



## Critical random forests

James B. Martin and Dominic Yeo

Department of Statistics,  
University of Oxford,  
24-29 St Giles',  
Oxford OX1 3LB, UK.  
*E-mail address:* [martin@stats.ox.ac.uk](mailto:martin@stats.ox.ac.uk)

Faculty of Industrial Engineering and Management  
Technion – Israel Institute of Technology,  
Haifa, Israel.  
*E-mail address:* [yeo@technion.ac.il](mailto:yeo@technion.ac.il)

**Abstract.** Let  $F(N, m)$  denote a random forest on a set of  $N$  vertices, chosen uniformly from all forests with  $m$  edges. Let  $F(N, p)$  denote the forest obtained by conditioning the Erdős-Rényi graph  $G(N, p)$  to be acyclic. We describe scaling limits for the largest components of  $F(N, p)$  and  $F(N, m)$ , in the *critical window*  $p = N^{-1} + O(N^{-4/3})$  or  $m = N/2 + O(N^{2/3})$ . Aldous (1997) described a scaling limit for the largest components of  $G(N, p)$  within the critical window in terms of the excursion lengths of a reflected Brownian motion with time-dependent drift. Our scaling limit for critical random forests is of a similar nature, but now based on a reflected diffusion whose drift depends on space as well as on time.

### 1. Introduction

Let  $G(N, p)$  be the Erdős-Rényi random graph with vertex set  $[N]$ , in which each of the  $\binom{N}{2}$  possible edges appears independently with probability  $p$ . In a seminal paper, Aldous (1997) gave a scaling limit for the joint distribution of the sizes of the largest components of  $G(N, p)$  within the critical window  $p(N) = \frac{1}{N} + O(N^{-4/3})$ . In this regime, the largest components are of order  $N^{2/3}$  (as was shown first by Bollobás, 1984 up to a logarithmic correction, and then by Luczak, 1990); after rescaling by  $N^{2/3}$ , their sizes converge to the lengths of the excursions of a reflected Brownian motion with time-dependent drift. The central result of Aldous (1997) may be written as follows:

---

*Received by the editors September 22th, 2017; accepted July 8th, 2018.*

*2010 Mathematics Subject Classification.* 05C80, 60C05.

*Key words and phrases.* Random forest, random graph, critical window, exploration process.

**Proposition 1.1.** *Let  $\lambda \in \mathbb{R}$ , and consider the sequence of random graphs  $G(N, p)$  with  $p = p(N) = N^{-1} + \lambda N^{-4/3}$ . Let  $C_1^{G(N,p)}, C_2^{G(N,p)}, \dots$  be the sequence of component sizes of  $G(N, p)$  written in non-increasing order, augmented by zeros.*

*Let  $B^\lambda(t), t \geq 0$  be a (time-inhomogeneous) reflected Brownian motion, with drift  $\lambda - t$  at time  $t$ , and  $B^\lambda(0) = 0$ . Let  $\mathcal{C}^{B^\lambda}$  be the lengths of the excursions of  $B^\lambda$ , written in non-increasing order. Then*

$$N^{-2/3} \left( C_1^{G(N,p)}, C_2^{G(N,p)}, \dots \right) \xrightarrow{d} \mathcal{C}^{B^\lambda} \text{ as } N \rightarrow \infty \quad (1.1)$$

*with respect to the  $\ell^2$  topology.*

A similar result may be written for the random graph  $G(N, m)$  (that is, a graph chosen uniformly from all those with vertex set  $[N]$  and with  $m$  edges), in the regime  $m = N/2 + O(N^{2/3})$ .

Aldous's result has been extended in multiple ways. The same Brownian scaling limit has been shown to arise in more general settings including configuration models and inhomogeneous random graphs, provided the tail of the degree distribution is sufficiently light (Turova, 2013; Riordan, 2012; Joseph, 2014; Bhamidi et al., 2010, 2014b; Dhara et al., 2017). In some such cases, finer scaling limits describing the metric structure of the large components, as well as their size, have been obtained, in terms of objects related to the Brownian continuum random tree (Addario-Berry et al., 2010, 2012; Bhamidi et al., 2014a, 2017). When the third moment of the vertex degrees is infinite, different scaling limits arise (Joseph, 2014; Bhamidi et al., 2012; Dhara et al., 2016) which can be described in terms of excursion lengths of the "thinned Lévy processes" introduced in Aldous and Limic (1998). Finally, dynamic models have been studied in which the developing component structure of a random graph process (or more generally a multiplicative coalescent) is described (as a process) by excursions of a Brownian motion or thinned Lévy process whose drift changes with time (Armendariz, 2001; Broutin and Marckert, 2016; Limic, 2016; Martin and Ráth, 2017).

In this paper we develop in a new direction, to consider the sizes of trees in random forests. Write  $F(N, m)$  for a graph chosen uniformly at random from all forests on  $[N]$  with  $m$  edges (equivalently, those forests consisting of  $N - m$  trees). Write also  $F(N, p)$  for the graph  $G(N, p)$  conditioned to be acyclic. Our main results give a scaling limit for the joint distribution of the sizes of the largest trees in  $F(N, m)$  or in  $F(N, p)$  in the critical regime (this critical regime coincides with that for  $G(N, m)$  and  $G(N, p)$  above). The limit is given by the collection of excursion lengths of a diffusion, as at (1.1) but now the limiting diffusion is inhomogeneous in space as well as in time.

Just as for Aldous's proof of Proposition 1.1, these convergence results are proved by analysing the *graph exploration process*, which encodes enough of the graph structure to recover the sequence of component sizes. We discuss the exploration process, and its scaling limit, in Section 1.3; before that, we introduce the notation needed to state our main results.

**1.1. Definition of the diffusion.** Luczak and Pittel (1992) studied the model  $F(N, m)$ , and identified subcritical, critical and supercritical regimes. Within the critical window, specifically for  $m = N/2 + (\lambda + o(1))N^{2/3}$ , their Theorem 4.1 establishes convergence in distribution for  $N^{-2/3}C_k$ , where  $C_k$  is the size of the  $k$ th largest tree in  $F(N, m)$ .

Luczak and Pittel’s analysis relies on enumeration results of Britikov (1988), which we also use extensively. Britikov gives asymptotics for  $f(N, m)$ , the number of forests on  $[N]$  with exactly  $m$  edges, via Bell polynomials. The relevant regime of Britikov’s result is summarised by Lemma 2.1(ii) of Luczak and Pittel (1992):

**Lemma 1.2.** *For any constant  $c > 0$ , as  $N \rightarrow \infty$ ,*

$$f(N, m) = (1 + o(1)) \frac{\sqrt{2\pi} N^{N-1/6}}{2^{N-m} (N-m)!} g\left(\frac{2m-N}{N^{2/3}}\right), \tag{1.2}$$

*uniformly for  $m \in [N/2 - cN^{2/3}, N/2 + cN^{2/3}]$ , where*

$$g(x) = \frac{1}{\pi} \int_0^\infty \exp(-\frac{4}{3}t^{3/2}) \cos(xt + \frac{4}{3}t^{3/2}) dt, \tag{1.3}$$

*is the density of a stable distribution with parameter  $3/2$ .*

As we shall see in Lemma 2.10, it follows that the asymptotic probability that  $G(N, p)$  is acyclic in this regime is  $\Theta(N^{-1/6})$ .

**Definition 1.3.** For  $b > 0$  and  $\lambda \in \mathbb{R}$ , let

$$\alpha(b, \lambda) := \frac{\int_0^\infty a^{-1/2} g(\lambda - a) \exp\left(\frac{(\lambda-a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da}{\int_0^\infty a^{-3/2} g(\lambda - a) \exp\left(\frac{(\lambda-a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da}. \tag{1.4}$$

**Lemma 1.4.** *The function  $g$  defined in (1.3) is positive, bounded, and uniformly continuous, and satisfies  $g(x) \rightarrow 0$  as  $x \rightarrow \pm\infty$ . The integrals in the numerator and denominator of (1.4) both converge for all  $b$  and  $\lambda$ . The function  $\alpha$  is continuous and increasing in its first argument, and satisfies  $\alpha(b, \lambda) \rightarrow 0$  as  $b \downarrow 0$ , uniformly on  $\lambda$  in compact intervals.*

We will prove Lemma 1.4 in Section 5, after the main probabilistic arguments.

We now define a reflected diffusion,  $Z^\lambda$ , whose drift at time  $s$  and height  $b$  is  $\lambda - s - \alpha(b, \lambda - s)$ . The excursion lengths of  $Z^\lambda$  will describe the scaling limits of the largest trees in our critical random forests. Comparing with the definition of  $B^\lambda$  in Proposition 1.1, we see that the function  $\alpha$  provides the correction to the drift which is required to account for the acyclicity condition.

**Proposition 1.5.** *Consider a standard Brownian motion  $W(\cdot)$  with natural filtration  $\mathcal{F}^W$ . For each  $\lambda \in \mathbb{R}$ , there exists a unique pair of non-negative  $\mathcal{F}^W$ -adapted processes  $Z^\lambda, K^\lambda$  satisfying:*

$$\begin{cases} Z^\lambda(0) = 0, \\ dZ^\lambda(t) = [\lambda - t - \alpha(Z^\lambda(t), \lambda - t)] dt + dW(t) + K^\lambda(t), \end{cases} \tag{1.5}$$

*where  $K^\lambda$  is the local-time process of  $Z^\lambda$  at zero. That is,  $K^\lambda(\cdot)$  is continuous and increasing, with  $K^\lambda(0) = 0$ , and  $\int_0^\infty Z^\lambda(t) dK^\lambda(t) = 0$ .*

Since the drift term in (1.5) is dominated by  $\lambda - t$ ,  $Z^\lambda$  almost surely has a well-defined largest excursion, and second-largest excursion, and so on.

**Definition 1.6.** Let  $\mathcal{C}^\lambda := (C_1^\lambda, C_2^\lambda, \dots)$  be the sequence of lengths of the excursions of  $Z^\lambda$ , written in non-increasing order.

1.2. *Main results.* We can now state the main results of the paper.

**Theorem 1.7.** Fix  $\lambda \in \mathbb{R}$  and suppose that  $m(N)$  is a sequence of integers such that  $m = N/2 + (\lambda + o(1))N^{2/3}$  as  $N \rightarrow \infty$ . Consider the sequence of random forests  $F(N, m)$ . Let  $C_1^{F(N,m)} \geq C_2^{F(N,m)} \geq \dots$  be the sequence of tree sizes in  $F(N, m)$ , in non-increasing order, augmented with zeros. Then

$$N^{-2/3} \left( C_1^{F(N,m)}, C_2^{F(N,m)}, \dots \right) \xrightarrow{d} \mathcal{C}^\lambda \tag{1.6}$$

as  $N \rightarrow \infty$ , with respect to the  $\ell^2$  topology.

**Theorem 1.8.** Fix  $\lambda \in \mathbb{R}$  and suppose that  $p(N)$  is a sequence such that  $p = N^{-1} + (\lambda + o(1))N^{-4/3}$  as  $N \rightarrow \infty$ . Consider the sequence of random forests  $F(N, p)$ . Let  $C_1^{F(N,p)} \geq C_2^{F(N,p)} \geq \dots$  be the sequence of tree sizes in  $F(N, p)$ , in non-increasing order, augmented with zeros. Then

$$N^{-2/3} \left( C_1^{F(N,p)}, C_2^{F(N,p)}, \dots \right) \xrightarrow{d} \mathcal{C}^\lambda \tag{1.7}$$

as  $N \rightarrow \infty$ , with respect to the  $\ell^2$  topology.

We will work mostly in the context of the model  $F(N, p)$ . In Section 6 we will deduce separately that Theorem 1.8 for  $F(N, p)$  implies Theorem 1.7 for  $F(N, m)$ .

1.3. *Exploration processes.* As was the case for Aldous’s Proposition 1.1, our proof of Theorem 1.8 is based on an analysis of the *exploration process* of the graph. To be specific, we will work with a breadth-first ordering (although the argument would work equally well with various other orderings).

Let  $G$  be any graph with vertex set  $[N]$ . We define the breadth-first ordering  $v_1, v_2, \dots, v_N$  in the following way. For a vertex  $v$ , let  $\Gamma(v)$  be the set of neighbours of  $v$  in  $G$ . For each  $n = 0, 1, \dots, N$ , denote

$$\mathcal{Z}_n := \Gamma(v_1) \cup \dots \cup \Gamma(v_n) \setminus \{v_1, \dots, v_n\}.$$

Now recursively, for each  $n = 0, 1, \dots, N - 1$ :

- if  $|\mathcal{Z}_n| = 0$ , then let  $v_{n+1}$  be the smallest element of  $[N] \setminus \{v_1, \dots, v_n\}$ , and let  $v_{n+2}, \dots, v_{n+1+|\Gamma(v_{n+1})|}$  be the elements of  $\Gamma(v_{n+1})$ , in increasing order;
- if  $|\mathcal{Z}_n| = r > 0$ , then let  $\{v_{n+r+1}, \dots, v_{n+r+a}\}$  be the elements of  $\mathcal{Z}_{n+1} \setminus \mathcal{Z}_n$  in increasing order, where  $a = |\mathcal{Z}_{n+1} \setminus \mathcal{Z}_n|$ .

Note that  $\mathcal{Z}_0 = \emptyset$ ,  $v_1 = 1$ , and  $\mathcal{Z}_1 = \Gamma(1)$ .

We can interpret the construction as follows. We imagine exploring the graph one vertex at a time, revealing neighbours as we proceed.  $\mathcal{Z}_n$  is the *stack* after step  $n$ , consisting of the vertices that we have seen but not yet processed. At the next step  $n$ , if there are any vertices on the stack, we process the one which was added earliest (namely  $v_{n+1}$ ), removing it from the stack and adding to the stack all its neighbours that have not previously been seen. If instead the stack is empty, we select a new vertex (the smallest-labelled vertex that has not previously been seen) and process that vertex in the same way.

We define the *reflected exploration process*  $(Z_n)_{n \geq 0}$  by  $Z_n = |\mathcal{Z}_n|$ . If we define  $0 = n_0, n_1, \dots, n_C = N$  to be the times  $n$  such that  $Z_n = 0$ , written in increasing order, then the components of  $G$  are  $\{v_{n_i+1}, v_{n_i+2}, \dots, v_{n_{i+1}}\}$  for  $i = 0, \dots, C - 1$ . In this way we can interpret the component sizes of  $G$  as the lengths of excursions from 0 of the reflected exploration process.

The strategy of proof of Theorem 1.8 is now to show that the exploration processes  $Z^{N,p}$  of the forests  $F(N,p)$ , in the regime of Theorem 1.8, converge as  $N \rightarrow \infty$ , when suitably rescaled, to the diffusion  $Z^\lambda$ , in such a way that the rescaled lengths of the longest excursions of  $Z^{N,p}$  converge to the lengths of the longest excursions of  $Z^\lambda$ . Our main convergence result is the following:

**Theorem 1.9.** *With  $\lambda, (p(N))$  as in Theorem 1.8, let  $(Z_n^{N,p})_{n \geq 0}$  be the reflected exploration process of  $F(N,p)$ . For  $s \geq 0$ , set*

$$\tilde{Z}_s^{N,p} := N^{-1/3} Z_{\lfloor N^{2/3}s \rfloor}^{N,p}. \tag{1.8}$$

*Then we have  $\tilde{Z}^{N,p} \xrightarrow{d} Z^\lambda$ , uniformly on compact time-intervals.*

Let us try to give some intuition for this result and for the role played by the function  $\alpha$ , by comparing the behaviour of the exploration processes for  $G(N,p)$  and for  $F(N,p)$ , for  $p = N^{-1} + \lambda N^{-4/3}$ .

We first recall the heuristic for the scaling in the  $G(N,p)$  case. The exploration process  $Z_n, n \geq 0$  is a Markov chain. Condition on  $Z_n$ , the size of the stack after  $n$  steps, being equal to  $r > 0$  and consider the distribution of  $Z_{n+1} - Z_n + 1$ , which is 1 more than the next increment of the process. This quantity is the number of neighbours that the vertex  $v_n$  has in  $[N] \setminus \{v_1, \dots, v_{n+r}\}$ , and it has  $\text{Bin}(N - n - r, p)$  distribution, with mean  $(N - n - r)p$ .

If we write  $n = tN^{2/3}$  and  $r = bN^{1/3}$ , the mean of that increment is then

$$(N - tN^{2/3} - bN^{1/3})(N^{-1/3} + \lambda N^{-4/3}) - 1$$

which (for  $t$  and  $b$  of constant order) is approximately  $(\lambda - t)N^{-1/3}$ . Meanwhile the variance is  $1 + O(N^{-1/3})$ . If we rescale time by a factor  $N^{2/3}$  and space by a factor  $N^{1/3}$ , we converge to a process with drift  $\lambda - t$  and variance 1 per unit time, namely the diffusion  $B^\lambda$  of Proposition 1.1.

Now consider instead the exploration process for  $F(N,p)$ , which is  $G(N,p)$  conditioned to be acyclic. We will see in Section 2.2 that the exploration process is still a Markov chain, but the acyclicity condition changes the distribution of the increments. Suppose again  $Z_n = r$ , so that the current stack is  $\{v_{n+1}, \dots, v_{n+r}\}$ . These stack vertices are already known to be in the same component of the graph. For the graph to remain acyclic, we now require that the subgraph induced by the vertices  $\{v_{n+1}, \dots, v_N\}$  is a forest, and furthermore no two of the stack vertices are in the same tree of this forest. As a result of this conditioning, the quantity  $Z_{n+1} - Z_n + 1$  no longer has  $\text{Bin}(N - n - r, p)$  distribution as in the  $G(N,p)$  case just discussed, but is stochastically dominated by  $\text{Bin}(N - n - r, p)$ ; furthermore, the downward bias produced is stronger when the stack size  $r$  is higher.

What we establish is that, in the same regime as above, this bias produces a change in the expected increment which is again of order  $N^{-1/3}$  and depends on the size of the stack. After rescaling as above, the drift obtained is now instead  $\lambda - t - \alpha(b, \lambda - t)$ , leading to the diffusion  $Z^\lambda$  defined at (1.5).

The particular convergence properties that we need in order to obtain Theorem 1.9 are collected in the following result:

**Proposition 1.10.** *Fix  $\lambda \in \mathbb{R}$ , and let  $p = p(N)$  satisfy the conditions of Theorem 1.8. For each  $N \in \mathbb{N}$ , the reflected exploration process  $Z^{N,p}$  of  $F(N,p)$  is a Markov chain. Further, fix any  $T, K < \infty$  and  $\delta > 0$ . Then, uniformly on*

$n \in [0, TN^{2/3}]$  and  $r \in [1, KN^{1/3}]$ ,

$$N^{1/3} \mathbb{E} \left[ Z_{n+1}^{N,p} - Z_n^{N,p} \mid Z_n^{N,p} = r \right] - \left[ \lambda - \frac{n}{N^{2/3}} + \alpha \left( \frac{r}{N^{1/3}}, \lambda - \frac{n}{N^{2/3}} \right) \right] \rightarrow 0, \quad (1.9)$$

$$\mathbb{E} \left[ \left[ Z_{n+1}^{N,p} - Z_n^{N,p} \right]^2 \mid Z_n^{N,p} = r \right] \rightarrow 1, \quad (1.10)$$

$$N^{2/3} \mathbb{P} \left( \left| Z_{n+1}^{N,p} - Z_n^{N,p} \right| > \delta N^{1/3} \mid Z_n^{N,p} = r \right) \rightarrow 0, \quad (1.11)$$

as  $N \rightarrow \infty$ . In addition,

$$\liminf_{N \rightarrow \infty} \inf_{n \in [0, TN^{2/3}]} \mathbb{E} \left[ \left[ Z_{n+1}^{N,p} \right]^2 \mid Z_n^{N,p} = 0 \right] > 0. \quad (1.12)$$

Here (1.9) and (1.10) give the required convergence of the mean and variance of the increments respectively. Then (1.11) will imply that the limit process does not have jumps, and finally (1.12) ensures that the limit process reflects appropriately at zero.

In Section 5.4, we show that Proposition 1.10 is sufficient to imply Theorem 1.9. The main ingredient will be Theorem 5.7, a special case of the very general results of Stroock and Varadhan (1971) on the convergence of Markov processes to reflected diffusions.

We mention one further technical point which causes extra complication in the proof of Theorem 1.8, compared to that of Aldous's Proposition 1.1. We will need to go slightly beyond Theorem 1.9 in showing that the excursions of the discrete exploration process (whose lengths are the tree sizes of the forest) converge appropriately, after rescaling, to the excursions of  $Z^\lambda$ . To do so, we need to exclude the possibility that zeros of  $Z^\lambda$  arise only as the limits of small *positive* local minima of the discrete processes; for this we will use the fact that, conditional on its vertex set, a tree appearing in  $F(N, p)$  is a uniform random tree, whose exploration process we can approximate by a Brownian excursion. (In the case of Proposition 1.1, the limiting diffusion  $B^\lambda$  is homogeneous in space; hence Aldous was able to work instead with the unreflected process, and correspondingly with a slightly different version of the exploration process, whose height at step  $m$  is equal to the stack size minus the number of complete components already explored. Then one only needs to show that excursions above the running minimum of the discrete processes converge to excursions above the running minimum of the diffusion, which follows easily from the uniform convergence of the paths.)

1.4. *Discussion.* Before embarking on the proof of our main results, we discuss various aspects of the ensembles  $F(N, p)$  and  $G(N, p)$ , the limiting diffusion processes  $Z^\lambda$  and  $B^\lambda$ , and other related models.

1.4.1. *Excursions of  $B^\lambda$  and of  $Z^\lambda$ .* Just as for the process  $B^\lambda$ , the excursions of  $Z^\lambda$  occur in size-biased order. This property is inherited from the discrete exploration processes – since the graph is exchangeable, the exploration visits the components in size-biased order.

We also have that, conditional on their length, the excursions of  $Z^\lambda$  are Brownian excursions. That is, if we condition on the set of excursion intervals of the process, the paths of the process on these intervals are independent Brownian excursions. This follows from the fact that the trees of  $F(N, p)$  are uniformly distributed,

given their vertex sets, and the fact that the exploration process of a uniform tree converges in distribution to a Brownian excursion (Le Gall, 2005). For the process  $B^\lambda$ , in contrast, the excursions are Brownian excursions weighted by the exponential of their area (see Aldous, 1997); relative to Brownian excursion, the higher drift at the beginning and lower drift at the end of the interval favours excursions with higher area in  $B^\lambda$ , but in  $Z^\lambda$  this bias turns out to be precisely cancelled by the negative contribution to the drift from the  $\alpha$  term. These properties of size-biased ordering and Brownian excursions are not at all obvious from the definition of  $Z^\lambda$ . It is interesting to ask whether there are other diffusions which have both these properties.

1.4.2. *Monotonicity properties.* It's straightforward that the edge set of  $G(N, p)$  is stochastically dominated by that of  $G(N, p')$  when  $p < p'$ ; similarly we have stochastic domination of  $G(N, m)$  by  $G(N, m')$  for  $m < m'$ .

However, there seems no obvious argument leading to analogous properties to hold for the families  $F(N, p)$  and  $F(N, m)$ ; as far as we know, the question of whether these monotonicity properties hold is open.

From the combinatorial calculations that we use to estimate the probability that  $G(N, p)$  is acyclic, we obtain that in the critical window, the number of edges in the forest  $F(N, p)$  typically behaves like  $N^2 p/2 + O(N^{1/2})$ . (In fact, much more strongly, one could immediately obtain that the local central limit theorem for the number of edges is the same in  $F(N, p)$  as in  $G(N, p)$ ). If we also had a monotonicity result, it would then be easy to deduce Theorem 1.7 for  $F(N, m)$  from Theorem 1.8 for  $F(N, p)$ , using a simple sandwiching argument. Without it, we need to work a little harder. In Lemma 6.2 in Section 6, we prove an ‘almost monotonicity result’: for parameters in an appropriate range, we can couple a sequence of random forests with different numbers of edges in such a way that, with high probability, the edge sets are indeed monotonic.

1.4.3.  $\lambda \rightarrow \infty$ , *the supercritical phase, and random planar graphs.* In the subcritical regime, the behaviours of  $G(N, m)$  and  $F(N, m)$  are very similar. Consider for example  $m \sim cN$  where  $0 < c < 1/2$ . Then with probability bounded away from 1 as  $N \rightarrow \infty$ , the graph  $G(N, m)$  is itself acyclic. In both models, the size of the largest component is on the order of  $\log N$ . More broadly, the results of Łuczak and Pittel (1992) indicate that the scaling limit for the largest components is the same for the two models whenever  $N/2 - m \gg N^{2/3}$ .

However the supercritical behaviour of  $G(N, m)$  and  $F(N, m)$  is very different. First consider the regime where  $m \sim cN$  where  $c > 1/2$ . For both models, we see a single “giant component” of linear size, and the second-largest component has sub-linear size. In  $G(N, m)$ , the second-largest component has size  $O(\log N)$ ; we have the well-known “duality” property whereby, once the giant component is removed, the rest of the graph looks like a subcritical random graph. For  $F(N, m)$ , on the other hand, Łuczak and Pittel (1992) show that the size of the second-largest tree (and, in fact, of the  $k$ th-largest for any  $k \geq 2$ ) is on the order of  $N^{2/3}$ . The number of vertices outside the giant tree is sufficiently large that the remainder of the graph looks *critical* rather than sub-critical.

Łuczak and Pittel (1992) also show (in Theorem 5.1) a distributional scaling limit for the  $O(N^{2/3})$  fluctuations of the size of the giant tree around its mean. A

supercritical random forest without its giant tree can then be treated as a critical forest with random criticality parameter  $\lambda$ .

As a consequence, for the models  $F(N, m)$  and  $F(N, p)$ , the scaling limits described in terms of the excursions of the diffusion remain relevant in describing the  $k$ th largest components for  $k \geq 2$  in the supercritical regime as well as in the critical window. Although we do not state such a result here, one can show that the scaling limit for these components is (up to a uniform multiplicative correction) a *mixture* of the distributions obtained in our main theorem.

The different behaviour between the random graph model and the random forest model is already visible “at the top of the scaling window”. Suppose  $m = N/2 + s$  where  $N^{2/3} \ll s \ll N$ . For the random graph case (Bollobás, 1984), one has  $|C_1^{G(N, m)}| \sim (4 + o(1))s$  and  $|C_2^{G(N, m)}| = o(N^{2/3})$ . In the random forest, the largest component grows approximately half as quickly, with  $|C_1^{F(N, m)}| \sim (2 + o(1))s$ , and for  $k \geq 2$ ,  $|C_k^{F(N, m)}|$  remains on the order of  $N^{2/3}$ .

We can also see the difference between the two models reflected in the behaviour of the diffusion processes  $B^\lambda$  and  $Z^\lambda$ , as  $\lambda \rightarrow \infty$ . When  $\lambda$  becomes large,  $B^\lambda$  typically has a single large excursion, which begins at time  $o(1)$  and ends at time  $2\lambda \pm o(1)$ . At the end of the excursion, the drift of the process is  $-\lambda + o(1)$ , and all subsequent excursions are very small. On the other hand, one can show from Theorem 1.9 that the large excursion of  $Z^\lambda$  again begins at time  $o(1)$ , but it is roughly half as long as for  $B^\lambda$ , ending at time  $\lambda \pm O(1)$ . At this time, the drift of the process is  $O(1)$ , and the next largest excursions remain of constant order as  $\lambda \rightarrow \infty$ .

Another related model is that of the *random planar graph*  $P(N, m)$ , uniformly chosen from all planar graphs on  $[N]$  with  $m$  edges, which in a sense interpolates between  $F(N, m)$  and  $G(N, m)$ . Kang and Łuczak (2012) analysed the behaviour of  $P(N, m)$  in various regimes, including the critical window  $m = N/2 + O(N^{2/3})$ . In this window, the largest components of  $P(N, m)$  are again on the scale of  $N^{2/3}$ , and towards the top of the window, the scaling is similar to that of  $F(N, m)$  rather than  $G(N, m)$ ; if  $m = N/2 + s$  with  $N^{2/3} \ll s \ll N$ , then  $|C_1^{P(N, m)}| \sim (2 + o(1))s$ , and for  $k \geq 2$ ,  $|C_k^{P(N, m)}| = \Theta(N^{2/3})$ . It’s interesting to speculate about whether one could also obtain a scaling limit for the joint distribution of the sizes of the largest components of  $P(N, m)$  in terms of the excursion lengths of a diffusion. However, it’s not clear whether one can formulate an exploration process of the graph  $P(N, m)$  which has the Markov property; without this, it would perhaps be less plausible to obtain suitable convergence to a diffusion.

1.5. *Plan of the paper.* Section 2 is devoted to proving Proposition 1.10. We show the Markov property for the exploration process, and establish the estimates on the expectation and variance of its jumps, and the necessary properties concerning continuity and reflection at 0 of the limiting process.

In order to maintain the flow of the argument as much as possible, some of the more involved combinatorial calculations required for Section 2 are postponed to Section 4.

In Section 3 we prove that Theorem 1.9 (giving convergence of the exploration process on compact time-intervals) implies Theorem 1.8 (our main scaling limit result for  $F(N, p)$ ).



Section 5 covers various technical aspects, first justifying the regularity properties in Lemma 1.4 and the existence of the diffusion  $Z^\lambda$ , and then applying Stroock and Varadhan’s general theory for the convergence of Markov processes to diffusions in order to show that Theorem 1.9 follows from Proposition 1.10.

At this point we have completed the proof of Theorem 1.8. Finally Section 6 is devoted to the coupling arguments needed to deduce the result for  $F(N, m)$  in Theorem 1.7 from that for  $F(N, p)$  in Theorem 1.8.

## 2. Convergence of the reflected exploration process

This section is devoted to the proof of Proposition 1.10. After collecting a few basic results concerning couplings and expected component sizes for the models  $G(N, p)$  and  $F(N, p)$ , we turn to the Markov property for the exploration process of  $F(N, p)$ , and give its transition probabilities. Then we embark on various combinatorial calculations concerning the probability of acyclicity in various critical random graphs. Some of the more involved calculations will be completed in Section 4.

2.1. *Proper couplings and estimates.* First, we state two standard results, which we will use regularly. The first couples  $G(N, p)$  as  $p$  varies. The second relates  $G(N, p)$  and  $F(N, p)$ , and follows from Strassen’s theorem (Strassen, 1965) and the Harris inequality (Harris, 1960), since acyclicity is a decreasing event.

**Lemma 2.1.** *For all  $N \in \mathbb{N}$ ,  $p \leq q \in [0, 1]$ , there exists a coupling of  $G(N, p)$  and  $G(N, q)$  such that  $E(G(N, p)) \subseteq E(G(N, q))$  almost surely.*

**Lemma 2.2.** *For all  $N \in \mathbb{N}$ ,  $p \in [0, 1]$ , there exists a coupling of  $G(N, p)$  and  $F(N, p)$  such that  $E(F(N, p)) \subseteq E(G(N, p))$  almost surely.*

The following result, adapted from Janson and Spencer (2007), controls the expected size of the component of a uniformly-chosen vertex from  $G(N, p)$  in the critical window.

**Lemma 2.3.** *(Janson and Spencer, 2007, Corollary 5.2). Fix  $\lambda \in \mathbb{R}$ , and let  $(p(N))$  satisfy  $N^{1/3}(Np(N) - 1) \rightarrow \lambda$ . Let  $|C_{G(N,p)}(v)|$  be the size of the component containing a uniformly-chosen vertex  $v$  in  $G(N, p)$ . Then there exists  $\Theta^\lambda \in (0, \infty)$  such that*

$$N^{-1/3} \mathbb{E} [|C_{G(N,p)}(v)|] \rightarrow \Theta^\lambda \tag{2.1}$$

as  $N \rightarrow \infty$ . Thus by Lemma 2.2, if we now let  $|C_{F(N,p)}(v)|$  be the size of the component containing a uniformly-chosen vertex in  $F(N, p)$ , we have

$$\limsup_{N \rightarrow \infty} N^{-1/3} \mathbb{E} [|C_{F(N,p)}(v)|] \leq \Theta^\lambda. \tag{2.2}$$

**Lemma 2.4.**  $\Theta^\lambda$  is increasing as a function of  $\lambda$ , and  $\Theta^\lambda \rightarrow 0$  as  $\lambda \rightarrow -\infty$ .

*Proof:* The increasing property follows from Lemma 2.1. Then, take  $\lambda < 0$  and  $p = \frac{1+\lambda N^{-1/3}}{N}$ . It is well-known (see van der Hofstad, 2017 for details) that the exploration process of  $G(N, p)$  is stochastically dominated by the exploration process of  $\mathcal{T}^{Np}$ , the Galton–Watson tree with Poisson( $Np$ ) offspring distribution. From this, we obtain

$$\mathbb{E} [|C_{G(N,p)}(v)|] \leq \mathbb{E} [|\mathcal{T}^{Np}|] = \frac{1}{1 - Np} = \frac{N^{1/3}}{|\lambda|},$$

and the result follows on taking  $\lambda \rightarrow -\infty$ . □

Let  $S^2(G)$  be the sum of the squares of the component sizes in a graph  $G$ . Then the expectation in (2.1) is simply  $\mathbb{E} [S^2(G(N, p))] / N$  (since  $C(v)$  is a size-biased choice from the components of  $G$ ), and the following corollary is an immediate consequence of Lemma 2.3.

**Corollary 2.5.** *Suppose that  $p(N) = 1/N + O(N^{-4/3})$  as  $N \rightarrow \infty$ . Then the quantities  $N^{-4/3}S^2(F(N, p))$  and  $N^{-4/3}S^2(G(N, p))$  are bounded in expectation.*

2.2. *Stack forests.*

**Definition 2.6.** For a graph  $G$ , we say a set  $A \subseteq V(G)$  is *separated* in  $G$  if no pair of vertices in  $A$  lie in the same component of  $G$ .

Recall from Section 1.3 that we are considering a breadth-first exploration process of  $F(N, p)$ . For the remainder of this short section, we suppress notation on  $N$  and  $p$  in the exploration process, since the result to follow holds for all  $p \in (0, 1)$ . Then  $\mathcal{Z}_n$  is the *stack* of vertices which have been seen but not explored yet. Note that all the vertices in  $\mathcal{Z}_n$  are in the same component of  $F(N, p)$ , since components are explored one-by-one. In particular, in the graph restricted to  $[N] \setminus \{v_1, \dots, v_n\}$ , no pair of vertices in  $\mathcal{Z}_n$  lie in the same component, as otherwise there would be a cycle in  $F(N, p)$ . We refer to the  $\mathcal{Z}_n$  trees on  $[N] \setminus \{v_1, \dots, v_n\}$  containing each  $v \in \mathcal{Z}_n$  as the *stack forest*, as in Figure 2.1. We can see that the vertices in  $\mathcal{Z}_n$  are *separated* in the restricted graph on  $[N] \setminus \{v_1, \dots, v_n\}$ .

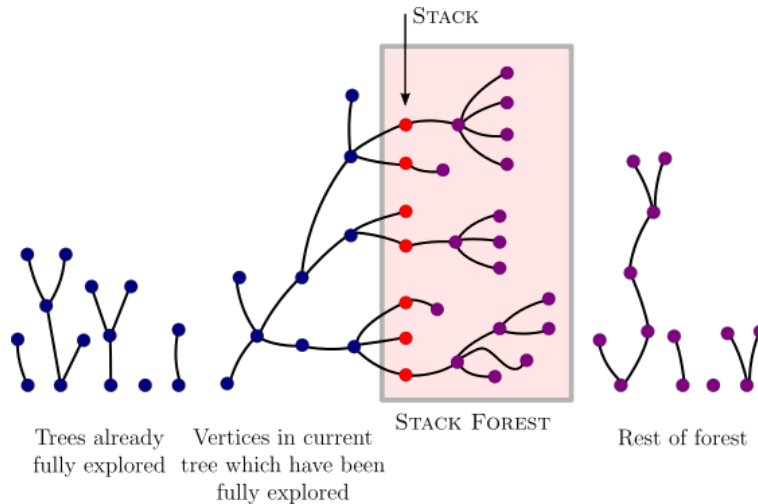


FIGURE 2.1. Illustration of the definition of *stack forest*

Now, suppose we condition on  $\{v_1, \dots, v_n\} \cup \mathcal{Z}_n$ , and the structure of  $F(N, p)$  on these  $n + \mathcal{Z}_n$  vertices. Then, the graph restricted to  $[N] \setminus \{v_1, \dots, v_n\}$  has the same distribution as

$$F([N] \setminus \{v_1, \dots, v_n\}, p),$$

with the extra condition that no pair of vertices from  $\mathcal{Z}_n$  lie in the same component.

We expand this explanation considerably in the proof of the following lemma, which formalises the claim that  $(Z_n)_{n \geq 0}$  is Markov, and characterises its transition probabilities via separation of the current stack in the remainder of the graph.

**Lemma 2.7.** *Let  $(Z_n)_{n \geq 0}$  be the exploration process of  $F(N, p)$ . Then  $Z$  is a Markov chain, and for  $n \geq 0$  and  $r \geq 1$ ,*

$$\begin{aligned} & \mathbb{P}(Z_{n+1} = r + \ell - 1 \mid Z_n = r) \\ & \propto \binom{N - n - r}{\ell} p^\ell (1 - p)^{N - n - r - \ell} \mathbb{P}([r + \ell - 1] \text{ separated in } F(N - n - 1, p)) \end{aligned} \tag{2.3}$$

as  $\ell$  varies over  $\{0, 1, \dots, N - n - r\}$ .

The distribution of  $Z_{n+1}$  given  $Z_n = 0$  is the same as the distribution of  $Z_{n+1}$  given  $Z_n = 1$ .

*Proof:* In the  $F(N, p)$  model, each forest  $H$  appears with probability proportional to  $\left(\frac{p}{1-p}\right)^{|E(H)|}$ , where  $|E(H)|$  is the number of edges of  $H$ .

Consider the first  $n$  steps of the exploration process. As well as conditioning on  $Z_n = r$ , consider conditioning further on the history  $(Z_1, \dots, Z_{n-1})$ , on the identity of the processed vertices  $v_1, \dots, v_n$  and on the vertices  $v_{n+1}, \dots, v_{n+r}$  currently on the stack. Let us write  $\mathcal{V}_n = \{v_1, \dots, v_n\}$  for the processed vertices,  $\mathcal{Z}_n = \{v_{n+1}, \dots, v_{n+r}\}$  for the stack, and  $\mathcal{U}_n = [N] \setminus (\mathcal{V}_n \cup \mathcal{Z}_n)$  for the remaining vertices. Assume for the moment that  $r \geq 1$ , i.e. that  $\mathcal{Z}_n$  is non-empty.

The conditioning determines the edges of the graph  $H$  restricted to the vertex set  $\mathcal{V}_n \cup \mathcal{Z}_n$ . Furthermore, under this condition there are no edges between  $\mathcal{V}_n$  and  $\mathcal{U}_n$ , with probability 1. So to specify  $H$  fully it is now enough to give the restriction  $\bar{H} = H|_{\mathcal{Z}_n \cup \mathcal{U}_n}$  of  $H$  to the vertex set  $\mathcal{Z}_n \cup \mathcal{U}_n$ .

The set of  $\bar{H}$  which are consistent with the conditioning is the set of  $\bar{H}$  for which  $H$  is a forest; for this, we require precisely that  $\bar{H}$  is a forest in which the vertices of the stack  $\mathcal{Z}_n$  are separated. Subject to this constraint, each  $\bar{H}$  appears with probability proportional to  $\left(\frac{p}{1-p}\right)^{|E(\bar{H})|}$  where  $|E(\bar{H})|$  is the number of edges of  $\bar{H}$ . After a suitable relabelling of the vertices, this gives the model  $F(N - n, p)$  subject to the condition that the vertices of  $[r]$  are separated.

The event in (2.3) occurs if the next increment  $Z_{n+1} - Z_n$  of the exploration process has size  $\ell - 1$ . This occurs if vertex  $v_{n+1}$  (the next vertex to be processed, which is currently on the stack) has  $\ell$  neighbours in  $\mathcal{U}_n$ . The conditional probability of an increment of size  $\ell - 1$  is then equal to

$$\alpha_{N,n,r,\ell} := \mathbb{P}(\deg_{\bar{H}}(1) = \ell \mid [r] \text{ separated in } \bar{H}) \text{ where } \bar{H} \sim F(N - n - r, p).$$

Note that this  $\alpha_{N,n,r,\ell}$  depends on the history we conditioned on only through the value of  $r$ ; hence in particular, given  $Z_n = r$ , the next increment is independent of the history  $(Z_1, \dots, Z_{n-1}) = (r_1, \dots, r_{n-1})$  of the exploration process, as required for the Markov property to hold, and the conditional probability on the left of (2.3) is also equal to  $\alpha_{N,n,r,\ell}$ .

There are  $\binom{N - n - r}{\ell}$  ways to choose  $\ell$  neighbours outside  $[r]$  for vertex 1. Without loss of generality, consider the case where these  $\ell$  neighbours are  $r + 1, r + 2, \dots, r + \ell$ . Then the property that  $[r]$  is separated is equivalent to the property that the set  $2, 3, \dots, r + \ell$  is separated in the graph with vertex 1 removed.

Reasoning in this way, and omitting factors which are constant in  $\ell$ , we obtain

$$\begin{aligned} \alpha_{N,n,r,\ell} &\propto \mathbb{P}(F(N-n-r,p) \text{ has } [r] \text{ separated and } \deg(1) = \ell) \\ &\propto \mathbb{P}(G(N-n-r,p) \text{ is a forest with } [r] \text{ separated and } \deg(1) = \ell) \\ &= \binom{N-n-r}{\ell} p^\ell (1-p)^{N-n-r-\ell} \\ &\quad \times \mathbb{P}(G(N-n-r-1,p) \text{ is a forest with } [r+\ell-1] \text{ separated}) \\ &\propto \binom{N-n-r}{\ell} p^\ell (1-p)^{N-n-r-\ell} \\ &\quad \times \mathbb{P}(F(N-n-r-1,p) \text{ has } [r+\ell-1] \text{ separated}). \end{aligned}$$

This is equal to the right-hand side of (2.3) as desired.

Observe that in the argument above, if  $r = |\mathcal{Z}_n| = 1$  then the property that the vertices of  $\mathcal{Z}_n$  are separated in  $\bar{H}$  becomes vacuously true; all forests  $\bar{H}$  are consistent with the history of the exploration process, and the conditional distribution of  $\bar{H}$  above becomes simply that of  $F(N-n,p)$ . In the case  $r = 0$ , where the stack is empty, we start exploring again from a new vertex (specifically, the vertex in  $\mathcal{U}_n$  with smallest label). Again all forests  $\bar{H}$  are consistent with the history, and so in fact the law of the rest of the process in the case  $Z_n = 0$  is the same as that in the case  $Z_n = 1$ , as desired.  $\square$

We want to quantify exactly how large a probabilistic penalty is incurred by adding an extra vertex to the stack, and so will consider limits of the quantity

$$\frac{\mathbb{P}([r+\ell] \text{ separated in } F(N-n-1,p))}{\mathbb{P}([r+\ell-1] \text{ separated in } F(N-n-1,p))}.$$

Given a graph in which  $[r+\ell-1]$  are separated, the conditional probability that  $r+\ell$  is also separated depends on the size of the stack forest rooted by  $[r+\ell-1]$ . So we will calculate the expected size of a stack forest in Section 2.4. We need precise asymptotics for the probability that  $G(N,p)$  is acyclic, which we derive in Section 2.3. We then use this to calculate the probability that the stack forest has a particular size.

2.3. *Enumerating weighted stack forests.* In this section, we consider the probability that  $G(N,p)$  is acyclic.

**Definition 2.8.** Let  $f(N,m)$  be the number of forests with vertex set  $[N]$  and exactly  $m$  edges. With a mild abuse of notation, we also define

$$f(N,p) := \mathbb{P}(G(N,p) \text{ acyclic}) = (1-p)^{\binom{N}{2}} \sum_{m=0}^{N-1} f(N,m) \left(\frac{p}{1-p}\right)^m, \quad p \in [0,1]. \tag{2.4}$$

**Lemma 2.9.** For any  $N \geq 0$  and any  $p \in (0,1)$ ,

$$f(N,p) \geq f(N+1,p) \geq f(N,p) \left[1 - \frac{1}{2} N p^2 \mathbb{E}[|C_{G(N,p)}(v)|]\right], \tag{2.5}$$

where  $C_{G(N,p)}(v)$  is the component containing a uniformly-chosen vertex  $v$  in  $G(N,p)$ .

*Proof:* Graphs with zero, one or two vertices are certainly acyclic, so  $f(0, p) = f(1, p) = f(2, p) = 1$ , the statement is true for  $N = 0, 1$ . We assume from now on that  $N \geq 2$ . We can define a forest on  $[N + 1]$  via the restriction to  $[N]$  (which is clearly also a forest) and the neighbourhood of vertex  $N + 1$ , where the latter must obey some conditions to avoid cycles. We take  $\mathbb{P}$  to be a probability distribution which couples  $G(N, p)$  and  $G(N + 1, p)$  such that  $E(G(N, p)) \subseteq E(G(N + 1, p))$ ,  $\mathbb{P}$ -a.s. Recall that in a graph  $G$ , for  $v \in V(G)$ ,  $\Gamma(v)$  is the set of vertices connected to  $v$  by an edge in  $E(G)$ . Then

$$f(N + 1, p) = f(N, p)\mathbb{P}(\Gamma(N + 1) \text{ separated in } G(N, p) \mid G(N, p) \text{ acyclic}),$$

and so the first inequality in (2.5) certainly holds. Now, for any set  $A \subset [N]$ , the event that  $A$  is separated in  $G$  is decreasing, while the event that  $G$  is acyclic is also decreasing. So, again by the Harris inequality,

$$f(N + 1, p) \geq f(N, p)\mathbb{P}(\Gamma(N + 1) \text{ separated in } G(N, p)),$$

and so

$$1 - \frac{f(N + 1, p)}{f(N, p)} \leq \mathbb{P}(\Gamma(N + 1) \text{ not separated in } G(N, p)). \tag{2.6}$$

Observe that the event that  $\Gamma(N + 1)$  is not separated in  $G(N, p)$  is the union over  $i, j \in [N]$  of the events

$$\{i, j \text{ both in } \Gamma(N + 1) \text{ and both in the same component of } G(N, p)\}.$$

Thus, by exchangeability of the vertices in  $[N]$ ,

$$\begin{aligned} \mathbb{P}(\Gamma(N + 1) \text{ not separated in } G(N, p)) \\ \leq \binom{N}{2} p^2 \mathbb{P}(1 \text{ and } 2 \text{ in same component of } G(N, p)). \end{aligned}$$

Then, if  $|C_{G(N,p)}(1)|$  is the size of the component of  $G(N, p)$  which contains vertex 1,

$$\mathbb{P}(1 \text{ and } 2 \text{ in same component of } G(N, p)) = \frac{\mathbb{E}[|C_{G(N,p)}(1)|] - 1}{N - 1}.$$

We conclude that

$$\begin{aligned} \mathbb{P}(\Gamma(N + 1) \text{ not separated in } G(N, p)) &\leq \binom{N}{2} p^2 \cdot \frac{\mathbb{E}[|C_{G(N,p)}(1)|] - 1}{N - 1} \\ &\leq \frac{1}{2} N p^2 \mathbb{E}[|C_{G(N,p)}(1)|], \end{aligned}$$

from which the result follows, using (2.6) and the fact that the vertices in  $G(N, p)$  are exchangeable.  $\square$

Now, using the asymptotics for  $f(N, m)$  in (1.2), we may obtain asymptotics for  $f(N, p)$ . Here, and in subsequent sections, some straightforward but lengthy calculations are required, and in some places, various expansions have to be taken to fifth order. To avoid breaking the flow of the main argument, we postpone this proof until Section 4.1.

**Lemma 2.10.** *Fix  $\lambda^- < \lambda^+ \in \mathbb{R}$ . Given  $p \in (0, 1)$ , define  $\Lambda = \Lambda(N, p) = N^{1/3}(Np - 1)$ . Then*

$$f(N, p) = \mathbb{P}(G(N, p) \text{ acyclic}) = (1 + o(1))g(\Lambda)e^{3/4\sqrt{2\pi}N^{-1/6}}, \tag{2.7}$$

uniformly for  $\Lambda \in [\lambda^-, \lambda^+]$  as  $N \rightarrow \infty$ .

Motivated by the definition of stack forests, for each  $0 \leq r \leq N$ , let  $\mathcal{A}_{N,r} \subseteq \mathcal{F}_N$  denote the set of forests where the vertices  $1, \dots, r$  are separated. Furthermore, given a forest  $F \in \mathcal{A}_{N,r}$ , let  $k_r(F)$  be the sum of the sizes of the components containing vertices  $1, \dots, r$ . We also define

$$\mathcal{A}_{N,r,k} := \{F \in \mathcal{A}_{N,r}, k_r(F) = k\}, \tag{2.8}$$

the set of forests where  $1, \dots, r$  are separated, and their stack forest has size  $k$ .

**Definition 2.11.** Given  $p \in (0, 1)$  and  $N, N', r, k \in \mathbb{N}$  satisfying  $N' \leq N$ , and  $r \leq k \leq N$ , we will use the following rescalings:

$$\begin{aligned} \Lambda = \Lambda(N, p) &:= N^{1/3}(Np - 1), & a = a(N, k) &:= \frac{k}{N^{2/3}}, \\ b = b(N, r) &:= \frac{r}{N^{1/3}}, & s = s(N, N') &:= \frac{N - N'}{N^{2/3}}. \end{aligned} \tag{2.9}$$

Here,  $b$  represent the rescaled size of the stack and  $a$  represents the rescaled size of the stack forest. When analysing the exploration process of  $G(N, p)$ , we require estimates for the graph structure on the  $N' \leq N$  vertices which have not yet been explored. Then  $s$  represents the rescaled number of vertices already explored in the exploration process.

**Note.** Observe that for  $p(N)$  satisfying the conditions of Theorem 1.8, we have  $\Lambda(N, p(N)) \rightarrow \lambda$ .

**Definition 2.12.** For much of this and the following sections, it will be necessary to make estimates uniformly across several variables. For constants  $T < \infty$ , and  $\lambda^- < \lambda^+$ , and  $0 < \epsilon < K < \infty$ , we let

$$\begin{aligned} \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T) &:= \left\{ (N', p, r, k) \in \mathbb{N} \times (0, 1) \times \mathbb{N} \times \mathbb{N} : s(N, N') \in [0, T], \right. \\ &\quad \left. \Lambda(N, p) \in [\lambda^-, \lambda^+], b(N, r) \in [\epsilon, K], k \in [r, KN^{2/3}] \right\}. \end{aligned}$$

In addition, we define the projection this set onto its first three entries

$$\begin{aligned} \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T) &:= \left\{ (N', p, r) : s(N, N') \in [0, T], \Lambda(N, p) \in [\lambda^-, \lambda^+], b(N, r) \in [\epsilon, K] \right\}, \end{aligned}$$

and a variant with a broader range of  $r$

$$\begin{aligned} \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T) &:= \left\{ (N', p, r) : s(N, N') \in [0, T], \Lambda(N, p) \in [\lambda^-, \lambda^+], r \in [1, KN^{1/3}] \right\}. \end{aligned}$$

The following lemma gives uniform asymptotics for the probability that  $G(N', p)$  lies in  $\mathcal{A}_{N',r,k}$ . The proof is postponed until Section 4.2.

**Lemma 2.13.** Fix constants  $\lambda^-, \lambda^+, \epsilon, K, T$  as in Definition 2.12. Then,

$$\begin{aligned} \mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k}) &= (1 + o(1))e^{3/4}g(\Lambda - s - a)N^{-5/6}ba^{-3/2} \\ &\quad \times \exp\left(-b(\Lambda - s) - \frac{b^2}{2a} - \frac{(\Lambda - s - a)^3 - (\Lambda - s)^3}{6}\right), \end{aligned} \tag{2.10}$$

uniformly on  $(N', p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$ , as  $N \rightarrow \infty$ .

2.4. *Expected size of the stack forest.* We now condition on  $[r]$  being separated in  $F(N', p)$ , and obtain an estimate for the expected size of the corresponding stack forest. Recall from (2.9) the definitions  $b = b(N, r)$  and  $s = s(N, N')$ , the rescaled stack size, and graph vertex count deficit, respectively.

**Lemma 2.14.** *Fix constants  $\lambda^-, \lambda^+, K, T$  as in Definition 2.12. Then,*

$$N^{-2/3} \mathbb{E} [k_r(F(N', p)) \mid F(N', p) \in \mathcal{A}_{N', r}] - \alpha(b, \Lambda - s) \rightarrow 0, \tag{2.11}$$

uniformly on  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$ , as  $N \rightarrow \infty$ .

*Proof:* We can rewrite the expectation in (2.11) in terms of the unconditioned random graphs  $G(N', p)$  as follows.

$$\begin{aligned} \mathbb{E} [k_r(F(N', p)) \mid F(N', p) \in \mathcal{A}_{N', r}] &= \frac{\sum_{k=r}^{N'} k \mathbb{P}(F(N', p) \in \mathcal{A}_{N', r, k})}{\sum_{k=r}^{N'} \mathbb{P}(F(N', p) \in \mathcal{A}_{N', r, k})} \\ &= \frac{\sum_{k=r}^{N'} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})}{\sum_{k=r}^{N'} \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})}. \end{aligned} \tag{2.12}$$

We shall see that both of the sums in (2.12) are dominated by contributions from  $k = \Theta(N^{2/3})$ .

In order to use Lemma 2.13, we assume  $\epsilon \in (0, K)$  is given. We will first show that (2.11) holds uniformly on  $\Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$ . Then, at the end, we will take  $\epsilon \rightarrow 0$ . We also select  $M > K$ , which we will take to  $\infty$  shortly.

We write  $h(a, b) := a^{-3/2} g(\Lambda - s - a) \exp\left(\frac{(\Lambda - s - a)^3 - (\Lambda - s)^3}{6}\right) \exp(-b^2/2a)$ . Since  $g$  is bounded,  $h(a, b) \rightarrow 0$  as  $a \rightarrow 0$  (indeed uniformly on  $b \in [\epsilon, K]$ ,  $\Lambda \in \mathbb{R}$ ,  $s \in \mathbb{R}_{\geq 0}$ ), so  $\int_0^M h(a, b) da < \infty$  for all  $M < \infty$ . On compact intervals in  $(a, b, \Lambda, s)$ ,  $h$  is uniformly continuous and bounded away from zero. We may now use Lemma 2.13 to approximate every summand in (2.12), uniformly over the required range. (Recall from (2.9) that  $a$  is a linear function of  $k$ .) So

$$\begin{aligned} \sum_{k=r}^{\lceil MN^{2/3} \rceil} \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}) &= (1 + o(1)) b N^{-5/6} \exp\left(-b(\Lambda - s) - \frac{(\Lambda - s)^3}{6} + \frac{3}{4}\right) \\ &\quad \times N^{2/3} \int_0^M a^{-3/2} g(\Lambda - s - a) \exp\left(\frac{(\Lambda - s - a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da, \\ \sum_{k=r}^{\lceil MN^{2/3} \rceil} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}) &= (1 + o(1)) b N^{-5/6} \exp\left(-b(\Lambda - s) - \frac{(\Lambda - s)^3}{6} + \frac{3}{4}\right) \\ &\quad \times N^{4/3} \int_0^M a^{-1/2} g(\Lambda - s - a) \exp\left(\frac{(\Lambda - s - a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da, \end{aligned} \tag{2.13}$$

uniformly on  $(N', p, r) \in \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)$ , as  $N \rightarrow \infty$ .

Observe, by comparison with the definition of  $\alpha$  in (1.4), that

$$\lim_{M \rightarrow \infty} N^{-2/3} \frac{\sum_{k=r}^{\lceil MN^{2/3} \rceil} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})}{\sum_{k=r}^{\lceil MN^{2/3} \rceil} \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})} = (1 + o(1)) \alpha(b, \Lambda - s),$$

uniformly on  $(N', p, r) \in \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)$ .

Therefore, to apply (2.12) to verify (2.11), we must check that the contribution to the expectation from the event that the size of the stack forest is larger than

$MN^{2/3}$  vanishes as  $M \rightarrow \infty$ . From (2.13), the contribution to the numerator of (2.12) from summands for which  $k \in [r, \lceil MN^{2/3} \rceil]$  has order  $N^{-5/6} \times N^{4/3} = N^{1/2}$ . So to verify (2.11) uniformly on  $\Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)$ , it will suffice to check that the following statement holds:

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{(N', p, r) \in \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)} N^{-1/2} \sum_{k=\lceil MN^{2/3} \rceil}^{N'} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}) = 0. \tag{2.14}$$

*The stack forest is not too large.* To show (2.14), we will show that the sequence  $(k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}))_{k \geq r}$  is eventually bounded by a geometric series. From the definition of  $F(N, p)$  in (2.4), we have that

$$\mathbb{P}(G(N, p) \in \mathcal{A}_{N, r, k}) = (1-p)^{\binom{N}{2} - \binom{N-k}{2}} \binom{N-r}{k-r} \left(\frac{p}{1-p}\right)^{k-r} r k^{k-r-1} F(N-k, p). \tag{2.15}$$

An explanation of where each term in this expression comes from is given in the proof of Lemma 2.13 in Section 4.2. We will use this to control the ratio of the probabilities  $\mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})$  in the following lemma.

**Lemma 2.15.** *Given the same constants as in Lemma 2.14, there exist constants  $M < \infty$  and  $\gamma > 0$  such that for all large enough  $N$ ,*

$$\frac{(k+1) \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k+1})}{k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k})} \leq 1 - \gamma N^{-2/3} \tag{2.16}$$

whenever  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$  and  $k \in [MN^{2/3}, N' - 1]$ .

This lemma is proved in Section 4.3. But then, we can bound (2.14) via a geometric series as

$$\begin{aligned} N^{-1/2} \sum_{k=MN^{2/3}}^{N'} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}) \\ \leq N^{-1/2} \frac{\lceil MN^{2/3} \rceil \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, \lceil MN^{2/3} \rceil})}{1 - (1 - \gamma N^{-2/3})}. \end{aligned}$$

By Lemma 2.13, this RHS is

$$\begin{aligned} (1 + o(1)) N^{-1/2} \frac{1}{\gamma} N^{2/3} \cdot MN^{2/3} e^{3/4} g(\Lambda - s - M) N^{-5/6} b M^{-3/2} \\ \times \exp\left(-b(\Lambda - s) - \frac{b^2}{2M} + \frac{(\Lambda - s - M)^3 - (\Lambda - s)^3}{6}\right) \\ = (1 + o(1)) M^{-1/2} e^{-b^2/2M} \exp\left(\frac{(\Lambda - s - M)^3 - (\Lambda - s)^3}{6}\right) \times g(\Lambda - s - M) \\ \times \frac{e^{3/4}}{\gamma} b \exp\left(-b(\Lambda - s) - \frac{(\Lambda - s)^3}{6}\right). \end{aligned}$$



Recall that  $g$  is uniformly bounded above and  $\exp\left(\frac{(\Lambda-s-M)^3-(\Lambda-s)^3}{6}\right) \leq 1$ . Then observe that  $M^{-1/2}e^{b^2/2M} \rightarrow 0$  as  $M \rightarrow \infty$ . Therefore

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{(N',p,r) \in \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)} N^{-1/2} \sum_{k=\lfloor MN^{2/3} \rfloor}^{N'} k \mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k}) = 0.$$

So we have finished the proof of (2.14), and thus we have shown that (2.11) holds uniformly on  $\Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)$ .

*Small stacks.* To finish this proof of Lemma 2.14, it remains to extend the convergence to uniformity on  $r \in [1, \lceil Kn^{1/3} \rceil]$ , rather than on  $[\lceil \epsilon N^{1/3} \rceil, \lceil Kn^{1/3} \rceil]$ .

Recall from Lemma 1.4 that  $\alpha(b, \Lambda) \rightarrow 0$  as  $b \downarrow 0$  uniformly on compact intervals in  $\Lambda$ . In particular

$$\lim_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \sup_{\substack{\Lambda \in [\lambda^-, \lambda^+] \\ s \in [0, T], r \in [1, \epsilon N^{1/3}]}} \alpha\left(\frac{r}{N^{1/3}}, \Lambda - s\right) = 0. \tag{2.17}$$

Before Definition 2.12, we defined  $k_r(F)$  for a forest  $F$ , but we can extend the definition to a general graph  $G$  with vertex set  $[N]$ . If  $|C(i)|$  is the size of the component containing vertex  $i \in [N]$ , then set  $k_r(G) := |C(1)| + \dots + |C(r)|$ , so some components may be counted at least twice. In particular,  $k_r(G)$  is an increasing function of graphs. However, for any  $r$ , the set  $\mathcal{A}_{N,r}$  is a decreasing family of graphs. Therefore

$$\mathbb{E}[k_r(G(N', p)) \mid G(N', p) \in \mathcal{A}_{N',r}] \leq \mathbb{E}[k_r(G(N', p))] \leq r \mathbb{E}[|C_{G(N',p)}(1)|], \tag{2.18}$$

where  $|C_{G(N',p)}(1)|$  is the size of the component containing vertex 1 in  $G(N', p)$ . From Lemma 2.3, for the range of  $N', p$  under consideration,

$$\limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N - TN^{2/3}, N] \\ \Lambda(N,p) \in [\lambda^-, \lambda^+]}} N^{-1/3} \mathbb{E}[|C_{G(N',p)}(1)|] \leq \Theta^{\lambda^+} < \infty. \tag{2.19}$$

We now take  $r \leq \epsilon N^{1/3}$  in (2.18), and apply (2.19) to obtain

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N - TN^{2/3}, N] \\ \Lambda(N,p) \in [\lambda^-, \lambda^+] \\ r \in [1, \epsilon N^{1/3}]}} N^{-2/3} \mathbb{E}[k_r(G(N', p)) \mid G(N', p) \in \mathcal{A}_{N',r}] &\leq \lim_{\epsilon \rightarrow 0} \epsilon \Theta^{\lambda^+} \\ &= 0. \end{aligned}$$

So, with (2.17), this gives

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N - TN^{2/3}, N] \\ \Lambda(N,p) \in [\lambda^-, \lambda^+] \\ r \in [1, \epsilon N^{1/3}]}} \left| N^{-2/3} \mathbb{E}[k_r(G(N', p)) \mid G(N', p) \in \mathcal{A}_{N',r}] \right. \\ \left. - \alpha\left(\frac{r}{N^{1/3}}, \Lambda - s\right) \right| = 0. \tag{2.20} \end{aligned}$$

We already know that (2.11) holds uniformly on  $\Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$ . So, combining with (2.20) and taking  $\epsilon$  small shows that (2.11) does hold uniformly on  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$ , as required for the full statement of Lemma 2.14.  $\square$

2.5. *Proof of Proposition 1.10: convergence of the drift.* Recall that  $\mathcal{A}_{N,r} \subseteq \mathcal{F}_N$  is the set of forests on  $[N]$  where vertices  $1, \dots, r$  are separated. Let  $F$  be a uniform choice from  $\mathcal{A}_{N,r}$ . Then

$$\mathbb{P}(F \in \mathcal{A}_{N,r+1} \mid F \in \mathcal{A}_{N,r,k}) = \frac{N - k}{N - r},$$

as the labels of the  $k - r$  other vertices in the stack forest containing vertices  $[r]$  are uniformly chosen from  $\{r + 1, \dots, N\}$ . Furthermore,  $\mathcal{A}_{N,r+1} \subseteq \mathcal{A}_{N,r}$ , and so

$$\begin{aligned} \frac{\mathbb{P}(F(N, p) \in \mathcal{A}_{N,r+1})}{\mathbb{P}(F(N, p) \in \mathcal{A}_{N,r})} &= \mathbb{P}(F(N, p) \in \mathcal{A}_{N,r+1} \mid F(N, p) \in \mathcal{A}_{N,r}) \\ &= \sum_{k=r}^N \mathbb{P}(F(N, p) \in \mathcal{A}_{N,r+1} \mid F(N, p) \in \mathcal{A}_{N,r,k}) \\ &\quad \times \mathbb{P}(F(N, p) \in \mathcal{A}_{N,r,k} \mid F(N, p) \in \mathcal{A}_{N,r}) \\ &= \frac{N - \mathbb{E}[k_r(F(N, p)) \mid F(N, p) \in \mathcal{A}_{N,r}]}{N - r}. \end{aligned}$$

It follows that uniformly on  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$ , as in Lemma 2.14, as  $N \rightarrow \infty$ ,

$$N^{1/3} \left[ 1 - \frac{\mathbb{P}(F(N', p) \in \mathcal{A}_{N',r+1})}{\mathbb{P}(F(N', p) \in \mathcal{A}_{N',r})} \right] - \alpha\left(\frac{r}{N^{1/3}}, \Lambda - s\right) \rightarrow 0.$$

The sequence  $p(N)$  satisfies the conditions in the statement of Theorem 1.8, that is  $\Lambda(N, p(N)) \rightarrow \lambda \in \mathbb{R}$  in the notation of Definition 2.11. So in fact we may replace  $\Lambda$  with  $\lambda$ , obtaining, again uniformly on  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$ ,

$$N^{1/3} \left[ 1 - \frac{\mathbb{P}(F(N', p) \in \mathcal{A}_{N',r+1})}{\mathbb{P}(F(N', p) \in \mathcal{A}_{N',r})} \right] - \alpha\left(\frac{r}{N^{1/3}}, \lambda - s\right) \rightarrow 0. \tag{2.21}$$

Now we can return to the increments of  $Z^{N,p}$ , the exploration process of  $F(N, p)$ . Recall Lemma 2.7, which asserts that

$$\begin{aligned} \mathbb{P}\left(Z_{n+1}^{N,p} - Z_n^{N,p} = \ell - 1 \mid Z_n^{N,p} = r\right) &\propto \mathbb{P}(B^{N-n-r,p} = \ell) \\ &\quad \times \mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1,r+\ell-1}), \quad \ell \geq 0, \end{aligned}$$

where  $B^{N-n-r,p} \sim \text{Bin}(N - n - r, p)$ . So we define

$$q_\ell^{N,n,r} := \mathbb{P}(B^{N-n-r,p} = \ell) \times \frac{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1,r+\ell-1})}{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1,r-1})}. \tag{2.22}$$

Therefore we also have  $\mathbb{P}\left(Z_{n+1}^{N,p} - Z_n^{N,p} = \ell - 1 \mid Z_n^{N,p} = r\right) \propto q_\ell^{N,n,r}$ . Heuristically, from (2.21), this quotient, which we will think of as a *weight*, should be approximately

$$\left(1 - \alpha\left(\frac{r}{N^{1/3}}, \Lambda - \frac{n}{N^{2/3}}\right) N^{-1/3}\right)^\ell,$$

and so we will be able to approximate  $\sum q_\ell^{N,n,r}$  by the probability generating function of  $B^{N-n-r,p}$ . Indeed, this approximation only breaks down when  $r + \ell - 1 \geq$

$KN^{1/3}$ , that is, outside the range of (2.21). Therefore, for any  $\delta > 0$ , for large enough  $N$ , we have, for all  $n \in [0, TN^{2/3}]$ ,  $r \in [1, \frac{K}{2}N^{1/3}]$ , and  $\ell \leq N^{1/4}$ .

$$\frac{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r+\ell-1})}{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r-1})} \leq \prod_{i=0}^{\ell-1} \left(1 - (\alpha(\frac{r+i-1}{N^{1/3}}, \lambda - \frac{n-1}{N^{2/3}}) - \delta)N^{-1/3}\right).$$

The function  $\alpha$  is uniformly continuous. Since the range of  $i$  in this product is asymptotically negligible relative to  $N^{1/3}$ , for large enough  $N$ , for large enough  $N$  we may replace  $r + i - 1$  by  $r$ , and  $\Lambda = \Lambda(N, p)$  by  $\lambda$ . That is,

$$\frac{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r+\ell-1})}{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r-1})} \leq \left(1 - (\alpha(\frac{r}{N^{1/3}}, \lambda - \frac{n}{N^{2/3}}) - \delta)N^{-1/3}\right)^\ell.$$

An identical argument gives

$$\frac{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r+\ell-1})}{\mathbb{P}(F(N - n - 1, p) \in \mathcal{A}_{N-n-1, r-1})} \geq \left(1 - (\alpha(\frac{r}{N^{1/3}}, \lambda - \frac{n}{N^{2/3}}) + \delta)N^{-1/3}\right)^\ell,$$

under the same conditions. From now on, we write  $\alpha_{n,r}^N = \alpha(\frac{r}{N^{1/3}}, \lambda - \frac{n}{N^{2/3}})$  for brevity.

Keeping  $\delta > 0$  fixed, we now address the sums  $\sum_{\ell=0}^\infty q_\ell^{N,n,r}$  and  $\sum_{\ell=0}^\infty (\ell-1)q_\ell^{N,n,r}$ . (Note first that both  $q_0^{N,n,r}$  and  $q_1^{N,n,r} \rightarrow 1/e$ , so these sums are uniformly bounded below.) For large enough  $N$ , we have, again for all  $n \in [0, TN^{2/3}]$ ,  $r \in [1, \frac{K}{2}N^{1/3}]$ ,

$$\begin{aligned} & \sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r} \\ & \leq \sum_{\ell=0}^{\lceil N^{1/4} \rceil} \mathbb{P}(B^{N-n-r,p} = \ell) \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)^\ell + \mathbb{P}(B^{N-n-r,p} \geq N^{1/4}) \\ & \leq \left[(1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)\right]^{N-n-r} + \mathbb{P}(B^{N-n-r,p} \geq N^{1/4}) \end{aligned}$$

Now, note that

$$\begin{aligned} & \left[(1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)\right]^{N-n-r} \\ & = \left[1 - (\alpha_{n,r}^N - \delta)N^{-4/3} + O(N^{-5/3})\right]^{N-n-r}, \end{aligned}$$

from which we find that

$$N^{1/3} \left[1 - \left[(1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)\right]^{N-n-r}\right] + (\alpha_{n,r}^N - \delta) \rightarrow 0, \tag{2.23}$$

uniformly as  $N \rightarrow \infty$ . The probability  $\mathbb{P}(B^{N-n-r,p} \geq N^{1/4})$  decays exponentially with some positive power of  $N$ , so we have shown that for large enough  $N$ ,

$$\sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r} \leq 1 - (\alpha_{n,r}^N - 2\delta)N^{-1/3}. \tag{2.24}$$

Under the same conditions,

$$\sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r} \geq 1 - (\alpha_{n,r}^N + 2\delta)N^{-1/3}.$$

Now we consider the sum  $\sum \ell q_\ell^{N,n,r}$ .

$$\begin{aligned} \sum_{\ell=0}^{N-n-r} \ell q_\ell^{N,n,r} &\leq \sum_{\ell=0}^{N-n-r} \ell \mathbb{P}(B^{N-n-r,p} = \ell) \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)^\ell \\ &\quad + N \mathbb{P}(B^{N-n-r,p} \geq N^{1/4}) \\ &\leq (N-n-r)p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right) \\ &\quad \times \left[ (1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right) \right]^{N-n-r-1} \\ &\quad + N \mathbb{P}(B^{N-n-r,p} \geq N^{1/4}). \end{aligned} \tag{2.25}$$

We can treat the term  $\left[ (1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right) \right]^{N-n-r-1}$  as in (2.23). We also have

$$(N-n-r)p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right) = 1 + \left(\lambda - \frac{n}{N^{2/3}} - (\alpha_{n,r}^N - \delta)\right) N^{-1/3} + O(N^{-2/3}).$$

So, in a similar fashion to (2.24), we establish

$$\sum_{\ell=0}^{N-n-r} \ell q_\ell^{N,n,r} \leq 1 + \left(\lambda - 2\alpha_{n,r}^N + 3\delta - \frac{n}{N^{2/3}}\right) N^{-1/3}, \tag{2.26}$$

and

$$\sum_{\ell=0}^{N-n-r} \ell q_\ell^{N,n,r} \geq 1 + \left(\lambda - 2\alpha_{n,r}^N - 3\delta - \frac{n}{N^{2/3}}\right) N^{-1/3}.$$

Therefore (where each successive statement holds whenever  $(N-n, p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, \frac{K}{2}, T)$  for large enough  $N$ )

$$\begin{aligned} \mathbb{E} \left[ Z_{n+1}^{N,p} - Z_n^{N,p} \mid Z_n^{N,p} = r \right] &= \frac{\sum_{\ell=0}^{N-n-r} \ell q_\ell^{N,n,r} - \sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r}}{\sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r}} \\ &\leq \frac{\left(\lambda - 2\alpha_{n,r}^N + 3\delta - \frac{n}{N^{2/3}}\right) N^{-1/3} + \left(\alpha_{n,r}^N + 2\delta\right) N^{-1/3}}{1 + \left(\lambda - \alpha_{n,r}^N - 2\delta - \frac{n}{N^{2/3}}\right) N^{-1/3}} \\ &\leq \left(\lambda - \alpha_{n,r}^N - \frac{n}{N^{2/3}} + 6\delta\right) N^{-1/3}. \end{aligned}$$

Similarly

$$\mathbb{E} \left[ Z_{n+1}^{N,p} - Z_n^{N,p} \mid Z_n^{N,p} = r \right] \geq \left(\lambda - \alpha_{n,r}^N - \frac{n}{N^{2/3}} - 6\delta\right) N^{-1/3},$$

and so since  $\delta > 0$  was arbitrary, after replacing  $\frac{K}{2}$  with  $K$ , we have completed the proof of (1.9) in Proposition 1.10.

2.6. Proof of Proposition 1.10: variance, jumps and reflection.

*Variance of increments.* We can show (1.10) using the estimates from Section 2.5. Recall the definition of  $q_\ell^{N,n,r}$  from (2.22). As in (2.25), we have

$$\begin{aligned} \sum_{\ell=0}^{N-n-r} \ell(\ell-1)q_\ell^{N,n,r} &\leq (N-n-r)(N-n-r-1)p^2 \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right)^2 \\ &\quad \times \left[ (1-p) + p \left(1 - (\alpha_{n,r}^N - \delta)N^{-1/3}\right) \right]^{N-n-r-2} \\ &\quad + N\mathbb{P}\left(B^{N-n-r,p} \geq N^{1/4}\right). \end{aligned}$$

Again, we use (2.23) and similarly to (2.26), we have

$$\begin{aligned} 1 + (2\lambda - 3\alpha_{n,r}^N - 4\delta - \frac{2n}{N^{2/3}}) N^{-1/3} \\ \leq \sum_{\ell=0}^{N-n-r} \ell(\ell-1)q_\ell^{N,n,r} \leq 1 + (2\lambda - 3\alpha_{n,r}^N + 4\delta - \frac{2n}{N^{2/3}}) N^{-1/3}. \end{aligned}$$

In particular, we obtain

$$\sum_{\ell=0}^{N-n-r} (\ell-1)^2 q_\ell^{N,n,r} = \sum_{\ell=0}^{N-n-r} \ell(\ell-1)q_\ell^{N,n,r} - \sum_{\ell=0}^{N-n-r} \ell q_\ell^{N,n,r} + \sum_{\ell=0}^{N-n-r} q_\ell^{N,n,r} \rightarrow 1,$$

uniformly, which is exactly (1.10).

*Jumps in the limit.* For any  $n \in [N]$ ,

$$\begin{aligned} \mathbb{P}\left(|Z_{n+1}^{N,p} - Z_n^{N,p}| > \delta N^{1/3}\right) &\leq \mathbb{P}\left(\exists v \in [N], \deg_{F(N,p)}(v) > \delta N^{1/3}\right) \\ &\stackrel{\text{Prop 2.2}}{\leq} \mathbb{P}\left(\exists v \in [N], \deg_{G(N,p)}(v) > \delta N^{1/3}\right) \\ &\leq N\mathbb{P}\left(\deg_{G(N,p)}(1) > \delta N^{1/3}\right). \end{aligned}$$

But  $\deg_{G(N,p)}(1) \sim \text{Bin}(N-1, p)$ , and so for any  $\delta > 0$ , this final term vanishes exponentially fast. So (1.11) follows.

*Speed at the boundary.* Finally, we check that the discrete processes  $(Z^{N,p})$  do not get stuck at zero. By Lemma 2.7, we have

$$\frac{\mathbb{P}\left(Z_{n+1}^{N,p} = 1 \mid Z_n^{N,p} = 0\right)}{\mathbb{P}\left(Z_{n+1}^{N,p} = 0 \mid Z_n^{N,p} = 0\right)} = \frac{\mathbb{P}\left(B^{N-n-1,p} = 1\right)}{\mathbb{P}\left(B^{N-n-1,p} = 0\right)} = \frac{(N-n-1)p}{1-p}.$$

Therefore

$$\liminf_{N \rightarrow \infty} \inf_{n \in [0, TN^{2/3}]} \frac{\mathbb{P}\left(Z_{n+1}^{N,p} = 1 \mid Z_n^{N,p} = 0\right)}{\mathbb{P}\left(Z_{n+1}^{N,p} = 0 \mid Z_n^{N,p} = 0\right)} \geq 1,$$

and so

$$\liminf_{N \rightarrow \infty} \inf_{n \in [0, TN^{2/3}]} \mathbb{E}\left[\left[Z_{n+1}^{N,p}\right]^2 \mid Z_n^{N,p} = 0\right] \geq \frac{1}{2},$$

as required for (1.12).

This completes the proof of Proposition 1.10 (subject to the proofs of Lemmas 2.10, 2.13, and 2.15 in Section 4).

### 3. Excursions and component sizes

In this section, we will prove that Theorem 1.8 follows from Theorem 1.9.

As in Aldous (1997), we must check that excursions of the limiting reflected SDE are matched by excursions of the discrete exploration processes. In particular, it must happen with vanishing probability that a zero of the limiting process  $Z^\lambda$  appears only as the limit of small *positive* local minima of the discrete processes  $Z^{N,p}$ . In addition, we must show that there are with high probability no large discrete components which appear late enough in the exploration that they are not represented in the limit. Several stages of the argument will be based on a comparison of  $F(N,p)$  and the original model  $G(N,p)$ , for which some of the results are easier, or known.

3.1. *Large components are explored early.* Theorem 1.9 establishes convergence of the exploration processes on compact time intervals. To use this to study the sizes of the largest components in  $F(N,p)$ , we need to ensure that these largest components appear early in the exploration process. We establish this in the following series of lemmas.

**Lemma 3.1.** *Fix  $\lambda^+ \in \mathbb{R}$ . Then*

$$\lim_{\gamma \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} \mathbb{P} \left( C_1(F(N,p)) \geq \gamma N^{2/3} \right) = 0. \tag{3.1}$$

*Proof:* We have

$$\begin{aligned} \mathbb{E} [|C_{F(N,p)}(v)|] &= \mathbb{E} \left[ \frac{1}{N} \sum_{i \geq 1} C_i(F(N,p)) \right] \\ &\geq \frac{1}{N} \mathbb{E} [|C_1(F(N,p))|^2] \geq \frac{(\gamma N^{2/3})^2}{N} \mathbb{P} \left( C_1(F(N,p)) \geq \gamma N^{2/3} \right). \end{aligned} \tag{3.2}$$

Result (3.1) then follows by using Lemma 2.3 and the coupling of Lemma 2.2 to control the first expectation in (3.2).  $\square$

The following lemma shows that critical components will with high probability include a vertex with label  $O(N^{1/3})$ .

**Lemma 3.2.** *Fix  $\epsilon > 0$ , and  $\lambda^+ \in \mathbb{R}$ . Then*

$$\lim_{\Gamma \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\Lambda \leq \lambda^+} \mathbb{P} \left( \exists \text{ cpt } C \text{ in } F(N,p) : |C| \geq \epsilon N^{2/3} \right) \tag{3.3}$$

$$\text{and } C \cap \{1, \dots, \lfloor \Gamma N^{1/3} \rfloor\} = \emptyset = 0.$$

*Proof:* Applying Markov’s inequality to (2.2), and summing over all vertices,

$$\limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} N^{-2/3} \mathbb{E} \left[ \left| \left\{ v \in [N] : |C_{F(N,p)}(v)| \geq \epsilon N^{2/3} \right\} \right| \right] \leq \frac{\Theta^{\lambda^+}}{\epsilon}.$$

Therefore,

$$\limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} \mathbb{E} \left[ \# \text{cpts } C \text{ in } F(N,p) \text{ s.t. } |C| \geq \epsilon N^{2/3} \right] \leq \frac{\Theta^{\lambda^+}}{\epsilon^2}.$$

Then, since the labelling is independent of the component sizes in  $F(N, p)$ ,

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} \mathbb{E} \left[ \# \text{cpts } C \text{ in } F(N, p) \text{ s.t. } |C| \geq \epsilon N^{2/3} \right. \\ & \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \left. \text{and } C \cap \{1, \dots, \lfloor \Gamma N^{1/3} \rfloor\} = \emptyset \right] \\ & \leq \frac{\Theta^{\lambda^+}}{\epsilon^2} \times \lim_{N \rightarrow \infty} \frac{\binom{N - \lfloor \Gamma N^{1/3} \rfloor}{\lfloor \epsilon N^{2/3} \rfloor}}{\binom{N}{\lfloor \epsilon N^{2/3} \rfloor}} \\ & \leq \frac{\Theta^{\lambda^+}}{\epsilon^2} \exp(-\Gamma \epsilon), \end{aligned}$$

where the final limit can be evaluated using Stirling’s approximation. (3.3) follows.  $\square$

We now use the previous result to show that the largest components will typically appear near the start of the exploration process. This will be important later, since if large critical components appear arbitrarily late in the exploration process, then they cannot be treated via convergence on compact intervals.

**Lemma 3.3.** *Fix  $\epsilon > 0$  and  $\lambda^+ \in \mathbb{R}$  as before. Then*

$$\lim_{T \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} \mathbb{P} \left( \exists \text{ cpt } C \text{ in } F(N, p) : |C| \geq \epsilon N^{2/3} \right) \quad (3.4)$$

$$\text{and } C \cap \{v_1, \dots, v_{\lfloor TN^{2/3} \rfloor}\} = \emptyset = 0,$$

where  $(v_1, v_2, \dots, v_N)$  is the exploration process of  $F(N, p)$ .

*Proof:* Fix  $\Gamma > 0$ , and let  $C_{F(N,p)}(k)$  be the component of vertex  $k$  in  $F(N, p)$ . We define the events

$$A^{\Gamma, T}(F(N, p)) := \{|C_{F(N,p)}(1)| + \dots + |C_{F(N,p)}(\lfloor \Gamma N^{1/3} \rfloor)| > TN^{2/3}\},$$

$$B^{\epsilon, \Gamma}(F(N, p)) := \left\{ \exists \text{ cpt } C \text{ in } F(N, p) : |C| \geq \epsilon N^{2/3}, C \cap \{1, \dots, \lfloor \Gamma N^{1/3} \rfloor\} = \emptyset \right\},$$

as in Lemma 3.2. Then, by Markov’s inequality,

$$\mathbb{P} (A^{\Gamma, T}(F(N, p))) \leq \frac{\Gamma N^{1/3} \mathbb{E} [|C_{F(N,p)}(1)|]}{TN^{2/3}},$$

So by Lemma 2.3

$$\limsup_{N \rightarrow \infty} \sup_{\Lambda(N,p) \leq \lambda^+} \mathbb{P} (A^{\Gamma, T}(F(N, p))) \leq \frac{\Theta^{\lambda^+} \Gamma}{T}. \quad (3.5)$$

Whenever  $F(N, p)$  contains a component of size at least  $\epsilon N^{2/3}$  which is not exhausted during the first  $TN^{2/3}$  steps of the exploration process, at least one of  $A^{\Gamma, T}(F(N, p))$  and  $B^{\epsilon, \Gamma}(F(N, p))$  must hold. So take  $\Gamma = \sqrt{T}$ , then let  $T \rightarrow \infty$ . By (3.5) and Lemma 3.2, the result follows.  $\square$

3.2. *Components and excursions up to time  $T$  - notation and goal.* Throughout this section, we fix  $\lambda \in \mathbb{R}$  and work with a sequence  $p(N)$  for which  $\Lambda(N, p) := N^{1/3}[Np - 1] \rightarrow \lambda$  as before. We will mostly suppress notational dependence on  $p$  and  $\lambda$ .

First, we establish the notation we will use to describe the sequence of rescaled component sizes in  $F(N, p)$ . Fix  $T > 0$ , then:

- Let  $\mathcal{C}^N := (C_1^N, C_2^N, \dots)$  be the sequence of sizes of components of  $F(N, p)$ , in non-increasing order.
- Analogously, let  $\mathcal{C}^{N,T} := (C_1^{N,T}, C_2^{N,T}, \dots)$  be the sequence of sizes of the components of  $F(N, p)$  which intersect with  $\{v_1, \dots, v_{\lfloor TN^{2/3} \rfloor}\}$ , an initial segment of the breadth-first ordering introduced in Section 1.3. That is, least one vertex has been seen by step  $\lfloor TN^{2/3} \rfloor$  of the exploration process. Again, we assume the sequence is ordered such that  $C_1^{N,T} \geq C_2^{N,T} \geq \dots$

We first show that for any  $T < \infty$  the excursion lengths in the exploration processes on the interval  $[0, T]$  appear correctly in the limit.

In everything that follows, we work on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  whose existence is guaranteed by the Skorohod representation theorem, where  $\tilde{Z}^{N,p} \xrightarrow{\mathbb{P}\text{-a.s.}} Z^\lambda$  with respect to the topology of uniform convergence on compact intervals.

In a mild abuse of notation, let  $C_1^T \geq C_2^T \geq \dots$  be the lengths of excursions of  $Z^\lambda$  above zero which have non-empty intersection with  $[0, T]$ , in non-increasing order. Set  $\mathcal{C}^T := (C_1^T, C_2^T, \dots)$ . We will prove the following convergence result for the components seen within the first  $TN^{2/3}$  steps of the exploration process.

**Proposition 3.4.** *Fix  $T > 0$  and  $k \geq 1$ . Then as  $N \rightarrow \infty$ ,*

$$N^{-2/3}(C_1^{N,T}, C_2^{N,T}, \dots, C_k^{N,T}) \xrightarrow{d} (C_1^T, C_2^T, \dots, C_k^T). \tag{3.6}$$

The concern is that the reflected exploration process might regularly approach zero without actually hitting zero, and thus starting a new component. To show that this effect does not appear in the limit, we use the fact that the components of  $F(N, p)$  have the structure of uniform random trees. Then we can approximate the exploration process within a component by a Brownian excursion, and show that the probability of zeros in the limit which do not correspond to the start or end of a component is small.

**Definition 3.5.** Given two sequences  $a = (a_1, \dots, a_k)$ ,  $b = (b_1, \dots, b_k)$ , let  $a^\downarrow, b^\downarrow$  denote the sequences rearranged into non-increasing order. Then, we say  $a \succeq b$  or  $a$  *weakly majorises*  $b$  if for every  $\ell \leq k$ ,

$$\sum_{i=1}^{\ell} a_i^\downarrow \geq \sum_{i=1}^{\ell} b_i^\downarrow.$$

It is easy to check that this gives a pre-order on  $(\mathbb{R} \cup \{\infty\})^k$ , and a partial order on non-increasing sequences finer than the standard ordering.

We will prove Proposition 3.4 by stochastically sandwiching  $\mathcal{C}^T$  between any weak limit of  $\mathcal{C}^{N,T}$ , and any weak limit of a related sequence of lengths  $\mathcal{C}^{N,T,\delta}$  associated with  $\tilde{Z}^N$ , which will be defined shortly. This stochastic ordering will be with respect to weak majorisation. The two directions of this sandwiching argument occupy the next two sections. Finally, we show that for small enough  $\delta$ , these outer distributions are close in the sense of the Lévy–Prohorov metric.



3.3. *Limits of component sizes stochastically majorise excursion lengths.* We show that limit points of  $\mathcal{C}^{N,T}$  majorise  $\mathcal{C}^T$ ,  $\mathbb{P}$ -almost surely.

For any reference time  $s \in [0, T]$ , we define

$$\ell(s) := \sup\{t \leq s : Z(t) = 0\}, \quad \ell^N(s) := \sup\{t \leq s : \tilde{Z}^N(t) = 0\},$$

$r(s) := \inf\{t \in [s, \infty) : Z(t) = 0\}$ ,  $r^N(s) := \beta(T) \wedge \inf\{t \in [s, r(T)] : \tilde{Z}^N(t) = 0\}$ , so that  $r(s) - \ell(s)$  is the width of the excursion of  $Z$  around time  $s$ . It will be convenient to avoid values of  $s$  where  $\ell^N$  and  $r^N$  are non-constant, so we define

$$\bar{\mathbb{Q}} := \bigcup_{N \in \mathbb{N}} N^{-2/3} \mathbb{Z}.$$

We also define the event

$$\Psi^T := \left\{ \tilde{Z}^N \rightarrow Z \text{ uniformly on } [0, r(T)], Z \text{ continuous on } [0, r(T)] \right\}.$$

Since  $r(T) < \infty$  almost surely, and  $\tilde{Z}^N \rightarrow Z$  uniformly on compact intervals, we have  $\mathbb{P}(\Psi^T) = 1$ . It follows easily that on  $\Psi^T$ ,

$$\limsup_{N \rightarrow \infty} \ell^N(s) \leq \ell(s), \quad \liminf_{N \rightarrow \infty} r^N(s) \geq r(s), \quad \forall s \in [0, T]. \tag{3.7}$$

Now, on  $\Psi^T$ , given  $Z$ , choose  $s_1, \dots, s_k \in [0, T] \setminus \bar{\mathbb{Q}}$  such that each  $s_i$  lies in the  $i$ th longest excursion of  $Z$ , which has non-empty intersection with  $[0, T]$ . That is,  $r(s_i) - \ell(s_i) = C_i^T$ . Now consider any limit point

$$(\bar{\ell}(s_1), \dots, \bar{\ell}(s_k), \bar{r}(s_1), \dots, \bar{r}(s_k), \bar{C}_1^T, \dots, \bar{C}_k^T), \tag{3.8}$$

of  $(\ell^N(s_1), \dots, \ell^N(s_k), r^N(s_1), \dots, r^N(s_k), C_1^{N,T}, \dots, C_k^{N,T})$ , as  $N \rightarrow \infty$ , where we allow  $\bar{C}_1^T$  and at most one of the  $\bar{r}(s_i)$  to be  $\infty$ . By compactness, we can be sure that there are such limit points. To avoid introducing extra notation, we will assume that (3.8) is a true limit, rather than a subsequential limit.

By (3.7), for any  $m \leq k$ ,

$$\bigcup_{i=1}^m [\bar{\ell}(s_i), \bar{r}(s_i)] \supseteq \bigcup_{i=1}^m [\ell(s_i), r(s_i)],$$

where the sets in the union on the right-hand side have disjoint interiors. By construction of  $\ell^N(s_i), r^N(s_j)$ , any pair of intervals  $[\ell^N(s_i), r^N(s_i)]$  and  $[\ell^N(s_j), r^N(s_j)]$  are either equal or disjoint. Therefore the intervals in the union on the left-hand side are either equal or have disjoint interiors. So for any limit point (3.8), let  $\Gamma_m \subseteq [m]$  be some set of indices such that

$$[\bar{\ell}(s_i), \bar{r}(s_i)] \neq [\bar{\ell}(s_j), \bar{r}(s_j)], \quad \forall i \neq j \in \Gamma_m$$

and

$$\bigcup_{i \in \Gamma_m} [\bar{\ell}(s_i), \bar{r}(s_i)] \supseteq \bigcup_{i=1}^m [\ell(s_i), r(s_i)].$$

Furthermore, we may demand  $\Gamma_1 \subseteq \Gamma_2 \subseteq \dots \subseteq \Gamma_k$ . Thus

$$\sum_{i \in \Gamma_m} (\bar{r}(s_i) - \bar{\ell}(s_i)) \geq \sum_{i=1}^m (r(s_i) - \ell(s_i)).$$

That is,

$$\begin{aligned} &(\bar{r}(s_1) - \bar{\ell}(s_1), \dots, \bar{r}(s_{|\Gamma_k|}) - \bar{\ell}(s_{|\Gamma_k|}), 0, \dots, 0) \\ &\succeq (r(s_1) - \ell(s_1), \dots, r(s_k) - \ell(s_k)). \end{aligned} \tag{3.9}$$

For any  $N$ , and any  $s \in [0, T] \setminus \bar{\mathbb{Q}}$ , the interval  $[\ell^N(s), r^N(s)]$  is associated via the reflected exploration process with exactly one component of  $F(N, p)$ . The size of this component is at least  $(r^N(s) - \ell^N(s))N^{2/3}$ .

**Note.** The two cases where the size of the component is not exactly equal to  $(r^N(s) - \ell^N(s))N^{2/3}$  are: 1) when  $r^N(s) = r(T)$ ; 2) when  $\tilde{Z}^N(s) = 0$ . In the latter case, since we have excluded the possibility  $s \in N^{-2/3}\mathbb{Z}$ , it must hold that  $\tilde{Z}^N$  is locally constant and equal to zero around  $s$ , so the component has size 1.

For large enough  $N$ , the intervals  $\{[\bar{\ell}^N(s_i), \bar{r}^N(s_i)] : i \in \Gamma_k\}$  are disjoint, and so  $N^{-2/3}(C_1^{N,T}, \dots, C_k^{N,T}) \succeq (r^N(s_1) - \ell^N(s_1), \dots, r^N(s_{|\Gamma_k|}) - \ell^N(s_{|\Gamma_k|}), 0, \dots, 0)$ . Since majorisation is preserved under limits (as the relation is a finite union of closed sets in  $\mathbb{R}^k \times \mathbb{R}^k$ ), we obtain

$$(\bar{C}_1^T, \dots, \bar{C}_k^T) \succeq (\bar{r}(s_1) - \bar{\ell}(s_1), \dots, \bar{r}(s_{|\Gamma_k|}) - \bar{\ell}(s_{|\Gamma_k|}), 0, \dots, 0).$$

So, combining with (3.9), we obtain

$$(\bar{C}_1^T, \dots, \bar{C}_k^T) \succeq (C_1^T, \dots, C_k^T), \tag{3.10}$$

which holds for every limit point  $(\bar{C}_1^T, \dots, \bar{C}_k^T)$  of  $N^{-2/3}(C_1^{N,T}, \dots, C_k^{N,T})$  on the event  $\Psi^T$  and so, in particular,  $\mathbb{P}$ -almost surely.

**3.4. Stochastic sandwiching via excursions above  $\delta$ .** We now give a stochastic lower bound for  $\mathcal{C}^T$  (again in the sense of weak majorisation).

Fix some  $\delta > 0$ . For any realisation of the path  $\tilde{Z}^N$ , the set  $\mathcal{D}^{N,\delta,T} := \{s \in [0, T] : \tilde{Z}^N(s) > \delta\}$  is a finite union of left-closed, right-open intervals. Let  $N^{-2/3}(C_1^{N,\delta,T} \geq \dots \geq C_k^{N,\delta,T})$  be the sequence of the  $k$  largest lengths of those intervals which are contained within the support of some excursion of  $\tilde{Z}^N$  (above zero) which has non-empty intersection with  $[0, T]$ . As before, augment with zeros if necessary. (Note that the  $N^{-2/3}$  ensures that  $C_1^{N,\delta,T}$  has the same scaling as  $C_1^{N,T}$ .) Certainly, for any  $\delta$ ,  $(C_1^{N,T}, \dots, C_k^{N,T}) \succeq (C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$  for each trajectory of  $\tilde{Z}^N$ . We will show that  $\mathcal{C}^T$  majorises limit points of  $N^{-2/3}(C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$ , again  $\mathbb{P}$ -almost surely.

Again, we work on the event  $\Psi^T$ . Then, consider  $\mathcal{D}^T := \{s \in [0, T] : Z(s) > 0\}$ , the collection of open intervals where the limit process  $Z$  is positive. On  $\Psi^T$ , for large enough  $N$ , we have  $\tilde{Z}^N(s) \leq \delta/2$  whenever  $Z(s) = 0$ , and so  $\mathcal{D}^{N,\delta,T} \subseteq \mathcal{D}^T$ . Therefore the sequence of all interval lengths in  $\mathcal{D}^{N,\delta,T}$  in non-increasing order is majorised by the corresponding ordered sequence of interval lengths in  $\mathcal{D}^T$ . So in particular

$$(C_1^T, \dots, C_k^T) \succeq N^{-2/3}(C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T}),$$

for large enough  $N$ , and hence on  $\Psi^T$  any limit point  $(\bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T})$  of  $N^{-2/3}(C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$  satisfies

$$(C_1^T, \dots, C_k^T) \succeq (\bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T}).$$

By (3.1), the collection  $N^{-2/3}(C_1^{N,T}, \dots, C_k^{N,T}, C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})_{N \geq 1}$  is tight in  $\mathbb{R}^k \times \mathbb{R}^k$ . Let  $(\bar{C}_1^T, \dots, \bar{C}_k^T, \bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T})$  be any joint weak limit of the sequence  $N^{-2/3}(C_1^{N,T}, \dots, C_k^{N,T}, C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$ . Since  $\mathbb{P}(\Psi^T) = 1$ , by combining with (3.10) we have shown that

$$(\bar{C}_1^T, \dots, \bar{C}_k^T) \succeq_{st} (C_1^T, \dots, C_k^T) \succeq_{st} (\bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T}). \tag{3.11}$$

3.5. *Comparing  $C^{N,T}$  and  $C^{N,\delta,T}$  via uniform trees.* We will now show for small  $\delta$ , any weak limits  $(\bar{C}_1^T, \dots, \bar{C}_k^T)$  and  $(\bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T})$  are themselves close in distribution in the sense of the Lévy–Prohorov metric on  $\mathbb{R}^k$ . To do this, we have to bound above the probability that the exploration process drops below height  $\delta N^{1/3}$  in the middle of an excursion above zero of width  $\Theta(N^{2/3})$ . The components of  $F(N, p)$  are, conditional on their sizes, uniform trees. Aldous (1991) explains how to view the uniform tree as an example of a Galton–Watson tree, here with Poisson offspring distribution, conditioned on its total progeny. From this, large excursions of  $\tilde{Z}^N$  are well-approximated by Brownian excursions. We then can then bound the probability that  $\tilde{Z}^N$  hits  $\delta$  without hitting zero using standard estimates.

Let  $\mathcal{T}_K$  be a uniform choice from the  $K^{K-2}$  unordered trees with vertex labels given by  $[K]$ . Then, let  $1 = S_0^{\mathcal{T}_K}, S_1^{\mathcal{T}_K}, \dots, S_K^{\mathcal{T}_K} = 0$ , be the corresponding breadth-first exploration process. The appropriate rescaling to consider is then  $\tilde{S}^{\mathcal{T}_K}(s) := \frac{1}{\sqrt{K}} S_{\lfloor Ks \rfloor}^{\mathcal{T}_K}$ , for  $s \in [0, 1]$ . From the description of  $\mathcal{T}_K$  as a conditioned Galton–Watson process, we follow Le Gall (see Lemma 1.16 of Le Gall, 2005) in using Kaigh’s scaling limit result for conditioned random walks (Kaigh, 1976) to obtain

$$\left( \tilde{S}^{\mathcal{T}_K}(s), s \in [0, 1] \right) \xrightarrow{d} (B^{\text{ex}}(s), s \in [0, 1]), \tag{3.12}$$

where  $B^{\text{ex}}$  is a standard normalised Brownian excursion on  $[0, 1]$ , and convergence is in the uniform topology.

We say the event  $\chi^{N,T}(\delta, \epsilon, \gamma)$  holds if  $\exists M, K \in \mathbb{Z}_{\geq 0}$  with  $\frac{K}{N^{2/3}} \geq \gamma$ , and  $\frac{M}{N^{2/3}} \leq T$ , such that  $\{v_M, \dots, v_{M+K-1}\}$  is a component of  $F(N, p)$ , and

$$\exists n \in [\epsilon K, (1 - \epsilon)K] \text{ s.t. } \tilde{Z}^N \left( \frac{M+n}{N^{2/3}} \right) \leq \delta. \tag{3.13}$$

That is,  $F(N, p)$  has a component of size at least  $\gamma N^{2/3}$  which is seen, at least partially, in the exploration process before time  $TN^{2/3}$ , and for which the exploration process takes a small value in the macroscopic interior of the interval defining the component. Now, given any  $M, K$ , and conditional on the vertices  $\{v_M, \dots, v_{M+K-1}\}$ , and the statement that they form a component, the structure of this component is a uniform tree. That is,

$$(Z_M^N, \dots, Z_{M+K-1}^N) \stackrel{d}{=} (S_1^{\mathcal{T}_K}, \dots, S_K^{\mathcal{T}_K}).$$

Therefore the following processes on  $s \in [0, 1]$  can be identified in distribution:

$$\left( \tilde{Z}^N \left( \frac{M+sK}{N^{2/3}} \right) \right) = \left( N^{-1/3} Z_{\lfloor M+sK \rfloor}^N \right) \stackrel{d}{=} \left( N^{-1/3} S_{\lfloor sK \rfloor}^{\mathcal{T}_K} \right) = \left( \frac{K^{1/2}}{N^{1/3}} \tilde{S}^{\mathcal{T}_K}(s) \right).$$

Therefore, for every  $M, K$ , conditional on any choice of vertices  $\{v_M, \dots, v_{M+K-1}\}$ , the probability that (3.13) holds is equal to the probability that

$$\inf_{s \in [\epsilon, 1-\epsilon]} \tilde{S}^{\mathcal{T}_K}(s) \leq \frac{N^{1/3}}{K^{1/2}} \delta. \tag{3.14}$$

By assumption  $\frac{N^{1/3}}{K^{1/2}}\delta \leq \gamma^{-1/2}\delta$ , and by (3.12), and the Portmanteau lemma,

$$\begin{aligned} \limsup_{K \rightarrow \infty} \mathbb{P} \left( \inf_{s \in [\epsilon, 1-\epsilon]} \tilde{S}^{\mathcal{T}_K}(s) \leq \gamma^{-1/2}\delta \right) &\leq \limsup_{K \rightarrow \infty} \mathbb{P} \left( \inf_{s \in [\epsilon, 1-\epsilon]} \tilde{S}^{\mathcal{T}_K}(s) < 2\gamma^{-1/2}\delta \right) \\ &\leq \mathbb{P} \left( \min_{s \in [\epsilon, 1-\epsilon]} B^{\text{ex}}(s) < 2\gamma^{-1/2}\delta \right). \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \limsup_{N \rightarrow \infty} \mathbb{P} (\chi^{N,T}(\delta, \epsilon, \gamma)) &\leq \mathbb{E} \left[ \# \text{ cpts size } \geq \gamma N^{2/3} \text{ seen before } TN^{2/3} \text{ in } Z^N \right] \\ &\quad \times \mathbb{P} \left( \min_{s \in [\epsilon, 1-\epsilon]} B(s) < 2\gamma^{-1/2}\delta \right) \\ \limsup_{N \rightarrow \infty} \mathbb{P} (\chi^{N,T}(\delta, \epsilon, \gamma)) &\leq \left( \frac{T}{\gamma} + 1 \right) \mathbb{P} \left( \min_{s \in [\epsilon, 1-\epsilon]} B(s) < 2\gamma^{-1/2}\delta \right). \end{aligned} \tag{3.15}$$

Given  $\epsilon, \gamma$ , we can choose  $\delta > 0$  so that the RHS of (3.15) is arbitrarily small. Now, fix some  $\gamma > 2\epsilon$ , and consider the event  $\chi^{N,T}(\delta, \frac{\epsilon}{2\gamma}, \epsilon)$ . Then, when  $\chi^{N,T}(\delta, \frac{\epsilon}{2\gamma}, \epsilon)$  does not hold, for every component with size  $K \geq \epsilon N^{2/3}$ , there is a unique excursion of  $Z^N$  above  $\delta N^{1/3}$  of length at least  $K(1 - \frac{\epsilon}{\gamma})$ . We call such an excursion above  $\delta N^{1/3}$  a *principal excursion*. If we also have  $C_1^N \leq \gamma N^{2/3}$ , then the length of any principal excursion is at least  $K - \epsilon N^{2/3}$ . Thus, any other excursion above  $\delta N^{1/3}$  within the component of size  $K$ , has length at most  $\epsilon N^{2/3}$ .

So, consider any  $i \leq k$  such that  $C_1^{N,T} \geq \dots \geq C_i^{N,T} \geq \epsilon N^{2/3}$ . Then, if both  $\chi^{N,T}(\delta, \frac{\epsilon}{2\gamma}, \epsilon)^c$  and  $\{C_1^N \leq \gamma N^{2/3}\}$  occur, at most  $i - 1$  elements of  $(C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$  can be larger than  $C_i^{N,T}$ . These are the principal excursions obtained from each of  $C_1^{N,T}, \dots, C_{i-1}^{N,T}$ . No other excursions above  $\delta N^{2/3}$  obtained from  $C_1^{N,T}, \dots, C_{i-1}^{N,T}$  are relevant, since they have lengths at most  $\epsilon N^{2/3}$ . However, these principal excursions from  $C_1^{N,T}, \dots, C_i^{N,T}$  all have length at least  $C_i^{N,T}(1 - \frac{\epsilon}{\gamma})$ . Thus we obtain

$$C_i^{N,T} \geq C_i^{N,\delta,T} \geq C_i^{N,T} \left(1 - \frac{\epsilon}{\gamma}\right) \geq C_i^{N,T} - \epsilon N^{2/3}. \tag{3.16}$$

And so

$$\begin{aligned} \limsup_{N \rightarrow \infty} \mathbb{P} \left( \max_{i \in [k]} \left| C_i^{N,T} - C_i^{N,\delta,T} \right| > \epsilon N^{2/3} \right) \\ \leq \limsup_{N \rightarrow \infty} \mathbb{P} \left( C_1^N > \gamma N^{2/3} \right) + \limsup_{N \rightarrow \infty} \mathbb{P} \left( \chi^{N,T}(\delta, \frac{\epsilon}{2\gamma}, \epsilon) \right). \end{aligned}$$

For fixed  $\epsilon > 0$ , letting  $\gamma \rightarrow \infty$  we can make the first term on the RHS small, and then by letting  $\delta \downarrow 0$  we can make the second term small. In particular, we can demand

$$\limsup_{N \rightarrow \infty} \mathbb{P} \left( \max_{i \in [k]} \left| C_i^{N,T} - C_i^{N,\delta,T} \right| > \epsilon N^{2/3} \right) \leq \epsilon. \tag{3.17}$$

Now, recall  $(\bar{C}_1^T, \dots, \bar{C}_k^T, \bar{C}_1^{\delta,T}, \dots, \bar{C}_k^{\delta,T})$  is some joint weak limit of  $N^{-2/3}(C_1^{N,T}, \dots, C_k^{N,T}, C_1^{N,\delta,T}, \dots, C_k^{N,\delta,T})$ . Let  $\pi$  be the usual Lévy–Prohorov metric for probability measures on  $\mathbb{R}^k$ , with respect to the  $\ell_\infty$  norm on  $\mathbb{R}^k$ . From

(3.11) and (3.17), we have for each  $\epsilon > 0$ ,

$$\begin{aligned} \pi(\mathcal{L}(\bar{C}_1^T, \dots, \bar{C}_k^T), \mathcal{L}(\bar{C}_1^{\delta, T}, \dots, C_k^{\delta, T})) &\leq \epsilon, \\ (\bar{C}_1^{\delta, T}, \dots, \bar{C}_k^{\delta, T}) &\preceq (C_1^T, \dots, C_k^T) \preceq (\bar{C}_1^T, \dots, \bar{C}_k^T). \end{aligned}$$

From this, it is easy to see that  $\pi(\mathcal{L}(C_1^T, \dots, C_k^T), \mathcal{L}(\bar{C}_1^T, \dots, \bar{C}_k^T)) \leq k\epsilon$ . Since  $\epsilon > 0$  is arbitrary, we find  $(\bar{C}_1^T, \dots, \bar{C}_k^T) \stackrel{d}{=} (C_1^T, \dots, C_k^T)$ , and thus the required convergence in distribution (3.6) follows, completing the proof of Proposition 3.4.

3.6. Proof of Theorem 1.8.

*Convergence in the product topology.* In both the discrete exploration processes and the limiting SDEs, we would expect the  $k$  largest components/excursions to appear early. From (3.4),

$$\limsup_{T \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P} \left( \max_{i \in [k]} |C_i^{N, T} - C_i^N| \leq \epsilon N^{2/3} \right) = 1.$$

Recall again that  $\lambda$  is fixed. By comparing the drifts, we can couple  $Z^\lambda$  and  $B^\lambda$ , as defined in Proposition 1.1 and Proposition 1.5, such that  $Z^\lambda(t) \leq B^\lambda(t)$  for all  $t \geq 0$ . The largest excursion of  $B^\lambda$  above zero is almost surely finite, and so the same holds for  $Z^\lambda$ . Thus, now turning to  $(C_1^\lambda, C_2^\lambda, \dots)$ , the sequence of excursion lengths of  $Z^\lambda$  in decreasing order,

$$\limsup_{T \rightarrow \infty} \mathbb{P}((C_1^T, \dots, C_k^T) = (C_1^\lambda, \dots, C_k^\lambda)) = 1.$$

So we can lift (3.6) and conclude that

$$N^{-2/3}(C_1^N, \dots, C_k^N) \xrightarrow{d} (C_1^\lambda, \dots, C_k^\lambda), \tag{3.18}$$

as  $N \rightarrow \infty$ .

*Convergence in  $\ell^2$ .* To lift (3.18) to convergence in  $\ell^2$ , it is enough to show that for every  $\epsilon > 0$

$$\lim_{k \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P} \left( \sum_{i > k} (C_i^N)^2 \geq \epsilon N^{4/3} \right) = 0. \tag{3.19}$$

The corresponding result for  $G(N, p)$ , namely that for every  $\epsilon > 0$ ,

$$\lim_{k \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P} \left( \sum_{i > k} C_i(G(N, p))^2 \geq \epsilon N^{4/3} \right) = 0, \tag{3.20}$$

is implied by Proposition 1.1; the argument establishing this property for  $G(N, p)$  is given in the proof of Proposition 15 of Aldous (1997). We can again use the coupling of  $F(N, p)$  and  $G(N, p)$  to derive (3.19) from (3.20). We require the following lemma.

**Lemma 3.6.** *Consider a family of finite non-increasing sequences*

$$\mathbf{x}_1 := (x_{1,1}, \dots, x_{1,\ell_1}), \mathbf{x}_2 := (x_{2,1}, \dots, x_{2,\ell_2}), \dots, \mathbf{x}_M := (x_{M,1}, \dots, x_{M,\ell_M})$$

with sums  $x_1 \geq x_2 \geq \dots \geq x_M$ . Now let  $y_1 \geq y_2 \geq \dots \geq y_{\sum \ell_i}$  be the permutation of all the terms  $x_{i,j}$  in descending order. Then, for any  $k \geq 1$  and  $\eta > 0$ ,

$$\sum_{i > k/\eta} y_i^2 \leq \eta \sum_{i=1}^k x_i^2 + \sum_{i > k} x_i^2. \tag{3.21}$$

*Proof of Lemma 3.6:* Noting that each sequence  $\mathbf{x}_i$  is non-increasing, set  $m_i := \max\{m : x_{i,m} \geq \eta x_i\}$  (with  $\max \emptyset := 0$ ). So  $m_i \leq \frac{1}{\eta}$  for all  $i$ . It follows that for any  $k \geq 1$ ,

$$\sum_{i=1}^k \sum_{j=1}^{m_i} x_{i,j}^2 \leq \sum_{i=1}^{k/\eta} y_i^2.$$

Thus

$$\sum_{i > k/\eta} y_i^2 \leq \sum_{i=1}^k \sum_{j \geq m_i} x_{i,j}^2 + \sum_{i > k} x_i^2.$$

But  $\sum_{j \geq m_i} x_{i,j}^2 \leq \sum_{j \geq m_i} x_{i,j}(\eta x_i) \leq \eta x_i^2$ , and so (3.21) follows.  $\square$

*Proof of (3.19):* Under the coupling of Lemma 2.2, each component of  $G(N, p)$  is the disjoint union of components of  $F(N, p)$ . So we apply Lemma 3.6 with  $x_1 \geq x_2 \geq \dots$  as the component sizes of  $G(N, p)$ , and  $y_1 \geq y_2 \geq \dots$  as the components of  $F(N, p)$ , obtaining, for every  $\eta > 0$ ,

$$\begin{aligned} \mathbb{P} \left( N^{-4/3} \sum_{i \geq k/\eta} (C_i^N)^2 \geq \epsilon \right) &\leq \mathbb{P} \left( N^{-4/3} \sum_{i \geq 1} C_i(G(N, p))^2 \geq \frac{\epsilon}{2\eta} \right) \\ &\quad + \mathbb{P} \left( N^{-4/3} \sum_{i \geq k} C_i(G(N, p))^2 \geq \frac{\epsilon}{2} \right). \end{aligned}$$

So, using (3.20) to eliminate the second term on the RHS, we have

$$\begin{aligned} \limsup_{k \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{P} \left( N^{-4/3} \sum_{i \geq k/\eta} (C_i^N)^2 \geq \epsilon \right) &\leq \limsup_{N \rightarrow \infty} \mathbb{P} \left( N^{-4/3} \sum_{i \geq 1} C_i(G(N, p))^2 \geq \frac{\epsilon}{2\eta} \right) \\ &\leq \frac{2\eta}{\epsilon} \Theta^\lambda, \end{aligned}$$

applying Lemma 2.3 and Markov’s inequality in the final step. Taking  $\eta \rightarrow 0$  completes the proof of (3.19), and of Theorem 1.8.  $\square$

#### 4. Detailed combinatorial calculations

4.1. *Proof of Lemma 2.10.* For convenience, we recall the statement of Lemma 2.10:

**Lemma.** Fix  $\lambda^- < \lambda^+ \in \mathbb{R}$ . Given  $p \in (0, 1)$ , let  $\Lambda = \Lambda(N, p) = N^{1/3}(Np - 1)$ . Then

$$\mathbb{P}(G(N, p) \text{ acyclic}) = (1 + o(1))g(\Lambda)e^{3/4\sqrt{2\pi}N^{-1/6}}, \tag{4.1}$$

uniformly for  $\Lambda \in [\lambda^-, \lambda^+]$  as  $N \rightarrow \infty$ .

*Proof:* For this range of  $p$ , we will see that the sum in (2.4) is dominated by contributions on the scale  $m = \frac{N}{2} + \frac{\Lambda N^{2/3}}{2} + \Theta(N^{1/2})$ . Shortly we will be required to approximate these relevant contributions in detail, but first we show that contributions from outside this regime vanish as  $N \rightarrow \infty$ . We consider those  $m$  for which

$$\left| m - \frac{N}{2} - \frac{\Lambda N^{2/3}}{2} \right| \geq N^{3/5}.$$

Let  $B \sim \text{Bin}\left(\binom{N}{2}, p\right)$ . Since  $f(N, m) \leq \binom{N}{m}$ ,

$$\begin{aligned} (1-p)^{\binom{N}{2}} & \left[ \sum_{m=0}^{\lceil N/2 + \Lambda N^{2/3}/2 - N^{3/5} \rceil} f(N, m) \left(\frac{p}{1-p}\right)^m \right. \\ & \left. + \sum_{\lfloor N/2 + \Lambda N^{2/3}/2 + N^{3/5} \rfloor}^{N-1} f(N, m) \left(\frac{p}{1-p}\right)^m \right] \\ & \leq \mathbb{P}\left(|B - \binom{N}{2}p| \geq N^{3/5}\right) \\ & \leq \frac{\text{Var}(B)}{N^{6/5}} \leq \frac{N^2 p}{2N^{6/5}} \leq \frac{1 + \lambda^+ N^{-1/3}}{2N^{1/5}} \ll N^{-1/6}. \end{aligned} \quad (4.2)$$

Here we used Chebyshev's inequality, which is sufficient for our purposes, but note that the probability of this moderate deviation event for  $B$  decays exponentially in some positive power of  $N$ .

Given  $\Lambda \in \mathbb{R}$  and  $m \leq N \in \mathbb{N}$ , define  $x = x(N, m, \Lambda) = \frac{\sqrt{2}}{N^{1/2}} \left[ m - \frac{N}{2} - \frac{\Lambda N^{2/3}}{2} \right]$ . Then, we consider the set of  $m$  satisfying

$$\left| m - \frac{N}{2} - \frac{\Lambda N^{2/3}}{2} \right| \leq N^{3/5}, \text{ that is, } |x| \leq \sqrt{2}N^{1/10}. \quad (4.3)$$

Thus

$$N - m = \frac{N}{2} - \frac{\Lambda}{2}N^{2/3} - \frac{x}{\sqrt{2}}N^{1/2}, \text{ and so } \frac{2(N - m)}{N} = 1 - \Lambda N^{-1/3} - \sqrt{2}xN^{-1/2}.$$

From this, we obtain

$$\begin{aligned} \log\left(\frac{2(N - m)}{N}\right) & = -\Lambda N^{-1/3} - \sqrt{2}xN^{-1/2} - \frac{\Lambda^2}{2}N^{-2/3} - \sqrt{2}\Lambda xN^{-5/6} \\ & \quad - \frac{\Lambda^3}{3}N^{-1} - x^2N^{-1} + O(N^{-16/15}), \end{aligned}$$

uniformly on the set of  $m$  defined at (4.3). In calculating the scale of this final error term, we use that  $|x| \leq \sqrt{2}N^{1/10}$ . Then

$$\begin{aligned} (N - m) \log\left(\frac{2(N - m)}{N}\right) & = -\left[ \frac{\Lambda}{2}N^{2/3} + \frac{x}{\sqrt{2}}N^{1/2} + \frac{\Lambda^2}{4}N^{1/3} + \frac{\Lambda x}{\sqrt{2}}N^{1/6} + \frac{\Lambda^3}{3} + \frac{x^2}{2} \right] \\ & \quad + \left[ \frac{\Lambda^2}{2}N^{1/3} + \frac{\Lambda x}{\sqrt{2}}N^{1/6} + \frac{\Lambda^3}{4} \right] \\ & \quad + \left[ \frac{\Lambda x}{\sqrt{2}}N^{1/6} + x^2 \right] + O\left(N^{-1/15}\right) \\ & = -\frac{\Lambda}{2}N^{2/3} - \frac{x}{\sqrt{2}}N^{1/2} + \frac{\Lambda^2}{4}N^{1/3} + \frac{\Lambda x}{\sqrt{2}}N^{1/6} \\ & \quad - \frac{\Lambda^3}{12} + \frac{x^2}{2} + O\left(N^{-1/15}\right). \end{aligned}$$

We now return to (1.2) and use Stirling’s approximation and the expression we have just shown, as well as continuity of  $g$ . Uniformly on the set of  $m$  in (4.3), (for which, recall,  $N - m = (1 + o(1))N/2$ ),

$$\begin{aligned}
 f(N, m) &= (1 + o(1)) \frac{\sqrt{2\pi} N^{N-1/6}}{2^{N-m} (N-m)!} g\left(\frac{2m-N}{N^{2/3}}\right), \\
 &= (1 + o(1)) \frac{g(\Lambda) \sqrt{2\pi} N^{N-1/6}}{2^{N-m}} \cdot \frac{1}{\sqrt{2\pi} \sqrt{N-m}} \left(\frac{e}{N-m}\right)^{N-m} \\
 &= (1 + o(1)) g(\Lambda) \sqrt{2} N^{m-2/3} \exp(N-m) \exp\left(- (N-m) \log\left(\frac{2(N-m)}{N}\right)\right) \\
 &= (1 + o(1)) g(\Lambda) \sqrt{2} N^{m-2/3} \exp\left(\frac{N}{2} - \frac{\Lambda}{2} N^{2/3} - \frac{x}{\sqrt{2}} N^{1/2}\right) \\
 &\quad \times \exp\left(\frac{\Lambda}{2} N^{2/3} + \frac{x}{\sqrt{2}} N^{1/2} - \frac{\Lambda^2}{4} N^{1/3} - \frac{\Lambda x}{\sqrt{2}} N^{1/6} + \frac{\Lambda^3}{12} - \frac{x^2}{2}\right) \\
 &= (1 + o(1)) g(\Lambda) \sqrt{2} N^{m-2/3} \exp\left(\frac{N}{2} - \frac{\Lambda^2}{4} N^{1/3} - \frac{\Lambda x}{\sqrt{2}} N^{1/6} + \frac{\Lambda^3}{12} - \frac{x^2}{2}\right). \tag{4.4}
 \end{aligned}$$

Now, we have

$$\begin{aligned}
 \binom{N}{2} \log(1-p) &= \binom{N}{2} \left[-\frac{1+\Lambda N^{-1/3}}{N} - \frac{1}{2} N^{-2} + O(N^{-7/3})\right] \\
 &= -\frac{N}{2} - \frac{\Lambda}{2} N^{2/3} + \frac{1}{4} + O\left(N^{-1/3}\right), \tag{4.5}
 \end{aligned}$$

and also

$$\begin{aligned}
 \log\left(\frac{Np}{1-p}\right) &= \log(1 + \Lambda N^{-1/3}) - \log(1-p) \\
 &= \Lambda N^{-1/3} - \frac{\Lambda^2}{2} N^{-2/3} + \frac{\Lambda^3}{3} N^{-1} + N^{-1} + O\left(N^{-4/3}\right).
 \end{aligned}$$

At this point, recall the definition

$$m = \frac{N}{2} + \frac{\Lambda}{2} N^{2/3} + \frac{x}{\sqrt{2}} N^{1/2}.$$

So, uniformly on the set of  $m$  for which  $|x| \leq \sqrt{2} N^{1/10}$ , as before,

$$\begin{aligned}
 m \log\left(\frac{Np}{1-p}\right) &= \left[\frac{\Lambda}{2} N^{2/3} - \frac{\Lambda^2}{4} N^{1/3} + \frac{\Lambda^3}{6} + \frac{1}{2}\right] \\
 &\quad + \left[\frac{\Lambda^2}{2} N^{1/3} - \frac{\Lambda^3}{4}\right] + \frac{\Lambda x}{\sqrt{2}} N^{1/6} + O\left(N^{-1/6}\right), \tag{4.6}
 \end{aligned}$$

where each bracket corresponds to a term in the definition of  $m$ .

Therefore, combining (4.5) and (4.6), uniformly in the same sense,

$$(1-p)^{\binom{N}{2}} \left(\frac{p}{1-p}\right)^m = (1 + o(1)) N^{-m} \exp\left(-\frac{N}{2} + \frac{\Lambda^2}{4} N^{1/3} + \frac{\Lambda x}{\sqrt{2}} N^{1/6} - \frac{\Lambda^3}{12} + \frac{3}{4}\right). \tag{4.7}$$

Combining (4.4) and (4.7), we obtain

$$(1-p)^{\binom{N}{2}} \left(\frac{p}{1-p}\right)^m f(N, m) = (1 + o(1)) g(\Lambda) \sqrt{2} N^{-2/3} \exp\left(-\frac{x^2}{2} + \frac{3}{4}\right). \tag{4.8}$$

We now fix  $N$  and  $\Lambda$ , and sum this quantity over the range of  $m$  given by (4.3). Recall that  $x$  is linear in  $m$ , with scaling factor  $\frac{N^{1/2}}{\sqrt{2}}$ , and so as  $N \rightarrow \infty$ , the sum of (4.8) over this range of  $m$  converges after rescaling to an integral. That is,



$$\begin{aligned}
 & (1-p)^{\binom{N}{2}} \sum_{m=\lfloor N/2+\Lambda N^{2/3}/2-N^{3/5} \rfloor}^{\lceil N/2+\Lambda N^{2/3}/2+N^{3/5} \rceil} f(N, m) \left(\frac{p}{1-p}\right)^m \\
 &= (1+o(1))e^{3/4}g(\Lambda)\sqrt{2}N^{-2/3} \sum_{m=\lfloor N/2+\Lambda N^{2/3}/2-N^{3/5} \rfloor}^{\lceil N/2+\Lambda N^{2/3}/2+N^{3/5} \rceil} e^{-x^2/2} \\
 &= (1+o(1))e^{3/4}g(\Lambda)\sqrt{2}\frac{N^{-1/6}}{\sqrt{2}} \int_{-\infty}^{\infty} e^{-x^2/2} dx, \\
 &= (1+o(1))e^{3/4}g(\Lambda)\sqrt{2\pi}N^{-1/6}.
 \end{aligned}$$

Combining with (4.2), which showed that contributions to the sum (2.4) outside this range of  $m$  are  $o(N^{-1/6})$ , we obtain the required result.  $\square$

4.2. *Proof of Lemma 2.13.* We recall the statement of Lemma 2.13:

**Lemma.** Fix constants  $\lambda^-, \lambda^+, \epsilon, K, T$  as in Definition 2.12. Then,

$$\begin{aligned}
 \mathbb{P}(G(N', p) \in \mathcal{A}_{N', r, k}) &= (1+o(1))g(\Lambda-s-a)e^{3/4}N^{-5/6}ba^{-3/2} \quad (4.9) \\
 &\quad \times \exp\left(-b(\Lambda-s) - \frac{b^2}{2a} + \frac{(\Lambda-s-a)^3 - (\Lambda-s)^3}{6}\right),
 \end{aligned}$$

uniformly on  $(N', p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$ , as  $N \rightarrow \infty$ .

*Proof:* We will add the required uniformity in  $N'$  at the end of this proof. First, we show

$$\begin{aligned}
 \mathbb{P}(G(N, p) \in \mathcal{A}_{N, r, k}) &= (1+o(1))g(\lambda-a)e^{3/4}N^{-5/6}ba^{-3/2} \quad (4.10) \\
 &\quad \times \exp\left(-b\lambda - \frac{b^2}{2a} + \frac{(\lambda-a)^3 - \lambda^3}{6}\right),
 \end{aligned}$$

uniformly on  $(p, r, k)$  such that  $(N, p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, 0)$ , as  $N \rightarrow \infty$ .

Subject to the constraint that vertices  $1, \dots, r$  are in different tree components, with sum equal to  $k$ , there are  $\binom{N-r}{k-r}$  ways to choose which remaining vertices are part of this stack forest. Given this choice, we can view the trees as rooted at the vertices  $[r]$ . In particular, Cayley’s formula states that there are  $rk^{k-r-1}$  such labelled rooted forests. Hence

$$\begin{aligned}
 & \mathbb{P}(G(N, p) \in \mathcal{A}_{N, r, k}) \\
 &= (1-p)^{\binom{N}{2}} \binom{N-r}{k-r} \left(\frac{p}{1-p}\right)^{k-r} rk^{k-r-1} \sum_{m=0}^{N-k-1} f(N-k, m) \left(\frac{p}{1-p}\right)^m. \quad (4.11)
 \end{aligned}$$

By Lemma 2.10, uniformly on  $(p, k)$  and for any  $r$  such that  $(N, p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, 0)$  (in fact  $r$  is arbitrary),

$$(1-p)^{\binom{N-k}{2}} \sum_{m=0}^{N-k-1} f(N-k, m) \left(\frac{p}{1-p}\right)^m = (1+o(1))g(\Lambda(N-k, p))e^{3/4}\sqrt{2\pi}N^{-1/6}.$$

Recall that this final sum is, up to a power of  $(1-p)$ , the probability that  $G(N-k, p)$  is acyclic. We also have

$$\begin{aligned} \Lambda(N-k, p) &= (N - aN^{2/3})^{1/3} \left[ (N - aN^{2/3})p - 1 \right] \\ &= (1 + o(1)) N^{1/3} \left[ (Np - 1) - aN^{-1/3} \right] \\ &= (1 + o(1)) [\Lambda(N, p) - a + o(1)]. \end{aligned}$$

So, again uniformly on  $(p, r, k)$  such that  $(N, p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, 0)$ ,

$$(1-p)^{\binom{N-k}{2}} \sum_{m=0}^{N-k-1} f(N-k, m) \left(\frac{p}{1-p}\right)^m = (1+o(1))g(\Lambda-a)e^{3/4}\sqrt{2\pi}N^{-1/6}. \tag{4.12}$$

We now carefully address the other terms in (4.11), starting with the term  $(1-p)^{\binom{N}{2}-\binom{N-k}{2}} \left(\frac{p}{1-p}\right)^{k-r}$ . Recall that  $Np = 1 + \Lambda N^{-1/3}$ . Firstly

$$\begin{aligned} \log \left[ \left(1 + \Lambda N^{-1/3}\right)^{k-r} \right] &= \left[ aN^{2/3} - bN^{1/3} \right] \left[ \Lambda N^{-1/3} - \frac{\Lambda^2}{2} N^{-2/3} + O(N^{-1}) \right] \\ &= \Lambda a N^{1/3} - \Lambda b - \frac{\Lambda^2 a}{2} + O(N^{-1/3}). \end{aligned}$$

Also

$$\begin{aligned} \binom{N}{2} - \binom{N-k}{2} - k + r &= \frac{N^2}{2} - \frac{(N-k)^2}{2} + \frac{k}{2} - k + r \\ &= Nk - \frac{k^2}{2} + O(N^{2/3}) \\ &= aN^{5/3} - \frac{a^2}{2}N^{4/3} + O(N^{2/3}), \end{aligned}$$

from which

$$\begin{aligned} \log \left[ (1-p)^{\binom{N}{2}-\binom{N-k}{2}-k+r} \right] &= \log \left[ \left(1 - N^{-1} - \Lambda N^{-4/3}\right)^{\binom{N}{2}-\binom{N-k}{2}-k+r} \right] \\ &= \left[ aN^{5/3} - \frac{a^2}{2}N^{4/3} + O(N^{2/3}) \right] \left[ -N^{-1} - \Lambda N^{-4/3} + O(N^{-2}) \right] \\ &= -aN^{2/3} - \Lambda a N^{1/3} + \frac{a^2}{2}N^{1/3} + \frac{\Lambda a^2}{2} + O(N^{-1/3}). \end{aligned}$$

From this,

$$\begin{aligned} (1-p)^{\binom{N}{2}-\binom{N-k}{2}} \left(\frac{p}{1-p}\right)^{k-r} &= (1+o(1))N^{-(k-r)} \exp \left( -aN^{2/3} + \frac{a^2}{2}N^{1/3} - \Lambda b + \frac{\Lambda a}{2}(a-\Lambda) \right). \tag{4.13} \end{aligned}$$

Turning now to the binomial coefficient  $\binom{N-r}{k-r}$  in (4.11), we treat each factorial separately. First observe that

$$\begin{aligned} \log \left[ \left(1 - bN^{-2/3}\right)^{N-bN^{1/3}} \right] &= \left[ N - bN^{1/3} \right] \left[ -bN^{-2/3} + O\left(N^{-4/3}\right) \right] \\ &= -bN^{1/3} + O\left(N^{-1/3}\right) \\ \log \left[ \left(1 - aN^{-1/3}\right)^{N-aN^{2/3}} \right] &= -aN^{2/3} + \frac{a^2}{2}N^{1/3} + \frac{a^3}{6} + O\left(N^{-1/3}\right) \\ \log \left[ \left(1 - \frac{b}{a}N^{-1/3}\right)^{aN^{2/3}-bN^{1/3}} \right] &= -bN^{1/3} + \frac{b^2}{2a} + O\left(N^{-1/3}\right). \end{aligned}$$

Then Stirling’s approximation gives

$$\begin{aligned} \left(N - bN^{1/3}\right)! &= (1 + o(1)) \frac{\sqrt{2\pi N}}{e^{N-bN^{1/3}}} \left(N - bN^{1/3}\right)^{N-bN^{1/3}} \\ &= (1 + o(1)) \frac{\sqrt{2\pi N}}{e^{N-bN^{1/3}}} N^{N-bN^{1/3}} \exp\left(-bN^{1/3}\right) \\ \left(N - aN^{2/3}\right)! &= (1 + o(1)) \frac{\sqrt{2\pi N}}{e^{N-aN^{2/3}}} N^{N-aN^{2/3}} \exp\left(-aN^{2/3} + \frac{a^2}{2}N^{1/3} + \frac{a^3}{6}\right) \\ \left(aN^{2/3} - bN^{1/3}\right)! &= \\ &= (1 + o(1)) \frac{\sqrt{2\pi} \sqrt{aN^{2/3}}}{e^{aN^{2/3}-bN^{1/3}}} a^{aN^{2/3}-bN^{1/3}} N^{\frac{2}{3}[aN^{2/3}-bN^{1/3}]} \exp\left(-bN^{1/3} + \frac{b^2}{2a}\right). \end{aligned}$$

So we obtain

$$\begin{aligned} \binom{N-r}{k-r} &= (1 + o(1)) \frac{1}{\sqrt{2\pi}} a^{-(aN^{2/3}-bN^{1/3}+1/2)} N^{\frac{1}{3}(aN^{2/3}-bN^{1/3}-1)} \\ &\quad \times \exp\left(aN^{2/3} - \frac{a^2}{2}N^{1/3} - \frac{b^2}{2a} - \frac{a^3}{6}\right). \end{aligned} \tag{4.14}$$

The final ingredient of (4.11) is the term

$$rk^{k-r-1} = ba^{aN^{2/3}-bN^{1/3}-1} N^{\frac{2}{3}[aN^{2/3}-bN^{1/3}]-\frac{1}{3}}. \tag{4.15}$$

To recover (4.11), we study the product of (4.12), (4.13), (4.14) and (4.15). Note that  $\exp\left(-\frac{\Lambda^2 a}{2} + \frac{\Lambda a^2}{2} - \frac{a^3}{6}\right) = \exp\left(\frac{(\Lambda-a)^3 - \Lambda^3}{6}\right)$ . So we can treat the terms in (4.11) uniformly on  $(p, r, k)$  such that  $(N, p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, 0)$ , as  $N \rightarrow \infty$  and obtain (4.10) as required.

We now finish the proof of (4.9), where in addition we require a uniform estimate over  $N' \in [N - TN^{2/3}, N]$ . We consider  $(N', p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$  as  $N \rightarrow \infty$ . Observe that

$$\Lambda' := \Lambda(N', p) = (1 + o(1))(\Lambda(N, p) - s), \quad N' = (1 + o(1))N, \tag{4.16}$$

$$b' := b(N', r) = (1 + o(1))b(N, r), \quad a' = a(N', k) = (1 + o(1))a(N, k). \tag{4.17}$$

Now fix  $\delta \in (0, \epsilon)$ . Then, for large enough  $N$ ,

$$\begin{aligned} (N', p, r, k) &\in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T) \\ \Rightarrow (N', p, r, k) &\in \Psi^{N'}(\lambda^- - T - \delta, \lambda^+ + \delta, \epsilon - \delta, K + \delta, 0). \end{aligned} \tag{4.18}$$

Certainly  $N - TN^{2/3} \rightarrow \infty$  as  $N \rightarrow \infty$ , so by (4.10) and (4.18),

$$\begin{aligned} \mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k}) &= (1 + o(1))g(\Lambda' - a')e^{3/4}N'^{-5/6}b'a'^{-3/2} \\ &\quad \times \exp\left(-b'\Lambda' - \frac{b'^2}{2a'} + \frac{(\Lambda' - a')^3 - \Lambda'^3}{6}\right), \end{aligned}$$

uniformly on  $(N', p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$  as  $N \rightarrow \infty$ . Finally, using (4.16), (4.17), and the fact that  $g$  is uniformly continuous, we may conclude

$$\begin{aligned} \mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k}) &= (1 + o(1))g(\Lambda - s - a)e^{3/4}N^{-5/6}ba^{-3/2} \\ &\quad \times \exp\left(-b(\Lambda - s) - \frac{b^2}{2a} + \frac{(\Lambda - a - s)^3 - (\Lambda - s)^3}{6}\right), \end{aligned}$$

as required, uniformly on  $(N', p, r, k) \in \Psi^N(\lambda^-, \lambda^+, \epsilon, K, T)$ . □

4.3. *Proof of Lemma 2.15.* We repeat the statement of Lemma 2.15:

**Lemma.** Given the same constants as in Lemma 2.14, there exist constants  $M < \infty$  and  $\gamma > 0$  such that

$$\frac{(k + 1)\mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k+1})}{k\mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k})} \leq 1 - \gamma N^{-2/3}, \tag{4.19}$$

for large enough  $N$ , whenever  $(N', p, r) \in \bar{\Psi}_0^N(\lambda^-, \lambda^+, K, T)$  and  $k \in [MN^{2/3}, N' - 1]$ .

*Proof:* Again, we will use (4.11), which for convenience we recall here.

$$\begin{aligned} \mathbb{P}(G(N, p) \in \mathcal{A}_{N,r,k}) &= (1 - p)^{\binom{N}{2}} \binom{N-r}{k-r} \left(\frac{p}{1-p}\right)^{k-r} r k^{k-r-1} \sum_{m=0}^{N-k-1} f(N - k, m) \left(\frac{p}{1-p}\right)^m \\ &= (1 - p)^{\binom{N}{2} - \binom{N-k}{2}} \binom{N-r}{k-r} \left(\frac{p}{1-p}\right)^{k-r} r k^{k-r-1} F(N - k, p). \end{aligned}$$

We apply this to (4.19) (with  $N$  replaced by  $N'$ ). Note that  $\binom{N'-r}{k+1-r} / \binom{N'-r}{k-r} = \frac{N'-k}{k+1-r}$ , and  $\binom{N'-k}{2} - \binom{N'-k-1}{2} = N' - k - 1$ . We obtain

$$\begin{aligned} &\frac{(k + 1)\mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k+1})}{k\mathbb{P}(G(N', p) \in \mathcal{A}_{N',r,k})} \\ &= \frac{(k + 1)(1 - p)^{-\binom{N'-k-1}{2}} \binom{N'-r}{k+1-r} \left(\frac{p}{1-p}\right)^{k-r} r (k + 1)^{k-r} F(N' - k - 1, p)}{k(1 - p)^{-\binom{N'-k}{2}} \binom{N'-r}{k-r} r k^{k-r-1} F(N' - k, p)} \\ &= \frac{k + 1}{k + 1 - r} (1 - p)^{N' - k - 2} \frac{N' - k}{N} \left(1 + \Lambda N^{-1/3}\right) \\ &\quad \times \left(\frac{k + 1}{k}\right)^{k-r} \cdot \frac{F(N' - k - 1, p)}{F(N' - k, p)}. \tag{4.20} \end{aligned}$$

We proceed in two parts. First we control the ratio of the  $F(N' - k, p)$  terms using (2.5). Then, we control the ratio of the remaining terms with an elementary but long Taylor expansion.

First, note that from the second inequality in (2.5), that for  $k \leq N' - 1$ ,

$$1 - \frac{F(N' - k, p)}{F(N' - k - 1, p)} \leq \frac{1}{2}(N' - k - 1)p^2 \mathbb{E} \left[ |C^{N' - k - 1, p}(v)| \right].$$

where  $|C^{N,p}(v)|$  is the size of the component containing a uniformly-chosen vertex  $v$  in  $G(N, p)$ . Now, via (4.16),

$$\limsup_{N \rightarrow \infty} \Lambda \left( N - \lfloor MN^{2/3} \rfloor, p \right) \leq \lambda^+ - M.$$

When  $k \geq MN^{2/3}$ , we have

$$N^{-1/3} \mathbb{E} \left[ |C^{N'-k-1,p}(v)| \right] \leq N^{-1/3} \mathbb{E} \left[ |C^{N-\lfloor MN^{2/3} \rfloor,p}(v)| \right],$$

and so from (2.2),

$$\limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N-TN^{2/3}, N] \\ k \geq MN^{2/3}}} N^{-1/3} \mathbb{E} \left[ |C^{N'-k-1,p}(v)| \right] \leq \Theta^{\lambda^+ - M}.$$

We obtain

$$\limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N-TN^{2/3}, N] \\ \lambda(N,p) \in [\lambda^-, \lambda^+] \\ 0 \leq k \leq N'-1}} N^{2/3} \left[ 1 - \frac{F(N'-k,p)}{F(N'-k-1,p)} \right] \leq \frac{1}{2} \Theta^{\lambda^+ - M},$$

from which it follows that

$$\limsup_{N \rightarrow \infty} \sup_{\substack{N' \in [N-TN^{2/3}, N] \\ \lambda(N,p) \in [\lambda^-, \lambda^+] \\ 0 \leq k \leq N'-1}} N^{2/3} \left[ \frac{F(N'-k-1,p)}{F(N'-k,p)} - 1 \right] \leq \frac{1}{2} \Theta^{\lambda^+ - M}. \tag{4.21}$$

We now treat the remaining terms in the ratio (4.20), that is

$$\frac{k+1}{k+1-r} \cdot (1-p)^{N'-k-2} \cdot \frac{N'-k}{N} \cdot \left( 1 + \lambda N^{-1/3} \right) \cdot \left( \frac{k+1}{k} \right)^{k-r}.$$

We split the calculation into several steps. Recall the rescalings  $a = \frac{k}{N^{2/3}}$  and  $b = \frac{r}{N^{1/3}}$ . Since we assume  $k \geq MN^{2/3}$ , we have  $\frac{1}{a} = O(1)$ .

$$\begin{aligned} \log \left( \frac{k+1}{k+1-r} \right) &= -\log \left( 1 - \frac{r}{k+1} \right) = \frac{r}{k+1} + \frac{1}{2} \left( \frac{r}{k+1} \right)^2 + O(N^{-1}) \\ &= \frac{b}{a} N^{-1/3} + \frac{b^2}{2a^2} N^{-2/3} + O(N^{-1}), \\ \log \left( 1 + \lambda N^{-1/3} \right) &= \lambda N^{-1/3} - \frac{\lambda^2}{2} N^{-2/3} + O(N^{-1}), \\ \log \left[ \left( \frac{k+1}{k} \right)^{k-r} \right] &= \left[ aN^{2/3} - bN^{1/3} \right] \left[ \frac{1}{a} N^{-2/3} - \frac{1}{2a^2} N^{-4/3} + O(N^{-2}) \right] \\ &= 1 - \frac{b}{a} N^{-1/3} - \frac{1}{2a} N^{-2/3} + O(N^{-1}). \end{aligned}$$

The final two terms in the product require extra care, because there is no finite upper bound on  $a$ . However, since  $a \leq N^{1/3}$ , we can still handle the error in the following term:

$$\begin{aligned} \log \left[ (1-p)^{N'-k-2} \right] &= \left[ N - (s+a)N^{2/3} - 2 \right] \left[ -N^{-1} - \lambda N^{-4/3} + O(N^{-2}) \right] \\ &= -1 + (s - \lambda + a)N^{-1/3} + \lambda(a+s)N^{-2/3} + O(N^{-1}). \end{aligned}$$

Finally, we have

$$\log\left(\frac{N' - k}{N}\right) = \log\left(1 - sN^{-1/3} - aN^{-1/3}\right) \leq -(a + s)N^{-1/3} - \frac{1}{2}(a + s)^2N^{-2/3}.$$

So there exists a constant  $C = C(\lambda^-, \lambda^+, \epsilon, K, T) < \infty$  such that

$$\begin{aligned} & \log\left[\frac{(k + 1)(1 - p)^{-\binom{N' - k - 1}{2}} \binom{N' - r}{k + 1 - r} \left(\frac{p}{1 - p}\right) r(k + 1)^{k - r}}{k(1 - p)^{-\binom{N' - k}{2}} \binom{N' - r}{k - r} r k^{k - r - 1}}\right] \\ & \leq N^{-2/3} \left[-\frac{1}{2}(\Lambda - (a + s))^2 + \frac{b^2}{2a^2} - \frac{1}{2a}\right] + \frac{C}{N}, \end{aligned} \tag{4.22}$$

uniformly on  $(N', p, r) \in \Psi_0^N(\lambda^-, \lambda^+, \epsilon, K, T)$  and  $k \geq MN^{2/3}$ , as  $N \rightarrow \infty$ . Recall that  $b \in [\epsilon, K]$ , and that  $k \geq MN^{2/3}$  is equivalent to  $a \geq M$ . So for large enough  $M$ , the term  $\frac{b^2}{2a^2}$  is dominated by the term  $-\frac{1}{2a}$  in (4.22). Then it holds that for large enough  $N$ ,

$$\frac{(k + 1)(1 - p)^{-\binom{N' - k - 1}{2}} \binom{N' - r}{k + 1 - r} \left(\frac{p}{1 - p}\right) r(k + 1)^{k - r}}{k(1 - p)^{-\binom{N' - k}{2}} \binom{N' - r}{k - r} r k^{k - r - 1}} \leq 1 - \frac{1}{3K}N^{-2/3}.$$

Using Lemma 2.4, we now also demand that  $M$  be large enough that  $\Theta^{\lambda^+ - M} \leq \frac{1}{6K}$ . So combining with (4.21), we can now approximate the LHS of (4.19) as required. Now take  $\gamma \in (0, \frac{1}{6K})$ , and we find that for large enough  $N$

$$\frac{(k + 1)\mathbb{P}(G(N, p) \in \mathcal{A}_{N', r, k + 1})}{k\mathbb{P}(G(N, p) \in \mathcal{A}_{N', r, k})} \leq 1 - \gamma N^{-2/3}.$$

□

### 5. Regularity of $g$ and $\alpha$

In this section, we prove various regularity properties of the function  $g$  defined in (1.3), and from this the technical properties we require about  $\alpha$ . In particular, the content of Lemma 1.4 is a subset of what follows.

5.1. *Properties of  $g$ .* Recall the definition of  $g$  from (1.3):

$$g(x) := \frac{1}{\pi} \int_0^\infty \exp(-\frac{4}{3}t^{3/2}) \cos(xt + \frac{4}{3}t^{3/2}) dt.$$

Britikov (1988) observes that  $g$  is, after stretching by a factor  $(2/3)^{2/3}$ , the density of the canonical stable distribution with self-similarity exponent  $\alpha = 3/2$  and skewness  $\beta = -1$ . The following lemma, which restates the regularity properties of  $g$  required for Lemma 1.4, follows from standard properties of such distributions, as stated, for example, by Zolotarev (1986).

**Lemma 5.1.** *The function  $g$  defined in (1.3) is smooth and positive and has finite integral. Furthermore, it is bounded, uniformly continuous, and satisfies  $g(x) \rightarrow 0$  as  $x \rightarrow \pm\infty$ .*

5.1.1. *The function  $\alpha$  is well-defined.* For  $k \in \mathbb{N}$ , we define

$$J_k(b, \lambda) := \int_0^\infty a^{-k/2} g(\lambda - a) \exp\left(\frac{(\lambda - a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da, \quad b > 0, \lambda \in \mathbb{R}. \quad (5.1)$$

**Lemma 5.2.** *For each  $k \in \mathbb{N}$ , this function  $J_k$  is well-defined and continuous, and has partial derivative with respect to  $b$  given by*

$$\frac{\partial}{\partial b} J_k(b, \lambda) = -b J_{k+2}(b, \lambda). \quad (5.2)$$

Furthermore, the function  $\alpha(b, \lambda) := \frac{J_1(b, \lambda)}{J_3(b, \lambda)}$  defined in (1.4) is also well-defined, continuous and differentiable with respect to  $b$ .

*Proof:* To show that  $J_k(b, \lambda) < \infty$ , we consider the integral in (5.1) separately over the ranges  $a \in (0, 1]$  and  $a \in [1, \infty)$ . We have

$$\int_1^\infty a^{-k/2} g(\lambda - a) \exp\left(\frac{(\lambda - a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da < e^{\lambda^3/6} \int_1^\infty g(\lambda - a) da < \infty, \quad (5.3)$$

and

$$\int_0^1 a^{-k/2} g(\lambda - a) \exp\left(\frac{(\lambda - a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) da < e^{\lambda^3/6} g_{\max} \int_0^1 a^{-k/2} \exp\left(-\frac{b^2}{2a}\right) da < \infty. \quad (5.4)$$

Thus we have  $J_k(b, \lambda) < \infty$ .

Since the bounds (5.3) and (5.4) hold locally uniformly in  $(b, \lambda)$ , continuity of  $J_k$  follows from the dominated convergence theorem.

We can check that we may differentiate (5.1) inside the integral to obtain that (5.2) holds for all  $k \geq 1$ . Well-definedness and continuity of  $\alpha(b, \lambda) := \frac{J_1(b, \lambda)}{J_3(b, \lambda)}$  follow immediately, since  $J_3(b, \lambda) > 0$  for all  $b > 0, \lambda \in \mathbb{R}$ , and furthermore  $\alpha(b, \lambda)$  is differentiable in its first argument as required, with

$$\frac{\partial}{\partial b} \alpha(b, \lambda) = \frac{b J_1(b, \lambda) J_5(b, \lambda)}{J_3(b, \lambda)^2} - b, \quad (5.5)$$

through two applications of (5.2). □

5.2. *Monotonicity of  $\alpha$ .* Heuristically, we can view (1.4) as the expectation of  $a$  with respect to the measure with density  $a^{-3/2} g(\lambda - a)$ , weighted by a factor  $\exp(-\frac{b^2}{2a})$ . Increasing  $b$  reweights in favour of larger values of  $a$ , so  $\alpha(b, \lambda)$  is increasing in  $b$ . We make this formal with the following straightforward lemma.

**Lemma 5.3.** *Let  $f, h$  be functions  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $h$  is strictly increasing, and the integrals*

$$\int_0^\infty a f(a) h(a) da, \quad \int_0^\infty f(a) da,$$

*exist and are finite. Then*

$$\frac{\int_0^\infty a f(a) h(a) da}{\int_0^\infty f(a) h(a) da} > \frac{\int_0^\infty a f(a) da}{\int_0^\infty f(a) da}.$$

**Corollary 5.4.**  *$\alpha(b, \lambda)$  is increasing as a function of  $b$ .*

*Proof:* Fix  $\lambda \in \mathbb{R}$  and  $b' > b$ , then set

$$f(a) := a^{-3/2}g(\lambda - a) \exp\left(\frac{(\lambda-a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right), \text{ and } h(a) := \exp\left(-\frac{b'^2-b^2}{2a}\right),$$

in Lemma 5.3. □

5.3. *Lipschitz property of  $\alpha$ .* The following proposition establishes the behaviour of  $\alpha(b, \lambda)$  as  $b \downarrow 0$  in the sense required to complete the proof of Lemma 2.14. It also establishes a Lipschitz condition for  $\alpha$ , required in Proposition 1.5 for the well-posedness of the reflected SDE (1.5).

**Proposition 5.5.** *Given  $-\infty < \lambda^- < \lambda^+ < \infty$ , we have*

$$\lim_{b \downarrow 0} \sup_{\lambda \in [\lambda^-, \lambda^+]} \alpha(b, \lambda) = 0. \tag{5.6}$$

Furthermore, given  $\rho < \infty$ , there exists a constant  $C < \infty$  such that  $\alpha$  satisfies the Lipschitz condition

$$|\alpha(b, \lambda) - \alpha(b', \lambda)| \leq C|b - b'|, \quad b, b' \in (0, \rho], \lambda \in [\lambda^-, \lambda^+]. \tag{5.7}$$

*Proof:* To show (5.7), it suffices to prove the following:

$$\sup_{b \in (0, \rho], \lambda \in [\lambda^-, \lambda^+]} \left| \frac{\partial}{\partial b} \alpha(b, \lambda) \right| < \infty. \tag{5.8}$$

The steps we take to prove (5.7) will also allow us to read off (5.6). Recall the expression (5.5) from the proof of Lemma 5.2:

$$\frac{\partial}{\partial b} \alpha(b, \lambda) = \frac{bJ_1(b, \lambda)J_5(b, \lambda)}{J_3(b, \lambda)^2} - b. \tag{5.5}$$

From Lemma 5.2, we know that  $\frac{\partial}{\partial b} \alpha(b, \lambda)$  is continuous, and so to verify (5.8), it remains to consider the limit as  $b \downarrow 0$ . We examine the behaviour of each of  $J_1(b, \lambda), J_3(b, \lambda), J_5(b, \lambda)$  in this limit.

First, we consider  $J_1$ . We define

$$\gamma_1(\lambda) := e^{\lambda^3/6} \int_0^\infty a^{-1/2}g(\lambda - a)da,$$

which is seen to be finite by a similar decomposition to (5.3) and (5.4). Then

$$\begin{aligned} \gamma_1(\lambda) - J_1(b, \lambda) &\leq g_{\max} \int_0^\infty a^{-1/2} \left[ e^{\lambda^3/6} - \exp\left(\frac{(\lambda-a)^3}{6}\right) \exp\left(-\frac{b^2}{2a}\right) \right] da \\ &\leq e^{\lambda^3/6} g_{\max} \int_0^\infty a^{-1/2} \left[ 1 - \exp\left(-\frac{b^2}{2a}\right) \right] da, \end{aligned}$$

and so by monotone convergence we have as  $b \downarrow 0$ ,

$$\sup_{\lambda \in (-\infty, \lambda^+]} |J_1(b, \lambda) - \gamma_1(\lambda)| \rightarrow 0. \tag{5.9}$$

Substituting  $u = \frac{b^2}{2a}$  into (5.1) gives

$$J_3(b, \lambda) = \frac{\sqrt{2}}{b} \int_0^\infty u^{-1/2}g\left(\lambda - \frac{b^2}{2u}\right) \exp\left(\frac{(\lambda-\frac{b^2}{2u})^3}{6}\right) \exp(-u)du.$$

So we define

$$\gamma_3(\lambda) := \sqrt{2}g(\lambda)e^{\lambda^3/6} \int_0^\infty u^{-1/2} \exp(-u)du,$$



and then by dominated convergence and uniform continuity of  $g$ ,

$$\lim_{b \downarrow 0} \sup_{\lambda \in (-\infty, \lambda^+]} |bJ_3(b, \lambda) - \gamma_3(\lambda)| = 0. \tag{5.10}$$

A very similar argument can be deployed to obtain

$$\lim_{b \downarrow 0} \sup_{\lambda \in (-\infty, \lambda^+]} |b^3 J_5(b, \lambda) - \gamma_5(\lambda)| = 0,$$

where

$$\gamma_5(\lambda) := 2\sqrt{2}g(\lambda)e^{\lambda^3/6} \int_0^\infty u^{1/2} \exp(-u)du.$$

So we can return to (5.5), which we rewrite as

$$\frac{\partial}{\partial b} \alpha(b, \lambda) = \frac{J_1(b, \lambda) \cdot b^3 J_5(b, \lambda)}{(bJ_3(b, \lambda))^2} - b.$$

We now take the limit  $b \downarrow 0$ , for  $\lambda \in [\lambda^-, \lambda^+]$ . This denominator is uniformly bounded away from zero for  $\lambda \in [\lambda^-, \lambda^+]$ . So we obtain

$$\lim_{b \downarrow 0} \sup_{\lambda \in [\lambda^-, \lambda^+]} \left| \frac{\partial}{\partial b} \alpha(b, \lambda) - \frac{\gamma_1(\lambda)\gamma_5(\lambda)}{\gamma_3(\lambda)^2} \right| = 0. \tag{5.11}$$

Now,  $\gamma_3, \gamma_5$  are clearly continuous, and  $\gamma_1$  is also continuous by the same argument as given for continuity of  $J_k$  in the proof of Lemma 5.2. Furthermore,  $\gamma_3$  is positive, and so we have

$$\max_{\lambda \in [\lambda^-, \lambda^+]} \gamma_1(\lambda) < \infty, \quad \min_{\lambda \in [\lambda^-, \lambda^+]} \gamma_3(\lambda) > 0.$$

Taken with (5.10), the latter shows that

$$\lim_{b \downarrow 0} \inf_{\lambda \in [\lambda^-, \lambda^+]} J_3(b, \lambda) = \infty.$$

Therefore, since  $\alpha(b, \lambda) = \frac{J_1(b, \lambda)}{J_3(b, \lambda)}$ , using (5.9) as well, we obtain precisely the first required statement (5.6).

For similar reasons, we have

$$\max_{\lambda \in [\lambda^-, \lambda^+]} \frac{\gamma_1(\lambda)\gamma_5(\lambda)}{\gamma_3(\lambda)^2} < \infty. \tag{5.12}$$

Since  $\frac{\partial}{\partial b} \alpha(b, \lambda)$  is continuous on  $(0, \rho] \times [\lambda^-, \lambda^+]$ , from (5.11) and (5.12), it's clear that

$$\sup_{b \in (0, \rho], \lambda \in [\lambda^-, \lambda^+]} \left| \frac{\partial}{\partial b} \alpha(b, \lambda) \right| < \infty,$$

from which (5.7) follows. This completes the proof of Proposition 5.5. □

5.4. *Existence of  $Z^\lambda$ .* First we prove Proposition 1.5, which asserts that  $Z^\lambda$  is well-defined. The short proof considers a limit of *localised* reflected SDEs, whose existence is given by the following theorem, which assumes a *global* Lipschitz and boundedness condition on the coefficients of the reflected SDE.

**Theorem 5.6.** (*Revuz and Yor, 1991, §IX 2.14*). *Let  $\sigma(s, x)$  and  $b(s, x)$  be functions  $\mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , and  $W$  a Brownian motion. For  $z_0 \geq 0$ , we call a solution to the SDE with reflection  $e_{z_0}(\sigma, b)$  a pair  $(Z, K)$  of processes such that*

(1) the process  $Z$  is continuous, positive,  $\mathcal{F}^W$ -adapted, and

$$Z(t) = z_0 + \int_0^t \sigma(s, Z(s))dW(s) + \int_0^t b(s, Z(s))ds + K(t), \tag{5.13}$$

(2) the process  $K$  is continuous, non-decreasing, vanishing at zero,  $\mathcal{F}^W$ -adapted, and

$$\int_0^\infty Z(s)dK(s) = 0. \tag{5.14}$$

If  $\sigma$  and  $b$  are bounded and satisfy the global Lipschitz condition

$$|\sigma(s, x) - \sigma(s, y)| + |b(s, x) - b(s, y)| \leq C|x - y|, \tag{5.15}$$

for every  $s, x, y \in (0, \infty)$  and some constant  $C$ , then there exists a solution to  $e_{z_0}(\sigma, b)$ , and furthermore this solution is unique.

5.4.1. *Proof of Proposition 1.5.* We now return to the existence of  $Z^\lambda$  as in (1.5), for fixed  $\lambda \in \mathbb{R}$ . In this setting  $\sigma(s, x) \equiv 1$ , but

$$b(s, x) := \lambda - s - \alpha(x, \lambda - s), \tag{5.16}$$

is neither bounded below nor satisfies the global Lipschitz property. However, by Proposition 5.5, for any  $R > 0$ , we can define  $b^R(s, x)$  such that  $b^R(s, x)$  is bounded and globally Lipschitz in  $x$ ; and  $b^R(s, x) = b(s, x)$  whenever  $(s, x) \in [0, R] \times [0, R]$ . Then Theorem 5.6 asserts that there is a unique pair of processes  $(Z^{\lambda, R}, K^{\lambda, R})$  corresponding to this drift, where  $Z^{\lambda, R}(0) = 0$ .

Let  $\tau^{\lambda, R}$  be the time at which  $Z^{\lambda, R}$  first hits  $R$ . Take  $R' \geq R$ . Then, it is clear that  $Z^{\lambda, R}$  is equal to  $Z^{\lambda, R'}$  up to time  $R \wedge \tau^{\lambda, R}$  almost surely. Also, since  $b(s, x)$  is bounded above by  $\lambda$ , it follows that  $\tau^{\lambda, R} \rightarrow \infty$  as  $R \rightarrow \infty$  almost surely. Therefore, we may define

$$Z^\lambda(t) = \lim_{R \rightarrow \infty} Z^{\lambda, R}(t),$$

for almost all paths of  $W$ , and  $Z^\lambda$ . It is immediate that  $Z^\lambda$  satisfies (1.5). Furthermore, any solution  $(Z^\lambda, K^\lambda)$  to (1.5) must coincide with  $(Z^{\lambda, R}, K^{\lambda, R})$  up to  $\tau^{\lambda, R}$ , and so uniqueness of  $(Z^\lambda, K^\lambda)$  follows as well, as required for Proposition 1.5.

5.5. *Convergence of non-negative Markov processes.* It remains to show that Theorem 1.9 follows from Proposition 1.10 as claimed.

A general framework for showing convergence of Markov processes to the solutions of SDEs was introduced by Stroock and Varadhan in the 60s (see, for example, Stroock and Varadhan, 2006). The convergence of Markov processes to reflected diffusions is treated by Stroock and Varadhan (1971) in high generality, allowing for general boundaries in  $\mathbb{R}^d$ , and inhomogeneous stickiness at the boundaries.

We assume that a sequence of Markov chains  $Z^N$ ,  $N \in \mathbb{N}$  is given, where  $Z^N$  has discrete state space  $\mathcal{S}^N \subseteq \mathbb{R}_{\geq 0}$ , with  $0 \in \mathcal{S}^N$ , and initial condition  $Z^N(0) = 0$ . We define the time-inhomogeneous transition operator  $\pi^N$  as

$$\pi_n^N(x, y) = \mathbb{P}(Z^N(n + 1) = y \mid Z^N(n) = x), \quad n \in \mathbb{N}, x, y \in \mathcal{S}^N.$$

We consider a time-rescaling  $(h(N))_{N \in \mathbb{N}}$  for which  $h(N) \rightarrow 0$  as  $N \rightarrow \infty$ .

**Remark.** In our specific example, we have  $\mathcal{S}^N = N^{-1/3}\mathbb{Z}_{\geq 0}$ , and  $h(N) = N^{-2/3}$ .

Then, for every  $x \in \mathcal{S}^N$ , we define the following rescaling transition quantities corresponding to drift, diffusivity, and macroscopic jump probabilities, respectively,

$$\begin{aligned}
 b^N(t, x) &= \frac{1}{h(N)} \sum_{y \in \mathcal{S}^N} (y - x) \pi_{\lfloor t/h(N) \rfloor}^N(x, y), \\
 a^N(t, x) &= \frac{1}{h(N)} \sum_{y \in \mathcal{S}^N} (y - x)^2 \pi_{\lfloor t/h(N) \rfloor}^N(x, y), \\
 \Delta_\epsilon &= \frac{1}{h(N)} \sum_{\substack{y \in \mathcal{S}^N \\ |y-x| > \epsilon}} \pi_{\lfloor t/h(N) \rfloor}^N(x, y).
 \end{aligned}$$

The following theorem, which is a special case of Theorem 6.3 from [Stroock and Varadhan \(1971\)](#), gives conditions under which time-rescaled versions of  $Z^N$  converge to SDEs with reflection.

**Theorem 5.7.** *Suppose we have that for any  $T, M > 0$ , and any  $\epsilon > 0$ ,*

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \sup_{\substack{x \in \mathcal{S}^N \\ x \leq M}} \Delta_\epsilon^N(t, x) = 0, \quad \liminf_{N \rightarrow \infty} \inf_{t \in [0, T]} a^N(t, 0) > 0$$

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \sup_{\substack{x \in \mathcal{S}^N \\ 0 < x \leq M}} |a^N(t, x) - 1| = 0, \quad \lim_{N \rightarrow \infty} \sup_{t \in [0, T]} \sup_{\substack{x \in \mathcal{S}^N \\ 0 < x \leq M}} |b^N(t, x) - b(t, x)| = 0,$$

and that furthermore  $b(\cdot, \cdot)$  satisfies the global Lipschitz condition (5.15) of the previous theorem. Then

$$Z^N \left( \lfloor \frac{t}{h(N)} \rfloor \right)_{t \geq 0} \Rightarrow (Z(t))_{t \geq 0},$$

as  $N \rightarrow \infty$  with respect to the topology of uniform convergence on  $\mathbb{D}[0, T]$  for each  $T < \infty$ , where  $Z$  is the unique solution to  $e_0(1, b)$ , as given by Theorem 5.6.

5.5.1. *Proof of Theorem 1.9.* Now let  $Z^{N,p}$  be the exploration process of  $F(N, p)$ , satisfying the conditions of Theorem 1.8. Again, in our setting, we must account for the fact that the drift of  $Z^\lambda$  is neither bounded nor globally Lipschitz.

Recall from (5.16) and the following paragraph the definitions of  $b(s, x)$  and  $b^R(s, x)$ . For any  $R \in \mathbb{N}$ , we can construct a Markov process  $(Z_n^{N,p,R}, n \geq 0)$  whose transition probabilities coincide with those of  $Z^{N,p}$  whenever  $n \in [0, TN^{2/3}]$  and  $Z_n^{N,p,R} \leq RN^{1/3}$ , and for which, by Proposition 1.10,

$$N^{1/3} \mathbb{E} \left[ Z_{tN^{2/3}+1}^{N,p,R} - Z_{tN^{2/3}}^{N,p,R} \mid Z_{tN^{2/3}}^{N,p,R} = xN^{1/3} \right] \rightarrow b^R(t, x),$$

uniformly for  $t \in [0, T]$  and  $x$  in any compact interval in  $(0, \infty)$ . We define the rescaled process  $\tilde{Z}^{N,\lambda,R}$  from  $Z^{N,p,R}$  analogously to (1.8). Then we have  $\tilde{Z}^{N,p,R} \xrightarrow{d} Z^{\lambda,R}$  uniformly on  $[0, T]$ .

From this,

$$\mathbb{P} \left( \sup_{n \in [0, TN^{2/3}]} Z_n^{N,p,R} > RN^{1/3} \right) \rightarrow 0,$$

as  $R \rightarrow \infty$ , and so as processes on  $[0, T]$ , the law of  $\tilde{Z}^{N,p,R}$  converges to the law of  $\tilde{Z}^{N,p}$  as  $R \rightarrow \infty$ , and the law of  $Z^{\lambda,R}$  converges to the law of  $Z^\lambda$ . Thus we have proved Theorem 1.9.

Combining with the results of Sections 2 and 3, the proof of our main result Theorem 1.8 is now also complete.

**6. Lifting from  $F(N, p)$  to  $F(N, m)$**

So far we have worked in the context of the model  $F(N, p)$  (since in that case the transition probabilities in the exploration process are rather more straightforward to work with than in the case of the model  $F(N, m)$ ). In this section we show that Theorem 1.8 for  $F(N, p)$  implies Theorem 1.7 for  $F(N, m)$ .

As discussed in Section 1.4.2, if we had natural monotonicity properties for the families  $F(N, p)$  and  $F(N, m)$ , then it would be straightforward to deduce Theorem 1.7 from Theorem 1.8 by a sandwiching argument. Instead, we will construct an “almost monotonic” coupling. The idea of Lemma 6.2 below is that, within the scaling window, if the difference between  $m^-$  and  $m^+$  is small compared to  $N^{2/3}$  as  $N \rightarrow \infty$ , then we can couple  $F(N, m^-)$  and  $F(N, m^+)$  so that with high probability, the former is contained in the latter. This coupling is achieved, informally speaking, by adding edges one by one uniformly at random, unless doing so would create a cycle. The next lemma will provide an upper bound on the probability that a cycle does in fact appear.

**Lemma 6.1.** *Let  $H$  be a forest on  $[N]$ , and let  $S^2 = S^2(H)$  be the sum of the squares of the component sizes of  $H$ . Let  $k$  edges, chosen independently and uniformly at random from  $[N] \times [N]$ , be added to  $H$ . (For convenience we allow self-edges and repeated edges). The probability that the resulting graph contains a cycle (including a self-edge or a repeated edge) is at most  $\frac{2kS^2/N^2}{1 - 2kS^2/N^2}$ . In particular if  $k = o(N^2/S^2)$ , then the graph is a forest with high probability as  $N \rightarrow \infty$ .*

*Proof:* Let the components of  $H$  be  $C_1, \dots, C_a$  with sizes  $x_1, \dots, x_a$ .

To create a cycle, for some  $r \geq 1$ , and some distinct  $b_1, b_2, \dots, b_r$ , we have to add an edge between  $C_{b_i}$  and  $C_{b_{i+1}}$  for each  $1 \leq i \leq r - 1$ , and an edge between  $C_{b_r}$  and  $C_{b_1}$ . This creates a cycle containing  $r$  new edges (and also perhaps some further edges which were already part of  $H$ ).

The probability that a given edge has endpoints in  $C_b$  and  $C_{b'}$  is  $2x_b x_{b'}/N^2$  if  $b \neq b'$ , and  $x_b^2/N^2$  if  $b = b'$ , so by a union bound, the probability that at least one of the  $k$  new edges created has endpoints in  $C_b$  and  $C_{b'}$  is at most  $2kx_b x_{b'}/N^2$ . For fixed  $r$  and  $b_1, \dots, b_r$ , a simple conditional probability argument then gives a bound on the probability of creating a collection of edges as specified, of  $(2k/N^2)^r (x_{b_1} x_{b_2}) \dots (x_{b_{r-1}} x_{b_r})(x_{b_r} x_{b_1})$ , which is  $(2k/N^2)^r x_{b_1}^2 \dots x_{b_r}^2$ .

Summing over  $r$  and over distinct  $b_1, \dots, b_r$ , we obtain that the probability of creating a cycle is at most  $\sum_{r=1}^{\infty} \left(\frac{2k}{N^2} S^2\right)^r$ , which gives the claimed bound.  $\square$

We don't know whether  $F(N, m + 1)$  stochastically dominates  $F(N, m)$  in general; that is, whether there is a coupling such that  $F(N, m) \subset F(N, m + 1)$  with probability 1. We get round this by introducing a method to create a coupling which is “monotone with high probability”.

Let  $H_m \sim F(N, m)$ , and consider generating  $\bar{H}_{m+1}$  by adding an edge chosen uniformly at random (from  $[N] \times [N]$ ) to  $H_m$ . Let  $\mathcal{A}$  be the event that  $\bar{H}_{m+1}$  is a forest.

We claim that conditional on  $\mathcal{A}$ , the distribution of  $\bar{H}_{m+1}$  is  $F(N, m + 1)$ . For

$$\begin{aligned} \mathbb{P}(\bar{H}_{m+1} = H') &= \sum_{e \in E(H')} \frac{1}{N^2} \mathbb{P}(H_m = H' \setminus \{e\}) \\ &= \sum_{e \in E(H')} \frac{1}{N^2} \frac{1}{f(N, m)} \\ &= \frac{m + 1}{N^2 f(N, m)}, \end{aligned}$$

which is indeed constant over  $H'$ .

Define also  $\check{H}_{m+1}$  to be distributed according to  $F(N, m + 1)$ , independently from  $H_m$  and the added edge. Now define

$$H_{m+1} = \begin{cases} \bar{H}_{m+1} & \text{on } \mathcal{A} \\ \check{H}_{m+1} & \text{on } \mathcal{A}^c \end{cases}.$$

Then indeed  $H_{m+1} \sim F(N, m + 1)$ , and  $\mathbb{P}(H_m \subset H_{m+1}) \geq \mathbb{P}(\mathcal{A})$ .

We may extend this; starting from  $H_m$ , sequentially add  $k$  edges independently and uniformly, to give graphs  $\bar{H}_{m+1}, \bar{H}_{m+2}, \dots, \bar{H}_{m+k}$ .

Let  $\mathcal{A}_j$  be the event that adding the first  $j$  edges does not create a cycle (including a self-edge or repeated edge).

Let  $\check{H}_{m+1}, \dots, \check{H}_{m+k}$  be independent samples from  $F(N, m + 1), \dots, F(N, m + k)$  respectively, and independent of  $H_m, \bar{H}_{m+1}, \dots, \bar{H}_{m+k}$ .

Now define

$$H_{m+j} = \begin{cases} \bar{H}_{m+j} & \text{on } \mathcal{A}_j \\ \check{H}_{m+j} & \text{on } \mathcal{A}_j^c \end{cases}.$$

Then  $H_{m+j} \sim F(N, m + j)$  for  $j = 0, 1, \dots, k$ , and  $\mathbb{P}(H_m \subset H_{m+1} \subset \dots \subset H_{m+k}) \geq \mathbb{P}(\mathcal{A}_k)$ .

**Lemma 6.2.** *Let  $m = N/2 + O(N^{2/3})$  as  $N \rightarrow \infty$ . Define  $p^-$  and  $p^+$  by  $N^2 p^- / 2 = \lfloor m - N^{3/5} \rfloor$  and  $N^2 p^+ / 2 = \lceil m + N^{3/5} \rceil$ .*

*Then there is a coupling of  $F^- \sim F(N, p^-)$ ,  $F \sim F(N, m)$ , and  $F^+ \sim F(N, p^+)$  such that with high probability as  $N \rightarrow \infty$ ,  $F^- \subseteq F \subseteq F^+$ .*

*Proof:* Let  $M^-$  have the distribution of the number of edges of  $F(N, p^-)$ , and independently let  $M^+$  have the distribution of the number of edges of  $F(N, p^+)$ .

Note that if, conditional on  $M^-$  and  $M^+$ ,  $F^- \sim F(N, M^-)$  and  $F^+ \sim F(N, M^+)$  then the unconditional distributions of  $F^-$  and  $F^+$  are  $F(N, p^-)$  and  $F(N, p^+)$ , respectively.

(2.7) and (4.2) tell us that in this regime, the probability that the number of edges of the graph  $F(N, p)$  deviates from  $N^2 p / 2$  by  $N^{3/5}$  or more goes to 0 as  $N \rightarrow \infty$ .

So with high probability as  $N \rightarrow \infty$ , we have  $M^- \in (m - 2N^{3/5}, m)$  and  $M^+ \in (m, m + 2N^{3/5})$ . If either of these fails, we give up trying to do anything smart and simply set  $F^- \sim F(N, M^-)$ ,  $F \sim F(N, m)$  and  $F^+ \sim F(N, M^+)$  independently.

Otherwise, we have  $M^- < m < M^+$ , and we use the above idea of adding edges sequentially. Throughout the construction below we condition on  $M^-$  and  $M^+$  and regard them as fixed.

Let  $H_{M^-} \sim F(N, M^-)$ , and sequentially add  $K = M^+ - M^- \leq 4N^{3/5}$  edges independently and uniformly, to give graphs  $\bar{H}_{M^-+1}, \bar{H}_{M^-+2}, \dots, \bar{H}_{M^+}$ . As before, let  $\mathcal{A}_j$  be the event that adding the first  $j$  edges does not create a cycle.

Let  $\check{H}_{M^-+1}, \dots, \check{H}_{M^+}$  be independent samples from  $F(N, M^- + 1), \dots, F(N, M^+)$  respectively, and independent of  $H_{M^-}, \bar{H}_{M^-+1}, \dots, \bar{H}_{M^+}$ .

Now for  $1 \leq j \leq K$ , define

$$H_{M^-+j} = \begin{cases} \bar{H}_{M^-+j} & \text{on } \mathcal{A}_j \\ \check{H}_{M^-+j} & \text{on } \mathcal{A}_j^c \end{cases}.$$

Then  $H_{M^-+j} \sim F(N, M^- + j)$  for  $j = 0, 1, \dots, K$ , and  $H_{M^-} \subset H_{M^-+1} \subset \dots \subset H_{M^+}$  whenever  $\mathcal{A}_K$  occurs.

In particular define  $F^- = H_{M^-}$ ,  $F = H_m$  and  $F^+ = H_{M^+}$ . Then (unconditionally),  $F^-$ ,  $F$  and  $F^+$  have the desired marginal distributions, and will be ordered as desired whenever the event  $\mathcal{A}_K$  occurs. So to complete the proof it suffices to show that  $\mathcal{A}_K$  occurs with high probability as  $N \rightarrow \infty$ .

Let  $S^2(H_{M^-})$  be the sum of squares of the component sizes of  $H_{M^-}$ . We know that, averaging over  $M^-$ , the distribution of  $H_{M^-}$  is that of  $F(N, p^-)$ . This is stochastically dominated by  $G(N, p^-)$ , and so Corollary 2.5 tells us that  $\mathbb{E}[S^2(H_{M^-})] \leq \mathbb{E}[S^2(G(N, p^-))] = O(N^{4/3})$ . In particular,  $S^2(H_{M^-}) \leq N^{4/3+\epsilon}$  with high probability as  $N \rightarrow \infty$ , for any  $\epsilon > 0$ .

But the number of edges  $K$  that we add in the sequential construction is at most  $4N^{3/5}$ . So Lemma 6.1 tells us that if indeed  $S^2(H_{M^-}) \leq N^{4/3+\epsilon}$ , then (if  $\epsilon$  is taken sufficiently small) with high probability no cycle is created by adding  $K$  edges to  $H_{M^-}$ . Hence the event  $\mathcal{A}_K$  occurs with high probability as desired.  $\square$

Finally, we can deduce our main scaling limit result for the model  $F(N, m)$ .

*Proof of Theorem 1.7:* If  $m$  has the given asymptotics, and  $p^-$  and  $p^+$  are defined in terms of  $m$  as in Lemma 6.2, then  $p^- = 1/N + (\lambda + o(1))N^{-4/3}$ , and the same is true for  $p^+$ .

From Theorem 1.8, this means that the rescaled component sizes of both  $F(N, p^-)$  and  $F(N, p^+)$  have the limit in distribution given by  $\mathcal{C}^\lambda$ .

But from Lemma 6.2, if the component sizes of  $F(N, p^-)$  and  $F(N, p^+)$  both have this distributional limit, then the same must be true of  $F(N, m)$ , and we are done.  $\square$

### Acknowledgments

We are grateful to Christina Goldschmidt for many valuable discussions during the course of this work. We thank Oliver Riordan for many insightful comments, especially concerning a simplification of Lemma 2.9, Balázs Ráth for a valuable conversation about the form of the diffusion  $Z^\lambda$  at an early stage of the project, and Tom Kurtz for helpful advice about the methods of Section 5.5. We thank the referee for pointing out an oversight in the proof, and for several further helpful comments. The second author was supported by EPSRC doctoral training grant EP/K503113, ISF grant 1325/14, and in part by the Joan and Reginald Coleman-Cohen Fund, and the work was also supported by EPSRC grant EP/J019496/1.

## References

- L. Addario-Berry, N. Broutin and C. Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electron. J. Probab.* **15**, no. 25, 741–775 (2010). [MR2650781](#).
- L. Addario-Berry, N. Broutin and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields* **152** (3-4), 367–406 (2012). [MR2892951](#).
- D. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge Univ. Press, Cambridge (1991). [MR1166406](#).
- D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (2), 812–854 (1997). [MR1434128](#).
- D. Aldous and V. Limic. The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.* **3**, No. 3, 59 pp. (1998). [MR1491528](#).
- M. I. Armendariz. *Brownian excursions and coalescing particle systems*. ProQuest LLC, Ann Arbor, MI (2001). ISBN 978-0493-36160-4. Thesis (Ph.D.)—New York University. [MR2702471](#).
- S. Bhamidi, N. Broutin, S. Sen and X. Wang. Scaling limits of random graph models at criticality: Universality and the basin of attraction of the Erdős-Rényi random graph. *ArXiv Mathematics e-prints* (2014a). [arXiv: 1411.3417](#).
- S. Bhamidi, Amarjit Budhiraja and Xuan Wang. The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probab. Theory Related Fields* **160** (3-4), 733–796 (2014b). [MR3278920](#).
- S. Bhamidi, R. van der Hofstad and J. S. H. van Leeuwaarden. Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.* **15**, no. 54, 1682–1703 (2010). [MR2735378](#).
- S. Bhamidi, R. van der Hofstad and J. S. H. van Leeuwaarden. Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.* **40** (6), 2299–2361 (2012). [MR3050505](#).
- S. Bhamidi, S. Sen and X. Wang. Continuum limit of critical inhomogeneous random graphs. *Probab. Theory Related Fields* **169** (1-2), 565–641 (2017). [MR3704776](#).
- B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** (1), 257–274 (1984). [MR756039](#).
- V. E. Britikov. Asymptotics of the number of forests made up of nonrooted trees. *Mat. Zametki* **43** (5), 672–684, 703 (1988). [MR954351](#).
- N. Broutin and J.-F. Marckert. A new encoding of coalescent processes: applications to the additive and multiplicative cases. *Probab. Theory Related Fields* **166** (1-2), 515–552 (2016). [MR3547745](#).
- S. Dhara, R. van der Hofstad, J. van Leeuwaarden and S. Sen. Heavy-tailed configuration models at criticality. *ArXiv Mathematics e-prints* (2016). [arXiv: 1612.00650](#).
- S. Dhara, R. van der Hofstad, J. van Leeuwaarden and S. Sen. Critical window for the configuration model: finite third moment degrees. *Electron. J. Probab.* **22**, Paper No. 16, 33 (2017). [MR3622886](#).
- T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.* **56**, 13–20 (1960). [MR0115221](#).

- R. van der Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge (2017). ISBN 978-1-107-17287-6. [MR3617364](#).
- S. Janson and J. Spencer. A point process describing the component sizes in the critical window of the random graph evolution. *Combin. Probab. Comput.* **16** (4), 631–658 (2007). [MR2334588](#).
- A. Joseph. The component sizes of a critical random graph with given degree sequence. *Ann. Appl. Probab.* **24** (6), 2560–2594 (2014). [MR3262511](#).
- W. D. Kaigh. An invariance principle for random walk conditioned by a late return to zero. *Ann. Probability* **4** (1), 115–121 (1976). [MR0415706](#).
- M. Kang and T. Łuczak. Two critical periods in the evolution of random planar graphs. *Trans. Amer. Math. Soc.* **364** (8), 4239–4265 (2012). [MR2912453](#).
- J.-F. Le Gall. Random trees and applications. *Probab. Surv.* **2**, 245–311 (2005). [MR2203728](#).
- V. Limic. Eternal multiplicative coalescent is encoded by its Lévy-type processes. *ArXiv Mathematics e-prints* (2016). [arXiv: 1601.01325](#).
- T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures Algorithms* **1** (3), 287–310 (1990). [MR1099794](#).
- T. Łuczak and B. Pittel. Components of random forests. *Combin. Probab. Comput.* **1** (1), 35–52 (1992). [MR1167294](#).
- J. B. Martin and B. Ráth. Rigid representations of the multiplicative coalescent with linear deletion. *Electron. J. Probab.* **22**, Paper No. 83, 47 (2017). [MR3718711](#).
- D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (1991). ISBN 3-540-52167-4. [MR1083357](#).
- O. Riordan. The phase transition in the configuration model. *Combin. Probab. Comput.* **21** (1-2), 265–299 (2012). [MR2900063](#).
- V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439 (1965). [MR0177430](#).
- D. W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24**, 147–225 (1971). [MR0277037](#).
- D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin (2006). ISBN 978-3-540-28998-2; 3-540-28998-4. [MR2190038](#).
- T. S. Turova. Diffusion approximation for the components in critical inhomogeneous random graphs of rank 1. *Random Structures Algorithms* **43** (4), 486–539 (2013). [MR3124693](#).
- V. M. Zolotarev. *One-dimensional stable distributions*, volume 65 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI (1986). ISBN 0-8218-4519-5. [MR854867](#).