SOME ASYMPTOTIC PROBLEMS FOR DYNAMICAL RANDOM GRAPHS

Xuan Wang

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research in the University of North Carolina at Chapel Hill

Chapel Hill
2014

Approved by:

Shankar Bhamidi

Amarjit Budhiraja

Ross Leadbetter

Peter Mucha

Vladas Pipiras

ABSTRACT

XUAN WANG: Some Asymptotic Problems for Dynamical Random Graphs
(Under the direction of Shankar Bhamidi and Amarjit Budhiraja)

This dissertation consists of two parts. In the first part we study the phase transition of a class of dynamical random graph processes, that evolve via the addition of new edges in a manner that incorporates both randomness as well as limited choice. As the density of edges increases, the graphs display a phase transition from the subcritical regime, where all components are small, to the supercritical regime, where a "giant" component emerges. We are interested in the behavior at criticality. First, we consider the simplest model of this kind, namely the Bohman-Frieze process. We show that the stochastic process of component sizes, in the critical window for the Bohman-Frieze process after proper scaling, converges to the standard multiplicative coalescent. Next, we study a more general family of dynamical random graph models, namely, the bounded-size-rule processes. We prove a useful upper bound on the size of the largest component in the barely subcritical regime. We then use this upper bound to study both sizes and surplus of the components of the bounded-size-rule processes in the critical window. In order to describe the joint evolution of sizes and surplus, we introduce the augmented multiplicative coalescent. Our main result shows that the vector of suitably scaled component sizes and surplus converges in distribution to the augmented multiplicative coalescent.

In the second part of this dissertation, we study a large deviation problem related to the configuration model with a given degree distribution. We define a random walk associated with the depth-first-exploration of the random graph constructed from the configuration model. The large deviation principle of this random walk is studied using weak convergence techniques. Some large deviation bounds on the probabilities related to the sizes of the largest component are proved.

ACKNOWLEDGEMENTS

This is a great opportunity to express my gratitude to a number of individuals who have helped me in achieving the task of completing this dissertation.

I owe my deepest gratitude to my advisors, Dr. Amarjit Budhiraja and Dr. Shankar Bhamidi, for their supervision, encouragement, patience, and all their pleasant training and fruitful discussions during my PhD study. In particular, I want to thank Dr. Budhiraja for his great patience in checking over all the technical details of my writing and countless hours of face-to-face instructions, and Dr. Bhamidi for sharing with me all the inspirational stories and tips about how to become a good researcher. Finally, I am also thankful for their financial support, with which I was able to fully concentrate on my study and research.

Also, I would like to acknowledge my committee members. I am especially grateful to Dr. Ross Leadbetter, who has taught me a lot through his wisdom and humor during many courses on various topics in probability theory. I thank Dr. Vladas Pipiras for his encouragement, suggestions, and wonderful jokes both inside and outside the classroom. I am also indebted to Dr. Peter Mucha for his kind help and advice on this dissertation.

I would like to acknowledge Dr. Chuanshu Ji, Dr. Yufeng Liu, Dr. Haipeng Shen and Dr. Jan Hannig for their lectures on the fundamentals in statistics and probability that helped me improve my knowledge and broaden my views. I also thank Dr. Edward Carlstein for his valuable questions and suggestions during my talk on topics related this dissertation.

I am also thankful to the staff members in the STOR Department at UNC-Chapel Hill, especially to Ms. Alison Kieber, for all her generous help and heart-warming smiles.

I am grateful to many friends. Their support and care helped me overcome setbacks and stay focused on my graduate study. In particular, my gratitude especially goes to Dr. Jiang Chen and Mr. Ruoyu Wu, for numerous discussions, strong supports, and wonderful memories at board-game nights. I also thank Mr. Ritwik Chaudhuri for many soul-reaching conversations and his inspiring “laundry machine problem”.

Most importantly, none of this would have been possible without the love and patience of my family. My family has been a constant source of love, care, support and strength all these years. I would like to express my heart-felt gratitude to my family.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

This dissertation consists of two parts. The first part, that forms Chapters 3, 4 and 5, studies the percolation phase transition for a family of dynamical random graph processes. We develop the limit theory for the sizes and complexity of the largest connected components of these models near criticality. In the second part, we study some large deviation problems for the configuration random graph model with a fixed degree distribution.

The percolation phase transition for random graphs refers to the phenomena that, as the density of edges increases, the network transitions from a configuration with many small components to one unique "giant" component. This has been a topic of great interest in many different communities ranging from statistical physics, combinatorics, computer science, social networks and probability theory. The Erdős-Rényi (ER) process is one of the most basic examples. In this model, the graph starts with n isolated vertices. Then edges are added to the graph step by step. At each step, two vertices are chosen uniformly at random among all vertices and an edge is placed between them. Scaling time such that at time t, $\lfloor nt/2 \rfloor$ edges have been added, the classical work of Erdős and Rényi [16] shows that a phase transition occurs at the *critical time* $t_c = 1$. Denoting by $\mathcal{C}_i^{(n)}(t)$ the size (the number of vertices in the component) of the i-th largest component at time t, the phase transition can be described as follows: When $0 < t < t_c$, $\mathcal{C}_1^{(n)}(t) \sim \log n$, and when $t > t_c$, $\mathcal{C}_1^n(t) \sim n$, as $n \to \infty$. The three regimes corresponding to $t < t_c$, $t = t_c$ and $t > t_c$ are called the *subcritical*, *critical* and *supercritical* regimes, respectively. The behavior of the graph near t_c is also of great interest. Aldous [2] shows that for fixed $\lambda \in \mathbb{R}$,

$$\left(\frac{1}{n^{2/3}}\mathcal{C}_i^{(n)}\left(t_c + \frac{\lambda}{n^{1/3}}\right) : i \in \mathbb{N}\right) \xrightarrow{d} \boldsymbol{X}_S(\lambda) \text{ as } n \to \infty, \tag{1.0.1}$$

where $\{\boldsymbol{X}_S(\lambda) : \lambda \in \mathbb{R}\}$ is the standard multiplicative coalescent, which is a stochastic process with values in

$$l^2_\downarrow := \left\{(x_1, x_2, ...) : x_1 \geq x_2 \geq ... \geq 0, \text{ and } \sum_{i=1}^{\infty} x_i^2 < \infty\right\},$$

and $\xrightarrow{d}$ denotes convergence in distribution with respect to the l^2-metric induced on $l^2_\downarrow$. The time interval $\left\{t_c + \lambda/n^{1/3} : \lambda \in B\right\}$ is referred to as the *critical window*, where B is an arbitrary bounded interval.

The first part of this dissertation studies some variants of the ER process, the so-called bounded-size-rule (BSR) processes. In order to introduce the BSR processes, we first introduce a more general family of processes, the Achlioptas processes. Achlioptas processes are random graph processes whose evolution is similar to that of the ER process, except that at each step, two pairs of vertices are picked uniformly at random, and only one pair of vertices is linked with an edge. The decision is based on a rule which only depends on the sizes of the components containing the four chosen vertices. One is led to the study of such models in trying to understand how choice interplays with randomness to delay or accelerate the phase transition. The BSR processes are defined to be Achlioptas processes with an extra constraint: There is some integer K such that all components of size greater than K must receive the same treatment. One example of the BSR processes is the Bohman-Frieze (BF) process, in which an edge is placed between the first pair of vertices if and only if both of them are singletons (isolated vertices). The BF process is a BSR process with $K = 1$, which encourages edge formation between isolated vertices.

Much recent interest has been drawn to a special Achlioptas process, the so-called Product Rule, wherein one connects the edge that minimizes the product of the component sizes on the two end points of the edge. Simulation of the Product Rule ([1]) and related such rules suggest seem to suggest a new phenomenon called "explosive percolation", wherein the phase transition appears much more abruptly, in the sense that the largest component seems to increase from a size smaller than $\sqrt{n}$ to a size larger than $n/2$ in a very small scaling window. For more reference on the explosive percolation, please see [28, 18, 13]. In contrast to the simulation results, recently in [30] it was shown that the phase transition of the Product Rule is actually continuous. However, the phase transition of such models has very different behavior as what one sees in the Erdős-Rényi random graph model. The Product Rule is an Achlioptas process with an "unbounded-size" rule. It is hoped that the techniques of analyzing bounded-size rules in this dissertation can be extended to the regime of unbounded-size rules which we shall attempt to do in the future.

Chapter 3 studies the BF process, which was first introduced in Bohman and Frieze [8]. The paper in particular shows that in the BF model the emergence of the giant component is delayed in comparison with the ER process. The paper [22] in fact shows that the critical time for the BF process is $t_c \equiv t_c(BF) \approx 1.176$. A rigorous proof of the existence of phase transitions for general BSR processes was first given by Spencer and Wormald [31]. In this chapter, we study the evolution of component sizes through the critical window for the BF process. Our precise result is as follows. Denote by $\mathcal{C}_i^{(n)}(t)$ the size of the i-th largest component in the BF process at time t. There exist constants $\alpha \approx 1.063$ and $\beta \approx 0.764$ such that for all fixed $\lambda \in \mathbb{R}$, we have

$$\left(\frac{\beta^{1/3}}{n^{2/3}}\mathcal{C}_i^{(n)}\left(t_c + \frac{\alpha\beta^{2/3}}{n^{1/3}}\lambda\right) : i \geq 1\right) \xrightarrow{d} \boldsymbol{X}_S(\lambda) \text{ as } n \to \infty, \tag{1.0.2}$$

where $\boldsymbol{X}_S(\lambda)$ is the same object as in (1.0.1). In fact we prove process level weak convergence in the space of RCLL (right-continuous-left-limit) functions from $\mathbb{R}$ to $l^2_\downarrow$. A paper [7] based on this chapter (joint work with S. Bhamidi and A. Budhiraja) has appeared in *Random Structures & Algorithms*.

Chapters 4 and 5 study general BSR processes. In Chapter 4, we focus on the subcritical regime for BSR processes. The key result of this chapter is an upper bound (that hold with high probability) of order $n^{2\gamma}\log^4 n$ on the size of the largest component at time $t_c - n^{-\gamma}$ for $\gamma \in (0, 1/4)$. This time scale is also called the *barely subcritical* regime. The proof uses a coupling of BSR processes with a certain family of inhomogeneous random graph models introduced in [11]. This coupling construction also gives an alternative characterization of the critical time for all BSR processes. A paper [6] based on this chapter (joint work with S. Bhamidi and A. Budhiraja) has been accepted in *Combinatorics, Probability & Computing*.

Chapter 5 studies the sizes and surplus (surplus of a component is the number of edges that need to be removed in order to obtain a tree) of the components in general BSR processes in the critical window. Let $\mathcal{C}_i^{(n)}(t)$ and $\xi_i^{(n)}(t)$, respectively, be the size and surplus of the i-th largest component in a BSR process. Our main result shows that there exist some rule-dependent constants $\alpha > 0$ and $\beta > 0$ such that for fixed $\lambda \in \mathbb{R}$, denoting $t_\lambda^{(n)} := t_c + \frac{\alpha\beta^{2/3}}{n^{1/3}}\lambda$, we have

$$\left(\left(\frac{\beta^{1/3}}{n^{2/3}}\mathcal{C}_i^{(n)}(t_\lambda^{(n)}), : i \geq 1\right), (\xi_i^{(n)}(t_\lambda^{(n)}) : i \geq 1)\right) \xrightarrow{d} (\boldsymbol{X}_S(\lambda), \boldsymbol{Y}_S(\lambda)) \text{ as } n \to \infty,$$

where $\boldsymbol{X}_S(\cdot)$ is the same process as in (1.0.1) and (1.0.2), and $\boldsymbol{Y}_S(\cdot)$ is a $\mathbb{N}_0^\infty$-valued stochastic process. The pair process $(\boldsymbol{X}_S, \boldsymbol{Y}_S)$ can be viewed as an extension of Aldous's multiplicative coalescent and we refer to it as the augmented multiplicative coalescent. We show that this process is "nearly Feller" which then allow us to argue that the convergence in (1.0.2) holds jointly at multiple instants λ_1, ..., $\lambda_m \in \mathbb{R}$, for $m \in \mathbb{N}$. This limit theorem can be seen as an universality result that says that all bounded-size-rule processes, with suitable (rule dependent) scaling, are governed asymptotically in the critical window by the augmented multiplicative coalescent. A paper [5] based on this chapter (joint work with S. Bhamidi and A. Budhiraja) has appeared in *Probability Theory and Related Fields*.

The second part (Chapter 6) of this dissertation focuses on some large deviation problems for configuration random graph models. The theory of large deviations is concerned with the asymptotic exponential decay rate of probabilities of rare events. In a typical setting, one is given a sequence of random variables $\{X_n\}$ with values in some Polish metric space (M, d) such that as $n \to \infty$, X_n converges to a non-random limit $x \in M$. The main problem of interest is to obtain the exponential rate of decay of $\mathbb{P}\{d(X_n, x) > \epsilon\}$ for $\epsilon > 0$. A systematic treatment of such an asymptotic study is given by establishing a Large Deviation Principle (LDP) which gives precise exponential decay rates for probabilities of the above form in terms of a rate function.

Chapter 6 studies a large deviation problem related to the sizes of components in a random graph model with fixed degree distribution. It is well known that the asymptotic degree distribution for the Erdős-Rényi random graph process is Poisson. In contrast for many real-world networks, the spread of degrees is often very large and the degree distributions have heavy-tails. In general, one can define a random graph model with a specified degree distribution as follows. Given a degree sequence $\{d_i\}_{i=1}^n$ satisfying $d_i \in \mathbb{N}$ and $\sum_{i=1}^n d_i$ is even, one starts with a vertex set $[n] := \{1, 2, ..., n\}$ with vertex i having d_i half-edges. A uniform random matching is constructed between half-edges to obtain the edges for the graph. Conditioned on the graph being simple (no self-loops or multi-edges), its distribution is uniform over the collection of all simple graphs with vertex set $[n]$ such that vertex i has exactly d_i neighbors. This random graph model is referred to as the configuration model ([9, 26]).

Let $\mathcal{C}_1^{(n)}$ be the size of the largest component in the configuration model on n vertices with a given degree sequence. The goal of Chapter 6 is to argue that the following asymptotic approximation is valid and to characterize the exponent $I(B)$ for $B \subset [0, 1]$:

$$\mathbb{P}\left\{\frac{1}{n}\mathcal{C}_1^{(n)} \in B\right\} \approx e^{-nI(B)}.$$

Towards this goal, we study the large deviation behavior of a random walk associated with a depth-first-exploration of the random graph generated by the configuration model. We use the weak convergence approach developed in [14] for studying large deviation properties of this random walk. The main challenge is that the transition kernel of the random walk is degenerate, and standard conditions that are used in [14] and related works are not satisfied. We give a conjecture on the form of the LDP rate function. Only upper bounds associated with the conjectured rate function are proved rigorously. This is an ongoing work together with S. Bhamidi and A. Budhiraja.

CHAPTER 2: BACKGROUND

2.1 Basic definitions

Define a graph $\mathcal{G}$ to be a pair $(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \neq \emptyset$ is the vertex set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the edge set. Usually the vertex set is taken to be $\mathcal{V} = [n] := \{1, 2, ..., n\}$. For $i, j \in \mathcal{V}$, (i, j) and (j, i) are treated identically in $\mathcal{E}$ and we only consider undirected graphs. $\mathcal{E}$ is treated as a multi-set in the sense that multiple copies of the same pairs are allowed in $\mathcal{E}$. By this definition, multi-edges and self-loops are allowed in a graph. For example, $\mathcal{G} = (\mathcal{V} = \{1, 2\}, \mathcal{E} = \{(1, 1), (1, 2), (1, 2)\}$ denotes a graph with two vertices and three edges.

A component of a graph $\mathcal{G}$ is a maximal connected subgraph of $\mathcal{G}$. Define the size of a component to be the number of vertices in the component. Define the surplus of a component to be the number of edges in the component that need to be removed in order to obtain a tree.

We use $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{d}$ to denote convergence in probability and in distribution respectively. All the unspecified limits are taken as $n \to \infty$. Given a sequence of events $\{E_n\}_{n \geq 1}$, we say E_n occurs with high probability (whp) if $\mathbb{P}\{E_n\} \to 1$.

For a Polish space (complete separable metric space) S, $\mathcal{D}(\mathbb{R} : S)$ (resp. $\mathcal{D}([0, \infty) : S)$) denote the space of right continuous functions with left limits (RCLL) from $\mathbb{R}$ (resp. $[0, \infty)$) to S, equipped with the usual Skorohod topology. Given a metric space S, we denote by $\mathcal{B}(S)$ the Borel σ-field on S and by $\mathrm{BM}(S), C_b(S), \mathcal{P}(S)$, the space of bounded (Borel) measurable functions, continuous and bounded function, and probability measures, on S, respectively.

2.2 Phase transition for the Erdős-Rényi random graph

The Erdős-Rényi random graph $\mathcal{G}(n,t)$ is defined as follows. Let $[n] = \{1, 2, ..., n\}$ be the vertex set for $\mathcal{G}(n,t)$. Then each pair of vertices $i, j \in [n]$, $i \neq j$, are connected with probability $1 - e^{-t/n}$, independently across different pairs. This definition is a variant of the classical Erdős-Rényi random graph, in which the probability of putting an edge is $\max 1, t/n$. The two versions are asymptotically equivalent as $n \to \infty$ in all concerns in this dissertation. Denote by $\mathcal{C}_1^{(n)}(t)$ the size of the largest component in $\mathcal{G}(n,t)$. The phase transition of the Erdős-Rényi random graph can be described as follows:

Subcritical regime: When $t < 1$, there exists some constant $C = C(t) > 0$ such that $\mathcal{C}_1^{(n)}(t) < C \log n$ with high probability.

Supercritical regime: When $t > 1$, there exists some constant $\rho = \rho(t) > 0$ such that $\mathcal{C}_1^{(n)}(t)/n \xrightarrow{\mathbb{P}} \rho$ as $n \to \infty$. In addition, ρ is the unique solution of $1 - \rho = e^{-\rho t}$ in $(0, 1)$.

The following construction gives a natural coupling of $\mathcal{G}(n,t)$ for different t. Define the Erdős-Rényi process $\left\{\mathcal{G}_{\mathbf{ER}}^{(n)}(t) : t \geq 0\right\}$ as follows. Initially, $\mathcal{G}_{\mathbf{ER}}^{(n)}(0)$ is the graph with vertex set $[n]$ and no edges. Consider a Poisson clock with rate $n/2$ (i.e. a Poisson process with rate $n/2$). We add an uniform random edge to the graph whenever the Poisson clock rings. This construction gives a continuous-time random graph process with the addition of edges at rate $n/2$. Note that the number of edges between any given pair of vertices in $\mathcal{G}_{\mathbf{ER}}^{(n)}(t)$ is a Poisson random variable with mean t/n. Thus the probability that there exist at least one edge between two fixed vertices is $1 - e^{-t/n}$. Therefore the distributions of component sizes in $\mathcal{G}(n,t)$ and $\mathcal{G}_{\mathbf{ER}}^{(n)}(t)$ are the same, in particular the phase transition of $\mathcal{G}_{\mathbf{ER}}^{(n)}(t)$ occurs at the critical time $t_c = 1$.

The scaling limit of the sizes of the components in the critical window was proved by Aldous in the seminal paper [2]. Denote by $\mathcal{C}_i^{(n)}(t)$, $i = 1, 2, ...$, the size of the i-th largest component in $\mathcal{G}_{\mathbf{ER}}^{(n)}(t)$. Define the rescaled component sizes vector $\bar{\boldsymbol{C}}_{\mathbf{ER}}^{(n)}(\lambda)$ as

$$\bar{\boldsymbol{C}}_{\mathbf{ER}}^{(n)}(\lambda) := \left(\frac{1}{n^{2/3}}\mathcal{C}_i^{(n)}\left(1 + \frac{\lambda}{n^{1/3}}\right) : i \geq 1\right).$$

Theorem 2.2.1 (Aldous [2]).

$$\bar{\boldsymbol{C}}_{\mathbf{ER}}^{(n)}(\cdot) \xrightarrow{d} \boldsymbol{X}_S(\cdot), \text{ as } n \to \infty,$$

where $\boldsymbol{X}_S(\cdot)$ is the standard multiplicative coalescent (see Section 2.3), which is a Markov process on $l^2_\downarrow = \{(x_1, x_2, \ldots) : x_1 \geq x_2 \geq \cdots \geq 0, \sum_i x_i^2 < \infty\}$. Here $l^2_\downarrow$ is equipped with the l^2-metric, and the weak convergence is in $\mathcal{D}(\mathbb{R} : l^2_\downarrow)$.

2.3 The standard multiplicative coalescent

In this section, we will introduce general *multiplicative coalescent* processes, and then introduce one special version of the process, namely, the *standard multiplicative coalescent*.

The multiplicative coalescent is a continuous-time Markov process on the state space $l^2_\downarrow$ equipped with the l^2-metric $\mathbf{d}(\mathbf{x}, \mathbf{y}) = (\sum_{i=1}^{\infty}(x_i - y_i)^2)^{1/2}$, where $\mathbf{x} = (x_1, x_2, ...)$ and $\mathbf{y} = (y_1, y_2, ...)$.

Dynamics of the multiplicative coalescent can be described as follows. Given $\mathbf{x} = (x_1, x_2, ...) \in l^2_\downarrow$ and $i < j \in \mathbb{N}$, define

$$\mathbf{x}^{ij} := (x_1, ..., x_l, x_i + x_j, x_{l+1}, ..., x_{i-1}, x_{i+1}, ..., x_{j-1}, x_{j+1}, ...),$$

where $x_l \geq x_i + x_j \geq x_{l+1}$. Then the infinitesimal generator for the multiplicative coalescent $\mathcal{A}_{MC}$ can be formally written as

$$(\mathcal{A}_{MC}f)(\mathbf{x}) = \sum_{j>i>0} x_i x_j [f(\mathbf{x}^{ij}) - f(\mathbf{x})], \text{ where } f \in C_b(l^2_\downarrow).$$

The form of $\mathcal{A}_{MC}$ says that, the multiplicative coalescent describes a coalescing dynamics where any two clusters merge at rate proportional to the product of the sizes of the two clusters. The existence of such a stochastic process with path in $\mathcal{D}([0,\infty) : l^2_\downarrow)$ was proved in [2].

The Feller property: Denote by $\{\boldsymbol{X}(\mathbf{x},t) : t \geq 0\}$ the multiplicative coalescent with the initial state $\mathbf{x} \in l^2_\downarrow$. Suppose $\{\mathbf{x}^{(n)} : n \in \mathbb{N}\} \subset l^2_\downarrow$ and $\mathbf{x}^{(n)} \to \mathbf{x}$ as $n \to \infty$, then the paper [2] shows that for any fixed $t > 0$, we have $\boldsymbol{X}(\mathbf{x}^{(n)}, t) \xrightarrow{d} \boldsymbol{X}(\mathbf{x}, t)$, as $n \to \infty$. This convergence says that the multiplicative coalescent is a Feller process. The Feller property plays an important role in proving the existence of the *standard multiplicative coalescent*.

The *standard multiplicative coalescent*, denoted by $\boldsymbol{X}_S(\cdot)$, is a Markov process with sample path in $\mathcal{D}(\mathbb{R} : l^2_\downarrow)$ with the infinitesimal generator $\mathcal{A}_{MC}$, for which the marginal distributions of $\boldsymbol{X}_S(\lambda)$ for fixed $\lambda \in \mathbb{R}$ can be characterized as follows. Define $W_\lambda(t) := W(t) + \lambda t - \frac{t^2}{2}$, where $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion. Let $\hat{W}_\lambda$ be the reflected version of W_λ, i.e.,

$$\hat{W}_\lambda(t) = W_\lambda(t) - \inf_{0\leq s\leq t} W_\lambda(s),\ t \geq 0. \tag{2.3.1}$$

Define an excursion of $\hat{W}_\lambda$ as an interval $(l, u) \subset [0, +\infty)$ such that $\hat{W}_\lambda(l) = \hat{W}_\lambda(u) = 0$ and $\hat{W}_\lambda(t) > 0$ for all $t \in (l, u)$. Define $u - l$ as the size of the excursion. Order the sizes of excursions of $\hat{W}_\lambda$ as $\theta_1(\lambda) > \theta_2(\lambda) > \ldots$ and write $\Xi(\lambda) = (\theta_i(\lambda) : i \geq 1)$. It can be shown that $\Xi(\lambda)$ is a $l^2_\downarrow$-valued random variable. Then the marginal distribution of $\boldsymbol{X}_S(\lambda)$ is the same as the distribution of $\Xi(\lambda)$. The paper [2] proves the existence of a standard multiplicative coalescent.

2.4 Construction of bounded-size-rule processes

Fix $K \in \mathbb{N}$ and let $\Omega_0 = \{\varpi\}$ and $\Omega_K = \{1, 2, \ldots, K, \varpi\}$ for $K \geq 1$, where ϖ represents components of sizes greater than K. Given $F \subset \Omega_K^4$. We now define the F-bounded-size-rule (F-BSR) process $\left\{\mathcal{G}_{BSR}^{(n)}(t) : t \geq 0\right\}$ as follows.

Let $\{\mathcal{P}_{\vec{v}} : \vec{v} \in [n]^4\}$ be i.i.d. Poisson point processes on $[0, \infty)$ with rate $1/2n^3$. Let $\{t_1, t_2, ...\} := \cup_{\vec{v}} \mathcal{P}_{\vec{v}}$ be such that $0 = t_0 < t_1 < t_2 < ....$ For a graph $\mathcal{G}$, let $\mathcal{C}_v(\mathcal{G})$ denote the size of the component in $\mathcal{G}$ containing the vertex v. Then $\mathcal{G}_{BSR}^{(n)}(t)$ is constructed as follows.

- Initially, $\mathcal{G}_{BSR}^{(n)}(t) := (\mathcal{V} = [n], \mathcal{E} = \emptyset)$, for $t \in [t_0, t_1)$.
- Given $\mathcal{G}_{BSR}^{(n)}(t)$ for $t \in [0, t_k)$, $k \geq 1$, suppose $t_k \in \mathcal{P}_{\vec{v}}$ for some $\vec{v} = (v_1, v_2, v_3, v_4)$, then

$$\mathcal{G}_{BSR}^{(n)}(t) := \begin{cases} \mathcal{G}_{BSR}^{(n)}(t_{k-1}) \cup (v_1, v_2), & \text{if } c(\vec{v}) \in F \\ \mathcal{G}_{BSR}^{(n)}(t_{k-1}) \cup (v_3, v_4), & \text{otherwise} \end{cases} \quad \text{for } t \in [t_k, t_{k+1}),$$

  where $c(\vec{v}) := (c(v_i) : i = 1, 2, 3, 4)$, and

$$c(v) := \begin{cases} \mathcal{C}_v(\mathcal{G}_{BSR}^{(n)}(t_{k-1})), & \text{if } \mathcal{C}_v(\mathcal{G}_{BSR}^{(n)}(t_{k-1})) \leq K, \\ \varpi, & \text{if } \mathcal{C}_v(\mathcal{G}_{BSR}^{(n)}(t_{k-1})) > K. \end{cases}$$

The rationale behind this scaling for the rate of the Poisson point process is that the total rate of adding edges is

$$\frac{n^4}{2n^3} = \frac{n}{2}.$$

Thus at time t, $\mathcal{G}_{BSR}^{(n)}(t)$ has approximately the same number of edges as the Erdős-Rényi process $\mathcal{G}_{ER}^{(n)}(t)$ defined in previous sections. The notation in the above construction follows from Spencer and Wormald [31]. They shows that $\mathcal{G}_{BSR}^{(n)}(t)$ displays a similar phase transition as the Erdős-Rényi process at a critical time $t_c \equiv t_c(F)$ which depends on the rule F.

The Erdős-Rényi process and the Bohman-Frieze process are special cases of the F-BSR processes:

- The Erdős-Rényi process: $K = 0$ and $F = \Omega_K^4$.
- The Bohman-Frieze process: $K = 1$ and $F = \{(1, 1, i, j) : i, j \in \Omega_K\}$.

Note that the representations (K, F) of these processes are not unique.

2.5 The configuration model

Given $\{d_i : i = 1, 2, ..., n\}$ satisfying $d_i \in \mathbb{N}$ and $\sum_{i=1}^{n} d_i$ is even, consider the collection of all graphs with the vertex set $[n]$ such that the degree of vertex i is d_i, and let $\mathcal{G}(n, \{d_i\})$ be a random member from this collection. Further assume that there exists a probability distribution on $\mathbb{N}$, $\{p_k : k \in \mathbb{N}\}$, such that $\sum_{k=1}^{\infty} kp_k < \infty$ and for each $k \in \mathbb{N}$,

$$\frac{|\{i \in [n] : d_i = k\}|}{n} \to p_k, \text{ as } n \to \infty.$$

Note that the degree sequence $\{d_i\} = \left\{d_i^{(n)}\right\}$ also depends on n. Then the sequence $\left\{\mathcal{G}(n, \left\{d_i^{(n)}\right\}) : n \in \mathbb{N}\right\}$ is referred to as the configuration model with degree distribution $\{p_k\}$. The phase transition of the configuration model can be described as follows. Define $\nu := \sum_{k=1}^{\infty} k(k-2)p_k \in (-\infty, +\infty]$. Denote by $\mathcal{C}_1^{(n)}$ the size of the largest component in $\mathcal{G}(n, \{d_i\})$.

Subcritical regime: When $\nu < 0$, $\mathcal{C}_1^{(n)}/n \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$.

Supercritical regime: When $\nu > 0$, $\mathcal{C}_1^{(n)}/n \xrightarrow{\mathbb{P}} \rho_0 > 0$ as $n \to \infty$. Here $\rho_0 \in (0, 1]$ can be determined by the following equations:

$$\rho_0 = \sum_{k=1}^{\infty} p_k \rho_1^k \text{ and } \rho_1 = \sum_{k=0}^{\infty} q_k \rho_1^k, \text{ where } q_k = \frac{(k+1)p_{k+1}}{\sum_{k=1}^{\infty} kp_k}.$$

In this dissertation, we will study some large deviation problems related to the random variable $\mathcal{C}_1^{(n)}/n$. See the next section for a brief introduction to large deviation theory.

2.6 Large deviation principle

Let $\{X^{(n)} : n \in \mathbb{N}\}$ be a family of random variables taking values in a Polish space $\mathcal{X}$. The theory of large deviations concerns with the probability of events $\{X^{(n)} \in A\}$ for which $\mathbb{P}\{X^{(n)} \in A\}$ converge to zero exponentially fast as $n \to \infty$. The exponential decay rate of such probabilities is of interest, which is typically expressed in term of the large deviation principle.

The Large Deviation Principle: Let $I : \mathcal{X} \to [0, \infty]$ be such that for each $M < \infty$, the level set $\{x \in \mathcal{X} : I(x) \leq M\}$ is compact. We call $I(\cdot)$ a rate function. The sequence $\{X^{(n)} : n \in \mathbb{N}\}$ is said to satisfy the large deviation principle on $\mathcal{X}$ with rate function I if the following two conditions hold:

1. *Large deviation upper bound.* For each closed set $F \subset \mathcal{X}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{X^{(n)} \in F\} \leq -\inf_{x\in F} I(x).$$

2. *Large deviation lower bound.* For each open set $G \subset \mathcal{X}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\{X^{(n)} \in G\} \geq -\inf_{x\in G} I(x).$$

Formally, the above definition says that if $X^{(n)}$ satisfy the large deviation principle with rate function I, then

$$\mathbb{P}\{X^{(n)} \in A\} \approx \exp\left\{-n \inf_{x\in A} I(x)\right\}.$$

CHAPTER 3: THE BOHMAN-FRIEZE PROCESS

3.1 Introduction

The Bohman-Frieze process $\{\mathcal{G}_{BF}^{(n)}(t)\}_{t\geq 0}$ can be described as follows: Let $\mathcal{G}_{BF}^{(n)}(0)$ be the graph of n isolated vertices. Consider a Poisson clock of rate $n/2$ (i.e. a Poisson process with rate $n/2$), and whenever the clock rings, we pick two candidate edges (e_1, e_2) uniformly among all $\binom{n}{2}$ possible edges and decide which one to add: if e_1 links two isolated vertices, then we add e_1, otherwise we add e_2.

In this chapter, our goal is to establish the scaling limits for the sizes of the largest components of the Bohman-Frieze process in the critical window. In the rest of this chapter $t_c = t_c(\mathbf{BF})$ will denote the critical time for $\mathcal{G}_{BF}^{(n)}(t)$. The main result of this chapter is the following theorem. Recall the definition of the state space $l^2_\downarrow$, the standard multiplicative coalescent $\boldsymbol{X}_S$, and the ordered excursion lengths $\Xi(\lambda)$ from Chapter 2.

Theorem 3.1.1. For some absolute constants $\alpha, \beta > 0$, which will be defined in (3.2.5) and (3.2.6), and for $\lambda \in \mathbb{R}$, let

$$\bar{\boldsymbol{C}}_{BF}^{(n)}(\lambda) = \left(\frac{\beta^{1/3}}{n^{2/3}}\mathcal{C}_i^{(n)}\left(t_c + \beta^{2/3}\alpha\frac{\lambda}{n^{1/3}}\right) : i \geq 1\right) \tag{3.1.1}$$

be the rescaled component sizes of the Bohman-Frieze process in the critical window. Then

$$\bar{\boldsymbol{C}}_{BF}^{(n)}(\cdot) \xrightarrow{d} \boldsymbol{X}_S(\cdot),$$

as $n \to \infty$, where $\xrightarrow{d}$ denotes weak convergence in the space $\mathcal{D}(\mathbb{R} : l^2_\downarrow)$. In particular, for each fixed $\lambda \in \mathbb{R}$, $\bar{\boldsymbol{C}}_{BF}^{(n)}(\lambda)$ converge in distribution (as $l^2_\downarrow$-valued random variables) to $\Xi(\lambda)$.

Organization of this chapter: We begin in Section 3.2 with the construction of the continuous time version of the BF-process and introduce the two constants, $\alpha, \beta > 0$, that show up in the main theorem. Next, in Section 3.3 we give an intuitive sketch of the proof of Theorem 3.1.1 and also provide details on the organization of the various steps in the proof that are carried out from Section 3.4 to Section 3.6. Finally Section 3.7 combines all these ingredients to complete the proof.

3.2 The Bohman-Frieze process

In this section we will define the precise version of the BF process that will be treated in this chapter. We will also introduce some key quantities of interest defined on the BF process, and give the constants α and β in Theorem 3.1.1.

Denote the vertex set by $[n] = \{1, 2, \dots, n\}$ and the edge set by $\mathcal{E}_n = \{\{v_1, v_2\} : v_1 \neq v_2 \in [n]\}$. To simplify notation we shall suppress n in the notation unless required. Denote by $\mathbf{BF}(t) = \mathbf{BF}_n(t)$, $t \in [0, \infty)$, the continuous time Bohman-Frieze random graph process, constructed as follows:

Let $\mathcal{E}^2 = \mathcal{E} \times \mathcal{E}$ be the set of all ordered pairs of edges. For every ordered pair of edges $\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2$ let $\mathcal{P}_{\mathbf{e}}$ be a Poisson process on $[0, \infty)$ with rate $2/n^3$, and let these processes be independent as $\mathbf{e}$ ranges over $\mathcal{E}^2$. We order the points generated by all the $\binom{n}{2} \times \binom{n}{2}$ Poisson processes by their natural order as $0 < t_1 < t_2 < \dots$. Then we can define the BF-process iteratively as follows:
(a) When $t \in [0, t_1)$, $\mathbf{BF}(t) = \mathbf{0}_n$, the empty graph with n vertices;
(b) Consider $t \in [t_k, t_{k+1})$, $k \in \mathbb{N}^+$, where t_k is a point in $\mathcal{P}_{\mathbf{e}}$ and $\mathbf{e} = (e_1, e_2) = (\{v_1, v_2\}, \{v_3, v_4\})$. If v_1, v_2 are both singletons (i.e. not connected to any other vertex) in $\mathbf{BF}(t_k-)$, then $\mathbf{BF}(t) = \mathbf{BF}(t_k-) \cup \{e_1\}$, else let $\mathbf{BF}(t) = \mathbf{BF}(t_k-) \cup \{e_2\}$. Note that multiple edges are allowed between two given vertices, however this has no significance in our analysis which is primarily concerned with component sizes.

Consider the same construction but with the modification that we always add e_1 to the graph and disregard the second edge e_2. Note that the total rate of adding new edges is:

$$\binom{n}{2} \times \binom{n}{2} \frac{2}{n^3} \approx \frac{n}{2}.$$

Then this random graph process is just a continuous time version of the standard Erdős-Rényi (ER) process wherein $t_c = 1$ is the critical time for the emergence of the giant component in this model.

As proved in [31], the Bohman-Frieze model also displays a phase transition and the critical time $t_c \approx 1.1763$. We now summarize some results from [31] that characterize this critical parameter in terms of the behavior of certain differential equations.

The following notations and definitions mostly follow [22]. Let $\mathcal{C}_n^{(i)}(t)$ denote the size of the i^{th} largest component in $\mathbf{BF}_n(t)$, and $\boldsymbol{C}_n(t) = (\mathcal{C}_n^{(i)}(t) : i \geq 1)$ the component size vector. For convenience, we define $\mathcal{C}_n^{(i)}(t) = 0$ whenever $t < 0$.

For fixed time $t \geq 0$, let $X_n(t)$ denote the number of singletons at this time and $\bar{x}(t) = X_n(t)/n$ denote the density of singletons. For simplicity, we have suppressed the dependence on n in the notation. For $k = 2, 3$, let

$$\mathcal{S}_k(t) = \sum_{i \geq 1} (\mathcal{C}_n^{(i)}(t))^k \tag{3.2.1}$$

and let $\bar{s}_k(t) = \mathcal{S}_k(t)/n$. Then from [31], there exist deterministic functions $x(t)$, $s_2(t)$ and $s_3(t)$ such that for each fixed $t \geq 0$:

$$\bar{x}(t) \xrightarrow{\mathbb{P}} x(t), \qquad \bar{s}_k(t) \xrightarrow{\mathbb{P}} s_k(t) \qquad \text{for } k = 2, 3,$$

as $n \to \infty$. The limiting function $x(t)$ is continuous and differentiable for all $t \in \mathbb{R}_+$. For $k \geq 2$, there exists $1 < t_c < \infty$ such that $s_k(t)$ is finite, continuous and differentiable for $0 \leq t < t_c$, and $s_k(t) = \infty$ for $t \geq t_c$. Furthermore, x, s_2, s_3 solve the following differential equations.

$$
\begin{aligned}
&x'(t) = -x^2(t) - (1 - x^2(t))x(t) && \text{for } t \in [0, \infty,) && x(0) = 1 && (3.2.2)\\
&s_2'(t) = x^2(t) + (1 - x^2(t))s_2^2(t) && \text{for } t \in [0, t_c), && s_2(0) = 1 && (3.2.3)\\
&s_3'(t) = 3x^2(t) + 3(1 - x^2(t))s_2(t)s_3(t) && \text{for } t \in [0, t_c), && s_3(0) = 1. && (3.2.4)
\end{aligned}
$$

This constant $t_c = t_c(\mathbf{BF})$ is the critical time such that whp, for $t < t_c$, the size of the largest component in $\mathbf{BF}(t)$ is $O(\log n)$, while for $t > t_c$ there exists a giant component of size $\Theta(n)$ in $\mathbf{BF}(t)$. Furthermore from [22] (Theorem 3.2) there exist constants

$$\alpha = (1 - \bar{x}^2(t_c))^{-1} \approx 1.063 \tag{3.2.5}$$

$$\beta = \lim_{t \uparrow t_c} \frac{s_3(t)}{[s_2(t)]^3} \approx .764 \tag{3.2.6}$$

such that as $t \uparrow t_c$

$$s_2(t) \sim \frac{\alpha}{t_c - t} \tag{3.2.7}$$

$$s_3(t) \sim \beta(s_2(t))^3 \sim \beta\frac{\alpha^3}{(t_c - t)^3}. \tag{3.2.8}$$

The two constant α and β are precisely the constants show up in Theorem 3.1.1.

3.3 Proof idea

Let us now give an idea of the proof. We begin by showing in Proposition 3.3.1 below that, just before the critical window, the configuration of the components satisfies some important regularity properties. This proposition will be used in Section 3.7 in order to apply a result of [2] that gives sufficient conditions for convergence to the multiplicative coalescent.

Proposition 3.3.1. Let $\gamma \in (1/6, 1/5)$ and define $t_n = t_c - n^{-\gamma}$. Then we have

$$\frac{n^2 \mathcal{S}_3(t_n)}{\mathcal{S}_2^3(t_n)} \xrightarrow{\mathbb{P}} \beta \tag{3.3.1}$$

$$\frac{n^{4/3}}{\mathcal{S}_2(t_n)} - \frac{n^{-\gamma+1/3}}{\alpha} \xrightarrow{\mathbb{P}} 0 \tag{3.3.2}$$

$$\frac{n^{2/3} \mathcal{C}_n^{(1)}(t_n)}{\mathcal{S}_2(t_n)} \xrightarrow{\mathbb{P}} 0. \tag{3.3.3}$$

Now note that t_n can be written as

$$t_n = t_c + \beta^{2/3} \alpha \frac{\lambda_n}{n^{1/3}}$$

where

$$\lambda_n = -\frac{n^{-\gamma+1/3}}{\alpha \beta^{2/3}} \to -\infty$$

as $n \to \infty$. The above proposition implies that the configuration of rescaled component sizes, for large n at time "$-\infty$", satisfy the regularity conditions for the standard multiplicative coalescent (see Proposition 4 in [2]).

Once the above has been proved, the second step is to show that through the critical window, the component sizes merge as in the multiplicative coalescent, at rate proportional to the product of the rescaled component sizes. This together with arguments similar to [4] will complete the proof of the main result.

Let us now outline the framework of the proof:

- In Section 3.4 we introduce some more convenient notation that will be used in this chapter.
- The bound on the largest component $\mathcal{C}_n^{(1)}(t)$ when $t \uparrow t_c$ (Proposition 3.5.1) plays a crucial role in proving the statements in Proposition 3.3.1. In order to achieve this, we introduce a series of related models from Section 3.5.1 through Section 3.5.3.

- Section 3.6 uses these models to prove asymptotically tight bounds on the size of the largest component through the subcritical window. The main goal of this Section is to prove Proposition 3.5.1.
- Proposition 3.3.1 can be proved by analyzing the sum of squares and cubes of component sizes near the critical window. The bound on the component sizes in Proposition 3.5.1 plays a key role this this argument. We delay this argument to Chapter 5 in a more general setting for all bounded-size-rule processes.
- Finally, in Section 3.7 we use Proposition 3.3.1 and a coupling with the standard multiplicative coalescent, in a manner similar to [4], to prove the main result.

3.4 Notation

3.4.1 Graphs and random graphs

A graph $\mathbf{G} = \{\mathcal{V}, \mathcal{E}\}$ consists of a vertex set $\mathcal{V}$ and an edge set $\mathcal{E}$, where $\mathcal{V}$ is a subset of some type space $\mathcal{X}$ and $\mathcal{E}$ is a subset of all possible edges $\{\{v_1, v_2\} : v_1 \neq v_2 \in \mathcal{V}\}$. An example of a type space is $[n] = \{1, 2, ..., n\}$. Frequently we will assume $\mathcal{X}$ to have additional structure, for example to be a measure space $(\mathcal{X}, \mathcal{T}, \mu)$. When $\mathcal{V}$ is a finite set, we write $|\mathcal{V}|$ for its cardinality.

$\mathbf{G}$ is called **null graph** if $\mathcal{V} = \emptyset$, and we write $\mathbf{G} = \emptyset$. $\mathbf{G}$ is called an **empty graph** if $|\mathcal{V}| = n$ and $\mathcal{E} = \emptyset$, and we write $\mathbf{G} = \mathbf{0}_n$.

Given two graphs, $\mathbf{G}_i = \{\mathcal{V}_i, \mathcal{E}_i\}$ for $i = 1, 2$, $\mathbf{G}_1$ is said to be a **subgraph** of $\mathbf{G}_2$ if and only if $\mathcal{V}_1 \subset \mathcal{V}_2$ and $\mathcal{E}_1 \subset \mathcal{E}_2$ and we denote this as $\mathbf{G}_1 \leq \mathbf{G}_2$ (or equivalently $\mathbf{G}_2 \geq \mathbf{G}_1$). We write $\mathbf{G}_1 = \mathbf{G}_2$ if $\mathbf{G}_1 \leq \mathbf{G}_2$ and $\mathbf{G}_1 \geq \mathbf{G}_2$.

A connected component $\mathcal{C} = \{\mathcal{V}_0, \mathcal{E}_0\}$ of a graph $\mathbf{G} = \{\mathcal{V}, \mathcal{E}\}$ is a subgraph which is connected (i.e. there is a path between any two vertices in $\mathcal{C}$). The number of vertices in $\mathcal{C}$ will be called the size of the component and frequently we will denote the size and the component by the same symbol.

Let $\mathcal{G}$ be the set of all possible graphs $(\mathcal{V}, \mathcal{E})$ on a given type space $\mathcal{X}$. When $\mathcal{V}$ is countable, we will consider $\mathcal{G}$ to be endowed with the discrete topology and the corresponding Borel sigma field and refer to a random element of $\mathcal{G}$ as a random graph. All random graphs in this chapter are given on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which will usually be suppressed in our proofs.

3.4.2 Probability and analysis

All the unspecified limits are taken as $n \to +\infty$. Given a sequence of events $\{E_n\}_{n\geq 1}$, we say E_n (or E) occurs with high probability (whp) if $\mathbb{P}\{E_n\} \to 1$. For functions $f, g : \mathbb{N} \to \mathbb{R}$, we write $g = O(f)$ if for some $C \in (0, \infty)$, $\limsup g(n)/f(n) < C$ and $g = \Theta(f)$ if $g = O(f)$ and $f = O(g)$. Given two sequences of random variables $\{\xi_n\}$ and $\{\zeta_n\}$, we say $\xi_n = O(\zeta_n)$ whp if there is a $C \in (0, \infty)$ such that $\xi_n < C\zeta_n$ whp, and write $\xi_n = \Theta(\zeta_n)$ whp if there exist $0 < C_1 \leq C_2 < \infty$ such that $C_1\zeta_n < \xi_n < C_2\zeta_n$ whp. Occasionally, when clear from the context, we suppress 'whp' in the statements.

We also use the following little o notation: For a sequence of real numbers $g(n)$, we write $g = o(f)$ if $\limsup |g(n)/f(n)| = 0$. For a sequence of random variables ξ_n, we write "$\xi_n = o_p(f)$" if $\xi_n/f(n)$ converges to 0 in probability.

For a real measurable function ψ on a measure space $(\mathcal{X}, \mathcal{T}, \mu)$, the norms $\|\psi\|_2$ and $\|\psi\|_\infty$ are defined in the usual way. We use $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{d}$ to denote the convergence in probability and in distribution respectively.

We use $=_d$ to denote the equality of random elements in distribution. Suppose that $(S, \mathcal{S})$ is a measurable space and we are given a partial ordering on S. Given two S valued random variables ξ_1, ξ_2, we say a pair of S valued random variables ξ_1^*, ξ_2^*

given on a common probability space define a coupling of (ξ_1, ξ_2) if $\xi_i =_d \xi_i^*$, $i = 1, 2$. We say the S valued random variable ξ_1 **stochastically dominates** ξ_2, and write $\xi_1 \geq_d \xi_2$ if there exists a coupling between the two random variables, say ξ_1^* and ξ_2^*, such that $\xi_1^* \geq \xi_2^*$ a.s.

For two sequences of S valued random elements ξ_n and $\tilde{\xi}_n$, we say "$\xi_n \leq_d \tilde{\xi}_n$ whp." if there exist a coupling between ξ_n and $\tilde{\xi}_n$ for each n (denote as ξ_n^* and $\tilde{\xi}_n^*$) such that $\xi_n^* \leq \tilde{\xi}_n^*$ whp.

Two examples of S that are relevant to this chapter are $\mathcal{D}([0, T] : \mathbb{R})$ and $\mathcal{D}([0, T] : \mathcal{G})$ with the natural associated partial ordering.

3.4.3 Other conventions

We always use n, m, k, i, j to denote non-negative integers unless specified otherwise. We use s, t, T to denote the time parameter for continuous time (stochastic) processes. The scaling parameter is denoted by n. Throughout this chapter $T = 2t_c$ which is a convenient upper bound for the time parameters of interest.

We use $d_1, d_2, ...$ for constants whose specific value are not important. Some of them may appear several times and the values might not be the same. We use $C_1, C_2, ...$ for constants that appear in the statement of theorems.

3.5 An estimate on the largest component

The following estimate on the largest component is the key ingredient in our analysis. Recall that t_c denotes the critical time for the BF process.

Proposition 3.5.1. Let $\gamma \in (0, 1/5)$ and let $I_n(t) \equiv \mathcal{C}_n^{(1)}(t)$ be the largest component of $\mathbf{BF}_n(t)$. Then, for some $B \equiv B(\gamma) \in (0, \infty)$,

$$\mathbb{P}\{I_n(t) \leq m(n, t), \forall t < t_c - n^{-\gamma}\} \to 1, \text{ when } n \to \infty$$

where

$$m(n,t) = B\frac{(\log n)^4}{(t_c - t)^2}. \tag{3.5.1}$$

The proof of Proposition 3.5.1 will be completed in Section 3.6.3. In the current section we will give constructions of some auxiliary random graph processes that are key to our analysis. Although not pursued here, we believe that analogous constructions will be key ingredients in treatment of more general random graph models as well. The section is organized as follows.

- In Section 3.5.1 we will carry out a preliminary analysis of the BF process and identify three deterministic maps a_0, b_0, c_0 from $[0, \infty)$ to $[0, 1]$ that play a fundamental role in our analysis.
- Guided by these deterministic maps, in Section 3.5.2 we will define a random graph process with immigrating vertices and attachments (RGIVA) which is simpler to analyze than, and is suitably 'close' to, the Bohman-Frieze process. A precise estimate on the approximation error introduced through this model is obtained in Section 3.6.3.
- In Section 3.5.3 we will introduce an inhomogeneous random graph (IRG) model associated with a given RGIVA model such that the two have identical component volumes at all times. This allows for certain functional analytic techniques to be used in estimating the maximal component size. We will also make an additional approximation to the IRG model which will facilitate the analysis.
- In Section 3.5.4 we summarize connections between the various models introduced above.

3.5.1 A preliminary analysis of Bohman-Frieze process

Recall that $\mathbf{BF}_n(t)$ denotes the BF process at time t and note that $\mathbf{BF}_n$ defines a stochastic process with sample paths in $\mathcal{D}([0,T]:\mathcal{G})$. Also recall that $\mathcal{C}_n^{(i)}(t)$ denotes the size of the i^{th} largest component in $\mathbf{BF}_n(t)$, $\boldsymbol{C}_n(t) = (\mathcal{C}_n^{(i)}(t) : i \geq 1)$ is the vector of component sizes and $X_n(t)$ denotes the number of singletons in $\mathbf{BF}_n(t)$. We let $\mathcal{F}_t \equiv \mathcal{F}_t^n = \sigma\{\mathbf{BF}_n(s), s \leq t\}$ and refer to it as the natural filtration for the BF process.

At any fixed time $t > 0$, let $\mathcal{COM}(t)$ denote the collection of all non-singleton components

$$\mathcal{COM}(t) = \{\mathcal{C}_n^{(i)}(t) : |\mathcal{C}_n^{(i)}(t)| \geq 2\}.$$

Recall that $\bar{x}(t) = X_n(t)/n$. We will now do an informal calculation of the rate at which an edge $e = \{v_1, v_2\}$ is added to the graph $\mathbf{BF}(t)$. There are three different ways an edge can be added: (i) both v_1 and v_2 are singletons, (ii) only one of them is a singleton, (iii) neither of them is a singleton.

Analysis of the three types of events:

(i) Both v_1 and v_2 are singletons. We will refer to such a component that is formed by connecting two singletons as a **doubleton**. This will happen at rate

$$\frac{2}{n^3}\left[\binom{X_n(t)}{2}\binom{n}{2} + \left(\binom{n}{2} - \binom{X_n(t)}{2}\right)\binom{X_n(t)}{2}\right] \stackrel{def}{=} n \cdot a_n^*(\bar{x}(t)). \tag{3.5.2}$$

The first product in the squared brackets is the count of all possible $\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2$ such that e_1 joins up two singletons and thus will be added to the graph, while the second product is the count of all $\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2$ such that the first edge e_1 does **not** connect two singletons while e_2 connects two singletons and will be added.

Define $a_0 : [0,1] \to [0,1]$ as

$$a_0(y) = 2\left(\frac{y^2}{2}\cdot\frac{1}{2} + \left(\frac{1}{2} - \frac{y^2}{2}\right)\frac{y^2}{2}\right) = \frac{1}{2}(y^2 + (1-y^2)y^2). \tag{3.5.3}$$

It is easy to check that

$$a_n^*(\bar{x}(t)) = a_0(\bar{x}(t)) + r_a(t), \text{ where, } \sup_t |r_a(t)| \le 5/n. \tag{3.5.4}$$

Recall that $x(t)$ is the solution of the differential equation (3.2.2). To simplify notation we will write $a_n^*(\bar{x}(t)) = a^*(\bar{x}) = a^*(t) = a^*$ and $a_0(t) = a_0(x(t))$ exchangeably. Similar conventions will be followed for the functions c_n^*, c_0 and b_n^*, b_0 that will be introduced below. We shall later show that $\sup_{t\le T} |\bar{x}_n(t) - x(t)| \to 0$ in probability (see Lemma 3.6.4, also see [31]). This in particular implies that $\sup_{t\le T} |a_n^*(t) - a_0(t)| \to 0$ in probability.

(ii) Only one of them is a singleton: This will happen if and only if e_1 does not connect two singletons while e_2 connects a singleton and a non-singleton, thus at the rate

$$\frac{2}{n^3}\left(\binom{n}{2} - \binom{X_n(t)}{2}\right)(n - X_n(t))X_n(t). \tag{3.5.5}$$

We are also interested in the rate that a given non-singleton vertex (say, v_0) is connected to any singleton, which is

$$\frac{2}{n^3}\left(\binom{n}{2} - \binom{X_n(t)}{2}\right)X_n(t) \stackrel{def}{=} c_n^*(\bar{x}(t)). \tag{3.5.6}$$

Thus at time t a singleton will be added to $\mathcal{COM}(t)$ during the small time interval $(t, t + dt]$, by attaching to a given vertex $v_0 \in \mathcal{COM}(t)$, with the rate $c^*(t)$.Define $c_0 : [0, 1] \to [0, 1]$ as

$$c_0(y) = (1 - y^2)y, \ y \in [0, 1]. \tag{3.5.7}$$

Then

$$c^*(\bar{x}(t)) = c_0(\bar{x}(t)) + r_c(t) \qquad \text{and } \sup_t |r_c(t)| \le 2/n. \tag{3.5.8}$$

(iii) Neither of them is a singleton: This will happen at the rate

$$\frac{2}{n^3}\left(\binom{n}{2} - \binom{X_n(t)}{2}\right)\binom{n - X_n(t)}{2}. \tag{3.5.9}$$

Also, the event that two fixed non-singleton vertices are connected has the rate

$$\frac{2}{n^3}\left(\binom{n}{2}-\binom{X_n(t)}{2}\right) \stackrel{def}{=} \frac{1}{n} b_n^*(\bar{x}(t)). \tag{3.5.10}$$

Let $b_0 : [0,1] \to [0,1]$ be defined as

$$b_0(y) = 1 - y^2,\ y \in [0,1]. \tag{3.5.11}$$

Then

$$b^*(\bar{x}(t)) = b_0(\bar{x}(t)) + r_b(t) \qquad \text{and } \sup_t |r_b(t)| \leq 2/n. \tag{3.5.12}$$

Note that for the study of the largest component one may restrict attention to the subgraph $\mathcal{COM}(t)$. The evolution of this subgraph is described in terms of stochastic processes $a^*(\bar{x}(t)), b^*(\bar{x}(t))$ and $c^*(\bar{x}(t))$. In the next subsection, we will introduce a random graph process that is "close" to $\mathcal{COM}(t)$ but easier to analyze. Intuitively, we replace $a^*(t), b^*(t), c^*(t)$ with deterministic functions $a(t), b(t), c(t)$ which are close to $a_0(t), b_0(t), c_0(t)$ (and thus, from Lemma 3.6.4, whp close to $a^*(\bar{x}(t)), b^*(\bar{x}(t)), c^*(\bar{x}(t))$) and construct a random graph with similar dynamics as $\mathcal{COM}(t)$.

3.5.2 Immigrating vertices and attachment

In this subsection, we introduce a random graph process with immigrating vertices and attachment (RGIVA). This construction is inspired by [4] where a random graph with immigrating vertices (RGIV) is constructed – we generalize this construction by including attachments. RGIVA process will be governed by three continuous maps a, b, c from $[0,T] \to [0,1]$ (referred to as **rate functions**) and the graph at time t will be denoted by $\mathbf{IA}_n(t) = \mathbf{IA}_n(a,b,c)_t$. When (a,b,c) is sufficiently close to (a_0,b_0,c_0) , the RGIVA model well approximates the BF model in a sense that will be made precise in Section 3.6.3.

The RGIVA process $\mathbf{IA}_n(t) = \mathbf{IA}_n(a,b,c)_t$. Given the rate functions a, b, c, define $\mathbf{IA}_n(t)$ as follows:

(a) $\mathbf{IA}_n(0) = \emptyset$, the null graph;

(b) For $t \in [0,T)$, conditioned on $\mathbf{IA}_n(t)$, during the small time interval $(t, t+dt]$,

- (immigration) a doubleton (consisting of two vertices and a joining edge) will be born at rate $n \cdot a(t)$,
- (attachment) for any given vertex v_0 in $\mathbf{IA}_n(t)$, a new vertex will be created and connected to v_0 at rate $c(t)$,
- (edge) for any given pair of vertices v_1, v_2 in $\mathbf{IA}_n(t)$, an edge will be added between them at rate $\frac{1}{n} \cdot b(t)$.

The events listed above occur independently of each other.

In the special case where $a(t) \equiv b(t) \equiv 1$, $c(t) \equiv 0$, and doubletons are replaced by singletons, the above model reduces to the RGIV model of [4]. We note that the above construction closely follows our analysis of three types of events in Section 3.5.1, replacing stochastic processes $a^*(\bar{x}_n(t)), b^*(\bar{x}_n(t)), c^*(\bar{x}_n(t))$ with deterministic maps $a(t), b(t), c(t)$.

The following lemma establishes a connection between the Bohman-Frieze process and the RGIVA process. Recall the partial order on the space $\mathcal{D}([0,T] : \mathbf{G})$.

Lemma 3.5.2. Let (a_L, b_L, c_L) and (a_U, b_U, c_U) be rate functions. Further, let $U \equiv U_n$ be the event that $\{a^*(t) \le a_U(t), b^*(t) \le b_U(t), c^*(t) \le c_U(t)$ for all $t \in [0,T]\}$ and $L \equiv L_n$ be the event that $\{a^*(t) \ge a_L(t), b^*(t) \ge b_L(t), c^*(t) \ge c_L(t)$ for all $t \in [0,T]\}$. Define for $t \in [0,T]$

$$\mathcal{COM}_n^U(t) = \begin{cases} \emptyset & \text{on } U^C \\ \mathcal{COM}_n(t) & \text{on } U \end{cases} \quad ; \quad \mathcal{COM}_n^L(t) = \begin{cases} \mathbf{IA}_n(a_L, b_L, c_L)_T & \text{on } L^C \\ \mathcal{COM}_n(t) & \text{on } L \end{cases}$$

Then

(i)Upper bound: $\mathcal{COM}_n^U \le_d \mathbf{IA}_n^U \equiv \mathbf{IA}_n(a_U, b_U, c_U)$.

(ii)Lower bound: $\mathcal{COM}_n^L \ge_d \mathbf{IA}_n^L \equiv \mathbf{IA}_n(a_L, b_L, c_L)$.

Proof: We only argue the upper bound. The lower bound is proved similarly. Construct $\mathbf{IA}_n^U(t)$ iteratively on $[0, T]$ as described in the definition, and construct $\mathcal{COM}_n^U(t)$ simultaneously by rejecting the proposed change on the graph with probabilities $(1 - a^*/a_U)^+$, $(1 - b^*/b_U)^+$ and $(1 - c^*/c_U)^+$ according to the three types of the events. Let $\tau = \inf\{0 \le t \le T : a^*(t) > a_U(t) \text{ or } b^*(t) > b_U(t) \text{ or } c^*(t) > c_U(t)\}$ and set $\mathcal{COM}_n^U(t)$ to be the null graph whenever $t \ge \tau$. This construction defines a coupling of $\mathbf{IA}_n^U$ and $\mathcal{COM}_n^U$ such that $\mathcal{COM}_n^U \le \mathbf{IA}_n^U$ a.s. The result follows. ∎

3.5.3 An inhomogeneous random graph with a weight function

In this section we introduce a inhomogeneous random graph (IRG) associated with $\mathbf{IA}_n(a, b, c)$ for given rate functions a, b, c. For a general treatment of IRG models we refer the reader to [11], which our presentation largely follows. We generalize the setting of [11] somewhat by including a weight function and considering the volume of a component instead of the number of vertices of a component. We begin with a description and some basic definitions for a general IRG model.

A **type space** is a measure space $(\mathcal{X}, \mathcal{T}, \mu)$ where $\mathcal{X}$ is a complete separable metric space (i.e. a Polish space), $\mathcal{T}$ is the Borel σ-field and μ is a finite measure.

A **kernel** on the type space $(\mathcal{X}, \mathcal{T}, \mu)$ is a measurable function $\kappa : \mathcal{X} \times \mathcal{X} \to [0, \infty)$. The kernel κ is said to be symmetric if $\kappa(\mathbf{x}, \mathbf{y}) = \kappa(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$. We will also use x, y instead of $\mathbf{x}, \mathbf{y}$ for elements in $\mathcal{X}$ when there is no confusion between an $x \in \mathcal{X}$ and the function $x(t)$ defined in (3.2.2).

A **weight function** ϕ is a measurable, non-negative function on $(\mathcal{X}, \mathcal{T}, \mu)$.

A **basic structure** is a triplet $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$, which consists of a type space, a kernel and a weight function.

The IRG model: Given a type space $(\mathcal{X}, \mathcal{T}, \mu)$, symmetric kernels $\{\kappa_n\}_{n\geq 1}$, and a weight function ϕ, a random graph $\mathbf{RG}_n(\kappa_n)$ ($\equiv \mathbf{RG}_n(\kappa_n, \mu) \equiv \mathbf{RG}_n(\kappa_n, \mu, \phi)$), for any integer $n > 0$, is constructed as follows:
(a) The vertex set $\mathcal{V}$ are the points of a Poisson point process on $(\mathcal{X}, \mathcal{T})$ with intensity $n \cdot \mu$.
(b) Given $\mathcal{V}$, for any two vertices $x, y \in \mathcal{V}$, place an edge between them with probability $\left(\frac{1}{n} \cdot \kappa_n(x, y)\right) \wedge 1$.

One can similarly define an IRG associated with a basic structure $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$, where κ is a symmetric kernel, by letting $\kappa_n = \kappa$ for all n in the above definition.

The weight function ϕ is used in defining the volume of a connected component in the above construction of a random graph. Given a component of $\mathbf{RG}_n(\kappa, \mu, \phi)$ whose vertex set is $\mathcal{V}_0$, define $\sum_{x\in\mathcal{V}_0} \phi(x)$ as the **volume** of the component.

One can associate κ with an integral opertor $\mathcal{K} : L^2(\mu) \to L^2(\mu)$ defined as

$$\mathcal{K}f(x) = \int_{\mathcal{X}} \kappa(x, y) f(y) \mu(dy) \tag{3.5.13}$$

Denote by $\rho = \rho(\kappa)$ the operator norm of $\mathcal{K}$. Then $\rho = \rho(\kappa) = \|\mathcal{K}\| = \sup_{\|f\|_2=1} \|\mathcal{K}f\|_2$.

Given rate functions a, b, c, there is a natural basic structure and the corresponding IRG model associated with $\mathbf{IA}_n(a, b, c)$, which we now describe.

Fix $t \in [0, T]$. Then the following two stage construction describes an equivalent (in law) procedure for obtaining $\mathbf{IA}_n(a, b, c)_t$:

Stage I: Recall that transitions in $\mathbf{IA}_n(a, b, c)$ are caused by three types of events:

immigration, *attachment* (to an existing vertex) and *edge formation* (between existing vertices). Consider the random graph obtained by including all the immigration and attachment events until time t but ignoring the edge formation events. We call the components resulting from this construction as clusters. Note that each cluster consists of exactly one doubleton (which starts the formation of the cluster) and possibly other vertices obtained through later attachments. Note that doubletons immigrate at rate $a(s)$ and supposing that a doubleton is born at time s, the size of the cluster at time $s \le u \le t$ denoted by $w(u)$ evolves according to a integer-valued time-inhomogeneous jump Markov process starting at $w(s) = 2$ and infinitesimal generator $\mathcal{A}(u)$ given as

$$\mathcal{A}(u)f(r) = c(u)r \cdot (f(r+1) - f(r))\,,\ f : \mathbb{N} \to \mathbb{R}, s \le u \le t. \tag{3.5.14}$$

We set $w(u) = 0$ for $0 \le u < s$ and denote this cluster which starts at instant s by (s, w).

Stage II: Given a realization of the random graph of Stage I, we add edges to the graph. Each pair of vertices will be connected during $(s, s + ds]$ with rate $\frac{1}{n}b(s)$. Thus the number of edges between two clusters $\mathbf{x} = (s, w), \mathbf{y} = (r, \tilde{w})$ at time instant t is a Poisson random variable with mean $\frac{1}{n}\int_0^t w(u)\tilde{w}(u)b(u)du$. Consequently,

$$\mathbb{P}\{\mathbf{x} \text{ and } \mathbf{y} \text{ is connected} \mid \text{Stage I}\} = 1 - \exp\{-\frac{1}{n}\int_0^t w(u)\tilde{w}(u)b(u)du\} \tag{3.5.15}$$

$$\le \frac{1}{n}\int_0^t w(u)\tilde{w}(u)b(u)du. \tag{3.5.16}$$

It is easy to see that the graph resulting from this two stage construction has the same distribution as $\mathbf{IA}_n(a, b, c)_t$.

We now introduce an IRG model associated with the above construction in which each cluster is treated as a single point in a suitable type space and the size of the cluster is recorded using an appropriate weight function. Let $\mathcal{X} = [0, T] \times \mathcal{W}$, where $\mathcal{W} = \mathcal{D}([0,T] : \mathbb{N})$ is the Skorohod D-space with the usual Skorohod topology. Denote by $\mathcal{T}$ the Borel sigma field on $[0,T] \times \mathcal{W}$. For future use, we will refer to this particular choice of type space $(\mathcal{X}, \mathcal{T})$ as the *cluster space*. For a fixed time $t \geq 0$, consider a weight function defined as

$$\phi_t(\mathbf{x}) = w(t), \quad \mathbf{x} = (s, w) \in [0,T] \times \mathcal{W}. \tag{3.5.17}$$

Then this weight function associates with each 'cluster' $\mathbf{x}$ its size at time t. We now describe the finite measure μ that governs the intensity of the Poisson point process $\mathcal{P}_t(a,b,c)$ of clusters (regarded as points in $\mathcal{X}$). Denote by ν_s the unique probability measure on the space $\mathcal{W}$ under which, a.s., $w(u) = 0$ for all $u < s$, $w(s) = 2$ and $w(u), u \in [s,T]$ has the probability law of the time inhomogeneous Markov process with generator $\{\mathcal{A}(u), s \leq u \leq T\}$ defined in (3.5.14). Let μ be a finite measure on $\mathcal{X}$ defined as $\mu(dsdw) = \nu_s(dw)a(s)ds$, namely, for a non-negative real measurable function f on $\mathcal{X}$

$$\int_{\mathcal{X}} f(\mathbf{x}) d\mu(\mathbf{x}) = \int_0^T a(s) \left(\int_{\mathcal{W}} f(s,w) d\nu_s(w) \right) ds.$$

We also define for each $t \in [0,T]$, a finite measure μ_t on $\mathcal{X}$ by the relation $\mu_t(A) = \mu(A \cap ([0,t] \times \mathcal{W}))$. Then for f as above,

$$\int_{\mathcal{X}} f(\mathbf{x}) d\mu_t(\mathbf{x}) = \int_0^t a(s) \left(\int_{\mathcal{W}} f(s,w) d\nu_s(w) \right) ds. \tag{3.5.18}$$

The measure μ_t will be the intensity of the Poisson point process on $\mathcal{X}$ which will be used in our construction of the IRG model associated with $\mathbf{IA}_n(a,b,c)_t$. Now we describe the kernel that will govern the edge formation amongst the points. Define

$$\kappa_{n,t}(\mathbf{x}, \mathbf{y}) = \kappa_{n,t}((s,w),(r,\tilde{w})) = n \left(1 - \exp\{-\frac{1}{n} \int_0^t w(u)\tilde{w}(u)b(u)du\} \right). \tag{3.5.19}$$

We will also use the following modification of the kernel $\kappa_{n,t}$.

$$\kappa_t(\mathbf{x}, \mathbf{y}) = \kappa_t((s,w),(r,\tilde{w})) = \int_0^t w(u)\tilde{w}(u)b(u)du. \tag{3.5.20}$$

With the above definitions we can now define IRG models $\mathbf{RG}_n(\kappa_{n,t}, \mu_t, \phi_t)$ and $\mathbf{RG}_n(\kappa_t, \mu_t, \phi_t)$ associated with the type space $(\mathcal{X}, \mathcal{T}, \mu_t)$.

Denote the size of the largest component [resp. the component containing the first immigrating doubleton] in $\mathbf{IA}_n(a,b,c)_t$ by $\mathcal{C}^{(1)}(a,b,c)_t$ [resp. $\mathcal{C}^{(0)}(a,b,c)_t$]. Also, denote the volume of the largest component [resp. the component containing the first cluster] in $\mathbf{RG}_n(\kappa_t, \mu_t, \phi_t)$ by $\mathcal{C}^{(1)}(\kappa_t, \mu_t, \phi_t)$ [resp. $\mathcal{C}^{(0)}(\kappa_t, \mu_t, \phi_t)$]. Then define $\mathcal{C}^{(1)}(\kappa_{n,t}, \mu_t, \phi_t)$, and $\mathcal{C}^{(0)}(\kappa_{n,t}, \mu_t, \phi_t)$ in a similar fashion. The following is an immediate consequence of the above construction.

Lemma 3.5.3. We have

$$(\mathcal{C}^{(1)}(a,b,c)_t, \mathcal{C}^{(0)}(a,b,c)_t) =_d (\mathcal{C}^{(1)}(\kappa_{n,t}, \mu_t, \phi_t), \mathcal{C}^{(0)}(\kappa_{n,t}, \mu_t, \phi_t))$$

and

$$\mathcal{C}^{(1)}(\kappa_{n,t}, \mu_t, \phi_t) \leq_d \mathcal{C}^{(1)}(\kappa_t, \mu_t, \phi_t),\ \mathcal{C}^{(0)}(\kappa_{n,t}, \mu_t, \phi_t) \leq_d \mathcal{C}^{(0)}(\kappa_t, \mu_t, \phi_t).$$

For future use we will write $\mathbf{RG}_n(\kappa_t, \mu_t, \phi_t) \equiv \mathbf{RG}_{n,t}(a,b,c)$.

3.5.4 A summary of the models

As noted earlier, the key step in the proof of Proposition 3.3.1 is a good estimate on the size of the largest component in the Bohman-Frieze process $\mathbf{BF}_n(t)$ as in Proposition 3.5.1. For this we have introduced a series of approximating models. We summarize the relationship between these models below.

- We can decompose the Bohman-Frieze process as $\mathbf{BF}_n = \mathcal{COM}_n \cup X_n$, namely the non-singleton components and singleton components at any time t.
- We shall show that $\mathcal{COM}_n \approx \mathbf{IA}_n(a_0, b_0, c_0)$, where a_0, b_0, c_0 are defined in (3.5.3), (3.5.11), (3.5.7). More precisely we shall show that as $n \to \infty$, for any

fixed $\delta > 0$, we have, whp.

$$\mathbf{IA}_n((a_0-\delta)^+, (b_0-\delta)^+, (c_0-\delta)^+) \leq_d \mathcal{COM}_n$$
$$\leq_d \mathbf{IA}_n((a_0+\delta)\wedge 1, (b_0+\delta)\wedge 1, (c_0+\delta)\wedge 1).$$

This is a consequence of Lemma 3.5.2.

- Given rate functions (a, b, c), for all $t \in [0, T]$,

$$\mathcal{C}^{(i)}(a,b,c)_t =_d \mathcal{C}^{(i)}(\kappa_{n,t}, \mu_t, \phi_t) \leq_d \mathcal{C}^{(i)}(\kappa_t, \mu_t, \phi_t),\ i = 0, 1.$$

  Here $\kappa_{n,t}, \kappa_t, \mu_t, \phi_t$ and a, b, c are related through (3.5.19), (3.5.20), (3.5.18) (see also (3.5.14)), (3.5.17), respectively.

3.6 Analysis of the largest component at sub-criticality

This section proves Proposition 3.5.1. The section is organized as follows:

- In Section 3.6.1 we reduce the problem to proving Proposition 3.6.3. We give the proof of Proposition 3.5.1 using this result. Rest of Section 3.6 is devoted to the proof of Proposition 3.6.3.
- In preparation for this proof, in Section 3.6.2 we present some key lemmas that allow us to estimate the errors between various models summarized in Section 3.5.4. Proofs of Lemmas 3.6.6 will be delayed in Section 3.6.4. Proofs of Lemmas 3.6.9 and 3.6.10 will be omitted since they will be proved in a more general setting for all bounded-size-rule processes in Chapter 4.
- Using these lemmas, in Section 3.6.3 we prove the key proposition, Proposition 3.6.3. The rest of Section 3.6 proves the supporting Lemmas 3.6.6, 3.6.9 and 3.6.10.
- In Section 3.6.4 we introduce a branching process related to the IRG model, and prove Lemma 3.6.6. A key step in the proof is Lemma 3.6.13 whose proof is left to Section 3.6.5.

3.6.1 From the largest component to the first component

In this section we will reduce the problem of proving the estimate on the largest component in Proposition 3.5.1 to an estimate on the *first* component as in Proposition 3.6.3. This reduction, although somewhat different, is inspired by a similar idea used in [4].

Recall that $\mathcal{C}_n^{(1)}(t) \equiv I_n(t)$ denotes the largest component in $\mathbf{BF}_n(t)$. Let $\mathcal{C}_n^s(t)$, $0 \le s \le t$, denote the component whose first doubleton is born at time s in $\mathbf{BF}_n(t)$. In particular $\mathcal{C}_n^s(t) = \emptyset$ if there is no doubleton born at time s. Without loss of generality, we assume that the first doubleton is born at time 0. Then $\mathcal{C}_n^0(t)$ denotes the component of the *first doubleton* at time t of the BF process. The following lemma estimates the size of the largest component $I_n(t)$ in terms of the size of the first component.

Lemma 3.6.1. For any $n \in \mathbb{N}$, $t_0 \in [0,T]$ and deterministic function $\alpha : [0,T] \to [0,\infty)$

$$\mathbb{P}\{I_n(t) > \alpha(t), \text{ for some } t < t_0\} \le nT\mathbb{P}\{\mathcal{C}_n^0(t) > \alpha(t), \text{ for some } t < t_0\}.$$

Proof: Let $\{\mathbf{BF}_n^{(i)}(t), t \ge 0\}_{i \in \mathbb{N}_0}$ be an i.i.d. family of $\{\mathbf{BF}_n(t), t \ge 0\}$ processes on the same vertex set $[n]$. Let N be a rate n Poisson process independent of the above collection. Denote by $\{\tau_i, i \in \mathbb{N}\}$ the jump times of the Poisson process. Set $\tau_0 = 0$. Denote the first component of $\mathbf{BF}_n^{(i)}$ at time t by $\mathcal{J}_n^{(i)}(t)$. Consider the random graph

$$\mathbf{G}_n^t = \cup_{i \in \mathbb{N}_0 : \tau_i \le t} \mathcal{J}_n^{(i)}(t)$$

and let $I_n^{\mathbf{G}}(t)$ denote the size of the largest component in $\mathbf{G}_n^t$. Then since $a_n^*(t) \le 1$

for all t, $I_n \leq_d I_n^{\mathbf{G}}$. Thus

$$\begin{aligned}
&\mathbb{P}\{I_n(t) > \alpha(t), \text{ for some } t < t_0\} \\
\leq &\mathbb{P}\{I_n^{\mathbf{G}}(t) > \alpha(t), \text{ for some } t < t_0\} \\
= &\sum_{k \in \mathbb{N}_0} \mathbb{P}\{I_n^{\mathbf{G}}(t) > \alpha(t), \text{ for some } t < t_0, N(T) = k\} \\
\leq &\sum_{k \in \mathbb{N}_0} \mathbb{P}\{\mathcal{J}_n^{(i)}(t) > \alpha(t), \text{ for some } t < t_0, \text{ for some } i \leq k\}\mathbb{P}\{N(T) = k\} \\
\leq &\sum_{k \in \mathbb{N}_0} k\mathbb{P}\{\mathcal{C}_n^0(t) > \alpha(t), \text{ for some } t < t_0\}\mathbb{P}\{N(T) = k\}.
\end{aligned}$$

The result follows. ■

Next, in the following lemma, we reduce an estimate on the probability of the event $\{\mathcal{C}_n^0(t) > \alpha(t), \text{ for some } t < t_0\}$ to an estimate on $\sup_{t \in [0,t_0]} \alpha(t)\mathbb{P}\{\mathcal{C}_n^0(t) > \alpha(t)\}$.

Lemma 3.6.2. There exists an $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$, $t_0 \in [0, T]$ and continuous $\alpha : [0, T] \to [0, \infty)$

$$\mathbb{P}\{\mathcal{C}_n^0(t) > 2\alpha(t), \text{ for some } 0 < t \leq t_0\} \leq 16nT^2 \sup_{0 \leq s \leq t_0} \left\{\alpha(s)\mathbb{P}\{\mathcal{C}_n^0(s) > \alpha(s)\}\right\}. \tag{3.6.1}$$

Proof: Fix $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$, $\sup_{s \in [0,T]}\{a_n^*(s) \vee b_n^*(s)\} \leq 2$. Consider now $n \geq N_0$. Define $\tau = \inf\{t > 0 : \mathcal{C}_n^0(t) > 2\alpha(t)\}$. Then

$$\mathbb{P}\{\mathcal{C}_n^0(t) > 2\alpha(t) \text{ for some } t \in [0, t_0]\} = \mathbb{P}\{\tau \leq t_0\}. \tag{3.6.2}$$

Denote by $\mathcal{C}_n^0 \leftrightarrow_t \mathcal{C}_n^s$ the event that components $\mathcal{C}_n^0$ and $\mathcal{C}_n^s$ merge at time t. By convention this event is taken to be an empty set if no doubleton is born at time instant s. Then

$$\{\tau = t\} = \{\mathcal{C}_n^0(t-) < 2\alpha(t)\} \cap \{\mathcal{C}_n^0(t-) + \mathcal{C}_n^s(t-) \geq 2\alpha(t); \mathcal{C}_n^0 \leftrightarrow_t \mathcal{C}_n^s, \text{ for some } s < t\}.$$

Next note that

- Since $a^*_n(s) \le 2$, the rate at which doubletons are born can be bounded by $2n$.

- Given a doubleton was born at instant s, the event $\{\mathcal{C}^0_n \leftrightarrow_u \mathcal{C}^s_n,$ for some $u \in (t, t+dt]\}$ occurs, conditionally on $\mathcal{F}_t$, with probability $\frac{1}{n}\mathcal{C}^0_n(t)\mathcal{C}^s_n(t)b^*_n(t)dt$. This probability, using the fact that $b^*_n(s) \le 2$ and $\mathcal{C}^s_n(t) \le n$, on the event $\{\mathcal{C}^0_n(t) < 2\alpha(t)\}$ is bounded by $4\alpha(t)dt$.

- $\mathbb{P}\{\mathcal{C}^0_n(t) + \mathcal{C}^s_n(t) \ge 2\alpha(t)\}$ is bounded by $2\mathbb{P}\{\mathcal{C}^0_n(t) \ge \alpha(t)\}$.

Using these observations we have the following estimate

$$\begin{aligned}
\mathbb{P}\{\tau \le t_0\} &\le \mathbb{E}\int_{[0,t_0]} \mathbf{1}_{\{\mathcal{C}^0_n(t) < 2\alpha(t)\}} \left[\int_{[0,t]} n a^*_n(s) \cdot (\frac{1}{n}\mathcal{C}^0_n(t)\mathcal{C}^s_n(t)) \cdot (b^*_n(t))ds\right] dt \\
&\le \int_{[0,t_0]} \left[\int_{[0,t]} 2n \cdot 2\mathbb{P}(\mathcal{C}^0_n(t) \ge \alpha(t)) \cdot (4\alpha(t))ds\right] dt \\
&\le \int_{[0,t_0]} (2nt) \cdot 2\mathbb{P}(\mathcal{C}^0_n(t) \ge \alpha(t)) \cdot (4\alpha(t))dt \\
&\le 16nT^2 \sup_{t\in[0,t_0]} \left\{\alpha(t)\mathbb{P}\{\mathcal{C}^0_n(t) > \alpha(t)\}\right\}.
\end{aligned}$$

Result follows on combining this estimate with (3.6.2). ■

The following proposition will be proved in Section 3.6.3.

Proposition 3.6.3. Given $\eta \in (0,\infty)$ and $\gamma \in (0, 1/5)$, there exist $B, C, N_1 \in (0,\infty)$ such that for all $n \ge N_1$

$$\mathbb{P}\left\{\mathcal{C}^0_n(t) \ge m(n,t)/2\right\} \le Cn^{-\eta} \text{ for all } 0 < t < t_c - n^{-\gamma}, \tag{3.6.3}$$

where $m(n,t)$ is as defined in (3.5.1).

Remark: Intuitively, one has that in the subcritical regime, i.e. when $t < t_c$, $\mathbb{P}\{\mathcal{C}^0_n(t) > m\} < d_1 e^{-d_2 m}$ for some constants d_1, d_2. This suggests a bound as in (3.6.3) for each fixed $t < t_c$. However, the constants d_1 and d_2 depend on t, and in fact one expects that, $d_2(t) \to 0$ when $t \uparrow t_c$. On the other hand, in order to prove the above proposition one requires estimates that are uniform for all $t < t_c - n^{-\gamma}$ as

$n \to \infty$. This analysis is substantially more delicate as will be seen in subsequent sections.

We now prove Proposition 3.5.1 using the above results.

Proof of Proposition 3.5.1: Fix $\gamma \in (0, 1/5)$ and fix $\eta > 2 + 2\gamma$. Let B, C, N_1 be as determined in Proposition 3.6.3 for this choice of η, γ and let $m(n,t)$ be as defined in (3.5.1). Without loss of generality we can assume that $N_1 \geq N_0$ where N_0 is as in Lemma 3.6.2. Then applying Lemmas 3.6.1 and 3.6.2 with $t_0 = t_c - n^{-\gamma}$ and $\alpha(t) = m(n,t)$, we have

$$\begin{aligned}
&\mathbb{P}(I_n(t) \geq m(n,t), \text{ for some } 0 < t < t_c - n^{-\gamma}\} \\
\leq& nT\mathbb{P}(\mathcal{C}_n^0(t) \geq m(n,t), \text{ for some } 0 < t < t_c - n^{-\gamma}\} \\
\leq& 16n^2T^3 \sup_{s \in [0, t_c - n^{-\gamma}]} \left\{ m(n,s)\mathbb{P}\{\mathcal{C}_n^0(s) \geq m(n.s)/2\} \right\} \\
\leq& 16CBn^{2-\eta+2\gamma}T^3(\log n)^4.
\end{aligned}$$

Since $\eta > 2 + 2\gamma$, the above probability converges to 0 as $n \to \infty$. The result follows. ■

3.6.2 Some preparatory results

This section collects some results that are helpful in estimating the errors between various models described in Section 3.5.4.

The first lemma estimates the error between $\bar{x}_n(t) \equiv \bar{x}(t) = X_n(t)/n$ and its deterministic limit $x(t)$ defined in (3.2.2).

Lemma 3.6.4. For any $T > 0$, there exists a $C(T) \in (0, \infty)$ such that, for all $\gamma_1 \in [0, 1/2)$,

$$\mathbb{P}\{\sup_{0 \leq t \leq T} |\bar{x}_n(t) - x(t)| > \frac{1}{n^{\gamma_1}}\} \leq \exp\{-C(T)n^{1-2\gamma_1}\}.$$

Proof: Recall that $[n] = \{1, 2, \dots, n\}$. Let $E_n = n^{-1}[n]$ and let $E = [0, 1]$. Recall the three types of events described in Section 3.5.1 that lead to edge formation in the BF model. Of these only events of type (i) and (ii) lead to a change in the number of singletons. For the events of type (i), i.e. in the case when a doubleton is created, $\bar{x}$ decreases by $2/n$. Two key functions (see (3.5.2)) for this case are

$$\begin{aligned} f^*_{-2}(y) &= a^*_n(y) \\ f_{-2}(y) &= a_0(y) = \frac{1}{2}\left(y^2 + (1-y^2)y\right). \end{aligned}$$

For the events of type (ii), i.e. in the case when a singleton attaches to a non-singleton component, $\bar{x}$ decreases by $1/n$. Two key functions (see (3.5.5)) for this case are

$$\begin{aligned} f^*_{-1}(y) &= (1-y)c^*_n(y) \\ f_{-1}(y) &= (1-y)c_0(y) = y(1-y^2)(1-y). \end{aligned}$$

Note that $0 \le f^*_l(\bar{x}) \le 1$ for $l = -1, -2$, and that $\bar{x}(t)$ is a Markov process on the state space E_n for which at time t we have the transitions $\bar{x}(t) \rightsquigarrow \bar{x}(t) - 1/n$ at rate $nf^*_{-1}(\bar{x}(t))$ and $\bar{x}(t) \rightsquigarrow \bar{x}(t) - 2/n$ at rate $nf^*_{-2}(\bar{x}(t))$. Furthermore

$$|f^*_{-1}(y) - f_{-1}(y)| \le \frac{2}{n} \qquad |f^*_{-2}(y) - f_{-2}(y)| \le \frac{5}{n}, \text{ for all } y \in [0, 1]. \tag{3.6.4}$$

Let $Y_{-1}(\cdot), Y_{-2}(\cdot)$ be independent rate one Poisson processes. Then the process $\bar{x}(t)$ started with $\bar{x}(0) = 1$ can be constructed (see eg. [24], [17]) as the unique solution of the stochastic equation

$$\bar{x}(t) = 1 - \frac{1}{n}Y_{-1}\left(n\int_0^t f^*_{-1}(\bar{x}(s))ds\right) - \frac{2}{n}Y_{-2}\left(n\int_0^t f^*_{-2}(\bar{x}(s))ds\right). \tag{3.6.5}$$

By Equation (3.2.2), the limiting function $x(\cdot)$ is the unique solution of the integral equation

$$x(t) = 1 - \int_0^t f_{-1}(x(s))ds - \int_0^t 2f_{-2}(x(s))ds. \tag{3.6.6}$$

Also note that $\forall y, z \in E$

$$|(f_{-1}(y) + 2f_{-2}(y)) - (f_{-1}(z) + 2f_{-2}(z))| \leq 6|y - z|. \tag{3.6.7}$$

Using (3.6.6) and (3.6.5) we get

$$|\bar{x}(t) - x(t)| \leq A_1^n(t) + A_2^n(t) + A_3^n(t)$$

where

$$\begin{aligned} A_1^n(t) =& \left| \sum_{l=-1,-2} l \left[\frac{1}{n} Y_l \left(n \int_0^t f_l^*(\bar{x}(s)) ds \right) - \int_0^t f_l^*(\bar{x}(s)) ds \right] \right| \\ \leq& 4 \sup_{l=-1,-2} \sup_{t<T} \left| \frac{Y_l(nt)}{n} - t \right|. \end{aligned}$$

and by (3.6.4)

$$A_2^n(t) = \left| \int_0^t \sum_{l=-1,-2} l \left[f_l^*(\bar{x}(s)) - f_l(\bar{x}(s)) \right] ds \right| \leq \frac{7}{n} T.$$

and finally by (3.6.7)

$$\begin{aligned} A_3^n(t) =& \left| \int_0^s \sum_{l=-1,-2} l \left[f_l(\bar{x}(s)) - f_l(x(s)) \right] ds \right| \\ \leq& \ 6 \int_0^t |\bar{x}(s) - x(s)| ds. \end{aligned}$$

Combining these estimates we get

$$|\bar{x}(t) - x(t)| \leq \left(\frac{7}{n} + 4 \sup_{l=-1,-2} \sup_{t \leq T} \left| \frac{Y_l(nt)}{n} - t \right| \right) + 6 \int_0^t |\bar{x}(s) - x(s)| ds.$$

This implies, by Gronwall's lemma (see e.g. [17], p498)

$$\sup_{s \leq T} |\bar{x}(s) - x(s)| \leq \left(\frac{7}{n} + 4 \sup_{l=-1,-2} \sup_{t \leq T} \left| \frac{Y_l(nt)}{n} - t \right| \right) e^{6T}.$$

Proof is completed using standard large deviations estimates for Poisson processes.

■

In the next lemma we note some basic properties of the integral operator associated with a kernel κ on a finite measure space.

Lemma 3.6.5. Let κ, κ' be kernels given on a finite measure space $(\mathcal{X}, \mathcal{T}, \mu)$. Assume that $\kappa, \kappa' \in L^2(\mu \times \mu)$. Denote the associated integral operators by $\mathcal{K}$ and $\mathcal{K}'$ (see (3.5.13)) and there norms by $\rho(\kappa), \rho(\kappa')$ respectively. Then

(i) $\mathcal{K}$ is a compact operator. In particular

$$\rho(\kappa) = \|\mathcal{K}\| \le \|\kappa\|_2 = \left(\int_{\mathcal{X}\times\mathcal{X}} \kappa^2(x,y)\mu(dx)\mu(dy)\right)^{1/2} < \infty.$$

(ii) If $\kappa \le \kappa'$, then $\rho(\kappa) \le \rho(\kappa')$.

(iii) $\rho(\kappa + \kappa') \le \rho(\kappa) + \rho(\kappa')$ and $\rho(t\kappa) = t\rho(\kappa)$ for $t \ge 0$.

(iv) $|\rho(\kappa) - \rho(\kappa')| \le \rho(|\kappa - \kappa'|)$.

(v) $\rho(\kappa) \le \|\kappa\|_\infty \mu(\mathcal{X})$.

Proof: (i) is a standard result, see Theorem VI.23 of [29].

(ii) For any nonnegative f in $L^2(\mu)$, $\mathcal{K}f(x) \le \mathcal{K}'f(x)$ pointwise. Thus for such f, $\|\mathcal{K}f\|_2 \le \|\mathcal{K}'f\|_2$. Result follows on observing that the suprema of $\|\mathcal{K}f\|_2, \|\mathcal{K}'f\|_2$ over $\{f \in L^2 : \|f\|_2 = 1\}$ is the same as the suprema over $\{f \in L^2 : \|f\|_2 = 1, f \ge 0\}$

.

(iii) This follows immediately from the facts that $\|(\mathcal{K} + \mathcal{K}')f\|_2 \le \|\mathcal{K}f\|_2 + \|\mathcal{K}'f\|_2$ and $\mathcal{K}(tf) = t\mathcal{K}f$.

(iv) Note that $\kappa \le \kappa' + |\kappa - \kappa'|$ and $\kappa' \le \kappa + |\kappa - \kappa'|$. Result follows on combining this observation with (ii) and (iii).

(v) This follows immediately from (i) and the fact that $\|\kappa\|_2 \le \|\kappa\|_\infty \mu(\mathcal{X})$.

■

We now present some auxiliary estimates for the IRG model from Section 3.5.3. The following lemma will be proved in Section 3.6.4. Recall the definition of a *basic structure* from Section 3.5.3.

Lemma 3.6.6. Let $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$ be a basic structure, where κ is symmetric. Suppose that μ is non-atomic and $\rho(\kappa) = \|\mathcal{K}\| < 1$. For fixed $x_0 \in \mathcal{X}$, denote by $\mathcal{C}_n^{RG}(x_0)$ the volume of the component of $\mathbf{RG}_n(\kappa)$ that contains x_0. Define $\mathcal{C}_n^{RG}(x_0) = 0$ if x_0 is not a vertex in $\mathbf{RG}_n(\kappa)$. Then for all $m \in \mathbb{N}$

$$\mathbb{P}\{\mathcal{C}_n^{RG}(x_0) > m\} < 2\exp\{-C_1\Delta^2 m\} \tag{3.6.8}$$

where

$$\Delta = 1 - \rho(\kappa), \quad C_1 = \frac{1}{8\|\phi\|_\infty(1 + 3\|\kappa\|_\infty \mu(\mathcal{X}))}. \tag{3.6.9}$$

The above result will be useful for estimating the size of a given component in $\mathbf{RG}_n(a,b,c)$. One difficulty in directly using this result is that the kernel κ_t and the weight function ϕ_t defined in (3.5.20) and (3.5.17) are not bounded. We will overcome this by using a truncation argument. In order to control the error caused by truncation, the following two results will be useful. For rest of this subsection the type space $(\mathcal{X}, \mathcal{T})$ will be taken to be the *cluster space* introduced above (3.5.17).

Lemma 3.6.7. Given rate functions (a,b,c) and $t \in [0,T]$, let μ_t be the finite measure on $(\mathcal{X}, \mathcal{T})$ defined as in (3.5.18). Let $\mathcal{P}_n$ be a Poisson point process on $(\mathcal{X}, \mathcal{T})$ with intensity $n \cdot \mu_t$. Define

$$Y_n \stackrel{def}{=} \sup_{(s,w) \in \mathcal{P}_n} w(t).$$

Then for every $A \in (0, \infty)$

$$\mathbb{P}\{Y_n > A\} < 2T \cdot n(1 - e^{-T})^{A/2}.$$

Proof: Let N be the number of points in $\mathcal{P}_n$, then N is Poisson with mean $\int_0^t na(s)ds \le nT$. Let $\{Z_2^{(i)}\}_{i\ge 1}$ be independent copies of Z_2 (also independent of N), where Z_2 is a pure jump Markov process on $\mathbb{N}$ with initial condition $Z_2(0) = 2$ and infinitesimal generator $\mathcal{A}_0$ defined as

$$\mathcal{A}_0 f(k) = k(f(k+1) - f(k)), k \in \mathbb{N}, \quad f : \mathbb{N} \to \mathbb{R}.$$

Thus Z_2 is just a Yule process started with two individuals at time zero. Note that

$$Y_n \leq \sup_{(s,w)\in\mathcal{P}_n} w(T) \leq_d \sup_{1\leq i\leq N} Z_2^{(i)}(T),$$

where the first inequality holds a.s and the second inequality uses the fact that $c \leq 1$. Standard facts about the Yule process (see e.g.[27]) imply that $Z_2^{(i)}(T)$ is distributed as sum of two independent Geom$\{e^{-T}\}$. Thus

$$\begin{aligned}\mathbb{P}\{Y_n > A\} &\leq \mathbb{E}(N)\cdot\mathbb{P}\{Z_2(T) > A\}\\ &\leq nT\cdot 2(1-e^{-T})^{A/2}.\end{aligned}$$

This completes the proof of the lemma. ■

The following corollary follows on taking $A = C\log n$ in the above lemma.

Corollary 3.6.8. Let Y_n be as in the above lemma and fix $\eta \in (0,\infty)$. Then there exist $C_1(\eta), C_2(\eta) \in (0,\infty)$ such that for any rate functions (a,b,c)

$$\mathbb{P}\{Y_n > C_1(\eta)\log n\} < C_2(\eta)n^{-\eta}, \quad \text{for all } n \in \mathbb{N}.$$

From Section 3.5.1, recall the definitions of the functions a_0, b_0, c_0 associated with the BF model. The following lemma will allow us to argue that $\mathbf{RG}_n(a_0,b_0,c_0)$ is well approximated by $\mathbf{RG}_n(a,b,c)$ if the rate functions (a,b,c) are sufficiently close to (a_0,b_0,c_0). Let $((\mathcal{X},\mathcal{T},\mu_t),\kappa_t,\phi_t)$ be the basic structure associated with rate functions (a,b,c). Let $\mathcal{K}_t$ be the integral operator defined by (3.5.13), replacing (μ,κ) there by (μ_t,κ_t). Let $\rho_t = \rho(\kappa_t)$. In order to emphasize the dependance on rate functions (a,b,c), we will sometimes write $\rho_t = \rho_t(a,b,c)$. Similar notation will be used for κ_t, μ_t, ϕ_t and $\mathcal{K}_t$.

Lemma 3.6.9. Fix rate functions (a,b,c). Suppose that $\inf_{s\in[0,T]} a(s) > 0$ and for some $\theta \in (0,\infty)$, $c(s) \geq \theta s$, for all $s \in [0,T]$. Given $\delta > 0$ and $t \in [0,T]$, let

$$\rho_{+,t} = \rho_t((a+\delta)\wedge 1, (b+\delta)\wedge 1, (c+\delta)\wedge 1),\ \rho_{-,t} = \rho_t((a-\delta)^+, (b-\delta)^+, (c-\delta)^+).$$

Then there exists $C_2 \in (0, \infty)$ and $\delta_0 \in (0, 1)$ such that for all $\delta \le \delta_0$ and $t \in [0, T]$

$$\max\{|\rho_t - \rho_{+,t}|, |\rho_t - \rho_{-,t}|\} \le C_2(-\log \delta)^3 \delta^{1/2}.$$

The proof of the above lemma is quite technical and is omitted.

The next lemma gives some basic properties of $\rho_t(a_0, b_0, c_0)$. Recall that t_c denotes the critical time for the emergence of the giant component in the BF model.

Lemma 3.6.10. Let $\rho(t) = \rho_t(a_0, b_0, c_0)$. Then:

(i) $\rho(t)$ is strictly increasing in $t \in [0, T]$;

(ii) $\rho(t_c) = 1$;

(iii) $\lim_{s \to 0^+}(\rho(t_c) - \rho(t_c - s))/s = \rho'_-(t_c) > 0$.

The proof of the lemma is also omitted.

3.6.3 Proof of Proposition 3.6.3

This section is devoted to the proof of Proposition 3.6.3. Fix $\eta \in (0, \infty)$ and $\gamma \in (0, 1/5)$.

Step 1: from $\mathbf{BF}_n$ to $\mathbf{IA}_{n,\delta}$

Let $\gamma_1 = 2/5$ and define $E_n = \{\sup_{0 \le t \le T} |\bar{x}_n(t) - x(t)| \le n^{-\gamma_1}\}$.

From Lemma 3.6.4,

$$\mathbb{P}\{E_n^c\} \le \exp\{-C(T)n^{1-2\gamma_1}\} = \exp\{-C(T)n^{1/5}\}. \tag{3.6.10}$$

From (3.5.4) and recalling that the Lipschitz norm of a_0 is bounded by 2 (see (3.5.3)), we have that on E_n

$$|a^*(t) - a_0(t)| \le 5n^{-1} + 2n^{-\gamma_1}, \text{ for all } t \in [0, T].$$

Similar bounds can be shown to hold for b^* and c^*. Thus we can find $n_1 \in \mathbb{N}$ and $d_1 \in (0, \infty)$ such that, for $n \geq n_1$, on E_n

$$a_n^*(t) \leq a_0(t) + \delta_n, b_n^*(t) \leq b_0(t) + \delta_n, c_n^*(t) \leq c_0(t) + \delta_n, \text{ for all } t \in [0, T],$$

where $\delta_n = d_1 n^{-\gamma_1}$. Since a_n^*, b_n^*, c_n^* are all bounded by 1, setting $(a_0(t)+\delta_n) \wedge 1 = a_{n,\delta}$ and similarly defining $b_{n,\delta}, c_{n,\delta}$, we in fact have that

$$a_n^*(t) \leq a_{n,\delta}(t), b_n^*(t) \leq b_{n,\delta}(t), c_n^*(t) \leq c_{n,\delta}(t), \text{ for all } t \in [0, T].$$

Let $\mathcal{C}_{n,\delta}^{IA}(t)$ denote the size of the first component in $\mathbf{IA}_n(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})_t$. From Lemma 3.5.2, we have for any $m \in \mathbb{N}$

$$\mathbb{P}\{\mathcal{C}_n^0(t) > m, E_n\} \leq \mathbb{P}\{\mathcal{C}_{n,\delta}^{IA}(t) > m, E_n\} \leq \mathbb{P}\{\mathcal{C}_{n,\delta}^{IA}(t) > m\}. \tag{3.6.11}$$

Step 2: from $\mathbf{IA}_{n,\delta}$ to $\mathbf{RG}_{n,\delta,A}$

For $t \in [0, T]$, and rate functions $a_{n,\delta}, b_{n,\delta}, c_{n,\delta}$, consider the inhomogeneous random graph model $\mathbf{RG}_n(\kappa_{t,\delta}, \mu_{t,\delta}, \phi_t)$, where $\kappa_{t,\delta} = \kappa_t(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ and $\mu_{t,\delta}$ is the measure for the IRG model corresponding to these rate functions as defined in (3.5.18). Let $A_n = C_1(\eta) \log n$, where $C_1(\eta)$ is as in Corollary 3.6.8. Consider the following truncation of the kernel $\kappa_{t,\delta}$ and weight function $\phi_t(s, w) = w(t)$:

$$\kappa_{t,\delta,A}(\mathbf{x}, \mathbf{y}) = \kappa_{t,\delta}(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\{w(T) \leq A_n\}} \mathbf{1}_{\{\tilde{w}(T) \leq A_n\}}, \ \mathbf{x} = (s, w), \mathbf{y} = (r, \tilde{w})$$

and

$$\phi_{t,A}(s, w) = \phi_t(s, w) \mathbf{1}_{\{w(T) \leq A_n\}}.$$

Then $\|\phi_{t,A}\|_\infty \leq A_n$, $\|\kappa_{t,\delta,A}\|_\infty \leq T A_n^2$.

Recall the Poisson point process $\mathcal{P}_t(a, b, c)$ associated with rate functions (a, b, c), introduced below (3.5.17) and write $\mathcal{P}_{t,\delta} = \mathcal{P}_t(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$. Let

$$Y_{n,\delta} := \sup_{(s,w) \in \mathcal{P}_{t,\delta}} w(T).$$

From Corollary 3.6.8

$$\mathbb{P}\{Y_{n,\delta} > A_n\} < C_2(\eta)n^{-\eta}. \tag{3.6.12}$$

Let $\mathcal{C}^{RG}_{n,\delta}(t) = \mathcal{C}^{RG}_{n,t}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ be the volume of the 'first' component in

$$\mathbf{RG}_{n,t}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta}) \equiv \mathbf{RG}_n(\kappa_{t,\delta}, \mu_{t,\delta}, \phi_t).$$

Then from Lemma 3.5.3

$$\mathbb{P}\{\mathcal{C}^{IA}_{n,\delta}(t) > m\} \leq \mathbb{P}\{\mathcal{C}^{RG}_{n,\delta}(t) > m\}. \tag{3.6.13}$$

Letting $\mathcal{C}^{RG}_{n,\delta,A}(t)$ denote the volume of the first component in $\mathbf{RG}_n(\kappa_{t,\delta,A}, \mu_{t,\delta}, \phi_{t,A})$, namely the random graph formed using the truncated kernel. Then

$$\begin{aligned}
\mathbb{P}\{\mathcal{C}^{RG}_{n,\delta}(t) > m\} &\leq \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}^{RG}_{n,\delta}(t) > m, Y_{n,\delta} \leq A_n\} \\
&= \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}^{RG}_{n,\delta,A}(t) > m, Y_{n,\delta} \leq A_n\} \\
&\leq \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}^{RG}_{n,\delta,A}(t) > m\} \qquad (3.6.14)
\end{aligned}$$

Step 3: Estimating $\mathcal{C}^{RG}_{n,\delta,A}$

We will apply Lemma 3.6.6, replacing $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$ by $\{(\mathcal{X}, \mathcal{T}, \mu_{t,\delta}), \kappa_{t,\delta,A}, \phi_{t,A}\}$, where $t \in (0, t_c - n^{-\gamma})$. From (3.6.8) we have

$$\mathbb{P}\{\mathcal{C}^{RG}_{n,\delta,A}(t) > m\} \leq 2\exp\{-C_1\Delta^2 m\}, \tag{3.6.15}$$

where

$$C_1 = \frac{1}{8\|\phi_{t,A}\|_\infty(1 + 3\|\kappa_{t,\delta,A}\|_\infty \mu_{t,\delta}(\mathcal{X}))},$$

and $\Delta = 1 - \rho(\kappa_{t,\delta,A})$. We now estimate $\rho(\kappa_{t,\delta,A})$. Since $\kappa_{t,\delta,A} \leq \kappa_{t,\delta}$, by (ii) of Lemma 3.6.5, we have $\rho(\kappa_{t,\delta,A}) \leq \rho(\kappa_{t,\delta})$. Note that rate functions (a_0, b_0, c_0) satisfy conditions of Lemma 3.6.9. Thus, recalling that $\delta_n = d_1 n^{-2/5}$, we have from this result, that for some $d_2 \in (0, \infty)$, $\rho(\kappa_{t,\delta}) < \rho(\kappa_t) + d_2(\log n)^3 n^{-1/5}$, for all $t \leq T$. Here $\kappa_t = \kappa_t(a_0, b_0, c_0)$.

Next, by Lemma 3.6.10, there exists $d_3 \in (0,\infty)$ such that $\rho(\kappa_t) < 1 - d_3(t_c - t)$ for all $t \in [0, t_c)$. Combining these estimates, we have for $t < t_c - n^{-\gamma}$,

$$\rho(\kappa_{t,\delta,A}) < 1 - d_3(t_c - t) + d_2(\log n)^3 n^{-1/5}.$$

Recalling that $\gamma \in (0, 1/5)$ we have that, for some $n_2 \in (n_1, \infty)$ and $d_4 \in (0,\infty)$,

$$\rho(\kappa_{t,\delta,A}) \le 1 - d_4(t_c - t), \text{ for all } t \in (0, t_c - n^{-\gamma}) \text{ and } n \ge n_2.$$

Using this estimate in (3.6.15) and recalling that $\|\phi_{t,A}\|_\infty \le A_n$, $\|\kappa_{t,\delta,A}\|_\infty \le TA_n^2$, we have that for some $d_5 \in (0,\infty)$

$$\mathbb{P}\{\mathcal{C}^{RG}_{n,\delta,A}(t) > m\} \le 2\exp\{-\frac{d_5}{(\log n)^3}(t_c - t)^2 m\}, \tag{3.6.16}$$

for all $m \in \mathbb{N}$, $t \in (0, t_c - n^{-\gamma})$ and $n \ge n_2$.

Step 4: Collecting estimates:

Combining (3.6.10), (3.6.11), (3.6.13), (3.6.12), (3.6.14) and (3.6.16), we have

$$\begin{aligned} \mathbb{P}\{\mathcal{C}^0_n(t) > m\} &\le \mathbb{P}\{E^c_n\} + \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}^{RG}_{n,\delta,A}(t) > m\} \\ &\le e^{-C(T)n^{1/5}} + C_2(\eta)n^{-\eta} + 2\exp\{-d_5\frac{(t_c - t)^2}{(\log n)^3}m\}. \end{aligned} \tag{3.6.17}$$

Finally, result follows on replacing m in the above display with $\frac{\eta(\log n)^4}{d_5(t_c - t)^2}$. ■

The following lemma will be used in the proof of Lemma 3.6.10. We will use notation and arguments similar to that in the proof of Proposition 3.6.3 above.

Lemma 3.6.11. Let (a, b, c) be rate functions. Fix $t \in [0, T]$. Let $I^{IA}_n(t)$ denote the largest component in $\mathbf{IA}_n(a, b, c)_t$. Suppose that $\rho_t(a, b, c) < 1$. Then for some $C_0 \in (0,\infty)$

$$\mathbb{P}\{I^{IA}_n(t) > C_0(\log n)^4\} \to 0 \text{ when } n \to \infty.$$

Proof: Let $\mathcal{C}^{IA}_n(t)$ be the first component of $\mathbf{IA}_n(a, b, c)_t$. Then an elementary argument (cf. proof of Lemma 3.6.1) shows that for $m > 0$

$$\mathbb{P}\{I^{IA}_n(t) > m\} \le Tn\mathbb{P}\{\mathcal{C}^{IA}_n(t) > m\}.$$

By an argument as in (3.6.14), we have

$$\mathbb{P}\{\mathcal{C}_n^{IA}(t) > m\} \le \mathbb{P}\{\mathcal{C}_n^{RG}(t) > m\} \le \mathbb{P}\{Y_n > A_n\} + \mathbb{P}\{\mathcal{C}_{n,A}^{RG}(t) > m\},$$

where $\mathcal{C}_n^{RG}$, Y_n and $\mathcal{C}_{n,A}^{RG}$ correspond to $\mathcal{C}_{n,\delta}^{RG}$, $Y_{n,\delta}$ and $\mathcal{C}_{n,\delta,A}^{RG}$ introduced above in the proof of Proposition 3.6.3, with $(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ replaced with (a, b, c). From Corollary 3.6.8 we can find $d_1 \in (0, \infty)$ such that $\mathbb{P}(Y_n \ge d_1 \log n) = O(n^{-2})$. Let $A_n = d_1 \log n$. Then, recalling that $\rho_t(a, b, c) < 1$, we gave by Lemma 3.6.6 that, for some $d_2 \in (0, \infty)$,

$$\mathbb{P}\{\mathcal{C}_{n,A}^{RG}(t) > m\} < 2 \exp\{-d_2 m/(\log n)^3\}.$$

Taking $m = \frac{2}{d_3}(\log n)^4$, we have $\mathbb{P}\{\mathcal{C}_{n,A}^{RG}(t) > m\} = O(n^{-2})$. Combining the above estimates we have $\mathbb{P}\{I_n^{IA}(t) > \frac{2}{d_3}(\log n)^4\} = O(n^{-1})$. The result follows. ■

3.6.4 Proof of Lemma 3.6.6: A branching process construction

The key idea in the proof of Lemma 3.6.6 is the coupling of the breadth first exploration of components in the IRG model with a certain continuous type branching process. This coupling will reduce the problem of establishing the estimate in Lemma 3.6.6 to a similar bound on the total volume of the branching process (Lemma 3.6.13). We refer the reader to [11] where a similar coupling in a setting where the type space $\mathcal{X}$ is finite using a finite-type branching process is constructed. In this subsection we will give the proof of Lemma 3.6.6 using Lemma 3.6.13. Proof of the latter result is given in Section 3.6.5.

Throughout this section we will fix a basic structure $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$, where κ is a symmetric kernel, and a $x_0 \in \mathcal{X}$. Let $\mathbf{RG}_n(\kappa)$ be the IRG constructed using this structure as in Section 3.5.3. We now describe a branching process associated with the above basic structure. The process starts in the 0-th generation with a single vertex of type $x_0 \in \mathcal{X}$ and in the k-th generation, a vertex x will have offspring, independently of the remaining k-th generation vertices, according to a Poisson point process on $\mathcal{X}$ with intensity $\kappa(x,y)\mu(dy)$ to form the $(k+1)^{th}$ generation. We denote this branching process as $\mathbf{BP}(x_0)$.

Denote by $\{\xi_i^{(k)}\}_{i=1}^{N_k} \subset \mathcal{X}$ the k^{th} generation of the branching process. Define the volume of the k-th generation as $G_k = \sum_{i=1}^{N_k} \phi(\xi_i^{(k)})$. The **total volume** of $\mathbf{BP}(x_0)$ is defined as $G = G(x_0) = \sum_{k=0}^{\infty} G_k$.

The following lemma, proved at the end of the section, shows that $\mathcal{C}_n^{RG}(x_0)$ is stochastically dominated by $G(x_0)$.

Lemma 3.6.12. For all $m_1 \in \mathbb{N}$,

$$\mathbb{P}\{\mathcal{C}_n^{RG}(x_0) > m_1\} \leq \mathbb{P}\{G(x_0) > m_1\}.$$

Next lemma, proved in Section 3.6.5 shows that the estimate in Lemma 3.6.6 holds with $\mathcal{C}_n^{RG}(x_0)$ replaced by $G(x_0)$.

Lemma 3.6.13. Suppose that $\rho(\kappa) = \|\mathcal{K}\| < 1$. Then for all $m \in \mathbb{N}$

$$\mathbb{P}\{G > m\} < 2\exp\{-C_1\Delta^2 m\} \tag{3.6.18}$$

where Δ and C_1 are as in (3.6.9).

Using the above lemmas we can now complete the proof of Lemma 3.6.6.

Proof of Lemma 3.6.6: Proof is immediate from Lemmas 3.6.13 and 3.6.12 . ■

We conclude this section with the proof of Lemma 3.6.12.

Proof of Lemma 3.6.12: Without loss of generality assume that $\mathcal{C}_n^{RG}(x_0) \neq 0$. We now explore the component $\mathcal{C}_n^{RG}(x_0)$ in the standard breadth first manner.

Define the sequence of **unexplored sets** $\{U_m\}_{m\geq 0}$ and the set of **removed vertices** $\{R_m\}_{m\geq 0}$ iteratively as follows: Let $R_0 = \emptyset, U_0 = \{x_0\}$ and $y_1 = x_0$. Suppose we have defined R_j, U_j, $j = 0, 1, \cdots, m-1$ and $U_{m-1} = \{y_m, y_{m+1}, \cdots y_{t_m}\}$. Then set

$$R_m = R_{m-1} \cup \{y_m\}$$

$$U_m = U_{m-1} \cup E_m \setminus \{y_m\}$$

where E_m denotes the set

$$\{x \in \mathcal{X} : x \text{ is a neighbor of } y_m \text{ in } \mathbf{RG}_n(\kappa) \text{ and } x \notin R_{m-1} \cup U_{m-1}\}.$$

If $U_{m-1} = \emptyset$ we set $U_j = E_j = \emptyset$ and $R_j = R_{m-1}$ for all $j \geq m$. Thus U_m are the vertices at step m that have been revealed by the exploration but whose neighbors have not been explored yet. Note that the number of vertices in $R_{m-1} \cup U_{m-1}$ equals t_m. Label the vertices in E_m as $y_{t_m+1}, y_{t_m+2}, \ldots y_{t_m+|E_m|}$. With this labeling we have a well defined specification of the sequence $\{R_j, U_j, E_{j+1}\}_{j\in\mathbb{N}_0}$. Note that $\mathcal{C}_n^{RG}(x_0) = m_0$ if and only if $U_{m_0-1} \neq \emptyset, U_{m_0} = \emptyset$ and $|R_{m_0}| = m_0$.

We will now argue that for every $m \in \mathbb{N}$, conditioned on $\{U_{m-1}, R_{m-1}\}$, E_m is a Poisson point process on the space $\mathcal{X}$ with intensity

$$\Lambda_m^*(dx) = \beta_m(x)(\kappa(y_m, x) \wedge n)\mu(dx),$$

where $\beta_m : \mathcal{X} \to [0, 1]$ is given as $\beta_1 \equiv 1$ and, for $m > 1$,

$$\beta_m(x) = \Pi_{y\in R_{m-1}} \left[1 - \left(\frac{\kappa(y, x)}{n} \wedge 1\right)\right], \ x \in \mathcal{X}.$$

Consider first the case $m = 1$. Denote the Poisson point process on $(\mathcal{X}, \mathcal{T})$ used in the construction of $\mathbf{RG}_n(\kappa, \mu)$ by $N_n(\kappa, \mu)$. From the complete independence property of Poisson point processes and the non-atomic assumption on μ, conditioned on the existence of a vertex x_0 in $N_n(\kappa, \mu)$, $N_n(\kappa, \mu) \setminus \{x_0\}$ is once again a Poisson point process with intensity $n \cdot \mu(dx)$ on $\mathcal{X}$. Also, conditioned on $N_n(\kappa, \mu)$, a given type

x vertex in $N_n(\kappa,\mu)$ would be connected to x_0 with probability $(\kappa(x_0,x)/n) \wedge 1$. Thus the neighbors of x_0, namely E_1, define a Poisson point process with intensity $(\kappa(x_0,x) \wedge n)\mu(dx)$. This proves the above statement on E_m with $m = 1$.

Consider now $m > 1$. Since μ is non-atomic and $U_{m-1} \cup R_{m-1}$ consists of only finitely many elements, it follows that conditioned on vertices in $R_{m-1} \cup U_{m-1}$ belonging to $N_n(\kappa,\mu)$, $N_n(\kappa,\mu) \setminus (R_{m-1} \cup U_{m-1})$ is once again a Poisson point process on $\mathcal{X}$ with intensity $n \cdot \mu(dx)$. Note that a vertex $x \in N_n(\kappa,\mu) \setminus (R_{m-1} \cup U_{m-1})$ is in E_m if and only if x is a neighbor of y_m and x is not a neighbor of any vertex in R_{m-1}. So conditioned on $\{R_{m-1}, U_{m-1}\}$, the probability that x is in E_m equals

$$(\kappa(y_m,x)/n \wedge 1) \cdot \Pi_{y \in R_{m-1}}[1 - (\kappa(y,x)/n) \wedge 1].$$

From this and the fact that the edges in $\mathbf{RG}_n(\kappa,\mu)$ are placed in a mutually independent fashion, it follows that the points in E_m, conditioned on $\{R_{m-1}, U_{m-1}\}$, describe a Poisson point process with intensity

$$\begin{aligned}
&n\mu(dx) \cdot (\kappa(y_m,x)/n \wedge 1) \cdot \Pi_{y \in R_{m-1}}[1 - (\kappa(y,x)/n) \wedge 1]\\
&= \Pi_{y \in R_{m-1}}[1 - (\kappa(y,x)/n) \wedge 1] \cdot (\kappa(y_m,x) \wedge n)\mu(dx)\\
&= \Lambda^*_m(dx).
\end{aligned}$$

Thus conditioned on $\{R_{m-1}, U_{m-1}\}$, E_m is a Poisson point process with the claimed intensity.

Next note that one can carry out an analogous breadth first exploration of $\mathbf{BP}(x_0)$. Denoting the corresponding vertex sets once more by

$$\{R_j, U_j, E_{j+1}\}_{j \in \mathbb{N}_0}$$

we see that conditioned on $\{R_{m-1}, U_{m-1}\}$, E_m is a Poisson point process with intensity $\kappa(y_m,x)\mu(dx)$.

As $0 \leq \beta_m(x) \leq 1$ and $\kappa \wedge n \leq \kappa$, we can now construct a coupling between $\mathbf{BP}(x_0)$ and $\mathcal{C}_n^{RG}(x_0)$ by first constructing $\mathbf{BP}(x_0)$ and then by iteratively rejecting each offspring of type x in E_m (and all of its descendants) with probability

$$1 - \frac{\beta_m(x)(\kappa(y_m, x) \wedge n)}{\kappa(y_m, x)}.$$

The lemma is now immediate. ■

3.6.5 Proof of Lemma 3.6.13

Assume throughout this subsection, without loss of generality, that

$$\max\{\|\phi\|_\infty, \|\kappa\|_\infty, \mu(\mathcal{X})\} < \infty.$$

Recall that κ is a symmetric kernel. Define, for $k \in \mathbb{N}$, the kernels $\kappa^{(k)}$ recursively as follows. $\kappa^{(1)} = \kappa$ and for all $k \geq 1$

$$\kappa^{(k+1)}(x, y) = \int_{\mathcal{X}} \kappa^{(k)}(x, u)\kappa(u, y)\mu(du).$$

Recall that $\{\xi_i^{(k)}\}_{i=1}^{N_k}$ denotes the k-th generation of $\mathbf{BP}(x_0)$ and note that it describes a Poisson point process with intensity $\kappa^{(k)}(x_0, y)\mu(dy)$. This observation allows us to compute exponential moments of the form in the lemmas below.

Lemma 3.6.14. Let $g : \mathcal{X} \to \mathbb{R}_+$ be a bounded measurable map. Fix $\delta > 0$ and let $0 < \epsilon < \log(1 + \delta)/\|g\|_\infty$. Then

$$\mathbb{E} \exp\{\epsilon \sum_{i=1}^{N_1} g(\xi_i^{(1)})\} \leq \exp\{\epsilon(1 + \delta)(\mathcal{K}g)(x_0)\}.$$

Proof: Fix δ, ϵ as in the statement of the lemma. By standard formulas for Poisson point processes

$$\begin{aligned}
\mathbb{E} \exp\{\epsilon \sum_{i=1}^{N_1} g(\xi_i^{(1)})\} &= \exp\{\int_{\mathcal{X}} \kappa(x_0, u)(e^{\epsilon g(u)} - 1)\mu(du)\} \\
&\leq \exp\{\int_{\mathcal{X}} \kappa(x_0, u)(1 + \delta)\epsilon g(u)\mu(du)\} \\
&= \exp\{\epsilon(1 + \delta)(\mathcal{K}g)(x_0)\},
\end{aligned}$$

where the middle inequality follows on noting that $e^{\epsilon g(u)} - 1 \leq (1+\delta)\epsilon g(u)$, whenever $\epsilon g(u) < \log(1+\delta)$. ■

Using the above lemma and a recursive argument, we obtain the following result. Recall that $G_k = \sum_{i=1}^{N_k} \phi(\xi_i^{(k)})$ denoted the volume of generation k where volume is measured using the function ϕ.

Lemma 3.6.15. Fix $k \in \mathbb{N}$ and $\delta > 0$. Given a weight function ϕ, define $\phi_0 = \phi + \sum_{i=1}^{k}(1+\delta)^i \mathcal{K}^i \phi$. Then for all $\epsilon \in (0, \frac{\log(1+\delta)}{\|\phi_0\|_\infty})$

$$\mathbb{E}\exp\{\epsilon \sum_{i=0}^{k} G_i\} \leq \exp\{\epsilon[\phi(x_0) + \sum_{i=1}^{k}(1+\delta)^i \mathcal{K}^i \phi(x_0)]\} = \exp\{\epsilon \phi_0(x_0)\}. \quad (3.6.19)$$

Proof: Define $\{\phi_i\}_{i=0}^{k}$ using a backward recursion, as follows. Let $\phi_k = \phi$. For $0 \leq i < k$

$$\phi_i = \phi + (1+\delta)\mathcal{K}\phi_{i+1}.$$

Let $\mathcal{F}_l = \sigma\{\{\xi_i^{(k)}\}_{i=1}^{N_k}, k = 1, \cdots l\}$. We will show recursively, as l goes from k to 0, that

$$\mathbb{E}[\exp\{\epsilon \sum_{i=l}^{k} G_i\} | \mathcal{F}_l] \leq \exp\{\epsilon \sum_{i=1}^{N_l} \phi_l(\xi_i^{(l)})\}\}. \quad (3.6.20)$$

The lemma is then immediate on setting $l = 0$ in the above equation.

When $l = k$, (3.6.20) is in fact an equality, and so (3.6.20) holds trivially for k.

Suppose now that (3.6.20) is true for $l+1$, for some $l \in \{0, 1, \cdots, k-1\}$. Then

$$\begin{aligned}
\mathbb{E}[\exp\{\epsilon \sum_{i=l}^{k} G_i\}|\mathcal{F}_l] &= \exp\{\epsilon G_l\}\mathbb{E}[\mathbb{E}[\exp\{\epsilon \sum_{i=l+1}^{k} G_i\}|\mathcal{F}_{l+1}]|\mathcal{F}_l] \\
&\leq \exp\{\epsilon G_l\}\mathbb{E}[\exp\{\epsilon \sum_{i=1}^{N_{l+1}} \phi_{l+1}(\xi_i^{(l+1)})\}|\mathcal{F}_l] \\
&\leq \exp\{\epsilon G_l\} \exp\{\epsilon(1+\delta) \sum_{i=1}^{N_l} \mathcal{K}\phi_{l+1}(\xi_i^{(l)})\} \\
&= \exp\{\epsilon \sum_{i=1}^{N_l} \phi(\xi_i^{(l)})\} \exp\{\epsilon(1+\delta) \sum_{i=1}^{N_l} \mathcal{K}\phi_{l+1}(\xi_i^{(l)})\} \\
&= \exp\{\epsilon \sum_{i=1}^{N_l} [\phi(\xi_i^{(l)}) + (1+\delta)\mathcal{K}\phi_{l+1}(\xi_i^{(l)})]\} \\
&= \exp\{\epsilon \sum_{i=1}^{N_l} \phi_l(\xi_i^{(l)})\}.
\end{aligned}$$

For the first inequality above we have used the fact that by assumption (3.6.20) holds for $l+1$ and for the second inequality we have applied Lemma 3.6.14 along with the observation that $\epsilon\|\phi_l\|_\infty < \log(1+\delta)$ holds for all $l = 1, 2, ..., k$, since for all l, $\phi_l \leq \phi_0$ and $\epsilon\|\phi_0\|_\infty < \log(1+\delta)$.

This completes the recursion and the result follows. ■

To emphasize that ϕ_0 in the above lemma depends on δ and k, write $\phi_0 = \phi_\delta^{(k)}$. Note that $\phi_\delta^{(k)}$ is increasing in k. Let $\phi_\delta^* = \lim_{k\to\infty} \phi_\delta^{(k)}$. The following corollary follows on sending $k \to \infty$ in (3.6.19).

Corollary 3.6.16. Fix $\delta > 0$ and $\epsilon \in (0, \log(1+\delta)/\|\phi_\delta^*\|_\infty)$. Then

$$\mathbb{E}\{\exp \epsilon G\} \leq \exp\{\epsilon \phi_\delta^*(x_0)\}. \tag{3.6.21}$$

Lemma 3.6.17. For $n \in \mathbb{N}$ and $x \in \mathcal{X}$

$$\mathcal{K}^n\phi(x) \leq \rho^{n-1}\|f_x\|_2\|\phi\|_2, \text{ where } f_x(\cdot) = \kappa(x, \cdot) \text{ and } \rho = \rho(\kappa).$$

Proof: Note that

$$\begin{aligned}\mathcal{K}^n\phi(x) &= \int_{\mathcal{X}} \kappa^{(n)}(x,u)\phi(u)\mu(du) \le \| \int_{\mathcal{X}} f_x(u)\kappa^{(n-1)}(u,\cdot)\mu(du)\|_2\|\phi\|_2 \\ &= \|\mathcal{K}^{n-1}f_x\|_2\|\phi\|_2 \le \rho^{n-1}\|f_x\|_2\|\phi\|_2.\end{aligned}$$

■

Now we can finish the proof of Lemma 3.6.13.

Proof of Lemma 3.6.13: Observing that $\|\phi\|_2 \le \|\phi\|_\infty \mu(\mathcal{X})^{1/2}$ and $\|f_x\|_2 < \|\kappa\|_\infty \mu(\mathcal{X})^{1/2}$, we have for $\delta \in (0,\infty)$ such that $(1+\delta)\rho < 1$, and $x \in \mathcal{X}$

$$\begin{aligned}\phi^*_\delta(x) &= \phi(x) + \sum_{i=1}^{\infty}(1+\delta)^i\mathcal{K}^i\phi(x) \\ &\le \|\phi\|_\infty + \|f_x\|_2\|\phi\|_2(\sum_{i=1}^{\infty}(1+\delta)^i\rho^{i-1}) \\ &\le \|\phi\|_\infty + \|\kappa\|_\infty\|\phi\|_\infty\mu(\mathcal{X})\frac{(1+\delta)}{1-(1+\delta)\rho},\end{aligned}$$

where the first inequality above follows from Lemma 3.6.17. Setting $\delta = \frac{\Delta}{2}$, we see

$$(1+\delta)\rho = (1+\Delta/2)(1-\Delta) < 1-\Delta/2.$$

Using this and that $\Delta < 1$, we have

$$\phi^*_\delta(x) \le \|\phi\|_\infty\left(1 + \frac{3\|\kappa\|_\infty\mu(\mathcal{X})}{\Delta}\right) \equiv d_1.$$

Let $\epsilon = \log(1+\delta)/(2d_1)$. Clearly $\epsilon \in (0, \log(1+\delta)/\|\phi^*_\delta\|_\infty)$. Using Corollary 3.6.16 we now have that

$$\begin{aligned}\mathbb{P}\{G > m\} &\le \exp\{-\epsilon m\}\exp\{\epsilon\phi^*_\delta(x_0)\} \\ &\le \exp\{-\epsilon m\}\exp\{\frac{\log(1+\delta)}{2}\} \\ &\le 2\exp\{-\frac{\log(1+\delta)}{2d_1}m\}.\end{aligned}$$

Finally, noting that $\log(1+\delta) \ge \frac{\delta}{2}$, we have

$$\frac{\log(1+\delta)}{2d_1} \ge \frac{\Delta^2}{8\|\phi\|_\infty(1+3\|\kappa\|_\infty\mu(\mathcal{X}))}.$$

The result follows. ■

3.7 Proof of Theorem 3.1.1

We will now complete the proof of Theorem 3.1.1. As always, we write the component sizes as

$$\boldsymbol{C}_n^{BF}(t) \equiv (\mathcal{C}_n^{(i)}(t) : i \geq 1) \equiv (\mathcal{C}_i(t) : i \geq 1);$$

and write the scaled component sizes as

$$\bar{\boldsymbol{C}}_n^{BF}(\lambda) \equiv \left(\frac{\beta^{1/3}}{n^{2/3}}\mathcal{C}_i^{(n)}\left(t_c + \beta^{2/3}\alpha\frac{\lambda}{n^{1/3}}\right) : i \geq 1\right) \equiv \left(\bar{\mathcal{C}}_i(\lambda) : i \geq 1\right) \tag{3.7.1}$$

Then Proposition 3.3.1 proves that with

$$\lambda_n = -\frac{n^{-\gamma+1/3}}{\alpha\beta^{2/3}}$$

and $\gamma \in (1/6, 1/5)$ we have, as $n \to \infty$,

$$\frac{\sum_i \left(\bar{\mathcal{C}}_i(\lambda_n)\right)^3}{\left[\sum_i \left(\bar{\mathcal{C}}_i(\lambda_n)\right)^2\right]^3} \xrightarrow{\mathbb{P}} 1, \quad \frac{1}{\sum_i \left(\bar{\mathcal{C}}_i(\lambda_n)\right)^2} + \lambda_n \xrightarrow{\mathbb{P}} 0, \quad \frac{\bar{\mathcal{C}}_1(\lambda_n)}{\sum_i \left(\bar{\mathcal{C}}_i(\lambda_n)\right)^2} \xrightarrow{\mathbb{P}} 0. \tag{3.7.2}$$

We shall now give an idea of the proof of the main result, and postpone precise arguments to the next two sections. The first step is to observe that the asymptotics in (3.7.2) imply that the $\bar{\boldsymbol{C}}_{bf}$ process at time λ_n satisfies the regularity conditions of Proposition 4 of [2]. The second key observation is that the scaled components merge in the critical window at a rate close to that for the multiplicative coalescent. Indeed, note that for any given time t components $i < j \in \mathbf{BF}(t)$ merge in a small time interval $[t, t+dt)$ at rate

$$\frac{1}{n}(1 - \bar{x}^2(t))\mathcal{C}_i(t)\mathcal{C}_j(t).$$

Thus letting $\lambda = (t - t_c)n^{1/3}/(\alpha\beta^{2/3})$ be the scaled time parameter, in the time interval $[\lambda, \lambda + d\lambda)$, these two components merge at rate

$$\begin{aligned}
\gamma_{ij}(\lambda) &= \frac{(1 - \bar{x}^2(t_c + \beta^{2/3}\alpha\frac{\lambda}{n^{1/3}}))}{n}\frac{\beta^{2/3}\alpha}{n^{1/3}}\mathcal{C}_i\left(t_c + \frac{\beta^{2/3}\alpha\lambda}{n^{1/3}}\right)\mathcal{C}_j\left(t_c + \frac{\beta^{2/3}\alpha\lambda}{n^{1/3}}\right) \\
&= \alpha\left(1 - \bar{x}^2\left(t_c + \beta^{2/3}\alpha\frac{\lambda}{n^{1/3}}\right)\right)\bar{\mathcal{C}}_i(\lambda)\bar{\mathcal{C}}_j(\lambda).
\end{aligned}$$

Now since, for large n,

$$\bar{x}^2\left(t_c + \beta^{2/3}\alpha\frac{\lambda}{n^{1/3}}\right) \approx x^2(t_c)$$

and from [22], $\alpha(1 - x^2(t_c)) = 1$ (see (3.2.5)) we get

$$\gamma_{ij}(\lambda) \approx \bar{\mathcal{C}}_i(\lambda)\bar{\mathcal{C}}_j(\lambda)$$

which is exactly the rate of merger for the multiplicative coalescent. The above two facts allow us to complete the proof using ideas similar to those in [4]. Let us now make these statements precise.

As before, throughout this section $t_n = t_c - n^{-\gamma} = t_c + \beta^{2/3}\alpha\frac{\lambda_n}{n^{1/3}}$, where γ is fixed in $(1/6, 1/5)$. We will first show that $\bar{\boldsymbol{C}}_n^{BF}(\lambda) \xrightarrow{d} \boldsymbol{X}(\lambda)$ in $l^2_\downarrow$ for each $\lambda \in \mathbb{R}$ and at the end of the section show that, in fact, $\bar{\boldsymbol{C}}_n^{BF} \xrightarrow{d} \boldsymbol{X}$ in $\mathcal{D}((-\infty, \infty) : l^2_\downarrow)$. Now fix $\lambda \in \mathbb{R}$. By choosing n large enough we can ensure that $\lambda \geq \lambda_n$. Henceforth consider only such n. Recall that $\mathcal{COM}_n(t)$ denotes the subgraph of $\mathbf{BF}_n(t)$ obtained by deleting all the singletons. Let $\sum_{i\in\mathcal{COM}}$ denote the summation over all components in $\mathcal{COM}_n$, and $\sum_i$ denote the summation over all components in $\mathbf{BF}_n$. Since

$$\sum_i \left(\bar{\mathcal{C}}_i(\lambda)\right)^2 - \sum_{i\in\mathcal{COM}} \left(\bar{\mathcal{C}}_i(\lambda)\right)^2 \leq \frac{d_1}{n^{4/3}} \sum_{i=1}^{X_n(t)} 1 = O(1/n^{1/3}), \tag{3.7.3}$$

it suffices to prove Theorem 3.1.1 and verify Proposition 3.3.1 with $\mathbf{BF}_n(t)$ replaced by $\mathcal{COM}_n(t)$. We write $\sum_i$ instead of $\sum_{i\in\mathcal{COM}}$ for simplicity of the notation from now on. We begin in Section 3.7.1 by constructing a coupling of $\{\mathcal{COM}_n(t)\}_{t\geq t_n}$ with two other random graph processes, sandwiching our process between these two processed, and proving statements analogous to those in Theorem 3.1.1 for scaled component vectors associated with these processes. Proof of Theorem 3.1.1 will then be completed in Section 3.7.2.

3.7.1 Coupling with the multiplicative coalescent

Lower bound coupling: Let, for $t \geq t_n$, $\mathcal{COM}_n^-(t)$ be a modification of $\mathcal{COM}_n(t)$ such that $\mathcal{COM}_n^-(t_n) = \mathcal{COM}_n(t_n)$, and when $t > t_n$, we change the dynamics of the random graph to the Erdős-Rényi type. More precisely, recall from Section 3.5.1 that a jump in $\mathbf{BF}_n(t)$ can be produced by three different kinds of events. These are described in items (i), (ii) and (iii) in Section 3.5.1. $\mathcal{COM}_n^-(t)$, $t \geq t_n$ is constructed from $\mathcal{COM}_n^-(t_n)$ by erasing events of type (i) and (ii) (i.e. immigrating doubletons and attaching singletons) and changing the probability of edge formation between two non-singletons (from that given in (3.5.10)) to the fixed value $b_n^*(t_n)/n$. Since $b_n^*(t)$ is nondecreasing in t, we have that $\mathcal{COM}_n(t_n + \cdot) \geq_d \mathcal{COM}_n^-(t_n + \cdot)$. Denote by $\bar{\boldsymbol{C}}_n^-(\lambda) = \left(\bar{\mathcal{C}}_i^-(\lambda) : i \geq 1\right)$ the scaled (as in (3.7.1)) component size vector for $\mathcal{COM}_n^-(t)$. From Proposition 4 of [2], it follows that for any $\lambda \in \mathbb{R}$,

$$\bar{\boldsymbol{C}}_n^-(\lambda) \xrightarrow{d} \boldsymbol{X}(\lambda) \tag{3.7.4}$$

in $l^2_\downarrow$. Indeed, note that the first and third convergence statements in (3.7.2) hold with $\bar{\mathcal{C}}_i$ replaced with $\bar{\mathcal{C}}_i^-$ since the contributions made by singletons to the scaled sum of squares is $O(n^{-1/3})$ (see (3.7.3)) and to the sum of cubes is even smaller. This shows that the first and third requirements in Proposition 4 of [2] (see equations (8), (10) therein) are met. To show the second requirement in Proposition 4 of [2], using the second convergence in (3.7.2),

$$\begin{aligned} & \lim_{n\to\infty} \left(\left(n^{2/3}\beta^{-1/3}\right)^2 \frac{b_n^*(t_n)}{n} \frac{\beta^{2/3}\alpha(\lambda - \lambda_n)}{n^{1/3}} - \frac{1}{\sum_i \left(\bar{\mathcal{C}}_i^-(\lambda_n)\right)^2} \right) \qquad (3.7.5) \\ = \; & \lim_{n\to\infty} \alpha b_n^*(t_n)\lambda - \lambda_n(\alpha b_n^*(t_n) - 1) \\ = \; & \lambda - \lim_{n\to\infty} \lambda_n(\alpha b_n^*(t_n) - 1), \end{aligned}$$

where the last equality follows on observing that, as $n \to \infty$, $b_n^*(t_n) \xrightarrow{\mathbb{P}} 1 - x^2(t_c)$ and $\alpha(1 - x^2(t_c)) = 1$. Also,

$$
\begin{aligned}
\lim_{n\to\infty} \lambda_n |\alpha b_n^*(t_n) - 1| &= \lim_{n\to\infty} \frac{n^{-\gamma+1/3}}{\beta^{2/3}} |b_n^*(t_n) - \alpha^{-1}| \\
&= \lim_{n\to\infty} \frac{n^{-\gamma+1/3}}{\beta^{2/3}} |b_0(\bar{x}(t_n)) - b_0(x(t_c))| \\
&\le d_1 \lim_{n\to\infty} n^{-\gamma+1/3} |\bar{x}(t_n) - x(t_c)| \\
&\le \lim_{n\to\infty} d_2 \left(n^{-\gamma+1/3} |\bar{x}(t_n) - x(t_n)| + n^{-\gamma+1/3} |t_n - t_c| \right),
\end{aligned}
$$

where the second equality follows from (3.5.12). The first term on the last line converges to 0 using Lemma 3.6.4. For the second term note that $n^{-\gamma+1/3}|t_n - t_c| = n^{-\gamma+1/3} n^{-\gamma}$ which converges to 0 since $\gamma > 1/6$. Thus we have shown that the expression in (3.7.5) converges to λ as $n \to \infty$ and therefore the second requirement in Proposition 4 of [2] (see equation 9 therein) is met as well. This proves that $\bar{\boldsymbol{C}}_n^-(\lambda) \xrightarrow{d} \boldsymbol{X}(\lambda)$ in $l_\downarrow^2$, for every $\lambda \in \mathbb{R}$. Although Proposition 4 of [2] only proves convergence at any fixed point λ, from the Feller property of the multiplicative coalescent process proved in Proposition 6 of the same paper it now follows that, in fact, $\bar{\boldsymbol{C}}_n^- \xrightarrow{d} \boldsymbol{X}$ in $\mathcal{D}((-\infty, \infty) : l_\downarrow^2)$.

Upper bound coupling: Let us construct $\{\mathcal{COM}_n^+(t) : t \ge t_n\}$ in the following way. Let $t_n^+ = t_c + n^{-\gamma}$ and let

$$
\lambda_n^+ = (t_n^+ - t_c) n^{1/3} / (\alpha \beta^{2/3}) = n^{1/3-\gamma} / (\alpha \beta^{2/3}).
$$

Let $\mathcal{COM}_n^+(t_n)$ be the graph obtained by including all immigrating doubleton and attachments during time $t \in [t_n, t_n^+]$ to the graph of $\mathcal{COM}_n(t)$, along with all the attachment edges. Namely, we construct $\mathcal{COM}_n^+(t_n)$ by including in $\mathcal{COM}_n(t_n)$ all events of type (i) and (ii) of Section 3.5.1 that occur over $[t_n, t_n^+]$. For $t > t_n$ the graph evolves in the Erdős-Rényi way such that edges are added between each pair of vertices in the fixed rate $b_n^*(t_n^+)/n$. The coupling between $\mathcal{COM}_n^+(\cdot + t_n)$ and $\mathcal{COM}_n(\cdot + t_n)$ can be achieved as follows: Construct a realization of $\{\mathcal{COM}_n(t) : t_n \le t \le t_n^+\}$

first, then use $b_n^*(t_n^+) - b_n^*(t)$ to make up for all the additional edges in $\mathcal{COM}_n^+(t)$ for $t_n \le t \le t_n^+$. Note that $\mathcal{COM}_n(t_n + \cdot) \le_d \mathcal{COM}_n^+(t_n + \cdot)$ over $[0, t_n^+ - t_n]$.

Let $\bar{\boldsymbol{C}}_n^+(\lambda) = \big(\bar{\mathcal{C}}_i^+(\lambda) : i \ge 1\big)$ be the scaled (as in (3.7.1)) component size vector for $\mathcal{COM}_n^+$. We will once more apply Proposition 4 of [2]. We first show that the three convergence statements in (3.7.2) hold with $\bar{\mathcal{C}}_i$ replaced with $\bar{\mathcal{C}}_i^+$. For this it will be convenient to consider processes under the original time scale. Write $\mathcal{C}_n^{(i)}(t_n) \equiv \mathcal{C}_i$. Also denote by $\{\mathcal{C}_i^+\}$ the component vector obtained by adding all events of type (ii) only, to $\mathcal{COM}_n(t_n)$ (i.e. attachment of singletons to components in $\mathcal{COM}_n(t_n)$), over $[t_n, t_n^+]$. Since c^* is bounded by 1, $\mathcal{C}_i^+$ is stochastically dominated by the sum of $\mathcal{C}_i$ independent copies of Geometric(p), with $p = e^{t_n - t_n^+} = e^{-2n^{-\gamma}}$. Thus

$$u_i \overset{def}{=} \mathcal{C}_i^+ - \mathcal{C}_i \le_d \text{Negative-binomial}(r, p) \text{ with } r = \mathcal{C}_i, p = e^{-2n^{-\gamma}}.$$

The random graph $\mathcal{COM}_n^+(t_n)$ contains components other than $\{\mathcal{C}_i^+\}$. These additional components correspond to the ones obtained from doubletons immigrating over $[t_n, t_n^+]$. Since there are at most n vertices, the number N of such doubletons is bounded by $n/2$. Denote by $\{\tilde{\mathcal{C}}_i^+\}_{i=1}^N$ the components corresponding to such doubletons. Once again using the fact that $c^* \le 1$, we have that

$$\tilde{\mathcal{C}}_i^+ \le_d 2 + \text{Negative-binomial}(2, p) \text{ with } p = e^{-2n^{-\gamma}}.$$

Write for $k = 2, 3$,

$$\mathcal{S}_k = \sum_i (\mathcal{C}_i)^k, \; \mathcal{S}_k^+ = \sum_i (\mathcal{C}_i^+)^k + \sum_{i=1}^N (\tilde{\mathcal{C}}_i^+)^k,$$

$$I = \max_i \mathcal{C}_i, \quad I^+ = \max\{\max_i \mathcal{C}_i^+, \max_i \tilde{\mathcal{C}}_i^+\}.$$

We shows that Propostion 3.3.1 holds with $(\mathcal{S}_2(t_n), \mathcal{S}_3(t_n), \mathcal{C}_n^{(1)}(t_n))$ replaced with $(\mathcal{S}_2^+(t_n), \mathcal{S}_3^+(t_n), I^+(t_n))$ in the following proposition.

Proposition 3.7.1. As $n \to \infty$,

$$\begin{aligned} I^+ &= \Theta(I) \\ \frac{\mathcal{S}_2^+}{\mathcal{S}_2} &\xrightarrow{\mathbb{P}} 1 \\ \frac{\mathcal{S}_3^+}{\mathcal{S}_3} &\xrightarrow{\mathbb{P}} 1 \\ n^{4/3}\left(\frac{1}{\mathcal{S}_2} - \frac{1}{\mathcal{S}_2^+}\right) &\xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Proof: Note that if U is Negative-binomial$(r, e^{-2n^{-\gamma}})$ then for some $d_1 \in (0, \infty)$

$$\mathbb{P}(U \geq 3\gamma^{-1} r) \leq \frac{d_1}{n^3}$$

and thus, as $n \to \infty$,

$$\mathbb{P}(\max_i \mathcal{C}_i^+ \geq (1 + 3\gamma^{-1})I) \leq \mathbb{P}(u_i \geq 3\gamma^{-1}\mathcal{C}_i \text{ for some } i = 1, \cdots n) \to 0.$$

A similar calculation shows that, for some $d_2 \in (0, \infty)$, as $n \to \infty$.

$$\mathbb{P}(\max_{i=1,\cdots N} \tilde{\mathcal{C}}_i^+ \geq d_2) \to 0.$$

The first statement in the proposition now follows on combining the above two displays.

Next, note that for Negative-binomial(r, p), the first, second and third moments are

$$\begin{aligned} M_1 &= \frac{1}{p} r(1-p) \\ M_2 &= \frac{1}{p^2}[r^2(1-p)^2 + r(1-p)] \\ M_3 &= \frac{1}{p^3}[r^3(1-p)^3 + 3r^2(1-p)^2 + r(4 - 9p + 7p^2 - 2p^3)]. \end{aligned}$$

From

$$\frac{\mathcal{S}_2(t_n)}{\alpha n^{1+\gamma}} \xrightarrow{\mathbb{P}} 1 \tag{3.7.6}$$

and (3.3.1) it follows that $\mathcal{S}_2 = \Theta(n^{1+\gamma})$ and $\mathcal{S}_3 = \Theta(n^{1+3\gamma})$. Also, clearly, $\sum_i \mathcal{C}_i = O(n)$.

Write $D_2 \stackrel{def}{=} \mathcal{S}_2^+ - \mathcal{S}_2 = \sum_{i=1}^N (\tilde{\mathcal{C}}_i^+)^2 + \sum_i (2\mathcal{C}_i u_i + u_i^2)$, then

$$\mathbb{E}[D_2 | \{\mathcal{C}_i\}_i] \leq d_2 \left(n \cdot n^{-\gamma} + \sum_i [(\mathcal{C}_i)^2 n^{-\gamma} + (\mathcal{C}_i)^2 n^{-2\gamma} + \mathcal{C}_i n^{-\gamma}] \right) = O(n)$$

thus $D_2/\mathcal{S}_2 \xrightarrow{\mathbb{P}} 0$ and consequently $\mathcal{S}_2^+/\mathcal{S}_2 \xrightarrow{\mathbb{P}} 1$.

Write $D_3 \stackrel{def}{=} \mathcal{S}_3^+ - \mathcal{S}_3 = \sum_{i=1}^N (\tilde{\mathcal{C}}_i^+)^3 + \sum_i [3(\mathcal{C}_i)^2 u_i + 3\mathcal{C}_i u_i^2 + u_i^3]$. One can similarly show that

$$\mathbb{E}[D_3 | \{\mathcal{C}_i\}_i] = O(n^{1+2\gamma})$$

thus $D_3/\mathcal{S}_3 \xrightarrow{\mathbb{P}} 0$ and so $\mathcal{S}_3^+/\mathcal{S}_3 \xrightarrow{\mathbb{P}} 1$.

To prove the third convergence, it suffices to prove

$$\frac{n^{4/3} D_2}{(\mathcal{S}_2)^2} \xrightarrow{\mathbb{P}} 0. \tag{3.7.7}$$

By the asymptotics shown above, we have

$$\frac{n^{4/3} D_2}{(\mathcal{S}_2)^2} = O(n^{4/3+1-2(1+\gamma)}) = O(n^{1/3-2\gamma})$$

As $\gamma > 1/6$, (3.7.7) follows and thus the proof is completed. ■

For scaled component size vector of $\mathcal{COM}_n^+$, the above proposition shows that the statements in (3.7.2) hold with $\bar{\mathcal{C}}_i$ replaced with $\bar{\mathcal{C}}_i^+$. In particular, the first and third requirements in Proposition 4 of [2] are met by $\{\bar{\mathcal{C}}_i^+\}$ Also, using the second convergence in (3.7.2), a calculation similar to that for (3.7.5) shows that

$$\lim_{n\to\infty} \left(\left(n^{2/3}\beta^{-1/3}\right)^2 \frac{b_n^*(t_n^+)}{n} \frac{\beta^{2/3}\alpha(\lambda - \lambda_n)}{n^{1/3}} - \frac{1}{\sum_i \left(\bar{\mathcal{C}}_i^+(\lambda_n)\right)^2} \right) \to \lambda.$$

Therefore the second requirement in Proposition 4 of [2] is satisfied. This proves that

$$\bar{\boldsymbol{C}}_n^+(\lambda) \xrightarrow{d} \boldsymbol{X}(\lambda). \tag{3.7.8}$$

in $l^2_\downarrow$, for every $\lambda \in \mathbb{R}$. Using Proposition 6 of [2] once again it now follows that $\bar{\boldsymbol{C}}_n^+ \xrightarrow{d} \boldsymbol{X}$ in $\mathcal{D}((-\infty,\infty) : l^2_\downarrow)$.

3.7.2 Completing the proof of Theorem 3.1.1

By [3, 4], there is a natural partial order $\preceq$ on $l^2_\downarrow$. Informally, interpreting an element of $l^2_\downarrow$ as a sequence of cluster sizes, $\mathbf{x}, \mathbf{y} \in l^2_\downarrow$, $\mathbf{x} \preceq \mathbf{y}$ if $\mathbf{y}$ can be obtained from $\mathbf{x}$ by adding new clusters and coalescing together clusters. The coupling constructed in Section 3.7.1 gives that, for every, $\lambda \in (\lambda_n, \lambda_n^+)$

$$\bar{\boldsymbol{C}}_n^-(\lambda) \preceq \bar{\boldsymbol{C}}_n^{BF}(\lambda) \preceq \bar{\boldsymbol{C}}_n^+(\lambda).$$

Since, as $n \to \infty$, $\lambda_n \to -\infty$ and $\lambda_n^+ \to +\infty$, (3.7.4), (3.7.8) along with Lemma 15 of [4] yield that

$$\bar{\boldsymbol{C}}_n^{BF}(\lambda) \xrightarrow{d} \boldsymbol{X}(\lambda)$$

for all $\lambda \in \mathbb{R}$.

Finally we argue convergence in $\mathcal{D}((-\infty,\infty) : l^2_\downarrow)$. For $\mathbf{x}, \mathbf{y} \in l^2_\downarrow$, let $\mathbf{d}^2(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^\infty (x_i - y_i)^2$, $\mathbf{x} = \{x_i\}$, $\mathbf{y} = \{y_i\}$. Then $\mathbf{d}^2(\mathbf{x}, \mathbf{y}) < \sum_i y_i^2 - \sum_i x_i^2$ whenever $\mathbf{x} \preceq \mathbf{y}$. To prove that $\bar{\boldsymbol{C}}_n^{BF} \to \boldsymbol{X}$ in $\mathcal{D}((-\infty,\infty) : l^2_\downarrow)$ it suffices to prove that

$$\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathbf{d}(\bar{\boldsymbol{C}}_n^{BF}, \bar{\boldsymbol{C}}_n^-) \xrightarrow{\mathbb{P}} 0, \text{ for all } -\infty < \lambda_1 < \lambda_2 < \infty. \tag{3.7.9}$$

Fix λ_1, λ_2 as above. Then

$$\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathbf{d}(\bar{\boldsymbol{C}}_n^{BF}, \bar{\boldsymbol{C}}_n^-) \leq \sup_{\lambda \in [\lambda_1, \lambda_2]} [\sum_i (\bar{\mathcal{C}}_i^+(\lambda))^2 - \sum_i (\bar{\mathcal{C}}_i^-(\lambda))^2]. \tag{3.7.10}$$

Let, for $\lambda \in \mathbb{R}$,

$$\mathcal{U}_+(\lambda) = \sum_i (\bar{\mathcal{C}}_i^+(\lambda))^2, \ \mathcal{U}_-(\lambda) = \sum_i (\bar{\mathcal{C}}_i^-(\lambda))^2 \text{ and } \mathcal{V}(\lambda) = \mathcal{U}_+(\lambda) - \mathcal{U}_-(\lambda).$$

From Lemma 15 of [4], $\mathcal{V}(\lambda) \xrightarrow{\mathbb{P}} 0$ for every $\lambda \in \mathbb{R}$. Thus it suffices to show that $\mathcal{V}$ is tight in $\mathcal{D}((-\infty,\infty) : \mathbb{R}_+)$. Note that both $\mathcal{U}_+$ and $\mathcal{U}_-$ are tight in $\mathcal{D}((-\infty,\infty) : \mathbb{R}_+)$.

Although, in general difference of relatively compact sequences in the $\mathcal{D}$-space need not be relatively compact, in the current setting due to properties of the multiplicative coalescent this difficulty does not arise. Indeed, if $\{\boldsymbol{X}^{\mathbf{x}}(t), t \geq 0\}$ denotes the multiplicative coalescent on the positive real line with initial condition $\mathbf{x} \in l^2_{\downarrow}$ then, for δ sufficiently small

$$\begin{aligned}\sup_{\tau \in \mathcal{T}(\delta)} \mathbb{E}\left(\mathbf{d}^2(\boldsymbol{X}^{\mathbf{x}}(\tau), \mathbf{x}) \wedge 1\right) \leq & \mathbb{E}\left[\sum_i (X_i^{\mathbf{x}}(\delta))^2 - \sum_i x_i^2\right] \\ \leq & 2\sum_{i<j} \delta x_i x_j \cdot 2x_i x_j \leq 2\delta ||\mathbf{x}||^4,\end{aligned}$$

where, $||\mathbf{x}|| = (\sum x_i^2)^{1/2}$, $\mathcal{T}(\delta)$ is the family of all stopping times (with the natural filtration) bounded by δ. Using the above property, the Markov property of the coalescent process and the tightness of $\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathcal{U}_+(\lambda)$, $\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathcal{U}_-(\lambda)$ one can verify Aldous's tightness criteria (see Theorem VI.4.5 in [20]) for $\mathcal{V}$ thus proving the desired tightness. ■

CHAPTER 4: THE BOUNDED-SIZE-RULE PROCESSES

4.1 Introduction

In recent years, there has been an increasing interest in understanding the role of limited choice along with randomness in the evolution of the network, in particular how they affect the time and nature of emergence of the giant component. The simplest such model that has been rigorously analyzed is the Bohman-Frieze process which was studied in Chapter 3. In Chapter 5 we will show that all bounded-size-rule processes have the same limiting behavior as the Bohman-Frieze model in the critical window. In order to establish such a limit theorem, the first step is to understand the asymptotics of the process right before it enters the critical window. This is the goal of the current chapter. More precisely, in this chapter we study the bounded-size-rule processes in the barely subcritical regime, i.e., when $t = t_c - n^{-\gamma}$ for $\gamma \in (0, 1/4)$, and prove a useful upper bound of order $n^{2\gamma} \log^4 n$ on the sizes of components in this regime. Furthermore, using a coupling between the bounded-size-rule processes and some inhomogeneous random graph models, we give a new characterization of the critical time for the phase transitions for all bounded-size-rule processes. These results will form the key ingredients in the proof of the results in Chapter 5.

Organization of this chapter: In Section 4.2 we give a precise construction of the bounded-size-rule processes and state our main theorems. Section 4.3 collects some notation used in this chapter. In Section 4.4 we introduce and analyze certain inhomogeneous random graphs associated with bounded-size-rule processes. Finally, Section 4.5 completes the proofs of the main results, Theorems 4.2.2 and 4.2.3.

4.2 Model and main theorems

The bounded-size K-rule process $\{\mathbf{BSR}^{(n)}(t)\}_{t\geq 0}$. Fix $K \geq 0$, and let $\Omega_K = \{1, 2, \ldots, K, \varpi\}$. Conceptually ϖ represents components of size greater than K. Given a graph $\mathbf{G}$ and a vertex $v \in \mathbf{G}$, write $\mathcal{C}_v(\mathbf{G})$ for the component that contains v. Define

$$c_{\mathbf{G}}(v) := \begin{cases} |\mathcal{C}_v(\mathbf{G})| & \text{if } |\mathcal{C}_v(\mathbf{G})| \leq K \\ \varpi & \text{if } |\mathcal{C}_v(\mathbf{G})| > K \end{cases} \tag{4.2.1}$$

For a quadruple of (not necessarily distinct) vertices v_1, v_2, v_3, v_4, write $\vec{v}$ for the ordered quadruple $\vec{v} = (v_1, v_2, v_3, v_4)$. Let $c_{\mathbf{G}}(\vec{v}) = (c_{\mathbf{G}}(v_1), c_{\mathbf{G}}(v_2), c_{\mathbf{G}}(v_3), c_{\mathbf{G}}(v_4))$. Fix $F \subseteq \Omega_K^4$. The set F will be another parameter in the construction of the process. The F-bounded-size rule(F-BSR) is defined as follows:

(a) At time $k = 0$ start with the empty graph $\mathbf{BSR}_0^{(n)} := \mathbf{0}_n$ on $[n]$ vertices.

(b) For $k \geq 0$, having constructed the graph $\mathbf{BSR}_k^{(n)}$, construct $\mathbf{BSR}_{k+1}^{(n)}$ as follows. Choose 4 vertices $\vec{v} = (v_1, v_2, v_3, v_4)$ uniformly at random amongst all n^4 possible quadruples and let $c_k(\vec{v}) = c_{\mathbf{BSR}_k}(\vec{v})$. If $c_k(\vec{v}) \in F$ then $\mathbf{BSR}_{k+1}^{(n)} = \mathbf{BSR}_k^{(n)} \cup (v_1, v_2)$ else $\mathbf{BSR}_{k+1}^{(n)} = \mathbf{BSR}_k^{(n)} \cup (v_3, v_4)$.

Mathematically it is more convenient to work with a formulation in which edges are added at Poissonian time instants rather than at fixed discrete times. More precisely, we will consider the following random graph process (denoted once more as $\mathbf{BSR}^{(n)}(t)$). For every quadruple of vertices $\vec{v} = (v_1, v_2, v_3, v_4) \in [n]^4$, let $\mathcal{P}_{\vec{v}}$ be a Poisson process with rate $1/2n^3$, independent between quadruples. Note that this implies that the rate of creation of edges is $n^4 \times 1/2n^3 = n/2$. Thus we have sped up time by a factor $n/2$ as in the above discrete time construction. Start with $\mathbf{BSR}^{(n)}(0) = \mathbf{0}_n$. For any $t \geq 0$, at which there is a point in $\mathcal{P}_{\vec{v}}$ for a quadruple

$\vec{v} \in [n]^4$, define

$$\mathbf{BSR}^{(n)}(t) = \begin{cases} \mathbf{BSR}^{(n)}(t-) \cup (v_1, v_2) & \text{if } c_{t-}(\vec{v}) \in F \\ \mathbf{BSR}^{(n)}(t-) \cup (v_3, v_4) & \text{otherwise,} \end{cases}$$

where $c_{t-}(\vec{v}) = c_{\mathbf{BSR}(t-)}(\vec{v})$.

Two examples of such processes are Erdős-Rényi process (here $K = 0$, $\Omega_K = \{\varpi\}$ and $F = \{(\varpi, \varpi, \varpi, \varpi)\}$) and Bohman-Frieze process (here $K = 1$, $\Omega_K = \{1, \varpi\}$ and $F = \{(1, 1, j_3, j_4) : j_3, j_4 \in \Omega_K\}$). Spencer and Wormald[31] showed that every bounded-size rules exhibits a phase transition similar to the Erdős-Rényi random graph process. More precisely, write $\mathcal{C}_i^{(n)}(t)$ for the i-th largest component in $\mathbf{BSR}^{(n)}(t)$, and $|\mathcal{C}_i^{(n)}(t)|$ for the size of this component. Define the *susceptibility* function

$$\mathcal{S}_2(t) = \sum_{i=1}^{\infty} |\mathcal{C}_i^{(n)}(t)|^2. \tag{4.2.2}$$

Then [31] proves the following result.

Theorem 4.2.1 (Theorem 1.1, [31]). Fix $F \in \Omega_K^4$. Then for the random graph process associated with the F-BSR, there exists deterministic monotonically increasing function $s_2(t)$ and a critical time t_c such that $\lim_{t \uparrow t_c} s_2(t) = \infty$ and

$$\frac{\mathcal{S}_2(t)}{n} \xrightarrow{\mathbb{P}} s_2(t) \text{ as } n \to \infty, \text{ for all } t \in [0, t_c).$$

For fixed $t < t_c$, $|\mathcal{C}_1^{(n)}(t)| = O(\log n)$ while for $t > t_c$, $|\mathcal{C}_1^{(n)}(t)| = \Theta_P(n)$.

Here we use o, O, Θ in the usual manner. Given a sequence of random variables $\{\xi_n\}_{n \geq 1}$ and a function $f(n)$, we say $\xi_n = O(f)$ if there is a constant C such that $\xi_n \leq Cf(n)$ with high probability (whp), and we say $\xi_n = \Omega(f)$ if there is a constant C such that $\xi_n \geq Cf(n)$ whp. Say that $\xi_n = \Theta(f)$ if $\xi_n = O(f)$ **and** $\xi_n = \Omega(f)$. In addition, we say $\xi_n = o(f)$ if $\xi_n / f(n) \xrightarrow{\mathbb{P}} 0$.

Thus as t transitions from less than t_c to greater than t_c, the size of the largest component jumps from size $O(\log n)$ to a giant component $\Theta(n)$. The aim of this chapter is to study the barely subcritical regime, i.e. to analyze the behavior of the size of the largest component at times $t = t_c - \varepsilon_n$ where $\varepsilon_n \to 0$. The following is the main result.

Theorem 4.2.2 (**Barely subcritical regime**). Fix $F \subset \Omega_K^4$ and $\gamma \in (0, 1/4)$. Then there exists $B \in (0, \infty)$ such that,

$$\mathbb{P}\left\{ |\mathcal{C}_1^{(n)}(t)| \le B \frac{(\log n)^4}{(t_c - t)^2}, \ \forall t \le t_c - n^{-\gamma} \right\} \to 1,$$

as $n \to \infty$.

As another consequence of our proofs, we obtain an alternative characterization of the critical time for a bounded-size rule given in Theorem 4.2.3 below. Let $\mathcal{X} = [0, \infty) \times \mathcal{D}([0, \infty) : \mathbb{N}_0)$ where $\mathcal{D}([0, \infty) : \mathbb{N}_0)$ is the Skorohod D-space of functions that are right continuous and have left limits with values in the space of nonnegative integers, equipped with the usual Skorohod topology. Given a finite measure μ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and a measurable map $\kappa : \mathcal{X} \times \mathcal{X} \to [0, \infty)$ satisfying $\int_{\mathcal{X} \times \mathcal{X}} \kappa^2(\mathbf{x}, \mathbf{y}) \mu(d\mathbf{x}) \mu(d\mathbf{y}) < \infty$, define the integral operator $\mathcal{K} : L^2(\mathcal{X}, \mu) \to L^2(\mathcal{X}, \mu)$ as

$$\mathcal{K} f(\mathbf{x}) = \int_{\mathcal{X}} \kappa(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) \mu(d\mathbf{y}), f \in L^2(\mathcal{X}, \mu), \ \mathbf{x} \in \mathcal{X}.$$

We refer to κ as a kernel on $\mathcal{X} \times \mathcal{X}$ and $\mathcal{K}$ as the integral operator associated with (κ, μ). We will show the following result.

Theorem 4.2.3 (**Characterization of the critical time**). Fix $F \subset \Omega_K^4$. Then there exists a collection of kernels $\{\kappa_t\}_{t \ge 0}$ on $\mathcal{X} \times \mathcal{X}$ and finite measures $\{\mu_t\}_{t \ge 0}$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that the integral operators $\mathcal{K}_t$ associated with (κ_t, μ_t), $t > 0$, have the property that the operator norms $\rho(t) = \|\mathcal{K}_t\|$ are continuous and strictly increasing in t. Furthermore, t_c is the unique time instant such that $\rho(t_c) = 1$.

See Section 4.4.3 for a precise definition of κ_t and μ_t. Using arguments similar to [23] for the Bohman-Frieze model, one can show that for any fixed $\varepsilon > 0$, the size of the largest component at time $t = t_c - \varepsilon$ can be lower bounded as $|\mathcal{C}_1^{(n)}(t)| \geq A \log n/(t_c - t)^2$ where A is a constant independent of ε. Thus up to logarithmic factors one expects the upper bound in Theorem 4.2.2 to be tight.

Theorem 4.2.2 plays a central role in the study of the asymptotic dynamic behavior of the process describing the vector of component sizes and associated surpluses for BSR processes in the critical scaling window and its connections with the augmented multiplicative coalescent process. This study is the subject of Chapter 5 to which we refer the reader for details.

4.3 Notation

We collect some notation used through the rest of the chapter. All unspecified limits are taken as $n \to \infty$. We use $\xrightarrow{\mathbb{P}}$ and $\xrightarrow{d}$ to denote convergence in probability and in distribution respectively. Given a sequence of events $\{E_n\}_{n\geq 1}$, we say E_n occurs with high probability (whp) if $\mathbb{P}\{E_n\} \to 1$.

For a set S and a function $\mathbf{g} : S \to \mathbb{R}^k$, we write $||\mathbf{g}||_\infty = \sum_{i=1}^k \sup_{s\in S} |g_i(s)|$, where $\mathbf{g} = (g_1, \cdots g_k)$. For a Polish space S, we denote by BM(S), the space of bounded measurable functions on S (equipped with the Borel sigma-field $\mathcal{B}(S)$.) For a finite set S, $|S|$ denotes the number of elements in the set. $\mathbb{N}_0$ is the set of nonnegative integers. For ease of notation, we shall often suppress the dependence on n and shall write for example $\mathbf{BSR}(t) = \mathbf{BSR}^{(n)}(t)$. Recall the Poisson processes $\mathcal{P}_{\vec{v}}$ used to construct $\mathbf{BSR}(\cdot)$ in Section 4.1. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be the associated filtration: $\mathcal{F}_t = \sigma\{\mathcal{P}_{\vec{v}}(s) : s \leq t, \vec{v} \in [n]^4\}$. We shall often deal with $\{\mathcal{F}_t\}$-semimartingales $\{J(t)\}_{t\geq 0}$ of the form

$$dJ(t) := \alpha(t)dt + dM(t), \tag{4.3.1}$$

where M is a $\{\mathcal{F}_t\}$ local martingale. We shall denote $\alpha = \mathbf{d}(J)$ and $M = \mathbf{M}(J)$. For a local martingale $M(t)$, we shall write $\langle M, M\rangle(t)$ for the predictable quadratic variation process namely the predictable process of bounded variation such that $M(t)^2 - \langle M, M\rangle(t)$ is a local martingale.

4.4 Inhomogeneous random graphs

Fix $K \geq 0$ and a general bounded-size rule $F \subseteq \Omega_K^4$ and recall that $\{\mathbf{BSR}(t)\}_{t\geq 0}$ denotes the continuous time bounded-size rule process started with the empty graph at $t = 0$. Note that $K = 0$ case corresponds to the Erdős-Rényi random graph process for which results such as Theorem 4.2.2 are well known. Thus, henceforth we shall assume $K \geq 1$. We begin in Section 4.4.1 by analyzing the proportion of vertices in components of size i for $i \leq K$. As shown in [31], these converge to a set of deterministic functions which can be characterized as the unique solution of a set of differential equations. We will need precise rates of convergence for these proportions which we establish in Lemma 4.4.2. We then study the evolution of components of size larger than K in Section 4.4.2. Finally, we relate the evolution of these components to an inhomogeneous random graph (IRG) model in Section 4.4.3.

4.4.1 Density of vertices in components of size bounded by K

Recall from Section 4.1, (4.2.1) that $c_t(v) = c_{\mathbf{BSR}(t)}(v)$, for $v \in [n]$. For $t \geq 0$ and $i \in \Omega_K$, define

$$X_i(t) = |\{v \in [n] : c_t(v) = i\}| \text{ and } \bar{x}_i(t) = X_i(t)/n. \tag{4.4.1}$$

Following [31], the first step in analyzing bounded-size rules is understanding the evolution of $\bar{x}_i(\cdot)$ as functions of time as $n \to \infty$. Although [31] proves the convergence of $\bar{x}_i(t)$ as $n \to \infty$, we give here a self contained proof of this convergence with precise rates of convergence that will be needed in the proof of Theorem 4.2.2.

Note that the BSR process changes values at the occurrence of points in the

Poisson processes $\mathcal{P}_{\vec{v}}$, $\vec{v} \in [n]^4$. We call each such occurrence as a 'round' and call a round **redundant** if the added edge in that round joins two vertices in the same component. Note that such rounds do not have any effect on component sizes or on the vector $\bar{\mathbf{x}}(t) = (\bar{x}_1(t), \bar{x}_2(t), \ldots, \bar{x}_K(t), \bar{x}_\varpi(t))$. We will in fact observe that such rounds are quite rare. We now describe the effect of non-redundant rounds on $\bar{\mathbf{x}}(\cdot)$. For $\vec{j} \in \Omega_K^4$ and $i \in \Omega_K$, write $\Delta(\vec{j}; i)$ for the change $\Delta X_i(t) := X_i(t) - X_i(t-)$ at an occurrence time t if the chosen quadruple $\vec{v} \in [n]^4$ satisfies $c_{t-}(\vec{v}) = \vec{j}$ and the round is not redundant. It is easy to check (see Section 2.1, [31]) that, when $\vec{j} = (j_1, j_2, j_3, j_4) \in F$,

$$\begin{aligned}
\Delta(\vec{j}; i) &= i \cdot (\mathbf{1}_{\{j_1+j_2=i\}} - \mathbf{1}_{\{j_1=i\}} - \mathbf{1}_{\{j_2=i\}}), \text{ for } 1 \le i \le K, \\
\Delta(\vec{j}; \varpi) &= \mathbb{1}_{\{j_1+j_2=\varpi\}}(j_1 \mathbb{1}_{\{j_1 \le K\}} + j_2 \mathbb{1}_{\{j_2 \le K\}}),
\end{aligned}$$

with the convention $j_1 + j_2 = \varpi$ when the sum of j_1, j_2 is greater than K, and $j_1 + \varpi = \varpi + j_1 = \varpi$ for all $j_1 \in \Omega_K$. For $\vec{j} = (j_1, j_2, j_3, j_4) \in F^c$ one uses the second edge $\{v_3, v_4\}$ and the expressions for $\Delta(\vec{j}; i)$ are identical to the above, with (j_3, j_4) replacing (j_1, j_2). Note that the corresponding change in the density $\bar{x}_i(t) = X_i(t)/n$ is given by $\Delta(\vec{j}; i)/n$. For $\vec{j} \in \Omega_K^4$ and $t > 0$, write

$$\mathcal{Q}(t; \vec{j}) := \left\{ \vec{v} \in [n]^4 : c_t(\vec{v}) = \vec{j} \right\}.$$

Since each quadruple $\vec{v} \in [n]^4$ is selected according to the Poisson process $\mathcal{P}_{\vec{v}}$ with rate $1/2n^3$, the above description of the jumps of $X_i(\cdot)$ leads to a semi-martingale decomposition of $\bar{x}_i$ of the form (4.3.1) with

$$\mathbf{d}(\bar{x}_i)(t) = \sum_{\vec{j} \in F} \sum_{\vec{v} \in \mathcal{Q}(t; \vec{j})} \frac{\Delta(\vec{j}; i)}{2n^4} \mathbb{1}\left\{\mathcal{C}_{v_1}(t) \ne \mathcal{C}_{v_2}(t)\right\} + \sum_{\vec{j} \in F^c} \sum_{\vec{v} \in \mathcal{Q}(t; \vec{j})} \frac{\Delta(\vec{j}; i)}{2n^4} \mathbb{1}\left\{\mathcal{C}_{v_3}(t) \ne \mathcal{C}_{v_4}(t)\right\}, \tag{4.4.2}$$

where $\mathcal{C}_v(t) := \mathcal{C}_v(\mathbf{BSR}(t))$ denotes the component containing v in $\mathbf{BSR}(t)$.

Define for $i \in \Omega_K$, the functions $F_i : [0,1]^{K+1} \to \mathbb{R}$ mapping the vector $\mathbf{x} =$

$(x_1, x_2, ..., x_K, x_\varpi) \in \mathbb{R}^{K+1}$ to

$$F_i^x(\mathbf{x}) = \frac{1}{2} \sum_{\vec{j} \in F} \Delta(\vec{j}; i) x_{j_1} x_{j_2} x_{j_3} x_{j_4} + \frac{1}{2} \sum_{\vec{j} \in F^c} \Delta(\vec{j}; i) x_{j_1} x_{j_2} x_{j_3} x_{j_4}. \tag{4.4.3}$$

Note that $|\Delta(\vec{j}; i)| \le 2K$ for all $\vec{j} \in \Omega_K^4$. Also,

$$\max \left\{ \sum_{\vec{j} \in F} \sum_{\vec{v} \in \mathcal{Q}(t;\vec{j})} \mathbb{1} \left\{ \mathcal{C}_{v_1}(t) = \mathcal{C}_{v_2}(t) \right\}, \sum_{\vec{j} \in F^c} \sum_{\vec{v} \in \mathcal{Q}(t;\vec{j})} \mathbb{1} \left\{ \mathcal{C}_{v_3}(t) = \mathcal{C}_{v_4}(t) \right\} \right\} \le n^3 K.$$

Thus we have

$$|\mathbf{d}(\bar{x}_i)(t) - F_i^x(\bar{\mathbf{x}}(t))| \le \frac{2K}{2n^4} \cdot 2Kn^3 = \frac{2K^2}{n}. \tag{4.4.4}$$

Note that $\bar{x}_1(0) = 1$ while $\bar{x}_i(0) = 0$ for other $i \in \Omega_K$. Guided by equations (4.4.2) – (4.4.4), [31] considered the system of differential equations for $\mathbf{x}(t) := (x_j(t) : j \in \Omega_K)$

$$x_i'(t) = F_i^x(\mathbf{x}(t)), \;\; i \in \Omega_K, \;\; t \ge 0, \; \mathbf{x}(0) = (1, 0, ..., 0), \tag{4.4.5}$$

and showed the following result.

Theorem 4.4.1 (Theorem 2.1, [31]). Equation (4.4.5) has a unique solution. For all $i \in \Omega_K$ and $t > 0$, $x_i(t) > 0$. Furthermore $\sum_{i \in \Omega_K} x_i(t) = 1$ and $\lim_{t \to \infty} x_\varpi(t) = 1$.

[31] also showed that the functions $\bar{x}_i(t) \xrightarrow{\mathbb{P}} x_i(t)$ for each fixed $t \ge 0$. We will need precise rates of convergence for our proofs for which we establish the following result.

Lemma 4.4.2. Fix $\delta \in (0, 1/2)$ and $T > 0$. There exist $C_1, C_2 \in (0, \infty)$ such that for all n,

$$\mathbb{P} \left(\sup_{i \in \Omega_K} \sup_{s \in [0,T]} |\bar{x}_i(t) - x_i(t)| > n^{-\delta} \right) < C_1 \exp\left(-C_2 n^{1-2\delta} \right).$$

Proof. Note that $F_i^x(\cdot)$ is a Lipchitz function, indeed for $\mathbf{x}, \tilde{\mathbf{x}} \in [0,1]^{K+1}$,

$$|F_i^x(\mathbf{x}) - F_i^x(\tilde{\mathbf{x}})| \le 4K(K+1)^4 \sum_{i \in \Omega_K} |x_i - \tilde{x}_i| \le 4K(K+1)^5 \sup_{i \in \Omega_K} |x_i - \tilde{x}_i|.$$

Write $D(t) := \sup_{i\in\Omega_K} |\bar{x}_i(t) - x_i(t)|$ and $M_i(t) := \mathbf{M}(\bar{x}_i)(t)$. Using (4.4.4), we get for all $i \in \Omega_K$ and $t \in [0, T]$,

$$\begin{aligned}|\bar{x}_i(t) - x_i(t)| \leq& \int_0^t |F_i^x(\bar{\mathbf{x}}(s)) - F_i^x(\mathbf{x}(s))| ds + T \cdot \frac{2K^2}{n} + |M_i(t)| \\ \leq& 4K(K+1)^5 \int_0^t D(s) ds + T \cdot \frac{2K^2}{n} + |M_i(t)|.\end{aligned}$$

Taking $\sup_{i\in\Omega_K}$ on both sides and using Gronwall's lemma we have

$$\sup_{t\in[0,T]} D(t) \leq \left(\sup_{i\in\Omega_K} \sup_{t\in[0,T]} |M_i(t)| + \frac{2TK^2}{n} \right) e^{4K(K+1)^5 T}.$$

Thus, for a suitable $d_1 \in (0, \infty)$,

$$\mathbb{P}\left\{ \sup_{t\in[0,T]} D(t) > n^{-\delta} \right\} \leq \sum_{i\in\Omega_K} \mathbb{P}\left\{ \sup_{t\in[0,T]} |M_i(t)| > d_1 n^{-\delta} \right\}. \tag{4.4.6}$$

To complete the proof we will use exponential tail bounds for martingales. From Theorem 5 in Section 4.13 of [25] we have that, for a square integrable martingale M with $M(0) = 0$, $|\Delta M(t)| \leq c$ for all t, and $\langle M, M\rangle(T) \leq Q$, a.s., for some $c, Q \in (0, \infty)$,

$$\mathbb{P}\left\{ \sup_{0\leq t\leq T} |M(t)| > \alpha \right\} \leq 2\exp\left\{ -\sup_{\lambda>0} [\alpha\lambda - Q\psi(\lambda)] \right\}, \text{ for all } \alpha > 0,$$

where $\psi(\lambda) = \frac{e^{\lambda c} - 1 - \lambda c}{c^2}$. Optimizing over λ, we get the bound

$$\mathbb{P}\left\{ \sup_{0\leq t\leq T} |M(t)| > \alpha \right\} \leq 2\exp\left\{ -\frac{\alpha}{c}\log\left(1 + \frac{\alpha c}{Q}\right) + \left[\frac{\alpha}{c} - \frac{Q}{c^2}\log\left(1 + \frac{\alpha c}{Q}\right) \right] \right\}. \tag{4.4.7}$$

In our context, note that for any $i \in \Omega_K$, $|\Delta M_i(t)| = |\Delta\bar{x}_i(t)| \leq 2K/n$. Also, the total rate of jumps is bounded by $n^4 \cdot \frac{1}{2n^3}$. Thus for all $i \in \Omega_K$, the quadratic variation process

$$\langle M_i, M_i\rangle(T) \leq \int_0^T \left(\frac{2K}{n}\right)^2 \times \frac{n^4}{2n^3} dt = \frac{2K^2 T}{n}.$$

Taking $\alpha = d_1 n^{-\delta}, Q = 2K^2T/n$ and $c = 2K/n$ in (4.4.7) completes the proof. ■

4.4.2 Evolution of components of size larger than K

Let $\mathbf{BSR}^*(t)$ denote the subgraph of $\mathbf{BSR}(t)$ consisting of components of size greater than K. In this section, we will focus on the dynamics and evolution of $\mathbf{BSR}^*(t)$. Note that $\mathbf{BSR}^*(0) = \emptyset$, i.e. a graph with no vertices or edges. As time progresses components of size less than K merge and components of size greater than K emerge. Three distinct types of events affect the evolution $\mathbf{BSR}^*(t)$:

1. **Immigration:** This occurs when two components of size $\leq K$ merge into a single component of size $> K$. We view the resulting component as a new immigrant into $\mathbf{BSR}^*(t)$. Note that the first component to appear in $\mathbf{BSR}^*(t)$ is an immigrant.
2. **Attachments**: This occurs when a component of size $\leq K$ gets linked to a component of size larger than K. The former component enters $\mathbf{BSR}^*(t)$ via attaching itself to a component of size larger than K.
3. **Edge formation**: Two distinct components of size larger than K merge into a single component via formation of an edge between these components. In this case, the vertex set of $\mathbf{BSR}^*(t)$ remains unchanged.

We now introduce some functions that describe the rate of occurrence for each of the three types of events. For $i_1, i_2 \in \Omega_K$, define $F^x_{i_1,i_2} : [0,1]^{K+1} \to \mathbb{R}$ as

$$F^x_{i_1,i_2}(\mathbf{x}) = \frac{1}{2}\left(\sum_{\vec{j}\in F:\{j_1,j_2\}=\{i_1,i_2\}} x_{j_1}x_{j_2}x_{j_3}x_{j_4} + \sum_{\vec{j}\in F^c:\{j_3,j_4\}=\{i_1,i_2\}} x_{j_1}x_{j_2}x_{j_3}x_{j_4} \right). \quad (4.4.8)$$

For $i_1, i_2 \leq K$, denote $n \cdot R_{i_1,i_2}(t)$ as the rate at which two components of size i_1, i_2 merge. When $i_1 \neq i_2$, this rate is precisely

$$(2n^3)^{-1}[\sum_{\substack{\vec{j}\in F\\ \{j_1,j_2\}=\{i_1,i_2\}}} X_{j_1}(t)X_{j_2}(t)X_{j_3}(t)X_{j_4}(t) + \sum_{\substack{\vec{j}\in F\\ \{j_3,j_4\}=\{i_1,i_2\}}} X_{j_1}(t)X_{j_2}(t)X_{j_3}(t)X_{j_4}(t)] := n \cdot F^x_{i_1,i_2}(\bar{\mathbf{x}}(t)).$$

Thus $R_{i_1,i_2}(t) = F^x_{i_1,i_2}(\bar{\mathbf{x}}(t))$. The case $i_1 = i_2 \le K$ is more subtle due to redundant rounds linking vertices in the same component. The rate of redundant rounds can be bounded by $\frac{1}{2n^3} \cdot Kn^3 \cdot 2 = K$, from which it follows that

$$|R_{i,i}(t) - F^x_{i,i}(\bar{\mathbf{x}}(t))| \le \frac{K}{n}.$$

The case $i_1 = i_2 = \varpi$ corresponds to creation of edges in $\mathbf{BSR}^*(t)$ and $n \cdot F^x_{\varpi,\varpi}(\bar{\mathbf{x}}(t))$ is the rate of creation of such edges.

We now give expressions for the rates for the three types of events that govern the evolution of $\mathbf{BSR}^*(t)$. The convention followed for the rest of this section is that for $i_1, i_2 \in \Omega_K$, $i_1 + i_2 = \varpi$ when the sum of is greater than K, and $\varpi + i_1 = i_1 + \varpi = \varpi$ for all $i_1 \in \Omega_K$.

I. Immigrating vertices: For $1 \le i \le K$, write $\mathfrak{I}_i(t) := n \cdot a^*_i(t)$ for the rate at which components of size $K + i$ immigrate into $\mathbf{BSR}^*(t)$ at time t. Using the above expressions for the rate of merger of components of various sizes we have

$$\left| a^*_i(t) - \sum_{1 \le i_i, i_2 \le K : i_1 + i_2 = K + i} F^x_{i_1,i_2}(\bar{\mathbf{x}}(t)) \right| \le \frac{K}{n}. \tag{4.4.9}$$

As before the error is due to redundant rounds which can only occur for $i_1 = i_2 = (K + i)/2$ (and when $(K + i)/2$ is an integer). Now define functions $F^a_i : [0,1]^{K+1} \to \mathbb{R}_+$, and $a_i(\cdot) : [0,\infty) \to [0,\infty)$ by

$$F^a_i(\mathbf{x}) = \sum_{\substack{1 \le i_i, i_2 \le K, \\ i_1 + i_2 = K + i}} F^x_{i_1,i_2}(\mathbf{x}), \qquad a_i(t) = F^a_i(\mathbf{x}(t)), \tag{4.4.10}$$

where $\mathbf{x}(t)$ is as in (4.4.5). Then (4.4.9) says that

$$\sup_{t \in [0,\infty)} |a^*_i(t) - F^a_i(\bar{\mathbf{x}}(t))| \le K/n. \tag{4.4.11}$$

Note that for any $\delta < 1$, the error term in (4.4.11) is $o(n^{-\delta})$. Using this observation along with the Lipschitz property of F_{i_1,i_2}, we have from Lemma 4.4.2 that for any fixed $T > 0$ and $\delta < 1/2$,

$$\mathbb{P}(\sup_{1\le i\le K}\sup_{s\in[0,T]}|a_i^*(t) - a_i(t)| > n^{-\delta}) \le C_1 \exp(-C_2 n^{1-2\delta}). \tag{4.4.12}$$

The constants C_1, C_2 here may be different from those in Lemma 4.4.2, however for notational ease we use the same symbols.

II. Attachments: Fix $1 \le i \le K$ and a vertex v contained in a component in $\mathbf{BSR}^*(t)$. Let, for $i \le K$, $c_i^*(t)$ denote the rate at which a component of size i attaches itself to the component of v through an edge connecting the former component to v. This rate can be calculated as follows. First note that the total rate of merger between a component of size i with a component in $\mathbf{BSR}^*(t)$ is $n \cdot F^x_{i,\varpi}(\bar{\mathbf{x}}(t))$. Since there are $X_\varpi(t)$ vertices in $\mathbf{BSR}^*(t)$ each of which has the same probability of being the vertex through which this attachment event happens, the rate at which a component of size i attaches to v is given by $nF^x_{i,\varpi}(\bar{\mathbf{x}}(t))/X_\varpi(t) = F^x_{i,\varpi}(\bar{\mathbf{x}}(t))/\bar{x}_\varpi(t)$. Since x_ϖ is a factor of $F^x_{i,\varpi}(\mathbf{x})$, $c_i^*(t)$ is a polynomial in $\bar{\mathbf{x}}(t)$. Define the functions $F_i^c : [0,1]^{K+1} \to \mathbb{R}$ and $c_i(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ as

$$F_i^c(\mathbf{x}) = F^x_{i,\varpi}(\mathbf{x})/x_\varpi, \qquad c_i(t) = F_i^c(\mathbf{x}(t)). \tag{4.4.13}$$

Then $c_i^*(t) = F_i^c(\bar{\mathbf{x}}(t))$. Once again using Lemma 4.4.2 we get for any $\delta < 1/2$ and $T > 0$,

$$\mathbb{P}(\sup_{1\le i\le K}\sup_{s\in[0,T]}|c_i^*(t) - c_i(t)| > n^{-\delta}) \le C_1 \exp(-C_2 n^{1-2\delta}). \tag{4.4.14}$$

III. Edge formation: Note that the rate of creation of an edge between vertices in $\mathbf{BSR}^*(t)$ is $nF^x_{\varpi,\varpi}(\bar{\mathbf{x}}(t))$. Since such an edge is equally likely to be between any of the $X^2_\varpi(t)$ pairs of vertices in $\mathbf{BSR}^*(t)$, we have that the rate of creation of an edge between specified vertices $\{v_1, v_2\}$ with $v_1, v_2 \in \mathbf{BSR}^*(t)$ is $b^*(t)/n$ where $b^*(t) = F^x_{\varpi,\varpi}(\bar{\mathbf{x}}(t))/x^2_\varpi(t)$. Define $F^b : [0,1]^{K+1} \to \mathbb{R}$ and $b(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ as

$$F^b(\mathbf{x}) = F^x_{\varpi,\varpi}(\mathbf{x})/x^2_\varpi \text{ and } b(t) = F^b(\mathbf{x}(t)). \tag{4.4.15}$$

Once more it is clear that $F^b(\mathbf{x})$ is a polynomial and furthermore $b^*(t) = F^b(\bar{\mathbf{x}}(t))$, so by Lemma 4.4.2, for any $\delta < 1/2$ and $T > 0$,

$$\mathbb{P}(\sup_{s\in[0,T]} |b^*(t) - b(t)| > n^{-\delta}) \leq C_1 \exp(-C_2 n^{1-2\delta}). \tag{4.4.16}$$

Write $\mathbf{a}(t) := \{a_i(t)\}_{1\leq i\leq K}$ and $\mathbf{c}(t) := \{c_i(t)\}_{1\leq i\leq K}$. We refer to $(\mathbf{a}, b, \mathbf{c})$ as rate functions. In the proposition below we collect some properties of these rate functions. These properties are easy consequences of Theorem 4.4.1.

Proposition 4.4.3. (a) For all $1 \leq i \leq K$ and $t > 0$, $b(t), a_i(t), c_i(t) > 0$.

(b) We have

$$\|\mathbf{a}\|_\infty := \sup_{t\geq 0} \sum_{i=1}^{K} a_i(t),\ \|\mathbf{c}\|_\infty := \sup_{t\geq 0} \sum_{i=1}^{K} c_i(t),\ \|b\|_\infty := \sup_{t\geq 0} b(t),$$

and $\max\{\|\mathbf{a}\|_\infty, \|\mathbf{c}\|_\infty, \|b\|_\infty\} \leq 1/2$.

(c) $\lim_{t\to\infty} b(t) = 1/2$.

Proof: Part(a) follows from Theorem 4.4.1 and the definition of the functions. For (b) observe that

$$\sum_{i=1}^{K} a_i(t) = \sum_{i=1}^{K} F_i^a(\mathbf{x}(t)) \leq \frac{1}{2} \sum_{\vec{j}\in\Omega_K} x_{j_1}(t)x_{j_2}(t)x_{j_3}(t)x_{j_4}(t) = \frac{1}{2}\left[\sum_{i\in\Omega_K} x_i(t)\right]^4 = \frac{1}{2}.$$

Statements on $\|c\|_\infty, \|b\|_\infty$ follow similarly.

For (c), note that $F^x_{\varpi,\varpi}(\mathbf{x}) \geq x_\varpi^4/2$ since when all the four vertices selected are from components of size greater than K, two components of size greater than K will surely be linked. From Theorem 4.4.1 $\lim_{t\to\infty} x_\varpi(t) = 1$ and thus $\limsup_{t\to\infty} b(t) \geq x_\varpi^2(t)/2$. The result now follows on combining this with (b). ■

4.4.3 Connection to inhomogeneous random graphs

In this section, we describe the inhomogeneous random graph (IRG) models that have been studied extensively in [11], and then approximate $\mathbf{BSR}^*(t)$ by a special class of such models. We will in fact use a variation of the models in [11] which uses a suitable weight function to measure the volume of a component. We begin by defining the basic ingredients in this model. Let $\mathcal{X}$ be a Polish space, equipped with the Borel σ-field $\mathcal{B}(\mathcal{X})$. We shall sometimes refer to this as the **type space**. Let μ be a non-atomic finite measure on $\mathcal{X}$ which we shall call the **type measure** on $\mathcal{X}$. A **kernel** will be a symmetric non-negative product measurable function $\kappa : \mathcal{X}\times\mathcal{X} \to \mathbb{R}$ and a **weight function** $\phi : \mathcal{X} \to \mathbb{R}$ will be a non-negative measurable function on $\mathcal{X}$. We call such a quadruple $\{\mathcal{X}, \mu, \kappa, \phi\}$ a **basic structure**.

The inhomogeneous random graph with weight function (IRG): Associated with a basic structure $\{\mathcal{X}, \mu, \kappa, \phi\}$, the IRG model $\mathbf{RG}^{(n)}(\mathcal{X}, \mu, \kappa, \phi)$ is a random graph described as follows:

(a) **Vertices:** the vertex set $\mathcal{V}$ of this random graph is a Poisson point process on the space $\mathcal{X}$ with intensity measure $n\mu$.

(b) **Edges:** an edge is added between vertices $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ with probability $1 \wedge \frac{\kappa(\mathbf{x},\mathbf{y})}{n}$, independent across different pairs. This defines the random graph.

(c) **Volume:** The volume of a component $\mathcal{C}$ of $\mathbf{RG}^{(n)}(\mathcal{X}, \mu, \kappa, \phi)$ is defined as

$$\mathrm{vol}_\phi(\mathcal{C}) := \sum_{x\in\mathcal{C}} \phi(x). \tag{4.4.17}$$

For the rest of this section we take

$$\mathcal{X} := [0, \infty) \times \mathcal{W} \text{ where } \mathcal{W} := \mathcal{D}([0, \infty) : \mathbb{N}_0). \tag{4.4.18}$$

We first describe how, for each $t > 0$, $\mathbf{BSR}^*(t)$ can be identified with a random graph with vertex set in $\mathcal{X}$. Recall the three types of events governing the evolution of $\mathbf{BSR}^*(t)$, described in Section 4.4.2. Each component in $\mathbf{BSR}^*(t)$ contains at least one group of $K + i$ vertices, $i = 1, \cdots, K$ which appeared at instant $s \le t$ in

$\mathbf{BSR}^*(\cdot)$, as an immigrant. We denote the collection of all such groups as $\mathrm{Imm}(t)$. For $\mathcal{C} \in \mathrm{Imm}(t)$, we denote by $\tau_{\mathcal{C}} \in (0, t]$ the instant this immigrant appears. Also, to each $\mathcal{C} \in \mathrm{Imm}(t)$, we associate a path in $\mathcal{D}([0, \infty) : \mathbb{N}_0)$, denoted as $w_{\mathcal{C}}$, such that $w_{\mathcal{C}}(s) = 0$ for all $s < \tau_{\mathcal{C}}$; $w_{\mathcal{C}}(s) = w_{\mathcal{C}}(t)$ for all $s \in [t, \infty)$; and for $s \in [\tau_{\mathcal{C}}, t]$, $w_{\mathcal{C}}(s) = |\mathcal{A}_{\mathcal{C}}(s)|$, where $\mathcal{A}_{\mathcal{C}}(s)$ denotes the component that is formed by $\mathcal{C}$ and all the attachment components that link to $\mathcal{C}$ over the time interval $[\tau_{\mathcal{C}}, s]$. Then $\{(\tau_{\mathcal{C}}, w_{\mathcal{C}}) : \mathcal{C} \in \mathrm{Imm}(t)\}$ is a point process on $\mathcal{X}$ and forms the vertex set of a random graph which we denote by $\Gamma(t)$. We form edges between any two vertices $(\tau_{\mathcal{C}}, w_{\mathcal{C}})$, $(\tau'_{\mathcal{C}}, w'_{\mathcal{C}})$ in $\Gamma(t)$ if the components $\mathcal{A}_{\mathcal{C}}(t)$ and $\mathcal{A}_{\mathcal{C}'}(t)$ are directly linked by some edge in $\mathbf{BSR}^*(t)$.

Define, for $t > 0$, the weight function $\phi_t : \mathcal{X} \to [0, \infty)$ as

$$\phi_t(\mathbf{x}) = \phi_t(s, w) = w(t), \qquad \mathbf{x} = (s, w) \in \mathcal{X}. \tag{4.4.19}$$

Note that by construction there is a one to one correspondence between the components in $\mathbf{BSR}^*(t)$ and the components in $\Gamma(t)$. For a component $\mathcal{C}_0$ in $\mathbf{BSR}^*(t)$, denote by $I_{\mathcal{C}_0}$ the corresponding component in $\Gamma(t)$. Then note that

$$|\mathcal{C}_0| = \mathrm{vol}_{\phi_t}(I_{\mathcal{C}_0}). \tag{4.4.20}$$

We will now describe inhomogeneous random graph models that approximate $\Gamma(t)$ (and hence $\mathbf{BSR}^*(t)$). Given a set of nonnegative continuous bounded functions $\boldsymbol{\alpha} = \{\alpha_i\}_{1 \le K}$, β and $\boldsymbol{\gamma} = \{\gamma_i\}_{1 \le i \le K}$ on $[0, \infty)$ we construct, for each $t > 0$, type measures $\mu_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ and kernels $\kappa_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ on $\mathcal{X}$ as follows. For $i = 1, \cdots K$ and $s \in [0, \infty)$, denote by $\tilde{\nu}_{s,i}$ the probability law on $\mathcal{D}([s, \infty) : \mathbb{N}_0)$ of the Markov process $\{\tilde{w}(r)\}_{r \in [s, \infty)}$ with infinitesimal generator

$$(\mathcal{A}(u)f)(k) = \sum_{j=1}^{K} k\gamma_j(u)(f(k+j) - f(k)), \; f \in \mathrm{BM}(\mathbb{N}_0) \tag{4.4.21}$$

and initial condition $\tilde{w}(s) = K + i$. In words, this is a pure jump Markov process which starts at time s at state $K + i$ and then at any time instant $u > s$, has jumps of size j at rate $\gamma_j(u)$. Denote by $\nu_{s,i}$ the probability law on $\mathcal{D}([0,\infty):\mathbb{N}_0)$ of the stochastic process $\{w(r)\}_{r\in[0,\infty)}$, defined as

$$w(r) = \tilde{w}(r) \quad \text{for } r \geq s\ , \qquad w(r) = 0 \text{ otherwise.} \tag{4.4.22}$$

Now define the finite measure $\mu_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) \equiv \mu_t$ as

$$\int_{\mathcal{X}} f(\mathbf{x}) d\mu_t(\mathbf{x}) = \sum_{i=1}^{K} \int_0^t \alpha_i(u) \left(\int_{\mathcal{W}} f(u,w)\nu_{u,i}(dw) \right) du,\ f \in \mathrm{BM}(\mathcal{X}). \tag{4.4.23}$$

Next, define the kernel $\kappa_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) \equiv \kappa_t$ on $\mathcal{X} \times \mathcal{X}$ as

$$\kappa_t(\mathbf{x}, \mathbf{y}) = \kappa_t((s,w),(r,\tilde{w})) = \int_0^t w(u)\tilde{w}(u)\beta(u)du,\ \mathbf{x} = (s,w), \mathbf{y} = (r,\tilde{w}) \in \mathcal{X}. \tag{4.4.24}$$

With the above choice of μ_t, κ_t and with weight function ϕ_t as in (4.4.19) we now construct the random graph $\mathbf{RG}^{(n)}(\mathcal{X}, \mu_t, \kappa_t, \phi_t)$ which we denote by $\mathbf{RG}^{(n)}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})(t)$. We will refer to the set of functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ as above, as **rate functions**. These rate functions will typically arise as small perturbations of the functions $(\mathbf{a}, b, \mathbf{c})$, thus in view of Proposition 4.4.3(b) it will suffice to consider $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ such that $\max\{||\boldsymbol{\alpha}||_\infty, ||\beta||_\infty, ||\boldsymbol{\gamma}||_\infty\} < 1$. Throughout this chapter we will assume that all rate functions (and their perturbations) satisfy this bound.

The following key result says that for large n, $\Gamma(t)$ is suitably close to $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})(t)$, where $(\mathbf{a}, b, \mathbf{c})$ are the rate functions introduced below (4.4.16). In order to state the result precisely, we extend the notion of a "subgraph" to the setting with type space $\mathcal{X}$ and weight function ϕ. For $i = 1, 2$, consider graphs $\mathbf{G}_i$, with finite vertex set $\mathcal{V}_i \subset \mathcal{X}$ and edge set $\mathcal{E}_i$. We say $\mathbf{G}_1$ is a subgraph of $\mathbf{G}_2$, and write $\mathbf{G}_1 \subset \mathbf{G}_2$, if there exists a one to one mapping $\Psi : \mathcal{V}_1 \to \mathcal{V}_2$ such that

(i) $\phi(\mathbf{x}) \leq \phi(\Psi(\mathbf{x}))$, for all $\mathbf{x} \in \mathcal{V}_1$.

(ii) $\{\mathbf{x}_1, \mathbf{x}_2\} \in \mathcal{E}_1$ implies $\{\Psi(\mathbf{x}_1), \Psi(\mathbf{x}_2)\} \in \mathcal{E}_2$.

Lemma 4.4.4. Fix $\delta \in (0, 1/2)$ and let $\varepsilon_n = n^{-\delta}$, $n \geq 1$. Define, for $t > 0$, the set of functions $\mathbf{a}^-(t) := \{(a_i(t) - \varepsilon_n) \vee 0\}_{1 \leq i \leq K}$, $\mathbf{a}^+(t) := \{a_i(t) + \varepsilon_n\}_{1 \leq i \leq K}$ and similarly $\mathbf{c}^-(t), \mathbf{c}^+(t)$ and $b^-(t), b^+(t)$. Define the inhomogeneous random graphs (IRG) with the above rate functions by

$$\mathbf{RG}^-(t) := \mathbf{RG}(\mathbf{a}^-, b^-, \mathbf{c}^-)(t), \ \ \mathbf{RG}^+(t) := \mathbf{RG}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t).$$

Then for every $T > 0$ there exist $C_3, C_4 \in (0, \infty)$, such that for all $t \in [0, T]$, there is a coupling of $\mathbf{RG}^-(t)$, $\mathbf{RG}^+(t)$ and $\Gamma(t)$ such that,

$$\mathbb{P}\left\{\mathbf{RG}^-(t) \subset \Gamma(t) \subset \mathbf{RG}^+(t)\right\} > 1 - C_3 \exp\left\{-C_4 n^{1-2\delta}\right\}.$$

Proof: The coupling between the three graphs is done in a manner such that $\Gamma(t)$ is obtained by a suitable thinning of vertices and edges in $\mathbf{RG}^+(t)$ and $\mathbf{RG}^-(t)$ is obtained by a thinning of $\Gamma(t)$. We will only provide details of the first thinning step. We first construct the vertex sets $\mathcal{V}^+$ and $\mathcal{V}^*$ in $\mathbf{RG}^+(t)$ and $\Gamma(t)$ respectively.

Let $\mathcal{V}^+$ be a Poisson point process on $\mathcal{X}$ with intensity $n\mu_t^+$, where $\mu_t^+ := \mu_t(\mathbf{a}^+, b^+, \mathbf{c}^+)$. For a fixed realization of $\mathcal{V}^+$, denote by $(x_1^+, \cdots, x_N^+)$, the points in $\mathcal{V}^+$, with $x_i^+ = (s_i^+, w_i^+)$ and $0 < s_1^+ < s_2^+ \cdots s_N^+ < t$. Write $\mathbf{w}^+ = (w_1^+, \cdots w_N^+)$. We now construct vertices in the corresponding realization of $\Gamma(t)$ (denoted as $\{x_1, \cdots x_{N_0}\}$), along with the realizations of $\bar{x}_i(s)$, $i \in \Omega_K$, $0 \leq s \leq t$, which then defines

$$(a_j^*(s), b^*(s), c_j^*(s)) \text{ for } 0 \leq s \leq t, j = 1, \cdots, K,$$

as functions of $\bar{\mathbf{x}}(s) = (\bar{x}_i(s))_{i \in \Omega_K}$ in Section 4.4.2. For that, we will construct functions $w_j : [0, t] \to \mathbb{N}_0$, $1 \leq j \leq N$ and $\bar{x}_i : [0, t] \to [0, 1]$, $i \in \Omega_K$. We will only describe the construction of $w_j, \bar{x}_i$ until the first time instant $s \in (0, t]$, when the property

$$a_j^*(s) \leq a_j^+(s), \ b^*(s) \leq b^+(s), \ c_j^*(s) \leq c_j^+(s) \text{ for all } 1 \leq k \leq K \tag{4.4.25}$$

is violated. Denote σ for the first time that (4.4.25) is violated with σ taken to be t if the property holds for all $s \in [0,t]$. Subsequent to that time instant the construction can be done in any fashion that yields the correct probability law for $\Gamma(t)$. For simplicity, we assume henceforth that $\sigma = t$. After obtaining the functions $w_j, \bar{x}_i$, we set $x_i^* = (\tau_i^*, w_i^*)$, where τ_i^* is the first jump instant of w_i (taken to be $+\infty$ if there are no jumps) and $w_i^* \in \mathcal{D}([0,\infty) : \mathbb{N}_0)$ is defined as $w_i^*(s) = w_i(s)1_{[0,t]}(s) + w_i(t)1_{(t,\infty)}(s)$. The vertex set $\mathcal{V}^*$ is then defined as

$$\mathcal{V}^* = \{x_1, \cdots x_{N_0}\} = \{x_i^* : \tau_i^* < t, i = 1, \cdots N\}.$$

We now give the construction of $\mathbf{w}(s) = (w_1(s), \cdots w_N(s))$ and $\bar{\mathbf{x}}(s)$ for $s \leq t$. Denote by $\{t_i\}_{i=1}^M$, $0 = t_0 < t_1 < t_2 < ...t_M < t$, the collection of all time instants of jumps of $\{w_i^+\}_{i=1}^N$ before time t. Denote by i_k the coordinate of $\mathbf{w}^+$ that has a jump at time t_k, and denote the corresponding jump size by j_k. We set $\mathbf{w}(0) = 0$, $\bar{x}_i(0) = 0$ for $i \neq 1$ and $\bar{x}_1(0) = 1$. The construction proceeds recursively over the time intervals $(t_{k-1}, t_k]$, $k = 1, \cdots M+1$, where $t_{M+1} = t$. Suppose that $(\mathbf{w}(s), \bar{\mathbf{x}}(s))$ have been defined for $s \in [0, t_{k-1}]$, for some $k \geq 1$. We now define these functions over the interval $(t_{k-1}, t_k]$.

Step 1: $s \in (t_{k-1}, t_k)$. Set $\mathbf{w}(s) = \mathbf{w}(t_{k-1})$. The values of $\bar{\mathbf{x}}(s)$ over the interval will be given as a realization of a jump process, for which jumps at time instant s occur at rates $n \cdot R_{i_1,i_2}(s)$, $i_1, i_2 \in \{1, \cdots K\}$, $i_1 + i_2 \leq K$, where the function $R_{i_1,i_2}(s)$, given as a function of $\bar{\mathbf{x}}(s)$ is defined as in Section 4.4.2. A jump at time instant s, corresponding to the pair (i_1, i_2) as above, changes the values of $\bar{\mathbf{x}}$ as: when $i_1 \neq i_2$

$$\bar{x}_{i_1}(s) = \bar{x}_{i_1}(s-) - \frac{i_1}{n},\ \bar{x}_{i_2}(s) = \bar{x}_{i_2}(s-) - \frac{i_2}{n},\ \bar{x}_{i_1+i_2}(s) = \bar{x}_{i_1+i_2}(s-) + \frac{i_1+i_2}{n},$$

and when $i_1 = i_2$,

$$\bar{x}_{i_1}(s) = \bar{x}_{i_1}(s-) - \frac{2i_1}{n},\ \bar{x}_{i_1+i_2}(s) = \bar{x}_{2i_1}(s-) + \frac{2i_1}{n}.$$

Remaining $\bar{x}_i$ stay unchanged. The values of $a_i^*(s)$, $b^*(s)$, $c_i^*(s)$ are determined accordingly.

Step 2: $s = t_k$. Recall that $w_{i_k}^+(t_k) - w_{i_k}^+(t_k-) = j_k$. We define $w_i(t_k) = w_i(t_k-)$ for all $i \neq i_k$. The values of $w_{i_k}(t_k)$ and $\bar{\mathbf{x}}(t_k)$ are determined as follows.

Case 1: $w_{i_k}^+(t_k-) = 0$. In this case $K+1 \leq j_k \leq 2K$ and t_k is the first jump of $w_{i_k}^+$. Define for $1 \leq l \leq K$, $Q_k(l) := \sum_{i=1}^{l} R_{i,j_k-i}(t_k-)$, where by definition $R_{i,i'} = 0$ if $i' > K$. Note that $Q_k(K) = a_{j_k}^*(t_k-)$. We set $Q_k(0) = 0$. The values of $w_{i_k}(t_k)$ and $\bar{\mathbf{x}}(t_k)$ are now determined according to the realization of an independent Uniform $[0,1]$ random variable u_k as follows.

- If $u_k > Q_k(K)/a_{j_k}(t_k-)$, define $(w_{i_k}(t_k), \bar{\mathbf{x}}(t_k)) = (w_{i_k}(t_k-), \bar{\mathbf{x}}(t_k-))$.
- Otherwise, suppose $1 \leq l_k \leq K$ is such that $Q_k(l_k - 1) < u_k \leq Q_k(l_k)$. Then define $w_{i_k}(t_k) = j_k$, $\bar{x}_\varpi(t_k) = \bar{x}_\varpi(t_k-) + \frac{j_k}{n}$ and

$$\bar{x}_{l_k}(t_k) = \bar{x}_{l_k}(t_k-) - \frac{l_k}{n}, \;\; \bar{x}_{j_k-l_k}(t_k) = \bar{x}_{j_k-l_k}(t_k-) - \frac{j_k - l_k}{n}, \quad \text{if } l_k \neq j_k - l_k,$$
$$\bar{x}_{l_k}(t_k) = \bar{x}_{l_k}(t_k-) - \frac{2l_k}{n}, \qquad \text{if } l_k = j_k - l_k.$$

  The value of all other x_i processes at t_k stay the same as their values at t_k-.

Case 2: $w_{i_k}^+(t_k-) \neq 0$. In this case $1 \leq j_k \leq K$. Once again, the values of $w_{i_k}(t_k)$ and $\bar{\mathbf{x}}(t_k)$ are determined according to the realization of an independent Uniform $[0,1]$ random variable u_k as follows.

- If $u_k > \frac{w_{i_k}(t_k-)c_{j_k}^*(t_k-)}{w_{i_k}^+(t_k-)c_{j_k}^+(t_k-)}$, define $(w_{i_k}(t_k), \bar{\mathbf{x}}(t_k)) = (w_{i_k}(t_k-), \bar{\mathbf{x}}(t_k-))$.
- Otherwise,

$$w_{i_k}(t_k) = w_{i_k}(t_k-) + j_k, \;\; \bar{x}_{j_k}(t_k) = \bar{x}_{j_k}(t_k-) - \frac{j_k}{n}, \;\; \bar{x}_\varpi(t_k) = \bar{x}_\varpi(t_k-) + \frac{j_k}{n},$$

  and the value of all other x_i processes stay the same as their value at t_k-.

This completes the construction of $(\mathbf{w}(s), \bar{\mathbf{x}}(s))$ for $s \in (t_{k-1}, t_k]$ and thus by this recursive procedure and our earlier discussion we obtain the vertex set

$$\mathcal{V}^* = \{x_1, \cdots x_{N_0}\} = \{x_i^* : \tau_i^* < t, i = 1, \cdots N\},$$

which will be used to construct $\Gamma(t)$.

Having constructed vertex sets $\mathcal{V}^+$ and $\mathcal{V}^*$, we now construct edges. For this we take realizations of independent Uniform $[0,1]$ random variables $\{u_{i,j}\}_{1 \le i < j < \infty}$ and construct edge between vertices x_i^+ and x_j^+ in $\mathcal{V}^+$ if

$$u_{i,j} \le \frac{1}{n} \int_0^t b^+(s) w_i^+(s) w_j^+(s) ds.$$

This completes the construction of $\mathbf{RG}^+(t)$. Finally construct an edge between x_i^* and x_j^* if both vertices are in $\mathcal{V}^*$ and

$$u_{i,j} \le 1 - \exp\left\{ -\frac{1}{n} \int_0^t b^*(s) w_i(s) w_j(s) ds \right\}.$$

This completes the construction of $\Gamma(t)$. By construction $\Gamma(t) \subset \mathbf{RG}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)$ on the set $\sigma = t$. Also, from (4.4.12), (4.4.16) and (4.4.14) it follows that $\mathbb{P}(\sigma < t) \le C_3 \exp\left\{-C_4 n^{1-2\delta}\right\}$ for suitable constant C_3, C_4. The result follows. ■

The following is an immediate corollary of Lemma 4.4.4.

Corollary 4.4.5. Fix $T > 0$. Then with $C_3, C_4 \in (0, \infty)$ and , for $t \in [0, T]$, a coupling of $\mathbf{RG}^-(t)$, $\mathbf{RG}^+(t)$ and $\Gamma(t)$ as in Lemma 4.4.4:

$$\begin{aligned} &\mathbb{P}\left\{ \mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^-}(t)) \le \mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\Gamma}(t)) \le \mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^+}(t)) \right\} \\ &\quad \ge 1 - C_3 \exp(-C_4 n^{1-2\delta}), \end{aligned} \tag{4.4.26}$$

where $\mathcal{I}_1^{\Gamma}(t)$ denotes the component in $\Gamma(t)$ with the largest volume with respect to the weight function ϕ_t, and $\mathcal{I}_1^{\mathbf{RG}^-}(t)$, $\mathcal{I}_1^{\mathbf{RG}^+}(t)$ are defined similarly.

4.5 Proof of the main results

In this section, we will complete the proof of Theorems 4.2.2 and 4.2.3. Proof of Theorem 4.2.3 is given in Section 4.5.4 while proof of Theorem 4.2.2 is given in Section 4.5.5. Recall that Lemma 4.4.4 says that $\mathbf{BSR}^*$ can be approximated by $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})$. Sections 4.5.2 and 4.5.3 analyze properties of integral operators associated with $\mathbf{RG}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ for a general family of rate functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. We begin in Section 4.5.1 by presenting some results about an IRG model $\mathbf{RG}^{(n)}(\mathcal{X}, \mu, \kappa, \phi)$ on a general type space $\mathcal{X}$.

4.5.1 Preliminary lemmas

In this section, we collect some results about the general inhomogeneous random graph model $\mathbf{RG}^{(n)}(\mathcal{X}, \mu, \kappa, \phi)$. Let $\mathcal{K}$ be the integral operator associated with (κ, μ), as defined in Section 4.1. Recall that the operator norm of $\mathcal{K}$, denoted as $\|\mathcal{K}\|$, is defined as

$$\|\mathcal{K}\| = \sup_{f \in L^2(\mathcal{X},\mu), f \neq 0} \frac{\|\mathcal{K}f\|_2}{\|f\|_2}, \tag{4.5.1}$$

where for $f \in L^2(\mathcal{X}, \mu)$, $\|f\|_2 = \left(\int_{\mathcal{X}} |f(\mathbf{x})|^2 \mu(d\mathbf{x})\right)^{1/2}$.

Lemma 4.5.1. Fix $(\mathcal{X}, \mu, \kappa, \phi)$. Denote the vertex set of $\mathbf{RG}^{(n)}(\mathcal{X}, \mu, \kappa, \phi) \equiv \mathbf{RG}^{(n)}$ by $\mathcal{P}_n$ which is a rate $n\mu$ Poisson point process on $\mathcal{X}$. Let $\mathcal{K}$ be the integral operator associated with (κ, μ). Suppose that $\|\mathcal{K}\| < 1$ and let $\Delta = 1 - \|\mathcal{K}\|$. Denote by $\mathcal{I}_1^{\mathbf{RG}}$ the component in $\mathbf{RG}^{(n)}$ with the largest volume (with respect to the weight function ϕ).

Then the following hold.

(i) If $\|\phi\|_\infty < \infty$ and $\|\kappa\|_\infty < \infty$, then for all $m \in \mathbb{N}$ and $D \in (0, \infty)$

$$\mathbb{P}\{\mathrm{vol}_\phi(\mathcal{I}_1^{\mathbf{RG}}) > m\} \leq 2nD \exp\{-C\Delta^2 m\} + \mathbb{P}(|\mathcal{P}_n| \geq nD),$$

where $C = (8\|\phi\|_\infty(1+3\|\kappa\|_\infty\mu(\mathcal{X})))^{-1}$.

(ii) Let for $n \geq 1$, $\Lambda_n \in \mathcal{B}(\mathcal{X})$ be such that

$$g(n) := 8 \sup_{\mathbf{x}\in\Lambda_n} |\phi(\mathbf{x})| \left(1 + 3\mu(\mathcal{X}) \sup_{(\mathbf{x},\mathbf{y})\in\Lambda_n\times\Lambda_n} |\kappa(\mathbf{x},\mathbf{y})|\right) < \infty.$$

Then for all $m \in \mathbb{N}$,

$$\mathbb{P}\{\mathrm{vol}_\phi(\mathcal{I}_1^{\mathbf{RG}}) > m\} \leq n\mu(\Lambda_n^c) + 2nD\exp(-\Delta^2 m/g(n)) + \mathbb{P}(|\mathcal{P}_n| \geq nD).$$

Proof: Part (i) has been proved in Chapter 3. We now prove (ii). Consider the truncated version of $\mathbf{RG}^{(n)}$ constructed using the basic structure $\{\mathcal{X}, \bar{\mu}, \bar{\kappa}, \bar{\phi}\}$ where $\bar{\mu} := \mu|_{\Lambda_n}$ (i.e. the restriction of μ to Λ_n), $\bar{\kappa}(\mathbf{x},\mathbf{y}) = \kappa(\mathbf{x},\mathbf{y})\mathbb{1}_{\Lambda_n}(\mathbf{x})\mathbb{1}_{\Lambda_n}(\mathbf{y})$ and $\bar{\phi}(\mathbf{x}) = \phi(\mathbf{x})\mathbb{1}_{\Lambda_n}(\mathbf{x})$. Note that $||\bar{\kappa}||_\infty < \infty$ and $||\bar{\phi}||_\infty < \infty$. Denote by $\bar{\mathcal{K}}$ the integral operator associated with $(\bar{\kappa}, \bar{\mu})$. Clearly $\|\bar{\mathcal{K}}\| \leq \|\mathcal{K}\|$ and thus $\bar{\Delta} = 1 - \|\bar{\mathcal{K}}\| \geq \Delta$. Consider the natural coupling between the truncated and original model by using the vertex set $\bar{\mathcal{P}}_n := \mathcal{P}_n \cap \Lambda_n$. Write $\bar{\mathcal{I}}_1^{\mathbf{RG}}$ for the component with the largest volume in the truncated model. Then we have

$$\begin{aligned}
\mathbb{P}\{\mathrm{vol}_\phi(\mathcal{I}_1^{\mathbf{RG}}) > m\} \leq& \mathbb{P}\{\mathcal{P}_n \cap \Lambda_n^c \neq \emptyset\} + \mathbb{P}\{\mathcal{P}_n \subset \Lambda_n, \mathrm{vol}_\phi(\mathcal{I}_1^{\mathbf{RG}}) > m\} \\
=& \mathbb{P}\{\mathcal{P}_n \cap \Lambda_n^c \neq \emptyset\} + \mathbb{P}\{\mathrm{vol}_\phi(\bar{\mathcal{I}}_1^{\mathbf{RG}}) > m\} \\
\leq& (1 - \exp\{-n\mu(\Lambda_n^c)\}) + 2nD\exp\{-\Delta^2 m/g(n)\} + \mathbb{P}(|\mathcal{P}_n| \geq nD),
\end{aligned}$$

where the last inequality follows from part (i) and the fact that $\Delta \leq \bar{\Delta}$. ■

We state the following two elementary lemmas, whose proof is omitted.

Lemma 4.5.2. Let κ, κ' be kernels on a common finite measure space $(\mathcal{X}, \mu)$, with the associated integral operators $\mathcal{K}, \mathcal{K}'$ respectively. Then

(a) $\|\mathcal{K}\| \leq \|\kappa\|_{2,\mu} := \left(\int_{\mathcal{X}\times\mathcal{X}} \kappa^2(\mathbf{x},\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y})\right)^{1/2}$.

(b) If $\kappa \leq \kappa'$, then $||\mathcal{K}|| \leq ||\mathcal{K}'||$.

(c) $||\mathcal{K}-\mathcal{K}'|| \leq ||\bar{\mathcal{K}}||$, where $\bar{\mathcal{K}}$ is the integral operator associated with $(|\kappa-\kappa'|,\mu)$.

Lemma 4.5.3. Let $\tilde{\mu},\mu$ be two finite measures on the space $\mathcal{X}$. Assume $\tilde{\mu} \ll \mu$ and let $g = d\tilde{\mu}/d\mu$ be the Radon-Nikodym derivative. Let $\tilde{\kappa}$ be a kernel on $\mathcal{X}\times\mathcal{X}$, and define κ as

$$\kappa(\mathbf{x},\mathbf{y}) := \sqrt{g(\mathbf{x})g(\mathbf{y})}\tilde{\kappa}(\mathbf{x},\mathbf{y}), \ \ \mathbf{x},\mathbf{y} \in \mathcal{X}.$$

Denote by $\mathcal{K}$ [resp. $\tilde{\mathcal{K}}$] the integral operator on $L^2(\mathcal{X},\mu)$ [resp. $L^2(\mathcal{X},\tilde{\mu})$] associated with (κ,μ) [resp. $(\tilde{\kappa},\tilde{\mu})$]. Then $\|\mathcal{K}\|_{L^2(\mu)} = \|\tilde{\mathcal{K}}\|_{L^2(\tilde{\mu})}$, where $\|\mathcal{K}\|_{L^2(\mu)}$ [resp. $\|\tilde{\mathcal{K}}\|_{L^2(\tilde{\mu})}$] is the norm of the operator $\mathcal{K}$ [resp. $\tilde{\mathcal{K}}$] on $L^2(\mu)$ [resp. $L^2(\tilde{\mu})$].

We end this section with a lemma drawing a connection between the Yule process and the pure jump Markov processes with distribution $\nu_{s,i}$ that arose in the construction of the inhomogeneous random graphs $\mathbf{RG}(\boldsymbol{\alpha},\beta,\boldsymbol{\gamma})$, see (4.4.22). Fix $j \geq 1$ and recall that a rate one Yule process started at time $t=0$ with j individuals is a pure birth Markov process $Y(t)$ with $Y(0)=j$ and the rate of going from state i to $i+1$ given by $\lambda(i) := i$. Also recall from (4.4.18) that $\mathcal{W}$ denotes the Skorohod space $\mathcal{W} := \mathcal{D}([0,\infty):\mathbb{N}_0)$.

Lemma 4.5.4. Fix $1 \leq i \leq K$ and $s \geq 0$ and rate functions $\boldsymbol{\alpha},\beta,\boldsymbol{\gamma}$. Let $\{w(t)\}_{t\geq 0}$ be a pure jump Markov process with law $\nu_{s,i} := \nu_{s,i}(\boldsymbol{\alpha},\beta,\boldsymbol{\gamma})$ as in (4.4.22). Then

(i) The process $w^*(t) := w(t/K\|\boldsymbol{\gamma}\|_\infty)/K$ can be stochastically dominated by a Yule processes $Y(\cdot)$ starting with two particles (i.e. $Y(0)=2$).

(ii) Fix $t>0$, $s\in[0,t]$ and $1\leq i\leq K$. Then we have

$$\int_{\mathcal{W}} [w(t)]^2 \nu_{s,i}(dw) \leq 6K^2 e^{2tK\|\boldsymbol{\gamma}\|_\infty},$$

and for any $A>0$ we have

$$\nu_{s,i}(w(t)>A) \leq 2(1-e^{-tK\|\boldsymbol{\gamma}\|_\infty})^{A/2K}.$$

Proof: Let us first prove (i). Note that under the law $\nu_{s,i}$, the process w satisfies $w(u) = 0$ for $u < s$ and $w(s) = K + i \leq 2K$. Further for times $t > s$, by (4.4.21), the jumps of the w can be bounded as $\Delta w(t) := w(t) - w(t-) \leq K$ at rate at most $w(t)\|\boldsymbol{\gamma}\|_\infty$. The process $w^*(\cdot)$ is obtained by rescaling time and space for the process $w(\cdot)$. This is once again a pure jump Markov process with jump sizes $\Delta w^*(t) \leq 1$ which happen at rate at most one. Further $w^*(0) \leq 2$. This immediately implies that this process is stochastically dominated by a Yule process with $Y(0) = 2$. This completes the proof.

We now consider (ii). We will use the result in part (i). Note that a Yule process started with two individuals at time $t = 0$ has the same distribution as the sum of two independent Yule processes $\{Y_1(t)\}_{t\geq 0}$ and $\{Y_2(t)\}_{t\geq 0}$ with $Y_1(0) = Y_2(0) = 1$. Now fix $t > 0$, $s \leq t$ and $1 \leq i \leq K$. Let $w(\cdot)$ have distribution $\nu_{s,i}$. From (i) we have

$$w(t) \leq_d K \cdot (Y_1(tK\|\boldsymbol{\gamma}\|_\infty) + Y_2(tK\|\boldsymbol{\gamma}\|_\infty)). \tag{4.5.2}$$

For simplicity write $X_1 = Y_1(tK\|\boldsymbol{\gamma}\|_\infty)$ and $X_2 = Y_2(tK\|\boldsymbol{\gamma}\|_\infty)$. Well known results about Yule processes ([27, Chapter 2]) say that the random variables X_1 and X_2 have a Geometric distribution with $p := e^{-tK\|\boldsymbol{\gamma}\|_\infty}$. The first bound in (ii) follows from the Geometric distribution and the fact

$$\int_{\mathcal{W}} [w(t)]^2 \nu_{s,i}(dw) \leq K^2 \mathbb{E}[(X_1 + X_2)^2].$$

The second bound in (ii) follows from

$$\nu_{s,i}(\{w(t) > A\}) \leq 2\mathbb{P}\{X_1 > A/2K\}.$$

This completes the proof. ■

4.5.2 Some perturbation estimates for $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})$.

Recall that Lemma 4.4.4 coupled the evolution of $\Gamma(t)$ (equivalently $\mathbf{BSR}^*(t)$) with two inhomogeneous random graphs $\mathbf{RG}(\mathbf{a}^+, b, \mathbf{c}^+)(t)$ and $\mathbf{RG}(\mathbf{a}^-, b, \mathbf{c}^-)(t)$ which can be considered as perturbations of $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})(t)$. The aim of this section is to understand the effect of such perturbations on the associated operator norms. Throughout this section $\mathcal{X}$ and ϕ_t are as in (4.4.18) and (4.4.19), respectively. Given the basic structure $\{\mathcal{X}, \mu_t, \kappa_t, \phi_t\}$, $t > 0$, associated with rate functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$, we denote the norm of the integral operator $\mathcal{K}_t$ associated with (κ_t, ϕ_t) as $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$.

The following proposition which is the main result of this section studies the effect of perturbations of $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ on this norm. For a K-dimensional vector $\mathbf{v} = (v_1, \cdots v_K)$ and a scalar θ, $\mathbf{v} + \theta$ denotes the vector $(v_1 + \theta, \cdots v_K + \theta)$ and $(\mathbf{v} + \theta)^+$ denotes the vector $((v_1 + \theta)^+, \cdots (v_K + \theta)^+)$.

Proposition 4.5.5. For $\varepsilon > 0$ let $\rho_t^+ = \rho_t(\boldsymbol{\alpha}+\varepsilon, \beta+\varepsilon, \boldsymbol{\gamma}+\varepsilon)$ and $\rho_t^- = \rho_t((\boldsymbol{\alpha}-\varepsilon)^+, (\beta-\varepsilon)^+, (\boldsymbol{\gamma}-\varepsilon)^+)$, where $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ are rate functions. Assume that $\max\{||\boldsymbol{\alpha}+\varepsilon||_\infty, ||\beta+\varepsilon||_\infty, ||\boldsymbol{\gamma}+\varepsilon||_\infty\} < 1$. For every $T > 0$, there is a $C_5 \in (0, \infty)$ such that for all $\varepsilon > 0$ and $t \in [0, T]$,

$$\max\{|\rho_t - \rho_t^+|, |\rho_t - \rho_t^-|\} \le C_5\sqrt{\varepsilon} \cdot (-\log \varepsilon)^2 .$$

Proof of Proposition 4.5.5 relies on Lemmas 4.5.6 – 4.5.10 below, and is given at the end of the section. We analyze the effect of perturbation of β, $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ separately in Lemmas 4.5.6, 4.5.8) and 4.5.10, respectively.

Lemma 4.5.6 (Perturbations of β). Let $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ be rate functions and β^ε be be a nonnegative function on $[0, \infty)$ with $\sup_{0 \le s < \infty} |\beta^\varepsilon(s) - \beta(s)| \le \varepsilon$. Then

$$|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}, \beta^\varepsilon, \boldsymbol{\gamma})| \le C\varepsilon,$$

where $C = 6t^2K^3\|\boldsymbol{\alpha}\|_\infty e^{2t\|\boldsymbol{\gamma}\|_\infty}$.

Proof: Let (μ_t, κ_t) be the type measure and kernel associated with $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ Note that a perturbation in β only affects the kernel κ_t and not μ_t. Recall the representation of μ_t in terms of probability measures $\{\nu_{u,i}, u \in [0,t], i = 1, \cdots, K\}$. From Lemma 4.5.4(ii)

$$\int_{\mathcal{W}} [w(t)]^2 \nu_{u,i}(dw) \leq 6K^2 e^{2tK\|\boldsymbol{\gamma}\|_\infty}, \text{ for all } u \in [0,t], i = 1, \cdots K. \tag{4.5.3}$$

Denote the kernel obtained by replacing β by β^ε in (4.4.24) by κ_t^ε. Since $\|\beta - \beta^\varepsilon\|_\infty < \varepsilon$, we have from (4.4.24) that

$$|\kappa_t(\mathbf{x}, \mathbf{y}) - \kappa_t^\varepsilon(\mathbf{x}, \mathbf{y})| \leq \varepsilon \int_0^t w(u)\tilde{w}(u)du \leq \varepsilon t w(t)\tilde{w}(t),$$

$\mu_t \otimes \mu_t$ a.e. $(\mathbf{x}, \mathbf{y}) = ((s,w),(r,\tilde{w}))$.

By Lemma 4.5.2 (a) and (c) we now have

$$\begin{aligned}
|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}, \beta^\varepsilon, \boldsymbol{\gamma})| \leq & \left(\int_{\mathcal{X}\times\mathcal{X}} |\kappa_t(\mathbf{x}, \mathbf{y}) - \kappa_t^\varepsilon(\mathbf{x}, \mathbf{y})|^2 d\mu_t(\mathbf{x}) d\mu_t(\mathbf{y}) \right)^{1/2} \\
\leq & \left(\int_{\mathcal{X}\times\mathcal{X}} (\varepsilon t w(t)\tilde{w}(t))^2 d\mu_t(\mathbf{x}) d\mu_t(\mathbf{y}) \right)^{1/2} \\
= & \varepsilon t \sum_{i=1}^K \int_0^t \alpha_i(s) \left[\int_{\mathcal{W}} [w(t)]^2 \nu_{s,i}(dw) \right] ds \\
\leq & \varepsilon t \cdot t \|\boldsymbol{\alpha}\|_\infty \cdot K \cdot 6K^2 e^{2tK\|\boldsymbol{\gamma}\|_\infty},
\end{aligned}$$

where the last inequality follows from (4.5.3). The result follows. ■

When $\boldsymbol{\alpha}$ or $\boldsymbol{\gamma}$ is perturbed, the underlying measure μ_t changes as well and thus one needs to analyze the corresponding Radon-Nikodym derivatives. This is done in the following two lemmas. We denote by $\{\mathcal{G}_s\}_{0 \leq s < \infty}$ the canonical filtration on $\mathcal{D}([0,\infty) : \mathbb{N}_0)$. In the following we follow the convention that $0/0 = 1$.

Lemma 4.5.7. Fix $\varepsilon > 0$ and let $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$, $(\tilde{\boldsymbol{\alpha}}, \tilde{\beta}, \tilde{\boldsymbol{\gamma}})$ be two sets of rate functions such that for all $1 \leq i \leq K$ and $s \geq 0$,

$$\alpha_i(s) - \varepsilon \leq \tilde{\alpha}_i(s) \leq \alpha_i(s), \text{ and } \gamma_i(s) - \varepsilon \leq \tilde{\gamma}_i(s) \leq \gamma_i(s).$$

Fix $t \geq 0$ and let μ_t and $\tilde{\mu}_t$ be the corresponding type measures on $\mathcal{X}$. For $(s, w) \in \mathcal{X}$ and $j \geq 1$ let τ_j^s for the time of the j-th jump of $w(\cdot)$ after time s (μ_t a.s.). Also write $\Delta(u) = w(u) - w(u-)$ for $u \geq 0$. Then there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$, $\tilde{\mu} \ll \mu$ and

$$
\begin{aligned}
&\frac{d\tilde{\mu}_t}{d\mu_t}(s, w) \\
=&\frac{\tilde{\alpha}_{\Delta(s)-K}(s)}{\alpha_{\Delta(s)-K}(s)} \times \Pi_{j:\tau_j^s \leq t} \frac{\tilde{\gamma}_{\Delta(\tau_j^s)}(\tau_j^s)}{\gamma_{\Delta(\tau_j^s)}(\tau_j)} \times \exp\left\{ -\int_s^t w(u) \left[\sum_{i=1}^K \tilde{\gamma}_i(u) - \sum_{i=1}^K \gamma_i(u) \right] du \right\}.
\end{aligned}
$$

Proof: For $i = 1, \cdots K$, define finite measures μ_t^i, $\tilde{\mu}_t^i$ on $\mathcal{X}$ as

$$\mu_t^i(du\, dw) = \alpha_i(u)\nu_{u,i}(dw)1_{[0,t]}(u)du, \ \tilde{\mu}_t^i(du\, dw) = \tilde{\alpha}_i(u)\tilde{\nu}_{u,i}(dw)1_{[0,t]}(u)du,$$

where $\nu_{u,i}$ is defined above (4.4.23) and $\tilde{\nu}_{u,i}$ is defined similarly on replacing γ_i with $\tilde{\gamma}_i$. We will show that

$$\text{for all } 1 \leq k \leq K \text{ and } s \in [0, T], \tilde{\nu}_{s,k} \ll \nu_{s,k} \text{ and } \frac{d\tilde{\nu}_{s,k}}{d\nu_{s,k}}(w) = L_s^t(w), \tag{4.5.4}$$

where

$$L_s^t := \Pi_{j\geq 1} \left(\frac{\tilde{\gamma}_{\Delta(\tau_j^s)}(\tau_j^s)}{\gamma_{\Delta(\tau_j^s)}(\tau_j^s)} \mathbb{1}_{\{\tau_j^s \leq t\}} \right) \times \exp\left\{ -\int_s^t w(u) \left[\sum_{i=1}^K \tilde{\gamma}_i(u) - \sum_{i=1}^K \gamma_i(u) \right] du \right\}. \tag{4.5.5}$$

The lemma is an immediate consequence of (4.5.4) on observing that μ_t^i and μ_t^j are mutually singular when $i \neq j$, and the relation $\mu_t = \sum_{i=1}^K \mu_t^i$, $\tilde{\mu}_t = \sum_{i=1}^K \tilde{\mu}_t^i$.

We now show (4.5.4). From the construction of $\nu_{s,k}$ it follows that, there are counting processes $\{N_i(u)\}_{u\in[s,t]}$, $i = 1, \cdots K$, on $\mathcal{W}$ such that

$$w(u) = w(s) + \sum_{i=1}^K iN_i(u), \text{ for } u \in [s, t], \ a.s. \ \nu_{s,k} \tag{4.5.6}$$

and

$$M_i(u) := N_i(u) - \int_s^u w(r)\gamma_i(r)dr \tag{4.5.7}$$

under $\nu_{s,k}$ is a $\{\mathcal{G}_u\}_{u\in[s,t]}$ local martingale for $u \in [s,t]$. From standard results it follows that L_s^t is a local-martingale and super-martingale (see Theorem VI.T2, p.165 of [12]). In order to show (4.5.4), it suffices to show that $\{L_s^u\}_{u\in[s,t]}$ is a martingale. By a change of variable formula it follows that (see e.g. Theorem A4.T4, p. 337 of [12])

$$L_s^v = 1 + \sum_{i=1}^K \int_s^v L_s^{u-} \cdot \left(\frac{\tilde{\gamma}_i(u)}{\gamma_i(u)} - 1\right) dM_i(u), \; v \in [s,t]. \tag{4.5.8}$$

In order to show L_s^t is a martingale, it then suffices, in view of (4.5.7), to show that (see e.g. Theorem II.T8 in [12]) for all $1 \le i \le K$,

$$\int_{\mathcal{W}} \left[\int_s^t L_s^u \cdot |\tilde{\gamma}_i(u) - \gamma_i(u)| w(u) du\right] d\nu_{s,k}(w) < \infty.$$

Finally note that $L_s^u \le e^{\varepsilon t w(t)}$. Using Lemma 4.5.4(i) and standard estimates for Yule processes, it follows that for ε sufficiently small

$$\sup_{s\in[0,t]} \sup_{1\le k\le K} \int_{\mathcal{W}} w(t) e^{\varepsilon t w(t)} d\nu_{s,k}(w) < \infty.$$

The result follows. ■

We will now use the above lemma to study the effect of perturbations in $\boldsymbol{\alpha}$ on $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$.

Lemma 4.5.8 (Perturbations of $\boldsymbol{\alpha}$). Fix $\varepsilon > 0$. Let $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ be rate functions and let $\boldsymbol{\alpha}^\varepsilon = (\alpha_1^\varepsilon, \cdots, \alpha_K^\varepsilon)$, where α_i^ε are continuous nonnegative bounded functions on $[0, \infty)$ such that for all $1 \le i \le K$ and $s \in [0,T]$

$$\alpha_i(s) - \varepsilon \le \alpha_i^\varepsilon(s) \le \alpha_i(s).$$

Then for every $t > 0$,

$$|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}^\varepsilon, \beta, \boldsymbol{\gamma})| \le C\sqrt{\varepsilon},$$

where $C = t\|\beta\|_\infty \cdot 6K^2 e^{2tK\|\boldsymbol{\gamma}\|_\infty} \cdot 4tK\sqrt{\|\boldsymbol{\alpha}\|_\infty}$.

Proof: Let (μ_t, κ_t) be the type measure and kernel associated with $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. Also, let μ_t^ε be the type measure associated with $(\boldsymbol{\alpha}^\varepsilon, \beta, \boldsymbol{\gamma})$. By Lemma 4.5.7,

$$g(s,w) := \frac{d\mu_t^\varepsilon}{d\mu_t}(s,w) = \frac{\alpha^\varepsilon_{\Delta(s)-K}(s)}{\alpha_{\Delta(s)-K}(s)} \text{ for } (s,w) \in [0,t] \times \mathcal{W}.$$

Using Lemma 4.5.2 (c), (a), Lemma 4.5.3, and the fact that $|\kappa_t(\mathbf{x}, \mathbf{y})| \leq t\|\beta\|_\infty w(t)\tilde{w}(t)$, $\mu_t \otimes \mu_t$ a.e. $(\mathbf{x}, \mathbf{y}) = ((s,w), (\tilde{s}, \tilde{w}))$, we have

$$\begin{aligned}
&|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}^\varepsilon, \beta, \boldsymbol{\gamma})| \\
\leq &\left(\int_{\mathcal{X}\times\mathcal{X}} |\sqrt{g(\mathbf{x})g(\mathbf{y})} - 1|^2 |\kappa_t(\mathbf{x}, \mathbf{y})|^2 d\mu_t(\mathbf{x}) d\mu_t(\mathbf{y}) \right)^{1/2} \\
\leq &t\|\beta\|_\infty \left(\int_{\mathcal{X}\times\mathcal{X}} |\sqrt{g(\mathbf{x})g(\mathbf{y})} - 1|^2 w^2(t)\tilde{w}^2(t) d\mu_t(\mathbf{x}) d\mu_t(\mathbf{y}) \right)^{1/2} \\
\leq &t\|\beta\|_\infty d_1 \left(\sum_{i,j=1}^K \int_{[0,t]^2} \left(\sqrt{\frac{\alpha_i^\varepsilon(s)\alpha_j^\varepsilon(u)}{\alpha_i(s)\alpha_j(u)}} - 1 \right)^2 \alpha_i(s)\alpha_j(u) ds du \right)^{1/2},
\end{aligned} \tag{4.5.9}$$

where

$$d_1 = \sup_{s\in[0,T]} \sup_{1\leq i \leq K} \int_{\mathcal{W}} |w(t)|^2 \nu_{s,i}(dw) \leq 6K^2 e^{2tK\|\gamma\|_\infty},$$

and the last inequality follows from (4.5.3). In order to bound (4.5.9), note that:

$$\begin{aligned}
&\left| \sqrt{\alpha_i(s)\alpha_j(u)} - \sqrt{\alpha_i^\varepsilon(s)\alpha_j^\varepsilon(u)} \right| \\
= &\left| \sqrt{\alpha_i(s)} \left(\sqrt{\alpha_j(u)} - \sqrt{\alpha_j^\varepsilon(u)} \right) + \left(\sqrt{\alpha_i(s)} - \sqrt{\alpha_i^\varepsilon(s)} \right) \sqrt{\alpha_j^\varepsilon(u)} \right| \\
\leq &2\sqrt{\varepsilon} \left(\sqrt{\alpha_i(s)} + \sqrt{\alpha_j(u)} \right).
\end{aligned}$$

Plugging the above bound in (4.5.9) gives the desired result. ■

We will now analyze the effect of perturbations in $\boldsymbol{\gamma}$ on $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. We need the following preliminary truncation lemma.

Lemma 4.5.9. For every $T > 0$, there exist $C_6, C_7, A_0 \in (0, \infty)$ such that for any $t \in [0,T]$ and rate functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ the following holds: Let μ_t, κ_t be the type measure and kernel associated with $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. Define, for $A \in (0, \infty)$, the kernel $\kappa_{A,t}$ as

$$\kappa_{A,t}(\mathbf{x}, \mathbf{y}) = \kappa_t(\mathbf{x}, \mathbf{y}) \mathbb{1}_{\{w(t) \le A, \tilde{w}(t) \le A\}} \text{ where } \mathbf{x} = (s, w), \mathbf{y} = (r, \tilde{w}). \tag{4.5.10}$$

Then for all $A > A_0$,

$$\rho(\kappa_t) - C_6 e^{-C_7 A} \le \rho(\kappa_{A,t}) \le \rho(\kappa_t),$$

where $\rho(\kappa_t)$ [resp. $\rho(\kappa_{A,t})$] denotes the norm of the operator associated with (κ_t, μ_t) [resp. $(\kappa_{A,t}, \mu_t)$].

Proof: The upper bound in the lemma is immediate from Lemma 4.5.2 (b). We now consider the lower bound. For the rest of the proof, we suppress the dependence of $\kappa_t, \kappa_{A,t}, \mu_t$ on t. Note that, from Lemma 4.5.2 (a,c)

$$\begin{aligned}
\rho(\kappa) - \rho(\kappa_A) \le & \left(\int_{\mathcal{X} \times \mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) - \kappa_A(\mathbf{x}, \mathbf{y}))^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}) \right)^{1/2} \\
\le & 2 \left(\int_{\mathcal{X} \times \mathcal{X}} (t \|\beta\|_\infty w(t) \tilde{w}(t) \mathbb{1}\{\tilde{w}(t) > A\})^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}) \right)^{1/2} \\
\le & 2t \|\beta\|_\infty \left(d_1 \sum_{i=1}^K \sum_{j=1}^K \int_{[0,t] \times [0,t]} \alpha_i(s) \alpha_j(u) ds du \right)^{1/2} \\
\le & 2t \|\beta\|_\infty \cdot t \|\boldsymbol{\alpha}\|_\infty \cdot \sqrt{d_1},
\end{aligned} \tag{4.5.11}$$

where

$$d_1 = \int_{\mathcal{W}} [w(t)]^2 \nu_{s,i}(dw) \int_{\mathcal{W}} [w(t)]^2 \mathbb{1}_{\{w(t) > A\}} \nu_{s,i}(dw). \tag{4.5.12}$$

By (4.5.2), $w(t) \le_d K(X_1 + X_2)$ where X_1, X_2 are independent and identically distributed with Geometric $p = e^{-tK\|\boldsymbol{\gamma}\|_\infty}$ distribution.

$$\begin{aligned}
& \int_{\mathcal{W}} [w(t)]^2 \mathbb{1}_{\{w(t) > A\}} \nu_{s,i}(dw) \\
\le & K^2 \mathbb{E} \left[(X_1 + X_2)^2 \mathbb{1}_{\{X_1 + X_2 > A/K\}} \right] \\
= & K^2 \mathbb{E} \left[(X_1 + X_2)^2 (\mathbb{1}_{\{X_1 + X_2 > C, X_1 \ge X_2\}} + \mathbb{1}_{\{X_1 + X_2 > C, X_1 < X_2\}}) \right] \\
\le & 4K^2 \mathbb{E} \left[X_1^2 \mathbb{1}_{\{X_1 > A/2K\}} + X_2^2 \mathbb{1}_{\{X_2 > A/2K\}} \right],
\end{aligned}$$

The above quantity can be bounded by

$$d_2(1 - e^{-2TK\|\boldsymbol{\gamma}\|_\infty})^{A/2K} \leq d_2 \exp\left\{-\frac{e^{-2TK\|\boldsymbol{\gamma}\|_\infty}}{K}A\right\}$$

for some constant d_2. The result now follows on using the above bound and (4.5.3) in (4.5.12) and (4.5.11). ■

Lemma 4.5.10 (Perturbations of $\boldsymbol{\gamma}$). For every $T > 0$, there exists $C_8 \in (0, \infty)$ and $\varepsilon_0 \in (0, 1)$ such that for all $t \in [0, T]$ and rate functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ the following holds: Suppose $\varepsilon \in (0, \varepsilon_0)$ and $\boldsymbol{\gamma}^\varepsilon = (\gamma_1^\varepsilon, \cdots \gamma_K^\varepsilon)$, where γ_i^ε are continuous, nonnegative maps on $[0, T]$ such that for all $1 \leq i \leq K$

$$\gamma_i(s) - \varepsilon \leq \ \gamma_i^\varepsilon(s) \leq \gamma_i(s), \text{ for all } s \in [0, T].$$

Then

$$|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^\varepsilon)| \leq C_8\sqrt{\varepsilon} \cdot (-\log \varepsilon)^2.$$

Proof: Let (μ_t, κ_t) [resp. $(\mu_t^\varepsilon, \kappa_t^\varepsilon)$] be the type measure and kernel associated with $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ [resp. $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^\varepsilon)$]. By Lemma 4.5.7, for $(s, w) \in [0, t] \times \mathcal{W}$

$$\frac{d\mu_t^\varepsilon}{d\mu_t}(s, w) = \Pi_{j\geq 1}\left(\frac{\gamma^\varepsilon_{\Delta(\tau_j^s)}(\tau_j^s)}{\gamma_{\Delta(\tau_j^s)}(\tau_j^s)}\mathbb{1}_{\{\tau_j^s \leq t\}}\right) \times \exp\left\{-\int_s^t w(u)\left[\sum_{i=1}^K \gamma_i^\varepsilon(u) - \sum_{i=1}^K \gamma_i(u)\right] du\right\}.$$

Denote the right side as $L_s^t(w)$. Then, as in the proof of Lemma 4.5.7, it follows that $\{L_s^u(w)\}_{u\in[s,t]}$ is a $\{\mathcal{G}_u\}_{u\in[s,t]}$ martingale under $\nu_{s,k}$ for every $k = 1, \cdots K$. Fix $A \in (A_0, \infty)$, where A_0 is as in Lemma 4.5.9, and let $\kappa_{A,t}$ be defined by (4.5.10). Similarly define $\kappa_{A,t}^\varepsilon$ by replacing κ_t with κ_t^ε in (4.5.10). Denote the operator norm of the integral operators associated with $(\kappa_{A,t}, \mu_t)$ and $(\kappa_{A,t}^\varepsilon, \mu_t^\varepsilon)$ by $\rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ and

$\rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^{\varepsilon})$, respectively. Then, by Lemma 4.5.3 and Lemma 4.5.2 (a,c),

$$\begin{aligned}
&|\rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^{\varepsilon})| \\
\leq& \left(\int_{\mathcal{X}\times\mathcal{X}} \left| \sqrt{\frac{d\mu_t^{\varepsilon}}{d\mu_t}(s,w)\frac{d\mu_t^{\varepsilon}}{d\mu_t}(u,\tilde{w})} - 1 \right|^2 (\kappa_{A,t}(\mathbf{x},\mathbf{y}))^2 d\mu_t(\mathbf{x})d\mu_t(\mathbf{y}) \right)^{1/2} \\
\leq& tA^2\|\beta\|_\infty \left(\sum_{i,j=1}^{K} \int_{[0,t]\times[0,t]} \alpha_i(s)\alpha_j(u) \int_{\mathcal{W}\times\mathcal{W}} \left| \sqrt{L_s^t(w)L_u^t(w)} - 1 \right|^2 \nu_{s,i}(dw)\nu_{u,j}(d\tilde{w}) \right)^{1/2}.
\end{aligned} \tag{4.5.13}$$

Next, using the martingale property of L_s^t, we have

$$\begin{aligned}
&\int_{\mathcal{W}\times\mathcal{W}} \left| \sqrt{L_s^t(w)L_u^t(w)} - 1 \right|^2 \nu_{s,i}(dw)\nu_{u,j}(d\tilde{w}) \\
=&2 - 2\int_{\mathcal{W}} \sqrt{L_s^t(w)}\nu_{s,i}(dw) \int_{\mathcal{W}} \sqrt{L_u^t(w)}\nu_{u,j}(dw) \\
\leq&4 - 2\int_{\mathcal{W}} \sqrt{L_s^t(w)}\nu_{s,i}(dw) - 2\int_{\mathcal{W}} \sqrt{L_u^t(w)}\nu_{u,j}(dw),
\end{aligned} \tag{4.5.14}$$

where the inequality on the last line follows on observing that from Jensen's inequality the two integrals on the second line are bounded by 1 and using the elementary inequality $a_1 + a_2 \leq a_1a_2 + 1$, for $a_1, a_2 \in [0,1]$. We will now estimate the two integrals on the last line of (4.5.14) by using the martingale properties of $\{L_s^u\}_{u\in[s,t]}$ and the representations (4.5.6) and (4.5.8) in the proof of Lemma 4.5.7. For the rest of the proof we write L_s^u as $L_s(u)$. By an application of Ito's formula, we have that for every $k = 1, \cdots K$, $\nu_{s,k}$ a.s.

$$\begin{aligned}
&\sqrt{L_s(t)} - 1 - \sum_{i=1}^{K}\int_s^t \frac{\sqrt{L_s(u-)}}{2}\left(\frac{\gamma_i^{\varepsilon}(u)}{\gamma_i(u)} - 1\right)dM_i(u) \\
=&\sum_{s<u\leq t}\left(\sqrt{L_s(u)} - \sqrt{L_s(u-)}\right) - \sum_{i=1}^{K}\int_s^t \frac{\sqrt{L_s(u-)}}{2}\left(\frac{\gamma_i^{\varepsilon}(u)}{\gamma_i(u)} - 1\right)dN_i(u) \\
=&\sum_{i=1}^{K}\int_s^t \sqrt{L_s(u-)}\left(\sqrt{\frac{\gamma_i^{\varepsilon}(u)}{\gamma_i(u)}} - 1\right)dN_i(u) - \sum_{i=1}^{K}\int_s^t \frac{\sqrt{L_s(u-)}}{2}\left(\frac{\gamma_i^{\varepsilon}(u)}{\gamma_i(u)} - 1\right)dN_i(u) \\
=&-\frac{1}{2}\sum_{i=1}^{K}\int_s^t \sqrt{L_s(u-)}\left(\sqrt{\frac{\gamma_i^{\varepsilon}(u)}{\gamma_i(u)}} - 1\right)^2 dN_i(u),
\end{aligned} \tag{4.5.15}$$

where the second equality follows on observing that for $u \in (s,t]$,

$$\sqrt{L_s(u)} = \sum_{i=1}^{K} \sqrt{L_s(u-)}\sqrt{\frac{\gamma_i^\varepsilon(u)}{\gamma_i(u)}}\Delta N_i(u).$$

As in the proof of Lemma 4.5.7, we can check that for all i, k,

$$\int_{\mathcal{W}} \left[\int_s^t \sqrt{L_s(u)} \cdot |\gamma_i^\varepsilon(u) - \gamma_i(u)| w(u) du\right] d\nu_{s,k}(w) < \infty,$$

and consequently the last term on the left side of (4.5.15) is a martingale. Denoting the expectation operator corresponding to the probability measure $\nu_{s,k}$ on $\mathcal{W}$ by $\mathbb{E}_{s,k}$, we have

$$\begin{aligned}
1 - \mathbb{E}_{s,k}[\sqrt{L_s(t)}] =& \frac{1}{2}\sum_{i=1}^{K} \mathbb{E}_{s,k}\left[\int_s^t \sqrt{L_s(u-)}\left(\sqrt{\frac{\gamma_i^\varepsilon(u)}{\gamma_i(u)}} - 1\right)^2 dN_i(u)\right] \\
=& \frac{1}{2}\sum_{i=1}^{K} \mathbb{E}_{s,k}\left[\int_s^t \sqrt{L_s(u)}\left(\sqrt{\frac{\gamma_i^\varepsilon(u)}{\gamma_i(u)}} - 1\right)^2 w(u)\gamma_i(u) du\right] \\
=& \frac{1}{2}\sum_{i=1}^{K} \int_s^t \mathbb{E}_{s,k}\left[\sqrt{L_s(u)} w(u)\left(\sqrt{\gamma_i^\varepsilon(u)} - \sqrt{\gamma_i(u)}\right)^2\right] du \\
\leq& \frac{1}{2}\int_s^t K\varepsilon \cdot \mathbb{E}_{s,k}[\sqrt{L_s(u)} w(u)] du \\
\leq& \frac{K\varepsilon}{2}\int_s^t \left(\mathbb{E}_{s,k}[L_s(u)]\mathbb{E}_{s,k}[w^2(u)]\right)^{1/2} du \\
\leq& \frac{K\varepsilon}{2} \cdot t \cdot (6K^2 e^{2TK\|\boldsymbol{\gamma}\|_\infty})^{1/2},
\end{aligned}$$

where the last inequality follows from (4.5.3). Using the above bound in (4.5.13) we now have

$$|\rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^\varepsilon)| \leq t A^2 \|\beta\|_\infty \cdot t\|\boldsymbol{\alpha}\|_\infty \cdot \left[2K\varepsilon t(6K^2 e^{2TK\|\boldsymbol{\gamma}\|_\infty})^{1/2}\right]^{1/2}.$$

Finally, by Lemma 4.5.9, we have

$$\begin{aligned}
|\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^\varepsilon)| \leq& |\rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) - \rho_{A,t}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}^\varepsilon)| + 2C_6 e^{-C_7 A} \\
<& d_1 A^2 \varepsilon^{1/2} + 2C_6 e^{-C_7 A},
\end{aligned}$$

where $d_1 = tA^2\|\beta\|_\infty \cdot t\|\boldsymbol{\alpha}\|_\infty \cdot \left[2Kt(6K^2e^{2TK\|\boldsymbol{\gamma}\|_\infty})^{1/2}\right]^{1/2}$. The result now follows on taking $A = -\log\varepsilon$ in the above display and taking ε_0 sufficiently small (in particular such that $-\log(\varepsilon_0) > A_0$). ■

Now we combine all the above ingredients to complete the proof of Proposition 4.5.5.

Proof of Proposition 4.5.5: Using Lemma 4.5.10, 4.5.6 and 4.5.8, we get

$$\begin{aligned}|\rho_t^+ - \rho_t| \leq& |\rho_t(\boldsymbol{\alpha}+\varepsilon, \beta+\varepsilon, \boldsymbol{\gamma}+\varepsilon) - \rho_t(\boldsymbol{\alpha}+\varepsilon, \beta+\varepsilon, \boldsymbol{\gamma})| \\ &+ |\rho_t(\boldsymbol{\alpha}+\varepsilon, \beta+\varepsilon, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}+\varepsilon, \beta, \boldsymbol{\gamma})| + |\rho_t(\boldsymbol{\alpha}+\varepsilon, \beta, \boldsymbol{\gamma}) - \rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})| \\ \leq& C_8\varepsilon^{1/2}(-\log\varepsilon)^2 + d_1\varepsilon + d_2\varepsilon^{1/2},\end{aligned}$$

where $d_1 = 6T^2K^3e^{2TK}$ and $d_2 = 24K^3T^2e^{2TK}$. A similar bound holds for $|\rho_t^- - \rho_t|$. The result follows. ■

4.5.3 Effect of time perturbation on ρ_t

Throughout this section we fix rate functions $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. The aim of this section is to understand the evolution of the operator norm $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ as t changes. The main result of the section is Proposition 4.5.11 which studies continuity and differentiability properties of the function $\rho(t) := \rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$, $t \geq 0$.

Proposition 4.5.11. Suppose that $\beta(t) > 0$ for $t > 0$ and $\liminf_{t\to\infty} \beta(t) > 0$. Then

(i) ρ is a continuous strictly increasing function on $\mathbb{R}_+$ with

$$\rho(0) = 0 \text{and} \lim_{t\to\infty} \rho(t) = \infty.$$

(ii) There is a unique value $t_c' = t_c'(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ such that $\rho(t_c') = 1$.

The proof of the proposition relies on the following lemma and is given after the proof of the lemma.

Lemma 4.5.12. Let $0 < t_1 \le t_2 < \infty$. Then

$$|t_2 - t_1| \cdot \frac{\inf_{t_1 \le u \le t_2} \beta(u)}{t_1 \|\beta\|_\infty} \cdot \rho(t_1) \le \rho(t_2) - \rho(t_1) \le |t_2 - t_1| \cdot 6 t_2 K^2 \|\beta\|_\infty \|\boldsymbol{\alpha}\|_\infty e^{2t_2 K \|\boldsymbol{\gamma}\|_\infty}.$$

Proof: Letting $\mu := \mu_{t_2}$ we have

$$\begin{aligned} |\rho(t_2) - \rho(t_1)| \le & \left(\int_{\mathcal{X}\times\mathcal{X}} (\kappa_{t_2}(\mathbf{x},\mathbf{y}) - \kappa_{t_1}(\mathbf{x},\mathbf{y}))^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}) \right)^{1/2} \\ \le & \left(\int_{\mathcal{X}\times\mathcal{X}} (\|\beta\|_\infty w(t_2)\tilde{w}(t_2)|t_2 - t_1|)^2 d\mu(\mathbf{x}) d\mu(\mathbf{y}) \right)^{1/2} \\ \le & |t_2 - t_1| \cdot \|\beta\|_\infty \cdot t_2 \|\boldsymbol{\alpha}\|_\infty \cdot 6K^2 e^{2t_2 K \|\boldsymbol{\gamma}\|_\infty}, \end{aligned}$$

where the last inequality once again follows from (4.5.3). This proves the upper bound.

Next note that, for $\mu \otimes \mu$ a.e. $(\mathbf{x},\mathbf{y})$ such that $\kappa_{t_1}(\mathbf{x},\mathbf{y}) \neq 0$, we have

$$\begin{aligned} \frac{\kappa_{t_2}(\mathbf{x},\mathbf{y})}{\kappa_{t_1}(\mathbf{x},\mathbf{y})} =& 1 + \frac{\int_{t_1}^{t_2} w(u)\tilde{w}(u)\beta(u)du}{\int_0^{t_1} w(u)\tilde{w}(u)\beta(u)du} \\ \ge & 1 + \frac{w(t_1)\tilde{w}(t_1) \inf_{t_1 \le u \le t_2} \beta(u) \cdot (t_2 - t_1)}{w(t_1)\tilde{w}(t_1)\|\beta\|_\infty t_1}. \end{aligned}$$

Thus $\kappa_{t_2}(\mathbf{x},\mathbf{y}) \ge \left(1 + |t_2 - t_1| \cdot \frac{\inf_{t_1 \le u \le t_2} \beta(u)}{t_1 \|\beta\|_\infty}\right) \kappa_{t_1}(\mathbf{x},\mathbf{y})$ which from Lemma 4.5.2 (b) implies

$$\rho(t_2) - \rho(t_1) \ge |t_2 - t_1| \cdot \frac{\inf_{t_1 \le u \le t_2} \beta(u)}{t_1 \|\beta\|_\infty} \cdot \rho(t_1).$$

This completes the proof of the lower bound. ■

Proof of Proposition 4.5.11: Since $\kappa_0 = 0$, the property $\rho(0) = 0$ is immediate. Also Lemma 4.5.12 shows that ρ is continuous and strictly increasing. Finally since $\inf_{t\to\infty} \beta(t) > 0$, there exists $\delta > 0$ and a $t^* \in (0,\infty)$ such that for all $t \ge t^*$, $\beta(t) \ge \delta$. From Lemma 4.5.12 we then have, for $t \ge t^*$, $\rho(t) - \rho(t^*) \ge \frac{(t-t^*)\delta}{t^* \|\beta\|_\infty}$. This proves that $\rho(t) \to \infty$ as $t \to \infty$ and completes the proof of (i). Part (ii) is immediate from (i). ■

4.5.4 Operator norm of $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})$ and critical time of BSR

In this section we will prove Theorem 4.2.3. Recall that by Lemma 4.4.4, for any fixed time t, $\mathbf{BSR}^*(t)$ (more precisely, $\Gamma(t)$) can be approximated by perturbations of $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})(t)$. To estimate the volume of the largest component in $\mathbf{RG}(\mathbf{a}, b, \mathbf{c})(t)$ we will use Lemma 4.5.1. In order to identify suitable Λ_n as in part (ii) of the lemma, we start with the following lemma.

Lemma 4.5.13. Let $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ be rate functions and let μ_t be the associated type measure. Fix $T > 0$. Define $\Lambda \in \mathcal{B}(\mathcal{X})$ as $\Lambda = \{(s, w) \in \mathcal{X} : w(T) \leq l\}$ for $l \in \mathbb{R}_+$. Then, for every $l \in \mathbb{R}_+$

$$\mu_t(\Lambda^c) < 2T\|\boldsymbol{\alpha}\|_\infty \cdot \exp\left(-l\frac{e^{-TK\|\boldsymbol{\gamma}\|_\infty}}{2K}\right).$$

Proof: Note that

$$\mu_t(\Lambda^c) = \sum_{i=1}^{K}\int_0^t \alpha_i(u)\nu_{u,i}(\Lambda^c) \leq ||\boldsymbol{\alpha}||_\infty T \sup_{u\in[0,T]}\sup_{1\leq i\leq K} \nu_{u,i}(\Lambda^c). \tag{4.5.16}$$

By (4.5.2),

$$\nu_{u,i}(\Lambda^c) = \nu_{u,i}(\{w : w(T) \geq l\}) \leq \mathbb{P}(X_1 + X_2 \geq l/K) \leq 2(1 - e^{-TK\|\boldsymbol{\gamma}\|_\infty})^{l/2K}.$$

where X_i are iid with $\mathrm{Geom}(e^{-T\|\boldsymbol{\gamma}\|_\infty})$ distribution. Using this estimate in (4.5.16), we have

$$\mu_t(\Lambda^c) \leq ||\boldsymbol{\alpha}||_\infty T \cdot 2(1 - e^{-TK\|\boldsymbol{\gamma}\|_\infty})^{l/2K}.$$

The result follows. ■

We will now use the above lemma along with Lemma 4.5.1 to estimate the largest component in $\mathbf{RG}^{(n)}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})(t)$. Recall the notation $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ from Section 4.5.2.

Lemma 4.5.14. Let $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$ be rate functions and denote by $\mathcal{I}_1^{\mathbf{RG}}(t)$ the component with the largest volume, with respect to the weight function ϕ_t, in $\mathbf{RG}^{(n)}(t) :=$

$\mathbf{RG}^{(n)}(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})(t)$. Then, for every $t > 0$ such that $\rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) < 1$, there exists $A \in (0, \infty)$ such that

$$\mathbb{P}(\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}}(t)) > A \log^4 n) \to 0, \text{ as } n \to \infty.$$

Proof: We will use Lemma 4.5.1(ii). Define

$$\Lambda_n := \{(s, w) \in \mathcal{X} : w(t) < B \log n\},$$

where B will be chosen appropriately later in the proof. Now consider the function $g(n)$ in Lemma 4.5.1(ii) with Λ_n defined as above and (μ, ϕ, κ) there replaced by $(\mu_t, \phi_t, \kappa_t)$, where (μ_t, κ_t) is the type measure and kernel associated with $(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma})$. Note that

$$\kappa_t(\mathbf{x}, \mathbf{y}) = \int_0^t \beta(u) w(u) \tilde{w}(u) du \leq t \|\beta\|_\infty w(t) \tilde{w}(t)$$

and therefore

$$g(n) \leq 8B \log n (1 + 3\mu_t(\mathcal{X}) \cdot t \|\beta\|_\infty B^2 \log^2 n). \tag{4.5.17}$$

Writing $m_n = A \log^4 n$, the bound in Lemma 4.5.1(ii) then gives

$$\mathbb{P}(\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}}(t)) > m_n) \leq n\mu_t(\Lambda_n^c) + 2n\mu_t(\mathcal{X}) \exp\Big(-\Delta^2 A \log^4 n / g(n)\Big), \tag{4.5.18}$$

where $\Delta = 1 - \rho_t(\boldsymbol{\alpha}, \beta, \boldsymbol{\gamma}) > 0$. Using Lemma 4.5.13 with $l = B \log n$ gives

$$n\mu_t(\Lambda_n^c) \leq nt \|\boldsymbol{\alpha}\|_\infty \cdot n^{-Be^{-T\|\boldsymbol{\gamma}\|_\infty}/2K} = o(1) \tag{4.5.19}$$

for $B > 2Ke^{T\|\boldsymbol{\gamma}\|_\infty}$. Now fix $B > e^{T\|\boldsymbol{\gamma}\|_\infty/2K}$, and choose A large such that

$$n\mu_t(\mathcal{X}) \exp\Big(-\Delta^2 A \log^4 n / g(n)\Big) \to 0$$

as $n \to \infty$. The result follows. ∎

Proof of Theorem 4.2.3: Let, for $t \geq 0$, (μ_t, κ_t) be the type measure and the kernel associated with rate functions $(\mathbf{a}, b, \mathbf{c})$. We will prove Theorem 4.2.3 with this choice of (μ_t, κ_t). From Proposition 4.5.11 we have that $\rho(t) = \rho_t(\mathbf{a}, b, \mathbf{c})$ is continuous and strictly increasing in t and there is a unique $t'_c \in (0, \infty)$ such that $\rho(t'_c) = 1$. It now suffices to show that: (a) For $t < t'_c$, $|\mathcal{C}_1(t)|$ (the size of the largest component in $\mathbf{BSR}^*(t)$) is $O(\log^4 n)$; and (b) for $t > t'_c$, $|\mathcal{C}_1(t)| = \Omega(n)$.

Consider first (a). Fix $t < t'_c$. For $\delta > 0$, define rate functions $(\mathbf{a}^+, b^+, \mathbf{c}^+) = (\mathbf{a} + \delta, b + \delta, \mathbf{c} + \delta)$. Since $\rho(t) < 1$, by Proposition 4.5.5, we can choose δ sufficiently small so that $\rho_t(\mathbf{a}^+, b^+, \mathbf{c}^+) < 1$. Denote $\mathcal{I}_1^{\mathbf{RG}^+}(t)$ for the component of the largest volume in $\mathbf{RG}^+(t) := \mathbf{RG}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)$. From Lemma 4.5.14 there exists $A \in (0, \infty)$ such that

$$\mathbb{P}(\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^+}(t)) > A \log^4 n) \to 0, \text{ as } n \to \infty.$$

Combining this result with Corollary 4.4.5 we see that

$$\mathbb{P}(\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\Gamma}(t)) > A \log^4 n) \to 0, \text{ as } n \to \infty,$$

where $\mathcal{I}_1^{\Gamma}(t)$ is the component with the largest volume in $\Gamma(t)$. Part (a) is now immediate from the one to one correspondence between the components in $\Gamma(t)$ and $\mathbf{BSR}^*(t)$ (see (4.4.20)).

We now consider (b). Fix $t > t'_c$. Then $\rho(t) > 1$. From Proposition 4.5.5 we can find $\delta > 0$ such that $\rho_t(\mathbf{a}^-, b^-, \mathbf{c}^-) > 1$, where $(\mathbf{a}^-, b^-, \mathbf{c}^-) = ((\mathbf{a} - \delta)^+, (b - \delta)^+, (\mathbf{c} - \delta)^+)$. Let $\mathcal{C}_1^{\mathbf{RG}^-}(t)$ be the component in $\mathbf{RG}^-(t) := \mathbf{RG}^{(n)}(\mathbf{a}^-, b^-, \mathbf{c}^-)(t)$ with the largest number of vertices. By Theorem 3.1 of [11], $|\mathcal{C}_1^{\mathbf{RG}^-}(t)| = \Theta(n)$. Since $\mathrm{vol}_{\phi_t}(\mathcal{C}_1^{\mathbf{RG}^-}(t)) \geq |\mathcal{C}_1^{\mathbf{RG}^-}(t)|$, we have $\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^-}(t)) = \Omega(n)$, where $\mathcal{I}_1^{\mathbf{RG}^-}(t)$ is the component with the largest volume in $\mathbf{RG}^-(t)$. Finally, in view of Corollary 4.4.5, we have the same result with $\mathcal{I}_1^{\mathbf{RG}^-}(t)$ replaced by $\mathcal{I}_1^{\Gamma}(t)$ and the result follows once more from the one to one correspondence between the components in $\Gamma(t)$ and $\mathbf{BSR}^*(t)$. ■

4.5.5 Barely subcritical regime for bounded-size rules

In this section we complete the proof of Theorem 4.2.2. Throughout this section we fix $\gamma \in (0, 1/4)$ and let $t_n = t_c - n^{-\gamma}$. The main ingredient in the proof is the following proposition.

Proposition 4.5.15. There exist $\bar{B}, \bar{C}, \bar{N} \in (0, \infty)$ such that for all $n \geq \bar{N}$ and all $0 \leq t \leq t_n$

$$\mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| \geq \bar{m}(n,t)\} \leq \frac{\bar{C}}{n^2}, \text{ where } \bar{m}(n,t) = \frac{\bar{B}(\log n)^4}{(t_c - t)^2}.$$

Let us first prove Theorem 4.2.2 assuming the above proposition.

Proof of Theorem 4.2.2: Write $\tau = \inf\{t \geq 0 : |\mathcal{C}_1^{(n)}(t)| \geq m(n,t)\}$, where $m(n,t) = \frac{2\bar{B}(\log n)^4}{(t_c - t)^2}$. Then

$$\mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| \geq m(n,t) \text{ for some } t \leq t_n\} = \mathbb{P}\{\tau \leq t_n\}. \tag{4.5.20}$$

Note that

$$\{\tau = t\} \subset \bigcup_{v,v' \in [n], v \neq v'} E^{v,v'}, \tag{4.5.21}$$

where, denoting the component in $\mathbf{BSR}(t)$ that contains the vertex $v \in [n]$ by $\mathcal{C}_v^{(n)}(t)$ and its size by $|\mathcal{C}_v^{(n)}(t)|$,

$$\begin{aligned} E^{v,v'} = \Big\{\max\big\{|\mathcal{C}_v^{(n)}(t-)|, |\mathcal{C}_{v'}^{(n)}(t-)|\big\} < m(n,t); \mathcal{C}_v^{(n)}(t-) \neq \mathcal{C}_{v'}^{(n)}(t-)\Big\} \\ \bigcap \big\{|\mathcal{C}_v^{(n)}(t-)| + |\mathcal{C}_{v'}^{(n)}(t-)| \geq m(n,t)\big\} \bigcap \{\mathcal{C}_v^{(n)}(t) = \mathcal{C}_{v'}^{(n)}(t)\}. \end{aligned} \tag{4.5.22}$$

Note that

$$\mathbb{P}\{|\mathcal{C}_v^{(n)}(t)| + |\mathcal{C}_{v'}^{(n)}(t)| \geq m(n,t)\} \leq 2\mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| \geq m(n,t)/2\} \tag{4.5.23}$$

and, on the set, $\{\max\big\{|\mathcal{C}_v^{(n)}(t)|, |\mathcal{C}_{v'}^{(n)}(t)|\big\} < m(n,t)\}$, the rate at which $\mathcal{C}_v^{(n)}(t)$ and $\mathcal{C}_{v'}^{(n)}(t)$ merge can be bounded by

$$\frac{1}{2n^3} \cdot 4|\mathcal{C}_v^{(n)}(t)||\mathcal{C}_{v'}^{(n)}(t)|n^2 \leq \frac{2m^2(n,t)}{n}.$$

Combining this observation with (4.5.21) and (4.5.23), we have

$$
\begin{aligned}
\mathbb{P}\{\tau \leq t_n\} \leq & \sum_{v,v' \in [n], v \neq v'} \int_0^{t_n} \mathbb{P}\{|\mathcal{C}_v^{(n)}(t)| + |\mathcal{C}_{v'}^{(n)}(t)| \geq m(n,t)\} \cdot \frac{2m^2(n,t)}{n} dt \\
\leq & 2n^2 \int_0^{t_n} \mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| \geq m(n,t)/2\} \cdot \frac{2m^2(n,t)}{n} dt \\
\leq & 4nt_c \sup_{t \leq t_n} \left\{ m^2(n,t) \mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| \geq \bar{m}(n,t)\} \right\} \\
= & O(n \cdot n^{4\gamma} (\log n)^8 \cdot n^{-2}) = O(n^{-1+4\gamma} (\log n)^8) = o(1),
\end{aligned}
$$

where the last line follows from Proposition 4.5.15 and the fact that $\gamma < 1/4$. Using the above estimate in (4.5.20) we have the result. ■

We will need the following lemma in the proof of Proposition 4.5.15.

Lemma 4.5.16. Let $(\mathbf{a}^+, b^+, \mathbf{c}^+) = (\mathbf{a} + \delta_n, b + \delta_n, \mathbf{c} + \delta_n)$, where $\delta_n = n^{-2\gamma_0}$ and $\gamma_0 \in (\gamma, 1/4)$. Let $\rho_t^{(n),+} = \rho_t(\mathbf{a}^+, b^+, \mathbf{c}^+)$. Then there exists $C_9, N_0 \in (0, \infty)$ such that for all $n \geq N_0$,

$$\rho_t^{(n),+} < 1 - C_9(t_c - t) \text{ for all } 0 \leq t \leq t_n.$$

Proof of Lemma 4.5.16: From Proposition 4.5.5, there is a $d_1 \in (0, \infty)$ such that

$$\rho_t^{(n),+} \leq \rho_t(\mathbf{a}, b, \mathbf{c}) + d_1 n^{-\gamma_0} \log^2 n, \text{ for all } t \leq t_c.$$

By Lemma 4.5.12 and since $\rho_{t_c}(\mathbf{a}, b, \mathbf{c}) = 1$, there exists $d_2 \in (0, \infty)$ such that

$$\rho_t(\mathbf{a}, b, \mathbf{c}) \leq 1 - d_2(t_c - t), \text{ for all } t \leq t_n.$$

Thus, since $\gamma < \gamma_0$, we have for some $N_0 > 0$

$$\rho_t^{(n),+} \leq 1 - d_2(t_c - t) + d_1 n^{-\gamma_0} (\log n)^2 < 1 - \frac{d_2}{2}(t_c - t),$$

for all $n \geq N_0$ and $0 \leq t \leq t_c - n^{-\gamma}$. The result follows. ■

Proof of Proposition 4.5.15: Recall the rate functions $(\mathbf{a}, b, \mathbf{c})$ introduced in Section 4.4.2. Choose $\gamma_0 \in (\gamma, 1/4)$ and let $(\mathbf{a}^+, b^+, \mathbf{c}^+)$ be as in Lemma 4.5.16. Fix

$t < t_n$ and consider the random graph $\mathbf{RG}^{(n)}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)$. From Lemma 4.4.4, we can couple $\Gamma(t)$ and $\mathbf{RG}^{(n)}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)$ such that

$$\mathbb{P}(\Gamma(t) \subseteq \mathbf{RG}^{(n)}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)) \geq 1 - C_3 \exp(-C_4 n^{1-4\gamma_0}), \text{ for all } t \in [0, T].$$

Recalling the one to one correspondence between components in $\mathbf{BSR}^*(t)$ and $\Gamma(t)$, and (4.4.20), we have for any $m \geq 1$,

$$\mathbb{P}\{|\mathcal{C}_1^{(n)}(t)| > m\} \leq \mathbb{P}\{\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^+}(t)) \geq m\} + C_3 \exp\{-C_4 n^{1-4\gamma_0}\}, \tag{4.5.24}$$

where $\mathcal{I}_1^{\mathbf{RG}^+}(t)$ is the component with the largest volume in $\mathbf{RG}^{(n)}(\mathbf{a}^+, b^+, \mathbf{c}^+)(t)$. From Lemma 4.5.16, there is a $N_0 > 0$ such that $\Delta_t^{(n),+} = 1 - \rho_t(\mathbf{a}^+, b^+, \mathbf{c}^+)$ satisfies

$$\Delta_t^{(n),+} \geq C_9(t_c - t), \text{ for all } t \leq t_n, \ n \geq N_0. \tag{4.5.25}$$

Using Lemma 4.5.1 and arguing as in equation (4.5.18) we have for all $t \in [0, T]$ and all $m \geq 1$,

$$\mathbb{P}\{\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^+}(t)) \geq m\} \leq n d_1 \exp\{-(\Delta_t^{(n),+})^2 m/(d_2 \log^3 n)\} + d_3 n^{-2}, \tag{4.5.26}$$

where d_1, d_2, d_3 are suitable constants. Using (4.5.25) in (4.5.26) we get

$$\mathbb{P}\{\mathrm{vol}_{\phi_t}(\mathcal{I}_1^{\mathbf{RG}^+}(t)) \geq m\} \leq n d_1 \exp\{-d_4 (t_c - t)^2 m / \log^3 n\} + d_3 n^{-2}.$$

The result now follows on substituting $m = m(n,t) = \frac{\bar{B}(\log n)^4}{(t_c - t)^2}$, with $\bar{B} > 3/d_4$, in the above inequality. ■

CHAPTER 5: THE AUGMENTED MC AND BSR

5.1 Introduction

In this chapter we introduce the *augmented multiplicative coalescent* (AMC), which captures the evolution of the sizes and surplus of the components of the bounded-size-rule processes in the critical window. The augmented multiplicative coalescent is an extension of Aldous's multiplicative coalescent. Recall the typical state space for the multiplicative coalescent is

$$l^2_\downarrow = \left\{ (x_1, x_2, ...) : x_1 \geq x_2 \geq ... \geq 0, \sum_{i=1}^{\infty} x_i^2 < \infty \right\}.$$

The dynamics of Aldous's multiplicative coalescent can be described as follows: Two clusters of sizes x_i and x_j merge into one cluster of size $x_i + x_j$ at the rate $x_i x_j$.

The augmented multiplicative coalescent is defined as a continuous-time Markov process on the state space

$$\mathbb{U}_\downarrow := \left\{ (\mathbf{x}, \mathbf{y}) : \mathbf{x} \in l^2_\downarrow, \mathbf{y} \in \mathbb{N}^\infty, \sum_{i=1}^{\infty} x_i y_i < \infty \right\},$$

where $\mathbf{x} = (x_1, x_2, ...)$ and $\mathbf{y} = (y_1, y_2, ...)$. $\mathbb{U}_\downarrow$ is equipped with a suitable metric which is defined in (5.2.5). To describe the dynamics of the augmented multiplicative coalescent, we treat (x_i, y_i) as the label on the i-th cluster, for $i \in \mathbb{N}$. The evolution of the augmented multiplicative coalescent follows the following two rules:

- For $i \neq j \in \mathbb{N}$, two clusters with labels (x_i, y_i) and (x_j, y_j) merge into one cluster with label $(x_i + x_j, y_i + y_j)$ at rate $x_i x_j$.
- For $i \in \mathbb{N}$, the cluster with label (x_i, y_i) changes into a cluster with label $(x_i, y_i + 1)$ at rate $x_i^2/2$.

The $\mathbf{x}$-coordinate represents the component sizes while the $\mathbf{y}$-coordinate corresponds to their surplus. The first bullet above describes the change in state when the i-th and j-th clusters merge while the second bullet corresponds to an edge formation between two vertices in the i-th component.

In this chapter, we give a concrete construction of the augmented multiplicative coalescent. We show that this process is well-defined and satisfies a certain regularity which can be viewed as a form of Feller property. Furthermore we show that there exists a special version of the augmented multiplicative coalescent, whose first coordinate is the standard multiplicative coalescent. We call this special version the *standard augmented multiplicative coalescent*. Finally we revisit the bounded-size-rule process that was introduced in Chapter 4, and show that the stochastic process of the sizes and surplus of its components in the critical window converge in the sense of finite dimensional distributions, to the standard augmented multiplicative coalescent.

Organization of this chapter: Section 5.2 reviews the definition of the bounded-size-rule processes and gives a formal description of the augmented multiplicative coalescent (AMC). In Section 5.3 we state our main results. Sections 5.4 and 5.5 are devoted to proving the existence and the near-Feller property of the AMC. In particular Section 5.5 contains the proof of Theorem 5.3.1. Section 5.6 studies general bounded-size rules in the critical window and proves Theorems 5.3.2. Finally in Section 5.7 we complete the proof of Theorem 5.3.3.

5.2 Definitions and notation

5.2.1 Notation

We collect some common notation and conventions used in this work. For a finite set A write $|A|$ for its cardinality. A graph $\mathbf{G}$ with no vertices and edges will be called a **null graph**. For graphs $\mathbf{G}_1, \mathbf{G}_2$, if $\mathbf{G}_1$ is a subgraph of $\mathbf{G}_2$ we shall write this as $\mathbf{G}_1 \subset \mathbf{G}_2$. Denote by $|\mathcal{C}|$ the size of a connected component $\mathcal{C}$. Denote by $\mathbf{spls}(\mathcal{C})$ the surplus of the component $\mathcal{C}$. Let $\mathcal{G}$ be the set of all possible graphs $(\mathcal{V}, \mathcal{E})$ on a given type space $\mathcal{X}$. When $\mathcal{V}$ is finite, we will consider $\mathcal{G}$ to be endowed with the discrete topology and the corresponding Borel sigma field and refer to a random element of $\mathcal{G}$ as a random graph.

For a RCLL (right continuous functions with left limits) function $f : [0, \infty) \to \mathbb{R}$, we write $\Delta f(t) = f(t) - f(t-)$, $t > 0$. Suppose that $(S, \mathcal{S})$ is a measurable space and we are given a partial ordering on S. We say the S-valued random variable ξ **stochastically dominates** $\tilde{\xi}$, and write $\xi \geq_d \tilde{\xi}$ if there exists a coupling between the two random variables on a common probability space such that $\xi^* \geq \tilde{\xi}^*$ a.s., where $\xi^* =_d \xi$ and $\tilde{\xi}^* =_d \tilde{\xi}$. For probability measures $\mu, \tilde{\mu}$ on S, we say μ stochastically dominates $\tilde{\mu}$, and write $\mu \geq_d \tilde{\mu}$ if $\xi \geq_d \tilde{\xi}$ where ξ has distribution μ and $\tilde{\xi}$ has distribution $\tilde{\mu}$.

5.2.2 The continuous-time bounded-size-rule processes

Fix $K \in \mathbb{N}$ and let $\Omega_0 = \{\varpi\}$ and $\Omega_K = \{1, 2, \ldots, K, \varpi\}$ for $K \geq 1$, where ϖ will represent components of size greater than K. Given a graph $\mathbf{G}$ and a vertex $v \in \mathbf{G}$, write $\mathcal{C}_v(\mathbf{G})$ for the component that contains v. Let

$$c(v) = \begin{cases} |\mathcal{C}_v(\mathbf{G})| & \text{if } |\mathcal{C}_v(\mathbf{G})| \leq K \\ \varpi & \text{if } |\mathcal{C}_v(\mathbf{G})| > K. \end{cases} \tag{5.2.1}$$

For a quadruple of vertices v_1, v_2, v_3, v_4, write $\vec{v} = (v_1, v_2, v_3, v_4)$ and let $c(\vec{v}) = (c(v_1), c(v_2), c(v_3), c(v_4))$.

Fix $F \subseteq \Omega_K^4$. We construct the F-BSR process $\{\mathbf{BSR}^{(n)}(t)\}_{t \geq 0}$ as follows. Define $\mathbf{BSR}^{(n)}(0) = \mathbf{0}_n$, the graph on $[n]$ with no edges. For every quadruple of vertices $\vec{v} =$

$(v_1, v_2, v_3, v_4) \in [n]^4$, let $\mathcal{P}_{\vec{v}}$ be a Poisson process with rate $\frac{1}{2n^3}$, independent between quadruples. Denote the function $c(v)$ [resp. $c(\vec{v})$] associated with $\mathbf{BSR}^{(n)}(t-)$ as $c_{t-}(v)$ [resp. $c_{t-}(\vec{v})$]. Given $\mathbf{BSR}^{(n)}(t-)$, and that for some $\vec{v} \in [n]^4$, $\mathcal{P}_{\vec{v}}$ has a point at the time instant t, we define

$$\mathbf{BSR}^{(n)}(t) = \begin{cases} \mathbf{BSR}^{(n)}(t-) \cup (v_1, v_2) & \text{if } c_{t-}(\vec{v}) \in F \\ \mathbf{BSR}^{(n)}(t-) \cup (v_3, v_4) & \text{otherwise.} \end{cases} \tag{5.2.2}$$

To simplify notation, when there is no scope for confusion, we will suppress n in the notation. For example, we write $\mathbf{BSR}_t := \mathbf{BSR}^{(n)}(t)$.

Denote $\mathcal{C}_i^{(n)}(t)$ for the i-th largest component of $\mathbf{BSR}_t$. Spencer and Wormald [31] shows that for given F-BSR, there exists a (model dependent) **critical time** $t_c > 0$ such that for $t < t_c$, $|\mathcal{C}_1^{(n)}(t)| = O(\log n)$ and for $t > t_c$, $|\mathcal{C}_1^{(n)}(t)| \sim f(t)n$ where $f(t) > 0$.

Along with the size of the components, another key quantity of interest is the surplus of the components. Denote $\xi_i^{(n)}(t) := \mathbf{spls}(\mathcal{C}_i^{(n)}(t))$ for the surplus of the component $\mathcal{C}_i^{(n)}(t)$. We will be interested in the joint vector of ordered component sizes and corresponding surplus

$$((|\mathcal{C}_i(t)|, \xi_i(t)) : i \geq 1).$$

5.2.3 The augmented multiplicative coalescent

5.2.3.1 Aldous's multiplicative coalescent

Let $l^2 = \{x = (x_1, x_2, \ldots) : \sum_i x_i^2 < \infty\}$. Then l^2 is a separable Hilbert space with the inner product $\langle x, y\rangle = \sum_{i=1}^\infty x_i y_i$, $x = (x_i), y = (y_i) \in l^2$. Let

$$l^2_\downarrow = \{(x_1, x_2, \ldots) : x_1 \geq x_2 \geq \cdots \geq 0, \sum_i x_i^2 < \infty\}. \tag{5.2.3}$$

Then $l^2_\downarrow$ is a closed subset of l^2 which we equip with the metric inherited from l^2. In [2] Aldous introduced a $l^2_\downarrow$ valued continuous time Markov process, called the *standard multiplicative coalescent*, that can be used to describe the asymptotic behavior of suitably scaled component size vector in Erdős-Rényi random graph evolution, near criticality. Subsequently, similar results have been shown to hold for the Bohman-Frieze process (Chapter 3) and other random graph models (see [4] and reference therein). We now give a brief description of this Markov process.

Fix $x = (x_i)_{i\in\mathbb{N}}$. Let $\{\xi_{i,j}, i, j \in \mathbb{N}\}$ be a collection of independent rate one Poisson processes. Given $t \geq 0$, consider the random graph with vertex set $\mathbb{N}$ in which there exist $\xi_{i,j}([0, tx_ix_j/2]) + \xi_{j,i}([0, tx_ix_j/2])$ edges between (i, j), $1 \leq i < j < \infty$, and there are $\xi_{i,i}([0, tx_i^2/2])$ self-loops at vertex $i \in \mathbb{N}$. The volume of a component $\mathcal{C}$ of this graph is defined to be

$$\text{vol}(\mathcal{C}) := \sum_{i\in\mathcal{C}} x_i.$$

Let $X_i(x, t)$ be the volume of the i-th largest (by volume) component. It can be shown that $X(x, t) = (X_i(x, t), i \geq 1) \in l^2_\downarrow$, a.s. (see Lemma 20 in [2]). Define

$$T_t : \text{BM}(l^2_\downarrow) \to \text{BM}(l^2_\downarrow),$$

as $T_tf(x) = \mathbb{E}(f(X(x, t)))$. It is easily checked that $(T_t)_{t\geq0}$ satisfies the semigroup property $T_{t+s} = T_tT_s$, $s, t \geq 0$, and [2] shows that (T_t) is Feller in the sense that $T_t(C_b(l^2_\downarrow)) \subset C_b(l^2_\downarrow)$ for all $t \geq 0$. The paper [2] also shows that the semigroup (T_t) along with an initial distribution $\mu \in \mathcal{P}(l^2_\downarrow)$ determines a Markov process with values in $l^2_\downarrow$ and RCLL sample paths. Denoting by P^μ the probability distribution of this Markov process on $\mathcal{D}([0, \infty) : l^2_\downarrow)$, the Feller property says that $\mu \mapsto P^\mu$ is a continuous map. One special choice of initial distribution for this Markov process is particularly relevant for the study of asymptotics of random graph models. We now describe this distribution. Let $\{W(t)\}_{t\geq0}$ be a standard Brownian motion, and for a fixed $\lambda \in \mathbb{R}$, define

$$W_\lambda(t) = W(t) + \lambda t - \frac{t^2}{2},\ t \geq 0.$$

Let $\hat{W}_\lambda$ denote the reflected version of W_λ, i.e.,

$$\hat{W}_\lambda(t) = W_\lambda(t) - \min_{0 \leq s \leq t} W_\lambda(s),\ t \geq 0. \tag{5.2.4}$$

An excursion of $\hat{W}_\lambda$ is an interval $(l, u) \subset [0, +\infty)$ such that $\hat{W}_\lambda(l) = \hat{W}_\lambda(u) = 0$ and $\hat{W}_\lambda(t) > 0$ for all $t \in (l, u)$. Define $u - l$ as the length of the excursion. Order the lengths of excursions of $\hat{W}_\lambda$ as

$$X_1^*(\lambda) > X_2^*(\lambda) > X_3^*(\lambda) > \cdots$$

and write $\boldsymbol{X}^*(\lambda) = (X_i^*(\lambda) : i \geq 1)$. Then $\boldsymbol{X}^*(\lambda)$ defines a $l^2_\downarrow$ valued random variable (see Lemma 25 in [2]) and let μ_λ be its probability distribution. Using the Feller property and connections with certain inhomogeneous random graph models, the paper [2] shows that $\mu_\lambda T_t = \mu_{\lambda+t}$, for all $\lambda \in \mathbb{R}$ and $t \geq 0$, where for $\mu \in \mathcal{P}(l^2_\downarrow)$, $\mu T_t \in \mathcal{P}(l^2_\downarrow)$ is defined in the usual way: $\mu T_t(A) = \int T_t(\mathbb{1}_A)(x)\mu(dx),\ A \in \mathcal{B}(l^2_\downarrow)$. Using this consistency property one can determine a unique probability measure $\mu_{\text{MC}} \in \mathcal{P}(\mathcal{D}((-\infty, \infty) : l^2_\downarrow))$ such that, denoting the canonical coordinate process on $\mathcal{D}((-\infty, \infty) : l^2_\downarrow)$ by $\{\pi_t\}_{-\infty < t < \infty}$,

$$\mu_{\text{MC}} \circ (\pi_{t+\cdot})^{-1} = P^{\mu_t}, \quad \text{for all } t \in \mathbb{R},$$

where $\pi_{t+\cdot}$ is the process $\{\pi_{t+s}\}_{s \geq 0}$. The measure μ_{MC} is known as the *standard multiplicative coalescent*. This measure plays a central role in characterizing asymptotic distribution of component size vectors in the critical window for random graph models [2, 4, 7].

5.2.3.2 The augmented multiplicative coalescent

We will now augment the above construction and introduce a measure on a larger space that can be used to describe the joint asymptotic behavior of the component size vector and the associated surplus vector, for a broad family of random graph models.

Let $\mathbb{N}^\infty = \{y = (y_1, \cdots) : y_i \in \mathbb{N}, \text{ for all } i \geq 1\}$ and define

$$\mathbb{U}_\downarrow = \{(x_i, y_i)_{i\geq 1} \in l^2_\downarrow \times \mathbb{N}^\infty : \sum_{i=1}^{\infty} x_i y_i < \infty \text{ and } y_m = 0 \text{ whenever } x_m = 0, m \geq 1\}.$$

We will view x_i as the volume of the i-th component and y_i the surplus of the i-th component of a graph with vertex set $\mathbb{N}$. Writing $x = (x_i)$ and $y = (y_i)$, we will sometimes denote (x_i, y_i) as $z = (x, y)$. We equip $\mathbb{U}_\downarrow$ with the metric

$$\mathbf{d}_{\mathbb{U}}((x, y), (x', y')) = \left(\sum_{i=1}^{\infty} (x_i - x'_i)^2\right)^{1/2} + \sum_{i=1}^{\infty} |x_i y_i - x'_i y'_i|. \tag{5.2.5}$$

The choice of this metric is discussed in Remark 5.4.15.

Let $\mathbb{U}^0_\downarrow = \{(x_i, y_i)_{i\geq 1} \in \mathbb{U}_\downarrow : \text{ if } x_k = x_m, k \leq m, \text{ then } y_k \geq y_m\}$. We now introduce the *augmented multiplicative coalescent* (AMC). This is a continuous time Markov process with values in $(\mathbb{U}^0_\downarrow, \mathbf{d}_{\mathbb{U}})$, whose dynamics can heuristically be described as follows: The process jumps at any given time instant from state $(x, y) \in \mathbb{U}^0_\downarrow$ to:

- (x^{ij}, y^{ij}) at rate $x_i x_j$, $i \neq j$, where (x^{ij}, y^{ij}) is obtained by merging components i and j into a component with volume $x_i + x_j$ and surplus $y_i + y_j$ and reordering the coordinates to obtain an element in $\mathbb{U}^0_\downarrow$.
- (x, y^i) at rate $x_i^2/2$, $i \geq 1$, where (x, y^i) is the state obtained by increasing the surplus in the i-th component from y_i to $y_i + 1$ and reordering the coordinates (if needed) to obtain an element in $\mathbb{U}^0_\downarrow$.

Whenever $z = (x, y) \in \mathbb{U}^0_\downarrow$ is such that $\sum_{i=1}^\infty x_i < \infty$, it is easy to construct a well defined Markov process $\{\boldsymbol{Z}(z, \lambda)\}_{\lambda \geq 0}$ that corresponds to the above transition mechanism, starting at time $\lambda = 0$ in the state z. However when $\sum_{i=1}^\infty x_i = \infty$, the existence of such a process requires more work. We show in Section 5.4 (see also Theorem 5.3.1) that in fact there is a well defined Markov process $\{Z(z,\lambda)\}_{\lambda\geq 0}$ corresponding to the above dynamical description for *any* $z \in \mathbb{U}_{\downarrow}^{0}$. Define, for $\lambda \geq 0$, $\mathcal{T}_\lambda : \mathrm{BM}(\mathbb{U}_{\downarrow}^{0}) \to \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ as

$$(\mathcal{T}_\lambda f)(z) = \mathbb{E} f(\boldsymbol{Z}(z,\lambda)).$$

As for Aldous's multiplicative coalescent, there is one particular family of distributions that plays a special role. Recall the reflected parabolic Brownian motion $\hat{W}_\lambda(t)$ from (5.2.4). Let $\mathcal{P}$ be a Poisson point process on $[0,\infty)\times[0,\infty)$ with intensity $\lambda_\infty^{\otimes 2}$ (where λ_∞ is the Lebesgue measure on $[0,\infty)$) independent of $\hat{W}_\lambda$. Let (l_i, r_i) be the i-th largest excursion of $\hat{W}_\lambda$. Define

$$X_i^*(\lambda) = r_i - l_i \text{ and } Y_i^*(\lambda) = |\mathcal{P} \cap \{(t,z) : 0 \leq z \leq \hat{W}_\lambda(t), l_i \leq t \leq r_i\}|.$$

Then $\boldsymbol{Z}^*(\lambda) = (\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda))$ is a.s. a $\mathbb{U}_{\downarrow}^{0}$ valued random variable, where $\boldsymbol{X}^* = (X_i^*)_{i\geq 1}$ and $\boldsymbol{Y}^* = (Y_i^*)_{i\geq 1}$. Let ν_λ be its probability distribution. In Theorem 5.3.1 we will show that there exists a $\mathbb{U}_{\downarrow}^{0}$ valued stochastic process $(\boldsymbol{Z}(\lambda))_{-\infty<\lambda<\infty}$ such that $\boldsymbol{Z}(\lambda)$ has probability distribution ν_λ for every $\lambda \in (-\infty,\infty)$ and for all $f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$, and $\lambda_1 < \lambda_2$, we have

$$\mathbb{E}[f(\boldsymbol{Z}(\lambda_2)) | \{\boldsymbol{Z}(\lambda)\}_{\lambda\leq\lambda_1}] = (\mathcal{T}_{\lambda_2-\lambda_1} f)(\boldsymbol{Z}(\lambda_1)).$$

The process $\boldsymbol{Z}$ will be referred to as the *standard augmented multiplicative coalescent*. We will also show that $\{\mathcal{T}_\lambda\}_{\lambda\geq 0}$ is a semigroup, which is *nearly Feller*, in the sense made precise in the statement of Theorem 5.3.1. It will be seen that this process plays a similar role in characterizing the asymptotic joint distributions of the component size and surplus vector in the critical window as Aldous's standard multiplicative coalescent does in the study of asymptotics of the component size vector.

5.3 Main results

Our first result establishes the existence of the standard augmented coalescent process. Let $\mathbb{U}_{\downarrow}^{1} = \{z = (x,y) \in \mathbb{U}_{\downarrow}^{0} : \sum_i x_i = \infty\}$.

Theorem 5.3.1. There is a collection of maps $\{\mathcal{T}_t\}_{t\geq 0}$, $\mathcal{T}_t : \mathrm{BM}(\mathbb{U}_{\downarrow}^{0}) \to \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ and a $\mathbb{U}_{\downarrow}^{0}$ valued stochastic process $\{\boldsymbol{Z}(\lambda)\}_{-\infty<\lambda<\infty} = \{(\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda))\}_{-\infty<\lambda<\infty}$ such that the following hold.

(i) $\{\mathcal{T}_t\}$ is a semigroup: $\mathcal{T}_t \circ \mathcal{T}_s = \mathcal{T}_{t+s}$, $s,t \geq 0$.

(ii) $\{\mathcal{T}_t\}$ is nearly Feller: For all $t > 0$, $f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ and $\{z_n\} \subset \mathbb{U}_{\downarrow}^{0}$, such that f is continuous at all points in $\mathbb{U}_{\downarrow}^{1}$ and $z_n \to z$ for some $z \in \mathbb{U}_{\downarrow}^{1}$, we have $\mathcal{T}_t f(z_n) \to \mathcal{T}_t f(z)$.

(iii) The marginal distribution of $\boldsymbol{Z}(\lambda)$ is characterized through the parabolic reflected Brownian motion $\hat{W}_{\lambda}$: For each $\lambda \in \mathcal{R}$, $\boldsymbol{Z}(\lambda)$ has the probability distribution ν_{λ}.

(iv) The stochastic process $\{\boldsymbol{Z}(\lambda)\}$ satisfies the Markov property with semigroup $\{\mathcal{T}_t\}$: For all $f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$, and $\lambda_1 < \lambda_2$, we have

$$\mathbb{E}[f(\boldsymbol{Z}(\lambda_2)) | \{\boldsymbol{Z}(\lambda)\}_{\lambda\leq\lambda_1}] = (\mathcal{T}_{\lambda_2-\lambda_1} f)(\boldsymbol{Z}(\lambda_1)).$$

(v) If $f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ is such that $f(x,y) = g(x)$ for some $g \in \mathrm{BM}(l_{\downarrow}^{2})$, then

$$(\mathcal{T}_t f)(z) = (T_t g)(x), \ \ \forall z = (x,y) \in \mathbb{U}_{\downarrow}^{0}.$$

Furthermore, $\{\boldsymbol{X}(\lambda)\}_{-\infty<\lambda<\infty}$ is Aldous's standard multiplicative coalescent.

A precise definition of $\mathcal{T}_t$ can be found in Section 5.4. Theorem 5.3.1 will be proved in Section 5.5.

Throughout this work we fix $K \in \mathbb{N}_0$, $F \in \Omega_K^4$ and consider a F -BSR as introduced in Section 5.2.2.

The result below considers the asymptotics of the 'susceptibility functions'. For any given time t and fixed $k \geq 1$ define the k-susceptibility function

$$\mathcal{S}_k^{(n)}(t) \equiv \mathcal{S}_k(t) := \sum_{i \geq 1} \left| \mathcal{C}_i^{(n)}(t) \right|^k.$$

Define the scaled susceptibility functions by, for $k \geq 1$,

$$\bar{s}_k(t) := \frac{\mathcal{S}_k(t)}{n}. \tag{5.3.1}$$

Then Theorem 1.1 of [31] shows that for any bounded-size rule, there exists a monotonically increasing function $s_2 : [0, t_c) \to [0, \infty)$ satisfying $s_2(0) = 1$ and $\lim_{t \uparrow t_c} s_2(t) = \infty$, such that

$$\bar{s}_2(t) \xrightarrow{\mathbb{P}} s_2(t) \quad \forall t \in [0, t_c).$$

Part (iii) of the following result gives a similar result for $\bar{s}_3$. Part (ii) in fact gives convergence of $\bar{s}_2, \bar{s}_3$ in a stronger sense. Part (i) of the theorem gives precise asymptotics of $s_2(t)$ and $s_3(t)$ as $t \uparrow t_c$.

Theorem 5.3.2. There exist monotonically increasing functions $s_k : [0, t_c) \to [0, \infty)$, $k = 2, 3$, such that $s_2(0) = s_3(0) = 1$ and $\lim_{t \uparrow t_c} s_2(t) = \lim_{t \uparrow t_c} s_3(t) = \infty$, having the following properties.

(i) There exist $\alpha, \beta \in (0, \infty)$ such that

$$s_2(t) = (1 + O(t_c - t)) \frac{\alpha}{t_c - t}, \quad s_3(t) = \beta [s_2(t)]^3 (1 + O(t_c - t)), \text{ as } t \uparrow t_c. \tag{5.3.2}$$

(ii) For every $\gamma \in (1/6, 1/5)$,

$$\sup_{t \in [0, t_n]} \left| \frac{n^{1/3}}{\bar{s}_2(t)} - \frac{n^{1/3}}{s_2(t)} \right| \xrightarrow{\mathbb{P}} 0 \tag{5.3.3}$$

$$\sup_{t \in [0, t_n]} \left| \frac{\bar{s}_3(t)}{(\bar{s}_2(t))^3} - \frac{s_3(t)}{(s_2(t))^3} \right| \xrightarrow{\mathbb{P}} 0, \tag{5.3.4}$$

where $t_n = t_c - n^{-\gamma}$.

(iii) For all $t \in [0, t_c)$, $\bar{s}_2(t) \xrightarrow{\mathbb{P}} s_2(t)$, $\bar{s}_3(t) \xrightarrow{\mathbb{P}} s_3(t)$ as $n \to \infty$.

We now state the main result which gives the asymptotic behavior in the critical scaling window as well as merging dynamics for all bounded-size rules.

Theorem 5.3.3 (**Bounded-size rules: Convergence at criticality**). Let $\alpha, \beta \in (0, \infty)$ be as in Theorem 5.3.2. For $\lambda \in \mathbb{R}$ define

$$\bar{\boldsymbol{C}}^{(n)}(\lambda) := \left(\frac{\beta^{1/3}}{n^{2/3}} \left| \mathcal{C}_i \left(t_c + \frac{\alpha\beta^{2/3}}{n^{1/3}} \lambda \right) \right| : i \geq 1 \right),$$
$$\bar{\boldsymbol{Y}}^{(n)}(\lambda) := \left(\xi_i \left(t_c + \frac{\alpha\beta^{2/3}}{n^{1/3}} \lambda \right) : i \geq 1 \right).$$

Then $\bar{\boldsymbol{Z}}^{(n)} = (\bar{\boldsymbol{C}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$ is a stochastic process with sample paths in $\mathcal{D}((-\infty, \infty) : \mathbb{U}_\downarrow)$ and for any set of times $-\infty < \lambda_1 < \lambda_2 < ... < \lambda_m < \infty$

$$\left(\bar{\boldsymbol{Z}}^{(n)}(\lambda_1), \ldots, \bar{\boldsymbol{Z}}^{(n)}(\lambda_m) \right) \xrightarrow{d} (\boldsymbol{Z}(\lambda_1), \ldots, \boldsymbol{Z}(\lambda_m)) \tag{5.3.5}$$

as $n \to \infty$, where $\boldsymbol{Z}$ is as in Theorem 5.3.1.

Organization of the proofs: The two main results in this chapter are Theorems 5.3.1 and 5.3.3. In Section 5.4 we introduce the semigroup $\{\mathcal{T}_t\}_{t \geq 0}$ and, as a first step towards Theorem 5.3.1, establish in Theorem 5.4.1 the existence of a $\mathbb{U}_\downarrow^0$ valued Markov process associated with this semigroup, starting from an arbitrary initial value. Then in Section 5.5 we complete the proof of Theorem 5.3.1. We then proceed to the analysis of bounded-size rules in Section 5.6 where we study the differential equation systems associated with the BSR process and prove Theorems 5.3.2. Finally in Section 5.7 we complete the proof of Theorem 5.3.3.

5.4 The augmented multiplicative coalescent

We begin by making precise the formal dynamics of the augmented multiplicative coalescent process given in Section 5.2.3.2. Fix $(x, y) \in \mathbb{U}_\downarrow^0$. Let $\{\xi_{i,j}\}_{i,j \in \mathbb{N}}$ be a collection of i.i.d. rate one Poisson processes. Let $\mathbf{G}(z, t)$, where $z = (x, y)$, be the random graph on vertex set $\mathbb{N}$ given as follows:

(I) For $i \in \mathbb{N}$, put y_i initial self-loops to the vertex i.

(II) For $i < j \in \mathbb{N}$, put $\xi_{i,j}([0, tx_ix_j/2]) + \xi_{j,i}([0, tx_ix_j/2])$ edges between vertices i and j. Also, for $i \in \mathbb{N}$, put additional $\xi_{i,i}([0, tx_i^2/2])$ self-loops to the vertex i.

Note that the total number of self-loops at a vertex i at time instant t is $y_i + \xi_{i,i}([0, tx_i^2/2])$. The self-loops coming from (I) and (II) will later be termed as "type I" and "type II" surplus.

Let $\mathcal{F}_t^x = \sigma\{\xi_{i,j}([0, sx_ix_j/2]) : 0 \leq s \leq t,\ i, j \in \mathbb{N}\}$, $t \geq 0$.

Recall the volume of a component $\mathcal{C}$ is defined to be $\mathrm{vol}(\mathcal{C}) = \sum_{i \in \mathcal{C}} x_i$. We have also defined surplus for finite graphs. For infinite graphs the definition requires some care. We define the surplus for a connected graph $\mathbf{G}$ with vertex set a subset of $\mathbb{N}$ as

$$\mathbf{spls}(\mathbf{G}) := \lim_{k \to \infty} \mathbf{spls}(\mathbf{G}^{[k]}),$$

where $\mathbf{G}^{[k]}$ is the **induced subgraph** that has the vertex set $[k]$ (the subgraph with vertex set $[k]$ and all edges between vertices in $[k]$ that are present in $\mathbf{G}$). It is easy to check that this definition of surplus does not depend on the labeling of the vertices. Further note that the surplus of a connected graph might be infinite with this definition.

Thus letting $\tilde{\mathcal{C}}_i(t)$ be the i-th largest component (in volume) in $\mathbf{G}(z,t)$, define $X_i(z,t) := \mathrm{vol}(\tilde{\mathcal{C}}_i(t))$ and $Y_i(z,t) := \mathbf{spls}(\tilde{\mathcal{C}}_i(t))$ to be the **volume** and the **surplus** of the i-th largest component at time t. In case two components have the same volume, the ordering of $(\tilde{\mathcal{C}}_i(t) : i \geq 1)$ is taken to be such that $Y_m(z,t) \geq Y_k(z,t)$ whenever $m \leq k$ and $X_m(z,t) = X_k(z,t)$.

Let $\boldsymbol{X}^z(t) := (X_i(z,t) : i \geq 1)$ and $\boldsymbol{Y}^z(t) := (Y_i(z,t) : i \geq 1)$. The paper [2] shows that $\boldsymbol{X}^z(t) \in l^2_{\downarrow}$ a.s. for all $t \geq 0$. The following result shows that $\boldsymbol{Z}^z(t) = (\boldsymbol{X}^z(t), \boldsymbol{Y}^z(t)) \in \mathbb{U}^0_{\downarrow}$ a.s., for all t.

Theorem 5.4.1. Fix $z = (x, y) \in \mathbb{U}^0_{\downarrow}$ and let $(\boldsymbol{X}^z(t), \boldsymbol{Y}^z(t))_{t \geq 0}$ be the stochastic process described above, then for any fixed $t \geq 0$, $(\boldsymbol{X}^z(t), \boldsymbol{Y}^z(t)) \in \mathbb{U}^0_{\downarrow}$.

The above theorem will be proved in Section 5.4.1. For $t \geq 0$, define $\mathcal{T}_t$:

$\mathrm{BM}(\mathbb{U}_{\downarrow}^{0}) \to \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ as

$$\mathcal{T}_t f(z) = \mathbb{E} f(\boldsymbol{Z}^z(t)),\ z \in \mathbb{U}_{\downarrow}^{0},\ f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0}).$$

The following result shows that $\{\mathcal{T}_t\}$ is a semigroup that is (nearly) Feller.

Theorem 5.4.2. For $t, s \geq 0$, $\mathcal{T}_t \circ \mathcal{T}_s = \mathcal{T}_{t+s}$. For all $t > 0$, $f \in \mathrm{BM}(\mathbb{U}_{\downarrow}^{0})$ and $\{z_n\} \subset \mathbb{U}_{\downarrow}^{0}$, such that f is continuous at all points in $\mathbb{U}_{\downarrow}^{1}$ and $z_n \to z$ for some $z \in \mathbb{U}_{\downarrow}^{1}$, we have $\mathcal{T}_t f(z_n) \to \mathcal{T}_t f(z)$.

The above theorem will be proved in Section 5.4.2. Throughout we will assume, without loss of generality, that for all $z \in \mathbb{U}_{\downarrow}^{0}$, $\boldsymbol{Z}^z$ is constructed using the same set of Poisson processes $\{\xi_{i,j}\}$. This coupling of $\boldsymbol{Z}^z$ for different values of z will not be noted explicitly in the statement of various results.

We begin with the following elementary lemma.

Lemma 5.4.3. Let $\{\mathcal{F}_m\}_{m \in \mathbb{N}_0}$ be a filtration given on some probability space.
(i) Let $\{Z_m\}_{m \geq 0}$ be a $\{\mathcal{F}_m\}$ adapted sequence of nondecreasing random variables such that $Z_0 = 0$. Let $\lim_{m \to \infty} Z_m = Z_\infty$. Suppose there exists a nonnegative random variable U such that $U < \infty$ a.s. and $\sum_{m=1}^{\infty} \mathbb{E}[Z_m - Z_{m-1} | \mathcal{F}_{m-1}] \leq U$. Then for any $\epsilon \in (0, 1)$,

$$\mathbb{P}\{Z_\infty > \epsilon\} \leq \frac{1+\epsilon}{\epsilon} \mathbb{E}[U \wedge 1].$$

(ii) Let $\{A_m\}$ be a sequence of events such that $A_m \in \mathcal{F}_m$. Suppose there exists a random variable $U < \infty$ a.s. such that $\sum_{m=1}^{\infty} \mathbb{E}[\mathbb{1}_{A_m} | \mathcal{F}_{m-1}] \leq U$. Then $\mathbb{P}\{A_m \text{ i.o.}\} = 0$. Furthermore,

$$\mathbb{P}\{\cup_{m=1}^{\infty} A_m\} \leq 2\mathbb{E}[U \wedge 1].$$

Proof: (i) Define $B_0 = 0$ and $B_m := \sum_{i=1}^{m} \mathbb{E}[Z_i - Z_{i-1} | \mathcal{F}_{i-1}]$ for $m = 1, 2, ...$ Note that B_m is nondecreasing and $\mathcal{F}_{m-1}$-measurable. Define $\tau = \inf\{l : B_{l+1} > 1\}$ where the infimum over an empty set is taken to be ∞. Since B_m is predictable, τ is a stopping time and, for all m, $B_{m\wedge\tau} \le 1$. Let $B_\infty = \lim_{m\to\infty} B_m$. Since $Z_{m\wedge\tau} - B_{m\wedge\tau}$ is a martingale, by the optimal stopping theorem and monotone convergence,

$$\mathbb{E}[Z_\tau] = \lim_{m\to\infty} \mathbb{E}[Z_{m\wedge\tau}] = \lim_{m\to\infty} \mathbb{E}[B_{m\wedge\tau}] \le \lim_{m\to\infty} \mathbb{E}[B_m \wedge 1] = \mathbb{E}[B_\infty \wedge 1].$$

Thus

$$\mathbb{P}\{Z_\infty > \epsilon\} \le \mathbb{P}\{\tau < \infty\} + \frac{1}{\epsilon}\mathbb{E}[B_\infty \wedge 1] = \mathbb{P}\{B_\infty > 1\} + \frac{1}{\epsilon}\mathbb{E}[B_\infty \wedge 1] \le \frac{1+\epsilon}{\epsilon}\mathbb{E}[U \wedge 1].$$

(ii) The first statement is immediate from the Borel-Cantelli lemma (cf. [15, Theorem 5.3.2]). For the second statement note that for any $\epsilon \in (0,1)$, we have $\cup_{m=1}^\infty A_m = \{\sum_{m=1}^\infty \mathbb{1}_{A_m} > \epsilon\}$. Now applying part (i) to $Z_m = \sum_{k=1}^m \mathbb{1}_{A_k}$ and taking $\epsilon \to 1$ yields the desired result. ■

Next, we present a result from [2] that will be used here. We begin with some notation. For $x \in l^2_\downarrow$, we write $x^{[k]} = (x_1, ..., x_k, 0, 0, ...)$ for the k-truncated version of x. Similarly, for a sequence $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, ...)$ of elements in $l^2_\downarrow$, $x^{(n)[k]}$ is the k-truncation of $x^{(n)}$. For $z = (x, y), z^{(n)} = (x^{(n)}, y^{(n)}) \in \mathbb{U}^0_\downarrow$ $z^{[k]}, y^{[k]}, z^{(n)[k]}, y^{(n)[k]}$ are defined similarly.

Recall the construction of $\mathbf{G}(z,t)$ described in items (I) and (II) at the beginning of the section. We will distinguish the surplus created in $\tilde{\mathcal{C}}_i(t)$ by the action in item (I) and that in item (II). The former will be referred to as the type I surplus and denoted by $\tilde{Y}_i(z,t)$ while the latter will be referred to as the type II surplus and denoted by $\hat{Y}_i(z,t) \equiv \hat{Y}_i(x,t)$. More precisely,

$$\tilde{Y}_i(z,t) = \sum_{j \in \tilde{\mathcal{C}}_i(t)} y_j \quad \text{and} \quad \hat{Y}_i(z,t) = Y_i(z,t) - \tilde{Y}_i(z,t).$$

Also define

$$\tilde{R}(z,t) := \sum_{i=1}^\infty X_i(z,t)\tilde{Y}_i(z,t), \quad \hat{R}(x,t) \equiv \hat{R}(z,t) := \sum_{i=1}^\infty X_i(z,t)\hat{Y}_i(z,t)$$

and

$$R(z,t) := \sum_{i=1}^\infty X_i(z,t)Y_i(z,t), \quad S(x,t) \equiv S(z,t) := \sum_{i=1}^\infty (X_i(x,t))^2.$$

The following properties of S and $\boldsymbol{X}$ have been established in [2, Proposition 5, Corollary 18, Lemma 22].

Theorem 5.4.4. (i) For every $x \in l^2_\downarrow$ and $t \geq 0$, we have $S(x,t) < \infty$ a.s. and $S(x^{[k]}, t) \uparrow S(x,t)$ as $k \to \infty$.
(ii) If $x^{(n)} \to x$ in $l^2_\downarrow$, then $\boldsymbol{X}(x^{(n)}, t) \xrightarrow{\mathbb{P}} \boldsymbol{X}(x,t)$ in $l^2_\downarrow$, as $n \to \infty$. In particular, $\{S(x^{(n)}, t)\}_{n \geq 1}$ is tight.

5.4.1 Existence of the AMC

This section proves Theorem 5.4.1. We begin by considering the type I surplus.

Proposition 5.4.5. For any $t \geq 0$ and $z \in \mathbb{U}^0_\downarrow$, $\tilde{R}(z,t) = \sum_{i=1}^\infty X_i(z,t)\tilde{Y}_i(z,t) < \infty$ a.s.

Proof of Proposition 5.4.5 is given below Lemma 5.4.7. The basic idea is to bound the truncated version $\tilde{R}^{[k]} = \tilde{R}(z^{[k]}, t)$ using a martingale argument, and then let $k \to \infty$. The truncation error is controlled using Lemma 5.4.6 below and a suitable supermartingale is constructed in Lemma 5.4.7.

Lemma 5.4.6. For every $z \in \mathbb{U}^0_\downarrow$ and $t \geq 0$, as $k \to \infty$, $\tilde{R}(z^{[k]}, t) \to \tilde{R}(z,t) \leq \infty$ a.s.

Proof: Fix $t \geq 0$. Denote by E_{ij} [resp. $E^{[k]}_{ij}$] the event that there exists a path from i to j in $\mathbf{G}(z,t)$ [resp. $\mathbf{G}(z^{[k]}, t)$], with the convention that $\mathbb{P}\{E_{ii}\} = \mathbb{P}\{E^{[k]}_{ii}\} = 1$. Let

$$f_i = \sum_{j=1}^\infty y_j \mathbb{1}_{E_{ij}}, \quad f^{[k]}_i = \sum_{j=1}^k y_j \mathbb{1}_{E^{[k]}_{ij}}.$$

Then

$$\tilde{R}(z,t) = \sum_{i=1}^\infty f_i x_i, \quad \tilde{R}(z^{[k]}, t) = \sum_{i=1}^\infty f^{[k]}_i x_i.$$

Since $E^{[k]}_{ij} \uparrow E_{ij}$, we have $f^{[k]}_i \uparrow f_i$. The result now follows from an application of monotone convergence theorem. ■

Lemma 5.4.7. Suppose that $z = (x, y) = z^{[k]}$ for some $k \geq 1$ and that $\sum_j y_j \neq 0$. Then

$$A_t \equiv A(z,t) := \log \tilde{R}(z,t) - \int_0^t S(z,u)du$$

is a supermartingale with respect to the filtration $\mathcal{F}_t^x = \sigma\{\xi_{i,j}([0, sx_ix_j/2]); 0 \leq s \leq t,\ i,j \in \mathbb{N}\}$.

Proof: From the construction of $\boldsymbol{Z}(z,\cdot)$ we see that $\tilde{R}(z,t)$ is a pure jump, nondecreasing process that at any time instant t, jumps at rate $X_i(z,t-)X_j(z,t-)$, $1 \leq i < j \leq k$, with jump sizes $B_{ij}(t-) = X_i(z,t-)\tilde{Y}_j(z,t-) + X_j(z,t-)\tilde{Y}_i(z,t-)$. Consequently $\log \tilde{R}(z,t)$ jumps at the same rate, with corresponding jump size $\log(1+\frac{B_{ij}(t-)}{\tilde{R}(z,t-)})$. From this and elementary properties of Poisson processes it follows that

$$\log \tilde{R}(z,t) = \log \tilde{R}(z,0) + \sum_{1\leq i<j\leq k} \int_0^t \log\left(1 + \frac{B_{ij}(u)}{\tilde{R}(z,u)}\right) X_i(z,u)X_j(z,u)du + M(t),$$

where M is a $\mathcal{F}_t^x$ martingale. Consequently, for $0 \leq s < t < \infty$

$$\begin{aligned}&\log \tilde{R}(z,t) - \log \tilde{R}(z,s)\\ &= \sum_{1\leq i<j\leq k} \int_s^t \log\left(1 + \frac{B_{ij}(u)}{\tilde{R}(z,u)}\right) X_i(z,u)X_j(z,u)du + M(t) - M(s).\end{aligned} \tag{5.4.1}$$

Next note that, for $u \geq 0$

$$\begin{aligned}&\sum_{1\leq i<j\leq k} \log\left(1 + \frac{B_{ij}(u)}{\tilde{R}(z,u)}\right) X_i(z,u)X_j(z,u)\\ &\leq \sum_{1\leq i<j\leq k} \frac{B_{ij}(u)}{\tilde{R}(z,u)} X_i(z,u)X_j(z,u)\\ &= \sum_{1\leq i<j\leq k} \frac{X_i(z,u)\tilde{Y}_j(z,u) + X_j(z,u)\tilde{Y}_i(z,u)}{\tilde{R}(z,u)} X_i(z,u)X_j(z,u)\\ &\leq S(z,u).\end{aligned}$$

Using this observation in (5.4.1) we now have

$$\mathbb{E}\left[\log \tilde{R}(z,t) - \log \tilde{R}(z,s) \mid \mathcal{F}_s^x\right] \leq \mathbb{E}\left[\int_s^t S(z,u)du \mid \mathcal{F}_s^x\right].$$

The result follows. ■

Proof of Proposition 5.4.5: Fix $z = (x, y) \in \mathbb{U}_{\downarrow}^0$. The result is trivially true if $\sum_i y_i = 0$. Assume now that $\sum_i y_i \neq 0$. For $k \geq 1$ and $a \in (0, \infty)$, define $T_a^{[k]} = \inf\{s \geq 0 : S(z^{[k]}, s) \geq a\}$. Fix $k \geq 1$ and assume without loss of generality that $\sum_{i=1}^k y_i > 0$. Write $R^{[k]}(t) = R(z^{[k]}, t)$, and $A^{[k]}(t) = A(z^{[k]}, t)$ where A is as in Lemma 5.4.7. From the supermartingale property $\mathbb{E}[A^{[k]}(T_a^{[k]} \wedge t)] \leq \mathbb{E}[A^{[k]}(0)] = \log \tilde{R}^{[k]}(0)$. Therefore

$$\mathbb{E}\left[\log \frac{\tilde{R}^{[k]}(T_a^{[k]} \wedge t)}{\tilde{R}^{[k]}(0)}\right] \leq \mathbb{E}\left[\int_0^{T_a^{[k]} \wedge t} S(z^{[k]}, u) du\right] \leq ta.$$

Thus

$$\begin{aligned}\mathbb{P}\{\tilde{R}^{[k]}(t) > m\} \leq & \mathbb{P}\{\tilde{R}^{[k]}(t) > m, T_a^{[k]} > t\} + \mathbb{P}\{T_a^{[k]} \leq t\} \\ \leq & \frac{ta}{\log m - \log \tilde{R}^{[k]}(0)} + \mathbb{P}\{T_a^{[k]} \leq t\}.\end{aligned}$$

By Lemma 5.4.6, $\tilde{R}^{[k]}(t) \to \tilde{R}(z, t)$, and by Theorem 5.4.4 (i), $S(z^{[k]}, t) \to S(z, t)$ when $k \to \infty$. Therefore letting $k \to \infty$ on both sides of the above inequality, we have

$$\mathbb{P}\{\tilde{R}(z, t) > m\} \leq \frac{ta}{\log m - \log \tilde{R}(z, 0)} + \mathbb{P}\{S(z, t) \geq a\}. \tag{5.4.2}$$

The result now follows on first letting $m \to \infty$ and then $a \to \infty$ in the above inequality. ■

The following result is an immediate consequence of the estimate in (5.4.2) and Theorem 5.4.4(ii).

Corollary 5.4.8. If $z^{(n)} \to z$ in $\mathbb{U}_{\downarrow}^0$, then for every $t \geq 0$, $\{\tilde{R}(z^{(n)}, t)\}_{n \geq 1}$ is tight.

Next we consider the type II surplus. Let, for $x \in l_{\downarrow}^2$

$$\mathcal{G}_t^x := \sigma\{\{\xi_{i,j}([0, s x_i x_j/2]) = 0\} : \ 0 \leq s \leq t, \ i, j \in \mathbb{N}\}.$$

The σ-field $\mathcal{G}_t^x$ records whether or not i and j are in the same component at time s, for all i, j and for all $s \leq t$. In particular, components $\{\tilde{\mathcal{C}}_i(s), i \geq 1, s \leq t\}$ can be determined from the information in $\mathcal{G}_t^x$ and consequently, $\boldsymbol{X}(x,t)$ is $\mathcal{G}_t^x$ measurable.

Lemma 5.4.9. (i) Fix $x \in l^2_\downarrow$ and $t \geq 0$. Then $\hat{R}(x,t) < \infty$ a.s.

(ii) Let $x^{(n)} \to x$ in $l^2_\downarrow$. Then the sequence $\{\hat{R}(x^{(n)},t)\}_{n\geq 1}$ is tight.

Proof: Note that (i) is an immediate consequence of (ii). Consider now (ii). For fixed $x \in l^2_\downarrow$ and $t \geq 0$, let $\hat{\mu}_i(x,t)$ denote the conditional law of $\hat{Y}_i(x,t)$, conditioned on $\mathcal{G}_t^x$. Then, for a.e. ω, $\hat{\mu}_i(x,t)$ is Poisson distribution with parameter

$$\int_0^t \sum_{j=1}^{\infty} \Big(\sum_{k,k' \in \tilde{\mathcal{C}}_j(s)} \frac{1}{2} x_k x_{k'} \Big) \mathbb{1}_{\{\tilde{\mathcal{C}}_j(s) \subset \tilde{\mathcal{C}}_i(t)\}} ds$$

$$= \int_0^t \frac{1}{2} \sum_{j=1}^{\infty} (X_j(x,s))^2 \mathbb{1}_{\{\tilde{\mathcal{C}}_j(s) \subset \tilde{\mathcal{C}}_i(t)\}} ds \leq \frac{t}{2} (X_i(x,t))^2,$$

where the last inequality is a consequence of the inequality $\sum_{j:\tilde{\mathcal{C}}_j(s)\subset\tilde{\mathcal{C}}_i(t)} (X_j(x,s))^2 \leq (X_i(x,t))^2$. Therefore $\hat{\mu}_i(x,t) \leq_d \hat{\nu}_i(x,t)$, a.s., where $\hat{\nu}_i(x,t)$ is a random probability measure on $\mathbb{N}$ such that for a.e. ω, $\hat{\nu}_i(x,t)$ is Poisson distribution with parameter $\frac{t}{2}(X_i(x,t,\omega))^2$.

A similar argument shows that the conditional distribution of $\sum_{i=1}^{\infty} \hat{Y}_i(x,t)$, given $\mathcal{G}_t^x$ is a.s. stochastically dominated by a random measure on $\mathbb{N}$ that, for a.e. ω has a Poisson distribution with parameter $\sum_{i=1}^{\infty} \frac{t}{2}(X_i(x,t,\omega))^2 = \frac{t}{2} S(x,t)$. Also, if $x^{(n)}$ is a sequence converging to x in $l^2_\downarrow$, we have that for each n, the conditional distribution of $\sum_{i=1}^{\infty} \hat{Y}_i(x^{(n)},t)$, given $\mathcal{G}_t^{x^{(n)}}$ is a.s. stochastically dominated by a Poisson random variable with parameter $\frac{t}{2} S(x^{(n)},t)$. From Theorem 5.4.4(ii), $\{S(x^{(n)},t)\}_{n\geq 1}$ is tight. Combining these facts we have that $\{\sum_{i=1}^{\infty} \hat{Y}_i(x^{(n)},t)\}_{n\geq 1}$ is a tight family. Finally, note that $\hat{R}(x^{(n)},t) \leq X_1(x^{(n)},t) \left(\sum_{i=1}^{\infty} \hat{Y}_i(x^{(n)},t) \right)$. The tightness of $\{\hat{R}(x^{(n)},t)\}_{n\geq 1}$ now follows on combining the above established tightness of $\{\sum_{i=1}^{\infty} \hat{Y}_i(x^{(n)},t)\}_{n\geq 1}$ and the tightness of $\{X_1(x^{(n)},t)\}_{n\geq 1}$, where the latter is once again a consequence of Theorem 5.4.4(ii). ■

We now complete the proof of Theorem 5.4.1.

Proof of Theorem 5.4.1. Fix $z = (x, y) \in \mathbb{U}_{\downarrow}^{0}$ and $t \geq 0$. From Lemma 5.4.9 (i) $\hat{R}(x,t) < \infty$ a.s. Also, from Proposition 5.4.5, $\tilde{R}(z,t) < \infty$ a.s. The result now follows on recalling that $R(z,t) = \hat{R}(x,t) + \tilde{R}(z,t)$. ■

We also record the following consequence of Lemma 5.4.9 and Corollary 5.4.8 for future use.

Corollary 5.4.10. If $z^{(n)} \to z$ in $\mathbb{U}_{\downarrow}^{0}$, then $\{R(z^{(n)}, t)\}_{n \geq 1}$ is tight.

5.4.2 Feller property of the AMC

In this section, we will prove Theorem 5.4.2. In fact we will show that if $z^{(n)} = (x^{(n)}, y^{(n)})$ converges to $z = (x, y)$ in $\mathbb{U}_{\downarrow}^{0}$, and $z \in \mathbb{U}_{\downarrow}^{1}$, then

$$(\boldsymbol{X}(z^{(n)}, t), \boldsymbol{Y}(z^{(n)}, t)) \xrightarrow{\mathbb{P}} (\boldsymbol{X}(z, t), \boldsymbol{Y}(z, t)). \tag{5.4.3}$$

We start with the following elementary lemma.

Lemma 5.4.11. Suppose $(x, y), (x^{(n)}, y^{(n)}) \in \mathbb{U}_{\downarrow}$ for $n \geq 1$. Then

$$\lim_{n\to\infty} \mathbf{d}_{\mathbb{U}}((x, y), (x^{(n)}, y^{(n)})) = 0$$

if and only if the following three conditions hold:

(i) $\lim_{n\to\infty} \sum_{i=1}^{\infty} (x_i^{(n)} - x_i)^2 = 0$.

(ii) $y_i^{(n)} = y_i$ for n sufficiently large, for all $i \geq 1$.

(iii) $\lim_{n\to\infty} \sum_{i=1}^{\infty} x_i^{(n)} y_i^{(n)} = \sum_{i=1}^{\infty} x_i y_i$.

Proof: The "only if" part is immediate. To see the "if" part, note that the first two conditions imply $\lim_{n\to\infty} x_i^{(n)} y_i^{(n)} = x_i y_i$ for all $i \geq 1$. By the third condition and Scheffe's lemma, we now have $\lim_{n\to\infty} \sum_i |x_i^{(n)} y_i^{(n)} - x_i y_i| = 0$. The result follows. ■

The key ingredient in the proof is the following lemma the proof of which is given after Lemma 5.4.14.

Lemma 5.4.12. Let $z^{(n)} = (x^{(n)}, y^{(n)})$ converge to $z = (x, y)$ in $\mathbb{U}^0_\downarrow$. Suppose that $z \in \mathbb{U}^1_\downarrow$. Then

(i) $Y_i(z^{(n)}, t) \xrightarrow{\mathbb{P}} Y_i(z, t)$ for all $i \geq 1$.

(ii) $\sum_{i=1}^\infty X_i(z^{(n)}, t) Y_i(z^{(n)}, t) \xrightarrow{\mathbb{P}} \sum_{i=1}^\infty X_i(z, t) Y_i(z, t)$.

Proof of Theorem 5.4.2 can now be completed as follows.

Proof of Theorem 5.4.2. The first part of the theorem is immediate from the construction given at the beginning of Section 5.4 and elementary properties of Poisson processes. For the second part, consider $z^{(n)} = (x^{(n)}, y^{(n)})$, $z = (x, y)$ as in the statement of the theorem. It suffices to prove (5.4.3). From Theorem 5.4.4(ii), $\boldsymbol{X}(z^{(n)}, t) \to \boldsymbol{X}(z, t)$ in probability, in $l^2_\downarrow$. The result now follows on combining this convergence with the convergence in Lemma 5.4.12 (on noting that $(\boldsymbol{X}(z, t), \boldsymbol{Y}(z, t)) \in \mathbb{U}^1_\downarrow$ a.s.) and applying Lemma 5.4.11. ∎

Rest of this section is devoted to the proof of Lemma 5.4.12. The key idea of the proof is as follows. Consider the induced subgraphs on the first k vertices $\mathbf{G}^{[k]} = \mathbf{G}(z^{[k]}, t)$ and $\mathbf{G}^{(n)[k]} = \mathbf{G}(z^{(n)[k]}, t)$. Since there are only finite number of vertices in $\mathbf{G}^{[k]}$, when $n \to \infty$, $\mathbf{G}^{(n)[k]}$ will eventually be identical to $\mathbf{G}^{[k]}$ almost surely. The main step in the proof is to control the difference between $\mathbf{G}^{(n)[k]}$ and $\mathbf{G}^{(n)}$ when k is large, uniformly for all n. For this we first analyze the difference between $\mathbf{G}^{(n)[k]}$ and $\mathbf{G}^{(n)[k+1]}$ in the lemma below.

Consider the set of vertices $[k+1] = \{1, 2, ..., k, k+1\}$, and for every $i \in [k+1]$, let vertex i have label (x_i, y_i) representing its size and surplus, respectively. Suppose $x_1 \geq x_2 \geq ... \geq x_{k+1}$. Fix $t > 0$. Define a random graph $\mathbf{G}^*$ on the above vertex set as follows. For $i \leq k$, the number of edges, N_i, between i and $k+1$ is distributed as Poisson(tx_ix_{k+1}). In addition, there are $N_0 = \text{Poisson}(tx_{k+1}^2/2)$ self-loops to the vertex $k+1$. All the Poisson random variables are taken to be mutually independent. Denote X_i and Y_i for the component volumes and surplus of the resulting star-like graph if i is the smallest labeled vertex in its component; otherwise let $X_i = Y_i = 0$. A precise definition of (X_i, Y_i) is as follows. Write $i \sim k+1$ if there is an edge between i and $k+1$ in $\mathbf{G}^*$. By convention $(k+1) \sim (k+1)$. Let $\mathcal{J}_k = \{i \in [k+1] : i \sim k+1\}$, and $i_0 = \min\{i : i \in \mathcal{J}_k\}$. Then

$$(X_i, Y_i) = \begin{cases} \left(\sum_{i\in\mathcal{J}_k} x_i, \sum_{i\in\mathcal{J}_k} y_i\right) & \text{if } i = i_0 \\ \\ (0,0) & \text{if } i \in \mathcal{J}_k \setminus \{i_0\} \\ \\ (x_i, y_i) & \text{if } i \in [k+1] \setminus \mathcal{J}_k. \end{cases}$$

Define $R_k = \sum_{i=1}^k x_iy_i$, $S_k = \sum_{i=1}^k x_i^2$, $R_{k+1} = \sum_{i=1}^{k+1} X_iY_i$. Then we have the following result.

Lemma 5.4.13. (i) $\mathbb{P}\{Y_i \neq y_i\} \leq tx_{k+1}y_{k+1}x_1 + tx_{k+1}^2\left(1 + itx_1^2 + tS_k + tR_kx_1\right)$.
(ii) $\mathbb{E}[R_{k+1} - R_k] \leq x_{k+1}y_{k+1}(1 + tS_k) + x_{k+1}^2(tR_k + t^2S_kR_k + t^2S_kx_1) + tx_{k+1}^3(1 + 2tS_k + t^2S_k^2)$.

Proof: (i) It is easy to see that, for $i = 1, \cdots k$,

$$\{Y_i \neq y_i\} \subset (\{y_{k+1} > 0\} \cap \{i \in \mathcal{J}_k\}) \cup \{N_0 \neq 0\} \cup_{j=1}^k \{N_j > 1\}$$
$$\cup_{j<i} \{N_jN_i \neq 0\} \cup_{j:y_j>0} \{N_jN_i \neq 0\}.$$

Using the observation that for a Poisson(λ) random variable Z, $\mathbb{P}\{Z \geq 1\} < \lambda$ and $\mathbb{P}\{Z \geq 2\} < \lambda^2$, we now have that

$$\begin{aligned}\mathbb{P}\{Y_i \neq y_i\} \leq & tx_i x_{k+1} \cdot y_{k+1} + \frac{tx_{k+1}^2}{2} + \sum_{j=1}^{k} (tx_j x_{k+1})^2 \\ & + \sum_{j=1}^{i-1} tx_j x_{k+1} \cdot tx_i x_{k+1} + \sum_{j=1}^{k} tx_j x_{k+1} \cdot tx_i x_{k+1} \cdot y_j.\end{aligned}$$

The proof is now completed on collecting all the terms and using the fact that $x_i \leq x_1$ for every i.

(ii) Note that

$$X_0 = x_{k+1} + \sum_{j=1}^{k} x_j \mathbb{1}_{\{N_j \geq 1\}},\ Y_0 = y_{k+1} + \sum_{j=1}^{k} y_j \mathbb{1}_{\{N_j \geq 1\}} + N_0 + \sum_{j=1}^{k} (N_j - 1)^+.$$

Then

$$\begin{aligned}& R_{k+1} - R_k \\ =& X_0 Y_0 - \sum_{j \in \mathcal{J}_k} x_j y_j \\ =& x_{k+1} y_{k+1} + \sum_{j=1}^{k} (x_j y_{k+1} + x_{k+1} y_j) \mathbb{1}_{\{N_j \geq 1\}} + \sum_{1 \leq j < l \leq k} (x_j y_l + x_l y_j) \mathbb{1}_{\{N_j \geq 1\}} \mathbb{1}_{\{N_l \geq 1\}} \\ +& N_0 X_0 + x_{k+1} \sum_{j=1}^{k} (N_j - 1)^+ + \sum_{j=1}^{k} x_j (N_j - 1)^+ \\ +& \sum_{1 \leq j < l \leq k} (x_j \mathbb{1}_{\{N_j \geq 1\}} (N_l - 1)^+ + x_l \mathbb{1}_{\{N_l \geq 1\}} (N_j - 1)^+).\end{aligned}$$

The result now follows on taking expectations in the above equation and using the fact that $\mathbb{E}[(N_j - 1)^+] < (tx_j x_{k+1})^2$. ■

Recall that, by construction, $X_i(z,t) \geq X_{i+1}(z,t)$ for all $z \in \mathbb{U}_\downarrow$, $t \geq 0$ and $i \in \mathbb{N}$. The following lemma which is a key ingredient in the proof of Lemma 5.4.12 says that if $z \in \mathbb{U}_\downarrow^1$, ties do not occur, a.s.

Lemma 5.4.14. Let $z \in \mathbb{U}_\downarrow^1$. Then for every $t > 0$ and $i \in \mathbb{N}$, $X_i(z,t) > X_{i+1}(z,t)$ a.s.

Proof: Fix $t > 0$. Consider the graph $\mathbf{G}(z,t)$ and write $\mathcal{C}_{x_i} \equiv \mathcal{C}_{x_i}(t)$ for the component of vertex (x_i, y_i) at time t. It suffices to show for all $i \neq j$

$$\mathbb{P}\left\{|\mathcal{C}_{x_i}| = |\mathcal{C}_{x_j}|, \mathcal{C}_{x_i} \neq \mathcal{C}_{x_j}\right\} = 0. \tag{5.4.4}$$

The key property we shall use is that for $z = (x, y) \in \mathbb{U}_{\downarrow}^1$, $\sum_{i=1}^{\infty} x_i = \infty$. Now fix $i \geq 1$. It is enough to show that $|\mathcal{C}_{x_i}|$ has no atom i.e for all $(x, y) \in \mathbb{U}_{\downarrow}^1$

$$\mathbb{P}(|\mathcal{C}_{x_i}| = a) = 0, \qquad \text{for any } a \geq 0. \tag{5.4.5}$$

To see this, first note that since $|\mathcal{C}_{x_i}| < \infty$ a.s., conditional on $\mathcal{C}_{x_i}$ the vector $z^* = ((x_k, y_k) : x_k \notin \mathcal{C}_{x_i}) \in \mathbb{U}_{\downarrow}^1$ almost surely. Thus on the event $x_j \notin \mathcal{C}_{x_i}$, conditional on $\mathcal{C}_{x_i}$, using (5.4.5) with $a = |\mathcal{C}_{x_i}|$ implies that $\mathbb{P}(|\mathcal{C}_{x_j}| = |\mathcal{C}_{x_i}| \mid \mathcal{C}_{x_i}) = 0$ and this completes the proof. Thus it is enough to prove (5.4.5). For the rest of the argument, to ease notation let $i = 1$. Let us first show the simpler assertion that the volume of direct neighbors of x_1 has a continuous distribution. More precisely, let $N_{i,j}(t) := \xi_{i,j}([0, tx_ix_j/2]) + \xi_{j,i}([0, tx_ix_j/2])$, $1 \leq i < j$, denote the number of edges between any two vertices x_i and x_j by time t. Then the volume of *direct* neighbors of the vertex x_1 is $L := \sum_{i=2}^{\infty} x_i \mathbb{1}_{\{N_{1,i}(t) \geq 1\}}$ and we will first show that L has no atom, namely

$$\mathbb{P}(L = a) = 0, \qquad \text{for all } a \geq 0. \tag{5.4.6}$$

For any random variable X define the maximum atom size of X by

$$\mathbf{atom}(X) := \sup_{a \in \mathbb{R}} \mathbb{P}\left\{X = a\right\}.$$

For two independent random variables X_1 and X_2 we have $\mathbf{atom}(X_1 + X_2) \leq \min\left\{\mathbf{atom}(X_1), \mathbf{atom}(X_2)\right\}$. For $m \geq 2$, define $L_m = \sum_{i=m}^{\infty} x_i \mathbb{1}_{\{N_{1,i}(t) \geq 1\}}$. Since L_m and $L - L_m$ are independent, we have $\mathbf{atom}(L) \leq \mathbf{atom}(L_m)$. Define the event

$$E_m := \left\{N_{1,i}(t) \leq 1 \text{ for all } i \geq m\right\},$$

and write

$$L_m^*(t) := \sum_{i=m}^{\infty} x_i N_{1,i}(t).$$

Then $L_m^*(t)$ is a pure jump Levy process with Levy measure $\nu(du) = \sum_{i=m}^{\infty} x_1 x_i \delta_{x_i}(du)$. By [19], such a Levy process has continuous marginal distribution since the Levy measure is infinite $(\nu(0,\infty) = (\sum_{i=m}^{\infty} x_i)x_1 = \infty)$. Thus $L_m^*(t)$ has no atom. Next, for any $a \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\{L_m = a\} \leq & \mathbb{P}\{E_m^c\} + \mathbb{P}\{E_m, L_m = a\} = \mathbb{P}\{E_m^c\} + \mathbb{P}\{E_m, L_m^*(t) = a\} \\ \leq & \sum_{i=m}^{\infty} \frac{(tx_1x_i)^2}{2} + 0 = \frac{t^2x_1^2}{2}\sum_{i=m}^{\infty} x_i^2.\end{aligned}$$

Thus $\mathbf{atom}(L) \leq \mathbf{atom}(L_m) \leq \frac{t^2x_1^2}{2}\sum_{i=m}^{\infty} x_i^2$. Since m is arbitrary, we have $\mathbf{atom}(L) = 0$. Thus L is a continuous variable, and (5.4.6) is proved.

Let us now strengthen this to prove (5.4.5). Let $\tilde{\mathbf{G}}$ be the subgraph of $\mathbf{G}(z,t)$ obtained by deleting the vertex x_1 and all related edges. Let $\tilde{X}_i$ be the volume of the i-th largest component of $\tilde{\mathbf{G}}$. Note that $\sum_{i=1}^{\infty} \tilde{X}_i = \sum_{i=2}^{\infty} x_i = \infty$ a.s. Conditional on $(\tilde{X}_i)_{i\geq 1}$, let $\tilde{N}_{1,i}$ have Poisson distribution with parameter $tx_1\tilde{X}_i$. Then

$$\mathcal{C}_{x_1} \overset{d}{=} x_1 + \sum_{i=1}^{\infty} \tilde{X}_i \mathbb{1}_{\{\tilde{N}_{1,i}\geq 1\}},$$

where the second term has the same form as the random variable L. Using (5.4.6) completes the proof. ■

We now proceed to the proof of Lemma 5.4.12.

Proof of Lemma 5.4.12. Fix $t > 0$ and $z^{(n)}, z$ as in the statement of the lemma. Denote $Y^{[k]} = Y(z^{[k]}, t)$, $Y^{(n)[k]} = Y(z^{(n)[k]}, t)$. Similarly, denote $\mathcal{C}_i^{[k]}$ and $\mathcal{C}_i^{(n)[k]}$ for the corresponding i-th largest component; and $X_i^{[k]}$ and $X_i^{(n)[k]}$ for their respective sizes. Also, write $X^{(n)} = X(x^{(n)}, t)$ and define $Y^{(n)}, R^{(n)}, S^{(n)}$ similarly.

For $i \in \mathbb{N}$, define the event $E_i^{(n)[k]}$ as,

$$E_i^{(n)[k]} := \{\omega : X_j^{(n)[k]}(\omega) > X_{j+1}^{(n)}(\omega), \text{ for } j = 1, 2, ..., i\},$$

and define $E_i^{[k]}$ similarly. Then

$$\begin{aligned}\mathbb{P}\{Y_i^{(n)} \neq Y_i(t)\} \leq & \mathbb{P}\{Y_i^{(n)} \neq Y_i^{(n)[k]}\} + \mathbb{P}\{Y_i^{(n)[k]} \neq Y_i^{[k]}\} + \mathbb{P}\{Y_i^{[k]} \neq Y_i(t)\} \\ \leq & \mathbb{P}\{Y_i^{(n)} \neq Y_i^{(n)[k]}, E_i^{(n)[k]}\} + \mathbb{P}\{(E_i^{(n)[k]})^c\} \\ & + \mathbb{P}\{Y_i^{(n)[k]} \neq Y_i^{[k]}\} + \mathbb{P}\{Y_i^{[k]} \neq Y_i(t)\}. \end{aligned} \tag{5.4.7}$$

Note that

$$E_i^{(n)[k]} \subset \{\omega : \mathcal{C}_j^{(n)[k]}(\omega) \subset \mathcal{C}_j^{(n)[m]}(\omega) \subset \mathcal{C}_j^{(n)}(\omega), \text{ for all } j = 1, 2, ..., i \text{ and } m \geq k\}.$$

Thus the probability of the event $\{Y_i^{(n)[m+1]} \neq Y_i^{(n)[m]}, E_i^{(n)[k]}\}$, for $m \geq k$, can be estimated using Lemma 5.4.13 (i). More precisely, let $\mathcal{F}^{[m]} = \sigma\{\xi_{i,j}; i, j \leq m\}$ for $m \geq 1$. Then by Lemma 5.4.13 (i),

$$\begin{aligned} & \mathbb{P}\{Y_i^{(n)[m+1]} \neq Y_i^{(n)[m]}, E_i^{(n)[k]} | \mathcal{F}^{[m]}\} \\ \leq & t x_{m+1}^{(n)} y_{m+1}^{(n)} X_1^{(n)[m]} + t (x_{m+1}^{(n)})^2 \left(1 + it (X_1^{(n)[m]})^2 + t S^{(n)[m]} + t R^{(n)[m]} X_1^{(n)[m]}\right), \end{aligned}$$

where $S^{(n)[m]} = \sum_i (X_i^{(n)[m]})^2$ and $R^{(n)[m]} = \sum_i (X_i^{(n)[m]} Y_i^{(n)[m]})$.

Note that $X_1^{(n)[k]} \leq X_1^{(n)}$, $R^{(n)[k]} \leq R^{(n)}$ and $S^{(n)[k]} \leq S^{(n)}$. Thus we have

$$\begin{aligned} & \sum_{m=k}^{\infty} \mathbb{P}\{Y_i^{(n)[m+1]} \neq Y_i^{(n)[m]}, E_i^{(n)[k]} | \mathcal{F}^{[m]}\} \\ \leq & t \left(\sum_{m=k+1}^{\infty} x_m^{(n)} y_m^{(n)} \right) X_1^{(n)} + t \left(\sum_{m=k+1}^{\infty} (x_m^{(n)})^2 \right) \left(1 + it (X_1^{(n)})^2 + t S^{(n)} + t R^{(n)} X_1^{(n)}\right). \end{aligned}$$

Denote the right hand side of the above inequality as $U^{(n)[k]}$. Then by Lemma 5.4.3(ii), we have

$$\begin{aligned}\mathbb{P}\{Y_i^{(n)} \neq Y_i^{(n)[k]}, E_i^{(n)[k]}\} = & \mathbb{P}\left(\cup_{m=k}^{\infty} \{Y_i^{(n)[m+1]} \neq Y_i^{(n)[m]}, E_i^{(n)[k]}\}\right) \\ \leq & 2\mathbb{E}[U^{(n)[k]} \wedge 1] \end{aligned} \tag{5.4.8}$$

and therefore

$$\mathbb{P}\{Y_i^{(n)} \neq Y_i(t)\} \leq 2\mathbb{E}[U^{(n)[k]} \wedge 1] + \mathbb{P}\{(E_i^{(n)[k]})^c\} + \mathbb{P}\{Y_i^{(n)[k]} \neq Y_i^{[k]}\} + \mathbb{P}\{Y_i^{[k]} \neq Y_i(t)\}. \tag{5.4.9}$$

Next note that $X_1^{(n)}$, $S^{(n)}$ and $R^{(n)}$ are all tight sequences by Corollary 5.4.10 and Theorem 5.4.4(ii). Thus $(1 + it(X_1^{(n)})^2 + tS^{(n)} + tR^{(n)}X_1^{(n)})$ is also tight. Also, since $z^{(n)} \to z$,

$$\limsup_{k\to\infty}\limsup_{n\to\infty}\sum_{i=k+1}^{\infty} x_i^{(n)} y_i^{(n)} = 0 \text{ and } \limsup_{k\to\infty}\limsup_{n\to\infty}\sum_{i=k+1}^{\infty} (x_i^{(n)})^2 = 0.$$

Combining the above observations we have that $\limsup_{k\to\infty}\limsup_{n\to\infty}\mathbb{P}\{U^{(n)[k]} > \epsilon\} = 0$ for all $\epsilon > 0$. From the inequality

$$\mathbb{E}[U^{(n)[k]} \wedge 1] \le \mathbb{P}\{U^{(n)[k]} > \epsilon\} + \epsilon$$

we now see that

$$\limsup_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}[U^{(n)[k]} \wedge 1] = 0. \tag{5.4.10}$$

Next, from a straightforward extension of Proposition 5 of Aldous [2] we have that $(\boldsymbol{X}^{(n)}, X_1^{(n)[k]}, ..., X_i^{(n)[k]}) \xrightarrow{d} (\boldsymbol{X}(t), X_1^{[k]}, ..., X_i^{[k]})$ in $l^2_\downarrow \times \mathbb{R}^i$ when $n \to \infty$, for each fixed i and k. Combining this with Lemma 5.4.14 we now see that for fixed i

$$\limsup_{k\to\infty}\limsup_{n\to\infty}\mathbb{P}\{(E_i^{(n)[k]})^c\} = 0.$$

Also, for each fixed k

$$\limsup_{n\to\infty}\mathbb{P}\{Y_i^{(n)[k]} \neq Y_i^{[k]}\} = 0.$$

Observing that $\lim_{k\to\infty} Y_i^{[k]} = Y_i(t)$ and the last term in (5.4.9) does not depend on n, we have that

$$\limsup_{k\to\infty}\limsup_{n\to\infty}\mathbb{P}\{Y_i^{[k]} \neq Y_i(t)\} = 0.$$

Part (i) of the lemma now follows on combining the above observations and taking limit as $n \to \infty$ and then $k \to \infty$ in (5.4.9).

We now prove part (ii) of the lemma. Note that

$$\liminf_{n\to\infty} R^{(n)} \ge \lim_{n\to\infty} R^{(n)[k]} = R^{[k]}.$$

With a similar argument as in Lemma 5.4.6, we have $R^{[k]} \to R(z,t)$ as $k \to \infty$. Thus sending $k \to \infty$ in the above display we have

$$\liminf_{n\to\infty} R^{(n)} \geq R(z,t). \tag{5.4.11}$$

To complete the proof, it suffices to show that

$$\text{For any } \epsilon > 0,\ \lim_{n\to\infty} \mathbb{P}\{R^{(n)} > R(z,t) + \epsilon\} = 0. \tag{5.4.12}$$

Note that

$$\begin{aligned}\mathbb{P}\{R^{(n)} - R(z,t) > \epsilon\} \leq & \mathbb{P}\{R^{(n)} - R^{(n)[k]} > \epsilon/2\} + \mathbb{P}\{R^{(n)[k]} - R(z,t) > \epsilon/2\} \\ \leq & \mathbb{P}\{R^{(n)} - R^{(n)[k]} > \epsilon/2\} + \mathbb{P}\{R^{(n)[k]} - R^{[k]} > \epsilon/2\}. \end{aligned} \tag{5.4.13}$$

The second term on the right side above goes to zero for each fixed k, as $n \to \infty$. For the first term, note that by Lemma 5.4.13(ii), for all $m \geq k$

$$\mathbb{E}[R^{(n)[m+1]} - R^{(n)[m]} | \mathcal{F}^{[m]}] \leq x_m^{(n)} y_m^{(n)} U_1^{(n)} + (x_m^{(n)})^2 U_2^{(n)} + (x_{m+1}^{(n)})^3 U_3^{(n)},$$

where $U_1^{(n)} = 1 + tS^{(n)}$, $U_2^{(n)} = tR^{(n)} + t^2 S^{(n)} R^{(n)} + t^2 S^{(n)} X_1^{(n)}$ and $U_3^{(n)} = t(1 + 2tS^{(n)} + t^2 (S^{(n)})^2)$. Thus by Lemma 5.4.3 (i),

$$\mathbb{P}\{R^{(n)} - R^{(n)[k]} > \epsilon\} \leq (1 + 1/\epsilon)\mathbb{E}[U^{(n)[k]} \wedge 1],$$

where $U^{(n)[k]} = (\sum_{m=k+1}^{\infty} x_m^{(n)} y_m^{(n)}) U_1^{(n)} + (\sum_{m=k+1}^{\infty} (x_m^{(n)})^2) U_2^{(n)} + (\sum_{m=k+1}^{\infty} (x_{m+1}^{(n)})^3) U_3^{(n)}$. Note that $U_1^{(n)}$, $U_2^{(n)}$ and $U_3^{(n)}$ are all tight sequences and $z^{(n)} \to z$. An argument similar to the one used to prove (5.4.10) now shows that, for all $\epsilon > 0$,

$$\limsup_{k\to\infty} \limsup_{n\to\infty} \mathbb{P}\{R^{(n)} - R^{(n)[k]} > \epsilon\} \leq \left(1 + \frac{1}{\epsilon}\right) \limsup_{k\to\infty} \limsup_{n\to\infty} \mathbb{E}[U^{(n)[k]} \wedge 1] = 0.$$

The statement in (5.4.12) now follows on using the above convergence in (5.4.13) and combining it with the observation below (5.4.13). This completes the proof of part (ii). ■

Remark 5.4.15. Lemma 5.4.12 is at the heart of the (near) Feller property in Theorem 5.4.2 which is crucial for the proof of the joint convergence in (5.3.5). The proof of the lemma reveals the reason for considering the metric $\mathbf{d}_{\mathbb{U}}$ on $\mathbb{U}_{\downarrow}$.

One natural metric on $\mathbb{U}_{\downarrow}$, denoted by $\mathbf{d}_1$, is the one obtained by replacing the second term in (5.2.5) with

$$\sum_{i=1}^{\infty} \frac{|y_i - y_i'|}{2^i} \wedge 1.$$

This metric corresponds to the topology on $\mathbb{U}_{\downarrow}$ inherited from $\ell^2 \times \mathbb{N}^{\infty}$ taking the topology generated by the inner product $\langle \cdot, \cdot \rangle$ on ℓ^2 and the product topology on $\mathbb{N}^{\infty}$; and then considering the product topology on $\ell^2 \times \mathbb{N}^{\infty}$.

Another metric (which we denote by $\mathbf{d}_2$) that can be considered on $\mathbb{U}_{\downarrow}$ corresponds to replacing the second term in (5.2.5) with $\mathbf{d}_{vt}(\mu_z, \mu_{z'})$, where $\mu_z = \sum_{i=1}^{\infty} \delta_{z_i}$, $\mu_{z'} = \sum_{i=1}^{\infty} \delta_{z_i'}$ and $\mathbf{d}_{vt}$ is the metric corresponding to the vague topology on the space of $\mathbb{N} \cup \{\infty\}$ valued locally finite measures on $(0, \infty) \times \mathbb{N}$.

The proof of Lemma 5.4.12 hinges upon the convergence of $\sum_{m=1}^{\infty} x_m^{(n)} y_m^{(n)}$ to $\sum_{m=1}^{\infty} x_m y_m$, as $n \to \infty$, even for the proof of convergence of $Y_i(z^{(n)}, t) \xrightarrow{\mathbb{P}} Y_i(z, t)$. Since $\mathbf{d}_1$ and $\mathbf{d}_2$ give no control over sums of the form $\sum_{m=1}^{\infty} x_m y_m$, this suggests that the convergence in $\mathbf{d}_1$ or $\mathbf{d}_2$ is "too weak" to yield the desired Feller property.

5.5 The standard augmented multiplicative coalescent.

In this section we prove Theorem 5.3.1. The Proposition 4 of [2] proves a very useful result on convergence of component size vectors of a general family of non-uniform random graph models to the ordered excursion lengths of $\hat{W}_{\lambda}$. We begin in this section by extending this result to the joint convergence of component size and component surplus vectors in $\mathbb{U}_{\downarrow}$, under a slight strengthening of the conditions assumed in [2].

Recall the excursion lengths and mark count process $\boldsymbol{Z}^*(\lambda) = (\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda))$ defined in Section 5.2.3.2. Our first result below shows that, for fixed $\lambda \in \mathbb{R}$, $\boldsymbol{Z}^*(\lambda)$ arises as a limit of $\boldsymbol{Z}(z^{(n)}, q^{(n)})$ in $\mathbb{U}_\downarrow$ for all sequences $\{z^{(n)}\} \subset \mathbb{U}_\downarrow$ and $q^{(n)} = q_\lambda^{(n)} \subset (0, \infty)$ that satisfy certain regularity conditions.

For $n \geq 1$, let $z^{(n)} = (x^{(n)}, y^{(n)}) \in \mathbb{U}_\downarrow^0$. Writing $z_i^{(n)} = (x_i^{(n)}, y_i^{(n)})$, $i \geq 1$, define

$$x^{*(n)} = \sup_{i \geq 1} x_i^{(n)}, \; s_r^{(n)} = \sum_{i=1}^{\infty} (x_i^{(n)})^r, \; r \geq 1.$$

Note that $x^{*(n)} = x_1^{(n)}$ since the sequence is ordered. Let $\{q^{(n)}\}$ be a nonnegative sequence. We will suppress (n) from the notation unless needed.

Theorem 5.5.1. Let $z^{(n)} = (z_1^{(n)}, \cdots) \in \mathbb{U}_\downarrow^0$ be such that $x_i^{(n)} = 0$ for all $i > n$ and $y_i^{(n)} = 0$ for all $i \geq 1$. Suppose that, as $n \to \infty$,

$$\frac{s_3}{(s_2)^3} \to 1, \quad q - \frac{1}{s_2} \to \lambda, \quad \frac{x^*}{s_2} \to 0, \tag{5.5.1}$$

and, for some $\varsigma \in (0, \infty)$,

$$s_1 \cdot \left(\frac{x^*}{s_2} \right)^\varsigma \to 0. \tag{5.5.2}$$

Then $\boldsymbol{Z}^{(n)} = \boldsymbol{Z}(z^{(n)}, q^{(n)})$ converges in distribution in $\mathbb{U}_\downarrow$ to $\boldsymbol{Z}^*(\lambda)$.

Remark: The convergence assumption in (5.5.1) is the same as that in Proposition 4 of [2]. The additional assumption in (5.5.2) is not very stringent as will be seen in Section 5.7 when this result is applied to a general family of bounded-size rules.

Given Theorem 5.5.1, the proof of Theorem 5.3.1 can now be completed as follows.

Proof of Theorem 5.3.1. The first two parts of the theorem were shown in Theorem 5.4.2. Also, part (v) of the theorem is immediate from the definition of $\{T_t\}$ in Section 5.2.3.1. Recall the definition of ν_λ from Section 5.2.3.2. In order to prove parts (iii)-(iv) it suffices to show that

$$\text{for any } \lambda_1, \lambda_2 \in \mathbb{R},\ \lambda_1 \leq \lambda_2,\ \nu_{\lambda_1} \mathcal{T}_{\lambda_2 - \lambda_1} = \nu_{\lambda_2}. \tag{5.5.3}$$

Indeed, using the semigroup property of $(\mathcal{T}_\lambda)$ and the above relation, it is straightforward to define a consistent family of finite dimensional distributions $\mu_{\lambda_1, \cdots \lambda_k}$ on $(\mathbb{U}_\downarrow)^{\otimes k}$, $-\infty < \lambda_1 < \lambda_2, \cdots \lambda_k < \infty$, $k \geq 1$, such that $\mu_\lambda = \nu_\lambda$ for every $\lambda \in \mathbb{R}$. The desired result then follows from Kolmogorov's consistency theorem.

We now prove (5.5.3). Let

$$z^{(n)} = (x^{(n)}, y^{(n)}),\ x_i^{(n)} = n^{-2/3},\ y_i^{(n)} = 0,\ i = 1, \cdots n,\ q_{\lambda_j}^{(n)} = \lambda_j + n^{1/3},\ j = 1, 2.$$

We set $z_i^{(n)} = 0$ for $i > n$. Note that with this choice of $x^{(n)}$, $s_1 = n^{1/3}, s_2 = n^{-1/3}, s_3 = n^{-1}$ and so clearly (5.5.1) and (5.5.2) (with any $\varsigma > 1$) are satisfied with $q = q_{\lambda_j}$, $\lambda = \lambda_j$, $j = 1, 2$. Thus, denoting the distribution of $\boldsymbol{Z}(z^{(n)}, q_{\lambda_j}^{(n)})$ by $\nu_{\lambda_j}^{(n)}$, we have by Theorem 5.5.1 that

$$\nu_{\lambda_j}^{(n)} \to \nu_{\lambda_j},\ \text{as } n \to \infty. \tag{5.5.4}$$

Also, from the construction of $\boldsymbol{Z}(z, t)$ in Section 5.4, it is clear that $\nu_{\lambda_1}^{(n)} \mathcal{T}_{\lambda_2 - \lambda_1} = \nu_{\lambda_2}^{(n)}$. The result now follows on combining the convergence in (5.5.4) with Theorem 5.4.2 and observing that $\boldsymbol{Z}^*(\lambda) \in \mathbb{U}_\downarrow^1$ a.s. for every $\lambda \in \mathbb{R}$. ∎

Rest of this section is devoted to the proof of Theorem 5.5.1 and is organized as follows. Recall the random graph process $\mathbf{G}(z, q)$, for $z \in \mathbb{U}_\downarrow$, $q \geq 0$, defined at the beginning of Section 5.4. In Section 5.5.1 we will give an equivalent in law construction of $\mathbf{G}(z, q)$, from [2], that defines the random graph simultaneously with a certain breadth-first-exploration random walk. The excursions of the reflected version of this walk encode the component sizes of the random graph while the area under the excursions gives the parameter of the Poisson distribution describing the (conditional) law of the surplus associated with the corresponding component. Using this construction, in Theorem 5.5.2, we will first prove a weaker result than Theorem 5.5.1 which proves the convergence in distribution of $\boldsymbol{Z}^{(n)}$ to $\boldsymbol{Z}^*(\lambda)$ in $l^2_\downarrow \times \mathbb{N}^\infty$, where we consider the product topology on $\mathbb{N}^\infty$. This result is proved in Section 5.5.2. In Section 5.5.3 we will give the proof of Theorem 5.5.1 using Theorem 5.5.2 and an auxiliary tightness lemma (Lemma 5.5.4). Finally, proof of Lemma 5.5.4 is given in Section 5.5.4.

5.5.1 Breadth-first-exploration random walk

In this section, following [2], we will give an equivalent in law construction of $\mathbf{G}(z,q)$ that defines the random graph simultaneously with a certain breadth-first-exploration random walk. Given $q \in (0,\infty)$ and $z \in \mathbb{U}^0_\downarrow$ such that $x_i = 0$ for all $i > n$ and $y_i = 0$ for all i, we will construct a random graph $\bar{\mathbf{G}}(z,q)$ that is equivalent in law to $\mathbf{G}(z,q)$, in two stages, as follows. We begin with a graph on $[n]$ with no edges. Let $\{\eta_{i,j}\}_{i,j\in\mathbb{N}}$ be independent Poisson point processes on $[0,\infty)$ such that η_{ij} for $i \neq j$ has intensity qx_j; and for $i = j$ has intensity $qx_i/2$.

Stage I: The breadth-first-search forest and associated random walk: Choose a vertex $v(1) \in [n]$ with $\mathbb{P}(v(1) = i) \propto x_i$. Let

$$\mathbb{I}_1 = \{j \in [n] : j \neq v(1) \text{ and } \eta_{v(1),j} \cap [0, x_{v(1)}] \neq \emptyset\}.$$

Form an edge between $v(1)$ and each $j \in \mathbb{I}_1$. Let $c(1) = |\mathbb{I}_1|$. Let $m_{v(1),j}$ be the first point in $\eta_{v(1),j}$ for each $j \in \mathbb{I}_1$. Order the vertices in $\mathbb{I}_1$ according to increasing values of $m_{v(1),j}$ and label these as $v(2), \cdots v(c(1)+1)$. Let

$$\mathcal{V}_1 = \{v(1)\},\ \mathcal{N}_1 = \{v(2), \cdots, v(c(1)+1)\},\ l_1 = x_{v(1)} \text{ and } d_1 = c(1).$$

Having defined $\mathcal{V}_{i'}$, $\mathcal{N}_{i'}$, $l_{i'}$, $d_{i'}$ and the edges up to step i', with $\mathcal{V}_{i'} = \{v(1), \cdots v(i')\}$, $\mathcal{N}_{i'} = \{v(i'+1), v(i'+2), \cdots v(d_{i'}+1)\}$ for $1 \leq i' \leq i-1$, define, if $\mathcal{N}_{i-1} \neq \emptyset$

$$\mathbb{I}_i = \{j \in [n] : j \notin \mathcal{N}_{i-1} \cup \mathcal{V}_{i-1} \text{ and } \eta_{v(i),j} \cap [0, x_{v(i)}] \neq \emptyset\}$$

and form an edge between $v(i)$ and each $j \in \mathbb{I}_i$. Let $c(i) = |\mathbb{I}_i|$ and let $m_{v(i),j}$ be the first point in $\eta_{v(i),j}$ for each $j \in \mathbb{I}_i$. Order the vertices in $\mathbb{I}_i$ according to increasing values of $m_{v(i),j}$ and label these as $v(d_{i-1}+2), \cdots v(d_i+1)$, where $d_i = d_{i-1} + c(i)$. Set

$$l_i = l_{i-1} + x_{v(i)},\ \mathcal{V}_i = \{v(1), \cdots v(i)\},\ \mathcal{N}_i = \{v(i+1), v(i+2), \cdots v(d_i+1)\}.$$

In case $\mathcal{N}_{i-1} = \emptyset$, we choose $v(i) \in [n] \setminus \mathcal{V}_{i-1}$ with probability proportional to x_j, $j \in [n] \setminus \mathcal{V}_{i-1}$ and define $\mathbb{I}_i, c(i), d_i, l_i, \mathcal{V}_i, \mathcal{N}_i$ and the edges at step i exactly as above.

This procedure terminates after exactly n steps at which point we obtain a forest-like graph with no surplus edges. We will include surplus to this graph in stage II below.

Associate with the above construction an (interpolated) random walk process $H^{(n)}(\cdot)$ defined as follows. $H^{(n)}(0) = 0$ and

$$H^{(n)}(l_{i-1}+u) = H^{(n)}(l_{i-1}) - u + \sum_{j \notin \mathcal{V}_i \cup \mathcal{N}_{i-1}} x_j \mathbb{1}_{\{m_{v(i),j} < u\}} \quad \text{for } 0 < u < x_{v(i)},\ i = 1, \cdots n, \tag{5.5.5}$$

where by convention $l_0 = 0$ and $\mathcal{N}_0 = \emptyset$. This defines $H^{(n)}(t)$ for all $t \in [0, l_n)$. Define $H^{(n)}(t) = H^{(n)}(l_n-)$ for all $t \geq l_n$.

Stage II: Construction of surplus edges: For each $i = 1, \cdots, n$, we construct surplus edges on the graph obtained in Stage I and a point process $\mathcal{P}_x$ on $[0, l_n]$, simultaneously, as follows.

(i) For each $v \in \mathbb{I}_i$ and $\tau \in \eta_{v(i),v} \cap [0, x_{v(i)}] \setminus \{m_{v(i),v}\}$, construct an edge between $v(i)$ and v. This corresponds to multi-edges between the two vertices $v(i)$ and v.

(ii) For each $\tau \in \eta_{v(i),v(i)} \cap [0, x_{v(i)}]$, construct an edge between $v(i)$ and itself. This corresponds to self-loops at the vertex $v(i)$.

(iii) For each $v(j) \in \mathcal{N}_{i-1} \setminus \{v(i)\}$ and $\tau \in \eta_{v(i),v(j)} \cap [0, x_{v(i)}]$, construct an edge between $v(i)$ and $v(j)$. This corresponds to additional edges between two vertices,

$v(i)$ and $v(j)$, that were indirectly connected in stage I.

For each of the above cases, we also construct points for the point process $\mathcal{P}_x$ at time $l_{i-1}+\tau \in [0, l_n]$.

This completes the construction of the graph $\bar{\mathbf{G}}(z,q)$ and the random walk $H^{(n)}(\cdot)$. This graph has the same law as $\mathbf{G}(z,q)$, so the associated component sizes and surplus vector denoted by $(\bar{\boldsymbol{X}}(z,q), \bar{\boldsymbol{Y}}(z,q))$ has the same law as that of $(\boldsymbol{X}(z,q), \boldsymbol{Y}(z,q))$. Furthermore, conditioned on $H^{(n)}$, $\mathcal{P}_x$ is Poisson point process on $[0, l_n]$ whose intensity we denote by $r_x(t)$.

Using the above construction we will show in next section, as a first step, a weaker result than Theorem 5.5.1.

5.5.2 Convergence in $l^2_\downarrow \times \mathbb{N}^\infty$.

The following is the main result of this section.

Theorem 5.5.2. Let $z^{(n)} \in \mathbb{U}^0_\downarrow$ and $q^{(n)} \in (0,\infty)$ be sequences that satisfy the conditions in Theorem 5.5.1. Then

$$(\boldsymbol{X}^{(n)}, \boldsymbol{Y}^{(n)}) \xrightarrow{d} (\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda)) \tag{5.5.6}$$

in the space $l^2_\downarrow \times \mathbb{N}^\infty$ as $n \to \infty$, where we consider the product topology on $\mathbb{N}^\infty$.

The key ingredient in the proof is the following result. With $z^{(n)}$ and $q^{(n)}$ as in the above theorem, define $\bar{\boldsymbol{X}}^{(n)} = \bar{\boldsymbol{X}}(z^{(n)}, q^{(n)})$, $\bar{\boldsymbol{Y}}^{(n)} = \bar{\boldsymbol{Y}}(z^{(n)}, q^{(n)})$ and $r^{(n)}(t) = r_{x^{(n)}}(t)\mathbb{1}_{[0,l_n]}(t)$, $t \geq 0$. Denote the random walk process from Section 5.5.1 constructed using $(x^{(n)}, q^{(n)})$ (rather than (x,q)), once more, by $H^{(n)}(\cdot)$.

Define the rescaled process $\bar{H}^{(n)}(\cdot)$ and its reflected version $\hat{H}^{(n)}(\cdot)$ as follows

$$\bar{H}^{(n)}(t) := \sqrt{\frac{s_2}{s_3}} H^{(n)}(t), \quad \hat{H}^{(n)}(t) := \bar{H}^{(n)}(t) - \min_{0 \leq u \leq t} \bar{H}^{(n)}(u). \tag{5.5.7}$$

Lemma 5.5.3. (i) As $n \to \infty$, the process $\bar{H}^{(n)} \xrightarrow{d} W_\lambda$ in $\mathcal{D}([0,\infty):\mathbb{R})$.

(ii) For $n \geq 1$,

$$\sup_{t\geq 0}\left| r^{(n)}(t) - q\sqrt{\frac{s_3}{s_2}}\hat{H}^{(n)}(t)\right| \leq \frac{3}{2}qx_*. \tag{5.5.8}$$

Given Lemma 5.5.3, the proof of Theorem 5.5.2 can be completed as follows.

Proof of Theorem 5.5.2: The paper [2] shows that the vector $\bar{\boldsymbol{X}}^{(n)}$ can be represented as the ordered sequence of excursion lengths of the process $\hat{H}^{(n)}$. Also, weak convergence of $\bar{H}^{(n)}$ to W_λ in Lemma 5.5.3 (i) implies the convergence of $\hat{H}^{(n)}$ to $\hat{W}_\lambda$. Using these facts, Proposition 4 of [2] shows that $\bar{\boldsymbol{X}}^{(n)}$ converges in distribution to the ordered excursion length sequence of $\hat{W}_\lambda$, namely $\boldsymbol{X}^*(\lambda)$, in $l^2_\downarrow$. Also, conditional on $\hat{H}^{(n)}$, $\mathcal{P}_x$ is a Poisson point process on $[0,\infty)$ with rate $r^{(n)}(t)$ and for $i \geq 1$, $\bar{Y}_i^{(n)}$ has a Poisson distribution with parameter $\int_{[a_i^{(n)},b_i^{(n)}]} r^{(n)}(s)ds$, where $a_i^{(n)}, b_i^{(n)}$ are the left and right endpoints of the i-th ordered excursion of $\hat{H}^n$. From conditions in (5.5.1) it follows that $x^* \to 0$ and $s_2 \to 0$, further more we have $qx^* \to 0$ and $q\sqrt{s_3/s_2} \to 1$. Lemma 5.5.3 (ii) then shows that $\int_{[a_i^{(n)},b_i^{(n)}]} r^{(n)}(s)ds$ converges in distribution to $\int_{[a_i,b_i]} \hat{W}_\lambda(s)ds$, where a_i, b_i are the left and right endpoints of the i-th ordered excursion of $\hat{W}_\lambda$. In fact we have the joint convergence of $\left(\hat{H}^{(n)}, \left(\int_{[a_i^{(n)},b_i^{(n)}]} r^{(n)}(s)ds\right)_{i\geq 1}\right)$ to $\left(\hat{W}_\lambda, \left(\int_{[a_i,b_i]} \hat{W}_\lambda(s)ds\right)_{i\geq 1}\right)$. This proves the convergence of $(\bar{\boldsymbol{X}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$ to $(\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda))$ in $l^2_\downarrow \times \mathbb{N}^\infty$. The result follows since $(\bar{\boldsymbol{X}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$ has the same law as $(\boldsymbol{X}^{(n)}, \boldsymbol{Y}^{(n)})$. ■

Proof of Lemma 5.5.3 Part (i) was proved in Proposition 4 of [2]. Consider now (ii).

For $j = 1, 2, \dots$, define $\delta_{v(j)} = 1_{\{\mathcal{N}_{j-1}=\emptyset\}}$, i.e. $\delta_{v(j)}$ is 1 if $v(j)$ is the first vertex that is explored in its component during the breadth-first-search, and is 0 otherwise. It is easy to verify that $H^{(n)}$ satisfies

$$H^{(n)}(l_i) = -\sum_{j=1}^{i} \delta_{v(j)} x_{v(j)} + \sum_{v \in \mathcal{N}_i} x_v,\ i = 1, \cdots n.$$

The above equation implies that for all $k \leq i$, $H^{(n)}(l_k) \geq -\sum_{j=1}^{i} \delta_{v(j)} x_{v(j)}$. In addition, taking $k_0 = \sup\left\{j \leq i : \delta_{v(j)} = 1\right\}$ we have $H^{(n)}(l_{k_0}) = -\sum_{j=1}^{i} \delta_{v(j)} x_{v(j)}$. In particular, this implies that $\inf_{j \leq i} H^{(n)}(l_j) = -\sum_{j=1}^{i} \delta_{v(j)} x_{v(j)}$. Also, from (5.5.5) we have that for $t \in (l_{i-1}, l_i]$, $H^{(n)}(t) \geq H^{(n)}(l_{i-1}) - x^*$. Consequently

$$\left| \inf_{0 \leq u \leq t} H^{(n)}(u) + \sum_{j=1}^{i-1} \delta_{v(j)} x_{v(j)} \right| = \left| \inf_{0 \leq u \leq t} H^{(n)}(u) - \inf_{\{j : l_j \leq t\}} H^{(n)}(l_j) \right| \leq x^*. \qquad (5.5.9)$$

Let $\mathcal{N}_{i-1} = \{v(i), v(i+1), ..., v(i+l)\}$. From the above expression for $H^{(n)}(l_i)$, we have that for $t \in (l_{i-1}, l_i]$

$$H^{(n)}(t) = \left(-\sum_{j=1}^{i-1} \delta_{v(j)} x_{v(j)} + \sum_{j=i}^{i+l} x_{v(j)} \right) - (t - l_{i-1}) + \sum_{j \notin \mathcal{V}_i \cup \mathcal{N}_{i-1}} x_j \mathbb{1}_{\{m_{v(i),j} < t - l_{i-1}\}}, \qquad (5.5.10)$$

Also, accounting for the three sources of surplus described in Stage II of the construction, one has the following formula for $r^{(n)}(t)$ at time $t \in (l_{i-1}, l_i]$:

$$r^{(n)}(t) = q \cdot \left(\frac{x_{v(i)}}{2} + \sum_{j=i+1}^{i+l} x_{v(j)} + \sum_{j \notin \mathcal{V}_i \cup \mathcal{N}_{i-1}} x_j \mathbb{1}_{\{m_{v(i),j} < t - l_{i-1}\}} \right).$$

The three terms in the above expression correspond to self-loops; edges between vertices that in stage I were only connected indirectly; and additional edges between two vertices that were directly connected in stage I. Combining the above expression with (5.5.10) and (5.5.9), we have

$$\begin{aligned} &\left| r^{(n)}(t) - q \cdot \left(H^{(n)}(t) - \min_{0 \leq s \leq t} H^{(n)}(s) \right) \right| \\ \leq & q \cdot \left(\left| \inf_{0 \leq s \leq t} H^{(n)}(s) + \sum_{j=1}^{i-1} \delta_{v(j)} x_{v(j)} \right| + \frac{x_{v(i)}}{2} \right) \leq \frac{3}{2} q x_*. \end{aligned} \qquad (5.5.11)$$

The result follows. ■

5.5.3 Proof of Theorem 5.5.1.

In this section we complete the proof of Theorem 5.5.1. The key step in the proof is the following lemma whose proof is given in Section 5.5.4.

Lemma 5.5.4. Let $z^{(n)} \in \mathbb{U}_{\downarrow}^{0}$ and $q^{(n)} \in (0, \infty)$ be as in Theorem 5.5.1. Let $\hat{H}^{(n)}$ be as introduced in (5.5.7). Then $\{\sup_{t \geq 0} \hat{H}^{(n)}(t)\}_{n \geq 1}$ is a tight family of $\mathbb{R}_{+}$ valued random variables.

Remark 5.5.5. In fact one can establish a stronger statement, namely

$$\sup_{u \geq t} \sup_{n \geq 1} \hat{H}_{u}^{(n)} \xrightarrow{\mathbb{P}} 0$$

as $t \to \infty$. Also, although not used in this work, using very similar techniques as in the proof of Lemma 5.5.4, it can be shown that $\sup_{u \geq t} \hat{W}_{\lambda}(u)$ converges a.s. to 0, as $t \to \infty$.

Proof of Theorem 5.5.1. Since $(\boldsymbol{X}^{(n)}, \boldsymbol{Y}^{(n)})$ has the same distributions as $(\bar{\boldsymbol{X}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$, we can equivalently consider the convergence of the latter sequence. From Theorem 5.5.2 we have that $(\bar{\boldsymbol{X}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$ converges to $(\boldsymbol{X}^{*}(\lambda), \boldsymbol{Y}^{*}(\lambda))$, in distribution, in $l_{\downarrow}^{2} \times \mathbb{N}^{\infty}$ (with product topology on $\mathbb{N}^{\infty}$). By appealing to Skorohod representation theorem, we can assume without loss of generality that the convergence is almost sure. By the definition of $\mathbf{d}_{\mathbb{U}_{\downarrow}}$ it now suffices to argue that

$$\sum_{i=1}^{\infty} \left| \bar{X}_{i}^{(n)} \bar{Y}_{i}^{(n)} - X_{i}^{*}(\lambda) Y_{i}^{*}(\lambda) \right| \xrightarrow{\mathbb{P}} 0.$$

Fix $\epsilon > 0$. Then, for any $k \in \mathbb{N}$,

$$\begin{aligned} & \mathbb{P}\left\{ \sum_{i=1}^{\infty} \left| \bar{X}_{i}^{(n)} \bar{Y}_{i}^{(n)} - X_{i}^{*}(\lambda) Y_{i}^{*}(\lambda) \right| > \epsilon \right\} \\ \leq & \mathbb{P}\left\{ \sum_{i=1}^{k} \left| \bar{X}_{i}^{(n)} \bar{Y}_{i}^{(n)} - X_{i}^{*}(\lambda) Y_{i}^{*}(\lambda) \right| > \frac{\epsilon}{3} \right\} + \mathbb{P}\left\{ \sum_{i=k+1}^{\infty} \bar{X}_{i}^{(n)} \bar{Y}_{i}^{(n)} > \frac{\epsilon}{3} \right\} \\ & + \mathbb{P}\left\{ \sum_{i=k+1}^{\infty} X_{i}^{*}(\lambda) Y_{i}^{*}(\lambda) > \frac{\epsilon}{3} \right\}. \end{aligned} \tag{5.5.12}$$

From the convergence of $(\bar{\boldsymbol{X}}^{(n)}, \bar{\boldsymbol{Y}}^{(n)})$ to $(\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda))$ in $l^2_\downarrow \times \mathbb{N}^\infty$ we have that

$$\lim_{n\to\infty} \mathbb{P}\left\{\sum_{i=1}^{k} \left|\bar{X}_i^{(n)}\bar{Y}_i^{(n)} - X_i^*(\lambda)Y_i^*(\lambda)\right| > \frac{\epsilon}{3}\right\} = 0.$$

Consider now the second term in (5.5.12). Let $E_L^{(n)} = \{\sup_{t\geq 0} r_t^{(n)} \leq L\}$. Then

$$\mathbb{P}\left\{\sum_{i=k+1}^{\infty} \bar{X}_i^{(n)}\bar{Y}_i^{(n)} > \frac{\epsilon}{3}\right\} \leq \mathbb{P}\left\{(E_L^{(n)})^c\right\} + \frac{3}{\epsilon}\mathbb{E}\left[\mathbb{1}_{E_L^{(n)}}\left(\sum_{i=k+1}^{\infty} \bar{X}_i^{(n)}\bar{Y}_i^{(n)} \wedge 1\right)\right].$$

Let $\mathcal{G} = \sigma\{\hat{H}^{(n)}(t) : t \geq 0\}$. Since $r_t^{(n)}$ is $\mathcal{G}$ measurable for all $t \geq 0$, $E_L^{(n)} \in \mathcal{G}$. Then

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{E_L^{(n)}}\left(\sum_{i=k+1}^{\infty} \bar{X}_i^{(n)}\bar{Y}_i^{(n)} \wedge 1\right)\right] =&\mathbb{E}\left[\mathbb{1}_{E_L^{(n)}}\mathbb{E}\left[\sum_{i=k+1}^{\infty} \bar{X}_i^{(n)}\bar{Y}_i^{(n)} \wedge 1 \mid \mathcal{G}\right]\right] \\
\leq&\mathbb{E}\left[\mathbb{1}_{E_L^{(n)}}\left(\sum_{i=k+1}^{\infty} \mathbb{E}\left[\bar{X}_i^{(n)}\bar{Y}_i^{(n)} \mid \mathcal{G}\right] \wedge 1\right)\right] \\
\leq& L\mathbb{E}\left[\left(\sum_{i=k+1}^{\infty} (\bar{X}_i^{(n)})^2\right) \wedge 1\right],
\end{aligned}$$

where the last inequality follows on observing that, conditionally on $\mathcal{G}$, $\bar{Y}_i^{(n)}$ has a Poisson distribution with rate that is dominated by $\bar{X}_i^{(n)} \cdot (\sup_{t\geq 0} r_t^{(n)})$. Using the convergence of $\bar{\boldsymbol{X}}^{(n)}$ to $\boldsymbol{X}^*$, we now have

$$\limsup_{n\to\infty} \mathbb{E}\left[\mathbb{1}_{E_L^{(n)}}\left(\sum_{i=k+1}^{\infty} \bar{X}_i^{(n)}\bar{Y}_i^{(n)}\right) \wedge 1\right] \leq L\mathbb{E}\left[\left(\sum_{i=k+1}^{\infty} (X_i^*(\lambda))^2\right) \wedge 1\right].$$

Let $\delta > 0$ be arbitrary. Using Lemma 5.5.4 and Lemma 5.5.3 (ii) we can choose $L \in (0, \infty)$ such that $\mathbb{P}\left\{(E_L^{(n)})^c\right\} \leq \delta$. Finally, taking limit as $n \to \infty$ in (5.5.12) we have that

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{P}\left\{\sum_{i=1}^{\infty} \left|\bar{X}_i^{(n)}\bar{Y}_i^{(n)} - X_i^*(\lambda)Y_i^*(\lambda)\right| > \epsilon\right\} \\
\leq&\delta + L\mathbb{E}\left[\left(\sum_{i=k+1}^{\infty} (X_i^*(\lambda))^2\right) \wedge 1\right] + \mathbb{P}\left\{\sum_{i=k+1}^{\infty} X_i^*(\lambda)Y_i^*(\lambda) > \frac{\epsilon}{3}\right\}. \qquad (5.5.13)
\end{aligned}$$

The result now follows on sending $k \to \infty$ in the above display and recalling that $\sum_{i=1}^{\infty}(X_i^*(\lambda))^2 < \infty$ and $\sum_{i=1}^{\infty} X_i^*(\lambda)Y_i^*(\lambda) < \infty$ a.s. and $\delta > 0$ is arbitrary. ∎

5.5.4 Proof of Lemma 5.5.4.

In this section we prove Lemma 5.5.4. We will only treat the case $\lambda = 0$. The general case can be treated similarly. The key step in the proof is the following proposition whose proof is given at the end of the section.

Note that $\sup_{t\geq 0} |\bar{H}^{(n)}(t) - \bar{H}^{(n)}(t-)| \leq x^* \sqrt{s_2/s_3} \to 0$ as $n \to \infty$. Also, as $n \to \infty$, $qs_2 \to 1$. Thus, without loss of generality, we will assume that

$$\sup_{n\geq 1}\sup_{t\geq 0} |\bar{H}^{(n)}(t) - \bar{H}^{(n)}(t-)| \leq 1, \ \sup_{n\geq 1} q^{(n)} s_2^{(n)} \leq 2. \tag{5.5.14}$$

Fix $\vartheta \in (0, 1/2)$ and define $t^{*(n)} = \left(\frac{s_2}{x^*}\right)^{\vartheta}$. Denote by $\{\mathcal{F}_t^{(n)}\}$ the filtration generated by $\{\bar{H}^{(n)}(t)\}_{t\geq 0}$. For ease of notation, we write $\sup_{t\in[a,b]} = \sup_{[a,b]}$. We will suppress (n) in the notation, unless needed.

Proposition 5.5.6. There exist $\Theta \in (0, \infty)$, events $G^{(n)}$, increasing $\mathcal{F}_t^{(n)}$-stopping times $1 = \sigma_0^{(n)} < \sigma_1^{(n)} < ...$, and a real positive sequence $\{\kappa_i\}$ with $\sum_{i=1}^{\infty} \kappa_i < \infty$, such that the following hold:

(i) For every $i \geq 1$, $\left\{\sigma_i^{(n)}\right\}_{n\geq 1}$ is tight.

(ii) For every $i \geq 1$,

$$\mathbb{P}\left(\left\{\sup_{[\sigma_{i-1}^{(n)}, \sigma_i^{(n)}]} \hat{H}^{(n)}(t) > 2\Theta + 1\right\} \cap \left\{\sigma_{i-1}^{(n)} < t^{*(n)}\right\} \cap G^{(n)}\right) \leq \kappa_i.$$

(iii) As $n \to \infty$, $\mathbb{P}\left\{\sup_{[\sigma^{*(n)},\infty)} \hat{H}^{(n)}(t) > \Theta; G^{(n)}\right\} \to 0$, where $\sigma^{*(n)} = \inf\left\{\sigma_i^{(n)} : \sigma_i^{(n)} \geq t^{*(n)}\right\}$.

(iv) As $n \to \infty$, $\mathbb{P}(G^{(n)}) \to 1$.

Given Proposition 5.5.6, the proof of Lemma 5.5.4 can be completed as follows.

Proof of Lemma 5.5.4:

Fix $\epsilon \in (0, 1)$. Let $\Theta \in (0, \infty)$, $G^{(n)}$, $\sigma_i^{(n)}$, κ_i be as in Proposition 5.5.6. Choose $i_0 > 1$ such that $\sum_{i\geq i_0} \kappa_i \leq \epsilon$. Since $\left\{\sigma_{i_0-1}^{(n)}\right\}$ is tight, there exists $T \in (0, \infty)$ such that $\limsup_{n\to\infty} \mathbb{P}\left\{\sigma^{(n)}_{i_0-1} > T\right\} \le \epsilon$. Thus for any $M' > 2\Theta + 1$, we have

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{[1,\infty)} \hat{H}^{(n)}(t) > M'\right\} \\
\le &\mathbb{P}\left\{\sup_{[1,T]} \hat{H}^{(n)}(t) > M'\right\} + \mathbb{P}\left\{\sigma^{(n)}_{i_0-1} > T\right\} + \mathbb{P}\left\{(G^{(n)})^c\right\} \\
+&\mathbb{P}\left\{\sup_{[\sigma^{(n)}_{i_0-1}, \sigma^{*(n)}]} \hat{H}^{(n)}(t) > 2\Theta + 1; G^{(n)}\right\} + \mathbb{P}\left\{\sup_{[\sigma^{*(n)},\infty)} \hat{H}^{(n)}(t) > \Theta; G^{(n)}\right\}.
\end{aligned}$$

Taking $\limsup_{n\to\infty}$ on both sides

$$\limsup_{n\to\infty} \mathbb{P}\left\{\sup_{[1,\infty)} \hat{H}^{(n)}(t) > M'\right\} \le \limsup_{n\to\infty} \mathbb{P}\left\{\sup_{[1,T]} \hat{H}^{(n)}(t) > M'\right\} + \epsilon + 0 + \epsilon + 0.$$

Since $\left\{\sup_{[1,T]} \hat{H}^{(n)}(t)\right\}_{n\ge1}$ is tight, we have,

$$\limsup_{M'\to\infty} \limsup_{n\to\infty} \mathbb{P}\left\{\sup_{[1,\infty)} \hat{H}^{(n)}(t) > M'\right\} \le 2\epsilon.$$

Since $\epsilon > 0$ is arbitrary, the result follows. ■

We now proceed to the proof of Proposition 5.5.6. The following lemma is key.

Lemma 5.5.7. There are $\{\mathcal{F}^{(n)}_t\}$ adapted processes $\{A^{(n)}(t)\}$, $\{B^{(n)}(t)\}$ and $\mathcal{F}^{(n)}_t$-martingale $\{M^{(n)}(t)\}$ such that

(i) $A^{(n)}(\cdot)$ is a non-increasing function of t, a.s. For all $t \ge 0$, $\bar{H}^{(n)}(t) = \int_0^t A^{(n)}(u)du + M^{(n)}(t)$.

(ii) For $t \ge 0$, $\langle M^{(n)}, M^{(n)}\rangle_t = \int_0^t B^{(n)}(u)du$.

(iii) $\sup_{n\ge1} \sup_{u\ge0} B^{(n)}(u) \le 2$.

(iv) With $G^{(n)} = \left\{A(t) < -t/2 \text{ for all } t \in [1, t^{*(n)}]\right\}$, $\mathbb{P}(G^{(n)}) \to 1$ as $n \to \infty$.

(v) For any $\alpha \in (0,\infty)$ and $t > 0$,

$$\mathbb{P}\left\{\sup_{u\in[0,t]} |M^{(n)}(u)| > \alpha\right\} \le 2\exp\{\alpha\} \cdot \exp\left\{-\alpha \log\left(1 + \frac{\alpha}{2t}\right)\right\}. \tag{5.5.15}$$

Proof: Recall the notation from Section 5.5.1. Parts (i) and (ii) are proved in [2]. Furthermore, from Lemma 11 of [2] it follows that, for $t \in [l_{i-1}, l_i)$, writing

$Q_2(t) = \sum_{j=1}^{i}(x_{v(j)})^2$, we have

$$A(t) \leq \sqrt{\frac{s_2}{s_3}}(-1 + qs_2 - qQ_2(t)), \quad B(t) \leq qs_2.$$

Part (iii) now follows on recalling from (5.5.14) that $qs_2 \leq 2$. To prove (iv) it suffices to show that

$$\sup_{t \leq t^*} \left| \frac{s_2}{s_3} Q_2(t) - t \right| \xrightarrow{\mathbb{P}} 0. \tag{5.5.16}$$

To prove this we will use the estimate on Page 832, Lemma 13 of [2], which says that for any fixed $\epsilon \in (0,1)$, and $L \in (0, \infty)$

$$\begin{aligned} &\mathbb{P}\left\{ \sup_{t \in [0,L]} \left| \frac{s_2}{s_3} Q_2(t) - t \right| > \epsilon \right\} \\ =&O\left(\frac{L^2 x^*}{s_2} + \sqrt{\frac{L(x^*)^2 s_2}{s_3}} + \frac{L^2 s_3}{s_2^2} + \sqrt{\frac{L s_3}{s_2}} + \frac{s_2^2}{(1 - 2Ls_2)^+} \right). \end{aligned}$$

Note that the first term on the right hand side determine its order when $L \to \infty$. Taking $L = t^*$ in the above estimate we see that, since $\vartheta \in (0, 1/2)$, the expression on the right side above goes to 0 as $n \to \infty$. This proves (5.5.16) and thus completes the proof of (iv). Finally, proof of (v) uses standard concentration inequalities for martingales. Indeed, recalling that the maximal jump size of $\bar{H}$, and consequently that of M, is bounded by 1 and $\langle M, M \rangle_t \leq 2t$, we have from Section 4.13, Theorem 5 of [25] that, for any fixed $\alpha > 0$ and $t > 0$,

$$\mathbb{P}\{ \sup_{u \in [0,t]} |M_u| > \alpha \} \leq 2 \exp\left\{ - \sup_{\lambda > 0} [\alpha\lambda - 2t\phi(\lambda)] \right\},$$

where $\phi(\lambda) = (e^\lambda - 1 - \lambda)$. A straightforward calculation shows

$$\sup_{\lambda > 0}[\alpha\lambda - 2t\phi(\lambda)] = \alpha \log\left(1 + \frac{\alpha}{2t}\right) - \left(\alpha - 2t \log\left(1 + \frac{\alpha}{2t}\right)\right) \geq \alpha \log\left(1 + \frac{\alpha}{2t}\right) - \alpha.$$

The result follows. ■

The bound (5.5.15) continues to hold if we replace $M(u)$ with $M(\tau+u)-M(\tau)$ for any finite stopping time τ. From this observation we immediately have the following corollary.

Corollary 5.5.8. Let M be as in Lemma 5.5.7. Then, for any finite stopping time τ:

(i) $\mathbb{P}\left\{\sup_{u\in[0,t]} |M(\tau+u)-M(\tau)| > \alpha\right\} \leq 2e^{-\alpha}$, whenever $\alpha > 2(e^2-1)t$.

(ii) $\mathbb{P}\left\{\sup_{u\in[0,t]} |M(\tau+u)-M(\tau)| > \alpha\right\} \leq 2(2e/\alpha)^{\alpha}t^{\alpha}$, for all $t>0$ and $\alpha>0$.

Part (i) of the corollary is useful when α is large and part (ii) is useful when t is small. Finally we now give the proof of Proposition 5.5.6.

Proof of Proposition 5.5.6: From Lemma 5.5.3 (i) we have that $\hat{H}^{(n)}$ converges in distribution to $\hat{W}_0$ (recall we assume that $\lambda=0$) as $n\to\infty$. Let $\{\epsilon_i\}_{i\geq 1}$ be a positive real sequence bounded by 1 and fix $\Theta\in(2,\infty)$. Choice of Θ and ϵ_i will be specified later in the proof. Let $\sigma_0^{(n)} < \tau_1^{(n)} \leq \sigma_1^{(n)} < \tau_2^{(n)} \leq \sigma_2^{(n)} < \ldots$ be a sequence of stopping times such that $\sigma_0^{(n)}=1$, and for $i\geq 1$,

$$\tau_i^{(n)} = \inf\{t\geq \sigma_{i-1}^{(n)}+\epsilon_i : \hat{H}^{(n)}(t)\geq\Theta\}\wedge(\sigma_{i-1}^{(n)}+1),\ \ \sigma_i^{(n)} = \inf\{t\geq\tau_i^{(n)} : \hat{H}^{(n)}(t)\leq 1\}. \tag{5.5.17}$$

Similarly define stopping times $1=\bar{\sigma}_0<\bar{\tau}_1\leq\bar{\sigma}_1<\bar{\tau}_2\leq\bar{\sigma}_2<\ldots$ by replacing $\hat{H}^{(n)}$ in (5.5.17) with $\hat{W}_0$. Due to the negative quadratic drift in the definition of W_0 it follows that $\bar{\sigma}_i<\infty$ for every i and from the weak convergence of $\hat{H}^{(n)}$ to $\hat{W}_0$ it follows that $\sigma_i^{(n)}\to\bar{\sigma}_i$ and $\tau_i^{(n)}\to\bar{\tau}_i$, in distribution, as $n\to\infty$. Here we have used the fact that if ζ denotes the first time W_0 hits the level $\alpha\in(0,\infty)$ then, a.s., for any $\delta>0$, there are infinitely many crossings of the level α in $(\zeta,\zeta+\delta)$. In particular we have that $\{\sigma_i^{(n)}\}_{n\geq 1}$ is a tight sequence, and this proves part (i) of Proposition 5.5.6.

For the rest of the proof we suppress (n) from the notation. Since the jump size of $\hat{H}$ is bounded by 1, we have that $\sup_{[\sigma_{i-1},\sigma_{i-1}+\epsilon_i]}\hat{H}(t)\leq\Theta$ implies $\sup_{[\sigma_{i-1},\tau_i]}\hat{H}(t)\leq$

$\Theta + 1$ and thus, in this case, when $t \in [\tau_i, \sigma_i]$, we have $\hat{H}(t) = \hat{H}(\tau_i) + (\bar{H}(t) - \bar{H}(\tau_i)) \leq \Theta + 1 + (\bar{H}(t) - \bar{H}(\tau_i))$. Let $G \equiv G^{(n)}$ be as in Lemma 5.5.7 (iv) and let $H_i = G \cap \{\sigma_{i-1} < t^*\}$, then writing $\mathbb{P}(\cdot \cap H_i)$ as $\mathbb{P}_i(\cdot)$,

$$\mathbb{P}_i \left\{ \sup_{[\sigma_{i-1},\sigma_i]} \hat{H}(t) > 2\Theta + 1 \right\} \leq \mathbb{P}_i \left\{ \sup_{[\sigma_{i-1},\sigma_{i-1}+\epsilon_i]} \hat{H}(t) > \Theta \right\} \tag{5.5.18}$$

$$+ \mathbb{P}_i \left\{ \sup_{[\tau_i,\sigma_i]} \left[\Theta + 1 + (\bar{H}(t) - \bar{H}(\tau_i))\right] > 2\Theta + 1 \right\}. \tag{5.5.19}$$

Denote the two terms on the right side by $\mathbb{T}_1$ and $\mathbb{T}_2$ respectively. Recalling that $\hat{H}(\sigma_{i-1}) \leq 2$, we have from the decomposition in Lemma 5.5.7 (i) and Corollary 5.5.8(ii) that

$$\mathbb{T}_1 \leq \mathbb{P} \left\{ \sup_{[\sigma_{i-1},\sigma_{i-1}+\epsilon_i]} |M(t) - M(\sigma_{i-1})| > \frac{\Theta - 2}{2} \right\} \leq C_{\frac{\Theta-2}{2}} \epsilon_i^{(\Theta-2)/2}, \tag{5.5.20}$$

Here, for $\alpha > 0$, $C_\alpha = 2(2e/\alpha)^\alpha$ and we have used the fact that on H_i, $A(t) \leq -t/2 \leq 0$ for all $t \in [\sigma_{i-1}, \sigma_{i-1} + \epsilon_i]$.

Next, let $\{\delta_i\}_{i\geq 1}$ be a sequence of positive reals bounded by 1. Setting $d_i = \sum_{j=1}^{i-1} \epsilon_i$, we have

$$\begin{aligned}
\mathbb{T}_2 \leq & \mathbb{P}_i \left\{ \sup_{[\tau_i,\tau_i+\delta_i]} (\bar{H}(t) - \bar{H}(\tau_i)) > \Theta \right\} + \mathbb{P}_i \left\{ \sup_{[\tau_i+\delta_i,\tau_i+1]} (\bar{H}(t) - \bar{H}(\tau_i)) > \Theta \right\} \\
& + \mathbb{P}\left\{\sigma_i > \tau_i + 1\right\} \\
\leq & \mathbb{P} \left\{ \sup_{[\tau_i,\tau_i+\delta_i]} (M(t) - M(\tau_i)) > \Theta \right\} + \mathbb{P} \left\{ \sup_{[\tau_i+\delta_i,\tau_i+1]} (M(t) - M(\tau_i)) > \Theta + \frac{\delta_i d_i}{2} \right\} \\
& + \mathbb{P} \left\{ M(\tau_i + 1) - M(\tau_i) > -\Theta + \frac{d_i}{2} \right\} \\
\leq & C_\Theta \delta_i^\Theta + 2e^{-\delta_i d_i/2} + 2e^{\Theta - d_i/2},
\end{aligned} \tag{5.5.21}$$

whenever

$$\min\{\delta_i d_i/2, d_i/2 - \Theta\} > 2(e^2 - 1). \tag{5.5.22}$$

Fix $\Theta > 14$. Then $\max\left\{C_\Theta, C_{(\Theta-2)/2}\right\} \le 2$. We will impose additional conditions on Θ later in the proof. Combining (5.5.20) and (5.5.21), we have

$$\mathbb{P}_i\left\{\sup_{[\sigma_{i-1},\sigma_i]} \hat{H}(t) > 2\Theta + 1\right\} \le 2({\epsilon_i}^{(\Theta-2)/2} + \delta_i^\Theta + e^{-\delta_i d_i/2} + e^{\Theta - d_i/2}) \equiv \kappa_i. \quad (5.5.23)$$

Let

$$\epsilon_i = i^{-1/2}, \;\; d_i = \sum_{j=1}^{i-1} \epsilon_i \sim i^{1/2}, \;\; \delta_i = 1/\sqrt{d_i} \sim i^{-1/4}.$$

Then, (5.5.22) holds for i large enough, and

$$\kappa_i \sim 2(i^{-(\Theta-2)/4} + i^{-\Theta/4} + e^{-i^{1/4}/2} + e^{\Theta} e^{-i^{1/2}/2}),$$

which, since $\Theta > 14$, is summable. This proves part (ii) of the Proposition.

Now we consider part (iii). We will construct another sequence of stopping times with values in $[t^*, \infty)$, as follows. Define

$$\sigma_0^* := \inf\{\sigma_i : \sigma_i \ge t^*\} = \inf\left\{t \ge t^* : \hat{H}(t) \le 1\right\}.$$

Then define τ_i^*, σ_i^* for $i \ge 1$ similarly as in (5.5.17). Similar arguments as before give a bound as (5.5.23) with d_i replaced by t^*, δ_i replaced by $1/\sqrt{t^*}$, ϵ_i replaced with $1/t^*$ and Θ replaced by any $\Theta_0 > 14$. Namely,

$$\begin{aligned}&\mathbb{P}\left\{\sup_{[\sigma_{i-1}^*,\sigma_i^*]} \hat{H}(t) > 2\Theta_0 + 1;\; G^{(n)}\right\}\\ \le& 2((1/t^*)^{(\Theta_0-2)/2} + (1/\sqrt{t^*})^{\Theta_0} + e^{-\sqrt{t^*}/2} + e^{\Theta_0 - t^*/2}).\end{aligned} \quad (5.5.24)$$

Here we have used the fact that since $A(t)$ is non-increasing, on $G^{(n)}$, $A(t) \le -t^*/2$ for all $t \ge t^*$.

Recall that, by construction, $\hat{H}(t) = 0$ when $t \ge s_1$. So there exist i_0 such that $\tau_{i_0}^* = \infty$, in fact since $\sigma_i^* \ge \sigma_{i-1}^* + \epsilon$, we have that $i_0 \le s_1/\epsilon$. Thus, we have from the above display that

$$\mathbb{P}\left\{\sup_{[\sigma_0^*,\infty)} \hat{H}(t) > 2\Theta_0 + 1\right\} \le \frac{2s_1}{\epsilon}((1/t^*)^{(\Theta_0-2)/2} + (1/\sqrt{t^*})^{\Theta_0} + e^{-\sqrt{t^*}/2} + e^{\Theta_0 - t^*/2}).$$

Taking $\Theta > 29$, we have on setting $\Theta_0 = \frac{\Theta-1}{2}$ in the above display

$$
\begin{aligned}
&\mathbb{P}\left\{\sup_{[\sigma_0^*,\infty)} \hat{H}(t) > \Theta\right\} \\
\leq & 2s_1 \left(\left(\frac{1}{t^*}\right)^{(\Theta-1)/4-2} + \left(\frac{1}{t^*}\right)^{(\Theta-1)/4-1} + \frac{1}{t^*}e^{-\sqrt{t^*}/2} + \frac{1}{t^*}e^{(\Theta-1)/2-t^*/2} \right).
\end{aligned}
$$

From (5.5.2) we have that $s_1 \cdot (\frac{1}{t^*})^{\varsigma/\vartheta} \to 0$. So if $\Theta \geq 4(\frac{\varsigma}{\vartheta} + 2) + 1$, the above expression approaches 0 as $n \to \infty$. The result now follows on taking $\Theta = \max\{29, 4(\frac{\varsigma}{\vartheta} + 2) + 1\}$. ■

5.6 Bounded-size rules at time $t_c - n^{-\gamma}$

Throughout Sections 5.6 and 5.7 we take $T = 2t_c$ which is a convenient upper bound for the time parameters of interest. In this section we prove Theorem 5.3.2.

We begin with some notation associated with BSR processes, which closely follows [31]. Recall from Section 5.2.2 the set Ω_K and the random graph process $\mathbf{BSR}^{(n)}(t)$ associated with a given K-BSR $F \subset \Omega_K^4$. Frequently we will suppress n in the notation. Also recall the definition of $c_t(v)$ from Section 5.2.2.

For $i \in \Omega_K$, define

$$X_i(t) = |\{v \in \mathbf{BSR}_t^{(n)} : c_t(v) = i\}| \text{ and } \bar{x}_i(t) = X_i(t)/n. \tag{5.6.1}$$

Denote by $\mathbf{BSR}^*(t)$ the subgraph of $\mathbf{BSR}(t)$ consisting of all components of size greater than K, and define, for $k = 1, 2, 3$

$$\mathcal{S}_{k,\varpi}(t) := \sum_{\{\mathcal{C} \subset \mathbf{BSR}^*(t)\}} |\mathcal{C}|^k \text{ and } \bar{s}_{k,\varpi}(t) = \mathcal{S}_{k,\varpi}(t)/n,$$

where $\{\mathcal{C} \subset \mathbf{BSR}^*(t)\}$ denotes the collection of all components in $\mathbf{BSR}^*(t)$. For notational convenience in long formulae, we sometimes write $\mathbf{BSR}(t) = \mathbf{BSR}_t$ and similarly $\mathbf{BSR}^*(t) = \mathbf{BSR}_t^*$. Similar notation will be used throughout this chapter.

Clearly

$$\mathcal{S}_k(t) = \mathcal{S}_{k,\varpi} + \sum_{i=1}^{K} i^{k-1} X_i(t), \;\; \bar{s}_k(t) = \bar{s}_{k,\varpi} + \sum_{i=1}^{K} i^{k-1} \bar{x}_i(t). \tag{5.6.2}$$

Also note that $\mathcal{S}_1(t) = n$ and $\mathcal{S}_{1,\varpi}(t) = X_\varpi(t)$.

Recall the Poisson processes $\mathcal{P}_{\vec{v}}$ introduced in Section 5.2.2. Let $\mathcal{F}_t = \sigma\{\mathcal{P}_{\vec{v}}(s) : s \le t, \vec{v} \in [n]^4\}$. For $T_0 \in [0,T]$ and a $\{\mathcal{F}_t\}_{0 \le t < T_0}$ semi-martingale $\{J(t)\}_{0 \le t < T_0}$ of the form

$$dJ(t) = \alpha(t)dt + dM(t), \;\; \langle M, M \rangle_t = \int_0^t \gamma(s) ds, \tag{5.6.3}$$

where M is a $\{\mathcal{F}_t\}$ local martingale and γ is a progressively measurable process, we write $\alpha = \mathbf{d}(J)$, $M = \boldsymbol{M}(J)$ and $\gamma = \mathbf{v}(J)$.

Organization: The rest of this section is organized as follows. In Section 5.6.1, we state a recent result on BSR models and certain deterministic maps associated with the evolution of $\mathbf{BSR}_t^*$ from Chapter 4 that will be used in this work. In Section 5.6.2, we will study the asymptotics of $\bar{s}_{2,\varpi}$ and $\bar{s}_{3,\varpi}$. In Section 5.6.3, we will complete the proof of Theorem 5.3.2(i). In Section 5.6.4, we will obtain some useful semi-martingale decompositions for certain functionals of $\bar{s}_2$ and $\bar{s}_3$. In Section 5.6.5, we will prove parts (ii) and (iii) of Theorem 5.3.2.

5.6.1 Evolution of $\mathbf{BSR}_t^*$.

We begin with the following lemma from Chapter 4 (see also [31]).

Lemma 5.6.1. (a) For each $i \in \Omega_K$, there exists a continuously differentiable function $x_i : [0,T] \to [0,1]$ such that for any $\delta \in (0, 1/2)$, there exist $C_1, C_2 \in (0,\infty)$ such that for all n,

$$\mathbb{P}\left(\sup_{i \in \Omega_K} \sup_{s \in [0,T]} |\bar{x}_i(t) - x_i(t)| > n^{-\delta} \right) < C_1 \exp\left(-C_2 n^{1-2\delta}\right).$$

(b) There exist polynomials $\{F_i^x(\mathbf{x})\}_{i\in\Omega_K}$, $\mathbf{x} = (x_i)_{i\in\Omega_K} \in \mathbb{R}^{K+1}$, such that $\mathbf{x}(t) = (x_i(t))_{i\in\Omega_K}$ is the unique solution to the differential equations:

$$x_i'(t) = F_i^x(\mathbf{x}(t)), \;\; i \in \Omega_K, \;\; t \in [0,T] \text{ with initial values } \mathbf{x}(0) = (1,0,...,0). \quad (5.6.4)$$

Furthermore, $\bar{x}_i$ is a $\{\mathcal{F}_t\}_{0\le t<T}$ semi-martingale of the form (5.6.3) and

$$\sup_{0\le t<T} |\mathbf{d}(\bar{x}_i)(t) - F_i^x(\bar{\mathbf{x}}(t))| \le \frac{K^2}{n}.$$

Also, for all $i \in \Omega_K$ and $t \in (0,T]$, we have $x_i(t) > 0$ and $\sum_{i\in\Omega_K} x_i(t) = 1$.

Recall that $\mathbf{BSR}^*(t)$ is the subgraph of $\mathbf{BSR}(t)$ consisting of all components of size greater than K. The evolution of this graph is governed by three type of events:

Type 1 (Immigrating vertices): This corresponds to the merger of two components of size bounded by K into a component of size larger than K. Such an event leads to the appearance of a new component in $\mathbf{BSR}^*(t)$ which we view as the immigration of a 'vertex' into $\mathbf{BSR}^*(t)$. Denote by $na_i^*(t)$ the rate at which a component of size $K+i$ immigrates into $\mathbf{BSR}_t^*$ at time t. In Chapter 4 it is shown that there are polynomials $F_i^a(\mathbf{x})$ for $1 \le i \le K$ such that, with $\bar{\mathbf{x}}(t) = (\bar{x}_i(t))_{i\in\Omega_K}$

$$\sup_{t\in[0,\infty)} |a_i^*(t) - F_i^a(\bar{\mathbf{x}}(t))| \le \frac{K}{n}. \quad (5.6.5)$$

We define, with $\mathbf{x}(t)$ as in Lemma 5.6.1,

$$a_i(t) := F_i^a(\mathbf{x}(t)), \; i = 1, \cdots K. \quad (5.6.6)$$

Type 2 (Attachments): This event corresponds to a component of size at most K getting linked with some component of size larger than K. For $1 \le i \le K$, denote by $|\mathcal{C}|c_i^*(t)$ the rate at which a component of size i attaches to a component $\mathcal{C}$ in $\mathbf{BSR}_{t-}^*$. Then there exist polynomials $F_i^c(\mathbf{x})$ for $1 \le i \le K$, such that $c_i^*(t) = F_i^c(\bar{\mathbf{x}}(t))$. Define

$$c_i(t) := F_i^c(\mathbf{x}(t)), i = 1, \cdots K. \quad (5.6.7)$$

Type 3 (Edge formation): This event corresponds to the addition of an edge between components in $\mathbf{BSR}_t^*$. The occurrence of this event adds one edge between two vertices in $\mathbf{BSR}_{t-}^*$, the vertex set stays unchanged, whereas the edge set has one additional element. From Chapter 4, there is a polynomial $F^b(\mathbf{x})$ such that, defining $b^*(t) = F^b(\bar{\mathbf{x}}(t))$, the rate at which each pair of components $\mathcal{C}_1 \neq \mathcal{C}_2 \in \mathbf{BSR}_t^*$ merge at time t, equals $|\mathcal{C}_1||\mathcal{C}_2|b^*(t)/n$. Furthermore, define

$$b(t) := F^b(\mathbf{x}(t)). \tag{5.6.8}$$

From Chapter 4, $F_i^a(\mathbf{x})$, $F_i^c(\mathbf{x})$ and $F^b(\mathbf{x})$ are polynomials with positive coefficients, thus from the last statement in Lemma 5.6.1 we have that $b(t_c) \in (0, \infty)$.

5.6.2 Analysis of $\bar{s}_{2,\varpi}(t)$ and $\bar{s}_{3,\varpi}(t)$

Define functions $F_{2,\varpi}^s : [0,1]^{K+1} \times \mathbb{R} \to \mathbb{R}$ and $F_{3,\varpi}^s : [0,1]^{K+1} \times \mathbb{R}^2 \to \mathbb{R}$ as

$$F_{2,\varpi}^s(\mathbf{x}, s_2) := \sum_{j=1}^{K} (K+j)^2 F_j^a(\mathbf{x}) + s_2 \sum_{j=1}^{K} 2j F_j^c(\mathbf{x}) + x_\varpi \sum_{j=1}^{K} j^2 F_j^c(\mathbf{x}) + (s_2)^2 F^b(\mathbf{x}), \tag{5.6.9}$$

for $(\mathbf{x}, s_2) \in [0,1]^{K+1} \times \mathbb{R}$ and, for $(\mathbf{x}, s_2, s_3) \in [0,1]^{K+1} \times \mathbb{R}^2$

$$\begin{aligned} F_{3,\varpi}^s(\mathbf{x}, s_2, s_3) := &\sum_{j=1}^{K} (K+j)^3 F_j^a(\mathbf{x}) + s_3 \sum_{j=1}^{K} 3j F_j^c(\mathbf{x}) + 3s_2 \sum_{j=1}^{K} j^2 F_j^c(\mathbf{x}) \\ &+ x_\varpi \sum_{j=1}^{K} j^3 F_j^c(\mathbf{x}) + 3 s_2 s_3 F^b(\mathbf{x}). \end{aligned} \tag{5.6.10}$$

The following lemma relates the evolution of $\bar{s}_{j,\varpi}(\cdot)$ to that of $F_{j,\varpi}^s(\bar{\mathbf{x}}(\cdot), \bar{s}_{2,\varpi}(\cdot))$. By definition, $\bar{s}_{j,\varpi}(t)$ is a non-decreasing process with RCLL paths, and therefore a semi-martingale. For $T_0 \in [0, T]$, a stochastic process $\{\xi(t)\}_{0 \le t < T_0}$, and a nonnegative sequence $\alpha(n)$, the quantity $O_{T_0}(\xi(t)\alpha(n))$ will represent a stochastic process $\{\eta(t)\}_{0 \le t < T_0}$ such that for some $d_1 \in (0, \infty)$, $\eta(t) \le d_1 \xi(t) \alpha(n)$, for all $0 \le t < T_0$ and $n \ge 1$.

Lemma 5.6.2. The processes $\bar{s}_{j,\varpi}$, $j = 2, 3$, are $\{\mathcal{F}_t\}_{0\le t<t_c}$ semi-martingales of the form (5.6.3) and

$$|\mathbf{d}(\bar{s}_{2,\varpi})(t) - F^s_{2,\varpi}(\bar{\mathbf{x}}(t), \bar{s}_{2,\varpi}(t))| = O_{t_c}(\mathcal{S}_4(t)/n^2)$$
$$|\mathbf{d}(\bar{s}_{3,\varpi})(t) - F^s_{3,\varpi}(\bar{\mathbf{x}}(t), \bar{s}_{2,\varpi}(t), \bar{s}_{3,\varpi}(t))| = O_{t_c}(\mathcal{S}_5(t)/n^2).$$

Proof: Note that $\mathcal{S}_{2,\varpi}$ and $\mathcal{S}_{3,\varpi}$ have jumps at time instant t with rates and values $\Delta\mathcal{S}_{2,\varpi}(t)$, $\Delta\mathcal{S}_{3,\varpi}(t)$, respectively, given as follows.

- for each $1 \le i \le K$, with rate $na^*_i(t)$,

$$\Delta\mathcal{S}_{2,\varpi}(t) = (K+i)^2, \;\; \Delta\mathcal{S}_{3,\varpi}(t) = (K+i)^3.$$

- for each $1 \le i \le K$ and $\mathcal{C} \subset \mathbf{BSR}^*_{t-}$, at rate $|\mathcal{C}|c^*_i(t)$,

$$\Delta\mathcal{S}_{2,\varpi}(t) = 2|\mathcal{C}|i + i^2, \;\; \Delta\mathcal{S}_{3,\varpi}(t) = 3|\mathcal{C}|^2 i + 3|\mathcal{C}|i^2 + i^3.$$

- for all unordered pair $\mathcal{C}, \tilde{\mathcal{C}} \subset \mathbf{BSR}^*_{t-}$, such that $\mathcal{C} \ne \tilde{\mathcal{C}}$, at rate $|\mathcal{C}||\tilde{\mathcal{C}}|b^*(t)/n$,

$$\Delta\mathcal{S}_{2,\varpi}(t) = 2|\mathcal{C}||\tilde{\mathcal{C}}|, \;\; \Delta\mathcal{S}_{3,\varpi}(t) = 3|\mathcal{C}|^2|\tilde{\mathcal{C}}| + 3|\mathcal{C}||\tilde{\mathcal{C}}|^2.$$

Thus

$$\begin{aligned}
\mathbf{d}(\mathcal{S}_{2,\varpi})(t) =& \sum_{j=1}^{K}(K+j)^2 na^*_j(t) + \sum_{j=1}^{K}\sum_{\mathcal{C}\subset\mathbf{BSR}^*_t}(2j|\mathcal{C}| + j^2)|\mathcal{C}|c^*_j(t) \\
&+ \sum_{\mathcal{C}\ne\tilde{\mathcal{C}}\subset\mathbf{BSR}^*_t} 2|\mathcal{C}||\tilde{\mathcal{C}}|\frac{b^*(t)|\mathcal{C}||\tilde{\mathcal{C}}|}{n} \\
=& \sum_{j=1}^{K}(K+j)^2 na^*_j(t) + \sum_{j=1}^{K} 2jc^*_j(t)\mathcal{S}_{2,\varpi}(t) + \sum_{j=1}^{K} j^2 c^*_j(t)X_\varpi(t) \\
&+ \frac{b^*(t)}{n}(\mathcal{S}^2_{2,\varpi}(t) - \mathcal{S}_{4,\varpi}(t)) \\
=& n\left(F^s_{2,\varpi}(\bar{\mathbf{x}}, \bar{s}_{2,\varpi}) + O(1/n) + O_{t_c}(S_{4,\varpi}(t)/n^2)\right)
\end{aligned} \tag{5.6.11}$$

and

$$
\begin{aligned}
&\mathbf{d}(\mathcal{S}_{3,\varpi})(t) \\
=&\sum_{j=1}^{K}(K+j)^3 n a_j^*(t) + \sum_{j=1}^{K}\sum_{\mathcal{C}\subset\mathbf{BSR}_t^*}(3j|\mathcal{C}|^2+3j^2|\mathcal{C}|+j^3)|\mathcal{C}|c_j^*(t) \\
&+\sum_{\mathcal{C}\neq\tilde{\mathcal{C}}\subset\mathbf{BSR}_t^*}(3|\mathcal{C}|^2|\tilde{\mathcal{C}}|+3|\mathcal{C}||\tilde{\mathcal{C}}|^2)\frac{b^*(t)|\mathcal{C}||\tilde{\mathcal{C}}|}{n} \\
=&\sum_{j=1}^{K}(K+j)^3 n a_j^*(t) + \sum_{j=1}^{K}3jc_j^*(t)\mathcal{S}_{3,\varpi}(t) + \sum_{j=1}^{K}3j^2c_j^*(t)\mathcal{S}_{2,\varpi}(t) \\
+&\sum_{j=1}^{K}j^3c_j^*(t)X_{\varpi}(t) + \frac{3b^*(t)}{n}(\mathcal{S}_{3,\varpi}(t)\mathcal{S}_{2,\varpi}(t)-\mathcal{S}_{5,\varpi}(t)) \\
=&n\left(F_{3,\varpi}^s(\bar{\mathbf{x}}(t),\bar{s}_{2,\varpi}(t),\bar{s}_{3,\varpi}(t)) + O(1/n) + O_{t_c}(S_{4,\varpi}(t)/n^2)\right).
\end{aligned}
$$

The result follows. ■

5.6.3 Proof of Theorem 5.3.2(i)

In this section we prove Theorem 5.3.2(i). We begin with the following differential equations, whose solutions play an important role in defining $s_2(\cdot)$ and $s_3(\cdot)$ that appear in Theorem 5.3.2.

Consider the equations

$$s'_{2,\varpi}(t) = F_{2,\varpi}^s(\mathbf{x}(t), s_{2,\varpi}(t)),\ s_{2,\varpi}(0)=0,\ t\geq 0 \tag{5.6.12}$$

$$s'_{3,\varpi}(t) = F_{3,\varpi}^s(\mathbf{x}(t), s_{2,\varpi}(t), s_{3,\varpi}(t)),\ s_{3,\varpi}(0)=0,\ t\geq 0. \tag{5.6.13}$$

The following lemma describes some properties of solutions to the above differential equations.

Lemma 5.6.3. Equation (5.6.12) and (5.6.13) have unique solutions $s_{2,\varpi}(t)$ and $s_{3,\varpi}(t)$ for $t\in[0,t_c)$. Furthermore $s_{2,\varpi}(t)$ and $s_{3,\varpi}(t)$ are non-negative, increasing and $\lim_{t\uparrow t_c} s_{2,\varpi}(t) = \lim_{t\uparrow t_c} s_{3,\varpi}(t) = \infty$.

Proof: The existence and uniqueness of solution to (5.6.12) follows from Lemma 5.6.1(b) and Theorem 2.2 in [31]. Furthermore, Theorem 2.2 in [31] also implies $\lim_{t\uparrow t_c} s_{2,\varpi}(t) = \infty$.

Note that $F^s_{3,\varpi}$ is a polynomial and the right hand side of (5.6.13) is linear in $s_{3,\varpi}$, thus (5.6.13) has a unique solution on $[0,t_c)$ as well. Since $s_{2,\varpi}(t)$ explodes at t_c, we have $\lim_{t\uparrow t_c} s_{3,\varpi}(t) = \infty$. The monotonicity of $s_{2,\varpi}(t)$ and $s_{3,\varpi}(t)$ follow from the positivity of the functions $\left\{F^a_j : 1 \le j \le K\right\}, F^b, \left\{F^c_j : 1 \le j \le K\right\}$ that appear in the definition of the functions $F^s_{2,\varpi}$ and $F^s_{3,\varpi}$ in (5.6.9) and (5.6.10) respectively. ∎

Define $s_k : [0,t_c) \to [0,\infty)$, $k = 2,3$, as follows.

$$s_k(t) := s_{k,\varpi}(t) + \sum_{i=1}^{K} i^{k-1} x_i(t), \text{ for } k = 2,3. \tag{5.6.14}$$

Then using (5.6.12) - (5.6.13) we get the following differential equations for s_2 and s_3.

Lemma 5.6.4. The functions s_2, s_3 are continuously differentiable on $[0,t_c)$ with

$$\lim_{t\uparrow t_c} s_2(t) = \lim_{t\uparrow t_c} s_3(t) = \infty,$$

and can be characterized as the unique solutions of the following differential equations

$$\begin{aligned} s_2'(t) &= F^s_2(\mathbf{x}(t), s_2(t)), \quad s_2(0) = 1, \\ s_3'(t) &= F^s_3(\mathbf{x}(t), s_2(t), s_3(t)), \quad s_3(0) = 1. \end{aligned}$$

where the function $F^s_2(\cdot)$ and $F^s_3(\cdot)$ are defined as

$$\begin{aligned} F^s_2(\mathbf{x}, s_2) &:= F^s_{2,\varpi}\left(\mathbf{x}, s_2 - \sum_{i=1}^{K} i x_i\right) + \sum_{i=1}^{K} i F^x_i(\mathbf{x}), \\ F^s_3(\mathbf{x}, s_2, s_3) &:= F^s_{3,\varpi}\left(\mathbf{x},\ s_2 - \sum_{i=1}^{K} i x_i,\ s_3 - \sum_{i=1}^{K} i^2 x_i\right) + \sum_{i=1}^{K} i^2 F^x_i(\mathbf{x}). \end{aligned}$$

Proof: The differentiability of $s_2(\cdot)$, $s_3(\cdot)$ on $[0,t_c)$ and the form of F^s_2 and F^s_3 follow from (5.6.14) and the differential equations (5.6.12), (5.6.13) and (5.6.4). The uniqueness of the solution follows from the fact that F_2^s and F_3^s are polynomials. The limit behavior as $t \uparrow t_c$ follows from the definition of s_k. ∎

The following lemma defines the two parameters α and β that appear in Theorems 5.3.2 and 5.3.3. Recall from Section 5.6.1 that $b(t_c) \in (0, \infty)$.

Lemma 5.6.5. The following two limits exist,

$$\alpha := \lim_{t \to t_c -} (t_c - t) s_2(t), \quad \beta := \lim_{t \to t_c -} \frac{s_3(t)}{(s_2(t))^3}.$$

Furthermore, $\alpha, \beta \in (0, \infty)$ and $\alpha = 1/b(t_c)$.

Proof: By (5.6.14), for $k = 2, 3$, $|s_k(t) - s_{k,\varpi}(t)| \le K^k$. Since $\lim_{t \to t_c -} s_{k,\varpi}(t) = \infty$, we thus have that $\lim_{t \to t_c -} s_k(t)/s_{k,\varpi}(t) = 1$. Write $y_\varpi(t) = 1/s_{2,\varpi}(t)$ and $z_\varpi(t) = y_\varpi^3(t) s_{3,\varpi}(t)$, it suffices to show that:

$$\lim_{t \to t_c -} \frac{t_c - t}{y_\varpi(t)} = \lim_{t \to t_c -} -\frac{1}{y'_\varpi(t)} = \frac{1}{b(t_c)}, \text{ and } \lim_{t \to t_c -} z_\varpi(t) \in (0, \infty). \tag{5.6.15}$$

Define $A_l(t) = \sum_{i=1}^K (K+i)^l a_i(t)$ and $C_l(t) = \sum_{i=1}^K i^l c_i(t)$ for $l = 1, 2, 3$. Then by (5.6.12), (5.6.13), (5.6.9) and (5.6.10), the derivative of $y_\varpi(t)$ and $z_\varpi(t)$ can be written as follows (we omit t from the notation):

$$y'_\varpi = -(A_2 + C_2 x_\varpi) y_\varpi^2 - 2C_1 y_\varpi - b, \tag{5.6.16}$$

$$\begin{aligned} z'_\varpi =& y_\varpi^3 \left[A_3 + 3C_1 s_{3,\varpi} + 3C_2 s_{2,\varpi} + C_3 x_\varpi + 3b s_{2,\varpi} s_{3,\varpi}\right] \\ & -3y_\varpi^2 s_{3,\varpi} \left[(A_2 + C_2 x_\varpi) y_\varpi^2 + 2C_1 y_\varpi + b\right] \\ =& -(3y_\varpi A_2 + 3y_\varpi C_2 x_\varpi + 3C_1) z_\varpi + (y_\varpi^3 A_3 + 3y_\varpi^2 C_2 + y_\varpi^3 C_3 x_\varpi) \\ =& -B_1 z_\varpi + B_2, \end{aligned} \tag{5.6.17}$$

where $B_1(t) = (3y_\varpi(t) A_2(t) + 3y_\varpi(t) C_2(t) x_\varpi(t) + 3C_1(t))$ and $B_2(t) = (y_\varpi^3(t) A_3(t) + 3y_\varpi^2(t) C_2(t) + y_\varpi^3(t) C_3(t) x_\varpi(t))$. Since $\lim_{t \to t_c -} y_\varpi(t) = 0$, we have $\lim_{t \to t_c -} y'_\varpi(t) = -b(t_c)$ which proves the first statement in (5.6.15).

Choose $t_1 \in (0, t_c)$ such that $y_\varpi(t), z_\varpi(t) \in (0, \infty)$ for all $t \in (t_1, t_c)$. Then from (5.6.17), for all such t

$$z_\varpi(t) = \int_{t_1}^{t} e^{-\int_s^t B_1(u)du} B_2(s)ds + z_\varpi(t_1)e^{-\int_{t_1}^t B_1(u)du}.$$

Since B_1, B_2 are nonnegative and $\sup_{t\in[t_1,t_c]}\{B_1(t) + B_2(t)\} < \infty$, we have

$$\lim_{t\to t_c-} z_\varpi(t) \in (0, \infty).$$

This completes the proof of (5.6.15). The result follows. ■

We now complete the proof of Theorem 5.3.2(i).

Proof of Theorem 5.3.2(i): Let α, β be as introduced in Lemma 5.6.5. From Lemma 5.6.4 it follows that $y(t) = 1/s_2(t)$ and $z(t) = y^3(t)s_3(t)$, for $0 \le t < t_c$, solve the differential equations

$$y'(t) = F^y(\mathbf{x}(t), y(t)),\ z'(t) = F^z(\mathbf{x}(t), y(t), z(t)),\ y(0) = z(0) = 1, \tag{5.6.18}$$

where $F^y : [0,1]^{K+2} \to \mathbb{R}$ and $F^z : [0,1]^{K+2} \times \mathbb{R} \to \mathbb{R}$ are defined as

$$F^y(\mathbf{x}, y) := -y^2 F_2^s(\mathbf{x}, 1/y), \quad F^z(\mathbf{x}, y, z) := 3zF^y(\mathbf{x}, y)/y + y^3 F_3^s(\mathbf{x}, 1/y, z/y^3), \tag{5.6.19}$$

$(\mathbf{x}, y, z) \in [0,1]^{K+2} \times \mathbb{R} \to \mathbb{R}$. We claim that $F^y(\mathbf{x}, y)$ and $F^z(\mathbf{x}, y, z)$ are polynomials. To see this note that, by (5.6.9), for any $u \in \mathbb{R}$

$$y^2 F_{2,\varpi}^s(\mathbf{x}, u/y) = y^2 \sum_{j=1}^{K} (K+j)^2 F_j^a(\mathbf{x}) + uy \sum_{j=1}^{K} 2jF_j^c(\mathbf{x}) + y^2 x_\varpi \sum_{j=1}^{K} j^2 F_j^c(\mathbf{x}) + u^2 F^b(\mathbf{x}).$$

From the definition of F_2^s in Lemma 5.6.4 it is now clear that $F^y(\mathbf{x}, y)$ is a polynomial. To check that $F^z(\mathbf{x}, y, z)$ is a polynomial, one only needs to examine the expression

$$-3zyF_{2,\varpi}^s(\mathbf{x}, \frac{1}{y} - \beta_1) + y^3 F_{3,\varpi}^s(\mathbf{x}, \frac{1}{y} - \beta_1, \frac{z}{y^3} - \beta_2), \text{ where } \beta_k = \sum_{i=1}^{K} i^k x_i,\ i = 1, 2.$$

Note that the non-polynomial part in the first term is $-3yz \cdot \frac{1}{y^2} F^b(\mathbf{x})$, which gets cancelled with the non-polynomial part of the second term, namely $y^3 \cdot \frac{3}{y} \cdot \frac{z}{y^3} F^b(\mathbf{x})$.

This proves the claim. Also, from (5.6.10) it is clear that the order of z in $F^z(\mathbf{x}, y, z)$ is at most 1.

Thus (5.6.18) has a unique solution. Also, defining $y(t_c) = \lim_{t\to t_c-} y(t) = \lim_{t\to t_c-} y_\varpi(t)$ and $z(t_c) = \lim_{t\to t_c-} z(t) = \lim_{t\to t_c-} z_\varpi(t)$, we see that y, z are twice continuously differentiable (from the left) at t_c. Furthermore, $y'(t_c-) = -\alpha^{-1}$ and $z(t_c-) = \beta$. Thus we have

$$y(t) = \frac{1}{\alpha}(t_c - t)(1 + O(t_c - t)), \;\; z(t) = \beta(1 + O(t_c - t)), \text{ as } t \uparrow t_c.$$

The result follows. ■

5.6.4 Asymptotic analysis of $\bar{s}_2(t)$ and $\bar{s}_3(t)$

In preparation for the proof of Theorem 5.3.2(ii), in this section we will obtain some useful semi-martingale decompositions for $Y(t) := \frac{1}{\bar{s}_2(t)}$ and $Z(t) := \frac{\bar{s}_3(t)}{(\bar{s}_2(t))^3}$. Throughout this section and next we will denote $|\mathcal{C}_1^{(n)}(t)|$ as $I(t)$. Recall the functions F_2^s, F_3^s introduced in Lemma 5.6.4.

Lemma 5.6.6. The processes $\bar{s}_2$ and $\bar{s}_3$ are $\{\mathcal{F}_t\}_{0\le t<t_c}$ semi-martingales of the form (5.6.3) and the following equations hold.

(a) $\mathbf{d}(\bar{s}_2)(t) = F_2^s(\bar{\mathbf{x}}(t), \bar{s}_2(t)) + O_{t_c}\left(I^2(t)\bar{s}_2(t)/n\right).$

(b) $\mathbf{d}(\bar{s}_3)(t) = F_3^s(\bar{\mathbf{x}}(t), \bar{s}_2(t), \bar{s}_3(t)) + O_{t_c}\left(I^3(t)\bar{s}_2(t)/n\right).$

(c) $\mathbf{v}(\bar{s}_2)(t) = O_{t_c}(I^2(t)\bar{s}_2^2(t)/n).$

Proof: Parts (a) and (b) are immediate from (5.6.2), Lemma 5.6.1(b) and Lemma 5.6.2. For part (c), recall the three types of events described in Section 5.6.1. For type 1, $\Delta\bar{s}_2(t)$ is bounded by $2K^2/n$ and the total rate of such events is bounded by $n/2$. For type 2, the attachment of a size j component, $1 \le j \le K$, to a component $\mathcal{C}$ in $\mathbf{BSR}^*_{t-}$ occurs at rate $|\mathcal{C}|c^*_j(t)$ and produces a jump $\Delta\bar{s}_2(t) = 2j|\mathcal{C}|/n$. For type 3, components $\mathcal{C}$ and $\tilde{\mathcal{C}}$ merge at rate $|\mathcal{C}||\tilde{\mathcal{C}}|b^*(t)/n$ and produce a jump $\Delta\bar{s}_2(t) = 2|\mathcal{C}||\tilde{\mathcal{C}}|/n$. Thus for $t \in [0, t_c)$, $\mathbf{v}(\bar{s}_2)(t)$ can be estimated as

$$
\begin{aligned}
&\mathbf{v}(\bar{s}_2)(t)\\
\le& \frac{n}{2}\left(\frac{2K^2}{n}\right)^2 + \sum_{j=1}^{K}\sum_{\mathcal{C}\subset\mathbf{BSR}^*_t}\left(\frac{2j|\mathcal{C}|}{n}\right)^2 |\mathcal{C}|c^*_j(t) + \sum_{\mathcal{C}\neq\tilde{\mathcal{C}}\subset\mathbf{BSR}^*_t}\left(\frac{2|\mathcal{C}||\tilde{\mathcal{C}}|}{n}\right)^2 \frac{b^*(t)|\mathcal{C}||\tilde{\mathcal{C}}|}{n}\\
\le& \frac{2K^4}{n} + \frac{4K^2\mathcal{S}_3}{n^2} + \frac{4(\mathcal{S}_3)^2}{n^3} = O_{t_c}\left(\frac{I^2(t)\bar{s}_2^2(t)}{n}\right).
\end{aligned}
$$

This proves (c). ■

In the next lemma, we obtain a semi-martingale decomposition for Y.

Lemma 5.6.7. The process $Y(t) = 1/\bar{s}_2(t)$ is a $\{\mathcal{F}_t\}_{0\le t<t_c}$ semi-martingale of the form (5.6.3) and

(i) With $F^y(\cdot)$ as defined in (5.6.19),

$$\mathbf{d}(Y)(t) = F^y(\bar{\mathbf{x}}(t), Y(t)) + O_{t_c}\left(\frac{I^2(t)Y(t)}{n}\right). \tag{5.6.20}$$

(ii)

$$\mathbf{v}(Y)(t) = O_{t_c}\left(\frac{I^2(t)Y^2(t)}{n}\right).$$

Proof: Note that

$$\Delta Y(t) = \frac{1}{\bar{s}_2 + \Delta\bar{s}_2} - \frac{1}{\bar{s}_2} = -\frac{\Delta\bar{s}_2}{\bar{s}_2^2} + \frac{(\Delta\bar{s}_2)^2}{\bar{s}_2^2(\bar{s}_2+\Delta\bar{s}_2)} = -\frac{\Delta\bar{s}_2}{\bar{s}_2^2} + O_{t_c}\left(\frac{(\Delta\bar{s}_2)^2}{\bar{s}_2^3}\right). \tag{5.6.21}$$

Thus by Lemma 5.6.6(a), we have,

$$
\begin{aligned}
&\mathbf{d}(Y)(t)\\
=& -\frac{1}{(\bar{s}_2(t))^2}\mathbf{d}(\bar{s}_2)(t) + O_{t_c}\left(\frac{1}{(\bar{s}_2(t))^3}\mathbf{v}(\bar{s}_2)(t)\right)\\
=& \left(-\frac{1}{(\bar{s}_2(t))^2}\right)\left(F^s_2(\bar{\mathbf{x}}(t), \bar{s}_2(t)) + O_{t_c}\left(\frac{I^2(t)\bar{s}_2(t)}{n}\right)\right) + O_{t_c}\left(\frac{1}{(\bar{s}_2(s))^3}\cdot\frac{I^2(t)\bar{s}_2^2(t)}{n}\right)\\
=& F^y(\bar{\mathbf{x}}(t), Y(t)) + O_{t_c}\left(\frac{I^2(t)Y(t)}{n}\right).
\end{aligned}
$$

This proves (i). For (ii), note that (5.6.21) also implies $(\Delta Y(t))^2 \leq \frac{(\Delta \bar{s}_2)^2}{\bar{s}_2^4}$. We then have

$$\mathbf{v}(Y)(t) \leq \frac{2\mathbf{v}(\bar{s}_2)(t)}{\bar{s}_2^4} = O_{t_c}\left(\frac{I^2(t)Y^2(t)}{n}\right).$$

The result follows. ■

We now give a semi-martingale decomposition for $Z(t) = \bar{s}_3(t)/(\bar{s}_2(t))^3$.

Lemma 5.6.8. The process $Z(t) = \bar{s}_3(t)/(\bar{s}_2(t))^3$ is a $\{\mathcal{F}_t\}_{0\leq t<t_c}$ semi-martingale of the form (5.6.3) and

(i) With $F^z(\cdot)$ as defined in (5.6.19),

$$\mathbf{d}(Z)(t) = F^z(\bar{\mathbf{x}}(t), Y(t), Z(t)) + O_{t_c}\left(\frac{I^3(t)Y^2(t)}{n}\right).$$

(ii)

$$\mathbf{v}(Z)(t) = O_{t_c}\left(\frac{I^4(t)Y^4(t)}{n} + \frac{I^6(t)Y^6(t)}{n}\right).$$

Proof: Note that

$$\Delta Z = Y^3\Delta\bar{s}_3 + 3Y^2\bar{s}_3\Delta Y + R(\Delta Y, \Delta\bar{s}_3),$$

where $R(\Delta Y, \Delta\bar{s}_3)$ is the error term which, using the observations that $\bar{s}_3 \leq I\bar{s}_2$, $\Delta\bar{s}_3 \leq 3I\Delta\bar{s}_2$ and $|\Delta Y| \leq Y^2\Delta\bar{s}_2$, can be bounded as follows.

$$\begin{aligned}|R(\Delta Y, \Delta\bar{s}_3)| \leq& 3Y^2|\Delta Y||\Delta\bar{s}_3| + 3Y\bar{s}_3|\Delta Y|^2\\ &\leq 3Y^2\cdot Y^2\Delta\bar{s}_2\cdot 3I\Delta\bar{s}_2 + 3I\cdot\left(Y^2\Delta\bar{s}_2\right)^2 = 12IY^4\cdot(\Delta\bar{s}_2)^2.\end{aligned}$$

From Lemma 5.6.6(b), Lemma 5.6.7(i) and Lemma 5.6.6(c), we have

$$\begin{aligned}\mathbf{d}(Z)(t) =& Y^3(t)\mathbf{d}(\bar{s}_3)(t) + 3Y^2(t)\bar{s}_3(t)\mathbf{d}(Y)(t) + O_{t_c}\left(I(t)Y^4(t)\mathbf{v}(\bar{s}_2)(t)\right)\\ =& Y^3(t)\left(F_3^s(\bar{\mathbf{x}}(t), \bar{s}_2(t), \bar{s}_3(t)) + O_{t_c}\left(\frac{I^3(t)\bar{s}_2(t)}{n}\right)\right)\\ &+ 3Y^2(t)\bar{s}_3(t)\left(F^y(\bar{\mathbf{x}}(t), Y(t)) + O_{t_c}\left(\frac{I^2(t)Y(t)}{n}\right)\right) + O_{t_c}\left(\frac{I^3(t)Y^2(t)}{n}\right)\\ =& F^z(\bar{\mathbf{x}}(t), Y(t), Z(t)) + O_{t_c}\left(\frac{I^3(t)Y^2(t)}{n}\right).\end{aligned}$$

This proves (i). For (ii), note that

$$Y^3|\Delta \bar{s}_3| + 3Y^2\bar{s}_3|\Delta Y| \leq Y^3 \cdot 3I|\Delta \bar{s}_2| + 3Y^2 \cdot I\bar{s}_2 \cdot Y^2|\Delta \bar{s}_2| = 6Y^3 I|\Delta \bar{s}_2|.$$

Thus,

$$|\Delta Z| \leq 6Y^3 I|\Delta \bar{s}_2| + 12IY^4 \cdot (\Delta \bar{s}_2)^2.$$

Applying Lemma 5.6.6(c) we now have,

$$\mathbf{v}(Z)(t) = O_{t_c}\left(Y^6 I^2 \mathbf{v}(\bar{s}_2)(t)\right) + O_{t_c}\left(\frac{I^6Y^6}{n}\right) = O_{t_c}\left(\frac{I^4Y^4}{n} + \frac{I^6Y^6}{n}\right).$$

The result follows. ■

5.6.5 Proof of Theorem 5.3.2(ii)

We begin with an upper bound on the size of the largest component at time $t \leq t_n = t_c - n^{-\gamma}$ for $\gamma \in (0, 1/4)$, which has been proved in Chapter 4, and will play an important role in the proof of Theorem 5.3.2(ii).

Theorem 5.6.9 (**Barely subcritical regime**). Fix $\gamma \in (0, 1/4)$. Then there exists $C_3 \in (0, \infty)$ such that, as $n \to \infty$,

$$\mathbb{P}\left\{ I^{(n)}(t) \leq C_3 \frac{(\log n)^4}{(t_c - t)^2},\ \forall t < t_c - n^{-\gamma} \right\} \to 1.$$

The next lemma is an elementary consequence of Gronwall's inequality.

Lemma 5.6.10. Let $\{t_n\}$ be a sequence of positive reals such that $t_n \in [0, t_c)$ for all n. Suppose that $U^{(n)}$ is a semi-martingale of the form (5.6.3) with values in $\mathbb{D} \subset \mathbb{R}$. Let $g : [0, t_c) \times \mathbb{D} \to \mathbb{R}$ be such that, for some $C_4(g) \in (0, \infty)$,

$$\sup_{t \in [0, t_c)} |g(t, u_1) - g(t, u_2)| \leq C_4(g)|u_1 - u_2|,\ u_1, u_2 \in \mathbb{D}. \tag{5.6.22}$$

Let $\{u(t)\}_{t \in [0, t_c)}$ be the unique solution of the differential equation

$$u'(t) = g(t, u(t)), \quad u(0) = u_0.$$

Further suppose that there exist positive sequences:

(i) $\{\theta_1(n)\}$ such that, whp, $|U^{(n)}(0) - u_0| \le \theta_1(n)$.

(ii) $\{\theta_2(n)\}$ such that, whp,

$$\int_0^{t_n} |\mathbf{d}(U^{(n)})(t) - g(t, U^{(n)}(t))|\, dt \le \theta_2(n).$$

(iii) $\{\theta_3(n)\}$ such that, whp, $\langle \boldsymbol{M}(U^{(n)}), \boldsymbol{M}(U^{(n)})\rangle_{t_n} \le \theta_3(n)$.

Then, whp,

$$\sup_{0\le t\le t_n} |U^{(n)}(t) - u(t)| \le e^{C_4(g)t_c}(\theta_1(n) + \theta_2(n) + \theta_4(n)),$$

where $\theta_4 = \theta_4(n)$ is any sequence satisfying $\sqrt{\theta_3(n)} = o(\theta_4(n))$.

Proof: We suppress n from the notation unless needed. Using the Lipschitz property of g, we have, for all $t \in [0, t_n]$,

$$\begin{aligned}
&|U(t) - u(t)| \\
&\le |U(0) - u_0| + \int_0^t |\mathbf{d}(U)(s) - g(s, U(s))|ds \\
&\qquad\qquad + \int_0^t |g(s, U(s)) - g(s, u(s))|ds + |\boldsymbol{M}(U)(t)| \\
&\le |U(0) - u_0| + \int_0^t |\mathbf{d}(U)(s) - g(s, U(s))|ds + |\boldsymbol{M}(U)(t)| + C_4 \int_0^t |U(s) - u(s)|ds.
\end{aligned}$$

Then by Gronwall's lemma

$$\begin{aligned}
&\sup_{0\le t\le t_n} |U(t) - u(t)| \qquad (5.6.23)\\
&\le \left(|U(0) - u_0| + \int_0^{t_n} |\mathbf{d}(U)(s) - g(s, U(s))|ds + \sup_{0\le t\le t_n} |\boldsymbol{M}(U)(t)|\right) e^{C_4 t_c}.
\end{aligned}$$

Let $\tau^{(n)} = \inf\{t \ge 0 : \langle \boldsymbol{M}(U), \boldsymbol{M}(U)\rangle_t > \theta_3(n)\}$. By Doob's inequality

$$\mathbb{E}[\sup_{0\le t\le t_n} |\boldsymbol{M}(U)(t\wedge\tau)|^2] \le 4\mathbb{E}[|\boldsymbol{M}(U)(t_n\wedge\tau)|^2] = 4\mathbb{E}\left[\langle \boldsymbol{M}(U), \boldsymbol{M}(U)\rangle_{t_n\wedge\tau}\right] \le 4\theta_3(n).$$

Then for any $\theta_4(n)$ such that $\theta_3 = o((\theta_4)^2)$, we have

$$\begin{aligned}\mathbb{P}\{\sup_{0\le t\le t_n} |\boldsymbol{M}(U)(t)| > \theta_4(n)\} \le& \mathbb{P}\{\tau^{(n)} < t_n\} + \mathbb{P}\left\{\sup_{0\le t\le t_n} |\boldsymbol{M}(U)(t\wedge\tau)| > \theta_4(n)\right\}\\ \le& \mathbb{P}\{\langle \boldsymbol{M}(U), \boldsymbol{M}(U)\rangle_{t_n} > \theta_3(n)\} + 4\theta_3(n)/\theta_4^2(n) \to 0.\end{aligned}$$

The result now follows on using the above observation in (5.6.23). ■

Proof of Theorem 5.3.2(ii): Let y and z be as in the proof of Theorem 5.3.2(i). It suffices to show

$$\sup_{0\le t\le t_n} |Y(t) - y(t)|\, n^{1/3} \xrightarrow{\mathbb{P}} 0 \tag{5.6.24}$$

$$\sup_{0\le t\le t_n} |Z(t) - z(t)| \xrightarrow{\mathbb{P}} 0. \tag{5.6.25}$$

We begin by proving the following weaker result than (5.6.24):

$$\sup_{0\le t\le t_n} |Y(t) - y(t)| = O(n^{-1/5}), \text{ whp.} \tag{5.6.26}$$

Recalling from the proof of Theorem 5.3.2(i) that $\mathbf{x} \mapsto F^y(\mathbf{x}, y)$ is Lipschitz for $\mathbf{x} \in [0,1]^{K+1}$, uniformly for all $y \in [0,1]$, we get for some $d_1 \in (0,\infty)$

$$\sup_{0\le t\le t_c} |F^y(\bar{\mathbf{x}}(t), Y(t)) - F^y(\mathbf{x}(t), Y(t))| \le d_1 \sup_{i\in\Omega_K} \sup_{0\le t\le t_c} |\bar{x}_i(t) - x_i(t)|.$$

From Lemma 5.6.7(i) and Lemma 5.6.1(a) we now get for some $d_2 \in (0,\infty)$, whp,

$$|\mathbf{d}(Y)(t) - F^y(\mathbf{x}(t), Y(t))| \le d_2 \left(\frac{I^2(t)Y(t)}{n} + n^{-2/5}\right), \text{ for all } t \in [0, t_n].$$

Thus, from Theorem 5.6.9 and recalling that $\gamma < 1/5$, we have whp,

$$\begin{aligned}\int_0^{t_n} |\mathbf{d}(Y)(t) - F^y(\mathbf{x}(t), Y(t))|dt =& O\left(\int_0^{t_n} \frac{(\log n)^8}{n(t_c - t)^4} dt + n^{-2/5}\right)\\ =& O((\log n)^8 n^{3\gamma - 1}) + O(n^{-2/5}) = O(n^{-2/5}).\end{aligned}$$

Next, by Lemma 5.6.7(ii) and using the fact $Y(t) \le 1$ for all $t \in [0, t_c)$,

$$\begin{aligned}\langle \boldsymbol{M}(Y), \boldsymbol{M}(Y)\rangle_{t_n} =& O\left(\int_0^{t_n} \frac{I^2(t)Y^2(t)}{n} dt\right) = O\left(\int_0^{t_n} \frac{I^2(t)}{n} dt\right)\\ =& O\left(\int_0^{t_n} \frac{(\log n)^8}{n(t_c - t)^4} dt\right) = O((\log n)^8 n^{3\gamma - 1}).\end{aligned} \tag{5.6.27}$$

The statement in (5.6.26) now follows on observing that $((\log n)^8 n^{3\gamma-1})^{1/2} = o(n^{-1/5})$ and applying Lemma 5.6.10 with $\mathbb{D} := [0,1]$, $g(t,y) := F^y(\mathbf{x}(t), y)$, $\theta_1 = 0$, $\theta_2 = n^{-2/5}$ and $\theta_3 = (\log n)^8 n^{3\gamma-1}$.

We now strengthen the estimate in (5.6.26) to prove (5.6.24). From Theorem 5.3.2(i) it follows that $y(t_n) = 1/s_2(t_n) = \Theta(n^{-\gamma})$. Since $\gamma < 1/5$, from (5.6.26) we have, whp, $Y(t) \le 2y(t)$ for all $t \le t_n$. Thus from the first equality in (5.6.27) and Theorem 5.3.2(i) we get, whp,

$$\langle \boldsymbol{M}(Y), \boldsymbol{M}(Y)\rangle_{t_n} = O\left(\int_0^{t_n} \frac{I^2(t)y^2(t)}{n}\right) = O\left(\int_0^{t_n} \frac{(\log n)^8}{n(t_c-t)^2} dt\right) = O((\log n)^8 n^{\gamma-1}).$$

Since $((\log n)^8 n^{\gamma-1})^{1/2} = o(n^{-2/5})$, applying Lemma 5.6.10 again gives

$$\sup_{0\le t\le t_n} |Y(t) - y(t)| = O(n^{-2/5}), \text{ whp.} \tag{5.6.28}$$

This proves (5.6.24).

We now prove (5.6.25). We will apply Lemma 5.6.10 to $\mathbb{D} := \mathbb{R}$ and $g(t,z) := F^z(\mathbf{x}(t), y(t), z)$. As noted in the proof of Theorem 5.3.2(i), g defined as above satisfies (5.6.22).

We now verify the three assumptions in Lemma 5.6.10. Note that (i) is satisfied with $\theta_1 = 0$, since $Z(0) = z(0) = 1$. Next, by Lemma 5.6.8(ii) and the fact $Y(t) \le 2y(t)$ for $t \le t_n$, whp, we have

$$\begin{aligned}\langle \boldsymbol{M}(Z), \boldsymbol{M}(Z)\rangle_{t_n} =& O\left(\int_0^{t_n} \left(\frac{I^4(t)Y^4(t)}{n} + \frac{I^6(t)Y^6(t)}{n}\right) dt\right) \\ =& O\left(\int_0^{t_n} \left(\frac{(\log n)^{16}}{n(t_c-t)^4} + \frac{(\log n)^{24}}{n(t_c-t)^6}\right) dt\right) = O((\log n)^{24} n^{5\gamma-1}).\end{aligned}$$

Since $\gamma < 1/5$, we can find $\theta_4(n) \to 0$ such that $\sqrt{(\log n)^{24} n^{5\gamma-1}} = o(\theta_4(n))$ Thus (iii) in Lemma 5.6.10 is satisfied. Next recall from the proof of Theorem 5.3.2(i) that $g(t,z)$ is linear in z. Also, $Z(t) \le I(t)$. Thus from Lemma 5.6.1 and (5.6.28), for some $d_3 \in (0,\infty)$ whp, for all $t \le t_n$

$$
\begin{aligned}
&|F^z(\bar{\mathbf{x}}(t), Y(t), Z(t)) - g(t, Z(t))| \\
\le & d_3(1+Z(t)) \left(\sup_{1\le i\le K} \sup_{0\le t\le t_n} |\bar{x}_i(t) - x_i(t)| + \sup_{0\le t\le t_n} |Y(t) - y(t)| \right) = I(t) O(n^{-2/5}).
\end{aligned}
$$

By Lemma 5.6.8(i) and the above bound,

$$
\begin{aligned}
\int_0^{t_n} |\mathbf{d}(Z)(t) - g(t, Z(t))| dt = & O\left(\int_0^{t_n} n^{-2/5} I(t) dt \right) + O\left(\int_0^{t_n} \frac{y^2(t) I^3(t)}{n} dt \right) \\
= & O\left(\int_0^{t_n} \frac{(\log n)^4 n^{-2/5}}{(t_c - t)^2} dt \right) + O\left(\int_0^{t_n} \frac{(\log n)^{12}}{(t_c - t)^4} dt \right) \\
= & O((\log n)^4 n^{\gamma - 2/5}) + O((\log n)^{12} n^{3\gamma - 1}). \qquad (5.6.29)
\end{aligned}
$$

This verifies (ii) in Lemma 5.6.10 with $\theta_2(n) = O((\log n)^{12} n^{3\gamma-1})$. From Lemma 5.6.10 we now have

$$
\sup_{0\le t\le t_n} |Z(t) - z(t)| = O(\theta_1(n) + \theta_2(n) + \theta_4(n)) = o(1).
$$

The result follows. ∎

Part (iii) of Theorem 5.3.2 is a simple consequence of part (ii).

Proof of Theorem 5.3.2(iii): The convergence when $t = 0$ is trivial. Consider now $t > 0$. Since $t_n \to t_c$ as $n \to \infty$, we have from part (ii) of the theorem that, for fixed $t \in (0, t_c)$,

$$
\frac{1}{\bar{s}_2(t)} \xrightarrow{\mathbb{P}} \frac{1}{s_2(t)}, \quad \frac{\bar{s}_3(t)}{(\bar{s}_2(t))^3} \xrightarrow{\mathbb{P}} \frac{s_3(t)}{(s_2(t))^3}.
$$

Also, since $t > 0$, we have that $x_i(t) > 0$ for all $i \in \Omega_K$, thus $s_2(t) > 0$. Theorem 5.3.2(iii) is now immediate. ∎

5.7 Coupling with the multiplicative coalescent

In this section we prove Theorem 5.3.3. Throughout this section we fix $\gamma \in (1/6, 1/5)$. The basic idea of the proof is as follows. Recall $\alpha, \beta \in (0,\infty)$ from Theorem 5.3.2 (see also Lemma 5.6.5). We begin by approximating the BSR random graph process by a process which until time $t_n := t_c - n^{-\gamma}$ is identical to the BSR process and in the time interval $[t_n, t_c + \alpha\beta^{2/3}\frac{\lambda}{n^{1/3}}]$ evolves as an Erdős-Rényi process, namely over this interval edges between any pair of vertices appear at rate $1/\alpha n$, and self loops at any given vertex appear at rate $1/2\alpha n$. Asymptotic behavior of this random graph is analyzed using Theorem 5.5.1. Theorems 5.3.2 and 5.6.9 help in verifying the conditions (5.5.1) and (5.5.2) in the statement of Theorem 5.5.1. We then complete the proof of Theorem 5.3.3 by arguing that the 'difference' between the BSR process and the modified random graph process is asymptotically negligible.

Let

$$t_n = t_c - \alpha\beta^{2/3}\frac{\lambda_n}{n^{1/3}} \text{ where } \lambda_n = \frac{n^{1/3-\gamma}}{\alpha\beta^{2/3}}. \tag{5.7.1}$$

Throughout this section, for $\lambda \in \mathbb{R}$, we denote $t^\lambda = t_c + \alpha\beta^{2/3}\lambda/n^{1/3}$. Recall the random graph process $\mathbf{BSR}^*(t)$ introduced in Section 5.6. Denote by $(|\mathcal{C}_i^*(t)|, \xi_i^*(t))_{i\geq 1}$ the vector of ordered component size and corresponding surplus in $\mathbf{BSR}^*(t)$ (the components are denoted by $\mathcal{C}_i^*(t)$). Let, for $\lambda \in \mathbb{R}$,

$$\bar{\boldsymbol{C}}^{(n),*}(\lambda) = \left(\frac{\beta^{1/3}}{n^{2/3}}\left|\mathcal{C}_i^*(t^\lambda)\right| : i \geq 1\right), \quad \bar{\boldsymbol{Y}}^{(n),*}(\lambda) = \left(\xi_i^*(t^\lambda) : i \geq 1\right).$$

For $i \geq 1$, denote $\bar{\boldsymbol{C}}_i^{(n),*}(\lambda)$ and $\bar{\boldsymbol{Y}}_i^{(n),*}(\lambda)$ for the i-th coordinate of $\bar{\boldsymbol{C}}^{(n),*}(\lambda)$ and $\bar{\boldsymbol{Y}}^{(n),*}(\lambda)$ respectively. Write $\bar{\boldsymbol{Y}}_i^{(n),*} = \tilde{\xi}_i^{(n)} + \hat{\xi}_i^{(n)}$ where $\tilde{\xi}_i^{(n)}(\lambda)$ represents the surplus in $\mathbf{BSR}^*(t^\lambda)$ that is created before time t_n, namely

$$\tilde{\xi}_i^{(n)}(\lambda) = \sum_{j:\mathcal{C}_j^*(t_n)\subset\mathcal{C}_i^*(t^\lambda)} \bar{\boldsymbol{Y}}_j^{(n),*}(-\lambda_n).$$

In Section 5.7.2 we will show that the contribution from $\tilde{\xi}^{(n)}(\lambda) := (\tilde{\xi}_i^{(n)}(\lambda) : i \geq 1)$ is asymptotically negligible. First, in Section 5.7.1 below we analyze the contribution from the 'new surplus', i.e. $\hat{\xi}^{(n)} := (\hat{\xi}_i^{(n)} : i \geq 1)$.

5.7.1 Surplus created after time t_n.

The main result of this section is as follows. Recall $\boldsymbol{Z}(\lambda) = (\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda))$ introduced in Theorem 5.3.1.

Theorem 5.7.1. For every $\lambda \in \mathbb{R}$, as $n \to \infty$, $(\bar{\boldsymbol{C}}^{(n),*}(\lambda), \hat{\xi}^{(n)}(\lambda))$ converges in distribution, in $\mathbb{U}_{\downarrow}$, to $(\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda))$.

The basic idea in the proof of the above theorem is to argue that $\mathbf{BSR}^*(t^\lambda)$ 'lies between' two Erdős-Rényi random graph processes $\mathbf{G}^{(n),-}(t^\lambda)$ and $\mathbf{G}^{(n),+}(t^\lambda)$, whp, and then apply Theorem 5.5.1 to each of these processes. For a graph $\mathbf{G}$, denote by $|\mathcal{C}_i(\mathbf{G})|$ and $\xi_i(\mathbf{G})$ the size and surplus, respectively, of the i-th largest component, $\mathcal{C}_i(\mathbf{G})$ of graph $\mathbf{G}$. We begin with the following lemma. Recall λ_n from (5.7.1).

Lemma 5.7.2. There exists a construction of $\{\mathbf{BSR}^*(t)\}_{t\geq 0}$ along with two other random graph processes $\{\mathbf{G}^{(n),-}(t)\}_{t\geq 0}$ and $\{\mathbf{G}^{(n),+}(t)\}_{t\geq 0}$ such that:
(i) With high probability,

$$\mathbf{G}^{(n),-}(t^\lambda) \subset \mathbf{BSR}^*(t^\lambda) \subset \mathbf{G}^{(n),+}(t^\lambda) \qquad \text{for all } \lambda \in [-\lambda_n, \lambda_n]. \tag{5.7.2}$$

(ii) Let for $i \geq 1$, $\bar{\boldsymbol{C}}_i^{(n),\mp}(\lambda) = \frac{\beta^{1/3}}{n^{2/3}} |\mathcal{C}_i\left(\mathbf{G}^{(n),\mp}(t^\lambda)\right)|$ and

$$\bar{\boldsymbol{Y}}_i^{(n),\mp}(\lambda) = \xi_i\left(\mathbf{G}^{(n),\mp}(t^\lambda)\right) - \sum_{j:\mathcal{C}_j\left(\mathbf{G}^{(n),\mp}(t_n)\right)\subset \mathcal{C}_i\left(\mathbf{G}^{(n),\mp}(t^\lambda)\right)} \xi_j\left(\mathbf{G}^{(n),\mp}(t_n)\right).$$

Then, for all $\lambda \in \mathbb{R}$

$$(\bar{\boldsymbol{C}}^{(n),\bullet}(\lambda), \bar{\boldsymbol{Y}}^{(n),\bullet}(\lambda)) \xrightarrow{d} (\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda)), \ \bullet = -, +,$$

where $\xrightarrow{d}$ denotes weak convergence in $\mathbb{U}_{\downarrow}$.

We remark that $\bar{\boldsymbol{Y}}^{(n),\mp}(\lambda)$ represents the surplus in $\mathbf{G}^{(n),\mp}(t^\lambda)$ created after time instant t_n. Proof of the lemma relies on the following proposition which is an immediate consequence of Theorem 5.3.2 and Theorem 5.6.9.

Proposition 5.7.3. There exists a $\kappa \in (0, \frac{1}{3} - \gamma)$ such that

$$\frac{\bar{s}_3(t_n)}{(\bar{s}_2(t_n))^3} \xrightarrow{\mathbb{P}} \beta, \quad \frac{n^{1/3}}{\bar{s}_2(t_n)} - \frac{n^{1/3-\gamma}}{\alpha} \xrightarrow{\mathbb{P}} 0, \quad \frac{I^{(n)}(t_n)}{n^{2\gamma+\kappa}} \xrightarrow{\mathbb{P}} 0.$$

We now prove Lemma 5.7.2.

Proof of Lemma 5.7.2: We suppress n in the notation for the random graph processes. Write $t_n^+ := t_c + n^{-\gamma}$. Let $\mathbf{BSR}(t)$ for $t \in [0, t_n^+]$ be constructed as in Section 5.2.2 and define $\mathbf{BSR}^*(t)$ for $t \in [0, t_n)$ as in Section 5.6. Set

$$\mathbf{G}^{(n),-}(t) = \mathbf{G}^{(n),+}(t) = \mathbf{BSR}^*(t), \text{ for } t \in [0, t_n).$$

We now give the construction of these processes for $t \in [t_n, t_n^+]$.

The construction is done in two rounds. In the first round, we construct processes $\mathbf{G}^{I,-}(t)$, $\mathbf{BSR}^{I,*}(t)$ and $\mathbf{G}^{I,+}(t)$ for $t \in [t_n, t_n^+]$ by using only the information about immigrations and attachments in $\mathbf{BSR}(t)$, while the edge formation between large components is ignored. We first construct the process $\{\mathbf{BSR}^I(t)\}_{t \in [t_n, t_n^+]}$ as follows. Let $\mathbf{BSR}^I(t_n) := \mathbf{BSR}(t_n)$. For $t > t_n$, $\mathbf{BSR}^I(t)$ is constructed along with and same as $\mathbf{BSR}(t)$, except for when

$$c_{t-}(\vec{v}) \in \{\vec{j} \in \Omega_K^4 : \vec{j} \in F, j_1 = j_2 = \varpi \quad \text{or} \quad \vec{j} \notin F, j_3 = j_4 = \varpi\},$$

in which case no edge is added to $\mathbf{BSR}^I(t)$.

Let $\bar{x}_i(t), a_i^*(t), b^*(t), c_i^*(t)$, $1 \le i \le K$, $t \in [t_n, t_n^+]$, be the processes determined from $\{\mathbf{BSR}(t)\}_{t \in [t_n, t_n^+]}$ as in Section 5.6. These processes will be used in the second round of the construction.

Now define $\mathbf{BSR}^{I,*}(t)$ to be the subgraph that consists of all large components (components of size greater than K) in $\mathbf{BSR}^I(t)$, and then define $\mathbf{G}^{I,-}(t)$ and $\mathbf{G}^{I,+}(t)$ for $t \in [t_n, t_n^+]$ as follows:

$$\mathbf{G}^{I,-}(t) \equiv \mathbf{BSR}^{I,*}(t_n), \text{ and } \mathbf{G}^{I,+}(t) \equiv \mathbf{BSR}^{I,*}(t_n^+).$$

Then

$$\mathbf{G}^{I,-}(t) \subset \mathbf{BSR}^{I,*}(t) \subset \mathbf{G}^{I,+}(t) \text{ for all } t \in [t_n, t_n^+].$$

We now proceed to the second round of the construction. Let

$$E_n = \left\{ b(t_c) - n^{-1/6} < b^*(t) < b(t_c) + n^{-1/6}, \text{ for all } t \in [t_n, t_n^+] \right\}.$$

Note that Lemma 5.6.1 and (5.6.8) implies that with probability at least $1 - C_1 e^{-C_2 n^{1/5}}$,

$$\begin{aligned} \sup_{t \in (t_n, t_n^+)} |b^*(t) - b(t_c)| \leq & \sup_{t \in (t_n, t_n^+)} |b^*(t) - b(t)| + \sup_{t \in (t_n, t_n^+)} |b(t) - b(t_c)| \\ \leq & d_1 n^{-2/5} + d_2 n^{-\gamma} = o(n^{-1/6}). \end{aligned}$$

Thus $\mathbb{P}\{E_n^c\} \to 0$ as $n \to \infty$. Since we only need the coupling to be good with high probability, it suffices to construct the coupling of the three processes until the first time $t \in [t_n, t_n^+]$ when $b^*(t) \notin [b(t_c) - n^{-1/6}, b(t_c) + n^{-1/6}]$. Equivalently, we can assume without loss of generality that $b^*(t) \in [b(t_c) - n^{-1/6}, b(t_c) + n^{-1/6}]$, for all $t \in [t_n, t_n^+]$, a.s.

We will construct $\mathbf{G}^+(t)$, $\mathbf{BSR}^*(t)$ and $\mathbf{G}^-(t)$ by adding new edges between components in the three random graph processes $\mathbf{G}^{I,-}(t)$, $\mathbf{BSR}^{I,*}(t)$ and $\mathbf{G}^{I,+}(t)$ such that, at time $t \in [t_n, t_n^+]$ edges are added between each pair of vertices in $\mathbf{G}^{I,-}(t)$, $\mathbf{BSR}^{I,*}(t)$ and $\mathbf{G}^{I,+}(t)$, at rates $\frac{1}{n}(b(t_c) - n^{-1/6})$, $\frac{1}{n} b^*(t)$ and $\frac{1}{n}(b(t_c) + n^{-1/6})$, respectively. The precise mechanism is as follows.

We first construct $\mathbf{G}^+(t)$ for $t \in (t_n, t_n^+]$ by adding edges between every pair of vertices in $\mathbf{G}^{I,+}(t)$ at the rate $\frac{1}{n}(b(t_c) + n^{-1/6})$ and creating self-loops at the rate $\frac{1}{2n}(b(t_c) + n^{-1/6})$ for each vertex in $\mathbf{G}^{I,+}(t)$.

Next, we construct $\mathbf{BSR}^*(t)$ and $\mathbf{G}^-(t)$ through successive thinning of $\mathbf{G}^+(t)$, thus obtaining the desired coupling. Let $(e_1, e_2, ...)$ be the sequence of edges that are added to $\mathbf{G}^{I,+}(t)$ to obtain $\mathbf{G}^{+}(t)$. Let $(u_1, u_2, ...)$ be i.i.d Uniform$[0,1]$ random variables that are also independent of the random variables used to construct $\mathbf{G}^{I,-}, \mathbf{BSR}^{I,*}, \mathbf{G}^{I,+}, \mathbf{G}^{+}$. Suppose at time t_k, we have $\mathbf{G}^{+}(t_k) = \mathbf{G}^{+}(t_k-) \cup \{e_k\}$, where $e_k = \{v_1, v_2\}$. We set $\mathbf{BSR}^{*}(t_k) = \mathbf{BSR}^{*}(t_k-) \cup \{e_k\}$ if and only if

$$v_1, v_2 \in \mathbf{BSR}^{I,*}(t_k-) \text{ and } u_k \leq \frac{b^*(t_k)}{b(t_c) + n^{-1/6}},$$

otherwise let $\mathbf{BSR}^{*}(t_k) = \mathbf{BSR}^{*}(t_k-)$. This defines the process $\mathbf{BSR}^{*}(t)$ (with the correct probability law) such that the second inclusion in (5.7.2) is satisfied. Finally, construct $\mathbf{G}^{-}(t)$ by a thinning of $\mathbf{BSR}^{*}(t)$ exactly as above by replacing $\frac{b^*(t_k)}{b(t_c)+n^{-1/6}}$ with $\frac{b(t_c)-n^{-1/6}}{b^*(t_k)}$. Then $\mathbf{G}^{-}(t)$, for $t \in [t_n, t_n^+]$ is an Erdős-Rényi type processes and the first inclusion in (5.7.2) is satisfied. This completes the proof of the first part of the lemma.

We now prove (ii). Consider first the case $\bullet = -$. We will apply Theorem 5.5.1. With notation as in that theorem, it follows from the Erdős-Rényi dynamics of $\mathbf{G}^{(n),-}(t)$ that, the distribution of $(\bar{\boldsymbol{C}}^{(n),-}(\lambda), \bar{\boldsymbol{Y}}^{(n),-}(\lambda))$, conditioned on $\{\mathcal{P}_{\vec{v}}(t), t \leq t_n \; \vec{v} \in [n]^4\}$, for each $\lambda \in [-\lambda_n, \lambda_n]$, is same as the distribution of $\boldsymbol{Z}(z^{(n)}, q^{(n)})$, where $z^{(n)} = (\bar{\boldsymbol{C}}^{(n),-}(-\lambda_n), \mathbf{0})$, $\mathbf{0}$ denotes the vector $(0, 0, \cdots)$ and $q^{(n)}$ is determined by the equality

$$\begin{aligned} & q^{(n)} \bar{\boldsymbol{C}}_i^{(n),-}(-\lambda_n) \bar{\boldsymbol{C}}_j^{(n),-}(-\lambda_n) \\ = & \frac{\alpha\beta^{2/3}}{n^{1/3}}(\lambda + \lambda_n) \frac{(b(t_c) - n^{-1/6})}{n} |\mathcal{C}_i(\mathbf{G}^{(n),-}(t_n))||\mathcal{C}_j(\mathbf{G}^{(n),-}(t_n))|, \end{aligned}$$

for $i \neq j$. Recalling that $\alpha b(t_c) = 1$ it then follows that $q^{(n)} = \lambda + \frac{n^{1/3-\gamma}}{\alpha\beta^{2/3}} + O(n^{1/6-\gamma})$. We now verify the conditions of Theorem 5.5.1. Taking $x^{(n)} = \bar{\boldsymbol{C}}^{(n),-}(-\lambda_n)$ we see with, x^*, s_k, $k = 1, 2, 3$ as in Theorem 5.5.1,

$$s_1^{(n)} \leq \beta^{1/3} n^{1/3}, \;\; s_2^{(n)} = \frac{\beta^{2/3}}{n^{4/3}} \sum_{\mathcal{C} \subset \mathbf{BSR}^*(t_n)} |\mathcal{C}|^2, \;\; s_3^{(n)} = \frac{\beta}{n^2} \sum_{\mathcal{C} \subset \mathbf{BSR}^*(t_n)} |\mathcal{C}|^3.$$

Recall the definition of $\bar{s}_k$ and $\bar{s}_{k,\varpi}$ from (5.3.1) and Section 5.6. Then

$$s_2^{(n)} = \frac{\beta^{2/3} \bar{s}_{2,\varpi}(t_n)}{n^{1/3}}, \;\; s_3^{(n)} = \frac{\beta \bar{s}_{3,\varpi}(t_n)}{n}, \;\; x^{*(n)} = \beta^{1/3} \frac{I(t_n)}{n^{2/3}}.$$

From the first two convergences in Proposition 5.7.3 and recalling that, for $k = 1, 2$, $|\bar{s}_{k,\varpi} - \bar{s}_k| \le K^k$, we immediately get that the first two convergences in (5.5.1) hold. Also,

$$\frac{x^*}{s_2} = \frac{I(t_n)}{\beta^{1/3} n^{1/3} \bar{s}_{2,\varpi}(t_n)} = \frac{I(t_n)}{\beta^{2/3} n^{\gamma+1/3}} O(1) \to 0, \text{ in probability,}$$

where the second equality is consequence of the second convergence in Proposition 5.7.3, and the convergence of the last term follows from the third convergence in Proposition 5.7.3. This proves the third convergence in (5.5.1).

Finally we note that the convergence in (5.5.2) holds with $\varsigma = \frac{1}{1-3(\gamma+\kappa)}$, where κ is as in Proposition 5.7.3, since

$$s_1 \left(\frac{x^*}{s_2}\right)^{\varsigma} \le O(1) n^{1/3} \left(\frac{I(t_n)}{n^{\gamma+1/3}}\right)^{\varsigma} = O(1) \left(\frac{I(t_n)}{n^{2\gamma+\kappa}}\right)^{\varsigma} \to 0,$$

where the last equality follows from our choice of ς and the convergence is a consequence of Proposition 5.7.3. Thus we have verified all the conditions in Theorem 5.5.1 and therefore we have from this result that $(\bar{\boldsymbol{C}}^{(n),-}(\lambda), \bar{\boldsymbol{Y}}^{(n),-}(\lambda))$ converges in distribution, in $\mathbb{U}_{\downarrow}$, to $(\boldsymbol{X}^*(\lambda), \boldsymbol{Y}^*(\lambda))$ proving part (ii) of the lemma for $\bullet = -$.

To prove part (ii) of the lemma for $\bullet = +$, one needs slightly more work. Once more we will apply Theorem 5.5.1. As before, conditioned on $\{\bar{\boldsymbol{C}}^{(n),+}(\lambda_0) : \lambda_0 \le -\lambda_n\}$, for each $\lambda \in [-\lambda_n, \lambda_n]$, the distribution of $(\bar{\boldsymbol{C}}^{(n),+}(\lambda), \bar{\boldsymbol{Y}}^{(n),+}(\lambda))$ is same as the distribution of $\boldsymbol{Z}(\bar{z}^{(n)}, \bar{q}^{(n)})$, where $\bar{z}^{(n)} = (\bar{\boldsymbol{C}}^{(n),+}(-\lambda_n), \mathbf{0})$ and $\bar{q}^{(n)} = \lambda + \frac{n^{1/3-\gamma}}{\alpha\beta^{2/3}} + O(n^{1/6-\gamma})$. Taking $x^{(n)} = \bar{\boldsymbol{C}}^{(n),+}(-\lambda_n)$ we see with, $x^*, s_k,\ k = 1, 2, 3$ as in Theorem 5.5.1,

$$s_1^{(n)} \le \beta^{1/3} n^{1/3}, \quad s_2^{(n)} = \frac{\beta^{2/3}}{n^{4/3}} \sum_{\mathcal{C} \subset \mathbf{BSR}^{I,*}(t_n^+)} |\mathcal{C}|^2, \quad s_3^{(n)} = \frac{\beta}{n^2} \sum_{\mathcal{C} \subset \mathbf{BSR}^{I,*}(t_n^+)} |\mathcal{C}|^3.$$

Next note that for any component $\mathcal{C} \subset \mathbf{G}^-(t_n) = \mathbf{BSR}^{I,*}(t_n)$ there is a unique component $\mathcal{C}^+ \subset \mathbf{G}^+(t_n) = \mathbf{BSR}^{I,*}(t_n^+)$, such that $\mathcal{C} \subset \mathcal{C}^+$. Denote by $\mathcal{C}_i$ the i-th largest component in $\mathbf{BSR}^{I,*}(t_n)$, and let $\mathcal{C}_i^+$ be the corresponding component in

$\mathbf{BSR}^{I,*}(t_n^+)$ such that $\mathcal{C}_i \subset \mathcal{C}_i^+$. Denote by N the number of immigrations that occur during $[t_n, t_n^+]$ in $\mathbf{BSR}^{I,*}$, and denote by $\{\tilde{\mathcal{C}}_i^+\}_{i=1}^N$ the components in $\mathbf{BSR}^{I,*}(t_n^+)$ resulting from these immigrations. Then

$$s_2^{(n)} = \frac{\beta^{2/3}\bar{s}_2^+}{n^1/3}, \quad s_3^{(n)} = \frac{\beta \bar{s}_3^+}{n}, \quad x^{*(n)} = \beta^{1/3}\frac{I^+}{n^{2/3}},$$

where

$$\begin{aligned}
\bar{s}_2^+ :=& \frac{1}{n}\left(\sum_{i=1}^{\infty} |\mathcal{C}_i^+|^2 + \sum_{i=1}^{N} |\tilde{\mathcal{C}}_i^+|^2\right), \\
\bar{s}_3^+ :=& \frac{1}{n}\left(\sum_{i=1}^{\infty} |\mathcal{C}_i^+|^3 + \sum_{i=1}^{N} |\tilde{\mathcal{C}}_i^+|^3\right), \\
I^+ :=& \max\left\{\max_i |\mathcal{C}_i^+|, \max_i |\tilde{\mathcal{C}}_i^+|\right\}.
\end{aligned}$$

To complete the proof it suffices to show that the statement in Proposition 5.7.3 holds with $(\bar{s}_2(t_n), \bar{s}_3(t_n), I^{(n)}(t_n))$ replaced with $(\bar{s}_2^+, \bar{s}_3^+, I^+)$. This follows from Proposition 5.7.4 given below and hence completes the proof of the lemma. ■

Proposition 5.7.4. With notation as in the proof of Lemma 5.7.2, as $n \to \infty$, we have

$$I^+ = O(I), \quad \frac{\bar{s}_2^+}{\bar{s}_2(t_n)} \xrightarrow{\mathbb{P}} 1, \quad \frac{\bar{s}_3^+}{\bar{s}_3(t_n)} \xrightarrow{\mathbb{P}} 1, \quad \frac{n^{1/3}}{\bar{s}_2(t_n)} - \frac{n^{1/3}}{\bar{s}_2^+} \xrightarrow{\mathbb{P}} 0.$$

Proof. The proof is similar to that of Proposition 3.7.1 in Chapter 3 thus we only give a sketch.

Observe that the total rate of attachments is $\sum_{i=1}^K c_i^*(t) \leq 1$ and each attachment has size no bigger than K. Recall that $\mathcal{C}_i$ denotes the i-th largest component in $\mathbf{BSR}^{I,*}(t_n)$. Denote by $V_i(t)$, $t \in [t_n, t_n^+]$, the stochastic process defining the size of the component containing $\mathcal{C}_i$ in $\mathbf{BSR}^{I,*}(t)$. Note that $V_i(t_n) = |\mathcal{C}_i|$ and $V_i(t_n^+) = |\mathcal{C}_i^+|$. Then $V_i(t)/K$ can be stochastically dominated by a Yule process starting with $\lceil |\mathcal{C}_i|/K \rceil$ particles and birth rate K. Using this and an argument similar to Chapter 3, it follows that,

$$|\mathcal{C}_i^+| - |\mathcal{C}_i| \leq_d K \cdot \text{Negative-Binomial}(\lceil |\mathcal{C}_i|/K \rceil, e^{-2Kn^{-\gamma}}).$$

Next, note that the immigrations are of size no bigger than $2K$, and thus for the same reason, we have the bound,

$$|\tilde{\mathcal{C}}_i^+| \leq_d 2K + K \cdot \text{Negative-Binomial}(2, e^{-2Kn^{-\gamma}}).$$

Since the total number of vertices is n, the number of immigrations N can be bounded by n/K.

With the above three bounds the proof of the proposition follows exactly as the proof of Proposition 3.7.1 in 3 with obvious changes needed due to the constant K that appears in the above bounds. Details are omitted. ■

The next proposition says that the inclusion in (5.7.2) can be strengthened to component-wise inclusion.

Proposition 5.7.5. Fix $\lambda \in \mathbb{R}$ and $i_0 \geq 1$. Then, as $n \to \infty$,

$$\mathbb{P}\left\{\mathcal{C}_i(\mathbf{G}^{(n),-}(t^\lambda)) \subset \mathcal{C}_i(\mathbf{BSR}^*(t^\lambda)) \subset \mathcal{C}_i(\mathbf{G}^{(n),+}(t^\lambda)) \quad \forall\, 1 \leq i \leq i_0\right\} \to 1.$$

Proof: From Lemma 5.7.2 and Lemma 15 in [4] (see Chapter 3 for a similar argument), we have, as $n \to \infty$,

$$(\bar{\boldsymbol{C}}^{(n),-}(\lambda), \bar{\boldsymbol{C}}^{(n),*}(\lambda), \bar{\boldsymbol{C}}^{(n),+}(\lambda)) \xrightarrow{d} (\boldsymbol{X}(\lambda), \boldsymbol{X}(\lambda), \boldsymbol{X}(\lambda)), \tag{5.7.3}$$

in $l^2_\downarrow \times l^2_\downarrow \times l^2_\downarrow$, where $\boldsymbol{X}$ is as in Theorem 5.3.1. Define events E_n, F_n as

$$E_n = \left\{\bar{\boldsymbol{C}}_i^{(n),-}(\lambda) > \bar{\boldsymbol{C}}_{i+1}^{(n),+}(\lambda) : 1 \leq i \leq i_0\right\}, F_n = \left\{\mathbf{G}^{(n),-}(\lambda) \subset \mathbf{BSR}^*(\lambda) \subset \mathbf{G}^{(n),+}(\lambda)\right\}.$$

Then on the set $E_n \cap F_n$

$$\mathcal{C}_i(\mathbf{G}^{(n),-}(\lambda)) \subset \mathcal{C}_i(\mathbf{BSR}^*(\lambda)) \subset \mathcal{C}_i(\mathbf{G}^{(n),+}(\lambda)), \quad \forall\, 1 \leq i \leq i_0.$$

From Lemma 5.7.2 (i) $\mathbb{P}\{F_n^c\} \to 1$. Also

$$\begin{aligned}\limsup_n \mathbb{P}(E_n^c) \leq & \limsup_n \mathbb{P}\left\{\bar{\boldsymbol{C}}_i^{(n),-}(\lambda) \leq \bar{\boldsymbol{C}}_{i+1}^{(n),+}(\lambda) \text{ for some } 1 \leq i \leq i_0\right\} \\ \leq & \mathbb{P}\left\{\boldsymbol{X}_i(\lambda) \leq \boldsymbol{X}_{i+1}(\lambda) \text{ for some } 1 \leq i \leq i_0\right\} = 0.\end{aligned}$$

This shows that $\mathbb{P}(E_n \cap F_n) \to 1$ as $n \to \infty$. The result follows. ■

We will also need the following elementary lemma. Proof is omitted.

Lemma 5.7.6. Let $\eta^{(n),-}, \eta^{(n),+}, \eta^*$ be real random variables such that $\eta^{(n),-} \le \eta^{(n),+}$ with high probability. Further assume $\eta^{(n),-} \xrightarrow{d} \eta^*$ and $\eta^{(n),+} \xrightarrow{d} \eta^*$. Then $\eta^{(n),+} - \eta^{(n),-} \xrightarrow{\mathbb{P}} 0$. Furthermore, if $\eta^{(n)}$ are random variables such that $\eta^{(n),-} \le \eta^{(n)} \le \eta^{(n),+}$ with high probability, then $\eta^{(n)} \xrightarrow{d} \eta^*$ and $\eta^{(n)} - \eta^{(n),-} \xrightarrow{\mathbb{P}} 0$.

We now complete the proof of Theorem 5.7.1.

Proof of Theorem 5.7.1: From Lemma 5.7.2 (ii) we have that

$$\left(\bar{\boldsymbol{C}}^{(n),-}(\lambda), \bar{\boldsymbol{Y}}^{(n),-}(\lambda), \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),-}(\lambda) \boldsymbol{Y}_i^{(n),-}(\lambda)\right) \xrightarrow{d} \left(\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda), \sum_{i=1}^{\infty} \boldsymbol{X}_i(\lambda) \boldsymbol{Y}_i(\lambda)\right), \tag{5.7.4}$$

in $l_{\downarrow}^2 \times \mathbb{N}^\infty \times \mathbb{R}$, where on $\mathbb{N}^\infty$ we consider the product topology.

In order to prove the theorem it suffices, in view of Lemma 5.4.11, to show that

$$\left(\bar{\boldsymbol{C}}^{(n),*}(\lambda), \hat{\xi}^{(n)}(\lambda), \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \hat{\xi}_i^{(n)}(\lambda)\right) \xrightarrow{d} \left(\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda), \sum_{i=1}^{\infty} \boldsymbol{X}_i(\lambda) \boldsymbol{Y}_i(\lambda)\right), \tag{5.7.5}$$

in $l_{\downarrow}^2 \times \mathbb{N}^\infty \times \mathbb{R}$.

From Proposition 5.7.5, we have for any $i_0 \in \mathbb{N}$, with high probability

$$\bar{\boldsymbol{Y}}_i^{(n),-}(\lambda) \le \hat{\xi}_i^{(n)}(\lambda) \le \bar{\boldsymbol{Y}}_i^{(n),+} \text{ for } 1 \le i \le i_0.$$

Also, from Lemma 5.7.2 (i), whp,

$$\sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),-}(\lambda) \bar{\boldsymbol{Y}}_i^{(n),-}(\lambda) \le \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n)}(\lambda) \bar{\boldsymbol{Y}}_i^{(n)}(\lambda) \le \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),+}(\lambda) \bar{\boldsymbol{Y}}_i^{(n),+}(\lambda).$$

From Lemma 5.7.6 and Lemma 5.7.2 (ii), we then have

$$\left(\left|\hat{\xi}^{(n)}(\lambda) - \bar{\boldsymbol{Y}}^{(n),-}(\lambda)\right|, \ \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \hat{\xi}_i^{(n)}(\lambda) - \sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n),-}(\lambda) \bar{\boldsymbol{Y}}_i^{(n),-}(\lambda)\right) \xrightarrow{\mathbb{P}} 0,$$

in $\mathbb{N}^\infty \times \mathbb{R}$, where for $y = (y_1, y_2, \cdots) \in \mathbb{Z}^\infty$, $|y| = (|y_1|, |y_2|, \cdots)$. The convergence in (5.7.5) now follows on combining (5.7.4) and (5.7.3). The result follows. ■

5.7.2 Proof of Theorem 5.3.3.

As a first step towards the proof we show the following convergence result for one dimensional distributions.

Theorem 5.7.7. For every $\lambda \in \mathbb{R}$, as $n \to \infty$, $(\bar{\boldsymbol{C}}^{(n)}(\lambda), \bar{\boldsymbol{Y}}^{(n)}(\lambda))$ converges in distribution, in $\mathbb{U}_\downarrow$, to $(\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda))$.

Proof. Fix $\lambda \in \mathbb{R}$. We first argue that

$$(\bar{\boldsymbol{C}}^{(n),*}(\lambda), \bar{\boldsymbol{Y}}^{(n),*}(\lambda)) \xrightarrow{d} (\boldsymbol{X}(\lambda), \boldsymbol{Y}(\lambda)), \text{ in } \mathbb{U}_\downarrow. \tag{5.7.6}$$

For this, it suffices to show that

$$\sum_{i=1}^{\infty} \tilde{\xi}_i^{(n)}(\lambda) \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \xrightarrow{\mathbb{P}} 0. \tag{5.7.7}$$

Define

$$E_n = \left\{ I^{(n)}(s) \le C_3 \frac{(\log n)^4}{(t_c - s)^2} \text{ for } s \le t_c - n^{-\gamma} \right\}.$$

By Theorem 5.6.9, $\mathbb{P}\{E_n^c\} \to 0$ and $E_n \in \tilde{\mathcal{F}}(\lambda) := \sigma\{|\mathcal{C}_i(s)| : i \ge 1, s \le t^\lambda\}$ for all $\lambda \ge -\lambda_n$. We begin by showing that there exists $d_1 \in (0, \infty)$ such that, for all $i \in \mathbb{N}$,

$$\mathbb{E}\left[\tilde{\xi}_i^{(n)}(\lambda) \mid \tilde{\mathcal{F}}_\lambda\right] \mathbb{1}_{E_n} \le d_1 \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) n^{\gamma - 1/3} (\log n)^4. \tag{5.7.8}$$

Note that at any time $s < t^\lambda$, for a component $\mathcal{C} \subset \mathbf{BSR}^{(n)}(s)$, there are at most $2|\mathcal{C}|^2 n^2$ quadruples of vertices which may provide a surplus edge within $\mathcal{C}$. Since edges are formed at rate $2/n^3$, we have that

$$\begin{aligned}
\mathbb{E}\left[\tilde{\xi}_i^n(\lambda) \mid \tilde{\mathcal{F}}_\lambda\right] \le & \int_0^{t_n} \left[\sum_{j:\mathcal{C}_j(\mathbf{BSR}^{(n)}(s)) \subset \mathcal{C}_i(\mathbf{BSR}^*(t^\lambda))} \frac{1}{2n^3} 2n^2 |\mathcal{C}_j(\mathbf{BSR}^{(n)}(s))|^2 \right] ds \\
\le & \frac{1}{n} |\mathcal{C}_i(\mathbf{BSR}^*(t^\lambda))| \int_0^{t_n} I(s) ds.
\end{aligned}$$

Thus, for some $d_0, d_1 \in (0,\infty)$

$$\mathbb{E}\left[\tilde{\xi}_i^{(n)}(\lambda) \mid \tilde{\mathcal{F}}(\lambda)\right] \mathbb{1}_{E_n} \le d_0 \frac{\bar{\boldsymbol{C}}_i^{(n),*}(\lambda)}{n^{1/3}} \int_0^{t_c - n^{-\gamma}} \frac{(\log n)^4}{(t_c - s)^2} ds \le d_1 \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) n^{\gamma - 1/3} (\log n)^4.$$

This proves (5.7.8). As an immediate consequence of this inequality we have that

$$\begin{aligned}\mathbb{E}\left[\sum_i \tilde{\xi}_i^{(n)}(\lambda) \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \mid \tilde{\mathcal{F}}(\lambda)\right] \mathbb{1}_{E_n} =& \sum_i \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \mathbb{1}_{E_n} \mathbb{E}\left[\tilde{\xi}_i^{(n)}(\lambda) \mid \tilde{\mathcal{F}}(\lambda)\right] \\ \le& d_1 n^{\gamma - 1/3} (\log n)^4 \sum_i \left(\bar{\boldsymbol{C}}_i^{(n),*}(\lambda)\right)^2.\end{aligned}$$

Observing that $\gamma - 1/3 < 0$ and, from Theorem 5.7.1, that $\sum_i \left(\bar{\boldsymbol{C}}_i^{(n),*}(\lambda)\right)^2$ converges in distribution, we have that

$$\mathbb{E}\left[\sum_i \tilde{\xi}_i^{(n)}(\lambda) \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \mid \tilde{\mathcal{F}}(\lambda)\right] \mathbb{1}_{E_n} \xrightarrow{\mathbb{P}} 0.$$

Since $\mathbb{P}\{E_n\} \to 1$, letting $\eta^{(n)} = \sum_i \tilde{\xi}_i^{(n)}(\lambda) \bar{\boldsymbol{C}}_i^{(n),*}(\lambda)$, we have that $\mathbb{E}(\eta^{(n)} \mid \tilde{\mathcal{F}}(\lambda)) \to 0$ in probability. Convergence in (5.7.7) now follows from Markov's inequality on noting that, as $n \to \infty$,

$$\mathbb{E}[\eta^{(n)} \wedge 1] = E\left[E[\eta^{(n)} \wedge 1 \mid \tilde{\mathcal{F}}(\lambda)]\right] \le \mathbb{E}\left[\mathbb{E}[\eta^{(n)} \mid \tilde{\mathcal{F}}(\lambda)] \wedge 1\right] \to 0.$$

This proves (5.7.6). Next note that

$$\sum_{i=1}^{\infty} |\bar{\boldsymbol{C}}_i^{(n)}(\lambda) - \bar{\boldsymbol{C}}_i^{(n),*}(\lambda)|^2 \le \frac{n}{n^{4/3}} O(1) \to 0, \text{ as } n \to \infty. \tag{5.7.9}$$

Also,

$$\begin{aligned}\mathbb{E}\left[\bar{\boldsymbol{Y}}_i^{(n)}(\lambda) \mid \tilde{\mathcal{F}}(\lambda)\right] \mathbb{1}_{\{|\mathcal{C}_i(t^\lambda)| \le K\}} \le& \left[\int_0^{t^\lambda} \sum_{j: \mathcal{C}_j(s) \subset \mathcal{C}_i(t^\lambda)} \frac{1}{2n^3} 2n^2 |\mathcal{C}_j(s)|^2 ds\right] \mathbb{1}_{\{|\mathcal{C}_i(t^\lambda)| \le K\}} \\ \le& \frac{K^2}{n} \cdot t^\lambda = O(n^{-1}).\end{aligned}$$

Thus, as $n \to \infty$,

$$\begin{aligned}&\mathbb{E}\sum_{i=1}^{\infty} |\bar{\boldsymbol{C}}_i^{(n)}(\lambda) \bar{\boldsymbol{Y}}_i^{(n)}(\lambda) - \bar{\boldsymbol{C}}_i^{(n),*}(\lambda) \bar{\boldsymbol{Y}}_i^{(n),*}(\lambda)| \\ =& \sum_{i=1}^{\infty} \mathbb{E}\left[\bar{\boldsymbol{C}}_i^{(n)}(\lambda) \bar{\boldsymbol{Y}}_i^{(n)}(\lambda) \mathbb{1}_{\{|\mathcal{C}_i(t^\lambda)| \le K\}}\right] \le O(n^{-1}) \mathbb{E}\left[\sum_{i=1}^{\infty} \bar{\boldsymbol{C}}_i^{(n)}(\lambda)\right] = O(n^{-2/3}) \to 0.\end{aligned}$$

The result now follows on combining the above convergence with (5.7.9) and (5.7.6).

■

Remark 5.7.8. The proofs of Theorems 5.7.1 and 5.7.7 in fact establish the following stronger statement: For all $\lambda \in \mathbb{R}$,

$$\Bigg(|\bar{\boldsymbol{Y}}^{(n),-}(\lambda) - \bar{\boldsymbol{Y}}^{(n)}(\lambda)|, \sum_{i=1}^{\infty} |\bar{\boldsymbol{C}}_i^{(n),-}(\lambda) - \bar{\boldsymbol{C}}_i^{(n)}(\lambda)|^2,$$
$$\sum_{i=1}^{\infty} |\bar{\boldsymbol{C}}_i^{(n),-}(\lambda)\bar{\boldsymbol{Y}}_i^{(n),-}(\lambda) - \bar{\boldsymbol{C}}_i^{(n)}(\lambda)\bar{\boldsymbol{Y}}_i^{(n)}(\lambda)| \Bigg) \to (\mathbf{0}, 0, 0),$$

in probability, in $\mathbb{N}^\infty \times \mathbb{R} \times \mathbb{R}$.

Proof of Theorem 5.3.3: For simplicity we present the proof for the case $m = 2$. The general case can be treated similarly. Fix $-\infty < \lambda_1 < \lambda_2 < \infty$. Denote, for $\lambda \in \mathbb{R}$, $\bar{\boldsymbol{Z}}^{(n),-}(\lambda) = (\bar{\boldsymbol{C}}^{(n),-}(\lambda), \bar{\boldsymbol{Y}}^{(n),-}(\lambda))$. In view of Remark 5.7.8 it suffices to show that, as $n \to \infty$,

$$(\bar{\boldsymbol{Z}}^{(n),-}(\lambda_1), \bar{\boldsymbol{Z}}^{(n),-}(\lambda_2)) \xrightarrow{d} (\boldsymbol{Z}(\lambda_1), \boldsymbol{Z}(\lambda_2)),$$

for which it is enough to show that for all $f_1, f_2 \in C_b(\mathbb{U}_\downarrow^0)$

$$\mathbb{E}\left[f_1(\bar{\boldsymbol{Z}}^{(n),-}(\lambda_1)) f_2(\bar{\boldsymbol{Z}}^{(n),-}(\lambda_2)) \right] \to \mathbb{E}\left[f_1(\boldsymbol{Z}(\lambda_1)) f_2(\boldsymbol{Z}(\lambda_2)) \right]. \tag{5.7.10}$$

Note that the left side of (5.7.10) equals

$$\mathbb{E}\left[f_1(\bar{\boldsymbol{Z}}^{(n),-}(\lambda_1)) \mathcal{T}_{\lambda_2-\lambda_1} f_2(\bar{\boldsymbol{Z}}^{(n),-}(\lambda_1)) \right],$$

which using Theorem 5.3.1 (2), Lemma 5.7.2 (ii) and the fact that $\boldsymbol{Z}(\lambda) \in \mathbb{U}_\downarrow^1$ a.s., converges to

$$\mathbb{E}\left[f_1(\boldsymbol{Z}(\lambda_1)) \mathcal{T}_{\lambda_2-\lambda_1} f_2(\boldsymbol{Z}(\lambda_1)) \right] = \mathbb{E}\left[f_1(\boldsymbol{Z}(\lambda_1)) f_2(\boldsymbol{Z}(\lambda_2)) \right],$$

where the last equality follows from Theorem 5.3.1 (3). This proves (5.7.10) and the result follows. ■

CHAPTER 6: LDP FOR THE CONFIGURATION MODEL

6.1 Introduction

In this chapter we study some large deviation problems related to the configuration model. The configuration model random graph $\mathcal{G}(n, \{d_i\})$ is defined as follows. Given $\{d_i : i = 1, 2, ..., n\}$ satisfying $d_i \in \mathbb{N}$ and $\sum_{i=1}^{n} d_i$ is even, consider the collection of all graphs with the vertex set $[n]$ such that the degree of vertex i is d_i, for $i \in [n]$. Define $\mathcal{G}(n, \{d_i\})$ to be a uniformly random member in this collection. Further assume that there exists a probability distribution on $\mathbb{N}$, $\{p_k : k \in \mathbb{N}\}$, such that $\sum_{k=1}^{\infty} kp_k < \infty$ and for each $k \in \mathbb{N}$,

$$\frac{|\{i \in [n] : d_i = k\}|}{n} \to p_k, \text{ as } n \to \infty.$$

Let $\mathcal{C}_1^{(n)}$ be the size of the largest component in $\mathcal{G}(n, \{d_i\})$. The goal of this chapter is to argue that $\mathbb{P}\left\{\frac{1}{n}\mathcal{C}_1^{(n)} \in B\right\} = e^{-nI(B)+o(n)}$ as $n \to \infty$, and to identify the exponent $I(B)$ in the expression for $B \subset [0, 1]$. Towards this goal, we construct a random walk associated with the depth-first-exploration of $\mathcal{G}(n, \{d_i\})$, and then study the large deviation properties of this random walk. Our techniques are based on the weak convergence approach developed in [14]. However, due to certain singularities in the transition kernel of this random walk, the conditions imposed in [14] are not satisfied for this example, and we have only been able to prove a large deviation upper bound. Establishing the corresponding lower bound is an open problem.

Organization of this chapter: Section 6.2 gives some preliminary results, including the definition of the configuration model, the assumptions, and the random walk associated with the depth-first-exploration. Section 6.3 defines the rate function and conjectures a large deviation principle (LDP) for the scaled random walk process. We also present our main result which establishes a large deviation upper bound. In Section 6.4, the proof of the main LDP is reduced to the LDP of some modified process in Theorem 6.4.1. Theorem 6.4.1 is proved in Section 6.5 and Section 6.6.

6.2 Preliminaries

6.2.1 Notation

The following notation will be used. Given a Polish space $\mathcal{E}$ and $T > 0$, $C([0,T]) : \mathcal{E})$ (resp. $C([0,\infty) : \mathcal{E})$) will denote the space of $\mathcal{E}$ valued continuous functions on $[0,T]$ (resp. $[0,\infty)$) which is equipped with the usual uniform topology (resp. local uniform topology). Also, $D([0,T] : \mathcal{E})$ will denote the space of right continuous functions with left limits on $[0,T]$ with values in $\mathcal{E}$. This space is equipped with the usual Skorohod topology. $\mathcal{B}(\mathcal{E})$ will denote the Borel σ-field on $\mathcal{E}$. $\mathcal{P}(\mathcal{E})$ will denote the space of probability measures on $\mathcal{E}$ which is considered with the usual topology of weak convergence. We say a sequence of $\mathcal{E}$ valued random variables is tight if the corresponding sequence of induced measures on $\mathcal{E}$ is tight. For a function $f : \mathcal{E} \to \mathbb{R}$, define $\|f\|_\infty = \sup_{x\in\mathcal{E}} |f(x)|$.

We denote by $\mathbb{R}^\infty$ the space of all real sequences which is identified with the countable product of copies of $\mathbb{R}$. This space is equipped with the usual product topology. A function I from a Polish space $\mathcal{E}$ to $[0,\infty]$ is called a rate function if for each $M < \infty$ the set $\{x \in \mathcal{E} : I(x) \le M\}$ is compact. Given a measurable space $(\Omega, \mathcal{F})$ and a Polish space $\mathcal{E}$ a family of probability measures $\nu(dy \mid x)$ on $\mathcal{E}$ parametrized by $x \in \Omega$ is called a stochastic kernel on $\mathcal{E}$ given Ω if for every Borel set B in $\mathcal{E}$, the map $x \mapsto \nu(B \mid x)$ is measurable.

6.2.2 The configuration model and assumptions

Fix $n \geq 1$. We start by describing the construction of the configuration model of random networks [10] with vertex set $[n] := \{1, 2, \ldots, n\}$. Let $\mathbf{d}(n) = (d_i : i \in [n])$ be a degree sequence, namely a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. Let $2m(n) := \sum_{i=1}^n d_i$. Each d_i may depend on n but we suppress this dependence in the notation. Now start with the n vertices with vertex $i \in [n]$ having d_i half-edges. Perform a uniform random matching on these $2m$ half-edges to form m edges so that every edge is composed of two half-edges. This procedure creates a random multigraph $G([n], \mathbf{d}(n))$ with $m(n)$ edges, allowing for multiple edges and self-loops, and is called the configuration model with degree sequence $\mathbf{d}(n)$. Since we are concerned with connectivity properties of the resulting graph, vertices with degree zero play no role in our analysis, and therefore we assume that $d_i > 0$ for all $i \in [n]$, $n \geq 1$. We make the following further assumptions on the collection $\{\mathbf{d}(n), n \in \mathbb{N}\}$.

Assumption 6.2.1. There exists a probability mass function $\mathbf{p} := \{p_k\}_{k\geq 1}$ on $\mathbb{Z}_+ := \{1, 2, \ldots\}$ such that, $p_2 \neq 1$ and writing $n_k^{(n)} := |\left\{i : d_i^{(n)} = k\right\}|$ for the number of vertices with degree k,

$$\frac{n_k^{(n)}}{n} \to p_k \text{ as } n \to \infty, \text{ for all } k \geq 1$$

We will make the following exponential integrability assumption on the degree distribution.

Assumption 6.2.2. For all $\lambda \in \mathbb{R}$, $\sup_{n\in\mathbb{N}} \sum_{k=1}^{\infty} \frac{n_k^{(n)}}{n} e^{\lambda k} < \infty$.

Remark 6.2.3. (i) Note that Assumptions 6.2.1 and 6.2.2, along with Fatou's lemma say that $\sum_{k=1}^{\infty} p_k e^{\lambda k} < \infty$ for every $\lambda \in \mathbb{R}$. Conversely, if $\sum_{k=1}^{\infty} p_k e^{\lambda k} < \infty$ for every $\lambda \in \mathbb{R}$ and $\{D_i, i \geq 1\}$ is an i.i.d. sequence with common distribution $\{p_k\}_{k\geq 1}$, then for a.e. ω, Assumptions 6.2.1 and 6.2.2 are satisfied with $d_i^{(n)} = D_i(\omega)$, $i = 1, \cdots n$, $n \geq 1$.

(ii) Under Assumptions 6.2.1 and 6.2.2, $\mu := \sum_{k=1}^{\infty} kp_k < \infty$ and the total number of edges $m^{(n)} = \frac{1}{2}\sum_{i=1}^{n} d_i$ satisfies $\frac{m^{(n)}}{n} \to \frac{1}{2}\sum_{k=1}^{\infty} kp_k$ as $n \to \infty$.

6.2.3 Edge-exploration algorithm

One can construct $G(n, (d_i)_1^n)$ whilst simultaneously exploring its component structure [21] which we now describe. This algorithm traverses the graph by exploring all its edges, unlike typical graph exploration algorithms, which sequentially explore vertices. At each stage of the algorithm, every vertex is in one of two possible states, sleeping or awake, while each half-edge is in one of three states: sleeping (unexplored), active or dead (removed). Write $\mathcal{A}_{\mathbb{V}}(j), \mathcal{S}_{\mathbb{V}}(j)$ for the set of awake and sleeping vertices at time j and similarly let $\mathcal{S}_{\mathbb{E}}(j), \mathcal{A}_{\mathbb{E}}(j), \mathcal{D}_{\mathbb{E}}(j)$ be the set of sleeping, active and dead half-edges at time j. We call a half-edge “living” if it is either sleeping or active. Initialize by setting all vertices and half-edges to be in the sleeping state. For step $j \geq 0$, write $A(j) := |\mathcal{A}_{\mathbb{E}}(j)|$ for the number of active half-edges and $V_k(j)$ for the number of sleeping vertices $v \in \mathcal{S}_{\mathbb{V}}(j)$ with degree k. Write $\mathbf{V}(j) = (V_k(j) : k \in \mathbb{N})$ for the corresponding vector in $\mathbb{R}_+^{\infty}$ and let $S(j) := \sum_{k=1}^{\infty} V_k(j)$ for the total number of sleeping vertices at time j. $A(j), \mathbf{V}(j)$ are regarded as random variables given on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which we introduce the filtration $\mathcal{F}_n(k) = \sigma\{A(j), \mathbf{V}(j), j \leq k\}$, $k \geq 0$.

At time $j = 0$, all vertices and half-edges are asleep hence $A(0) = 0$ and for $k \geq 1$ and $V_k(0) = n_k$. The exploration process proceeds as follows:

Edge exploration algorithm (EEA):

(a) If the number of active half-edges and sleeping vertices $A(j) = 0$ and $S(j) = 0$, all vertices and half-edges have been explored and we terminate the algorithm.

(b) If $A(j) = 0$, and $S(j) > 0$ so there exist sleeping vertices, pick one such vertex with probability proportional to its degree (alternatively pick a sleeping half-edge uniformly at random) and mark the vertex as awake and all its half-edges as active. Thus on the set $\{A(j) = 0, S(j) > 0\}$,

$$\mathbb{P}\left(A(j+1) = k|\mathcal{F}(j)\right) = \frac{kV_k(j)}{\sum_{k=1}^{\infty} kV_k(j)}, \qquad \text{for } k \geq 1.$$

(c) If $A(j) > 0$, pick an active half-edge uniformly at random, pair it with another uniformly chosen living half-edge (either active or sleeping), say e^*, merge both half edges to form a full edge and kill both half-edges. If e^* was sleeping when picked, wake the vertex corresponding to half-edge e^*, and mark all its other half-edges active. Thus we have

$$\begin{aligned}\mathbb{P}\left\{A(j+1) - A(j) = -2|\mathcal{F}(j)\right\} &= \frac{A(j) - 1}{\sum_{k=1}^{\infty} kV_k(t) + A(j) - 1},\\ \mathbb{P}\left\{A(j+1) - A(j) = k - 2|\mathcal{F}(j)\right\} &= \frac{kV_k(j)}{\sum_{k=1}^{\infty} kV_k(j) + A(j) - 1} \text{ for } k \geq 1.\end{aligned} \tag{6.2.1}$$

The first expression covers the case when e^* was active before being killed, while the second expression corresponds to the case where e^* was sleeping and belonged to a degree-k vertex.

Construction of $G(n, \mathbf{d}(n))$: The random graph formed at the termination of the above algorithm has the same distribution as the configuration model.

Note that at each step in the EEA, either a new vertex is woken up or two half-edges are killed. Since there are a total of n vertices and $2m$ half-edges, we have from Assumptions 6.2.1 and 6.2.2 that the algorithm terminates in at most $m + n \leq nT$ steps where $T := 1 + \sup_n \frac{1}{2} \sum_{k=1}^{\infty} k \frac{n_k^{(n)}}{n}$. We define $A(j) \equiv 0$ and $\mathbf{V}(j) \equiv \mathbf{0}$ for all $j \geq j_0$ where j_0 is the time instant at which the algorithm terminates.

Scaling and Linear Interpolation: Note that for $j \in \mathbb{N}_0$, $(\mathbf{V}(j))$ is a $\mathbb{R}^\infty$ valued random variable. Define continuous time process $\mathbf{V}^n$ by scaling and linear interpolation of the discrete time sequence $\{\mathbf{V}(j)\}_{j\in\mathbb{N}_0}$ as follows.

$$\mathbf{V}^n(t) := \frac{1}{n}\mathbf{V}(\lfloor nt\rfloor) + \left(t - \frac{\lfloor nt\rfloor}{n}\right)[\mathbf{V}(\lfloor nt\rfloor + 1) - \mathbf{V}(\lfloor nt\rfloor)], \qquad t \in [0,T]. \quad (6.2.2)$$

Thus $\mathbf{V}^n(\cdot)$ is a random variable with values in $C([0,T] : \mathbb{R}^\infty)$.

6.3 Main result and the rate function

The main result of this chapter gives a large deviation upper bound for $\{\mathbf{V}^n\}_{n\in\mathbb{N}}$ in the path space $C([0,T] : \mathbb{R}^\infty)$. First, we define the corresponding rate function for the large deviation principle.

Let

$$\nu^* = \sup_n \sum_k k\frac{n_k^{(n)}}{n}, \quad \nu_p^* = \sup_n \sum_k e^{pk}\frac{n_k^{(n)}}{n}, \ p \geq 1.$$

Define

$$\mathfrak{D}_{\text{deg}} := \left\{\boldsymbol{\beta} = (\beta_k) \in \mathbb{R}^\infty : \beta_k \geq 0\,\forall k, \sum_k k\beta_k \leq \nu^*, \sum_k e^{pk}\beta_k \leq \nu_p^*, \forall p \geq 1\right\}$$

and let $\mathfrak{D}^*_{\text{deg}} := [0,\nu^*] \times \mathfrak{D}_{\text{deg}}$, $\bar{\mathfrak{D}}_{\text{deg}} := [0,\infty) \times \mathfrak{D}_{\text{deg}}$.

Define the stochastic kernel $\mu(\cdot \mid \cdot)$ on $\mathbb{N}_0$ given $\bar{\mathfrak{D}}_{\text{deg}}$ as follows. For $\mathbf{x} = (a,\boldsymbol{\beta}) \in \bar{\mathfrak{D}}_{\text{deg}}$ with $\sum_k k\beta_k + a > 0$, define

$$\mu(k \mid \mathbf{x}) := \begin{cases} \frac{a}{\sum_k k\beta_k + a} \text{ when } k = 0 \\ \frac{k\beta_k}{\sum_k k\beta_k + a} \text{ when } k > 0 \end{cases} \quad (6.3.1)$$

If $\sum_k k\beta_k + a = 0$, define $\mu(\cdot \mid \mathbf{x}) = \delta_0(\cdot)$.

Let $\mathfrak{D}_{\text{vel}} \subset \mathbb{R}^\infty$ be the set

$$\mathfrak{D}_{\text{vel}} := \left\{\boldsymbol{\gamma} = (\gamma_k)_{k\geq 1} \in \mathbb{R}^\infty : \gamma_k \leq 0, \text{ for all } k \in \mathbb{N}, \sum_{k=1}^\infty \gamma_k \geq -1\right\}.$$

For fixed $\boldsymbol{\gamma} \in \mathfrak{D}_{\text{vel}}$, define the probability measure $\nu(\cdot \mid \boldsymbol{\gamma})$ on $\mathbb{N}_0$,

$$\nu(k \mid \boldsymbol{\gamma}) := \begin{cases} 1 + \sum_{k=1}^{\infty} \gamma_k & \text{for } k = 0, \\ -\gamma_k & \text{for } k > 0. \end{cases} \tag{6.3.2}$$

For $\mu, \nu \in \mathcal{P}(\mathbb{N}_0)$, the relative entropy of μ with respect to ν is defined as

$$R(\nu \| \mu) := \int_{\mathbb{N}_0} \log \frac{d\nu}{d\mu}(y)\nu(dy) = \sum_{k=0}^{\infty} \nu(k) \log \frac{\nu(k)}{\mu(k)}, \tag{6.3.3}$$

when $\nu \ll \mu$, and $R(\nu \| \mu) := \infty$ otherwise. For $\mathbf{x} = (a, \boldsymbol{\beta}) \in \mathfrak{D}^*_{\text{deg}}$ and $\boldsymbol{\gamma} \in \mathfrak{D}_{\text{vel}}$ define

$$L(\mathbf{x}, \boldsymbol{\gamma}) := R\big(\nu(\cdot \mid \boldsymbol{\gamma}) \big\| \mu(\cdot \mid \mathbf{x})\big). \tag{6.3.4}$$

Recall the probability mass function $\mathbf{p} := \{p_k\}_{k \geq 1}$ introduced in Assumption 6.2.1. Define $\mathfrak{P} := C([0,T] : \mathbb{R}^\infty)$. Consider the subset $\mathfrak{P}_I$ of $\mathfrak{P}$ consisting of those functions $\mathbf{v}(\cdot) = (v_k(\cdot))_{k \geq 1}$ such that

(a) $v_k(0) = p_k$, $v_k(t) \geq 0$ for all $t \in [0,T]$, and $v_k(\cdot)$ is absolutely continuous for all $k \geq 1$.

(b) Writing $\dot{\mathbf{v}}(\cdot) = (\dot{v}_k(\cdot))_{k \geq 1}$ for the corresponding derivatives, for a.e. $t \in [0,T]$ $\dot{\mathbf{v}}(t) \in \mathfrak{D}_{\text{vel}}$.

We recall that the one dimensional Skorohod map $\Gamma : D([0,T] : \mathbb{R}) \to D([0,T] : \mathbb{R})$ is defined as follows. Given a function $b \in D([0,T] : \mathbb{R})$,

$$\Gamma(b)(t) := b(t) - \inf_{0 \leq s \leq t} (b(s) \wedge 0),\ t \in [0,T]. \tag{6.3.5}$$

For $\mathbf{v} \in \mathfrak{P}$ with $\mathbf{v}(\cdot) = (v_k(\cdot))_{k \geq 1}$, define two new functions $b^{\mathbf{v}}(\cdot), a^{\mathbf{v}}(\cdot) \in C([0,T] : \mathbb{R})$ via the operations,

$$b^{\mathbf{v}}(t) := \sum_{k=1}^{\infty} k v_k(0) - 2t - \sum_{k=1}^{\infty} k v_k(t), \qquad a^{\mathbf{v}}(t) := \Gamma(b^{\mathbf{v}})(t),\ t \geq 0. \tag{6.3.6}$$

and let $\mathbf{x}^{\mathbf{v}} = (a^{\mathbf{v}}(\cdot), \mathbf{v}(\cdot))$. Note that for every $\mathbf{v}(\cdot) \in \mathfrak{P}_I$, $\mathbf{x}^{\mathbf{v}}(t) \in \mathfrak{D}^*_{\text{deg}}$. Indeed, for $\mathbf{v} \in \mathfrak{P}_I$ and $t \geq 0$, $v_k(t) \in [0, p_k]$ for every $k \geq 1$ and so from Fatou's lemma $\mathbf{v}(t) \in \mathfrak{D}_{\text{deg}}$. Also,

$$\begin{aligned} a^{\mathbf{v}}(t) = & b^{\mathbf{v}}(t) - \inf_{0 \leq s \leq t} b^{\mathbf{v}}(s) = -\sum_k k v_k(t) - 2t - \inf_{0 \leq s \leq t} \left(-\sum_k k v_k(s) - 2s \right) \\ & \leq -\sum_k k v_k(t) - 2t + 2t + \sum_k k v_k(0) \leq \sum_k k v_k(0) \leq \nu^*. \end{aligned}$$

This shows $\mathbf{x}^{\mathbf{v}}(t) = (a^{\mathbf{v}}(t), \mathbf{v}(t)) \in \mathfrak{D}^*_{\text{deg}}$ for every $t \geq 0$.

Define the function $I : \mathfrak{P} \to [0, \infty]$ as

$$I(\mathbf{v}) := \begin{cases} \int_0^T L(\mathbf{x}^{\mathbf{v}}(t), \dot{\mathbf{v}}(t)) dt & \text{when } \mathbf{v} \in \mathfrak{P}_I, \\ \infty \quad \text{when } \mathbf{v} \notin \mathfrak{P}_I, \end{cases} \tag{6.3.7}$$

With the usual convention $0 \cdot \infty = 0$, this has the explicit form

$$\begin{aligned} I(\mathbf{v}) := \int_0^T \Bigg[& \left(1 + \sum_k \dot{v}_k(t) \right) \log \frac{1 + \sum_k \dot{v}_k(t)}{a^{\mathbf{v}}(t) / [\sum_k k v_k(t) + a^{\mathbf{v}}(t)]} \\ & - \sum_k \dot{v}_k(t) \log \frac{-\dot{v}_k(t)}{k v_k(t) / [\sum_k k v_k(t) + a^{\mathbf{v}}(t)]} \Bigg] \mathbf{1}_{\{\sum_k k v_k(t) + a^{\mathbf{v}}(t) > 0\}} dt. \end{aligned} \tag{6.3.8}$$

Recall from (6.2.2), the process $\mathbf{V}^n(\cdot)$ that keeps track of the degree sequence of sleeping vertices. The large deviation principle that we want to study is the following:

Conjecture 6.3.1. The function I in (6.3.7) is a rate function on $\mathfrak{P}$ and the sequence $\{\mathbf{V}^n\}_{n \in \mathbb{N}}$ satisfies a large deviation principle in $\mathfrak{P}$ with rate function I.

Our main result is the following theorem.

Theorem 6.3.2. (i) The function I in (6.3.7) is a rate function.

(ii) *(Large deviation upper bound)* For all closed set $F \subset \mathfrak{P}$,

$$\limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P} \{ \mathbf{V}^n \in F \} \leq \inf_{\mathbf{v} \in F} I(\mathbf{v}).$$

Whether the large deviation lower bound holds with the rate function I is an open problem.

6.4 Proof of the large deviation upper bound

The map $\mathbf{x} \mapsto \mu(\cdot \mid \mathbf{x})$ is not continuous at $\mathbf{x} = (0, \mathbf{0})$. This causes difficulty in proving various weak convergence results needed in the proof of the large deviations principle. To deal with this we first regularize the stochastic kernel $\mu(\cdot \mid \cdot)$ as follows.

6.4.1 Regularization of the stochastic kernel.

For fixed $\epsilon > 0$ and $a, \mathbf{v} := (v_k)_{k\geq 1}$ such that $\mathbf{x} = (a, \mathbf{v}) \in \bar{\mathfrak{D}}_{\text{deg}}$, define the stochastic kernel,

$$\mu^\epsilon(k \mid \mathbf{x}) := \begin{cases} \frac{a^\epsilon}{(\sum_k kv_k + a)\vee\epsilon} & \text{for } k = 0 \\ \frac{kv_k}{(\sum_k kv_k + a)\vee\epsilon} & \text{for } k > 0 \end{cases}. \tag{6.4.1}$$

Here $a^\epsilon := (\epsilon - \sum_k kv_k) \vee a$. Note that $a^\epsilon + \sum_k kv_k = \epsilon$ when $a + \sum_k kv_k < \epsilon$, thus the above always defines a probability measure. Define

$$\Lambda_\epsilon := \left\{ (a, \mathbf{v}) \in \mathfrak{D}^*_{\text{deg}} : a + \sum_k kv_k \leq \epsilon \right\}.$$

Then $\mu^\epsilon(\cdot \mid \mathbf{x}) = \mu(\cdot \mid \mathbf{x})$ for $\mathbf{x} \in \Lambda_\epsilon^c$. Also, for every $\epsilon > 0$, $\mathbf{x} \mapsto \mu^\epsilon(\cdot \mid \mathbf{x})$ is a continuous map on $\bar{\mathfrak{D}}_{\text{deg}}$. We will now argue that it suffices to prove a large deviation principle for the case when the dynamics is driven by the regularized kernel μ^ϵ. In order to make this statement precise we begin by observing the following dynamical description of the sequence $\{A(j), \mathbf{V}(j)\}_{j\in\mathbb{N}_0}$. Recall that $A(0) = 0$ and $\mathbf{V}(0) = (n_k)_{k\geq 1}$. Define the sequence $\{\mathbf{X}(j)\}_{j\in\mathbb{N}_0}$ of $\mathbb{R}_+ \times \mathbb{R}^\infty$ valued random variables as

$$\mathbf{X}(j) := ((A(j) - 1)^+, \mathbf{V}(j)), \qquad j \geq 0. \tag{6.4.2}$$

For $j \geq 0$, let $\xi(j + 1)$ denote the degree of the vertex chosen at time $j + 1$, with $\xi(j + 1) = 0$ if two active half-edges are merged at time instant $j + 1$ thus resulting in no new vertex removed from the set of sleeping vertices. From (6.2.1) and (6.3.1),

$$\mathbb{P}(\xi(j+1)=k \mid \mathcal{F}(j)) = \mu(k \mid \frac{1}{n}\mathbf{X}(j)) \text{ for all } k \in \mathbb{N}_0, j \in \mathbb{N}_0,\ a.s. \tag{6.4.3}$$

and

$$V_k(j+1) = V_k(j) - \mathbf{1}_{\{\xi(j+1)=k\}},\ A(j+1) = A(j) + (\xi(j+1)-2) + 2\mathbf{1}_{\{A(j)\le 0\}}. \tag{6.4.4}$$

Thus $V_k(j) = n_k - \sum_{i=1}^{j} \mathbf{1}_{\{\xi(i)=k\}}$ and

$$A(j) = \sum_{i=1}^{j}(\xi(i)-2) + 2\sum_{i=1}^{j}\mathbf{1}_{\{A(i-1)\le 0\}}. \tag{6.4.5}$$

Now, for $\epsilon > 0$ the sequence $\{\mathbf{X}^\epsilon(j), \mathbf{V}^\epsilon(j), A^\epsilon(j)\}_{j\in\mathbb{N}_0}$ is defined as in (6.4.2)-(6.4.5) above by replacing $\mu(k \mid \frac{1}{n}\mathbf{X}(j))$ with $\mu^\epsilon(k \mid \frac{1}{n}\mathbf{X}^\epsilon(j))$. Define the continuous time process $\mathbf{V}^{n,\epsilon}$ through (6.2.2) by replacing $\{\mathbf{V}(j)\}$ on the right side of the equation with $\{\mathbf{V}^\epsilon(j)\}$. ext, for $\mathbf{x} \in \mathfrak{D}^*_{\text{deg}}$ and $\boldsymbol{\gamma} \in \mathfrak{D}_{\text{vel}}$, let

$$L^\epsilon(\mathbf{x}, \boldsymbol{\gamma}) := R\big(\nu(\cdot \mid \boldsymbol{\gamma}) \big\| \mu^\epsilon(\cdot \mid \mathbf{x})\big) \tag{6.4.6}$$

where ν is as in (6.3.2) and consider the following function analogous to (6.3.7) given in terms of the kernel μ^ϵ: For $\boldsymbol{\varphi} \in \mathfrak{P}$,

$$I^\epsilon(\boldsymbol{\varphi}) := \begin{cases} \int_0^T L^\epsilon(\mathbf{x}^{\boldsymbol{\varphi}}(t), \dot{\boldsymbol{\varphi}}(t))dt & \text{when } \boldsymbol{\varphi} \in \mathfrak{P}_I, \\ \infty & \text{when } \boldsymbol{\varphi} \notin \mathfrak{P}_I. \end{cases} \tag{6.4.7}$$

where $\mathbf{x}^{\boldsymbol{\varphi}}(\cdot) = (a^{\boldsymbol{\varphi}}(\cdot), \boldsymbol{\varphi}(\cdot))$ and $a^{\boldsymbol{\varphi}}(\cdot)$ is as in (6.3.6). The following conjecture is the main ingredient in the proof of Theorem 6.3.2.

Theorem 6.4.1. For every $\epsilon > 0$, the function I^ϵ in (6.4.7) is a rate function on $\mathfrak{P}$ and the sequence $\{\mathbf{V}^{n,\epsilon}\}_{n\in\mathbb{N}}$ satisfies a large deviation principle upper bound in $\mathfrak{P}$ with rate function I^ϵ.

Partial proofs of Theorem 6.4.1 is given at the end of this section. We now prove Theorem 6.3.2.

6.4.2 Proof of Theorem 6.3.2

Recall that $\mathfrak{P} = C([0,T] : \mathbb{R}^\infty)$ is a Polish space that can be metrized, for example using the following metric: for $\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}} \in \mathfrak{P}$ with $\boldsymbol{\varphi}(\cdot) = (\varphi_k(\cdot) : k \in \mathbb{N})$ and $\tilde{\boldsymbol{\varphi}} = (\tilde{\varphi}_k(\cdot) : k \in \mathbb{N})$

$$\mathbf{d}_{\mathfrak{P}}(\boldsymbol{\varphi}, \tilde{\boldsymbol{\varphi}}) := \sum_{k=1}^{\infty} \frac{\|\varphi_k - \tilde{\varphi}_k\|_\infty \wedge 1}{2^k}, \tag{6.4.8}$$

where $\|\cdot\|_\infty$ denotes the usual sup-norm on $C([0,T] : \mathbb{R})$. The following result shows that $\mathbf{V}^n$ is well approximated by $\mathbf{V}^{n,\epsilon}$.

Lemma 6.4.2. Let $h : \mathfrak{P} \to \mathbb{R}$ be bounded and Lipschitz with Lipschitz constant $C_h < \infty$. Then for all $n \geq 1$ and $\epsilon > 0$

$$\left| \frac{1}{n} \log \mathbb{E} \exp\left[-nh(\mathbf{V}^n)\right] - \frac{1}{n} \log \mathbb{E} \exp\left[-nh(\mathbf{V}^{n,\epsilon})\right] \right| \leq 2C_h\epsilon.$$

Proof: Without loss of generality we assume that the sequences $\{\mathbf{X}(j)\}_{j\in\mathbb{N}_0}$ and $\{\mathbf{X}^\epsilon(j)\}_{j\in\mathbb{N}_0}$ are constructed simultaneously using the same noise sequence $\{\xi(j)\}$ until the first time instant j^* when $\frac{1}{n}\mathbf{X}(j^*) \in \Lambda_\epsilon$. Denoting $\tau = j^*/n$ we see that $\mathbf{V}^n(t) = \mathbf{V}^{n,\epsilon}(t)$ for all $t \in [0,\tau]$ and $\sup_k\{V_k^n(t) \vee V_k^{n,\epsilon}(t)\} \leq \epsilon$ for all $t > \tau$. Thus, for all $k \geq 1$,

$$\sup_{t\in[0,T]} |V_k^n(t) - V_k^{n,\epsilon}(t)| = \sup_{t\in[\tau,T]} |V_k^n(t) - V_k^{n,\epsilon}(t)| \leq 2\epsilon.$$

Consequently $\mathbf{d}_{\mathfrak{P}}(\mathbf{V}^n, \mathbf{V}^{n,\epsilon}) \leq 2\epsilon$ and therefore $|h(\mathbf{V}^n) - h(\mathbf{V}^{n,\epsilon})| \leq 2C_h\epsilon$. The result follows. ■

Our next result shows that for every h as in Lemma 6.4.2, $\inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\}$ is well approximated by $\inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\}$.

Proposition 6.4.3. Let $h : \mathfrak{P} \to \mathbb{R}$ be bounded and Lipschitz with Lipschitz constant $C_h < \infty$. Then, for all $\epsilon > 0$,

$$\left| \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} - \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} \right| \leq 2C_h\epsilon.$$

The following proposition says that I is a rate function.

Proposition 6.4.4. For every $M > 0$, $S_M = \{\mathbf{v} \in \mathfrak{P} : I(\mathbf{v}) \le M\}$ is a compact subset of $\mathfrak{P}$.

Proof of Propositions 6.4.3 and 6.4.4 will be given in Sections 6.4.2.1 and 6.4.2.2 respectively. First we complete the proof of Theorem 6.3.2. **Proof of Theorem 6.3.2:** Note that Proposition 6.4.4 shows that I is a rate function. It then suffices to show(cf. Corollary 1.2.5 [14]) that for every bounded and Lipschitz function $h : \mathfrak{P} \to \mathbb{R}$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\{\exp[-nh(\mathbf{V}^n)]\} = -\inf_{\mathbf{v}\in\mathfrak{P}} \{I(\mathbf{v}) + h(\mathbf{v})\}.$$

From Theorem 6.4.1 the above equality holds with $\mathbf{V}^n$ replaced by $\mathbf{V}^{n,\epsilon}$ and I replaced by I^ϵ. The result now follows on combining this observation with Lemma 6.4.2 and Proposition 6.4.3. ■

6.4.2.1 Proof of Proposition 6.4.3.

We start with two auxiliary results needed for the proof. Proofs of these results are deferred to Section 6.4.3.

Lemma 6.4.5. Suppose $\boldsymbol{\varphi}(\cdot) = (\varphi_k(\cdot) : k \in \mathbb{N}) \in \mathfrak{P}$ is such that φ_k is nonnegative and nonincreasing for each k. Fix $a(0) \ge 0$ and define for $t > 0$,

$$a(t) = a(t; \boldsymbol{\varphi}, a(0)) := \Gamma(b)(t), \tag{6.4.9}$$

where

$$b(t) = b(t; \boldsymbol{\varphi}, a(0)) := a(0) - 2t + \sum_{k=1}^{\infty} k\varphi_k(0) - \sum_{k=1}^{\infty} k\varphi_k(t). \tag{6.4.10}$$

Then $t \mapsto \sum_{k=1}^{\infty} k\varphi_k(t) + a(t)$ is nonincreasing on $[0, T]$.

The next result deals with properties of an associated system of differential equations. Fix constants $a_0 \geq 0$ and a $\mathbf{v}_0 = \{v_{k,0}\}_{k\geq 1} \in \mathfrak{D}_{\text{deg}}$. Consider the sequence of $\mathbf{v}(\cdot) = \{v_k(\cdot)\}_{k\geq 1}$ that solves the system of integral equations

$$v_k(t) = v_{k,0} - \int_0^t \frac{k v_k(s)}{\sum_{l=1}^{\infty} l v_l(s) + a(s)} ds, \text{ for } k \geq 1. \tag{6.4.11}$$

where $a(t) = a(t; \mathbf{v}, a_0)$ is as in (6.4.9) (with $(a(0), \boldsymbol{\varphi})$ replaced by $(a_0, \mathbf{v})$).

Theorem 6.4.6. There exists a collection of functions $\mathbf{v} = \{v_k\}_{k\geq 1} \subset C([0,T] : \mathbb{R}^\infty)$ such that writing $\tau := \inf\{t \leq T : \sum_{l=1}^{\infty} l v_l(t) + a(t) = 0\}$, the following conditions are satisfied:

(i) The integral equations (6.4.11) hold on $[0, \tau)$.

(ii) For $t \geq \tau$, $a(t) = 0$ and $v_k(t) = 0$ for all $k \geq 1$.

(iii) The functions $t \mapsto \sum_{l=1}^{\infty} l v_l(t) + a(t)$ and $t \mapsto v_k(t)$, $k \geq 1$ are nonincreasing on $[0, T]$.

(iv) $\tau \leq \sum_{l=1}^{\infty} v_{l,0} + \frac{1}{2}(\sum_{l=1}^{\infty} l v_{l,0} + a_0)$.

Let us now proceed with the proof of the Proposition 6.4.3. We first show

$$\inf_{\boldsymbol{\varphi} \in \mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} \leq \inf_{\boldsymbol{\varphi} \in \mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} + 2C_h \epsilon. \tag{6.4.12}$$

For any fixed $\sigma > 0$, we can find $\boldsymbol{\varphi}^* \in \mathfrak{P}_I$ such that

$$I(\boldsymbol{\varphi}^*) + h(\boldsymbol{\varphi}^*) \leq \inf_{\boldsymbol{\varphi} \in \mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} + \sigma < \infty. \tag{6.4.13}$$

Let $\tau^\epsilon := \inf\left\{t \in [0,T] : \sum_{k=1}^{\infty} k \varphi_k^*(t) + a^{\boldsymbol{\varphi}^*}(t) \leq \epsilon\right\}$. Given a non-negative sequence $\boldsymbol{\varsigma} = (\varsigma_k)_{k\geq 1}$, denote by $\boldsymbol{\vartheta}^\epsilon(\boldsymbol{\varsigma}) = (\vartheta_k^\epsilon(\varsigma_k; t), t \geq 0)_{k\geq 1}$ the unique solution of

$$\dot{\vartheta}_k^\epsilon(\varsigma_k; t) = -\frac{k \vartheta_k^\epsilon(\varsigma_k; t)}{\epsilon}, \;\; t \geq 0, \;\; \vartheta_k^\epsilon(\varsigma_k; 0) = \varsigma_k, \;\; k \geq 1. \tag{6.4.14}$$

Let $\boldsymbol{\psi} = \boldsymbol{\vartheta}(\boldsymbol{\varphi}(\tau^\epsilon))$ and define $\tilde{\boldsymbol{\varphi}}^*(t) := \boldsymbol{\varphi}^*(t)$ for $t \in [0, \tau^\epsilon]$ and $\tilde{\boldsymbol{\varphi}}^*(t) := \boldsymbol{\psi}(t - \tau^\epsilon)$ for $t \in [\tau^\epsilon, T]$. Note that by construction $\tilde{\boldsymbol{\varphi}}^* \in \mathfrak{P}_I$. Also, by Lemma 6.4.5, $\sum_{k=1}^\infty k\tilde{\varphi}_k^*(t) + a^{\tilde{\boldsymbol{\varphi}}^*}(t) \leq \epsilon$ for $t \geq \tau^\epsilon$. From (6.3.2), (6.4.1) and (6.4.14) it follows that for all $t \geq \tau^\epsilon$, $\nu(\cdot \mid \dot{\tilde{\boldsymbol{\varphi}}}^*(t)) = \mu^\epsilon(\cdot \mid \mathbf{x}^{\tilde{\boldsymbol{\varphi}}^*}(t))$ and consequently from (6.4.6) we have

$$\int_{\tau^\epsilon}^T L^\epsilon(\mathbf{x}^{\tilde{\boldsymbol{\varphi}}^*}(t), \dot{\tilde{\boldsymbol{\varphi}}}^*(t))dt = 0.$$

Thus,

$$I^\epsilon(\tilde{\boldsymbol{\varphi}}^*) = \int_0^{\tau^\epsilon} L^\epsilon(\mathbf{x}^{\tilde{\boldsymbol{\varphi}}^*}(t), \dot{\tilde{\boldsymbol{\varphi}}}^*(t))dt = \int_0^{\tau^\epsilon} L(\mathbf{x}^{\boldsymbol{\varphi}^*(t)}, \dot{\boldsymbol{\varphi}}^*(t))dt \leq I(\boldsymbol{\varphi}^*). \tag{6.4.15}$$

Noting that $\sup_{0\leq t\leq T} |\tilde{\varphi}_k^*(t) - \varphi_k^*(t)| \leq 2\epsilon$ for all $k \geq 1$, we see that

$$|h(\tilde{\boldsymbol{\varphi}}^*) - h(\boldsymbol{\varphi}^*)| \leq 2C_h\epsilon. \tag{6.4.16}$$

Thus

$$I^\epsilon(\tilde{\boldsymbol{\varphi}}^*) + h(\tilde{\boldsymbol{\varphi}}^*) \leq I(\boldsymbol{\varphi}^*) + h(\boldsymbol{\varphi}^*) + 2C_h\epsilon \leq \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} + \sigma + 2C_h\epsilon.$$

Letting $\sigma \to 0$ proves (6.4.12). Let us now prove the reverse inequality, namely

$$\inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} \leq \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} + 2C_h\epsilon. \tag{6.4.17}$$

For $\sigma > 0$ fix $\boldsymbol{\varphi}^*$ now satisfying

$$I^\varepsilon(\boldsymbol{\varphi}^*) + h(\boldsymbol{\varphi}^*) \leq \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\} + \sigma.$$

Define τ^ϵ as before and consider now the sequence of functions $\mathbf{v} := (v_k(t) : t \geq 0)_{k\geq 1}$ as in Theorem 6.4.6 with $v_{k,0} = \varphi_k^*(\tau^\epsilon)$ and $a_0 = a^{\boldsymbol{\varphi}^*}(\tau^\epsilon)$. Define $\tilde{\boldsymbol{\varphi}}^*$ as in the first part of the proof by replacing $\boldsymbol{\psi}$ with $\mathbf{v}$. Note that, by construction, $L(\mathbf{x}^{\tilde{\boldsymbol{\varphi}}^*}(t), \dot{\tilde{\boldsymbol{\varphi}}}^*(t)) = 0$ for all $t \geq \tau^\epsilon$. So

$$I(\tilde{\boldsymbol{\varphi}}^*) = \int_0^{\tau^\epsilon} L(\boldsymbol{x}^{\tilde{\boldsymbol{\varphi}}^*}(t), \dot{\tilde{\boldsymbol{\varphi}}}^*(t))dt = \int_0^{\tau^\epsilon} L^\epsilon(\boldsymbol{x}^{\boldsymbol{\varphi}^*}(t), \dot{\boldsymbol{\varphi}}^*(t))dt \leq I^\epsilon(\boldsymbol{\varphi}^*).$$

Also, as before it follows that (6.4.16) holds. The inequality in (6.4.17) now follows on combining the last two observations as in the proof of (6.4.12). ■

6.4.2.2 Proof of Proposition 6.4.4.

From the Theorem 6.4.1 we have that for every $\epsilon > 0$, I^ϵ is a rate function. We will now use this fact to show that I is a rate function as well. Fix $M > 0$ and suppose $\{\boldsymbol{\varphi}_n\} \subset S_M$ is a sequence such that $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$. Note that we must have that $\boldsymbol{\varphi} \in \mathfrak{P}_I$. We need to show that $I(\boldsymbol{\varphi}) \le M$. We argue via contradiction. Suppose that, for some $\delta > 0$, $I(\boldsymbol{\varphi}) > M + \delta$. Let $\tau = \inf\{t \ge 0 : \sum_k k\boldsymbol{\varphi}_k(t) + a^{\boldsymbol{\varphi}}(t) = 0\}$. From the definition of I we see that there is a $\delta_0 > 0$ and $t_0 \in (0, \tau)$ such that

$$\int_{[0,t_0]} L(\mathbf{x}^{\boldsymbol{\varphi}}(t), \dot{\boldsymbol{\varphi}}(t))dt > M + \delta_0. \tag{6.4.18}$$

Let $M_{\boldsymbol{\varphi}}(t) = \sum_k k\varphi_k(t) + a^{\boldsymbol{\varphi}}(t)$ and let $\gamma = \inf_{t\in[0,t_0]} M_{\boldsymbol{\varphi}}(t)$. Note that $\gamma > 0$. From the definition of L^ϵ we see that

$$\int_{[0,t_0]} L^{\gamma/2}(\mathbf{x}^{\boldsymbol{\varphi}}(t), \dot{\boldsymbol{\varphi}}(t))dt = \int_{[0,t_0]} L(\mathbf{x}^{\boldsymbol{\varphi}}(t), \dot{\boldsymbol{\varphi}}(t))dt > M + \delta_0. \tag{6.4.19}$$

Since $\boldsymbol{\varphi}_n \to \boldsymbol{\varphi}$, we can find $n_0 \in \mathbb{N}$ such that

$$M_{\boldsymbol{\varphi}_n}(t) \ge \gamma/2, \quad \text{for all } t \in [0, t_0],\ n \ge n_0. \tag{6.4.20}$$

Let $\tau_n = \inf\{t \in [0, T] : M_{\boldsymbol{\varphi}_n}(t) \le \gamma/2\}$ and define $\boldsymbol{\varphi}_n^\gamma \in \mathfrak{P}_I$ as

$$\boldsymbol{\varphi}_n^\gamma(t) = \boldsymbol{\varphi}_n(t)\mathbf{1}_{[0,\tau_n]}(t) + \boldsymbol{\vartheta}^{\gamma/2}(\boldsymbol{\varphi}_n(\tau_n); t - \tau_n)\mathbf{1}_{(\tau_n,T]}(t),\ t \in [0, T],$$

where $\boldsymbol{\vartheta}^{\gamma/2}$ is as in (6.4.14) (with ϵ replaced by $\gamma/2$). As in (6.4.15) we have

$$I^{\gamma/2}(\boldsymbol{\varphi}_n^\gamma) \le I(\boldsymbol{\varphi}_n) \le M. \tag{6.4.21}$$

From Theorem 6.4.6(iv) $\sup_n \tau_n \le T$, and so we can find a subsequence (labeled again as $\{n\}$) such that $\tau_n \to \tau$ for some $\tau \in [0, T]$. Also, since $\varphi_{n,k}$ is nonincreasing for every k, we have that

$$\varphi_{n,k}^\gamma(t) = (\varphi_{n,k}(t) - \varphi_{n,k}(\tau_n))^+ + \vartheta_k^{\gamma/2}(\varphi_{n,k}(\tau_n); (t - \tau_n)^+),\ t \in [0, T].$$

By uniform convergence of $\varphi_{n,k}$ to φ_k and continuity of $(t,\zeta) \mapsto \boldsymbol{\vartheta}^{\gamma/2}(\zeta;t)$ we have that for every $k \geq 1$, $\varphi_{n,k}^{\gamma}$ converges uniformly on $[0,T]$, as $n \to \infty$, to

$$\varphi_k^{\gamma}(t) = (\varphi_k(t) - \varphi_k(\tau))^+ + \vartheta_k^{\gamma/2}(\varphi_k(\tau),(t-\tau)^+),\ t \in [0,T]$$

and thus $\boldsymbol{\varphi}_n^{\gamma} \to \boldsymbol{\varphi}^{\gamma}$. Combining this with (6.4.21) and recalling that $I^{\gamma/2}$ is a rate function we now have that $I^{\gamma/2}(\boldsymbol{\varphi}^{\gamma}) \leq M$. Also note that $\boldsymbol{\varphi}_n^{\gamma}(t) = \boldsymbol{\varphi}_n(t)$ for all $t \leq \tau_n$ and from (6.4.20) $\tau_n \geq t_0$. Therefore $\boldsymbol{\varphi}^{\gamma}(t) = \boldsymbol{\varphi}(t)$ for all $t \leq t_0$. Thus, from (6.4.19)

$$M + \delta_0 < \int_{[0,t_0]} L^{\gamma/2}(\mathbf{x}^{\boldsymbol{\varphi}}(t), \dot{\boldsymbol{\varphi}}(t))dt = \int_{[0,t_0]} L^{\gamma/2}(\mathbf{x}^{\boldsymbol{\varphi}^{\gamma}}(t), \dot{\boldsymbol{\varphi}}^{\gamma}(t))dt \leq I^{\gamma/2}(\boldsymbol{\varphi}^{\gamma}) \leq M,$$

which is a contradiction. Thus we must have that $I(\boldsymbol{\varphi}) \leq M$. The result follows. ■

6.4.3 Proofs of Lemma 6.4.5 and Theorem 6.4.6

Let us now prove the auxiliary results used in the proof of Proposition 6.4.3.

6.4.3.1 Proof of Lemma 6.4.5:

Note that for $0 \leq t_1 \leq t_2 \leq T$

$$\begin{aligned} a(t_2) =& a(t_1) - 2(t_2 - t_1) - \sum_{k=1}^{\infty} k(\varphi_k(t_2) - \varphi_k(t_1)) \\ & - \inf_{t_1 \leq s \leq t_2} \left(a(t_1) - 2(s - t_1) - \sum_{k=1}^{\infty} k\left(\varphi_k(s) - \varphi_k(t_1)\right)\right) \wedge 0. \end{aligned}$$

Consider first the case when the infimum in the second line of the display is zero. In that case

$$a(t_2) + \sum_{k=1}^{\infty} k\varphi_k(t_2) = a(t_1) + \sum_{k=1}^{\infty} k\varphi_k(t_1) - 2(t_2 - t_1) \leq a(t_1) + \sum_{k=1}^{\infty} k\varphi_k(t_1).$$

Next consider the case when the infimum is negative. Then

$$\begin{aligned} a(t_2) + \sum_{k=1}^{\infty} k\varphi_k(t_2) =& -2t_2 + \sup_{t_1 \leq s \leq t_2} \left(2s + \sum_{k=1}^{\infty} k\varphi_k(s)\right) \\ \leq & -2t_2 + 2t_2 + \sum_{k=1}^{\infty} k\varphi_k(t_1) \leq a(t_1) + \sum_{k=1}^{\infty} k\varphi_k(t_1). \end{aligned}$$

This completes the proof. ■

6.4.3.2 Proof of Theorem 6.4.6:

Note that in this theorem, infinitely many functions are related via integral equations. We first consider the following approximation involving finitely many functions. For $m \geq 1$, let $[m] = \{1, 2, \ldots, m\}$. Recall a_0 and $\mathbf{v}_0 = (v_{k,0})_{k=1}^{\infty}$ introduced above Theorem 6.4.6. Consider the integral equations

$$v_k(t) = v_{k,0} - \int_0^t \frac{k v_k(s)}{\sum_{l=1}^m l v_l(s) + a^{[m]}(s)} ds, \qquad k \in [m],\ t > 0 \tag{6.4.22}$$

where analogous to (6.4.9)–(6.4.10),

$$b^{[m]}(t) = b^{[m]}(t; \mathbf{v}, a_0) := a_0 - 2t + \sum_{k=1}^m k v_{k,0} - \sum_{k=1}^m k v_k(t), \tag{6.4.23}$$

$$a^{[m]}(t) = a^{[m]}(t; \mathbf{v}, a_0) := \Gamma(b^{[m]})(t). \tag{6.4.24}$$

Lemma 6.4.7. Fix $m \geq 1$. Then there exists $c = c(m) \in (0, \infty)$ such that for all $\delta \leq c(m) \min\{d_0, d_0^2\}$, where $d_0 = a_0 + \sum_{k=1}^m k v_{k,0}$, there is a unique $\mathbf{v} = (v_k)_{k=1}^m \in C([0, \delta] : \mathbb{R}^m)$ that solves (6.4.22)–(6.4.24).

Proof: We use the contraction mapping theorem. Assume $\delta < d_0/(m+2)$. Additional restrictions on δ will be introduced later in the proof. Define $\mathfrak{P}_{m,\delta} \subset C([0, \delta] : \mathbb{R}^m)$ as

$$\mathfrak{P}_{m,\delta} := \left\{\mathbf{v} \in C([0, \delta] : \mathbb{R}^m) : v_k(0) = v_{k,0},\ v_k(\cdot) \geq 0 \text{ and nonincreasing } \forall\, k \in [m]\right\}. \tag{6.4.25}$$

Note that $\mathfrak{P}_{m,\delta}$ is a closed subset of $C([0, \delta] : \mathbb{R}^m)$. Write

$$D(t) := \sum_{l=1}^m l v_l(t) + a^{[m]}(t).$$

Note that for any $\mathbf{v} \in \mathfrak{P}_{m,\delta}$,

$$D(t) = \sum_{l=1}^{m} l v_l(t) + b^{[m]}(t) - \inf_{s\in[0,t]} b^{[m]}(s) \wedge 0 = d_0 - 2t - \inf_{s\in[0,t]} b^{[m]}(s) \wedge 0 \geq d_0 - 2t. \tag{6.4.26}$$

Define the map $\mathcal{T} : \mathfrak{P}_{m,\delta} \to C([0,\delta] : \mathbb{R}^m)$

$$(\mathcal{T}\mathbf{v})_k(t) := v_{k,0} - \int_0^t \frac{k v_k(s)}{\sum_{l=1}^m l v_l(s) + a^{[m]}(s; \mathbf{v}, a_0)} ds, \qquad k \in [m],\ t \in [0,\delta].$$

We now claim that $\mathcal{T}$ maps $\mathfrak{P}_{m,\delta}$ to itself. It is immediate that $(\mathcal{T}\mathbf{v})_k(\cdot)$ is nonincreasing. Also, using the inequalities $v_k(t) \leq v_{k,0}$, $k \in [m]$ and (6.4.26),

$$(\mathcal{T}\mathbf{v})_k(\delta) \geq v_{k,0} - \int_0^\delta \frac{m v_{k,0}}{d_0 - 2\delta} ds = \left(1 - \frac{m\delta}{d_0 - 2\delta}\right) v_{k,0} \geq 0,$$

where the last inequality uses the fact that $d_0 - (m+2)\delta > 0$. This proves the claim.

Next we show that $\mathcal{T}$ is a contraction for δ small. Define

$$\|\mathbf{v}\|_{m,\delta} := \sup_{1\leq k\leq m} \sup_{t\in[0,\delta]} |v_k(t)|,\ \mathbf{v} \in \mathfrak{P}_{m,\delta}.$$

Consider $\mathbf{v}, \tilde{\mathbf{v}} \in \mathfrak{P}_{m,\delta}$. Let $\tilde{a}^{[m]}(t) = a^{[m]}(t; \tilde{\mathbf{v}}, a_0)$, $\tilde{b}^{[m]}(t) = b^{[m]}(t; \tilde{\mathbf{v}}, a_0)$ and $\tilde{D}(t) = \sum_{l=1}^m l\tilde{v}_l(t) + \tilde{a}^{[m]}(t)$. From the second equality in (6.4.26), for $t \in [0,\delta]$

$$\begin{aligned}
|D(t) - \tilde{D}(t)| &= \left| \inf_{s\in[0,t]} b^{[m]}(s) \wedge 0 - \inf_{s\in[0,t]} \tilde{b}^{[m]}(s) \wedge 0 \right| \\
&\leq \sup_{s\in[0,t]} |b^{[m]}(s) - \tilde{b}^{[m]}(s)| \leq \sum_{k=1}^m k \sup_{s\in[0,\delta]} |v_k(s) - \tilde{v}_k(s)| \leq m^2 \|\mathbf{v} - \tilde{\mathbf{v}}\|_{m,\delta}.
\end{aligned}$$

Therefore, for $k \in \mathbb{N}, m$ and $t \in [0,\delta]$,

$$\begin{aligned}
|(\mathcal{T}\mathbf{v})_k(t) - (\mathcal{T}\tilde{\mathbf{v}})_k(t)| &\leq \int_0^t \left| \frac{k v_k(s)}{D(s)} - \frac{k\tilde{v}_k(s)}{\tilde{D}(s)} \right| ds \\
&\leq \int_0^t \left(\frac{k v_k(s)}{D(s)\tilde{D}(s)} \left| D(s) - \tilde{D}(s) \right| + \frac{k}{\tilde{D}(s)} |v_k(s) - \tilde{v}_k(s)| \right) ds \\
&\leq \delta \left(\frac{m v_{k,0}}{(d_0 - 2\delta)^2} \cdot m^2 \|\mathbf{v} - \tilde{\mathbf{v}}\|_{m,\delta} + \frac{m}{d_0 - 2\delta} \|\mathbf{v} - \tilde{\mathbf{v}}\|_{m,\delta} \right) \\
&= \delta \left(\frac{m^3 v_{k,0}}{(d_0 - 2\delta)^2} + \frac{m}{d_0 - 2\delta} \right) \|\mathbf{v} - \tilde{\mathbf{v}}\|_{m,\delta}.
\end{aligned}$$

Taking supremum over $t \in [0,\delta]$ and $k \in [m]$,

$$\|\mathcal{T}\mathbf{v} - \mathcal{T}\tilde{\mathbf{v}}\|_{m,\delta} \le \delta \left(\frac{m^3 \max_{1\le k\le m} v_{k,0}}{(d_0 - 2\delta)^2} + \frac{m}{d_0 - 2\delta} \right) \|\mathbf{v} - \tilde{\mathbf{v}}\|_{m,\delta}.$$

Noting that $\max_{1\le k\le m} v_{k,0} \le \nu^*$ we see that there is a $c(m) \in (0,\infty)$ such that for all $\delta \le c(m)\min\{d_0, d_0^2\}$, $\mathcal{T} : \mathfrak{P}_{m,\delta} \to \mathfrak{P}_{m,\delta}$ is a contraction. This completes the proof. ■

Using Lemma 6.4.7 we can now prove the following finite dimensional analogue of Theorem 6.4.6.

Theorem 6.4.8. Fix $m \ge 1$ and suppose $\sum_{k=1}^m k v_{k,0} + a_0 > 0$. Then there exists a unique $\mathbf{v} = (v_k)_{k=1}^m \in C([0,\infty),\mathbb{R}^m)$ with the following properties.

(i) $\tau^{[m]} := \inf\{t \ge 0 : \sum_{l=1}^m l v_l(t) + a^{[m]}(t) = 0\} > 0$.

(ii) $(v_k)_{k=1}^m$ satisfy the integral equations (6.4.22) on $[0,\tau^{[m]})$.

(iii) For all $t \ge \tau^{[m]}$, $v_k(t) = 0$ for all $k \in [m]$ and $a^{[m]}(t) = 0$.

(iv) The functions $t \mapsto v_k(t)$ for all $k \in [m]$ and $t \mapsto \sum_{l=1}^m l v_l(t) + a^{[m]}(t)$ are nonincreasing on $[0,\infty)$.

(v) $\tau^{[m]} \le \sum_{l=1}^m v_{l,0} + \frac{1}{2}(\sum_{l=1}^m l v_{l,0} + a_0)$.

Proof: Using Lemma 6.4.7 we can recursively construct a unique solution of (6.4.22) – (6.4.24) on $[0,\sigma_1], [0,\sigma_2], \cdots$ where $0 = \sigma_0 < \sigma_1 < \cdots$ In fact denoting $\alpha_i = \sum_{l=1}^m l v_l(\sigma_i) + a^{[m]}(\sigma_i)$, we can have $\{\sigma_i\}$ such that for all $i \ge 0$, $\alpha_i > 0$ and $\sigma_{i+1} - \sigma_i = c(m)\min\{d_i, d_i^2\}$. This in particular proves (i).

The recursive construction gives a unique solution of (6.4.22) – (6.4.24) on $[0,\sigma_\infty)$, where $\sigma_\infty = \lim_{j\to\infty}\sigma_j$. Furthermore, $t \mapsto v_k(t)$ is nonincreasing on $[0,\sigma_\infty)$ for all $k \in [m]$ and combining this fact with Lemma 6.4.5, $t \mapsto \sum_{l=1}^m l v_l(t) + a^{[m]}(t)$ is nonincreasing on $[0,\sigma_\infty)$ as well.

Parts (ii) -(iv) are now immediate if $\sigma_\infty = \infty$. Consider now the case when $\sigma_\infty < \infty$. Then, noting that

$$\infty > \sigma_\infty = \sum_{j=0}^{\infty}(\sigma_{j+1} - \sigma_j) = c(m)\sum_{j=1}^{\infty}\alpha_j,$$

we see that $\alpha_j \to 0$ as $j \to \infty$. Recalling that $t \mapsto D(t)$ is nonincreasing on $[0, \sigma_\infty)$ we now see that $v_k(t) \to 0$ for all $k \in [m]$ and $a^{[m]}(t) \to 0$, as $t \to \sigma_\infty$. This proves parts (ii)-(iv).

Finally, we prove (v). Note that if $\eta : [0, \infty) \to \mathbb{R}$ is an absolutely continuous function then $\zeta = \Gamma(\eta)$ satisfies

$$\zeta(t) = \eta(0) + \int_0^t \dot{\eta}(s)\mathbf{1}_{\{\zeta(s)>0\}}ds, \quad \text{for all } t \geq 0.$$

Thus for $t < \tau^{[m]}$

$$\begin{aligned} D(t) &= \sum_{l=1}^{m} l v_l(t) + a^{[m]}(t) = \sum_{l=1}^{m} l v_l(t) + a_0 - \int_0^t \left(\sum_{l=1}^{m} l\dot{v}_l(s) + 2\right)\mathbf{1}_{\left\{a^{[m]}(s)>0\right\}}ds \\ &\leq \sum_{l=1}^{m} l v_l(t) + a_0 - 2\int_0^t \mathbf{1}_{\left\{a^{[m]}(s)>0\right\}}ds - \int_0^t \sum_{l=1}^{m} l\dot{v}_l(s)ds \end{aligned}$$

where the second inequality uses the fact that $\dot{v}_k(t) \leq 0$ a.s. Taking limit as $t \to \tau^{[m]}$ we have

$$\int_0^{\tau^{[m]}} \mathbf{1}_{\left\{a^{[m]}(s)>0\right\}}ds \leq \frac{1}{2}d_0. \tag{6.4.27}$$

One the other hand, for $t \in [0, \tau^{[m]})$

$$\sum_{l=1}^{m} v_l(t) = \sum_{l=1}^{m} v_{l,0} - \int_0^t \frac{\sum_{l=1}^{m} l v_l(s)}{\sum_{l=1}^{m} l v_l(s) + a^{[m]}(s)}ds \leq \sum_{l=1}^{m} v_{l,0} - \int_0^t \mathbf{1}_{\left\{a^{[m]}(s)=0\right\}}ds.$$

Thus we have $\int_0^{\tau^{[m]}} \mathbf{1}_{\left\{a^{[m]}(s)=0\right\}}ds \leq \sum_{l=1}^{m} v_{l,0}$. Combining this with (6.4.27) we have (v). ■

Proof of Theorem 6.4.6: Since $\sum_{l=1}^{\infty} l v_{l,0} + a_0 > 0$, we have $\sum_{l=1}^{m} l v_{l,0} + a_0 > 0$ for all m large enough. Without loss of generality, we assume this for all $m \in \mathbb{N}$. For

$m \geq 1$, denote by $(\varphi_k^{[m]})_{k=1}^m$ the unique collection of functions obtained from Theorem 6.4.8. We set $\varphi_k^{[m]}(t) \equiv v_{k,0}$ for all $t \geq 0$, $k > m$. Noting that

$$|\varphi_k^{[m]}(t) - \varphi_k^{[m]}(s)| \leq |t-s| \text{ for all } m, k \in \mathbb{N} \text{ and } s, t \geq 0,$$

we see that for each $k \geq 1$, $\{\varphi_k^{[m]}, m \geq 1\}$ is relatively compact in $C([0,\infty) : \mathbb{R})$. Denote by v_k the limit, as $m \to \infty$, along a convergent subsequence. By a diagonalization argument we can take the same subsequence for all k and without loss of generality we assume that $\left\{\varphi_k^{[m]}\right\}$ converges to v_k as $m \to \infty$ for every k. Clearly $\mathbf{v} = (v_k)_{k\geq 1} \in \mathfrak{P}_I$. We claim that $\mathbf{v}$ satisfies all the properties stated in Theorem 6.4.6. Let τ be as in the statement of the theorem. From Theorem 6.4.8(v) it follows that $\tau < T$. Also, since $\sup_{t\in[0,T]} \varphi_k^{[m]}(t) \leq v_{k,0}$ and $\sum k v_{k,0} \leq \nu^* < \infty$, we have

$$\sum_{l=1}^{\infty} l\varphi_l^{[m]}(t) \to \sum_{l=1}^{\infty} l v_l(t) \text{ uniformly on } [0,T] \text{ as } m \to \infty.$$

Let $a^{[m]}(t) := a^{[m]}(t; \boldsymbol{\varphi}^{[m]}, a_0)$, $a(t) := a(t; \mathbf{v}, a_0)$, $b^{[m]}(t) := b^{[m]}(t; \boldsymbol{\varphi}^{[m]}, a_0)$, and $b(t) := b(t; \mathbf{v}, a_0)$ (cf. (6.4.9) - (6.4.10) and (6.4.23) - (6.4.24)). By the continuity of Skorokhod map Γ, we have

$$a^{[m]}(t) \to a(t) \text{ and } b^{[m]}(t) \to b(t) \text{ uniformly on } [0,T] \text{ as } m \to \infty.$$

Note that for each fixed $t < \tau$, there is a $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$, we have $\varphi_l^{[m]}(s) + a^{[m]}(s) > 0$ for all $s \in [0,t]$. Thus, for each $k \in \mathbb{N}$ and $m \geq m_0$,

$$\varphi_k^{[m]}(s) = v_{k,0} - \int_0^s \frac{k\varphi_k^{[m]}(u)}{\sum_{l=1}^{\infty} l\varphi_l^{[m]}(u) + a^{[m]}(u)} du, \ s \in [0,t].$$

Sending $m \to \infty$ in the above equation, we see that $\mathbf{v}$ solves (6.4.11) on $[0,t]$ for all $t < \tau$. We extend $\mathbf{v}$ to $[\tau, T]$ by continuity. This proves (i)-(ii). The other two parts follow from analogous properties of $\boldsymbol{\varphi}^{[m]}$. ■

6.4.4 From Theorem 6.4.1 to the Laplace principle

In order to prove the theorem it suffices to show:

(a) For all $\epsilon > 0$, I^ϵ is a rate function.

(b) **(Upper Bound)** For all $\epsilon > 0$ and bounded and Lipschitz functions $h : \mathfrak{P} \to \mathbb{R}$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\{\exp[-nh(\mathbf{V}^{n,\epsilon})]\} \leq -\inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\}.$$

Proofs of (a) and (b) are given in Sections 6.6 and 6.5 respectively.

Due to the singularity of the transition kernel of the process $\mathbf{V}^{n,\epsilon}(\cdot)$, we are unable to prove the Laplace principle lower bound:

Conjecture 6.4.9. for all $\epsilon > 0$ and bounded and Lipschitz functions $h : \mathfrak{P} \to \mathbb{R}$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{E}\{\exp[-nh(\mathbf{V}^{n,\epsilon})]\} \geq -\inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I^\epsilon(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\}.$$

We leave the proof of the lower bound as an open problem.

Note: *For rest of this chapter $\epsilon > 0$ will be fixed and therefore we suppress it from notation, in particular we write $\mathbf{V}^{n,\epsilon}, \mu^\epsilon$ as $\mathbf{V}^n, \mu$, respectively.*

6.5 Laplace principle upper bound

In this section we will prove the Laplace principle upper bound, namely item (b) of Section 6.4.4. Write

$$W^n := -\frac{1}{n} \log \mathbb{E}\{\exp[-nh(\mathbf{V}^n)]\}.$$

Then the goal is to show

$$\liminf_{n\to\infty} W^n \geq \inf_{\boldsymbol{\varphi}\in\mathfrak{P}} \{I(\boldsymbol{\varphi}) + h(\boldsymbol{\varphi})\}. \tag{6.5.1}$$

We begin with a variational characterization that will play a key role in proofs of both the upper and the lower bounds.

6.5.1 The variational formula

Consider a sequence of random variables

$$\left\{\bar{\Xi}(j) : j \geq 0\right\} := \left\{(\bar{A}(j), \bar{\mathbf{V}}(j), \bar{\mathbf{X}}(j), \bar{B}(j)) : j \geq 0\right\}$$

with values in $(\mathbb{R}_+ \times \mathbb{R}_+^\infty \times (\mathbb{R}_+ \times \mathbb{R}_+^\infty) \times \mathbb{R})$ and a sequence $\{\nu_j\}_{j\geq 0}$ of random probability measures on $\mathbb{N}_0$, that is constructed recursively on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as follows:

1. Define $\bar{A}(0) = 0$, $\bar{B}(0) = 0$, $\bar{V}_k(0) = n_k$, $k \in \mathbb{N}$, $\bar{\mathbf{X}}(0) = (0, \bar{\mathbf{V}}(0))$. Let $\bar{\mathcal{F}}_0 = \{\emptyset, \Omega\}$ be the trivial σ - field.
2. For $j \geq 0$, having defined $\{\bar{A}(m), \bar{\mathbf{V}}(m), \bar{\mathbf{X}}(m), \bar{B}(m)\}_{m=0}^{j}$ and $\bar{\mathcal{F}}_j$, select a random $\bar{\mathcal{F}}_j$ measurable probability measure ν_j on $\mathbb{N}_0$ and let $\bar{\xi}(j+1)$ be a $\mathbb{N}_0$ valued random variable such that

$$\mathbb{P}\left\{\bar{\xi}(j+1) = k \mid \bar{\mathcal{F}}_j\right\} = \nu_j(k),\ k \in \mathbb{N}_0. \tag{6.5.2}$$

3. Let

$$\bar{V}_k(j+1) = \bar{V}_k(j) - \mathbf{1}_{\left\{\bar{\xi}(j+1)=k\right\}},\ k \in \mathbb{N}, \tag{6.5.3}$$

and

$$\bar{B}(j+1) = \bar{B}(j) + (\bar{\xi}(j+1) - 2), \qquad \bar{A}(j+1) = \bar{A}(j) + (\bar{\xi}(j+1) - 2) + 2 \cdot \mathbf{1}_{\left\{\bar{A}(j) \leq 0\right\}}. \tag{6.5.4}$$

Let $\bar{\mathbf{V}}(j) = (\bar{V}_k(j))_{k\geq 1}$ and set $\bar{\mathbf{X}}(j) = ((\bar{A}(j) - 1)^+, \bar{\mathbf{V}}(j))$.

Define the continuous time process $\bar{\mathbf{V}}^n$ through (6.2.2) by replacing $\{\mathbf{V}(j)\}$ on the right side of the equation with $\{\bar{\mathbf{V}}(j)\}$.

In the variational representation below, we will take infimum over all sequences $\{\nu_j\}_{j\in\mathbb{N}_0}$ of random probability measures of the form above and we refer to any such sequence as an admissible control and the corresponding sequence $\{\bar{\Xi}(j)\}$ as a controlled sequence.

The following variational characterization is proved exactly as Theorem 4.4.1 in [14].

Lemma 6.5.1. Let $h : \mathfrak{P} \to \mathbb{R}$ be bounded and measurable. Then for all $n \geq 1$,

$$W^n = \inf_{\{\nu_j\}} \mathbb{E} \left\{ \frac{1}{n} \sum_{j=1}^{\lfloor nT \rfloor} R\big(\nu_j(\cdot) \big\| \mu(\cdot \mid \frac{1}{n}\bar{\mathbf{X}}(j))\big) + h(\bar{\mathbf{V}}^n) \right\},$$

where the inf is taken over all admissible controls.

Now fix $\varepsilon \in (0,1)$ and $n \geq 1$. Then we can find an admissible control sequence $\{\nu_j^n\}_{j\geq 0}$ such that

$$W^n + \varepsilon \geq \mathbb{E} \left\{ \frac{1}{n} \sum_{j=1}^{\lfloor nT \rfloor} R\big(\nu_j^n(\cdot) \big\| \mu(\cdot \mid \frac{1}{n}\bar{\mathbf{X}}(j))\big) + h(\bar{\mathbf{V}}^n) \right\}. \tag{6.5.5}$$

Consider the collection $\{\nu^n(\cdot \mid t)\}_{t\in[0,T]}$ of random probability measures on $\mathbb{N}_0$ defined as $\nu^n(\cdot \mid t) := \nu^n_{\lfloor tn \rfloor}(\cdot)$, $t \in [0,T]$. Define a random probability measure $\bar{\nu}^n(\cdot)$ on $\mathbb{N}_0 \times [0,T]$ as

$$\bar{\nu}^n(A \times B) = \frac{1}{T} \int_B \nu^n(A|t)dt, \ A \in \mathcal{B}(\mathbb{N}_0), \ B \in \mathcal{B}([0,T]). \tag{6.5.6}$$

Define the stochastic process $\{\tilde{\mathbf{X}}^n(t)\}_{t\in[0,T]}$ with sample paths in $D([0,T] : \mathbb{R}_+ \times \mathbb{R}_+^\infty)$ as $\tilde{\mathbf{X}}^n(t) = \frac{1}{n}\bar{\mathbf{X}}(\lfloor nt \rfloor)$, $0 \leq t \leq T$. Then (6.5.5) can be rewritten as

$$W^n + \varepsilon \geq \mathbb{E} \left\{ \int_{[0,T]} R\big(\nu^n(\cdot \mid t) \big\| \mu(\cdot \mid \tilde{\mathbf{X}}^n(t))\big)dt + h(\bar{\mathbf{V}}^n) \right\}. \tag{6.5.7}$$

The rest of this section is organized as follows.

(i) In Section 6.5.2, we show $\{\bar{\nu}^n : n \geq 1\}$ is tight and has a certain uniform integrability property.

(ii) Using this result, in Section 6.5.3 we will argue that $\{\mathbf{Z}^n = (\bar{\nu}^n, \tilde{\mathbf{X}}^n, \bar{\mathbf{V}}^n) : n \geq 1\}$ is tight as well and then characterize weak limit points of the sequence.

(iii) Finally, in Section 6.5.4 we will complete the proof of the Laplace principle upper bound by taking limits as $n \to \infty$ in (6.5.7).

6.5.2 Tightness of $\{\bar{\nu}^n\}$

The goal of this section is to prove the following proposition.

Proposition 6.5.2. $\{\bar{\nu}^n\}_{n\geq 1}$ defined in (6.5.6) is a tight sequence of $\mathcal{P}(\mathbb{N}_0 \times [0,T])$ valued random variables. Furthermore, it has the following uniform integrability property

$$\lim_{K\to\infty} \sup_{n\in\mathbb{N}_0} \mathbb{E}\left\{ \sum_{k=K+1}^{\infty} k\bar{\nu}^n(\{k\} \times [0,T]) \right\} = 0.$$

We will make use of the following Feller property of the stochastic kernel $\mu(\cdot \mid \cdot)$. We recall that we are suppressing ϵ in the notation, namely $\mu = \mu^\epsilon$ for some fixed $\epsilon > 0$, where μ^ϵ is as defined in (6.4.1).

Lemma 6.5.3. Let $\{\mathbf{x}^{(n)}\}_{n=1}^{\infty} \subset \bar{\mathfrak{D}}_{\text{deg}}$ be such that $\mathbf{x}^{(n)} \to \mathbf{x}$ as $n \to \infty$. Then, $\mu(\cdot \mid \mathbf{x}^{(n)}) \to \mu(\cdot \mid \mathbf{x})$.

Proof: Write $\mathbf{x}^{(n)} = (a^{(n)}, (v_k^{(n)})_{k\geq 1})$ and $\mathbf{x} = (a, (v_k)_{k\geq 1})$. We need to show for each fixed $k \geq 0$, $\lim_{n\to\infty} \mu(k \mid \mathbf{x}^{(n)}) = \mu(k \mid \mathbf{x})$. Noting that the denominator in the definition of $\mu(k|\mathbf{x}^{(n)})$ is bounded away from 0, it suffices to show $\lim_{n\to\infty} \sum_{k=1}^{\infty} kv_k^{(n)} = \sum_{k=1}^{\infty} kv_k$. This follows from convergence of $v_k^{(n)}$ to v_k for each k and noting from the definition of $\mathfrak{D}_{\text{deg}}$ that for any $m \geq 1$ $\sup_n \sum_{k\geq m} kv_k^{(n)} \leq \frac{2\nu_1^*}{m}$. ■

In order to prove Proposition 6.5.2 we will prove a more general result of the following form.

Theorem 6.5.4. Let $\{\pi_j^n\}_{j\in\mathbb{N}_0,n\geq 1}$ and $\{\gamma_j^n\}_{j\in\mathbb{N}_0,n\geq 1}$ be collections of $\mathcal{P}(\mathbb{N}_0)$ valued random variables such that

$$\sup_{n\in\mathbb{N}_0} \mathbb{E}\left\{\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\sum_{k\in\mathbb{N}_0} e^{\lambda k}\pi_j^n(k)\right\} < \infty \text{ for all } \lambda > 0 \tag{6.5.8}$$

$$\sup_{n\in\mathbb{N}_0} \mathbb{E}\left\{\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor} R(\gamma_j^n\|\pi_j^n)\right\} < \infty \tag{6.5.9}$$

Then the sequence $\{\bar{\gamma}^n : n \geq 1\}$ of $\mathcal{P}(\mathbb{N}_0 \times [0,T])$ valued random variables, defined as

$$\bar{\gamma}^n(A\times B) = \frac{1}{T}\int_B \gamma_{\lfloor nt\rfloor}^n(A)dt,\ A\subset\mathbb{N}_0, B\in\mathcal{B}([0,T]),$$

is tight and furthermore has the following uniform integrability property.

$$\lim_{K\to\infty}\sup_{n\in\mathbb{N}_0}\mathbb{E}\sum_{k=K+1}^{\infty} k\bar{\gamma}^n(\{k\}\times[0,T]) = 0. \tag{6.5.10}$$

Proof: From (6.5.9) we see that for all $j\in\mathbb{N}_0$, $\gamma_j^n \ll \pi_j^n$ a.s. Let $f_j^n(y) := \frac{d\gamma_j^n}{d\pi_j^n}(y)$, $y\in\mathbb{N}_0$. Using the elementary inequality

$$ab \leq e^{\lambda a} + \frac{1}{\lambda}(b\log b - b + 1),\ \text{for all } a\geq 0, b\geq 0, \lambda\geq 1$$

with $a = y$ and $b = f_j^n(y)$, we have

$$yf_j^n(y) \leq e^{\lambda y} + \frac{1}{\lambda}[f_j^n(y)\log f_j^n(y) - f_j^n(y) + 1] \text{ for all } \lambda\geq 1,\ j\in\mathbb{N}_0,\ n\in\mathbb{N}.$$

Then, for every $\lambda\geq 1$,

$$\begin{aligned}
T\mathbb{E}\int_{\{y\in\mathbb{N}_0:y>K\}\times[0,T]} y\bar{\gamma}^n(dy\times dt) &\leq \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\{y\in\mathbb{N}_0,y>K\}} y\gamma_j^n(dy)\\
&= \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\{y\in\mathbb{N}_0,y>K\}} yf_j^n(y)\cdot\pi_j^n(dy)\\
&\leq I_1(\lambda) + I_2(\lambda),
\end{aligned}$$

where

$$I_1(\lambda) := \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\{y\in\mathbb{N}_0, y>K\}} e^{\lambda y}\pi_j^n(dy) \le \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\{y\in\mathbb{N}_0, y>K\}} e^{-\lambda K}\cdot e^{2\lambda y}\pi_j^n(dy)$$

$$\le e^{-\lambda K}\mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\mathbb{N}_0} e^{2\lambda y}\pi_j^n(dy) \le e^{-\lambda K}M_1(2\lambda),$$

and $M_1(\lambda)$ denotes the supremum in (6.5.8); and

$$I_2(\lambda) := \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\{y\in\mathbb{N}_0, y>K\}} \frac{1}{\lambda}[f_j^n(y)\log f_j^n(y) - f_j^n(y) + 1]\pi_j^n(dy)$$

Since $b\log b - b + 1 \ge 0$ for all $b \ge 0$, we have

$$\begin{aligned} I_2(\lambda) \le& \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\mathbb{N}_0} \frac{1}{\lambda}[f_j^n(y)\log f_j^n(y) - f_j^n(y) + 1]\pi_j^n(dy) \\ =& \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor}\int_{\mathbb{N}_0} \frac{1}{\lambda}f_j^n(y)\log f_j^n(y)\pi_j^n(dy) = \frac{1}{\lambda}\mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT\rfloor} R(\gamma_j^n\|\pi_j^n) \le \frac{1}{\lambda}M_2, \end{aligned}$$

where M_2 is the supremum in (6.5.9). Combining the bounds on I_1 and I_2, we have, for all $n \in \mathbb{N}_0$ and $\lambda \ge 1$,

$$\sup_n T\,\mathbb{E}\int_{\{y\in\mathbb{N}_0 : y>K\}\times[0,T]} y\bar{\gamma}^n(dy\times dt) \le e^{-\lambda K}M_1(2\lambda) + \frac{1}{\lambda}M_2.$$

The uniform integrability stated in (6.5.10) now follows on first sending $K \to \infty$ and then let $\lambda \to \infty$ in the above display.

Finally, in order to prove the tightness of $\{\bar{\gamma}^n\}$ it suffices to show that $\sigma^n = \mathbb{E}\bar{\gamma}^n$ is a tight sequence of probability measures in $\mathbb{N}_0 \times [0,T]$. However the tightness of the latter sequence is immediate on noting that

$$\limsup_{K\to\infty}\sup_n \sigma^n\left([K+1,\infty)\times[0,T]\right) \le \limsup_{K\to\infty}\frac{1}{K}\sup_n\sum_{k=K+1}^{\infty} k\sigma^n(\{k\}\times[0,T]) = 0,$$

where the last equality is from (6.5.10). ■

Proof of Proposition 6.5.2: It suffices to verify (6.5.8) and (6.5.9) with $\pi_j^n = \mu(\cdot \mid \frac{1}{n}\bar{\mathbf{X}}(j))$ and $\gamma_j^n = \nu_j^n$. The proof of (6.5.9) is immediate from (6.5.5) on noting that h is bounded, $\varepsilon \in (0,1)$ and recalling the definition of W^n.

We now verify (6.5.8). Note that for every $k \geq 1$, $\bar{V}_k(0) = n_k$ and from (6.5.3), $\{\bar{V}_k(j\}$ is a non-decreasing sequence a.s., since $\bar{V}_k(j+1) - \bar{V}_k(j)$ can only take values in $\{0, -1\}$. Also note from (6.4.1) that on the set $\{\bar{V}_k(j) = 0\}$ $\mu(k \mid \frac{1}{n}\bar{\mathbf{X}}(j)) = 0$ and since as argued above $R(\nu_j^n \| \mu(\cdot \mid \frac{1}{n}\bar{\mathbf{X}}(j))) < \infty$ a.s., we must have that $\nu_j^n(k) = 0$ a.s. Thus on the set $\{\bar{V}_k(j) = 0\}$ we have $\bar{V}_k(l) = 0$ for all $l \geq j$, a.s. Combining the above observations we have that $\frac{1}{n}\bar{V}_k(j) \leq n_k/n$ for all $k, n, j \in \mathbb{N}$. Thus, for all $\lambda > 0$

$$\begin{aligned}\sup_{n \in \mathbb{N}_0} \mathbb{E}\frac{1}{n}\sum_{j=0}^{\lfloor nT \rfloor}\sum_{k \in \mathbb{N}_0} e^{\lambda k}\mu(k \mid \frac{1}{n}\bar{\mathbf{X}}(j)) &\leq T + \frac{1}{n\epsilon}\sum_{j=0}^{\lfloor nT \rfloor}\sum_{k \in \mathbb{N}_0} k e^{\lambda k}\frac{1}{n}\bar{V}_k(j) \\ &\leq T\left(1 + \frac{1}{\lambda\epsilon}\sup_n \sum_{k \in \mathbb{N}_0} e^{2\lambda k}\frac{n_k}{n}\right) < \infty.\end{aligned}$$

This verifies (6.5.8) and completes the proof of Proposition 6.5.2. ∎

6.5.3 Characterizing limit points.

The following is the main result of this section.

Proposition 6.5.5. (a) $\{\mathbf{G}^n = (\bar{\nu}^n, \tilde{\mathbf{X}}^n, \bar{\mathbf{V}}^n) : n \geq 1\}$ is a tight sequence of $\mathcal{P}(\mathbb{N}_0 \times [0,T]) \times D([0,T] : \mathbb{R}_+ \times \mathbb{R}_+^\infty) \times C([0,T] : \mathbb{R}_+^\infty)$ valued random variables.

(b) Suppose $\mathbf{G}^n$ converges in distribution along a subsequence to $\mathbf{G}^* = (\bar{\nu}^*, \mathbf{X}^*, \mathbf{V}^*)$. Then the following hold

(i) Writing $\tilde{\mathbf{X}}^* = (\tilde{A}^*, \tilde{\mathbf{V}}^*)$, we have $\bar{\mathbf{V}}^* = \tilde{\mathbf{V}}^*$ a.s.

(ii) For all $k \geq 1$ and $t \in [0,T]$

$$\bar{V}_k^*(t) = p_k - T\bar{\nu}^*(\{k\} \times [0,t]) = p_k - \int_{[0,t]} \bar{\nu}^*(k \mid s)ds, \tag{6.5.11}$$

where $\bar{\nu}^*(\cdot \mid \cdot)$ is a stochastic kernel on $\mathbb{N}_0$ given $[0,T]$.

(iii) Letting

$$\hat{B}^*(t) = \int_{\mathbb{N}_0 \times [0,t]} y\bar{\nu}^*(dyds) - 2t,\ t \in [0,T],$$

$\tilde{A}^* = \Gamma(\hat{B}^*)$ and $\tilde{\mathbf{X}}^*(t) = \mathbf{x}^{\bar{\mathbf{V}}^*(t)}$, for a.e. $t \in [0,T]$, a.s.

The following lemma will be needed in the proof of the proposition. Define

$$\tilde{\mathbf{V}}^n(t) = \frac{1}{n}\bar{\mathbf{V}}(\lfloor nt \rfloor),\ \tilde{A}^n(t) = \frac{1}{n}(\bar{A}(\lfloor nt \rfloor) - 1)^+,\ \tilde{B}^n(t) = \frac{1}{n}\bar{B}(\lfloor nt \rfloor),\ t \in [0,T]. \tag{6.5.12}$$

Note that $\tilde{\mathbf{X}}^n = (\tilde{A}^n, \tilde{\mathbf{V}}^n)$.

Lemma 6.5.6. For every $n \geq 1$,

$$\sup_{0 \leq t \leq T} \left| \tilde{A}^n(t) - \Gamma(\tilde{B}^n)(t) \right| \leq \frac{3}{n}.$$

Proof: For $j \in \mathbb{N}_0$, let $\mathcal{S}_j(\bar{B}) = \bar{B}(j) - \inf_{i \leq j}(\bar{B}(i) \wedge 0)$. It suffices to show that for all $j \in \mathbb{N}_0$

$$|\bar{A}(j) - \mathcal{S}_j(\bar{B})| \leq 2. \tag{6.5.13}$$

We prove the statement by induction on j. When $j = 0$, $A(0) = \gamma_0(\bar{B}) = 0$ and so (6.5.13) holds. Suppose now that (6.5.13) holds for some $j \in \mathbb{N}_0$ and consider $j+1$. If $(\bar{A}(j), \mathcal{S}_j(\bar{B})) \in \{(0,1),(0,0),(-1,1),(-1,0)\}$, one can check by a direct calculation that (6.5.13) holds with j replaced by $j+1$.

Consider now the case where either $\bar{A}(j) \geq 1$ or $\mathcal{S}_j(\bar{B}) \geq 2$. Letting $A_1(j) = \bar{A}(j) \vee \mathcal{S}_j(\bar{B})$ and $A_2(j) = \bar{A}(j) \wedge \mathcal{S}_j(\bar{B})$, we have that

$$A_1(j+1) - A_1(j) = \bar{\xi}(j+1) - 2 \text{ and } A_2(j+1) - A_2(j) \geq \bar{\xi}(j+1) - 2.$$

Thus

$$\begin{aligned} |\bar{A}(j+1) - \mathcal{S}_{j+1}(\bar{B})| &= A_1(j+1) - A_2(j+1) \\ &= (A_1(j+1) - A_1(j)) - (A_2(j+1) - A_2(j)) + (A_1(j) - A_2(j)) \\ &\leq (\bar{\xi}(j+1) - 2) - (\bar{\xi}(j+1) - 2) + 2 = 2. \end{aligned}$$

The result follows. ∎

We now proceed to the proof of Proposition 6.5.5.

Proof of Proposition 6.5.5: Tightness of $\bar{\nu}^n$ in $\mathcal{P}(\mathbb{N}_0 \times [0,T])$ was shown in Proposition 6.5.2. To prove tightness of $\{(\tilde{\mathbf{X}}^n, \bar{\mathbf{V}}^n)\}$ we will consider auxiliary processes $(\hat{B}^n, \hat{\mathbf{V}}^n)$ with sample paths in $C([0,T] : \mathbb{R} \times \mathbb{R}_+^\infty)$ defined as follows. For $t \in [0,T]$,

$$\hat{V}_k^n(t) = p_k - T \int_{\{k\}\times[0,t]} \bar{\nu}^n(dy \times ds) \text{ for } k \geq 1 \tag{6.5.14}$$

$$\hat{B}^n(t) = T \int_{\mathbb{N}_0 \times [0,t]} y \bar{\nu}^n(dy \times ds) - 2t. \tag{6.5.15}$$

We first argue that

$$\{\hat{(}B^n, \hat{\mathbf{V}}^n)\}_{n\geq 1} \text{ is tight in } C([0,T] : \mathbb{R} \times \mathbb{R}_+^\infty). \tag{6.5.16}$$

Tightness of $\hat{\mathbf{V}}^n$ is immediate on noting that

$$|\hat{V}_k^n(t) - \hat{V}_k^n(s)| \leq |t-s|, \text{ for all } 0 \leq s \leq t \leq T,\ n \geq 1,\ k \geq 1. \tag{6.5.17}$$

Next, for $0 \leq s \leq t \leq T$ and $K \in \mathbb{N}$,

$$|\hat{B}^n(t) - \hat{B}^n(s)| = \left| \int_{\mathbb{N}_0 \times (s,t]} y \bar{\nu}^n(dydu) \right| \leq (K+2)|t-s| + \mathcal{U}^n(K), \tag{6.5.18}$$

where $\mathcal{U}^n(K) = \sum_{k=K+1}^{\infty} k \bar{\nu}^n(\{k\} \times [0,T])$. Also, from Proposition 6.5.2

$$\sup_{n\in\mathbb{N}} \mathbb{E}(\mathcal{U}^n(K)) \to 0 \text{ as } K \to \infty.$$

Tightness of $\{\hat{B}^n\}$ is now immediate. We now argue that, for all $k \geq 1$,

$$\max\left\{ \|\bar{V}_k^n - \hat{V}_k^n\|_\infty, \|\tilde{V}_k^n - \hat{V}_k^n\|_\infty \right\} \xrightarrow{\mathbb{P}} 0 \text{ as } n \to \infty, \tag{6.5.19}$$

$$\text{and } \|\tilde{B}^n - \hat{B}^n\|_\infty \xrightarrow{\mathbb{P}} 0 \text{ as } n \to \infty. \tag{6.5.20}$$

From (6.5.17) we have that

$$\max\left\{ \sup_{t\in[0,T]} |\bar{V}_k^n(t) - \hat{V}_k^n(t)|, \sup_{t\in[0,T]} |\tilde{V}_k^n(t) - \hat{V}_k^n(t)| \right\}$$
$$\leq \sup_{0\leq j \leq \lfloor nT \rfloor} |\bar{V}_k^n(j/n) - \hat{V}_k^n(j/n)| + \frac{1}{n}. \tag{6.5.21}$$

Next

$$
\begin{aligned}
&\sup_{0\le j\le \lfloor nT\rfloor} |\bar{V}^n_k(j/n) - \hat{V}^n_k(j/n)| \\
=&\sup_{0\le j\le \lfloor nT\rfloor} \left| \left(\frac{n_k}{n} - \frac{1}{n}\sum_{i=1}^{j} \mathbf{1}_{\{\bar{\xi}(i)=k\}}\right) - \left(p_k - \frac{1}{n}\sum_{i=1}^{j} \nu^n_{i-1}(k)\right)\right| \\
\le&|n_k/n - p_k| + \sup_{0\le j\le \lfloor nT\rfloor} \left|\frac{1}{n}\sum_{i=1}^{j}(\mathbf{1}_{\{\bar{\xi}(i)=k\}} - \nu^n_{i-1}(k))\right| \\
=&|n_k/n - p_k| + \frac{1}{n}\sup_{0\le j\le \lfloor nT\rfloor} |Y(j)|,
\end{aligned}
$$

where $Y(0) = 0$ and $Y(j) := \sum_{i=1}^{j}(\mathbf{1}_{\{\bar{\xi}(i)=k\}} - \nu^n_{i-1}(k))$, $j = 1, 2, ...$ From (6.5.2) we see that $\{Y(j)\}$ is a martingale and so for any $\eta > 0$,

$$
\begin{aligned}
\mathbb{P}\left\{\frac{1}{n}\sup_{0\le j\le \lfloor nT\rfloor} |Y(j)| > \eta\right\} \le& \frac{4}{n^2\eta^2}\mathbb{E}(|Y(\lfloor nT\rfloor)|)^2 \\
=&\frac{4}{n^2\eta^2}\sum_{j=1}^{\lfloor nT\rfloor}\mathbb{E}\left[(\mathbf{1}_{\{\bar{\xi}(i)=k\}} - \nu^n_{i-1}(k))^2\right] \le \frac{4T}{n\eta^2}.
\end{aligned}
$$

Using the above two displays in (6.5.21) and recalling that $n_k/n \to p_k$ as $n \to \infty$, we have (6.5.19).

Next consider (6.5.20). Note that

$$
\|\tilde{B}^n - \hat{B}^n\|_\infty \le \sup_{0\le j\le \lfloor nT\rfloor} |\tilde{B}^n(j/n) - \hat{B}^n(j/n)| + \Theta_n,
$$

where $\Theta_n = \sup_{s,t\in[0,T],|t-s|\le 1/n} |\hat{B}^n(t) - \hat{B}^n(s)|$. From (6.5.18) and recalling that $\sup_{n\in\mathbb{N}} \mathbb{E}(\mathcal{U}^n(K)) \to 0$ as $K \to \infty$, we see that $\Theta_n \to 0$ in L^1 as $n \to \infty$. Also,

$$
\begin{aligned}
&\sup_{1\le j\le \lfloor nT\rfloor} |\tilde{B}^n(j/n) - \hat{B}^n(j/n)| \\
=&\max_{1\le j\le \lfloor nT\rfloor} \left|\frac{1}{n}\sum_{i=1}^{j}(\bar{\xi}(i) - 2) - \frac{1}{n}\sum_{i=1}^{j}\left(\int_{\mathbb{N}_0} y\nu^n_{i-1}(dy) - 2\right)\right| \qquad (6.5.22)
\end{aligned}
$$

For any $K > 0$, the right hand side can be bounded by

$$
\begin{aligned}
&\frac{1}{n}\sup_{0\le j\le \lfloor nT\rfloor}\left|\sum_{i=1}^{j}\left(\bar{\xi}(i)\wedge K - \int_{\mathbb{N}_0}(y\wedge K)\nu^n_{i-1}(dy)\right)\right| + \frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}|\bar{\xi}(i) - (\bar{\xi}(i)\wedge K)| \\
&\quad + \frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\left|\int_{\mathbb{N}_0}(y\wedge K)\nu^n_{i-1}(dy) - \int_{\mathbb{N}_0} y\nu^n_{i-1}(dy)\right| := I_1(K) + I_2(K) + I_3(K).
\end{aligned}
$$

Note that $\sum_{i=1}^{j}\left(\bar{\xi}(i)\wedge K-\int_{\mathbb{N}_0}(y\wedge K)\nu_{i-1}^n(dy)\right)$, $j=1,2,...$, is a martingale and so

$$\mathbb{E}[(I_1(K))^2]\leq\frac{4}{n^2}\mathbb{E}\left[\sum_{i=1}^{\lfloor nT\rfloor}\left[\bar{\xi}(i)\wedge K-\int_{\mathbb{N}_0}(y\wedge K)\nu_{i-1}^n(dy)\right]^2\right]\leq\frac{4}{n^2}\cdot nTK^2=\frac{4TK^2}{n}.$$

Also,

$$\begin{aligned}\mathbb{E}[I_2(K)]\leq&\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\mathbb{E}|\bar{\xi}(i)-(\bar{\xi}(i)\wedge K)|\\ \leq&\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\mathbb{E}\left[\int_{\mathbb{N}_0}y\mathbf{1}_{\{y>K\}}\nu_{i-1}^n(dy)\right]=\mathbb{E}\left[\int_{\{y>K\}\times[0,T]}y\bar{\nu}^n(dy\times ds)\right],\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[I_3(K)]=&\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\mathbb{E}\left[\left(\int_{\mathbb{N}_0}y\nu_{i-1}^n(dy)-\int_{\mathbb{N}_0}(y\wedge K)\nu_{i-1}^n(dy)\right)\right]\\ =&\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\mathbb{E}\left[\left(\int_{\{y>K\}}y\nu_{i-1}^n(dy)-K\nu_{i-1}^n(\{y>K\})\right)\right]\\ \leq&\frac{1}{n}\sum_{i=1}^{\lfloor nT\rfloor}\mathbb{E}\left[\int_{\mathbb{N}_0}y\mathbf{1}_{\{y>K\}}\nu_{i-1}^n(dy)\right]=\mathbb{E}\left[\int_{\{y>K\}\times[0,T]}y\bar{\nu}^n(dy\times ds)\right].\end{aligned}$$

Using these estimates in (6.5.22), we have for any $\eta>0$

$$\begin{aligned}\mathbb{P}\left\{\sup_{0\leq j\leq\lfloor nT\rfloor}|\tilde{B}^n(j/n)-\hat{B}^n(j/n)|>\eta\right\}\leq&\sum_{j=1}^{3}\mathbb{P}\left\{I_j(K)>\eta/3\right\}\\ \leq&\frac{9}{\eta^2}\mathbb{E}[(I_1(K))^2]+\frac{3}{\eta}\mathbb{E}[I_2(K)]+\frac{3}{\eta}\mathbb{E}[I_3(K)]\\ \leq&\frac{9}{\eta^2}\cdot\frac{4TK^2}{n}+\frac{6}{\eta}E\left[\mathcal{U}^n(K)\right].\end{aligned}$$

Proof of (6.5.20) now follows on recalling that $\sup_{n\in\mathbb{N}}\mathbb{E}(\mathcal{U}^n(K))\to 0$ as $K\to\infty$ and first sending $n\to\infty$ and then $K\to\infty$ in the above display.

Tightness of $\bar{\mathbf{V}}^n$ and $\tilde{\mathbf{V}}^n$ is now immediate from (6.5.19) and the tightness of $\hat{\mathbf{V}}^n$ established earlier. Also, from Lemma 6.5.6, tightness of $\{\tilde{B}^n\}$ established earlier and the continuity of the Skorohod map, we have that $\tilde{A}^n$ is tight in $D([0,T]:\mathbb{R}_+)$. This proves the tightness of $\tilde{\mathbf{X}}^n$ and completes the proof of part (a).

Next suppose $\mathbf{G}^n = (\bar{\nu}^n, \tilde{\mathbf{X}}^n, \bar{\mathbf{V}}^n)$ converges in distribution to $\mathbf{G}^* = (\bar{\nu}^*, \tilde{\mathbf{X}}^*, \bar{\mathbf{V}}^*)$. Without loss of generality we can assume that the convergence is a.s. From (6.5.14) it follows that, for all $k \geq 1$, $\hat{V}_k^n$ converges a.s. in $C([0,T] : \mathbb{R}_+)$ to $\hat{V}_k^*$ given as

$$\hat{V}_k^*(t) = p_k - T\bar{\nu}^*(\{k\} \times [0,t]), t \in [0,T]. \tag{6.5.23}$$

Also, using Proposition 6.5.2 and (6.5.15) we have that $\hat{B}^n$ converges a.s. in $C([0,T] : \mathbb{R})$ to $\hat{B}^*$ given as

$$\hat{B}^*(t) = T\int_{\mathbb{N}_0\times[0,t]} y\bar{\nu}^*(dydt) - 2t,\ t \in [0,T]. \tag{6.5.24}$$

From (6.5.19) we obtain that $\bar{\mathbf{V}}^* = \hat{\mathbf{V}}^* = \tilde{\mathbf{V}}^*$, proving (i). Also, combining with (6.5.23), this proves the first equality in (6.5.11). The second equality in (6.5.11) is immediate on noting that $\bar{\nu}^n(\mathbb{N}_0 \times \cdot)$ is the normalized Lebesgue measure on $[0,T]$ for every n. This completes the proof of (ii).

Next, from (6.5.20) and the above established convergence for $\hat{B}^n$ we have that $\tilde{B}^n \to \hat{B}^*$ in probability. Using Lemma 6.5.6 again, we now have that $\tilde{A}^n$ converges a.s. to $\Gamma(\hat{B}^*)$ which proves $\tilde{A}^* = \Gamma(\hat{B}^*)$. Also, recall that for $\mathbf{v} \in \mathfrak{P}$, $\mathbf{x}^{\mathbf{v}} = (a^{\mathbf{v}}, \mathbf{v})$, where $a^{\mathbf{v}}$ is as in (6.3.6). Thus, from (6.5.24)

$$\begin{aligned}
b^{\bar{\mathbf{V}}^*}(t) =& \sum_{k=1}^{\infty} k\bar{V}_k^*(0) - 2t - \sum_{k=1}^{\infty} k\bar{V}_k^*(t) \\
=& \sum_{k=1}^{\infty} kp_k - 2t - \sum_{k=1}^{\infty} kp_k + \int_{\mathbb{N}_0\times[0,t]} y\bar{\nu}^*(dyds) \\
=& \hat{B}^*(t).
\end{aligned}$$

Combining this with the equality $\tilde{A}^* = \Gamma(\hat{B}^*)$ we now have that $\tilde{\mathbf{X}}^*(t) = \mathbf{x}^{\bar{\mathbf{V}}^*}(t)$ a.s. for all $t \in [0,T]$. This proves part (iii). ■

6.5.4 Completing the proof for the Laplace upper bound

Recall from (6.5.5) that

$$W^n + \varepsilon \geq \mathbb{E}\left\{TR\left(\bar{\nu}^n \middle\| \bar{\mu}^n\right) + h(\bar{\mathbf{V}}^n)\right\}, \tag{6.5.25}$$

where $\bar{\mu}^n$ is a $\mathcal{P}(\mathbb{N}_0 \times [0,T])$ valued random variable defined as

$$\bar{\mu}^n(A \times B) := \frac{1}{T}\int_B \mu(A \mid \tilde{\mathbf{X}}^n(t))dt,\ A \in \mathcal{B}(\mathbb{N}_0),\ B \in \mathcal{B}([0,T]). \tag{6.5.26}$$

Let $\mathbf{G}^n$ be as in Proposition 6.5.5 and fix a subsequence along which $\mathbf{G}^n$ converges in distribution to $\mathbf{G}^*$. It suffices to prove (6.5.1) along such a subsequence. We assume without loss of generality that $\mathbf{G}^n \to \mathbf{G}^*$ a.s. Recall that $\mathbf{G}^*$ satisfies properties (i)-(iii) in Proposition 6.5.5 and note that $\tilde{\mathbf{X}}^n(t) \in \mathfrak{D}^*_{\text{deg}}$ a.s. for all $t \in [0,T]$. From the convergence of $\tilde{\mathbf{X}}^n$ to $\mathbf{X}^*$ and Lemma 6.5.3 we now have that

$$\mu(\cdot \mid \tilde{\mathbf{X}}^n(t)) \to \mu(\cdot \mid \mathbf{X}^*(t)),\ \text{for all } t \in [0,T],\ a.s.$$

Consequently, $\bar{\mu}^n \to \mu^*$ a.s., where μ^* is defined through (6.5.26) by replacing $\tilde{\mathbf{X}}^n$ on the right side with $\mathbf{X}^*$.

Using the above result and Fatou's lemma in (6.5.25) we have

$$\begin{aligned}\liminf_{n\to\infty} W^n + \varepsilon \geq & T\,\mathbb{E}\left[\liminf_{n\to\infty} R(\bar{\nu}^n \| \bar{\mu}^n)\right] + \mathbb{E}\left[h(\mathbf{V}^*)\right] \\ \geq & T\,\mathbb{E}\left[R(\nu^*(\cdot)\|\mu^*(\cdot))\right] + \mathbb{E}\left[h(\mathbf{V}^*)\right],\end{aligned} \tag{6.5.27}$$

where the second line follows from the lower-semicontinuity of the relative entropy. From Proposition 6.5.5(b)(ii) we see that

$$\dot{V}^*_k(t) = -\nu^*(k \mid t),\ \text{a.e. } t \in [0,T],\ a.s.$$

and so using notation from (6.3.2) we have $\nu^*(\cdot \mid t) = \nu(\cdot \mid \dot{\mathbf{V}}^*(t))$, a.e. t, a.s.. Also, from Proposition 6.5.5(b)(iii), $\mathbf{X}^*(t) = \mathbf{x}^{\mathbf{V}^*}(t)$ a.s. for all $t \in [0,T]$. Thus recalling the definition of L (see (6.4.6)) we have

$$R(\nu^*\|\mu^*) = \int_{[0,T]} R\big(\nu^*(\cdot \mid t)\big\|\mu(\cdot \mid \mathbf{X}^*(t))\big)dt = \int_{[0,T]} L(x^{\mathbf{V}^*}(t), \dot{\mathbf{V}}^*(t))dt = I(\mathbf{V}^*).$$

Finally, combining this with (6.5.27) we have

$$\liminf_{n\to\infty} W^n + \varepsilon \geq \mathbb{E}\left[I(\mathbf{V}^*) + h(\mathbf{V}^*)\right] \geq \inf_{\mathbf{v}\in\mathfrak{P}} \{I(\mathbf{v}) + h(\mathbf{v})\}.$$

Since ε is arbitrary, we have (6.5.1) and the proof of the Laplace principle upper bound (i.e. item (b) in Section 6.4.4) is complete. ■

6.6 I^ϵ is a rate function.

In this section we show that for every $\epsilon > 0$, I^ϵ is a rate function, namely the set $S^\epsilon_M = \{\boldsymbol{\varphi} \in \mathfrak{P} : I^\epsilon(\boldsymbol{\varphi}) \le M\}$ is compact for every $M < \infty$. As in the previous section we suppress ϵ from the notation.

Let $\{\boldsymbol{\varphi}^n\}_{n\ge1} \subset S_M$ be such that $\boldsymbol{\varphi}^n \to \boldsymbol{\varphi}^0$. It suffices to show that $I(\boldsymbol{\varphi}^0) \le M$. Note that, since $I(\boldsymbol{\varphi}^n) < \infty$, $\boldsymbol{\varphi}^n \in \mathfrak{P}_I$ for each n, namely

$$\varphi^n_k(0) = p_k, \varphi^n_k(t) \ge 0, \text{ for all } t \in [0,T];\ \varphi^n_k \text{ is absolutely continuous on } [0,T],\ k \ge 1$$
$$\text{and } \dot{\boldsymbol{\varphi}}^n = (\dot{\varphi}^n_k)_{k\ge1} \in \mathfrak{D}_{\text{vel}}. \tag{6.6.1}$$

It then follows that $\boldsymbol{\varphi}^0 \in \mathfrak{P}_I$ as well. Next, for $n \ge 0$,

$$I(\boldsymbol{\varphi}^n) = \int_0^T R\big(\nu(\cdot \mid \dot{\boldsymbol{\varphi}}^n(t)) \| \mu(\cdot \mid \mathbf{x}^{\boldsymbol{\varphi}^n}(t))\big)dt = TR(\bar{\nu}^n \| \bar{\mu}^n),$$

where $\bar{\nu}^n, \bar{\mu}^n \in \mathcal{P}(\mathbb{N}_0 \times [0,T])$ are defined as

$$\bar{\nu}^n(A \times B) = \frac{1}{T}\int_B \nu(A \mid \dot{\boldsymbol{\varphi}}^n(t))dt,$$
$$\bar{\mu}^n(A \times B) = \frac{1}{T}\int_B \mu(A \mid \mathbf{x}^{\boldsymbol{\varphi}^n}(t))dt,$$

for $A \in \mathcal{B}(\mathbb{N}_0)$ and $B \in \mathcal{B}([0,T])$. In order to prove the result it suffices from the lower semicontinuity property of relative entropy to argue that $\bar{\nu}^n \to \bar{\nu}^0$ and $\bar{\mu}^n \to \bar{\mu}^0$. For this, note that since $\varphi^n_k(t) \le p_k$ for all $n \in \mathbb{N}_0, k \in \mathbb{N}$, we get by Assumption 6.2.2 and dominated convergence that $b^{\boldsymbol{\varphi}^n} \to b^{\boldsymbol{\varphi}^0}$ and combining with the continuity property of the Skorohod map, we have $a^{\boldsymbol{\varphi}^n} \to a^{\boldsymbol{\varphi}^0}$. This in view of Lemma 6.5.3 shows that $\mu(\cdot \mid \mathbf{x}^{\boldsymbol{\varphi}^n}(t)) \to \mu(\cdot \mid \mathbf{X}^{\boldsymbol{\varphi}^0}(t))$ for all $t \in [0,1]$ and thus $\bar{\mu}^n \to \bar{\mu}^0$ as $n \to \infty$.

Consider now $\bar{\nu}^n$. Note that for any $\lambda \geq 1$

$$\begin{aligned}\int_{\mathbb{N}_0 \times [0,T]} e^{\lambda y} d\bar{\mu}^n &= \frac{1}{T} \int_{[0,T]} \sum_k e^{\lambda k} \mu(k \mid \mathbf{x}^{\boldsymbol{\varphi}^n}(t)) dt \\ &\leq 1 + \frac{1}{T\epsilon} \int_{[0,T]} \sum_k k e^{\lambda k} \varphi_k^n(t) dt \\ &\leq 1 + \frac{1}{\epsilon} \sum_k k e^{\lambda k} p_k < \infty,\end{aligned}$$

where the last inequality is from Assumption 6.2.2. Thus an argument similar to that in the proof of Theorem 6.5.4 (see also Lemma 1.4.3 in [14]) shows that $\{\bar{\nu}^n\}_{n \geq 1}$ is a tight sequence and $\sup_n \sum_{k=K+1}^{\infty} k \bar{\nu}^n(\{k\} \times [0,T]) \to 0$ as $K \to \infty$. Finally let $\bar{\nu}$ be a limit point of $\bar{\nu}^n$. Noting that for all $k \geq 1$ and $n \in \mathbb{N}_0$

$$\varphi_k^n(t) = p_k - T\bar{\nu}^n(\{k\} \times [0,t]), \text{ for all } t \in [0,T],$$

we now have on sending $n \to \infty$ along the convergent subsequence that

$$p_k - T\bar{\nu}^0(\{k\} \times [0,t]) = \varphi_k^0(t) = p_k - T\bar{\nu}(\{k\} \times [0,t]), \text{ for all } k \geq 1.$$

Also, $\bar{\nu}^0(\mathbb{N}_0 \times [0,t]) = \bar{\nu}(\mathbb{N}_0 \times [0,t]) = t$ for all $t \in [0,1]$. Thus $\bar{\nu}^0 = \bar{\nu}$. Thus we have proved that $\bar{\nu}^n \to \bar{\nu}^0$. The result follows. ■

BIBLIOGRAPHY

[1] D. Achlioptas, R. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science*, 323(5920):1453, 2009.

[2] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25(2):812–854, 1997.

[3] D. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab*, 3:1–59, 1998.

[4] D. Aldous and B. Pittel. On a random graph with immigrating vertices: Emergence of the giant component. *Random Structures and Algorithms*, 17(2):79–102, 2000.

[5] S. Bhamidi, A. Budhiraja, and X. Wang. The augmented multiplicative coalescent and critical dynamic random graph models. *To appear in Probability Theory and Related Fields. arXiv preprint arXiv:1212.5493*, 2012.

[6] S. Bhamidi, A. Budhiraja, and X. Wang. Bounded-size rules: The barely subcritical regime. *arXiv preprint arXiv:1212.5480*, 2012.

[7] S. Bhamidi, A. Budhiraja, and X. Wang. Aggregation models with limited choice and the multiplicative coalescent. *Random Structures & Algorithms*, 2013.

[8] T. Bohman and A. Frieze. Avoiding a giant component. *Random Structures and Algorithms*, 19(1):75–85, 2001.

[9] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[10] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[11] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31(1):3–122, 2007.

[12] P. Brémaud. *Point processes and queues*. Springer-Verlag, New York, 1981. Martingale dynamics, Springer Series in Statistics.

[13] W. Chen and R. M. DSouza. Explosive percolation with multiple giant components. *Physical review letters*, 106(11):115701, 2011.

[14] P. Dupuis and R. S. Ellis. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1997. A Wiley-Interscience Publication.

[15] R. Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.

[16] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.

[17] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[18] P. Grassberger, C. Christensen, G. Bizhani, S.-W. Son, and M. Paczuski. Explosive percolation is continuous, but with unusual finite size behavior. *Physical review letters*, 106(22):225701, 2011.

[19] P. Hartman and A. Wintner. On the infinitesimal generators of integral convolutions. *American Journal of Mathematics*, 64(1):273–298, 1942.

[20] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.

[21] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.

[22] S. Janson and J. Spencer. Phase Transitions for Modified Erdos–Rényi Processes. *Arxiv preprint arXiv:1005.4494*, 2010.

[23] M. Kang, , W. Perkins, and J. Spencer. The Bohman-Frieze Process Near Criticality. *Arxiv preprint arXiv:1106.0484v1*, 2010.

[24] T. G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1977/78.

[25] R. S. Liptser and A. N. Shiryayev. *Theory of martingales*, volume 49 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1989. Translated from the Russian by K. Dzjaparidze [Kacha Dzhaparidze].

[26] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.

[27] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.

[28] F. Radicchi and S. Fortunato. Explosive percolation in scale-free networks. *Physical review letters*, 103(16):168701, 2009.

[29] M. Reed and B. Simon. *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York, 1972.

[30] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Arxiv preprint arXiv:1102.5306*, 2011.

[31] J. Spencer and N. Wormald. Birth control for giants. *Combinatorica*, 27(5):587–628, 2007.