

# A PROBABILISTIC APPROACH TO THE LEADER PROBLEM IN RANDOM GRAPHS

LOUIGI ADDARIO-BERRY<sup>1</sup>, SHANKAR BHAMIDI<sup>2</sup>, AND SANCHAYAN SEN<sup>3</sup>

**ABSTRACT.** We study the fixation time of the identity of the leader, i.e., the most massive component, in the general setting of Aldous’s multiplicative coalescent [4, 5], which in an asymptotic sense describes the evolution of the component sizes of a wide array of near-critical coalescent processes, including the classical Erdős-Rényi process.

We show tightness of the fixation time in the “Brownian” regime, explicitly determining the median value of the fixation time to within an optimal  $O(1)$  window. This generalizes Łuczak’s result [31] for the Erdős-Rényi random graph using completely different techniques.

In the heavy-tailed case, in which the limit of the component sizes can be encoded using a thinned pure-jump Lévy process, we prove that only one-sided tightness holds. This shows a genuine difference in the possible behavior in the two regimes.

The solution to the leader problem in the setting of the Erdős-Rényi random graph played an important role in the study of the scaling limit of the minimal spanning tree on the complete graph [2]. We believe that analogous results, such as those proved herein, will be useful in establishing universality of the intrinsic geometry of the minimal spanning tree across a large class of models.

## 1. INTRODUCTION AND MAIN RESULTS

The foundational work of Erdős-Rényi [23, 24] motivated an enormous amount of work on the study of dynamically evolving random graph models in the ensuing decades. We now briefly describe one of the motivating questions of this paper and then discuss renewed interest on this problem over the last few years.

One of the main models studied in the original work of Erdős-Rényi in [23] is the following “random graph process”  $(ER(n, M), M \geq 0)$  on  $[n] := \{1, 2, \dots, n\}$ . Set  $ER(n, 0)$  to be the empty graph. For  $M \geq 1$ ,  $ER(n, M)$  is obtained from  $ER(n, M - 1)$  by choosing one of the  $\binom{n}{2} - M + 1$  edges not present in  $ER(n, M)$  uniformly at random and placing this in the system. Write  $\mathcal{C}_{(k)}(M)$  (respectively  $|\mathcal{C}_{(k)}(M)|$ ) for the  $k$ -th largest component in  $ER(n, M)$  (respectively the size of this component), breaking ties arbitrarily. Here we have suppressed dependence on  $n$  to simplify notation. In [24], the following “double jump” was identified where it was shown that for  $M \ll n/2$ ,  $|\mathcal{C}_{(1)}(M)| = O_P(\log n)$ , if  $M = n/2$  (which corresponds to the so-called **critical regime**), then  $|\mathcal{C}_{(1)}(M)| = \Theta_P(n^{2/3})$ , whilst if  $M = cn/2$  with  $c > 1$ ,

<sup>1</sup>DEPARTMENT OF MATHEMATICS AND STATISTICS, MCGILL UNIVERSITY, MONTREAL, CANADA

<sup>2</sup>DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH, UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL, USA

<sup>3</sup>DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF SCIENCE, BANGALORE, INDIA

*E-mail addresses:* `louigi.addario@mcgill.ca`, `bhamidi@email.unc.edu`, `sanchayan.sen1@gmail.com`

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

*Key words and phrases.* Multiplicative coalescent, critical random graphs, Erdős-Rényi random graph, inhomogeneous random graphs, entrance boundary, Markov processes.

then  $|\mathcal{C}_{(1)}(M)| \sim f(c)n$  for a deterministic function  $f$  satisfying  $f(c) > 0$  for  $c > 1$ . This stimulated an enormous amount of work (see [15, 28, 31, 32] and the references therein) on understanding the behavior close to the critical regime, and the dynamic properties of the above construction wherein components merge via the addition of new edges. This resulted in the following fundamental result of Aldous [4]. Fix  $\lambda \in \mathbb{R}$  and consider the process of normalized component sizes close to the critical value arranged in decreasing order:

$$\bar{\mathbf{C}}_n(\lambda) = \left( n^{-2/3} |\mathcal{C}_{(k)}(n/2 + \lambda n^{2/3})|, k \geq 1 \right).$$

For any  $p \geq 1$ , consider the metric space

$$l_{\downarrow}^p := \left\{ \mathbf{x} = (x_i : i \geq 1) : x_1 \geq x_2 \geq \dots \geq 0, \sum_{i=1}^{\infty} x_i^p < \infty \right\} \quad (1.1)$$

equipped with the natural metric inherited from  $l^p$ .

**Theorem 1.1** ([4]). *View the process  $\{\bar{\mathbf{C}}_n(\lambda) : -\infty < \lambda < \infty\}$  as a Markov process on  $l_{\downarrow}^2$ . Then as  $n \rightarrow \infty$ , the finite dimensional distributions of the above process converges weakly to that of a Markov process on  $l_{\downarrow}^2$  which is referred to as the standard multiplicative coalescent.*

We will describe this result (as well as the entrance boundary of the Markov process) in more detail in Section 3.4; much more extensive discussions of this process and the relationship to the evolution of the Erdős-Rényi random graph can be found in [4]. We are now in a position to state the main problem motivating this paper.

**Leader problem:** Erdős suggested that one should view the original random graph process  $(\text{ER}(n, M), M \geq 0)$  as a “race of components.” One fascinating aspect of this view was studied in [31]. First we need some definitions. For a graph  $G$ , we call any connected component of  $G$  with the maximum number of vertices a leader. Now consider the Erdős-Rényi process  $(\text{ER}(n, M), M \geq 0)$ . For  $0 \leq M_1 < M_2$ , we say that a *change of leader does not occur* in the interval  $[M_1, M_2]$  if there exists a leader  $\mathcal{C}$  in  $\text{ER}(n, M_1)$  such that for all  $M \in [M_1, M_2]$ , the component in  $\text{ER}(n, M)$  containing  $\mathcal{C}$  is a leader in  $\text{ER}(n, M)$ . Define

$$L_{\text{er}}^{(n)} := \min \{s \geq 0 : \text{a change of leader does not occur in the process } (\text{ER}(n, M), M \geq s)\}. \quad (1.2)$$

Then Łuczak in [31, Theorem 7] showed that

$$\text{the sequence of random variables } \{n^{-2/3}(L_{\text{er}}^{(n)} - n/2)\}_{n \geq 1} \text{ is tight.} \quad (1.3)$$

**Aim of this paper:** The original proof in [31] is highly intricate using careful and refined combinatorial analysis of the number of components of various complexities including trees of various sizes, coupled with a “symmetry rule” relating properties of the process below and above the critical threshold. These estimates are combined with a “scanning method” to prove (1.3). This paper is motivated by the following two threads:

(1) In the last few years, a host of random graph models have been shown to belong to the Erdős-Rényi or more precisely, the multiplicative coalescent universality class [6, 7, 9, 10, 12, 13, 20, 25, 29, 35]. This includes the configuration model [14, 34], a large sub-class of the inhomogeneous random graph models as formulated in [16], and the so-called bounded size rules [36]. It is hard to generalize Łuczak’s result in (1.2) to these random graph models via the beautiful counting arguments in [31]. Thus the aim of this paper is to give a short probabilistic proof of the above result that is robust and applies to the general setting of

the multiplicative coalescent. The classical Erdős-Rényi case can be recovered from our result; moreover the techniques in this paper apply to a number of other entrance boundary conditions for the multiplicative coalescent that have arisen in the study of *heavy-tailed* critical random graphs [13, 20, 29].

(2) Coupled with renewed interest in the critical regime, the last few years have also witnessed an explosion in the study of various models of information propagation on network models. In this context one major model that has been explored in great detail is the minimal spanning tree (MST) problem [18, 19]. Here one typically starts with a network model in the supercritical regime (having a giant component). Each edge is assumed to have a random positive length sampled in an *i.i.d.* fashion across edges from a continuous distribution on  $(0, \infty)$ . The aim then is to understand the (metric) structure of the MST on the giant component, e.g., the typical distance between points on the MST. To date the only rigorous results in this context are those in [2, 3]. In [2], the following was shown: Consider the MST  $\mathcal{M}_n$  on the complete graph on  $[n]$  and view this as a tree with edge length one. Rescale each edge of  $\mathcal{M}_n$  by  $n^{-1/3}$  and view this a compact metric space  $\tilde{\mathcal{M}}_n$ . Then as  $n \rightarrow \infty$ ,  $\tilde{\mathcal{M}}_n$  converges in the Gromov-Hausdorff sense to a limiting random compact metric space  $\mathcal{M}_{(\infty)}$ .

A key ingredient in the proof of the main result in [2] is the following theorem proved in [1]: Consider the Erdős-Rényi random graph in the critical regime as in the setting of Theorem 1.1. Fix  $\lambda \in \mathbb{R}$  and  $k \geq 1$ . Consider  $\mathcal{C}_{(k)}(n/2 + \lambda n^{2/3})$  as a metric space where each edge has length one. Write  $\tilde{\mathcal{C}}_{(k)}(\lambda)$  for the resulting metric space where each edge is rescaled by  $n^{-1/3}$ . Then there exist limiting random compact metric spaces  $\mathbf{M}_{(k)}$  such that  $\tilde{\mathcal{C}}_{(k)}(\lambda)$  converge in the Gromov-Hausdorff sense to  $\mathbf{M}_{(k)}$ .

The  $n^{-1/3}$  scaling in the results of both [1] and [2] is not a coincidence. A key step in the proof in [2] is showing that the MST  $\tilde{\mathcal{M}}_n$  is close (in a strong sense) to the restriction of  $\tilde{\mathcal{M}}_n$  to the maximal component in the critical regime  $\tilde{\mathcal{C}}_{(1)}(\lambda)$  “for a large  $\lambda$ .” The proof of this statement relies on the leader result of Łuczak [31] implying that for large  $\lambda$ , the identity of the maximal component does **not** change.

It is believed that under some very general assumptions on the underlying discrete structure, the scaling limit of the minimal spanning tree exists and is universal up to some constants. The problem of establishing universality of the scaling limit of the minimal spanning tree is currently open. One way to approach the problem of universality that has proven to be useful [6, 8, 10, 11] would be to first obtain the scaling limit of the MST constructed on an inhomogeneous random graph that is closely related to the multiplicative coalescent, and then extending it to more general random graph models (including those with heavy tailed degrees) by suitable coupling methods. Carrying out this program requires the extension of the leader result to the setting of the multiplicative coalescent, which is accomplished in this paper.

**1.1. The multiplicative coalescent.** The *multiplicative coalescent* is a Markov process with state space  $l_{\downarrow}^2$ . For a state  $\mathbf{x} = (x_i, i \geq 1) \in l_{\downarrow}^2$ ,  $x_i$  represents the weight or mass of the  $i$ -th largest cluster. In the evolution of the multiplicative coalescent, each pair  $\{i, j\}$  of clusters merges at rate  $x_i x_j$ ; such a merger results in the state obtained by removing  $x_i$  and  $x_j$  from  $\mathbf{x}$  and adding a new entry  $x_i + x_j$  inserted at the location corresponding to its rank. We will

write

$$(\mathbf{x}(t), t \geq 0) = ((x_1(t), x_2(t), \dots), t \geq 0)$$

for the multiplicative coalescent process started from  $\mathbf{x}(0) = \mathbf{x}$ ; thus  $x_1(t) \geq x_2(t) \geq \dots \geq 0$  for all  $t \geq 0$ .

The following is a specific construction of the multiplicative coalescent with initial weights  $\mathbf{x}$ . Define a random graph process  $(G(\mathbf{x}, t), t \geq 0)$  as follows. For  $t \geq 0$ , the vertices of  $G(\mathbf{x}, t)$  are the positive integers  $\mathbb{N}$ .  $G(\mathbf{x}, 0)$  is the empty graph on  $\mathbb{N}$ , and for each  $i, j \in \mathbb{N}$ , edges between  $i$  and  $j$  arrive according to a Poisson point process with rate  $x_i x_j$ . In this process, two distinct connected components with masses  $a$  and  $b$  merge at rate  $ab$ , where mass of a component  $\mathcal{C}$  is  $\sum_{i \in \mathcal{C}} x_i$ . Thus, we may couple the processes  $G(\mathbf{x}, \cdot)$  and  $\mathbf{x}(\cdot)$  so that for all  $t \geq 0$ ,  $\mathbf{x}(t) = (x_i(t), i \geq 1)$  is the ordered sequence of masses of connected components of  $G(\mathbf{x}, t)$ . We work with this construction throughout.

For  $t \geq 0$ , any component in  $G(\mathbf{x}, t)$  with mass  $x_1(t)$  is called a leader in  $G(\mathbf{x}, t)$ . For  $0 \leq t_1 < t_2$  with  $t_2 \in \mathbb{R}_{>0} \cup \{\infty\}$ , we say that a *change of leader does not occur* in  $(G(\mathbf{x}, t), t \in [t_1, t_2])$  if there exists a leader  $\mathcal{C}$  in  $G(\mathbf{x}, t_1)$  such that for all  $t \in (t_1, t_2)$ , the component in  $G(\mathbf{x}, t)$  that contains  $\mathcal{C}$  is a leader in  $G(\mathbf{x}, t)$ . For  $k \geq 1$ , we say that the leader changes at least  $k$  times in  $[t_1, t_2)$  if there exist  $t_1 = u_1 < u_2 < \dots < u_{k+1} = t_2$  such that a change of leader occurs in  $[u_j, u_{j+1})$  for  $j = 1, \dots, k$ . We will simply write “a change of leader does (not) occur in  $G(\mathbf{x}, \cdot)$ ” to mean “a change of leader does (not) occur in  $(G(\mathbf{x}, t), t \geq 0)$ ” etc. Our results address the behavior of the random times:

$$L(\mathbf{x}) = \inf \{s \geq 0 : \text{a change of leader does not occur in } (G(\mathbf{x}, t), t \geq s)\}. \quad (1.4)$$

For any  $\mathbf{z} = (z_1, z_2, \dots) \in [0, \infty)^\mathbb{N}$  and  $r \geq 1$ , we will write

$$\sigma_r(\mathbf{z}) := \sum_{i \geq 1} z_i^r \in [0, \infty]. \quad (1.5)$$

In this paper, we will work with a sequence  $(\mathbf{x}^{(n)}, n \geq 1)$  of starting configurations, where  $\mathbf{x}^{(n)} = (x_i^{(n)}, i \geq 1) \in l_1^2$  with  $x_i^{(n)} > 0$  for only finitely many integers  $i$ . We assume throughout that there exists  $\mathbf{c} = (c_i, i \geq 1)$  such that for all  $i \geq 1$ ,

$$\frac{x_i^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \longrightarrow c_i \text{ as } n \rightarrow \infty. \quad (1.6)$$

We consider two conditions that correspond to two different regimes.

**Condition I (Brownian limit).** The sequence  $\mathbf{c}$  has  $c_i = 0$  for all  $i \geq 1$ , and

$$\frac{\sigma_3(\mathbf{x}^{(n)})}{\sigma_2(\mathbf{x}^{(n)})^3} \rightarrow 1 \text{ as } n \rightarrow \infty.$$

**Condition II (Pure jump limit).** The sequence  $\mathbf{c}$  is such that

$$K i^{-\alpha} \leq c_i \leq K' i^{-\alpha} \text{ for some } \alpha \in (1/3, 1/2), \quad (1.7)$$

where  $0 < K \leq K'$  are universal constants, and

$$\frac{\sigma_3(\mathbf{x}^{(n)})}{\sigma_2(\mathbf{x}^{(n)})^3} \rightarrow \sigma_3(\mathbf{c}) \text{ as } n \rightarrow \infty. \quad (1.8)$$

It is easy to show that under both Condition I and Condition II,

$$\sigma_2(\mathbf{x}^{(n)}) \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (1.9)$$

Indeed, since  $\sigma_3(\mathbf{x}^{(n)}) \leq x_1^{(n)} \sigma_2(\mathbf{x}^{(n)})$ , under Condition I we have

$$\sigma_2(\mathbf{x}^{(n)})^2 = (1 + o(1)) \frac{\sigma_3(\mathbf{x}^{(n)})}{\sigma_2(\mathbf{x}^{(n)})} \leq (1 + o(1)) x_1^{(n)} = o(1) \cdot \sigma_2(\mathbf{x}^{(n)}),$$

which implies (1.9). Under Condition II, for any  $M \in \mathbb{N}$ , by (1.6) we have

$$\sigma_2(\mathbf{x}^{(n)})^2 = (1 + o(1)) \frac{\sum_{i \leq M} (x_i^{(n)})^2}{\sum_{i \leq M} c_i^2} \leq (1 + o(1)) \frac{\sigma_2(\mathbf{x}^{(n)})}{\sum_{i \leq M} c_i^2}.$$

By assumption,  $\sigma_2(\mathbf{c}) = \infty$ , so the above bound again yields (1.9). The fact that  $\sigma_2(\mathbf{x}^{(n)}) \rightarrow 0$  will be used repeatedly throughout the paper.

**Remark 1.** In terms of random graphs, Condition I corresponds to the mean-field percolation regime and Condition II corresponds to the heavy-tailed regime. More precisely, the component sizes of barely-subcritical random graphs whose degree distribution obeys a power law with exponent  $\tau = (\alpha + 1)/\alpha \in (3, 4)$  satisfies (1.6) with  $\mathbf{c}$  as in Condition II; see [13, 21, 25, 29]. In that setting, in a graph with  $n$  nodes, the  $i$ -th largest degree will typically be of order  $(n/i)^\alpha$ . Limiting vectors  $\mathbf{c}$  of the form (1.7) arise in the description of the scaling limits of such random graphs at criticality.

## 2. MAIN RESULTS

We state our main results in this section.

**Theorem 2.1.** *Under Condition I, the collection of random variables*

$$\left\{ L(\mathbf{x}^{(n)}) - \frac{1}{\sigma_2(\mathbf{x}^{(n)})} : n \geq 1 \right\}$$

*is tight.*

This result says that the leader changes for the last time around time  $t = 1/\sigma_2(\mathbf{x}^{(n)})$  with  $O(1)$  fluctuations. One can recover (1.3) from Theorem 2.1. We will elaborate on this in Remark 4.

Under Condition II, we show that the leader does not change much after time  $1/\sigma_2(\mathbf{x}^{(n)})$ .

**Theorem 2.2.** *Under Condition II, the collection of random variables*

$$\left\{ \left( L(\mathbf{x}^{(n)}) - \frac{1}{\sigma_2(\mathbf{x}^{(n)})} \right)^+ : n \geq 1 \right\}$$

*is tight.*

There is a genuine difference in the possible behavior in the two regimes as shown in the next theorem.

**Theorem 2.3.** *Under Condition II, whenever  $c_1 > c_2$ ,*

$$\lim_{\lambda \rightarrow \infty} \liminf_{n \rightarrow \infty} \mathbb{P} \left( L(\mathbf{x}^{(n)}) \leq -\lambda + 1/\sigma_2(\mathbf{x}^{(n)}) \right) > 0.$$

**Remark 2.** It would be interesting to see if the following stronger result holds: Under Condition II,  $\liminf_{n \rightarrow \infty} \mathbb{P}(L(\mathbf{x}^{(n)}) = 0) > 0$  if  $c_1 > c_2$ . If this is not true, then the natural question would be the following: What is a threshold function  $a_n = a_n(\mathbf{x}^{(n)}) \rightarrow \infty$  such that under Condition II, if  $c_1 > c_2$ , then

$$\begin{aligned} \liminf_{n \rightarrow \infty} \mathbb{P}(L(\mathbf{x}^{(n)}) \leq -\lambda_n + 1/\sigma_2(\mathbf{x}^{(n)})) &> 0, \quad \text{if } \lambda_n \ll a_n, \quad \text{and} \\ \lim_{n \rightarrow \infty} \mathbb{P}(L(\mathbf{x}^{(n)}) \leq -\lambda_n + 1/\sigma_2(\mathbf{x}^{(n)})) &= 0, \quad \text{if } \lambda_n \gg a_n. \end{aligned}$$

In contrast, under Condition I, the following result holds.

**Theorem 2.4.** *Let  $N^{(n)}$  denote the number of times the leader changes in  $[0, (\sigma_2(\mathbf{x}^{(n)}))^{-1}]$ . Under Condition I,  $N^{(n)} \xrightarrow{\mathbb{P}} \infty$ .*

It is known [4, 5] that under both Condition I and Condition II, there exists an  $l_{\downarrow}^2$ -valued process  $(\zeta(\lambda), \lambda \in \mathbb{R})$  (called an eternal multiplicative coalescent) such that for every fixed  $\lambda \in \mathbb{R}$ ,  $\mathbf{x}^{(n)}(t_{\lambda}) \xrightarrow{d} \zeta(\lambda)$  as  $n \rightarrow \infty$ , where  $t_{\lambda} = t_{\lambda}(\mathbf{x}^{(n)}) = \lambda + 1/\sigma_2(\mathbf{x}^{(n)})$ ; this is discussed in more detail in Section 3.4. (Theorem 1.1 is a special case of this result.) The arguments in the proofs of Theorems 2.1, 2.2, 2.3, and 2.4 carry over in an identical way to the continuum yielding analogous results for  $\zeta(\cdot)$ . We collect these results in the next theorem; we omit the proof.

**Theorem 2.5.** (a) *If  $c_i = 0$  for all  $i \geq 1$ , or if  $\mathbf{c}$  satisfies (1.7), then*

$$\lim_{\lambda \rightarrow \infty} \mathbb{P}\left(\text{a change of leader occurs in } G(\zeta(\lambda), \cdot)\right) = 0.$$

(b) *If  $c_i = 0$  for all  $i \geq 1$ , then for all  $k \geq 1$ ,*

$$\lim_{\lambda \rightarrow \infty} \mathbb{P}\left(\text{the leader changes at least } k \text{ times in } \left(G(\zeta(-\lambda), t), t \in [0, \lambda]\right)\right) = 1.$$

(c) *If  $\mathbf{c}$  satisfies (1.7), then*

$$\lim_{\lambda \rightarrow \infty} \mathbb{P}\left(\text{a change of leader does not occur in } G(\zeta(-\lambda), \cdot)\right) > 0.$$

Now, Theorems 2.1 and 2.2 state results about tightness. A natural question about convergence in distribution arises here. In this context, we make the following conjecture.

**Conjecture 2.6.** *Under both Condition I and Condition II, for every  $\lambda \in \mathbb{R}$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(L(\mathbf{x}^{(n)}) > (\sigma_2(\mathbf{x}^{(n)}))^{-1} + \lambda\right) = \mathbb{P}\left(\text{a change of leader occurs in } G(\zeta(\lambda), \cdot)\right).$$

**Remark 3.** In [31], Łuczak studied the size of the leader in the setting of the Erdős-Rényi random graph. Informally, his results (in particular [31, Theorem 5]) give a law of large numbers for the mass of the largest component at time  $1/\sigma_2 + \lambda$ , as  $\lambda \rightarrow \infty$ , in the Erdős-Rényi setting. Łuczak's results in particular imply that under Condition I, the mass of the leader at time  $1/\sigma_2 + \lambda$  grows linearly in  $\lambda$ . Our Theorem 6.1 stated below, provides an analogous result under Condition II. Under Condition II, the growth rate is not  $\lambda$  but  $\lambda^{(1-2\alpha)/\alpha}$ . Related results for the barely supercritical configuration model with heavy-tailed degrees were proved in [26].

**Remark 4.** Let us briefly describe how to prove (1.3) using Theorem 2.1. Consider the modified Erdős-Rényi process  $(\overline{\text{ER}}(n, M), M \geq 0)$ , where at each step, edges are sampled uniformly and with replacement, i.e., multiple edges are allowed. A simple computation shows that the number of multiple edges created up to  $M \leq n$  is tight. Thus, it is enough to show that  $n^{-2/3}(\overline{L}_{\text{er}}^{(n)} - n/2)$  is tight, where  $\overline{L}_{\text{er}}^{(n)}$  denotes the minimum of all  $s \geq 0$  such that a change of leader does not occur in  $(\overline{\text{ER}}(n, M), M \geq s)$ .

Consider the weight sequence  $\mathbf{y}^{(n)} = (y_i^{(n)}, i \geq 1)$ , where

$$y_i^{(n)} = \begin{cases} n^{-2/3}, & \text{if } i \leq n \\ 0, & \text{if } i > n. \end{cases} \quad (2.1)$$

Then  $\mathbf{y}^{(n)}$  satisfies Condition I and  $\sigma_2(\mathbf{y}^{(n)}) = n^{-1/3}$ . Consider the random graph process  $(G(\mathbf{y}^{(n)}, t), t \geq 0)$ , and for  $k \geq 1$ , let  $\tau_k$  be the time when the  $k$ -th edge is added in this process. Clearly  $(\overline{\text{ER}}(n, M), M \geq 1)$  has the same distribution as the process  $(G(\mathbf{y}^{(n)}, \tau_M), M \geq 1)$  restricted to the vertex set  $\{1, \dots, n\}$ . For  $u \in \mathbb{R}$ , define  $M_n(u) = \lfloor n/2 + un^{2/3} \rfloor$ . (Note that for any  $u$ ,  $M_n(u) > 0$  for all large  $n$ .) Then,

$$\mathbb{P}(\overline{L}_{\text{er}}^{(n)} \geq M_n(u)) = \mathbb{P}(L(\mathbf{y}^{(n)}) \geq \tau_{M_n(u)}). \quad (2.2)$$

Now, edges appear at a rate of  $\lambda_n := n^{-4/3} \binom{n}{2}$  in the process  $G(\mathbf{y}^{(n)}, \cdot)$ . Hence, for any fixed  $u \in \mathbb{R}$ ,

$$\tau_{M_n(u)} = \lambda_n^{-1} M_n(u) + O_P(\lambda_n^{-1} \sqrt{M_n(u)}) = n^{1/3} + 2u + O_P(n^{-1/6}) = [\sigma_2(\mathbf{y}^{(n)})]^{-1} + 2u + O_P(n^{-1/6}).$$

(Here  $O_P(a_n)$  represents a sequence  $(Y_n, n \geq 1)$  of random variables such that  $(a_n^{-1} Y_n, n \geq 1)$  is tight.) So (2.2) together with Theorem 2.1 shows that  $n^{-2/3}(\overline{L}_{\text{er}}^{(n)} - n/2)$  is tight.

### 3. PRELIMINARIES

**3.1. Notation.** Throughout this paper, we make use of the following notation. We will write  $\xrightarrow{d}$ ,  $\xrightarrow{P}$ , and  $\xrightarrow{\text{a.s.}}$  to denote convergence in distribution, convergence in probability, and almost sure convergence respectively. For a sequence of random variables  $(X_n, n \geq 1)$ , we write  $X_n = o_P(b_n)$  when  $|X_n|/b_n \xrightarrow{P} 0$  as  $n \rightarrow \infty$ . For a non-negative function  $n \rightarrow g(n)$ , we write  $f(n) = O(g(n))$  when  $|f(n)|/g(n)$  is uniformly bounded, and  $f(n) = o(g(n))$  when  $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$ . Furthermore, we write  $f(n) = \Theta(g(n))$  if  $f(n) = O(g(n))$  and  $g(n) = O(f(n))$ .

Throughout this paper,  $C, C', K, K'$  will denote positive constants that depend only on  $\mathbf{c} = (c_i, i \geq 1)$ , and their values may change from line to line. Special constants will be indexed, e.g.,  $K_{6.4}$  etc. Given two functions  $f_1, f_2 : [0, \infty) \rightarrow [0, \infty)$ , we write  $f_1 \asymp f_2$  on  $A \subseteq [0, \infty)$  if there exist  $0 < K' \leq K < \infty$  such that

$$K' f_2(x) \leq f_1(x) \leq K f_2(x) \text{ for all } x \in A.$$

Similarly, for two sequences  $\{a_n\}_{n \geq 1}$  and  $\{b_n\}_{n \geq 1}$ , we will write  $\{a_n\} \asymp \{b_n\}$  or simply  $a_n \asymp b_n$  to mean that there exist  $0 < K' \leq K < \infty$  such that

$$K' b_n \leq a_n \leq K b_n \text{ for all } n \geq 1.$$

**3.2. Tail bounds for sums and suprema.** We will use the following bound on the suprema of collections of observables of empirical processes, which is a specialization of [30, Theorem 1.1 (b)](see also [17, Theorem 12.9]).

**Theorem 3.1** ([30]). *Let  $X_i$ ,  $i \geq 1$ , be independent real random variables. Fix  $A > 0$  and  $n \geq 1$ , and let  $\{f_{i,s} : 1 \leq i \leq n, s \in S\}$  be a countable collection of measurable functions from  $\mathbb{R}$  to  $[-A, A]$ . Write*

$$Z = \sup \left\{ \sum_{i=1}^n f_{i,s}(X_i) : s \in S \right\},$$

and let

$$V = 2\mathbb{E}[Z] + \sup_{s \in S} \text{Var} \left( \sum_{i \leq n} f_{i,s}(X_i) \right).$$

Then

$$\mathbb{P} \left( Z \geq \mathbb{E}Z + x \right) \leq \exp \left( -\frac{x}{4A} \log \left( 1 + 2 \log \left( 1 + \frac{x}{AV} \right) \right) \right).$$

**3.3. Random graph constructions.** We work with the construction of the multiplicative coalescent introduced in Section 1.1. Recall that in this construction,  $\mathbf{x}(t) = (x_i(t), i \geq 1)$  is the ordered sequence of masses of the connected components of random graph  $G(\mathbf{x}, t)$  defined in Section 1.1. We always assume

$$\mathbf{x} = (x_i, i \geq 1) = (x_i(0), i \geq 1) = \mathbf{x}(0) \in l_1^2.$$

In this case,  $(\mathbf{x}(t), t \geq 0)$  is a Markov process with state space  $l_1^2$ . Its generator  $\mathcal{A}$  is given by

$$\mathcal{A}f(\mathbf{z}) = \sum_i \sum_{j>i} z_i z_j (f(\mathbf{z}^{i,j}) - f(\mathbf{z})), \quad (3.1)$$

where  $\mathbf{z}^{i,j} \in l_1^2$  is formed from  $\mathbf{z}$  by removing entries  $z_i$  and  $z_j$  and inserting a new entry  $z_i + z_j$  in a location which preserves membership in  $l_1^2$ .

Let  $\mathcal{S}$  denote the space of all vertex-weighted multigraphs with vertex set  $\mathbb{N}$  that satisfy

$$\sum_{\mathcal{C} \text{ component of } G} \left( \sum_{i \in \mathcal{C}} w_i(G) \right)^2 < \infty,$$

where  $w_i(G)$  denotes the weight of vertex  $i$  in  $G$ . Then  $(G(\mathbf{x}, t), t \geq 0)$  is a Markov process with state space  $\mathcal{S}$ . Its generator, also denoted  $\mathcal{A}$ , is given by

$$\mathcal{A}f(G) = \sum_i \sum_{j>i} w_i(G) w_j(G) (f(G^{i,j}) - f(G)), \quad (3.2)$$

where  $G^{i,j}$  is formed from  $G$  by adding an edge between vertices  $i$  and  $j$ .

For  $G \in \mathcal{S}$  and any connected component  $\mathcal{C}$  of  $G$ , we write  $\mathcal{W}(\mathcal{C}) = \sum_{i \in \mathcal{C}} w_i(G)$  for the mass of  $\mathcal{C}$ . We write  $\mathcal{C}_i(G)$  for the connected component of  $G$  having the  $i$ -th largest mass. We also write  $\mathcal{C}(G; i)$  for the connected component of  $G$  containing vertex  $i$ . To simplify notation, we write  $\mathcal{C}_i(\mathbf{x}, t)$  and  $\mathcal{C}(\mathbf{x}, t; i)$  for the corresponding objects for  $G(\mathbf{x}, t)$ . Thus,  $x_i(t) = \mathcal{W}(\mathcal{C}_i(\mathbf{x}, t))$  for all  $i \geq 1$  and  $t \geq 0$ , and  $\sigma_p(\mathbf{x}(t)) = \sum_{\mathcal{C} \text{ component of } G(\mathbf{x}, t)} (\mathcal{W}(\mathcal{C}))^p$  for all  $t \geq 0$  and  $p \geq 1$ .



**3.4. Convergence of multiplicative coalescents.** Define a process  $(V(s), s \geq 0)$  as follows. Under Condition I, the limit sequence  $\mathbf{c}$  is identically zero. In this case let

$$V(s) = B(s) - s^2/2,$$

where  $(B(s), s \geq 0)$  is a standard Brownian motion. Under Condition II, let  $\xi_1, \xi_2, \dots$  be independent random variables with  $\xi_i \sim \text{Exp}(c_i)$ , and let

$$V(s) = \sum_{i \geq 1} c_i (\mathbb{1}_{\{\xi_i \leq s\}} - c_i s). \quad (3.3)$$

In both cases, define a process  $(W_\lambda(s), s \geq 0)$  by setting

$$W_\lambda(s) = \lambda s + V(s). \quad (3.4)$$

Let  $\overline{W}_\lambda(\cdot)$  denote the process  $W_\lambda$  reflected at zero, namely,

$$\overline{W}_\lambda(s) = W_\lambda(s) - \inf_{0 \leq u \leq s} W_\lambda(u), \quad s \geq 0. \quad (3.5)$$

Suppose  $\mathbf{x} \in l_1^2$  has finite length, i.e., has only finitely many nonzero entries. A standard tool in the study of the random graph  $G(\mathbf{x}, t)$  is the ‘‘breadth-first walk’’ process  $(B_{\mathbf{x}, t}(u), u \geq 0)$  associated with a breadth-first exploration of the random graph  $G(\mathbf{x}, t)$ . We recall this process from [5] briefly; the reader is referred to [5, Section 2.3] for a more detailed description. Let  $U_{i,j}$ ,  $1 \leq i, j \leq n$ , be independent random variables with  $U_{i,j} \sim \text{Exponential}(tx_j)$ . Choose  $\nu(1)$  by size-biased sampling, i.e., vertex  $\nu$  is chosen with probability proportional to  $x_\nu$ . Define  $\{v : U_{\nu(1), \nu} \leq x_{\nu(1)}\}$  to be the set of children of  $\nu(1)$ , and order these children as  $\nu(2), \nu(3), \dots$  so that  $U_{\nu(1), \nu(i)}$  is increasing. Set  $B_{\mathbf{x}, t}(0) = 0$ , and let

$$B_{\mathbf{x}, t}(u) = -u + \sum_{\nu} x_\nu \mathbb{1}_{\{U_{\nu(1), \nu} \leq u\}}, \quad 0 \leq u \leq x_{\nu(1)}.$$

Inductively, write  $\tau_i = \sum_{j \leq i} x_{\nu(j)}$ . If  $\nu(i)$  is in the same component as  $\nu(1)$ , then the set  $\{v \notin \{\nu(1), \dots, \nu(i-1)\} : v \text{ is a child of one of } \{\nu(1), \dots, \nu(i-1)\}\}$  consists of  $\nu(i), \dots, \nu(l(i))$  for some  $l(i) \geq i$ . Let the children of  $\nu(i)$  be  $\{v \notin \{\nu(1), \dots, \nu(l(i))\} : U_{\nu(i), \nu} \leq x_{\nu(i)}\}$ , and order them as  $\nu(l(i)+1), \nu(l(i)+2), \dots$  such that  $U_{\nu(i), \nu}$  is increasing. Set

$$B_{\mathbf{x}, t}(\tau_{i-1} + u) = B_{\mathbf{x}, t}(\tau_{i-1}) - u + \sum_{\nu \text{ child of } \nu(i)} x_\nu \mathbb{1}_{\{U_{\nu(i), \nu} \leq u\}}, \quad 0 \leq u \leq x_{\nu(i)}. \quad (3.6)$$

Once the exploration of the component of  $\nu(1)$  is complete, choose the next vertex from the set of the remaining vertices by size-biased sampling, and continue. In the end, this produces a random forest. The partition of the set of vertices into different connected components in this forest has the same distribution as the partition of the set of vertices into different connected components in  $G(\mathbf{x}, t)$ . Let  $\xi_j^n$  denote the ‘birth-time’ of  $j$ , i.e., the time when vertex  $j$  appears in the above exploration. Thus, if  $j$  is the first vertex in its component to appear in the above process, then  $\xi_j^n$  is the time when the exploration of that component started. If  $j$  is not the first vertex in its component to appear in the above process and  $j$  is the child of  $\nu(i)$ , then  $\xi_j^n = \tau_{i-1} + U_{\nu(i), j}$ .

Now consider  $\mathbf{x}^{(n)}$ ,  $n \geq 1$ , satisfying either Condition I or Condition II. Now and henceforth, for  $\lambda \in \mathbb{R}$ , we write

$$t_\lambda = t_\lambda(\mathbf{x}^{(n)}) = \lambda + 1/\sigma_2(\mathbf{x}^{(n)}).$$

(Note that (1.9) implies that for every  $\lambda \in \mathbb{R}$ ,  $t_\lambda(\mathbf{x}^{(n)}) \geq 0$  for all  $n$  sufficiently large.) Then under both Conditions I and II on the sequence of starting configurations  $(\mathbf{x}^{(n)}, n \geq 1)$ , by [5, Proposition 9],

$$\left( \frac{B_{\mathbf{x}^{(n)}, t_\lambda}(s)}{\sigma_2(\mathbf{x}^{(n)})}, s \geq 0 \right) \xrightarrow{d} \left( W_\lambda(s), s \geq 0 \right). \quad (3.7)$$

Let us now look at implication of this convergence. For an excursion  $\gamma$  of  $\overline{W}_\lambda$ , write  $|\gamma|$  for the length of  $\gamma$ . Let  $\gamma_i(\lambda)$ ,  $i \geq 1$ , denote the excursions of  $\overline{W}_\lambda$  from zero in decreasing order of length, i.e.,  $|\gamma_1(\lambda)| > |\gamma_2(\lambda)| > \dots$ . Then under both Condition I and Condition II, by [5, Proposition 7] (see also [4, Proposition 4]), for each fixed  $\lambda \in \mathbb{R}$ ,  $\zeta(\lambda) := (|\gamma_i(\lambda)|, i \geq 1) \in l_\downarrow^2$  almost surely, and further,

$$\mathbf{x}^{(n)}(t_\lambda) \xrightarrow{d} \zeta(\lambda), \quad (3.8)$$

with respect to the topology on  $l_\downarrow^2$ .

For any component  $\mathcal{C}$  of  $G(\mathbf{x}^{(n)}, t_\lambda)$ , let  $l_n(\mathcal{C})$  (resp.  $r_n(\mathcal{C})$ ) denote the time when the exploration of  $\mathcal{C}$  started (resp. concluded) in  $B_{\mathbf{x}^{(n)}, t_\lambda}$ . For  $u > 0$ , let  $\mathfrak{C}^{(n)}(\lambda; u)$  denote the component of  $G(\mathbf{x}^{(n)}, t_\lambda)$  being explored by  $B_{\mathbf{x}^{(n)}, t_\lambda}$  at time  $u$ . For an excursion  $\gamma$  of  $\overline{W}_\lambda$ , write  $l(\gamma)$  and  $r(\gamma)$  for the start and finish times of  $\gamma$ .

For the rest of Section 3.4, we work under Condition II. It follows from [5, Proposition 14 (b)] that for every  $u > 0$ ,  $\mathbb{P}(\overline{W}_\lambda(u) = 0) = 0$ . Since  $\overline{W}_\lambda(\cdot)$  is continuous at  $u$  with probability one for every  $u > 0$ , we conclude that for all  $u > 0$ ,

$$\mathbb{P}(\overline{W}_\lambda(u) = 0 \text{ or } \overline{W}_\lambda(u-) = 0) = 0. \quad (3.9)$$

Consequently, for every  $u > 0$ ,

$$\mathbb{P}(l(\gamma) < u < r(\gamma) \text{ for some excursion } \gamma \text{ of } \overline{W}_\lambda) = 1. \quad (3.10)$$

Denote this excursion by  $\gamma(\lambda; u)$ . Now (3.7) and [5, Proposition 14 (d)] imply (see [5, Page 27]) that the following convergence of point processes holds jointly with (3.7):

$$\left\{ (l_n(\mathcal{C}), \mathcal{W}(\mathcal{C})) : \mathcal{C} \text{ component of } G(\mathbf{x}^{(n)}, t_\lambda) \right\} \xrightarrow{d} \left\{ (l(\gamma), |\gamma|) : \gamma \text{ excursion of } \overline{W}_\lambda \right\}, \quad (3.11)$$

the underlying notion of convergence being that of vague convergence of counting measures on  $[0, a] \times [b, c]$  for every  $a > 0$  and  $c > b > 0$ . It follows from (3.11) and (3.10) that for every  $u > 0$ ,

$$\left( l_n(\mathfrak{C}^{(n)}(\lambda; u)), r_n(\mathfrak{C}^{(n)}(\lambda; u)) \right) \xrightarrow{d} \left( l(\gamma(\lambda; u)), r(\gamma(\lambda; u)) \right) \quad (3.12)$$

jointly with the convergence in (3.7).

Recall from (3.3) and (3.4) the definitions of the processes  $V(\cdot)$  and  $W_\lambda(\cdot)$  under Condition II. Fix  $i \geq 1$ . Define

$$W_\lambda^{i-}(s) = W_\lambda(s) - c_i \mathbb{1}_{\{\xi_i \leq s\}} = (\lambda - c_i^2)s + \sum_{\substack{j \geq 1 \\ j \neq i}} c_j (\mathbb{1}_{\{\xi_j \leq s\}} - c_j s), \quad s \geq 0.$$

Let  $\overline{W}_\lambda^{i-}(s) = W_\lambda^{i-}(s) - \inf_{0 \leq u \leq s} W_\lambda^{i-}(u)$ ,  $s \geq 0$ . Since  $\xi_i$  is independent of the process  $W_\lambda^{i-}(\cdot)$ , the argument leading to (3.9) shows that  $\mathbb{P}(\overline{W}_\lambda^{i-}(\xi_i-) = 0) = 0$ . Using the fact that

$W_\lambda^{i-}(s) = W_\lambda(s)$  for  $s \in [0, \xi_i)$ , we see that  $\overline{W}_\lambda(\xi_i-) = \overline{W}_\lambda^{i-}(\xi_i-) > 0$  with probability one. Since  $\overline{W}_\lambda(\xi_i) > 0$ , we get

$$\mathbb{P}(\forall j \geq 1, l(\gamma) < \xi_j < r(\gamma) \text{ for some excursion } \gamma \text{ of } \overline{W}_\lambda) = 1. \quad (3.13)$$

For an excursion  $\gamma$  of  $\overline{W}_\lambda$  and  $j \in \mathbb{N}$ , write  $j \in \gamma$  if  $l(\gamma) < \xi_j < r(\gamma)$ .

Consider  $\varepsilon > 0$  such that  $c_i > \varepsilon > c_{i+1}$  for some  $i \geq 1$ . Then (3.7), (3.12), and (3.13) imply that for any such  $\varepsilon > 0$ ,

$$\sum_j \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \mathbb{1}\{j \in \mathfrak{C}^{(n)}(\lambda; u); x_j^{(n)} > \varepsilon \sigma_2(\mathbf{x}^{(n)})\} \xrightarrow{d} \sum_{j \geq 1} c_j^2 \cdot \mathbb{1}\{j \in \gamma(\lambda; u); c_j > \varepsilon\} \quad (3.14)$$

jointly with the convergence in (3.12). (The convergence in (3.14) can be deduced using the following elementary fact: Suppose  $f_n \rightarrow f$ , as  $n \rightarrow \infty$ , in the Skorohod  $J_1$  topology on  $\mathbb{D}[0, T]$ . Then there exists a sequence  $(\lambda_n, n \geq 1)$  of strictly increasing, continuous functions on  $[0, T]$  satisfying (i)  $\lambda_n(0) = 0$ , and  $\lambda_n(T) = T$  for all  $n \geq 1$ , and (ii)  $\sup_{t \in [0, T]} |\lambda_n(t) - t| + \sup_{t \in [0, T]} |f_n \circ \lambda_n(t) - f(t)| \rightarrow 0$  as  $n \rightarrow \infty$ . In particular,  $f_n \circ \lambda_n(t) - f_n \circ \lambda_n(t-) \rightarrow f(t) - f(t-)$ , as  $n \rightarrow \infty$ , for all  $t \in [0, T]$ .)

Let  $\xi_j^n$  be the birth time of  $j$  in the breadth-first walk process as defined below (3.6). As explained in [5, Page 17], there exist independent random variables  $\tilde{\xi}_j^n, j = 1, \dots, n$ , such that  $\tilde{\xi}_j^n \sim \text{Exponential}(t_\lambda x_j^{(n)})$ , and  $\xi_j^n = \tilde{\xi}_j^n$  if  $j$  is not the first vertex in its component appearing in the breadth-first walk. Consequently,

$$\sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \mathbb{1}\{j \in \mathfrak{C}^{(n)}(\lambda; u)\} \leq \left( \frac{x_k^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 + \sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \mathbb{1}\{\tilde{\xi}_j^n \leq r_n(\mathfrak{C}^{(n)}(\lambda; u))\}. \quad (3.15)$$

Now, for any  $A > 0$ ,

$$\begin{aligned} \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{E} \left[ \sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \mathbb{1}\{\tilde{\xi}_j^n \leq A\} \right] &\leq \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \left[ \sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \cdot A t_\lambda x_j^{(n)} \right] \\ &= \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} A (1 + \lambda \sigma_2(\mathbf{x}^{(n)})) \left[ \sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^3 \right] = \lim_{k \rightarrow \infty} A \sum_{j \geq k} c_j^3 = 0, \end{aligned} \quad (3.16)$$

where the penultimate step uses (1.9), (1.8), and (1.6), and the final step uses the fact  $\mathbf{c} \in l^3$ . Since  $r_n(\mathfrak{C}^{(n)}(\lambda; u)) \leq u + x_1^{(n)}(t_\lambda)$ , it follows from (3.8) that for every  $u > 0$ , the sequence  $(r_n(\mathfrak{C}^{(n)}(\lambda; u)), n \geq 1)$  is tight. Combining this observation with (3.15) and (3.16), we see that for every  $\delta > 0$ ,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \sum_{j \geq k} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \mathbb{1}\{j \in \mathfrak{C}^{(n)}(\lambda; u)\} > \delta \right) = 0. \quad (3.17)$$

We can similarly show that  $\mathbb{P}(\sum_{j \geq k} c_j^2 \cdot \mathbb{1}\{j \in \gamma(\lambda; u)\} > \delta) \rightarrow 0$  as  $k \rightarrow \infty$ , which together with (3.17) and (3.14) yields that for every  $u > 0$  and  $\lambda \in \mathbb{R}$ ,

$$\sum_{j \in \mathcal{C}^{(n)}(\lambda; u)} \left( \frac{x_j^{(n)}}{\sigma_2(\mathbf{x}^{(n)})} \right)^2 \xrightarrow{d} \sum_{j \geq 1} c_j^2 \cdot \mathbb{1}\{j \in \gamma(\lambda; u)\} \quad (3.18)$$

jointly with the convergence in (3.12).

#### 4. PROOF OF THEOREM 2.1

We break up the claim in Theorem 2.1 into two propositions that we will prove separately.

**Proposition 4.1.** *Under Condition I,  $\left\{ \left( L(\mathbf{x}^{(n)}) - \frac{1}{\sigma_2(\mathbf{x}^{(n)})} \right)^+ : n \geq 1 \right\}$  is tight.*

**Proposition 4.2.** *Under Condition I,  $\left\{ \left( L(\mathbf{x}^{(n)}) - \frac{1}{\sigma_2(\mathbf{x}^{(n)})} \right)^- : n \geq 1 \right\}$  is tight.*

The heart of the proof that the family of random variables in Proposition 4.1 is tight is given in the following proposition.

**Proposition 4.3.** *Fix  $\mathbf{y} = (y_i, i \geq 1) \in l_1^2$ . Then for every  $\ell \geq 1$ ,  $(f_\ell(G(\mathbf{y}, t)), t \geq 0)$  is a submartingale, where*

$$f_\ell(G) = \frac{(\mathcal{W}(\mathcal{C}(G; \ell)))^2}{\sum_{\mathcal{C} \text{ component of } G} \mathcal{W}(\mathcal{C})^2}, \quad G \in \mathcal{S}.$$

Consequently, if  $y_1^2 / \sigma_2(\mathbf{y}) > 1 - \varepsilon$ , then

$$\mathbb{P}(\text{a change of leader does not occur in } G(\mathbf{y}, \cdot)) > 1 - 5\varepsilon. \quad (4.1)$$

**Proof:** Fix  $\mathbf{z} \in l_1^2$  such that only finitely many coordinates of  $\mathbf{z}$  are non-zero. Let  $G \in \mathcal{S}$  be such that  $\mathcal{W}(\mathcal{C}_i(G)) = z_i$  for all  $i \geq 1$ . Suppose  $\ell \in \mathcal{C}_k(G)$ . For convenience, we will write  $\sigma_2$  for  $\sigma_2(\mathbf{z})$ . We claim that

$$\mathcal{A} f_\ell(G) \geq z_k \sum_{i \geq 1, i \neq k} \frac{z_i^3}{\sigma_2 + 2z_i z_k}. \quad (4.2)$$

We will prove (4.2) shortly, but let us first note that since the right-hand side is nonnegative, it follows that the process

$$(f_\ell(G(\mathbf{y}^{[n]}, t)), t \geq 0) \text{ is a submartingale,} \quad (4.3)$$

where  $\mathbf{y}^{[n]} = (y_1, \dots, y_n, 0, 0, \dots)$ . To prove the claim for a general  $\mathbf{y} \in l_1^2$ , note that the graphs  $G(\mathbf{y}, t)$  and  $G(\mathbf{y}^{[n]}, t)$  can be coupled in an obvious way so that the latter is a subgraph of the former. By [4, Corollary 18 (c)], in this coupling,

$$\sigma_2(\mathbf{y}^{[n]}(t)) \xrightarrow{\text{a.s.}} \sigma_2(\mathbf{y}(t)) \text{ as } n \rightarrow \infty. \quad (4.4)$$

Now,  $\mathcal{W}(\mathcal{C}(\mathbf{y}, t; \ell)) = \sum_j y_j \cdot \mathbb{1}_{\{j \in \mathcal{C}(\mathbf{y}, t; \ell)\}}$ . Since  $j \in \mathcal{C}(\mathbf{y}, t; \ell)$  if and only if there is a finite path between  $j$  and  $\ell$  in  $G(\mathbf{y}, t)$ ,  $\mathbb{1}_{\{j \in \mathcal{C}(\mathbf{y}, t; \ell)\}}$  is the almost sure increasing limit of  $\mathbb{1}_{\{j \in \mathcal{C}(\mathbf{y}^{[n]}, t; \ell)\}}$  in the above coupling. An application of the monotone convergence theorem implies that

$$\mathcal{W}(\mathcal{C}(\mathbf{y}^{[n]}, t; \ell)) \xrightarrow{\text{a.s.}} \mathcal{W}(\mathcal{C}(\mathbf{y}, t; \ell)), \text{ as } n \rightarrow \infty. \quad (4.5)$$

Using (4.3) in conjunction with (4.4) and (4.5) and passing to the  $n \rightarrow \infty$  limit, it follows that  $(f_\ell(G(\mathbf{y}, t)), t \geq 0)$  is a submartingale.

To prove (4.2), note that

$$\begin{aligned} \mathcal{A} f_\ell(G) &= \sum_{i \neq k} z_k z_i \left( \frac{(z_k + z_i)^2}{\sigma_2 + 2z_k z_i} - \frac{z_k^2}{\sigma_2} \right) \\ &\quad + \sum_{\substack{i \geq 1 \\ i \neq k}} \sum_{\substack{j > i \\ j \neq k}} z_i z_j \left( \frac{z_k^2}{\sigma_2 + 2z_i z_j} - \frac{z_k^2}{\sigma_2} \right) =: T_1 + T_2. \end{aligned} \quad (4.6)$$

Now

$$\begin{aligned} T_1 &= \sum_{i \neq k} \frac{z_k z_i}{\sigma_2 (\sigma_2 + 2z_k z_i)} \left( (z_k^2 + z_i^2 + 2z_k z_i) \sigma_2 - z_k^2 (\sigma_2 + 2z_k z_i) \right) \\ &= \sum_{i \neq k} \frac{z_k z_i}{\sigma_2 (\sigma_2 + 2z_k z_i)} \left( z_i^2 \sigma_2 + 2z_k z_i (\sigma_2 - z_k^2) \right) \\ &= z_k \sum_{i \neq k} \frac{z_i^3}{\sigma_2 + 2z_k z_i} + 2z_k^2 \sum_{i \neq k} \frac{z_i^2 (\sum_{j \neq k} z_j^2)}{\sigma_2 (\sigma_2 + 2z_k z_i)}. \end{aligned}$$

Using  $2z_k z_i \leq \sigma_2$ , we get

$$T_1 \geq z_k \sum_{i \neq k} \frac{z_i^3}{\sigma_2 + 2z_k z_i} + \frac{z_k^2}{\sigma_2^2} \times \left( \sum_{i \neq k} z_i^2 \right)^2. \quad (4.7)$$

Next

$$-T_2 = z_k^2 \sum_{\substack{i \geq 1 \\ i \neq k}} \sum_{\substack{j > i \\ j \neq k}} \frac{2z_i^2 z_j^2}{\sigma_2 (\sigma_2 + 2z_i z_j)} \leq z_k^2 \sum_{\substack{i \geq 1 \\ i \neq k}} \sum_{\substack{j > i \\ j \neq k}} \frac{2z_i^2 z_j^2}{\sigma_2^2} \leq \frac{z_k^2}{\sigma_2^2} \times \left( \sum_{i \neq k} z_i^2 \right)^2. \quad (4.8)$$

Combining (4.6), (4.7), and (4.8) yields (4.2).

Now take  $\ell = 1$  and assume that  $f_1(G(\mathbf{y}, 0)) = y_1^2 / \sigma_2(\mathbf{y}) > 1 - \varepsilon$ . Define the stopping time

$$T = \inf \{s \geq 0 : f_1(G(\mathbf{y}, s)) \leq 4/5\},$$

where the infimum of an empty set is understood to be  $+\infty$ . Since  $f_1(G(\mathbf{y}, t))$  is a bounded submartingale,  $f_1(G(\mathbf{y}, \infty)) := \lim_{t \rightarrow \infty} f_1(G(\mathbf{y}, t))$  exists almost surely. Thus,

$$1 - \varepsilon \leq \mathbb{E} [f_1(G(\mathbf{y}, T))] \leq \frac{4}{5} \mathbb{P}(f_1(G(\mathbf{y}, T)) \leq 4/5) + \mathbb{P}(f_1(G(\mathbf{y}, T)) > 4/5).$$

This shows that  $\mathbb{P}(T = \infty) \geq 1 - 5\varepsilon$ . Then (4.1) follows upon noting that  $\mathcal{W}(\mathcal{C}(\mathbf{y}, t; 1)) > 2 \max_{i \geq 2} \mathcal{W}(\mathcal{C}_i(\mathbf{y}, t))$  if  $f_1(G(\mathbf{y}, t)) > 4/5$ . ■

**Proof of Proposition 4.1:** Combined with Proposition 4.3, the bound given in Proposition 4.4 below completes the proof. ■

**Proposition 4.4.** *Under Condition I, for all  $\varepsilon > 0$ , there exists  $\lambda > 0$  depending only on  $\varepsilon$  and the sequence  $\{\mathbf{x}^{(n)}\}_{n \geq 1}$  such that for all  $n \geq 1$ ,*

$$\mathbb{P} \left( \frac{(x_1^{(n)}(t_\lambda))^2}{\sigma_2(\mathbf{x}^{(n)}(t_\lambda))} \geq 1 - \varepsilon \right) \geq 1 - \varepsilon,$$

where  $t_\lambda = t_\lambda(\mathbf{x}^{(n)}) = \lambda + 1/\sigma_2(\mathbf{x}^{(n)})$ .

**Proof:** First consider the particular sequences  $\mathbf{y}^{(n)}$  defined in (2.1), for which  $\sigma_2(\mathbf{y}^{(n)}) = n^{-1/3}$ . We then have  $t_\lambda(\mathbf{y}^{(n)}) = \lambda + n^{1/3}$ . Ignoring vertex weights and multiple edges, the graph  $G(\mathbf{y}^{(n)}, \lambda + n^{1/3})$  has the same law as the Erdős-Rényi random graph on  $n$  vertices, where edges are placed independently between each pair of vertices with probability

$$p_\lambda = p_\lambda(n) = \mathbb{P}(\text{Poisson}(n^{-4/3}(\lambda + n^{1/3})) \geq 1) = (1 + o(1)) \left( \frac{1}{n} + \frac{\lambda}{n^{4/3}} \right).$$

It thus follows from [27, Theorem A.1] that

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left( \frac{(y_1^{(n)}(\lambda + n^{1/3}))^2}{\sigma_2(\mathbf{y}^{(n)}(\lambda + n^{1/3}))} \geq 1 - \varepsilon \right) \geq 1 - \varepsilon, \quad (4.9)$$

provided  $\lambda$  is sufficiently large.

Fix any sequence  $\{\mathbf{x}^{(n)}\}_{n \geq 1}$  satisfying Condition I. Using (3.8), we see that with  $t_\lambda = t_\lambda(\mathbf{x}^{(n)})$ ,

$$\mathbf{x}^{(n)}(t_\lambda) \xrightarrow{d} \zeta(\lambda) \text{ on } l_1^2, \text{ and} \quad (4.10)$$

$$\mathbf{y}^{(n)}(\lambda + n^{1/3}) \xrightarrow{d} \zeta(\lambda) \text{ on } l_1^2. \quad (4.11)$$

Now fix  $\lambda$  sufficiently large so that (4.9) holds. Then using (4.10) and (4.11),

$$\begin{aligned} \liminf_{n \rightarrow \infty} \mathbb{P} \left( \frac{(x_1^{(n)}(t_\lambda))^2}{\sigma_2(\mathbf{x}^{(n)}(t_\lambda))} > 1 - 2\varepsilon \right) &\geq \mathbb{P} \left( \frac{|\gamma_1(\lambda)|^2}{\sigma_2(\zeta(\lambda))} > 1 - 2\varepsilon \right) \\ &\geq \mathbb{P} \left( \frac{|\gamma_1(\lambda)|^2}{\sigma_2(\zeta(\lambda))} \geq 1 - \varepsilon \right) \\ &\geq \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{(y_1^{(n)}(\lambda + n^{1/3}))^2}{\sigma_2(\mathbf{y}^{(n)}(\lambda + n^{1/3}))} \geq 1 - \varepsilon \right) \geq 1 - \varepsilon, \end{aligned}$$

as desired. ■

It remains to prove tightness of the family of random variables in Proposition 4.2. We will need some properties of the distributional limits  $(\zeta(\lambda), \lambda \in \mathbb{R})$ . By [5, Proposition 18 and Equation (80)], for any  $\varepsilon > 0$ , we can choose  $\lambda_\varepsilon > 0$  such that with probability at least  $1 - \varepsilon$ , the following three assertions hold simultaneously:

$$\lambda_\varepsilon \sigma_2(\zeta(-\lambda_\varepsilon)) \geq 1/2, \quad \left| \frac{1}{\sigma_2(\zeta(-\lambda_\varepsilon))} - \lambda_\varepsilon \right| \leq 1, \text{ and } \lambda_\varepsilon \times |\gamma_1(-\lambda_\varepsilon)| \leq \varepsilon. \quad (4.12)$$

Writing  $t_\lambda = \lambda + 1/\sigma_2(\mathbf{x}^{(n)})$  as before, and using (4.10), we can choose  $n_0(\varepsilon)$  such that for all  $n \geq n_0(\varepsilon)$ ,

$$\lambda_\varepsilon \sigma_2(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon})) \geq 1/2, \quad \left| \frac{1}{\sigma_2(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}))} - \lambda_\varepsilon \right| \leq 1, \text{ and } \lambda_\varepsilon \times x_1^{(n)}(t_{-\lambda_\varepsilon}) \leq \varepsilon \quad (4.13)$$

with probability at least  $1 - 2\varepsilon$ .

We will also use a simple lemma leveraging the coupling between  $(\mathbf{x}(t), t \geq 0)$  and  $(G(\mathbf{x}, t), t \geq 0)$ . Recall that  $\mathcal{C}(\mathbf{x}, t; j)$  is the component containing vertex  $j$  in  $G(\mathbf{x}, t)$ .

**Lemma 4.5.** For  $\mathbf{x} = (x_i, i \geq 1) \in l_1^2$  and  $t > 0$ , if  $t\sigma_2(\mathbf{x}) < 1$  then for all  $j \geq 1$ ,

$$\mathbb{E}[\mathcal{W}(\mathcal{C}(\mathbf{x}, t; j))] \leq \frac{x_j}{1 - t\sigma_2(\mathbf{x})}.$$

**Proof:** By the definition of  $G(\mathbf{x}, t)$ , for each  $i \neq j$ , the probability that there is at least one edge between  $i$  and  $j$  in  $G(\mathbf{x}, t)$  is  $1 - \exp(-tx_i x_j)$ , and these events are mutually independent. It follows that for  $i \neq j$ ,

$$\begin{aligned} \mathbb{P}(i \in \mathcal{C}(\mathbf{x}, t; j)) &\leq \sum_{k \geq 0} \sum_{j_1 \neq j} \dots \sum_{j_k \neq j} (tx_j x_{j_1}) \cdot (tx_{j_1} x_{j_2}) \cdot \dots \cdot (tx_{j_{k-1}} x_{j_k}) \cdot (tx_{j_k} x_i) \\ &= tx_i x_j \sum_{k \geq 0} (t(\sigma_2(\mathbf{x}) - x_j^2))^k = \frac{tx_i x_j}{1 - t(\sigma_2(\mathbf{x}) - x_j^2)}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbb{E}[\mathcal{W}(\mathcal{C}(\mathbf{x}, t; j))] &= x_j + \sum_{i \neq j} x_i \cdot \mathbb{P}(i \in \mathcal{C}(\mathbf{x}, t; j)) \leq x_j + \frac{tx_j(\sigma_2(\mathbf{x}) - x_j^2)}{1 - t(\sigma_2(\mathbf{x}) - x_j^2)} \\ &= \frac{x_j}{1 - t(\sigma_2(\mathbf{x}) - x_j^2)} \leq \frac{x_j}{1 - t\sigma_2(\mathbf{x})}, \end{aligned} \tag{4.14}$$

as desired. ■

**Proof of Proposition 4.2:** Let us first describe the core idea in words. Using the above estimates and lemma, we will show that the maximal component from time  $t_{-\lambda_\varepsilon}$  is unlikely to become very large by time  $t_{-2}$ . On the other hand, by (4.10), the maximal component at time  $t_{-2}$  is reasonably large. This implies that with high probability the leader changes at least once in the time interval  $[t_{-\lambda_\varepsilon}, t_{-2}]$ . We now make this idea precise.

Let  $\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}) =: \mathbf{z} = (z_i, i \geq 1)$  and suppose we start the multiplicative coalescent with  $\mathbf{z}$  as the initial configuration. Run this process for  $\lambda_\varepsilon - 2$  units of time. Let  $E(n)$  be the event given in (4.13). Using Lemma 4.5, on the event  $E(n)$ ,

$$\begin{aligned} \mathbb{E}[\mathcal{W}(\mathcal{C}(\mathbf{z}, \lambda_\varepsilon - 2; 1)) \mid \mathbf{z}] &\leq \frac{z_1}{1 - (\lambda_\varepsilon - 2)\sigma_2(\mathbf{z})} \\ &= \frac{z_1}{\sigma_2(\mathbf{z})} \times \frac{1}{\frac{1}{\sigma_2(\mathbf{z})} - \lambda_\varepsilon + 2} \leq \frac{z_1}{\sigma_2(\mathbf{z})} \leq \frac{\varepsilon}{\lambda_\varepsilon \sigma_2(\mathbf{z})} \leq 2\varepsilon. \end{aligned}$$

Hence, for all  $\delta > 0$  and  $n \geq n_0(\varepsilon)$ , by an application of Markov's inequality,

$$\mathbb{P}(\mathcal{W}(\mathcal{C}(\mathbf{z}, \lambda_\varepsilon - 2; 1)) \geq \delta) \leq \mathbb{P}(E(n)^c) + 2\varepsilon/\delta \leq 2\varepsilon(1 + 1/\delta). \tag{4.15}$$

However by (4.10) for all  $\eta > 0$ , there exists  $\delta_\eta \in (0, 1)$  and  $n_1(\eta) \geq 1$  such that for all  $n \geq n_1(\eta)$ ,

$$\mathbb{P}(x_1^{(n)}(t_{-2}) \geq \delta_\eta) \geq 1 - \eta. \tag{4.16}$$

Taking  $\varepsilon = \eta\delta_\eta$  and combining (4.15) and (4.16), we see that for all  $n \geq n_0(\varepsilon) \vee n_1(\eta)$ ,

$$\mathbb{P}\left(\left(L(\mathbf{x}^{(n)}) - \frac{1}{\sigma_2(\mathbf{x}^{(n)})}\right)^- \geq \lambda_\varepsilon\right) \leq \eta + 2\varepsilon(1 + 1/\delta_\eta) \leq 5\eta,$$

which completes the proof. ■

## 5. PROOF OF THEOREM 2.4

The proof of Proposition 4.2 in fact shows that for any  $\eta > 0$ , there is  $\lambda_\eta^{(1)} > 0$  such that for all  $n$  sufficiently large,

$$\mathbb{P}(\text{The leader changes between times } t = t_{-\lambda_\eta^{(1)}} \text{ and } t = 0) \geq 1 - \eta.$$

By repeating the same argument, we can choose  $\lambda_\eta^{(1)} < \lambda_\eta^{(2)} < \lambda_\eta^{(3)} \dots$  such that for all  $j \geq 1$ , for all  $n$  sufficiently large,

$$\mathbb{P}(\text{The leader changes between times } t = t_{-\lambda_\eta^{(j+1)}} \text{ and } t = t_{-\lambda_\eta^{(j)}}) \geq 1 - \eta/2^j.$$

It thus follows that for any  $j \geq 1$  and  $\eta > 0$ ,

$$\limsup_n \mathbb{P}(N^{(n)} \leq j) \leq 2\eta,$$

which completes the proof.

## 6. PROOF OF THEOREM 2.2

In this section, we work under Condition II. Thus, throughout this section,  $\mathbf{c} = (c_i, i \geq 1)$  satisfies  $c_i \asymp i^{-\alpha}$  for some fixed  $\alpha \in (1/3, 1/2)$ , and  $(\mathbf{x}^{(n)}, n \geq 1)$  satisfies (1.6) and (1.8) with this limiting sequence  $\mathbf{c}$ .

Define  $\phi : [0, \infty) \rightarrow \mathbb{R}$  and  $\Phi_\lambda : [0, \infty) \rightarrow \mathbb{R}$  for each  $\lambda > 0$  by setting

$$\phi(s) = \sum_{i \geq 1} c_i^2 \left( \frac{c_i s - 1 + e^{-c_i s}}{c_i s} \right) \quad \text{and} \quad \Phi_\lambda(s) := \lambda s - s\phi(s).$$

Since  $u \mapsto (u - 1 + e^{-u})$  is convex, it follows that  $s\phi(s)$  is also convex, and consequently  $\Phi_\lambda$  is concave. Also observe that for each  $\lambda > 0$ ,  $\Phi_\lambda$  has a unique positive zero, which we denote by  $s_0(\lambda)$ . Further,  $s_0(\lambda)$  is strictly increasing on  $(0, \infty)$ , and  $s_0(\lambda) \uparrow \infty$  as  $\lambda \uparrow \infty$ .

Recall the definitions of the processes  $W_\lambda(\cdot)$  and  $\overline{W}_\lambda(\cdot)$  from (3.4) and (3.5), and of the vector  $\zeta(\lambda) := (|\gamma_i(\lambda)|, i \geq 1) \in l_\downarrow^2$  of ranked excursion lengths of  $\overline{W}_\lambda$ . Note that for all  $\lambda > 0$ ,

$$W_\lambda(s) - \Phi_\lambda(s) = \sum_{i \geq 1} c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)). \quad (6.1)$$

The next theorem bounds the length and squared sum of jump sizes of the longest excursion  $\gamma_1(\lambda)$ .

**Theorem 6.1.** *For all  $\eta > 0$ , there exist positive constants  $\lambda_0$  and  $C$  depending only on  $\eta$  and on  $(c_i, i \geq 1)$  such that for all  $\lambda \geq \lambda_0$ ,*

$$\mathbb{P}\left(|\gamma_1(\lambda)| \leq (1 - \eta)s_0(\lambda)\right) \leq \exp\left(-C\lambda^{\frac{1}{2(1-2\alpha)}} \log \log \lambda\right), \quad (6.2)$$

and

$$\mathbb{P}\left(|\gamma_1(\lambda)| \geq (1 + \eta)s_0(\lambda)\right) \leq C\lambda^{-\frac{1}{1-2\alpha}}. \quad (6.3)$$

Further, there exist positive constants  $\delta, C'$  and  $\lambda'$  depending only on  $\mathbf{c}$  such that for all  $\lambda \geq \lambda'$ ,

$$\mathbb{P}\left(\sum_{i \in \gamma_1(\lambda)} c_i^2 \leq (1 + 2\delta)\lambda\right) \leq C'\lambda^{-\frac{1}{1-2\alpha}}. \quad (6.4)$$

Theorem 6.1 will imply rather straightforwardly that the bound of Proposition 4.4 also holds under Condition II; this is the content of the next proposition.



**Proposition 6.2.** *Under Condition II, for all  $\varepsilon > 0$ , there exists  $\lambda_0 > 0$  depending only on  $\varepsilon$  and the sequence  $\{\mathbf{x}^{(n)}\}_{n \geq 1}$  such that for all  $\lambda \geq \lambda_0$ ,*

$$\liminf_{n \rightarrow \infty} \mathbb{P} \left( \frac{(x_1^{(n)}(t_\lambda))^2}{\sigma_2(\mathbf{x}^{(n)}(t_\lambda))} \geq 1 - \varepsilon \right) \geq 1 - \varepsilon,$$

where  $t_\lambda = t_\lambda(\mathbf{x}^{(n)}) = \lambda + 1/\sigma_2(\mathbf{x}^{(n)})$ .

Combined with Proposition 4.3, Theorem 2.2 follows immediately. Thus, the rest of the section is devoted to the proofs of Theorem 6.1 and Proposition 6.2.

We first list some simple asymptotics which will be useful in the proof. Note that  $u - 1 + e^{-u} \asymp u$  on  $u \geq 1$ , and  $u - 1 + e^{-u} \asymp u^2$  on  $u \in [0, 1]$ . For  $s > 0$ , let

$$i_0(s) = \min \{i \geq 1 : c_i s < 1\}. \quad (6.5)$$

Note that  $i_0(s) \geq 2$  when  $s \geq 1/c_1$ , and

$$s\phi(s) = \sum_{i=1}^{i_0(s)-1} c_i (c_i s - 1 + e^{-c_i s}) + \sum_{i=i_0(s)}^{\infty} c_i (c_i s - 1 + e^{-c_i s}) \asymp \sum_{i=1}^{i_0(s)-1} c_i (c_i s) + \sum_{i=i_0(s)}^{\infty} c_i (c_i s)^2 \quad (6.6)$$

on  $s \geq 1/c_1$ . From the relation  $c_i \asymp i^{-\alpha}$  it follows that

$$i_0(s) \asymp s^{1/\alpha} \quad \text{on } s \geq 1/c_1. \quad (6.7)$$

Combining (6.6) with (6.7) and the relation  $c_i \asymp i^{-\alpha}$ , we get

$$s\phi(s) = \sum_{i=1}^{\infty} c_i (c_i s - 1 + e^{-c_i s}) \asymp s^{(1-\alpha)/\alpha} \quad \text{on } s \geq 1/c_1. \quad (6.8)$$

Similarly, using the relations

$$1 - e^{-u} \asymp 1 \quad \text{on } u \geq 1 \quad \text{and} \quad 1 - e^{-u} \asymp u \quad \text{on } u \in [0, 1], \quad (6.9)$$

we get

$$\sum_{i=1}^{\infty} c_i^2 (1 - e^{-c_i s}) \asymp s^{(1-2\alpha)/\alpha} \quad \text{on } s \geq 1/c_1, \quad (6.10)$$

and, for any  $k \geq 2$  and  $j \geq 1$ ,

$$\sum_{i=j}^{\infty} c_i^k (1 - e^{-c_i s}) \asymp s j^{1-\alpha(k+1)} \quad \text{on } 0 \leq s \leq 1/c_j. \quad (6.11)$$

Finally, note that  $1 - e^{-u} - ue^{-u} \geq 0$  for  $u \in \mathbb{R}$ , and further,  $1 - e^{-u} - ue^{-u} \asymp 1$  on  $u \geq 1$  and  $1 - e^{-u} - ue^{-u} \asymp u^2$  on  $u \in [0, 1]$ . Hence,

$$\sum_{i=1}^{\infty} c_i (1 - e^{-c_i s} - c_i s e^{-c_i s}) \asymp s^{(1-\alpha)/\alpha} \quad \text{on } s \geq 1/c_1. \quad (6.12)$$

For fixed  $\eta \in (0, 1)$  and  $w \in [1 - \eta, 1]$ ,

$$\left| \sum_{i=1}^{\infty} c_i (c_i s e^{-c_i s w} - c_i s e^{-c_i s}) \right| \leq \sum_{i=1}^{\infty} c_i^2 s (1 - e^{-c_i s \eta}) \quad \text{for } s \geq 0. \quad (6.13)$$

If  $c_1 s \eta \geq 1$ , then

$$\sum_{i=1}^{\infty} c_i^2 s (1 - e^{-c_i s \eta}) = \sum_{i=1}^{i_0(s\eta)-1} c_i^2 s (1 - e^{-c_i s \eta}) + \sum_{i=i_0(s\eta)}^{\infty} c_i^2 s (1 - e^{-c_i s \eta}) \leq C s^{(1-\alpha)/\alpha} \eta^{(1-2\alpha)/\alpha}, \quad (6.14)$$

where  $i_0(\cdot)$  is as in (6.5), and the last inequality uses (6.9), (6.7), and the relation  $c_i \asymp i^{-\alpha}$ . If  $c_1 s \eta < 1$ , then

$$\sum_{i=1}^{\infty} c_i^2 s (1 - e^{-c_i s \eta}) \leq s^2 \eta \sum_{i=1}^{\infty} c_i^3 \leq s^{(1-\alpha)/\alpha} \cdot (c_1 \eta)^{(1-3\alpha)/\alpha} \cdot \eta \cdot \sum_{i=1}^{\infty} c_i^3 \leq C s^{(1-\alpha)/\alpha} \eta^{(1-2\alpha)/\alpha}. \quad (6.15)$$

Combining (6.13), (6.14), and (6.15) with (6.12), we see that for  $\eta > 0$  sufficiently small, uniformly over  $w \in [1 - \eta, 1]$ ,

$$\sum_{i=1}^{\infty} c_i (1 - e^{-c_i s} - c_i s e^{-c_i s w}) \asymp s^{(1-\alpha)/\alpha} \quad \text{on } s \geq 1/c_1. \quad (6.16)$$

For  $u > 0$ , define

$$Z_u^{(1)} = \sup_{s \leq u} \left| \sum_{i \geq 1} c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| = \sup_{s \leq u} |W_\lambda(s) - \Phi_\lambda(s)|, \quad (6.17)$$

where the second equality follows from (6.1). Note that

$$\begin{aligned} Z_u^{(1)} &= \lim_{k \rightarrow \infty} \sup_{s \leq u: ks \in \mathbb{N}} \left| \sum_{i \geq 1} c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \\ &= \lim_{k \rightarrow \infty} \left( \lim_{n \rightarrow \infty} \sup_{s \leq u: ks \in \mathbb{N}} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right), \end{aligned}$$

where the second step uses the fact that  $\sum_{i \geq 1} c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s))$  is almost surely convergent for any fixed  $s$  (which is a simple consequence of Kolmogorov's three-series theorem [22] and the fact that  $\mathbf{c} \in \ell^3$ ). Defining

$$Z_{u,n}^{(1)} = \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right|,$$

for all  $k$  we have  $\sup_{s \leq u: ks \in \mathbb{N}} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \leq Z_{u,n}^{(1)}$ . Combined with the above identities and an application of Fatou's lemma, this yields that for all  $z$ ,

$$\mathbb{P}(Z_u^{(1)} > z) \leq \sup_{n \geq 1} \mathbb{P}(Z_{u,n}^{(1)} > z). \quad (6.18)$$

The point of this is that it will shortly allow the use of discrete-time martingale bounds which *a priori* only hold for observables of finite-dimensional random variables. We likewise define

$$Z_u^{(2)} = \sup_{s \leq u} \left| \sum_{i \geq 1} c_i^2 (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \quad \text{and} \quad Z_{u,n}^{(2)} = \sup_{s \leq u} \left| \sum_{i=1}^n c_i^2 (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right|. \quad (6.19)$$

The same argument as above shows that for all  $z$ ,

$$\mathbb{P}(Z_u^{(2)} > z) \leq \sup_{n \geq 1} \mathbb{P}(Z_{u,n}^{(2)} > z). \quad (6.20)$$

**Lemma 6.3.** *There exists a constant  $C_{6.3}$  depending only on  $(c_i, i \geq 1)$  such that for all  $u \geq 1/C_{6.3}$  and all  $x \geq 1/C_{6.3}$ ,*

$$\mathbb{P}\left(Z_u^{(1)} \geq xu^{(1-2\alpha)/(2\alpha)}\right) \leq \exp\left(-C_{6.3}x \log \log x\right), \quad \text{and} \quad (6.21)$$

$$\mathbb{P}\left(Z_u^{(2)} \geq x\right) \leq \exp\left(-C_{6.3}x \log \log x\right). \quad (6.22)$$

**Proof:** By (6.18), it suffices to prove (6.21) with  $Z_{u,n}^{(1)}$  in place of  $Z_u^{(1)}$ . We consider positive and negative fluctuations separately, defining

$$U_{u,n} = \sup_{s \leq u} \left( \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right), \quad \text{and} \quad V_{u,n} = \sup_{s \leq u} \left( \sum_{i=1}^n c_i (\mathbb{P}(\xi_i \leq s) - \mathbb{1}_{\{\xi_i \leq s\}}) \right).$$

Note that  $Z_{u,n}^{(1)} \leq U_{u,n} + V_{u,n}$ .

We shall apply Theorem 3.1 to the random variable  $u^{-(1-2\alpha)/(2\alpha)} U_{u,n}$ . To this end, observe that for  $s \leq u$ ,

$$\begin{aligned} \text{Var} \left( u^{-(1-2\alpha)/(2\alpha)} \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right) &= u^{-(1-2\alpha)/\alpha} \sum_{i=1}^n c_i^2 \mathbb{P}(\xi_i \leq s) (1 - \mathbb{P}(\xi_i \leq s)) \\ &\leq u^{-(1-2\alpha)/\alpha} \sum_{i \geq 1} c_i^2 (1 - e^{-c_i u}) \leq K', \end{aligned}$$

where the last bound is due to (6.10). Now Lemma 6.4 given below implies that  $\mathbb{E} U_{u,n} \leq K_{6.4} u^{(1-2\alpha)/(2\alpha)}$  provided  $u \geq 1/c_1$ . Using this, an application of Theorem 3.1 with  $V = K'$  and  $A = c_1 u^{-(1-2\alpha)/(2\alpha)}$  gives

$$\mathbb{P} \left( \frac{U_{u,n}}{u^{(1-2\alpha)/(2\alpha)}} \geq K_{6.4} + x \right) \leq \exp \left( - \frac{x u^{(1-2\alpha)/(2\alpha)}}{4c_1} \log \left( 1 + 2 \log \left( 1 + \frac{x u^{(1-2\alpha)/(2\alpha)}}{c_1 K'} \right) \right) \right).$$

(Note that the supremum in the definition of  $U_{u,n}$  can be equivalently taken over the countable set  $[0, u] \cap \mathbb{Q}$ , where  $\mathbb{Q}$  denotes the set of rationals. Thus, Theorem 3.1 is applicable here.) An identical treatment yields a similar tail bound for  $V_{u,n}$ , which, combined with the preceding tail bound, establishes (6.21).

The proof of (6.22) is very similar to the proof of (6.21). The key difference comes from the improved bound

$$\text{Var} \left( \sum_{i=1}^n c_i^2 (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right) = \sum_{i=1}^n \text{Var} (c_i^2 (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s))) \leq \sum_{i=1}^{\infty} c_i^4 < \infty,$$

and the use of (6.24) rather than (6.23). This explains the gain of a factor  $u^{(1-2\alpha)/(2\alpha)}$  in the bound of (6.22) relative to that of (6.21). We omit the details to avoid repetition.  $\blacksquare$

**Lemma 6.4.** *There exists a constant  $K_{6.4}$  depending only on  $(c_i, i \geq 1)$  such that for all  $u \geq 1/c_1$  and all  $n \geq 1$ ,*

$$\mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right) \leq K_{6.4} u^{(1-2\alpha)/(2\alpha)}, \quad \text{and} \quad (6.23)$$

$$\mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i^2 (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right) \leq K_{6.4}. \quad (6.24)$$

**Proof:** Let  $(\xi'_1, \dots, \xi'_n)$  be an independent copy of  $(\xi_1, \dots, \xi_n)$ . Writing  $\mathbb{E}_\xi[\cdot] := \mathbb{E}[\cdot | \xi_1, \dots, \xi_n]$ , we have

$$\begin{aligned} & \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right) \\ &= \mathbb{E} \left( \sup_{s \leq u} \left| \mathbb{E}_\xi \left( \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{1}_{\{\xi'_i \leq s\}}) \right) \right| \right) \\ &\leq \mathbb{E} \mathbb{E}_\xi \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{1}_{\{\xi'_i \leq s\}}) \right| \right) = \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{1}_{\{\xi'_i \leq s\}}) \right| \right). \end{aligned}$$

Introducing i.i.d. random variables  $\varepsilon_1, \dots, \varepsilon_n$  with  $\mathbb{P}(\varepsilon_1 = 1) = 1/2 = \mathbb{P}(\varepsilon_1 = -1)$ , we see that

$$\begin{aligned} & \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right) \tag{6.25} \\ &\leq \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{1}_{\{\xi'_i \leq s\}}) \right| \right) \\ &\leq \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq s\}} \right| \right) + \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (-\varepsilon_i) \mathbb{1}_{\{\xi'_i \leq s\}} \right| \right) \\ &= 2 \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq s\}} \right| \right) = 2 \mathbb{E} \mathbb{E}_\xi \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq s\}} \right| \right). \end{aligned}$$

To bound the inner expectation, let  $\{k(1), \dots, k(M)\} \subseteq \{1, \dots, n\}$  be the indices  $k$  for which  $\xi_k \leq u$ , listed in increasing order. Then

$$\sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq s\}} \right| = \max_{1 \leq j \leq M} \left| \sum_{i=1}^j c_{k(i)} \varepsilon_{k(i)} \right|.$$

Write  $S(j) = \sum_{i=1}^j c_{k(i)} \varepsilon_{k(i)}$ . Then, conditionally on  $\xi$ ,  $(S(j), 0 \leq j \leq M)$  is a martingale. Hence,

$$\begin{aligned} \mathbb{E}_\xi \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq s\}} \right| \right) &= \mathbb{E}_\xi \left[ \max_{1 \leq j \leq M} |S(j)| \right] \leq \left( \mathbb{E}_\xi \left[ \max_{1 \leq j \leq M} S(j)^2 \right] \right)^{1/2} \\ &\leq 2 \left( \mathbb{E}_\xi [S(M)^2] \right)^{1/2} = 2 \left( \mathbb{E}_\xi \left[ \left( \sum_{i=1}^n c_i \varepsilon_i \mathbb{1}_{\{\xi_i \leq u\}} \right)^2 \right] \right)^{1/2} \leq 2 \left( \sum_{i \geq 1} c_i^2 \mathbb{1}_{\{\xi_i \leq u\}} \right)^{1/2}, \end{aligned}$$

where the second step uses the Cauchy-Schwarz inequality, the third step uses Doob's inequality, and the final inequality follows by directly expanding the square and using the independence of  $\varepsilon_i$ ,  $1 \leq i \leq n$ . Combined with (6.25), this gives

$$\begin{aligned} \mathbb{E} \left( \sup_{s \leq u} \left| \sum_{i=1}^n c_i (\mathbb{1}_{\{\xi_i \leq s\}} - \mathbb{P}(\xi_i \leq s)) \right| \right) &\leq 4 \mathbb{E} \left[ \left( \sum_{i \geq 1} c_i^2 \mathbb{1}_{\{\xi_i \leq u\}} \right)^{1/2} \right] \\ &\leq 4 \left( \mathbb{E} \left[ \sum_{i \geq 1} c_i^2 \mathbb{1}_{\{\xi_i \leq u\}} \right] \right)^{1/2} = 4 \left( \sum_{i \geq 1} c_i^2 (1 - e^{-c_i u}) \right)^{1/2} \leq K_{6.4} u^{(1-2\alpha)/(2\alpha)}, \end{aligned}$$

the final inequality holding by (6.10). This yields (6.23). The proof of (6.24) is essentially identical and is omitted.  $\blacksquare$

The final step before the proofs of Theorem 6.1 and Proposition 6.2 is to collect a few facts about the functions  $\phi$  and  $s_0$  defined at the start of the section, and about the function  $\psi : [0, \infty) \rightarrow \mathbb{R}$  defined by

$$\psi(s) = \sum_{i \geq 1} c_i^2 (1 - e^{-c_i s}).$$

**Lemma 6.5.** *The following properties are satisfied by  $\phi$ ,  $\psi$ , and  $s_0$ .*

- (a)  $s_0(\lambda) \asymp \lambda^{\alpha/(1-2\alpha)}$  on  $[s_0^{-1}(1/c_1), \infty)$ .
- (b) There exists  $\eta_0 \in (0, 1/2)$  such that for all  $\eta \in (0, \eta_0]$ ,

$$\sup_{s \geq 2/c_1} \frac{\phi(\eta s)}{\phi(s)} \leq \sup_{s \geq 2/c_1} \frac{\phi((1-\eta)s)}{\phi(s)} =: 1 - \varepsilon_\eta,$$

where  $\varepsilon_\eta > 0$  and  $\lim_{\eta \rightarrow 0} \varepsilon_\eta = 0$ .

- (c) There exist  $\delta_0 > 0$  and  $\lambda_0 > 0$  such that for all  $\lambda \geq \lambda_0$ ,

$$\psi(s_0(\lambda)) \geq (1 + \delta_0)\lambda.$$

- (d) For all  $s > 0$  and  $\eta \in (0, 1)$ ,  $\psi(\eta s) \geq \eta \psi(s)$ . Further, if

$$g(\eta) := \sup_{s \geq 1/c_1} \frac{\psi(\eta s)}{\psi(s)},$$

then  $g(\eta) \rightarrow 0$  as  $\eta \rightarrow 0$ .

**Proof:** Note that  $\lambda = \phi(s_0(\lambda)) \asymp (s_0(\lambda))^{(1-2\alpha)/\alpha}$  on  $\{\lambda : s_0(\lambda) \geq 1/c_1\}$ , where the final step uses (6.8). This proves (a).

Next, a direct calculation shows that the function  $u \mapsto (u - 1 + e^{-u})/u$  is increasing on  $\mathbb{R}$ . This implies that

$$\sup_{s \in [2/c_1, \infty)} \frac{\phi(\eta s)}{\phi(s)} \leq \sup_{s \in [2/c_1, \infty)} \frac{\phi((1-\eta)s)}{\phi(s)}$$

for  $0 < \eta < 1/2$ . Now

$$\begin{aligned} \frac{\phi(s)}{\phi((1-\eta)s)} - 1 &= \frac{\sum_{i \geq 1} c_i \left( \eta(1 - e^{-c_i s}) + e^{-c_i s} - e^{-c_i(1-\eta)s} \right)}{(1-\eta)s \cdot \phi((1-\eta)s)} \\ &\geq \frac{\sum_{i \geq 1} c_i \left( \eta(1 - e^{-c_i s}) + e^{-c_i s} - e^{-c_i(1-\eta)s} \right)}{K s^{(1-\alpha)/\alpha}} \\ &= \frac{\sum_{i \geq 1} c_i \int_{1-\eta}^1 \left( 1 - e^{-c_i s} - c_i s e^{-c_i s u} \right) du}{K s^{(1-\alpha)/\alpha}}, \end{aligned} \tag{6.26}$$

where the second step makes use of (6.8) and is valid whenever  $c_1(1-\eta)s \geq 1$ . By (6.16), there is  $\eta_0 \in (0, 1/2)$  such that uniformly over  $u \in [1-\eta_0, 1]$ ,

$$\sum_{i \geq 1} c_i (1 - e^{-c_i s} - c_i s e^{-c_i s u}) \asymp s^{(1-\alpha)/\alpha} \quad \text{on } s \geq 1/c_1.$$

This last observation combined with (6.26) shows that there exists  $K' > 0$  such that for every  $\eta \in (0, \eta_0]$ ,

$$\frac{\phi(s)}{\phi((1-\eta)s)} - 1 \geq K' \eta \quad \text{for all } s \geq 1/(c_1(1-\eta_0)).$$

which completes the proof of (b).

Since  $1 - e^{-u} \geq (u - 1 + e^{-u})/u$  for  $u \in (0, \infty)$ , it immediately follows that  $\psi(s) \geq \phi(s)$  for  $s \geq 0$ . Now

$$s(\psi(s) - \phi(s)) = \sum_{i \geq 1} c_i(1 - e^{-c_i s} - c_i s e^{-c_i s}) \asymp s^{(1-\alpha)/\alpha} \asymp s\phi(s) \text{ on } s \geq 1/c_1,$$

where we use (6.16) and (6.8) for the final two asymptotic equivalences. It follows that there exist  $\delta_0 > 0$  and  $s_\star > 0$  such that  $\psi(s) \geq (1 + \delta_0)\phi(s)$  for all  $s \geq s_\star$ , which in turn implies the existence of a  $\lambda_0$  such that

$$\psi(s_0(\lambda)) \geq (1 + \delta_0)\phi(s_0(\lambda)) = (1 + \delta_0)\lambda$$

for all  $\lambda \geq \lambda_0$ . This proves (c).

Finally, since  $1 - \exp(-\eta s) \geq \eta(1 - e^{-s})$ , it follows that  $\psi(\eta s) \geq \eta\psi(s)$ . Now note that

$$\begin{aligned} g(\eta) &\leq \sup_{\eta c_1 s \geq 1} \frac{\psi(\eta s)}{\psi(s)} + \sup_{1 \leq c_1 s \leq 1/\eta} \frac{\psi(\eta s)}{\psi(s)} \\ &\leq K \left( \eta^{(1-2\alpha)/\alpha} + \sup_{1 \leq c_1 s \leq 1/\eta} \frac{\eta s}{s^{(1-2\alpha)/\alpha}} \right) \leq K \left( \eta^{(1-2\alpha)/\alpha} + \frac{\eta^{(1-2\alpha)/\alpha}}{c_1^{(3\alpha-1)/\alpha}} \right), \end{aligned}$$

where the second step uses (6.10) and (6.11). This completes the proof of (d).  $\blacksquare$

**Proof of Theorem 6.1:** Let  $\delta_0$ ,  $\lambda_0$ ,  $\eta_0$ , and  $g(\cdot)$  be as in Lemma 6.5. Choose any  $\eta \in (0, \eta_0]$  small enough that

$$(1 - \eta - g(\eta))(1 + \delta_0) > 1 + \delta_0/2. \quad (6.27)$$

Now note that when  $s_0(\lambda) \geq 2/c_1$ ,

$$\begin{aligned} \Phi_\lambda((1 - \eta)s_0(\lambda)) &= (1 - \eta)s_0(\lambda) \left( \lambda - \phi((1 - \eta)s_0(\lambda)) \right) \\ &\geq (1 - \eta)s_0(\lambda) \left( \lambda - (1 - \varepsilon_\eta)\phi(s_0(\lambda)) \right) = (1 - \eta)s_0(\lambda) \lambda \varepsilon_\eta \geq K_\eta \lambda^{\frac{1-\alpha}{1-2\alpha}} \end{aligned}$$

for some  $K_\eta > 0$ , where the second step uses Lemma 6.5 (b) and the last step uses Lemma 6.5 (a). Similarly,  $\Phi_\lambda(\eta s_0(\lambda)) \geq K_\eta \lambda^{\frac{1-\alpha}{1-2\alpha}}$  when  $s_0(\lambda) \geq 2/c_1$ . Since  $\Phi_\lambda(\cdot)$  is concave, it follows that

$$\Phi_\lambda(s) \geq K_\eta \lambda^{\frac{1-\alpha}{1-2\alpha}}, \text{ for } \eta s_0(\lambda) \leq s \leq (1 - \eta)s_0(\lambda). \quad (6.28)$$

Recall the definitions of  $W_\lambda(\cdot)$  and  $Z_u^{(1)}$  from (3.4) and (6.17) respectively. Then

$$W_\lambda(s) \geq \Phi_\lambda(s) - Z_{s_0(\lambda)}^{(1)} \text{ for } 0 \leq s \leq s_0(\lambda). \quad (6.29)$$

Using (6.28), (6.29), Lemma 6.5(a), and applying (6.21) with  $u = s_0(\lambda)$  and  $x = \theta \lambda^{\frac{1}{2(1-2\alpha)}}$  where  $\theta > 0$  is very small, we see that for all  $\lambda$  sufficiently large,

$$\mathbb{P} \left( \inf \{ W_\lambda(s) : \eta s_0(\lambda) \leq s \leq (1 - \eta)s_0(\lambda) \} > 0 \right) \geq 1 - \exp(-C f_1(\lambda)), \quad (6.30)$$

where  $f_1(\lambda) = \lambda^{\frac{1}{2(1-2\alpha)}} \log \log \lambda$ . Write  $\gamma_\star$  for  $\gamma(\lambda; s_0(\lambda)/2)$ -the excursion of the reflected process  $\overline{W}_\lambda(\cdot)$  alive at time  $s_0(\lambda)/2$ . Then (6.30) shows that  $\gamma_\star$  is alive when  $\eta s_0(\lambda) \leq s \leq (1 - \eta)s_0(\lambda)$  with probability at least  $1 - \exp(-C f_1(\lambda))$ . Thus, with probability at least  $1 - \exp(-C f_1(\lambda))$ ,  $|\gamma_1(\lambda)| \geq (1 - 2\eta)s_0(\lambda)$ . This proves (6.2).

Let us make a note here of the following bound which we will use shortly: For all  $\lambda$  sufficiently large,

$$\mathbb{P}\left(\exists s \in [(1-\eta)s_0(\lambda), (1+\eta)s_0(\lambda)] \text{ such that } W_\lambda(s) < 0\right) \geq 1 - \exp(-Cf_1(\lambda)). \quad (6.31)$$

To see this, observe that Lemma 6.5 (b) implies that there exists  $\varepsilon'_\eta > 0$  such that  $\phi((1+\eta/2)s) \geq (1+\varepsilon'_\eta)\phi(s)$  for all  $s \geq 2/c_1$ . Consequently, for  $\lambda \geq s_0^{-1}(2/c_1)$ ,

$$\begin{aligned} \Phi_\lambda((1+\eta/2)s_0(\lambda)) &= (1+\eta/2)s_0(\lambda) \left( \lambda - \phi((1+\eta/2)s_0(\lambda)) \right) \\ &\leq (1+\eta/2)s_0(\lambda) \left( \lambda - (1+\varepsilon'_\eta)\phi(s_0(\lambda)) \right) = -(1+\eta/2)s_0(\lambda) \lambda \varepsilon'_\eta \leq -K'_\eta \lambda^{\frac{1-\alpha}{1-2\alpha}}, \end{aligned}$$

for some  $K'_\eta > 0$ , where the last step uses Lemma 6.5 (a). Since  $\Phi_\lambda$  is concave with  $\Phi_\lambda(0) = 0 = \Phi_\lambda(s_0(\lambda))$ ,  $\Phi_\lambda$  is non-increasing on  $[s_0(\lambda), \infty)$ . Hence, for  $\lambda \geq s_0^{-1}(2/c_1)$ ,  $\Phi_\lambda(s) \leq -K'_\eta \lambda^{\frac{1-\alpha}{1-2\alpha}}$  for  $(1+\eta/2)s_0(\lambda) \leq s \leq (1+\eta)s_0(\lambda)$ . Since  $W_\lambda(s) \leq \Phi_\lambda(s) + Z_{(1+\eta)s_0(\lambda)}^{(1)}$  for  $0 \leq s \leq (1+\eta)s_0(\lambda)$ , we can use (6.21) as before to get (6.31).

Now, for all sufficiently large  $\lambda$ ,

$$\sum_{i \in \gamma_\star} c_i^2 \geq Y := \sum_{i \geq 1} c_i^2 \mathbb{1}\{\eta s_0(\lambda) \leq \xi_i \leq (1-\eta)s_0(\lambda)\}$$

with probability at least  $1 - \exp(-Cf_1(\lambda))$ . Recall the definition of  $Z_u^{(2)}$  from (6.19). Then

$$Y \geq \psi((1-\eta)s_0(\lambda)) - \psi(\eta s_0(\lambda)) - 2Z_{s_0(\lambda)}^{(2)}.$$

It follows that for all sufficiently large  $\lambda$ , with probability at least  $1 - \exp(-Cf_1(\lambda))$ ,

$$\begin{aligned} \sum_{i \in \gamma_\star} c_i^2 &\geq (1-\eta)\psi(s_0(\lambda)) - g(\eta)\psi(s_0(\lambda)) - 2Z_{s_0(\lambda)}^{(2)} \\ &\geq (1-\eta-g(\eta))(1+\delta_0)\lambda - 2Z_{s_0(\lambda)}^{(2)} \geq (1+\delta_0/2)\lambda - 2Z_{s_0(\lambda)}^{(2)}, \end{aligned} \quad (6.32)$$

where the first, second, and third inequality use Lemma 6.5(d), Lemma 6.5(c), and (6.27) respectively. Using the tail bound (6.22) with  $x = \theta\lambda$  (where  $\theta > 0$  is very small), we see that for all  $\lambda$  sufficiently large,

$$\mathbb{P}\left(\sum_{i \in \gamma_\star} c_i^2 \leq (1+4\delta)\lambda\right) \leq \exp(-Kf_2(\lambda)) \quad (6.33)$$

where  $\delta := \delta_0/16$  and  $f_2(\lambda) := \lambda \log \log \lambda$ .

Let  $\lambda_0 > 0$  be large enough so that all the preceding bounds hold for every  $\lambda \geq \lambda_0$ . For the remainder of the proof, we work with a fixed  $\lambda \geq \lambda_0$ . None of the constants in the subsequent argument will depend on  $\lambda$ .

Recall from Section 3.4 the breadth-first walk process  $B^{(n)} = (B_{\mathbf{x}^{(n)}, t_\lambda}(s), s \geq 0)$  associated with  $G(\mathbf{x}^{(n)}, t_\lambda)$ , and write  $\mathfrak{C}_\star^{(n)}$  for  $\mathfrak{C}^{(n)}(\lambda; s_0(\lambda)/2)$ -the component of  $G(\mathbf{x}^{(n)}, t_\lambda)$  being explored by  $B^{(n)}$  at time  $s_0(\lambda)/2$ . Also, let  $\mathcal{F}^{(n)}$  and  $\mathcal{R}^{(n)}$  respectively denote the  $\sigma$ -field generated by the process  $B^{(n)}$  up to the time when the exploration of  $\mathfrak{C}_\star^{(n)}$  concludes and the set of vertices found up to that time.

Let  $E^{(n)}$  be the event that the following happen: **(1)** exploration of  $\mathfrak{C}_\star^{(n)}$  begins before time  $\eta s_0(\lambda)$ ; **(2)** exploration of  $\mathfrak{C}_\star^{(n)}$  ends between times  $(1 - \eta)s_0(\lambda)$  and  $(1 + \eta)s_0(\lambda)$ ; and **(3)**

$$\sum_{v \in \mathfrak{C}_\star^{(n)}} \frac{(x_v^{(n)})^2}{\sigma_2(\mathbf{x}^{(n)})^2} \geq \lambda(1 + 2\delta).$$

So in particular, if  $E^{(n)}$  occurs then

$$(1 - 2\eta)s_0(\lambda) \leq \mathcal{W}(\mathfrak{C}_\star^{(n)}) \leq (1 + \eta)s_0(\lambda), \text{ and } \mathcal{W}(\mathcal{C}(\mathbf{x}^{(n)}, t_\lambda; v)) \leq 2\eta s_0(\lambda) \text{ for } v \in \mathcal{R}^{(n)} \setminus \mathfrak{C}_\star^{(n)}. \quad (6.34)$$

Then (3.12) and (3.18) combined with (6.30), (6.31), and (6.33) imply that there exists  $n_0 = n_0(\lambda)$  such that

$$\mathbb{P}(E^{(n)}) \geq 1 - 2 \exp(-C f_2(\lambda)) \text{ for } n \geq n_0. \quad (6.35)$$

Let

$$\mathbf{x}^\star = \mathbf{x}^\star(n) = (x_v^{(n)}, v \notin \mathcal{R}^{(n)}),$$

and let

$$\sigma_2^\star = \sigma_2^\star(n) = \sigma_2(\mathbf{x}^{(n)}) - \sum_{v \in \mathcal{R}^{(n)}} (x_v^{(n)})^2.$$

Defining  $\lambda^\star$  by the identity  $\lambda + 1/\sigma_2(\mathbf{x}^{(n)}) = -\lambda^\star + 1/\sigma_2^\star$ , it follows that on the event  $E^{(n)}$ ,

$$\lambda^\star = \frac{\sigma_2(\mathbf{x}^{(n)}) - \sigma_2^\star}{\sigma_2(\mathbf{x}^{(n)})\sigma_2^\star} - \lambda \geq \frac{\lambda(1 + 2\delta)(\sigma_2(\mathbf{x}^{(n)}))^2}{\sigma_2(\mathbf{x}^{(n)})\sigma_2^\star} - \lambda \geq 2\delta\lambda. \quad (6.36)$$

Writing  $\mathcal{S}^{(n)}$  for the set of components of  $G(\mathbf{x}^{(n)}, t_\lambda)$  explored after exploring  $\mathfrak{C}_\star^{(n)}$ , we have

$$\sum_{\mathcal{C} \in \mathcal{S}^{(n)}} \mathcal{W}(\mathcal{C})^2 = \sum_{v \notin \mathcal{R}^{(n)}} x_v^{(n)} \cdot \mathcal{W}(\mathcal{C}(\mathbf{x}^{(n)}, t_\lambda; v)). \quad (6.37)$$

Write  $t_{-2\delta\lambda}^\star = -2\delta\lambda + 1/\sigma_2^\star$ . On account of (6.36), on the event  $E^{(n)}$ ,  $t_{-2\delta\lambda}^\star \geq t_\lambda$ . Consequently, for  $v \notin \mathcal{R}^{(n)}$ ,

$$\begin{aligned} \mathbb{1}_{E^{(n)}} \cdot \mathbb{E} \left[ \mathcal{W}(\mathcal{C}(\mathbf{x}^{(n)}, t_\lambda; v)) \mid \mathcal{F}^{(n)} \right] &= \mathbb{1}_{E^{(n)}} \cdot \mathbb{E} \left[ \mathcal{W}(\mathcal{C}(\mathbf{x}^\star, t_\lambda; v)) \mid \mathcal{F}^{(n)} \right] \\ &\leq \mathbb{E} \left[ \mathcal{W}(\mathcal{C}(\mathbf{x}^\star, t_{-2\delta\lambda}^\star; v)) \mid \mathcal{F}^{(n)} \right] \leq \frac{x_v^{(n)}}{1 - t_{-2\delta\lambda}^\star \sigma_2^\star} = \frac{x_v^{(n)}}{2\delta\lambda\sigma_2^\star}, \end{aligned} \quad (6.38)$$

where the third step uses Lemma 4.5. Combining (6.37) and (6.38), we get

$$\mathbb{E} \left[ \sum_{\mathcal{C} \in \mathcal{S}^{(n)}} \mathcal{W}(\mathcal{C})^2 \cdot \mathbb{1}_{E^{(n)}} \mid \mathcal{F}^{(n)} \right] \leq \sum_{v \notin \mathcal{R}^{(n)}} x_v^{(n)} \cdot \frac{x_v^{(n)}}{2\delta\lambda\sigma_2^\star} = \frac{1}{2\delta\lambda}. \quad (6.39)$$

It thus follows that

$$\mathbb{1}_{E^{(n)}} \cdot \mathbb{P} \left( \sum_{\mathcal{C} \in \mathcal{S}^{(n)}} \mathcal{W}(\mathcal{C})^2 \geq x \mid \mathcal{F}^{(n)} \right) \leq \frac{1}{2\delta x \lambda}.$$

Taking  $x = \theta' \lambda^{\frac{2\alpha}{1-2\alpha}}$  (where  $\theta' > 0$  is very small) and using Lemma 6.5 (a) together with (6.34) and (6.35), we see that for all large  $n$ ,

$$\mathbb{P}(\mathfrak{C}_\star^{(n)} \text{ is } \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)) \geq 1 - 2 \exp(-C f_2(\lambda)) - C' \lambda^{-\frac{1}{1-2\alpha}}. \quad (6.40)$$



(Recall that  $\mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)$  denotes the component of  $G(\mathbf{x}^{(n)}, t_\lambda)$  having the largest mass.) Using (6.40), (6.34), and (6.35), it follows that for large  $n$ ,

$$\mathbb{P}(\mathcal{W}(\mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)) > (1 + \eta)s_0(\lambda)) \leq 4 \exp(-Cf_2(\lambda)) + C'\lambda^{-\frac{1}{1-2\alpha}}.$$

Since (3.8) implies that  $\mathcal{W}(\mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)) \xrightarrow{d} |\gamma_1(\lambda)|$ , this proves (6.3).

Finally, by [11, Lemma 5.5], as  $n \rightarrow \infty$ ,

$$\sum_{v \in \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} \frac{(x_v^{(n)})^2}{\sigma_2(\mathbf{x}^{(n)})^2} \xrightarrow{d} \sum_{i \in \gamma_1(\lambda)} c_i^2. \quad (6.41)$$

Thus,

$$\begin{aligned} \mathbb{P}\left(\sum_{i \in \gamma_1(\lambda)} c_i^2 < \lambda(1 + 2\delta)\right) &\leq \liminf_n \mathbb{P}\left(\sum_{v \in \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} \frac{(x_v^{(n)})^2}{\sigma_2(\mathbf{x}^{(n)})^2} < \lambda(1 + 2\delta)\right) \\ &\leq \liminf_n \left[ \mathbb{P}\left(\sum_{v \in \mathcal{C}_\star^{(n)}} \frac{(x_v^{(n)})^2}{\sigma_2(\mathbf{x}^{(n)})^2} \leq \lambda(1 + 2\delta)\right) + \mathbb{P}(\mathcal{C}_\star^{(n)} \neq \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)) \right] \leq C\lambda^{-\frac{1}{1-2\alpha}}, \end{aligned}$$

where the last step follows from (6.35) and (6.40). This completes the proof of (6.4).  $\blacksquare$

**Proof of Proposition 6.2:** Write  $\sigma_2 = \sigma_2(\mathbf{x}^{(n)})$  for simplicity, and recall that  $t_\lambda = \lambda + 1/\sigma_2$ . Recall that  $\mathcal{C}_i(\mathbf{x}^{(n)}, t_\lambda)$  is the component of  $\mathcal{G}(\mathbf{x}^{(n)}, t_\lambda)$  with the  $i$ -th largest mass  $x_i^{(n)}(t_\lambda)$ .

Fix  $\varepsilon > 0$ . Using (3.8), (6.41), Theorem 6.1, and the asymptotic from Lemma 6.5 (a), we can choose  $\lambda(\varepsilon)$  positive and large such that  $\liminf_{n \rightarrow \infty} \mathbb{P}(\mathcal{B}_n(\lambda)) > 1 - \varepsilon$  for all  $\lambda \geq \lambda(\varepsilon)$ , where

$$\mathcal{B}_n(\lambda) := \left\{ x_1^{(n)}(t_\lambda) \geq C_{6.42} \lambda^{\frac{\alpha}{1-2\alpha}} \right\} \cap \left\{ \sum_{v \in \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} \frac{(x_v^{(n)})^2}{\sigma_2^2} \geq (1 + \delta)\lambda \right\}. \quad (6.42)$$

From now on we work with a fixed  $\lambda \geq \lambda(\varepsilon)$  and  $n$  large so that  $\mathbb{P}(\mathcal{B}_n(\lambda)) \geq 1 - \varepsilon$ .

We now reprise the argument leading to (6.39). Define

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}^{(n)} = (x_v, v \notin \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)), \text{ and } \hat{\sigma}_2 = \sigma_2(\hat{\mathbf{x}}^{(n)}) = \sigma_2 - \sum_{v \in \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} x_v^2.$$

We note here that  $\mathcal{B}_n(\lambda)$  is measurable w.r.t. the sigma-field generated by  $\hat{\mathbf{x}}$ . Further,  $x_1^{(n)}(t_\lambda) = \sigma_1(\mathbf{x}^{(n)}) - \sigma_1(\hat{\mathbf{x}}^{(n)})$ .

Define  $\lambda'$  by the identity  $\lambda + 1/\sigma_2 = -\lambda' + 1/\hat{\sigma}_2$ . Then on the event  $\mathcal{B}_n(\lambda)$ ,

$$\lambda' = \frac{\sigma_2 - \hat{\sigma}_2}{\sigma_2 \hat{\sigma}_2} - \lambda \geq \frac{(1 + \delta)\sigma_2^2 \lambda}{\sigma_2 \hat{\sigma}_2} - \lambda \geq \delta \lambda. \quad (6.43)$$

Write  $\hat{t}_{-\delta\lambda} = -\delta\lambda + 1/\hat{\sigma}_2$ . By (6.43), on the event  $\mathcal{B}_n(\lambda)$ ,

$$\hat{t}_{-\delta\lambda} \geq t_\lambda. \quad (6.44)$$

Note that

$$\begin{aligned} \mathbb{E}\left[\sigma_2(\hat{\mathbf{x}}(\hat{t}_{-\delta\lambda})) \mid \hat{\mathbf{x}}\right] &= \sum_{v \notin \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} x_v^{(n)} \mathbb{E}\left[\mathcal{W}(\mathcal{C}(\hat{\mathbf{x}}, \hat{t}_{-\delta\lambda}; v)) \mid \hat{\mathbf{x}}\right] \\ &\leq \sum_{v \notin \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)} x_v^{(n)} \frac{x_v^{(n)}}{1 - \hat{t}_{-\delta\lambda} \hat{\sigma}_2} = \frac{1}{\delta \lambda}, \end{aligned} \quad (6.45)$$

where the second step uses Lemma 4.5.

For  $z \geq 0$ , define the event  $F(z) = \{\hat{x}_1(t_\lambda) \leq z\}$ , i.e.,  $F(z)$  denotes the event that all components of  $\mathcal{G}(\hat{\mathbf{x}}, t_\lambda)$  have weight at most  $z$ . When  $\mathcal{B}_n(\lambda)$  occurs, by (6.44),  $\sigma_2(\hat{\mathbf{x}}(\hat{t}_{-\delta\lambda})) \geq (\hat{x}_1(t_\lambda))^2$ . Thus, on the event  $\mathcal{B}_n(\lambda)$ , for any  $z > 0$ ,

$$1 - \mathbb{P}(F(z) \mid \hat{\mathbf{x}}) \leq \mathbb{P}\left(\sigma_2(\hat{\mathbf{x}}(\hat{t}_{-\delta\lambda})) > z^2 \mid \hat{\mathbf{x}}\right) \leq \frac{1}{z^2 \delta \lambda}, \quad (6.46)$$

where the final step uses (6.45).

Now observe that, conditional on  $\hat{\mathbf{x}}$ , the graph  $\mathcal{G}(\mathbf{x}^{(n)}, t_\lambda) \setminus \mathcal{C}_1(\mathbf{x}^{(n)}, t_\lambda)$  has the same distribution as  $\mathcal{G}(\hat{\mathbf{x}}, t_\lambda)$  conditional on  $\hat{\mathbf{x}}$  and the event  $F(x_1^{(n)}(t_\lambda))$ . Hence, on the event  $\mathcal{B}_n(\lambda)$ ,

$$\begin{aligned} \mathbb{P}\left(\sum_{i \geq 2} x_i^{(n)}(t_\lambda)^2 \geq 1 \mid \hat{\mathbf{x}}\right) &= \mathbb{P}\left(\sigma_2(\hat{\mathbf{x}}(t_\lambda)) \geq 1 \mid \hat{\mathbf{x}}, F(x_1^{(n)}(t_\lambda))\right) \\ &\leq \frac{\mathbb{P}\left(\sigma_2(\hat{\mathbf{x}}(t_\lambda)) \geq 1 \mid \hat{\mathbf{x}}\right)}{\mathbb{P}\left(F(x_1^{(n)}(t_\lambda)) \mid \hat{\mathbf{x}}\right)} \leq \frac{\mathbb{P}\left(\sigma_2(\hat{\mathbf{x}}(\hat{t}_{-\delta\lambda})) \geq 1 \mid \hat{\mathbf{x}}\right)}{\mathbb{P}\left(F(C_{6.42} \lambda^{\frac{\alpha}{1-2\alpha}}) \mid \hat{\mathbf{x}}\right)} \leq \frac{1/(\delta\lambda)}{1 - 1/(\delta\lambda)}, \end{aligned}$$

where the third step uses (6.44), and the last step uses (6.46) and presumes that  $\lambda$  is large enough so that  $C_{6.42} \lambda^{\frac{\alpha}{1-2\alpha}} > 1$ . Hence,

$$\mathbb{P}\left(\sum_{i \geq 2} x_i^{(n)}(t_\lambda)^2 \geq 1\right) \leq \mathbb{P}(\mathcal{B}_n(\lambda)^c) + \frac{1/(\delta\lambda)}{1 - 1/(\delta\lambda)} \leq \varepsilon + \frac{1/(\delta\lambda)}{1 - 1/(\delta\lambda)}.$$

The rest is routine. ■

## 7. PROOF OF THEOREM 2.3

The proof relies on the following lemma.

**Lemma 7.1.** *Suppose  $\mathbf{y} = (y_i, i \geq 1) \in l_\downarrow^2$  with  $y_1 > y_2 > 0$ . Let  $y_1/y_2 = (1 + 4\delta)$ . Fix  $\eta > 0$  such that*

$$\eta \leq 1/10, \quad \eta \sigma_2(\mathbf{y}) \leq \min\{1, y_1^2\}/4, \quad \text{and} \quad 1 + \delta \leq (1 + 2\delta)(1 - \eta \sigma_2(\mathbf{y})). \quad (7.1)$$

*Suppose there exists  $k \geq 1$  such that*

$$\left(\sum_{j=1}^k y_j\right)^2 \geq \frac{\sigma_2(\mathbf{y})}{\eta} \quad \text{and} \quad \sum_{j \geq k+1} y_j^2 \leq \eta \sigma_2(\mathbf{y}). \quad (7.2)$$

*Then*

$$\begin{aligned} &\mathbb{P}(\text{a change of leader does not occur in } G(\mathbf{y}, \cdot)) \\ &\geq C_{7.3} \left(1 - \exp(-y_k^2)\right)^k \times \exp\left(- (1 + 2\delta) k^2 y_1^2\right) \times \left(\frac{\delta}{1 + \delta}\right)^k =: g(\mathbf{y}, k, \delta), \end{aligned} \quad (7.3)$$

where  $C_{7.3} > 0$  is an absolute constant.

We will make use of the following result in the proof of Lemma 7.1.

**Lemma 7.2** ([4], Lemma 20). *For  $\mathbf{z} \in l_\downarrow^2$  and  $t_0 \geq 0$ , for all  $s > \sigma_2(\mathbf{z})$ ,*

$$\mathbb{P}\left(\sigma_2(\mathbf{z}(t_0)) > s\right) \leq \frac{t_0 s \sigma_2(\mathbf{z})}{s - \sigma_2(\mathbf{z})}.$$

**Proof of Lemma 7.1:** Let us first describe the proof idea in words. Note that once the component of vertex 1 achieves an ' $l^2$ -lead' over the other components in the sense of Proposition 4.3, then it stays the leader with positive probability for the rest of the process. So it is enough to show that with positive probability the component of vertex 1 achieves this  $l^2$ -lead by some fixed time  $t_0 > 0$ , and it also stayed the leader in the time interval  $[0, t_0]$ . We will now make this argument precise. We can in fact take  $t_0 = 1$ , and will do so partway through the proof, but it is useful at the start to write  $t_0$  to avoid confusion between time 1 and the vertex label 1.

For  $A \subseteq \{1, 2, \dots\}$ ,  $\ell \in A$ , and  $\mathbf{u} = (y_i; i \in A)$ , write  $\mathcal{C}(\mathbf{u}, t_0; \ell)$  for the component of  $\ell$  in the subgraph of  $G(\mathbf{y}, t_0)$  that has  $A$  as its vertex set. (This notation agrees with the one introduced in Section 3.3. The point we emphasize here is that all such components are coupled in that they are all subgraphs of  $G(\mathbf{y}, t_0)$ .) Let

$$\mathbf{z}_2 = (y_i; i = 2 \text{ or } i \geq k + 1),$$

$$\mathbf{z}_j = (y_i; i \in \{j\} \cup \{i \geq k + 1 : i \notin \mathcal{C}(\mathbf{z}_\ell, t_0; \ell) \text{ for } 2 \leq \ell \leq j - 1\}), \quad j = 3, \dots, k, \text{ and}$$

$$\mathbf{z}_{k+1} = (y_i; i \geq k + 1 \text{ and } i \notin \mathcal{C}(\mathbf{z}_\ell, t_0; \ell) \text{ for } 2 \leq \ell \leq k).$$

Define the events

$$E_1 := \bigcap_{j=2}^k \{\text{there is an edge between 1 and } j \text{ in } G(\mathbf{y}, t_0)\},$$

$$E_2 := \{\mathcal{W}(\mathcal{C}(\mathbf{z}_2, t_0; 2)) \leq (1 + 2\delta)y_2\},$$

$$E_j := \left( \bigcap_{\ell=2}^{j-1} \{\text{no edge between } j \text{ and any vertex in } \mathcal{C}(\mathbf{z}_\ell, t_0; \ell)\} \right) \cap \{\mathcal{W}(\mathcal{C}(\mathbf{z}_j, t_0; j)) \leq (1 + 2\delta)y_j\}$$

for  $j = 3, \dots, k$ , and

$$E_{k+1} := \{\sigma_2(\mathbf{z}_{k+1}(t_0)) \leq 2\eta\sigma_2(\mathbf{y})\}.$$

Then

$$\mathbb{P}(E_2) = 1 - \mathbb{P}(E_2^c) \geq 1 - \left( \frac{1}{(1 + 2\delta)y_2} \right) \left( \frac{y_2}{1 - t_0 \sum_{j \geq k+1} y_j^2} \right) \geq \frac{\delta}{1 + \delta}, \quad (7.4)$$

where the second step uses Markov's inequality and the penultimate bound in (4.14), and the last step uses  $t_0 = 1$ , the second inequality in (7.2), and the third inequality in (7.1). Similarly, for  $j = 3, \dots, k$ ,

$$\mathbb{P}(E_j | E_2 \cap \dots \cap E_{j-1}) \geq \exp\left(-y_j \sum_{\ell=2}^{j-1} (1 + 2\delta)y_\ell\right) \times \left(\frac{\delta}{1 + \delta}\right) \geq \exp\left(-(1 + 2\delta)ky_1^2\right) \times \left(\frac{\delta}{1 + \delta}\right). \quad (7.5)$$

Using Lemma 7.2 with  $t_0 = 1$  and the second inequality in (7.2),

$$\mathbb{P}(E_{k+1} | E_2 \cap \dots \cap E_k) \geq 1 - \left( \frac{2\eta\sigma_2(\mathbf{y}) \cdot \eta\sigma_2(\mathbf{y})}{2\eta\sigma_2(\mathbf{y}) - \eta\sigma_2(\mathbf{y})} \right) = 1 - 2\eta\sigma_2(\mathbf{y}) \geq 1/2, \quad (7.6)$$

where the last step uses the second inequality in (7.1). Finally,  $E_1$  is independent of the collection of events  $\{E_2, \dots, E_{k+1}\}$ , and

$$\mathbb{P}(E_1) \geq (1 - \exp(-y_k^2))^k. \quad (7.7)$$

Combining (7.4), (7.5), (7.6), and (7.7), we get

$$\mathbb{P}(E_1 \cap \dots \cap E_{k+1}) \geq \frac{1}{2} \left(1 - \exp(-y_k^2)\right)^k \times \exp\left(- (1+2\delta)k^2 y_1^2\right) \times \left(\frac{\delta}{1+\delta}\right)^k. \quad (7.8)$$

Now, on the event  $E_1 \cap \dots \cap E_{k+1}$ , the component of 1 remains the unique leader all throughout the time interval  $[0, t_0]$ , and at time  $t_0$ , we have

$$\begin{aligned} \frac{(\mathcal{W}(\mathcal{C}(\mathbf{y}, t_0; 1)))^2}{\sigma_2(\mathbf{y}(t_0))} &\geq \frac{(\mathcal{W}(\mathcal{C}(\mathbf{y}, t_0; 1)))^2}{(\mathcal{W}(\mathcal{C}(\mathbf{y}, t_0; 1)))^2 + 2\eta\sigma_2(\mathbf{y})} \\ &\geq \frac{\sigma_2(\mathbf{y})/\eta}{\sigma_2(\mathbf{y})/\eta + 2\eta\sigma_2(\mathbf{y})} = \frac{1}{1+2\eta^2}, \end{aligned}$$

where the first step uses the definitions of the events  $E_1$  and  $E_{k+1}$ , and the second step uses the definition of the event  $E_1$ , and the first inequality in (7.2). Using Proposition 4.3,

$$\mathbb{P}(\text{the leader does not change in } (G(\mathbf{y}, t), t \geq t_0) \mid E_1 \cap \dots \cap E_{k+1}) \geq \frac{1-8\eta^2}{1+2\eta^2} \geq \frac{1-8/100}{1+2/100},$$

where the last step uses the fact  $\eta \leq 1/10$ . Combining this with (7.8) yields the claim.  $\blacksquare$

**Completing the proof of Theorem 2.3:** Consider  $\mathbf{x}^{(n)}$ ,  $n \geq 1$ , satisfying Condition II with  $c_1 > c_2$ . Let  $\theta \in (0, 1/3)$  be such that

$$c_1(1-\theta) > c_2(1+\theta). \quad (7.9)$$

Let  $\zeta(\cdot)$  be the eternal multiplicative coalescent that satisfies (3.8) for each fixed  $\lambda \in \mathbb{R}$ , and recall that  $\zeta(\lambda) = (|\gamma_1(\lambda)|, |\gamma_2(\lambda)|, \dots)$ .

Let  $\varepsilon = 1/10$ . By [5, Equation (65)], we can choose  $\lambda_\varepsilon > 0$  such that

$$\mathbb{P}\left(\forall \lambda \geq \lambda_\varepsilon, \lambda \cdot |\gamma_1(-\lambda)| \in (c_1(1-\theta), c_1(1+\theta)) \text{ and } \lambda \cdot |\gamma_2(-\lambda)| < c_2(1+\theta)\right) > 1 - \varepsilon. \quad (7.10)$$

Consider  $\lambda_\star > \lambda_\varepsilon$ . Let  $\lambda_1 > \lambda_2 > \dots > \lambda_\ell$  be such that  $\lambda_\star = \lambda_1$ ,  $\lambda_\varepsilon = \lambda_\ell$ ,

$$\frac{c_1(1-\theta)}{\lambda_j} > \frac{c_2(1+\theta)}{\lambda_{j+1}}, \text{ and } \frac{2(1-\theta)}{\lambda_j} > \frac{(1+\theta)}{\lambda_{j+1}} \text{ for } j = 1, \dots, \ell-1. \quad (7.11)$$

(It is possible to choose  $\lambda_1, \dots, \lambda_\ell$  like this because of the relations (7.9) and  $2(1-\theta) > (1+\theta)$ , the latter holding since  $\theta < 1/3$ .) By (3.8) and the Feller property of multiplicative coalescents [4, Proposition 5],  $(\mathbf{x}^{(n)}(t_{-\lambda_j}), 1 \leq j \leq \ell) \xrightarrow{d} (\zeta(-\lambda_j), 1 \leq j \leq \ell)$ , which combined with (7.10) implies that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\mathcal{A}_n) \geq 1 - \varepsilon, \quad (7.12)$$

where  $\mathcal{A}_n := \bigcap_{j=1}^{\ell} \left\{ \lambda_j x_1^{(n)}(t_{-\lambda_j}) \in (c_1(1-\theta), c_1(1+\theta)) \text{ and } \lambda_j x_2^{(n)}(t_{-\lambda_j}) < c_2(1+\theta) \right\}$ . Assume that  $\mathcal{A}_n$  holds, and observe that if for some  $\lambda \in [\lambda_2, \lambda_1]$ , in the random graph  $G(\mathbf{x}^{(n)}, t_{-\lambda})$ , the mass of a component other than the component containing the leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_1})$  becomes at least  $c_1(1-\theta)/\lambda_1$ , then in  $G(\mathbf{x}^{(n)}, t_{-\lambda_2})$ , either **(i)** that component will have merged with the component containing the leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_1})$ , or **(ii)** that component and the component containing the leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_1})$  will remain disjoint. However, (7.11) shows that under **(i)**,  $x_1^{(n)}(t_{-\lambda_2}) \geq 2c_1(1-\theta)/\lambda_1 > c_1(1+\theta)/\lambda_2$ , and under **(ii)**,  $x_2^{(n)}(t_{-\lambda_2}) \geq c_1(1-\theta)/\lambda_1 > c_2(1+\theta)/\lambda_2$ , leading to a contradiction. Proceeding inductively, we see that

on the event  $\mathcal{A}_n$ , throughout the time interval  $[t_{-\lambda_1}, t_{-\lambda_\ell}]$ , the component containing the leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_1})$  remains the unique leader. Hence, (7.12) implies that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\mathcal{B}_n) \geq 1 - \varepsilon, \quad (7.13)$$

where  $\mathcal{B}_n$  is the event that there is a unique leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_\star})$ , and the component containing the leader in  $G(\mathbf{x}^{(n)}, t_{-\lambda_\star})$  remains the unique leader throughout  $[t_{-\lambda_\star}, t_{-\lambda_\varepsilon}]$ . (Note that the lower bound in (7.13) holds for all  $\lambda_\star \in (\lambda_\varepsilon, \infty)$ .)

Since  $\mathbb{P}(|\gamma_2(-\lambda_\varepsilon)| > 0) = 1$ , using (3.8) we see that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(x_2^{(n)}(t_{-\lambda_\varepsilon}) > 0) = 1. \quad (7.14)$$

Let  $\delta_n$  satisfy  $x_1^{(n)}(t_{-\lambda_\varepsilon}) = (1 + 4\delta_n)x_2^{(n)}(t_{-\lambda_\varepsilon})$ . Since  $\mathbb{P}(|\gamma_1(-\lambda_\varepsilon)| > |\gamma_2(-\lambda_\varepsilon)|) = 1$  [21, Fact 1], (3.8) implies that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\delta_n > 0) = 1. \quad (7.15)$$

Let  $\eta_n$  be the maximum of all  $\eta$  that satisfy (7.1) when  $\mathbf{y}$  is replaced by  $\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon})$  and  $\delta$  is replaced by  $\delta_n$ . Since  $\sigma_2(\zeta(-\lambda_\varepsilon)) < \infty$  a.s. and  $|\gamma_1(-\lambda_\varepsilon)| > 0$  a.s., (3.8) and (7.15) imply that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\eta_n > 0) = 1. \quad (7.16)$$

Let  $K_n$  be the minimum of all  $k \geq 1$  that satisfy (7.2) when  $\mathbf{y}$  is replaced by  $\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon})$  and  $\eta$  is replaced by  $\eta_n$  with the convention that  $K_n = \infty$  if no such  $k$  exists. Since  $\zeta(-\lambda_\varepsilon) \in I_\downarrow^2 \setminus I_\downarrow^1$  a.s., (3.8) and (7.16) imply that we can choose  $k_0 \geq 1$  such that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(K_n \leq k_0) > 1 - \varepsilon. \quad (7.17)$$

Using (3.8), we see that

$$g(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}), k_0, \delta_n) \xrightarrow{d} g(\zeta(-\lambda_\varepsilon), k_0, \Delta), \quad \text{as } n \rightarrow \infty, \quad (7.18)$$

where  $g$  is as in (7.3), and  $\Delta$  satisfies  $|\gamma_1(-\lambda_\varepsilon)| = (1 + 4\Delta) \cdot |\gamma_2(-\lambda_\varepsilon)|$ . Since  $\Delta > 0$  a.s. and  $|\gamma_{k_0}(-\lambda_\varepsilon)| > 0$  a.s., the right side of (7.18) is positive a.s. Hence, we can choose  $\kappa > 0$  such that

$$\liminf_{n \rightarrow \infty} \mathbb{P}(g(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}), k_0, \delta_n) > \kappa) > 1 - \varepsilon. \quad (7.19)$$

Define  $\mathcal{B}'_n := \mathcal{B}_n \cap \{x_2^{(n)}(t_{-\lambda_\varepsilon}) > 0\} \cap \{K_n \leq k_0\} \cap \{g(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}), k_0, \delta_n) > \kappa\}$ . Combining (7.13), (7.14), (7.17), and (7.19), we get

$$\liminf_{n \rightarrow \infty} \mathbb{P}(\mathcal{B}'_n) \geq 1 - 3\varepsilon. \quad (7.20)$$

Now

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \mathbb{P}(L(\mathbf{x}^{(n)}) \leq t_{-\lambda_\star}) \\ & \geq \liminf_{n \rightarrow \infty} \mathbb{P}\left(\mathcal{B}'_n \cap \left\{ \text{a change of leader does not occur in } (G(\mathbf{x}^{(n)}, t), t \geq t_{-\lambda_\varepsilon}) \right\}\right) \\ & \geq \liminf_{n \rightarrow \infty} \mathbb{E}\left[\mathbb{1}_{\mathcal{B}'_n} \cdot g(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}), k_0, \delta_n)\right] \geq \kappa(1 - 3\varepsilon) = 7\kappa/10, \end{aligned} \quad (7.21)$$

where the second inequality uses Lemma 7.1 and the fact that  $K_n \leq k_0$  on the event  $\mathcal{B}'_n$ , and the third inequality uses (7.20) and the fact that  $g(\mathbf{x}^{(n)}(t_{-\lambda_\varepsilon}), k_0, \delta_n) > \kappa$  on the event  $\mathcal{B}'_n$ . Since the right side of (7.21) is free of  $\lambda_\star$ , letting  $\lambda_\star$  tend to infinity completes the proof.

## ACKNOWLEDGEMENTS

The authors thank James Martin and Bá-lazs Ráth for helpful discussions about the results of [33]. The authors also thank an anonymous referee for a careful reading of an earlier version of the paper and their detailed comments which led to significant improvements in the paper. LAB was partially supported by NSERC Discovery Grant 341845. SB was partially supported by NSF grants DMS-1613072, DMS-1606839 and ARO grant W911NF-17-1-0010. SS was partially supported by a CRM-ISM fellowship, MATRICS grant MTR/2019/000745 from SERB, and by the Infosys Foundation, Bangalore.

## REFERENCES

- [1] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probab. Theory Related Fields **152** (2012), no. 3-4, 367–406. MR2892951
- [2] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, Ann. Probab. **45** (2017), no. 5, 3075–3144. MR3706739
- [3] L. Addario-Berry and S. Sen, *Geometry of the minimal spanning tree of a random 3-regular graph*, arXiv preprint arXiv:1810.03802 (2018).
- [4] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Ann. Probab. **25** (1997), no. 2, 812–854. MR1434128 (98d:60019)
- [5] D. Aldous and V. Limic, *The entrance boundary of the multiplicative coalescent*, Electron. J. Probab. **3** (1998), No. 3, 59 pp. (electronic). MR1491528 (99d:60086)
- [6] 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 preprint arXiv:1411.3417 (2014).
- [7] S. Bhamidi, A. Budhiraja, and X. Wang, *The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs*, Probab. Theory Related Fields **160** (2014), no. 3-4, 733–796. MR3278920
- [8] S. Bhamidi, S. Dhara, R. van der Hofstad, and S. Sen, *Universality for critical heavy-tailed network models: metric structure of maximal components*, Electron. J. Probab. **25** (2020), no. 47, 1–57.
- [9] S. Bhamidi and S. Sen, *Geometry of the vacant set left by random walk on random graphs, Wright’s constants, and critical random graphs with prescribed degrees*, To appear in Random Structures Algorithms (2020+).
- [10] S. Bhamidi, S. Sen, and X. Wang, *Continuum limit of critical inhomogeneous random graphs*, Probab. Theory Related Fields **169** (2017), no. 1-2, 565–641. MR3704776
- [11] S. Bhamidi, R. van der Hofstad, and S. Sen, *The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs*, Probab. Theory Related Fields **170** (2018), no. 1-2, 387–474. MR3748328
- [12] 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** (2010), no. 54, 1682–1703. MR2735378
- [13] S. Bhamidi, R. van der Hofstad, and J. S. H. van Leeuwaarden, *Novel scaling limits for critical inhomogeneous random graphs*, Ann. Probab. **40** (2012), no. 6, 2299–2361. MR3050505
- [14] B. Bollobás, *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs*, European J. Combin. **1** (1980), no. 4, 311–316. MR595929
- [15] B. Bollobás, *Random graphs*, Second, Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001. MR1864966
- [16] B. Bollobás, S. Janson, and O. Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures Algorithms **31** (2007), no. 1, 3–122. MR2337396 (2008e:05124)
- [17] S. Boucheron, G. Lugosi, and P. Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR3185193
- [18] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley, *Optimal paths in disordered complex networks*, Physical review letters **91** (2003), no. 16, 168701.

- [19] L. A. Braunstein, Z. Wu, Y. Chen, S. V. Buldyrev, T. Kalisky, S. Sreenivasan, R. Cohen, E. Lopez, S. Havlin, and H. E. Stanley, *Optimal path and minimal spanning trees in random weighted networks*, International Journal of Bifurcation and Chaos **17** (2007), no. 07, 2215–2255.
- [20] S. Dhara, R. van der Hofstad, J. S. H. van Leeuwaarden, and S. Sen, *Critical window for the configuration model: finite third moment degrees*, Electron. J. Probab. **22** (2017), Paper No. 16, 33. MR3622886
- [21] S. Dhara, R. van der Hofstad, J. S. H. van Leeuwaarden, and S. Sen, *Heavy-tailed configuration models at criticality*, To appear in Ann. Inst. Henri Poincaré Probab. Stat. (2020+).
- [22] R. Durrett, *Probability—theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019. Fifth edition of [ MR1068527]. MR3930614
- [23] P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297. MR0120167
- [24] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61. MR0125031
- [25] R. van der Hofstad, *Critical behavior in inhomogeneous random graphs*, Random Structures Algorithms **42** (2013), no. 4, 480–508. MR3068034
- [26] R. van der Hofstad, S. Janson, and M. Łuczak, *Component structure of the configuration model: barely supercritical case*, Random Structures Algorithms **55** (2019), no. 1, 3–55. MR3974190
- [27] S. Janson and M. J. Łuczak, *Susceptibility in subcritical random graphs*, J. Math. Phys. **49** (2008), no. 12, 125207, 23. MR2484338
- [28] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. MR1782847
- [29] A. Joseph, *The component sizes of a critical random graph with given degree sequence*, Ann. Appl. Probab. **24** (2014), no. 6, 2560–2594. MR3262511
- [30] T. Klein and E. Rio, *Concentration around the mean for maxima of empirical processes*, Ann. Probab. **33** (2005), no. 3, 1060–1077. MR2135312 (2006c:60022)
- [31] T. Łuczak, *Component behavior near the critical point of the random graph process*, Random Structures Algorithms **1** (1990), no. 3, 287–310. MR1099794
- [32] T. Łuczak, B. Pittel, and J. C. Wierman, *The structure of a random graph at the point of the phase transition*, Trans. Amer. Math. Soc. **341** (1994), no. 2, 721–748. MR1138950
- [33] J. B. Martin and B. Ráth, *Rigid representations of the multiplicative coalescent with linear deletion*, Electron. J. Probab. **22** (2017), Paper No. 83, 47. MR3718711
- [34] M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, “Random Graphs ’93” (Poznań, 1993), 1995, pp. 161–179. MR1370952
- [35] O. Riordan, *The phase transition in the configuration model*, Combin. Probab. Comput. **21** (2012), no. 1-2, 265–299. MR2900063
- [36] J. Spencer and N. Wormald, *Birth control for giants*, Combinatorica **27** (2007), no. 5, 587–628. MR2375718