CONTINUUM LIMIT OF CRITICAL INHOMOGENEOUS RANDOM GRAPHS

SHANKAR BHAMIDI[1], SANCHAYAN SEN[2], AND XUAN WANG[3]

ABSTRACT. Motivated by applications, the last few years have witnessed tremendous interest in understanding the structure as well as the behavior of dynamics for inhomogeneous random graph models. In this study we analyze the maximal components at criticality of one famous class of such models, the rank-one inhomogeneous random graph model [37], [13, Section 16.4]. Viewing these components as measured random metric spaces, under finite moment assumptions for the weight distribution, we show that the components in the critical scaling window with distances scaled by $n^{-1/3}$ converge in the Gromov-Haussdorf-Prokhorov metric to rescaled versions of the limit objects identified for the Erdős-Rényi random graph components at criticality in [3]. A key step is the construction of connected components of the random graph through an appropriate tilt of a famous class of random trees called **p**-trees [8, 19]. This is the first step in rigorously understanding the scaling limits of objects such as the Minimal spanning tree and other strong disorder models from statistical physics [15] for such graph models. By asymptotic equivalence [28], the same results are true for the Chung-Lu model [20–22] and the Britton-Deijfen-Lof model [17]. A crucial ingredient of the proof of independent interest is tail bounds for the height of **p**-trees. The techniques developed in this paper form the main technical bedrock for proving continuum scaling limits in the critical regime for a wide array of other random graph models in [9] including the configuration model and inhomogeneous random graphs with general kernels [13].

1. INTRODUCTION

Motivated by applications and empirical data, the last few years have seen tremendous interest in formulating and studying inhomogeneous random graph models, estimating the parameters in the model from data, and studying dynamic processes such as epidemics on such models, see e.g. [6, 13, 22, 23, 36, 44] and the references therein. In such random graph models different vertices have different propensities for connecting to other vertices. To fix ideas consider the main model analyzed in this work:

Rank one model: This version of the model was introduced by Norros and Reittu [13, 37] (with a variant arising in the work of Aldous in the construction of the standard multiplicative coalescent [7]). We construct a random graph on the vertex set $[n] = \{1, 2, \ldots, n\}$ as follows. Each vertex $i \in [n]$ has an associated weight $w_i \geqslant 0$. Think of this as the propensity of the vertex to form friendships (form edges) in a network. Write $\mathbf{w} = (w_i)_{i\in[n]}$ for the vector of weights and let $l_n = \sum_{i=1}^{n} w_i$ be the sum of these weights. The weights actually form a triangular array

[1]DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH, 304 HANES HALL, UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL, NC 27599

[2]DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, EINDHOVEN UNIVERSITY OF TECHNOLOGY, THE NETHERLANDS

[3]DEPARTMENT OF MATHEMATICS, 307 RAWLES HALL, INDIANA UNIVERSITY, BLOOMINGTON, INDIANA

E-mail addresses: bhamidi@email.unc.edu, s.sen@tue.nl, xw45@indiana.edu.

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

Key words and phrases. Multiplicative coalescent, continuum random tree, critical random graphs, branching processes, **p**-trees, scaling limits.

$\mathbf{w} = \mathbf{w}^{(n)} = (w_i^{(n)} : i \in [n])$, but we will omit n in the notation. Taking the weight sequence $\mathbf{w}$ as input, the random graph is constructed as follows. Define probabilities $p_{ij} := 1 - \exp(-w_i w_j / l_n)$. Construct the random graph $\mathcal{G}_n^{\mathrm{nr}}(\mathbf{w})$ by putting an edge $\{i, j\}$ between vertices i, j with probability p_{ij}, independent across edges.

This is an important example of the general class of inhomogeneous random graphs analyzed by Bollobas, Janson and Riordan [13]. This model is also closely related to two other famous models of inhomogeneous random graphs (and in fact shown to be asymptotically equivalent in a number of settings [28])

(a) **Chung-Lu model** [20–22]**:** Given the set of weights $\mathbf{w}$ as above, here one attaches edges independently with probability

$$p_{ij} := \max\left\{\frac{w_i w_j}{l_n}, 1\right\}$$

(b) **Britton-Deijfen-Lof model** [17]**:** Here one attaches edges independently with probability

$$p_{ij} := \frac{w_i w_j}{l_n + w_i w_j}$$

These models are inhomogeneous in the sense that different vertices have different proclivity to form edges. Further assume that that the empirical distribution of weights $F_n = n^{-1} \sum_{i=1}^n \delta_{w_i}$ satisfies

$$F_n \xrightarrow{\mathrm{w}} F, \qquad \text{as } n \to \infty, \tag{1.1}$$

for a limiting cumulative distribution function F. Then by [37, Theorem 3.13] as $n \to \infty$, the degree distribution converges in the sense that for $k \geqslant 0$, writing $N_k(n)$ for the number of vertices with degree k,

$$\frac{N_k(n)}{n} \xrightarrow{\mathrm{P}} \mathbb{E}\left(e^{-W} \frac{W^k}{k!}\right), \qquad k \geqslant 0,$$

where $W \sim F$. Thus asymptotically one can get any desired tail behavior for the degree distribution by choosing the weight sequence appropriately. Also note that the Erdős-Rényi random graph $\mathcal{G}^{\mathrm{er}}(n, \lambda/n)$ is a special case where all the weights $w_i \equiv \lambda$.

Aside from applications, inhomogeneous random graph models have sparked a lot of interest in the statistical physics community, in particular understanding how the network structure affects weak and strong disorder models of flow, e.g. first passage percolation, minimal spanning tree etc. In the next section we describe what is known about how the network transitions from the subcritical to the supercritical regime and then describe these conjectures from statistical physics in more detail.

1.1. **Connectivity and phase transition.** The main aim of this study is the structural properties of the maximal components in the critical regime. In order to define the critical regime, we first recall known connectivity properties of the model. Let $W \sim F$ where F as before denotes the limiting weight distribution as in 1.1. Assume $0 < \mathbb{E}(W^2) < \infty$ and assume

$$\frac{\sum_{i=1}^n w_i^2}{\sum_{i=1}^n w_i} \to \frac{\mathbb{E}(W^2)}{\mathbb{E}(W)}, \qquad \text{as } n \to \infty. \tag{1.2}$$

Define the parameter

$$\nu = \frac{\mathbb{E}(W^2)}{\mathbb{E}(W)}. \tag{1.3}$$

Write $\mathscr{C}_n^{(i)}$ for the i-th largest component of $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w})$ (breaking ties arbitrarily) and write $|\mathscr{C}_n^{(i)}|$ for the number of vertices in this component. Then by [13, Theorem 3.1 and Section 16.4] (also see [17, 21, 37])

(i) **Subcritical regime:** If $\nu < 1$, then $|\mathscr{C}_n^{(1)}| = o_P(n)$.

(ii) **Supercritical regime:** If $\nu > 1$, then

$$\frac{|\mathscr{C}_n^{(1)}|}{n} \xrightarrow{\mathrm{P}} \rho(\nu, F) > 0,$$

where $\rho(\nu, F)$ is the survival probability of an associated multi-type branching process.

(iii) **Critical regime:** Thus $\nu = 1$ corresponds to the critical regime. This is the regime of interest for this paper. Similar to the Erdős-Rényi random graph, we will study the entire critical scaling window, by working with weights $(1+\lambda/n^{1/3})w_i$ so that the connection probability is

$$p_{ij}(\lambda) := 1 - \exp\left(-\left(1+\frac{\lambda}{n^{1/3}}\right)\frac{w_i w_j}{l_n}\right) \tag{1.4}$$

We write $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$ for the corresponding random graph on the vertex set $[n]$. In this critical regime (with $\nu = 1$), assuming finite third moments $\mathbb{E}(W^3) < \infty$, it is known [11, 27] that for any fixed i, the i-th largest component scales like $|\mathscr{C}_n^{(i)}(\lambda)| \sim n^{2/3}$. We shall give a precise description of this result in Section 4.

1.2. **Motivation and outline.** Let us now informally describe the motivations behind our work. The main aim of this paper is to study the maximal components $\mathscr{C}_n^{(i)}(\lambda)$ for the rank-one model above in the critical scaling window and show: *these maximal components viewed as metric spaces with edge lengths rescaled by $n^{-1/3}$ converge to random fractals related to the continuum random tree*. A natural question is why one should focus on this particular class of random graph models.

(a) **Universality for random graph processes at criticality:** The nature of emergence and scaling limits of **component sizes** of maximal components in the critical regime of the Norros-Reittu model (and the closely related multiplicative coalescent) have recently been observed in a number of other random graph models including the configuration model [30, 35] as well as a general class of dynamic random graph processes called Bounded size rules (see e.g. [10, 42]). The first step in understanding the metric structure of the maximal components for these models in the critical regime is the rank-one model. This paper forms the main technical bedrock for proving continuum scaling limits in the critical regime for a wide array of other random graph models in [9] including the configuration model and inhomogeneous random graphs with general kernels [13]. Using the technical tools developed in this paper, in particular Section 7.1, we paraphrase the main theme of [9] as follows:

The metric structure of the maximal components in the critical regime for a number of random graph models, including the configuration model under moment conditions, and inhomogeneous random graph models with finite state space, satisfy analogous results to Theorem 3.3.

Thus it is not just the main results in this paper that are of interest, rather the **proof section** of this paper sets out the tools required to prove this general program of universality.

(b) **Scaling limits at criticality and the Minimal spanning tree:** Our second main motivation was to rigorously understand predictions from statistical physics (see e.g.[16] and the references therein) which predict that most inhomogeneous random graph models in the critical regime satisfy a remarkably universal behavior in the following sense. Assume that the limiting degree distribution has finite third moments, then distances in the maximal components in the critical regime scale like $n^{1/3}$. Further consider the minimal spanning tree on the giant component in the supercritical regime where each edge is assigned a $U[0,1]$ edge length. It is conjectured that (graph) distances in this object also scale like $n^{1/3}$. For the case of the Erdős-Rényi random graph, this entire program has been carried forth in [4]. Proving this conjecture for general inhomogenous random graphs rigorously turns out to be technically quite challenging and in particular require a number of non-obvious assumptions, see Assumptions 3.1(d). To strengthen convergence in the stronger l^4 metric we needed to derive tail bounds for the height of **p**-trees (Theorem 3.7) which are of independent interest.

The rest of the paper is organized as follows. In Section 2, we start with the appropriate spaces and topology for convergence of a collection of metric spaces and define the Gromov-Haussdorf-Prokhorov metric. We state our main results in Section 3. We give a precise description of the limit objects in Section 4. We discuss our main results and their relevance as well as the technical challenges in extending these results in Section 5. Starting from Section 6 we prove the main results.

2. NOTATION AND PRELIMINARIES

We introduce some basic notation in Section 2.1. In Section 2.2 we define relevant notions of convergence of measured metric spaces. In Section 2.3 we recall basic graph theoretic definitions. In Section 2.4, we introduce a family of random trees called **p**-trees that play a crucial role in our proofs.

2.1. Notations and conventions.

For simplicity, we will often write "n^α terms" or "$\sum_{k=1}^{n/2}$" when really it should be "$\lfloor n^\alpha \rfloor$ terms" or "$\lceil n^\alpha \rceil$ terms" or "$\sum_{k=1}^{\lfloor n/2 \rfloor}$" or "$\sum_{k=1}^{\lceil n/2 \rceil}$" as appropriate but this will not affect the proofs. We use notations such as $K_{3.7}$ and $K_{7.13}$ to denote absolute constants. The subscripts indicate the theorems or lemmas where the constants are first introduced. For example, $K_{3.7}$ is the constant in Theorem 3.7 and $K_{7.13}$ is the constant in Corollary 7.13. Local constants are denoted by $C_1, C_2, \cdots$ or $B_1, B_2, \cdots$. We use the standard Landau notation of $o(\cdot)$, $O_{\mathbb{P}}(\cdot)$ and so forth.

2.2. Topology on the space of measured metric spaces.

We mainly follow [1,4,18]. All metric spaces under consideration will be measured compact metric spaces. Let us recall the Gromov-Haussdorf distance d_{GH} between metric spaces. Fix two metric spaces $X_1 = (X_1, d_1)$ and $X_2 = (X_2, d_2)$. For a subset $C \subseteq X_1 \times X_2$, the distortion of C is defined as

$$\operatorname{dis}(C) := \sup\left\{|d_1(x_1, y_1) - d_2(x_2, y_2)| : (x_1, x_2), (y_1, y_2) \in C\right\}. \tag{2.1}$$

A correspondence C between X_1 and X_2 is a measurable subset of $X_1 \times X_2$ such that for every $x_1 \in X_1$ there exists at least one $x_2 \in X_2$ such that $(x_1, x_2) \in C$ and vice-versa. The Gromov-Haussdorf distance between the two metric spaces (X_1, d_1) and (X_2, d_2) is defined as

$$d_{\mathrm{GH}}(X_1, X_2) = \frac{1}{2} \inf\left\{\operatorname{dis}(C) : C \text{ is a correspondence between } X_1 \text{ and } X_2\right\}. \tag{2.2}$$

We will need a metric that also keeps track of associated measures on the corresponding spaces. A compact measured metric space (X,d,μ) is a compact metric space (X,d) with an associated finite measure μ on the Borel sigma algebra $\mathscr{B}(X)$. Given two compact measured metric spaces (X_1,d_1,μ_1) and (X_2,d_2,μ_2) and a measure π on the product space $X_1\times X_2$, the discrepancy of π with respect to μ_1 and μ_2 is defined as

$$D(\pi;\mu_1,\mu_2) := ||\mu_1-\pi_1||+||\mu_2-\pi_2|| \tag{2.3}$$

where π_1,π_2 are the marginals of π and $||\cdot||$ denotes the total variation of signed measures. Then the Gromov-Haussdorf-Prokhorov distance between X_1 and X_2 is defined

$$d_{\mathrm{GHP}}(X_1,X_2) := \inf\left\{\max\left(\frac{1}{2}\operatorname{dis}(C),\ D(\pi;\mu_1,\mu_2),\ \pi(C^c)\right)\right\}, \tag{2.4}$$

where the infimum is taken over all correspondences C and measures π on $X_1\times X_2$.

Denote by $\mathscr{S}$ the collection of all measured metric spaces (X,d,μ). The function d_{GHP} is a pseudometric on $\mathscr{S}$, and defines an equivalence relation $X\sim Y \Leftrightarrow d_{\mathrm{GHP}}(X,Y)=0$ on $\mathscr{S}$. Let $\bar{\mathscr{S}} := \mathscr{S}/\sim$ be the space of isometry equivalent classes of measured compact metric spaces and $\bar{d}_{\mathrm{GHP}}$ be the induced metric. Then by [1], $(\bar{\mathscr{S}},\bar{d}_{\mathrm{GHP}})$ is a complete separable metric space. To ease notation, we will continue to use $(\mathscr{S},d_{\mathrm{GHP}})$ instead of $(\bar{\mathscr{S}},\bar{d}_{\mathrm{GHP}})$ and $X=(X,d,\mu)$ to denote both the metric space and the corresponding equivalence class.

Since we will be interested in not just one metric space but an infinite sequence of metric spaces, the relevant space of interest is a subset of $\mathscr{S}^{\mathbb{N}}$. For a fixed measured compact metric space (X,d,μ), define the diameter as $\operatorname{diam}(X) := \sup_{x,y\in X} d(x,y)$ and the total mass as $\operatorname{mass}(X) := \mu(X)$. Then the relevant space for our study will be

$$\mathscr{M} := \left\{(X_1,X_2,\ldots) : X_i=(X_i,d_i,\mu_i)\in\mathscr{S},\ \sum_{i=1}^{\infty}(\operatorname{diam}(X_i)^4+\operatorname{mass}(X_i)^4)<\infty\right\}. \tag{2.5}$$

The two relevant topologies on this space are

(i) **Product topology:** We shall denote the product topology inherited by d_{GHP} on a single co-ordinate by $\mathscr{T}_1$.

(ii) l^4 **metric** [3]**:** We shall let $\mathscr{T}_2$ denote the topology on $\mathscr{M}$ induced by the distance

$$\operatorname{dist}((X_1,X_2,\ldots),(X_1',X_2',\ldots)) := \left(\sum_{i=1}^{\infty} d_{\mathrm{GHP}}(X_i,X_i')^4\right)^{1/4}. \tag{2.6}$$

The aim of this paper is to study the limits of connected components of random graphs viewed as measured metric spaces. In order to state our results, both the metric and the corresponding measure need to be re-scaled appropriately. To make this precise, we introduce the scaling operator $\operatorname{scl}(\alpha,\beta)$, for $\alpha,\beta\in(0,\infty)$, as follows:

$$\operatorname{scl}(\alpha,\beta):\mathscr{S}\to\mathscr{S}, \qquad \operatorname{scl}(\alpha,\beta)[(X,d,\mu)] := (X,d',\mu'),$$

where $d'(x,y) := \alpha d(x,y)$ for all $x,y\in X$, and $\mu'(A) := \beta\mu(A)$ for $A\subset X$. For simplicity, we write the output of the above scaling operator as $\operatorname{scl}(\alpha,\beta)X$. Using the definition of d_{GHP}, it is easy to check that for $X\in\mathscr{S}$ and for fixed $\alpha,\beta>0$,

$$d_{\mathrm{GHP}}(\operatorname{scl}(\alpha,\beta)X,X) \leqslant |\alpha-1|\cdot\operatorname{diam}(X)+|\beta-1|\cdot\operatorname{mass}(X).$$

Note that $\operatorname{diam}(\cdot)$ and $\operatorname{mass}(\cdot)$ are both continuous functions on $(\mathscr{S},d_{\mathrm{GHP}})$. Using this fact and the above bound we have the following easy proposition.

Proposition 2.1. *Let* $\{\alpha_n : n \geqslant 1\}$ *and* $\{\beta_n : n \geqslant 1\}$ *be two sequences of positive numbers. Further assume* $\lim_{n\to\infty} \alpha_n = \alpha > 0$ *and* $\lim_{n\to\infty} \beta_n = \beta > 0$. *Let* $\{X_n : n \geqslant 1\} \subset \mathscr{S}$ *be a sequence of metric spaces such that* $X_n \to X \in \mathscr{S}$ *in the metric* d_{GHP} *as* $n \to \infty$. *Then we have*

$$\mathrm{scl}(\alpha_n, \beta_n) X_n \to \mathrm{scl}(\alpha, \beta) X, \text{ in } d_{\mathrm{GHP}} \text{ as } n \to \infty.$$

As in the above proposition and the rest of this paper, we will always equip $\mathscr{S}$ with the topology generated by d_{GHP}.

2.3. **Graphs, trees and ordered trees.** All graphs in this study will be simple undirected graphs $\mathscr{G}$. We will typically write $V(\mathscr{G})$ for the vertex set of $\mathscr{G}$ and $E(\mathscr{G})$ for the corresponding edge set. We will write an edge as $e = (u, v) \in E(\mathscr{G})$ with the understanding that (u, v) represents an undirected edge and is equivalent to (v, u). As before we write $[n] = \{1, 2, \ldots, n\}$. We will typically denote a connected component of $\mathscr{G}$ by $\mathscr{C} \subseteq \mathscr{G}$. A connected component $\mathscr{C}$, can be viewed as a metric space by imposing the usual graph distance d_{G} namely

$$d_{\mathrm{G}}(v, u) = \text{ number of edges on the shortest path between } v \text{ and } u, \qquad u, v \in \mathscr{C}.$$

Recall that to construct the random graph, we started with a collection of vertex weights $\{w_i : i \in [n]\}$. Thus there are two natural measures for a connected graph $\mathscr{G}$ with associated vertex weights $\mathbf{w} := \{w_v : v \in \mathscr{G}\}$:

(i) **Counting measure:** $\mu_{\mathrm{ct}}(A) := |A|$, for $A \subset V(\mathscr{G})$.
(ii) **Weighted measure:** $\mu_{\mathbf{w}}(A) := \sum_{v\in A} w_v$, for $A \subset V(\mathscr{G})$. If no weights are specified then the default convention is to take $w_v \equiv 1$ for all v thus resulting in $\mu_{\mathbf{w}} = \mu_{\mathrm{ct}}$.

For a fixed finite connected graph $\mathscr{G}$ equipped with vertex weights $\{w_v : v \in \mathscr{G}\}$ (by convention $w_v = 1$ if not pre-specified), one obtains a measured metric space $(V(\mathscr{G}), d_{\mathrm{G}}, \mu)$, where μ is either μ_{ct} or $\mu_{\mathbf{w}}$. For $\mathscr{G}$ finite and connected, the corresponding metric space is compact with finite measure. We use $\mathscr{G}$ for both the graph and the corresponding measured metric space.

A tree $\mathbf{t}$ is a connected graph with no cycle. A *rooted* tree is a pair $(\mathbf{t}, r)$ where $\mathbf{t}$ is a tree and $r \in V(\mathbf{t})$ is a distinguished vertex referred to as the root. All trees in the sequel will be rooted trees. Thinking of r as the original progenitor of a genealogy, for vertices in the tree the notions *parent, children, ancestors, siblings, generations* and *heights* have their usual interpretation. An *ordered* tree is a rooted tree in which an order is specified amongst the children of each vertex so that one can talk about the first child, the second child etc. Such trees will be represented as $(\mathbf{t}, \boldsymbol{\pi})$, where $\mathbf{t}$ is a rooted tree and $\boldsymbol{\pi}$ is the corresponding order. Such trees are also referred to as planar trees as they can be embedded on the plane, arranging children of each vertex from left to right in increasing value of the order.

For $n \in \mathbb{N}$, write $\mathbb{G}_n$, $\mathbb{G}_n^{\mathrm{con}}$, $\mathbb{T}_n$ and $\mathbb{T}_n^{\mathrm{ord}}$ for the collection of all graphs, connected graphs, rooted trees and ordered trees, respectively, with vertex set $[n]$. For ease of notation, we suppress $\boldsymbol{\pi}$ in the pair $(\mathbf{t}, \boldsymbol{\pi})$ for ordered trees and just write $\mathbf{t} \in \mathbb{T}_n^{\mathrm{ord}}$. Planar trees can be treated as connected graphs by forgetting about the root and order. Therefore all notations for graphs apply to rooted ordered trees as well. In particular, any tree $\mathbf{t}$ can be viewed as a measured metric space in $\mathscr{S}$.

2.4. **p-Trees.** In this section, we define a family of random tree models called $\mathbf{p}$-trees, which play a key role in the proof. See [38] for a comprehensive survey including their role in linking combinatorial objects such as the Abel-Cayley-Hurwitz multinomial expansions with probability. Fix $m \geqslant 1$, and a probability mass function $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ with $p_i > 0$ for all $i \in [m]$. A $\mathbf{p}$-tree is a random tree in $\mathbb{T}_m$, with law as follows. For any fixed $\mathbf{t} \in \mathbb{T}_m$ and $v \in \mathbf{t}$, write $d_v(\mathbf{t})$, for the number of children of v in the tree $\mathbf{t}$. Then the law of $\mathbf{p}$-tree, denoted by $\mathbb{P}_{\text{tree}}$, is defined as:

$$\mathbb{P}_{\text{tree}}(\mathbf{t}) = \mathbb{P}_{\text{tree}}(\mathbf{t};\mathbf{p}) = \prod_{v\in[m]} p_v^{d_v(\mathbf{t})}, \qquad \mathbf{t} \in \mathbb{T}_m. \tag{2.7}$$

Generating a random $\mathbf{p}$-tree $\mathcal{T}$ and then assigning a uniform random order on the children of every vertex $v \in \mathcal{T}$ gives a random element with law $\mathbb{P}_{\text{ord}}(\cdot;\mathbf{p})$ given by

$$\mathbb{P}_{\text{ord}}(\mathbf{t}) = \mathbb{P}_{\text{ord}}(\mathbf{t};\mathbf{p}) = \prod_{v\in[m]} \frac{p_v^{d_v(\mathbf{t})}}{(d_v(\mathbf{t}))!}, \qquad \mathbf{t} \in \mathbb{T}_m^{\text{ord}}. \tag{2.8}$$

Obviously a $\mathbf{p}$-tree can be constructed by first generating an ordered $\mathbf{p}$-tree with the above distribution and then forgetting about the order.

3. RESULTS

We are now in a position to describe our main results. Section 3.1 describes our main results for the Norros-Reittu model(and the associated Chung-Lu model and Britton-Deijfen-Lof model) in the critical regime. In Section 3.2 we describe tail bounds on the diameter of $\mathbf{p}$-trees which play a crucial role in the proof of convergence in the l^4 metric.

3.1. Scaling limits for the Norros-Reittu random graph at criticality.

We start by stating the assumptions on the weight sequence $\mathbf{w}$ used to construct the random graph $\mathcal{G}_n^{\text{nr}}(\mathbf{w},\lambda)$. Note that through out $w_i = w_i(n)$ might depend on n but we suppress this dependence. Define

$$\sigma_k(n) := n^{-1}\sum_{i=1}^{n} w_i^k \text{ for } k = 1,2,3$$
$$w_{\max} = \max_{i\in[n]} w_i \text{ and } w_{\min} = \min_{i\in[n]} w_i.$$

We make the following assumptions on the asymptotic behavior of the weight sequence.

Assumption 3.1.

(a) **Convergence of three moments:** *There exist constants* $\sigma_k > 0$ *for* $k = 1,2,3$ *such that*
$$\max\left\{n^{1/3}|\sigma_1(n) - \sigma_1|,\ n^{1/3}|\sigma_2(n) - \sigma_2|\ , |\sigma_3(n) - \sigma_3|\right\} \to 0 \textit{ as } n \to \infty.$$
(b) **Critical regime:** $\sigma_1 = \sigma_2$.
(c) **Bound on the maximum**: *There exists* $\eta_0 \in (0,1/6)$ *such that* $w_{\max} = o(n^{1/6-\eta_0})$.
(d) **Bound on the minimum**: *There exists* $\gamma_0 > 0$ *such that* $1/w_{\min} = o(n^{\gamma_0})$. *Thus minimal weights decrease at most polynomially quickly to zero.*

For convergence in $\mathcal{T}_2$ topology, we will need the following stronger assumption on $w_{\max}$.

Assumption 3.2. *There exists* $\eta_0 \in (0,1/48)$ *such that* $w_{\max} = o(n^{1/48-\eta_0})$.

For fixed $\lambda \in \mathbb{R}$, recall the Norros-Reittu random graph $\mathcal{G}_n^{\text{nr}}(\mathbf{w},\lambda)$ defined below (1.4). For $i \geqslant 1$, let $\mathcal{C}_n^{(i)}(\lambda)$, denote the i-th largest component of $\mathcal{G}_n^{\text{nr}}(\mathbf{w},\lambda)$. As described in Section 1.1, the number of vertices in $\mathcal{C}_n^{(i)}$ is of order $n^{2/3}$ (we describe the precise limit result in Section 4). Equipping $\mathcal{C}_n^{(i)}(\lambda)$ with the graph distance metric and assigning weight w_v for each vertex v, we view each of these components as measured metric spaces (see Section 2.3). Our main result is about the limit of the scaled metric spaces defined as

$$\mathbf{M}^n(\lambda) := (\text{scl}(n^{-1/3}, n^{-2/3}) \cdot \mathcal{C}_n^{(i)}(\lambda) : i \geqslant 1),$$

namely, rescaling graph distance by $n^{-1/3}$ and each of the weights by $n^{-2/3}$.

Theorem 3.3. *Fix $\lambda \in \mathbb{R}$. Consider the Norros-Reittu model $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$ with weight sequence $\mathbf{w}$.*

(i) *Under Assumption* 3.1*, then as $n \to \infty$,*

$$\mathbf{M}^n(\lambda) \xrightarrow{\mathrm{w}} \mathbf{M}(\lambda) \tag{3.1}$$

where $\mathbf{M}(\lambda) = (M_i(\lambda) : i \geqslant 1)$ is an $\mathscr{M}$-valued random variable and the convergence takes place with respect to $\mathscr{T}_1$ topology. The construction of $\mathbf{M}(\lambda)$ is given in Section 4*.*

(ii) *Under the additional Assumption* 3.2*, the above convergence takes place with respect to $\mathscr{T}_2$ topology as in* (2.6)*.*

Remark 1. Write $\mathscr{G}^{\mathrm{er}}(n,p)$ for the Erdős-Rényi random graph with vertex set $[n]$ and connection probability p. The critical scaling window corresponds to $p = 1/n + \lambda/n^{4/3}$ with $\lambda \in \mathbb{R}$ ([12, 24, 29, 32, 33]). Write $\mathscr{C}_n^{(i),\mathrm{er}}(\lambda)$ for the i-th largest component of $\mathscr{G}^{\mathrm{er}}(n, 1/n + \lambda/n^{4/3})$, and treat them as measured metric spaces. Define $\mathbf{M}^{n,\mathrm{er}}(\lambda)$ as

$$\mathbf{M}^{n,\mathrm{er}}(\lambda) := \left(\mathrm{scl}(n^{-1/3}, n^{-2/3}) \cdot \mathscr{C}_n^{(i),\mathrm{er}}(\lambda) : i \geqslant 1\right).$$

Building on the analysis by Aldous [7] on the size and surplus of components at criticality, [3, Theorem 2] and [4, Section 4] showed that, as $n \to \infty$,

$$\mathbf{M}^{n,\mathrm{er}}(\lambda) \xrightarrow{\mathrm{w}} \mathbf{M}^{\mathrm{er}}(\lambda) = (M_i^{\mathrm{er}}(\lambda) : i \geqslant 1) \in \mathscr{M}, \tag{3.2}$$

where $\mathbf{M}^{\mathrm{er}}(\lambda)$ is described in detail in Section 4, with convergence is with respect to the $\mathscr{T}_2$ topology. It will be shown in Lemma 4.2 that the limiting metric spaces in Theorem 3.3 satisfy

$$\mathbf{M}(\lambda) \stackrel{d}{=} \mathrm{scl}\left(\frac{\mu}{\sigma_3^{2/3}}, \frac{\mu}{\sigma_3^{1/3}}\right) \cdot \mathbf{M}^{\mathrm{er}}(\lambda\mu/\sigma_3^{2/3}).$$

This actually shows that under the assumptions of Theorem 3.3, critical rank-one inhomogeneous random graphs viewed as metric spaces belong to the Erdős-Rényi universality class.

The following Corollary gives a simple choice of weights which satisfy the relevant assumptions.

Corollary 3.4. *Let $\{w_i : i \in [n]\}$ be iid copies of a strictly positive random variable W that satisfies*

$$\mathbb{E}(W) = \mathbb{E}(W^2), \quad \lim_{x \downarrow 0} x^{-\epsilon}\, \mathbb{P}(W \leq x) = 0 \textit{ for some } \epsilon > 0.$$

Conditional on the weights $\mathbf{w} = \{w_i : i \in [n]\}$ construct $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$ as above.

(i) *Assume $\mathbb{E}\, W^{6+\epsilon} < \infty$, then the convergence in* (3.1) *holds in $\mathscr{T}_1$ topology.*

(ii) *Assume $\mathbb{E}\, W^{48+\epsilon} < \infty$, then the convergence in* (3.1) *holds in $\mathscr{T}_2$ topology.*

Now write D_n for the diameter of the graph $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$, namely the largest graph distance between two vertices in the same component in $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$. Once one is able to prove convergence in the l^4 metric, as in [3], one gets asymptotics for the diameter as well.

Theorem 3.5. *Assume that the weight sequence satisfies Assumptions* 3.1 *and* 3.2*. Then*

$$\frac{D_n}{n^{1/3}} \xrightarrow{\mathrm{w}} \Xi_\infty$$

where Ξ_∞ is a positive random variable that has an absolutely continuous distribution.

By [28, Corollary 2.12], the Norros-Reittu random graph model is asymptotically equivalent (in the sense of [28]) to the Chung-Lu model and the Britton-Deijfen-Lof model under Assumptions 3.1. Hence, an immediate consequence of Theorem 3.3 and Theorem 3.5 is the following corollary.

Theorem 3.6. *In the setup of Theorems* 3.3 *and* 3.5*, the conclusions hold for the Chung-Lu model and Britton-Deijfen-Lof model.*

3.2. **Height of p-trees.** Fix $m \geqslant 1$ and assume that we are given a positive probability vector $\mathbf{p} = \mathbf{p}^{(m)} = (p_i^{(m)} : i \in [m])$. We omit m in the notation. Define

$$\sigma(\mathbf{p}) = \sqrt{\sum_{i=1}^{m} p_i^2}, \ \ p_{\max} = \max_{1\leqslant i\leqslant m} p_i, \ \ p_{\min} = \min_{1\leqslant i\leqslant m} p_i.$$

We will prove the following tail bound for the height of **p**-trees. Let $\mathcal{T} \in \mathbb{T}_m$ be a random **p**-tree with distribution as in (2.7). Let $\mathrm{ht}(\mathcal{T})$ be the height of the tree $\mathcal{T}$.

Theorem 3.7 (Tail bounds)**.** *Assume that there exist* $\epsilon_0 \in (0,1/2)$ *and* $r_0 \in (2,\infty)$ *such that*

$$\sigma(\mathbf{p}) \leqslant \frac{1}{2^{10}}, \ \ \frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2+\epsilon_0}} \leqslant 1, \ \ \frac{[\sigma(\mathbf{p})]^{r_0}}{p_{\min}} \leqslant 1. \tag{3.3}$$

Then for any $r > 0$*, there exists some constant* $K_{3.7} = K_{3.7}(r) > 0$*, such that*

$$\mathbb{P}\left(\mathrm{ht}(\mathcal{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \frac{K_{3.7}}{x^r}, \ \textit{for } x \geqslant 1. \tag{3.4}$$

Remark 2. The assumptions in (3.3) are satisfied since the three quantities in (3.3), under the assumptions in this paper, converge to zero as $m \to \infty$. The above theorem gives a uniform tail bound for all **p** satisfying (3.3). This makes it possible to control the diameter for many components in $\mathcal{G}_n^{\mathrm{nr}}(\mathbf{w},\lambda)$ uniformly and prove convergence in $\mathcal{T}_2$ topology in Theorem 3.3.

4. Description of limit objects

In this section we describe the limit objects $\mathbf{M}(\lambda)$ arising in Theorem 3.3, first constructed for the Erdős-Rényi random graph in [3]. We need the following three ingredients:

(i) **Real trees:** An abstract notion of "tree-like" metric spaces.
(ii) **Shortcuts:** A notion of when and where to identify points in the real tree to take into account the fact that maximal components in the critical regime may not be trees and could have non-zero surplus or complexity.
(iii) **Tilted Brownian excursions:** We will need Brownian excursions whose length are described by the limit of component sizes (appropriately rescaled) of the rank-one model as proved in [11] tilted in favor of excursions with "large area".

4.1. **Real trees and shortcuts.** A compact metric space (X,d) is called a *real tree* [25, 31] if between every two points there is a unique geodesic such that this path is also the only non self-intersecting path between the two points. Functions encoding excursions can be used to construct such metric spaces which we now describe.

For $0 < a < b < \infty$, an *excursion* on $[a,b]$ is a continuous function $h \in C([a,b])$ with $h(a) = 0 = h(b)$ and $h(t) > 0$ for $t \in (a,b)$. The length of such an excursion is $b-a$. For $l \in (0,\infty)$, let $\mathcal{E}_l$ be the space of all excursions on the interval $[0,l]$. Given an excursion $h \in \mathcal{E}_l$, one can construct a real tree as follows. Define the pseudo-metric d_h on $[0,l]$ as follows:

$$d_h(s,t) := h(s) + h(t) - 2\inf_{u\in[s,t]} h(u), \ \ \text{for } s,t \in [0,l]. \tag{4.1}$$

Define the equivalence relation $s \sim t \Leftrightarrow d_h(s,t) = 0$. Let the $[0,l]/\sim$ denote the corresponding quotient space and consider the metric space $\mathcal{T}(h) := ([0,l]/\sim, \bar{d}_h)$, where $\bar{d}_h$ is the metric induced by d_h. Then $\mathscr{T}(h)$ is a real tree ([25, 31]). Let $q_h : [0,l] \to \mathscr{T}(h)$ be the canonical projection and write μ_h for the push-forward of the Lebesgue measure on $[0,l]$ onto $\mathscr{T}(h)$ via q_h. Equipped with μ_h, $\mathscr{T}(h)$ is now a measured metric space.

Since our limit objects will not necessarily be trees, we define a procedure that incorporates "short cuts" (more precisely identification of points) on a real tree. Let $h, g \in \mathscr{E}_l$ be two excursions, and $\mathscr{P} \subseteq \mathbb{R}_+ \times \mathbb{R}_+$ be a countable set with

$$g \cap \mathscr{P} := \left\{(x,y) \in \mathscr{P} : 0 \leqslant x \leqslant l,\ 0 \leqslant y < g(x)\right\} < \infty.$$

Using these three ingredients, we construct a metric space $\mathscr{G}(h,g,\mathscr{P})$ as follows. Let $\mathscr{T}(h)$ be the real tree associated with h and $q_h : [0,l] \to \mathscr{T}(h)$ be the canonical projection. Suppose $g \cap \mathscr{P} = \left\{(x_i, y_i) : 1 \leqslant i \leqslant k\right\}$ for some $k < \infty$. For each $i \leqslant k$, define

$$r(x_i, y_i) := \inf\left\{x : x \geqslant x_i,\ g(x) \leqslant y_i\right\}. \tag{4.2}$$

For the $i \leqslant k$ points, identify the points $q_h(x_i)$ and $q_h(r(x_i, y_i))$ in $\mathscr{T}(h)$. Call the resulting metric space $\mathscr{G}(h,g,\mathscr{P})$. Equipping this metric space with the push forward of the measure μ_h on $\mathscr{T}(h)$ makes $\mathscr{G}(h,g,\mathscr{P})$ a measured metric space. $\mathscr{G}(h,g,\mathscr{P})$ can be viewed as the metric space obtained by adding k shortcuts in $\mathscr{T}(h)$, with the location of the shortcuts determined by the excursion g and the collection of points $\mathscr{P}$. A shortcut between two points $u, v \in \mathscr{T}(h)$ identifies these two points as a single point.

4.2. **Scaling limits for component sizes of the Norros-Reittu model.** Let $\{B(s) : s \geqslant 0\}$ be a standard Brownian motion. For $\kappa, \sigma \in (0,\infty)$ and $\lambda \in \mathbb{R}$, define

$$W^{\lambda}_{\kappa,\sigma}(s) := \kappa B(s) + \lambda s - \sigma \frac{s^2}{2}, \qquad s \geqslant 0. \tag{4.3}$$

Define the reflected process

$$\bar{W}^{\lambda}_{\kappa,\sigma}(s) := W^{\lambda}_{\kappa,\sigma}(s) - \min_{0 \leqslant u \leqslant s} W^{\lambda}_{\kappa,\sigma}(u), \qquad s \geqslant 0. \tag{4.4}$$

Define the metric space

$$l^2_{\downarrow} := \left\{ \mathbf{x} = (x_i : i \geqslant 1) : x_1 \geqslant x_2 \geqslant \ldots \geqslant 0, \sum_{i=1}^{\infty} x_i^2 < \infty \right\},$$

equipped with the natural metric inherited from l^2. It was shown by Aldous in [7] that the excursions of $\bar{W}^{\lambda}_{\kappa,\sigma}$ from zero can be arranged in decreasing order of their lengths as

$$\boldsymbol{\xi}^{\lambda}_{\kappa,\sigma} = (\xi^{\lambda}_{\kappa,\sigma}(i) : i \geqslant 1), \tag{4.5}$$

where $\xi^{\lambda}_{\kappa,\sigma}(i)$ denotes the length of the i-th largest excursion, and further $\boldsymbol{\xi}^{\lambda}_{\kappa,\sigma} \in l^2_{\downarrow}$.

Consider the critical Norros-Reittu model $\mathscr{G}^{\mathrm{nr}}_n(\mathbf{w}, \lambda)$ with connection probabilities as in (1.4). Let $|\mathscr{C}^{(i)}_n(\lambda)|$ be the size of the i-th largest component, for $i \geqslant 1$. In [11] the following was shown about the normalized component sizes under finite third moment assumptions on the weight sequence.

Theorem 4.1 ([11])**.** *Fix $\lambda \in \mathbb{R}$. Under Assumption* 3.1 *(a), (b) and (c), as $n \to \infty$, we have*

$$\left(\frac{|\mathscr{C}^{(i)}_n(\lambda)|}{n^{2/3}} : i \geqslant 1\right) \xrightarrow{\mathrm{w}} \boldsymbol{\xi}^{\lambda}_{\sqrt{\sigma_3/\sigma_1}, \sigma_3/\sigma_1^2},$$

where the weak convergence is with respect to the topology generated by the l^2 norm.

Remark 3. It was assumed in [11] that $w_{\max} = o(n^{1/3})$ and that the empirical distribution $\frac{1}{n}\sum_{i=1}^{n}\delta_{w_i}$ converges in distribution to some limiting distribution F. This second assumption was not used in the proof of Theorem 4.1 other than to assert that express the limit constants σ_i in terms of moments of $W \sim F$ and thus can be removed as long as one has Assumptions 3.1 (a) and (b). We place stronger assumptions on $w_{\max}$ as in Assumption 3.1 (c).

To ease notation, for the rest of this section we shall suppress dependence on σ_1, σ_3 and λ and write the limiting component sizes as

$$\boldsymbol{\xi}^{\lambda}_{\sqrt{\sigma_3/\sigma_1},\sigma_3/\sigma_1^2} := \mathbf{Z} = (Z_i : i \geqslant 1). \tag{4.6}$$

Using Brownian scaling $\left\{a^{-1/2}B(as) : s \geqslant 0\right\} \overset{d}{=} \{B(s) : s \geqslant 0\}$ with $a = (\kappa/\sigma)^{2/3}$ implies that, for $s \geqslant 0$,

$$W^{\lambda}_{\kappa,\sigma}(as) = \frac{\kappa^{4/3}}{\sigma^{1/3}}\left[a^{-1/2}B(as) + \frac{\lambda}{\kappa^{2/3}\sigma^{1/3}}s - \frac{s^2}{2}\right] \overset{d}{=} \frac{\kappa^{4/3}}{\sigma^{1/3}} W^{\lambda/\kappa^{2/3}\sigma^{1/3}}_{1,1}(s).$$

Thus we have $\boldsymbol{\xi}^{\lambda}_{\kappa,\sigma} \overset{d}{=} (\kappa/\sigma)^{2/3}\boldsymbol{\xi}^{\lambda/\kappa^{2/3}\sigma^{1/3}}_{1,1}$. Therefore the limit object in Theorem 4.1 satisfies

$$\mathbf{Z} \overset{d}{=} \frac{\sigma_1}{\sigma_3^{1/3}}\boldsymbol{\xi}^{\lambda\sigma_1/\sigma_3^{2/3}}_{1,1}. \tag{4.7}$$

This relation will be useful in proving Lemma 4.2.

4.3. **Tilted Brownian excursions.** For fixed $l > 0$, recall that $\mathcal{E}_l$ denotes the space of excursions of length l. We can treat $\mathcal{E}_l$ as a subset of $C([0,\infty),[0,\infty))$ by identifying $h \in \mathcal{E}_l$ with $g \in C([0,\infty),[0,\infty))$ where $g(s) = h(s)$ for $s \in [0,l]$ and $g(s) = 0$ for $s > l$. Write $\mathcal{E} := \cup_{l>0}\mathcal{E}_l$ for the space of all finite length excursions from zero and equip $\mathcal{E}$ with the L^∞ norm, namely, $\|h\|_\infty = \sup_{s\in[0,\infty)}|h(s)|$.

Let $\{\mathbf{e}_l(s) : s \in [0,l]\}$ be a standard Brownian excursion of length l. For $l > 0$ and $\theta > 0$, define the tilted Brownian excursion $\tilde{\mathbf{e}}^{\theta}_l$ as an $\mathcal{E}$-valued random variable such that for all bounded continuous function $f : \mathcal{E} \to \mathbb{R}$,

$$\mathbb{E}[f(\tilde{\mathbf{e}}^{\theta}_l)] = \frac{\mathbb{E}\left[f(\mathbf{e}_l)\exp\left(\theta\int_0^l \mathbf{e}_l(s)ds\right)\right]}{\mathbb{E}\left[\exp\left(\theta\int_0^l \mathbf{e}_l(s)ds\right)\right]}. \tag{4.8}$$

Note that $\mathbf{e}_l$ and $\tilde{\mathbf{e}}^{\theta}_l$ are both supported on $\mathcal{E}_l$. Writing ν_l and $\tilde{\nu}^{\theta}_l$ respectively for the law of $\mathbf{e}_l$ and $\tilde{\mathbf{e}}^{\theta}_l$ on $\mathcal{E}_l$ the Radon-Nikodym derivative is given by

$$\frac{d\tilde{\nu}^{\theta}_l}{d\nu_l}(h) = \frac{\exp\left(\theta\int_0^l h(s)ds\right)}{\int_{\mathcal{E}_l}\exp\left(\theta\int_0^l h(s)ds\right)d\nu_l(dh)}, \qquad h \in \mathcal{E}_l.$$

When $l = 1$, we use $\mathbf{e}(\cdot)$ for the standard Brownian excursion. For fixed $l > 0$ and $\theta = 1$ we write $\tilde{\mathbf{e}}_l(\cdot)$ for the corresponding tilted excursion.

By Brownian scaling, $\left\{\sqrt{a}\mathbf{e}_l(s/a) : s \in [0,al]\right\} \overset{d}{=} \{\mathbf{e}_{al}(s) : s \in [0,al]\}$ for $a > 0$, thus $\int_0^{al}\mathbf{e}_{al}(s)ds \overset{d}{=} a^{3/2}\int_0^l \mathbf{e}_l(s)ds$. Taking $a = \theta^{2/3}$ and applying this to (4.8) gives

$$\mathbb{E}[f(\tilde{\mathbf{e}}^{\theta}_l)] = \frac{\mathbb{E}\left[f(\frac{1}{\sqrt{a}}\mathbf{e}_{al}(a\cdot))\exp\left(\int_0^{al}\mathbf{e}_{al}(s)ds\right)\right]}{\mathbb{E}\left[\exp\left(\int_0^{al}\mathbf{e}_{al}(s)ds\right)\right]} = \mathbb{E}\left[f\left(\frac{1}{\sqrt{a}}\tilde{\mathbf{e}}_{al}(a\cdot)\right)\right].$$

Thus the tilted excursions have the following scaling:

$$\left\{\tilde{\mathbf{e}}_l^{\theta}(s) : s \in [0,l]\right\} \stackrel{d}{=} \left\{\frac{1}{\theta^{1/3}}\tilde{\mathbf{e}}_{\theta^{2/3}l}(\theta^{2/3}s) : s \in [0,l]\right\}. \tag{4.9}$$

Note that the scaling relation below the equation (20) of [3] is not correct, which says $\|\tilde{\mathbf{e}}_l\|_\infty \stackrel{d}{=} \sqrt{l}\|\tilde{\mathbf{e}}\|_\infty$ under our notation. Based on (4.9), the correct version should be $\|\tilde{\mathbf{e}}_l\|_\infty \stackrel{d}{=} \sqrt{l}\|\tilde{\mathbf{e}}^{l^{3/2}}\|_\infty$.

In general, for any $l, \gamma, \theta > 0$, we have

$$\left\{\tilde{\mathbf{e}}_l^{\theta\gamma}(s) : s \in [0,l]\right\} \stackrel{d}{=} \left\{\frac{1}{\theta^{1/3}}\tilde{\mathbf{e}}^{\gamma}_{\theta^{2/3}l}(\theta^{2/3}s) : s \in [0,l]\right\}. \tag{4.10}$$

4.4. **Construction of the scaling limit.** We are now in a position to describe the scaling limits $\mathbf{M}(\lambda) = (M_i(\lambda) : i \geqslant 1)$ in Theorem 3.3. Let $\mathbf{Z}$ be a $l^2_{\downarrow}$-valued random variable as defined in (4.6) representing limits of normalized component sizes. Conditional on $\mathbf{Z}$, generate a sequence $\mathbf{h} := (h_i : i \geqslant 1)$ of independent random excursions in $\mathscr{E}$ via the prescription

$$h_i \stackrel{d}{=} \tilde{\mathbf{e}}_{Z_i}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \qquad i \geqslant 1.$$

Let $\mathbf{P} = (\mathscr{P}_i : i \in \mathbb{N})$ be a sequence of *iid* rate one Poisson point processes on $\mathbb{R}_+ \times \mathbb{R}_+$, independent of $(\mathbf{Z}, \mathbf{h})$. Define the metric spaces $M_i(\lambda) \in \mathscr{S}$, $i \geqslant 1$ as

$$M_i(\lambda) := \mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}}h_i, \frac{\sigma_3^{1/2}}{\sigma_1^{3/2}}h_i, \mathscr{P}_i\right) = \mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}}\tilde{\mathbf{e}}_{Z_i}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \frac{\sigma_3^{1/2}}{\sigma_1^{3/2}}\tilde{\mathbf{e}}_{Z_i}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \mathscr{P}_i\right), \tag{4.11}$$

where recall the construction of the metric spaces $\mathscr{G}(h, g, \mathscr{P})$ using the real trees encoded by excursions h and shortcuts generated by the excursion g and collection of points $\mathscr{P}$ as introduced in Section 4.1. Write $\mathbf{M}(\lambda) = (M_i(\lambda) : i \geqslant 1)$ for the sequence of random metric spaces so constructed. Then this is the asserted continuum limit of the critical components in the Norros-Reittu model in Theorem 3.3.

Now we compare this limit to the limit metric spaces for Erdős-Rényi random graphs as proved in [3].

Lemma 4.2. *The limit objects for the rank-one model satisfy the distributional equivalence*

$$\mathbf{M}(\lambda) \stackrel{d}{=} \mathrm{scl}\left(\frac{\sigma_1}{\sigma_3^{2/3}}, \frac{\sigma_1}{\sigma_3^{1/3}}\right) \cdot \mathbf{M}^{\mathrm{er}}\left(\frac{\lambda\sigma_1}{\sigma_3^{2/3}}\right),$$

where for any $\lambda' \in \mathbb{R}$, $\mathbf{M}^{\mathrm{er}}(\lambda')$ denote the limit objects for the Erdős-Rényi random graph $\mathscr{G}_n^{\mathrm{er}}(n, 1/n + \lambda'/n^{4/3})$ as constructed in [3].

Proof: Write

$$\boldsymbol{\xi}_{1,1}^{\lambda} := \boldsymbol{\gamma}(\lambda) = (\gamma_i(\lambda) : i \in \mathbb{N}). \tag{4.12}$$

In [3] it is shown that the scaling limits for the Erdős-Rényi model $\mathscr{G}^{\mathrm{er}}(n, 1/n + \lambda/n^{4/3})$ is

$$\mathbf{M}^{\mathrm{er}}(\lambda) = (M_i^{\mathrm{er}}(\lambda) : i \geqslant 1), \text{ where } M_i^{\mathrm{er}}(\lambda) := \mathscr{G}(2\tilde{\mathbf{e}}_{\gamma_i(\lambda)}, \tilde{\mathbf{e}}_{\gamma_i(\lambda)}, \mathscr{P}_i).$$

In order to compare $\mathbf{M}(\lambda)$ and $\mathbf{M}^{\mathrm{er}}(\lambda)$, we again use Brownian scaling. By (4.7), we have

$$Z_i = \frac{\sigma_1}{\sigma_3^{1/3}}\gamma_i(\lambda\sigma_1/\sigma_3^{2/3}). \tag{4.13}$$

By the definition of $\mathscr{G}(h,g,\mathscr{P})$ in Section 4.1 and $\mathrm{scl}(\alpha,\beta)$ in Section 2.2, we have for $\alpha,\beta>0$,

$$\mathrm{scl}(\alpha,\beta)\cdot\mathscr{G}(h,g,\mathscr{P})=\mathscr{G}(\alpha h(\cdot/\beta),\frac{1}{\beta}g(\cdot/\beta),\mathscr{P}^{\beta}), \tag{4.14}$$

where $\mathscr{P}^{\beta}:=\{(\beta x,y/\beta):(x,y)\in\mathscr{P}\}$. With the these ingredients, letting $\bar\theta:=\sigma_3^{1/2}/\sigma_1^{3/2}$, we have

$$\begin{aligned}M_i(\lambda)=&\mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}}\tilde{\mathbf{e}}^{\bar\theta}_{Z_i},\frac{\sigma_3^{1/2}}{\sigma_1^{3/2}}\tilde{\mathbf{e}}^{\bar\theta}_{Z_i},\mathscr{P}_i\right)\\ \stackrel{d}{=}&\mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}\bar\theta^{1/3}}\tilde{\mathbf{e}}_{\bar\theta^{2/3}Z_i}(\bar\theta^{2/3}\cdot),\frac{\sigma_3^{1/2}}{\sigma_1^{3/2}\bar\theta^{1/3}}\tilde{\mathbf{e}}_{\bar\theta^{2/3}Z_i}(\bar\theta^{2/3}\cdot),\mathscr{P}_i\right)\\ =&\mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}\bar\theta^{1/3}}\tilde{\mathbf{e}}_{\gamma_i(\lambda\sigma_1/\sigma_3^{2/3})}(\bar\theta^{2/3}\cdot),\frac{\sigma_3^{1/2}}{\sigma_1^{3/2}\bar\theta^{1/3}}\tilde{\mathbf{e}}_{\gamma_i(\lambda\sigma_1/\sigma_3^{2/3})}(\bar\theta^{2/3}\cdot),\mathscr{P}_i\right),\end{aligned}$$

where the first line is by definition, the second line is because of (4.9), and the third line follows from (4.13). To ease notation write $\gamma_i=\gamma_i(\lambda\sigma_1/\sigma_3^{2/3})$. Then taking $\bar\alpha=\sigma_3^{2/3}/\sigma_1$, we have

$$\begin{aligned}\mathrm{scl}(\bar\alpha,\bar\theta^{2/3})\cdot M_i(\lambda)=&\mathscr{G}\left(\frac{2\bar\alpha\sigma_1^{1/2}}{\sigma_3^{1/2}\bar\theta^{1/3}}\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\frac{\sigma_3^{1/2}}{\sigma_1^{3/2}\bar\theta}\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\mathscr{P}_i^{\bar\theta^{2/3}}\right)\\ =&\mathscr{G}\left(2\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\mathscr{P}_i^{\bar\theta^{2/3}}\right)\\ \stackrel{d}{=}&\mathscr{G}\left(2\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\tilde{\mathbf{e}}_{\gamma_i}(\cdot),\mathscr{P}_i\right),\end{aligned}$$

where the first line uses (4.14), the second line follows form the definition of $\bar\alpha$ and $\bar\theta$, and the third line follows from the fact about Poisson point processes $\mathscr{P}_i^{\beta}\stackrel{d}{=}\mathscr{P}_i$ for all $\beta>0$, and the independence between $\mathscr{P}_i$ and h_i in the construction. The proof of Lemma 4.2 is completed. ■

5. Discussion

Before proceeding to the proofs, let us briefly describe the relevance of these results and their connection to the existing literature on random graphs.

(a) **Connection to existing results:** As remarked in Section 1.2, the main aim of the paper was to rigorously understand conjectures in statistical physics on scaling limits of distances in the critical regime for inhomogeneous random graph models, which then predict distances for the minimal spanning tree (on the giant component) in the supercritical regime where each edge has $U[0,1]$ edge weights. See [15] and the references therein. This entire program has been rigorously carried out for the complete graph see [2,3,5] for a sequence of results including distance scaling for the maximal components in the critical regime for the Erdős-Rényi random graph, finally culminating in the scaling limit of the minimal spanning tree of the complete graph equipped with uniform edge weights. Most influential to this study is [3] which constructed the scaling limit of these components in the critical regime. Extending these results to the context of general inhomogeneous random graph models turned out to be challenging since the homogeneous nature of the Erdős-Rényi played an important role in various parts of the proof in [3].

While there have been few rigorous results on the actual structure of components in the critical regime for general random graph models, if one were interested in only sizes of the maximal components, then this has witnessed significant progress over the last few years, see [26, 30, 35, 39] for results on the configuration model, [10] for results on a class of dynamic random graph processes called the bounded-size-rules and most relevant for this work [11, 43] for results for the rank-one inhomogeneous random graph. See [14] for a recent survey.

(b) **Importance of the assumptions:** Consider the critical rank-one model studied in this paper. To prove our main results, we needed moment Assumptions 3.1. If one were interested in just proving results on the sizes of components (Theorem 4.1), finite third moment assumptions, namely Assumptions 3.1(a),(b) and (c) replaced by $w_{\max} = o(n^{1/3})$ suffice. However to understand the actual metric structure of the component, one is lead to these stronger assumptions owing to technical conditions in [8], required to show that in our setting the associated (untilted) $\mathbf{p}$-trees (properly rescaled by $n^{1/3}$) have scaling limits related to the continuum random tree. The results in [8] in fact assume exponential moments, however as in the Remark under [8, Theorem 3], these can be extended without much work following the same proof technique as in [8] to the setting of finite moment conditions, we state this extension in Theorem 7.6. We believe that in fact these results can be extended all the way to finite third moments and are in the process of understanding how to refine these results. Given the technical nature of the proof even with sufficient moment assumptions, we defer this to future work.

6. Proof preliminaries and outline

We now start on the proofs of the main results. In this section we start with an outline of the proof and describe some preliminary properties of the rank-one model.

6.1. Outline of proof.

Let us describe the main steps in the proof. We start in Section 6.2 where we describe how the the connected components of the rank-one model can be constructed in two steps. In particular this will imply that there are two major foci in understanding the maximal connected components:

(a) Constructing a rank-one model conditional on being connected. In Section 7.1, we explore such a graph in a randomized depth-first manner and show that the law of this depth-first tree is that of an *ordered tilted* $\mathbf{p}$-*tree* (Proposition 7.4). This will imply an alternate way of constructing a rank-one graph conditioned on being connected: first generate an ordered tilted $\mathbf{p}$-tree and then add the surplus edges independently (Proposition 7.4) with appropriate probability. Strengthening the results of [8], it follows that an ordinary $\mathbf{p}$-tree converges in Gromov-Hausdorff topology (after the tree distance has been properly rescaled) to a continuum random tree under some regularity conditions on the driving probability mass function $\mathbf{p}$. Provided we can show the corresponding tilt is “nice”, this would imply that rank-one random graphs conditioned on being connected converge to a tilted continuum tree where certain pairs of points have been identified. We achieve this with the help of Lemma 7.10 (which shows that the tilt converges pointwise) and Lemma 7.11 (which yields uniform integrability of the tilt).

(b) Regularity of vertex weights in the maximal components. To study this, we start in Section 8 by describing the exploration of the graph $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$ in a size-biased random order first used in [7] and later used in [11] to prove Theorem 4.1 on the scaling of the maximal components in the critical regime. We use this exploration to establish strong regularity properties of the *weights* of vertices in these maximal components (Proposition 8.1).

Now conditional on these regularity properties being satisfied within each maximal component, the internal structure of each maximal component is simply that of a rank-one inhomogeneous graph conditioned on being connected. We then combine these two aspects to prove convergence of the scaled components Section 9 where we first prove convergence in the product topology $\mathscr{T}_1$. To extend the convergence of the components in $\mathscr{T}_2$ topology (l^4 metric) amounts to proving a tail bound on the diameter of the components. Since the depth-first tree of each component spans the component and is distributed as a tilted **p**-tree, it is enough to get a tail bound on heights of **p**-trees. We achieve this in Section 10 by using techniques from [8, 19]. Finally in Section 11, we complete the proof of some lemmas which are essential ingredients in the proof of convergence in $\mathscr{T}_2$ topology.

6.2. **Connected components of the model.** Recall that $(\mathscr{C}_n^{(i)} : i \in \mathbb{N})$ are the components of $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$. Fix $\mathcal{V} \subset [n]$ and write $\mathbb{G}_{\mathcal{V}}^{\mathrm{con}}$ the space of all simple connected graphs with vertex set $\mathcal{V}$. For fixed $a > 0$, and probability mass function $\mathbf{p} = (p_v : v \in \mathcal{V})$, define probability distributions $\mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}, a, \mathcal{V})$ on $\mathbb{G}_{\mathcal{V}}^{\mathrm{con}}$ as follows: Define for $i, j \in \mathcal{V}$,

$$q_{ij} := 1 - \exp(-a p_i p_j). \tag{6.1}$$

Then

$$\mathbb{P}_{\mathrm{con}}(G; \mathbf{p}, a, \mathcal{V}) := \frac{1}{Z(\mathbf{p}, a)} \prod_{(i,j) \in E(G)} q_{ij} \prod_{(i,j) \notin E(G)} (1 - q_{ij}), \text{ for } G \in \mathbb{G}_{\mathcal{V}}^{\mathrm{con}}, \tag{6.2}$$

where $Z(\mathbf{p}, a)$ is the normalizing constant

$$Z(\mathbf{p}, a) := \sum_{G \in \mathbb{G}_{\mathcal{V}}^{\mathrm{con}}} \prod_{(i,j) \in E(G)} q_{ij} \prod_{(i,j) \notin E(G)} (1 - q_{ij}).$$

Now define $\mathcal{V}^{(i)} := \{v \in [n] : v \in \mathscr{C}_n^{(i)}\}$ for $i \in \mathbb{N}$ and note that $\{\mathcal{V}^{(i)} : i \geqslant 1\}$ denotes a random (finite) partition of the complete vertex set $[n]$. The next proposition characterizes the distribution of the random graphs $(\mathscr{C}_n^{(i)} : i \in \mathbb{N})$ conditioned on the partition $\{\mathcal{V}^{(i)} : i \in \mathbb{N}\}$.

Proposition 6.1. *For $i \geqslant 1$ define*

$$\mathbf{p}^{(i)} := \left(\frac{w_v}{\sum_{v \in \mathcal{V}^{(i)}} w_v} : v \in \mathcal{V}^{(i)} \right), \quad a^{(i)} := \left(1 + \frac{\lambda}{n^{1/3}}\right) \frac{(\sum_{v \in \mathcal{V}^{(i)}} w_v)^2}{l_n}.$$

Then for any $k \in \mathbb{N}$ and $G_i \in \mathbb{G}_{\mathcal{V}^{(i)}}^{\mathrm{con}}$, we have

$$\mathbb{P}\left(\mathscr{C}_n^{(i)} = G_i, \ \forall i \geqslant 1 \mid \{\mathcal{V}^{(i)} : i \geqslant 1\}\right) = \prod_{i \geqslant 1} \mathbb{P}_{\mathrm{con}}(G_i; \mathbf{p}^{(i)}, a^{(i)}, \mathcal{V}^{(i)}).$$

The above proposition says the random graph $\mathscr{G}_n^{nr}(\mathbf{w}, \lambda)$ can be generated in two stages.

(i) In the first stage generate the partition of the vertices into different components, i.e. $\{\mathcal{V}^{(i)} : i \in \mathbb{N}\}$.
(ii) In the second stage, given the partition, we generate the internal structure of each component following the law of $\mathbb{P}_{\mathrm{con}}(\cdot; \mathbf{p}^{(i)}, a^{(i)}, \mathcal{V}^{(i)})$, independently across different components.

The proof of Proposition 6.1 immediately follows from the expression of connection probability and independence structure of the Norros-Reittu model, which is omitted.

The plan of the next section is to study the technically more challenging question of the rank-one random graph conditioned on being connected.

7. Scaling limit of rank-one graphs conditioned on being connected

In this section we study the continuum limit of general rank-one random graphs conditioned on being connected. For $m \geqslant 1$ and let $\mathbf{p}^{(m)} = (p_i^{(m)} : i \in [m])$ be a probability mass function. We will suppress m in the notation and just write $\mathbf{p} = (p_i : i \in [m])$. Recall the definitions

$$\sigma(\mathbf{p}) := \sqrt{\sum_{i\in[m]} p_i^2}, \qquad p_{\max} := \max_{i\in[m]} p_i, \qquad p_{\min} := \min_{i\in[m]} p_i.$$

We will make the following assumptions on $\mathbf{p}$ as $m \to \infty$.

Assumption 7.1. *There exists $\epsilon > 0$ and $r > 0$ such that, as $m \to \infty$, we have*

$$\sigma(\mathbf{p}) \to 0, \ \ \frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2+\epsilon}} \to 0, \ \ \frac{[\sigma(\mathbf{p})]^r}{p_{\min}} \to 0.$$

Recall that $\mathbb{G}_m^{\mathrm{con}}$ was defined as the collection of all connected graphs with vertex set $[m]$. Let $\{a(m) : m \geqslant 1\}$ be a sequence of positive real numbers. We will use $a = a(m)$ and $\mathbf{p}$ to construct a probability measure $\mathbb{P}_{\mathrm{con}}$ on $\mathbb{G}_m^{\mathrm{con}}$ as follows:

$$\mathbb{P}_{\mathrm{con}}(G) := \mathbb{P}_{\mathrm{con}}(G; \mathbf{p}, a, [m]), \qquad G \in \mathbb{G}_m^{\mathrm{con}}, \tag{7.1}$$

where the latter is defined in (6.2). Let $\mathscr{G}_m$ be a $\mathbb{T}_m^{\mathrm{ord}}$-valued random variable with distribution $\mathbb{P}_{\mathrm{con}}$. Thus $\mathscr{G}_m$ has the same distribution as a rank-one random graph with vertex set $[m]$ and connection probabilities $q_{ij} = 1 - \exp(-ap_ip_j)$, *conditioned on being connected*. We will think of $\mathscr{G}_m \in \mathbb{G}_m^{\mathrm{con}}$ as a measured metric space as described in Section 2.3 using the graph distance, and further assigning mass p_i to vertex $i \in [m]$. The main result of this section shows that under some conditions, as $m \to \infty$, the metric space $\mathscr{G}_m$ with graph distance rescaled by $\sigma(\mathbf{p})$ converges to a measured (random) metric space with distribution as described in Section 4.1. In addition to Assumption 7.1, we need the following assumption on $a(m)$.

Assumption 7.2. *For some constant $\bar\gamma \in (0,\infty)$,*

$$\lim_{m\to\infty} a\sigma(\mathbf{p}) = \bar\gamma. \tag{7.2}$$

The main aim of this section is to prove the following result.

Theorem 7.3. *Let $\mathscr{G}_m$ be a $\mathbb{G}_m^{\mathrm{con}}$-valued random variable with law $\mathbb{P}_{\mathrm{con}}$. Under Assumptions 7.1 and 7.2, as $m \to \infty$,*

$$\mathrm{scl}\big(\sigma(\mathbf{p}), 1\big) \cdot \mathscr{G}_m \xrightarrow{\mathrm{w}} \mathscr{G}(2\tilde{\mathbf{e}}^{\bar\gamma}, \bar\gamma\tilde{\mathbf{e}}^{\bar\gamma}, \mathscr{P}),$$

where $\tilde{\mathbf{e}}^{\bar\gamma}$ is the tilted Brownian excursion as defined in (4.8), $\mathscr{P}$ is a rate one Poisson point process on $\mathbb{R}_+^2$ independent of $\tilde{\mathbf{e}}^{\bar\gamma}$, and $\mathscr{G}(2\tilde{\mathbf{e}}^{\bar\gamma}, \bar\gamma\tilde{\mathbf{e}}^{\bar\gamma}, \mathscr{P})$ is as defined in Section 4.1.

This section is organized as follows. We start in Section 7.1 where we will give an alternative construction of the law $\mathbb{P}_{\mathrm{con}}$ from an ordered $\mathbf{p}$-tree by tilting this distribution appropriately. In Section 7, we study the scaling limit of a random connected graph without applying the tilt. Finally in Section 7.3, we prove the tightness of the tilt and complete the proof of Theorem 7.3.

7.1. Distribution of connected components and tilted p-trees.

Recall that $\mathbb{T}_m^{\mathrm{ord}}$ denotes the space of ordered trees on m vertices. We start by introducing the following **randomized Depth First Search (rDFS)** procedure, which takes a graph $G \in \mathbb{G}_m^{\mathrm{con}}$ as the input and gives a random ordered tree in $\mathbb{T}_m^{\mathrm{ord}}$, denote by $\Gamma_{\mathbf{p}}(G)$, as its output. Given $G \in \mathbb{G}_m^{\mathrm{con}}$, the rDFS consists of two stages:

I. Selection of a root: Pick $v(1) \in [m]$ at random with the distribution $\mathbf{p}$. The vertex $v(1)$ is the starting point of the rDFS algorithm on the graph G and also the root of the ensuing tree $\Gamma_{\mathbf{p}}(G)$.

II. Depth-First-Search: At each step $1 \leqslant i \leqslant m$, we will keep track of three types of vertices.

(a) The set of already explored vertices, $\mathscr{O}(i)$.
(b) The set of active vertices$\mathscr{A}(i)$. We view $\mathscr{A}$ as a vertical *stack* with $\mathscr{A}(i)$ denoting the state of the stack in the *end* at the i-th step.
(c) The set of unexplored vertices $\mathscr{U}(i) := [m] \setminus (\mathscr{A}(i) \cup \mathscr{O}(i))$.

Initialize with $\mathscr{A}(0) = \{v(1)\}, \mathscr{O}(0) = \emptyset$. At step $i \geqslant 1$, let $v(i)$ denote the vertex on *top* of the stack $\mathscr{A}(i-1)$ and let

$$\mathscr{D}(i) := \{u \in \mathscr{U}(i-1) : (v(i), u) \in E(G)\},$$

namely $\mathscr{D}(i)$ is the set of unexplored neighbors of $v(i)$. Let $d_{v(i)} = |\mathscr{D}(i)|$ and suppose $\mathscr{D}(i) = \left\{u(j) : 1 \leqslant j \leqslant d_{v(i)}\right\}$. Then update the stack $\mathscr{A}(\cdot)$ in the following manner:

(i) **Randomization:** Generate $\boldsymbol{\pi} = \boldsymbol{\pi}(i)$ a uniform random permutation on $[d(i)]$.
(ii) Delete $v(i)$ from $\mathscr{A}(i-1)$.
(iii) Arrange the vertices $\mathscr{D}(i)$ on top of the stack $\mathscr{A}(i-1)$ using the order $\boldsymbol{\pi}$ generated in (i).

Define $\mathscr{A}(i)$ to be the state of the stack $\mathscr{A}$ after the above operations. As sets, $\mathscr{A}(i) = \mathscr{A}(i-1) \cup \mathscr{D}(i) \setminus \{v(i)\}$. Define $\mathscr{O}(i) := \mathscr{O}(i-1) \cup \{v(i)\}$ and $\mathscr{U}(i) := \mathscr{U}(i-1) \setminus \mathscr{D}(i)$.

Note that in the above rDFS algorithm, we have $|\mathscr{O}(i)| = i$ for $i \in [m]$. Thus after m steps we complete the exploration of all vertices in G. At the end of the procedure we are left with a rooted random tree $\Gamma_{\mathbf{p}}(G) \in \mathbb{T}_m^{\mathrm{ord}}$ with $v(1)$ as the root and with edge set $E(\Gamma_{\mathbf{p}}(G)) := \{(v(i), u) : i \in [m],\ u \in \mathscr{D}(i)\}$. Carrying the order $\{\boldsymbol{\pi}(i) : i \in [m]\}$ used to order the vertices at each stage of the procedure then makes the resulting tree an ordered tree that we explore in a depth first manner resulting in the order $(v(1), \ldots, v(m))$. This completes the construction of $\Gamma_{\mathbf{p}}(G) \in \mathbb{T}_m^{\mathrm{ord}}$. Note that for fixed G, $\Gamma_{\mathbf{p}}(G)$ is a $\mathbb{T}_m^{\mathrm{ord}}$-valued random variable.

The rDFS algorithm incorporates randomization in two places: First, the root is chosen randomly using the probability mass function $\mathbf{p}$; second, the children (unexplored vertices) of each vertex are explored in an uniform random order. Given an ordered tree $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$, one can run a depth first search on $\mathbf{t}$ in a deterministic manner starting from the root of the tree $\mathbf{t}$ and exploring the children using the associated order of the tree. Let $(\mathscr{O}(i), \mathscr{A}(i), \mathscr{U}(i), \mathscr{D}(i) : i \in [m])$ be the corresponding sets of vertices obtained from this deterministic depth first search of the tree $\mathbf{t}$. Write $\mathfrak{P}(\mathbf{t}, \boldsymbol{\pi})$ for the set of edges $\{u, v\}$ such that there exists $0 \leqslant i \leqslant n-1$ such that $u, v \in \mathscr{A}(i)$, namely both are active but have not yet been explored. Using terminology from [3] call this collection of edges, the set of *permitted edges*. By definition,

$$\mathfrak{P}(\mathbf{t}) := \left\{(v(i), j) : i \in [m],\ \ j \in \mathscr{A}(i-1) \setminus \{v(i)\}\right\}. \tag{7.3}$$

Write $[m]_2$ for the collection of all possible edges on a graph with vertex set $[m]$ and recall that $E(\mathbf{t})$ denotes the edge set of $\mathbf{t}$. Call the remaining edges,

$$\mathfrak{F}(\mathbf{t}) := [m]_2 \setminus (\mathfrak{P}(\mathbf{t}) \cup E(\mathbf{t})),$$

the set of *forbidden* edges.

For a fixed planar tree $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$, define the subset $\mathbb{G}(\mathbf{t}) \subset \mathbb{G}_m^{\mathrm{con}}$ as

$$\mathbb{G}(\mathbf{t}) := \left\{G \in \mathbb{G}_m^{\mathrm{con}} : E(\mathbf{t}) \subset E(G) \subset E(\mathbf{t}) \cup \mathfrak{P}(\mathbf{t})\right\}. \tag{7.4}$$

For fixed $G \in \mathbb{G}_m^{\text{con}}$, let $\nu^{\text{dfs}}(\cdot; G)$ be the probability distribution of $\Gamma_{\mathbf{p}}(G)$ on $\mathbb{T}_m^{\text{ord}}$. When $G \notin \mathbb{G}(\mathbf{t})$, by [3, Lemma 7], we have $\nu^{\text{dfs}}(\mathbf{t}; G) = 0$. When $G \in \mathbb{G}(\mathbf{t})$, from the construction, we have

$$\nu^{\text{dfs}}(\mathbf{t}; G) = p_{r(\mathbf{t})} \prod_{i \in [m]} \frac{1}{d_i(\mathbf{t})!}, \tag{7.5}$$

where $r(\mathbf{t})$ denotes the root of $\mathbf{t}$ and $d_i(\mathbf{t})$ denotes the number of children of i in $\mathbf{t}$.

Recall the law of an ordered $\mathbf{p}$-tree, denoted by $\mathbb{P}_{\text{ord}}(\cdot)$, as defined in (2.8). Define the function $L : \mathbb{T}_m^{\text{ord}} \to \mathbb{R}_+$ by

$$L(\mathbf{t}) := \prod_{(k,\ell) \in E(\mathbf{t})} \left[\frac{\exp(ap_k p_\ell) - 1}{ap_k p_\ell} \right] \exp\left(\sum_{(k,\ell) \in \mathfrak{P}(\mathbf{t})} ap_k p_\ell \right), \qquad \mathbf{t} \in \mathbb{T}_m^{\text{ord}}. \tag{7.6}$$

Using $L(\cdot)$ to tilt the distribution of the $\mathbf{p}$-tree results in the distribution

$$\tilde{\mathbb{P}}_{\text{ord}}(\mathbf{t}) := \mathbb{P}_{\text{ord}}(\mathbf{t}) \cdot \frac{L(\mathbf{t})}{\mathbb{E}_{\text{ord}}[L(\cdot)]}, \qquad \mathbf{t} \in \mathbb{T}_m^{\text{ord}}, \tag{7.7}$$

where $\mathbb{E}_{\text{ord}}[L(\cdot)]$ denotes the expectation of $L(\cdot)$ with respect to the law $\mathbb{P}_{\text{ord}}$.

Now note that given a fixed planar tree $\mathbf{t} \in \mathbb{T}_m^{\text{ord}}$, one can construct a connected random graph by adding each possible permitted edge $(i,j) \in \mathfrak{P}(\mathbf{t})$ independently with probability $q_{ij} = 1 - \exp(-ap_i p_j)$. Write $\nu^{\text{per}}(\cdot; \mathbf{t})$ for the probability distribution of this random graph, where "per" stands for "permitted edges". Obviously by construction, the support of $\nu^{\text{per}}(\cdot; \mathbf{t})$ is the set $\mathbb{G}(\mathbf{t})$ as defined in (7.4) and has the explicit form,

$$\nu^{\text{per}}(G; \mathbf{t}) := \mathbb{1}_{\{G \in \mathbb{G}(\mathbf{t})\}} \prod_{(i,j) \in \mathfrak{P}(\mathbf{t}) \cap E(G)} q_{ij} \prod_{(i,j) \in \mathfrak{P}(\mathbf{t}) \setminus E(G)} (1 - q_{ij}). \tag{7.8}$$

The main result of this section is the following proposition. In words what this result says is the following: one can sample a connected random graph $\mathscr{G} \sim \mathbb{P}_{\text{con}}$ with distribution in (7.1), in the following two step procedure:

(a) Generate a random planar tree $\tilde{\mathscr{T}}$ using the tilted $\mathbf{p}$-tree distribution $\tilde{\mathbb{P}}_{\text{ord}}(\cdot)$ given in (7.7) via the tilt $L(\cdot)$.
(b) Conditional on $\tilde{\mathscr{T}}$, add each of the permitted edges $(i,j) \in \mathfrak{P}(\tilde{\mathscr{T}})$ independently with the appropriate probability q_{ij}.

Proposition 7.4. *For all $G \in \mathbb{G}_m^{\text{con}}$ and $\mathbf{t} \in \mathbb{T}_m^{\text{ord}}$,*

$$\mathbb{P}_{\text{con}}(G) \nu^{\text{dfs}}(\mathbf{t}; G) = \tilde{\mathbb{P}}_{\text{ord}}(\mathbf{t}) \nu^{\text{per}}(G; \mathbf{t}). \tag{7.9}$$

In particular, we have $\mathbb{P}_{\text{con}}(G) = \sum_{\mathbf{t} \in \mathbb{T}_m^{\text{ord}}} \tilde{\mathbb{P}}_{\text{ord}}(\mathbf{t}) \nu^{\text{per}}(G; \mathbf{t})$.

Proof:. From the definition of $\nu^{\text{dfs}}(\mathbf{t}; G)$ and $\nu^{\text{per}}(G; \mathbf{t})$, the left hand side and right hand side of (7.9) are non zero if and only if $G \in \mathbb{G}(\mathbf{t})$. When $G \in \mathbb{G}(\mathbf{t})$, using (7.1) and (7.5) for the left hand side gives

$$\mathbb{P}_{\text{con}}(G) \nu^{\text{dfs}}(\mathbf{t}; G) = \frac{1}{Z(\mathbf{p})} \prod_{(i,j) \in E(G)} (1 - e^{-ap_i p_j}) \prod_{(i,j) \notin E(G)} e^{-ap_i p_j} \times p_{r(\mathbf{t})} \prod_{i \in [m]} \frac{1}{d_i(\mathbf{t})!}. \tag{7.10}$$

To ease notation write $d_i(\mathbf{t}) = d_i$ for the number of children of vertex i in $\mathbf{t}$. Let us now simplify the right hand side of (7.9). Using (7.6) and the fact that for a fixed tree $\mathbf{t}$, $\mathfrak{F}(\mathbf{t}), \mathfrak{P}(\mathbf{t})$ and $E(\mathbf{t})$ form a partition of all possible edges (denoted by $[m]_2$) on the vertex set $[m]$ gives

$$
\begin{aligned}
L(\mathbf{t}) &= \prod_{(i,j)\in E(\mathbf{t})}\left[\frac{e^{ap_ip_j}-1}{ap_ip_j}\right]\prod_{(i,j)\in\mathfrak{P}(\mathbf{t})} e^{ap_ip_j} \\
&= \frac{p_{r(\mathbf{t})}}{a^{m-1}}\prod_{i\in[m]}\frac{1}{p_i^{d_i+1}}\prod_{(i,j)\in E(\mathbf{t})}(e^{ap_ip_j}-1)\prod_{(i,j)\in\mathfrak{P}(\mathbf{t})} e^{ap_ip_j} \\
&= \frac{p_{r(\mathbf{t})}}{a^{m-1}}\prod_{i\in[m]}\frac{1}{p_i^{d_i+1}}\prod_{(i,j)\in[m]_2} e^{ap_ip_j}\prod_{(i,j)\in E(\mathbf{t})}(1-e^{-ap_ip_j})\prod_{(i,j)\in\mathfrak{F}(\mathbf{t})} e^{-ap_ip_j}.
\end{aligned}
$$

Using the above display, (2.8) and (7.8) we have

$$
\begin{aligned}
\tilde{\mathbb{P}}_{\mathrm{ord}}(\mathbf{t})\nu^{\mathrm{per}}(G;\mathbf{t}) &= \frac{1}{\mathbb{E}_{\mathrm{ord}}[L(\cdot)]}\prod_{i\in[m]}\frac{p_i^{d_i}}{d_i!}\times\prod_{(i,j)\in\mathfrak{P}(\mathbf{t})\cap E(G)}(1-e^{-ap_ip_j})\prod_{(i,j)\in\mathfrak{P}(\mathbf{t})\setminus E(G)} e^{-ap_ip_j}\times L(\mathbf{t}) \\
&= \frac{a^{-(m-1)}\prod_{(i,j)\in[m]_2}e^{ap_ip_j}}{\mathbb{E}_{\mathrm{ord}}[L(\cdot)]\prod_{i\in[m]}p_i}\times p_{r(\mathbf{t})}\prod_{i\in[m]}\frac{1}{d_i(\mathbf{t})!}\times\prod_{(i,j)\in E(G)}(1-e^{-ap_ip_j})\prod_{(i,j)\notin E(G)} e^{-ap_ip_j},
\end{aligned}
$$

where the last display is obtained by using $E(G)=E(\mathbf{t})\cup(\mathfrak{P}(\mathbf{t})\cap E(G))$ and $E(G)^c=\mathfrak{F}(\mathbf{t})\cup(\mathfrak{P}(\mathbf{t})\setminus E(G))$. Comparing the above expression with (7.10) we have

$$
\frac{\mathbb{P}_{\mathrm{con}}(G)\nu^{\mathrm{dfs}}(\mathbf{t};G)}{\tilde{\mathbb{P}}_{\mathrm{ord}}(\mathbf{t})\nu^{\mathrm{per}}(G;\mathbf{t})} = f(\mathbf{p},a,m),
$$

where $f(\mathbf{p},a,m)$ is a constant, independent of $\mathbf{t}$ or G. Since both the left and the right hand sides are probability distributions, $f(\mathbf{p},a,m)\equiv 1$. This completes the proof. ■

7.2. **Convergence of untilted graphs.** Using Proposition 7.4, define the probability distribution $\tilde{\nu}^{\mathrm{jt}}(\cdot,\cdot)$ on $\mathbb{T}_m^{\mathrm{ord}}\times\mathbb{G}_m^{\mathrm{con}}$ via the prescription,

$$
\tilde{\nu}^{\mathrm{jt}}(\mathbf{t},G) := \mathbb{P}_{\mathrm{con}}(G)\nu^{\mathrm{dfs}}(\mathbf{t};G) = \tilde{\mathbb{P}}_{\mathrm{ord}}(\mathbf{t})\nu^{\mathrm{per}}(G;\mathbf{t}),\ \text{for } \mathbf{t}\in\mathbb{T}_m^{\mathrm{ord}}, G\in\mathbb{G}_m^{\mathrm{con}}. \tag{7.11}
$$

This is the main object of interest. Let us first study the simpler object which does not incorporate the tilt. More precisely define the probability distribution $\nu^{\mathrm{jt}}(\cdot,\cdot)$ on $\mathbb{T}_m^{\mathrm{ord}}\times\mathbb{G}_m^{\mathrm{con}}$ as follows:

$$
\nu^{\mathrm{jt}}(\mathbf{t},G) := \mathbb{P}_{\mathrm{ord}}(\mathbf{t})\nu^{\mathrm{per}}(G;\mathbf{t}),\ \text{for } \mathbf{t}\in\mathbb{T}_m^{\mathrm{ord}}, G\in\mathbb{G}_m^{\mathrm{con}}. \tag{7.12}
$$

In this section, we will study the limit behavior of ν^{jt} and $L(\mathbf{t})$ under ν^{jt}. Write $(\mathscr{T}^{\mathbf{p}},\mathscr{G}^{\mathbf{p}})\sim\nu^{\mathrm{jt}}$ for the $\mathbb{T}_m^{\mathrm{ord}}\times\mathbb{G}_m^{\mathrm{con}}$-valued random variable with distribution ν^{jt}. The main aim of this section is the following result for the untilted object. The next section studies the tilted version.

Proposition 7.5. *Let $(\mathscr{T}^{\mathbf{p}},\mathscr{G}^{\mathbf{p}})$ be $\mathbb{T}_m^{\mathrm{ord}}\times\mathbb{G}_m^{\mathrm{con}}$-valued random variable with distribution ν^{jt} viewed as measured metric spaces using the vertex weights $\mathbf{p}$. Then under Assumptions 7.1 and 7.2, as $m\to\infty$ we have*

$$
\left(\mathrm{scl}\left(\sigma(\mathbf{p}),1\right)\mathscr{G}^{\mathbf{p}}, L(\mathscr{T}^{\mathbf{p}})\right) \xrightarrow{\mathrm{w}} \left(\mathscr{G}(2\mathbf{e},\bar{\gamma}\mathbf{e},\mathscr{P}), \exp\left(\bar{\gamma}\int_0^1\mathbf{e}(s)ds\right)\right).
$$

Before diving into the proof, we start by giving an explicit construction of $(\mathscr{T}^{\mathbf{p}},\mathscr{G}^{\mathbf{p}})$ from $(\mathbf{X},\mathscr{P})$, where $\mathbf{X}=(X_i : i\in[m])$ are *iid* Uniform[0,1] r.v.s and $\mathscr{P}$ is an rate one Poisson point process on $\mathbb{R}_+^2$, independent of $\mathbf{X}$. The construction is based on [8] which starts by setting up a

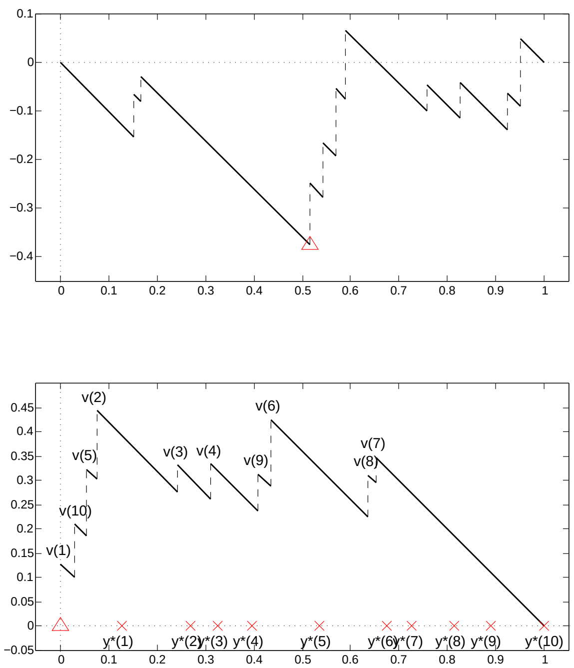

FIGURE 7.1. The functions $F^{\mathbf{p}}$ on the top and the corresponding function $F^{\mathrm{exc},\mathbf{p}}$ for a specific choice of $m = 10$ points and a pmf $\mathbf{p}$.

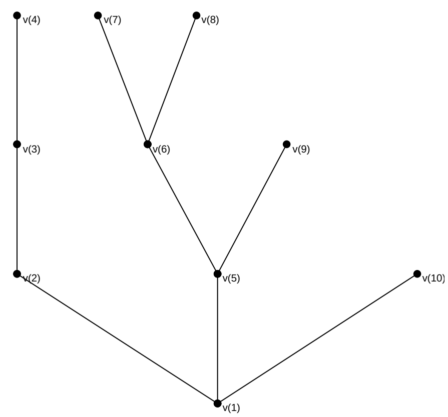

FIGURE 7.2. The tree obtained from the depth-first construction

map $\psi_{\mathbf{p}} : (0,1)^m \to \mathbb{T}_m^{\mathrm{ord}}$ as follows. Fix a collection of distinct points $\mathbf{x} = (x_i : i \in [m]) \in (0,1)^m$. Define

$$F^{\mathbf{p}}(u) := -u + \sum_{i=1}^{m} p_i \mathbb{1}\{x_i \leqslant u\}, \qquad u \in [0,1]. \tag{7.13}$$

Assume that there exists a unique point $v^* \in [m]$ such that $F^{\mathbf{p}}(x_{v^*}-) = \min_{u\in[0,1]} F^{\mathbf{p}}(u)$. Set v^* to be the root of the tree $\psi_{\mathbf{p}}(\mathbf{x})$. Define $y_i := x_i - x_{v^*}$ for $i \in [m]$, and

$$F^{\mathrm{exc},\mathbf{p}}(u) := F^{\mathbf{p}}(x_{v^*} + u \bmod 1) - F^{\mathbf{p}}(x_{v*}-), \qquad 0 \leqslant u < 1.$$

Then $F^{\mathrm{exc},\mathbf{p}}(1-) = 0$ and $F^{\mathrm{exc},\mathbf{p}}(u) > 0$ for $u \in [0,1)$. Extend the definition of $F^{\mathrm{exc},\mathbf{p}}$ to $u \in [0,1]$ by defining $F^{\mathrm{exc},\mathbf{p}}(1) = 0$. We will use $F^{\mathrm{exc},\mathbf{p}}$ to construct a depth-first-search of an ordered tree whose exploration in this depth first manner is encoded by the function $F^{\mathrm{exc},\mathbf{p}}$. This in turn defines the tree $\psi_{\mathbf{p}}(\mathbf{x})$. As before, in this construction we will carry along a set of explored vertices $\mathscr{O}(i)$, active vertices $\mathscr{A}(i)$ and unexplored vertices $\mathscr{U}(i) = [m] \setminus (\mathscr{A}(i) \cup \mathscr{O}(i))$, for $0 \leqslant i \leqslant m$. As before we will view $\mathscr{A}(i)$ as the state of a vertical stack $\mathscr{A}$ after i-th step in the depth-first-search.

Initialize with $\mathscr{O}(0) = \emptyset$, $\mathscr{A}(0) = \{v^*\}$, $\mathscr{U}(0) = [m] \setminus \{v(1)\}$, and define $y^*(0) = 0$. At step $i \in [m]$, let $v(i)$ be the value that is on the top of the stack $\mathscr{A}(i-1)$ and define $y^*(i) := y^*(i-1) + p_{v(i)}$. Define $\mathscr{D}(i) := \{i \in [m] : y^*(i-1) < y_i < y^*(i)\}$. Suppose $\mathscr{D}(i) = \{u(j) : 1 \leqslant i \leqslant k\}$ where we have ordered these vertices in the sequence that they are found in this interval namely

$$y^*(i-1) < y_{u(1)} < \ldots < y_{u(k)} < y^*(i).$$

Update the stack $\mathscr{A}$ as follows:

(i) Delete $v(i)$ from $\mathscr{A}$.
(ii) Push $u(j)$, $1 \leqslant j \leqslant k$, to the top of $\mathscr{A}$ sequentially (so that $u(k)$ will be on the **top** of the stack at the end).

Let $\mathscr{A}(i)$ be the state of the stack after the above operations. Our approach here is not exactly the same as the one in [8], where in their construction the vertices are pushed to the stack in the reverse order. However as remarked in [8] this does not effect the resulting distribution of the tree. Update $\mathscr{O}(i) := \mathscr{O}(i-1) \cup \{v(i)\}$ and $\mathscr{U} := \mathscr{U}(i-1) \setminus \mathscr{D}(i)$.

The tree $\psi_{\mathbf{p}}(\mathbf{x}) \in \mathbb{T}_m^{\mathrm{ord}}$ is constructed by putting the edges $\{(v(i), u) : i \in [m], u \in \mathscr{D}(i)\}$ and using the order prescribed in the above exploration to make the tree an ordered tree. The fact that this procedure actually produces a tree is proved in [8]. So far we have given the construction of a deterministic tree $\psi_{\mathbf{p}}(\mathbf{x})$ using $\mathbf{x} \in (0,1)^m$. Using the collection of uniform random variables $\mathbf{X}$ then results in $\psi_{\mathbf{p}}(\mathbf{X})$ being a random ordered tree in $\mathbb{T}_m^{\mathrm{ord}}$. It is further shown in [8] that $\psi_{\mathbf{p}}(\mathbf{X})$ has the same distribution as an ordered $\mathbf{p}$-tree, i.e., $\psi_{\mathbf{p}}(\mathbf{X})$ has the law $\mathbb{P}_{\mathrm{ord}}$ in (2.8).

We use the same notation to denote the various construct in the above construction when replacing $\mathbf{x}$ with $\mathbf{X}$, so that notations such as $\mathscr{A}(i)$, $\mathscr{D}(i)$ and $y^*(i)$ now correspond to random objects. Define

$$H^{\mathbf{p}}(u) := \text{ height of } v(i) \text{ in } \psi_{\mathbf{p}}(\mathbf{X}), \qquad u \in (y^*(i-1), y^*(i)], i \in [m]. \tag{7.14}$$

Extend $H^{\mathbf{p}}(u)$ to $u = 0$ continuously. $F^{\mathrm{exc},\mathbf{p}}$ in (7.2) and $H^{\mathbf{p}}$ are random elements in $\mathscr{D}([0,1],\mathbb{R})$.

The following result was proved [8, Proposition 3 and Theorem 3] under a set of assumptions stronger than Assumption 7.1. In [8] they also made the remark that their assumptions can be relaxed. It turns out Assumption 7.1 is a sufficient condition for the same result.

Theorem 7.6. *Under Assumptions* 7.1*, as $m \to \infty$ we have*

$$\left(\frac{F^{\mathrm{exc},\mathbf{p}}(\cdot)}{\sigma(\mathbf{p})}, \sigma(\mathbf{p}) H^{\mathbf{p}}(\cdot)\right) \xrightarrow{\mathrm{w}} (\mathbf{e}, 2\mathbf{e}) \tag{7.15}$$

where $\mathbf{e}$ *is a standard Brownian excursion.*

Proof: By [8, Equation (19)], under the assumptions $\lim_{m\to\infty} \sigma(\mathbf{p}) = 0$ and $\lim_{m\to\infty} p_{\max}/\sigma(\mathbf{p}) = 0$, we have as $m \to \infty$,

$$\frac{1}{\sigma(\mathbf{p})} F^{\mathrm{exc},\mathbf{p}}(\cdot) \xrightarrow{\mathrm{w}} \mathbf{e}(\cdot). \tag{7.16}$$

The following lemma can be proved by imitating the proof of [8, Proposition 4]. We relax the assumption about exponential moments used in [8, Proposition 4] with the bound on $p_{\max}$ as in Assumption 7.1, and the price is a stronger assumption on $p_{\min}$. We state the result without a proof.

Lemma 7.7. *Under Assumption* 7.1*, we have as $m \to \infty$,*

$$\sup_{u \in [0,1]} \left| \frac{1}{2} \sigma(\mathbf{p}) H^{\mathbf{p}}(u) - \frac{1}{\sigma(\mathbf{p})} F^{\mathrm{exc},\mathbf{p}}(u) \right| \xrightarrow{\mathrm{P}} 0.$$

The proof of Theorem 7.6 is completed by combining (7.16) and Lemma 7.7. ■

Remark 4. The proof of [8, Proposition 4] uses large deviation inequalities, this is where the assumption on exponential moments is used. The use of large deviation inequalities makes the proof simpler. However, as observed by the authors of [8] (see the remark after the statement of [8, Theorem 3]), it is possible to prove this result simply by making use of Markov inequality and Burkholder-Davis-Gundy inequality (instead of large deviation bounds) and it turns out that Assumption 7.1 is sufficient for this purpose. We will use similar techniques in the proof of Theorem 3.7, so we omit the proof of Lemma 7.7 to avoid repetition.

Next, we will construct a random graph $\psi^G_{\mathbf{p}}(\mathbf{X}) \in \mathbb{G}^{\text{con}}_m$ such that $(\psi^G_{\mathbf{p}}(\mathbf{X}), \psi_{\mathbf{p}}(\mathbf{X})) \stackrel{d}{=} (\mathscr{G}^{\mathbf{p}}, \mathscr{T}^{\mathbf{p}})$ as defined in Theorem 7.5. For $i \in [m]$, let $\mathscr{S}(i) := \mathscr{A}(i-1) \setminus \{v(i)\}$. Define the function $A_m(\cdot)$ on $[0,1]$ via

$$A_m(u) := \sum_{v \in \mathscr{S}(i)} p_v, \qquad \text{for } u \in (y^*(i-1), y^*(i)], i \in [m]. \tag{7.17}$$

Define $\bar{A}_m(u) := aA_m(u)$, $u \in [0,1]$, where a is the scaling constant in the definition of the edge probabilities q_{ij}. Recall that $\mathscr{P}$ is a rate one Poisson point process on $\mathbb{R}^2_+$, independent of $\mathbf{X}$. Let $\bar{A}_m \cap \mathscr{P} := \{(x,y) \in \mathscr{P} : y \leqslant \bar{A}_m(x)\}$. For each point $(x,y) \in \bar{A}_m \cap \mathscr{P}$, define

$$r_m(x,y) = \inf\{x' \geqslant x : \bar{A}_m(x') < y\}. \tag{7.18}$$

Conditioned on $\psi_{\mathbf{p}}(\mathbf{X})$, the graph $\psi^G_{\mathbf{p}}(\mathbf{X})$ is constructed as follows: Suppose $\bar{A}_m \cap \mathscr{P} = \{(x_l, y_l) : l \in [k]\}$. Then for each $l \in [k]$ define $i_l \in [m]$ to be such that $y^*(i_l - 1) < x_l < y^*(i_l)$, and define $j_l \in [m]$ to be such that $y^*(j_l) = r_m(x_l, y_l)$. Let $\psi^G_{\mathbf{p}}(\mathbf{X})$ be the graph obtained by adding edges $(v(i_l), v(j_l))$, $l \in [k]$, to $\psi_{\mathbf{p}}(\mathbf{X})$. There is a small probability that multiple edges are placed between two vertices if there are multiple points in $\mathscr{P}$ that are very close to each other. In that case, let $\psi^G_{\mathbf{p}}(\mathbf{X})$ be the simple graph obtained by replacing all multi-edges with simple edges.

The key observation is that for every edge in $\mathfrak{P}(\psi_{\mathbf{p}}(\mathbf{X}))$ of the form $(v(i), v(j))$ such that $v(j) \in \mathscr{A}(i-1) \setminus \{v(i)\}$, we can find a unique corresponding rectangle in $(\mathbb{R}^+)^2$ below the path $\bar{A}(\cdot)$:

$$R(i,j) := \{(x,y) \in (\mathbb{R}^+)^2 : y^*(i-1) \leqslant x < y^*(i),\ \bar{A}_m(y^*(j)) < y \leqslant \bar{A}_m(y^*(j)-)\}.$$

Notice that these rectangles have the following properties:

- They consist of a partition of $\{(x,y) \in (\mathbb{R}^+)^2 : 0 \leqslant x < 1,\ 0 < y \leqslant \bar{A}_m(x)\}$.
- $R(i,j)$ has width $p_{v(i)}$ and height $ap_{v(j)}$.
- $(v(i), v(j))$ is an edge in $\psi^G_{\mathbf{p}}(\mathbf{X})$ if and only if $R(i,j) \cap \mathscr{P} \neq \emptyset$.

Based on the above observation, since $\mathscr{P}$ is a Poisson point process, we have, for $(v(i), v(j)) \in \mathfrak{P}(\psi_{\mathbf{p}}(\mathbf{X}))$,

$$\mathbb{P}((v(i), v(j)) \text{ is added to } \psi_{\mathbf{p}}(\mathbf{X}) | \psi_{\mathbf{p}}(\mathbf{X})) = 1 - \exp\left(-ap_{v(i)}p_{v(j)}\right). \tag{7.19}$$

Further we have $\mathbb{P}(\psi^G_{\mathbf{p}}(\mathbf{X}) = G | \psi_{\mathbf{p}}(\mathbf{X}) = \mathbf{t}) = \nu^{\text{per}}(G;\mathbf{t})$ and thus

$$(\psi^G_{\mathbf{p}}(\mathbf{X}), \psi_{\mathbf{p}}(\mathbf{X})) \stackrel{d}{=} (\mathscr{G}^{\mathbf{p}}, \mathscr{T}^{\mathbf{p}}). \tag{7.20}$$

Proof of Proposition 7.5: Using (7.15) and Skorohod embedding, we can construct $\{F^{\text{exc},\mathbf{p}}, H^{\mathbf{p}} : m \in \mathbb{N}\}$ on a common probability space Ω_1 such that

$$\left(\frac{F^{\text{exc},\mathbf{p}}(\cdot)}{\sigma(\mathbf{p})}, \sigma(\mathbf{p})H^{\mathbf{p}}(\cdot)\right) \xrightarrow{\text{a.e.}} (\mathbf{e}, 2\mathbf{e}).$$

Let $\mathscr{P}$ be a rate one Poisson point process on $\mathbb{R}_+^2$, independent of $\{F^{\mathrm{exc},\mathbf{p}}, H^{\mathbf{p}} : m \geqslant 1\}$ and the almost sure limit $\mathbf{e}$. By (7.20), we can write

$$(\mathscr{G}^{\mathbf{p}}, \mathscr{T}^{\mathbf{p}}) := (\psi_{\mathbf{p}}^{G}(\mathbf{X}), \psi_{\mathbf{p}}(\mathbf{X})).$$

We start with a preliminary lemma analyzing asymptotics for $A_m(\cdot)$ in (7.17).

Lemma 7.8. *As $n \to \infty$, we have*

$$\sup_{t\in[0,1]} \left| \frac{F^{\mathrm{exc},\mathbf{p}}(t) - A_m(t)}{\sigma(\mathbf{p})} \right| \xrightarrow{\text{a.e.}} 0.$$

Proof: By the definition of $F^{\mathrm{exc},\mathbf{p}}$ we have

$$F^{\mathrm{exc},\mathbf{p}}(y^*(i)) = \sum_{v\in\mathscr{A}(i)} p_v, \quad \text{for } i \in [m]. \tag{7.21}$$

By (7.17), we have

$$A_m(t) = \sum_{v\in\mathscr{S}(i)} p_v = \sum_{v\in\mathscr{A}(i-1)} p_v - p_{v(i)}, \quad \text{for } t \in (y^*(i-1), y^*(i)].$$

Thus

$$\begin{aligned}
&\sup_{t\in(y^*(i-1),y^*(i)]} |A_m(t) - F^{\mathrm{exc},\mathbf{p}}(t)| \\
\leqslant &|A_m(y^*(i)) - F^{\mathrm{exc},\mathbf{p}}(y^*(i-1))| + \sup_{t\in(y^*(i-1),y^*(i)]} |F^{\mathrm{exc},\mathbf{p}}(t) - F^{\mathrm{exc},\mathbf{p}}(y^*(i-1))| \\
= &p_{v(i)} + \sup_{t\in(y^*(i-1),y^*(i)]} |F^{\mathrm{exc},\mathbf{p}}(t) - F^{\mathrm{exc},\mathbf{p}}(y^*(i-1))|
\end{aligned}$$

Denoting $\Delta_m(\delta) := \sup_{0\leqslant s<t\leqslant 1,|s-t|\leqslant\delta} |F^{\mathrm{exc},\mathbf{p}}(s) - F^{\mathrm{exc},\mathbf{p}}(t)|$, then we have

$$\sup_{t\in[0,1]} \left| \frac{F^{\mathrm{exc},\mathbf{p}}(t) - A_m(t)}{\sigma(\mathbf{p})} \right| \leqslant \frac{p_{\max}}{\sigma(\mathbf{p})} + \frac{\Delta_m(p_{\max})}{\sigma(\mathbf{p})}.$$

By Assumption 7.2, we have $p_{\max}/\sigma(\mathbf{p}) \to 0$ and $p_{\max} \to 0$ as $m \to \infty$. In addition, since $\sup_{t\in[0,1]} |F^{\mathrm{exc},\mathbf{p}}(t)/\sigma(\mathbf{p}) - \mathbf{e}(t)| \to 0$ and $\mathbf{e}(\cdot)$ is continuous on $[0,1]$, we have $\Delta_m(p_{\max})/\sigma(\mathbf{p}) \to 0$ as $m \to \infty$ as well. The proof of Lemma 7.8 is completed. ■

By Lemma 7.8 and the construction of the point process $\mathscr{P}$, since $a\sigma(\mathbf{p}) \to \bar{\gamma}$, we have

$$\left(\frac{1}{\sigma(\mathbf{p})} F^{\mathrm{exc},\mathbf{p}}, \sigma(\mathbf{p}) H^{\mathbf{p}}, \bar{A}_m, \mathscr{P} \right) \xrightarrow{\text{a.e.}} (\mathbf{e}, 2\mathbf{e}, \bar{\gamma}\mathbf{e}, \mathscr{P}).$$

Thus there exists $k \in \mathbb{N}_0$ such that for all m large enough

$$\bar{A}_m \cap \mathscr{P} = \{(x_l, y_l) : l = 1, 2, ..., k\}.$$

Recall from Section 4.1 that given any excursion h we can construct a real tree $\mathscr{T}(h)$. Let $(v(i_l), v(j_l))$ be as defined below (7.18), $r(x_l, y_l)$ be as defined in (4.2) by replacing g with $\bar{\gamma}\mathbf{e}$, and $q_{2\mathbf{e}}$ be the canonical map $[0,1] \to \mathscr{T}(2\mathbf{e})$. Then $\mathscr{G}^{\mathbf{p}}$ and $\mathscr{G}(2\mathbf{e}, \bar{\gamma}\mathbf{e}, \mathscr{P})$ are gained by identify the pairs $(v(i_l), v(j_l))$ and $(q_{2\mathbf{e}}(x_l), q_{2\mathbf{e}}(r(x_l, y_l)))$ respectively, for $1 \leqslant l \leqslant k$. Denote

$$\mathscr{G}_m^{\mathbf{p}} := \operatorname{scl}\big(\sigma(\mathbf{p}), 1\big) \cdot \mathscr{G}^{\mathbf{p}} \quad \text{and} \quad \mathscr{T}_m^{\mathbf{p}} := \operatorname{scl}\big(\sigma(\mathbf{p}), 1\big) \cdot \mathscr{T}^{\mathbf{p}}.$$

In order to complete the proof of Proposition 7.5, we will prove the following two lemmas.

Lemma 7.9. $\mathscr{G}_m^{\mathbf{p}} \xrightarrow{\text{a.e.}} \mathscr{G}(2\mathbf{e}, \bar{\gamma}\mathbf{e}, \mathscr{P})$*, as $m \to \infty$.*

Proof: By [4, Lemma 4.2], we need to construct, for each $m \in \mathbb{N}$, a correspondence C_m between $\mathcal{T}_m^{\mathbf{p}}$ and $\mathcal{T}(2\mathbf{e})$ and a measure ξ_m on the space $\mathcal{T}_m^{\mathbf{p}} \times \mathcal{T}(2\mathbf{e})$ such that

(i) $(v(i_l), q_{2\mathbf{e}}(x_l)) \in C_m$ and $(v(j_l), q_{2\mathbf{e}}(r(x_l, y_l))) \in C_m$, for $l = 1, 2, ..., k$.
(ii) $\xi_m(C_m^c) \to 0$ as $m \to \infty$.
(iii) $D(\xi_m) \to 0$ as $m \to \infty$, where $D(\xi_m)$ is the discrepancy defined in (2.3).
(iv) $\mathrm{dis}(C_m) \to 0$ as $m \to \infty$, where $\mathrm{dis}(C_m)$ is the distortion defined in (2.1).

Once the above conditions are verified, by [4, Lemma 4.2], we have

$$d_{\mathrm{GHP}}(\mathcal{G}_m^{\mathbf{p}}, \mathcal{G}(2\mathbf{e}, \bar{\gamma}\mathbf{e}, \mathcal{P})) \leqslant (k+1) \max\left(\frac{1}{2}\mathrm{dis}(C_m), D(\xi_m), \xi_m(C_m^c)\right) \to 0,$$

as $m \to \infty$ and therefore Lemma 7.9 is proved.

Now we describe the construction of C_m and ξ_m. Define

$$\epsilon_m := 2 \sup_{l=1,2,...,k} |r_m(x_l, y_l) - r(x_l, y_l)|. \tag{7.22}$$

By definition of $r(x, y)$, we have

$$\mathbf{e}(x) > \mathbf{e}(r(x_l, y_l)) \text{ for } x \in [x_l, r(x_l, y_l)), l = 1, 2, ..., k.$$

Further by the property of Brownian excursions, for each $\delta > 0$, there exists $x \in [r(x_l, x_l), r(x_l, x_l) + \delta)$ such that $\mathbf{e}(x) < \mathbf{e}(r(x_l, y_l))$. Since $\sup_{t\in[0,1]} |\bar{A}_m(t) - \bar{\gamma}\mathbf{e}(t)| \xrightarrow{\text{a.e.}} 0$, then

$$|r_m(x_l, y_l) - r(x_l, y_l)| \xrightarrow{\text{a.e.}} 0 \text{ as } m \to \infty, \text{ for } l = 1, 2, ..., k.$$

Thus we have $\epsilon_m \xrightarrow{\text{a.e.}} 0$ as $m \to \infty$.

Define the correspondence C_m as

$$C_m := \left\{(v(i), q_{2\mathbf{e}}(x)) : i \in [m], x \in [0 \vee (y^*(i-1) - \epsilon_m), 1 \wedge (y^*(i) + \epsilon_m)]\right\}.$$

By the definition of ϵ_m, the condition (i) is automatically satisfied. Define the measure ξ_m as

$$\xi_m(\{v(i)\} \times A) := \mathrm{Leb}\left(q_{2\mathbf{e}}^{-1}(A) \cap [y^*(i-1), y^*(i)]\right), \tag{7.23}$$

for $i \in [m]$, $A \subset [0,1]$ measurable. Since the map $i \mapsto v(i)$ is 1-1, C_m and ξ_m above are well defined. It is easy to check that $\xi_m(C_m) = 1$ and $D(\xi_m) = 0$, thus the conditions (ii) and (iii) are also satisfied. We only need to check the condition (iv). If $(v(i_1), q_{2\mathbf{e}}(u_1))$ and $(v(i_2), q_{2\mathbf{e}}(u_2))$ are two elements in C_m. Denote d_1 and d_2 for the metric on $\mathcal{T}_m^{\mathbf{p}}$ and $\mathcal{T}(2\mathbf{e})$ respectively. Observe that if either one is an ancestor of the other, we have

$$d_1(v(i_1), v(i_2))/\sigma(\mathbf{p}) = H^{\mathbf{p}}(y^*(i_1)) + H^{\mathbf{p}}(y^*(i_2)) - 2 \inf_{t\in[y^*(i_1), y^*(i_2)]} H^{\mathbf{p}}(t),$$

otherwise:

$$d_1(v(i_1), v(i_2))/\sigma(\mathbf{p}) = H^{\mathbf{p}}(y^*(i_1)) + H^{\mathbf{p}}(y^*(i_2)) - 2 \inf_{t\in[y^*(i_1), y^*(i_2)]} H^{\mathbf{p}}(t) + 2.$$

Thus we have

$$
\begin{aligned}
&|d_1(v(i_1), v(i_2)) - d_2(q_{2\mathbf{e}}(u_1), q_{2\mathbf{e}}(u_2))| \\
&\leqslant \Bigg| \sigma(\mathbf{p}) H^{\mathbf{p}}(y^*(\bar{i}_1)) + \sigma(\mathbf{p}) H^{\mathbf{p}}(y^*(\bar{i}_2)) - 2\sigma(\mathbf{p}) \inf_{t\in[y^*(\bar{i}_1), y^*(\bar{i}_2)]} H^{\mathbf{p}}(t) \\
&\quad -2\mathbf{e}(u_1) - 2\mathbf{e}(u_2) + 4\inf_{t\in[u_1,u_2]} \mathbf{e}(t)\Bigg| + 2\sigma(\mathbf{p}) \\
&\leqslant 4 \sup_{t\in[0,1]} \left|\sigma(\mathbf{p}) H^{\mathbf{p}}(t) - 2\mathbf{e}(t)\right| + 8\Delta_{\mathbf{e}}(\epsilon_m) + 2\sigma(\mathbf{p}),
\end{aligned}
$$

where $\Delta_{\mathbf{e}}(\delta) = \sup_{0\leqslant s<t\leqslant 1, |s-t|<\delta} |\mathbf{e}(s) - \mathbf{e}(t)|$, for $\delta > 0$. Thus the above expression, which is also a bound on $\mathrm{dis}(C_m)$, goes to zero as $m \to \infty$. Condition (iv) is verified. The proof of Lemma 7.9 is completed. ■

The last lemma that we need to complete the proof of Proposition 7.5 is the following:

Lemma 7.10. *As $m \to \infty$,*

$$
L(\mathcal{T}^{\mathbf{p}}) = \left[\prod_{(i,j)\in E(\mathcal{T}^{\mathbf{p}})} \frac{\exp(ap_ip_j) - 1}{ap_ip_j}\right] \exp\left(\sum_{(i,j)\in \mathfrak{P}(\mathcal{T}^{\mathbf{p}})} ap_ip_j\right) \xrightarrow{\text{a.e.}} \exp\left(\bar{\gamma}\int_0^1 \mathbf{e}(s)\,ds\right).
$$

Proof: By the basic inequality $(e^x - 1)/x \leqslant e^x$ for $x > 0$, we have for $\mathbf{t} \in \mathbb{T}_m^{\mathrm{ord}}$,

$$
\prod_{(i,j)\in E(\mathbf{t})} \frac{\exp(ap_ip_j) - 1}{ap_ip_j} \leqslant \exp\left(a \sum_{(i,j)\in E(\mathbf{t})} p_ip_j\right) \leqslant \exp(ap_{\max}),
$$

where the last inequality follows using the fact that $\mathbf{t}$ is a tree, thus for each $(i,j) \in E(\mathbf{t})$ such that i is the parent of j we have $p_ip_j \leqslant p_{\max}p_j$. By Assumption 7.2, we have $ap_{\max} \to 0$, thus the above display goes to one as $m \to \infty$. Then notice that

$$
\sum_{(i,j)\in \mathfrak{P}(\mathcal{T}^{\mathbf{p}})} ap_ip_j = a \sum_{i\in[m]} \sum_{j\in\mathcal{S}(i)} p_ip_j = \int_0^1 \bar{A}_m(s)\,ds \to \bar{\gamma}\int_0^1 \mathbf{e}(s)\,ds,
$$

as $m \to \infty$, where the last convergence follows since $\bar{A}_m \xrightarrow{\text{a.e.}} \bar{\gamma}\mathbf{e}$. The proof of Lemma 7.10 is thus completed. ■

Completing the proof of Proposition 7.5: The proof follows from Lemma 7.9 and 7.10. ■

7.3. **Uniform integrability of the tilt.** The last key ingredient in proving Theorem 7.3 we need is the tightness of $L(\mathcal{T}^{\mathbf{p}})$. We start with a concentration inequality on $\|F^{\mathrm{exc},\mathbf{p}}\|_\infty$ that allows us to control the tilt appearing on the right hand side of (7.6). A key step is a concentration inequality for partial sums when sampling without replacement, a problem studied in a slightly different setting in [41].

Lemma 7.11. *Recall that $\sigma(\mathbf{p}) = \sqrt{\sum_{i=1}^m p_i^2}$ and $p_{\max} = \max_{i\in[m]} p_i$. Assume that*

$$
4p_{\max} \leqslant x \leqslant \frac{16\sigma^2(\mathbf{p})}{p_{\max}}. \tag{7.24}
$$

Then we have

$$
\mathbb{P}(\|F^{\mathrm{exc},\mathbf{p}}\|_\infty > x) \leqslant 12\exp\left(-\frac{x^2}{1024(\sigma(\mathbf{p}))^2}\right).
$$

Proof: Write $X_1,\dots X_m$ for the *iid* $U(0,1)$ random variables used to construct $F^{\mathbf{p}}$ which is then used to construct $F^{\text{exc},\mathbf{p}}$ from (7.2). Let $X_{(1)} < X_{(2)} < \cdots X_{(m)}$ be the corresponding order statistics and let π denote the corresponding permutation of $[m]$ namely $X_{(i)} = X_{\pi(i)}$. Obviously π is a uniform random permutation. Now by definition

$$\|F^{\text{exc},\mathbf{p}}\|_\infty \leqslant \sup_{t\in[0,1]} F^{\mathbf{p}}(t) + \left|\inf_{t\in[0,1]} F^{\mathbf{p}}(t)\right|. \tag{7.25}$$

Let us analyze the first term. Define $\vartheta_i := -X_{(i)} + \sum_{j=1}^{i} p_{\pi(j)}$, namely the value $F^{\mathbf{p}}(\cdot)$ at each location with a positive jump. Since $\sup_{t\in[0,1]} F^{\mathbf{p}}(t) \leqslant \sup_{i\in[m]} |\vartheta_i|$, we consider

$$\begin{aligned}\mathbb{P}\left(\sup_{i\in[m]} |\vartheta_i| \geqslant \frac{x}{2}\right) &\leqslant \mathbb{P}\left(\sup_{i\in[m]} \left|-X_{(i)} + \frac{i}{m}\right| \geqslant \frac{x}{4}\right) + \mathbb{P}\left(\sup_{i\in[m]} \left|\sum_{j=1}^{i} p_{\pi(j)} - \frac{i}{m}\right| \geqslant \frac{x}{4}\right)\\ &:= T_1 + T_2 \end{aligned}\tag{7.26}$$

Let $F_m(u) := n^{-1}\sum_{i=1}^m \mathbb{1}_{\{X_i\leqslant u\}}$, $u \in [0,1]$, denote the empirical distribution function of $(X_i : 1 \leqslant i \leqslant n)$ so that $F_m(X_{(i)}) = i/m$. Thus by the DKW inequality [34]

$$T_1 = \mathbb{P}\left(\sup_{i\in[m]} \left|F_m(X_{(i)}) - X_{(i)}\right| \geqslant \frac{x}{4}\right) = \mathbb{P}\left(\sup_{u\in[0,1]} |F_m(u) - u| \geqslant \frac{x}{4}\right) \leqslant 2\exp(-mx^2/8). \tag{7.27}$$

We now analyze T_2. Since $\mathbf{p}$ is a probability distribution, for any $m/2 \leqslant k \leqslant m-1$, $|\sum_{j=1}^k p_{\pi(j)} - k/m| = |\sum_{j=k+1}^m p_{\pi(j)} - (m-k)/m|$. Without loss of generality, assume m is even. Define

$$p(m,x) := \mathbb{P}\left(\sup_{k\in[m/2]} \left|\sum_{j=1}^{k} p_{\pi(j)} - \frac{k}{m}\right| \geqslant \frac{x}{4}\right)$$

. Now

$$\begin{aligned} T_2 &\leqslant p(m,x) + \mathbb{P}\left(\sup_{m/2\leqslant k\leqslant m-1} \left|\sum_{j=1}^{k} p_{\pi(j)} - \frac{k}{m}\right| \geqslant \frac{x}{4}\right)\\ &\leqslant p(m,x) + \mathbb{P}\left(\sup_{m/2\leqslant k\leqslant m-1} \left|\sum_{j=k+1}^{m} p_{\pi(j)} - \frac{m-k}{m}\right| \geqslant \frac{x}{4}\right)\\ &= p(m,x) + \mathbb{P}\left(\sup_{k'\in[m/2]} \left|\sum_{l=1}^{k'} p_{\pi(m-l+1)} - \frac{k'}{m}\right| \geqslant \frac{x}{4}\right) = 2p(m,x). \end{aligned}\tag{7.28}$$

where the last line follows by noting that the permutation π' defined via $\pi(l) = \pi(m-l+1)$ is again a uniform permutation on $[m]$. We are now left with bounding $p(m,x)$. Assume that we generate π by sequentially drawing without replacement from $[m]$. For $k \geqslant 1$, let $\mathscr{F}_k$ denote the σ-field generated by $(\pi(1),\dots,\pi(k))$. Writing $S_0 = 0$ and $S_k := \sum_{j=1}^k p_{\pi(j)}$, $k \in [m]$, it is easy to check that $\{Y_k : k = 0,1,\dots,m-1\}$ defined by the following is an $\mathscr{F}_k$-martingale:

$$Y_k := \frac{S_k - k/m}{m-k}, \text{ for } k = 0,1,\dots,m-1.$$

Note that $\sup_{i\in[m/2]} |S_i - i/m| \leqslant m \sup_{i\in[m/2]} |Y_i|$, thus

$$p(m,x) \leqslant \mathbb{P}\left(\sup_{k\in[m/2]} |Y_k| \geqslant \frac{x}{4m}\right), \tag{7.29}$$

For $h > 0$, since $\exp(hx) > 0$ is convex in x, then $\exp(hY_k)$ is a sub-martingale. Hence,

$$\mathbb{P}\left(\sup_{k\in[m/2]} Y_k \geqslant \frac{x}{4m}\right) \leqslant \exp(-\frac{hx}{4m})\,\mathbb{E}\left[\exp(hY_{m/2})\right]. \tag{7.30}$$

By a similar bound on $\mathbb{P}\left(\inf_{k\in[m/2]} Y_k \leqslant -\frac{x}{4m}\right)$ and the fact $Y_{m/2} \overset{d}{=} -Y_{m/2}$, following (7.29) we have

$$p(m,x) \leqslant 2\exp(-\frac{hx}{4m})\,\mathbb{E}\left[\exp(hY_{m/2})\right] = 2\exp(-\frac{hx}{4m})\,\mathbb{E}\left[\exp\left(\frac{2h}{m}S_{m/2} - \frac{h}{m}\right)\right]. \tag{7.31}$$

Now we use the standard technique of bounding the moment generating function of $S_{m/2}$ by repeatedly conditioning on the previous time steps. Note that for $0 < \delta < 1/p_{\max}$ and $k \in [m/2]$ we have

$$\begin{aligned}
\mathbb{E}[\exp(\delta p_{\pi(k+1)}) \mid \mathscr{F}_k] =& \frac{1}{m-k} \sum_{j\notin\{v(i):i\in[k]\}} \exp(\delta p_j) \\
\leqslant& \frac{1}{m-k} \sum_{j\notin\{v(i):i\in[k]\}} (1 + \delta p_j + \delta^2 p_j^2) \\
\leqslant& 1 + \frac{\delta}{m-k}(1 - \sum_{j\in[k]} p_{\pi(j)}) + \frac{2\delta^2\sigma^2(\mathbf{p})}{m} \\
\leqslant& \exp\left(\frac{\delta}{m-k}(1 - S_k) + \frac{2\delta^2\sigma^2(\mathbf{p})}{m}\right),
\end{aligned} \tag{7.32}$$

where the second line uses the fact that $e^x < 1 + x + x^2$ for $x \in [0,1]$ and the third line uses the fact $\sum_{j\in[m]} p_{\pi(j)} = 1$ and $k \leqslant m/2$. Using (7.32) repeatedly in evaluating $\mathbb{E}[\exp(\delta S_k)]$ for $k \leqslant m/2$, we have

$$\begin{aligned}
\mathbb{E}[\exp(\delta S_k)] =& \mathbb{E}\left[\exp(\delta S_{k-1})\,\mathbb{E}\left[\exp(\delta p_{\pi(k)}) \mid \mathscr{F}_{k-1}\right]\right] \\
\leqslant& \mathbb{E}\left[\exp(\delta S_{k-1})\exp\left(\frac{\delta}{m-(k-1)}(1 - S_{k-1}) + \frac{2\delta^2\sigma^2(\mathbf{p})}{m}\right)\right] \\
=& \mathbb{E}\left[\exp\left(\frac{m-k}{m-k+1}\delta S_{k-1}\right)\right]\exp\left(\frac{2\delta^2\sigma^2(\mathbf{p})}{m}\right)\exp\left(\frac{\delta}{m-k+1}\right) \\
\leqslant& \mathbb{E}\left[\exp\left(\frac{m-k}{m-k+1}\cdot\frac{m-k+1}{m-k+2}\delta S_{k-2}\right)\right]\exp\left(2\cdot\frac{2\delta^2\sigma^2(\mathbf{p})}{m}\right) \\
&\times \exp\left(\frac{\delta(m-k)}{(m-k)(m-k+1)} + \frac{\delta(m-k)}{(m-k+1)(m-k+2)}\right).
\end{aligned}$$

Proceeding inductively, we have

$$\begin{aligned}
\mathbb{E}[\exp(\delta S_k)] \leqslant& \mathbb{E}\left[\frac{m-k}{m}\delta S_0\right]\exp\left(k\cdot\frac{2\delta^2\sigma^2(\mathbf{p})}{m} + (m-k)\delta\cdot\sum_{j=0}^{k-1}\frac{1}{(m-k+j)(m-k+j+1)}\right) \\
=& \exp\left(k\cdot\frac{2\delta^2\sigma^2(\mathbf{p})}{m} + (m-k)\delta\cdot\frac{k}{m(m-k)}\right).
\end{aligned}$$

Note that in the l-th iteration of applying (7.32), δ is replaced by $\delta(m-k)/(m-k+l-1)$, which is less than δ. Therefore, by assuming $\delta < 1/p_{\max}$, all iterative use of (7.32) are valid. Taking

$k = m/2$ in the above inequality, we have

$$\mathbb{E}[\exp(\delta S_{m/2})] \leqslant \exp(\delta^2\sigma^2(\mathbf{p}) + \delta/2).$$

Using the above bound with $\delta = 2h/m$ in (7.31), we have

$$p(m,x) \leqslant 2\exp\left(-\frac{hx}{4m} + \frac{4h^2\sigma^2(\mathbf{p})}{m^2} + \frac{h}{m} - \frac{h}{m}\right) \leqslant 2\exp\left(-\frac{x^2}{256\sigma^2(\mathbf{p})}\right), \tag{7.33}$$

where the last inequality follows from taking $h = mx/32\sigma^2(\mathbf{p})$. By our choice of δ and h, the restriction $\delta < 1/p_{\max}$ reduces to the upper bound in the assumption (7.24).

Now combining (7.27), (7.28) and (7.24), we have

$$\mathbb{P}\left(\sup_{t\in[0,1]} F^{\mathbf{p}}(t) \geqslant \frac{x}{2}\right) \leq \mathbb{P}\left(\sup_{i\in[m]} |\vartheta_i| \geqslant \frac{x}{2}\right) \leqslant 2\exp\left(-\frac{mx^2}{8}\right) + 4\exp\left(-\frac{x^2}{256\sigma^2(\mathbf{p})}\right). \tag{7.34}$$

This tackles the first term in (7.25). To deal with the second term, define $\vartheta'_i = -X_{(i)} + \sum_{j=1}^{i-1} p_{\pi(j)}$ so that for ϑ_i as defined after (7.25), $\vartheta_i = \vartheta'_i + p_{\pi(i)}$. Then

$$\left|\inf_{t\in[0,1]} F^{\mathbf{p}}(t)\right| = \sup_{i\in[m]} |\vartheta'_i| \leqslant \sup_{i\in[m]} |\vartheta_i| + p_{\max}.$$

By assumption we have $p_{\max} < x/4$, using we have

$$\begin{aligned}\mathbb{P}\left(|\inf_{t\in[0,1]} F^{\mathbf{p}}(t)| \geqslant \frac{x}{2}\right) &\leqslant \mathbb{P}\left(\sup_{i\in[m]} |\vartheta_i| \geqslant \frac{x}{4}\right)\\ &\leqslant 2\exp\left(-\frac{mx^2}{32}\right) + 4\exp\left(-\frac{x^2}{1024\sigma^2(\mathbf{p})}\right).\end{aligned}$$

This together with (7.25), (7.34) and the fact $m\sigma^2(\mathbf{p}) \geqslant 1$ completes the proof of Lemma 7.11. ■

Corollary 7.12. *For any $B > 0$ satisfying*

$$\frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2}} \leqslant \sqrt{1/8B}, \tag{7.35}$$

there exists $K_{7.12} = K_{7.12}(B)$ such that

$$\mathbb{E}[\exp(B\|F^{\mathrm{exc},\mathbf{p}}\|_\infty/\sigma(\mathbf{p}))] \leqslant K_{7.12}.$$

Proof: By the trivial bound $\|F^{\mathrm{exc},\mathbf{p}}\|_\infty \leqslant 1$, we have

$$\mathbb{E}\left[\exp\left(\frac{B\|F^{\mathrm{exc},\mathbf{p}}\|_\infty}{\sigma(\mathbf{p})}\right)\right] = \int_0^{1/\sigma(\mathbf{p})} B\exp(By)\,\mathbb{P}\left(\|F^{\mathrm{exc},\mathbf{p}}\|_\infty \geqslant y\sigma(\mathbf{p})\right) dy. \tag{7.36}$$

Decomposing the integral over the intervals $[0, 4p_{\max}/\sigma(\mathbf{p})]$, $[4p_{\max}/\sigma(\mathbf{p}), 16\sigma(\mathbf{p})/p_{\max}]$ and $[16\sigma(\mathbf{p})/p_{\max}, 1/\sigma(\mathbf{p})]$, applying Lemma 7.11 to the second interval gives

$$\begin{aligned}\mathbb{E}\left[\exp\left(\frac{B\|F^{\mathrm{exc},\mathbf{p}}\|_\infty}{\sigma(\mathbf{p})}\right)\right] \leqslant& \frac{4Bp_{\max}}{\sigma(\mathbf{p})}\exp\left(\frac{4p_{\max}}{\sigma(\mathbf{p})}\right) + \int_{4p_{\max}/\sigma(\mathbf{p})}^{16\sigma(\mathbf{p})/p_{\max}} B\exp\left(By - \frac{y^2}{1024}\right) dy\\ &+ \mathbb{P}\left(\|F^{\mathrm{exc},\mathbf{p}}\|_\infty > 16\sigma^2(\mathbf{p})/p_{\max}\right)\cdot\exp\left(\frac{B}{\sigma(\mathbf{p})}\right),\\ :=& \mathscr{B}_1 + \mathscr{B}_2 + \mathscr{B}_3.\end{aligned} \tag{7.37}$$

For the first two terms above, using (7.35) and $\sigma(\mathbf{p}) \leqslant 1$, we have

$$\mathscr{B}_1 \leqslant \frac{4B}{\sqrt{8B}} \exp\left(\frac{4}{\sqrt{8B}}\right), \quad \mathscr{B}_2 \leqslant \int_0^\infty B \exp\left(By - \frac{y^2}{1024}\right) dy.$$

For $\mathscr{B}_3$, using (7.35) and Lemma 7.11, we have

$$\mathscr{B}_3 \leqslant 12 \exp\left(-\frac{\sigma^2(\mathbf{p})}{4p_{\max}^2} + \frac{B}{\sigma(\mathbf{p})}\right) \leqslant 12 \exp\left(-\frac{B}{\sigma(\mathbf{p})}\right) = 12e^{-B}. \tag{7.38}$$

The proof of Corollary 7.12 is completed. ■

Corollary 7.13. *Assume that* $\gamma > 0$, $B_1 > 0$, *and* $B_2 \in (0, 1/\sqrt{8\gamma B_1}]$ *satisfy*

$$a\sigma(\mathbf{p}) \leqslant B_1, \quad \frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2}} \leqslant B_2.$$

Let $\mathscr{T}^{\mathbf{p}}$ *be a* $\mathbb{T}_m^{\text{ord}}$*-valued random variable with distribution* $\mathbb{P}_{\text{ord}}$*, and* $L(\cdot)$ *be as defined in* (7.6). *Then there exists a constant* $K_{7.13} = K_{7.13}(\gamma, B_1, B_2) > 0$ *such that*

$$\mathbb{E}[L^\gamma(\mathscr{T}^{\mathbf{p}})] < K_{7.13}.$$

In particular, when $p_{\max}/[\sigma(\mathbf{p})]^{3/2} \to 0$ *and* $a\sigma(\mathbf{p}) \to \bar{\gamma}$ *as* $m \to \infty$*, the sequence* $\left\{L(\mathscr{T}^{\mathbf{p}}) : m \geqslant 1\right\}$ *is uniformly integrable.*

Proof: Recall $F^{\text{exc},\mathbf{p}}$ from (7.13) and $\bar{A}_m$ from below (7.17). Let $\mathbf{X} = (X_i : i \in [m])$ be the *iid* Uniform[0,1] random variables used in the definition of $F^{\text{exc},\mathbf{p}}$ and $\bar{A}_m$. Define $\mathscr{T}^{\mathbf{p}} = \psi_{\mathbf{p}}(\mathbf{X})$ thus $\mathscr{T}^{\mathbf{p}}$ has the law $\mathbb{P}_{\text{ord}}$. We have

$$L(\mathscr{T}^{\mathbf{p}}) \leqslant \exp(ap_{\max}) \exp\left(\int_0^1 \bar{A}_m(s)\, ds\right) \leqslant \exp(B_1 B_2) \exp\left(\frac{B_1 \|F^{\text{exc},\mathbf{p}}\|_\infty}{\sigma(\mathbf{p})}\right),$$

where the last inequality uses the fact $\|A_m\|_\infty \leqslant \|F^{\text{exc},\mathbf{p}}\|_\infty$ (see the proof of Lemma 7.8). Then the corollary directly follows from Corollary 7.12, and we have $K_{7.13} = e^{\gamma B_1 B_2} K_{7.12}(\gamma B_1)$. Taking $\gamma > 1$ we have the uniform integrability of $L(\mathscr{T}^{\mathbf{p}})$. ■

Now we are ready to give the proof of Theorem 7.3.

Proof of Theorem 7.3: Denote scl_m for the scaling operator

$$\text{scl}_m = \text{scl}\big(\sigma(\mathbf{p}), 1\big).$$

Let $(\mathscr{T}^{\mathbf{p}}, \mathscr{G}^{\mathbf{p}})$ has the law ν^{jt} as in (7.12), and $(\tilde{\mathscr{T}}^{\mathbf{p}}, \tilde{\mathscr{G}}^{\mathbf{p}})$ has the law $\tilde{\nu}^{\text{jt}}$ as in (7.11). We want to show that for any bounded continuous function $f(\cdot)$ on $\mathscr{S}$,

$$\mathbb{E}[f(\text{scl}_m \cdot \tilde{\mathscr{G}}^{\mathbf{p}})] \to \mathbb{E}[f(\mathscr{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}}, \bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}}, \mathscr{P}))], \quad \text{as } m \to \infty.$$

Define $g_f(\mathbf{t})$ for $\mathbf{t} \in \mathbb{T}_m^{\text{ord}}$ as

$$g_f(\mathbf{t}) := \sum_{G \in \mathbb{G}_m^{\text{con}}} f(\text{scl}_m G) \nu^{\text{per}}(G; \mathbf{t}).$$

By the definition of ν^{jt} and $\tilde{\nu}^{\text{jt}}$, we have $\mathbb{E}[f(\text{scl}_m \mathscr{G}^{\mathbf{p}}) \mid \mathscr{T}^{\mathbf{p}}] = g_f(\mathscr{T}^{\mathbf{p}})$ and $\mathbb{E}[f(\text{scl}_m \cdot \tilde{\mathscr{G}}^{\mathbf{p}}) \mid \tilde{\mathscr{T}}^{\mathbf{p}}] = g_f(\tilde{\mathscr{T}}^{\mathbf{p}})$. Then by (7.7), we have

$$\mathbb{E}[f(\text{scl}_m \cdot \tilde{\mathscr{G}}^{\mathbf{p}})] = \mathbb{E}[g_f(\tilde{\mathscr{T}}^{\mathbf{p}})] = \frac{\mathbb{E}[g_f(\mathscr{T}^{\mathbf{p}}) L(\mathscr{T}^{\mathbf{p}})]}{\mathbb{E}[L(\mathscr{T}^{\mathbf{p}})]} = \frac{\mathbb{E}[f(\text{scl}_m \mathscr{G}^{\mathbf{p}}) L(\mathscr{T}^{\mathbf{p}})]}{\mathbb{E}[L(\mathscr{T}^{\mathbf{p}})]}. \tag{7.39}$$

By Proposition 7.5 we have the joint convergence

$$L(\mathscr{T}^{\mathbf{p}}) \xrightarrow{\mathrm{w}} \exp\left(\bar{\gamma}\int_0^1 \mathbf{e}(s)ds\right),$$

$$f(\mathrm{scl}_m \mathscr{G}^{\mathbf{p}})L(\mathscr{T}^{\mathbf{p}}) \xrightarrow{\mathrm{w}} f(\mathscr{G}(2\mathbf{e},\bar{\gamma}\mathbf{e},\mathscr{P}))\exp\left(\bar{\gamma}\int_0^1 \mathbf{e}(s)ds\right).$$

By (7.39), the above convergence and the uniform integrability of $L(\mathscr{T}^{\mathbf{p}})$ (Lemma 7.13), we have

$$\lim_{m\to\infty}\mathbb{E}[f(\mathrm{scl}_m\cdot\mathscr{G}^{\mathbf{p}})] = \frac{\mathbb{E}\left[f(\mathscr{G}(2\mathbf{e},\bar{\gamma}\mathbf{e},\mathscr{P}))\exp\left(\bar{\gamma}\int_0^1\mathbf{e}(s)ds\right)\right]}{\mathbb{E}\left[\exp\left(\bar{\gamma}\int_0^1\mathbf{e}(s)ds\right)\right]} = \mathbb{E}\left[f(\mathscr{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}},\bar{\gamma}\tilde{\mathbf{e}}^{\bar{\gamma}},\mathscr{P}))\right],$$

where $\tilde{\mathbf{e}}^{\bar{\gamma}}$ is the tilted Brownian excursion defined in (4.8). The proof of Theorem 7.3 is completed. ■

8. Size-biased reordering and component exploration

Recall the definition of $\mathbf{Z} = (Z_i : i \geqslant 1)$ as in (4.6). From Theorem 4.1, we have

$$\left(\frac{|\mathscr{C}_n^{(i)}|}{n^{2/3}} : i \geqslant 1\right) \xrightarrow{\mathrm{w}} \mathbf{Z},$$

as $n\to\infty$ in the l^2-topology therefore also in the product topology. The next proposition gives more asymptotic properties for the weights of vertices in each component.

Proposition 8.1. *Recall that $\mathscr{C}_n^{(i)}$ is the i-th largest component of $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w},\lambda)$ for $i \geqslant 1$. Assume that the conditions in Assumption* 3.1 *(a) and (b) hold. Further, assume that $w_{\max} = o(n^{1/3})$. Then, for fixed $i \ge 1$, we have*

$$\left(\frac{|\mathscr{C}_n^{(i)}|}{n^{2/3}}, \frac{\sum_{v\in\mathscr{C}_n^{(i)}} w_v}{n^{2/3}}, \frac{\sum_{v\in\mathscr{C}_n^{(i)}} w_v^2}{n^{2/3}}\right) \xrightarrow{\mathrm{w}} \left(Z_i, Z_i, \frac{\sigma_3 Z_i}{\sigma_1}\right) \text{ as } n\to\infty. \tag{8.1}$$

Proof: We start with the proof of the convergence of component sizes in Theorem 4.1 proved in [11]. Recall that given a set $[n]$ and an associated weight sequence $\{w_v : v\in[n]\}$ with $w_v > 0$, a size biased reordering is a random reordering of $[n]$ as $(v(1), v(2),\dots,v(n))$ using the weight sequence where

$$\mathbb{P}(v(1)=j) \propto w_j, \qquad j\in[n], \tag{8.2}$$

and having selected $\{v(1),\dots,v(j-1)\}$, $v(j)$ is selected from $[n]\setminus\{v(i) : 1\leqslant i\leqslant j-1\}$ with probability proportional to the corresponding weights w_v, $v\in[n]\setminus\{v(i):1\leqslant i\leqslant j-1\}$.

Now we describe the construction. We simultaneously construct the graph $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w},\lambda)$ and explore it in a breadth-first manner. For all ordered distinct pairs of vertices (u,v), $u,v\in[n]$, $u\neq v$ let $\{\xi_{uv} : u\neq v\in[n]\}$ be a collection of independent exponential random variables with rate

$$r_{uv} := \left(1+\frac{\lambda}{n^{1/3}}\right)\frac{w_v}{l_n}. \tag{8.3}$$

To initiate the exploration process, start by selecting the first vertex $v(1)\in[n]$ using (8.2). For $i\in[n]\setminus\{v(1)\}$ arrange the exponential random variables $\xi_{v(1)i}$ in increasing order as

$$\xi_{v(1),v'(1)} < \xi_{v(1),v'(2)} < \cdots.$$

The neighbors of $v(1)$ can then be constructed as the set $\mathcal{N}_1 = \{i : \xi_{v(1),i} \leqslant x_{v(1)}\}$. Let $c(1) = |\mathcal{N}_1|$ and label these $c(1)$ vertices as $v(2) = v'(1),\ldots,v(c(1)+1) = v'(c(1))$, thus in increasing order of the associated exponential random variables. Now explore the neighbors of $v(2)$ not yet found by the exploration process (i.e. using the collection of random variables $\xi_{v(2)u}$, $u \notin \{v(1),\ldots v(c(1)+1)\}$) and list them as $v(c(1)+2),\ldots,v(c(1)+1+c(2))$, as before in increasing order of the values of the associated exponential random variables. Repeat this process with $v(3),\ldots,v(c(1)+1)$. Next, move to $v(c(1)+2)$. In general, explore all vertices in a generation and then move to the next generation. If the next generation is empty, then we have finished exploring a component. Then, we select a new vertex v amongst the yet to be explored vertices with probability proportional to their weight w_v and continue as before. It can be easily checked that the resulting graph has the same distribution as $\mathcal{G}_n^{\mathrm{nr}}(\mathbf{w},\lambda)$.

This exploration process results in an ordering of the vertex set $[n]$ as $(v(1),v(2),\ldots,v(n))$. Consider the walk associated with the process

$$S_n(0) = 0, \qquad S_n(i) = S_n(i-1) + c(i) - 1. \tag{8.4}$$

The construction satisfies

(i) the ordering $(v(1),v(2),\ldots,v(n))$ has the same distribution as the size-biased re-ordering of the vertex set $[n]$ using the vertex weight sequence $\mathbf{w}$.

(ii) The walk $\{S_n(i) : i \geqslant 0\}$ encodes the sizes of components (see [7]) in the following sense. Write $T_{-k} = \min\{i : S_n(i) = -k\}$. The number of vertices in the first component explored by the walk (not necessarily the largest component) is given by $|\tilde{\mathscr{C}}_1| = T_{-1}$ the size of the second component explored by the walk is given by $|\tilde{\mathscr{C}}_1| := T_{-2} - T_{-1}$ and so on and further for any $j \geqslant 1$

$$S_n(T_{-j}) = -j, \qquad S_n(i) > -j \text{ for } \qquad T_{-(j-1)} < i < T_{-j} \tag{8.5}$$

Thus excursions beyond past minima encode sizes of components in the order seen by the walk. By [7], Theorem 4.1 was proven in [11] by showing that

$$\left\{\frac{1}{n^{1/3}} S_n(sn^{2/3}) : s \geqslant 0\right\} \xrightarrow{\mathrm{w}} \left\{W^{\lambda}_{\sqrt{\frac{\sigma_3}{\sigma_1}},\frac{\sigma_3}{\sigma_1^2}}(s) : s \geqslant 0\right\}. \tag{8.6}$$

where $W^{\lambda}_{\kappa,\sigma}(\cdot)$ is the inhomogeneous Brownian motion as in (4.3) and convergence is in the Skorohod metric $D(\mathbb{R}_+,\mathbb{R})$. By the techniques in [7], excursions beyond past minima of $S_n(\cdot)$ arranged in decreasing order converge to excursion beyond past minima of $W^{\lambda}_{\kappa,\sigma}$. In this construct, the vertices are ordered in an size-biased-ordering, and the vertices within each component consists of a consecutive subsequence of the size-biased-ordering. The next lemma studied partial sum of a size-biased-ordering in a general setting.

Lemma 8.2. *Let $\mathbf{w} = \mathbf{w}^{(n)} = \{w_i^{(n)} > 0 : i \in [n]\}$ be a set of weights, and $\mathbf{u} = \mathbf{u}^{(n)} = \{u_i^{(n)} : i \in [n]\}$ be a non-negative function on $[n]$. Let $m = m(n) \leqslant n$ be a increasing sequence of integers. We omit n in the notation in the rest of the lemma. Let $\{v(i) \in [n] : i \in [n]\}$ be a size-biased random reordering of the indexes based on the weight $\mathbf{w}$. Denote $w_{(i)} := w_{v(i)}$ and $u_{(i)} = u_{v(i)}$ for $i \in [n]$. Let $w_{\max} := \max_{i\in[n]} w_i$ and $u_{\max} := \max_{i\in[n]} u_i$. Define $c_n := \sum_{i\in[n]} w_i u_i / \sum_{i\in[n]} w_i$. Assume that*

$$\lim_n \frac{m w_{\max}}{\sum_{i\in[n]} w_i} = 0 \quad \textit{and} \quad \lim_n \frac{u_{\max}}{m c_n} = 0.$$

Define $Y(t) := (\sum_{i=1}^{\lfloor mt \rfloor} u_{(i)})/mc_n$*, for* $t \in [0,\infty)$*, with* $u_{(i)} := 0$ *for* $i > n$*. Then*

$$\sup_{t\in[0,1]} |Y(t)-t| \xrightarrow{\mathrm{P}} 0, \text{ as } n \to \infty.$$

Remark 5. The above lemma says that the average of the first m values of $u_{(i)}$ is approximately c_n. The proof is a generalization of [11, Lemma 2.3], which deals with the case when $u_i \equiv w_i^2$ and $m = n^{2/3}$.

Proof of Lemma 8.2: This follows via the introduction of an extra randomization trick developed in [7] and also used in [11]. We will give a full proof here. Define

$$\tau_k := \sum_{i\in[n]} w_i^k \text{ for } k = 1,2,3.$$

For $i \in [n]$, let $\zeta_i \sim \mathrm{Exp}(mw_i/\tau_1)$ be independent exponential random variables. Define the process $\{N(t) : t \in [0,\infty)\}$ as

$$N(t) := \sum_{i\in[n]} \mathbb{1}_{\{\zeta_i \leqslant t\}}, \quad \text{for} \quad t \in [0,\infty).$$

Define the process $\left\{\tilde{Y}(t) : t \in [0,\infty)\right\}$ as

$$\tilde{Y}(t) := \frac{1}{mc_n} \sum_{i\in[n]} u_{(i)} \mathbb{1}_{\{\zeta_i \leqslant t\}}, \quad \text{for} \quad t \in [0,\infty).$$

Note that by the construction, we have $\{Y(N(t)/m) : t \geqslant 0\} \stackrel{d}{=} \left\{\tilde{Y}(t) : t \geqslant 0\right\}$. Therefore when $\epsilon < 1$, on the event $\{|N(t)/m - t| < \epsilon, \forall t \in [0,2]\}$ we have $N(2)/m > 1$ and thus

$$\begin{aligned}
\sup_{t\in[0,1]} |Y(t)-t| &\leqslant \sup_{t\in[0,2]} \left| Y\left(\frac{N(t)}{m}\right) - \frac{N(t)}{m} \right| \\
&\leqslant \sup_{t\in[0,2]} \left| Y\left(\frac{N(t)}{m}\right) - t \right| + \sup_{t\in[0,2]} \left| \frac{N(t)}{m} - t \right|. \\
&\leqslant \sup_{t\in[0,2]} \left| Y\left(\frac{N(t)}{m}\right) - t \right| + \epsilon.
\end{aligned}$$

Thus

$$\mathbb{P}\left(\sup_{t\in[0,1]} |Y(t)-t| > 2\epsilon\right) \leqslant \mathbb{P}\left(\sup_{t\in[0,2]} \left|\tilde{Y}(t)-t\right| > \epsilon\right) + \mathbb{P}\left(\sup_{t\in[0,2]} \left|\frac{N(t)}{m} - t\right| > \epsilon\right). \tag{8.7}$$

Then we bound the first term on the right hand side of (8.7). Define the filtration $\mathscr{F}_t := \sigma(\{\zeta_i \leqslant t\} : i \in [n])$ for $t \geqslant 0$. Then we have for $t > s > 0$,

$$\begin{aligned}
\mathbb{E}[\tilde{Y}(t) \mid \mathscr{F}_s] &= \frac{1}{mc_n} \sum_{i\in[n]} \left[u_i \mathbb{1}_{\{\zeta_i \leqslant s\}} + u_i \mathbb{1}_{\{\zeta_i > s\}} (1 - \exp((t-s)mw_i/\tau_1) \right] \\
&\leqslant \tilde{Y}(s) + \frac{1}{mc_n} \sum_{i\in[n]} \frac{(t-s)mw_i u_i}{\tau_1} \\
&= \tilde{Y}(s) + (t-s).
\end{aligned}$$

Therefore, by a supermartingale inequality [40, Lemma 2.54.5] for the supermartingale $\left\{\tilde{Y}(t) - t : t \in [0,\infty)\right\}$, we have

$$\mathbb{P}(\sup_{t\in[0,2]} |Y(t)-t| > \epsilon) \leqslant \frac{9}{\epsilon} \left(|\mathbb{E}(\tilde{Y}(2)-2)| + \sqrt{\mathrm{Var}(\tilde{Y}(2))} \right). \tag{8.8}$$

Using the fact $x - x^2/2 \leqslant 1 - e^{-x} \leqslant x$, it is easy to see that

$$|\mathbb{E}[\tilde{Y}(2) - 2]| \leqslant \frac{1}{mc_n} \sum_{i\in[n]} u_i \frac{4m^2 w_i^2}{2\tau_1^2} = \frac{2m\sum_{i\in[n]} w_i^2 u_i}{\tau_1 \sum_{i\in[n]} w_i u_i}.$$

For the variance we have

$$\begin{aligned}\mathrm{Var}(\tilde{Y}(2)) &= \frac{1}{m^2 c_n^2} \sum_{i\in[n]} \left[u_i^2(1-\exp(-2mw_i/\tau_1))\exp(-2mw_i/\tau_1)\right] \\ &\leqslant \frac{1}{m^2 c_n^2} \sum_{i\in[n]} \frac{2mw_i u_i^2}{\tau_1} = \frac{2\tau_1(\sum_{i\in[n]} w_i u_i^2)}{m(\sum_{i\in[n]} w_i u_i)^2}.\end{aligned}$$

Similar bound holds for $\mathbb{P}\left(\sup_{t\in[0,2]}\left|\frac{N(t)}{m} - t\right| > \epsilon\right)$ by plugging in $f(x)$ with a special constant function $f(x) \equiv 1$. Thus from (8.7) we have

$$\begin{aligned}&\mathbb{P}\left(\sup_{t\in[0,1]} |Y(t) - t| > 2\epsilon\right) \\ &\leqslant \frac{9}{\epsilon}\left(\frac{2m\sum_{i\in[n]} w_i^2 u_i}{\tau_1\sum_{i\in[n]} w_i u_i} + \sqrt{\frac{2\tau_1(\sum_{i\in[n]} w_i u_i^2)}{m(\sum_{i\in[n]} w_i u_i)^2}} + \frac{2m\tau_2}{\tau_1^2} + \sqrt{\frac{2}{m}}\right) \\ &\leqslant \frac{9}{\epsilon}\left(\frac{2mw_{\max}}{\tau_1} + \sqrt{\frac{2\tau_1 u_{\max}}{m\sum_{i\in[n]} w_i u_i}} + \frac{2m\tau_2}{\tau_1^2} + \sqrt{\frac{2}{m}}\right)\end{aligned}$$

The first two terms in the above display goes to zero, because of the assumptions in the lemma. Since $w_{\max} \geqslant \tau_2/\tau_1$ and $u_{\max} \geqslant c_n$, the rest two terms also goes to zero. The proof of Lemma 8.2 is completed. ■

Completing the proof of Proposition 8.1: Fix $i \geqslant 1$ and let $L(n,i)$ denote the time when we start exploring the i-th largest component in the above size-biased construction of $\mathcal{G}_n(\lambda, \mathbf{w})$ and let $R(n,i)$ be the time when we complete the exploration of the i-th largest component so that the size of $|\mathcal{C}_n^{(i)}| = R(n,i) - L(n,i)$. Let $L(\infty, i)$ denote the time of the start of the i-th largest excursion from zero of $\bar{W}^{\lambda}_{\kappa,\sigma}(\cdot)$ and $R(\infty,i)$ denote the end of this excursion where κ,σ are as in Theorem 4.1. Thus the limiting component sizes are given by $Z_i = R(\infty,i) - L(\infty,i)$ and further by [7, 11] $L(n,i)/n^{2/3} \xrightarrow{\mathrm{w}} L(\infty,i)$ and $R(n,i)/n^{2/3} \xrightarrow{\mathrm{w}} R(\infty,i)$.

Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a monotone non-decreasing function. We will apply Lemma 8.2 to the case when $u_i \equiv f(w_i)$, $i \in [n]$, and $m = Tn^{2/3}$ for some large constant $T > 0$. Define $c_n(f) := \sum_{i\in[n]} w_i f(w_i)/\sum_{i\in[n]} w_i$. Notice that

$$\frac{1}{n^{2/3}} \sum_{v\in\mathcal{C}_n^{(i)}} f(w_v) = \frac{1}{n^{2/3}} \sum_{j=1}^{R(n,i)} f(w_{v(j)}) - \frac{1}{n^{2/3}} \sum_{j=1}^{L(n,i)} f(w_{v(j)}).$$

Thus for any fixed $T > 0$ and $\varepsilon > 0$

$$\begin{aligned}&\mathbb{P}\Big(\frac{1}{n^{2/3}}\Big|\sum_{v\in\mathcal{C}_n^{(i)}} f(w_v) - c_n(f)|\mathcal{C}_n^{(i)}|\Big| > \varepsilon\Big) \\ &\leqslant \mathbb{P}\left(R(n,i) > Tn^{2/3}\right) + \mathbb{P}\left(\sup_{u\leqslant T}\left|\frac{\sum_{i=1}^{n^{2/3}u} f(w_{v(i)})}{Tn^{2/3}c_n(f)} - u\right| > \frac{\varepsilon}{2Tc_n(f)}\right).\end{aligned}$$

Taking $f(x) = f_k(x) := x^k$ for $k = 1,2$. By Assumption 3.1 (a) we have $\lim_{n\to\infty} c_n(f_k) = \sigma_{k+1}/\sigma_1$. The assumptions in Lemma 8.2 reduce to $w_{\max} = o(n^{1/3})$. Thus we can apply Lemma 8.2 to the second term in the above inequality. Thus first letting $n \to \infty$ and then $T \to \infty$, we have for $k = 1,2$,

$$\left| \frac{\sum_{v\in\mathscr{C}_n^{(i)}} w_v^k}{n^{2/3}} - c_n(f_k)\frac{|\mathscr{C}_n^{(i)}|}{n^{2/3}} \right| \xrightarrow{\text{P}} 0, \text{ as } n \to \infty. \tag{8.9}$$

Notice that for $k = 1,2$, by Theorem 4.1 we have

$$c_n(f_k)\frac{|\mathscr{C}_n^{(i)}|}{n^{2/3}} \xrightarrow{\text{w}} \frac{\sigma_{k+1}}{\sigma_1} Z_1 \text{ as } n \to \infty.$$

Combining the above convergence, (8.9) and the assumption $\sigma_2 = \sigma_1$, we completes the proof of Proposition 8.1. ∎

9. Completing the proof of Theorem 3.3

We shall now combine the various ingredients of the last sections to complete the proof of Theorem 3.3. We start with the proof of convergence in the product topology.

9.1. **Convergence in the product topology.** We work under Assumption 3.1 in this section. Due to the conditional independence given the partition $\{\mathcal{V}^{(i)} : i \in \mathbb{N}\}$, as suggested by Proposition 6.1, we can work with each maximal component separately. To ease notation let us work with the largest component $\mathscr{C}_n^{(1)}(\lambda)$. Without loss of generality, we will work with the probability space on which the convergence in Proposition 8.1 holds almost surely:

$$\left(\frac{|\mathscr{C}_n^{(1)}|}{n^{2/3}}, \frac{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}{n^{2/3}}, \frac{\sum_{v\in\mathscr{C}_n^{(1)}} w_v^2}{n^{2/3}} \right) \xrightarrow{\text{a.e.}} \left(Z_i, Z_i, \frac{\sigma_3 Z_i}{\sigma_1} \right). \tag{9.1}$$

Recall the definition of $M_i(\lambda)$ in (4.11). Thus we need to prove

$$\text{scl}\left(\frac{1}{n^{1/3}}, \frac{1}{n^{2/3}}\right) \cdot \mathscr{C}_n^{(1)}(\lambda) \xrightarrow{\text{w}} \mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}} \tilde{\mathbf{e}}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \frac{\sigma_3^{1/2}}{\sigma_1^{3/2}} \tilde{\mathbf{e}}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \mathscr{P}_1 \right). \tag{9.2}$$

By Proposition 6.1, conditional on the vertices in $\mathscr{C}_n^{(1)}(\lambda)$, the random graph $\mathscr{C}_n^{(1)}(\lambda)$ has the same distribution as a connected rank-one random graph as in (6.2) using

$$\mathbf{p} = \left(\frac{w_v}{\sum_{u\in\mathscr{C}_n^{(1)}} w_u} : v \in \mathscr{C}_n^{(1)} \right), \qquad a = \left(1 + \frac{\lambda}{n^{1/3}}\right) \frac{(\sum_{v\in\mathscr{C}_n^{(1)}} w_v)^2}{l_n}.$$

Proof of Theorem 3.3 (i): Our aim is to use Theorem 7.3. Let us first verify the Assumptions 7.1 and 7.2. Notice that the relavant quantities are

$$\sigma(\mathbf{p}) = \frac{\sqrt{\sum_{v\in\mathscr{C}_n^{(1)}} w_v^2}}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}, \ p_{\max} \leqslant \frac{w_{\max}}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}, \ \text{ and } p_{\min} \geqslant \frac{w_{\min}}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}.$$

By (9.1) we have $\sigma(\mathbf{p}) = \Theta(n^{-1/3})$. Therefore, Assumption 7.1 can be verified with any $\epsilon \in (0, 3\eta_0)$ and $r \in (2 + 3\gamma_0, \infty)$. Assumption 7.2 is a consequence of (9.1):

$$\lim_{n\to\infty} a\sigma(\mathbf{p}) = \lim_{n\to\infty} \frac{(\sum_{v\in\mathscr{C}_n^{(1)}} w_v)^2}{l_n} \cdot \frac{\sqrt{\sum_{v\in\mathscr{C}_n^{(1)}} w_v^2}}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v} = \frac{\sigma_3^{1/2}}{\sigma_1^{3/2}} Z_1^{3/2} := \bar{\gamma}_1.$$

Thus Assumption 7.1 and 7.2 are satisfied. Now applying Theorem 7.3 we have

$$\operatorname{scl}\left(\sigma(\mathbf{p}), \frac{1}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}\right)\cdot \mathscr{C}_n^{(1)}(\lambda) \xrightarrow{\text{w}} \mathscr{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}_1}, \bar{\gamma}_1\tilde{\mathbf{e}}^{\bar{\gamma}_1}, \mathscr{P}_1). \tag{9.3}$$

By replacing (l,γ,θ) in the Brownian scaling (4.10) with $(1, l^{3/2}, \gamma l^{-3/2})$, we have, for all $\gamma > 0$ and $l > 0$,

$$\left\{\tilde{\mathbf{e}}^{\gamma}(s) : s\in[0,1]\right\} \stackrel{d}{=} \left\{\frac{1}{l^{1/2}}\tilde{\mathbf{e}}_l^{\gamma/l^{3/2}}(ls) : s\in[0,1]\right\}. \tag{9.4}$$

By comparing the two scaling operator in (9.2) and (9.3), we have

$$\operatorname{scl}\left(\frac{1}{n^{1/3}}, \frac{1}{n^{2/3}}\right) = \operatorname{scl}\left(\frac{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}{n^{1/3}\sqrt{\sum_{v\in\mathscr{C}_n^{(1)}} w_v^2}}, \frac{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}{n^{2/3}}\right)\cdot \operatorname{scl}\left(\sigma(\mathbf{p}), \frac{1}{\sum_{v\in\mathscr{C}_n^{(1)}} w_v}\right).$$

Therefore by Proposition 2.1 and the convergence in (9.1), we have

$$\begin{aligned}
\lim_{n\to\infty} \operatorname{scl}\left(\frac{1}{n^{1/3}}, \frac{1}{n^{2/3}}\right)\cdot\mathscr{C}_n^{(1)}(\lambda) &= \operatorname{scl}\left(\sqrt{\frac{Z_1\sigma_1}{\sigma_3}}, Z_1\right)\mathscr{G}(2\tilde{\mathbf{e}}^{\bar{\gamma}_1}, \bar{\gamma}_1\tilde{\mathbf{e}}^{\bar{\gamma}_1}, \mathscr{P}_1)\\
&\stackrel{d}{=} \operatorname{scl}\left(\sqrt{\frac{Z_1\sigma_1}{\sigma_3}}, Z_1\right)\mathscr{G}\left(\frac{2}{Z_1^{1/2}}\tilde{\mathbf{e}}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}(Z_1\cdot), \frac{\bar{\gamma}_1}{Z_1^{1/2}}\mathbf{e}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}(Z_1\cdot), \mathscr{P}_1\right)\\
&\stackrel{d}{=} \mathscr{G}\left(\frac{2\sigma_1^{1/2}}{\sigma_3^{1/2}}\tilde{\mathbf{e}}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \frac{\bar{\gamma}_1}{Z_1^{3/2}}\mathbf{e}_{Z_1}^{\sigma_3^{1/2}/\sigma_1^{3/2}}, \mathscr{P}_1\right),
\end{aligned}$$

where the limit in the first line denote the limit of weak convergence, the second line uses the scaling in (9.4) with $l = Z_1$ and $\gamma = \bar{\gamma}_1$, and the third line use the scaling in (4.14) and the scaling invariance of $\mathscr{P}$. Collecting the terms in the last display gives (9.2). The proof of Theorem 3.3 (i) is completed. ■

9.2. **Convergence in the l^4 metric.** We will now strengthen the convergence in (3.1) to convergence in $\mathscr{T}_2$ topology. Since $\lambda\in\mathbb{R}$ is fixed, we will subsequently drop it from our notation. We consider the Norros-Reittu model $\mathscr{G}_n^{\text{nr}}(\mathbf{w},\lambda)$ in this section. We first require some notation. As usual, let $\mathscr{C}_n^{(i)}$ be the i-th largest component. Denote by $|\mathscr{C}_n^{(i)}|$, the number of vertices $\mathscr{C}_n^{(i)}$. For $v\in[n]$, let $\mathscr{C}_n(v)$ denote the component that contains v. For $k = 1,2$ and $i \geqslant 1$, let

$$X_n(v;k) := \sum_{j\in\mathscr{C}_n(v)} w_j^k \text{ and } X_{n,i}(k) := X_n(v;k) \text{ for any } v\in\mathscr{C}_n^{(i)}. \tag{9.5}$$

For $i\geqslant 1$ Define

$$\mathbf{p}^{(i)} = \left(w_j/X_{n,i}(1) : j\in\mathscr{C}_n^{(i)}\right).$$

Let $\mathscr{F}_{ptn} = \sigma\left(\{w_v : v\in\mathscr{C}_n^{(i)}\}_{i\geqslant 1}\right)$ be the σ-field generated by the partition of weights into different components. Note that $X_{n,i}(k)$ is measurable with respect to $\mathscr{F}_{ptn}$.

In the proof of the l^4 convergence, the plan is to treat small components and large components differently. Precisely, define $\alpha_0 = 1/12 - \eta_0$. For components with $|\mathscr{C}_n^{(i)}| < n^{\alpha_0}$, we will use trivial bounds on its diameter and total mass; for components with $|\mathscr{C}_n^{(i)}| \geqslant n^{\alpha_0}$, the following two lemmas provide the bound we need for these large components.

Lemma 9.1. *Let $X_n(v;k)$ be as above and recall the definition of σ_k from Section* 3.1. *Then the following hold under Assumptions* 3.1 *and* 3.2. *For any $r > 0$, there exists constants $n_0 > 0$ and $K_{9.1} = K_{9.1}(r,\mathbf{w}) > 0$ such that*

(a) *For all* $v \in [n]$, $n^{1/12-2\eta_0} \leqslant m \leqslant n^{47/48}$, $k = 1,2$ *and* $n > n_0$,

$$\mathbb{P}\left(X_n(v;k) \geqslant \frac{32\sigma_{k+1}m}{\sigma_1} \text{ and } |\mathscr{C}_n(v)| \leqslant m\right) \leqslant \frac{K_{9.1}}{n^r}. \tag{9.6}$$

(b) *For all* $v \in [n]$, $n^{1/12-2\eta_0} \leqslant m \leqslant n^{45/48}$, $k = 1,2$ *and* $n > n_0$

$$\mathbb{P}\left(X_n(v;k) \leqslant \frac{\sigma_{k+1}m}{16\sigma_1} \text{ and } |\mathscr{C}_n(v)| \geqslant m\right) \leqslant \frac{K_{9.1}}{n^r}. \tag{9.7}$$

The proof of Lemma 9.1 is deferred to Section 11.1.

Let $\underline{A} := 1/16$ and $\overline{A} := 32\sigma_3/\sigma_1$. Recall $\alpha_0 = 1/12 - 2\eta_0$. Define the events

$$E_n(\alpha_0) := \left\{\text{for } k = 1,2 \text{ and } v \in [n], |\mathscr{C}_n(v)| \geqslant n^{\alpha_0} \text{ implies } \underline{A}|\mathscr{C}_n(v)| \leqslant X_n(v;k) \leqslant \overline{A}|\mathscr{C}_n(v)|\right\}. \tag{9.8}$$

Lemma 9.2. *Assume that Assumptions* 3.1 *and* 3.2 *hold. Let* $\alpha_0 = 1/12 - 2\eta_0$. *Then there exists constants* $K_{9.2} > 0$ *and* $n_0 > 0$ *such that for all* $n \geqslant n_0$ *and* $\eta \in (0, 2\sigma_3/\sigma_1^{1/3})$ *we have,*

$$\mathbb{1}_{E_n(\alpha_0)}\mathbb{1}_{\left\{n^{\alpha_0} \leqslant |\mathscr{C}_n^{(i)}| \leqslant \eta n^{2/3}\right\}} \mathbb{E}\left[(\operatorname{diam}(\mathscr{C}_n^{(i)}))^4 \mid \mathscr{F}_{ptn}\right] \leqslant \frac{K_{9.2}}{[\sigma(\mathbf{p}^{(i)})]^4} \text{ for all } i \geqslant 1.$$

The proof of Lemma 9.2 will be given in Section 11.2.

Proof of Theorem 3.3 (ii): Since we have

$$\begin{aligned} &d_{\mathrm{GHP}}\left(\operatorname{scl}\left(\frac{1}{n^{1/3}}, \frac{1}{n^{2/3}}\right)\mathscr{C}_n^{(i)}, M_i(\lambda)\right) \\ &\qquad \leqslant \frac{\operatorname{diam}\left(\mathscr{C}_n^{(i)}\right)}{n^{1/3}} + \operatorname{diam}(M_i(\lambda)) + \frac{\operatorname{mass}\left(\mathscr{C}_n^{(i)}\right)}{n^{2/3}} + \operatorname{mass}(M_i(\lambda)), \end{aligned}$$

to prove convergence in $\mathscr{T}_2$ topology, it is enough to show that for any $\epsilon > 0$

$$\limsup_{N\to\infty}\limsup_{n\to\infty}\left[\mathbb{P}\left(\sum_{i\geqslant N}\frac{\operatorname{diam}^4(\mathscr{C}_n^{(i)})}{n^{4/3}} > \epsilon\right) + \mathbb{P}\left(\sum_{i\geqslant N}\frac{X_{n,i}(1)^4}{n^{8/3}} > \epsilon\right)\right] = 0. \tag{9.9}$$

First we consider the first term in (9.9). Fix $\alpha_0 := 1/12 - 2\eta_0$. Notice that

$$\frac{1}{n^{4/3}}\sum_{i\geq 1}\mathbb{1}_{\left\{|\mathscr{C}_n^{(i)}| < n^{\alpha_0}\right\}}\operatorname{diam}^4(\mathscr{C}_n^{(i)}) \leq \frac{n \cdot n^{4\alpha_0}}{n^{4/3}} = \frac{1}{n^{8\eta_0}}.$$

So it is enough to focus on the components with size at least n^{α_0}. Recall $E_n(\alpha_0)$ as in (9.8). We will first show

$$\mathbb{P}\left(E_n(\alpha_0)^c\right) \to 0 \text{ as } n \to \infty. \tag{9.10}$$

By Lemma 9.1 (b), there exist n_1 such that when $n > n_1$, for $k = 1,2$,

$$\begin{aligned} &\mathbb{P}\left(\exists v \in [n] \text{ with } |\mathscr{C}_n(v)| \geqslant n^{\alpha_0} \text{ and } X_n(v;k) \geqslant \frac{32\sigma_{k+1}}{\sigma_1}|\mathscr{C}_n(v)|\right) \\ &\leqslant \mathbb{P}\left(|\mathscr{C}_n^{(1)}| > n^{3/4}\right) + \sum_{v\in[n]}\sum_{m=n^{\alpha_0}}^{n^{3/4}}\mathbb{P}\left(|\mathscr{C}_n(v)| = m \text{ and } X_n(v;k) \geqslant \frac{32\sigma_{k+1}}{\sigma_1}|\mathscr{C}_n(v)|\right) \\ &\leqslant \mathbb{P}\left(|\mathscr{C}_n^{(1)}| > n^{3/4}\right) + n\cdot n^{3/4}\cdot\frac{K_{9.1}}{n^2} = o(1), \end{aligned}$$

where the third line is a consequence of (9.6) with $r = 2$. By a similar argument and an application of (9.7), we can show that

$$\mathbb{P}\left(\exists v \in [n] \text{ with } |\mathscr{C}_n(v)| \geqslant n^{\alpha_0} \text{ and } X_n(v;k) \leqslant \frac{\sigma_{k+1}}{16\sigma_1}|\mathscr{C}_n(v)|\right) = o(1).$$

Noticing $\sigma_3 \geq \sigma_2 = \sigma_1$, we then have (9.10).

Fix $\eta \in (0, 2\sigma_3/\sigma_1^{1/3})$ (the upper bound of η is due to Lemma 9.2) and by Theorem 4.1 we can find N_η such that $\mathbb{P}\left(\sum_{i\geqslant N_\eta} n^{-4/3}|\mathscr{C}_n^{(i)}|^2 > \eta\right) \leqslant \eta$ for all $n \geqslant 1$. Let $G_n(\alpha_0,\eta) := E_n(\alpha_0) \cap \left\{\sum_{i\geqslant N_\eta} n^{-4/3}|\mathscr{C}_n^{(i)}|^2 \leqslant \eta\right\}$. Let $\sum_1$ denote sum over all components $\mathscr{C}_n^{(i)}$ for which $i \geqslant N_\eta$ and $|\mathscr{C}_n^{(i)}| \geqslant n^{\alpha_0}$. Then

$$\begin{aligned}
&\mathbb{P}\left(\textstyle\sum_1 n^{-4/3}\operatorname{diam}^4(\mathscr{C}_n^{(i)}) > \varepsilon\right) &(9.11)\\
\leqslant&\mathbb{E}\left[\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\mathbb{P}\left(\textstyle\sum_1 n^{-4/3}\operatorname{diam}^4(\mathscr{C}_n^{(i)}) > \varepsilon \,\middle|\, \mathscr{F}_{ptn}\right)\right] + \mathbb{P}(E_n(\alpha_0)^c) + \eta.\\
\leqslant&\frac{1}{\varepsilon n^{4/3}}\mathbb{E}\left[\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 \mathbb{E}\left(\operatorname{diam}^4(\mathscr{C}_n^{(i)}) \,\middle|\, \mathscr{F}_{ptn}\right)\right] + \mathbb{P}(E_n(\alpha_0)^c) + \eta.
\end{aligned}$$

By Lemma 9.2, there exist n_2 such that for $n \geqslant n_2$,

$$\begin{aligned}
&\frac{1}{\varepsilon n^{4/3}}\mathbb{E}\left[\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 \mathbb{E}\left(\operatorname{diam}^4(\mathscr{C}_n^{(i)}) \,\middle|\, \mathscr{F}_{ptn}\right)\right] &(9.12)\\
\leqslant&\frac{1}{\varepsilon n^{4/3}}\mathbb{E}\left[\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 \dfrac{K_{9.2}}{[\sigma(\mathbf{p}^{(i)})]^4}\right] \leqslant \frac{K_{9.2}}{\varepsilon n^{4/3}}\mathbb{E}\left[\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 |\mathscr{C}_n^{(i)}|^2\right] \leqslant \frac{\eta K_{9.2}}{\varepsilon},
\end{aligned}$$

where the last line uses the fact $[\sigma(\mathbf{p}^{(i)})]^2|\mathscr{C}_n^{(n)}| \geq 1$ and the definition of $G_n(\alpha_0,\eta_0)$.

Combining (9.10), (9.11), and (9.12), we arrive at

$$\limsup_n \mathbb{P}\left(n^{-4/3}\textstyle\sum_1 \operatorname{diam}^4(\mathscr{C}_n^{(i)}) > \varepsilon\right) \leqslant \eta + \frac{\eta K_{9.2}}{\varepsilon}$$

Since η can be arbitrarily small, we conclude that

$$\limsup_{N\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(n^{-4/3}\sum_{i\geqslant N}\operatorname{diam}^4(\mathscr{C}_n^{(i)}) > \varepsilon\right) = 0. \tag{9.13}$$

Next, we consider the second term in (9.9). For components with $|\mathscr{C}_n^{(i)}| < n^{\alpha_0}$, on the event

$$\left\{\text{for all } v \in [n],\ |\mathscr{C}_n(v)| < n^{\alpha_0} \text{ implies } X_n(v,1) \leq 32n^{\alpha_0}\right\}, \tag{9.14}$$

we have

$$\frac{1}{n^{8/3}}\sum_{i\geq 1}\mathbb{1}_{\left\{|\mathscr{C}_n^{(i)}|<n^{\alpha_0}\right\}}X_{n,i}(1)^4 \leq 32^4 n \cdot \frac{n^{4\alpha_0}}{n^{8/3}} \leq \frac{32^4}{n^{4/3}}. \tag{9.15}$$

By (9.6), the event in (9.14) occurs with high probability and this take care of the small component. For components with $|\mathscr{C}_n^{(i)}| \geq n^{\alpha_0}$ we have

$$\mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 \dfrac{X_{n,i}(1)^4}{n^{8/3}} \leqslant \mathbb{1}\left\{G_n(\alpha_0,\eta)\right\}\textstyle\sum_1 \dfrac{\overline{A}^4|\mathscr{C}_n^{(i)}|^4}{n^{8/3}} \leqslant \overline{A}^4\eta\dfrac{|\mathscr{C}_n^{(1)}|^2}{n^{4/3}} \tag{9.16}$$

Since η is arbitrary and $|\mathscr{C}_n^{(1)}|/n^{2/3}$ is tight, combining (9.15) and (9.16), we conclude that

$$\limsup_{N\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(n^{-8/3}\sum_{i\geqslant N}X_{n,i}(1)^4 > \varepsilon\right) = 0.$$

This together with (9.13) yields (9.9) and completes the proof of Theorem 3.3 (ii). ■

10. TAIL BOUNDS FOR HEIGHT OF $\mathbf{p}$ TREES: PROOF OF THEOREM 3.7

For the convenence of reference, we restate the assumptions in Theorem 3.7 as follows.

Assumption 10.1. *There exists* $\epsilon_0 \in (0, 1/2)$ *and* $r_0 \in (2, \infty)$ *such that*

$$\sigma(\mathbf{p}) \leqslant \frac{1}{2^{10}}, \quad \frac{p_{\max}}{[\sigma(\mathbf{p})]^{3/2+\epsilon_0}} \leqslant 1, \quad \frac{[\sigma(\mathbf{p})]^{r_0}}{p_{\min}} \leqslant 1,$$

We will prove the following lemma in this section.

Lemma 10.2. *Assume the setting of Theorem* 3.7*. Then for any integer* $r \geqslant \lfloor r_0/2\epsilon_0 \rfloor + 1$*, there exists a constant* $K_{10.2} = K_{10.2}(r) > 0$ *such that*

$$\mathbb{P}\left(\mathrm{ht}(\mathscr{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \frac{K_{10.2}}{x^r}, \textit{ for } 1 \leqslant x \leqslant [\sigma(\mathbf{p})]^{-2\epsilon_0}.$$

Using Lemma 10.2, we prove Theorem 3.7 as follows.

Proof of Theorem 3.7: Note that $\mathbb{P}(\mathrm{ht}(\mathscr{T}) > m) = 0$. Take any $r \geqslant \lfloor r_0/2\epsilon_0 \rfloor + 1$, we only need to show that (3.4) holds for all $[\sigma(\mathbf{p})]^{-2\epsilon_0} < x \leqslant m\sigma(\mathbf{p})$. Define $r' := \lfloor (r_0 - 1)r/(2\epsilon_0) \rfloor + 1$. Then we have $r' \geq r \geq \lfloor r_0/2\epsilon_0 \rfloor + 1$, thus we can apply Lemma 10.2 with r'. For $[\sigma(\mathbf{p})]^{-2\epsilon_0} < x \leqslant m\sigma(\mathbf{p})$, we have

$$\mathbb{P}\left(\mathrm{ht}(\mathscr{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \mathbb{P}\left(\mathrm{ht}(\mathscr{T}) \geqslant \frac{[\sigma(\mathbf{p})]^{-2\epsilon_0}}{\sigma(\mathbf{p})}\right) \leqslant K_{10.2}(r')[\sigma(\mathbf{p})]^{2\epsilon_0 r'} \leqslant K_{10.2}(r')\frac{[\sigma(\mathbf{p})]^{r_0 r}}{[\sigma(\mathbf{p})]^r}.$$

By Assumption 10.1, $[\sigma(\mathbf{p})]^{r_0} \leqslant p_{\min} \leqslant 1/m$. Then we have

$$\mathbb{P}\left(\mathrm{ht}(\mathscr{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant K_{10.2}(r')\frac{1}{[m\sigma(\mathbf{p})]^r} \leqslant \frac{K_{10.2}(r')}{x^r},$$

for all $[\sigma(\mathbf{p})]^{-2\epsilon_0} < x \leqslant m\sigma(\mathbf{p})$. Defining $K_{3.7}(r) := K_{10.2}(\lfloor (r_0 - 1)r/(2\epsilon_0) \rfloor + 1)$, we complete the proof of Theorem 3.7. ■

The goal of the rest of this section is to prove Lemma 10.2. We will derive quantitative versions of some of the results of [8]. We will also use the techniques developed in [19]. Recall that $\mathscr{T}$ is a rooted tree with vertex set labelled by $[m]$ and so given a vertex $v \in \mathscr{T}$, we can let $\mathbb{A}(v)$ be the set of ancestors of v. More precisely, writing $\mathrm{ht}(v)$ for the height of vertex $v \in \mathscr{T}$ and the path from the root ρ to v as $u_0 = \rho, u_1, \ldots, u_{\mathrm{ht}(v)-1}, u_{\mathrm{ht}(v)} = v$, then $\mathbb{A}(v) = \{u_0, \ldots, u_{\mathrm{ht}(v)-1}\}$. Let $\mathscr{G}(v) := \sum_{u \in \mathbb{A}(v)} p_u$. Recall the function $F^{\mathrm{exc},\mathbf{p}}$ in (7.2) used to construct $\mathscr{T}$. In particular recall that for each vertex $v \in \mathscr{T}$, there was an i such that we find the children of v in the interval $[y^*(i-1), y^*(i))$. Define $e(v) = y^*(i)$. Fix $x > 0$ and define the events

$$\mathscr{B}_1 := \left\{\max_{v \in [m]} \frac{F^{\mathrm{exc},\mathbf{p}}(e(v))}{\sigma(\mathbf{p})} \geqslant \frac{x}{8}\right\}, \tag{10.1}$$

$$\mathscr{B}_2 := \left\{\max_{v \in [m]} \frac{F^{\mathrm{exc},\mathbf{p}}(e(v))}{\sigma(\mathbf{p})} \leqslant \frac{x}{8}, \max_{v \in [m]}\left(\frac{\mathscr{G}(v)}{2\sigma(\mathbf{p})} - \frac{F^{\mathrm{exc},\mathbf{p}}(e(v))}{\sigma(p)}\right) \geqslant \frac{x}{8}\right\}, \tag{10.2}$$

and finally

$$\mathscr{B}_3 := \left\{\max_{v \in [m]} \frac{\mathscr{G}(v)}{2\sigma(\mathbf{p})} \leqslant \frac{x}{4}, \mathrm{ht}(\mathscr{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right\}. \tag{10.3}$$

Thus

$$\mathbb{P}\left(\mathrm{ht}(\mathscr{T}) \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \mathbb{P}(\mathscr{B}_1) + \mathbb{P}(\mathscr{B}_2) + \mathbb{P}(\mathscr{B}_3). \tag{10.4}$$

We will bound each one of the terms on the right individually in Lemmas 10.3, 10.6 and 10.4. Using these three lemmas, we give the proof of Lemma 10.2 as follows:

Proof of Lemma 10.2: This directly follows from the bounds in Lemmas 10.3, 10.6 and 10.4. ■

10.1. **Analysis of the event $\mathscr{B}_1$.** We start by bounding $\mathbb{P}(\mathscr{B}_1)$ in the following lemma.

Lemma 10.3. *Under Assumption* 10.1*,*

$$\mathbb{P}(\mathscr{B}_1) \leqslant 12e^{-x^2/2^{16}} \text{ for } 1 \leqslant x \leqslant 128[\sigma(\mathbf{p})]^{-2\epsilon_0}. \tag{10.5}$$

Proof: Replacing x by $x\sigma(\mathbf{p})/8$ in Lemma 7.11, we have the same bound as in (10.5), but for all x such that

$$\frac{32p_{\max}}{\sigma(\mathbf{p})} \leqslant x \leqslant \frac{128\sigma(\mathbf{p})}{p_{\max}}.$$

Then by Assumption 10.1, we have $32p_{\max}/\sigma(\mathbf{p}) \leq 32[\sigma(\mathbf{p})]^{1/2} \leq 1$ and $128\sigma(\mathbf{p})/p_{\max} \geq 128[\sigma(\mathbf{p})]^{-1/2-\epsilon_0} \geq 128[\sigma(\mathbf{p})]^{-2\epsilon_0}$. This completes the proof of Lemma 10.3. ■

10.2. **Analysis of the event $\mathscr{B}_3$.** We will prove the following bound on $\mathbb{P}(\mathscr{B}_3)$

Lemma 10.4. *Under Assumption* 10.1*, for each integer* $r \geqslant \lfloor r_0/2\epsilon_0 \rfloor + 1$*, there exist a constant* $K_{10.4} = K_{10.4}(r)$ *such that*

$$\mathbb{P}(\mathscr{B}_3) \leqslant \frac{K_{10.4}}{x^r} \text{ for } x \geqslant 1. \tag{10.6}$$

Proof: Note that on the set $\mathscr{B}_3$, there exists a vertex $v \in \mathscr{T}$ such that

(a) The height of this vertex satisfies $x/\sigma(\mathbf{p}) \leqslant \mathrm{ht}(v) \leqslant x/\sigma(\mathbf{p}) + 1$.
(b) For this v,

$$\sigma(\mathbf{p})\,\mathrm{ht}(v) - \frac{\mathscr{G}(v)}{\sigma(\mathbf{p})} \geqslant \frac{x}{2}.$$

Thus

$$\begin{aligned}\mathbb{P}(\mathscr{B}_3) &\leqslant \frac{1}{p_{\min}} \sum_{v\in[m]} p_v \mathbb{E}\left(\mathbb{1}\left\{\sigma(\mathbf{p})\,\mathrm{ht}(v) - \frac{\mathscr{G}(v)}{\sigma(\mathbf{p})} \geqslant \frac{x}{2}, \mathrm{ht}(v) \leqslant \frac{x}{\sigma(\mathbf{p})} + 1\right\}\right) \\ &= \frac{1}{p_{\min}} \mathbb{P}\left(\sigma(\mathbf{p})\,\mathrm{ht}(\mathbb{V}) - \frac{\mathscr{G}(\mathbb{V})}{\sigma(\mathbf{p})} \geqslant \frac{x}{2}, \mathrm{ht}(\mathbb{V}) \leqslant \frac{x}{\sigma(\mathbf{p})} + 1\right) \\ &=: \frac{1}{p_{\min}} \mathbb{P}(\mathscr{B}_4),\end{aligned} \tag{10.7}$$

where $\mathbb{V}$ with distribution independent of $\mathscr{T}$ is a vertex selected from $\mathscr{T}$ with $\mathbb{P}(\mathbb{V} = j) = p_j$. By [19, Corollary 3],

$$(\mathrm{ht}(\mathbb{V}), \mathscr{G}(\mathbb{V})) \stackrel{d}{=} \left(T - 2, \sum_{i=1}^{T-1} p_{\xi_i}\right), \tag{10.8}$$

where $(\xi_i : i \geqslant 1)$ are *iid* with distribution $\mathbf{p}$ namely $\mathbb{P}(\xi_i = j) = p_j$ for $j \in [m]$ and T is the first repeat time of this sequence namely

$$T = \min\left\{j \geqslant 2 : \xi_j = \xi_i \text{ for some } 1 \leqslant i < j\right\}.$$

Hence

$$\begin{aligned}\mathbb{P}(\mathscr{B}_4) &= \mathbb{P}\left((T-2)\sigma(\mathbf{p}) - \frac{\sum_{i=1}^{T-1} p_{\xi_i}}{\sigma(\mathbf{p})} \geqslant \frac{x}{2},\ \sigma(\mathbf{p})(T-2) \leqslant x + \sigma(\mathbf{p})\right) \\ &\leqslant \mathbb{P}\left((T-1) - \frac{\sum_{i=1}^{T-1} p_{\xi_i}}{(\sigma(\mathbf{p}))^2} \geqslant \frac{x}{2\sigma(\mathbf{p})},\ (T-1) \leqslant \frac{x + 2\sigma(\mathbf{p})}{\sigma(\mathbf{p})}\right).\end{aligned} \tag{10.9}$$

Define the random variables $X_j = p_{\xi_j}/(\sigma(\mathbf{p}))^2 - 1$. The we have

$$\mathbb{P}(\mathscr{B}_4) \leqslant \mathbb{P}\left(\max_{1\leqslant j\leqslant 2+x/\sigma(\mathbf{p})} |S_j| \geqslant \frac{x}{2\sigma(\mathbf{p})}\right). \tag{10.10}$$

Notice that the sequence $\{S_j : j \geqslant 1\}$ obtained by setting $S_j = \sum_{i=1}^{j} X_j$ is a martingale. The following lemma is a basic concentration result about S_j.

Lemma 10.5. *For each integer $r \geqslant 1$, there exists a constant $K_{10.5} = K_{10.5}(r) > 0$, such that for all $k \geqslant 1$ and $t > 0$, we have*

$$\mathbb{P}\left(\max_{1\leqslant j\leqslant k} |S_j| \geqslant t\right) \leqslant K_{10.5} \cdot \frac{k^r p_{\max}^{2r}}{t^{2r}[\sigma(\mathbf{p})]^{4r}}.$$

Proof: By the Markov inequality and the Burkholder-Davis-Gundy inequality, for any integer $r > 0$, we have

$$\mathbb{P}\left(\max_{1\leqslant j\leqslant k} |S_j| \geqslant t\right) \leqslant \frac{\mathbb{E}\left[\max_{1\leqslant j\leqslant k} |S_j|^{2r}\right]}{t^{2r}} \leqslant C_1(r)\frac{\mathbb{E}\left[(\sum_{j=1}^{k} X_j^2)^r\right]}{t^{2r}}, \tag{10.11}$$

where $C_1(r)$ is the constant that shows up in the Burkholder-Davis-Gundy inequality and only depends on r. Notice that $|X_1| \leq \max\left\{\frac{p_{\xi_1}}{\sigma^2(\mathbf{p})}, 1\right\} \leq p_{\max}/\sigma^2(\mathbf{p})$. We have

$$\mathbb{E}\left[(\sum_{j=1}^{k} X_j^2)^r\right] \leqslant k^r\, \mathbb{E}[|X_1|^{2r}] \leqslant k^r \frac{p_{\max}^{2r}}{[\sigma(\mathbf{p})]^{4r}}. \tag{10.12}$$

Combining (10.11) and (10.12) proves the bound in Lemma 10.5 with $K_{10.5}(r) = C_1(r)$. ■

Applying Lemma 10.5 to (10.10) with $t = x/2\sigma(\mathbf{p})$ and $k = 2x/\sigma(\mathbf{p}) > 2 + x/\sigma(\mathbf{p})$, we have for $r \geqslant 1$

$$\mathbb{P}(\mathscr{B}_4) \leqslant K_{10.5}(r)\left(\frac{2\sigma(\mathbf{p})}{x}\right)^{2r} \cdot \left(\frac{2x}{\sigma(\mathbf{p})}\right)^r \cdot \frac{p_{\max}^{2r}}{[\sigma(\mathbf{p})]^{4r}} = \frac{K_{10.5}(r)2^{3r}}{x^r} \cdot \frac{p_{\max}^{2r}}{[\sigma(\mathbf{p})]^{3r}}.$$

Taking $r \geqslant \lfloor r_0/2\epsilon_0\rfloor + 1$, by Assumption 10.1, (10.7) and the above bound we have

$$\mathbb{P}(\mathscr{B}_3) \leqslant \frac{K_{10.5}(r)2^{3r}}{x^r} \cdot \frac{1}{[\sigma(\mathbf{p})]^{r_0}} \cdot [\sigma(\mathbf{p})]^{2r\epsilon_0} \leqslant \frac{K_{10.5}(r)2^{3r}}{x^r} \text{ for } x \geqslant 1,$$

The proof of Lemma 10.4 is completed with $K_{10.4}(r) := K_{10.5}(r)2^{3r}$. ■

10.3. **Analysis of the event $\mathscr{B}_2$.** Let us now analyze $\mathscr{B}_2$. In this section, we will prove:

Lemma 10.6. *Under Assumption* 10.1, *for each integer $r \geqslant \lfloor r_0/2\epsilon_0\rfloor + 1$, there exists a constant $K_{10.6} = K_{10.6}(r)$ such that*

$$\mathbb{P}(\mathscr{B}_2) \leqslant \frac{K_{10.6}}{x^r} \textit{ for } 1 \leqslant x \leqslant 8[\sigma(\mathbf{p})]^{-2\epsilon_0}. \tag{10.13}$$

Proof: The event $\mathscr{B}_2$ consists of two events happening concurrently: (a) the sum of weights p_j on the path to some vertex v from the root, namely $\mathscr{G}(v) = \sum_{j\in\mathbb{A}(v)} p_j$ is "large", where $\mathbb{A}(v)$ are the set of ancestors of v; (b) the maximum value of the excursion $F^{\mathrm{exc},\mathbf{p}}$ being small. We start with the following proposition.

Proposition 10.7. *Under Assumption* 10.1*, for each integer $r \geqslant r_0$, there exists a constant $K_{10.7} = K_{10.7}(r)$ such that*

$$\mathbb{P}\left(\max_{v\in[m]} \mathscr{G}^2(v) \geqslant x\sigma(\mathbf{p})\right) \leqslant \frac{K_{10.7}}{x^r} \text{ for all } \frac{1}{8} \leqslant x \leqslant [\sigma(\mathbf{p})]^{-2\epsilon_0}.$$

Proof: We have

$$\mathbb{P}\left(\max_{v\in[m]} \mathscr{G}^2(v) \geqslant x\sigma(\mathbf{p})\right) \leqslant \mathbb{P}(\mathscr{B}_5) + \mathbb{P}(\mathscr{B}_6) \tag{10.14}$$

where

$$\mathscr{B}_5 := \left\{\max_{v\in[m]} \mathscr{G}(v) \geqslant \sqrt{x\sigma(\mathbf{p})},\ \mathrm{ht}(\mathscr{T}) \leqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right\},$$

and

$$\mathscr{B}_6 := \left\{\mathrm{ht}(\mathscr{T}) \geqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right\}.$$

Arguing as in (10.7), we see that,

$$\mathbb{P}(\mathscr{B}_5) \leqslant \frac{1}{p_{\min}} \mathbb{P}\left(\mathscr{G}(\mathbb{V}) \geqslant \sqrt{x\sigma(\mathbf{p})},\ \mathrm{ht}(\mathbb{V}) \leqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right) \tag{10.15}$$

where as before $\mathbb{V}$ is selected from $[m]$ independent of $\mathscr{T}$ using the probability vector $\mathbf{p}$. Using the distributional representation (10.8) we get, when $x \geqslant 1/8$

$$\begin{aligned}
\mathbb{P}(\mathscr{B}_5) &= \frac{1}{p_{\min}} \mathbb{P}\left(\sum_{i=1}^{T-1} \frac{p_{\xi_i}}{(\sigma(\mathbf{p}))^2} \geqslant \frac{\sqrt{x}}{[\sigma(\mathbf{p})]^{3/2}},\ T-2 \leqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right) \\
&\leqslant \frac{1}{p_{\min}} \mathbb{P}\left(\sum_{i=1}^{T-1} \left(\frac{p_{\xi_i}}{(\sigma(\mathbf{p}))^2} - 1\right) \geqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}},\ T-1 \leqslant \frac{\sqrt{x}}{[\sigma(\mathbf{p})]^{3/2}}\right) \\
&\leqslant \frac{1}{p_{\min}} \mathbb{P}\left(\max_{1\leqslant k\leqslant \sqrt{x}/[\sigma(\mathbf{p})]^{3/2}} S_k \geqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right),
\end{aligned}$$

where the second line uses the fact that $\sqrt{x}/(2[\sigma(\mathbf{p})]^{3/2}) \geqslant 1/(16[\sigma(\mathbf{p})]^{3/2}) > 1$, and S_k in the third line is as defined after (10.9). Using Lemma 10.5 with $k = \sqrt{x}/[\sigma(\mathbf{p})]^{3/2}$ and $t = \sqrt{x}/(2[\sigma(\mathbf{p})]^{3/2})$ in last display, for $x \geqslant 1/8$ and $r' \geqslant 2r_0$ we have

$$\begin{aligned}
\mathbb{P}(\mathscr{B}_5) &\leqslant \frac{K_{10.5}(r')}{p_{\min}} \left(\frac{2[\sigma(\mathbf{p})]^{3/2}}{\sqrt{x}}\right)^{2r'} \cdot \left(\frac{\sqrt{x}}{[\sigma(\mathbf{p})]^{3/2}}\right)^{r'} \cdot \frac{p_{\max}^{2r'}}{[\sigma(\mathbf{p})]^{4r'}} = \frac{K_{10.5}(r')2^{2r'}}{p_{\min}} \cdot \frac{1}{x^{r'/2}} \cdot \frac{p_{\max}^{2r'}}{[\sigma(\mathbf{p})]^{5r'/2}} \\
&\leqslant \frac{K_{10.5}(r')2^{2r'}}{[\sigma(\mathbf{p})]^{r_0}} \cdot \frac{1}{x^{r'/2}} \cdot \frac{[\sigma(\mathbf{p})]^{3r'}}{[\sigma(\mathbf{p})]^{5r'/2}} \leqslant \frac{K_{10.5}(r')2^{2r'}}{x^{r'/2}},
\end{aligned}$$

where the second line uses $[\sigma(\mathbf{p})]^{r_0}/p_{\min} \leqslant 1$ and $p_{\max} \leqslant [\sigma(\mathbf{p})]^{3/2}$. Hence, let $r' = 2r$ in the above display, we have when $x \geqslant 1/8$, $r \geqslant r_0$,

$$\mathbb{P}(\mathscr{B}_5) \leqslant \frac{K_{10.5}(2r)2^{4r}}{x^r}. \tag{10.16}$$

To finish the proof for Proposition 10.7, we need to bound $\mathbb{P}(\mathscr{B}_6)$. We will in fact give exponential tail bounds for this event. Arguing as before and using the distributional representation in (10.8) we first get

$$\mathbb{P}(\mathscr{B}_6) \leqslant \frac{1}{p_{\min}} \mathbb{P}\left(\mathrm{ht}(\mathbb{V}) \geqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right) \leqslant \frac{1}{p_{\min}} \mathbb{P}\left(T \geqslant \frac{\sqrt{x}}{2[\sigma(\mathbf{p})]^{3/2}}\right), \tag{10.17}$$

where as before T is the first repeat time of the sequence $\{\xi_j : j \geqslant 1\}$ where $\xi_j \sim \mathbf{p}$ are *iid*. We will prove the following tail bound for T.

Lemma 10.8. *For* $0 < t < 1/p_{\max}$*, we have*

$$\mathbb{P}(T \geqslant t) \leqslant 2\exp\left(-\frac{t^2\sigma^2(\mathbf{p})}{24}\right).$$

Proof: We will need an alternative construction of the random variable T, (see [19, Section 4]), where we essentially construct the sequence $\{\xi_j : j \geqslant 1\}$ in continuous time. The advantage of this construction is reflected in (10.20) below. Using $\mathbf{p} = (p_1, \ldots, p_j)$, partition the unit interval $[0,1]$ as $\{I_j : j \in [m]\}$ where I_j has length p_j. Consider a rate one poisson process $\mathscr{N}$ on $\mathbb{R}_+ \times [0,1]$. We can represent $\mathscr{N} = \{(S_0, U_0), (S_1, U_1), \ldots\}$ where $S_0 < S_1 < \cdots$ are points of a rate one Poisson process on $\mathbb{R}_+$ and U_j are *iid* uniform random variables. Abusing notation, write $\mathscr{N}(t)$ for the number of points in $(0,t] \times [0,1]$ and $\mathscr{N}(t^-)$ for the number of points in $(0,t) \times [0,1]$. Now write $\xi_j = \sum_{i=1}^m i \mathbb{1}\{U_j \in I_i\}$. In this continuous time construction, as before let T denote the first repeat time of the sequence $\{\xi_j : j \geqslant 1\}$ and write $\mathscr{S}$ for the actual "time" namely $\mathscr{S} = \inf\{s : \mathscr{N}(s) > T\}$. Thus $\mathscr{N}(\mathscr{S}^-) = \mathscr{N}((0, \mathscr{S}^-) \times [0,1]) = T$. Then we have

$$\mathbb{P}(T \geqslant t) \leqslant \mathbb{P}(\mathscr{S} \leqslant t/2, T \geqslant t) + \mathbb{P}(\mathscr{S} \geqslant t/2). \tag{10.18}$$

Let us analyze $\mathbb{P}(\mathscr{S} \leqslant t/2, T \geqslant t)$. Note that this event implies that $\mathscr{N}(t/2) \geqslant t$. Standard tail bounds for the Poisson distribution then give

$$\mathbb{P}(\mathscr{N}(t/2) \geqslant t) \leqslant \exp\left(-\frac{t}{2}(2\log 2 - 1)\right) < e^{-t/6}. \tag{10.19}$$

Next, we bound $\mathbb{P}(\mathscr{S} \geqslant t/2)$. By [19, Equations (26) and (29)], for $0 < t < 1/2p_{\max}$ we have

$$\log \mathbb{P}(\mathscr{S} > t) \leqslant -\frac{t^2}{2}\sigma^2(\mathbf{p}) + \frac{t^3 p_{\max}\sigma^2(\mathbf{p})}{3(1 - tp_{\max})} \leqslant -\frac{t^2\sigma^2(\mathbf{p})}{6}. \tag{10.20}$$

replacing t by $t/2$ in the above expression, we have for all $0 < t < 1/p_{\max}$, $\mathbb{P}(\mathscr{S} > t/2) \leqslant \exp(-t^2\sigma^2(\mathbf{p})/24)$. Combining this bound, (10.19) and the fact $t\sigma^2(\mathbf{p}) \leq \sigma^2(\mathbf{p})/p_{\max} \leqslant 1$ we completes the proof of Lemma 10.8. ■

Applying Lemma 10.8 with $t = \sqrt{x}/(2[\sigma(\mathbf{p})]^{3/2})$ to (10.17), when $1/8 \leqslant x \leqslant [\sigma(\mathbf{p})]^{-2\epsilon_0}$, we have $tp_{\max} \leqslant p_{\max}/(2[\sigma(\mathbf{p})]^{3/2+\epsilon_0}) < 1$ and $x > x/2 + 1/16$ and therefore we have

$$\mathbb{P}(\mathscr{B}_6) \leqslant \frac{2}{p_{\min}}\exp\left(-\frac{x}{96\sigma(\mathbf{p})}\right) \leqslant \frac{2}{[\sigma(\mathbf{p})]^{r_0}}\exp\left(-\frac{1}{2^{11}\sigma(\mathbf{p})}\right)\exp\left(-\frac{x}{192}\right) \leqslant Ce^{-x/192},$$

where $C := \sup_{y \geq 0} 2y^{r_0}e^{-y/2^{11}}$. This combined with (10.16) finishes the proof of Proposition 10.7. ■

To complete the analysis of the event $\mathscr{B}_2$, we need to strengthen [8, Lemma 10]. We first setup some notation. Let $A \subset [m]$. We will use this later for the set $\mathbb{A}(v)$, the set of ancestors of a fixed vertex v. Let $\mathbf{q}$ be the probability distribution obtained by merging the elements of A into a single point. More precisely $\mathbf{q} = (q_1, \ldots, q_{m-|A|+1})$ where $q_1 = \sum_{v \in A} p_v := p(A)$ and $\{q_i : i \geqslant 2\}$ corresponds to the set $\{p_i : i \in [m] \setminus A\}$. Let $\mathscr{T}(1, \mathbf{q})$ be a $\mathbf{q}$-tree constructed as in (2.7) with the probability mass function $\mathbf{q}$, conditional on vertex 1 being the root. Denote by $\mathscr{H}_1$ the children of vertex 1. For each $v \in \mathscr{H}_1$ flip a fair coin (independent for different v) let $c(v)$ denote the outcome of this flip. Define the random variable

$$X = \sum_{v \in \mathscr{H}_1} q_v \mathbb{1}\{c(v) \text{ is Head}\}. \tag{10.21}$$

Proposition 10.9. *Let* $A \subseteq [m]$ *and* $q_1 = p(A)$ *and define* $K_{10.9} := \sup_{y \geqslant 0} ye^{-y/512}$. *Then*

$$\mathbb{P}\left(\frac{q_1}{2} - \frac{q_1^2}{2} - X \geqslant x\sigma(\mathbf{p})\right) \leqslant K_{10.9} \exp\left(-\frac{x^2}{2q_1}\right) \text{ for } x \geqslant 1/16.$$

Proof: Let $\{U_i : i \in [m]\}$ be *iid* Uniform(0,1) random variables (independent of the tree and the coin tosses as well). Define the random variables

$$Y := \sum_{i \notin A} p_i \mathbb{1}\left\{U_i \leqslant q_1/2\right\}.$$

By [8, Equation 40]

$$\mathbb{P}(X \in \cdot) \leqslant \frac{1}{q_1} \mathbb{P}(Y \in \cdot). \tag{10.22}$$

Consider the centered version

$$\tilde{Y} = \sum_{i \notin A} p_i \Big(\mathbb{1}\left\{U_i \leqslant q_1/2\right\} - q_1/2\Big)$$

Then note that

$$\mathbb{P}\left(\frac{q_1}{2} - \frac{q_1^2}{2} - Y \geqslant x\sigma(\mathbf{p})\right) = \mathbb{P}\left(-\tilde{Y} \geqslant x\sigma(\mathbf{p})\right). \tag{10.23}$$

By Markov's inequality, for any $\lambda > 0$

$$\begin{aligned}\mathbb{P}\left(-\frac{\tilde{Y}}{(\sigma(\mathbf{p}))^2} \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \exp\left(-\frac{\lambda x}{\sigma(\mathbf{p})}\right) \prod_{i \notin A} \exp\left(\frac{\lambda q_1 p_i}{2(\sigma(\mathbf{p}))^2}\right)\\ \times \prod_{i \notin A}\left[1 - \frac{q_1}{2}\left(1 - \exp\left(-\frac{\lambda p_i}{\sigma^2(\mathbf{p})}\right)\right)\right]\end{aligned}$$

The simple inequality $1 - x \leqslant \exp(-x)$ for $x \geqslant 0$ and some algebra gives

$$\mathbb{P}\left(-\frac{\tilde{Y}}{(\sigma(\mathbf{p}))^2} \geqslant \frac{x}{\sigma(\mathbf{p})}\right) \leqslant \exp\left(-\frac{\lambda x}{\sigma(\mathbf{p})}\right) \prod_{i \notin A} \exp\left[\frac{q_1}{2}\left(\frac{\lambda p_i}{\sigma^2(\mathbf{p})} - 1 + \exp\left(-\frac{\lambda p_i}{\sigma^2(\mathbf{p})}\right)\right)\right]$$

Since $e^{-u} - 1 + u \leqslant u^2/2$ for all $u \geqslant 0$, we finally get

$$\mathbb{P}\left(-\tilde{Y} \geqslant x\sigma(\mathbf{p})\right) \leqslant \exp\left(-\frac{\lambda x}{\sigma(\mathbf{p})}\right) \exp\left(\frac{q_1 \lambda^2}{4\sigma^2(\mathbf{p})}\right)$$

Taking $\lambda = 2x\sigma(\mathbf{p})/q_1$, we get

$$\mathbb{P}\left(-\tilde{Y} \geqslant x\sigma(\mathbf{p})\right) \leqslant \exp\left(-\frac{x^2}{q_1}\right).$$

Using (10.22), (10.23) and $x^2 \geqslant x^2/2 + 1/512$, we arrive at

$$\begin{aligned}\mathbb{P}\left(\frac{q_1}{2} - \frac{q_1^2}{2} - X \geqslant x\sigma(\mathbf{p})\right) &\leqslant \frac{1}{q_1} \exp\left(-\frac{x^2}{q_1}\right) \leqslant \frac{1}{q_1} \exp\left(-\frac{1}{512 q_1}\right) \exp\left(-\frac{x^2}{2q_1}\right)\\ &\leqslant K_{10.9} \exp\left(-\frac{x^2}{2q_1}\right).\end{aligned} \tag{10.24}$$

The proof of Proposition 10.9 is completed. ■

For $v\in[m]$, set $\mathscr{A}(v)$ equal to $\mathscr{A}(i)$ (recall the definition of active vertices $\mathscr{A}(i)$ from Section 7.1) if $v=v(i)$ is the i-th vertex vertex explored in the depth-first exploration of the tree. If we fix any $A\subset[m]$ and $v\in[m]$ and condition a random $\mathbf{p}$-tree on $\mathbb{A}(v)=A$, then by [8, Lemma 11] we have $X\leqslant p(\mathscr{A}(v))$ where X is as in (10.21). Since $p(\mathscr{A}(v))=F^{\mathrm{exc},\mathbf{p}}(e(v))$, we conclude from Proposition 10.9 that for $x\geqslant 1$ and $v\in[m]$,

$$\mathbb{P}\left(\frac{\mathscr{G}(v)}{2}-\frac{\mathscr{G}^2(v)}{2}-F^{\mathrm{exc},\mathbf{p}}(e(v))\geqslant\frac{x\sigma(\mathbf{p})}{16}\Big|\mathbb{A}(v)\right)\leqslant K_{10.9}\exp\left(-\frac{x^2}{512p(\mathbb{A}(v))}\right).$$

On the set $\{\mathscr{G}(v)\leqslant x\sigma(\mathbf{p})\}$ we have $p(\mathbb{A}(v))=\mathscr{G}(v)\leqslant x\sigma(\mathbf{p})$. Hence, for $x\geqslant 1$ and $v\in[m]$

$$\mathbb{P}\left(\frac{\mathscr{G}(v)}{2}-\frac{\mathscr{G}^2(v)}{2}-F^{\mathrm{exc},\mathbf{p}}(e(v))\geqslant\frac{x\sigma(\mathbf{p})}{16},\mathscr{G}(v)\leqslant x\sigma(\mathbf{p})\right)\leqslant K_{10.9}\exp\left(-\frac{x}{512\sigma(\mathbf{p})}\right).$$

Since $x\geqslant 1$ and $m\leq 1/p_{\min}\leq[\sigma(\mathbf{p})]^{-r_0}$, this yields

$$\begin{aligned}&\mathbb{P}\left(\frac{\mathscr{G}(v)}{2}-\frac{\mathscr{G}^2(v)}{2}-F^{\mathrm{exc},\mathbf{p}}(e(v))\geqslant\frac{x\sigma(\mathbf{p})}{16}\text{ and }\mathscr{G}(v)\leqslant x\sigma(\mathbf{p})\text{ for some }v\in[m]\right)\\&\leqslant K_{10.9}m\exp\left(-\frac{x}{512\sigma(\mathbf{p})}\right)\leqslant\left[\frac{K_{10.9}}{[\sigma(\mathbf{p})]^{r_0}}\exp\left(-\frac{1}{2^{10}\sigma(\mathbf{p})}\right)\right]\exp\left(-\frac{x}{2^{10}\sigma(\mathbf{p})}\right)\\&\leqslant C_1e^{-x/2^{10}},\end{aligned}$$

where $C_1:=K_{10.9}\sup_{y\geqslant 0}y^{r_0}e^{-y/1024}$. Combine the above bound and Proposition 10.7, we have we have, for $1\leqslant x\leqslant 8[\sigma(\mathbf{p})]^{-2\epsilon_0}$ and $r\geqslant r_0$,

$$\begin{aligned}&\mathbb{P}\left(\frac{\mathscr{G}(v)}{2}-F^{\mathrm{exc},\mathbf{p}}(e(v))\geqslant\frac{x\sigma(\mathbf{p})}{8}\text{ and }\mathscr{G}(v)\leqslant x\sigma(\mathbf{p})\text{ for some }v\in[m]\right)\\&\leqslant C_1e^{-x/2^{10}}+\mathbb{P}\left(\frac{1}{2}\max_{v\in[m]}\mathscr{G}(v)^2\geqslant\frac{x\sigma(\mathbf{p})}{16}\right)\leqslant C_1e^{-x/2^{10}}+\frac{8^rK_{10.7}}{x^r}.\end{aligned}\tag{10.25}$$

Define

$$E:=\left\{\frac{\|F^{\mathrm{exc},\mathbf{p}}\|_{L^\infty}}{\sigma(\mathbf{p})}\leqslant\frac{x}{8}\right\}\bigcap\left\{\frac{\mathscr{G}(v)}{2\sigma(\mathbf{p})}-\frac{F^{\mathrm{exc},\mathbf{p}}(e(v))}{\sigma(\mathbf{p})}\geqslant\frac{x}{8}\text{ and }\frac{\mathscr{G}(v)}{\sigma(\mathbf{p})}\leqslant x\text{ for some }v\in[m]\right\}.$$

Then E and $\mathscr{B}_2$ are same provided $x\geqslant 2p_{\max}/\sigma(\mathbf{p})$. Indeed, if $\left\{\|F^{\mathrm{exc},\mathbf{p}}\|_{L^\infty}\leqslant x\sigma(\mathbf{p})/8\right\}$ holds and $\{\mathscr{G}(v_0)/2-F^{\mathrm{exc},\mathbf{p}}(e(v_0))\geqslant x\sigma(\mathbf{p})/8\}$ holds for some v_0 with $\mathscr{G}(v_0)\leqslant x\sigma(\mathbf{p})$, then E is true. On the other hand, if $\mathscr{G}(v_0)>x\sigma(\mathbf{p})$, then there is an ancestor v_1 of v_0 satisfying $x\sigma(\mathbf{p})/2\leqslant\mathscr{G}(v_1)\leqslant x\sigma(\mathbf{p})$ (this is true since $x\sigma(\mathbf{p})\geqslant 2p_{\max}$). For this v_1, we have

$$\frac{\mathscr{G}(v_1)}{2}-F^{\mathrm{exc},\mathbf{p}}(e(v_1))\geqslant\frac{x\sigma(\mathbf{p})}{4}-\frac{x\sigma(\mathbf{p})}{8}=\frac{x\sigma(\mathbf{p})}{8}.$$

Thus, the event E is still true. Since $2p_{\max}/\sigma(\mathbf{p})\leqslant 1$ under Assumption 10.1, we conclude from (10.25) that for $r\geqslant r_0$, there exists some constant $K_{10.6}(r)$ depending only on r such that for $1\leqslant x\leqslant 8[\sigma(\mathbf{p})]^{-2\epsilon_0}$, we have

$$\mathbb{P}(\mathscr{B}_2)=\mathbb{P}(E)\leqslant\frac{K_{10.6}}{x^r}.$$

This completes the proof of Lemma 10.6. ■

Proof of Lemma 10.2: Combining (10.4), (10.5), (10.6) and (10.13), we conclude that, for any $r > \max\{r_0, \lfloor r_0/2\epsilon_0 \rfloor\}$, there exists some constant $K(r) > 0$ such that for all $1 \leqslant x \leqslant 8[\sigma(\mathbf{p})]^{-2\epsilon_0}$, we have

$$\mathbb{P}(\mathrm{ht}(\mathscr{T}) \geqslant x\sigma(\mathbf{p})) \leqslant \frac{K(r)}{x^r}. \tag{10.26}$$

This completes the proof of Lemma 10.2. ■

11. Proof of Lemma 9.1 and Lemma 9.2

11.1. Proof of Lemma 9.1.

Proof of (9.6)**:** For each $v \in [n]$, define the random permutation π_v as follows: $\pi_v(1) = v$ and $(\pi_v(2), \ldots, \pi_v(n))$ is a size-biased permutation of $[n]\setminus\{v\}$ where size of j is w_j. Then $(\pi_v(i) : i \geq 1)$ has the same law as the sequence of vertices of the random graph $\mathscr{G}_n^{\mathrm{nr}}(\mathbf{w}, \lambda)$ appear in a size-biased order during a breadth-first search starting from the vertex v. For ease of notation, we fix v and write $\bar{w}_i := w_{\pi_v(i)}$ in the rest of the proof.

Hence,

$$Q_v := \mathbb{P}\left(X_n(v;k) \geqslant \frac{32\sigma_{k+1}m}{\sigma_1} \text{ and } |\mathscr{C}_n(v)| \leqslant m\right) \leq \mathbb{P}\left(\sum_{i=1}^{m} \bar{w}_i^k \geqslant \frac{32\sigma_{k+1}m}{\sigma_1}\right).$$

By Assumption 3.1 (a), there exists $n_1 > 0$ such that when $n \geq n_1$, we have

$$\frac{\sigma_k}{2} < \frac{\sum_{i=1}^n w_i^k}{n} < 2\sigma_k, \text{ for } k = 1,2,3. \tag{11.1}$$

We only give the proof when $k = 2$, and the case when $k = 1$ is similar. Let $\mathscr{F}_j^v = \sigma\{\pi_v(i) : 1 \leqslant i \leqslant j\}$, for $1 \leq j \leq m$. Note that, for $2 \leqslant j \leqslant m$ and $n \geq n_1$, we have

$$\mathbb{E}\left(\bar{w}_j^2 \middle| \mathscr{F}_{j-1}^v\right) = \frac{\sum_{i=j}^n \bar{w}_i^3}{\sum_{i=j}^n \bar{w}_i} \leq \frac{\sum_{i=1}^n w_i^3}{\sum_{i=1}^n w_i - m w_{\max}} \leq \frac{2\sigma_3 n}{\sigma_1 n/2 - m w_{\max}}. \tag{11.2}$$

By Assumption 3.2, there exists $n_2 > 0$ such that when $n \geq n_2$,

$$m w_{\max} \leq n^{47/48} w_{\max} < \frac{\sigma_1 n}{4}, \text{ and } w_{\max}^2 < \frac{8\sigma_3 n^{1/12-2\eta_0}}{\sigma_1} \leq \frac{8\sigma_3 m}{\sigma_1}.$$

By the above bound and (11.2),

$$\bar{w}_1^2 + \sum_{j=2}^m \mathbb{E}[\bar{w}_j^2 \mid \mathscr{F}_{j-1}^v] \leq w_{\max}^2 + m \cdot \frac{8\sigma_3}{\sigma_1} < \frac{16\sigma_3 m}{\sigma_1}.$$

Thus writing $\Delta_j := \bar{w}_j^2 - \mathbb{E}\left[\bar{w}_j^2 \middle| \mathscr{F}_{j-1}^v\right]$, by the Burkholder-Davis-Gundy inequality, we have for any integer $r' > 0$,

$$Q_v \leqslant \mathbb{P}\left(|\sum_{j=2}^m \Delta_j| \geqslant \frac{16\sigma_3 m}{\sigma_1}\right) \leq \left(\frac{\sigma_1}{16\sigma_3 m}\right)^{2r'} \mathbb{E}\left(\sum_{j=2}^m \Delta_j^2\right)^{r'}.$$

Since $\Delta_j \le w_{\max}^2$, and by Assumption 3.2, there exists $n_3 > 0$ such that when $n > n_3$, $w_{\max} < 1/n^{1/48-\eta_0}$, therefore

$$Q_v \leqslant \left(\frac{\sigma_1}{16\sigma_3}\right)^{2r'} \cdot \frac{w_{\max}^{4r'}}{m^{r'}} \le \left(\frac{\sigma_1}{16\sigma_3}\right)^{2r'} \cdot \left(\frac{n^{1/12-4\eta_0}}{m^{1/12-2\eta_0}}\right)^{r'} = \left(\frac{\sigma_1}{16\sigma_3}\right)^{2r'} \cdot \frac{1}{n^{2\eta_0 r'}}. \tag{11.3}$$

Then taking $n_0 = \max\{n_1, n_2, n_3\}$ and $r' = \lfloor r/2\eta_0 \rfloor + 1$, we have proved (9.6).

Proof of (9.7)**:** The idea is similar to the proof of (9.6). Using the same notation, we have

$$\mathbb{P}\left(X_n(v;k) \leqslant \frac{\sigma_{k+1} m}{16\sigma_1} \text{ and } |\mathscr{C}_n(v)| \geqslant m\right) \le \mathbb{P}\left(\sum_{i=1}^m \bar{w}_i^k \leqslant \frac{\sigma_{k+1} m}{16\sigma_1}\right).$$

Then when n is large such that (11.1) is true and

$$m w_{\max}^3 \le n^{45/48} w_{\max}^3 \le \frac{\sigma_3 n}{4},$$

we have

$$\mathbb{E}\left(\bar{w}_j^2 \middle| \mathscr{F}_{j-1}^v\right) \ge \frac{\sum_{i=1}^n w_i^3 - m w_{\max}^3}{\sum_{i=1}^n w_i} \ge \frac{\sigma_3}{8\sigma_1}.$$

Then we can use a bound similar to (11.3) to complete the proof. Then we choose $K_{9.1}$ to be the largest constant shown up in the four bounds. The proof of Lemma 9.1 is completed. ■

11.2. **Proof of Lemma 9.2.** Note the that in this section, the constant γ_0 comes from Assumption 3.1, and the constant η_0 comes from Assumption 3.2.

Proof of Lemma 9.2: For convenience, we will write $(\overline{w}_1, \dots, \overline{w}_m)$ for $(w_j : j \in \mathscr{C}_n^{(i)})$, where $m = m^{(i)} := |\mathscr{C}_n^{(i)}|$. Let $p_j^{(i)} = \overline{w}_j / X_{n,i}(1)$ for $1 \leqslant j \leqslant m$ and let $a^{(i)}$ be as in the statement of Proposition 6.1. Define $p_{\max}^{(i)} := \max_{j\in[m]} p_j^{(i)}$ and $p_{\min}^{(i)} := \min_{j\in[m]} p_j^{(i)}$. Further, let $L^{(i)}(\mathbf{t})$ be as in in (7.6) with $a^{(i)}, p_k^{(i)}, p_\ell^{(i)}$ replacing a, p_k, p_ℓ respectively. In the rest of the proof, we will hide i from the notation occasionally.

Note that, $L(\mathbf{t}) \geqslant 1$ for any $\mathbf{t} \in \mathbb{T}_m^{\text{ord}}$. Define $\mathbb{P}^{(i)}(\cdot) := \mathbb{P}_{\text{ord}}(\cdot; \mathbf{p}^{(i)})$ where the latter is defined in (2.8). Thus, it follows from Proposition 6.1 and Proposition 7.4 that

$$\begin{aligned} \mathbb{E}\left[(\text{diam}(\mathscr{C}_n^{(i)}))^4 \mid \mathscr{F}_{ptn}\right] &\leqslant \frac{\int \text{ht}^4(\mathbf{t}) L(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t})}{\int L(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t})} \leqslant \int \text{ht}^4(\mathbf{t}) L(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t}) \\ &\leqslant \left(\int \text{ht}^8(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t})\right)^{1/2} \left(\int L^2(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t})\right)^{1/2} \end{aligned} \tag{11.4}$$

Define $r_0 := 2\gamma_0/\alpha_0 + 2$ and $\epsilon_0 := 6\eta_0$. Define the events

$$H_n^{(i)} := \left\{\sigma(\mathbf{p}^{(i)}) \leqslant \frac{1}{2^{10}}, \frac{p_{\max}^{(i)}}{[\sigma(\mathbf{p}^{(i)})]^{3/2+\epsilon_0}} \leqslant 1, \frac{[\sigma(\mathbf{p}^{(i)})]^{r_0}}{p_{\min}^{(i)}} \leqslant 1, a^{(i)} \sigma(\mathbf{p}^{(i)}) \leqslant \frac{1}{16}\right\}. \tag{11.5}$$

Then restricted to $H_n^{(i)}$, applying Theorem 3.7 with $r = 9$, we have

$$[\sigma(\mathbf{p})]^8 \int \text{ht}^8(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t}) \leqslant \int_0^\infty 8x^7\, \mathbb{P}^{(i)}\left(\sigma(\mathbf{p})\,\text{ht}(\mathbf{t}) \geqslant x\right) dx \leqslant 8 + \int_1^\infty 8x^7 \cdot \frac{K_{3.7}(9)}{x^9} dx, \tag{11.6}$$

Restricted to $H_n^{(i)}$, applying Corollary 7.13 with $B_1 = 1/16$, $B_2 = 1$, and $\gamma = 2$, we have

$$\int L^2(\mathbf{t}) d\,\mathbb{P}^{(i)}(\mathbf{t}) \leqslant K_{7.13}\left(2, \frac{1}{16}, 1\right). \tag{11.7}$$

By (11.4), (11.6) and (11.7), the proof is completed once we show the following: there exist n_0 such that for all $n \geqslant n_0$, we have

$$E_n(\alpha_0) \cap \left\{n^{\alpha_0} \leqslant |\mathscr{C}_n^{(i)}| \leqslant \eta n^{2/3}\right\} \subset H_n^{(i)} \text{ for all } i \geqslant 1. \tag{11.8}$$

Restricted on $E_n(\alpha_0)$, there exist absolute constants $C_1, C_2, C_3, C_4 > 0$ such that for all $i \geqslant 1$ and $n \geqslant 1$, we have

$$\frac{C_1}{\sqrt{m}} \leqslant \sigma(\mathbf{p}^{(i)}) = \frac{\sqrt{X_{n,i}(2)}}{X_{n,i}(1)} \leqslant \frac{C_2}{\sqrt{m}},$$
$$p_{\max} \leqslant \frac{C_3 w_{\max}}{m}, \ \ p_{\min} \geqslant \frac{C_4 w_{\min}}{m}.$$

The following calculation will be restricted to $E_n(\alpha_0) \cap \left\{n^{\alpha_0} \leqslant |\mathscr{C}_n^{(i)}| \leqslant \eta n^{2/3}\right\}$. Note that

$$\begin{aligned}\frac{p_{\max}^{(i)}}{[\sigma(\mathbf{p}^{(i)})]^{3/2+\epsilon_0}} &\leqslant \frac{C_3 w_{\max}/m}{(C_1/\sqrt{m})^{3/2+\epsilon_0}} = \frac{C_2}{C_1^{3/2+\epsilon_0}} \cdot \frac{w_{\max}}{m^{1/4-\epsilon_0/2}} \\ &\leqslant \frac{C_2}{C_1^{3/2+\epsilon_0}} \cdot \frac{n^{1/48-\eta_0}}{n^{(1/12-2\eta_0)(1/4-3\eta_0)}} \leq \frac{C_2}{C_1^{3/2+\epsilon_0}} \cdot \frac{1}{n^{\eta_0/4}},\end{aligned} \tag{11.9}$$

Similarly

$$\frac{[\sigma(\mathbf{p}^{(i)})]^{r_0}}{p_{\min}^{(i)}} \leqslant \frac{(C_2/\sqrt{m})^{r_0}}{C_4 w_{\min}/m} = \frac{C_2^{r_0}}{C_4} \cdot \frac{1}{w_{\min} m^{r_0/2-1}} \leq \frac{C_2^{r_0}}{C_4} \cdot \frac{1}{w_{\min} n^{\gamma_0}} \tag{11.10}$$

By, (11.9), (11.10) and Assumption 3.1 (d), there exist n_1 such that when $n \geqslant n_1$, the first three conditions in (11.5) hold uniformly for all $i \geqslant 1$. Now we verify the last condition in (11.5). Let n_2 be such that when $n \geqslant n_2$, $|\lambda|/n^{1/3} < 1$ and $l_n > n\sigma_1/2$, then when $n \geqslant n_2$ we have,

$$a^{(i)}\sigma(\mathbf{p}^{(i)}) = \left(1 + \frac{\lambda}{n^{1/3}}\right)\frac{(X_{n,i}(1))^2}{l_n} \cdot \frac{\sqrt{X_{n,i}(2)}}{X_{n,i}(1)} \leqslant \frac{4(\overline{A})^{3/2}}{\sigma_1} \cdot \frac{(m)^{3/2}}{n} \leqslant \frac{4(\overline{A})^{3/2}}{\sigma_1} \cdot \eta^{3/2} \leqslant \frac{1}{16},$$

since $\overline{A} = 32\sigma_3/\sigma_1$ and $\eta < 2\sigma_3/\sigma_1^{1/3}$. Therefore, when $n \geqslant n_0 := \max\{n_1, n_2\}$, the claim (11.8) is true. The proof of Lemma 9.2 is completed. ■

Acknowledgements

We would like to thank Amarjit Budhiraja for many stimulating conversations. SB has been partially supported by NSF-DMS grants 1105581 and 1310002 and SES grant 1357622. SS has been supported in part by NSF grant DMS-1007524 and Netherlands Organisation for Scientific Research (NWO). XW has been supported in part by the National Science Foundation (DMS-1004418, DMS-1016441), the Army Research Office (W911NF-0-1-0080, W911NF-10-1-0158) and the US-Israel Binational Science Foundation (2008466).

References

[1] R. Abraham, J.-F. Delmas, and P. Hoscheit, *A note on the gromov-hausdorff-prokhorov distance between (locally) compact metric measure spaces*, Electron. J. Probab. **18** (2013), no. 14, 1–21.

[2] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *Critical random graphs: limiting constructions and distributional properties*, Electron. J. Probab **15** (2010), no. 25, 741–775.

[3] L. Addario-Berry, N. Broutin, and C. Goldschmidt, *The continuum limit of critical random graphs*, Probability Theory and Related Fields **152** (2012), no. 3-4, 367–406.

[4] L. Addario-Berry, N. Broutin, C. Goldschmidt, and G. Miermont, *The scaling limit of the minimum spanning tree of the complete graph*, arXiv preprint arXiv:1301.1664 (2013).

[5] L. Addario-Berry, N. Broutin, and B. Reed, *Critical random graphs and the structure of a minimum spanning tree*, Random Structures & Algorithms **35** (2009), no. 3, 323–347.

[6] R. Albert and A.-L. Barabási, *Statistical mechanics of complex networks*, Reviews of modern physics **74** (2002), no. 1, 47.

[7] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, The Annals of Probability (1997), 812–854.

[8] D. Aldous, G. Miermont, and J. Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of jeulin's local time identity*, Probability theory and related fields **129** (2004), no. 2, 182–218.

[9] 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 (2014).

[10] S. Bhamidi, A. Budhiraja, and X. Wang, *The augmented multiplicative coalescent and critical dynamic random graph models*, Accepted in Probability Theory and Related Fields (2012).

[11] 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 (2011m:60021)

[12] B. Bollobás, *The evolution of random graphs*, Transactions of the American Mathematical Society **286** (1984), no. 1, 257–274.

[13] B. Bollobás, S. Janson, and O. Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures & Algorithms **31** (2007), no. 1, 3–122.

[14] B. Bollobás and O. Riordan, *The phase transition in the erdős-rényi random graph process*, Erdős centennial, 2013, pp. 59–110.

[15] 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.

[16] L. A. Braunstein, S. V. Buldyrev, R. Cohen, S. Havlin, and H. E. Stanley, *Optimal paths in disordered complex networks*, Phys. Rev. Lett. **91** (2003Oct), 168701.

[17] T. Britton, M. Deijfen, and A. Martin-Löf, *Generating simple random graphs with prescribed degree distribution*, Journal of Statistical Physics **124** (2006), no. 6, 1377–1397.

[18] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Vol. 33, American Mathematical Society Providence, 2001.

[19] M. Camarri and J. Pitman, *Limit distributions and random trees derived from the birthday problem with unequal probabilities*, Electron. J. Probab **5** (2000), no. 2, 1–18.

[20] F. Chung and L. Lu, *The average distances in random graphs with given expected degrees*, Proceedings of the National Academy of Sciences **99** (2002), no. 25, 15879–15882.

[21] F. Chung and L. Lu, *Connected components in random graphs with given expected degree sequences*, Annals of combinatorics **6** (2002), no. 2, 125–145.

[22] F. R. K Chung and L. Lu, *Complex graphs and networks*, AMS Bookstore, 2006.

[23] R. Durrett, *Random graph dynamics*, Vol. 20, Cambridge university press, 2007.

[24] P. Erdős, *On the evolution of random graphs*, Publ. Math. Inst. Hungar. Acad. Sci **5** (1960), 17–61.

[25] S. N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008. Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR2351587 (2009d:60014)

[26] H. Hatami and M. Molloy, *The scaling window for a random graph with a given degree sequence*, Random Structures & Algorithms **41** (2012), no. 1, 99–123.

[27] R. v. d. Hofstad, *Critical behavior in inhomogeneous random graphs*, Preprint (2009).

[28] S. Janson, *Asymptotic equivalence and contiguity of some random graphs*, Random Structures & Algorithms **36** (2010), no. 1, 26–45.

[29] S. Janson, D. E Knuth, T. Łuczak, and B. Pittel, *The birth of the giant component*, Random Structures & Algorithms **4** (1993), no. 3, 233–358.

[30] A. Joseph, *The component sizes of a critical random graph with given degree sequence*, arXiv preprint arXiv:1012.2352 (2010).

[31] J.-F. Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR2203728 (2007h:60078)

[32] T. Łuczak, *Component behavior near the critical point of the random graph process*, Random Structures & Algorithms **1** (1990), no. 3, 287–310.

[33] T. Łuczak, B. Pittel, and J. C Wierman, *The structure of a random graph at the point of the phase transition*, Transactions of the American Mathematical Society **341** (1994), no. 2, 721–748.

[34] P. Massart, *The tight constant in the dvoretzky-kiefer-wolfowitz inequality*, The Annals of Probability **18** (199007), no. 3, 1269–1283.

[35] A. Nachmias and Y. Peres, *Critical percolation on random regular graphs*, Random Structures & Algorithms **36** (2010), no. 2, 111–148.

[36] M. E. Newman, *The structure and function of complex networks*, SIAM review **45** (2003), no. 2, 167–256.

[37] I. Norros and H. Reittu, *On a conditionally poissonian graph process*, Advances in Applied Probability **38** (2006), no. 1, 59–75.

[38] J. Pitman, *Random mappings, forests, and subsets associated with abel-cayley-hurwitz multinomial expansions*, Sém. Lothar. Combin **46** (2001), 02.

[39] O. Riordan, *The phase transition in the configuration model*, Combinatorics, Probability & Computing **21** (2012), no. 1-2, 265–299.

[40] L C. G Rogers and D. Williams, *Diffusions, markov processes, and martingales: Volume 1, foundations*, Cambridge university press, 2000.

[41] R. J Serfling et al., *Probability inequalities for the sum in sampling without replacement*, The Annals of Statistics **2** (1974), no. 1, 39–48.

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

[43] T. S Turova, *Diffusion approximation for the components in critical inhomogeneous random graphs of rank 1*, arXiv preprint arXiv:0907.0897 (2009).

[44] R. Van Der Hofstad, *Random graphs and complex networks*, Available on http://www. win. tue. nl/rhofstad/NotesRGCN. pdf (2009).