

Bohman-Frieze processes at criticality and emergence of the giant component

Shankar Bhamidi*

Amarjit Budhiraja[†]

Xuan Wang[‡]

Department of Statistics and Operations Research,
304 Hanes Hall,
University of North Carolina,
Chapel Hill, NC.

June 6, 2011

Abstract

The evolution of the usual Erdős-Rényi random graph model on n vertices can be described as follows: At time 0 start with the empty graph, with n vertices and no edges. Now at each time k , choose 2 vertices uniformly at random and attach an edge between these two vertices. Let $\mathbf{G}_n(k)$ be the graph obtained at step k . Following [22] and [25], the work of Aldous in [4] shows that for any fixed $t \in \mathbb{R}$, when $k(n) = n/2 + n^{2/3}t/2$, the sizes of the components in $\mathbf{G}_n(k(n))$ scale like $n^{2/3}$ and, as $n \rightarrow \infty$, the scaled component size vector converges to the standard multiplicative coalescent at time t . Furthermore, when $k(n) = sn/2$, for any $s < 1$, the size of the largest component is $O(\log n)$ while for $s > 1$, the size of the largest component scales like $f(s)n$, where $f(s) > 0$. Thus $t_c = 1$ is the critical parameter for this model at which time a giant component emerges.

The last decade has seen variants of this process introduced, loosely under the name *Achlioptas processes*, to understand the effect of simple changes in the edge formation scheme on the emergence of the giant component. Stimulated by a question of Achlioptas, one of the simplest and most popular of such models is the Bohman-Frieze (BF) model wherein at each stage k , 2 edges $e_1(k) = (v_1, v_2)$ and $e_2(k) = (v_3, v_4)$ are chosen uniformly at random. If at this time v_1, v_2 are both isolated then this edge is added, otherwise e_2 is added. Then [12] (and further analysis in [40]) show that once again there is a critical parameter, which is larger than 1, above and below which the asymptotic behavior is as in the Erdős-Rényi setting, in particular this simple modification in the attachment scheme delays the time of emergence of the giant component. While an intense study for this and related models seems to suggest that at criticality, this model should be in the same universality class as the original Erdős-Rényi process, a precise mathematical treatment of the dynamics in the critical window has to date escaped analysis. In this work we study the component structure of the BF model in the critical window and show that at criticality the sizes of components properly rescaled and re-centered converge to the same limits as for the original Erdős-Rényi process. Our proofs rely on a series of careful approximations of the original random graph model leading eventually to the treatment of certain near critical multitype branching processes with general type spaces through functional and stochastic analytic tools.

Key words: critical random graphs, multiplicative coalescent, entrance boundary, giant component, branching processes, inhomogeneous random graphs.

MSC2000 subject classification. 60C05, 05C80, 90B15.

*bhamidi@email.unc.edu

†budhiraj@email.unc.edu

‡wangxuan@email.unc.edu

1 Introduction

Random graph models of various systems in the real world have witnessed a tremendous growth in the last decade. The availability of a large amount of empirical data on real world networks and systems such as road and rail networks, bio-chemical networks, social networks and data transmission networks such as the internet has stimulated an inter-disciplinary effort to formulate models to understand such systems ranging from (largely) static systems such as road and rail networks, to more dynamic systems such as social networks, information networks and models of coagulation and aggregation in physics and colloidal chemistry.

The classical Erdős-Rényi(ER) random graph can be thought of as a network evolving in time via the following prescription: Start with the empty graph at time zero. Then at each discrete time point, choose an edge uniformly at random and place it in the network (irrespective of whether it was present before or not). This simple model exhibits a tremendous amount of complexity, see [15, 27]. As one adds more and more edges, eventually the system transitions from a “sub-critical” regime wherein the largest component is of order $\log n$, to the super-critical regime wherein there exists a unique giant component of order n and the next largest component is $O(\log n)$. Understanding the emergence of this giant component and the properties of near critical Erdős-Rényi random graph have stimulated an enormous amount of work, see e.g [34, 14, 19] and the references therein.

This model has been modified in various ways the last few years, to understand the effect of choice in the emergence of the giant component. For example, in [1] and [20] simulation based studies are carried out for an attachment scheme where at each stage one chooses two edges uniformly at random, and then one uses the edge which delays the emergence of a giant component (for example by choosing the edge that minimizes the product of the two adjoining components). A number of fascinating conjectures about the “explosive emergence” of the giant component are stated in these articles.

Such modifications of the ER random graph model have generated a tremendous amount of interest in many different communities, ranging from combinatorics to statistical physics. Although a number of such models have been analyzed in the sub-critical and super-critical regime, see for example the marvelous and comprehensive [40], understanding how the giant component emerges, even for very simple modifications to the ER setting, for example for the Bohman-Frieze model (which we describe below) has escaped a mathematical understanding. The aim of this paper is to prove the first rigorous results about such models at criticality and obtain precise asymptotics for the merging dynamics in the critical window that lead to the emergence of the giant component.

The Bohman-Frieze (BF) model is as follows: Let $\mathbf{0}_n$ be the empty graph on the vertex set $[n] \stackrel{def}{=} \{1, 2, \dots, n\}$ with no edges. The random graph process $\{\mathbf{G}_n^{BF}(k)\}_{k \in \mathbb{N}} \equiv \{\mathbf{G}_n(k)\}_{k \in \mathbb{N}}$ evolves as follows:

- At time $k = 0$ let $\mathbf{G}_n(0) = \mathbf{0}_n$,
- The process evolves in discrete steps. For $k \geq 0$ let the state of the graph at time k be $\mathbf{G}_n(k)$. Then at time $k + 1$, the graph $\mathbf{G}_n(k + 1)$ is constructed as follows: Choose two edges $e_1 = (v_1, v_2)$ and $e_2 = (v_3, v_4)$ uniformly at random amongst all the possible $\binom{n}{2}$ edges. If v_1, v_2 are isolated vertices then let $\mathbf{G}_n(k + 1)$ be the graph $\mathbf{G}_n(k)$ with edge e_1 added else let $\mathbf{G}_n(k + 1)$ be the graph $\mathbf{G}_n(k)$ with the edge e_2 added.

Note that the standard Erdős-Rényi (ER) process can be thought of as a variant of the above process wherein we add edge e_1 at each stage (irrespective of whether the end points are isolated vertices or not). We shall use $\mathbf{G}_n^{ER}(k)$ to denote the usual Erdős-Rényi process.

For the standard Erdős-Rényi process classical results (see e.g. [22], [21], [15]) tell us that by scaling time by n and looking at the process $\mathbf{G}_n^{ER}(\lfloor nt/2 \rfloor)$, a phase transition occurs at $t_c(er) = 1$, namely

- (a) **Subcritical regime:** For fixed $t < 1$, the largest component in $\mathbf{G}_n^{ER}(\lfloor nt/2 \rfloor)$ is $O(\log n)$.
- (b) **Supercritical regime:** For fixed $t > 1$, the largest component has size $\sim f(t)n$ for some positive function f (namely there is a giant component) while the second largest component is of size $O(\log n)$.
- (c) **Critical regime:** For $t = 1$, the first and the second largest components both are $\Theta(n^{2/3})$.

For the BF model, the original work in [12] shows that there exists $t_0 > 1$ such that the size of the largest component in $\mathbf{G}_n^{BF}(\lfloor nt_0/2 \rfloor)$ is $o_p(n)$. Thus this model “avoids” the giant component for more time than the original Erdős-Rényi model. The comprehensive paper [40] showed for the BF model and for a number of other related processes, that there exists a critical value t_c , such that conclusions (a) and (b) hold for these models as well with 1 replaced by t_c . For the BF model, the supercritical regime, namely the regime $t > t_c \equiv t_c(bf)$ was studied further in [26]. Numerical calculations from this paper show that $t_c(bf) \approx 1.17$.

Understanding what happens in the critical regime for these models is of tremendous interest for many different fields, as it turns out that much of the “action” happens at criticality. For the Erdős-Rényi processes, Aldous’s result in [4] shows that if we denote the component size vector (components arranged in decreasing order), in the graph $\mathbf{G}_n^{ER}(\lfloor nt/2 \rfloor)$, by

$$\mathbf{C}_n^{ER}(t) = (\mathcal{C}_n^{(i)}(t) : i \geq 1),$$

then, for any fixed $\lambda \in \mathbb{R}$, as $n \rightarrow \infty$,

$$\bar{\mathbf{C}}_n^{ER}(\lambda) \stackrel{\text{def}}{=} n^{-2/3} \mathbf{C}_n^{ER} \left(1 + \frac{\lambda}{n^{1/3}} \right) \xrightarrow{d} \mathbf{X}(\lambda), \quad (1.1)$$

where the process $(\mathbf{X}(\lambda) : -\infty < \lambda < \infty)$ is a Markov process called the *eternal standard multiplicative coalescent* on the space

$$l_{\downarrow}^2 = \{(x_1, x_2, \dots) : x_1 \geq x_2 \geq \dots \geq 0, \sum_i x_i^2 < \infty\}. \quad (1.2)$$

The space l_{\downarrow}^2 is endowed with the same topology as inherited from l^2 , and the above \xrightarrow{d} denotes weak convergence in this space. In fact the paper [4] shows weak convergence of $\bar{\mathbf{C}}_n^{ER}$ to \mathbf{X} in $\mathcal{D}((-\infty, \infty), l_{\downarrow}^2)$ (the space of RCLL functions from $(-\infty, \infty)$ to l_{\downarrow}^2 endowed with the usual Skorohod topology). We defer a full definition of the above Markov process to Section 2.1.

Despite the tremendous amount of interest in the critical regime of such models and the significant progress in analyzing these models above and below criticality (see e.g. [40] for a general treatment of such models and [26] for an analysis of the BF model), analysis of the component sizes at criticality and an understanding of the dynamic behavior of these components and how they merge through the critical scaling window for the giant component to emerge, for settings beyond the classical Erdős-Rényi model has remained an extremely challenging problem. The most commonly used technique in getting refined results about the sizes of the largest component at criticality, which is to show that the breadth-first exploration processes of components converge to a certain ‘inhomogeneous Brownian motion’ does not seem to extend to such settings and one needs to develop a rather different machinery. In this work we take a significant step towards the understanding of the behavior of such models by treating the asymptotics in the critical regime for the Bohman-Frieze process. Although this model makes a simple modification to the attachment scheme for the Erdős-Rényi setting, the analysis requires many new constructions and mathematical ideas. Our main result, Theorem 2.1, shows that at criticality the sizes of components properly rescaled and re-centered converge, as for the original Erdős-Rényi process, to the eternal standard multiplicative coalescent. The proof proceeds through a sequence of approximations that reduce the study to that for a certain inhomogeneous random graph model which allows us to bring to bear techniques from the theory of multitype branching processes on general type spaces and integral operators on Hilbert spaces. These culminate in an estimate, obtained in Proposition 5.1, on the size of the largest component through the “subcritical” window. Using this estimate and other careful estimates on quadratic variations of certain martingales associated with the Bohman-Frieze process we then obtain (see Proposition 4.1) precise asymptotics on the sum of squares and cubes of component sizes just below the critical window. The latter result is a key ingredient to the proof as it allows us to show that just

before the critical window, the component sizes satisfy the regularity conditions required for convergence to the standard multiplicative coalescent, and thus allow us to couple the BF process with a standard multiplicative coalescent through the critical window to prove the main result. Although not pursued here, the techniques developed in this work for establishing the asymptotic behavior in Proposition 4.1 are potentially of use for the analysis of the critical regime for a very general class of Achlioptas processes (called bounded size rules).

Organization of the paper: Due to the length of the paper, let us now briefly guide the reader through rest of this work. We begin in Section 2 with the construction of the continuous time version of the BF-process and then state the main result (Theorem 2.1) in Section 2.1. Next, in Section 3 we give a wide ranging discussion of related work and the relevance of our results. We shall then start with the proof of the main result by giving an intuitive sketch in Section 4 wherein we also provide details on the organization of the various steps in the proof carried out in Sections 5-7. Finally Section 8 combines all these ingredients to complete the proof.

2 The Bohman-Frieze process

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 process and the convergence in (1.1) continues to hold, where $\mathbf{C}_n^{ER}(t)$ now represents the component size vector at time t for this continuous time ER process and once more, $t_c = 1$ is the critical parameter for the model.

As proved in [40], the Bohman-Frieze model also displays a phase transition and the critical time $t_c \approx 1.1763$. We now summarize some results from the latter paper that characterize this critical parameter in terms of the behavior of certain differential equations.

The following notations and definitions mostly follow [26]. Let $\mathcal{C}_n^{(i)}(t)$ denote the size of the i^{th} largest component in $\mathbf{BF}_n(t)$, and $\mathbf{C}_n(t) = (\mathcal{C}_n^{(i)}(t) : i \geq 1)$ 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

$$\mathcal{S}_k(t) = \sum_{i \geq 1} (\mathcal{C}_n^{(i)}(t))^k \tag{2.1}$$

and let $\bar{s}_k(t) = \mathcal{S}_k(t)/n$. Then from [40], there exist deterministic functions $x(t), s_2(t), s_3(t)$ such that for each fixed $t \geq 0$:

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

as $n \rightarrow \infty$. 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.

$$x'(t) = -x^2(t) - (1 - x^2(t))x(t) \quad \text{for } t \in [0, \infty,) \quad x(0) = 1 \quad (2.2)$$

$$s_2'(t) = x^2(t) + (1 - x^2(t))s_2^2(t) \quad \text{for } t \in [0, t_c), \quad s_2(0) = 1 \quad (2.3)$$

$$s_3'(t) = 3x^2(t) + 3(1 - x^2(t))s_2(t)s_3(t) \quad \text{for } t \in [0, t_c), \quad s_3(0) = 1. \quad (2.4)$$

This constant $t_c = t_c(bf)$ 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 [26] (Theorem 3.2) there exist constants

$$\alpha = (1 - \bar{x}^2(t_c))^{-1} \approx 1.063 \quad (2.5)$$

$$\beta = \lim_{t \uparrow t_c} \frac{s_3(t)}{[s_2(t)]^3} \approx .764 \quad (2.6)$$

such that as $t \uparrow t_c$

$$s_2(t) \sim \frac{\alpha}{t_c - t} \quad (2.7)$$

$$s_3(t) \sim \beta (s_2(t))^3 \sim \beta \frac{\alpha^3}{(t_c - t)^3}. \quad (2.8)$$

2.1 Main results

Our goal in this work is to establish a limit theorem of the form in (1.1) for the BF process.

We begin with a precise definition of the standard eternal multiplicative coalescent process \mathbf{X} introduced in Section 1. For $x \in l^2$, let $\text{ord}(x) \in l^2_{\downarrow}$ be a reordering of the vector x in the decreasing order. For $x \in l^2_{\downarrow}$, $1 \leq i < j < \infty$, let

$$x^{ij} = \text{ord}(x + \langle x, e_j \rangle (e_i - e_j)),$$

where $\{e_i\}$ is the canonical basis in l^2 . Namely, x^{ij} is the vector obtained by merging the i -th and j -th ‘components’ in x and reordering the vector. Aldous([4]) showed that there is a Feller Markov process with sample paths in $D([0, \infty) : l^2_{\downarrow})$ with infinitesimal generator

$$\mathcal{A}_{\text{MC}}f(x) = \sum_{i < j} x_i x_j (f(x^{ij}) - f(x)), \quad x \in l^2_{\downarrow}, \quad f : l^2_{\downarrow} \rightarrow \mathbb{R}. \quad (2.9)$$

This Markov process describes a coalescence dynamics where at any time instant the i -th and j -th clusters merge at rate equal to the product of the sizes of the two clusters. 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 $\{W(t)\}_{t \geq 0}$ be a standard Brownian motion, and for a fixed $\lambda \in \mathbb{R}$, define

$$W_{\lambda}(t) = W(t) + \lambda t - \frac{t^2}{2}, \quad t \geq 0.$$

Let \bar{W}_{λ} denote the reflected version of W_{λ} , i.e.,

$$\bar{W}_{\lambda}(t) = W_{\lambda}(t) - \min_{0 \leq s \leq t} W_{\lambda}(s), \quad t \geq 0. \quad (2.10)$$

Define an excursion of \bar{W}_λ as an interval $(l, u) \subset [0, +\infty)$ such that $\bar{W}_\lambda(l) = \bar{W}_\lambda(u) = 0$ and $\bar{W}_\lambda(t) > 0$ for all $t \in (l, u)$. Define $u - l$ as the size of the excursion. Order the sizes of excursions of \bar{W}_λ as

$$\xi_1(\lambda) > \xi_2(\lambda) > \xi_3(\lambda) > \dots$$

and write $\Xi(\lambda) = (\xi_i(\lambda) : i \geq 1)$. Then $\Xi(0)$ defines a l_\downarrow^2 valued random variable. Denote by $\{\mathbf{X}(\lambda)\}_{\lambda \geq 0}$ the l_\downarrow^2 valued Markov process with initial distribution as the probability law of $\Xi(0)$ and infinitesimal generator as in (2.9). Then [4] shows that for each $\lambda \in (0, \infty)$, $\mathbf{X}(\lambda)$ has the same law as $\Xi(\lambda)$. In fact, [4] shows that the process \mathbf{X} can be extended for $\lambda \in (-\infty, \infty)$, namely there is a process with sample paths in $\mathcal{D}((-\infty, \infty) : l_\downarrow^2)$, denoted once more as \mathbf{X} , such that for every $\lambda \in (-\infty, \infty)$, $\mathbf{X}(\lambda) =_d \Xi(\lambda)$ and $\{\mathbf{X}(\lambda + t)\}_{t \geq 0}$ is a Markov process with generator \mathcal{A}_{MC} . This description uniquely characterizes a probability measure on $\mathcal{D}((-\infty, \infty) : l_\downarrow^2)$ representing the probability law of \mathbf{X} . The process \mathbf{X} is called the *eternal standard multiplicative coalescent*.

We are now in a position to state our main result. For the rest of this work $t_c = t_c(bf)$ will denote the critical point for the continuous time Bohman-Frieze process as defined above. Recall the constants α and β defined in (2.5) and (2.6).

Theorem 2.1 *For $\lambda \in \mathbb{R}$, let*

$$\bar{\mathbf{C}}_n^{BF}(\lambda) = \left(\frac{\beta^{1/3}}{n^{2/3}} \mathcal{C}_n^{(i)} \left(t_c + \beta^{2/3} \alpha \frac{\lambda}{n^{1/3}} \right) : i \geq 1 \right) \quad (2.11)$$

be the rescaled component sizes of the Bohman-Frieze process at time $t_c + \alpha \beta^{2/3} \frac{\lambda}{n^{1/3}}$. Then

$$\bar{\mathbf{C}}_n^{BF} \xrightarrow{d} \mathbf{X}$$

as $n \rightarrow \infty$, where \mathbf{X} is the eternal standard multiplicative coalescent, and \xrightarrow{d} denotes weak convergence in the space $\mathcal{D}((-\infty, \infty) : l_\downarrow^2)$. In particular for each fixed $\lambda \in \mathbb{R}$, the rescaled component sizes in the critical window satisfy $\bar{\mathbf{C}}_n^{BF}(\lambda) \xrightarrow{d} \Xi(\lambda)$ where $\Xi(\lambda)$ are the sizes of the excursions for the reflected inhomogeneous Brownian motion in (2.10).

3 Discussion

Let us now give a wide ranging discussion of the implications of the above result and proof techniques.

- (a) **Dynamic network models:** The formulation and study of dynamic networks models, such as the Bohman-Frieze process wherein one has both randomness and choice, are relatively new. Motivated by a question by Achlioptas, that whether simple modifications of the Erdős-Rényi random graph model could allow one to avoid the giant component for a larger amount of time, such models now fall loosely under the term *Achlioptas processes* (see e.g. [12, 13, 40, 30]). The models considered in the above papers use *bounded size* rules, wherein there is a fixed K such that when making a choice between the two randomly selected edges, all component sizes greater than some K are treated identically (e.g. $K = 1$ for the model considered in this study). Much recent interest has been generated by recent models formulated in [1], wherein interesting examples of *unbounded size* rules have been analyzed. Simulations of such models seem to suggest a new phenomenon called “explosive percolation”, wherein the phase transition from a subcritical to a supercritical regime with a giant component appears much more “abruptly”, in 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. Examples of such rules include the product rule wherein one chooses two edges at random and connects the edge that *minimizes* the product of the component sizes on the 2 end points of the

edge. See [20, 18] for additional models of this type. Recently in [39] it was shown that for a number of such models, the phase transition is actually “continuous”, so for such models one expects similar behavior as what one sees in the Erdős-Rényi random graph model. However, understanding what happens at criticality and the behavior of the scaling window is a challenging math program. The techniques in this paper have the potential to be extended to the analysis of a number of such models, which we attempt to do in work in progress.

- (b) **Multiplicative coalescent:** As shown in the current work and several other papers, see e.g [7, 11, 10, 35], multiplicative coalescent arises as the limit object in a large number of random graph models at criticality. Thus asymptotics of large random graphs has intimate connections with the general theory of coalescent (and fragmentation) processes that is currently a very active area in probability, see e.g. [5, 9, 37].
- (c) **Starting from an arbitrary configuration:** In this work we have only considered the case where we start with the empty configuration $\mathbf{0}_n$. One can imagine starting from a different configuration and then attempt to analyze the emergence of the giant component. Along the lines of [6], under the assumption that starting configuration satisfies suitable regularity conditions, it should be possible to study the asymptotic behavior of this model in terms of the entrance boundary of the standard multiplicative coalescent. This will be pursued in future work.
- (d) **Proof techniques:** The study of critical random graphs have generated an enormous amount of interest in the probabilistic combinatorics community and a large number of techniques have been developed for the fine scale asymptotics of such models, ranging from generating function and counting arguments (see e.g. [25], [15]); branching processes exploring local neighborhoods of the graph (see e.g [28], and the references therein, [33]); multi-type branching processes (see e.g. [16] and the references therein); differential equation based techniques (see e.g. [40], [12] and for a comprehensive survey [41]); and the the breadth first walk approach coupled with the martingale central limit theorem (see e.g. [29], [32] for one of the first few studies using this technique; [4] where this was used to explore the fine scale structure of the classical Erdős-Rényi random graph at criticality and connections with the multiplicative coalescent; see [35], [11] for further results using this technique). Our methods are inspired by [7] which uses estimates on the size of largest component in the subcritical window to analyze the asymptotic behavior of the the sum of squares and cubes of component sizes near criticality. Estimates on the latter in turn allow the verification of the sufficient conditions given in [4] for convergence to the standard multiplicative coalescent. Although our general outline of the proof is similar to [7] as described above, it turns out that for general Achlioptas processes, completing this program is a rather challenging problem and for the BF model treated in the current work we need to develop quite a bit of machinery, which is done in Sections 5 and 6, in order to obtain the required estimates.
- (e) **Other structural properties of components:** In this study we focus on the sizes of the components in the critical scaling window. Developing probabilistic methodology for understanding the actual structure of the large components is also of great interest, see for example [17] for one of the first studies at a non-rigorous level, exploring the connection between the structure of these components at criticality and the internal structure of the minimal spanning tree in various random graph models (the “strong disorder” regime in statistical physics). In particular, it was conjectured in [17] that the diameter of the minimal spanning trees in various random graph models scales like $n^{1/3}$. Rigorous studies have now been carried out for the Erdős-Rényi random graph model and fascinating connections have been discovered between the critical random graphs and the famous continuum random tree of Aldous (see [3], [2]). Proving such structural convergence of the entire components in dynamic network settings such as Achlioptas process at criticality, to random fractals such as continuum random trees, would be of great interest in a number of different fields.

4 Proof idea

Let us now give an idea of the proof. We begin by showing in Proposition 4.1 below that, just before the critical window, the configuration of the components satisfies some important regularity properties. This proposition will be used in Section 8 in order to apply a result of [4] that gives sufficient conditions for convergence to the multiplicative coalescent.

Proposition 4.1 *Let $\gamma \in (1/6, 1/5)$ and define $t_n = t_c - n^{-\gamma}$. Then we have*

$$\frac{n^2 \mathcal{S}_3(t_n)}{\mathcal{S}_2^3(t_n)} \xrightarrow{\mathbb{P}} \beta \quad (4.1)$$

$$\frac{n^{4/3}}{\mathcal{S}_2(t_n)} - \frac{n^{-\gamma+1/3}}{\alpha} \xrightarrow{\mathbb{P}} 0 \quad (4.2)$$

$$\frac{n^{2/3} \mathcal{C}_n^{(1)}(t_n)}{\mathcal{S}_2(t_n)} \xrightarrow{\mathbb{P}} 0. \quad (4.3)$$

Now note that t_n can be written as

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

where

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

as $n \rightarrow \infty$. 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 [4]).

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 [7] will complete the proof of the main result.

Let us now outline the framework of the proof:

- The bound on the largest component $\mathcal{C}_n^{(1)}(t)$ when $t \uparrow t_c$ (Proposition 5.1) plays a crucial role in proving the statements in Proposition 4.1. In order to achieve this, we introduce a series of related models from Section 5.2 through Section 5.4.
- Section 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 5.1.
- Section 7 uses the bounds on the largest component from Proposition 5.1 to analyze the sum of squares and cubes of component sizes near the critical window. As one can imagine when working so close to the critical window, one needs rather careful estimates. Our arguments are based on a precise analysis of quadratic variation processes for certain martingales associated with the BF model. This section completes the proof of Proposition 4.1.
- Finally, in Section 8 we use Proposition 4.1 and a coupling with the standard multiplicative coalescent, in a manner similar to [7], to prove the main result.

Without further ado let us now start with the proofs.

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 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 $B \equiv B(\gamma) \in (0, \infty)$,*

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

where

$$m(n, t) = B \frac{(\log n)^4}{(t_c - t)^2}. \quad (5.1)$$

The proof of Proposition 5.1 will be completed in Section 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 5.1 we give the basic notation and mathematical conventions used in this paper.
- In Section 5.2 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 5.3 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 6.3.
- In Section 5.4 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 5.5 we summarize connections between the various models introduced above.

5.1 Notation

5.1.1 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 $[n] = \{1, 2, \dots, n\}$. 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 work are given on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which will usually be suppressed in our proofs.

5.1.2 Probability and analysis

All the unspecified limits are taken as $n \rightarrow +\infty$. Given a sequence of events $\{E_n\}_{n \geq 1}$, we say E_n (or E) occurs with high probability (whp) if $\mathbb{P}\{E_n\} \rightarrow 1$. For functions $f, g : \mathbb{N} \rightarrow \mathbb{R}$, 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 work are $\mathcal{D}([0, T] : \mathbb{R})$ and $\mathcal{D}([0, T] : \mathbf{G})$ with the natural associated partial ordering.

5.1.3 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 work $T = 2t_c$ which is a convenient upper bound for the time parameters of interest.

We use d_1, d_2, \dots 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, \dots for constants that appear in the statement of theorems.

5.2 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 $\mathbf{BF}_n(t)$, $\mathbf{C}_n(t) = (\mathcal{C}_n^{(i)}(t) : i \geq 1)$ 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

$$\frac{2}{n^3} \left[\binom{X_n(t)}{2} \binom{n}{2} + \left(\binom{n}{2} - \binom{X_n(t)}{2} \right) \binom{X_n(t)}{2} \right] \stackrel{def}{=} n \cdot a_n^*(\bar{x}(t)). \quad (5.2)$$

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 $a_0 : [0, 1] \rightarrow [0, 1]$ as

$$a_0(y) = 2 \left(\frac{y^2}{2} \cdot \frac{1}{2} + \left(\frac{1}{2} - \frac{y^2}{2} \right) \frac{y^2}{2} \right) = \frac{1}{2}(y^2 + (1 - y^2)y^2). \quad (5.3)$$

It is easy to check that

$$a_n^*(\bar{x}(t)) = a_0(\bar{x}(t)) + r_a(t), \text{ where, } \sup_t |r_a(t)| \leq 5/n. \quad (5.4)$$

Recall that $x(t)$ is the solution of the differential equation (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 $\sup_{t \leq T} |\bar{x}_n(t) - x(t)| \rightarrow 0$ in probability (see Lemma 6.4 of this paper, also see [40]). This in particular implies that $\sup_{t \leq T} |a_n^*(t) - a_0(t)| \rightarrow 0$ 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

$$\frac{2}{n^3} \left(\binom{n}{2} - \binom{X_n(t)}{2} \right) (n - X_n(t)) X_n(t). \quad (5.5)$$

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

$$\frac{2}{n^3} \left(\binom{n}{2} - \binom{X_n(t)}{2} \right) X_n(t) \stackrel{def}{=} c_n^*(\bar{x}(t)). \quad (5.6)$$

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 $c_0 : [0, 1] \rightarrow [0, 1]$ as

$$c_0(y) = (1 - y^2)y, \quad y \in [0, 1]. \quad (5.7)$$

Then

$$c^*(\bar{x}(t)) = c_0(\bar{x}(t)) + r_c(t) \quad \text{and} \quad \sup_t |r_c(t)| \leq 2/n. \quad (5.8)$$

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

$$\frac{2}{n^3} \left(\binom{n}{2} - \binom{X_n(t)}{2} \right) \binom{n - X_n(t)}{2}. \quad (5.9)$$

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

$$\frac{2}{n^3} \left(\binom{n}{2} - \binom{X_n(t)}{2} \right) \stackrel{def}{=} \frac{1}{n} b_n^*(\bar{x}(t)). \quad (5.10)$$

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

$$b_0(y) = 1 - y^2, \quad y \in [0, 1]. \quad (5.11)$$

Then

$$b^*(\bar{x}(t)) = b_0(\bar{x}(t)) + r_b(t) \quad \text{and} \quad \sup_t |r_b(t)| \leq 2/n. \quad (5.12)$$

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 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 $\mathcal{COM}(t)$.

5.3 The random graph process with immigrating vertices and attachment

In this subsection, we introduce a random graph process with immigrating vertices and attachment (RGIVA). This construction is inspired by [7] 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 $[0, T] \rightarrow [0, 1]$ (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 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 [7]. We note that the above construction closely follows our analysis of three types of events in Section 5.2, 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 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 $\{a^*(t) \leq a_U(t), b^*(t) \leq b_U(t), c^*(t) \leq c_U(t) \text{ for all } t \in [0, T]\}$ and $L \equiv L_n$ be the event that $\{a^*(t) \geq a_L(t), b^*(t) \geq b_L(t), c^*(t) \geq c_L(t) \text{ for all } t \in [0, T]\}$. Define for $t \in [0, T]$

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

Then

- (i) Upper bound: $\mathcal{COM}_n^U \leq_d \mathbf{IA}_n^U \equiv \mathbf{IA}_n(a_U, b_U, c_U)$.
- (ii) Lower bound: $\mathcal{COM}_n^L \geq_d \mathbf{IA}_n^L \equiv \mathbf{IA}_n(a_L, b_L, c_L)$.

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 $\tau = \inf\{0 \leq t \leq T : a^*(t) > a_U(t) \text{ or } b^*(t) > b_U(t) \text{ or } c^*(t) > c_U(t)\}$ and set $\mathcal{COM}_n^U(t)$ to be the null graph whenever $t \geq \tau$. This construction defines a coupling of \mathbf{IA}_n^U and \mathcal{COM}_n^U such that $\mathcal{COM}_n^U \leq \mathbf{IA}_n^U$ a.s. The result follows. ■

5.4 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 [16], which our presentation largely follows. We generalize the setting of [16] 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 $\kappa : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty)$. 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 (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 $(\frac{1}{n} \cdot \kappa_n(x, y)) \wedge 1$.

One can similarly define an IRG model 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 operator $\mathcal{K} : L^2(\mu) \rightarrow L^2(\mu)$ defined as

$$\mathcal{K}f(x) = \int_{\mathcal{X}} \kappa(x, y)f(y)\mu(dy) \tag{5.13}$$

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 $s \leq u \leq t$ 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

$$\mathcal{A}(u)f(r) = c(u)r \cdot (f(r+1) - f(r)), \quad f : \mathbb{N} \rightarrow \mathbb{R}, s \leq u \leq t. \quad (5.14)$$

We set $w(u) = 0$ for $0 \leq u < s$ 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,

$$\mathbb{P}\{\mathbf{x} \text{ and } \mathbf{y} \text{ is connected} \mid \text{Stage I}\} = 1 - \exp\left\{-\frac{1}{n} \int_0^t w(u)\tilde{w}(u)b(u)du\right\} \quad (5.15)$$

$$\leq \frac{1}{n} \int_0^t w(u)\tilde{w}(u)b(u)du. \quad (5.16)$$

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

$$\phi_t(\mathbf{x}) = w(t), \quad \mathbf{x} = (s, w) \in [0, T] \times \mathcal{W}. \quad (5.17)$$

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 (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,

$$\int_{\mathcal{X}} f(\mathbf{x}) d\mu_t(\mathbf{x}) = \int_0^t a(s) \left(\int_{\mathcal{W}} f(s, w) d\nu_s(w) \right) ds. \quad (5.18)$$

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

$$\kappa_{n,t}(\mathbf{x}, \mathbf{y}) = \kappa_{n,t}((s, w), (r, \tilde{w})) = n \left(1 - \exp\left\{-\frac{1}{n} \int_0^t w(u) \tilde{w}(u) b(u) du\right\} \right). \quad (5.19)$$

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

$$\kappa_t(\mathbf{x}, \mathbf{y}) = \kappa_t((s, w), (r, \tilde{w})) = \int_0^t w(u) \tilde{w}(u) b(u) du. \quad (5.20)$$

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. $\mathcal{C}^{(0)}(\kappa_t, \mu_t, \phi_t)$]. Define $\mathcal{C}^{(1)}(\kappa_{n,t}, \mu_t, \phi_t)$, $\mathcal{C}^{(0)}(\kappa_{n,t}, \mu_t, \phi_t)$ in a similar fashion. The following is an immediate consequence of the above construction.

Lemma 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

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

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

5.5 A summary of the models

As noted earlier, the key step in the proof of Proposition 4.1 is a good estimate on the size of the largest component in the Bohman-Frieze process $\mathbf{BF}_n(t)$ as in Proposition 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 (5.3), (5.11), (5.7). More precisely we shall show that as $n \rightarrow \infty$, 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 5.2.

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

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

Here $\kappa_{n,t}, \kappa_t, \mu_t, \phi_t$ and a, b, c are related through (5.19), (5.20), (5.18) (see also (5.14)), (5.17), respectively.

6 Analysis of the largest component at sub-criticality: Proof of Proposition 5.1

This section proves Proposition 5.1. The section is organized as follows:

- In Section 6.1 we reduce the problem to proving Proposition 6.3. We give the proof of Proposition 5.1 using this result. Rest of Section 6 is devoted to the proof of Proposition 6.3.
- In preparation for this proof, in Section 6.2 we present some key lemmas that allow us to estimate the errors between various models summarized in Section 5.5. Proofs of some lemmas (Lemmas 6.6, 6.9 and 6.10) are deferred to later sections.
- Using these lemmas, in Section 6.3 we prove the key proposition, Proposition 6.3. The rest of Section 6 proves the supporting Lemmas 6.6, 6.9 and 6.10.
- In Section 6.4 we introduce a branching process related to the IRG model, and prove Lemma 6.6. A key step in the proof is Lemma 6.13 whose proof is left to Section 6.5.
- Section 6.6 analyzes the kernel κ_t associated with $\mathbf{RG}_{n,t}(a, b, c)$ and proves Lemma 6.9 .
- Finally, in Section 6.7 we give the proof of Lemma 6.10.

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 5.1 to an estimate on the *first* component as in Proposition 6.3. This reduction, although somewhat different, is inspired by a similar idea used in [7].

Recall that $\mathcal{C}_n^{(1)}(t) \equiv I_n(t)$ denotes the largest component in $\mathbf{BF}_n(t)$. Let $\mathcal{C}_n^s(t)$, $0 \leq s \leq t$, 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 6.1 *For any $n \in \mathbb{N}$, $t_0 \in [0, T]$ and deterministic function $\alpha : [0, T] \rightarrow [0, \infty)$*

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

Proof: Let $\{\mathbf{BF}_n^{(i)}(t), t \geq 0\}_{i \in \mathbb{N}_0}$ be an i.i.d. family of $\{\mathbf{BF}_n(t), t \geq 0\}$ 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

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

and let $I_n^{\mathbf{G}}(t)$ denote the size of the largest component in \mathbf{G}_n^t . Then since $a_n^*(t) \leq 1$ 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 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 $\alpha : [0, T] \rightarrow [0, \infty)$*

$$\mathbb{P}\{\mathcal{C}_n^0(t) > 2\alpha(t), \text{ for some } 0 < t \leq t_0\} \leq 16nT^2 \sup_{0 \leq s \leq t_0} \{\alpha(s) \mathbb{P}\{\mathcal{C}_n^0(s) > \alpha(s)\}\}. \quad (6.1)$$

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

$$\mathbb{P}\{\mathcal{C}_n^0(t) > 2\alpha(t) \text{ for some } t \in [0, t_0]\} = \mathbb{P}\{\tau \leq t_0\}. \quad (6.2)$$

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 $a_n^*(s) \leq 2$, the rate at which doubletons are born can be bounded by $2n$.
- Given a doubleton was born at instant s , the event $\{\mathcal{C}_n^0 \leftrightarrow_u \mathcal{C}_n^s, \text{ for some } u \in (t, t + dt)\}$ occurs, conditionally on \mathcal{F}_t , with probability $\frac{1}{n} \mathcal{C}_n^0(t) \mathcal{C}_n^s(t) b_n^*(t) dt$. This probability, using the fact that $b_n^*(s) \leq 2$ and $\mathcal{C}_n^s(t) \leq n$, on the event $\{\mathcal{C}_n^0(t) < 2\alpha(t)\}$ is bounded by $4\alpha(t) dt$.
- $\mathbb{P}\{\mathcal{C}_n^0(t) + \mathcal{C}_n^s(t) \geq 2\alpha(t)\}$ is bounded by $2\mathbb{P}\{\mathcal{C}_n^0(t) \geq \alpha(t)\}$.

Using these observations we have the following estimate

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

Result follows on combining this estimate with (6.2). ■

The following proposition will be proved in Section 6.3.

Proposition 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 $n \geq N_1$*

$$\mathbb{P}\{\mathcal{C}_n^0(t) \geq m(n, t)/2\} \leq Cn^{-\eta} \text{ for all } 0 < t < t_c - n^{-\gamma}, \quad (6.3)$$

where $m(n, t)$ is as defined in (5.1).

Remark: Intuitively, one has that in the subcritical regime, i.e. when $t < t_c$, $\mathbb{P}\{\mathcal{C}_n^0(t) > m\} < d_1 e^{-d_2 m}$ for some constants d_1, d_2 . This suggests a bound as in (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, $d_2(t) \rightarrow 0$ 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 $n \rightarrow \infty$. This analysis is substantially more delicate as will be seen in subsequent sections.

We now prove Proposition 5.1 using the above results.

Proof of Proposition 5.1: Fix $\gamma \in (0, 1/5)$ and fix $\eta > 2 + 2\gamma$. Let B, C, N_1 be as determined in Proposition 6.3 for this choice of η, γ and let $m(n, t)$ be as defined in (5.1). Without loss of generality we can assume that $N_1 \geq N_0$ where N_0 is as in Lemma 6.2. Then applying Lemmas 6.1 and 6.2 with $t_0 = t_c - n^{-\gamma}$ and $\alpha(t) = m(n, t)$, we have

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

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

6.2 Some Preparatory Results

This section collects some results that are helpful in estimating the errors between various models described in Section 5.5.

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 (2.2).

Lemma 6.4 *For any $T > 0$, there exists a $C(T) \in (0, \infty)$ such that, for all $\gamma_1 \in [0, 1/2)$,*

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

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 5.2 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 (5.2)) for this case are

$$\begin{aligned} f_{-2}^*(y) &= a_n^*(y) \\ f_{-2}(y) &= a_0(y) = \frac{1}{2}(y^2 + (1 - y^2)y). \end{aligned}$$

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 (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 $0 \leq f_l^*(\bar{x}) \leq 1$ 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

$$|f_{-1}^*(y) - f_{-1}(y)| \leq \frac{2}{n} \quad |f_{-2}^*(y) - f_{-2}(y)| \leq \frac{5}{n}, \quad \text{for all } y \in [0, 1]. \quad (6.4)$$

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. [31], [23]) as the unique solution of the stochastic equation

$$\bar{x}(t) = 1 - \frac{1}{n}Y_{-1} \left(n \int_0^t f_{-1}^*(\bar{x}(s))ds \right) - \frac{2}{n}Y_{-2} \left(n \int_0^t f_{-2}^*(\bar{x}(s))ds \right). \quad (6.5)$$

By Equation (2.2), the limiting function $x(\cdot)$ is the unique solution of the integral equation

$$x(t) = 1 - \int_0^t f_{-1}(x(s))ds - \int_0^t 2f_{-2}(x(s))ds. \quad (6.6)$$

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

$$|(f_{-1}(y) + 2f_{-2}(y)) - (f_{-1}(z) + 2f_{-2}(z))| \leq 6|y - z|. \quad (6.7)$$

Using (6.6) and (6.5) we get

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

where

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

and by (6.4)

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

and finally by (6.7)

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

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. [23], 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 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 (5.13)) and their norms by $\rho(\kappa), \rho(\kappa')$ respectively. Then*

- (i) \mathcal{K} is a compact operator. In particular $\rho(\kappa) = \|\mathcal{K}\| \leq \|\kappa\|_2 = \left(\int_{\mathcal{X} \times \mathcal{X}} \kappa^2(x, y) \mu(dx) \mu(dy) \right)^{1/2} < \infty$.
- (ii) If $\kappa \leq \kappa'$, then $\rho(\kappa) \leq \rho(\kappa')$.
- (iii) $\rho(\kappa + \kappa') \leq \rho(\kappa) + \rho(\kappa')$ and $\rho(t\kappa) = t\rho(\kappa)$ for $t \geq 0$.
- (iv) $|\rho(\kappa) - \rho(\kappa')| \leq \rho(|\kappa - \kappa'|)$.
- (v) $\rho(\kappa) \leq \|\kappa\|_\infty \mu(\mathcal{X})$.

Proof: (i) is a standard result, see Theorem VI.23 of [38].

(ii) For any nonnegative f in $L^2(\mu)$, $\mathcal{K}f(x) \leq \mathcal{K}'f(x)$ pointwise. Thus for such f , $\|\mathcal{K}f\|_2 \leq \|\mathcal{K}'f\|_2$. 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 $\{f \in L^2 : \|f\|_2 = 1, f \geq 0\}$.

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

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

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

■

We now present some auxiliary estimates for the IRG model from Section 5.4. The following lemma will be proved in Section 6.4. Recall the definition of a *basic structure* from Section 5.4.

Lemma 6.6 *Let $(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi$ be a basic structure, where κ is symmetric. Suppose that μ is nonatomic and $\rho(\kappa) = \|\mathcal{K}\| < 1$. For fixed $x_0 \in \mathcal{X}$, denote by $\mathcal{C}_n^{\text{RG}}(x_0)$ the volume of the component of $\mathbf{RG}_n(\kappa)$ that contains x_0 . Define $\mathcal{C}_n^{\text{RG}}(x_0) = 0$ if x_0 is not a vertex in $\mathbf{RG}_n(\kappa)$. Then for all $m \in \mathbb{N}$*

$$\mathbb{P}\{\mathcal{C}_n^{\text{RG}}(x_0) > m\} < 2 \exp\{-C_1 \Delta^2 m\} \quad (6.8)$$

where

$$\Delta = 1 - \rho(\kappa), \quad C_1 = \frac{1}{8\|\phi\|_\infty(1 + 3\|\kappa\|_\infty \mu(\mathcal{X}))}. \quad (6.9)$$

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 (5.20) and (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 (5.17).

Lemma 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 (5.18). Let \mathcal{P}_n be a Poisson point process on $(\mathcal{X}, \mathcal{T})$ with intensity $n \cdot \mu_t$. Define*

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

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 $\int_0^t na(s)ds \leq nT$. Let $\{Z_2^{(i)}\}_{i \geq 1}$ 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

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

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.[36]) imply that $Z_2^{(i)}(T)$ is distributed as sum of two independent $\text{Geom}\{e^{-T}\}$. 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 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 5.2, 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 (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 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*

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

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

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

The proof of the above lemma is quite technical and deferred to Section 6.6.

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

The proof of the lemma is given in Section 6.7.

6.3 Proof of Proposition 6.3

This section is devoted to the proof of Proposition 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 $E_n = \{\sup_{0 \leq t \leq T} |\bar{x}_n(t) - x(t)| \leq n^{-\gamma_1}\}$.

From Lemma 6.4,

$$\mathbb{P}\{E_n^c\} \leq \exp\{-C(T)n^{1-2\gamma_1}\} = \exp\{-C(T)n^{1/5}\}. \quad (6.10)$$

From (5.4) and recalling that the Lipschitz norm of a_0 is bounded by 2 (see (5.3)), we have that on E_n

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

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 5.2, we have for any $m \in \mathbb{N}$

$$\mathbb{P}\{\mathcal{C}_{n,\delta}^0(t) > m, E_n\} \leq \mathbb{P}\{\mathcal{C}_{n,\delta}^{IA}(t) > m, E_n\} \leq \mathbb{P}\{\mathcal{C}_{n,\delta}^{IA}(t) > m\}. \quad (6.11)$$

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 IRG 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 (5.18). Let $A_n = C_1(\eta) \log n$, where $C_1(\eta)$ is as in Corollary 6.8. Consider the following truncation of the kernel $\kappa_{t,\delta}$ and weight function $\phi_t(s, w) = w(t)$:

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

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 (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 6.8

$$\mathbb{P}\{Y_{n,\delta} > A_n\} < C_2(\eta) n^{-\eta}. \quad (6.12)$$

Let $\mathcal{C}_{n,\delta}^{RG}(t) = \mathcal{C}_{n,t}^{RG}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ 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 5.3

$$\mathbb{P}\{\mathcal{C}_{n,\delta}^{IA}(t) > m\} \leq \mathbb{P}\{\mathcal{C}_{n,\delta}^{RG}(t) > m\}. \quad (6.13)$$

Letting $\mathcal{C}_{n,\delta,A}^{RG}(t)$ 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

$$\begin{aligned} \mathbb{P}\{\mathcal{C}_{n,\delta}^{RG}(t) > m\} &\leq \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}_{n,\delta}^{RG}(t) > m, Y_{n,\delta} \leq A_n\} \\ &= \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}_{n,\delta,A}^{RG}(t) > m, Y_{n,\delta} \leq A_n\} \\ &\leq \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}_{n,\delta,A}^{RG}(t) > m\} \end{aligned} \quad (6.14)$$

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

We will apply Lemma 6.6, replacing $\{(\mathcal{X}, \mathcal{T}, \mu), \kappa, \phi\}$ there by $\{(\mathcal{X}, \mathcal{T}, \mu_{t,\delta}), \kappa_{t,\delta,A}, \phi_{t,A}\}$, where $t \in (0, t_c - n^{-\gamma})$. From (6.8) we have

$$\mathbb{P}\{\mathcal{C}_{n,\delta,A}^{RG}(t) > m\} \leq 2 \exp\{-C_1 \Delta^2 m\}, \quad (6.15)$$

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 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 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 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 $d_4 \in (0, \infty)$,

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

Using this estimate in (6.15) and recalling that $\|\phi_{t,A}\|_\infty \leq A_n$, $\|\kappa_{t,\delta,A}\|_\infty \leq T A_n^2$, we have that for some $d_5 \in (0, \infty)$

$$\mathbb{P}\{\mathcal{C}_{n,\delta,A}^{RG}(t) > m\} \leq 2 \exp\left\{-\frac{d_5}{(\log n)^3} (t_c - t)^2 m\right\}, \text{ for all } m \in \mathbb{N}, t \in (0, t_c - n^{-\gamma}) \text{ and } n \geq n_2. \quad (6.16)$$

Step 4: Collecting estimates:

Combining (6.10), (6.11), (6.13), (6.12), (6.14) and (6.16), we have

$$\begin{aligned} \mathbb{P}\{\mathcal{C}_n^0(t) > m\} &\leq \mathbb{P}\{E_n^c\} + \mathbb{P}\{Y_{n,\delta} > A_n\} + \mathbb{P}\{\mathcal{C}_{n,\delta,A}^{RG}(t) > m\} \\ &\leq e^{-C(T)n^{1/5}} + C_2(\eta)n^{-\eta} + 2 \exp\left\{-d_5 \frac{(t_c - t)^2}{(\log n)^3} m\right\}. \end{aligned} \quad (6.17)$$

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 6.10. We will use notation and arguments similar to that in the proof of Proposition 6.3 above.

Lemma 6.11 *Let (a, b, c) be rate functions. Fix $t \in [0, T]$. Let $I_n^{IA}(t)$ denote the largest component in $\mathbf{IA}_n(a, b, c)_t$. Suppose that $\rho_t(a, b, c) < 1$. Then for some $C_0 \in (0, \infty)$*

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

Proof: Let $\mathcal{C}_n^{IA}(t)$ be the first component of $\mathbf{IA}_n(a, b, c)_t$. Then an elementary argument (cf. proof of Lemma 6.1) shows that for $m > 0$

$$\mathbb{P}\{I_n^{IA}(t) > m\} \leq T n \mathbb{P}\{\mathcal{C}_n^{IA}(t) > m\}.$$

By an argument as in (6.14), we have

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

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 6.3, with $(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ replaced with (a, b, c) . From Corollary 6.8 we can find $d_1 \in (0, \infty)$ such that $\mathbb{P}(Y_n \geq d_1 \log n) = O(n^{-2})$. Let $A_n = d_1 \log n$. Then, recalling that $\rho_t(a, b, c) < 1$, we gave by Lemma 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. ■

6.4 Proof of Lemma 6.6: A branching process construction

The key idea in the proof of Lemma 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 6.6 to a similar bound on the total volume of the branching process (Lemma 6.13). We refer the reader to [16] 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 6.6 using Lemma 6.13. Proof of the latter result is given in Section 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 5.4. 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 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 6.5 shows that the estimate in Lemma 6.6 holds with $\mathcal{C}_n^{RG}(x_0)$ replaced by $G(x_0)$.

Lemma 6.13 *Suppose that $\rho(\kappa) = \|\mathcal{K}\| < 1$. Then for all $m \in \mathbb{N}$*

$$\mathbb{P}\{G > m\} < 2 \exp\{-C_1 \Delta^2 m\} \tag{6.18}$$

where Δ and C_1 are as in (6.9).

Using the above lemmas we can now complete the proof of Lemma 6.6.

Proof of Lemma 6.6: Proof is immediate from Lemmas 6.13 and 6.12. ■

We conclude this section with the proof of Lemma 6.12.

Proof of Lemma 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 $R_j, U_j, j = 0, 1, \dots, m-1$ and $U_{m-1} = \{y_m, y_{m+1}, \dots, y_{t_m}\}$. Then set

$$\begin{aligned} R_m &= R_{m-1} \cup \{y_m\} \\ U_m &= U_{m-1} \cup E_m \setminus \{y_m\} \end{aligned}$$

where

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

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 $y_{t_m+1}, y_{t_m+2}, \dots, y_{t_m+|E_m|}$. 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 $\beta_m : \mathcal{X} \rightarrow [0, 1]$ is given as $\beta_1 \equiv 1$ and, for $m > 1$,

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

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

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

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

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

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. ■

6.5 Proof of Lemma 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 6.14 *Let $g : \mathcal{X} \rightarrow \mathbb{R}_+$ be a bounded measurable map. Fix $\delta > 0$ and let $0 < \epsilon < \log(1 + \delta) / \|g\|_\infty$. Then*

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

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

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

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 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 $\epsilon \in (0, \frac{\log(1 + \delta)}{\|\phi_0\|_\infty})$*

$$\mathbb{E} \exp\left\{\epsilon \sum_{i=0}^k G_i\right\} \leq \exp\left\{\epsilon[\phi(x_0) + \sum_{i=1}^k (1 + \delta)^i \mathcal{K}^i \phi(x_0)]\right\} = \exp\{\epsilon \phi_0(x_0)\}. \quad (6.19)$$

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

$$\mathbb{E}[\exp\{\epsilon \sum_{i=l}^k G_i\} | \mathcal{F}_l] \leq \exp\left\{\epsilon \sum_{i=1}^{N_l} \phi_l(\xi_i^{(l)})\right\}. \quad (6.20)$$

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

When $l = k$, (6.20) is in fact an equality, and so (6.20) holds trivially for k . Suppose now that (6.20) is true for $l + 1$, for some $l \in \{0, 1, \dots, k - 1\}$. 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 (6.20) holds for $l + 1$ and for the second inequality we have applied Lemma 6.14 along with the observation that $\epsilon \|\phi_l\|_\infty < \log(1 + \delta)$ holds for all $l = 1, 2, \dots, k$, 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. \blacksquare

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 $\phi_\delta^* = \lim_{k \rightarrow \infty} \phi_\delta^{(k)}$. The following corollary follows on sending $k \rightarrow \infty$ in (6.19).

Corollary 6.16 *Fix $\delta > 0$ and $\epsilon \in (0, \log(1 + \delta) / \|\phi_\delta^*\|_\infty)$. Then*

$$\mathbb{E}\{\exp \epsilon G\} \leq \exp\{\epsilon \phi_\delta^*(x_0)\}. \quad (6.21)$$

Lemma 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

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

\blacksquare

Now we can finish the proof of Lemma 6.13.

Proof of Lemma 6.13: Observing that $\|\phi\|_2 \leq \|\phi\|_\infty \mu(\mathcal{X})^{1/2}$ 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 $x \in \mathcal{X}$

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

where the first inequality above follows from Lemma 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

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

Let $\epsilon = \log(1 + \delta)/(2d_1)$. Clearly $\epsilon \in (0, \log(1 + \delta)/\|\phi_\delta^*\|_\infty)$. Using Corollary 6.16 we now have that

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

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

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

The result follows. ■

6.6 Proof of Lemma 6.9

We begin with a general result for integral operators on general type spaces.

Lemma 6.18 *Let ν, μ be two mutually absolutely continuous finite measures on a measure space $(\mathcal{X}, \mathcal{T})$. Let $g = d\nu/d\mu$. Let $\kappa : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$ be a kernel. Define another kernel $\kappa' : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}_+$ as*

$$\kappa'(x, y) = \sqrt{\frac{g(x)}{g(y)}} \kappa(x, y), \quad x, y \in \mathcal{X}.$$

Denote by \mathcal{K} [resp. \mathcal{K}'] the integral operator associated with κ [resp. κ'] on $L^2(\mathcal{X}, \mathcal{T}, \nu)$ [resp. $L^2(\mathcal{X}, \mathcal{T}, \mu)$]. Then $\|\mathcal{K}\|_{L^2(\nu)} = \|\mathcal{K}'\|_{L^2(\mu)}$.

Proof: Note that the operator $\mathcal{A} : L^2(\mathcal{X}, \mathcal{T}, \nu) \rightarrow L^2(\mathcal{X}, \mathcal{T}, \mu)$ defined as $(\mathcal{A}f) = \sqrt{g}f$, $f \in L^2(\nu)$, is an isometry. Also, for $f \in L^2(\mu)$

$$(\mathcal{A}\mathcal{K}\mathcal{A}^{-1}f)(x) = \sqrt{g(x)} \int_{\mathcal{X}} \kappa(x, y) \frac{1}{\sqrt{g(y)}} f(y) \mu(dy) = \int_{\mathcal{X}} \kappa'(x, y) f(y) \mu(dy).$$

Thus $\mathcal{AKA}^{-1} = \mathcal{K}'$. The result now follows on noting that $\|\mathcal{AKA}^{-1}\|_{L^2(\mu)} = \|\mathcal{K}\|_{L^2(\nu)}$. ■

For the rest of this subsection we will take $(\mathcal{X}, \mathcal{T})$ to be the cluster space introduced in Section 5.4 (see above (5.17)). Given rate functions (a, b, c) , $\mu_t(a, b, c)$, $\kappa_t(a, b, c)$, $\phi_t(a, b, c)$, $\mathcal{K}_t(a, b, c)$, $\rho_t(a, b, c)$ are as introduced above Lemma 6.9.

Lemma 6.19 *Let (a_i, b_i, c_i) , $i = 1, 2$ be two sets of rate functions. Suppose that a_1, c_1 are strictly positive on $(0, T]$ and $a_2 \leq a_1, c_2 \leq c_1$ on $[0, T]$. Also suppose that for some $\delta \in (0, e^{-T}/T)$, $c_1 \leq c_2 + \delta$ on $[0, T]$. Fix $t \in [0, T]$. Let $\mu_i = \mu_t(a_i, b_i, c_i)$, $i = 1, 2$. Then $\mu_2 \ll \mu_1$ and*

$$\frac{d\mu_2}{d\mu_1}(s, w) = \frac{a_2(s)}{a_1(s)} \times \exp\left\{-\int_s^T w(u)[c_2(u) - c_1(u)]du\right\} \times \prod_{i=1}^{w(T)-2} \left(\frac{c_2(\tau_i)}{c_1(\tau_i)}\right), \quad (s, w) \in [0, t] \times \mathcal{W}.$$

where $\tau_i = \tau_i(s, w)$ is (μ_1 a.s.) the i^{th} jump of w after time s .

Proof: Recall the probability measure ν_s on \mathcal{W} introduced below (5.17). Write $\nu_s^i = \nu_s(a_i, b_i, c_i)$, $i = 1, 2$. Note that (see (5.18)), for $s \in [0, t]$, $\mu_i(ds dw) = \nu_s^i(dw)a_i(s)ds$, $i = 1, 2$. Thus to prove the result it suffices to show that, for all $s \in [0, t]$, $\nu_s^2 \ll \nu_s^1$ and $\frac{d\nu_s^2}{d\nu_s^1} = L_s^T$, where, for $t \in [s, T]$,

$$L_s^t(w) = \exp\left\{-\int_s^t w(u)[c_2(u) - c_1(u)]du\right\} \times \prod_{i \geq 1} \left(\frac{c_2(\sigma_i)}{c_1(\sigma_i)} \mathbf{1}_{\{\sigma_i \leq t\}}\right),$$

and $\sigma_i(w)$ is (ν_s^1 a.s.) the i^{th} jump of w . For this, it suffices in turn to show that $\int_{\mathcal{W}} L_s^t(w)\nu_s^1(dw) = 1$ for all $t \in [s, T]$ (see eg. Theorem T3, p.166 of [8]). The process $\{L_s^t\}_{t \in [s, T]}$ on \mathcal{W} with the canonical filtration is a local martingale under ν_s^1 (see Theorem T2, p.166, [8]) so to check the martingale property, it suffices to check (see (2.4) on page 166, and Theorem T8 on page 27 of [8]), that

$$\int_{\mathcal{W}} \left[\int_s^T L_s^u(w)|c_1(u) - c_2(u)|du \right] \nu_s^1(dw) < +\infty.$$

Note that $L_s^u(w) \leq \exp\{Tw(T)\delta\}$ since $c_1 - \delta \leq c_2 \leq c_1$. Thus

$$\int_{\mathcal{W}} \left[\int_s^T L_s^u(w)|c_1(u) - c_2(u)|du \right] \nu_s^1(dw) \leq 2T \int_{\mathcal{W}} e^{Tw(T)\delta} \nu_0^1(dw).$$

Note that, under ν_0^1 , $w(T)$ is stochastically bounded by the sum of two independent $\text{Geom}(e^{-T})$ (see proof of Lemma 6.7). Thus the last integral is bounded by $(\mathbb{E}e^{TZ\delta})^2$, where Z is a $\text{Geom}(e^{-T})$ random variable. This expectation is finite since $T\delta < e^{-T}$. The result follows. ■

Proof of Lemma 6.9 will make use of parts (iv) and (v) of Lemma 6.5. In order to use (v) we will need the kernels to be suitably bounded. For that we use a truncation of the kernels, the associated error of which is estimated through the following lemma.

Lemma 6.20 *For $A \in (0, \infty)$, $t \in [0, T]$ and rate functions (a, b, c) , define the kernel $\kappa_{A,t}(a, b, c) \equiv \kappa_{A,t} : \mathcal{X} \times \mathcal{X} \rightarrow [0, \infty)$ as*

$$\kappa_{A,t}(\mathbf{x}, \mathbf{y}) = \kappa_t(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\{w(T) \leq A, \tilde{w}(T) \leq A\}}, \quad \text{where } \kappa_t = \kappa_t(a, b, c), \mathbf{x} = (s, w), \mathbf{y} = (r, \tilde{w}).$$

Denote by $\mathcal{K}_{A,t}$ the integral operator corresponding to $\kappa_{A,t}$ on $L^2(\mathcal{X}, \mathcal{T}, \mu_t)$ and $\rho(\kappa_{A,t})$ its norm, where $\mu_t = \mu_t(a, b, c)$. Then there exist $A_0, C_3, C_4 \in (0, \infty)$ such that

$$\rho(\kappa_t) - C_3 e^{-C_4 A} \leq \rho(\kappa_{A,t}) \leq \rho(\kappa_t),$$

for all rate functions (a, b, c) , $t \in [0, T]$ and $A \geq A_0$.

Proof: We will suppress t in the notation. Since $\kappa_A \leq \kappa$, from Lemma 6.5 (ii) and (iii) we have

$$\rho(\kappa) - \rho(\kappa - \kappa_A) \leq \rho(\kappa_A) \leq \rho(\kappa).$$

Consequently,

$$\rho(\kappa) - \rho(\kappa_A) \leq \rho(\kappa - \kappa_A) \leq \|\kappa - \kappa_A\|_2,$$

where $\|\kappa\|_2$ denotes the $L^2(\mu \times \mu)$ norm of the kernel κ . Note that for $\mathbf{x}, \mathbf{y} \in \mathcal{X}$

$$\kappa(\mathbf{x}, \mathbf{y}) = \int_0^t w(u)\tilde{w}(u)b(u)du \leq Tw(T)\tilde{w}(T), \text{ for } \mu \times \mu \text{ a.e. } (\mathbf{x}, \mathbf{y}) = ((s, w), (r, \tilde{w})). \quad (6.22)$$

Let $\Lambda = \{(s, w) \in \mathcal{X} : w(T) > A\}$. Then for fixed $\mathbf{x} = (s, w) \in \Lambda^c$

$$\begin{aligned} \int_{\mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) - \kappa_A(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) &\leq \int_{\mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\Lambda}(\mathbf{y}))^2 \mu(d\mathbf{y}) \\ &\leq \int_{\mathcal{X}} [T^2 w^2(T) \tilde{w}^2(T) \mathbf{1}_{\{\tilde{w}(T) > A\}}] \mu(dr d\tilde{w}) \\ &\leq T^3 w^2(T) \mathbb{E}_0[w^2(T) \mathbf{1}_{\{w(T) > A\}}], \end{aligned}$$

where in the second line we have used (6.22) and \mathbb{E}_0 in the third line denotes the expectation corresponding to the probability measure ν_0 on \mathcal{W} . Next, for fixed $\mathbf{x} = (s, w) \in \Lambda$, we have in a similar manner

$$\int_{\mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) - \kappa_A(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \leq T^3 w^2(T) \mathbb{E}_0[w^2(T)].$$

Thus for any $\mathbf{x} = (s, w) \in \mathcal{X}$,

$$\int_{\mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) - \kappa_A(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \leq T^3 w^2(T) \mathbb{E}_0[w^2(T) \mathbf{1}_{\{w(T) > A\}}] + \mathbf{1}_{\{w(T) > A\}} T^3 w^2(T) \mathbb{E}_0[w^2(T)].$$

Integrating with respect to $\mathbf{x} \in \mathcal{X}$, and noting that $\nu_s(w(T) \geq \alpha) \leq \nu_0(w(T) \geq \alpha)$, $\alpha \geq 0$, we have

$$\begin{aligned} \|\kappa - \kappa_A\|_2^2 &= \int_{\mathcal{X}} \int_{\mathcal{X}} (\kappa(\mathbf{x}, \mathbf{y}) - \kappa_A(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \mu(d\mathbf{x}) \\ &\leq T^3 \mathbb{E}_0[w^2(T) \mathbf{1}_{\{w(T) > A\}}] \int_0^t a(s) \mathbb{E}_0[w^2(T)] ds + T^3 \mathbb{E}_0[w^2(T)] \int_0^t a(s) \mathbb{E}_0[\mathbf{1}_{\{w(T) > A\}} w^2(T)] ds \\ &\leq 2T^4 \mathbb{E}_0[w^2(T)] \mathbb{E}_0[w^2(T) \mathbf{1}_{\{w(T) > A\}}]. \end{aligned}$$

As noted in the proof of Lemma 6.7, under ν_0 , $w(T)$ is stochastically dominated by $Z_1^* + Z_2^*$ where Z_1^*, Z_2^* are two independent copies of $\text{Geom}(e^{-T})$. Therefore

$$\mathbb{E}_0[w_0^2(T)] \leq \mathbb{E}(Z_1^* + Z_2^*)^2 = 2e^{2T}(2 - e^{-T}) < 4e^{2T}$$

and for a suitable $A_0 \in (0, \infty)$

$$\mathbb{E}_0[w_0^2(T) \mathbf{1}_{\{w_0(T) > A\}}] \leq \sum_{k > A} k^2 \times (k-1)e^{-2T}(1 - e^{-T})^{k-2} \leq (1 - e^{-T})^{A/2}$$

for all $A \geq A_0$. Combining the above estimates, for all $A \geq A_0$

$$\|\kappa - \kappa_A\|_2 < 8T^4 e^{2T} (1 - e^{-T})^{A/2}.$$

The result follows. ■

In the proof of Lemma 6.9 we will apply Lemma 6.18 to measures μ_1, μ_2 of the form in Lemma 6.19. This latter lemma shows that (under the conditions of the lemma) $\mu_2 \ll \mu_1$. However, to use Lemma 6.18 we need the two measures to be mutually absolutely continuous. To treat this difficulty we will use an additional truncation introduced in the lemma below and the elementary fact in Lemma 6.22.

Lemma 6.21 *Given rate functions (a, b, c) and $A \in (0, \infty)$ and $t \in [0, T]$, let $\kappa_t, \kappa_{A,t}$ be as in Lemma 6.20. For $(s, w) \in \mathcal{X}$, let $\tau(s, w) = \inf\{u > s : w(u) - w(u-) \neq 0\}$. For $\delta > 0$, let $\Lambda_\delta = \{(s, w) \in \mathcal{X} : s \leq \delta \text{ and } \tau(s, w) \leq s + \delta\}$. Define the kernel $\kappa_{A,\delta,t}$ as*

$$\kappa_{A,\delta,t}(\mathbf{x}, \mathbf{y}) = \kappa_{A,t}(\mathbf{x}, \mathbf{y}) \mathbf{1}_{\Lambda_\delta^c}(\mathbf{x}) \mathbf{1}_{\Lambda_\delta^c}(\mathbf{y}).$$

Then there exists a $C_5 \in (0, \infty)$ such that for all rate functions (a, b, c) and $\delta, A \in (0, \infty)$

$$\rho(\kappa_{A,t}) - C_5 A^2 \delta \leq \rho(\kappa_{A,\delta,t}) \leq \rho(\kappa_{A,t}).$$

Proof: Once more we will suppress t from the notation. As in the proof of Lemma 6.20, we have

$$\rho(\kappa_A) - \|\kappa_A - \kappa_{A,\delta}\|_2 \leq \rho(\kappa_{A,\delta}) \leq \rho(\kappa_A).$$

For $\mathbf{x} \in \Lambda^c$,

$$\begin{aligned} \int_{\mathcal{X}} (\kappa_A(\mathbf{x}, \mathbf{y}) - \kappa_{A,\delta}(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) &\leq \int_{\mathcal{X}} \kappa_A^2(\mathbf{x}, (r, \tilde{w})) \mathbf{1}_{\{\tau \leq r + \delta\}} \mathbf{1}_{\{r \leq \delta\}} \mu(dr d\tilde{w}) \\ &\leq T^2 A^4 \int_0^\delta \nu_r \{\tau(r, w) \leq r + \delta\} dr, \end{aligned}$$

where the above inequality uses the bound $\kappa_A \leq T A^2$.

Also, for fixed $\mathbf{x} \in \Lambda$,

$$\int_{\mathcal{X}} (\kappa_A(\mathbf{x}, \mathbf{y}) - \kappa_{A,\delta}(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \leq T^2 A^4 \cdot T$$

Thus for all $\mathbf{x} \in \mathcal{X}$,

$$\int_{\mathcal{X}} (\kappa_A(\mathbf{x}, \mathbf{y}) - \kappa_{A,\delta}(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \leq T^2 A^4 \int_0^\delta \nu_r \{\tau(r, w) \leq r + \delta\} dr + \mathbf{1}_\Lambda(\mathbf{x}) T^3 A^4.$$

Finally

$$\|\kappa_A - \kappa_{A,\delta}\|_2^2 = \int_{\mathcal{X}} \int_{\mathcal{X}} (\kappa_A(\mathbf{x}, \mathbf{y}) - \kappa_{A,\delta}(\mathbf{x}, \mathbf{y}))^2 \mu(d\mathbf{y}) \mu(d\mathbf{x}) \leq 2T^3 A^4 \int_0^\delta \nu_r \{\tau(r, w) \leq r + \delta\} dr.$$

The result follows on observing that $\nu_r \{\tau(r, w) \leq r + \delta\} = 1 - \exp\{-\int_r^{r+\delta} 2c(u) du\} \leq 2\delta$. ■

We will use the following elementary lemma.

Lemma 6.22 *Let γ_0, γ_1 be finite measures on a measure space $(\mathcal{X}, \mathcal{T})$ such that $\gamma_0 \ll \gamma_1$. Let $G \in \mathcal{T}$ be such that $\{\mathbf{x} : d\gamma_0/d\gamma_1 > 0\} \supset G$. For $i = 0, 1$, let γ_i^G be restriction of γ_i to G : $\gamma_i^G(\cdot) = \gamma_i(\cdot \cap G)$. Then γ_0^G and γ_1^G are mutually absolutely continuous with $\frac{d\gamma_0^G}{d\gamma_1^G}(\mathbf{x}) = \frac{d\gamma_0}{d\gamma_1}(\mathbf{x}) \mathbf{1}_G(\mathbf{x})$ and $\frac{d\gamma_1^G}{d\gamma_0^G}(\mathbf{x}) = (\frac{d\gamma_0}{d\gamma_1}(\mathbf{x}))^{-1} \mathbf{1}_G(\mathbf{x})$, a.s. γ_1^G and γ_0^G .*

Lemma 6.9 requires establishing an estimate of the form $C_2(-\log \delta)^3 \delta^{1/2}$ for both $|\rho_t - \rho_{+,t}|$ and $|\rho_t - \rho_{-,t}|$. Proofs for the two cases are similar and so we only provide details for $|\rho_t - \rho_{-,t}|$ and leave the other case for the reader. Given $\delta \in (0, \infty)$ and rate functions (a, b, c) as in the statement of Lemma 6.9, we denote $(a_\delta, b_\delta, c_\delta) = ((a - \delta)^+, (b - \delta)^+, (c - \delta)^+)$. Note that

$$\begin{aligned} |\rho_t - \rho_{-,t}| &= |\rho_t(a, b, c) - \rho_t(a_\delta, b_\delta, c_\delta)| \\ &\leq |\rho_t(a, b, c) - \rho_t(a_\delta, b, c_\delta)| + |\rho_t(a_\delta, b, c_\delta) - \rho_t(a_\delta, b_\delta, c_\delta)|. \end{aligned} \tag{6.23}$$

We treat the first term on the right side in Lemma 6.23 while the second term is estimated in Lemma 6.24.

Lemma 6.23 *Let (a, b, c) be rate functions as in the statement of Lemma 6.9. There exists $C_6, \delta_2 \in (0, \infty)$ such that for all $\delta \in (0, \delta_2)$ and $t \in [0, T]$*

$$|\rho_t(a, b, c) - \rho_t(a_\delta, b, c_\delta)| < C_6(-\log \delta)^3 \delta^{1/2}.$$

Proof: Since the kernel $\kappa_t(a, b, c)$ does not depend on a, c , $\kappa_t(a, b, c) = \kappa_t(a_\delta, b, c_\delta) = \kappa_t$. Henceforth suppress t from the notation. Let $\rho = \rho_t(a, b, c)$ and $\rho_\delta = \rho_t(a_\delta, b, c_\delta)$. For $\varepsilon > 0$, let $D \subset \mathcal{X}$ be defined as

$$D \equiv D_\varepsilon = \{(s, w) : w(T) > A \text{ or } (\tau(s, w) \leq s + \varepsilon \text{ and } s \leq \varepsilon)\}^c$$

and define the kernel κ_D as

$$\kappa_D(\mathbf{x}, \mathbf{y}) = \kappa(\mathbf{x}, \mathbf{y}) \mathbf{1}_D(\mathbf{x}) \mathbf{1}_D(\mathbf{y}).$$

Using Lemmas 6.20 and Lemma 6.21,

$$\rho - C_3 e^{-C_4 A} - C_5 A^2 \varepsilon \leq \rho(\kappa_D) \leq \rho$$

and

$$\rho_\delta - C_3 e^{-C_4 A} - C_5 A^2 \varepsilon \leq \rho_\delta(\kappa_D) \leq \rho_\delta,$$

where $\rho(\kappa_D)$ [resp. $\rho_\delta(\kappa_D)$] is the norm of the corresponding integral operator on $L^2(\mu)$ [resp. $L^2(\mu_\delta)$], where $\mu = \mu(a, b, c)$ and $\mu_\delta = \mu(a_\delta, b, c_\delta)$. Thus we have

$$|\rho - \rho_\delta| < 2C_3 e^{-C_4 A} + 2C_5 A^2 \varepsilon + |\rho(\kappa_D) - \rho_\delta(\kappa_D)|. \quad (6.24)$$

We now estimate $|\rho(\kappa_D) - \rho_\delta(\kappa_D)|$. By Lemma 6.19, $\mu_\delta \ll \mu$ and

$$g(s, w) =_{def} \frac{d\mu_\delta}{d\mu}(s, w) = \frac{a_\delta(s)}{a_0(s)} \times \exp\left\{-\int_s^T w(u)[c_\delta(u) - c(u)]du\right\} \times \prod_{i=1}^{w(T)-2} \left(\frac{c_1(\tau_i)}{c_0(\tau_i)}\right),$$

μ a.s., where τ_i are as in the statement of Lemma 6.19. For μ a.e. $(s, w) \in D$ we have

$$w(t) \leq w(T) \leq A \text{ and } \tau_1(s, w) > \varepsilon,$$

consequently $c_\delta(\tau_i) \geq (\theta\varepsilon - \delta)^+$ for all i , where θ is as in the statement of Lemma 6.9. Also, since a is bounded away from 0, we can find $d_1 \in (0, \infty)$ such that $a_\delta(s) \geq (d_1 - \delta)^+$. Thus for μ a.e. $\mathbf{x} \in D$

$$(d_1 - \delta)^+ ((\theta\varepsilon - \delta)^+)^A < g(\mathbf{x}) < \exp\{TA\delta\}. \quad (6.25)$$

Denote by μ^D [resp. μ_δ^D] the restrictions of μ [resp. μ_δ] to D . Then from Lemma 6.22, whenever $\delta < \delta_0 = \min\{\theta\varepsilon, d_1\}$, μ^D and μ_δ^D are mutually absolutely continuous and

$$\frac{d\mu_\delta^D}{d\mu^D}(\mathbf{x}) = g(\mathbf{x}) \mathbf{1}_D(\mathbf{x}), \text{ a.e. } d\mu^D.$$

For the rest of the proof we consider only $\delta < \delta_0$. Then,

$$\left(1 - \frac{\delta}{d_1}\right) \left(1 - \frac{\delta}{\theta\varepsilon}\right)^A < g(\mathbf{x}) < \exp\{TA\delta\}. \quad (6.26)$$

Note that $\rho(\kappa_D, \mu) = \rho(\kappa_D, \mu^D)$ and $\rho(\kappa_D, \mu_\delta) = \rho(\kappa_D, \mu_\delta^D)$. Also by Lemma 6.18, $\rho(\kappa_D, \mu_\delta^D) = \rho(\kappa'_D, \mu^D)$, where

$$\kappa'_D(\mathbf{x}, \mathbf{y}) = \kappa_D(\mathbf{x}, \mathbf{y}) \sqrt{\frac{g(\mathbf{x})}{g(\mathbf{y})}} \mathbf{1}_D(\mathbf{x}) \mathbf{1}_D(\mathbf{y}), \mathbf{x}, \mathbf{y} \in \mathcal{X}.$$

Thus using Lemma 6.5

$$\begin{aligned}
|\rho(\kappa_D) - \rho_\delta(\kappa_D)| &= |\rho(\kappa_D, \mu) - \rho(\kappa_D, \mu_\delta)| \\
&= |\rho(\kappa_D, \mu^D) - \rho(\kappa_D, \mu_\delta^D)| \\
&= |\rho(\kappa_D, \mu^D) - \rho(\kappa'_D, \mu^D)| \\
&\leq \rho\left(\left|\sqrt{\frac{g(\mathbf{x})}{g(\mathbf{y})}} - 1\right| \kappa_D(\mathbf{x}, \mathbf{y}), \mu^D\right) \\
&\leq T \sup_{\mathbf{x}, \mathbf{y} \in \mathcal{X}} \left(\left|\sqrt{\frac{g(\mathbf{x})}{g(\mathbf{y})}} - 1\right| \kappa_D(\mathbf{x}, \mathbf{y}) \right).
\end{aligned}$$

From (6.26) we see that whenever $\delta \leq \delta_0$,

$$\sup_{\mathbf{x}, \mathbf{y} \in D} \left(\sqrt{\frac{g(\mathbf{x})}{g(\mathbf{y})}} \vee \sqrt{\frac{g(\mathbf{y})}{g(\mathbf{x})}} \right) < \left(1 + \frac{2}{d_1} \delta\right) \times \exp\{TA\delta\} \times \left(1 + \frac{2\delta}{\theta\epsilon}\right)^A \equiv d(\delta, \epsilon, A).$$

Noting that $\kappa_D \leq TA^2$, we have

$$\sup_{\mathbf{x}, \mathbf{y} \in \mathcal{X}} \left(\left|\sqrt{\frac{g(\mathbf{x})}{g(\mathbf{y})}} - 1\right| \kappa_D(\mathbf{x}, \mathbf{y}) \right) \leq TA^2 |d(\delta, \epsilon, A) - 1|.$$

Thus

$$|\rho(\kappa_D) - \rho_\delta(\kappa_D)| \leq T^2 A^2 |d(\delta, \epsilon, A) - 1|.$$

Combining this with (6.24), we have

$$|\rho - \rho_\delta| < 2C_3 e^{-C_4 A} + 2C_5 A^2 \epsilon + TA^2 |d(\delta, \epsilon, A) - 1|. \tag{6.27}$$

Take $A = -\frac{1}{C_4} \log \delta$ and $\epsilon = \delta^{1/2}$. Note that when δ is sufficiently small, $\delta \leq \frac{1}{2} \min\{\theta\epsilon, d_1\}$ and so the above inequality holds for such δ . Also, with this choice, we can find $\delta_1, d_2 \in (0, \infty)$ such that the sum of the first two terms on the right side in (6.27) is bounded by

$$2C_3 \delta + \frac{2C_5}{C_4^2} (-\log \delta)^2 \delta^{1/2} \leq d_2 (-\log \delta)^2 \delta^{1/2} \quad \text{for all } \delta \leq \delta_1.$$

Also note that with the above choice of ϵ, A , $d(\delta, \epsilon, A) \rightarrow 1$ as $\delta \rightarrow 0$. Furthermore

$$d(\delta, \epsilon, A) = (1 + O(\delta))(1 + O(\delta))(1 + O(\delta^{1/2}(-\log \delta))).$$

Thus, we can find $d_3, \delta_2 \in (0, \infty)$ such that whenever $\delta \leq \delta_2$

$$TA^2 |d(\delta, \epsilon, A) - 1| \leq d_3 (-\log \delta)^3 \delta^{1/2}.$$

The result follows on combining the above estimates. \blacksquare

We now estimate the second term in (6.23).

Lemma 6.24 *There exists $C_7 \in (0, \infty)$ and $\delta_3 \in (0, 1)$ such that for all $t \in [0, T]$*

$$|\rho_t(a_\delta, b, c_\delta) - \rho_t(a_\delta, b_\delta, c_\delta)| < C_7 (-\log \delta)^2 \delta$$

Proof: Denote $\kappa_t(a_\delta, b_\delta, c_\delta) = \kappa_{\delta,t} = \kappa_\delta$ and recall that $\kappa_t(a_\delta, b_\delta, c_\delta) = \kappa_t = \kappa$. We suppress t in rest of the proof. For $A > 0$, let κ_A [resp. $\kappa_{\delta,A}$] be the truncated kernels as in Lemma 6.20 associated with κ [resp. κ_δ]. Clearly

$$|\kappa_A - \kappa_{\delta,A}| \leq TA^2\delta.$$

Using this bound and Lemma 6.20, for all $A \geq A_0$

$$\begin{aligned} |\rho(a_\delta, b, c_\delta) - \rho(a_\delta, b_\delta, c_\delta)| &= |\rho(\kappa, \mu_\delta) - \rho(\kappa_\delta, \mu_\delta)| \\ &\leq TA^2\delta + |\rho(\kappa, \mu_\delta) - \rho(\kappa_A, \mu_\delta)| + |\rho(\kappa_\delta, \mu_\delta) - \rho(\kappa_{\delta,A}, \mu_\delta)| \\ &\leq TA^2\delta + 2C_3 \exp\{-C_4A\}, \end{aligned}$$

where $\mu_\delta = \mu(a_\delta, b_\delta, c_\delta) = \mu(a_\delta, b, c_\delta)$ is as in Lemma 6.23.

Result follows on taking $A = (-\log \delta)/C_4$ and δ sufficiently small. ■

Proof of Lemma 6.9. The proof is immediate on using Lemmas 6.23 and 6.24 in (6.23). ■

6.7 Proof of Lemma 6.10

We begin with an elementary lemma which allows one to regard the operators $\mathcal{K}_t(a, b, c)$, $t \in [0, T]$, to be defined on a common Hilbert space. Recall that for a kernel κ on $\mathcal{X} \times \mathcal{X}$ and a finite measure μ on \mathcal{X} , we denote by $\rho(\kappa, \mu)$ the norm of the integral operator associated with κ on $L^2(\mathcal{X}, \mathcal{T}, \mu) = L^2(\mu)$.

Lemma 6.25 *Let (a, b, c) be rate functions. Then for all $t \in [0, T]$*

$$\rho(\kappa_t(a, b, c), \mu_t(a, b, c)) = \rho(\kappa_t(a, b, c), \mu_T(a, b, c)).$$

Proof: Write $\kappa_t(a, b, c) = \kappa_t$, $\mu_t(a, b, c) = \mu_t$. Denote the integral operator corresponding to κ_t on $L^2(\mu_t)$ [resp. $L^2(\mu_T)$] by \mathcal{K}_t [resp. \mathcal{K}_t^T]. Let $\mathcal{X}_t = [0, t] \times \mathcal{W}$, then μ_t is supported on \mathcal{X}_t and κ_t is supported on $\mathcal{X}_t \times \mathcal{X}_t$. Thus for any $\psi \in L^2(\mu_T)$, $\mathcal{K}_t^T \psi = \mathcal{K}_t \psi$ is also supported on \mathcal{X}_t , and this implies $\|\mathcal{K}_t^T\| \leq \|\mathcal{K}_t\|$. On the other hand, for any $\psi \in L^2(\mu_t)$, $\mathcal{K}_t \psi = \mathcal{K}_t^T(\psi \mathbf{1}_{[0,t] \times \mathcal{W}})$

$$\|\mathcal{K}_t \psi\|_{L^2(\mu_t)} \leq \|\mathcal{K}_t^T\| \|\psi \mathbf{1}_{[0,t] \times \mathcal{W}}\|_{L^2(\mu_T)} = \|\mathcal{K}_t^T\| \|\psi\|_{L^2(\mu_t)}.$$

Thus $\|\mathcal{K}_t^T\| \geq \|\mathcal{K}_t\|$. ■

The following theorem concerning IRG models on a general type space is a corollary of Theorem 3.1 and Theorem 3.12 in [16].

Theorem 6.26 [16] *Let $(\mathcal{X}, \mathcal{T}, \mu)$ be a type space. Consider the weight function $\phi \equiv 1$ on this space. Let κ_n, κ be symmetric kernels on $\mathcal{X} \times \mathcal{X}$ such that*

$$x_n \rightarrow x \text{ and } y_n \rightarrow y \text{ implies } \kappa_n(x_n, y_n) \rightarrow \kappa(x, y).$$

Let $\mathcal{C}_n^{(1)}(\kappa)$ denote the size of the largest component in $\mathbf{RG}_n(\kappa, \mu, \phi)$. Then

(i) *If $\rho(\kappa, \mu) \leq 1$, then $\mathcal{C}_n^{(1)}(\kappa_n)/n \xrightarrow{\mathbb{P}} 0$. Furthermore, if $\|\kappa\|_\infty < +\infty$, then $\mathcal{C}_n^{(1)}(\kappa_n) = O(\log n)$.*

(ii) *If $\rho(\kappa, \mu) > 1$, then $\mathcal{C}_n^{(1)}(\kappa_n) = \Theta(n)$.*

Proof of Lemma 6.10: Let $\kappa_t(a_0, b_0, c_0) = \kappa_t$ and $\mu_t(a_0, b_0, c_0) = \mu_t$. Note that for $0 \leq t_1 \leq t_2 \leq T$, since $\kappa_{t_1} \leq \kappa_{t_2}$, we have from Lemma 6.25 that

$$\rho(t_1) = \rho(\kappa_{t_1}, \mu_{t_1}) = \rho(\kappa_{t_1}, \mu_T) \leq \rho(\kappa_{t_2}, \mu_T) = \rho(\kappa_{t_2}, \mu_{t_2}) = \rho(t_2).$$

Thus $\rho(t)$ is nondecreasing in t . Next note that, since $w(u)\tilde{w}(u)b_0(u)$ is non decreasing in u for $\mu_T \times \mu_T$ a.e. $((s, w), (r, \tilde{w}))$, $\kappa_t(\mathbf{x}, \mathbf{y})$ is convex in t for a.e. \mathbf{x}, \mathbf{y} , i.e. for $\mu_T \times \mu_T$ a.e. $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{X}$, and all $t_1, t_2 \in [0, T]$, and $\alpha, \beta \in [0, 1]$, $\alpha + \beta = 1$,

$$\kappa_{\alpha t_1 + \beta t_2}(\mathbf{x}, \mathbf{y}) \leq \alpha \kappa_{t_1}(\mathbf{x}, \mathbf{y}) + \beta \kappa_{t_2}(\mathbf{x}, \mathbf{y}).$$

Thus

$$\begin{aligned} \rho(\alpha t_1 + \beta t_2) &= \rho(\kappa_{\alpha t_1 + \beta t_2}, \mu_T) \\ &\leq \rho(\alpha \kappa_{t_1} + \beta \kappa_{t_2}, \mu_T) \\ &\leq \rho(\alpha \kappa_{t_1}, \mu_T) + \rho(\beta \kappa_{t_2}, \mu_T) \\ &= \alpha \rho(\kappa_{t_1}, \mu_T) + \beta \rho(\kappa_{t_2}, \mu_T) \\ &= \alpha \rho(t_1) + \beta \rho(t_2), \end{aligned}$$

where lines 3 and 4 above use parts (ii) and (iii) of Lemma 6.5 and line 2 uses the convexity of κ .

Thus ρ is convex on $[0, T]$. Also since $\rho(0) = 0$ and $\rho(t) > 0$ for $t > 0$, we have that ρ is strictly increasing on $[0, T]$ and has a strictly positive left derivative on $(0, T]$. This proves parts (i) and (iii) of Lemma 6.10.

We now consider part (ii). For $\delta > 0$, let

$$\rho^{\delta,+}(t) = \rho_t((a_0 + \delta) \wedge 1, (b_0 + \delta) \wedge 1, (c_0 + \delta) \wedge 1), \quad \rho^{\delta,-}(t) = \rho_t((a_0 - \delta)^+, (b_0 - \delta)^+, (c_0 - \delta)^+).$$

Similarly define $\mu_t^{\delta,-}$, $\kappa_t^{\delta,-}$, $\mu_t^{\delta,+}$ and $\kappa_t^{\delta,+}$. We will argue by contradiction.

Suppose first that $\rho(t_c) > 1$, then by Lemma 6.9 and the continuity of $\rho(t)$, there exist $\epsilon, \delta > 0$ such that $\rho^{\delta,-}(t_c - \epsilon) > 1$. Denote $\mathbf{RG}_t^{\delta,-}(\kappa_n) = \mathbf{RG}_n(\kappa_{n,t}^{\delta,-}, \mu_t^{\delta,-}, \phi_t)$, where $\kappa_{n,t}^{\delta,-}$ is defined as in (5.19), replacing b there with $(b_0 - \delta)^+$. Since $\kappa_{n,t}^{\delta,-}$ converges to $\kappa_t^{\delta,-}$ uniformly on compact subsets of $\mathcal{X} \times \mathcal{X}$, by Theorem 6.26, the size of the largest component of $\mathbf{RG}_t^{\delta,-}(\kappa_n)$, whp, is $\Theta(n)$ and consequently the volume of the largest component of $\mathbf{RG}_t^{\delta,-}(\kappa_n)$ is, whp, at least $\Theta(n)$. By Lemma 5.2, and as $\bar{x}_n(t) \rightarrow x(t)$, we have whp

$$\mathcal{COM}_n(t_c - \epsilon) \geq_d \mathbf{IA}_n((a_0 - \delta)^+, (b_0 - \delta)^+, (c_0 - \delta)^+)_{t_c - \epsilon}$$

and from Lemma lemma:rgiva-irg-def the largest component in $\mathbf{IA}_n((a_0 - \delta)^+, (b_0 - \delta)^+, (c_0 - \delta)^+)_{t_c - \epsilon}$ has the same distribution as that in $\mathbf{RG}_{t_c - \epsilon}^{\delta,-}(\kappa_n)$. However, by Theorem 1.1 of [40] the largest component size of Bohman-Frieze model for $t < t_c$ is, whp, $\Theta(\log n)$ which contradicts the fact that the volume of the largest component of $\mathbf{RG}_{t_c - \epsilon}^{\delta,-}(\kappa_n)$ is, whp, at least $\Theta(n)$. Thus we have shown that $\rho(t_c) \leq 1$.

Suppose now that $\rho(t_c) < 1$. then there exists $\epsilon, \delta > 0$ such that $\rho^{\delta,+}(t_c + \epsilon) < 1$. Then a similar argument as above shows that, whp,

$$\mathcal{COM}_n(t_c + \epsilon) \leq_d \mathbf{IA}_n((a_0 + \delta) \wedge 1, (b_0 + \delta) \wedge 1, (c_0 + \delta) \wedge 1)_{t_c + \epsilon} =_{def} \mathbf{IA}_{n,t_c + \epsilon}^\delta.$$

Lemma 6.11 implies that whp the largest component in $\mathbf{IA}_{n,t_c + \epsilon}^\delta$ is $O(\log^4 n)$. However from Theorem 1.1 of [40], for $t > t_c$, the largest component size of Bohman-Frieze model is whp $\Theta(n)$. This contradiction shows that $\rho(t_c) \geq 1$. Combining the above arguments we have $\rho(t_c) = 1$. ■

7 Proof of Proposition 4.1

We shall now study the asymptotics of \mathcal{S}_2 and \mathcal{S}_3 , namely the sum of squares and cubes of the component sizes. We first analyze \mathcal{S}_2 since the asymptotics for this will be required in the analysis of \mathcal{S}_3 .

7.1 Analysis of $\bar{s}_2(\cdot)$ near criticality

Let us start with the sum of squares. Recall from (2.1) that $\mathcal{S}_2(t)$ denotes the sum of squares of the component sizes in $\mathbf{BF}(t)$ and $\bar{s}_2(t) = \mathcal{S}_2(t)/n$. Also recall the limiting functions $s_k(t)$, $k = 2, 3$, introduced in (2.3) and (2.4). Note that these functions are non-decreasing and they blow up at the critical point t_c (see (2.7) and (2.8)). Let $y(t) = 1/s_2(t)$. Then this function satisfies the differential equation

$$y'(t) = -x^2(t)y^2(t) - (1 - x^2(t)), \quad y(0) = 1, \quad t \in [0, t_c]. \quad (7.1)$$

Note that y is a monotonically decreasing function with $y(t) \rightarrow 0$ as $t \rightarrow t_c$ and as shown in Theorem 3.2 of [26], the scaling behavior near t_c is

$$y(t) = \frac{1}{\alpha}(t_c - t) + O(t_c - t)^2 \quad (7.2)$$

as $t \uparrow t_c$, where α is as in (2.5). Let $Y_n(t) = 1/\bar{s}_2(t)$. To simplify notation, we suppress the dependence of the process Y on n when convenient. Note that (4.2) is equivalent to showing

$$n^{1/3} \left| Y(t_n) - \frac{1}{\alpha n^\gamma} \right| \xrightarrow{\mathbb{P}} 0. \quad (7.3)$$

Here, and throughout Section 7, $t_n = t_c - 1/n^\gamma$ with $\gamma \in (1/6, 1/5)$. From (7.2)

$$\left| y(t_n) - \frac{1}{\alpha n^\gamma} \right| = O\left(\frac{1}{n^{2\gamma}}\right) = o\left(\frac{1}{n^{1/3}}\right). \quad (7.4)$$

Thus to show (7.3) it is enough to prove the following

Proposition 7.1 *As $n \rightarrow \infty$*

$$n^{1/3} \sup_{s \leq t_n} |Y(s) - y(s)| \xrightarrow{\mathbb{P}} 0.$$

We shall prove this via a sequence of lemmas. We begin with some notation. Recall that $\mathbf{C}_n^{BF}(t) \equiv (\mathcal{C}_n^{(i)}(t) : i \geq 1) \equiv (\mathcal{C}_i(t) : i \geq 1)$ is the component size vector, $I_n(t)$ the size of the largest component, and $X_n(t)$ the number of singletons, in $\mathbf{BF}_n(t)$. Let \sum_i denote the summation over all components and $\sum_{i < j}$ denote the summation over all pairs of components (i, j) with $i < j$.

The first Lemma identifies the semimartingale decomposition of the process $Y(\cdot)$ as well as the predictable quadratic variation $\langle M \rangle$ of the martingale M in the decomposition. Recall the natural filtration associated with the BF process introduced in Section 5.2.

Lemma 7.2 *The process $Y_n(\cdot)$ can be decomposed as*

$$Y_n(t) = 1 + \int_0^t A_n(s) ds + M_n(t), \quad t \in [0, t_c] \quad (7.5)$$

where

- (a) M_n is a RCLL martingale with respect to the natural filtration $\{\mathcal{F}_t\}_{t \geq 0}$ of the BF process.
- (b) The process $A_n = A_1^n + R_1^n$ where (suppressing n)

$$A_1(u) = -Y^2(u)\bar{x}^2(u) - (1 - \bar{x}^2(u)) + (1 - \bar{x}^2(u)) \frac{Y^2(u)}{n^2} \sum_i \mathcal{C}_i^A(u), \quad u \leq t_c \quad (7.6)$$

and for some $C_7 \in (0, \infty)$,

$$|R_1^n(u)| \leq C_7 \left(\frac{1}{n} + \frac{I_n^2(u)}{n} \right), \quad \text{for all } n \in \mathbb{N} \text{ and } u \leq t_c. \quad (7.7)$$

(c) Predictable quadratic variation of M_n is given as

$$\langle M_n \rangle(t) = \int_0^t B_n(u) du$$

where B_n is such that

$$B_n(u) \leq \frac{4}{n} + \frac{4Y_n^2(u)I_n^2(u)}{n}, \quad u \leq t_c. \quad (7.8)$$

Proof: We will suppress n from the notation when convenient. Note that

$$Y(t) = 1 + \sum_{s \leq t} \Delta Y(s), \quad \text{where } \Delta Y(s) = Y(s) - Y(s-).$$

We now analyze the possible jumps of Y . Note that any jump in Y corresponds to a jump of one of the Poisson processes $\mathcal{P}_{\mathbf{e}}$, $\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2$ (recall the notation from Section 2). A jump of $\mathcal{P}_{\mathbf{e}}$ at a time instant u could result in two different kinds of jumps in Y .

(i) **Merger caused by the first edge e_1 :** In this case $\Delta \mathcal{S}_2(u) = \mathcal{S}_2(u) - \mathcal{S}_2(u-) = 2$ which implies

$$\Delta Y(u) \equiv \alpha_1(u-) = -\frac{2Y^2(u-)}{n} \left[1 - O\left(\frac{2Y(u-)}{n}\right) \right].$$

(ii) **Merger caused by the second edge e_2 :** In this case, suppose components i and j merge, then $\Delta \mathcal{S}_2(u) = 2\mathcal{C}_i(u-)\mathcal{C}_j(u-)$ and thus

$$\Delta Y_n(u) \equiv \alpha_2^{i,j}(u-) = -2\frac{\mathcal{C}_i(u-)\mathcal{C}_j(u-)}{n} Y_n^2(u-) \left[1 - O\left(2\frac{\mathcal{C}_i(u-)\mathcal{C}_j(u-)Y_n(u-)}{n}\right) \right].$$

With these observations we can represent Y in terms of stochastic integrals with respect to $\mathcal{P}_{\mathbf{e}}$ as follows. Define

$$\mathcal{H}_1(u) = \{\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2 : e_1 = (v_1, v_2) \text{ where both } v_1, v_2 \text{ are singletons at time } u\},$$

$$\mathcal{H}_2^{(i,j)}(u) = \{\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2 \setminus \mathcal{H}_1(u) : e_2 = (v_1, v_2) \text{ where one vertex is in } \mathcal{C}^i(u) \text{ while the other is in } \mathcal{C}^j(u)\}.$$

Also let

$$\mathcal{U}_{\mathbf{e}}(u) = \alpha_1(u)\mathbf{1}_{\mathcal{H}_1(u)}(\mathbf{e}), \quad \mathcal{U}_{\mathbf{e}}^{i,j}(u) = \alpha_2^{i,j}(u)\mathbf{1}_{\mathcal{H}_2^{(i,j)}(u)}(\mathbf{e}).$$

Then

$$Y(t) = 1 + \sum_{\mathbf{e} \in \mathcal{E}^2} \int_{(0,t]} \left(\mathcal{U}_{\mathbf{e}}(s-) + \sum_{i < j} \mathcal{U}_{\mathbf{e}}^{i,j}(s-) \right) \mathcal{P}_{\mathbf{e}}(ds). \quad (7.9)$$

Recalling that $\mathcal{P}_{\mathbf{e}}$ is a rate $2/n^3$ Poisson process, one can write Y as

$$Y(t) = 1 + \int_{[0,t]} A(s) ds + M(t),$$

where

$$A(s) = \frac{2}{n^3} \sum_{\mathbf{e} \in \mathcal{E}^2} \left(\mathcal{U}_{\mathbf{e}}(s) + \sum_{i < j} \mathcal{U}_{\mathbf{e}}^{i,j}(s) \right).$$

Note that

$$\sum_{\mathbf{e} \in \mathcal{E}^2} \mathbf{1}_{\mathcal{H}_1(s)}(\mathbf{e}) = \binom{X_n(s)}{2} \binom{n}{2} = \frac{n^4}{4} \bar{x}_n^2(s) \cdot (1 + O(1/n)), \quad (7.10)$$

$$\sum_{\mathbf{e} \in \mathcal{E}^2} \mathbf{1}_{\mathcal{H}_2^{(i,j)}(s)}(\mathbf{e}) = \left[\binom{n}{2} - \binom{X_n(s)}{2} \right] \mathcal{C}_i(s) \mathcal{C}_j(s) = \frac{n^2}{2} (1 - \bar{x}_n^2(s)) \mathcal{C}_i(s) \mathcal{C}_j(s) \cdot (1 + O(1/n)). \quad (7.11)$$

Thus we get $A(s) = A_1(s) + R_1(s)$, where A_1 represents the leading order terms:

$$\begin{aligned} A_1(s) &= -\frac{n}{2} \bar{x}_n^2(s) \cdot \frac{2Y^2(s)}{n} - \sum_{i < j} \frac{1}{n} (1 - \bar{x}_n^2(s)) \mathcal{C}_i(s) \mathcal{C}_j(s) \cdot \frac{2Y^2(s) \mathcal{C}_i(s) \mathcal{C}_j(s)}{n} \\ &= -\bar{x}_n^2(s) Y^2(s) - (1 - \bar{x}_n^2(s)) Y^2(s) \cdot \frac{1}{n^2} \sum_{i < j} 2\mathcal{C}_i^2(s) \mathcal{C}_j^2(s) \\ &= -\bar{x}_n^2(s) Y^2(s) - (1 - \bar{x}_n^2(s)) Y^2(s) \cdot \frac{1}{n^2} \left[\left(\sum_i \mathcal{C}_i^2(s) \right)^2 - \sum_i \mathcal{C}_i^4(s) \right] \\ &= -\bar{x}_n^2(s) Y^2(s) - (1 - \bar{x}_n^2(s)) + (1 - \bar{x}_n^2(s)) Y^2(s) \cdot \frac{1}{n^2} \sum_i \mathcal{C}_i^4(s) \end{aligned}$$

and the last line follows from the fact $Y = \frac{n}{\sum_i \mathcal{C}_i^2}$. The term R_1 consists of the lower order terms and using the observations $Y \leq 1$, $\bar{x}_n \leq 1$ and $|A_1| \leq \bar{x}_n^2 Y^2 + (1 - \bar{x}_n^2) \leq 2$, it can be estimated as follows.

$$\begin{aligned} |R_1(u)| &\leq |A_1(u)| \cdot \frac{d_1}{n} + \frac{n}{2} \bar{x}_n^2(u) \cdot \frac{2Y^2(u)}{n} \cdot \frac{d_2 Y(u)}{n} \\ &\quad + (1 - \bar{x}_n^2(u)) Y^2(u) \cdot \frac{1}{n^2} \sum_{i < j} \left[2\mathcal{C}_i^2(u) \mathcal{C}_j^2(u) \cdot \frac{d_3 \mathcal{C}_i(u) \mathcal{C}_j(u) Y(u)}{n} \right] \\ &\leq \frac{2d_1}{n} + \frac{d_2}{n} + \frac{d_3 I^2(u)}{n} \cdot \frac{Y^2(u)}{n^2} \sum_{i < j} 2\mathcal{C}_i^2(u) \mathcal{C}_j^2(u) \\ &\leq \frac{2d_1 + d_2}{n} + \frac{d_3 I^2(u)}{n}. \end{aligned}$$

Next, from (7.9) and using independence of Poisson processes $\mathcal{P}_{\mathbf{e}}$,

$$\langle M \rangle(t) = \frac{2}{n^3} \sum_{\mathbf{e} \in \mathcal{E}^2} \int_{(0,t]} \left((\mathcal{U}_{\mathbf{e}}(s))^2 + \sum_{i < j} (\mathcal{U}_{\mathbf{e}}^{i,j}(s))^2 \right) ds \equiv \int_{(0,t]} B(s) ds,$$

where, using (7.10) and (7.11) once more, we can estimate B as follows.

$$\begin{aligned} B(u) &\leq \frac{n}{2} \bar{x}_n^2(u) \cdot 2 \cdot \left(\frac{2Y^2(u)}{n} \right)^2 + (1 - \bar{x}_n^2(u)) \cdot 2 \cdot \sum_{i < j} \frac{\mathcal{C}_i(u) \mathcal{C}_j(u)}{n} \left(\frac{2Y^2(u) \mathcal{C}_i(u) \mathcal{C}_j(u)}{n} \right)^2 \\ &\leq \frac{4}{n} + \frac{4Y^2(u) I^2(u)}{n} \cdot \frac{Y^2(u)}{n^2} \sum_{i < j} 2\mathcal{C}_i^2(u) \mathcal{C}_j^2(u) \\ &\leq \frac{4}{n} + \frac{4Y^2(u) I^2(u)}{n}. \end{aligned}$$

This completes the proof of the lemma.

■

The following result bounds the difference $Y - y$ through an application of Gronwall's lemma.

Lemma 7.3 For every $n \in \mathbb{N}$,

$$\sup_{s \leq t_n} |Y_n(s) - y(s)| \leq \varepsilon_n e^{2t_c}$$

where

$$\varepsilon_n = 4t_c \sup_{s \leq t_c} |\bar{x}(s) - x(s)| + \sup_{s \leq t_n} |M_n(t)| + C_8 \int_0^{t_n} \frac{I_n^2(s)}{n} ds,$$

$C_8 = (2C_7 + 1)$ and C_7, M_n are as in Lemma 7.2.

Proof: Since y solves (7.1) and Y satisfies (7.5), using the fact that \bar{x}, x take values in $[0, 1]$, we have for any fixed $t \leq t_n$

$$\begin{aligned} |Y(t) - y(t)| \leq & \left| \int_0^t (Y^2(s)\bar{x}^2(s) - y^2(s)x^2(s)) ds \right| + \int_0^t |\bar{x}^2(s) - x^2(s)| ds + \int_0^t |R_1(s)| ds \\ & + \int_0^t \left(\frac{Y^2(s)}{n^2} \sum_i C_i^4(s) \right) ds + \sup_{s \leq t_n} |M(t_n)|. \end{aligned} \quad (7.12)$$

Let us now analyze each term individually. Writing

$$Y^2(s)\bar{x}^2(s) - y^2(s)x^2(s) = \bar{x}^2(s)(Y^2(s) - y^2(s)) + y^2(s)(\bar{x}^2(s) - x^2(s))$$

and using the fact that Y, y, x, \bar{x} take values in the interval $[0, 1]$ we get

$$\left| \int_0^t Y^2(s)\bar{x}^2(s) - y^2(s)x^2(s) ds \right| \leq 2 \int_0^t |Y(s) - y(s)| ds + 2t_c \sup_{s \leq t_c} |\bar{x}(s) - x(s)|$$

Similarly the second term in (7.12) can be estimated as

$$\int_0^t |\bar{x}^2(s) - x^2(s)| ds \leq 2t_c \sup_{s \leq t_c} |\bar{x}(s) - x(s)|.$$

From (7.7) the integrand in the third term in (7.12) can be bounded as

$$|R_1(s)| \leq C_7 \left(\frac{1}{n} + \frac{I_n^2(s)}{n} \right) \leq 2C_7 \frac{I_n^2(s)}{n}.$$

Similarly, the integrand in the fourth term in (7.12) can be bounded as

$$\frac{Y^2(s)}{n^2} \sum_i C_i^4(s) \leq \frac{I_n^2(s)}{n} Y^2(s) \frac{1}{n} \sum_i C_i^2(s).$$

Noting that by definition $Y(s) \frac{1}{n} \sum_i C_i^2(s) = 1$ and $Y(s) \leq 1$ we get

$$\int_0^t \left(\frac{Y^2(s)}{n^2} \sum_i C_i^4(s) \right) ds \leq \int_0^{t_n} \frac{I_n^2(s)}{n} ds.$$

Combining we get, for all $t \leq t_n$

$$|Y(t) - y(t)| \leq \varepsilon_n + \int_0^t 2|Y(s) - y(s)| ds.$$

Thus Gronwall's lemma (see Theorem 5.1, Appendix in [23]) proves the result. \blacksquare

One last ingredient in the proof of Proposition 7.1 is the following estimate on the martingale in Lemma 7.2.

Lemma 7.4 As $n \rightarrow \infty$,

$$n^{1/3} \sup_{s \leq t_n} |M(s)| \xrightarrow{\mathbb{P}} 0.$$

Proof of Lemma 7.4 is given in Section 7.2.

Proof of Proposition 7.1: In view of Lemma 7.3 it is enough to show $n^{1/3} \varepsilon_n \xrightarrow{\mathbb{P}} 0$ as $n \rightarrow \infty$, where ε_n is as defined in the Lemma 7.3. Let us analyze each of the terms in ε_n . First note that by Lemma 6.4, for any $\vartheta < 1/2$, and in particular for $\vartheta = 1/3$

$$\sup_{s \leq t_c} n^\vartheta |\bar{x}(s) - x(s)| \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \rightarrow \infty.$$

Next, from Proposition 5.1

$$\{I_n(t) \leq m(n, t), \forall t < t_c - n^{-\gamma}\} \text{ whp as } n \rightarrow \infty,$$

where $m(n, t) = B \frac{(\log n)^4}{(t_c - t)^2}$. Thus, recalling that $\gamma \in (1/6, 1/5)$, we have, whp,

$$\begin{aligned} \int_0^{t_n} \frac{I_n^2(s)}{n} ds &\leq \frac{B^2 (\log n)^8}{n} \int_0^{t_n} \frac{1}{(t_c - s)^4} ds \\ &= \frac{B^2 n^{3\gamma} (\log n)^8}{n} = o\left(\frac{1}{n^{1/3}}\right). \end{aligned} \quad (7.13)$$

The result now follows on combining these estimates with Lemma 7.4. ■

7.2 Proof of Lemma 7.4

We shall prove the lemma by first showing that $n^\vartheta \sup_{s \leq t_n} |M(s)| \xrightarrow{\mathbb{P}} 0$ for any $\vartheta < 1/5$ and then sharpening the estimates to allow for any $\vartheta \leq 1/3$.

Lemma 7.5 As $n \rightarrow \infty$,

$$\mathbb{P}\left(\sup_{s \leq t_n} |M_n(s)| > \frac{n^{3\gamma/2} (\log n)^6}{\sqrt{n}}\right) \rightarrow 0 \quad (7.14)$$

and

$$\mathbb{P}(Y_n(t) \leq 2y(t) \forall t < t_n) \rightarrow 1.$$

Proof: Fix $\gamma_1 \in (1/3, 1/2)$ and define stopping times τ_i , $i = 1, 2$ by

$$\tau_1 = \inf\{t : I_n^1(t) > m(n, t)\}, \quad \tau_2 = \inf\{t : |\bar{x}(t) - x(t)| > n^{-\gamma_1}\}.$$

From Proposition 5.1 and Lemma 6.4

$$\mathbb{P}(\tau_1 \wedge \tau_2 > t_n) \rightarrow 1 \quad (7.15)$$

as $n \rightarrow \infty$. Let $\tau^* = t_n \wedge \tau_1 \wedge \tau_2$. For (7.14), in view of (7.15), it suffices to prove the statement with t_n replaced by τ^* . By Doob's maximal inequality we have

$$\mathbb{E}(\sup_{s \leq \tau^*} |M(s)|^2) \leq 4\mathbb{E}(\langle M \rangle(\tau^*)) = 4\mathbb{E} \int_0^{\tau^*} B(u) du.$$

Furthermore, from (7.8),

$$B(s) \leq \frac{4}{n} + \frac{4Y^2(s)I_n^2(s)}{n} \leq \frac{4}{n} + \frac{4I_n^2(s)}{n}.$$

Since for $t < \tau^*$ we have $I_n(t) \leq B(\log n)^4/(t_c - t)^2$ (see (5.1)), we have

$$\mathbb{E}[\langle M \rangle(\tau^*)] \leq \int_0^{t_n} \left(\frac{4}{n} + \frac{4B^2(\log n)^8}{n(t_c - t)^4} \right) dt \leq d_1 \frac{n^{3\gamma}(\log n)^{10}}{n}.$$

Combining the estimates, we have

$$\mathbb{E}(\sup_{s \leq \tau^*} |M(s)|) \leq d_1^{1/2} \frac{n^{3\gamma/2}(\log n)^5}{\sqrt{n}}.$$

A simple application of Markov's inequality and (7.15) gives (7.14). To get the final assertion in the Lemma, note that on the set

$$B_n = \{\tau_1 > t_n\} \cap \{\tau_2 > t_n\} \cap \left\{ \sup_{s \leq \tau^*} |M(s)| < \frac{n^{3\gamma/2}(\log n)^6}{\sqrt{n}} \right\}$$

$|\bar{x}(t) - x(t)| < n^{-\gamma_1}$ and (see (7.13))

$$\int_0^{t_n} \frac{I_n^2(s)}{n} ds = o\left(\frac{1}{n^{1/3}}\right).$$

Therefore, since $\gamma < 1/5$, the error ε_n in Lemma 7.3 satisfies whp

$$\varepsilon_n < \frac{n^{3\gamma/2}(\log n)^7}{\sqrt{n}} = o\left(\frac{1}{n^\gamma}\right).$$

Noting that $y(t)$ is a monotonically decreasing function, the above along with (7.4) implies that $Y(t) = (1 + o_p(1))y(t)$ for $t \leq t_n$ on the set B_n . Since $\mathbb{P}(B_n) \rightarrow 1$, the result follows. ■

Proof of Lemma 7.4: Along with the stopping times τ_1, τ_2 introduced in Lemma 7.5, consider the stopping time

$$\tau_3 = \inf \{t : Y(t) > 2y(t)\}$$

Then Lemma 7.5 and (7.15) imply that

$$\mathbb{P}(\tau_1 \wedge \tau_2 \wedge \tau_3 > t_n) \rightarrow 1$$

as $n \rightarrow \infty$. Thus to complete the proof, it is enough to show that for the stopping time $\tau = t_n \wedge \tau_1 \wedge \tau_2 \wedge \tau_3$, $n^{1/3}\mathbb{E}(\sup_{s \leq \tau} M(t)) \rightarrow 0$ as $n \rightarrow \infty$. Once again by Doob's maximal inequality it is enough to show that

$$n^{2/3}\mathbb{E}(\langle M \rangle(\tau)) \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (7.16)$$

Now note that (7.8) implies that

$$\begin{aligned} \mathbb{E}[\langle M \rangle(\tau)] &\leq \mathbb{E} \int_0^\tau \left(\frac{4}{n} + \frac{4Y^2(s)I_n^2(s)}{n} \right) ds \\ &\leq \frac{4t_n}{n} + \frac{4}{n} \int_0^{t_n} 4y^2(s) \frac{2B^2(\log n)^8}{(t_c - s)^4} ds \\ &\leq d_1 \left(\frac{1}{n} + \frac{1}{n} \int_0^{t_n} \frac{B^2(\log n)^8}{\alpha^2(t_c - s)^2} ds + \frac{(\log n)^8}{n} \right). \end{aligned}$$

In the second line of the above display we have used the fact that $I_n(t) \leq m(n, t)$ for all $t \leq \tau_1$ and in the last line we have used (7.2). Thus

$$\mathbb{E}[n^{2/3}\langle M \rangle(\tau)] \leq d_2 \left(\frac{1}{n^{1/3}} + \frac{n^{2/3+\gamma}(\log n)^8}{n} \right).$$

Since $\gamma < 1/5$ we have (7.16) and this completes the proof. ■

7.3 Analysis of \bar{s}_3 near criticality

We will now analyze the sum of cubes of component sizes near criticality and consequently prove (4.1). Define

$$z(t) = \frac{s_3(t)}{s_2^3(t)}, \quad t \in [0, t_c].$$

Then differential equations (2.3) (2.4) imply (see [26]) that z solves the differential equation

$$z'(t) = 3x^2(t)y^3(t) - 3x^2(t)y(t)z(t), \quad z(0) = 1, \quad t \in [0, t_c] \quad (7.17)$$

and furthermore $z(t) \rightarrow \beta$ as $t \rightarrow t_c$. Now consider the process

$$Z_n(t) = \frac{\mathcal{S}_3(t)/n}{(\mathcal{S}_2(t)/n)^3} = Y_n^3(t) \frac{\mathcal{S}_3(t)}{n}.$$

Then to show (4.1), it is enough to show the following proposition:

Proposition 7.6 *Fix any $\gamma \in (1/6, 1/5)$ and let as before $t_n = t_c - n^{-\gamma}$. Then*

$$|Z(t_n) - z(t_n)| \xrightarrow{\mathbb{P}} 0$$

as $n \rightarrow \infty$.

The analysis is similar to that for $\mathcal{S}_2(\cdot)$ as carried out in the Section 7.1. We begin by writing the semimartingale decomposition for Z_n and identifying the predictable quadratic variation $\langle \tilde{M} \rangle$ of the martingale \tilde{M} in the decomposition.

Lemma 7.7 *The process Z_n can be decomposed as*

$$Z_n(t) = 1 + \int_0^t \tilde{A}_n(s) ds + \tilde{M}_n(t), \quad t \in [0, t_c] \quad (7.18)$$

where

(a) \tilde{M}_n is a RCLL martingale with respect to the natural filtration $\{\mathcal{F}_t\}_{t \geq 0}$ of the BF process.

(b) The process $\tilde{A}_n = \tilde{A}_1^n + \tilde{R}_1^n + \tilde{R}_2^n$ where (suppressing n), for $u \in [0, t_c]$,

$$\begin{aligned} \tilde{A}_1(u) &= 3\bar{x}^2(u)Y^3(u) - 3Z(u)Y(u)\bar{x}^2(u), \\ |\tilde{R}_1(u)| &= \left| 3(1 - \bar{x}^2(u)) \left[Z(u)Y(u) \sum_i \frac{\mathcal{C}_i^4(u)}{n^2} - Y^3(u) \sum_i \frac{\mathcal{C}_i^5(u)}{n^2} \right] \right| \leq \frac{6I^3(u)Y^2(u)}{n} \end{aligned} \quad (7.19)$$

and

$$|\tilde{R}_2(u)| \leq \frac{30Y^2(u)I_n^3(u)}{n} + \frac{12Y^3(u)I_n^5(u)}{n^2}.$$

(c) Predictable quadratic variation of \tilde{M}_n is given as

$$\langle \tilde{M}_n \rangle(t) = \int_0^t \tilde{B}_n(u) du,$$

and the process \tilde{B}_n can be bounded as

$$\tilde{B}(u) \leq C_9 \left(\frac{Y^4(u)I^4(u)}{n} + \frac{Y^5(u)I^6(u)}{n^2} + \frac{Y^6(u)I^8(u)}{n^3} + \frac{Y^7(u)I^{10}(u)}{n^4} + \frac{Y^8(u)I^{12}(u)}{n^5} \right) \quad (7.20)$$

for some $C_9 \in (0, \infty)$.

Proof: We will proceed as in the proof of Lemma 7.2. Note that

$$Z(t) = 1 + \sum_{s \leq t} \Delta Z(s), \text{ where } \Delta Z(s) = Z(s) - Z(s-).$$

Next note that if $\Delta \mathcal{S}_2(u) = a$ and $\Delta \mathcal{S}_3(u) = b$, then

$$Z(u) = \frac{\bar{s}_3(u-) + b/n}{(\bar{s}_2(u-) + a/n)^3} = \frac{Y^3(u-)(\bar{s}_3(u-) + b/n)}{(1 + aY(u-)/n)^3}.$$

Using the estimate $|(1+x)^{-3} - (1-3x)| \leq 6x^2$, for $0 < x < 1$, we get

$$Z(u) = Y^3(u-) \left(\bar{s}_3(u-) + \frac{b}{n} \right) \left(1 - \frac{3aY(u-)}{n} + \tilde{R}(a, u-) \right),$$

where

$$|\tilde{R}(a, u-)| \leq \frac{6a^2 Y^2(u-)}{n^2}.$$

Thus

$$\Delta Z(u) \equiv \tilde{\zeta}(a, b, u-) = -\frac{3aZ(u-)Y(u-)}{n} + \frac{Y^3(u-)b}{n} + \tilde{R}_{a,b}(u-),$$

where the remainder term

$$\begin{aligned} |\tilde{R}_{a,b}(u-)| &\leq Z(u-)|\tilde{R}(a, u-)| + \frac{3abY^4(u-)}{n^2} + \frac{bY^3(u-)|\tilde{R}(a, u-)|}{n} \\ &\leq \frac{6a^2 Z(u-)Y^2(u-)}{n^2} + \frac{3abY^4(u-)}{n^2} + \frac{6a^2 b Y^5(u-)}{n^3}. \end{aligned}$$

Any jump in $\mathcal{S}_2, \mathcal{S}_3$ or Z corresponds to a jump of one of the Poisson processes $\mathcal{P}_{\mathbf{e}}$, $\mathbf{e} = (e_1, e_2) \in \mathcal{E}^2$. A jump of $\mathcal{P}_{\mathbf{e}}$ at a time instant u could result in the following different values for a, b .

(i) **Merger caused by the first edge e_1 :** In this case $a = 2$ and $b = 6$ and $\tilde{R}_{a,b}$ can be estimated as

$$|\tilde{R}_{2,6}(u-)| \leq \frac{24Y^4(u-)I(u-)}{n^2} + \frac{36Y^4(u-)}{n^2} + \frac{144Y^5(u-)}{n^3} \leq \frac{204Y^4(u-)I(u-)}{n^2}.$$

(ii) **Merger caused by the second edge e_2 :** In this case, suppose components i and j merge, then

$$a \equiv \theta_{i,j}(u-) = 2\mathcal{C}_i(u-)\mathcal{C}_j(u-), \quad b \equiv \eta_{i,j}(u-) = 3\mathcal{C}_i^2(u-)\mathcal{C}_j(u-) + 3\mathcal{C}_i(u-)\mathcal{C}_j^2(u-)$$

and noting that $a \leq 2I^2$, $b \leq 6I^3$ and

$$Z(u) = Y^3(u)\bar{s}_3(u) \leq Y^2(u)I_n(u), \tag{7.21}$$

we have

$$|\tilde{R}_{a,b}(u-)| \leq \frac{6a^2 Z(u-)Y^2(u-)}{n^2} + \frac{3abY^4(u-)}{n^2} + \frac{6a^2 b Y^5(u-)}{n^3} \leq \frac{30aY^4(u-)I^3(u-)}{n^2} + \frac{72aY^5(u-)I^5(u-)}{n^3}.$$

With these observations we can represent Z in terms of stochastic integrals with respect to $\mathcal{P}_{\mathbf{e}}$ as follows. Recall $\mathcal{H}_1, \mathcal{H}_2^{(i,j)}$ introduced in the proof of Lemma 7.2. Define $\tilde{\alpha}_1(u) = \tilde{\zeta}(2, 6, u)$ and $\tilde{\alpha}_2^{i,j}(u) = \tilde{\zeta}(\theta_{i,j}(u), \eta_{i,j}(u), u)$. Also, let

$$\tilde{\mathcal{U}}_{\mathbf{e}}(u) = \tilde{\alpha}_1(u)\mathbf{1}_{\mathcal{H}_1(u)}(\mathbf{e}), \quad \tilde{\mathcal{U}}_{\mathbf{e}}^{i,j}(u) = \tilde{\alpha}_2^{i,j}(u)\mathbf{1}_{\mathcal{H}_2^{(i,j)}(u)}(\mathbf{e}).$$

Then

$$Z(t) = 1 + \sum_{\mathbf{e} \in \mathcal{E}^2} \int_{(0,t]} \left(\tilde{\mathcal{U}}_{\mathbf{e}}(s-) + \sum_{i < j} \tilde{\mathcal{U}}_{\mathbf{e}}^{i,j}(s-) \right) \mathcal{P}_{\mathbf{e}}(ds). \quad (7.22)$$

Recalling that $\mathcal{P}_{\mathbf{e}}$ is a rate $2/n^3$ Poisson process, one can write Z as

$$Z(t) = 1 + \int_{[0,t]} \tilde{A}(s) ds + \tilde{M}(t),$$

where

$$\tilde{A}(s) = \frac{2}{n^3} \sum_{\mathbf{e} \in \mathcal{E}^2} \left(\tilde{\mathcal{U}}_{\mathbf{e}}(s) + \sum_{i < j} \tilde{\mathcal{U}}_{\mathbf{e}}^{i,j}(s) \right).$$

Also, once again using independence of Poisson processes $\mathcal{P}_{\mathbf{e}}$,

$$\langle \tilde{M} \rangle(t) = \frac{2}{n^3} \sum_{\mathbf{e} \in \mathcal{E}^2} \int_{(0,t]} \left((\tilde{\mathcal{U}}_{\mathbf{e}}(s))^2 + \sum_{i < j} (\tilde{\mathcal{U}}_{\mathbf{e}}^{i,j}(s))^2 \right) ds \equiv \int_{(0,t]} \tilde{B}(s) ds.$$

The proof is now completed upon using (7.10) and (7.11) as for Lemma 7.2. ■

As is clear from the above lemma, a precise analysis of Z will involve considering several terms of the form $I^\theta Y^\vartheta$. The following lemma shows that such terms are asymptotically negligible for suitable θ, ϑ .

Lemma 7.8 *For any $\theta, \vartheta \geq 0$ and $p > 0$ satisfying $\gamma(2\theta - \vartheta - 1) < p$*

$$\int_0^{t_n} \frac{I_n^\theta(u) Y_n^\vartheta(u)}{n^p} du \xrightarrow{\mathbb{P}} 0,$$

as $n \rightarrow \infty$.

Proof: From Proposition 5.1,

$$\{I_n(t) \leq B(\log n)^4 / (t_c - t)^2 \text{ for all } t \leq t_n\} \text{ occurs whp.}$$

Also, by Proposition 7.1 and (7.4) we know that

$$\sup_{t \leq t_n} |Y(t) - y(t)| = o_p(y(t_n))$$

and from (7.2), for t near t_c , $y(t_c) \sim (t_c - t)/\alpha$. Using these bounds in the above integral proves the result. ■

Proof of Proposition 7.6. By the integral representation of the process Z given in Lemma 7.7 and the fact that z solves the differential equation (7.17), we have that

$$\begin{aligned} |Z(t) - z(t)| &\leq 3 \int_0^t |\bar{x}^2(u) Y^3(u) - x^2(u) y^3(u)| du + 3 \int_0^t |Z(u) Y(u) \bar{x}^2(u) - z(u) y(u) x^2(u)| du \\ &\quad + \int_0^{t_n} |\tilde{R}_1(u)| du + \int_0^{t_n} |\tilde{R}_2(u)| du + \sup_{t \leq t_n} |\tilde{M}(t)|. \end{aligned} \quad (7.23)$$

Now the integrand in the first term can be bounded as

$$\begin{aligned} 3|\bar{x}^2(u) Y^3(u) - x^2(u) y^3(u)| &\leq 3Y^3(u) |\bar{x}^2(u) - x^2(u)| + 3\bar{x}^2(u) |Y^3(u) - y^3(u)| \\ &\leq 6|\bar{x}(u) - x(u)| + 9|Y(u) - y(u)|. \end{aligned} \quad (7.24)$$

The integrand in the second term in (7.23) can be decomposed as

$$\begin{aligned} z(u)y(u)x^2(u) - Z(u)Y(u)\bar{x}^2(u) &= x^2(u)y(u)(z(u) - Z(u)) \\ &\quad + y(u)Z(u)(x^2(u) - \bar{x}^2(u)) + Z(u)\bar{x}^2(u)(y(u) - Y(u)). \end{aligned}$$

Thus the second integral in (7.23) can be bounded by

$$3 \int_0^t |(z(u) - Z(u))| du + 6 \int_0^{t_n} Z(u)|x(u) - \bar{x}(u)| du + 3 \int_0^{t_n} Z(u)|Y(u) - y(u)| du. \quad (7.25)$$

Combining (7.25) and (7.24) we get that

$$|Z(t) - z(t)| \leq \varepsilon_n + 3 \int_0^t |(Z(u) - z(u))| du$$

where

$$\begin{aligned} \varepsilon_n &= 9t_c \sup_{s \leq t_n} (|\bar{x}(s) - x(s)| + |Y(s) - y(s)|) + 6 \int_0^{t_n} Z(u)|x(u) - \bar{x}(u)| du + 3 \int_0^{t_n} Z(u)|Y(u) - y(u)| du \\ &\quad + \int_0^{t_n} |\tilde{R}_1(u)| du + \int_0^{t_n} |\tilde{R}_2(u)| du + \sup_{t \leq t_n} |\tilde{M}(t)| \\ &= \eta_1 + \eta_2 + \eta_3 + \eta_4 + \eta_5 + \eta_6. \end{aligned}$$

By Gronwall's lemma, it is enough to show that $\varepsilon_n \rightarrow 0$ in probability as $n \rightarrow \infty$. Let us show each of the six constituents of ε_n satisfy this asymptotics. By Lemma 6.4 and Proposition 7.1 $\eta_1 \rightarrow 0$ in probability. Again by Lemma 6.4, for any $\vartheta < 1/2$, whp,

$$\eta_2 \leq \frac{6}{n^\vartheta} \int_0^{t_n} Z(u) du \leq \frac{6}{n^\vartheta} \int_0^{t_n} Y^2(u) I_n(u) du.$$

Using Lemma 7.8, the last term converges to 0 in probability as $n \rightarrow \infty$. Thus $\eta_2 \rightarrow 0$ in probability.

An identical argument, using Proposition 7.1 instead of Lemma 6.4 shows that $\eta_3 \rightarrow 0$ in probability. For η_4 , note that from (7.19),

$$|R_1(u)| \leq \frac{6I^3(n, u)Y^2(u)}{n}.$$

Lemma 7.8 now shows that $\eta_4 \rightarrow 0$ in probability. A similar argument, using the bounds in Lemma 7.7 on $R_2(u)$ establishes that $\eta_5 \rightarrow 0$ in probability. For η_6 note that for an arbitrary stopping time τ

$$\mathbb{E} \sup_{t \leq t_n \wedge \tau} |\tilde{M}_t|^2 \leq 4\mathbb{E}[\langle M \rangle(t \wedge \tau)] = 4\mathbb{E} \int_0^{t \wedge \tau} B(u) du.$$

The bound on $\tilde{B}(u)$ in (7.20) along with Lemma 7.8 and a localization argument similar to the one used in the proof of Lemma 7.4 now shows that η_6 converges to 0 in probability.

7.4 Proof of Proposition 4.1

We now complete the proof of Proposition 4.1. Proof of (4.2) follows from Proposition 7.1 and the discussion immediately above the proposition. Proof of (4.1) is immediate from Proposition 7.6. Finally we consider (4.3). From Proposition 7.1 and (7.4)

$$\frac{\mathcal{S}_2(t_n)}{\alpha n^{1+\gamma}} \xrightarrow{\mathbb{P}} 1 \quad (7.26)$$

Also, from Proposition 5.1

$$\mathbb{P} \left(\frac{\mathcal{C}_n^{(1)}(t_n)}{n^{2\gamma} \log^4 n} \leq B \right) \rightarrow 1$$

as $n \rightarrow \infty$. Combining, and recalling that $\gamma \in (1/6, 1/5)$, we have

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

This completes the proof of Propostion 4.1. \blacksquare

8 Proof of Theorem 2.1

We will now complete the proof of Theorem 2.1. As always, we write the component sizes as

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

and write the scaled component sizes as

$$\bar{\mathbf{C}}_n^{BF}(\lambda) \equiv \left(\frac{\beta^{1/3}}{n^{2/3}} \mathcal{C}_n^{(i)} \left(t_c + \beta^{2/3} \alpha \frac{\lambda}{n^{1/3}} \right) : i \geq 1 \right) \equiv (\bar{\mathcal{C}}_i(\lambda) : i \geq 1) \quad (8.1)$$

Then Proposition 4.1 proves that with

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

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

$$\frac{\sum_i (\bar{\mathcal{C}}_i(\lambda_n))^3}{\left[\sum_i (\bar{\mathcal{C}}_i(\lambda_n))^2 \right]^3} \xrightarrow{\mathbb{P}} 1, \quad \frac{1}{\sum_i (\bar{\mathcal{C}}_i(\lambda_n))^2} + \lambda_n \xrightarrow{\mathbb{P}} 0, \quad \frac{\bar{\mathcal{C}}_1(\lambda_n)}{\sum_i (\bar{\mathcal{C}}_i(\lambda_n))^2} \xrightarrow{\mathbb{P}} 0. \quad (8.2)$$

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 (8.2) imply that the $\bar{\mathbf{C}}_{i_f}$ process at time λ_n satisfies the regularity conditions of Proposition 4 of [4]. 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 [26], $\alpha(1 - x^2(t_c)) = 1$ (see (2.5)) we get

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

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 [7]. 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 $\bar{C}_n^{BF}(\lambda) \xrightarrow{d} \mathbf{X}(\lambda)$ in l_{\downarrow}^2 for each $\lambda \in \mathbb{R}$ and at the end of the section show that, in fact, $\bar{C}_n^{BF} \xrightarrow{d} \mathbf{X}$ in $\mathcal{D}((-\infty, \infty) : l_{\downarrow}^2)$. 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

$$\sum_i (\bar{c}_i(\lambda))^2 - \sum_{i \in \mathcal{COM}} (\bar{c}_i(\lambda))^2 \leq \frac{d_1}{n^{4/3}} \sum_{i=1}^{X_n(t)} 1 = O(1/n^{1/3}), \quad (8.3)$$

it suffices to prove Theorem 2.1 and verify Proposition 4.1 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 8.1 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 processes, and proving statements analogous to those in Theorem 2.1 for scaled component vectors associated with these processes. Proof of Theorem 2.1 will then be completed in Section 8.2.

8.1 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 5.2 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 5.2. $\mathcal{COM}_n^-(t)$, $t \geq t_n$ 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 (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 $\bar{C}_n^-(\lambda) = (\bar{c}_i^-(\lambda) : i \geq 1)$ the scaled (as in (8.1)) component size vector for $\mathcal{COM}_n^-(t)$. From Proposition 4 of [4], it follows that for any $\lambda \in \mathbb{R}$,

$$\bar{C}_n^-(\lambda) \xrightarrow{d} \mathbf{X}(\lambda) \quad (8.4)$$

in l_{\downarrow}^2 . Indeed, note that the first and third convergence statements in (8.2) hold with \bar{c}_i replaced with \bar{c}_i^- since the contributions made by singletons to the scaled sum of squares is $O(n^{-1/3})$ (see (8.3)) and to the sum of cubes is even smaller. This shows that the first and third requirements in Proposition 4 of [4] (see equations (8), (10) therein) are met. To show the second requirement in Proposition 4 of [4], using the second convergence in (8.2),

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left(\left(n^{2/3} \beta^{-1/3} \right)^2 \frac{b_n^*(t_n)}{n} \frac{\beta^{2/3} \alpha (\lambda - \lambda_n)}{n^{1/3}} - \frac{1}{\sum_i (\bar{c}_i^-(\lambda_n))^2} \right) \\ &= \lim_{n \rightarrow \infty} \alpha b_n^*(t_n) \lambda - \lambda_n (\alpha b_n^*(t_n) - 1) \\ &= \lambda - \lim_{n \rightarrow \infty} \lambda_n (\alpha b_n^*(t_n) - 1), \end{aligned} \quad (8.5)$$

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

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

where the second equality follows from (5.12). The first term on the last line converges to 0 using Lemma 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 (8.5) converges to λ as $n \rightarrow \infty$ and therefore the second requirement in Proposition 4 of [4] (see equation 9 therein) is met as well. This proves that $\bar{\mathcal{C}}_n^-(\lambda) \xrightarrow{d} \mathbf{X}(\lambda)$ in l_{\downarrow}^2 , for every $\lambda \in \mathbb{R}$. Although Proposition 4 of [4] 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, $\bar{\mathcal{C}}_n^- \xrightarrow{d} \mathbf{X}$ in $\mathcal{D}((-\infty, \infty) : l_{\downarrow}^2)$.

Upper bound coupling: Let us construct $\{\mathcal{COM}_n^+(t) : t \geq t_n\}$ 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 5.2 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 $\{\mathcal{COM}_n(t) : t_n \leq t \leq t_n^+\}$ first, then use $b_n^*(t_n^+) - b_n^*(t)$ to make up for all the additional edges in $\mathcal{COM}_n^+(t)$ for $t_n \leq t \leq t_n^+$. Note that $\mathcal{COM}_n(t_n + \cdot) \leq_d \mathcal{COM}_n^+(t_n + \cdot)$ over $[0, t_n^+ - t_n]$.

Let $\bar{\mathcal{C}}_n^+(\lambda) = (\bar{c}_i^+(\lambda) : i \geq 1)$ be the scaled (as in (8.1)) component size vector for \mathcal{COM}_n^+ . We will once more apply Proposition 4 of [4]. We first show that the three convergence statements in (8.2) 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

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

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 $c^* \leq 1$, we have that

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

Write

$$\mathcal{S}_k = \sum_i (\mathcal{C}_i)^k, \mathcal{S}_k^+ = \sum_i (\mathcal{C}_i^+)^k + \sum_{i=1}^N (\tilde{\mathcal{C}}_i^+)^k \text{ for } k = 2, 3 \text{ and } I = \max_i \mathcal{C}_i, I^+ = \max\{\max_i \mathcal{C}_i^+, \max_i \tilde{\mathcal{C}}_i^+\}.$$

The following proposition shows that Propostion 4.1 holds with $(\mathcal{S}_2(t_n), \mathcal{S}_3(t_n), \mathcal{C}_n^{(1)}(t_n))$ replaced with $(\mathcal{S}_2^+(t_n), \mathcal{S}_3^+(t_n), I^+(t_n))$.

Proposition 8.1 *As $n \rightarrow \infty$,*

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

Proof: An elementary calculation shows 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 $n \rightarrow \infty$,

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

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

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

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 (7.26) and (4.1) 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

$$\frac{n^{4/3} D_2}{(\mathcal{S}_2)^2} \xrightarrow{\mathbb{P}} 0. \tag{8.6}$$

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 $\gamma > 1/6$, (8.6) follows and thus the proof is completed. ■

For scaled component size vector of \mathcal{COM}_n^+ , the above proposition shows that the statements in (8.2) hold with $\bar{\mathcal{C}}_i$ replaced with $\bar{\mathcal{C}}_i^+$. In particular, the first and third requirements in Proposition 4 of [4] are met by $\{\bar{\mathcal{C}}_i^+\}$. Also, using the second convergence in (8.2), a calculation similar to that for (8.5) shows that

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

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

$$\bar{\mathcal{C}}_n^+(\lambda) \xrightarrow{d} \mathbf{X}(\lambda). \quad (8.7)$$

in l_{\downarrow}^2 , for every $\lambda \in \mathbb{R}$. Using Proposition 6 of [4] once again it now follows that $\bar{\mathcal{C}}_n^+ \xrightarrow{d} \mathbf{X}$ in $\mathcal{D}((-\infty, \infty) : l_{\downarrow}^2)$.

8.2 Completing the proof of Theorem 2.1

By [6, 7], there is a natural partial order \preceq on l_{\downarrow}^2 . Informally, interpreting an element of l_{\downarrow}^2 as a sequence of cluster sizes, $\mathbf{x}, \mathbf{y} \in l_{\downarrow}^2$, $\mathbf{x} \preceq \mathbf{y}$ if \mathbf{y} can be obtained from \mathbf{x} by adding new clusters and coalescing together clusters. The coupling constructed in Section 8.1 gives that, for every, $\lambda \in (\lambda_n, \lambda_n^+)$

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

Since, as $n \rightarrow \infty$, $\lambda_n \rightarrow -\infty$ and $\lambda_n^+ \rightarrow +\infty$, (8.4), (8.7) along with Lemma 15 of [7] yield that

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

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

Finally we argue convergence in $\mathcal{D}((-\infty, \infty) : l_{\downarrow}^2)$. For $\mathbf{x}, \mathbf{y} \in l_{\downarrow}^2$, 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 $\bar{\mathcal{C}}_n^{BF} \rightarrow \mathbf{X}$ in $\mathcal{D}((-\infty, \infty) : l_{\downarrow}^2)$ it suffices to prove that

$$\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathbf{d}(\bar{\mathcal{C}}_n^{BF}, \bar{\mathcal{C}}_n^-) \xrightarrow{\mathbb{P}} 0, \text{ for all } -\infty < \lambda_1 < \lambda_2 < \infty. \quad (8.8)$$

Fix λ_1, λ_2 as above. Then

$$\sup_{\lambda \in [\lambda_1, \lambda_2]} \mathbf{d}(\bar{\mathcal{C}}_n^{BF}, \bar{\mathcal{C}}_n^-) \leq \sup_{\lambda \in [\lambda_1, \lambda_2]} \left[\sum_i (\bar{\mathcal{C}}_i^+(\lambda))^2 - \sum_i (\bar{\mathcal{C}}_i^-(\lambda))^2 \right]. \quad (8.9)$$

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

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

From Lemma 15 of [7], $\mathcal{V}(\lambda) \xrightarrow{\mathbb{P}} 0$ 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 $\{\mathbf{X}^{\mathbf{x}}(t), t \geq 0\}$ denotes the multiplicative coalescent on the positive real line with initial condition $\mathbf{x} \in l_{\downarrow}^2$ then, for δ sufficiently small

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

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' tightness criteria (see Theorem VI.4.5 in [24]) for \mathcal{V} thus proving the desired tightness. ■

Acknowledgements AB and XW have been supported in part by the National Science Foundation (DMS-1004418), the Army Research Office (W911NF-0-1-0080, W911NF-10-1-0158) and the US-Israel Binational Science Foundation (2008466). SB's research has been supported in part by the UNC research council and a UNC Junior Faculty development award. SB would like to thank SAMSI for many interesting discussions with the participants of the complex networks program held at SAMSI in the year 2010-2011.

References

- [1] D. Achlioptas, R.M. D'Souza, and J. Spencer. Explosive percolation in random networks. *Science*, 323(5920):1453, 2009.
- [2] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, pages 1–40, 2009.
- [3] L. Addario-Berry, N. Broutin, and B. Reed. Critical random graphs and the structure of a minimum spanning tree. *Random Structures & Algorithms*, 35(3):323–347, 2009.
- [4] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25(2):812–854, 1997.
- [5] D.J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.
- [6] DJ Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.*, 3:1–59, 1998.
- [7] D.J. 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.
- [8] François Baccelli and Pierre Brémaud. *Elements of queueing theory*, volume 26 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2003. Palm martingale calculus and stochastic recurrences, Stochastic Modelling and Applied Probability.
- [9] Jean Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.
- [10] S. Bhamidi, R. Van Der Hofstad, and J.S.H. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Arxiv preprint arXiv:0909.1472*, 2009.

- [11] Shankar Bhamidi, Remco van der Hofstad, and Johan S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.*, 15:no. 54, 1682–1703, 2010.
- [12] T. Bohman and A. Frieze. Avoiding a giant component. *Random Structures and Algorithms*, 19(1):75–85, 2001.
- [13] T. Bohman, A. Frieze, and N.C. Wormald. Avoidance of a giant component in half the edge set of a random graph. *Random Structures & Algorithms*, 25(4):432–449, 2004.
- [14] B. Bollobas and O. Riordan. Asymptotic normality of the size of the giant component via a random walk. *Arxiv preprint arXiv:1010.4595*, 2010.
- [15] Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
- [16] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31(1):3–122, 2007.
- [17] L.A. Braunstein, S.V. Buldyrev, R. Cohen, S. Havlin, and H.E. Stanley. Optimal paths in disordered complex networks. *Physical review letters*, 91(16):168701, 2003.
- [18] W. Chen and R.M. D’Souza. Explosive percolation with multiple giant components. *Arxiv preprint arXiv:1011.5854*, 2010.
- [19] J. Ding, J.H. Kim, E. Lubetzky, and Y. Peres. Diameters in supercritical random graphs via first passage percolation. *Combinatorics, Probability and Computing*, 19(5-6):729–751, 2010.
- [20] R.M. DSouza and M. Mitzenmacher. Local cluster aggregation models of explosive percolation. *Physical Review Letters*, 104(19):195702, 2010.
- [21] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.
- [22] P. Erdős and A. Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.
- [23] Stewart N. Ethier and Thomas 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.
- [24] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.
- [25] S. Janson, D. Knuth, T. Luczak, and B. Pittel. The birth of the giant component, with an introduction by the editors. *Random Struct. Alg*, 4:231–358, 1994.
- [26] S. Janson and J. Spencer. Phase Transitions for Modified Erdos–Rényi Processes. *Arxiv preprint arXiv:1005.4494*, 2010.
- [27] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [28] Svante Janson, Tomasz Luczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

- [29] R.M. Karp. The transitive closure of a random digraph. *Random Structures & Algorithms*, 1(1):73–93, 1990.
- [30] M. Krivelevich, E. Lubetzky, and B. Sudakov. Hamiltonicity thresholds in Achlioptas processes. *Random Structures & Algorithms*, 37(1):1–24, 2010.
- [31] Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1977/78.
- [32] Anders Martin-Löf. The final size of a nearly critical epidemic, and the first passage time of a Wiener process to a parabolic barrier. *J. Appl. Probab.*, 35(3):671–682, 1998.
- [33] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, “Random Graphs ’93” (Poznań, 1993)*, volume 6, pages 161–179, 1995.
- [34] A. Nachmias and Y. Peres. Component sizes of the random graph outside the scaling window. *Alea*, 3:133–142, 2007.
- [35] Asaf Nachmias and Yuval Peres. Critical percolation on random regular graphs. *Random Structures Algorithms*, 36(2):111–148, 2010.
- [36] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [37] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.
- [38] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York, 1972.
- [39] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Arxiv preprint arXiv:1102.5306*, 2011.
- [40] J. Spencer and N. Wormald. Birth control for giants. *Combinatorica*, 27(5):587–628, 2007.
- [41] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. *Lectures on approximation and randomized algorithms*, pages 73–155, 1999.