Electron. Commun. Probab. **21** (2016), no. 64, 1–15.
DOI: 10.1214/16-ECP20
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On the largest component in the subcritical regime of the Bohman-Frieze process

Sanchayan Sen*

Abstract

Kang, Perkins, and Spencer [7] conjectured that the size of the largest component of the Bohman-Frieze process at a fixed time t smaller than t_c, the critical time for the process, is $L_1(t) = O(\log n/(t_c - t)^2)$ with high probability. Bhamidi, Budhiraja, and Wang [3] have shown that a bound of the form $L_1(t_n) = O((\log n)^4/(t_c - t_n)^2)$ holds with high probability for $t_n \le t_c - n^{-\gamma}$ where $\gamma \in (0, 1/4)$. In this paper, we improve the result in [3] by showing that for any fixed $\lambda > 0$, $L_1(t_n) = O(\log n/(t_c - t_n)^2)$ with high probability for $t_n \le t_c - \lambda n^{-1/3}$. In particular, this settles the conjecture in [7].

Keywords: Bohman-Frieze process; Achlioptas process; bounded-size rules; branching process; subcritical regime.
AMS MSC 2010: 60C05; 05C80.
Submitted to ECP on May 4, 2016, final version accepted on August 30, 2016.

1 Introduction

Initiated by a question of Dimitris Achlioptas, the study of modified Erdős–Rényi processes (called Achlioptas processes) has grown into a large area of research in the past decade. At each step of an Achlioptas process, two edges are sampled at random, and one of these edges is added to the graph according to some selection rule. The Erdős–Rényi process, for example, is an Achlioptas process where the first edge sampled is always added to the random graph. One such Achlioptas process called the Bohman-Frieze process has received a lot of attention and will be the main focus of this paper. We first describe a continuous time version of the Bohman-Frieze process. Consider the complete graph $K_n = (V_n, E_n)$ on the vertex set $V_n = \{1, \ldots, n\}$. Consider independent Poisson processes $\mathcal{P}_e,\ e \in E_n \times E_n$, each having rate $2/n^3$. Let $\cup_{e \in E_n \times E_n} \mathcal{P}_e = \{u_1 < u_2 < \ldots\}$. Then the continuous time Bohman-Frieze process $BF_n(\cdot)$ evolves as follows:

• *$BF_n(u)$ is the empty graph on V_n for $0 \le u < u_1$.*

• *If the Poisson process $\mathcal{P}_e$ has a point at u_i, where $e = (e_1, e_2) \in E_n \times E_n$ and the endpoints of e_1 are isolated vertices in $BF_n(u_i-)$, set $BF_n(u) = BF_n(u_i-) \cup e_1$ for $u \in [u_i, u_{i+1})$. Otherwise, set $BF_n(u) = BF_n(u_i-) \cup e_2$ for $u \in [u_i, u_{i+1})$.*

The corresponding discrete time version $DBF_n(\cdot)$ evolves as follows:

• *$DBF_n(u)$ is the empty graph on V_n for $0 \le u < 2/n$.*

• *At time $2(k+1)/n$, two edges e_1 and e_2 are selected uniformly (with replacement) from E_n. If the endpoints of e_1 are isolated vertices in $DBF_n(2k/n)$, set $DBF_n(u) =$*

*Department of Mathematics and Statistics, McGill University. E-mail: sanchayan.sen1@gmail.com

$DBF_n(2k/n) \cup e_1$ *for* $u \in [2(k+1)/n, 2(k+2)/n)$. *Otherwise, set* $DBF_n(u) = DBF_n(2k/n) \cup e_2$ *for* $u \in [2(k+1)/n, 2(k+2)/n)$.

We will denote the size of the largest component of $BF_n(t)$ (resp. $DBF_n(t)$) by $L_1^{BF}(t)$ (resp. $L_1^{DBF}(t)$). Spencer and Wormald [8] have shown that there exists $t_c > 1$ (which we call the critical time for the process) such that

$$L_1^{DBF}(t) = \begin{cases} O_P(\log n), \text{ if } t < t_c, \\ \Theta_P(n), \text{ if } t > t_c. \end{cases} \tag{1.1}$$

It is easy to see that t_c is also the critical time for the continuous time Bohman-Frieze process, and hence we will refer to it as the critical time for the Bohman-Frieze process. In [7, Theorem 4], it was claimed that for any fixed $t \in (0, t_c)$,

$$\mathbb{P}(L_1^{DBF}(t) \geq K \log n/(t_c - t)^2) \to 1 \tag{1.2}$$

for some constant K free of t. Later the authors discovered a gap in the proof (Mihyun Kang, personal communication). So the bound in (1.2) seems to be an open problem now. In the same paper, Kang, Perkins, and Spencer conjectured [7, Conjecture 1] that the bound in (1.2) is of the right order, i.e.,

$$\mathbb{P}(L_1^{DBF}(t) \leq K' \log n/(t_c - t)^2) \to 1$$

for any fixed $t \in (0, t_c)$ and some constant K' free of t. Bhamidi, Budhiraja, and Wang [3, 2] independently showed that for $\gamma \in (0, 1/4)$, there exists a constant $C = C(\gamma)$ such that

$$\mathbb{P}(L_1^{BF}(t_n) \leq C(\log n)^4/(t_c - t_n)^2) \to 1 \text{ for } t_n \leq t_c - n^{-\gamma} \tag{1.3}$$

by connecting the dynamics of $BF_n(\cdot)$ to an inhomogeneous random graph model. In this work, we take the approach in [3] and go through a more careful analysis to prove [7, Conjecture 1] (Theorem 2.1 and Corollary 2.2). Our result is true for $t_n \leq t_c - \lambda n^{-1/3}$ (for any fixed $\lambda > 0$) and thus closes the gap between the critical window and the interval $0 < t \leq t_c - n^{-\gamma}, \gamma < 1/4$ where the bound in [3] is valid.

2 Main results

The following theorem is our main result.

Theorem 2.1. *Let t_c be the critical time for the Bohman-Frieze process. Let $t = t(n)$ satisfy $t \leq t_c - \lambda n^{-1/3}$ for a fixed $\lambda > 0$. Let $L_1^{BF}(t)$ denote the size of the largest component of $BF_n(t)$. Then there exists a universal positive constant $\overline{C}$ such that*

$$\mathbb{P}\Big(L_1^{BF}(t) > \overline{C} \log n/(t_c - t)^2\Big) \xrightarrow{n\to\infty} 0. \tag{2.1}$$

An immediate consequence of Theorem 2.1 is an analogous result for the discrete time Bohman-Frieze process.

Corollary 2.2. *Let t_c be the critical time for the Bohman-Frieze process. Let $t = t(n)$ satisfy $t \leq t_c - \lambda n^{-1/3}$ for a fixed $\lambda > 0$. Let $L_1^{DBF}(t)$ denote the size of the largest component of $DBF_n(t)$. Then there exists a universal positive constant $\overline{\overline{C}}$ such that*

$$\mathbb{P}\big(L_1^{DBF}(t) > \overline{\overline{C}} \log n/(t_c - t)^2\big) \xrightarrow{n\to\infty} 0.$$

Remark 2.3. Let $L_i^{BF}(t)$ denote size of the i-th largest component of $BF_n(t)$ for $i = 1, 2, \ldots$. It was shown in [3] that for fixed $\lambda \in \mathbb{R}$ and $t = t(n) = t_c + \lambda n^{-1/3}$ (the so-called critical window), the vector $n^{-2/3}(L_1^{BF}(t), L_2^{BF}(t), \ldots)$ converges in distribution to the ordered excursion lengths of a reflected Brownian motion with parabolic drift (see [1]).

So the results of Theorem 2.1 and Corollary 2.2 are sub-optimal for this particular choice of $t(n)$. Very precise results about the size of the largest component in the barely-subcritical regime of the Erdős-Rényi process is known; see, e.g., [6, Theorem 5.23]. We expect the Bohman-Frieze process to exhibit similar behavior. From this consideration, one would guess that the bounds in Theorem 2.1 and Corollary 2.2 are sub-optimal also when $t(n)$ is, for instance, of the form $t(n) = t_c - \log n / n^{1/3}$. However, we expect our bounds to be of the right order whenever $t(n) \leq t_c - n^{-\gamma}$ and $\gamma < 1/3$. But as mentioned before, a lower bound of the form (1.2) is currently unavailable in the literature.

3 The associated inhomogeneous random graph model

The most natural way of bounding component sizes of random graphs is by branching process approximations. Spencer and Wormald [8] studied the Bohman-Frieze (BF) process directly via the differential equation method. However, this method does not associate the BF process with any branching process. In [3], an interesting connection between the BF process and an inhomogeneous random graph model (IRG) was established which made studying the BF process simpler, as IRGs are more amenable to analysis via branching process approximations. We will now briefly describe this connection and introduce the necessary notation. The interested reader can find a detailed account of general IRG models in [4].

Let $X_n(v)$ be the number of singletons (i.e., isolated vertices) in $BF_n(v)$, and set $x_n(v) = X_n(v)/n$. An edge added in the BF process at time v can be of the following types:
(A) Both its endpoints were isolated vertices in $BF_n(v-)$.
(B) Only one of its endpoints was an isolated vertex in $BF_n(v-)$.
(C) None of its endpoints were isolated in $BF_n(v-)$.
Two singletons are added in the BF process (i.e., an edge of type A is created) if one of the following happens:
(i) the first edge selected connects two isolated vertices or
(ii) the first edge selected does not connect two isolated vertices, but the second edge selected joins two isolated vertices.
Hence, two singletons are added in the BF process at a rate

$$\frac{2}{n^3}\left[\binom{X_n(v)}{2}\binom{n}{2} + \left(\binom{n}{2} - \binom{X_n(v)}{2}\right)\binom{X_n(v)}{2}\right] =: n a_n(x_n(v)),$$

where $a_n(\cdot)$ is a bounded function that can be directly computed from the above expression. One can similarly show that a given non-singleton vertex (i.e., a vertex that is not isolated) in $BF_n(v)$ gets connected to some isolated vertex (i.e., an edge of type B is created) at a rate

$$\frac{2}{n^3}\left(\binom{n}{2} - \binom{X_n(v)}{2}\right) X_n(v) =: c_n(x_n(v))$$

for a function $c_n(\cdot)$. Finally, two given non-singleton vertices are joined (i.e., an edge of type C is created) at a rate

$$\frac{2}{n^3}\left(\binom{n}{2} - \binom{X_n(v)}{2}\right) =: \frac{1}{n} b_n(x_n(v)),$$

for a function $b_n(\cdot)$. Define the functions $a_0, b_0, c_0 : [0,1] \to \mathbb{R}_+$ as follows:

$$a_0(y) = y^2 - y^4/2,\ b_0(y) = 1 - y^2,\ \text{and}\ c_0(y) = (1 - y^2)y. \tag{3.1}$$

Then it is straightforward to check that

$$\sup_{y\in[0,1]} \left(|a_n(y) - a_0(y)| + |b_n(y) - b_0(y)| + |c_n(y) - c_0(y)|\right) = O(1/n).$$

Further, the function $x_n(\cdot)$ is highly concentrated around $x(\cdot)$ (see [8, 3]) which satisfies the ODE

$$x'(v) = -x(v) - x^2(v) + x^3(v),\ x(0) = 1. \tag{3.2}$$

As a result, the three rate functions $a_n(x_n(v)), b_n(x_n(v))$, and $c_n(x_n(v))$ lie very close to the functions

$$a(v) := a_0(x(v)),\ b(v) := b_0(x(v)),\ \text{and } c(v) := c_0(x(v)) \tag{3.3}$$

respectively with high probability. This motivates one to consider an IRG model (defined below) associated with the deterministic rate functions a, b, c given by (3.3) that approximates the BF process with high probability.

We need a few definitions to describe the IRG model. Let $T = 2t_c$, where t_c is as in (1.1), and let W be the space $D([0,T] : \mathbb{Z}_{\geq 0})$ equipped with the Skorohod topology. Suppose we are given three continuous functions $\bar{a}, \bar{b}, \bar{c} : [0,T] \to [0,\infty)$. For each $s \in [0,T]$, let $V_s^{(\bar{c})}$ be a random increasing function taking values in W defined as follows:

• $V_s^{(\bar{c})}(u) = 0$ *for* $u \in [0,s)$ *and* $V_s^{(\bar{c})}(s) = 2$ *almost surely.*

• *Conditioned on* $V_s^{(\bar{c})}(u),\ s \leq u \leq v$, $\bar{V}_s^{(\bar{c})}(\cdot)$ *increases by one in* $(v, v+dv]$ *with rate* $V_s^{(\bar{c})}(v)\bar{c}(v)$.

Let $\nu_s^{(\bar{c})}$ be the law of $V_s^{(\bar{c})}$. Let $X = [0,T] \times W$. Define a measure $\mu = \mu[\bar{a}, \bar{c}]$ and a function ϕ_v on X as follows:

$$\mu(d(s,w)) = \bar{a}(s)ds\ \nu_s^{(\bar{c})}(dw),\ \phi_v(s,w) = w(v)\ \text{ for }\ (s,w) \in X. \tag{3.4}$$

Finally, define a *kernel* $k_v = k_v[\bar{b}]$ on $X \times X$ by

$$k_v\big((s_1,w_1),(s_2,w_2)\big) := \int_0^v w_1(u)w_2(u)\bar{b}(u)\,du,\ \text{where } (s_i,w_i) \in X,\ i = 1,2. \tag{3.5}$$

Then the IRG $RG_{n,v}(\bar{a}, \bar{b}, \bar{c})$ can be described as follows: declare the vertex set to be points of a Poisson process $\mathcal{P}$ on X with intensity $n\mu(dx)$, and then join any two points $x, y \in \mathcal{P}$ with probability $1 \wedge k_v(x,y)/n$. The volume of a connected component C is given by

$$\text{volume}(C) := \sum_{x \in C} \phi_v(x).$$

Lemma 4.2 stated below gives a precise way of comparing the the random graph $BF_n(t)$ with the IRG $RG_{n,t}(a,b,c)$, where a, b, c are as in (3.3).

Associated with the kernel k_v there is an integral operator K_v on $L^2(X,\mu)$ given by $K_v f(x) := \int_X k_v(x,y)f(y)\mu(dy)$. We will denote by $k_v^{(i)}$ the i-fold convolution of k_v with itself. Hence the integral operator associated with $k_v^{(i)}$ is K_v^i. We will write $\rho_v[\bar{a}, \bar{b}, \bar{c}]$ to denote the norm, in $L^2(\mu[\bar{a}, \bar{c}])$, of the operator K_v associated with $k_v = k_v[\bar{b}]$.

Organization of the paper. In Section 4.1, we give the statements of some technical results needed in the proof. In Section 4.2, we first reduce the problem of bounding L_1^{BF} to the problem of bounding the total progeny of the branching process associated with the IRG model. Then we go through a careful analysis of this branching process in Lemma 4.5. The proof of Theorem 2.1 is then easily completed by using a Chernoff bound. In Section 4.3, we give the proof of a technical estimate (Lemma 4.1) used in the proof of Theorem 2.1. Finally, in Section 4.4, we give the proof of Corollary 2.2.

We point out here that the two key ingredients that allow us to improve the result in [3] are: (i) Lemma 4.5 which gives very precise control over exponential moments of the total progeny of the associated branching process, and (ii) Lemma 4.1 which gives a sharp bound on the change in the operator norm $\rho_v[a,b,c]$ when the functions a, b, c are slightly perturbed.

4 Proofs

Throughout this section, C, C' will denote positive universal constants whose values may change from line to line. Special constants will be indexed, as for example C_1, C_2 etc. Throughout this section $a_0(\cdot), b_0(\cdot), c_0(\cdot), x(\cdot), a(\cdot), b(\cdot)$ and $c(\cdot)$ will be as in (3.1), (3.2), and (3.3) respectively. Recall from Section 3 that $T = 2t_c$ where t_c is the critical time for the BF process.

Assume that $t = t(n)$ satisfies $t \leq t_c - \lambda n^{-1/3}$. Fix $\gamma \in (1/3, 1/2)$, and define $E_n := \{\sup_{0 \leq u \leq T} |x_n(u) - x(u)| > n^{-\gamma}\}$. Since $a(u) = a_0(x(u))$ is C^1 on $[0, T]$ and $|a_n(u) - a_0(u)| = O(1/n)$, it follows that on the event E_n^c,

$$a_n(x_n(u)) \leq a_0(x_n(u)) + \frac{C}{n} \leq a(u) + C'\big(\frac{1}{n} + |x_n(u) - x(u)|\big) \leq a(u) + \frac{C_1}{n^\gamma},$$

and similar upper bounds hold for $b_n(x_n(u))$ and $c_n(x_n(u))$. Set

$$\delta = \delta_n = C_1 n^{-\gamma} \text{ and define } a_{n,\delta}(u) = a(u) + \delta_n$$

for $u \in [0, T]$. Define $b_{n,\delta}(u)$ and $c_{n,\delta}(u)$ similarly. Hence, $a_n(x_n(u)) \leq a_{n,\delta}(u)$ on E_n^c and similar upper bounds hold for $b_n(x_n(u))$ and $c_n(x_n(u))$. Note also that the following bound is a simple consequence of (3.1):

$$\sup_{u \in [0,T]} \max\{a_{n,\delta}(u), b_{n,\delta}(u), c_{n,\delta}(u)\} \leq 1 \text{ for large } n. \tag{4.1}$$

4.1 Some preliminary results

The following lemma, whose proof is deferred to Section 4.3, gives us control over $\rho_u[a_{n,\delta}, b_{n,\delta}, c_{n,\delta}]$.

Lemma 4.1. *There exists a positive constant β_0 such that*

$$\Big|\rho_u[a_{n,\delta}, b_{n,\delta}, c_{n,\delta}] - \rho_u[a, b, c]\Big| \leq \beta_0 \delta \text{ for each } u \in [0, T].$$

A weaker version of the above result was proven in [3, Lemma 6.9], where a bound of the form $O(|(\log \delta)|^3 \sqrt{\delta})$ was established. This weaker bound allows one to perform the analysis in the regime $t(n) \leq t_c - n^{-\gamma}$, $\gamma \in (0, 1/4)$. The stronger result in Lemma 4.1 allows us to prove the bound in (2.1) for $t(n) \leq t_c - \lambda n^{-1/3}$.

We will collect some useful estimates from [3] in the next lemma. Write $\nu_{s,\delta}$ for $\nu_s^{(c_{n,\delta})}$ (recall the definition from around (3.4)). Consider the IRG model $RG_{n,t}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ conditioned on having a point $(0, w)$, where w is distributed according to $\nu_{0,\delta}$. Let $C_{n,\delta}^{RG}(t)$ be the volume of the component containing $(0, w)$. Further, let $C_n^0(t)$ denote the component of the first edge appearing in $BF_n(t)$.

Lemma 4.2. *The following hold.*
(i) Bound on $\mathbb{P}(E_n)$ [3, Lemma 6.4]: $\mathbb{P}(E_n) \leq \exp(-Cn^{1-2\gamma})$.
(ii) Connection between BF process, and $RG_{n,t}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$ [3, Lemma 5.1, Lemma 5.2, and Lemma 6.1]: We have

$$\mathbb{P}(L_1^{BF}(t) > m) \leq nT\mathbb{P}(C_n^0(t) > m).$$

Further,

$$\mathbb{P}(C_n^0(t) > m, E_n^c) \leq \mathbb{P}(C_{n,\delta}^{RG}(t) > m).$$

(iii) Properties of operator norms [3, Lemma 6.10]: The function $f(u) := \rho_u[a, b, c]$ is strictly increasing and satisfies $f(t_c) = 1$. Further, there exists $\eta > 0$ such that

$$(1 - f(u))/(t_c - u) \to \eta \text{ as } u \uparrow t_c.$$

(iv) Absolute continuity of measures [3, Lemma 6.19]: Write $\mu = \mu[a, c]$, and $\mu_\delta = \mu[a_{n,\delta}, c_\delta]$. Then $\mu << \mu_\delta$ and for $x = (s, w) \in X$,

$$h(x) := \frac{d\mu}{d\mu_\delta}(x) = \frac{a(s)}{a_{n,\delta}(s)} \exp\left(\int_s^T w(u)(c_{n,\delta}(u) - c(u))du\right) \prod_{i=1}^{w(T)-2} \left(\frac{c(\gamma_i)}{c_{n,\delta}(\gamma_i)}\right), \tag{4.2}$$

where $\gamma_i = \gamma_i(s, w)$ is the time of the ith jump of w after time s.

Let $\overline{\mu} := \mu[1,1]$, i.e., the measure on $[0,T] \times W$ associated with the functions that are identically equal to one, and similarly let $\overline{\nu}_s = \nu_s^{(1)}$. Thus $\overline{\mu}(d(s,w)) = ds\ \overline{\nu}_s(dw)$. The next lemma helps us control the cluster sizes.

Lemma 4.3. *The cluster sizes follow negative binomial distribution under $\overline{\nu}_s$, i.e.,*

$$\overline{\nu}_s(w(T) = k) = (k-1)e^{-2(T-s)}\left(1 - e^{-(T-s)}\right)^{k-2} \text{ for } k \geq 2 \text{ and } s \in [0,T). \tag{4.3}$$

As a consequence, $w(T)$ has exponentially decaying tail under $\overline{\nu}_0$. Further,

$$\overline{\nu}_0 \underset{\text{st}}{\geq} \overline{\nu}_s \underset{\text{st}}{\geq} \nu_{s,\delta} \text{ for } s \in [0,T],$$

where the second inequality is true for large n. Here "$\underset{\text{st}}{\geq}$" denotes stochastic domination.

Proof: (4.3) is a standard fact about birth processes and can be derived from, for example, [5, Equation (1)]. The first stochastic domination is immediate and the second one is a consequence of (4.1). ■

4.2 Proof of Theorem 2.1

We first make note of an inequality that will be used later in the proof. An application of Lemma 4.1 together with (iii) of Lemma 4.2 yields

$$\rho_t(a_{n,\delta}, b_{n,\delta}, c_{n,\delta}) \leq \rho_t(a,b,c) + \frac{\beta_0 C_1}{n^\gamma} \leq 1 - C(t_c - t) + \frac{\beta_0 C_1}{n^\gamma} \leq 1 - \beta(t_c - t) \tag{4.4}$$

for sufficiently large n and a positive universal constant β. Here we have used the fact that $t \leq t_c - \lambda n^{-1/3}$ and $\gamma > 1/3$. This shows that

$$\Delta_{n,t} \geq \beta(t_c - t) > 0 \tag{4.5}$$

for sufficiently large n, where

$$\Delta = \Delta_{n,t} := 1 - \rho_t(a_{n,\delta}, b_{n,\delta}, c_{n,\delta}). \tag{4.6}$$

The next few steps consist of reducing the problem to getting an upper bound on the total progeny of a suitable branching process. From Lemma 4.2, we get

$$\begin{aligned} \mathbb{P}(L_1^{BF}(t) > m) &\leq nT\left[\mathbb{P}(C_n^0(t) > m, E_n^c) + \mathbb{P}(E_n)\right] \\ &\leq nT\left[\mathbb{P}(C_{n,\delta}^{RG}(t) > m) + \exp(-Cn^{1-2\gamma})\right] \\ &= nT\Big[\int_W \mathbb{P}_{w_0}(C_{n,\delta}^{RG}(t) > m)\ \nu_{0,\delta}(dw_0) + \exp(-Cn^{1-2\gamma})\Big], \end{aligned} \tag{4.7}$$

where $\mathbb{P}_{w_0}(\cdot) = \mathbb{P}\left(\,\cdot\,\middle|\,(0,w_0) \in RG_{n,t}(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})\right)$ for every fixed $w_0 \in W$.

Write $k_{t,\delta}$ for $k_t[b_{n,\delta}]$ and recall that $\mu_\delta = \mu[a_{n,\delta}, c_{n,\delta}]$. Consider a branching process on $X_t := [0,t] \times W$ as follows: Define $x^0 = (0, w_0)$ to be generation zero of the branching process. For $k \geq 0$, denote the total number of points in generation k by N_k (thus $N_0 = 1$), and let $x_1^{(k)}, \ldots, x_{N_k}^{(k)}$ denote the points in generation k (note that $x^0 = x_1^{(0)}$). Then for $k \geq 0$ and $1 \leq i \leq N_k$, $x_i^{(k)}$ gives birth to its own offsprings according to a Poisson process with intensity $k_{t,\delta}(x_i^{(k)}, y)\ \mu_\delta(dy)$ independent of the other points in generation k.

Let $G(x^0) := \sum_{k=0}^{\infty} \sum_{i=1}^{N_k} \phi_t(x_i^{(k)})$ be the total progeny. By a breadth-first search argument (see [3, Lemma 6.12]),

$$\mathbb{P}_{w_0}(C_{n,\delta}^{RG}(t) > m) \leq \mathbb{P}(G(x^0) > m). \tag{4.8}$$

From (4.7) and (4.8), we have

$$\mathbb{P}(L_1^{BF}(t) > m) \le nT\Big[\int_W \mathbb{P}\big(G\left((0,w_0)\right) > m\big)\ \nu_{0,\delta}(dw_0) + \exp(-Cn^{1-2\gamma})\Big]. \tag{4.9}$$

We will use the next two lemmas to bound $\mathbb{P}\big(G\left((0,w_0)\right) > m\big)$. Let $K_{t,\delta}$ be the integral operator on $L^2(X,\mu_\delta)$ associated with the kernel $k_{t,\delta}$.

Lemma 4.4. *Let $f_\infty := \phi_t + \sum_{i=1}^\infty (1+\Delta/2)^i K_{t,\delta}^i \phi_t$. Then there exists a universal constant C_3 such that for every $(s,w) \in X_t = [0,t]\times W$,*

$$f_\infty(s,w) \le C_3 w(T)/\Delta.$$

Lemma 4.5. *Set $f_0 = \phi_t$ and define*

$$f_\ell := \phi_t + \sum_{i=1}^{\ell}(1+\Delta/2)^i K_{t,\delta}^i \phi_t \text{ for } \ell = 1,2,\dots.$$

Fix $x = (s,w) \in X_t$, and let $y_1,\dots,y_N$ be the points of a Poisson process in X_t having intensity $k_{t,\delta}(x,y)\mu_\delta(dy)$. Then there exists $\eta_0 > 0$ (independent of x) such that for every $\eta \in (0,\eta_0]$,

$$\mathbb{E}\exp\Big(\eta\Delta^2\sum_{i=1}^N f(y_i)\Big) \le \exp\Big(\eta\Delta^2(1+\Delta/2)K_{t,\delta}f(x)\Big)$$

for every $f \in \{f_\ell : \ell \ge 0\}$.

Lemma 4.5 will be used to get a sharp tail bound on $G((0,w_0))$. A restrictive version of this result was obtained in [3, Lemma 6.14], where the authors used a direct truncation argument. [3, Lemma 6.14] leads to a bound of the form $O((\log n)^4/(t_c-t)^2)$, whereas the stronger result in Lemma 4.5 will yield the bound $O(\log n/(t_c-t)^2)$.

To simplify notation, we will write $\rho_{t,\delta}$ for $\rho_t(a_{n,\delta}, b_{n,\delta}, c_{n,\delta})$, and a_δ (resp. b_δ, c_δ) for $a_{n,\delta}$ (resp. $b_{n,\delta}, c_{n,\delta}$) throughout the rest of this section.

Proof of Lemma 4.4: First note that

$$\begin{aligned}\|\phi_t\|^2_{L^2(\mu_\delta)} &= \int_{X_t} w_1(t)^2 \mu_\delta(d(s_1,w_1)) \qquad (4.10)\\ &= \int_0^t a_\delta(s_1)\mathbb{E}_{\nu_{s_1,\delta}}[w_1(T)^2]ds_1 \le \mathbb{E}_{\overline{\nu}_0}\left[w_1(T)^2\right]\int_0^T a_\delta(s_1)ds_1 < \infty,\end{aligned}$$

where the third step uses Lemma 4.3. Note also that

$$\begin{aligned}K_{t,\delta}^i\phi_t(s,w) &= \int_{X_t} k_{t,\delta}^{(i)}\big((s,w),(s_1,w_1)\big)\phi_t(s_1,w_1)\ \mu_\delta(d(s_1,w_1)) \qquad (4.11)\\ &\le \|k_{t,\delta}^{(i)}((s,w),\ \cdot\ )\|_{L^2(\mu_\delta)}\|\phi_t\|_{L^2(\mu_\delta)}\\ &\le C\|K_{t,\delta}^{i-1}k_{t,\delta}((s,w),\ \cdot\ )\|_{L^2(\mu_\delta)} \le C\rho_{t,\delta}^{i-1}\|k_{t,\delta}((s,w),\ \cdot\ )\|_{L^2(\mu_\delta)}.\end{aligned}$$

Now

$$k_{t,\delta}((s,w),(s_1,w_1)) = \int_0^t w(u)w_1(u)b_\delta(u)du \le Cw(T)w_1(T). \tag{4.12}$$

Hence

$$\begin{aligned}\|k_{t,\delta}((s,w),.)\|^2_{L^2(\mu_\delta)} &= \int_{X_t} k_{t,\delta}((s,w),(s_1,w_1))^2\mu_\delta(d(s_1,w_1)) \qquad (4.13)\\ &\le C\int_0^t a_\delta(s_1)ds_1\int_{w_1\in W} w^2(T)w_1^2(T)\nu_{s_1,\delta}(dw_1)\\ &\le Cw^2(T)\mathbb{E}_{\overline{\nu}_0}\left[w_1(T)^2\right]\int_0^T a_\delta(s_1)ds_1 \le C'w^2(T),\end{aligned}$$

where the penultimate step again uses Lemma 4.3. From (4.10), (4.11) and (4.13), we get

$$\begin{aligned} f_\infty(s,w) &= \phi_t(s,w) + \sum_{i=1}^{\infty}(1+\Delta/2)^i K^i_{t,\delta}\phi_t(s,w) \\ &\le w(T) + C\sum_{i=1}^{\infty}(1+\Delta/2)^i \rho^{i-1}_{t,\delta} w(T) \\ &\le C' w(T)\left(1+\frac{1}{1-(1-\Delta)(1+\Delta/2)}\right) \le \frac{C_3 w(T)}{\Delta}, \end{aligned}$$

which is the desired bound. ∎

Proof of Lemma 4.5: By properties of Poisson processes,

$$\mathbb{E}\exp\left(\eta\Delta^2\sum_{i=1}^{N} f(y_i)\right) = \exp\left(\int_{X_t} k_{t,\delta}(x,u)\big(\exp(\eta\Delta^2 f(u))-1\big)\mu_\delta(du)\right). \tag{4.14}$$

Now

$$\begin{aligned} &\int_{X_t} k_{t,\delta}(x,u)(\exp(\eta\Delta^2 f(u))-1)\mu_\delta(du) \qquad (4.15)\\ &= \int_{X_t} k_{t,\delta}(x,(s_1,w_1))(\exp(\eta\Delta^2 f(s_1,w_1))-1)\mu_\delta(d(s_1,w_1)) \\ &= \int_{s_1=0}^{t}\int_{w_1\in W}\left(\int_{s_1}^{t} w(v)w_1(v)b_\delta(v)dv\right)\left(\exp\left(\eta\Delta^2 f(s_1,w_1)\right)-1\right)\nu_{s_1,\delta}(dw_1)a_\delta(s_1)ds_1 \\ &= \int_{s_1=0}^{t}\int_{v=s_1}^{t} w(v)b_\delta(v)a_\delta(s_1)\mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(v)\left(\exp\left(\eta\Delta^2 f(s_1,w_1)\right)-1\right)\Big]dvds_1. \end{aligned}$$

Fix $s_1\in(0,t)$ and $v\in(s_1,t)$. Then

$$\begin{aligned} &\mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(v)(\exp(\eta\Delta^2 f(s_1,w_1))-1)\Big] \qquad (4.16)\\ &\quad= \eta\Delta^2\mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(v)\Big(f+\eta\Delta^2 f^2/2! + \sum_{j=2}^{\infty}(\eta\Delta^2)^j f^{j+1}/(j+1)!\Big)\Big] \\ &\quad=: \eta\Delta^2\mathbb{E}_{\nu_{s_1,\delta}}(T_1+T_2+T_3). \end{aligned}$$

We now have to get upper bounds on $\mathbb{E}_{\nu_{s_1,\delta}}(T_3)$ and $\mathbb{E}_{\nu_{s_1,\delta}}(T_2)$. To this end, note that $\Delta^{2j}\le\Delta^{j+2}$ for $j\ge 2$ and $f\le f_\infty$. Hence from Lemma 4.4,

$$\begin{aligned} \mathbb{E}_{\nu_{s_1,\delta}}(T_3) &\le \mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(v)\sum_{j=2}^{\infty}(\eta\Delta^2)^j\frac{(C_3 w_1(T))^{j+1}}{\Delta^{j+1}(j+1)!}\Big] \\ &\le \mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(T)\sum_{j=2}^{\infty}\eta^j\Delta\frac{(C_3 w_1(T))^{j+1}}{(j+1)!}\Big] \\ &\le C_3^2\eta\Delta\times\mathbb{E}_{\nu_{s_1,\delta}}\Big[w_1(T)^3\sum_{j=2}^{\infty}\frac{(C_3\eta w_1(T))^{j-1}}{(j+1)!}\Big] \\ &\le C_3^2\eta\Delta\times\mathbb{E}_{\overline{\nu}_0}\Big[w_1(T)^3\exp(C_3\eta w_1(T))\Big] \\ &\le C_3^2\eta\Delta\times\big[\mathbb{E}_{\overline{\nu}_0}(w_1(T)^6)\big]^{1/2}\big[\mathbb{E}_{\overline{\nu}_0}(\exp(2C_3\eta w_1(T)))\big]^{1/2}. \end{aligned}$$

Since w_1 has an exponentially decaying tail (Lemma 4.3), we can choose $\eta_1>0$ small so that

$$\mathbb{E}_{\overline{\nu}_0}\left[\exp(2C_3\eta_1 w_1(T))\right] < \infty,$$

and for $0 \le \eta \le \eta_1$,

$$\mathbb{E}_{\nu_{s_1,\delta}}(T_3) \le \Delta \le \frac{\Delta}{4}\mathbb{E}_{\nu_{s_1,\delta}}\big[w_1(v)f(s_1,w_1)\big]. \tag{4.17}$$

Here the last inequality uses the fact that if $v \in (s_1,t)$ and w_1 is distributed as $\nu_{s_1,\delta}$, then $w_1(v) \ge 2$ and $f(s_1,w_1) \ge \phi_t(s_1,w_1) = w_1(t) \ge 2$.

It follows from Lemma 4.4 that $f(s_1,w_1) \le f_\infty(s_1,w_1) \le C_3 w_1(T)/\Delta$ whenever $f \in \{f_\ell : \ell \ge 0\}$. Hence

$$\mathbb{E}_{\nu_{s_1,\delta}} T_2 = \mathbb{E}_{\nu_{s_1,\delta}}\big[w_1(v)\eta\Delta^2 f^2/2\big] \le \frac{C_3\eta\Delta}{2}\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2 f(s_1,w_1)\right]. \tag{4.18}$$

Suppose $f = f_\ell = \phi_t + \sum_{i=1}^{\ell}(1+\Delta/2)^i K_{t,\delta}^i\phi_t$. Choose an $\eta_2 > 0$ so that

$$C_3\eta_2\mathbb{E}_{\overline{\nu}_0}\left[w_1(T)^3\right] \le 2. \tag{4.19}$$

Then for $0 < \eta \le \eta_2$,

$$\begin{aligned}\frac{C_3\eta\Delta}{2}\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2\phi_t(s_1,w_1)\right] &\le \frac{C_3\eta\Delta}{2}\mathbb{E}_{\overline{\nu}_0}\left[w_1(T)^3\right] \\ &\le \Delta \le \frac{\Delta}{4}\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(v)\phi_t(s_1,w_1)\right].\end{aligned} \tag{4.20}$$

For $i \ge 1$, let $g_i = K_{t,\delta}^{i-1}\phi_t$. Then

$$\begin{aligned}&\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2 K_{t,\delta}^i\phi_t(s_1,w_1)\right] \\ &\quad = \mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2 K_{t,\delta} g_i(s_1,w_1)\right] \\ &\quad = \mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2\int_{X_t}\left(\int_{z=s_1}^t w_1(z)w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\mu_\delta(d(s_2,w_2))\right] \\ &\quad = \int_{X_t}\left[\left(\int_{z=s_1}^t \mathbb{E}_{\nu_{s_1,\delta}}\left(w_1(T)^2 w_1(z)\right)w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\right]\mu_\delta(d(s_2,w_2)) \\ &\quad \le \int_{X_t}\left[\left(\int_{z=s_1}^t \mathbb{E}_{\overline{\nu}_0}\left(w_1(T)^3\right)w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\right]\mu_\delta(d(s_2,w_2)),\end{aligned} \tag{4.21}$$

where the second equality follows from the definition of $K_{t,\delta}$ and the final step uses the stochastic domination relation between $\overline{\nu}_0$ and $\nu_{s_1,\delta}$ (Lemma 4.3). Hence for $0 < \eta \le \eta_2$,

$$\begin{aligned}&\frac{C_3\eta\Delta}{2}\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(T)^2 K_{t,\delta}^i\phi_t(s_1,w_1)\right] \\ &\quad \le \frac{C_3\eta\Delta}{2}\mathbb{E}_{\overline{\nu}_0}\left[w_1(T)^3\right]\int_{X_t}\left[\left(\int_{z=s_1}^t w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\right]\mu_\delta(d(s_2,w_2)) \\ &\quad \le \Delta\int_{X_t}\left[\left(\int_{z=s_1}^t w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\right]\mu_\delta(d(s_2,w_2)) \\ &\quad \le \frac{\Delta}{4}\int_{X_t}\left[\left(\int_{z=s_1}^t \mathbb{E}_{\nu_{s_1,\delta}}\left(w_1(v)w_1(z)\right)w_2(z)b_\delta(z)dz\right)g_i(s_2,w_2)\right]\mu_\delta(d(s_2,w_2)) \\ &\quad = \frac{\Delta}{4}\mathbb{E}_{\nu_{s_1,\delta}}\left[w_1(v)\int_{X_t}k_{t,\delta}((s_1,w_1),(s_2,w_2))g_i(s_2,w_2)\mu_\delta(d(s_2,w_2))\right] \\ &\quad = \frac{\Delta}{4}\mathbb{E}_{\nu_{s_1,\delta}}\left(w_1(v)K_{t,\delta}^i\phi_t(s_1,w_1)\right),\end{aligned} \tag{4.22}$$

where the second inequality follows from (4.19), and the third inequality is a consequence of the following observation: $w_1(v), w_1(z) \geq 2$ under $\nu_{s_1,\delta}$ when $v, z \in (s_1, t)$. Combining (4.18), (4.20), and (4.22), we get

$$\mathbb{E}_{\nu_{s_1,\delta}} T_2 \leq \frac{\Delta}{4} \mathbb{E}_{\nu_{s_1,\delta}} \left[w_1(v) f(s_1, w_1) \right]. \tag{4.23}$$

From (4.16), (4.17), and (4.23),

$$\mathbb{E}_{\nu_{s_1,\delta}} \left[w_1(v)(\exp(\eta\Delta^2 f(s_1, w_1)) - 1) \right] \leq \eta\Delta^2 \left(1 + \Delta/2\right) \mathbb{E}_{\nu_{s_1,\delta}} \left[w_1(v) f(s_1, w_1) \right]. \tag{4.24}$$

From (4.24) and (4.15), we get

$$\begin{aligned} &\int_{X_t} k_{t,\delta}(x, u) \left(\exp(\eta\Delta^2 f(u)) - 1 \right) \mu_\delta(du) \qquad (4.25) \\ &\leq \eta\Delta^2 \left(1 + \Delta/2\right) \int_{s_1=0}^{t} \int_{v=s_1}^{t} w(v) b_\delta(v) a_\delta(s_1) \mathbb{E}_{\nu_{s_1,\delta}} \left[w_1(v) f(s_1, w_1) \right] dv ds_1 \\ &= \eta\Delta^2 (1 + \Delta/2) \int_{s_1=0}^{t} a_\delta(s_1) \mathbb{E}_{\nu_{s_1,\delta}} \left[k_{t,\delta}(x, (s_1, w_1)) f(s_1, w_1) \right] ds_1 \\ &= \eta\Delta^2 (1 + \Delta/2) K_{t,\delta} f(x), \end{aligned}$$

for $0 < \eta \leq \eta_0 := \eta_1 \wedge \eta_2$. This together with (4.14) yields the result. ■

Completing the proof of Theorem 2.1. Now the proof of Theorem 2.1 becomes routine and can be finished by a Chernoff bound argument. Let $G_i := \sum_{j=1}^{N_i} \phi_t(x_j^{(i)})$ be the total volume of points in generation k, and let $\mathcal{F}_k$ be the σ-field generated by $\{x_i^{(j)} : i \leq N_j, j \leq k\}$.

Lemma 4.6. *For* $0 < \eta \leq \eta_0$*,*

$$\mathbb{E}\Big(\exp\big(\eta\Delta^2 \sum_{i=j}^{k} G_i\big) \Bigm| \mathcal{F}_j \Big) \leq \exp\Big(\eta\Delta^2 \sum_{i=1}^{N_j} f_{k-j}(x_i^{(j)})\Big). \tag{4.26}$$

Proof : The assertion is immediate for $j = k$. Assume that it is true for $\ell + 1 \leq j \leq k$. Then

$$\begin{aligned} &\mathbb{E}\Big(\exp\Big(\eta\Delta^2 \sum_{i=\ell}^{k} G_i\Big) \Bigm| \mathcal{F}_\ell \Big) \\ &= \exp(\eta\Delta^2 G_\ell) \mathbb{E}\Big[\mathbb{E}\Big(\exp\Big(\eta\Delta^2 \sum_{i=\ell+1}^{k} G_i\Big) \Bigm| \mathcal{F}_{\ell+1} \Big) \Bigm| \mathcal{F}_\ell \Big] \\ &\leq \exp(\eta\Delta^2 G_\ell) \mathbb{E}\Big[\exp\Big(\eta\Delta^2 \sum_{i=1}^{N_{\ell+1}} f_{k-\ell-1}(x_i^{(\ell+1)})\Big) \Bigm| \mathcal{F}_\ell \Big] \\ &\leq \exp(\eta\Delta^2 G_\ell) \exp\Big(\eta\Delta^2 (1 + \Delta/2) \sum_{i=1}^{N_\ell} K_{t,\delta} f_{k-\ell-1}(x_i^{(\ell)})\Big) = \exp\Big(\eta\Delta^2 \sum_{i=1}^{N_\ell} f_{k-\ell}(x_i^{(\ell)})\Big), \end{aligned}$$

where the second step follows from the induction hypothesis and the third step follows from Lemma 4.5. This proves our claim. ■

Setting $j = 0$ and letting k tend to infinity in (4.26), we get

$$\begin{aligned} \mathbb{E}\left[\exp(\eta_0 \Delta^2 G(x^0)) \right] &\leq \exp\left(\eta_0 \Delta^2 f_\infty\left((0, w_0)\right) \right) \\ &\leq \exp\left(\eta_0 \Delta C_3 w_0(T) \right) \leq \exp(C_3 \eta_0 w_0(T)), \end{aligned}$$

where the second inequality is a consequence of Lemma 4.4. Hence

$$\begin{aligned}\int_W \mathbb{P}\Big(G((0,w_0)) > m\Big)\nu_{0,\delta}(dw_0) &\le \exp(-\eta_0\Delta^2 m)\int_W \mathbb{E}\Big[\exp\big(\eta_0\Delta^2 G((0,w_0))\big)\Big]\nu_{0,\delta}(dw_0)\\ &\le \exp(-\eta_0\Delta^2 m)\int_W \exp\left(C_3\eta_0 w_0(T)\right)\ \overline{\nu}_0(dw_0)\\ &= C_4\exp(-\eta_0\Delta^2 m) \le C_4\exp(-\eta_0\beta^2(t_c-t)^2 m), \quad (4.27)\end{aligned}$$

where the last step follows from (4.5). The constant C_4 is finite by the choice of η_0. From (4.27) and (4.9), we see that

$$\mathbb{P}\Big(L_1^{BF}(t) > \frac{2\log n}{\eta_0\beta^2(t_c-t)^2}\Big) \overset{n\to\infty}{\longrightarrow} 0,$$

which completes the proof of Theorem 2.1.

4.3 Proof of Lemma 4.1

Recall the definitions of $a(\cdot)$, $b(\cdot)$, and $c(\cdot)$ from (3.3). As in the previous section, we will write $a_\delta(u), b_\delta(u)$, and $c_\delta(u)$ for $a_{n,\delta}(x(u)), b_{n,\delta}(x(u))$ and $c_{n,\delta}(x(u))$ respectively throughout this section. We will write k_u for $k_u[b]$ and $k_{u,\delta}$ for $k_u[b_\delta]$. Recall that μ, μ_δ, and $\overline{\mu}$ stand for $\mu[a,c]$, $\mu[a_\delta,c_\delta]$, and $\mu[1,1]$ respectively. Then

$$\begin{aligned}|\rho_u(a,b,c) - \rho_u(a_\delta,b_\delta,c_\delta)| &\le |\rho_u(a,b,c) - \rho_u(a_\delta,b,c_\delta)| + |\rho_u(a_\delta,b,c_\delta) - \rho_u(a_\delta,b_\delta,c_\delta)|\\ &=: T_1 + T_2. \quad (4.28)\end{aligned}$$

We first bound the term T_1, as bounding T_2 is considerably simpler. Writing $\mathfrak{n}(k^*,\mu^*)$ for the norm, in $L^2(\mu^*)$, of the integral operator associated with the kernel k^*, note that $\rho_u(a,b,c) = \mathfrak{n}(k_u,\mu)$ and $\rho_u(a_\delta,b,c_\delta) = \mathfrak{n}(k_u,\mu_\delta)$. Let $h = d\mu/d\mu_\delta$ be as in (iv) of Lemma 4.2. Since $\mathfrak{n}(k_u(x,y),\mu) = \mathfrak{n}(k_u(x,y)\sqrt{h(x)h(y)},\mu_\delta)$,

$$\begin{aligned}T_1 &= \left|\mathfrak{n}\left(k_u(x,y)\sqrt{h(x)h(y)},\mu_\delta\right) - \mathfrak{n}(k_u(x,y),\mu_\delta)\right| \quad (4.29)\\ &\le \mathfrak{n}\Big(k_u(x,y)\big(\sqrt{h(x)h(y)}-1\big),\mu_\delta\Big)\\ &\le \left(\int_X\int_X k_u^2(x,y)\big(\sqrt{h(x)h(y)}-1\big)^2\mu_\delta(dx)\mu_\delta(dy)\right)^{1/2},\end{aligned}$$

where the second step uses the triangle inequality, and the third step uses the fact

$$\mathfrak{n}(k^*,\mu^*)^2 \le \int\int k^*(x,y)^2\mu^*(dx)\mu^*(dy).$$

Recall the notation $\gamma_i(\cdot,\cdot)$ used in the expression of h from (4.2). Writing $x = (s_1,w_1)$, $y = (s_2,w_2)$ and $\tau_i = \gamma_i(x)$, $\xi_j = \gamma_j(y)$, we have $L \le \sqrt{h(x)h(y)} \le U$, where

$$L := \left(\frac{a(s_1)}{a_\delta(s_1)}\frac{a(s_2)}{a_\delta(s_2)}\prod_{i=1}^{w_1(T)-2}\frac{c(\tau_i)}{c_\delta(\tau_i)}\prod_{j=1}^{w_2(T)-2}\frac{c(\xi_j)}{c_\delta(\xi_j)}\right)^{\frac{1}{2}} \text{ and } U := \exp\left(\frac{\delta T}{2}\big(w_1(T)+w_2(T)\big)\right). \quad (4.30)$$

From (4.12) and (4.29),

$$\begin{aligned} T_1 &\le C\Big(\int_X \int_X w_1(T)^2 w_2(T)^2 \left[(U-1)^2 + (L-1)^2\right] \mu_\delta(dx)\mu_\delta(dy)\Big)^{1/2} \\ &\le C \left(\int_X \int_X w_1(T)^2 w_2(T)^2 (U-1)^2 \mu_\delta(dx)\mu_\delta(dy)\right)^{1/2} \\ &\quad + C \left(\int_X \int_X w_1(T)^2 w_2(T)^2 (L-1)^2 \mu_\delta(dx)\mu_\delta(dy)\right)^{1/2} \\ &=: C(T_3 + T_4). \end{aligned} \tag{4.31}$$

Note that

$$\begin{aligned} (U-1)^2 \le 2\Big(\exp\left(\frac{\delta T w_1(T)}{2}\right) - 1\Big)^2 \exp\left(\delta T w_2(T)\right) \\ + 2\Big(\exp\left(\frac{\delta T w_2(T)}{2}\right) - 1\Big)^2 \exp\left(\delta T w_1(T)\right). \end{aligned}$$

This together with the simple inequality $e^x - 1 \le x e^x$ yields

$$\begin{aligned} T_3^2 &\le 4 \int\int w_1^2(T) w_2^2(T) \Big(\exp\left(\frac{\delta T w_1(T)}{2}\right) - 1\Big)^2 \exp\left(\delta T w_2(T)\right) \mu_\delta(dx)\mu_\delta(dy) \\ &\le C\delta^2 \Big[\int w_1^4(T) \exp\big(\delta T w_1(T)\big) \mu_\delta(dx)\Big] \Big[\int w_2^2(T) \exp\left(\delta T w_2(T)\right) \mu_\delta(dy)\Big] \\ &\le C'\delta^2 \mathbb{E}_{\overline{\nu}_0}\left[w_1^4(T)\exp\left(\delta T w_1(T)\right)\right] \mathbb{E}_{\overline{\nu}_0}\left[w_2^2(T)\exp\left(\delta T w_2(T)\right)\right]. \end{aligned} \tag{4.32}$$

Since the tail of $w_1(T)$ decays exponentially, we conclude that for large n (so that $\delta = \delta_n$ is sufficiently small),

$$T_3 \le C\delta. \tag{4.33}$$

Next, from the definition of T_4 in (4.31) and the definition of L in (4.30), it follows that

$$T_4^2 \le C \int\int w_1^2(T) w_2^2(T) (L_1 + L_2 + L_3 + L_4) \mu_\delta(dx)\mu_\delta(dy), \tag{4.34}$$

where

$$L_1 = \left(1 - \sqrt{\frac{a(s_1)}{a_\delta(s_1)}}\right)^2, \quad L_2 = \left(1 - \sqrt{\frac{a(s_2)}{a_\delta(s_2)}}\right)^2,$$

$$L_3 = \left(1 - \prod_{i=1}^{w_1(T)-2} \sqrt{\frac{c(\tau_i)}{c_\delta(\tau_i)}}\right)^2, \text{ and } L_4 = \left(1 - \prod_{j=1}^{w_2(T)-2} \sqrt{\frac{c(\xi_j)}{c_\delta(\xi_j)}}\right)^2.$$

Note that $\inf_{[0,T]} a(s) =: m_1 > 0$ and $0 \le a_\delta(s) - a(s) \le \delta$. Hence

$$\begin{aligned} &\int\int w_1^2(T) w_2^2(T) L_1 \ \mu_\delta(dx)\mu_\delta(dy) \\ &\quad \le \int\int w_1^2(T) w_2^2(T) \left(1 - \frac{a(s_1)}{a_\delta(s_1)}\right)^2 \mu_\delta(dx)\mu_\delta(dy) \\ &\quad \le \frac{\delta^2}{m_1^2} \int\int w_1^2(T) w_2^2(T) \mu_\delta(dx)\mu_\delta(dy) \le C\delta^2 \left(\mathbb{E}_{\overline{\nu}_0} w_1^2(T)\right)^2. \end{aligned} \tag{4.35}$$

The integrand corresponding to L_2 can be handled in the same way. Next note that $\inf_{[0,\epsilon]}(c(t))' = \inf_{[0,\epsilon]} c_0'(x(t))x'(t) > 0$ (see [8]) whenever $\epsilon > 0$ is small enough. Hence

$$c_\delta(t) \ge \max(\delta, m_2 t) \text{ for } t \in [0,T] \tag{4.36}$$

for a positive constant m_2. Further, $c(t)$ is increasing in an interval $[0, t_0]$ and is bounded away from zero on $[t_0, T]$. Therefore,

$$c_\delta(\tau_i) \geq \min(c_\delta(\tau_1), m_3) \text{ for } 1 \leq i \leq w_1(T) - 2. \tag{4.37}$$

Hence, on the set $w_1(T) \geq 3$,

$$\begin{aligned}
L_3 &\leq w_1(T) \sum_{i=1}^{w_1(T)-2} \left(1 - \sqrt{\frac{c(\tau_i)}{c_\delta(\tau_i)}}\right)^2 \\
&\leq w_1(T) \sum_{i=1}^{w_1(T)-2} \left(1 - \frac{c(\tau_i)}{c_\delta(\tau_i)}\right)^2 \leq w_1(T) \sum_{i=1}^{w_1(T)-2} \frac{\delta^2}{c_\delta^2(\tau_i)} \\
&\leq w_1(T)\delta^2 \sum_{i=1}^{w_1(T)-2} \left(\frac{1}{m_3} + \frac{1}{c_\delta(\tau_1)}\right)^2 \leq C\delta^2 w_1(T)^2 \frac{1}{c_\delta(\tau_1)^2},
\end{aligned}$$

where the penultimate step follows from (4.37). We thus have

$$\begin{aligned}
&\int\int w_1^2(T) w_2^2(T) L_3 \; \mu_\delta(dx)\mu_\delta(dy) \qquad (4.38) \\
&\leq C\delta^2 \int\int w_2^2(T) \frac{w_1^4(T)}{c_\delta(\tau_1)^2} \mathbb{I}_{\{w_1(T)\geq 3\}} \mu_\delta(dx)\mu_\delta(dy) \\
&\leq C'\delta^2 \int \frac{w_1^4(T)}{c_\delta(\tau_1)^2} \mathbb{I}_{\{w_1(T)\geq 3\}} \mu_\delta(dx) \\
&\leq C'\delta^2 \left(\int w_1^{4p}(T)\mu_\delta(dx)\right)^{1/p} \left(\int \frac{1}{c_\delta(\tau_1)^{2q}} \mathbb{I}_{\{w_1(T)\geq 3\}} \mu_\delta(dx)\right)^{1/q} \\
&\leq C''\delta^2 \left(\int \frac{1}{c_\delta(\tau_1)^{2q}} \mathbb{I}_{\{w_1(T)\geq 3\}} \mu_\delta(dx)\right)^{1/q}.
\end{aligned}$$

Here we choose $p, q > 1$ so that $p^{-1} + q^{-1} = 1$ and $2q = 2 + \theta$ with $0 < \theta < 1$. Define

$$\overline{\tau}_1 := \tau_1 \cdot \mathbb{I}\{w_1(T) \geq 3\} + T \cdot \mathbb{I}\{w_1(T) = 2\}.$$

Then

$$\begin{aligned}
&\int \frac{1}{c_\delta(\tau_1)^{2q}} \mathbb{I}_{\{w_1(T)\geq 3\}} \mu_\delta(dx) \leq \int \frac{1}{c_\delta(\overline{\tau}_1)^{2+\theta}} \mu_\delta(dx) \qquad (4.39) \\
&= \int_{s_1=0}^T a_\delta(s_1) ds_1 \mathbb{E}_{\nu_{s_1,\delta}} \Big[\frac{1}{c_\delta(\overline{\tau}_1)^{2+\theta}}\Big] \\
&= \int_{s_1=0}^T a_\delta(s_1) ds_1 \Big[\int_{u=s_1}^T \frac{1}{c_\delta(u)^{2+\theta}} \exp\Big(-2\int_{s_1}^u c_\delta(z)dz\Big) 2c_\delta(u)du + \frac{\nu_{s_1,\delta}\{w_1(T)=2\}}{c_\delta(T)^{2+\theta}}\Big] \\
&\quad \leq C \int_{s_1=0}^T ds_1 \Big[\int_{u=s_1}^T \frac{du}{(m_2 u)^{1+\theta}} + \frac{1}{(m_2 T)^{2+\theta}}\Big].
\end{aligned}$$

In the third step we have simply used the density of τ_1 to evaluate the expectation, and the last inequality is a consequence of (4.36). Since $\theta < 1$, the last integral is finite. A similar analysis can be carried out for the integrand corresponding to L_4. Combining (4.33), (4.35), (4.38) and (4.39), we get

$$T_1 \leq C\delta. \tag{4.40}$$

Finally,

$$\begin{aligned} T_2^2 &\le \mathfrak{n}\big((k_{u,\delta} - k_u),\ \mu_\delta\big)^2 \\ &\le \int_X \int_X \Big(k_{u,\delta}(x,y) - k_u(x,y)\Big)^2 \mu_\delta(dx)\mu_\delta(dy) \\ &\le \int_X \int_X \Big(\int_0^T w_1(z)w_2(z)(b_\delta(z) - b(z))dz\Big)^2 \mu_\delta(dx)\mu_\delta(dy) \\ &\le C\delta^2 \Big(\int w_1^2(T)\overline{\mu}(dx)\Big)^2 = C'\delta^2. \end{aligned} \tag{4.41}$$

Combining (4.28), (4.40), and (4.41) completes the proof.

4.4 Proof of Corollary 2.2

Let $Y_n(s)$ denote the number of edges in $BF_n(s)$. Define a process $\overline{BF}_n(\cdot)$ by

$$\overline{BF}_n(2Y_n(s)/n) := BF_n(s) \text{ for } s \ge 0,$$

and extend the definition to $\mathbb{R}_+$ by right continuity. Then $\overline{BF}_n(\cdot)$ has the same distribution as $DBF_n(\cdot)$. Let $L_1^{\overline{BF}}(s)$ denote the size of the largest component of $\overline{BF}_n(s)$. Let us assume that $t_c/2 \le t \le t_c - \lambda n^{-1/3}$, since for $t \le t_c/2$ the desired bound will follow directly. We have

$$\begin{aligned} \mathbb{P}\Big(L_1^{BF}\Big(t + \frac{\log n}{\sqrt{n}}\Big) \ge m\Big) &\ge \mathbb{P}\Big(L_1^{BF}\Big(t + \frac{\log n}{\sqrt{n}}\Big) \ge m,\ Y_n\Big(t + \frac{\log n}{\sqrt{n}}\Big) \ge \frac{nt}{2}\Big) \\ &\ge \mathbb{P}\Big(L_1^{\overline{BF}}(t) \ge m,\ Y_n\Big(t + \frac{\log n}{\sqrt{n}}\Big) \ge \frac{nt}{2}\Big). \end{aligned}$$

Hence

$$\mathbb{P}\big(L_1^{\overline{BF}}(t) \ge m\big) \le \mathbb{P}\big(L_1^{BF}(t + \log n/\sqrt{n}) \ge m\big) + \mathbb{P}\big(Y_n(t + \log n/\sqrt{n}) < nt/2\big). \tag{4.42}$$

Let $Z_1, \dots, Z_n$ be i.i.d. Poisson random variables with mean $\mu_n := \frac{1}{2}(1 - 1/n)^2(t + \log n/\sqrt{n})$. Then $Y_n(t + \log n/\sqrt{n}) \overset{d}{=} Z_1 + \ldots + Z_n$. Since $nt/2 - n\mu_n = -\sqrt{n}\log n/2 + O(1)$,

$$\mathbb{P}\Big(Y_n\Big(t + \frac{\log n}{\sqrt{n}}\Big) < \frac{nt}{2}\Big) \le \mathbb{P}\Big(\sum_{j=1}^n (Z_j - \mu_n)/\sqrt{n\mu_n} < -C\log n\Big) \le \Phi(-C\log n) + \frac{C'}{\sqrt{n}},$$

where Φ is the distribution function of a standard Gaussian random variable. The last inequality is a consequence of the classical Berry-Esseen theorem. The result follows once we combine (4.42) and Theorem 2.1 with the last inequality.

References

[1] Aldous, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** no. 2, 812–854. MR-1434128

[2] Bhamidi, S., Budhiraja, A., and Wang, X. (2014). Bounded-size rules: The barely subcritical regime. *Combinatorics, Probability and Computing.* **23** no. 4, 505–538. MR-3217358

[3] Bhamidi, S., Budhiraja, A., and Wang, X. (2015). Aggregation models with limited choice and the multiplicative coalescent. *Random Struct. Alg.* **46** no. 1, 55–116. MR-3291294

[4] Bollobás, B., Janson, S., and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** no. 1, 3–122. MR-2337396

[5] de la Fortelle, A. (2006). Yule process sample path asymptotics. *Elect. Comm. in Probab.* **11**, 193–199. MR-2266709

[6] Janson, S., Łuczak, T., and Ruciński, A. (2000). *Random Graphs*. John Wiley & Sons. New York MR-1782847

[7] Kang, M., Perkins, W., and Spencer, J. (2012). The Bohman-Frieze process near criticality. *Random Struct. Alg.* **43** no. 2, 221–250. MR-3085769

[8] Spencer, J. and Wormald, N. (2007). Birth control for giants. *Combinatorica.* **27** no. 5, 587–628. MR-2375718

Acknowledgments. I thank Joel Spencer for drawing the problem to my attention and for many helpful comments and lively discussions. I also thank Shankar Bhamidi, Amarjit Budhiraja, and Xuan Wang for carefully reading the first draft of this paper and for many useful comments and suggestions; particularly for suggesting an improvement in an earlier version of Lemma 4.1. Finally, I thank an anonymous referee for suggestions on improving the presentation.