Submitted to Bernoulli
arXiv: arXiv:1704.02064

Scaling limit of random forests with prescribed degree sequences

TAO LEI[1,*]

[1] *805 Sherbrooke Street West, Montréal, Québec, Canada H3A 0B9*
E-mail: [*]tao.lei@mail.mcgill.ca

In this paper, we consider the random plane forest uniformly drawn from all possible plane forests with a given degree sequence. Under suitable conditions on the degree sequences, we consider the limit of a sequence of such forests with the number of vertices tends to infinity in terms of Gromov-Hausdorff-Prokhorov topology. This work falls into the general framework of showing convergence of random combinatorial structures to certain Gromov-Hausdorff scaling limits, described in terms of the *Brownian Continuum Random Tree* (BCRT), pioneered by the work of Aldous [6–8]. In fact we identify the limiting random object as a sequence of random real trees encoded by excursions of some first passage bridges reflected at minimum. We establish such convergence by studying the associated Lukasiewicz walk of the degree sequences. In particular, our work is closely related to and uses the results from the recent work of Broutin and Marckert [16] on scaling limit of random trees with prescribed degree sequences, and the work of Addario-Berry [3] on tail bounds of the height of a random tree with prescribed degree sequence.

Keywords: random forests, first passage bridge, Gromov-Hausdorff-Prokhorov distance.

1. Introduction

Scaling limits for finite graphs is a topic at the intersection of combinatorics and probability. In this paper, we investigate the Gromov-Hausdorff-Prokhorov convergence of random forests with prescribed degree sequence. Our work is a natural continuation of [16] where it is shown that under natural hypotheses on the degree sequences, after suitable normalization, uniformly random trees with given degree sequence converge to Brownian continuum random tree, with the size of trees going to infinity.

In a series of papers [6–8], Aldous introduced the concept of Brownian continuum random tree (BCRT) and showed that *critical* Galton-Watson tree conditioned on its size has BCRT as limiting objects. Since then, many families of graphs have been shown to have BCRT or random processes derived from BCRT as their limiting objects. For example, multi-type Galton-Watson trees [26], unordered binary trees [24], critical Erdös-Rényi random graph [4], random planar maps with a unique large face [22], random planar quadrangulations with a boundary [13].

As in [16], our combinatorial model is motivated by the metric structure of graphs with a prescribed degree sequence. This model was first introduced by Bender and Canfield [11] and by Bollobás [15] in the form of the configuration model. This model can give rise to graphs with any particular (legitimate) prescribed degree sequence (including, e.g., heavy tailed degree distributions, a feature which is observed in realistic networks but is not captured by the Erdös-Rényi random graph model).

Our main results, which are stated formally in Section 1.2, are that, under natural assumptions on degree sequences and after suitable normalization, large uniformly random forests with given degree sequence converge in distribution to the forests coded by Brownian first passage bridge, with respect to the Gromov-Hausdorff-Prokhorov topology. In order to present these results rigorously, we need the following subsection to introduce the necessary concepts and notations involved.

1.1. Definitions and Notations

Plane trees and forests

Let

$$\mathcal{U} = \bigcup_{n=0}^{\infty} \mathbb{N}^n,$$

where $\mathbb{N} = \{1, 2, \cdots\}$ and $\mathbb{N}^0 = \{\emptyset\}$ and *rooted plane trees* T are defined to be certain subset of $\mathcal{U}$, with root $\emptyset$. We refer the readers to [19] for the formal definition of plane trees. For $u \in \mathrm{T}$, let $k_{\mathrm{T}}(u)$ be the *degree* of u in T, that is, the number of children of u. We denote the lexicographic order on $\mathcal{U}$ by $<$ (e.g. $\emptyset < 11 < 21 < 22$). The lexicographic order on $\mathcal{U}$ induces a total order on the set of all rooted plane trees.

We call a finite sequence of finite rooted plane trees $\mathrm{F} = (\mathrm{T}_1, \mathrm{T}_2, \cdots, \mathrm{T}_m)$ a *rooted plane forest*. For forest F, we let $\mathrm{F}^{\downarrow}$ be the sequence of tree components of F in decreasing order of size, breaking ties lexicographically.

***Definition* 1.1.** A *degree sequence* is a sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$ of non-negative integers with $\sum_{i\geq 0} s^{(i)} < \infty$ such that $c(\mathbf{s}) := \sum_{i\geq 0} (1-i) s^{(i)} > 0$. For a plane tree T, the degree sequence $\mathbf{s}(\mathrm{T}) = (s^{(i)}(\mathrm{T}), i \geq 0)$ is given by

$$s^{(i)}(\mathrm{T}) = |\{u \in \mathrm{T} : k_{\mathrm{T}}(u) = i\}|.$$

For a plane forest $\mathrm{F} = (\mathrm{T}_1, \cdots, \mathrm{T}_m)$, the degree sequence $\mathbf{s}(\mathrm{F}) = (s^{(i)}(\mathrm{F}), i \geq 0)$ is given by

$$s^{(i)}(\mathrm{F}) = \sum_{j=1}^{m} s^{(i)}(\mathrm{T}_j).$$

Note that $c(\mathbf{s}(\mathrm{T})) = 1$ for any plane tree T. In general , the number of tree components in F is always $c(\mathbf{s}(\mathrm{F}))$. For any degree sequence $\mathbf{s}$, we adopt the notations

$$n(\mathbf{s}) := \sum_{i\geq 0} s^{(i)}, \quad \Delta(\mathbf{s}) := \max\{i : s^{(i)} > 0\}.$$

Figure 1.1, below, shows a plane forest with degree sequence $\mathbf{s} = (7, 2, 2, 1, 0, \cdots)$ with $s^{(i)} = 0$ for $i \geq 4$.

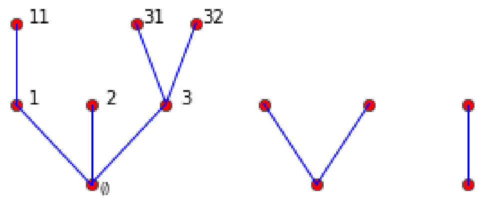

Figure 1. A plane forest (with labels for the first tree) with degree sequence $\mathbf{s} = (7, 2, 2, 1, 0, \cdots)$

For any degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$, we let $\mathrm{F}(\mathbf{s})$ denote the set of all plane forests with degree sequence $\mathbf{s}$. Let $\mathbb{P}_{\mathbf{s}}$ be the uniform measure on $\mathrm{F}(\mathbf{s})$ and let $\mathbb{F}(\mathbf{s})$ be a random plane forest with law $\mathbb{P}_{\mathbf{s}}$.

First passage bridge

We also need to recall the following definition of *first passage bridge* as in [10]. Informally, for $\lambda > 0$, the *first passage bridge* of unit length from 0 to $-\lambda$, denoted F^{br}_λ, is a $C[0, 1]-$valued random variable with law

$$(F^{br}_\lambda(t), 0 \leq t \leq 1) \stackrel{d}{=} (B(t), 0 \leq t \leq 1 \mid T_\lambda = 1)$$

where B is a standard Brownian motion and $T_\lambda := \inf\{t : B(t) < -\lambda\}$ is the first passage time below level $-\lambda < 0$.

For $l \geq 0$, we write B^{br}_l for the Brownian bridge of duration 1 from 0 to $-l$. As explained in Proposition 1 of [21], the law of the Brownian bridge B^{br}_l is characterized by $B^{br}_l(1) = -l$ and the formula

$$\mathbf{E}\left[f((B^{br}_l(t))_{0 \leq t \leq m})\right] = \mathbf{E}\left[f((B(t))_{0 \leq t \leq m})\frac{p_{1-m}(-l - B(m))}{p_1(-l)}\right] \tag{1.1}$$

for all bounded measurable function f, and all $0 \leq m < 1$, where p_a is the Gaussian density with variance a and mean 0, that is, $p_a(x) = \frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a}}$. In a similar way the law of F^{br}_λ can be defined as the law such that

$$\mathbf{E}\left[f((F^{br}_\lambda(t))_{0 \leq t \leq s})\right] = \mathbf{E}\left[(f(B(t))_{0 \leq t \leq s})\frac{p'_{1-s}(-\lambda - B(s))}{p'_1(-\lambda)} \mathbb{1}_{\{\inf\limits_{r \leq s} B(r) > -\lambda\}}\right] \tag{1.2}$$

for all bounded measurable functions f and all $0 \leq s < 1$ and $F^{br}_\lambda(1) = -\lambda$, where p'_a is the derivative of p_a. These formulae set the finite-dimensional laws of the first passage bridge. In [12] (see Section 5.1 for details) it is shown that it admits a continuous version, and that F_λ^{br} is the weak limit of F_λ^ϵ where $(F_\lambda^\epsilon(t), 0 \le t \le 1)$ has the law of B conditioned on the event $\{B(1) < -\lambda + \epsilon, \ \inf_{s\le 1} B(s) > -\lambda - \epsilon\}$, hence justifying the informal conditioning definition.

Gromov-Hausdorff-Prokhorov distance

We refer readers to standard literature (e.g. [17]) for the definition of the Gromov-Hausdorff distance d_{GH}. A *rooted measured metric space* $\mathcal{X} = (X, d, \emptyset, \mu)$ is a metric space (X, d) with a distinguished element $\emptyset \in X$ and a finite Borel measure μ. Note that the definitions work in more general settings, e.g. μ could be a boundedly finite Borel measure (see [2]), but for the purpose of this paper, finite measure μ is enough. We refer to [2] for a careful definition of the Gromov-Hausdorff-Prokhorov distance d_{GHP}. Let $\mathbb{K}$ denote the set of GHP-isometry classes of compact rooted measured metric spaces and we identify $\mathcal{X}$ with its GHP-isometry class. Then Theorem 2.5 in [2] shows that $(\mathbb{K}, d_{GHP})$ is a Polish metric space. We next define a distance between sequences of rooted measured metric spaces. For $\mathbf{X} = (\mathcal{X}_j, j \ge 1), \mathbf{X}' = (\mathcal{X}'_j, j \ge 1)$ in $\mathbb{K}^\mathbb{N}$, we let

$$d_{GHP}^\infty(\mathbf{X}, \mathbf{X}') = \sup_{j\ge 1} d_{GHP}(\mathcal{X}_j, \mathcal{X}'_j).$$

If $\mathbf{X} \in \mathbb{K}^n$ for some $n \in \mathbb{N}$, in order to view $\mathbf{X}$ as a member of $\mathbb{K}^\mathbb{N}$, we append to $\mathbf{X}$ an infinite sequence of zero metric spaces $\mathcal{Z}$. Here $\mathcal{Z}$ is the rooted measured metric space consisting of a single point with measure 0. Let $\mathbf{Z} = (\mathcal{Z}, \mathcal{Z}, \cdots)$ and

$$\mathbb{L}_\infty = \{\mathbf{X} \in \mathbb{K}^\mathbb{N} : \limsup_{j\to\infty} d_{GHP}(\mathcal{X}_j, \mathcal{Z}) = 0\}.$$

By definition of GHP distance it is not hard to see that $d_{GHP}(\mathcal{X}, \mathcal{Z}) = \frac{\mathrm{diam}(X)}{2} + \mu(X)$, hence $\mathbf{X} \in \mathbb{L}_\infty$ if and only if $\limsup_{j\to\infty} (\mathrm{diam}(X_j) + \mu_j(X_j)) = 0$. It is likewise straightforward to show that $(\mathbb{L}_\infty, d_{GHP}^\infty)$ is a complete separable metric space.

Real trees

Next we briefly recall the concepts of real trees and their continuous function encodings. A more lengthy presentation about the probabilistic aspects of real trees can be found in [20, 23]. Real trees are defined to be certain compact metric space [23]. A real tree (T, d) is *rooted* if there is a distinguished vertex (the root) $\emptyset \in T$ and we denote a rooted real tree by $(T, d, \emptyset)$. If there is a finite Borel measure μ on T, then $(T, d, \emptyset, \mu)$ is a measured rooted real tree. Real trees can be constructed from continuous functions. Let $g : [0, \infty) \to [0, \infty)$ be a continuous function with compact support and such that $g(0) = 0$. For every $s, t \ge 0$, let

$$d_g^\circ(s, t) = g(s) + g(t) - 2m_g(s, t)$$

where

$$m_g(s,t) = \min_{s\wedge t\le r\le s\vee t} g(r).$$

The function d_g° is a pseudometric on $[0,\infty)$. Define an equivalence relation $\sim$ on $[0,\infty)$ by setting $s\sim t$ iff $d_g^\circ(s,t)=0$. Then let $T_g=[0,\infty)/\sim$ and let d_g be the induced distance on T_g. Then (T_g,d_g) is a real tree (see, e.g. Theorem 2.2 in [23]).

Let $\emptyset_g$ denote the equivalence class of 0. Let p_g be the canonical projection from $[0,\infty)$ to T_g and $\sigma_g = sup\{t : g(t)>0\}$. Let $\mathbf{m}_g$ be the push forward of the Lebesgue measure on $[0,\sigma_g]$ ((σ_g,∞) has measure 0) by p_g. Then $\mathcal{T}_g=(T_g,d_g,\emptyset_g,\mathbf{m}_g)$ is a compact measured rooted real tree. In particular, $\mathcal{T}_g\in\mathbb{K}$. Let $\mathbf{e}$ denote the standard Brownian excursion, then $\mathcal{T}_{\mathbf{e}}$ is called the *Brownian continuum random tree* (BCRT for short).

1.2. Statement of main theorems

For $c>0$, let $c\mathbf{e}\in C[0,\infty)$ denote the Brownian excursion of length c, that is $(c\mathbf{e})(s) := \sqrt{c}\,\mathbf{e}(\frac{s}{c}\wedge 1)$ for $s\ge 0$. For any probability distribution $\mathbf{p}=(p^{(i)}, i\ge 0)$ on $\mathbb{N}$, let $\mu(\mathbf{p}) = \sum_{i\ge 0} ip^{(i)}$ and $\sigma^2(\mathbf{p}) = \sum_{i\ge 0} i^2p^{(i)} - 1$.

In this paper we consider a sequence of degree sequences $(\mathbf{s}_\kappa,\kappa\in\mathbb{N})$, where $\mathbf{s}_\kappa = (s_\kappa^{(i)}, i\ge 0)$. We assume $\mathbf{n}_\kappa := \sum_{i\ge 0} s_\kappa^{(i)}\to\infty$ and let $\mathbb{F}_\kappa := \mathbb{F}(\mathbf{s}_\kappa)$ and write $\mathbb{F}_\kappa^\downarrow = (\mathbb{T}_{\kappa,l},\ l\ge 1)$. We write $\mathbf{p}_\kappa = (p_\kappa^{(i)}, i\ge 0) := (\frac{s_\kappa^{(i)}}{\mathbf{n}_\kappa}, i\ge 0)$. For $\mathbb{F}_\kappa^\downarrow = (\mathbb{T}_{\kappa,l},\ l\ge 1)$, let $\mathcal{T}_{\kappa,l}$ denote the measured rooted real tree $(\mathbb{T}_{\kappa,l}, \frac{\sigma_\kappa}{2\mathbf{n}_\kappa^{1/2}}d_{gr}, \emptyset_{\kappa,l}, \mu_{\kappa,l})$ where $\sigma_\kappa=\sigma(\mathbf{p}_\kappa)$ and $\mu_{\kappa,l}$ denotes the uniform measure putting mass $\frac{1}{\mathbf{n}_\kappa}$ on each vertex of $\mathbb{T}_{\kappa,l}$. Let $\mathcal{F}_\kappa^\downarrow = (\mathcal{T}_{\kappa,l}, l\ge 1)$. Let $\Delta_\kappa := \max\{i : s_\kappa^{(i)} > 0\}$. We are now prepared to state our main theorems.

Theorem 1.2. *Suppose that there exists a distribution $\boldsymbol{p}=(p^{(i)}, i\ge 0)$ on $\mathbb{N}$ such that $\boldsymbol{p}_\kappa$ converges to $\boldsymbol{p}$ coordinatewise. Suppose also that $\sigma(\boldsymbol{p}_\kappa)\to\sigma(\boldsymbol{p})\in(0,\infty)$. If $\frac{c(\mathbf{s}_\kappa)}{\sigma(\mathbf{p}_\kappa)\mathbf{n}_\kappa^{1/2}}\to\lambda\in(0,\infty)$, then*

$$\mathcal{F}_\kappa^\downarrow \xrightarrow{d} (\mathcal{T}_{\gamma_l}, l\ge 1) \text{ as } \kappa\to\infty, \tag{1.3}$$

with respect to the product topology for d_{GHP} where $(\gamma_l, l\ge 1)$ are the excursions of the process $(F_\lambda^{br}(s) - \inf_{s'\in(0,s)} F_\lambda^{br}(s'))_{0\le s\le 1}$, listed in decreasing order of length.

Theorem 1.3. *Under the conditions of Theorem 1.2, suppose additionally that there exists $\epsilon>0$ such that $\Delta_\kappa = O(\mathbf{n}_\kappa^{\frac{1-\epsilon}{2}})$. Then the convergence (1.3) holds in $(\mathbb{L}_\infty, d_{GHP}^\infty)$.*

***Remark* 1.1.** The assumptions of Theorem 1.2 imply that $\mu(\mathbf{p}_\kappa)\to\mu(\mathbf{p})=1$ and that $\Delta_\kappa = o({\mathbf{n}_\kappa}^{1/2})$. We include the proof of these facts in Section 1 of the supplemental materials.

***Remark* 1.2.** The pair $((\gamma_l, l \geq 1), (\mathcal{T}_{\gamma_l}, l \geq 1))$ has the same law as $((\gamma_l, l \geq 1), (\mathcal{T}_{|\gamma_l|\mathbf{e}_l}, l \geq 1))$ where $(\mathbf{e}_l, l \geq 1)$ are standard Brownian excursions, independent of each other and of $(\gamma_l, l \geq 1)$.

1.3. Key ingredients of the paper

Here we summarize the two key ingredients of this paper. The first element is the convergence of the large trees in (1.3), which is essentially given by the following proposition. For all $l \geq 1$, let $X_{\kappa,l} = \frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa}$.

***Proposition* 1.4.** Under the conditions of Theorem 1.2, for any fixed $j \geq 1$,

$$((X_{\kappa,l})_{l \leq j}, (\mathcal{T}_{\kappa,l})_{l \leq j}) \xrightarrow{d} ((|\gamma_l|)_{l \leq j}, (\mathcal{T}_{|\gamma_l|\mathbf{e}_l})_{l \leq j}) \tag{1.4}$$

as $\kappa \to \infty$, where $(\mathbf{e}_l)_{l \leq j}$ are independent copies of $\mathbf{e}$, and $(\gamma_l, l \geq 1)$ are the excursions of $(F_\lambda^{br}(s) - \inf_{s' \in (0,s)} F_\lambda^{br}(s'))_{0 \leq s \leq 1}$ ranked in decreasing order of length.

There are two parts of the convergence in (1.4). One is the convergence of the normalized sizes of large trees to lengths of excursions. This will be given by the following proposition. To state this result, we need to first introduce some notions. Let $\mathcal{C}_0(1) = \{x \in C([0,1], \mathbb{R}) : x(0) = 0\}$ For a non-negative function $g^+ \in \mathcal{C}_0(1)$, an *excursion* γ of g^+ is the restriction of g^+ to a time interval $[l(\gamma), r(\gamma)]$ such that $g^+(l(\gamma)) = g^+(r(\gamma)) = 0$ and $g^+(s) > 0$ for $s \in (l(\gamma), r(\gamma))$. In this case $[l(\gamma), r(\gamma)]$ is called an *excursion interval* of g^+. The length of the excursion is denoted as $|\gamma| = r(\gamma) - l(\gamma)$. For a function g we write $g(s) - \min_{0 \leq s' < s} g(s')$ to denote $(g(s) - \min_{0 \leq s' < s} g(s'),\ 0 \leq s \leq 1)$. For $g \in \mathcal{C}_0(1)$, sometimes we refer the excursions of $g(s) - \min_{0 \leq s' < s} g(s')$ as excursions of g. Let $l_1^\downarrow = \{x = (x_1, x_2, \cdots) : x_1 \geq x_2 \geq \cdots \geq 0, \sum_i x_i \leq 1\}$ and endow $l_1^\downarrow$ with the topology induced by the l_1 distance: $d(x,y) = \sum_i |x_i - y_i|$.

***Proposition* 1.5.** Under the hypothesises of Theorem 1.2, we have

$$(|\mathbb{T}_{\kappa,l}|/\mathbf{n}_\kappa)_{l \geq 1} \xrightarrow{d} (|\gamma_l|)_{l \geq 1} \tag{1.5}$$

in $l_1^\downarrow$, where $(\gamma_l, l \geq 1)$ are the excursions of $F_\lambda^{br}(s) - \min_{0 \leq s' \leq s} F_\lambda^{br}(s')$ ranked in decreasing order of length.

This proposition will be a corollary of the following theorem, which is the main result of Section 4. For a plane forest F, let $u_1 < u_2 < \cdots < u_{|\mathrm{F}|}$ be the nodes of F listed according to their lexicographical order in $\mathcal{U}$ in each tree component, with nodes of first tree listed first, then the nodes of second tree and so on. The *depth-first walk, or Lukasiewicz path* S_{F} is defined as follows. First set $S_{\mathrm{F}}(0) = 0$ and then let

$$S_{\mathrm{F}}(i) = \sum_{j=1}^{i} (k_{\mathrm{F}}(u_j) - 1) \text{ for } i = 1, 2, \cdots, |\mathrm{F}|.$$

We extend the definition of S_{F} to the compact interval $[0, |\mathrm{F}|]$ by linear interpolation.

Theorem 1.6. *Under the conditions of Theorem 1.2, we have*

$$\left(\frac{S_{\mathbb{F}_\kappa}(t\mathbf{n}_\kappa)}{\sigma(\mathbf{p}_\kappa)\mathbf{n}_\kappa{}^{1/2}} \right)_{t\in[0,1]} \xrightarrow{d} F_\lambda^{br} \tag{1.6}$$

in $\mathcal{C}_0(1)$ *as* $\kappa \to \infty$.

The second part of the convergence of (1.4) is the convergence of the large trees, for which we will rely on the following result about random trees with given degree sequences from [16].

Theorem 1.7 (Theorem 1 in [16])**.** *Let* $\{\mathbf{s}_\kappa, \kappa \geq 1\}$ *be a degree sequence such that* $\mathbf{n}_\kappa := n(\mathbf{s}_\kappa) \to \infty, \Delta_\kappa := \Delta(\mathbf{s}_\kappa) = o(\mathbf{n}_\kappa^{1/2})$. *Suppose that there exists a distribution* $\boldsymbol{p}$ *on* $\mathbb{N}$ *with mean 1 such that* $\boldsymbol{p}_\kappa$ *converges to* $\boldsymbol{p}$ *coordinatewise and such that* $\sigma(\boldsymbol{p}_\kappa) \to \sigma(\boldsymbol{p}) \in (0, \infty)$. *Let* $\mathbb{T}_\kappa$ *be the random plane tree under* $\mathbb{P}_{\mathbf{s}_\kappa}$, *the uniform measure on the set of plane trees with degree sequence* $\mathbf{s}_\kappa$. *Let* $\mathcal{T}_\kappa$ *denote the measured rooted metric space* $(\mathbb{T}_\kappa, \frac{\sigma(\boldsymbol{p}_\kappa)}{2\mathbf{n}_\kappa^{1/2}} d_{gr}, \emptyset_\kappa, \mu_\kappa)$ *where* μ_κ *denotes the uniform measure putting mass* $\frac{1}{\mathbf{n}_\kappa}$ *on each vertex of* $\mathbb{T}_\kappa$. *Then when* $\kappa \to \infty, \mathcal{T}_\kappa \xrightarrow{d} \mathcal{T}_{\mathbf{e}}$ *in the Gromov-Hausdorff-Prokhorov sense.*

***Remark* 1.3.** In fact Theorem 1 in [16] is only stated in the Gromov-Hausdorff sense, that is, $(\mathbb{T}_\kappa, \frac{\sigma(\mathbf{p}_\kappa)}{2\mathbf{n}_\kappa^{1/2}} d_{gr}) \xrightarrow{d} (T_{\mathbf{e}}, d_{\mathbf{e}}, \emptyset_{\mathbf{e}})$. But the conclusion can be strengthened to GHP convergence easily. For completeness, we include a proof of this fact in Section 2 of the supplemental materials.

The following proposition contains the additional ingredient required to prove Theorem 1.3.

***Proposition* 1.8.** Under the conditions of Theorem 1.3, for all $a > 0$, we have

$$\lim_{j\to\infty} \limsup_{\kappa\to\infty} \mathbf{P} \left(\sup_{l>j} \operatorname{diam}(\mathcal{T}_{\kappa,l}) > a \right) = 0.$$

The key results leading to Proposition 1.8 include a height bound for random tree with prescribed degree sequence and a variance bound for uniformly permuted child sequences. The height bound of uniformly random tree with prescribed degree sequence is given in the following theorem.

Theorem 1.9 (Theorem 1 in [3]). *Fix a degree sequence* $\mathbf{s} = (s^{(i)}, i \geq 0)$ *such that* $\sum_{i\geq 0} is^{(i)} = |\mathbf{s}| - 1$*, and let* $\mathbb{T}(\mathbf{s})$ *be a uniformly random plane tree with degree sequence* $\mathbf{s}$*. Then for all* $m \geq 1$ *we have*

$$\mathbf{P}\left(h(\mathbb{T}(\mathbf{s})) \geq m\right) \leq 7\exp\left(-m^2/608\sigma^2(\mathbf{s})1_{\mathbf{s}}^2\right)$$

where $1_{\mathbf{s}} = \frac{|\mathbf{s}|-2}{|\mathbf{s}|-1-s^{(1)}}$.

The following probability bound on variances of uniformly permuted integer sequences allows us to control the variance of degrees of trees in random forests, and thereby apply Theorem 1.9 to prove Proposition 1.8.

***Proposition* 1.10.** Fix $c = (c_1, \cdots, c_n) \in \mathbb{N}^n$ and let π be a uniformly random permutation of $\{1, \cdots, n\}$. Set $C_i = c_{\pi(i)}$ for $1 \leq i \leq n$, and let $S_j = \sum_{i\leq j} C_i^2$ for $1 \leq i \leq n$. Then for all $\lambda \geq 2$ and $1 \leq k \leq n$, with $\Delta = \max_{1\leq i\leq n} C_i = \max_{1\leq i\leq n} c_i$, and $\sigma^2(c) = \sum_{i\leq n} c_i^2 = S_n$, we have

$$\mathbf{P}\left(S_k \geq \lambda\frac{k}{n}S_n\right) \leq \exp\left(-\frac{3\sigma^2(c)}{16n}\cdot\frac{\lambda k}{\Delta^2}\right).$$

Now let us prove our main theorems with these key results.

Proof of Theorem 1.2 and Theorem 1.3. By Skorokhod's representation theorem, we may work in a probability space in which the convergence in Proposition 1.4 is almost sure. Hence Proposition 1.4 yields that for any fixed j, $\sup_{l\leq j} d_{GHP}(\mathcal{T}_{\kappa,l}, \mathcal{T}_{|\gamma_l|\mathbf{e}_l}) \xrightarrow{d} 0$. This establishes Theorem 1.2. Now to prove the convergence in $(\mathbb{L}_\infty, d_{GHP}^\infty)$, it suffices to prove that for any $a > 0$,

$$\lim_{j\to\infty}\limsup_{\kappa\to\infty}\mathbf{P}\left(\sup_{l>j}\left(\mathrm{diam}(\mathcal{T}_{\kappa,l}) + \mathrm{mass}(\mathcal{T}_{\kappa,l}) + \mathrm{diam}(\mathcal{T}_{\gamma_l}) + \mathrm{mass}(\mathcal{T}_{\gamma_l})\right) > a\right) = 0.$$

It suffices to separately prove

$$\lim_{j\to\infty}\limsup_{\kappa\to\infty}\mathbf{P}\left(\sup_{l>j}\mathrm{diam}(\mathcal{T}_{\kappa,l}) > a\right) = 0, \quad \lim_{j\to\infty}\limsup_{\kappa\to\infty}\mathbf{P}\left(\sup_{l>j}\mathrm{mass}(\mathcal{T}_{\kappa,l}) > a\right) = 0$$

$$\lim_{j\to\infty}\mathbf{P}\left(\sup_{l>j}\mathrm{diam}(\mathcal{T}_{\gamma_l}) > a\right) = 0, \quad \lim_{j\to\infty}\mathbf{P}\left(\sup_{l>j}\mathrm{mass}(\mathcal{T}_{\gamma_l}) > a\right) = 0.$$

For this purpose, we need to control the probability that small trees having either large diameter or large mass. Note that for a tree its diameter is bounded by twice of its height.

In fact the mass of tree is easy to control since for any $a > 0$ and any κ,

$$\begin{aligned}\mathbf{P}\left(\sup_{l>j} \operatorname{mass}(\mathcal{T}_{\kappa,l}) > a\right) &= \mathbf{P}\left(\sup_{l>j} \frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} > a\right) \\ &\leq \mathbf{P}\left(|\mathbb{T}_{\kappa,j}| > a\mathbf{n}_\kappa\right) = 0 \text{ for } j > 1/a\end{aligned}$$

For the diameter we resort to Proposition 1.8.

We also need to bound $\operatorname{diam}(\mathcal{T}_{\gamma_l})$ and $\operatorname{mass}(\mathcal{T}_{\gamma_l})$ for l large. Note that $\operatorname{mass}(\mathcal{T}_{\gamma_l}) = |\gamma_l|$ and for any a, let $j > 1/a$, then $\mathbf{P}\left(\sup_{l>j} |\gamma_l| > a\right) = 0$.

For $\operatorname{diam}(\mathcal{T}_{\gamma_l})$, $\operatorname{diam}(\mathcal{T}_{\gamma_l}) \leq 2h(\mathcal{T}_{\gamma_l}) = 2\max(\gamma_l)$. For $0 \leq s \leq 1$, let

$$R(s) = F_\lambda^{br}(s) - \inf_{s' \in (0,s)} F_\lambda^{br}(s')$$

and the excursion interval of γ_l be $[g_l, d_l]$. Then

$$\begin{aligned}\operatorname{diam}(\mathcal{T}_{\gamma_l}) &\leq 2 \sup_{t\in[g_l,d_l]} R(t) = 2\left(\sup_{t\in[g_l,d_l]} F_\lambda^{br}(t) - \inf_{t\in[g_l,d_l]} F_\lambda^{br}(t)\right) \\ &\leq 2\sup\left(|F_\lambda^{br}(t) - F_\lambda^{br}(s)| : |t-s| \leq d_l - g_l\right)\end{aligned}$$

and $d_l - g_l = |\gamma_l| \leq 1/l$. So for any $j \geq 1/\epsilon$,

$$\sup_{l>j} \operatorname{diam}(\mathcal{T}_{\gamma_l}) \leq 2\sup\left(|F_\lambda^{br}(t) - F_\lambda^{br}(s)| : |t-s| \leq \epsilon\right) \to 0 \text{ as } \epsilon \to 0$$

since F_λ^{br} is uniformly continuous. Hence we have the tail insignificance for diameter of $\mathcal{T}_{\gamma_l}$ and the claim is proved. □

To conclude this section, we sketch how our paper is organized. In Section 2 we investigate a special rotation mapping, which connects the collection of lattice bridges corresponding to certain degree sequence $\mathbf{s}$ and the set of first passage lattice bridges corresponding to $\mathbf{s}$. This will be the key starting point of our work using depth-first walk process to code the structure of random forests with given degree sequences. The combinatorial argument in this section will be also useful for our later work on transferring results such as Proposition 1.10 to something similar which is applicable to random forests. This section will be purely combinatorial and only deal with fixed degree sequences. In Section 3, we collect some concentration results using martingale methods. These probability bounds will be useful for checking that the assumptions in Theorem 1.7 are satisfied for large trees of $\mathcal{F}_\kappa^\downarrow$. The second part of this section proves the variance bound in Proposition 1.10. Again all results in this section is non-asymptotic and hence are presented with regards to a fixed degree sequence. In Section 4, we prove Theorem 1.6, the convergence of scaled exploration processes to some random process related to first passage bridge, using the rotation mapping in Section 2. We will then get Proposition 1.5 as a corollary from this weak convergence result. Finally, in Section 5 we finish the proof of Proposition 1.4 and Proposition 1.8 using results from Section 3 and Section 4.

2. An n−to−1 map transforming lattice bridge to first passage lattice bridge

Given a degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$, let $d(\mathbf{s}) \in \mathbb{Z}_{\geq 0}^{n(\mathbf{s})}$ be the vector whose entries are weakly increasing and with $s^{(i)}$ entries equal to i, for each $i \geq 0$. For example, if $\mathbf{s} = (3, 2, 0, 1, 0, \cdots)$ with $s^{(i)} = 0$ for $i \geq 4$, then $d(\mathbf{s}) = (0, 0, 0, 1, 1, 3)$. Let $\mathrm{D}(\mathbf{s})$ be the collection of all possible child sequences corresponding to degree sequence $\mathbf{s}$, i.e., all possible result as a permutation of $d(\mathbf{s})$.

A *lattice bridge* is a function $b : [0, k] \to \mathbb{R}$ with $b(0) = 0$ and $b(i) \in \mathbb{Z}$, $\forall i \in [k]$, which is piecewise linear between integers. Here k is an arbitrary positive integer. We let

$$\Lambda(\mathbf{s}) = \{b : [0, n(\mathbf{s})] \to \mathbb{R} : b \text{ is a lattice bridge and } \forall i \geq 0, |\{j \in \mathbb{N} : b(j+1)-b(j) = i-1\}| = s^{(i)}\}$$

and call $\Lambda(\mathbf{s})$ the set of *lattice bridges corresponding to* $\mathbf{s}$. Note that if $b \in \Lambda(\mathbf{s})$, then $b(n(\mathbf{s})) = -c(\mathbf{s})$. It is straightforward to see

$$|\Lambda(\mathbf{s})| = \binom{n}{(s^{(i)}, i \geq 0)} = \frac{n!}{\prod_{i\geq 0} s^{(i)}!}.$$

We then let

$$F(\mathbf{s}) = \{b \in \Lambda(\mathbf{s}) : \inf_{j \leq n(\mathbf{s})-1} b(j) > -c(\mathbf{s})\}$$

and call $F(\mathbf{s})$ the collection of *first passage lattice bridges corresponding to* $\mathbf{s}$.

For $s > 0$, let $\mathcal{C}_0(s) = \{x \in C([0, s], \mathbb{R}) : x(0) = 0\}$. For $u \in [0, s]$, let $\theta_{u,s} : \mathcal{C}_0(s) \to \mathcal{C}_0(s)$ denote the *cyclic shift at* u, that is,

$$(\theta_{u,s}(x))(t) = \begin{cases} x(t+u) - x(u), & \text{if } t+u \leq s; \\ x(t+u-s) + x(s) - x(u), & \text{if } t+u \geq s. \end{cases}$$

For $x \in \mathcal{C}_0(s)$ and $y \in \mathbb{R}^-$, let $t(y, x) := \inf\{t \in [0, s] : x(t) \leq y\}$ be the first time the graph of x drops below y. Sometimes we drop the argument x for convenience and simply write $t(y)$. If $y < \min_{u\in[0,s]} x(u)$ we set $t(y, x) = 0$ by convention, so $\theta_{t(y)}(x) = x$. In what follows, for $k \in \mathbb{N}$ we write $[k] - 1 = \{0, 1, \cdots, k-1\}$. And when the context is clear, we simply drop the subscript s and write θ_u for $\theta_{u,s}$.

***Lemma* 2.1.** For $b \in \Lambda(\mathbf{s})$, and for each $j \in [c(\mathbf{s})] - 1$, we have $\theta_{t(\min(b)+j)}(b) \in F(\mathbf{s})$.

Proof. Let $m \leq 0$ be the minimum of b. Fix an integer i such that $m \leq i \leq m + c(\mathbf{s}) - 1$ and $u < n(\mathbf{s})$. We shall prove that $\theta_{t(i)}(b)(u) > -c(\mathbf{s})$, which proves the lemma. If $0 \leq u \leq n(\mathbf{s}) - t(i)$, then $\theta_{t(i)}(b)(u) = b(t(i)+u) - b(t(i)) \geq m - i > -c(\mathbf{s})$. If $n(\mathbf{s}) - t(i) \leq u < n(\mathbf{s})$, then $\theta_{t(i)}(b)(u) = b(t(i)+u-n(\mathbf{s})) + b(n(\mathbf{s})) - b(t(i)) = b(t(i)+u-n(\mathbf{s})) - c(\mathbf{s}) - i$. Since $u < n(\mathbf{s})$, $t(i) + u - n(\mathbf{s}) < t(i)$ and we must have $b(t(i) + u - n(\mathbf{s})) > i$ by our definition of t. Therefore in this case we also have $\theta_{t(i)}(b)(u) > -c(\mathbf{s})$. □

Next, define a function $f : \Lambda(\mathbf{s}) \times ([c(\mathbf{s})] - 1) \to F(\mathbf{s})$ by $f(b, j) := \theta_{t(\min(b)+j)}(b)$.

***Lemma* 2.2.** f is an $n(\mathbf{s})-$to-1 map from $\Lambda(\mathbf{s}) \times ([c(\mathbf{s})] - 1)$ to $F(\mathbf{s})$.

Proof. For $l \in F(\mathbf{s})$, if size of preimage of l under f is strictly large than $n(\mathbf{s})$, then we must have $b_1, b_2 \in \Lambda(\mathbf{s}), j_1, j_2 \in [c(\mathbf{s})] - 1$ such that $f(b_1, j_1) = f(b_2, j_2) = l$ and $t(\min(b_1) + j_1) = t(\min(b_2) + j_2)$, since t can only take values in $[n(\mathbf{s})]$. By the definition of f we must then have $b_1 = b_2$ and hence $j_1 = j_2$. Therefore each element in $F(\mathbf{s})$ can have at most $n(\mathbf{s})$ preimages in $\Lambda(\mathbf{s}) \times ([c(\mathbf{s})] - 1)$. On the other hand, we have (see, e.g., [28], page 128)

$$|F(\mathbf{s})| = \frac{c(\mathbf{s})}{n(\mathbf{s})}\binom{n(\mathbf{s})}{(s^{(i)}, i \geq 0)} = \frac{c(\mathbf{s})}{n(\mathbf{s})}\frac{n(\mathbf{s})!}{\prod\limits_{i\geq 0} s^{(i)}!}.$$

Hence $n(\mathbf{s}) \times |F(\mathbf{s})| = c(\mathbf{s}) \times |\Lambda(\mathbf{s})| = |\Lambda(\mathbf{s}) \times ([c(\mathbf{s})] - 1)|$, so it must in fact hold that each $l \in F(\mathbf{s})$ has exactly $n(\mathbf{s})$ preimages. $\square$

For a sequence $\mathbf{c} = (c_1, \cdots, c_n) \in \mathbb{R}^n$, write $W_{\mathbf{c}}(j) = \sum\limits_{i=1}^{j}(c_i - 1)$ for $j \in [n]$. We let $W_{\mathbf{c}}(0) = 0$ and make $W_{\mathbf{c}}$ a continuous function on $[0, n]$ by linear interpolation. Note that the depth-first walk S_{F} is precisely $W_{\mathbf{c}}$ where $\mathbf{c} = (k_{\mathrm{F}}(u_1), \cdots, k_{\mathrm{F}}(u_{|\mathrm{F}|}))$. For $\mathbf{c} = (c_1, \cdots, c_n) \in \mathbb{R}^n$ and a permutation π of $[n]$, write $\pi(\mathbf{c}) = (c_{\pi(1)}, \cdots, c_{\pi(n)})$.

***Corollary* 2.3.** Let $\mathbf{s}$ be a degree sequence. Let π be a uniformly random permutation of $[n(\mathbf{s})]$ and let ν be independent of π and drawn uniformly at random from $[c(\mathbf{s})] - 1$. Then

$$f(W_{\pi(d(\mathbf{s}))}, \nu) \stackrel{d}{=} S_{\mathbb{F}(\mathbf{s})},$$

and both are uniformly random elements of $F(\mathbf{s})$.

Proof. By definition, $(W_{\pi(d(\mathbf{s}))}, \nu)$ is uniformly at random in $\Lambda(\mathbf{s}) \times ([c(\mathbf{s})] - 1)$. By Lemma 2.2, it follows that $f(W_{\pi(d(\mathbf{s}))}, \nu)$ is uniformly random in $F(\mathbf{s})$. On the other hand, the map sending plane forest F to its Lukasiewicz path S_{F} restricts to an invertible map from $\mathrm{F}(\mathbf{s})$ to $F(\mathbf{s})$. Thus, $S_{\mathbb{F}(\mathbf{s})}$ is also uniformly distributed in $F(\mathbf{s})$. $\square$

First-passage bridges are naturally connected to plane forests. In a similar way, general lattice bridges are naturally connected to *marked* plane forests. This interpretation will be more convenient for some later proofs (Propositions 3.5, 3.7 and 3.8).

A *marked forest* is a pair (F, v) where F is a plane forest and $v \in v(F)$. We call v the *mark* of (F, v). Recall that $\mathrm{F}(\mathbf{s})$ denotes the collection of all plane forests with degree sequence $\mathbf{s}$. Let $\mathrm{MF}(\mathbf{s})$ be the collection of all marked forests with degree sequence $\mathbf{s}$ and for $1 \leq i \leq c(\mathbf{s})$, let $\mathrm{MF}^i(\mathbf{s})$ be the collection of marked forests $(F, v) \in \mathrm{MF}(\mathbf{s})$ such that the mark v lies within the $i-$th tree of F. We define a map $g : \mathrm{MF}(\mathbf{s}) \to \mathrm{D}(\mathbf{s})$ which lists the degrees of vertices of a marked forest starting from the mark in DFS order. Formally, for $(F, v) \in \mathrm{MF}(\mathbf{s})$, if the DFS ordering of $v(F)$ is $v_1, \cdots, v_{n(\mathbf{s})}$ and

$v = v_i$, then $g((F,v)) = (k_F(v_i), \cdots, k_F(v_{n(\mathbf{s})}), k_F(v_1), \cdots, k_F(v_{i-1}))$. Next define a map $h : \mathrm{MF}(\mathbf{s}) \to \mathrm{F}(\mathbf{s})$ by $h((F,v)) = F$. Then we have the following easy fact.

***Lemma* 2.4.** g is a $c(\mathbf{s})$−to−1 surjective map and for each $1 \le i \le c(\mathbf{s})$, $g^i := g|_{\mathrm{MF}^i(\mathbf{s})}$ is a bijection between $\mathrm{MF}^i(\mathbf{s})$ and $\mathrm{D}(\mathbf{s})$. Also, h is a $n(\mathbf{s})$−to−1 surjective map.

Proof. For $d \in \mathrm{D}(\mathbf{s}), |g^{-1}(\{d\}) \cap \mathrm{MF}^i(\mathbf{s})| = 1$ for all $1 \le i \le c(\mathbf{s})$. In fact, the element of each $g^{-1}(\{d\}) \cap \mathrm{MF}^i(\mathbf{s})$ can be obtained by cyclically permuting the tree components of the element of $g^{-1}(\{d\}) \cap \mathrm{MF}^1(\mathbf{s})$. This shows that g^i is a bijection. The other two claims are straightforward. □

The map g being surjective immediately gives the following result.

***Corollary* 2.5.** Let $\mathbb{MF}(\mathbf{s})$ be a uniformly random element of $\mathrm{MF}(\mathbf{s})$, then $g(\mathbb{MF}(\mathbf{s}))$ is a uniformly random element of $\mathrm{D}(\mathbf{s})$.

3. Concentration results

In the first part of this section, we deal with a martingale concerning the proportion of a fixed degree of uniformly permuted degree sequence. This will be useful for proving Proposition 1.4 in Section 5 where we need to first show that the degree proportions in each large trees of $\mathcal{F}_\kappa^\downarrow$ are more or less in line with the degree proportion of the given degree sequences. The second part of this section deals with the variance bound of uniformly permuted child sequences, which leads to a key technical proposition on the height of tree components of $\mathbb{F}(\mathbf{s})$. For both subsections we will use concentration results from [25].

Let $\mathbf{s} = (s^{(i)}, i \ge 0)$ with $|\mathbf{s}| = n$ be a fixed degree sequence and let $\mathbf{C} = (C_1, \cdots, C_n)$ denote the uniformly permuted child sequence $\pi(d(\mathbf{s}))$ (recall the notation from Section 2), where π is a uniform random permutation of $[n]$. For each $i \ge 0$, let $q^{(i)} = s^{(i)}/n$ be the degree proportion of degree i of $\mathbf{s}$.

3.1. Martingales of degree proportions of uniformly permuted degree sequence

In this subsection, we introduce some martingales concerning proportions of particular degree appeared at each step in a uniformly permuted degree sequence and use them and martingale concentration inequality from [25] as tools to prove Lemma 3.4 and Proposition 3.5, which are useful for eventually proving that the empirical degree distributions of large trees of $\mathbb{F}_\kappa$ behave well (Proposition 5.1). We first recall the following martingale bound in [25]. Let $\{X_i\}_{i=0}^n$ be a bounded martingale adapted to a filtration $\{\mathcal{F}_i\}_{i=0}^n$. Let $V = \sum_{i=0}^{n-1} var\{X_{i+1} \mid \mathcal{F}_i\}$, where

$$var\{X_{i+1} \mid \mathcal{F}_i\} := \mathbf{E}\left[(X_{i+1} - X_i)^2 \mid \mathcal{F}_i\right] = \mathbf{E}\left[X_{i+1}^2 \mid \mathcal{F}_i\right] - X_i^2.$$

Let

$$v = \text{ess sup } V, \text{ and } b = \max_{0\le i\le n-1} \text{ess sup}(X_{i+1} - X_i \mid \mathcal{F}_i).$$

Then we have the following bound.

Theorem 3.1 ([25], Theorem 3.15)**.** *For any $t \ge 0$,*

$$\mathbf{P}\left(\max_{0\le i\le n} X_i \ge t\right) \le \exp\left(-\frac{t^2}{2v(1+bt/(3v))}\right).$$

For fixed i, for $0 \le j \le n-1$, let $Y_j^{(i)} = |\{1 \le l \le j : C_l = i\}|$ and let $X_j^{(i)} = s^{(i)} - Y_j^{(i)}$. Note that for $j > 0$

$$X_j^{(i)} = \begin{cases} X_{j-1}^{(i)} - 1, & \text{if } C_j = i; \\ X_{j-1}^{(i)}, & \text{otherwise.} \end{cases}$$

Let $\mathcal{F}_j$ be the $\sigma-$field generated by $C_1, \cdots, C_j$.

***Lemma* 3.2.** Let $M_j^{(i)} := \frac{X_j^{(i)}}{n-j} - q^{(i)}$, then

(a) $M_j^{(i)}$ is an $\mathcal{F}_j-$martingale;

(b) The *predictable quadratic variation* of $M_{j+1}^{(i)}$ satisfies
$var\{M_{j+1}^{(i)} \mid \mathcal{F}_j\} := \mathbf{E}\left[{M_{j+1}^{(i)}}^2 \mid \mathcal{F}_j\right] - {M_j^{(i)}}^2 \le \frac{1}{4}\frac{1}{(n-(j+1))^2}$.

The proof of Lemma 3.2 is straightforward and we include the proof in Section 3 of the supplemental materials. Now we can apply Theorem 3.1.

***Proposition* 3.3.** For any $t > 0$ and $0 < s < n$, we have

$$\mathbf{P}\left(\max_{0\le j\le n-s} |q^{(i)} - \frac{X_j^{(i)}}{n-j}| \ge t\right) \le 2\exp\left(-\frac{3st^2}{3+2t}\right). \tag{3.1}$$

Proof. Fix $s < n$, and consider the martingale $\{M_j^{(i)}\}_{j=0}^{n-s}$. By Lemma 3.2(b), we know that

$$V = \sum_{j=1}^{n-s} var\{M_j^{(i)} \mid \mathcal{F}_{j-1}\} \le \frac{1}{4}\sum_{j=0}^{n-s-1} \frac{1}{(n-(j+1))^2} \le \frac{1}{4}\int_{s-1}^{n-1} \frac{1}{x^2}dx \le \frac{1}{2s}.$$

Hence $v = \text{ess sup } V \le \frac{1}{2s}$. Also, for $j \le n-s-1$, if $X_{j+1}^{(i)} = X_j^{(i)}$, then

$$|M_{j+1}^{(i)} - M_j^{(i)}| = \frac{X_j^{(i)}}{(n-j)(n-j-1)} \le \frac{1}{s},$$

and if $X_{j+1}^{(i)} = X_j^{(i)} - 1$, then

$$|M_{j+1}^{(i)} - M_j^{(i)}| = |\frac{X_j^{(i)} - 1}{n-(j+1)} - \frac{X_j^{(i)}}{n-j}| = |\frac{X_j^{(i)}}{(n-(j+1))(n-j)} - \frac{1}{n-(j+1)}| \le \frac{1}{s}.$$

Applying Theorem 3.1 to both $\{M_j^{(i)}\}_{j=0}^{n-s}$ and $\{-M_j^{(i)}\}_{j=0}^{n-s}$ gives

$$\mathbf{P}\left(\max_{0\le j\le n-s}\left|q^{(i)} - \frac{X_j^{(i)}}{n-j}\right| \ge t\right) \le 2\exp\left(-\frac{t^2}{\frac{1}{s}+\frac{2t}{3s}}\right),$$

as claimed. □

Now we give a probability bound of proportion of certain degree i deviates from $q^{(i)}$ by an error of at least ϵ.

***Lemma* 3.4.** For fixed $i \in \mathbb{N}$ and $\epsilon > 0$, let $B^{\epsilon,i} = \{\exists x \ge \log^3 n : |Y_x^{(i)} - q^{(i)}x| \ge \epsilon x\}$. Then for any n large enough such that $\frac{\sqrt{5}}{\log n} < \epsilon < 1$, $\mathbf{P}\left(B^{\epsilon,i}\right) \le \frac{2}{n^3}$.

Proof. By symmetry, the event $\{\exists j \ge \log^3 n : |Y_j^{(i)} - q^{(i)}j| \ge \epsilon j\}$ has the same distribution as the event $\{\exists l \le n - \log^3 n : |X_l^{(i)} - q^{(i)}(n-l)| \ge \epsilon(n-l)\}$. Hence we can write

$$\mathbf{P}\left(B^{\epsilon,i}\right) = \mathbf{P}\left(\max_{0\le l\le n-\log^3 n}|q^{(i)} - \frac{X_l^{(i)}}{n-l}| \ge \epsilon\right).$$

Taking $s = \log^3 n, t = \epsilon$ in (3.1), the result follows. □

Now we consider how degrees distribute among the tree components of the random forest $\mathbb{F}(\mathbf{s})$. Write $\mathbb{F}(\mathbf{s})^{\downarrow} = (\mathbb{T}_l,\ l \ge 1)$. Let $\mathbf{s}_l = (s_l^{(i)}, i \ge 0)$ denote the (empirical) degree sequence of the $l-$th largest tree $\mathbb{T}_l$, and let $\mathbf{n}_l = n(\mathbf{s}_l)$. Recall that $q^{(i)} = s^{(i)}/n$ and let $q_l^{(i)} = s_l^{(i)}/\mathbf{n}_l$ be the empirical proportion of degree i vertices of $\mathbb{T}_l$; if $\mathbb{F}(\mathbf{s})$ has fewer than l trees then $q_l^{(i)} = 0$. Note that $q^{(i)}$ is deterministic while $q_l^{(i)}$ is random.

***Proposition* 3.5.** For fixed $\epsilon > 0$ and i, l, let $B_l^{\epsilon,i} = \{|q_l^{(i)} - q^{(i)}| > \epsilon\}$. Then for fixed $\epsilon > 0, i \in \mathbb{N}$, we have

$$\mathbf{P}\left(\bigcup_{l:\ |\mathbb{T}_l| > n^{1/4}} B_l^{\epsilon,i}\right) \le n\mathbf{P}\left(B^{\epsilon,i}\right). \tag{3.2}$$

Proof. Let V be a uniformly random vertex of $\mathbb{F}(\mathbf{s})$, then $(\mathbb{F}(\mathbf{s}), V)$ is uniformly distributed in $\mathrm{MF}(\mathbf{s})$. List the nodes of $\mathbb{F}(\mathbf{s})$ in cyclic lexicographic order as $V = V_1, V_2, \cdots, V_n$, and for $i \le n$ let C_i be the degree of V_i. By Corollary 2.5, the sequence $(C_1, \cdots, C_n) = g(\mathbb{F}(\mathbf{s}), V)$ is uniformly distributed in $\mathrm{D}(\mathbf{s})$; in other words, it is distributed as a uniformly random permutation of $d(\mathbf{s})$. For any $1 \leq j \leq n$, let $\tilde{B}_j^{\epsilon,i}$ be the event that there exists $m > n^{1/4}$ such that

$$|\frac{\#\{1 \leq t \leq m : C_{j+t \pmod n} = i\}}{m} - q^{(i)}| > \epsilon.$$

Since $(C_1, \cdots, C_n)$ is uniformly distributed in $\mathrm{D}(\mathbf{s})$, it is immediate that $\mathbf{P}\left(\tilde{B}_1^{\epsilon,i}\right) = \cdots = \mathbf{P}\left(\tilde{B}_n^{\epsilon,i}\right)$. Suppose a tree $T \in \mathbb{F}(\mathbf{s})$ with $|T| > n^{1/4}$ has that

$$|\frac{\#\{u : k_T(u) = i\}}{|T|} - q^{(i)}| > \epsilon.$$

If V is not a node of T, then there exists $m > n^{1/4}, 0 < j \leq n - m$ such that

$$V(T) = \{V_{j+1}, \cdots, V_{j+m}\}, \ |\frac{\#\{1 \leq t \leq m : C_{j+t} = i\}}{m} - q^{(i)}| > \epsilon.$$

If V is a node of T, then there exists $m > n^{1/4}, j > n - m$ such that

$$V(T) = \{V_{j+1}, \cdots, V_n, V_1, \cdots, V_{j+m-n}\}, \ |\frac{\#\{t \geq j+1 \text{ or } t \leq j+m-n : C_t = i\}}{m} - q^{(i)}| > \epsilon.$$

In either case we must have $\tilde{B}_j^{\epsilon,i}$ true for some $1 \leq j \leq n$. Therefore

$$\mathbf{P}\left(\bigcup_{l:\ |\mathbb{T}_l| > n^{1/4}} B_l^{\epsilon,i}\right) \leq n\mathbf{P}\left(\tilde{B}_1^{\epsilon,i}\right) \leq n\mathbf{P}\left(B^{\epsilon,i}\right),$$

which gives the claim. □

3.2. Probability bound of trees of random forest having abnormally large height

In this subsection, we prove tail bounds on the heights of trees in $\mathbb{F}(\mathbf{s})$, by first proving tail bounds on the sums of squares of the child sequences. This will be used in proving Proposition 1.8 in Section 5. To be more specific, let $c = (c_1, c_2, \cdots, c_n) \in \mathrm{D}(\mathbf{s})$ be a child sequence with $\sigma^2(\mathbf{s}) := \sum_{i=1}^{n} c_i^2 = \sum_i i^2 s^{(i)}$ and write $M := \sigma^2(\mathbf{s})/n$ and $\Delta = \Delta(\mathbf{s}) := \max_i c_i$. Recall that $C_1, C_2, \cdots, C_n$ are the uniformly permuted child sequence and let $S_j := \sum_{i \leq j} C_i^2$. We will use the following lemma.

***Lemma* 3.6.** Let $X_1, \cdots, X_k$ be samples from finite population $x_1, \cdots, x_n$, without replacement, with $X_1 - \mathbf{E}[X_1] \leq b$. Let $S_k = \sum_{i=1}^{k} X_i, V = \sum_{i=1}^{k} \mathrm{Var}(X_i)$ and $\mu_k = \mathbf{E}[S_k]$. Then for any $t \geq 0$, with $\epsilon = bt/V$, we have

$$\mathbf{P}\left(S_k - \mu_k \geq t\right) \leq \exp\left(-\frac{t^2}{2V + 2bt/3}\right). \tag{3.3}$$

This lemma corresponds to Theorem 2.7 in [25] in the case of sampling with replacement. The proof of Theorem 2.7 in [25], proceeds by bounding Laplace transformations, which are convex functions. Hence by applying Proposition 20.6 in [5], the result still holds when sampling without replacement.

Now we get our probability bound on the deviations of $(S_k, k \le n)$.

Proof of Proposition 1.10. We apply (3.3); we have $\mu_k = \mathbf{E}[S_k] = \frac{k}{n} S_n, b = \Delta^2$,

$$V = \sum_{i=1}^{k} \mathrm{Var}(C_i^2) \le k\mathbf{E}\left[C_1^4\right] = \frac{k}{n}\sum_{i=1}^{n} c_i^4 \le \frac{k}{n}\Delta^2\sigma^2(c) = k\Delta^2 M,$$

where $M = \sigma^2(c)/n$. For $\lambda > 1$, taking $t = (\lambda-1)\frac{k}{n}\sigma^2(c)$, we obtain

$$\begin{aligned}\mathbf{P}\left(S_k \ge \lambda\frac{k}{n}S_n\right) &= \mathbf{P}\left(S_k - \mu_k \ge (\lambda-1)kM\right) \\ &\le \exp\left(-\frac{((\lambda-1)kM)^2}{2k\Delta^2 M + \frac{2}{3}\Delta^2(\lambda-1)kM}\right)\end{aligned}$$

Using the assumption $\lambda \ge 2$ twice, we have

$$\begin{aligned}\mathbf{P}\left(S_k \ge \lambda\frac{k}{n}S_n\right) &\le \exp\left(-\frac{((\lambda-1)kM)^2}{\frac{8}{3}(\lambda-1)\Delta^2 kM}\right) \\ &= \exp\left(-\frac{3(\lambda-1)kM}{8\Delta^2}\right) \le \exp\left(-\frac{3M}{16}\cdot\frac{\lambda k}{\Delta^2}\right) = \exp\left(-\frac{3\sigma^2(c)}{16n}\cdot\frac{\lambda k}{\Delta^2}\right),\end{aligned}$$

which finishes the proof. □

Using results from Section 2, we now have the following estimate on variance of tree components of $\mathbb{F}(\mathbf{s})$. For a tree T, we let $\sigma^2(T) = \sum_{u\in T} k_T(u)^2$.

Proposition 3.7. Let $\mathbf{s} = (s^{(i)}, i \ge 0)$ be a degree sequence with $|\mathbf{s}| = n$ and $M = \sigma^2(\mathbf{s})/n$. Then for $\lambda \ge 4, \Delta^2(\mathbf{s})/n < \alpha \le 1$,

$$\mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| \le \alpha n, \sigma^2(T) \ge \lambda\alpha\sigma^2(\mathbf{s})\right) \le \frac{2}{\alpha}\exp(-\frac{3M}{16}\lambda). \tag{3.4}$$

Proof. Let V be a uniformly random vertex of $\mathbb{F}(\mathbf{s})$, then $(\mathbb{F}(\mathbf{s}), V)$ is uniformly distributed in $\mathrm{MF}(\mathbf{s})$. List the nodes of $\mathbb{F}(\mathbf{s})$ in cyclic lexicographic order as $V = V_1, V_2, \cdots, V_n$, and for $i \le n$ let C_i be the degree of V_i. By Corollary 2.5, the sequence $(C_1, \cdots, C_n) = g(\mathbb{F}(\mathbf{s}), V)$ is uniformly distributed in $\mathrm{D}(\mathbf{s})$; in other words, it is distributed as a uniformly random permutation of $d(\mathbf{s})$. In what follows we omit some floor notations for readability. For $0 \le j \le \lfloor\frac{1}{\alpha}\rfloor$, let B_j be the event that

$$\sum_{i=j\alpha n+1}^{(j+2)\alpha n} C^2_{i \pmod n} \ge \lambda\alpha\sigma^2(\mathbf{s}).$$

Since $C_1, \cdots, C_n$ is distributed as a uniformly random permutation of $d(\mathbf{s})$, we clearly have

$$\mathbf{P}(B_0) = \mathbf{P}(B_1) = \cdots = \mathbf{P}\left(B_{\lfloor \frac{1}{\alpha} \rfloor}\right).$$

Suppose that a given tree $T \in \mathbb{F}(\mathbf{s})$ has $|T| \le \alpha n$ and $\sigma^2(T) \ge \lambda\alpha\sigma^2(\mathbf{s})$. Then there exist $0 \le l < n$ and $m \le \alpha n$ such that $V(T) = \{V_{l+t \pmod n} : 1 \le t \le m\}$. Hence there exists $0 \le j \le \lfloor \frac{1}{\alpha} \rfloor$ such that $V(T) \subset \{V_{i \pmod n},\ j\alpha n + 1 \le i \le (j+2)\alpha n\}$. This implies that

$$\sum_{i=j\alpha n+1}^{(j+2)\alpha n} C^2_{i \pmod n} \ge \sigma^2(T) \ge \lambda\alpha\sigma^2(\mathbf{s}),$$

i.e. B_j is true. Hence the probability in question is at most

$$\begin{aligned}(1 + \lfloor \frac{1}{\alpha} \rfloor)\mathbf{P}(B_0) \le \frac{2}{\alpha}\mathbf{P}\left(S_{\lfloor 2\alpha n \rfloor} \ge \lambda\alpha\sigma^2(\mathbf{s})\right) \quad &\le \quad \frac{2}{\alpha}\mathbf{P}\left(S_{\lfloor 2\alpha n \rfloor} \ge \frac{\lambda}{2} \cdot \frac{\lfloor 2\alpha n \rfloor}{n}\sigma^2(\mathbf{s})\right) \\ &\le \quad \frac{2}{\alpha}\exp\left(-\frac{3M}{16}\lambda\right),\end{aligned}$$

where we take $k = \lfloor 2\alpha n \rfloor$ in Proposition 1.10 and use $\alpha > \Delta^2(\mathbf{s})/n$ at the last step. $\square$

Now we finish this section by proving a key proposition on probability bound of $\mathbb{F}(\mathbf{s})$ containing trees with unusually large height.

Proposition 3.8. $\forall\, \epsilon, \rho \in (0, 1), \exists n_0 = n_0(\epsilon) \in \mathbb{N}$ and $\beta_0 > 0$ such that the following is true. Let $\mathbf{s}$ be any degree sequence with $|\mathbf{s}| = n \ge n_0$. Suppose that $\Delta(\mathbf{s}) \le n^{\frac{1-\epsilon}{2}}, s^{(1)} \le (1-\epsilon)|\mathbf{s}|$ and $\epsilon \le \sigma^2(\mathbf{s})/n \le 1/\epsilon$, then for any $0 < \beta < \beta_0$,

$$\mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| < \beta n, h(T) > \beta^{1/8} n^{1/2}\right) \le \rho.$$

Proof. Fix $\beta > 0$ small, let $\delta = \beta^{1/8}$, and consider the following four events.

- E_1 is the event that there exists a tree T (of $\mathbb{F}(\mathbf{s})$) with $\Delta^2(\mathbf{s}) < |T| < \beta n$ and $\sigma^2(T) > (\frac{|T|}{n})^{1/2}\sigma^2(\mathbf{s})$.
- E_2 is the event that there exists a tree T with $|T| \le n^{1-\epsilon}$ and $\sigma^2(T) > n^{1-\frac{\epsilon}{2}}$.
- E_3 is the event that there exists a tree T with $\Delta^2(\mathbf{s}) < |T| < \beta n$ and $\sigma^2(T) \le (\frac{|T|}{n})^{1/2}\sigma^2(\mathbf{s})$ such that $h(T) > \delta n^{1/2}$.
- E_4 is the event that there exists a tree T with $|T| \le n^{1-\epsilon}$ and $\sigma^2(T) \le n^{1-\frac{\epsilon}{2}}$ such that $h(T) > \delta n^{1/2}$.

If there is $T \in \mathbb{F}(\mathbf{s})$ with $|T| < \beta n$, and $h(T) > \delta n^{1/2}$, then one of E_1, E_2, E_3 or E_4 must occur, so it suffices to bound $\mathbf{P}(E_1) + \mathbf{P}(E_2) + \mathbf{P}(E_3) + \mathbf{P}(E_4)$. For E_1, we further decompose the interval $[\Delta^2(\mathbf{s}), \beta n]$ dyadically. In the next sum, we bound the $k-$th summand by taking $\alpha = \frac{\beta}{2^k}, \lambda = \frac{2^{\frac{k-1}{2}}}{\beta^{1/2}} \geq 4$ in Proposition 3.7.

$$\begin{aligned}\mathbf{P}(E_1) &\leq \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| \in \left[\frac{\beta n}{2^{k+1}}, \frac{\beta n}{2^k}\right], \sigma^2(T) > \left(\frac{\beta}{2^{k+1}}\right)^{1/2} \sigma^2(\mathbf{s})\right) \\ &\leq \sum_{k\geq 0} \frac{2^{k+1}}{\beta} \exp\left(-\frac{3\sigma^2(\mathbf{s})}{16n} \frac{2^{\frac{k-1}{2}}}{\beta^{1/2}}\right) = O\left(\frac{1}{\beta}\exp(-\frac{\epsilon}{\beta^{1/2}})\right) \end{aligned} \tag{3.5}$$

where we use that $\sigma^2(\mathbf{s})/n \geq \epsilon$ in the final step.

Next, note that $\mathbf{P}(E_2) \leq \sum_{j=1}^{n^{1-\epsilon}} \mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| = j, \sigma^2(T) > n^{1-\epsilon/2}\right)$. For any fixed j, using Corollary 2.5, with similar argument as in proof of Proposition 3.7, we have

$$\mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| = j, \sigma^2(T) > n^{1-\epsilon/2}\right) \leq n\mathbf{P}\left(S_j \geq n^{1-\epsilon/2}\right).$$

For any $j \leq n^{1-\epsilon}$, use Proposition 1.10 with $\lambda \frac{j}{n} \sigma^2(\mathbf{s}) = n^{1-\epsilon/2}$ and $\Delta(\mathbf{s}) \leq n^{\frac{1-\epsilon}{2}}$, we have

$$\mathbf{P}\left(S_j \geq n^{1-\epsilon/2}\right) \leq \exp\left(-\frac{3\sigma^2(\mathbf{s})}{16n} \cdot \frac{\lambda j}{\Delta(\mathbf{s})^2}\right) \leq \exp\left(-\frac{3}{16} n^{\epsilon/2}\right).$$

These give that

$$\mathbf{P}(E_2) \leq n^{2-\epsilon} \exp\left(-\frac{3}{16} n^{\epsilon/2}\right). \tag{3.6}$$

We bound $\mathbf{P}(E_3)$ as follows. For $k \geq 0$, let $E_{3,k}$ be the event that there exists $T \in \mathbb{F}(\mathbf{s})$ with $\frac{\beta n}{2^{k+1}} \leq |T| \leq \frac{\beta n}{2^k}$ and $\sigma^2(T) \leq (\frac{|T|}{n})^{1/2}\sigma^2(\mathbf{s})$ such that height $h(T) > \delta n^{1/2}$. Also, let B be the event that there exists $T \in \mathbb{F}(\mathbf{s})$ with $|T| \geq n^{1/4}$ such that

$$\left| \frac{s^{(1)}(T)}{|T|} - \frac{s^{(1)}}{n} \right| \geq \epsilon/2.$$

For n large enough, we have $\frac{\sqrt{5}}{\log n} < \epsilon/2 < 1$. Hence it is immediate from Lemma 3.4 and Proposition 3.5 that $\mathbf{P}(B) \leq \frac{2}{n^2}$ for n large. Also, for n large, if $h(T) \geq \delta n^{1/2}$ then $|T| \geq h(T) \geq n^{1/4}$, so

$$\mathbf{P}(E_3) \leq \mathbf{P}(B) + \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}(E_{3,k} \cap B^c) \leq \frac{2}{n^2} + \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}(E_{3,k} \cap B^c). \tag{3.7}$$

Let m be the number of trees $T \in \mathbb{F}(\mathbf{s})$ with $\frac{\beta n}{2^{k+1}} \leq |T| \leq \frac{\beta n}{2^k}$ and $\sigma^2(T) \leq (\frac{|T|}{n})^{1/2}\sigma^2(\mathbf{s})$, and list the *random* degree sequences of these trees as $\mathbf{R}_1, \cdots, \mathbf{R}_m$. Then for any degree sequences $\mathbf{r}_1, \cdots, \mathbf{r}_m$,

$$\begin{aligned}\mathbf{P}(E_{3,k} \cap B^c \cap \{(\mathbf{R}_1, \cdots, \mathbf{R}_m) = (\mathbf{r}_1, \cdots, \mathbf{r}_m)\}) &= \mathbf{P}(B^c \cap \{(\mathbf{R}_1, \cdots, \mathbf{R}_m) = (\mathbf{r}_1, \cdots, \mathbf{r}_m)\}) \\ &\quad \cdot \mathbf{P}(E_{3,k} \mid B^c \cap \{(\mathbf{R}_1, \cdots, \mathbf{R}_m) = (\mathbf{r}_1, \cdots, \mathbf{r}_m)\}).\end{aligned}$$

Moreover

$$\mathbf{P}\left(E_{3,k} \mid B^c \cap \{(\mathbf{R}_1, \cdots, \mathbf{R}_m) = (\mathbf{r}_1, \cdots, \mathbf{r}_m)\}\right) = \mathbf{P}\left(\exists i \leq m, h(\mathbb{T}(\mathbf{r}_i)) \geq \delta n^{1/2}\right),$$

where $\mathbb{T}(\mathbf{r}_i)$ is a uniformly random plane tree with degree sequence $\mathbf{r}_i$. It follows from these identities that

$$\mathbf{P}\left(E_{3,k} \cap B^c\right) \leq \sup \mathbf{P}\left(\exists i \leq m, h(\mathbb{T}(\mathbf{r}_i)) \geq \delta n^{1/2}\right), \tag{3.8}$$

where the supremum is over vectors $(\mathbf{r}_1, \cdots, \mathbf{r}_m)$ of degree sequences such that

$$\mathbf{P}\left(E_{3,k} \cap B^c \cap \{(\mathbf{R}_1, \cdots, \mathbf{R}_m) = (\mathbf{r}_1, \cdots, \mathbf{r}_m)\}\right) > 0.$$

The last condition implies that, for all $i \leq m$,

$$\left| \frac{\mathbf{r}_i^{(1)}}{n(\mathbf{r}_i)} - \frac{s^{(1)}}{n} \right| < \epsilon/2.$$

Recall that $s^{(1)} \leq (1-\epsilon)|\mathbf{s}| = (1-\epsilon)n$, so $\frac{\mathbf{r}_i^{(1)}}{n(\mathbf{r}_i)} < 1 - \epsilon/2$ and that

$$\sigma^2(\mathbf{r}_i) \leq (\frac{n(\mathbf{r}_i)}{n})^{1/2}\sigma^2(\mathbf{s}) \leq (\frac{\beta}{2^k})^{1/2}\sigma^2(\mathbf{s}).$$

Finally we must have $n(\mathbf{r}_i) \geq \frac{\beta}{2^{k+1}} n$ for all $i \leq m$, so $m \leq \frac{2^{k+1}}{\beta}$. Now recall Theorem 1.9, which states that for a degree sequence $\mathbf{r} = (r^{(i)}, i \geq 0)$ and for all $h \geq 1$,

$$\mathbf{P}\left(h(\mathbb{T}(\mathbf{r})) \geq h\right) \leq 7 \exp\left(-h^2/608\sigma^2(\mathbf{r})1_{\mathbf{r}}^2\right)$$

where $1_{\mathbf{r}} = \frac{|\mathbf{r}|-2}{|\mathbf{r}|-1-r^{(1)}}$; note that this is at most $4/\epsilon$ for all degree sequences under consideration (for n large enough such that $n^{1/4} \geq 4/\epsilon$).

Using a union bound in (3.8), and then applying Theorem 1.9, we obtain that

$$\mathbf{P}\left(E_{3,k} \cap B^c\right) \leq \frac{2^{k+1}}{\beta} \cdot 7 \exp\left(-\frac{\epsilon^3\delta^2}{9728}(\frac{2^k}{\beta})^{1/2}\right)$$

where we use the assumption $\sigma^2(\mathbf{s})/n \leq 1/\epsilon$. And summing over k in (3.7) yields that

$$\mathbf{P}\left(E_3\right) \leq \sum_{k \geq 0} \frac{2^{k+1}}{\beta} \cdot 7 \exp\left(-\frac{\epsilon^3\delta^2}{9728}(\frac{2^k}{\beta})^{1/2}\right) + \frac{2}{n^2} \leq C_5 \frac{1}{\beta} \exp\left(-\frac{C_6}{\beta^{1/4}}\right) + \frac{2}{n^2} \tag{3.9}$$

if we take $\delta = \beta^{1/8}$, where $C_5 > 0$ is some universal constant and $C_6 > 0$ is some constant depending on ϵ.

For $\mathbf{P}(E_4)$, similar to the previous treatment of $\mathbf{P}(E_3)$, for n large, we have

$$\mathbf{P}\left(E_4\right) \leq \frac{2}{n^2} + \mathbf{P}\left(E_4 \cap B^c\right).$$

There are at most n trees in total, so a reprise of the conditioning argument used to bound $\mathbf{P}(E_3)$ gives

$$\mathbf{P}(E_4 \cap B^c) \leq n \sup \mathbf{P}\left(h(\mathbb{T}(\mathbf{r})) \geq \delta n^{1/2}\right),$$

where the supremum is over degree sequences $\mathbf{r}$ with $n(\mathbf{r}) \leq n^{1-\epsilon}$, with $\sigma^2(\mathbf{r}) \leq n^{1-\epsilon/2}$, and with $r^{(1)} \leq (1-\epsilon/2)n(\mathbf{r})$. By Theorem 1.9, we obtain that

$$\begin{aligned}\mathbf{P}(E_4) \leq \frac{2}{n^2} + 7n\exp\left(-\frac{\delta^2 n}{608\sigma^2(\mathbf{r})1_{\mathbf{r}}^2}\right) &\leq \frac{2}{n^2} + 7n\exp\left(-\frac{\delta^2 n}{608 n^{1-\frac{\epsilon}{2}}\frac{16}{\epsilon^2}}\right) \\ &= \frac{2}{n^2} + 7n\exp\left(-\frac{\epsilon^2}{9728}n^{\epsilon/2}\beta^{1/4}\right) \quad (3.10)\end{aligned}$$

recall that we take $\delta = \beta^{1/8}$. Of the bounds on $\mathbf{P}(E_i), 1 \leq i \leq 4$ in (3.5), (3.6), (3.9) and (3.10), the largest is for $\mathbf{P}(E_3)$ (provided n is large enough). Hence by taking $\beta > 0$ small enough, we can make the bound less than any prescribed number $\rho > 0$, which yields the result. $\square$

4. Convergence of the Lukasiewicz walk of forest to first passage bridge

In this section, we aim to prove Theorem 1.6 and conclude Proposition 1.5 as a corollary of Theorem 1.6. Throughout the section, we fix a sequence $(\mathbf{s}_\kappa, \kappa \in \mathbb{N})$ of degree sequences, and let $\mathbf{n}_\kappa, \mathbf{p}_\kappa$ be as in Section 1 and the function d be as in Section 2. Write $\sigma_\kappa = \sigma(\mathbf{p}_\kappa), d_\kappa = d(\mathbf{s}_\kappa), \sigma = \sigma(\mathbf{p})$. Recall from Section 1 that for $l \geq 0$, we write B_l^{br} for the Brownian bridge of duration 1 from 0 to $-l$. Moreover, we simply write B^{br} for the case $l = 0$.

***Proposition* 4.1.** Assume $(\mathbf{s}_\kappa, \kappa \geq 0)$ satisfies the hypothesis of Theorem 1.2, and in particular that $c_\kappa = c(\mathbf{s}_\kappa) = (1+o(1))\lambda\sigma_\kappa\mathbf{n}_\kappa^{1/2}$ as $\kappa \to \infty$ for some $\lambda > 0$ and that $\sigma_\kappa \to \sigma$. For each $\kappa \geq 0$, fix a uniform random permutation π_κ of $[\mathbf{n}_\kappa]$, and define a $C[0,1]$ function $\widetilde{W}_\kappa$ by

$$\widetilde{W}_\kappa(t) := \frac{W_{\pi_\kappa(d_\kappa)}(t\mathbf{n}_\kappa)}{\sigma_\kappa {\mathbf{n}_\kappa}^{1/2}}.$$

Then

$$\widetilde{W}_\kappa \xrightarrow{d} B_\lambda^{br} \text{ in } C[0,1].$$

To prove this theorem, we make use of the following result, which is Corollary 20.10 (a) in [5].

Theorem 4.2. *Consider a triangular array $(Z_{q,i} : 1 \leq i \leq M_q, 1 \leq q)$ of random variables satisfying*

(a) For each q, the sequence $(Z_{q,1},\cdots,Z_{q,M_q})$ *is exchangeable;*
(b) $\max_i |Z_{q,i}| \xrightarrow{p} 0$ *as* $q\to\infty$.

Define $\mu_q=\sum_i Z_{q,i}$, $\tau_q^2=\sum_i (Z_{q,i}-\frac{\mu_q}{M_q})^2$ *and* $S^q(t)=\sum_{i=1}^{\lfloor tM_q\rfloor} Z_{q,i}$.
Let $X(t)=\tau B^{br}(t)+\mu t$ *where* (τ,μ) *is independent of* B^{br}. *Then*

$$S^q \xrightarrow{d} X \text{ in } D[0,1] \text{ iff } (\mu_q,\tau_q)\xrightarrow{d}(\mu,\tau).$$

With Theorem 4.2, the proof of Proposition 4.1 is straightforward and we include the proof in Section 4 of the supplemental materials.

Let $f:\mathcal{C}_0(1)\times[0,\infty)\to\mathcal{C}_0(1)$ be defined by $f(b,v):=\theta_u(b)$ where $u=\inf\{t: b(t)\le \min_{0\le s\le 1} b(s)+v\}$. Note that since b is continuous, the minimum of b exists. Also, for $v\le -\min_{0\le s\le 1} b(s)$, we have $u=\inf\{t:b(t)=\min_{0\le s\le 1} b(s)+v\}$ and for $v\ge -\min_{0\le s\le 1} b(s)$ we have $u=0$ so $f(b,v)=\theta_0(b)=b$.

Recall from Section 1 the *first passage bridge (of unit length from 0 to* $-\lambda$*)* F_λ^{br} is

$$(F_\lambda^{br}(t),0\le t\le 1)\stackrel{d}{=}(B(t),0\le t\le 1 \mid T_\lambda=1)$$

where $T_\lambda:=\inf\{t:B(t)<-\lambda\}$ is the first passage time below level $-\lambda<0$ and B is the standard Brownian motion. We are going to use the following result from [10].

Theorem 4.3 ([10], Theorem 7)**.** *Let* ν *be uniformly distributed over* $[0,\lambda]$ *and independent of* B_λ^{br}. *Define the r.v.* $U=\inf\{t: B_\lambda^{br}(t)=\inf_{0\le s\le 1} B_\lambda^{br}(s)+\nu\}$. *Then the process* $\theta_U(B_\lambda^{br})$ *has the law of the first passage bridge* F_λ^{br}. *Moreover,* U *is uniformly distributed over* $[0,1]$ *and independent of* $\theta_U(B_\lambda^{br})$.

***Remark* 4.1.** Note that [10] considers first passage times above positive levels, whereas we consider first passage below negative levels. But the two cases are clearly equivalent.

As preparation we begin with showing the almost sure continuity of the map f. We first show that for a fixed function b, the closeness of the location where b is cyclically shifted will guarantee the continuity of the map f.

***Lemma* 4.4.** For any $b\in\mathcal{C}_0(1)$, the function $g^b:[0,1]\to\mathcal{C}_0(1)$ with $g^b(u)=\theta_u(b)$ is uniformly continuous.

Proof. We want to show that $\|\theta_u-\theta_v\|$ is small when $|u-v|$ is small. Since $\theta_u\circ\theta_v=\theta_{u+v \mod 1}$, without loss of generality, we can assume that $v=0$. In other words we just aim to bound $\|\theta_u(b)-b\|$ for small u. Fix $\delta\in(0,1/2)$ and let $\epsilon=\epsilon(\delta)=\sup_{|t-s|<\delta}|b(t)-b(s)|$ be the modulus of continuity of b. Let $0<u<\delta$. If $t\in[0,1-u]$, then $|\theta_u(b)(t)-b(t)|=|b(t+u)-b(u)-b(t)|\le|b(u)-b(0)|+|b(t+u)-b(t)|\le 2\epsilon(u)$. If $t\in[1-u,1]$, then $|\theta_u(b)(t)-b(t)|=|b(t+u-1)+b(1)-b(u)-b(t)|\le|b(t+u-1)-b(u)|+|b(1)-b(t)|\le 2\epsilon(u)$. Since $\epsilon(u)\to 0$ as $u\to 0$, the result follows. ∎

***Lemma* 4.5.** Given $b \in \mathcal{C}_0(1)$ and $0 \le v \le -\min(b)$, if $f(b,v) = \theta_{t_{v+\min(b)}}(b)$ is not continuous at (b,v), then b attains a local minimum at $t_{v+\min(b)}$.

Proof. We first fix the first argument of f and view f as a function of the second argument. By Lemma 4.4, if $f(b,v)$ is not continuous at v, then $t_{v+\min(b)}$ is not continuous at v. The continuity of b clearly implies right-continuity of $t_{v+\min(b)}$ as a function of v. Moreover, for all $0 \le v \le -\min(b)$, b attains a left-local minimum at $t_{v+\min(b)}$. Letting $t^+ = \lim_{v' \uparrow v} t_{v'+\min(b)}$, then it follows that

$$b(x) \ge v + \min(b) \text{ for all } x \in [t_{v+\min(b)}, t^+].$$

This implies that if $t_{v+\min(b)}$ is not continuous at v, then $t^+ > t_{v+\min(b)}$, so b also attains a right-local minimum at $t_{v+\min(b)}$. This proves that b attains a local minimum at $t_{v+\min(b)}$.

Secondly, if we fix the second argument of f and consider the continuity in the first argument. If $f(b,v)$ is not continuous at b, that is, we have a sequence of functions b_n with $\|b_n - b\| \to 0$, but $\|\theta_{t_{v+\min(b_n)}}(b_n) - \theta_{t_{v+\min(b)}}(b)\|$ does not converge to 0. So it must be that $t_{v+\min(b_n)}$ does not converge to $t_{v+\min(b)}$. Use a similar argument as above we can show that b attains a local minimum at $t_{v+\min(b)}$. And the continuities in v and in b imply the lemma. □

For $\lambda > 0$, we next collect a few properties of Brownian bridge B_λ^{br} and first passage bridge F_λ^{br}:

***Lemma* 4.6.** Brownian bridge B_λ^{br} satisfies the following properties:

(a) Let $\tau_+ = \inf\{t > 0 : B_\lambda^{br}(t) > 0\}$, $\tau_- = \inf\{t > 0 : B_\lambda^{br}(t) < 0\}$, then almost surely $\tau_+ = \tau_- = 0$;

(b) Given two nonoverlapping closed intervals (which may share one common endpoint) in $[0,1]$, the minima of B_λ^{br} on these two intervals are almost surely different;

(c) Almost surely, every local minimum of B_λ^{br} is a strict local minimum;

(d) The set of times where local minima are attained is countable.

Moreover, these four properties also hold for first passage bridge F_λ^{br}.

Proof. First note that the four properties are satisfied by a standard Brownian motion B (e.g. see Theorem 2.8 and Theorem 2.11 in [27]). Let C_n be the set of functions $f \in C[0,1]$ such that all four properties in the lemma occur up to time $1 - 1/n$ (i.e. the restriction of f on $[0, 1-1/n]$ satisfies all four properties). Then $\mathbf{P}(B \in C_n) = 1$ for all $n \in \mathbb{N}$. By equation (1.1) and equation (1.2) we know that the law of B_λ^{br} and the law of F_λ^{br} are both absolutely continuous with respect to the law of B up to time $1 - 1/n$. Hence we must have $\mathbf{P}\left(B_\lambda^{br} \in C_n\right) = \mathbf{P}\left(F_\lambda^{br} \in C_n\right) = 1$ for any $n \in \mathbb{N}$. This immediately implies that properties (a), (c) and (d) hold for B_λ^{br} and F_λ^{br}. It also implies (b), except for the case where one of the intervals has the form $[s,1]$ and the minimum on $[s,1]$ is reached at 1. For F_λ^{br}, by definition the global minimum $-\lambda$ is uniquely achieved at 1, hence the minimum on $[s,1]$ will not be the same as the minimum on any nonoverlapping interval. For B_λ^{br}, consider $\tilde{B}_\lambda(t) = -B_\lambda^{br}(1-t) - \lambda$, then $\tilde{B}_\lambda \stackrel{d}{=} B_\lambda^{br}$, so $\tilde{B}_\lambda$ almost surely takes positive values on any interval $[0,\epsilon]$ by property (a). It follows that $\min_{t\in[s,1]} B_\lambda^{br}(t)$ is almost surely achieved at some $t \neq 1$. This completes the proof. $\square$

***Lemma* 4.7.** Let ν be $Unif[0,\lambda]$−distributed and independent of B_λ^{br}. Then the function $f:\mathcal{C}_0(1)\times[0,\infty)\to\mathcal{C}_0(1)$ satisfies $\mathbf{P}\left(f \text{ is continuous at } (B_\lambda^{br},\nu)\right)=1$.

Proof. By Lemma 4.5, we have

$$\mathbf{P}\left(f \text{ is not continuous at } (B_\lambda^{br},\nu)\right) \leq \mathbf{P}\left(B_\lambda^{br} \text{ attains a local minimum at } t_{\nu+\min(B_\lambda^{br})}\right)$$

Let $M=\{u\in[0,1]: B_\lambda^{br} \text{ attains local minimum at } u\}$ and let $\tilde{M}=\{B_\lambda^{br}(u): u\in M\}$. By Lemma 4.6, M is countable, hence $\tilde{M}$ is countable.

Next note that $\mathbf{P}\left(B_\lambda^{br} \text{ attains a local minimum at } t_{\nu+\min(B_\lambda^{br})}\right) \leq \mathbf{P}\left(\nu+\min(B_\lambda^{br})\in\tilde{M}\right)$. Moreover, ν is a continuous random variable, independent of B_λ^{br}, so the last probability equals zero. $\square$

Now we are ready to give the proof of Theorem 1.6.

Proof of Theorem 1.6. For each $\kappa\geq 1$ let ν_κ be a uniformly random element of $[c_\kappa]-1$ independent of π_κ, and let ν be $Unif[0,\lambda]$ and independent of B_λ^{br}. By Corollary 2.3,

$$f(\widetilde{W}_\kappa, \frac{\nu_\kappa}{\sigma_\kappa \mathbf{n}_\kappa{}^{1/2}}) = f\left(\frac{W_{\pi(d(\mathbf{s}_\kappa))}(t\mathbf{n}_\kappa)}{\sigma_\kappa\mathbf{n}_\kappa{}^{1/2}}, \frac{\nu_\kappa}{\sigma_\kappa\mathbf{n}_\kappa{}^{1/2}}\right) \stackrel{d}{=} \left(\frac{S_{\mathbb{F}_\kappa}(t\mathbf{n}_\kappa)}{\sigma_\kappa\mathbf{n}_\kappa{}^{1/2}}\right)_{t\in[0,1]}.$$

By Proposition 4.1, we have $\widetilde{W}_\kappa \xrightarrow{d} B_\lambda^{br}$, and clearly we have $\sigma_\kappa{}^{-1}\mathbf{n}_\kappa{}^{-1/2}\nu_\kappa \xrightarrow{d} \nu$. By independence we have $(\widetilde{W}_\kappa, \sigma_\kappa{}^{-1}\mathbf{n}_\kappa{}^{-1/2}\nu_\kappa) \xrightarrow{d} (B_\lambda^{br},\nu)$. Since by Lemma 4.7 we have

$$\mathbf{P}\left(f \text{ is continuous at } (B_\lambda^{br},\nu)\right)=1,$$

we can apply the mapping theorem (e.g. Theorem 2.7 in [14]) to conclude that

$$f(\widetilde{W}_\kappa, \sigma_\kappa{}^{-1}\mathbf{n}_\kappa{}^{-1/2}\nu_\kappa) \xrightarrow{d} f(B_\lambda^{br},\nu).$$

By Theorem 4.3, $F_\lambda^{br} \stackrel{d}{=} f(B_\lambda^{br},\nu)$, hence we conclude that

$$\left(\frac{S_{\mathbb{F}_\kappa}(t\mathbf{n}_\kappa)}{\sigma_\kappa\mathbf{n}_\kappa{}^{1/2}}\right)_{t\in[0,1]} \xrightarrow{d} F_\lambda^{br},$$

as required. $\square$

Now we begin with the preparation work to prove Proposition 1.5. We define the map $h:\mathcal{C}_0(1)\to l_1^\downarrow$ such that for $g\in\mathcal{C}_0(1)$, $h(g)$ equals to the decreasing ordering of excursion length of $g(s)-\min_{0\leq s'<s} g(s')$. (we append at most countably many zeros to make $h(g)$ an element of $l_1^\downarrow$). Define $h_k:\mathcal{C}_0(1)\to\mathbb{R}^k$ as $h_k=\pi_k\circ h$ where $\pi_k: l_1^\downarrow\to\mathbb{R}^k$ is the projection onto the subspace spanned by the first k coordinates. To prove Proposition 1.5, we use the following result from [18].

***Lemma* 4.8.** [Lemma 3.8 and Corollary 3.10 in [18]] Suppose $\zeta : [0,1] \to \mathbb{R}$ is continuous. Let E be the set of non-empty intervals $I = (l, r)$ such that

$$\zeta(l) = \zeta(r) = \min_{s \le l} \zeta(s), \qquad \zeta(s) > \zeta(l) \text{ for } l < s < r.$$

Suppose that for all intervals $(l_1, r_1), (l_2, r_2) \in E$ with $l_1 < l_2$, we have

$$\zeta(l_1) > \zeta(l_2). \tag{4.1}$$

Suppose also that the complement of $\cup_{I \in E} I$ has Lebesgue measure 0. Fix functions $(\zeta_m, m \ge 1)$ such that $\zeta_m \to \zeta$ uniformly on $[0,1]$, and real numbers $(t_{m,i},\ m, i \ge 1)$ which satisfy the following:

(i) $0 = t_{m,0} < t_{m,1} < \cdots < t_{m,k} = 1$;

(ii) $\zeta_m(t_{m,i}) = \min_{u \le t_{m,i}} \zeta_m(u)$;

(iii) $\lim_m \max_i (\zeta_m(t_{m,i}) - \zeta_m(t_{m,i+1})) = 0$.

Then the vector consisting of decreasingly ranked elements of $\{t_{m,i} - t_{m,i-1} : 1 \le i \le k\}$ (attaching zeroes if necessary to make the vector an element in $\mathbb{R}^{|E|}$) converges componentwise and in l_1 to the vector consisting of decreasingly ranked elements of $\{r - l : (l, r) \in E\}$.

***Lemma* 4.9.** Let $\mathcal{E}$ be the set of excursions γ of $F^{br}_\lambda(s) - \min_{0 \le s' < s} F^{br}_\lambda(s')$. Then almost surely for all $\gamma_1, \gamma_2 \in \mathcal{E}$ with $l(\gamma_1) < l(\gamma_2)$, we have $F^{br}_\lambda(l(\gamma_1)) > F^{br}_\lambda(l(\gamma_2))$.

Proof. Suppose to the contrary that for some $\gamma_1, \gamma_2 \in \mathcal{E}$ with $l(\gamma_1) < l(\gamma_2)$, we have $F^{br}_\lambda(l(\gamma_1)) \le F^{br}_\lambda(l(\gamma_2))$, then since γ_1, γ_2 are excursions of $F^{br}_\lambda(s) - \min_{0 \le s' < s} F^{br}_\lambda(s')$, we must in fact have $F^{br}_\lambda(l(\gamma_1)) = F^{br}_\lambda(l(\gamma_2))$. In this case then we can find $a, b, c \in \mathbb{Q}$ such that $a < l(\gamma_1) < b < l(\gamma_2) < c$, and F^{br}_λ achieves the same minima (at $l(\gamma_1)$ and $l(\gamma_2)$ respectively) on $[a, b]$ and $[b, c]$. This has probability zero by Lemma 4.6 (b). $\square$

To prove the next lemma, we introduce the following notation. Let $(S_{1/2}(\lambda), 0 \le \lambda < \infty)$ denote a stable subordinator of index $1/2$, which is the increasing process with stationary independent increments such that

$$\mathbf{E}\left[\exp\left(-\theta S_{1/2}(\lambda)\right)\right] = \exp\left(-\lambda\sqrt{2\theta}\right), \qquad \theta, \lambda \ge 0,$$

$$\mathbf{P}\left(S_{1/2}(1) \in dx\right) = (2\pi)^{-1/2} x^{-3/2} \exp\left(-\frac{1}{2x}\right)dx, \qquad x > 0.$$

***Lemma* 4.10.** Almost surely, the coordinates of $h(F^{br}_\lambda)$ sum to 1, and are all strictly positive.

Proof. By Proposition 5 of [10], $h(F^{br}_\lambda)$ has the law of the vector of ranked excursion lengths of $|B^{br}|$ conditioned to have total local time λ at 0, which in turn has the same law as ranked excursion lengths of Brownian bridge conditioned to have total local time λ at 0 (this vector has the same law as the random vector $Y(\lambda)$ in [9], see equation (36) there). The latter is distributed as the scaled ranked jump sizes of the stable subordinator $S_{1/2}(\cdot)$ conditioned to be $\frac{1}{\lambda^2}$ at time 1 (e.g. see Theorem 4 in [9]). By Lemma 10 in [9], the coordinates of $h(F_\lambda^{br})$ almost surely sum to 1. This immediately implies that the stable subordinator almost surely has infinitely many jumps, so almost surely all coordinates of $h(F_\lambda^{br})$ are strictly positive. Indeed, suppose to the contrary that the excursion intervals are $(l_1, r_1), \cdots, (l_k, r_k)$, where $r_i \le l_{i+1}, 1 \le i \le k-1$. Then since $\sum_{i=1}^{k} (r_i - l_i) = 1$, we must in fact have $r_i = l_{i+1}, \forall 1 \le i \le k-1$ and $l_1 = 0, r_k = 1$. But this implies that $0 = F_\lambda^{br}(l_1) = F_\lambda^{br}(r_1) = F_\lambda^{br}(l_2) = \cdots = F_\lambda^{br}(l_k) = F_\lambda^{br}(r_k) = F_\lambda^{br}(1)$, contradicting to the fact $F_\lambda^{br}(1) = -\lambda < 0$. □

Proof of Proposition 1.5. We first prove that for any fixed $j \ge 1$,

$$(|\mathbb{T}_{\kappa,l}|/\mathbf{n}_\kappa)_{1\le l\le j} \xrightarrow{d} (|\gamma_l|)_{1\le l\le j}. \tag{4.2}$$

Let $\zeta_\kappa = \left(\frac{S_{\mathbb{F}_\kappa}(t\mathbf{n}_\kappa)}{\sigma_\kappa \mathbf{n}_\kappa{}^{1/2}}\right)_{t\in[0,1]}$ and let $\zeta = \left(F_\lambda^{br}(t)\right)_{t\in[0,1]}$. By (1.6) and by Skorokhod's representation theorem, we may work in a probability space in which $\zeta_\kappa \xrightarrow{a.s.} \zeta$. Let E be the set of excursion intervals of ζ. Then Lemma 4.9 guarantees equation (4.1) in Lemma 4.8 is true and Lemma 4.10 guarantees that the complement of $\cup_{I\in E} I$ has Lebesgue measure 0, as required by Lemma 4.8. For each κ let $t_{\kappa,0} = 0$ and for $1 \le j \le c_\kappa$ let $t_{\kappa,j}$ be such that $\mathbf{n}_\kappa t_{\kappa,j}$ is the time the depth-first walk $S_{\mathbb{F}_\kappa}$ finishes visiting the $j-$th tree of $\mathbb{F}_\kappa$. Then almost surely, condition (i) of Lemma 4.8 is clearly true and condition (iii) is also true since for each $1 \le j \le c_\kappa$, $\zeta_\kappa(t_{\kappa,j}) = \zeta_\kappa(t_{\kappa,j-1}) - \frac{1}{\sigma_\kappa \mathbf{n}_\kappa^{1/2}}$. The definition of Lukasiewicz walk guarantees that the times at which $\frac{S_{\mathbb{F}_\kappa}(t\mathbf{n}_\kappa)}{\sigma_\kappa \mathbf{n}_\kappa{}^{1/2}}$ hits a new minimum coincide with the times at which the walk finishes exploring the trees of the forest. Hence almost surely condition (ii) of Lemma 4.8 is also satisfied. Also note that the vector consisting of decreasingly ranked elements of $\{t_{\kappa,j} - t_{\kappa,j-1}, 1 \le j \le c_\kappa\}$ is simply the scaled decreasing ordering of tree component sizes $(|\mathbb{T}_{\kappa,l}|/\mathbf{n}_\kappa)_{1\le l\le c_\kappa}$. Hence by Lemma 4.8 we know that

$$(|\mathbb{T}_{\kappa,l}|/\mathbf{n}_\kappa)_{1\le l\le j} \xrightarrow{a.s.} h_j(F_\lambda^{br})$$

which immediately implies weak convergence. Lemma 4.10 guarantees that this is true for any positive integer j. We also have $h_j(F_\lambda^{br}) \stackrel{d}{=} (|\gamma_l|)_{1\le l\le j}$ by definition, and (4.2) follows.

To prove (1.5) from (4.2), we only need to prove that for any $\epsilon > 0$, there exists $I_0 \in \mathbb{N}$ such that $\limsup_{\kappa\to\infty} \mathbf{P}\left(\sum_{l>I_0} \frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} > \epsilon\right) < \epsilon$. Since by Lemma 4.10 we have $\sum_l |\gamma_l| = 1$ almost surely, in particular, $\lim_{I\to\infty} \mathbf{P}\left(\sum_{l>I} |\gamma_l| > \epsilon\right) = 0$. So there exists I_0 such that $\mathbf{P}\left(\sum_{l>I_0} |\gamma_l| > \epsilon\right) < \epsilon/2$. Let A_κ be the event that $\sum_{l\le I_0} \frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} < 1 - \epsilon$ and A be the event that $\sum_{l\le I_0}|\gamma_l|<1-\epsilon$ (which has probability less than $\epsilon/2$ by our choice of I_0). By (4.2), we have $|\mathbf{P}(A_\kappa)-\mathbf{P}(A)|<\epsilon/2$ for κ large enough. Therefore

$$\begin{aligned}\limsup_{\kappa\to\infty}\mathbf{P}\left(\sum_{l>I_0}\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa}>\epsilon\right) &= \limsup_{\kappa\to\infty}\mathbf{P}(A_\kappa)\\ &\le \mathbf{P}(A)+\limsup_{\kappa\to\infty}|\mathbf{P}(A_\kappa)-\mathbf{P}(A)|\le\epsilon/2+\epsilon/2=\epsilon,\end{aligned}$$

as required. □

5. Proof of Proposition 1.4 and Proposition 1.8

We assume that we have the conditions of Theorem 1.2 hold. In particular, we have a probability distribution $\mathbf{p}$ on $\mathbb{N}$. Recall that $\sigma=\sigma(\mathbf{p}),\sigma_\kappa=\sigma(\mathbf{p}_\kappa)$. Let $\mathbf{s}_{\kappa,l}=(s^{(i)}_{\kappa,l},i\ge 0)$ denote the degree sequence of $\mathbb{T}_{\kappa,l}$ and let $\mathbf{n}_{\kappa,l}=n(\mathbf{s}_{\kappa,l})$. Recall that $p^{(i)}_\kappa=s^{(i)}_\kappa/\mathbf{n}_\kappa$ and let $p^{(i)}_{\kappa,l}=s^{(i)}_{\kappa,l}/\mathbf{n}_{\kappa,l}$ be the empirical proportion of degree i among all vertices of the $l-$th largest tree $\mathbb{T}_{\kappa,l}$. Note that $p^{(i)}_\kappa$ is deterministic while $p^{(i)}_{\kappa,l}$ is random.

First, we are going to prove Proposition 1.4 by using Theorem 1.7. To do so, we will have to first show that the assumptions of Theorem 1.7 are satisfied in our setting.

***Proposition* 5.1.** Under the assumption of Theorem 1.2, for all $l\ge 1$, as $\kappa\to\infty$ we have

(a) $\mathbf{p}_{\kappa,l}\xrightarrow{p}\mathbf{p}$ coordinatewise, that is, $p^{(i)}_{\kappa,l}\xrightarrow{p}p^{(i)}$ for all $i\ge 1$.

(b) $\sigma(\mathbf{p}_{\kappa,l})\xrightarrow{p}\sigma(\mathbf{p})$.

Proof. For (a), we know that by Lemma 3.4 and Proposition 3.5, for fixed $\epsilon>0,i,l\in\mathbb{N}$ and κ large enough, we have

$$\mathbf{P}\left(|p^{(i)}_{\kappa,l}-p^{(i)}_\kappa|>\epsilon\right)\le 2/\mathbf{n}_\kappa+\mathbf{P}\left(|\mathbb{T}_{\kappa,l}|\le\mathbf{n}_\kappa^{1/4}\right). \tag{5.1}$$

For any $\epsilon'>0$, there exists $\delta>0$ such that $\mathbf{P}(|\gamma_l|<\delta)<\epsilon'/2$ and by (4.2) we can find κ_0 such that for all $\kappa\ge\kappa_0$ we have $\mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa}<\delta\right)\le\mathbf{P}(|\gamma_l|<\delta)+\epsilon'/2$ and $\mathbf{n}_\kappa^{-3/4}<\delta$. Hence $\mathbf{P}\left(|\mathbb{T}_{\kappa,l}|\le\mathbf{n}_\kappa^{1/4}\right)=\mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa}\le\mathbf{n}_\kappa^{-3/4}\right)\le\mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa}\le\delta\right)<\epsilon'$. Hence $\mathbf{P}\left(|\mathbb{T}_{\kappa,l}|\le\mathbf{n}_\kappa^{1/4}\right)=o(1)$ as $\kappa\to\infty$. Therefore by (5.1) we know that $|p^{(i)}_{\kappa,l}-p^{(i)}_\kappa|\xrightarrow{p}0$ as $\kappa\to\infty$, which implies (a) since by assumption of Theorem 1.2 we have $\mathbf{p}_\kappa$ converges to $\mathbf{p}$ coordinatewise.

Now we proceed to prove (b). Fix $l\ge 1$ and $\delta>0$, and let $\epsilon>0$ be small enough that

$$\limsup_{\kappa\to\infty}\mathbf{P}(|\mathbb{T}_{\kappa,l}|<\epsilon\mathbf{n}_\kappa)<\delta.$$

Such ϵ exists by (4.2).

Then let M be large enough that $\sigma^2_{\kappa,>M} := \sum_{i>M} i^2 \frac{s_\kappa^{(i)}}{\mathbf{n}_\kappa} < \epsilon^2$ for all κ (such M exists since under the assumption of Theorem 1.2 σ^2_κ converges). And let $\sigma^2_{\kappa,l,>M} = \sum_{i>M} i^2 \frac{s_{\kappa,l}^{(i)}}{\mathbf{n}_{\kappa,l}}$ similarly. Note that

$$\sigma^2_{\kappa,l,>M} \leq \sum_{i>M} i^2 \frac{s_\kappa^{(i)}}{|\mathbb{T}_{\kappa,l}|} = \sigma^2_{\kappa,>M} \frac{\mathbf{n}_\kappa}{|\mathbb{T}_{\kappa,l}|},$$

so if $\sigma^2_{\kappa,l,>M} > \epsilon$ then $|\mathbb{T}_{\kappa,l}| < \epsilon \mathbf{n}_\kappa$. By the triangle inequality, we have

$$|\sigma^2(\mathbf{p}_{\kappa,l}) - \sigma^2(\mathbf{p}_\kappa)| \leq \sum_{i\leq M} i^2 |p_{\kappa,l}^{(i)} - p_\kappa^{(i)}| + \sum_{i>M} i^2 p_{\kappa,l}^{(i)} + \sum_{i>M} i^2 p_\kappa^{(i)}.$$

Since $|p_{\kappa,l}^{(i)} - p_\kappa^{(i)}| \to 0$ in probability for all i by part (a), and $\sum_{i>M} i^2 p_\kappa^{(i)} < \epsilon^2 < \epsilon$, and $\sigma(\mathbf{p}_\kappa) \to \sigma(\mathbf{p})$ by assumption of Theorem 1.2, this yields that

$$\limsup_{\kappa\to\infty} \mathbf{P}\left(|\sigma^2(\mathbf{p}_{\kappa,l}) - \sigma^2(\mathbf{p})| > 4\epsilon\right) \leq \limsup_{\kappa\to\infty} \mathbf{P}\left(\sum_{i>M} i^2 p_{\kappa,l}^{(i)} > \epsilon\right) \leq \limsup_{\kappa\to\infty} \mathbf{P}\left(|\mathbb{T}_{\kappa,l}| < \epsilon \mathbf{n}_\kappa\right) < \delta,$$

which proves part (b). □

***Lemma* 5.2.** Let $\Delta_{\kappa,l}$ be the largest degree of a vertex of $\mathbb{T}_{\kappa,l}$. For any fixed l, we have

$$\frac{\Delta_{\kappa,l}}{\sqrt{|\mathbb{T}_{\kappa,l}|}} \xrightarrow{p} 0 \text{ as } \kappa \to \infty.$$

Proof. For any $\delta > 0$, we need to prove $\lim_{\kappa\to\infty} \mathbf{P}\left(\frac{\Delta_{\kappa,l}}{\sqrt{|\mathbb{T}_{\kappa,l}|}} > \delta\right) = 0$. For any $\epsilon > 0$, by Lemma 4.10 we can choose $\epsilon' > 0$ such that $\mathbf{P}(|\gamma_l| < \epsilon') \leq \epsilon/2$. Then choose κ_0 such that when $\kappa \geq \kappa_0$ we have

$$\frac{\Delta_\kappa^2}{\mathbf{n}_\kappa} \cdot \frac{1}{\delta^2} < \epsilon' \text{ and } \mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} < \epsilon'\right) \leq \mathbf{P}(|\gamma_l| < \epsilon') + \frac{\epsilon}{2}.$$

This is possible since $\Delta_\kappa = o(\mathbf{n}_\kappa^{1/2})$ by Remark 1.1 and $|\mathbb{T}_{\kappa,l}|/\mathbf{n}_\kappa \xrightarrow{d} |\gamma_l|$ by (4.2). Therefore

$$\mathbf{P}\left(\frac{\Delta_{\kappa,l}}{\sqrt{|\mathbb{T}_{\kappa,l}|}} > \delta\right) \leq \mathbf{P}\left(\frac{\Delta_\kappa}{\sqrt{|\mathbb{T}_{\kappa,l}|}} > \delta\right) = \mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} < \frac{\Delta_\kappa^2}{\mathbf{n}_\kappa} \cdot \frac{1}{\delta^2}\right) \leq \mathbf{P}\left(\frac{|\mathbb{T}_{\kappa,l}|}{\mathbf{n}_\kappa} < \epsilon'\right) \leq \epsilon,$$

hence the claim. □

With Proposition 5.1 and Lemma 5.2, we are now ready to give the proof of Proposition 1.4.

Proof of Proposition 1.4. Let $\mathbf{s}_{\kappa,l}$ be the random degree sequence of the $l-$th largest tree in the forest $\mathbb{F}_\kappa$. Then by Proposition 1.5, we have

$$\left(\frac{n(\mathbf{s}_{\kappa,1})}{\mathbf{n}_\kappa}, \cdots, \frac{n(\mathbf{s}_{\kappa,j})}{\mathbf{n}_\kappa}\right) \xrightarrow{d} (|\gamma_1|, \cdots, |\gamma_j|)\,.$$

By Proposition 5.1 and Lemma 5.2, we know we can apply Theorem 1.7 to $\mathcal{T}_{\kappa,l}$ to conclude that for each fixed $l \le j$,

$$\frac{\mathbf{n}_\kappa^{1/2}}{n(\mathbf{s}_{\kappa,l})^{1/2}} \mathcal{T}_{\kappa,l} \xrightarrow{d} \mathcal{T}_{\mathbf{e}_l}$$

where $(\mathbf{e}_l)_{l \le j}$ are independent copies of $\mathbf{e}$. Since the trees $(\mathcal{T}_{\kappa,l}, l \le j)$ are conditionally independent given their degree sequences, it follows that

$$\left(\frac{\mathbf{n}_\kappa^{1/2}}{n(\mathbf{s}_{\kappa,l})^{1/2}} \mathcal{T}_{\kappa,l}, l \le j\right) \xrightarrow{d} (\mathcal{T}_{\mathbf{e}_l}, l \le j)\,.$$

The result follows by Brownian scaling. $\square$

Finally, we give the proof of Proposition 1.8 based on Proposition 3.8, with the assumptions of Theorem 1.3.

Proof of Proposition 1.8. By assumption we have $\sigma_\kappa \to \sigma \in (0,\infty)$ and $s_\kappa^{(1)}/|\mathbf{s}_\kappa| \to p^{(1)} < 1$. Fix $\rho > 0$ and let $\epsilon > 0$ be such that $2\epsilon < \sigma^2 < \frac{1}{2\epsilon}$. Then let $\beta_0 = \beta_0(\rho,\epsilon)$ be as in Proposition 3.8, so that for all n sufficiently large, if a degree sequence $\mathbf{s}$ satisfies $|\mathbf{s}| = n, \Delta(\mathbf{s}) \le n^{\frac{1-\epsilon}{2}}, s^{(1)} \le (1-\epsilon)|\mathbf{s}|$ and $\epsilon \le \sigma^2(\mathbf{s})/n \le 1/\epsilon$, then for any $0 < \beta < \beta_0$,

$$\mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}) : |T| < \beta n, h(T) > \beta^{1/8} n^{1/2}\right) \le \rho.$$

For κ sufficiently large, $\mathbf{s}_\kappa$ satisfies these conditions. Hence for any $0 < \beta < \beta_0$,

$$\mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}_\kappa) : |T| < \beta \mathbf{n}_\kappa, h(T) > \beta^{1/8} \mathbf{n}_\kappa^{1/2}\right) \le \rho. \tag{5.2}$$

Finally, taking $\beta = (a/\sigma_\kappa)^8$ in (5.2), since $\mathcal{T}_{\kappa,l} = \frac{\sigma_\kappa}{2\mathbf{n}_\kappa^{1/2}} \mathbb{T}_{\kappa,l}$ and for all $j > 1/\beta$ we have $|\mathbb{T}_{\kappa,j}| < \beta \mathbf{n}_\kappa$, it follows that for all κ sufficiently large,

$$\begin{aligned} \mathbf{P}\left(\sup_{l>j} h(\mathbb{T}_{\kappa,l}) > \frac{a\mathbf{n}_\kappa^{1/2}}{\sigma_\kappa}\right) &\le \mathbf{P}\left(\exists T \in \mathbb{F}(\mathbf{s}_\kappa) : |T| < \beta \mathbf{n}_\kappa, h(T) > \beta^{1/8} \mathbf{n}_\kappa^{1/2}\right) \\ &\le \rho. \end{aligned}$$

Since $\mathrm{diam}(\mathcal{T}_{\kappa,l}) \le 2h(\mathcal{T}_{\kappa,l})$, the result now follows easily. $\square$

6. acknowledgements

I would like to thank Louigi Addario-Berry for suggesting this project and numerous helpful discussions thereafter. I would also like to thank anonymous referees for many helpful suggestions. This work was partially supported by NSERC CGS and I thank the institution.

References

[1] Abraham, R., Delmas, J.-F. and Hoscheit, P. (2014). Exit times for an increasing Lévy tree-valued process. *Probab. Theory Related Fields* **159**, 357–403.

[2] Abraham, R., Delmas, J.-F. and Hoscheit, P. (2013). A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18**, 1–21.

[3] Addario-Berry, L. (2012). Tail bounds for the height and width of a random tree with a given degree sequence. *Random Structures and Algorithms* **41**, 253–261.

[4] Addario-Berry, L., Broutin, N. and Goldschmidt, C. (2010). Critical random graphs: limiting constructions and distributional properties, *Electronic Journal of Probability*, **15**, 741–775.

[5] Aldous, D. (1985). Exchangeability and related topics, Ecole d'étë de probabilités de Saint-Flour, XIII. *Lecture Notes in Mathematics,* **1117**, Springer, Berlin, 1–198.

[6] Aldous, D. (1991). The continuum random tree. I. *Annals of Probability*, **19**, 1–28.

[7] Aldous, D. (1991). The continuum random tree. II. An Overview. *Stochastic analysis (Durham, 1990)*, London Math. Soc. Lecture Note Ser., **167**, Cambridge Univ. Press, Cambridge, 23–70.

[8] Aldous, D. (1993). The continuum random tree. III. *Annals of Probability*, **21**, 248–289.

[9] Aldous, D. (1998). The standard additive coalescent. *Annals of Probability*, **26**, 1703–1726.

[10] Bertoin, J., Chaumont, L. and Pitman, J. (2003). Path transformations of first passage bridges. *Electron. Comm. Probab.* **8**, 155–166.

[11] Bender, E. and Canfield, E. (1978). The asymptotic number of labeled graphs with given degree sequences, *J. Combin. Theory, Ser. A*, **24**, 296–307.

[12] Bettinelli, J. (2010). Scaling limits for random quadrangulations of positive genus. *Electron. J. Probab.* **15**, 1594–1644.

[13] Bettinelli, J. (2015). Scaling limit of random planar quadrangulations with a boundary. *Ann. Inst. H. Poincaré Probab. Statist.* **51**, 432–477.

[14] Billingsley, P. (1999). *Convergence of Probability Measures*, Second Edition, John Wiley & Sons, New York.

[15] Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *Eur. J. Combin.*, **1**, 311–316.

[16] Broutin, N. and Marckert, J.-F. (2014). Asymptotics of trees with a prescribed degree sequence and applications. *Random Struct. Alg.*, **44**, 290–316.

[17] Burago, D., Burago, Y. D. and Ivanov, S. (2001). *A course in metric geometry*, AMS Providence.

[18] Chassaing, P. and Louchard, G. (2002). Phase Transition for Parking Blocks, Brownian Excursion and Coalescence. *Random Struct. Alg.* **21**, 76–119.

[19] Duquesne, T. and Le Gall, J.-F. (2002). Random trees, Lévy processes and spatial branching processes. *Astérisque,* **281**.

[20] Evans, S. N. (2008). Probability and real trees. *Lecture Notes in Mathematics.* Springer.

[21] Fitzsimmons, P., Pitman, J. and Yor, M. (1992). Markovian bridges: Construction, Palm Interpretation, and Splicing. *Seminar on Stochastic Processes*, 101–130, Birkhäuser.

[22] Janson, S. and Stefánsson, S. Ö. (2015). Scaling limits of random planar maps with a unique large face, *Annals of Probability*, **43**, 1045–1081.

[23] Le Gall, J.-F. (2005). Random trees and applications. *Probability Surveys*, **2**, 245–311.

[24] Marckert, J.-F. and Miermont, G. (2011). The CRT is the scaling limit of unordered binary trees, *Random Structures and Algorithms*, **38**, 467–501.

[25] McDiarmid, C. (1998). Concentration. *Probabilistic Methods for Algorithmic Discrete Mathematics,* 195–248, New York, Springer.

[26] Miermont, G. (2008). Invariance principles for spatial multitype galton-watson trees. *Ann. Inst. Henri Poincare Probab. Stat.*, **44**, 1128–1161.

[27] Mörters, P. and Peres, Y. (2010). *Brownian Motion*, Cambridge University Press.

[28] Pitman, J. (2006). Combinatorial stochastic processes, *Lecture Notes in Mathematics,* **1875**, Springer, Berlin.