Reversible Markov Chains and Random Walks on Graphs*. Unfinished monograph, recompiled 2014, available at the url http://www.stat.berkeley.edu/∼aldous/RWG/book.html, 2002.

[AL03] J. Adler and U. Lev. Bootstrap percolation: Visualizations and applications. *Brazilian Journal of Physics*, 33:641–644, 2003.

[AN72] K. B. Athreya and P. E. Ney. *Branching processes*. Springer, 1972.

[AS10] S. R. Athreya and J. M. Swart. Survival of contact processes on the hierarchical group. *Probab. Theory Related Fields*, 147(3-4):529–563, 2010.

[BBCS05] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 301–310. Society for Industrial and Applied Mathematics, 2005.

[BdSV20] S. A. Bethuelsen, G. B. da Silva, and D. Valesin. Graph constructions for the contact process with a prescribed critical rate. *arXiv preprint, arXiv:2005.00420*, 2020.

[BFdlV82] B. Bollobás and W. Fernandez de la Vega. *The diameter of random regular graphs*, volume 2. 1982.

[BG90] C. Bezuidenhout and G. Grimmett. The critical contact process dies out. *The Annals of Probability*, 18(4):1462–1482, 1990.

[BG91] M. Bramson and L. Gray. A useful renormalization argument. In *Random Walks, Brownian Motion, and Interacting Particle Systems. Progress in Probability*, volume 28, pages 113–152. Birkhäuser, Boston, 1991.

[BNNS18] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly. The contact process on random graphs. Conference Presentation at *Advances in Asymptotic Probability, in honor of Amir Dembo's 60th birthday*, Stanford, CA, Presented by A. Sly, 2018.

[BNNS19] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly. Survival and extinction of epidemics on random graphs with general degrees. Conference Presentation at *Columbia-Princeton Probability Day 2019*, Princeton, NJ, Presented by O. Nguyen, 2019.

[BNNS20] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly. The contact process on random graphs. Conference Presentation at *Special Session on Stochastic Spatial Models, Joint Mathematics Meetings*, Denver, CO, Presented by D. Nam, 2020.

[BNNS21] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly. Survival and extinction of epidemics on random graphs with general degree. *Ann. Probab.*, 49(1):244–286, 2021.

[Bol01] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2nd edition, 2001.

[Bov06] A. Bovier. Metastability: a potential theoretic approach. In *International Congress of Mathematicians. Vol. III*, pages 499–518. Eur. Math. Soc., Zürich, 2006.

[CD09] S. Chatterjee and R. Durrett. Contact process on random graphs with power law degree distribution have critical value 0. *Ann. Probab.*, 37(6):2332–2356, 2009.

[CD13] S. Chatterjee and R. Durrett. A first order phase transition in the threshold $\theta \geq 2$ contact process on random r-regular graphs and r-trees. *Stochastic Process. Appl.*, 123(2):561–578, 2013.

[Che92] H.-N. Chen. On the stability of a population growth model with sexual reproduction on $\mathbb{Z}^2$. *Ann. Probab.*, 20(1):232–285, 1992.

[Che94] H.-N. Chen. On the stability of a population growth model with sexual reproduction on $\mathbb{Z}^d$, $d \geq 2$. *Ann. Probab.*, 22(3):1195–1226, 1994.

[CMMV14] M. Cranston, T. Mountford, J.-C. Mourrat, and D. Valesin. The contact process on finite trees revisited. *ALEA Lat. Am. J. Probab. Math. Stat.*, 11(2):385–408, 2014.

[CS15] V. H. Can and B. Schapira. Metastability for the contact process on the configuration model with infinite mean degree. *Electron. J. Probab*, 20(26):1–22, 2015.

[CSW18] S. Chatterjee, D. Sivakoff, and M. Wascher. The effect of avoiding known infected neighbors on the persistence of a recurring infection process. *arXiv preprint, arXiv:1811.00627*, 2018.

[DG85] R. Durrett and L. Gray. Some peculiar properties of a particle system with sexual reproduction. *Unpublished manuscript*, 1985.

[DL88] R. Durrett and X.-F. Liu. The contact process on a finite set. *Ann. Probab.*, 16(3):1158–1173, 1988.

[DM10] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Brazilian Journal of Probability and Statistics*, 24(2):137–211, 2010.

[DR17] A. Deshayes and L. T. Rolla. Scaling limit of subcritical contact process. *Stochastic Process. Appl.*, 127(8):2630–2649, 2017.

[DS88] R. Durrett and R. H. Schonmann. The contact process on finite set. ii. *Ann. Probab.*, 16(3):1570–1583, 1988.

[Dur07] R. Durrett. *Random graph dynamics*, volume 200. Cambridge university press Cambridge, 2007.

[Dur10] R. Durrett. *Probability: Theory and Examples*. Cambridge U. Press, 4th edition, 2010.

[DY20] R. Durrett and D. Yao. The symbiotic contact process. *Electron. J. Probab.*, 25:Paper No. 4, 21, 2020.

[ER59] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[FK15] A. Frieze and M. Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2015.

[FMMV19] L. R. G. Fontes, D. H. U. Marchetti, T. S. Mountford, and M. E. Vares. Contact process under renewals I. *Stochastic Process. Appl.*, 129(8):2903–2911, 2019.

[FMUV21] L. R. Fontes, T. S. Mountford, D. Ungaretti, and M. E. Vares. Renewal contact processes: phase transition and survival. *arXiv preprint, arXiv:2101.06207*, 2021.

[FMV20] L. R. Fontes, T. S. Mountford, and M. E. Vares. Contact process under renewals II. *Stochastic Process. Appl.*, 130(2):1103–1118, 2020.

[FS08] L. R. G. Fontes and R. H. Schonmann. Threshold $\theta \geq 2$ contact processes on homogeneous trees. *Probab. Theory Related Fields*, 141(3-4):513–541, 2008.

[Har74] T. E. Harris. Contact interactions on a lattice. *Ann. Probab.*, 2(6):969–988, 1974.

[HD20a] X. Huang and R. Durrett. The contact process on random graphs and Galton Watson trees. *ALEA Lat. Am. J. Probab. Math. Stat.*, 17(1):159–182, 2020.

[HD20b] X. Huang and R. Durrett. A stochastic spatial model for the sterile insect control strategy. *arXiv preprint, arXiv:2009.09464*, 2020.

[HP20] M. Holmes and E. Perkins. On the range of lattice models in high dimensions. *Probab. Theory Related Fields*, 176(3-4):941–1009, 2020.

[Hua20] X. Huang. Exponential growth and continuous phase transitions for the contact process on trees. *Electron. J. Probab.*, 25:Paper No. 77, 21, 2020.

[Jan09] S. Janson. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, 18(1-2):205–225, 2009.

[JL07] S. Janson and M. J. Luczak. A simple solution to the k-core problem. *Random Structures Algorithms*, 30(1-2):50–62, 2007.

[JLM19] E. Jacob, A. Linker, and P. Mörters. Metastability of the contact process on fast evolving scale-free networks. *Ann. Appl. Probab.*, 29(5):2654–2699, 2019.

[JLR00] S. Janson, T. Luczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[Kim06] J. H. Kim. Poisson cloning model for random graphs, International Congress of Mathematicians. Vol. III. *Eur. Math. Soc.*, pages 873–897, 2006.

[Lig96] T. M. Liggett. Multiple transition points for the contact process on the binary tree. *Ann. Probab.*, 24(4):1675–1710, 1996.

[Lig99] T. M. Liggett. *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*. Springer-Verlag, New York, 1999.

[Lig05] T. M. Liggett. *Interacting Particle Systems*. Springer, Berlin, 2005.

[LMSV20] A. Linker, D. Mitsche, B. Schapira, and D. Valesin. The contact process on random hyperbolic graphs: metastability and critical exponents. *arXiv preprint, arXiv:2007.10017*, 2020.

[LP16] R. Lyons and Y. Peres. *Probability on trees and networks*, volume 42 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, New York, 2016.

[LS10] E. Lubetzky and A. Sly. Cutoff phenomena for random walks on random regular graphs. *Duke Math. J.*, 153(3):475–510, 2010.

[LS17] S. Lalley and W. Su. Contact processes on random regular graphs. *Ann. Appl. Probab.*, 27(4):2061–2097, 2017.

[MMVY16] T. Mountford, J.-C. Mourrat, D. Valesin, and Q. Yao. Exponential extinction time of the contact process on finite graphs. *Stochastic Process. Appl.*, 126(7):1974–2013, 2016.

[Mou93] T. S. Mountford. A metastable result for the finite multidimensional contact process. *Canad. Math. Bull.*, 36:216–226, 1993.

[MPV19] T. Mountford, P. L. B. Pantoja, and D. Valesin. The asymmetric multitype contact process. *Stochastic Process. Appl.*, 129(8):2783–2820, 2019.

[MR95] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures Algorithms*, 6(2-3):161–179, 1995.

[MS09] T. Mountford and R. H. Schonmann. The survival of large dimensional threshold contact processes. *Ann. Probab.*, 37(4):1483–1501, 2009.

[MS16] L. Ménard and A. Singh. Percolation by cumulative merging and phase transition for the contact process on some random graphs. *Ann. Sci. École Norm. Supér.*, 49(5):1189–1238, 2016.

[MV16] J.-C. Mourrat and D. Valesin. Phase transition of the contact process on random regular graphs. *Electron. J. Probab.*, 21(31):1–17, 2016.

[MVY13] T. Mountford, D. Valesin, and Q. Yao. Metastable densities for the contact process on power law random graphs. *Electron. J. Probab.*, 18(103):1–36, 2013.

[Nag79] S. V. Nagaev. Large deviations of sums of independent random variables. *Ann. Probab.*, 7(5):745–789, 1979.

[Nam19] D. Nam. A discontinuous phase transition in the threshold-$\theta \geq 2$ contact process on random graphs. *arXiv preprint, arXiv:1907.05005*, 2019.

[NNS21] D. Nam, O. Nguyen, and A. Sly. Critical value asymptotics for the contact process on random graphs. *arXiv:1910.13958, To appear in Trans. Amer. Math. Soc.*, 2021.

[Pem92] R. Pemantle. The contact process on trees. *Ann. Probab.*, 20(4):2089–2116, 1992.

[Pet11] J. Peterson. The contact process on the complete graph with random vertex-dependent infection rates. *Stochastic Process. Appl.*, 121(3):609–629, 2011.

[SS14] A. Sturm and J. M. Swart. Subcritical contact processes seen from a typical infected site. *Electron. J. Probab.*, 19:no. 53, 46, 2014.

[SS19] A. Sakai and G. Slade. Spatial moments for high-dimensional critical contact process, oriented percolation and lattice trees. *Electron. J. Probab.*, 24:Paper No. 65, 18, 2019.

[Sta96] A. M. Stacey. The existence of an intermediate phase for the contact process on trees. *Ann. probab.*, 24(4):1711–1726, 1996.

[Sta01] A. Stacey. The contact process on finite homogeneous trees. *Probab. Theory Relat. Fields*, 121(4):551–576, 2001.

[SV17] B. Schapira and D. Valesin. Extinction time for the contact process on general graphs. *Probab. Theory Related Fields*, 169(3-4):871–899, 2017.

[SV18] B. Schapira and D. Valesin. Exponential rate for the contact process extinction time. *arXiv preprint, arXiv:1806.04491*, 2018.

[Swa18] J. M. Swart. A simple proof of exponential decay of subcritical contact processes. *Probab. Theory Related Fields*, 170(1-2):1–9, 2018.

[Too74] A. L. Toom. Nonergodic multidimensional systems of automata. *Probl. Inf. Transm.*, 10:239–246, 1974.

[VD13] C. Varghese and R. Durrett. Phase transitions in the quadratic contact process on complex networks. *Phys. Rev. E*, 87:062819, 2013.

[vdH17] R. van der Hofstad. *Random Graphs and Complex Networks, Volume 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017.

[XP17] X. Xue and Y. Pan. Contact processes with random recovery rates and edge weights on complete graphs. *J. Stat. Phys.*, 169(5):951–971, 2017.

[Xue15] X. Xue. Contact processes with random vertex weights on oriented lattices. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1):245–259, 2015.

[Xue16] X. Xue. Critical value for the contact process with random recovery rates and edge weights on regular tree. *Physica A: Statistical Mechanics and its Applications*, 462:793–806, 2016.

[Xue18] X. Xue. An improved upper bound for the critical value of the contact process on $\mathbb{Z}^d$ with $d \geq 3$. *Electron. Commun. Probab.*, 23:Paper No. 77, 11, 2018.

[YD20] D. Yao and R. Durrett. Epidemics on evolving graphs. *arXiv preprint, arXiv:2003.08534*, 2020.