Critical and Supercritical Scaling Limits of Random Forests with Prescribed Degree Sequences

Tao Lei

Doctor of Philosophy

Department of Mathematics and Statistics

McGill University

Montréal, Québec

2018-03-16

A thesis submitted to McGill University in partial fulfilment of the requirements for the degree of Doctor of Philosophy

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis supervisor, Professor Louigi Addario-Berry, who patiently guided me through my Ph.D. program. He introduced me to the world of random graphs. Without him, this work will never materialize.

I would also like to thank members of the probability group at McGill University, in particular, Professor Linan Chen, Professor Luc Devorye and Professor Bruce Reed. I also appreciate the staffs at department of mathematics and statistics at McGill. I thank Professor James Martin and Professor Linan Chen for examining my thesis. They provided many helpful suggestions which improved the quality of the final version of the thesis.

My study was partially supported by NSERC CGS and by Schulich fellowship. I thank them for the generous financial support.

I would also like to thank my friends and fellow math students at McGill, in particular, Xiangwen Zhang, Eslava Laura, Luiz Takei, Siyuan Lu, Yuting Wen, Xianchao Wu, Shaodong Wang and Brody Causley for enlightening conversations, baked in everyday life.

I thank my parents for their unconditional love and support. Their support are critical for me. And finally I would like to dedicate this work to my beloved wife, Yishu, for her faith and love in me, and our time together at McGill.

In China, twelve years form a cycle (e.g. Chinese zodiac). It has been a cycle since I came to Canada to pursue graduate studies. I sincerely thank everyone involved in this cycle of my life and look forward to the next chapter of the journey.

ABSTRACT

In this paper, we consider random plane forests uniformly drawn from all possible plane forests with a given degree sequence. Under suitable conditions on the degree sequences, we consider the possible scaling limits, with respect to the Gromov-Hausdorff-Prokhorov topology, of a sequence of such forests as the number of vertices tends to infinity. 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, 7, 8]. We study the scaling limit in two regimes: critical and supercritical. In the critical regime we identify the limiting random object as a sequence of random real trees encoded by excursions of some first passage bridges reflected at their minima. We establish such convergence by studying the associated Lukasiewicz walk of the degree sequences. In the supercritical regime, where there is a unique "giant tree" containing all but a vanishing fraction of the nodes, we give a description of the limit of the forest of "small trees" obtained by removing the giant tree. We accomplish this by relating plane forests to marked cyclic forests and the corresponding lattice paths. Our work is closely related to and uses the results from the recent work of Broutin and Marckert [20] on scaling limit of random trees with prescribed degree sequences.

ABRÉGÉ

Dans cet article, nous considérons les forêts planes aléatoires tirées uniformément de toutes les forêts planes possibles avec une séquence de degrés donnée. Dans des conditions appropriées sur les séquences de degrés, nous considérons les limites d'échelle possibles, par rapport à la topologie de Gromov-Hausdorff-Prokhorov, d'une séquence de telles forêts, lorsque le nombre de sommets tend vers l'infini. Ce travail s'inscrit dans le cadre général de la preuve de la convergence de structures combinatoires aléatoires à certaines limites d'échelle de Gromov-Hausdorff, décrites en termes *de l'arbre aléatoire continu brownien* (AACB) introduit par Aldous [6, 7, 8]. Nous étudions la limite d'échelle dans deux régimes: critique et supercritique. Dans le régime critique, nous identifions l'objet aléatoire limitant comme une séquence d'arbres réels aléatoires codés par des excursions de certains ponts de premier passage réfléchis à leurs minimas. Nous établissons cette convergence en étudiant la marche de Lukasiewicz associée aux séquences de degrés. Dans le régime supercritique, où il y a un «arbre géant» unique contenant tout sauf une fraction disparaissant des nœuds, nous donnons une description de la limite de la forêt de «petits arbres» obtenue en enlevant l'arbre géant. Nous réalisons ceci en reliant les forêts planes aux forêts cycliques marquées et aux processus de codage correspondants. Notre travail est étroitement lié au travail récent de Broutin et Marckert [20] sur la limite d'échelle des arbres aléatoires avec des séquences de degrés prescrits, et s'appuie sur ces derniers.

PREFACE

Chapters 3 and 4 of this thesis are based on the papers [47] and [48], respectively. Both papers are authored by myself. My supervisor Prof. Louigi Addario-Berry gave a lot helpful suggestions during the writings of these two papers.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Introduction

In this thesis, we work on the problem of characterizing the scaling limits of uniformly random plane forests with prescribed sequences in different regimes.

In a lot of cases, real-world networks can be viewed as graphs. For example, the world-wide-web (WWW), where vertices are the webpages and edges are the hyperlinks between webpages. A social network is another example of a graph, where each person is a vertex and the friendship relation defines the edge between vertices. All these graphs are of enormous size and keep evolving. Hence an appropriate way to model these networks is graph-valued random process. This leads to the study of random graphs, which goes back to the seminal papers [30, 31] by Erdös and Rényi where they studied the binomial random graph models $G(n, p)$ and $G(n, m)$. Since then a lot of work has been carried out in this field.

Since we need to understand graphs of large size, it is natural to raise questions about "limits" of graphs. Convergence of finite graphs is an area at the intersection of combinatorics and probability. Convergence theorems are results about the asymptotic structures of sequences of graphs where the size of the graphs tends to infinity along the sequence. There are several different directions in this area, which capture different asymptotic properties of the graph. In the "dense" regime, Lovász and Szegedy [50, 49] introduced the concept of *graphons* and used the *cut metric* [19] as a key tool to implement the idea of graph limits, making links to Szemerédi's regularity lemma. In the "sparse" regime, there is the notion of *local weak convergence* initiated by Benjamini and Schramm [13]. A third approach is to consider the Gromov-Hausdorff convergence of the graphs, which is related to the global metric space structure of the graphs. In this work, we take the last approach to study the scaling limit of large random forests.

In the field of Gromov-Hausdorff limits of random graphs, one important body of work is on scaling limits of *random trees*. Seminal results in this area were proved in Aldous's trilogy of papers [6, 7, 8] in 1990's. Since then, people started considering the combinatorial structure as a metric space (giving each edge length 1) and tried finding limit random metric space under suitably scaling. Aldous in particular proved that if $(T_n, n \geq 1)$ is a sequence of trees, with T_n uniformly drawn from n^{n-1} trees on n labelled vertices, then after scaling by $n^{1/2}$, the sequence converge in distribution to a random compact metric space, called *Brownian Continuum Random Tree* (BCRT). Since then, a large class of random tree models have been shown to have the BCRT as scaling limits which showcases the universality of BCRT: e.g. critical multitype Galton-Watson trees [56], unordered binary trees [52], uniformly unordered trees [36], random trees with a prescribed degree sequences satisfying certain conditions [20] and uniform unlabelled unrooted trees [65]. Moreover, classical random graph models, such as the famous Erdös-Rényi random graph model, have also been shown to have Gromov-Hausdorff scaling limits, which can be described in terms of BCRT. For example, in the works [5, 4] by Addario-Berry, Broutin and Goldschmidt, they showed that the scaling limit of critical Erdös-Rényi random graph consists of rescaled BCRT's glued together at a finite number of points. More and more evidence have convinced the central role of BCRT in this field. More universality results about other random graphs and random discrete structures are also proved, such as random graphs from sub-critical classes [58], random dissections [25], random planar maps with a unique large face [41] and random planar quadrangulations with a boundary [16].

The objective of this thesis is to investigate Gromov-Hausdorff convergence for combinatorial models of random forests. In particular, my focus is the asymptotic metric structure of random forests, uniformly drawn from all plane forests with a prescribed degree sequence. This is motivated by the metric structure of graphs with a prescribed degree sequence, introduced by Bender and Canfield [12] and by Bollobás [18] in the form of the *configuration model*. This model can give rise to graphs with any particular (legitimate) prescribed degree sequence (e.g. heavy tailed degree distributions, observed in realistic network, but a feature not grasped by the Erdös-Rényi random graph model). This flexibility is perhaps why the "degree-sequence" models have become very popular and found use in diverse areas such as food webs [64], opinion dynamics [69], economic network effects [66] and academic career trajectories [51], giving one of the motives to our work. A brief history of the application of configuration model can be found in Section 1.4 of [35].

More precisely, the work of this thesis is a natural continuation and generalization of the work [20] by Broutin and Marckert, where they studied the asymptotic behavior of a tree chosen uniformly at random amongst the set of rooted plane trees with a prescribed degree sequence. They showed that under natural assumptions on the degree sequences and with suitable scaling, the random trees converge toward the BCRT. The authors predict that their work is a first step to setting up invariance principles for critical random graphs with a prescribed degree sequence, and our work can be viewed as one step in that direction.

The model studied in [20] and hence in this work is related to Galton-Watson trees [11, 38], (which is closely related to *simply generated trees* in the combinatorics literature), by conditioning on observing the prescribed degree sequence. In section 1.4.2 we will illustrate the concept of simply generated trees and a few examples of this class. An excellent survey of existing work, and more examples, can be found in [40]. The combinatorial approach often involves representing quantities of interest as coefficients of power series, and applying analytic tools [27, 34] such as singularity analysis or saddle-point methods to obtain the limit. With the more probabilistic approach, a lot of related work were done by Pavlov and summarized in [59]. In particular, he considered a random forest $\mathfrak{F}_{N,n}$ consisting of N simply generated random trees and n non-root vertices. He computed [60] the limit behavior of the maximum size of a tree in $\mathfrak{F}_{N,n}$ under different assumptions of relationships between N and n. To do this, he exploited the connection between $\mathfrak{F}_{N,n}$ and Galton-Watson process with N initial particles and found integral and local convergence of the distributions of sums of certain auxiliary independent random variables. We will present related work in this direction in more detail in Section 1.4.4.

Instead of focusing on random quantities such as the size of the largest tree, we aim to push the probabilistic approach further and view random forest as random metric space and prove certain weak convergence of stochastic processes. The convergence of interesting quantities will then be read off from the convergence of the processes. Establishing such convergence results often uses some functional encoding of the discrete structure under investigation. We will illustrate several common coding functions when dealing with limits of graphs in Section 1.2.2.

Other forest models have also been studied in the literature. For example, Luczak and Pittel [44] studied forest $\mathcal{F}(n, M)$, chosen uniformly from the family of all labelled *unrooted* forests with n vertices and M edges. They showed that this model exhibited three regimes of asymptotic behavior: subcritical, nearcritical and supercritical and the phase transition happened at $M = n/2$, just like Erdös-Rényi random graph $G(n, M)$. For each of the phases, they determined the limit distribution of the size of the k-th largest component of $\mathcal{F}(n, M)$. Along this direction very recently Martin and Yeo [53] gave a full description of the scaling limit of the largest component of $\mathcal{F}(n, M)$ inside the critical window $M = n/2 + O(n^{2/3})$.

1.2 Notations of trees and forests

1.2.1 Plane trees and forests

We recall the following definition of plane trees (as in e.g. [28]). Let

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

where $\mathbb{N} = \{1, 2, \cdots\}$ and $\mathbb{N}^0 = \{\emptyset\}$. If $u = (u_1, u_2, \cdots, u_n) \in \mathcal{U}$ we write $u = u_1 u_2 \cdots u_n$ for short and let $|u| = n$ be the *generation* of u, or the *height* of u. If $u = u_1 \cdots u_m, v = v_1 \cdots v_n$, we write $uv = u_1 \cdots u_m v_1 \cdots v_n$ for the *concatenation* of u and v.

Definition 1.2.1. *A* rooted plane tree T *is a subset of* $\mathcal{U}$ *satisfying the following conditions:*

(i) $\emptyset \in \mathrm{T}$*;*

(ii) If $v \in \mathrm{T}$ *and* $v = uj$ *for some* $u \in \mathcal{U}$ *and* $j \in \mathbb{N}$*, then* $u \in \mathrm{T}$*;*

(iii) For every $u \in \mathrm{T}$*, there exists a number* $k_{\mathrm{T}}(u) \geq 0$ *such that* $uj \in \mathrm{T}$ *if and only if* $1 \leq j \leq k_{\mathrm{T}}(u)$*. We call* $k_{\mathrm{T}}(u)$ *the* degree *of* u *in* T*.*

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 a forest F, we let $\mathrm{F}^{\downarrow}$ be the sequence of tree components of F in decreasing order of size, breaking ties lexicographically.

Degree sequence

Definition 1.2.2. *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 since

$$\sum_{i \geq 0} i s^{(i)}(\mathrm{F}) = \sum_{j=1}^{m} \sum_{u \in \mathrm{T}_j} k_{\mathrm{T}_j}(u) = \sum_{j=1}^{m} (|\mathrm{T}_j| - 1)$$

and $\sum_{i\geq 0} s^{(i)}(\mathrm{F}) = \sum_{j=1}^{m} |\mathrm{T}_j|$, the number of tree components in F is always $c(\mathbf{s}(\mathrm{F}))$. For any degree sequence $\mathbf{s}$, we adopt the notations

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

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

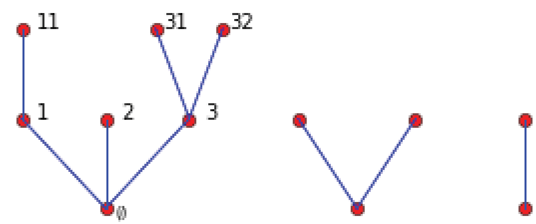

Figure 1–1: A plane forest (with labels for the first tree) with degree sequence $\mathbf{s} = (7, 2, 2, 1, 0, \cdots)$

1.2.2 Codings of trees and forests

To study the structure of plane trees and plane forests, it is common to instead study some related coding functions of the discrete structure. In this section, we introduce several commonly used coding functions of plane trees and plane forests.

Height process

For a rooted plane tree T, list the vertices of T in the lexicographic order as $u_1, u_2, \ldots, u_{|T|}$. Define the *height process* of T as the function $H : [0, |T| - 1] \to \mathbb{R}^+$ such that for integer k, $H(k) = |u_{k+1}|$ and the process at non-integer times is defined by linear interpolation. For a forest $F = (T_1, \ldots, T_n)$, the height process of F will be the concatenated height processes of $T_1, \ldots, T_n$. In Figure 1–2 we have the height process of the first tree component of the forest in Figure 1–1.

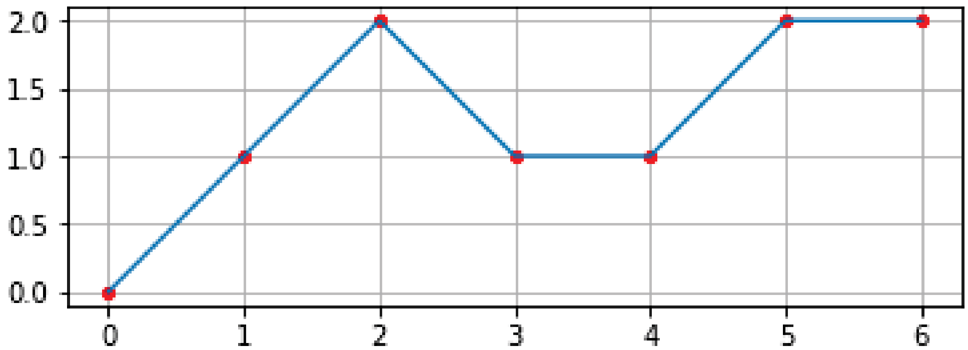

Figure 1–2: The height process of the first tree in Figure 1–1

Contour process

Let T be a rooted plane tree and view each edge of T as an interval of unit length. Imagine that a particle moves along the edges of T with unit speed, starting from the root at time 0. Each time the particle leaves a vertex u, it moves to the lexicographically next unvisited child of u, if such a child exists; otherwise it moves back to the parent of u. The exploration concludes the moment the particle has visited all vertices and returned to the root. Let $C(t)$ be the graph distance between the particle and the root at time t. Then $C : [0, 2(|T| - 1)] \to \mathbb{R}^+$ defines a function and is called the *contour process* of T. This is also referred as *Harris Walk* [37] or *Dyck path* [63]. For a forest, the contour process is simply the concatenation of individual contour processes for each tree. In Figure 1–3 we have the contour process of the first tree component of the forest in Figure 1–1.

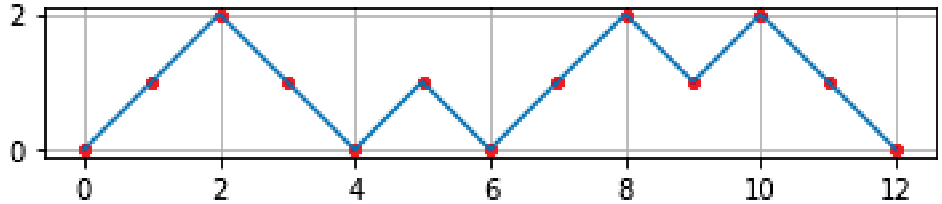

Figure 1–3: The contour process of the first tree in Figure 1–1

Depth-first walk

Let T be a finite tree and let $u_1, u_2, \ldots, u_{|T|}$ be the vertices of T listed in lexicographic order. The *depth-first walk* (or *Lukasiewicz path*) $\mathcal{W}(T) = \{\mathcal{W}_i(T) : 0 \le i \le |T|\}$ of T is defined by setting $\mathcal{W}_0(T) = 0$, and for $0 \le i \le |T| - 1$, letting

$$\mathcal{W}_{i+1}(T) = \mathcal{W}_i(T) + k_T(u_{i+1}) - 1.$$

When the context is clear, sometimes we omit the argument T for simplicity. Without confusion, sometimes we view $\mathcal{W}(T)$ as a function defined on $[0, |T|]$ where the values of the function at non-integer values are defined by linear interpolation. In Figure 1–4 we have the depth-first walk of the first tree component of the forest in Figure 1–1.

It is easy to see that for a rooted plane tree T, the depth-first walk $\mathcal{W}(T)$ satisfies the following properties:

- $\mathcal{W}_0 = 0, \mathcal{W}_{|T|} = -1$;
- $\mathcal{W}_i \ge 0$ for every $0 \le i \le |T| - 1$;
- $\mathcal{W}_i - \mathcal{W}_{i-1} \ge -1$ for every $1 \le i \le |T|$.

In particular, this implies that

$$|T| = \inf\{0 \le i \le |T| : \mathcal{W}_i = -1\}.$$

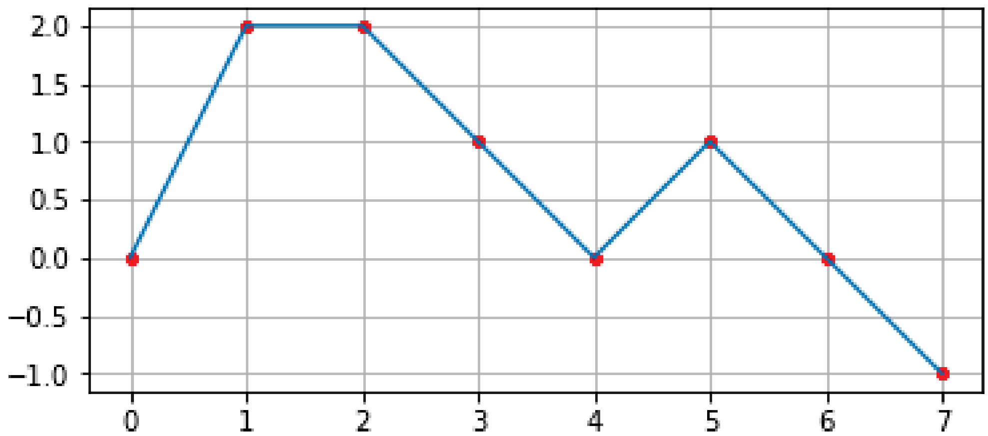

Figure 1–4: The depth-first walk of the first tree in Figure 1–1

In general, for a forest $F = (T_1, \cdots, T_n)$, we list the vertices of F as $u_1, \ldots, u_{|F|}$ by first listing the vertices of T_1 in lexicographic order, then the vertices of T_2 in lexicographic order, and so on. Then the same definition gives the depth-first walk $\mathcal{W}(F)$ of F. We always have $\mathcal{W}_0(F) = 0, \mathcal{W}_{|F|}(F) = -n$ and for any integer $1 \le j \le n$,

$$\inf\{l : \mathcal{W}_l(F) = -j\} = |T_1| + \cdots + |T_j|.$$

These coding functions often have similar asymptotic behaviour (with slightly different scalings), in the setting of random forests. Sometimes one coding function may be more directly related to the functionals under consideration (e.g. the metric structure is easily read off from the contour function) while other coding function (e.g. the depth-first walk) is more suitable for proving convergence theorem by applying Donsker's Theorem [17] or martingale techniques. For example, the following result showed that upon scaling, the concatenated contour processes of a sequence of infinitely many i.i.d. Galton-Watson trees converges to a reflected Brownian motion.

Theorem 1.2.3 ([46], Theorem 6.5 in [61])**.** *Let* $(C(t), t \geq 0)$ *be the continuous path obtained by concatenation of the contour processes of an infinite independent and identically distributed sequence of critical Galton-Watson trees with finite non-zero offspring variance* σ^2*. Then as* $n \to \infty$*,*

$$\left(\frac{C(2nt)}{\sqrt{n}}, t \geq 0\right) \xrightarrow{d} \left(\frac{2}{\sigma}|B_t|, t \geq 0\right)$$

in the sense of weak convergence in $C[0, \infty)$*, where* B *is a standard Brownian motion.*

In [46], Theorem 1.2.3 was proved by using standard results for depth-first walk and then relating the contour process to the depth-first walk. The height process was used as an intermediary in this comparison.

1.3 Our forest models and summary of main results

For any degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$, we let $\mathrm{T}(\mathbf{s})$ denote the set of all plane trees with degree sequence $\mathbf{s}$ and $\mathrm{F}(\mathbf{s})$ denote the set of all plane forests with degree sequence $\mathbf{s}$. Let $\mathbb{P}_{\mathbf{s}}$ be the uniform measure on $\mathrm{T}(\mathbf{s})$ and let $\mathbb{T}(\mathbf{s})$ be a random plane tree with law $\mathbb{P}_{\mathbf{s}}$. Let $\mathbb{Q}_{\mathbf{s}}$ be the uniform measure on $\mathrm{F}(\mathbf{s})$ and let $\mathbb{F}(\mathbf{s})$ be a random plane forest with law $\mathbb{Q}_{\mathbf{s}}$. $\mathbb{T}(\mathbf{s})$ is called the random tree with degree sequence $\mathbf{s}$. This model is studied in [20] where it is shown that under reasonable conditions on degree sequences and suitable scaling, $\mathbb{T}(\mathbf{s})$ converges to $\mathcal{T}$, the Brownian continuum random tree. Similarly, we call $\mathbb{F}(\mathbf{s})$ the random forest with degree sequence $\mathbf{s}$. This is the model we are going to study in this work.

Summary of main results

In this subsection we summarize the main results we obtained to give a better idea for comparison when reading the works of Janson [40] and Pavlov [59], which we are going to introduce in Section 1.4. The introduction of some of the concepts needed to make these statements rigorous and meaningful is postponed to Chapter 2.

In this work we consider a sequence of degree sequences $(\mathbf{s}_n, n \in \mathbb{N})$, where $\mathbf{s}_n = (s_n^{(i)}, i \geq 0)$. We assume $|\mathbf{s}_n| = n$. We only need this assumption for the easiness of notation. In general our results are still valid (if the degree sequences are indexed by κ) as long as $|\mathbf{s}_\kappa| \to \infty$ as $\kappa \to \infty$ and with $|\mathbf{s}_\kappa|$ in place of n.

For any probability distribution $\mathbf{p} = (p^{(i)}, i \geq 0)$ on $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $\mu(\mathbf{p}) = \sum_{i \geq 0} i p^{(i)}$ and $\sigma^2(\mathbf{p}) = \sum_{i \geq 0} i^2 p^{(i)} - 1$. Let $\mathbb{F}_n := \mathbb{F}(\mathbf{s}_n)$ and write $\mathbb{F}_n^{\downarrow} = (\mathbb{T}_{n,l},\ l \geq 1)$. We write $\mathbf{p}_n = (p_n^{(i)}, i \geq 0) := (\frac{s_n^{(i)}}{n}, i \geq 0)$. For $\mathbb{F}_n^{\downarrow} = (\mathbb{T}_{n,l},\ l \geq 1)$, let $\mathcal{T}_{n,l}$ denote the measured rooted real tree

$$\mathcal{T}_{n,l} = \left(\mathbb{T}_{n,l}, \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} d_{gr}, \emptyset_{n,l}, \mu_{n,l} \right)$$

where $\mu_{n,l}$ denotes the uniform measure putting mass $\frac{1}{n}$ on each vertex of $\mathbb{T}_{n,l}$, d_{gr} denotes the graph distance and $\emptyset_{n,l}$ denotes the root of $\mathbb{T}_{n,l}$. Let

$$\mathcal{F}_n^{\downarrow} = (\mathcal{T}_{n,l}, l \geq 1).$$

Let $\Delta_n := \Delta(\mathbf{s}_n) = \max\{i : s_n^{(i)} > 0\}$. Later in the work we sometimes write $c_n := c(\mathbf{s}_n), \sigma_n := \sigma(\mathbf{p}_n)$ for simpler notations.

Now we are ready to state our main theorems in the case of $c(\mathbf{s}_n) = \Theta(n^{1/2})$.

Theorem 1.3.1. *Suppose that there exists a distribution* $\mathbf{p} = (p^{(i)}, i \geq 0)$ *on* $\mathbb{N}_0$ *such that* $\mathbf{p}_n$ *converges to* $\mathbf{p}$ *coordinatewise. Suppose also that* $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p}) \in (0, \infty)$. *If* $\frac{c(\mathbf{s}_n)}{\sigma(\mathbf{p}_n) n^{1/2}} \to \lambda \in (0, \infty)$, *then*

$$\mathcal{F}_n^{\downarrow} \xrightarrow{d} (\mathcal{T}_{\gamma_l}, l \geq 1) \text{ as } n \to \infty, \tag{1.3.1}$$

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

Note that we postpone the rigorous description of Gromov-Hausdorff-Prokhorov convergence and function encodings of real trees (the construction of $\mathcal{T}_{\gamma_l}$) to Chapter 2. The definition of first passage bridge F_λ^{br} will be given in Section 3.1.

Theorem 1.3.2. *Under the conditions of Theorem 1.3.1, suppose additionally that there exists $\epsilon > 0$ such that $\Delta_n = O(n^{\frac{1-\epsilon}{2}})$. Then the convergence (1.3.1) holds in $(\mathbb{L}_\infty, d_{GHP}^\infty)$.*

Note that in proving these two theorems, we also prove the following convergence of the sizes of tree components $(\mathbb{T}_{n,l}, l \geq 1)$. 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 1.3.3. *Under the hypothesises of Theorem 1.3.1, we have*

$$(|\mathbb{T}_{n,l}|/n)_{l \geq 1} \xrightarrow{d} (|\gamma_l|)_{l \geq 1} \tag{1.3.2}$$

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.

Next we state our result in the regime of $c(\mathbf{s}_n) = o(n^{1/2})$. As before we let $\mathbb{F}_n = \mathbb{F}(\mathbf{s}_n)$ be the uniform plane forest with given degree sequence and let $\mathbb{F}_n^\downarrow := (\mathbb{T}_{n,l}, l \geq 1)$ be the decreasing reordering of $\mathbb{F}_n$. Let

$$\mathcal{T}_{n,1} = \left(\mathbb{T}_{n,1}, \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} d_{gr}, \emptyset_{n,1}, \mu_{n,1} \right)$$

be the real tree where $\mu_{n,1}$ denotes the measure putting mass $\frac{1}{n}$ on each vertex of $\mathbb{T}_{n,1}$. For $l \geq 2$, we let

$$\hat{\mathcal{T}}_{n,l} = \left(\mathbb{T}_{n,l}, \frac{\sigma(\mathbf{p}_n)}{2c_n} d_{gr}, \emptyset_{n,l}, \mu_{n,l}\right)$$

be the real tree where $\mu_{n,l}$ denotes the measure putting mass $\frac{1}{c_n^2}$ on each vertex. Let

$$\hat{\mathcal{F}}_n = \left(\hat{\mathcal{T}}_{n,l}, 2 \leq l \leq c_n\right).$$

Let $\mathcal{F}$ be a forest of real trees encoded by excursions of Brownian motion. The detailed construction of this limit $\mathcal{F}$ will be illustrated in Section 4.1. For standard Brownian motion B and $x \geq 0$, let $\tau(x) := \inf(t : B(t) \leq -x)$. We have the following main theorem.

Theorem 1.3.4. *Suppose that there exists a distribution $\mathbf{p} = (p^{(i)}, i \geq 0)$ on $\mathbb{N}_0$ such that $\mathbf{p}_n$ converges to $\mathbf{p}$ coordinatewise. Suppose also that $\sigma(\mathbf{p}_n) \to \sigma := \sigma(\mathbf{p}) \in (0, \infty)$. If $c_n := c(\mathbf{s}_n) = o(n^{1/2})$, then*

$$\left(\mathcal{T}_{n,1}, \hat{\mathcal{F}}_n, \frac{n - |\mathbb{T}_{n,1}|}{c_n^2}\right) \xrightarrow{d} \left(\mathcal{T}, \mathcal{F}, \tau\left(\frac{1}{\sigma}\right)\right)$$

where the first coordinate of the joint convergence is in the GHP sense, the second coordinate is in the sense of coordinatewise GHP convergence, and $\mathcal{T}$ and $\mathcal{F}$ are independent.

1.4 Motivations – conditioned Galton-Watson trees, simply generated trees and related works

In this section, we review the concept of Galton-Watson tree and simply generated tree, both of which are random tree models motivating our work on random forests. Then we present existing results on simply generated random forests related to our work.

1.4.1 Galton-Watson trees

Given any probability distribution $(\pi_i, i \geq 0)$ on $\mathbb{Z}_{\geq 0}$ and let ξ be a random variable with distribution $(\pi_i, i \geq 0)$. Starting from the root, giving each node an independent copy of ξ number of children. The realized tree $\mathcal{T}$ is called the Galton-Watson tree (with *offspring distribution* $(\pi_i, i \geq 0)$ (or ξ)). Let $\mathbf{E}\xi = \sum_{i=0}^{\infty} i\pi_i$. We know that [11] if $\mathbf{E}\xi \leq 1$ (subcritical or critical), then $\mathcal{T}$ is a finite tree with probability 1. And if $\mathbf{E}\xi > 1$ (supercritical), then $\mathcal{T}$ is infinite with positive probability.

The random tree $\mathbb{T}(\mathbf{s})$ is related to Galton-Watson tree by a simple conditioning. Given a degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$, let $\mathcal{T}$ be the Galton-Watson tree with offspring distribution $(\pi_i, i \geq 0)$ where for any $i \geq 0$, $\pi_i > 0$ if $s^{(i)} > 0$. Let $\mathcal{T}'$ be $\mathcal{T}$ conditioned on $\mathbf{s}(\mathcal{T}) = \mathbf{s}$. Then for any fixed tree T with $\mathbf{s}(T) = \mathbf{s}$,

$$\mathbf{P}\left\{\mathcal{T}' = T\right\} = \prod_{i \geq 0} \pi_i^{s^{(i)}},$$

which does not depend on T. Hence $\mathcal{T}'$ is a uniformly random model, which must be $\mathbb{T}(\mathbf{s})$.

1.4.2 Simply generated trees

Simply generated random tree is a random tree model which generalizes the idea of Galton-Watson tree. This model was first introduced in [55]. Fix a *weight sequence* $\mathbf{w} = (w_i, i \geq 0)$ of non-negative real numbers. For a finite tree T, the *weight* of T is:

$$w(T) = \prod_{v \in T} w_{k_T(v)}. \tag{1.4.1}$$

Trees with such weights are called *simply generated trees*. Let $\mathfrak{T}_n$ be the collection of all plane trees with n vertices. Let $\mathcal{T}_n$ be the random tree given by:

$$\mathbf{P}\{\mathcal{T}_n = T\} = \frac{w(T)}{Z_n}, \quad T \in \mathfrak{T}_n$$

where $Z_n = \sum_{T \in \mathfrak{T}_n} w(T)$ is called the *partition function*. We call $\mathcal{T}_n$ the *simply generated random tree* with weight sequence $\mathbf{w}$.

If $\mathbf{w} = (w_k, k \geq 0)$ is a probability distribution, that is, $\sum_{k=0}^{\infty} w_k = 1$, then if we let $\mathcal{T}$ be the Galton-Watson tree with offspring distribution $(w_k, k \geq 0)$, then for any finite tree T,

$$\mathbf{P}\{\mathcal{T} = T\} = w(T) \text{ and } Z_n = \mathbf{P}\{|\mathcal{T}| = n\}.$$

Hence the simply generated random tree with $\mathbf{w}$ is just the Galton-Watson tree $\mathcal{T}$ conditioned to have size n, which we denote as $\mathcal{T}_n$. In fact as pointed out in [40], as long as $\sum_{k=0}^{\infty} w_k z^k < \infty$ for some $z > 0$, the weight sequence $\mathbf{w}$ is equivalent to some probability distribution. Hence simply generated random tree with weight $\mathbf{w}$ gives rise to some conditioned Galton-Watson tree.

In [40] Janson gave a unified treatment of the limit of $\mathcal{T}_n$ and summarized the previous work in detail. In particular, we would like to highlight the main limit theorem for simply generated random trees there.

First, following Section 3 of [40], we collect a few useful pieces of notation. For a fixed weight sequence $\mathbf{w} = (w_k, k \geq 0)$, we let

$$\Phi(z) := \sum_{k=0}^{\infty} w_k z^k$$

be the *generating function* of the given weight sequence and let

$$\rho := 1/\limsup_{k\to\infty} w_k^{1/k} \in [0,\infty]$$

be the *radius of convergence*. For t such that $\Phi(t) < \infty$, define

$$\Psi(t) := \frac{t\Phi'(t)}{\Phi(t)} = \frac{\sum_{k=0}^{\infty} k w_k t^k}{\sum_{k=0}^{\infty} w_k t^k}.$$

$\Psi(t)$ is defined and finite at least for $0 \le t < \rho$ and if $\Phi(\rho) < \infty$, then $\Psi(\rho)$ is still defined with $\Psi(\rho) \le \infty$. Moreover, if $\Phi(\rho) = \infty$, we define $\Psi(\rho) := \lim_{t\nearrow\rho} \Psi(t) \le \infty$. And we write $\nu := \Psi(\rho)$.

Next, following Section 5 of [40], we describe a construction of *modified Galton-Watson tree* $\hat{\mathcal{T}}$, which will be the limit tree of the theorem we are going to present. Again let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and let $(\pi_k, k \ge 0)$ be a probability distribution on $\mathbb{N}_0$ and ξ be a random variable such that $\mathbf{P}\{\xi = k\} = \pi_k,\ k \in \mathbb{N}_0$. We assume that $\mu := \mathbf{E}\xi \le 1$. Let $\hat{\xi}$ be a random variable defined as

$$\mathbf{P}\left\{\hat{\xi} = k\right\} := \begin{cases} k\pi_k, & k = 0,1,2,\cdots \\ 1-\mu, & k = \infty \end{cases}$$

Then a modified Galton-Watson tree is defined as following (as in [43] and [42]). There are two kinds of nodes: *normal* and *special*, with the root being special. For normal nodes, they reproduce with offspring distribution ξ, while for special nodes, they reproduce with offspring distribution $\hat{\xi}$. For normal nodes and special nodes with infinitely many children, their children are all normal. For a special node with finite number of children, one of its uniformly chosen child is special, with all other children being normal. Since each special child has at most one special child, all special nodes form a path from the root, which we call the *spine* of $\hat{\mathcal{T}}$.

There are two cases for $\hat{T}$. When $\mu = 1$, almost surely the special node has finitely many children. Hence $\hat{T}$ has an infinite spine. The tree $\hat{\mathcal{T}}$ is infinite but locally finite. When $\mu < 1$, for each special node, it has probability $1 - \mu > 0$ no special child. So the spine has almost surely finite length L, which follows a geometric distribution $\mathrm{Ge}(1-\mu)$. The spine ends with a special node with infinitely many children, which we call an *explosion*. There are alternative ways of construction $\mathcal{T}$, details of which can be found in Section 5 of [40].

Finally we define a notion of (local) convergence for rooted plane trees. Let $(T_n, n \geq 1)$ and T be rooted plane trees, which by definition are just certain subsets of $\mathcal{U}$. We denote $T_n \to T$ if for each $u \in \mathcal{U}$: if $u \in T$, then $u \in T_n$ for n large enough; conversely, if $u \notin T$, then $u \notin T_n$ for n large enough.

With these notations, definition of $\hat{\mathcal{T}}$ and notion of convergence, we are ready to state one of the main results in [40].

Theorem 1.4.1 (Theorem 7.1 in [40])**.** *Let $\mathbf{w} = (w_k, k \geq 0)$ be any weight sequence with $w_0 > 0$ and $w_k > 0$ for some $k \geq 2$. Let $\mathcal{T}_n$ be the simply generated random tree with weight sequence $\mathbf{w}$.*

(i) If $\nu \geq 1$, let τ be the unique number in $[0, \rho]$ such that $\Psi(\tau) = 1$.

(ii) If $\nu < 1$, let $\tau := \rho$.

In both cases, $0 \leq \tau < \infty$ and $0 < \Phi(\tau) < \infty$. Let

$$\pi_k := \frac{\tau^k w_k}{\Phi(\tau)}, \quad k \geq 0;$$

then $(\pi_k, k \geq 0)$ *is a probability distribution, with expectation*

$$\mu = \Psi(\tau) = \min(\nu, 1) \leq 1$$

and variance $\sigma^2 = \tau\Psi'(\tau) \leq \infty$. *Let* $\hat{\mathcal{T}}$ *be the infinite modified Galton-Watson tree constructed before. Then* $\mathcal{T}_n \xrightarrow{d} \hat{\mathcal{T}}$ *as* $n \to \infty$*, in the topology mentioned before..*

Furthermore, in case (i), $\mu = 1$ *(the critical case) and* $\hat{\mathcal{T}}$ *is locally finite with an infinite spine; in case (ii),* $\mu = \nu < 1$ *(the subcritical case) and* $\hat{\mathcal{T}}$ *has a finite spine ending with an explosion.*

1.4.3 Balls-in-boxes model

In [40], Theorem 1.4.1 is proved by proving limit theorems for a model called *balls-in-boxes* (see Section 11 of [40] for details). In this subsection we define this model and present the limit theorem for this model as in [40].

Given integers $m \geq 0, n \geq 1$, considering the random allocation of m unlabelled balls in n labelled boxes. The set of all possible allocations is

$$\mathcal{B}_{m,n} = \{(y_1, \dots, y_n) \in \mathbb{N}_0^n : \sum_{i=1}^{n} y_i = m\}.$$

Here y_i is the number of balls in box i. Again we fix a weight sequence $\mathbf{w} = (w_k, k \geq 0)$ and define the weight of an allocation $\mathbf{y} = (y_1, \dots, y_n)$ as:

$$w(\mathbf{y}) = \prod_{i=1}^{n} w_{y_i}.$$

The random allocation $B_{m,n}$ is chosen from $\mathcal{B}_{m,n}$ with probability proportional to its weight:

$$\mathbf{P}\{B_{m,n} = \mathbf{y}\} = \frac{w(\mathbf{y})}{Z(m,n)}, \quad \mathbf{y} \in \mathcal{B}_{m,n},$$

where the *partition function* $Z_{m,n} = \sum_{\mathbf{y} \in \mathcal{B}_{m,n}} w(\mathbf{y})$. This $B_{m,n}$ is called the *balls-in-boxes* model.

In [40] the author aims to describe the asymptotic distribution of $B_{m,n}$ as $n \to \infty$. For a weight sequence $\mathbf{w}$, the *support* is $\operatorname{supp}(\mathbf{w}) := \{k : w_k > 0\}$. Let

$$\omega(\mathbf{w}) := \sup \operatorname{supp}(\mathbf{w}) = \sup\{k : w_k > 0\} \leq \infty.$$

Janson considers the case when $m/n \to \lambda$ for some λ. Assume for simplicity that $0 \leq \lambda < \omega = \omega(\mathbf{w})$. For any allocation $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{N}_0^n$ and any $k \geq 0$, let

$$N_k(\mathbf{y}) := |\{i : y_i = k\}|.$$

Since $B_{m,n} = (Y_1, \ldots, Y_n)$ is exchangeable, the distribution of $B_{m,n}$ is completely determined by $N_k(B_{m,n})$ for $k \in \mathbb{N}_0$. The following theorem gives the asymptotic behaviour of $N_k(B_{m,n})$.

Theorem 1.4.2 (Theorem 11.4 in [40])**.** *Let $\mathbf{w} = (w_k, k \geq 0)$ be any weight sequence with $w_0 > 0$ and $w_k > 0$ for some $k \geq 1$. Suppose that $n \to \infty$ and $m = m(n)$ with $m/n \to \lambda$ with $0 \leq \lambda < \omega$.*

(i) If $\lambda \leq \nu$, let τ be the unique number in $[0, \rho]$ such that $\Psi(\tau) = \lambda$.

(ii) If $\lambda > \nu$, let $\tau := \rho$.

In both cases, $0 \leq \tau < \infty$ and $0 < \Phi(\tau) < \infty$. Let

$$\pi_k := \frac{w_k \tau^k}{\Phi(\tau)}, k \geq 0.$$

Then $(\pi_k, k \geq 0)$ *is a probability distribution, with expectation* $\mu = \Psi(\tau) = \min(\lambda, \nu)$ *and variance* $\sigma^2 = \tau\Psi'(\tau) \leq \infty$. *Moreover, for every* $k \geq 0$,

$$N_k(B_{m,n}) \xrightarrow{p} \pi_k.$$

Next we present the connection between this framework developed in [40] and our random forest model $\mathbb{F}(\mathbf{s})$. This is partly explained as Example 12.8 in [40].

Simply generated forests and Galton-Watson forests. A *simply generated forest* is a sequence $(T_1, \ldots, T_n)$ of rooted trees with weight

$$w(T_1, \ldots, T_n) := \prod_{i=1}^{n} w(T_i), \tag{1.4.2}$$

where $w(T_i)$ is given by (1.4.1), for some fixed weight sequence $\mathbf{w}$. A *simply generated random forest* with n trees and m nodes, where $m > n$ are positive integers, is such a forest chosen at random, with probability proportional to its weight. In the special case $n = 1$, this is just the simply generated random tree defined before. In general, for any n, a simply generated random forest $(T_1, \cdots, T_n)$, conditioned on the sizes $(|T_1|, \ldots, |T_n|)$, is a sequence of independent simply generated random trees (all defined by the same weight sequence $\mathbf{w}$). And the sizes $(|T_1|, \ldots, |T_n|)$ is a random allocation $B_{m,n}$ defined by the weight sequence $(Z_k, k \geq 0)$, where Z_k is the partition function for simply generated random trees with weight sequence $\mathbf{w}$. Hence the simply generated random forest can be constructed from a two stage process: simply generated random trees and balls-in-boxes model. For given n, we define a *Galton-Watson forest* with n trees to be the sequence $(\mathcal{T}_1, \ldots, \mathcal{T}_n)$ of n i.i.d. Galton-Watson trees. In the case that the weight sequence $\mathbf{w}$ is a probability distribution, (1.4.2) is the same as the probability distribution of a Galton-Watson forest. Hence in this case the simply generated random forest is the same as a Galton-Watson forest with n trees conditioned on having m vertices.

Our random forest model $\mathbb{F}(\mathbf{s})$ can be viewed as a simply generated random forest (with n nodes and $c(\mathbf{s})$ trees) further conditioned to have degree sequence $\mathbf{s}$. Let $\mathbf{w} = (w_i, i \geq 0)$ be any weight sequence such that $w_i > 0 \iff s^{(i)} > 0$. Since for any forest F with degree sequence $\mathbf{s}$, $w(F) = \prod_{\{i:s^{(i)}>0\}} w_i^{s^{(i)}}$, which does not depend the structure of F. Hence this gives a uniform probability to any forest with degree sequence $\mathbf{s}$.

We present the following theorem about the size of the largest tree in a simply generated forest in [40]. The theorem describes the limit behavior of the size of the largest tree in simply generated random forest with n trees and m vertices. In [40] this served as an example of application of the general framework and tools developed there.

For two sequences of random variables X_n and X_n', we write $X_n \overset{d}{\approx} X_n'$ if there exists a coupling of X_n and X_n' such that $\mathbf{P}\{X_n = X_n'\} \to 1$ as $n \to \infty$. This is equivalent to $d_{\mathrm{TV}}(X_n, X_n') \to 0$ where d_{TV} denotes the *total variation distance*. For any weight sequence $\mathbf{w} = (w_i, i \geq 0)$, denote the *span* of $\mathbf{w}$ by $\mathrm{span}(\mathbf{w})$, that is,

$$\mathrm{span}(\mathbf{w}) = \max\{l \geq 1 : l \mid (i-j) \text{ whenever } w_i, w_j > 0\}.$$

Theorem 1.4.3 (Theorem 19.45 in [40])**.** *Consider a simply generated random forest with n trees and m vertices defined by a weight sequence $\mathbf{w}$, and assume that $m = \lambda n + O(1)$ where $1 < \lambda < \infty$. Suppose that $\nu(\mathbf{w}) \geq 1$ and $\mathrm{span}(\mathbf{w}) = 1$. Let $Y_{(j)}$*

be the size of the $j-th$ largest tree. Define $\tau_1 > 0$ by $\Psi(\tau_1) = 1$, and assume that $\sigma^2 := \tau_1 \Psi'(\tau_1) < \infty$ (which is automatic if $\nu(\mathbf{w}) > 1$). Define further $\tau_2 > 0$ by

$$\Psi(\tau_2) = 1 - 1/\lambda$$

and let

$$q := \frac{\tau_2}{\Phi(\tau_2)} \frac{\Phi(\tau_1)}{\tau_1}.$$

Then $0 < q < 1$ and

$$Y_{(1)} \overset{d}{\approx} \left\lfloor \frac{\log n - \frac{3}{2} \log\log n + \log b + W}{\log(1/q)} \right\rfloor,$$

where W has the Gumbel distribution

$$\mathbf{P}\{W \leq x\} = e^{-e^{-x}}, \quad -\infty < x < \infty$$

and

$$b := \frac{\tau_1 \log^{3/2}(1/q)}{\tau_2} \sqrt{2\pi\sigma^2}(1-q).$$

Furthermore, $Y_{(j)} = Y_{(1)} + O_p(1)$ for each fixed j.

Note that this theorem only applied to the case with n trees and m vertices where $m/n \to \lambda$ for $1 < \lambda < \infty$. In our setting we have c_n trees and n vertices, we deal with the cases $c_n = \Theta(n^{1/2})$ and $c_n = o(n^{1/2})$. In particular $n/c_n \to \infty$. Now we mentioned the work of Pavlov (and others) [60, 23] in the case of $m/n \to \infty$.

1.4.4 Pavlov's work

In [59], Pavlov specifically worked with simply generated random forest with N trees and $N + n$ vertices. Following his notation, he considered the model $\mathfrak{F}_{N,n}$ consisting of N simply generated trees with in total n non-rooted vertices. Pavlov's work addresses a wider range of regimes for the relationship between parameters N and n. In [60], Pavlov studied the limit distribution of the maximum size of a tree in the forest $\mathfrak{F}_{N,n}$. His approach is to exploit the correspondence between $\mathfrak{F}_{N,n}$ and some Galton-Watson branching process beginning with N particles and with offspring distribution $(p_k, k \geq 0)$. Here $p_k = \frac{t^k w_k}{\Phi(t)}$ is a probability weight sequence equivalent to the original weight sequence $\mathbf{w} = (w_k, k \geq 0)$. Let d be the span of $(p_k, k \geq 0)$ and let ξ be a random variable with probability distribution $(p_k, k \geq 0)$ and he assumes $\mathbf{E}\xi = 1, \mathbf{Var}\xi = \sigma^2$. Let η be the size of the largest tree of $\mathfrak{F}_{N,n}$. He proved theorems on distribution of η in different regimes such as (i) $n/N \to b$ for some constant $b > 0$; (ii) $n/N \to \infty, n/N^2 \to 0$; (iii) $\sigma n/N^2 \to \gamma$ for some constant $\gamma > 0$ and (iv) $n/N^2 \to \infty$. Our work in the regime that $c_n = \Theta(n^{1/2})$ and $c_n = o(n^{1/2})$ corresponds to the cases (iii) and (iv) respectively. Hence we are going to only state his results for these two regimes here. In Section 5.3 we will present Pavlov's result in case (ii), which relates to potential future works.

Theorem 1.4.4 (Theorem 2.1.4 in [59])**.** *Let $N, n \to \infty$ in such a way that n takes values which are divided by d, $\sigma n/N^2 \to \gamma$ where $\gamma > 0$ is a constant. Then for any fixed $z > 0$,*

$$\mathbf{P}\left\{\eta/n \leq z\right\} \to \gamma^{3/2} \exp\left(\frac{1}{2\gamma}\right) \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} I_k(\gamma z, \gamma),$$

where

$$I_0(u,v) = \left(v^3 \exp\left(\frac{1}{v}\right)\right)^{-1/2}$$

$$I_k(u,v) = \int_{X_k(u,v)} \frac{\exp(-1/(2(v - x_1 - \cdots - x_k)))dx_1 \cdots dx_k}{(2\pi)^{k/2}(x_1 \cdots x_k(v - x_1 - \cdots - x_k))^{3/2}}$$

$$X_k(u,v) = \{x_i \geq u, i = 1, \ldots, k, x_1 + \cdots + x_k \leq v\}, k = 1, 2, \ldots$$

Theorem 1.4.5 (Theorem 2.1.5 in [59])**.** *Let $n \to \infty$ in such a way that n takes values which are divided by d, $n/N^2 \to \infty$. Then for any fixed $z > 0$,*

$$\mathbf{P}\left\{\sigma \frac{n - \eta}{N^2} \leq z\right\} \to \frac{1}{\sqrt{2\pi}} \int_0^z y^{-3/2} \exp\left(-1/(2y)\right) dy$$

We will see in Section 4.3 that a corollary of Theorem 1.3.4 gives us the same limit probability distribution for size of the largest tree as Theorem 1.4.5. See the remark after Corollary 4.3.2 for details.

Similar problems of the random forest are studied in [23] and the following results on the size of $j-$th largest tree (for any fixed j) are obtained. For the same random forest model $\mathfrak{F}_{N,n}$, let $\nu_{(1)} \leq \nu_{(2)} \leq \ldots \leq \nu_{(N)}$ be the tree sizes listed in increasing order.

Theorem 1.4.6. *Under the conditions of Theorem 1.4.4, for any $h \in \mathbb{N}_0$,*

$$\mathbf{P}\left\{\nu_{(N-h)}/n \leq z\right\} \to \exp\left(\frac{1}{2\gamma}\right) \gamma^{3/2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sum_{g=0}^{h} \frac{1}{g!} I_{k+g}(\gamma z, \gamma).$$

Theorem 1.4.7. *Under the conditions of Theorem 1.4.5, for any $h \in \mathbb{N}$,*

$$\mathbf{P}\left\{\nu_{(N-h)} \leq zN^2\right\} \to \exp\left(-E(z)\right) \sum_{g=0}^{h-1} \frac{E^g(z)}{g!},$$

where $E(z) = \sqrt{2/(\sigma\pi z)}$.

CHAPTER 2
Relevant Concepts and Tools

In this chapter, we introduce the relevant concepts and useful tools for formulating and proving our results.

2.1 Real trees

To study scaling limits of sequences of random trees/forests, we need a continuous version of the notion of trees. Hence we are going to briefly recall the concepts of real trees of their encodings by continuous functions. This concept can be thought as a version of the codings of trees and forests introduced in Section 1.2.2. A more lengthy presentation about the probabilistic aspects of real trees can be found in [32, 45].

Definition 2.1.1. *A compact metric space* (T, d) *is a* real tree *if the following hold for every* $a, b \in T$*:*

(i) There exists a unique shortest path $[[a, b]]$ *from* a *to* b *(of length* $d(a, b)$*), that is, there is a unique isometric map* $f_{a,b}$ *from* $[0, d(a, b)]$ *into* T *such that* $f_{a,b}(0) = a$ *and* $f_{a,b}(d(a, b)) = b$ *(we denote* $f_{a,b}([0, d(a, b)])$ *by* $[[a, b]]$*).*

(ii) The only non-self-intersecting path from a *to* b *is* $[[a, b]]$*, that is, if* q *is a continuous injective map from* $[0, 1]$ *into* T*, such that* $q(0) = a$ *and* $q(1) = b$*, we have* $q([0, 1]) = [[a, b]]$*.*

For a real tree (T, d), an element $v \in T$ is called a *vertex*. A real tree (T, d) is *rooted* if there is a distinguished vertex (the root) $\emptyset \in T$; we denote a rooted real tree by $(T, d, \emptyset)$. The height of a vertex v is $d(\emptyset, v)$, its distance from the root. A *leaf* is a vertex v such that $v \notin [[\emptyset, w]]$ for any vertex $w \neq v$. If there is a finite Borel measure μ on T, then $(T, d, \emptyset, \mu)$ is a *measured* rooted real tree.

Next we show a way of constructing real trees from continuous functions. One can think of the continuous function playing the role of contour process for the real tree. Let $g : [0, \infty) \to [0, \infty)$ be a continuous function with compact support and such that $g(0) = 0$. For every $s, t \geq 0$, let

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

where

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

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

Theorem 2.1.2 (Theorem 2.2 in [45])**.** *(T_g, d_g) is a real tree.*

As pointed out in [45], any rooted real tree $(T, d, \emptyset)$ can be represented in the form T_g for some continuous function g.

To get an intuition of this construction, for a rooted plane tree T with graph distance d_{gr}, let $\hat{\mathrm{T}}$ be the metric space obtained from T by viewing each edge as an isometric copy of the unit interval $[0, 1]$. Let $C(t)$ be the contour function of T defined in Section 1.2.2, then the metric space $\mathcal{T}_C$ constructed from C is isometric to $\hat{\mathrm{T}}$.

The real tree coded by a continuous function g is naturally endowed with a root and a Borel measure, as follows. 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.

Let $e^{(x)}$ denote *Brownian excursion of length* x. Recall that Brownian excursions satisfy the *Brownian scaling* property:

$$(\sqrt{\lambda} e^{(x)}(t/\lambda), 0 \le t \le \lambda x) \stackrel{d}{=} (e^{(\lambda x)}(t), 0 \le t \le \lambda x).$$

Let $\mathbf{e}$ denote the standard Brownian excursion, that is, the Brownian excursion of length 1. Then $\mathcal{T}_{\mathbf{e}}$ is called the *Brownian continuum random tree* (BCRT for short). Note that this definition is used in works such as [45] whereas in the Aldous' work [6, 7, 8], it uses the real tree encoded by *twice* of a standard Brownian motion. The difference of these two definitions are notational and it only leads to a difference of extra factor of two in the scalings. In this work we simply write $\mathcal{T}$ to denote BCRT and use the definition of $\mathcal{T}_{\mathbf{e}}$.

With the preparation work in this section, we are able to rigorously interpret the objects of our study, i.e. $\mathcal{F}_n^{\downarrow}, \mathcal{T}_{n,1}$ and $\hat{\mathcal{F}}_n$ in Theorem 1.3.1, Theorem 1.3.2 and Theorem 1.3.4.

2.2 Gromov-Hausdorff-Prokhorov convergence

To measure the distance between two real trees, or more generally, two rooted measured metric spaces, we need to use the notion of Gromov-Hausdorff-Prokhorov distance. We first recall the definition of the Gromov-Hausdorff distance (see for example Definition 7.3.10 in [21]), used for measuring the distance between two compact metric spaces. The idea is to embed two spaces into a common larger metric space.

For a metric space (Z, d^Z), let d_H^Z be the *Hausdorff distance* between compact subsets of Z, that is, for non-empty subsets A, B of Z,

$$d_H^Z(A, B) = \inf\{\epsilon > 0 : A \subset B^\epsilon, B \subset A^\epsilon\},$$

where A^ϵ is the $\epsilon-$*enlargement* of A:

$$A^\epsilon = \{z \in Z : \inf_{y \in A} d^Z(y, z) < \epsilon\}.$$

Let (X, d) and (X', d') be compact metric spaces. Then the *Gromov-Hausdorff distance* between (X, d) and (X', d') is given by

$$d_{GH}((X, d), (X', d')) = \inf_{\phi, \phi', Z} d_H^Z(\phi(X), \phi'(X')),$$

where the infimum is taken over all isometric embeddings $\phi : X \hookrightarrow Z$ and $\phi' : X' \hookrightarrow Z$ into some common Polish metric space (Z, d^Z).

Note that strictly speaking d_{GH} is not a distance since different compact metric spaces can have GH distance zero.

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 coming definitions in this subsection work in more general settings, e.g. μ could be a boundedly finite Borel measure (see [1]), but for the purpose of this paper, the case of finite measures μ is enough.) For example, the compact measured rooted real tree $\mathcal{T}_g = (T_g, d_g, \emptyset_g, \mathbf{m}_g)$ defined in Section 2.1 is a rooted measured metric space.

Let $\mathcal{X} = (X, d, \emptyset, \mu)$ and $\mathcal{X}' = (X', d', \emptyset', \mu')$ be two compact rooted measured metric spaces. We say $\mathcal{X}$ and $\mathcal{X}'$ are *GHP-isometric* if there exists an isometric one-to-one map $\Phi : X \to X'$ such that $\Phi(\emptyset) = \emptyset'$ and $\Phi_*\mu = \mu'$ where $\Phi_*\mu$ is the *push forward* of measure μ to (X', d'), that is, $\Phi_*\mu(A) = \mu(\Phi^{-1}(A))$ for $A \in \mathcal{B}(X')$. In this case, call Φ a *GHP-isometry*.

Let d_P^Z denotes the *Prokhorov distance* between finite Borel measures on Z, that is, for two finite measures μ, ν on Z,

$$d_P^Z(\mu, \nu) = \inf\{\epsilon > 0 : \mu(A) \le \nu(A^\epsilon) + \epsilon, \nu(A) \le \mu(A^\epsilon) + \epsilon \text{ for any closed set } A\}.$$

If both $\mathcal{X}$ and $\mathcal{X}'$ are compact, then we define the *Gromov-Hausdorff-Prokhorov distance* as:

$$d_{GHP}(\mathcal{X}, \mathcal{X}') = \inf_{\Phi, \Phi', Z} (d^Z(\Phi(\emptyset), \Phi'(\emptyset')) + d_H^Z(\Phi(X), \Phi'(X')) + d_P^Z(\Phi_*\mu, \Phi'_*\mu'))$$

where the infimum is taken over all isometric embeddings $\Phi : X \hookrightarrow Z$ and $\Phi' : X' \hookrightarrow Z$ into some common Polish metric space (Z, d^Z). Let $\mathbb{K}$ denote the set of GHP-isometry classes of compact rooted measured metric spaces. We often identify such a metric space $\mathcal{X}$ with its GHP-isometry class. For example, for the real tree $\mathcal{T}_g$ constructed in Section 2.1, we have $\mathcal{T}_g \in \mathbb{K}$. We have the following results from [1].

Theorem 2.2.1 (Theorem 2.5 in [1]). *The function d_{GHP} defines a metric on $\mathbb{K}$ and the space $(\mathbb{K}, d_{GHP})$ is a Polish metric space.*

We next define a distance between sequences of rooted measured metric spaces. For $\mathbf{X} = (\mathcal{X}_j, j \geq 1), \mathbf{X}' = (\mathcal{X}'_j, j \geq 1)$ in $\mathbb{K}^{\mathbb{N}}$, we let

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

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 the 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 $(\mathbb{L}_{\infty}, d^{\infty}_{GHP})$ is a complete separable metric space.

We next define coordinatewise GHP convergence of sequences of measured metric spaces. For $\mathbf{X}_n = (\mathcal{X}_{n,j}, j \geq 1), \mathbf{X} = (\mathcal{X}_j, j \geq 1)$ in $\mathbb{K}^{\mathbb{N}}$, we say that $\mathbf{X}_n$ converges to $\mathbf{X}$ in *coordinatewise GHP sense* if for any $j \in \mathbb{N}$,

$$\sup_{1 \leq l \leq j} d_{GHP}(\mathcal{X}_{n,l}, \mathcal{X}_l) \to 0 \text{ as } n \to \infty.$$

Now we are able to understand the modes of convergence in our main theorems rigorously.

2.3 Concentration inequalities

Later in Chapter 3 and Chapter 4, in order to prove concentration results for depth-first walk of random forests, we are going to use the results of McDiarmid [54].

In [54] McDiarmid provided an excellent exposition on two methods of proving concentration inequalities and many application of the inequalities to problems of theoretical computer science. The first method is based on the martingale difference sequences; we will use results from this approach in our work. The second method in [54] is the more recent work [67, 68] of Talagrand on the concentration inequality involving *convex distance* on product spaces.

The paper [54] starts with bounding large deviations of the sums of independent binary random variables, which is the setting of Chernoff bound [24], and then present various extensions in this direction, including Hoeffding's work [39]. We are going to use the following tail bounds of sums in Section 3.3.2.

Theorem 2.3.1 (Theorem 2.7 in [54])**.** *Let random variables $X_1^*, \cdots, X_n^*$ be independent, with $X_k^* - \mathbf{E}X_k^* \leq b$ for each k. Let $S_n^* = \sum X_k^*$, and let S_n^* have expected value μ and variance V (the sum of the variances of X_k^*). Then for any $t \geq 0$, with $\epsilon = bt/V$, we have*

$$\mathbf{P}\left\{S_n^* - \mu \geq t\right\} \leq \exp\left(-\frac{V}{b^2}((1+\epsilon)\ln(1+\epsilon) - \epsilon)\right) \leq \exp\left(-\frac{t^2}{2V + 2bt/3}\right).$$

In Section 3 of [54], McDiarmid extends the results in more generality by adopting the martingale framework. In particular, the following theorem extends Theorem 2.3.1; we use this result in Section 3.3.1 and Section 4.4.

Let $\{X_j\}_{j=0}^n$ be a bounded martingale adapted to a filtration $\{\mathcal{F}_j\}_{j=0}^n$. Let $V = \sum_{j=0}^{n-1} \mathbf{Var}\left\{X_{j+1} \mid \mathcal{F}_j\right\}$, where

$$\mathbf{Var}\left\{X_{j+1} \mid \mathcal{F}_j\right\} := \mathbf{E}\left[(X_{j+1} - X_j)^2 \mid \mathcal{F}_j\right] = \mathbf{E}\left[X_{j+1}^2 \mid \mathcal{F}_j\right] - X_j^2.$$

Let

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

Then we have the following bound.

Theorem 2.3.2 ([54], Theorem 3.15)**.** *With v and b defined as above, for any $t \ge 0$,*

$$\mathbf{P}\left\{\max_{0\le j\le n}(X_j - X_0) \ge t\right\} \le \exp\left(-\frac{t^2}{2v(1+bt\backslash(3v))}\right).$$

For completeness we include a proof of Theorem 2.3.2, which is based on that given in [54].

Proof. We first prove a weaker inequality: for any $0 \le k \le n$ and $t \ge 0$,

$$\mathbf{P}\left\{X_k - X_0 \ge t\right\} \le \exp\left(-\frac{t^2}{2v(1+bt\backslash(3v))}\right). \tag{2.3.1}$$

Define a function g by $g(x) = \frac{e^x-1-x}{x^2}$. Then $g'(x) = \frac{(x-2)e^x+x+2}{x^3}$. By taking derivatives of $\tilde{g}(x) := (x-2)e^x + x + 2$ we see that $\tilde{g}(x) < 0$ for $x < 0$ and $\tilde{g}(x) > 0$ for $x > 0$. Hence g is increasing and in particular for $x \le b$, we have

$$e^x = 1 + x + x^2 g(x) \le 1 + x + x^2 g(b).$$

Since ess $\sup(X_{j+1} - X_j \mid \mathcal{F}_j) \le b$, for any $h > 0$ and $0 \le j < n$, we have

$$\begin{aligned}
\mathbf{E}\left[e^{h(X_{j+1}-X_j)} \mid \mathcal{F}_j\right] &\le \mathbf{E}\left[1 + h(X_{j+1}-X_j) + h^2(X_{j+1}-X_j)^2 g(hb) \mid \mathcal{F}_j\right] \\
&= 1 + h^2 g(hb)\mathbf{Var}\left\{X_{j+1} - X_j \mid \mathcal{F}_j\right\} \\
&= 1 + h^2 g(hb)\mathbf{Var}\left\{X_{j+1} \mid \mathcal{F}_j\right\} \\
&\le \exp\left(h^2 g(hb)\mathbf{Var}\left\{X_{j+1} \mid \mathcal{F}_j\right\}\right).
\end{aligned} \tag{2.3.2}$$

Hence for any $1 \leq j \leq k$, using tower law and the fact $e^{h(X_{j-1}-X_0)}$ is $\mathcal{F}_{j-1}-$measurable, we have

$$\begin{aligned} \mathbf{E}\left[e^{h(X_j - X_0)}\right] &= \mathbf{E}\left[\mathbf{E}\left[e^{h(X_j-X_0)} \mid \mathcal{F}_{j-1}\right]\right] \\ &= \mathbf{E}\left[e^{h(X_{j-1}-X_0)}\mathbf{E}\left[e^{h(X_j-X_{j-1})} \mid \mathcal{F}_{j-1}\right]\right] \\ &\leq \exp\left(h^2 g(hb)\mathbf{Var}\left\{X_j \mid \mathcal{F}_{j-1}\right\}\right)\mathbf{E}\left[e^{h(X_{j-1}-X_0)}\right], \qquad (2.3.3) \end{aligned}$$

where we use (2.3.2) in the last line. By Markov's inequality and using (2.3.3) recursively for $j = k, \ldots, 1$, it follows that for any $h > 0$,

$$\begin{aligned} \mathbf{P}\left\{X_k - X_0 \geq t\right\} &\leq e^{-ht}\mathbf{E}\left[e^{h(X_k-X_0)}\right] \\ &\leq e^{-ht}\prod_{j=1}^{k}\exp\left(h^2 g(hb)\mathbf{Var}\left\{X_j \mid \mathcal{F}_{j-1}\right\}\right) \\ &= e^{-ht+g(hb)h^2 v}, \end{aligned}$$

which is minimized at $h = \frac{1}{b}\ln\left(1+\frac{bt}{v}\right)$. This gives us the bound

$$\mathbf{P}\left\{X_k - X_0 \geq t\right\} \leq \exp\left(-\frac{v}{b^2}((1+\epsilon)\ln(1+\epsilon)-\epsilon)\right)$$

where $\epsilon = bt/v$. It is straightforward to show that, for all $x \geq 0$,

$$f(x) := (6+8x+2x^2)\ln(1+x) - 6x - 5x^2 \geq 0.$$

This implies that, for all $x \geq 0$,

$$(1+x)\ln(1+x) - x \geq 3x^2/(6+2x),$$

which gives our desired bound in (2.3.1).

Next, for any $h > 0$, let $T_j := e^{h(X_j - X_0)}$. Then T_j is a $\mathcal{F}_j$–submartingale. By Doob's submartingale inequality (e.g. section 14.6 in [70]), we have for any $k \leq n$,

$$\mathbf{P}\left\{\max_{0\leq j\leq k}(X_j - X_0) \geq t\right\} = \mathbf{P}\left\{\max_{0\leq j\leq k} T_j \geq e^{ht}\right\} \leq e^{-ht}\mathbf{E}T_k = e^{-ht}\mathbf{E}\left[e^{h(X_k - X_0)}\right].$$

Since (2.3.1) is proved by bounding the quantity $\mathbf{E}\left[e^{h(X_k - X_0)}\right]$, we obtain the same bound for the maximum. $\square$

CHAPTER 3
Critical case

3.1 Introduction

In this chapter, we aim to prove Theorem 1.3.1 and Theorem 1.3.2. These theorems tell 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 understand these results rigorously, we need to first introduce the concept of first passage bridge. This chapter is essentially tailored from the manuscript of [47].

First passage bridge

Recall the following definition of *first passage bridge* as in [14]. Informally, for $\lambda > 0$, the *first passage bridge* of unit length from 0 to $-\lambda$, denoted F_λ^{br}, is a $C[0,1]-$valued random variable with law

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

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

For $l \geq 0$, we write B_l^{br} for the Brownian bridge of duration 1 from 0 to $-l$. As explained in Proposition 1 of [33], the law of the Brownian bridge B_l^{br} is characterized by $B_l^{br}(1) = -l$ and the formula

$$\mathbf{E}\left[f((B_l^{br}(t))_{0\le t\le m})\right] = \mathbf{E}\left[f((B(t))_{0\le t\le m})\frac{p_{1-m}(-l-B(m))}{p_1(-l)}\right] \tag{3.1.1}$$

for all bounded measurable function f, and all $0 \le m < 1$, 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_λ^{br} can be defined as the law such that

$$\mathbf{E}\left[f((F_\lambda^{br}(t))_{0\le t\le s})\right] = \mathbf{E}\left[(f(B(t))_{0\le t\le s})\frac{p'_{1-s}(-\lambda-B(s))}{p'_1(-\lambda)}\mathbb{1}_{\{\inf\limits_{r\le s} B(r)>-\lambda\}}\right] \tag{3.1.2}$$

for all bounded measurable functions f and all $0 \le s < 1$ and $F_\lambda^{br}(1) = -\lambda$, where p'_a is the derivative of p_a. These formulae set the finite-dimensional laws of the first passage bridge. In [15] (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 as $\epsilon \to 0$, where $(F_\lambda^\epsilon(t), 0 \le t \le 1)$ has the law of B conditioned on the event $\{B(1) < -\lambda+\epsilon,\ \inf\limits_{s\le 1} B(s) > -\lambda-\epsilon\}$, hence justifying the informal conditioning definition.

We first recall Theorem 1.3.1 and Theorem 1.3.2 below.

Theorem 1.3.1. *Suppose that there exists a distribution $\mathbf{p} = (p^{(i)}, i \ge 0)$ on $\mathbb{N}_0$ such that $\mathbf{p}_n$ converges to $\mathbf{p}$ coordinatewise. Suppose also that $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p}) \in (0,\infty)$. If $\frac{c(\mathbf{s}_n)}{\sigma(\mathbf{p}_n)n^{1/2}} \to \lambda \in (0,\infty)$, then*

$$\mathcal{F}_n^\downarrow \xrightarrow{d} (\mathcal{T}_{\gamma_l}, l \ge 1) \text{ as } n \to \infty, \tag{1.3.1}$$

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

Theorem 1.3.2. *Under the conditions of Theorem 1.3.1, suppose additionally that there exists* $\epsilon > 0$ *such that* $\Delta_n = O(n^{\frac{1-\epsilon}{2}})$. *Then the convergence (1.3.1) holds in* $(\mathbb{L}_\infty, d_{GHP}^\infty)$.

We would like to first make the following two comments.

Remark 3.1.1. *The assumptions of Theorem 1.3.1 imply that* $\mu(\mathbf{p}_n) \to \mu(\mathbf{p}) = 1$ *and that* $\Delta_n = o(n^{1/2})$. *We include the proof of these facts as Lemma 3.6.1 in Section 3.6.*

For c > 0, let c$\mathbf{e} \in C[0,\infty)$ denote the Brownian excursion of length c, that is $e^{(\mathrm{c})}$, as we defined in Section 2.1. For an excursion γ, let $|\gamma|$ be its length (please refer to Section 3.1.1 for a careful definition).

Remark 3.1.2. *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* $(\gamma_l, l \geq 1)$. *This is true because of the Brownian scaling property of Brownian excursion.*

3.1.1 Key ingredients of the paper

Here we summarize the two key ingredients of this chapter. The first element is the convergence of the large trees in (1.3.1), which is essentially given by the following proposition. We first state this result and then give a careful introduction of the concepts involved. For all $l \geq 1$, let $X_{n,l} = \frac{|\mathbb{T}_{n,l}|}{n}$.

Proposition 3.1.1. *Under the conditions of Theorem 1.3.1, for any fixed* $j \geq 1$,

$$((X_{n,l})_{l \leq j}, (\mathcal{T}_{n,l})_{l \leq j}) \xrightarrow{d} ((|\gamma_l|)_{l \leq j}, (\mathcal{T}_{|\gamma_l|\mathbf{e}_l})_{l \leq j}) \tag{3.1.3}$$

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

There are two parts of the convergence in (3.1.3). One is the convergence of the normalized sizes of large trees to lengths of excursions. This is given by Proposition 1.3.3, which we recall below. We first introduce more notions which explain the proposition more rigorously. 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 $g(s) - \min_{0\le s'<s} g(s')$ to denote $(g(s) - \min_{0\le s'<s} g(s'),\ 0 \le s \le 1)$. For $g \in \mathcal{C}_0(1)$, sometimes we refer the excursions of $g(s) - \min_{0\le s'<s} g(s')$ as excursions of g. Recall that

$$l_1^\downarrow = \{x = (x_1, x_2, \cdots) : x_1 \ge x_2 \ge \cdots \ge 0, \sum_i x_i \le 1\}$$

and endow $l_1^\downarrow$ with the topology induced by the l_1 distance: $d(x,y) = \sum_i |x_i - y_i|$.

Proposition 1.3.3. *Under the hypothesises of Theorem 1.3.1, we have*

$$(|\mathbb{T}_{n,l}|/n)_{l\ge 1} \xrightarrow{d} (|\gamma_l|)_{l\ge 1} \tag{1.3.2}$$

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

This proposition will be a corollary of the following theorem, which is the main result of Section 3.4. For a plane forest F, let $u_1 < u_2 < \cdots < u_{|\mathrm{F}|}$ be the nodes of F listed according to their lexicographic order in $\mathcal{U}$ in each tree component, with nodes of first tree listed first, then the nodes of second tree and so on. Recall in Section 1.2.2, 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 3.1.2. *Under the conditions of Theorem 1.3.1, we have*

$$\left(\frac{S_{\mathbb{F}_n}(tn)}{\sigma(\mathbf{p}_n) n^{1/2}} \right)_{t \in [0,1]} \xrightarrow{d} F_\lambda^{br} \tag{3.1.4}$$

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

The second part of the convergence of (3.1.3) is the convergence of the large trees, for which we will rely on the following result about random trees with given degree sequences from [20].

Theorem 3.1.3 (Theorem 1 in [20])**.** *Let* $\{\mathbf{s}_n, n \geq 1\}$ *be a degree sequence such that* $|\mathbf{s}_n| = n \to \infty, \Delta_n := \Delta(\mathbf{s}_n) = o(n^{1/2})$*. Suppose that there exists a distribution* $\mathbf{p} = (p^{(i)}, i \geq 0)$ *on* $\mathbb{N}$ *with mean 1 such that* $\mathbf{p}_n = (s_n^{(i)}/n, i \geq 0)$ *converges to* $\mathbf{p}$ *coordinatewise and such that* $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p}) \in (0, \infty)$*. Let* $\mathbb{T}_n$ *be the random plane tree under* $\mathbb{P}_{\mathbf{s}_n}$*, the uniform measure on the set of plane trees with degree sequence* $\mathbf{s}_n$*. Let* $\mathcal{T}_n$ *denote the measured rooted metric space* $(\mathbb{T}_n, \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} d_{gr}, \emptyset_n, \mu_n)$ *where* μ_n *denotes the uniform measure putting mass* $\frac{1}{n}$ *on each vertex of* $\mathbb{T}_n$*. Then when* $n \to \infty, \mathcal{T}_n \xrightarrow{d} \mathcal{T}$ *in the Gromov-Hausdorff-Prokhorov sense, where* $\mathcal{T}$ *is BCRT.*

Remark 3.1.3. *In fact Theorem 1 in [20] is only stated in the Gromov-Hausdorff sense, that is,* $(\mathbb{T}_n, \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} d_{gr}, \emptyset_n) \xrightarrow{d} (\mathcal{T}_{\mathbf{e}}, d_{\mathbf{e}}, \emptyset_{\mathbf{e}})$*. But the conclusion can be strengthened to GHP convergence easily. For completeness, we include a proof of this fact in Section 3.6.*

The following proposition contains the additional ingredient required to prove Theorem 1.3.2.

Proposition 3.1.4. *Under the conditions of Theorem 1.3.2, for all* $a > 0$, *we have*

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

The key results leading to Proposition 3.1.4 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 3.1.5 (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 3.1.5 to prove Proposition 3.1.4.

Proposition 3.1.6. *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 $1 \le i \le n$, and let $S_j = \sum_{i \le j} C_i^2$ for $1 \le i \le n$. Then for all $\lambda \ge 2$ and $1 \le k \le n$, with $\Delta = \max_{1 \le i \le n} C_i = \max_{1 \le i \le n} c_i$, and $\sigma^2(c) = \sum_{i \le n} c_i^2 = S_n$, we have*

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

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

Proof of Theorem 1.3.1 and Theorem 1.3.2. By Skorokhod's representation theorem, we may work in a probability space in which the convergence in Proposition 3.1.1 is almost sure. Hence Proposition 3.1.1 yields that for any fixed j, $\sup_{l \le j} d_{GHP}(\mathcal{T}_{n,l}, \mathcal{T}_{|\gamma_l|\mathbf{e}_l}) \xrightarrow{d} 0$. This establishes Theorem 1.3.1. Now to prove the convergence in $(\mathbb{L}_\infty, d_{GHP}^\infty)$, it suffices to prove that for any $a > 0$,

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

It suffices to separately prove

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

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

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 n,

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

For the diameter we resort to Proposition 3.1.4.

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

For $\mathrm{diam}(\mathcal{T}_{\gamma_l})$, $\mathrm{diam}(\mathcal{T}_{\gamma_l}) \leq 2h(\mathcal{T}_{\gamma_l}) = 2\max(\gamma_l)$. 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

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

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

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

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. $\square$

To conclude this section, we sketch how this chapter is organized. In Section 3.2 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 3.1.6 to something similar which is applicable to random forests. This section will be purely combinatorial and only deal with fixed degree sequences. In Section 3.3, we collect some concentration results using martingale methods. These probability bounds will be useful for checking that the assumptions in Theorem 3.1.3 are satisfied for large trees of $\mathcal{F}_n^{\downarrow}$. The second part of this section proves the variance bound in Proposition 3.1.6. Again all results in this section is non-asymptotic and hence are presented with regards to a fixed degree sequence. In Section 3.4, we prove Theorem 3.1.2, the convergence of scaled exploration processes to some random process related to first passage bridge, using the rotation mapping in Section 3.2. We will then get Proposition 1.3.3 as a corollary from this weak convergence result. In Section 3.5 we finish the proof of Proposition 3.1.1 and Proposition 3.1.4 using results from Section 3.3 and Section 3.4. Finally, we prove Remark 3.1.1 and Remark 3.1.3 in Section 3.6.

3.2 An $|\mathbf{s}|-$to-1 map transforming lattice bridge to first passage lattice bridge

Given a degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$, let $d(\mathbf{s}) \in \mathbb{Z}_{\geq 0}^{|\mathbf{s}|}$ 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

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

and call $\Lambda(\mathbf{s})$ the set of *lattice bridges corresponding to* $\mathbf{s}$. Note that if $b \in \Lambda(\mathbf{s})$, then $b(|\mathbf{s}|) = -c(\mathbf{s})$. Furthermore, we have

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

since to determine $b \in \Lambda(\mathbf{s})$, it suffices to choose the $s^{(0)}$ positions with step size -1, $s^{(1)}$ positions with step size 0, $s^{(2)}$ positions with step size 1, etc.

We then let

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

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 $y < \min\limits_{u \in [0,s]} x(u)$ 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 3.2.1. *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 $u < |\mathbf{s}|$. We shall prove that $\theta_{t(i)}(b)(u) > -c(\mathbf{s})$, which proves the lemma. If $0 \leq u \leq |\mathbf{s}|-t(i)$, then $\theta_{t(i)}(b)(u) = b(t(i)+u)-b(t(i)) \geq m-i > -c(\mathbf{s})$. If $|\mathbf{s}|-t(i) \leq u < |\mathbf{s}|$, then $\theta_{t(i)}(b)(u) = b(t(i)+u-|\mathbf{s}|)+b(|\mathbf{s}|)-b(t(i)) = b(t(i)+u-|\mathbf{s}|)-c(\mathbf{s})-i$. Since $u < |\mathbf{s}|$, $t(i)+u-|\mathbf{s}| < t(i)$ and we must have $b(t(i)+u-|\mathbf{s}|) > i$ 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 3.2.2. *f is an $|\mathbf{s}|$−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 $|\mathbf{s}|$, 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 $[|\mathbf{s}|]$. 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 $|\mathbf{s}|$ preimages in $\Lambda(\mathbf{s}) \times ([c(\mathbf{s})]-1)$. On the other hand, we have (see, e.g., [61], page 128)

$$|F(\mathbf{s})| = \frac{c(\mathbf{s})}{|\mathbf{s}|}\binom{|\mathbf{s}|}{(s^{(i)}, i \geq 0)} = \frac{c(\mathbf{s})}{|\mathbf{s}|}\frac{|\mathbf{s}|!}{\prod\limits_{i\geq 0} s^{(i)}!}. \tag{3.2.1}$$

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

Recall the concept of depth-first walk S_{F} of a plane forest F. For a sequence $\mathbf{c} = (c_1, \cdots, c_n) \in \mathbb{R}^n$, we 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 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 $\pi(\mathbf{c}) = (c_{\pi(1)}, \cdots, c_{\pi(n)})$. Also, recall from the beginning of this section that for a degree sequence $\mathbf{s}$, $d(\mathbf{s})$ is a vector with $s^{(i)}$ entries equal to i for each $i \geq 0$.

Corollary 3.2.3. *Let* $\mathbf{s}$ *be a degree sequence. Let* π *be a uniformly random permutation of* $[|\mathbf{s}|]$ *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 3.2.2, 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 3.3.4, 3.3.7 and 3.3.8).

A *marked forest* is a pair (F, v) where F is a plane forest and $v \in v(F)$. Sometimes we refer v as 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 $v_1, \cdots, v_{|\mathbf{s}|}$ and $v = v_i$, then $g((F, v)) = (k_F(v_i), \cdots, k_F(v_{|\mathbf{s}|}), k_F(v_1), \cdots, k_F(v_{i-1}))$. 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 3.2.4. *g is a $c(\mathbf{s})-$to-1 surjective map and for each $1 \leq i \leq c(\mathbf{s})$, $g^i := g|_{\mathrm{MF}^i(\mathbf{s})}$ is a bijection between $\mathrm{MF}^i(\mathbf{s})$ and $\mathrm{D}(\mathbf{s})$. Also, h is a $|\mathbf{s}|-$to-1 surjective map.*

Proof. For $d \in \mathrm{D}(\mathbf{s})$, $|g^{-1}(\{d\}) \cap \mathrm{MF}^i(\mathbf{s})| = 1$ for all $1 \leq i \leq c(\mathbf{s})$. 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. $\square$

The map g being surjective immediately gives the following result.

Corollary 3.2.5. *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 $\mathrm{D}(\mathbf{s})$.*

3.3 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 3.1.1 in Section 3.5 where we need to first show that the degree proportions in each large trees of $\mathcal{F}_n^{\downarrow}$ 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 [54].

Let $\mathbf{s} = (s^{(i)}, i \geq 0)$ 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 3.2), where π is a uniform random permutation of $[n]$. For each $i \geq 0$, let $q^{(i)} = s^{(i)}/n$ be the degree proportion of degree i of $\mathbf{s}$.

3.3.1 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 Theorem 2.3.2 as tools to prove Lemma 3.3.3 and Proposition 3.3.4, which are useful for eventually proving that the empirical degree distributions of large trees of $\mathbb{F}_n$ behave well (Proposition 3.5.1).

For fixed i, for $0 \leq j \leq n-1$, let $Y_j^{(i)} = |\{1 \leq l \leq j : C_l = i\}|$ 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 3.3.1. *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*

$\mathbf{Var}\left\{M_{j+1}^{(i)} \mid \mathcal{F}_j\right\} := \mathbf{E}\left[{M_{j+1}^{(i)}}^2 \mid \mathcal{F}_j\right] - {M_j^{(i)}}^2 \leq \frac{1}{4}\frac{1}{(n-(j+1))^2}.$

Proof. (a) Since $q^{(i)}$ is a constant, it suffices to show that $\frac{X_j^{(i)}}{n-j}$ is an $\mathcal{F}_j-$martingale. In fact

$$\begin{aligned}
\mathbf{E}\left[X_{j+1}^{(i)} \mid \mathcal{F}_j\right] &= X_j^{(i)} - \mathbf{P}\left\{C_{j+1} = i \mid \mathcal{F}_j\right\} \\
&= X_j^{(i)} - \frac{X_j^{(i)}}{n-j},
\end{aligned}$$

so

$$\mathbf{E}\left[\frac{X_{j+1}^{(i)}}{n-(j+1)} \mid \mathcal{F}_j\right] = \frac{X_j^{(i)}}{n-j-1}(1 - \frac{1}{n-j}) = \frac{X_j^{(i)}}{n-j}.$$

Thus $\frac{X_j^{(i)}}{n-j}$ is an $\mathcal{F}_j-$martingale.

(b) By definition, we have

$$\begin{aligned}
\mathbf{Var}\left\{M_{j+1}^{(i)} \mid \mathcal{F}_j^{(i)}\right\} &= \mathbf{E}\left[{M_{j+1}^{(i)}}^2 \mid \mathcal{F}_j\right] - {M_j^{(i)}}^2 \\
&= \frac{\mathbf{E}\left[{X_{j+1}^{(i)}}^2 \mid \mathcal{F}_j\right]}{(n-(j+1))^2} - \frac{{X_j^{(i)}}^2}{(n-j)^2}
\end{aligned}$$

Now we substitute

$$\mathbf{E}\left[{X_{j+1}^{(i)}}^2 \mid \mathcal{F}_j\right] = (X_j^{(i)}-1)^2\frac{X_j^{(i)}}{n-j} + {X_j^{(i)}}^2 \cdot \frac{n-j-X_j^{(i)}}{n-j} = {X_j^{(i)}}^2 - \frac{2{X_j^{(i)}}^2}{n-j} + \frac{X_j^{(i)}}{n-j}$$

in the above result and obtain

$$\begin{aligned}
\mathbf{Var}\left\{M_{j+1}^{(i)} \mid \mathcal{F}_j\right\} &= \frac{{X_j^{(i)}}^2}{(n-(j+1))^2} - \frac{{X_j^{(i)}}^2}{(n-j)^2} - \frac{2{X_j^{(i)}}^2}{(n-j)(n-(j+1))^2} + \frac{X_j^{(i)}}{(n-j)(n-(j+1))^2} \\
&= \frac{(2(n-j)-1){X_j^{(i)}}^2}{(n-(j+1))^2(n-j)^2} - \frac{2{X_j^{(i)}}^2}{(n-j)(n-(j+1))^2} + \frac{X_j^{(i)}}{(n-j)(n-(j+1))^2} \\
&= \frac{X_j^{(i)}(n-j-X_j^{(i)})}{(n-(j+1))^2(n-j)^2} \le \frac{1}{4} \cdot \frac{1}{(n-(j+1))^2},
\end{aligned}$$

which gives the claim. $\square$

Now we can apply Theorem 2.3.2.

Proposition 3.3.2. *For any $t > 0$ and $0 < s < n$, we have*

$$\mathbf{P}\left\{\max_{0\le j\le n-s} |q^{(i)} - \frac{X_j^{(i)}}{n-j}| \ge t\right\} \le \exp\left(-\frac{3st^2}{3+2t}\right). \tag{3.3.1}$$

Proof. Fix $s < n$, and consider the martingale $\{M_j^{(i)}\}_{j=0}^{n-s}$. By Lemma 3.3.1(b), we know that

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

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

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

as claimed. $\square$

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

Lemma 3.3.3. *For fixed $i \in \mathbb{N}$ and $\epsilon > 0$, let*

$$B^{\epsilon,i} = \{\exists x \geq \log^3 n : |Y_x^{(i)} - q^{(i)}x| \geq \epsilon x\}.$$

Then for any n large enough such that $\frac{\sqrt{5}}{\log n} < \epsilon < 1$,

$$\mathbf{P}\left\{B^{\epsilon,i}\right\} \leq n^{-3}.$$

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

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

Taking $s = \log^3 n, t = \epsilon$ in (3.3.1), the result follows. □

Now we consider how degrees distribute among the tree components of the random forest $\mathbb{F}(\mathbf{s})$. Write $\mathbb{F}(\mathbf{s})^{\downarrow} = (\mathbb{T}_l,\ l \geq 1)$. Let $\mathbf{s}_l = (s_l^{(i)}, i \geq 0)$ denote the (empirical) degree sequence of the $l-$th largest tree $\mathbb{T}_l$. Recall that $q^{(i)} = s^{(i)}/n$ and let $q_l^{(i)} = s_l^{(i)}/|\mathbf{s}_l|$ 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 3.3.4. *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*

$$\mathbf{P}\left\{\bigcup_{l:\ |\mathbb{T}_l| > n^{1/4}} B_l^{\epsilon,i}\right\} \leq n\mathbf{P}\left\{B^{\epsilon,i}\right\}. \tag{3.3.2}$$

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 $i \leq n$ let C_i be the degree of V_i. By Corollary 3.2.5, 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. $\square$

3.3.2 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 3.1.4 in Section 3.5. 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 Theorem 2.3.1. Since $C_1, C_2, \cdots, C_k$ are sampled without replacement from the population $c_1, c_2, \cdots, c_n$, we may not directly apply Theorem 2.3.1. We address this issue as follows.

Recall (or see, e.g., [10]) that given real random variables U, V, we say U is a *dilation* of V if there exist random variables $\hat{U}, \hat{V}$ such that

$$\hat{U} \stackrel{d}{=} U, \ \hat{V} \stackrel{d}{=} V \text{ and } \mathbf{E}\left[\hat{U}|\hat{V}\right] = \hat{V}.$$

Proposition 3.3.5 (Proposition 20.6 in [10]). *Suppose $X_1, \cdots, X_k$ and $X_1^*, \cdots, X_k^*$ are samples from the same finite population $x_1, \cdots, x_n$, without replacement and with replacement, respectively. Let $S_k = \sum_{i=1}^{k} X_i, S_k^* = \sum_{i=1}^{k} X_i^*$. Then S_k^* is a dilation of S_k. In particular, $\mathbf{E}[\phi(S_k^*)] \geq \mathbf{E}[\phi(S_k)]$ for all continuous convex function $\phi : \mathbb{R} \to \mathbb{R}$.*

The proof of Theorem 2.3.1, in [54] (which is similar to the proof of Theorem 2.3.2, which we included in Section 2.3), proceeds by bounding the quantity $\mathbf{E}[\exp(h(S_n^* - \mu))]$, where h is any real number. By Proposition 3.3.5, we have $\mathbf{E}[\exp(h(S_n - \mu))] \leq \mathbf{E}[\exp(h(S_n^* - \mu)]$, which means that the proof applies mutatis mutandis in the setting of sampling without replacement.

Corollary 3.3.6. *Let $X_1, \cdots, X_k$ be samples from finite population $x_1, \cdots, x_n$, without replacement, with $X_1 - \mathbf{E}X_1 \le b$. Let $S_k = \sum_{i=1}^{k} X_i, V = \sum_{i=1}^{k} \mathbf{Var}X_i$ and $\mu_k = \mathbf{E}S_k$. Then for any $t \ge 0$, with $\epsilon = bt/V$, we have*

$$\mathbf{P}\left\{S_k - \mu_k \ge t\right\} \le \exp\left(-\frac{V}{b^2}((1+\epsilon)\ln(1+\epsilon) - \epsilon)\right) \le \exp\left(-\frac{t^2}{2V + 2bt/3}\right). \tag{3.3.3}$$

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

Proof of Proposition 3.1.6. We apply (3.3.3); we have $\mu_k = \mathbf{E}S_k = \frac{k}{n}S_n, b = \Delta^2$,

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

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 3.2, 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 3.3.7. *Let* $\mathbf{s} = (s^{(i)}, i \geq 0)$ *be a degree sequence with* $|\mathbf{s}| = n$ *and* $M = \sigma^2(\mathbf{s})/n$*. Then for* $\lambda \geq 4, \alpha > \Delta^2(\mathbf{s})/n$*,*

$$\mathbf{P}\left\{\exists T \in \mathbb{F}(\mathbf{s}) : |T| \leq \alpha n, \sigma^2(T) \geq \lambda\alpha\sigma^2(\mathbf{s})\right\} \leq \frac{2}{\alpha}\exp(-\frac{3M}{16}\lambda). \tag{3.3.4}$$

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 $i \leq n$ let C_i be the degree of V_i. By Corollary 3.2.5, 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 $0 \leq j \leq \lfloor\frac{1}{\alpha}\rfloor$, let B_j be the event that

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

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

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

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

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

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

$$(1+\lfloor\frac{1}{\alpha}\rfloor)\mathbf{P}\left\{B_0\right\} \leq \frac{2}{\alpha}\mathbf{P}\left\{S_{\lfloor 2\alpha n\rfloor} \geq \lambda\alpha\sigma^2(\mathbf{s})\right\} \leq \frac{2}{\alpha}\mathbf{P}\left\{S_{\lfloor 2\alpha n\rfloor} \geq \frac{\lambda}{2}\cdot\frac{\lfloor 2\alpha n\rfloor}{n}\sigma^2(\mathbf{s})\right\}$$
$$\leq \frac{2}{\alpha}\exp\left(-\frac{3M}{16}\lambda\right),$$

where we take $k = \lfloor 2\alpha n\rfloor$ in Proposition 3.1.6 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 3.3.8. *$\forall\ \epsilon,\rho \in (0,1), \exists n_0 = n_0(\epsilon) \in \mathbb{N}$ and $\beta_0 > 0$ such that the following is true. Let $\mathbf{s}$ be any degree sequence with $|\mathbf{s}| = n \geq n_0$. Suppose that $\Delta(\mathbf{s}) \leq n^{\frac{1-\epsilon}{2}}, s^{(1)} \leq (1-\epsilon)|\mathbf{s}|$ and $\epsilon \leq \sigma^2(\mathbf{s})/n \leq 1/\epsilon$, then for any $0 < \beta < \beta_0$,*

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

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 $|T| \leq n^{1-\epsilon}$ 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 $\sigma^2(T) \leq (\frac{|T|}{n})^{1/2}\sigma^2(\mathbf{s})$ such that $h(T) > \delta n^{1/2}$.
- E_4 is the event that there exists a tree T with $|T| \leq n^{1-\epsilon}$ and $\sigma^2(T) \leq n^{1-\frac{\epsilon}{2}}$ 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 $\mathbf{P}\left\{E_1\right\} + \mathbf{P}\left\{E_2\right\} + \mathbf{P}\left\{E_3\right\} + \mathbf{P}\left\{E_4\right\}$. 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 3.3.7.

$$\begin{aligned}
\mathbf{P}\{E_1\} &\leq \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}\left\{\exists T \in \mathbb{F}(\mathbf{s}) : |T| \in \left[\frac{\beta n}{2^{k+1}}, \frac{\beta n}{2^k}\right], \sigma^2(T) > \left(\frac{\beta}{2^{k+1}}\right)^{1/2} \sigma^2(\mathbf{s})\right\} \\
&\leq \sum_{k\geq 0} \frac{2^{k+1}}{\beta} \exp\left(-\frac{3\sigma^2(\mathbf{s})}{16n} \frac{2^{\frac{k-1}{2}}}{\beta^{1/2}}\right) \\
&= O\left(\frac{1}{\beta}\exp(-\frac{\epsilon}{\beta^{1/2}})\right) \qquad (3.3.5)
\end{aligned}$$

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

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 3.2.5, with similar argument as in proof of Proposition 3.3.7, 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 3.1.6 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

$$\mathbf{P}\{E_2\} \leq n^{2-\epsilon} \exp\left(-\frac{3}{16}n^{\epsilon/2}\right). \qquad (3.3.6)$$

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 3.3.3 and Proposition 3.3.4 that $\mathbf{P}\{B\} \leq n^{-2}$ 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

$$\mathbf{P}\{E_3\} \leq \mathbf{P}\{B\} + \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}\{E_{3,k} \cap B^c\} \leq \frac{1}{n^2} + \sum_{k=0}^{\lfloor \log_2 \frac{\beta n}{\Delta^2(\mathbf{s})} \rfloor} \mathbf{P}\{E_{3,k} \cap B^c\}. \quad (3.3.7)$$

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

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

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

$$\mathbf{P}\{E_{3,k} \cap B^c\} \leq \sup \mathbf{P}\left\{\exists i \leq m, h(\mathbb{T}(\mathbf{r}_i)) \geq \delta n^{1/2}\right\}, \quad (3.3.8)$$

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

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

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

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

and that

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

Finally we must have $|\mathbf{r}_i| \ge \frac{\beta}{2^{k+1}}n$ for all $i \le m$, so $m \le \frac{2^{k+1}}{\beta}$. Now recall Theorem 3.1.5, which states that for a degree sequence $\mathbf{r} = (r^{(i)}, i \ge 0)$ and for all $h \ge 1$,

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

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 $n^{1/4} \ge 4/\epsilon$). Using a union bound in (3.3.8), and then applying Theorem 3.1.5, we obtain that

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

where we use the assumption $\sigma^2(\mathbf{s})/n \le 1/\epsilon$. And summing over k in (3.3.7) yields that

$$\mathbf{P}\left\{E_3\right\} \le \sum_{k\ge 0}\frac{2^{k+1}}{\beta} \cdot 7\exp\left(-\frac{\epsilon^3\delta^2}{9728}(\frac{2^k}{\beta})^{1/2}\right) + \frac{1}{n^2} \le C_5\frac{1}{\beta}\exp\left(-\frac{C_6}{\beta^{1/4}}\right) + \frac{1}{n^2} \quad (3.3.9)$$

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

$$\mathbf{P}\left\{E_4\right\} \le \frac{1}{n^2} + \mathbf{P}\left\{E_4 \cap B^c\right\}.$$

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 3.1.5, we obtain that

$$\begin{aligned}
\mathbf{P}\{E_4\} &\leq \frac{1}{n^2} + 7n\exp\left(-\frac{\delta^2 n}{608\sigma^2(\mathbf{r})1_{\mathbf{r}}^2}\right) \\
&\leq \frac{1}{n^2} + 7n\exp\left(-\frac{\delta^2 n}{608 n^{1-\frac{\epsilon}{2}}\frac{16}{\epsilon^2}}\right) \\
&= \frac{1}{n^2} + 7n\exp\left(-\frac{\epsilon^2}{9728} n^{\epsilon/2}\beta^{1/4}\right); \qquad (3.3.10)
\end{aligned}$$

recall that we take $\delta = \beta^{1/8}$. Of the bounds on $\mathbf{P}\{E_i\}, 1 \leq i \leq 4$ in (3.3.5), (3.3.6), (3.3.9) and (3.3.10), 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. $\square$

3.4 Convergence of the Lukasiewicz walk of forest to first passage bridge

In this section, we aim to prove Theorem 3.1.2 and conclude Proposition 1.3.3 as a corollary of Theorem 3.1.2. Throughout the section, we fix a sequence $(\mathbf{s}_n, n \in \mathbb{N})$ of degree sequences, with $\mathbf{s}_n = (s_n^{(i)}, i \geq 0)$ and $|\mathbf{s}_n| = n$. Recall that $\mathbf{p}_n = (p_n^{(i)}, i \geq 0) = (s_n^{(i)}/n, i \geq 0)$ and $\sigma_n = \sigma(\mathbf{p}_n)$ as defined in Section 1.3. Write $\sigma = \sigma(\mathbf{p})$ and $c_n = c(\mathbf{s}_n)$. Recall the notation d as in Section 3.2 and write $d_n = d(\mathbf{s}_n)$. Recall from Section 3.1 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 3.4.1. *Assume* $(\mathbf{s}_n, n \in \mathbb{N})$ *satisfies the hypothesis of Theorem 1.3.1, and in particular that* $c_n = c(\mathbf{s}_n) = (1+o(1))\lambda\sigma_n n^{1/2}$ *as* $n \to \infty$ *for some* $\lambda > 0$ *and that* $\sigma_n \to \sigma$. *For each* $n \in \mathbb{N}$, *fix a uniform random permutation* π_n *of* $[n]$, *and define a* $C[0,1]$ *function* $\widetilde{W}_n$ *by*

$$\widetilde{W}_n(t) := \frac{W_{\pi_n(d_n)}(tn)}{\sigma_n n^{1/2}}.$$

Then as $n \to \infty$,

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

To prove this theorem, we make use of the following result, which is Corollary 20.10 (a) in [10].

Theorem 3.4.2. *Consider a triangular array* $(Z_{q,i} : 1 \le i \le M_q, 1 \le q)$ *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*

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

Proof of Proposition 3.4.1. Let $d_{n,i} := \pi_n(d_n)_i - 1$, for $1 \le i \le n$. We apply the above theorem directly with $Z_{n,i} = \frac{d_{n,i}}{\sigma_n n^{1/2}}$. Condition (a) is satisfied since π_n is a uniformly random permutation of $[n]$. Condition (b) is satisfied since $\Delta_n = o(n^{1/2})$ and $\sup \sigma_n < \infty$.

Next note that, since $\sum_i d_{n,i} = \sum_i (\pi_n(d_n)_i - 1) = -c_n$,

$$\mu_n = \sum_i Z_{n,i} = \frac{\sum_i d_{n,i}}{\sigma_n n^{1/2}} = \frac{-c_n}{\sigma_n n^{1/2}} \to -\lambda \text{ as } n \to \infty, \tag{3.4.1}$$

the final convergence holding by our assumption on c_n. We also have

$$\begin{aligned}
\tau_n^2 &= \sum_i \left(\frac{d_{n,i}}{\sigma_n n^{1/2}} - \frac{-c_n}{\sigma_n n^{1/2} n} \right)^2 \\
&= \frac{1}{\sigma_n^2 n} (\sum_i d_{n,i}^2 + 2\frac{c_n}{n} \sum_i d_{n,i} + \frac{c_n^2}{n}) = \frac{1}{\sigma_n^2 n} (\sum_i d_{n,i}^2 - \frac{c_n^2}{n}) \\
&= \frac{1}{\sigma_n^2 n} \sum_i d_{n,i}^2 + o(1),
\end{aligned}$$

the last equation holding since $c_n = O(n^{1/2})$.

Next note that

$$\begin{aligned}
\sum_i d_{n,i}^2 = \sum_i (\pi_n(d_n)_i - 1)^2 &= \sum_i ((d_n)_i)^2 + n - 2 \sum_i (d_n)_i \\
&= n(\sigma_n^2 + 1) + n - 2(n - c_n) \\
&= n\sigma_n^2 + 2c_n.
\end{aligned}$$

It follows that

$$\tau_n^2 = \frac{1}{\sigma_n^2 n} (n\sigma_n^2 + 2c_n) + o(1) \to 1 \tag{3.4.2}$$

as $n \to \infty$ by our assumption on $\mathbf{s}_n$.

Using equations (3.4.1) and (3.4.2), by Theorem 3.4.2 we conclude that

$$\left(\frac{W_{\pi_n(d_n)}(\lfloor tn \rfloor)}{\sigma_n n^{1/2}},\ 0 \le t \le 1 \right) \xrightarrow{d} \left(B^{br}(t) - \lambda t,\ 0 \le t \le 1 \right) \text{ in } D[0,1].$$

For all t,

$$\left|\frac{W_{\pi_n(d_n)}(\lfloor tn \rfloor)}{\sigma_n n^{1/2}} - \frac{W_{\pi_n(d_n)}(tn)}{\sigma_n n^{1/2}}\right| \le \frac{\Delta_n}{\sigma_n n^{1/2}} = o(1)$$

by assumption, so we must also have $\left(\widetilde{W}_n(t), 0 \le t \le 1\right) \xrightarrow{d} \left(B^{br}(t) - \lambda t,\ 0 \le t \le 1\right)$ in $D[0,1]$. Since the Skorohod topology relativized to $C[0,1]$ coincides with the uniform topology (see page 124 of [17]), the result follows. $\square$

Let $f : \mathcal{C}_0(1) \times [0, \infty) \to \mathcal{C}_0(1)$ be defined by

$$f(b, v) := \theta_u(b) \text{ where } u = \inf\{t : b(t) \le \min_{0 \le s \le 1} b(s) + v\}.$$

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

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

where $\tau(\lambda) := \inf\{t : B(t) < -\lambda\}$ 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 [14].

Theorem 3.4.3 ([14], 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 3.4.1. *Note that [14] 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 3.4.4. *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 $|\theta_u(b)(t) - b(t)| = |b(t+u) - b(u) - b(t)| \leq |b(u) - b(0)| + |b(t+u) - b(t)| \leq 2\epsilon(u)$. If $t \in [1-u, 1]$, then $|\theta_u(b)(t) - b(t)| = |b(t+u-1) + b(1) - b(u) - b(t)| \leq |b(t+u-1) - b(u)| + |b(1) - b(t)| \leq 2\epsilon(u)$. Since $\epsilon(u) \to 0$ as $u \to 0$, the result follows. $\square$

Lemma 3.4.5. *Given $b \in \mathcal{C}_0(1)$ and $0 \leq v \leq -\min(b)$, if $f(b, v) = \theta_{t_{v+\min(b)}}(b)$ is not continuous at v, then b attains a local minimum at $t_{v+\min(b)}$.*

Proof. By Lemma 3.4.4, 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 $0 \leq v \leq -\min(b)$, 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

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

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 the lemma. $\square$

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

Lemma 3.4.6. *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 [57]). 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 (3.1.1) and equation (3.1.2) 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 3.4.7. *Let ν be* $\mathrm{Unif}[0,\lambda]-$*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 3.4.5, 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 3.4.6, 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 3.1.2.

Proof of Theorem 3.1.2. For each $n \geq 1$ let ν_n be a uniformly random element of $[c_n]-1$ independent of π_n, and let ν be $\mathrm{Unif}[0,\lambda]$ and independent of B_λ^{br}. By

Corollary 3.2.3,

$$f(\widetilde{W}_n, \frac{\nu_n}{\sigma_n n^{1/2}}) = f\left(\frac{W_{\pi(d_n)}(tn)}{\sigma_n n^{1/2}}, \frac{\nu_n}{\sigma_n n^{1/2}}\right) \stackrel{d}{=} \left(\frac{S_{\mathbb{F}_n}(tn)}{\sigma_n n^{1/2}}\right)_{t\in[0,1]}.$$

By Proposition 3.4.1, we have $\widetilde{W}_n \xrightarrow{d} B_\lambda^{br}$, and clearly we have $\sigma_n^{-1}n^{-1/2}\nu_n \xrightarrow{d} \nu$. By independence we have $(\widetilde{W}_n, \sigma_n^{-1}n^{-1/2}\nu_n) \xrightarrow{d} (B_\lambda^{br}, \nu)$. Since by Lemma 3.4.7 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 [17]) to conclude that

$$f(\widetilde{W}_n, \sigma_n^{-1}n^{-1/2}\nu_n) \xrightarrow{d} f(B_\lambda^{br}, \nu).$$

By Theorem 3.4.3, $F_\lambda^{br} \stackrel{d}{=} f(B_\lambda^{br}, \nu)$, hence we conclude that

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

as required. □

Now we begin with the preparation work to prove Proposition 1.3.3. 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 $g(s) - \min_{0\le s'<s} g(s')$. (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 1.3.3, we use the following result from [22].

Lemma 3.4.8. *[Lemma 3.8 and Corollary 3.10 in [22]] 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

$$\zeta(l_1) > \zeta(l_2). \tag{3.4.3}$$

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

Proof. Suppose to the contrary that for some $\gamma_1, \gamma_2 \in \mathcal{E}$ with $l(\gamma_1) < l(\gamma_2)$, we have $F_\lambda^{br}(l(\gamma_1)) \le F_\lambda^{br}(l(\gamma_2))$, then since γ_1, γ_2 are excursions of $F_\lambda^{br}(s) - \min_{0 \le s' < s} F_\lambda^{br}(s')$, we must in fact have $F_\lambda^{br}(l(\gamma_1)) = F_\lambda^{br}(l(\gamma_2))$. 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_λ^{br} 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 3.4.6 (b).

□

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

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

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

Lemma 3.4.10. *Almost surely, the coordinates of $h(F_\lambda^{br})$ sum to 1, and are all strictly positive.*

Proof. By Proposition 5 of [14], $h(F_\lambda^{br})$ 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 $r_i \leq l_{i+1}, 1 \leq i \leq k-1$. Then since $\sum_{i=1}^{k}(r_i - l_i) = 1$, we must in fact have $r_i = l_{i+1}, \forall 1 \leq i \leq k-1$ 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 $F_\lambda^{br}(1) = -\lambda < 0$. $\square$

Proof of Proposition 1.3.3. We first prove that for any fixed $j \geq 1$, as $n \to \infty$,

$$(|\mathbb{T}_{n,l}|/n)_{1\leq l\leq j} \xrightarrow{d} (|\gamma_l|)_{1\leq l\leq j}. \tag{3.4.4}$$

Let $\zeta_n = \left(\frac{S_{\mathbb{F}_n}(tn)}{\sigma_n n^{1/2}}\right)_{t\in[0,1]}$ and let $\zeta = \left(F_\lambda^{br}(t)\right)_{t\in[0,1]}$. By (3.1.4) and by Skorokhod's representation theorem, we may work in a probability space in which $\zeta_n \xrightarrow{a.s.} \zeta$. Let E be the set of excursion intervals of ζ. Then Lemma 3.4.9 guarantees equation (3.4.3) in Lemma 3.4.8 is true and Lemma 3.4.10 guarantees that the complement of $\cup_{I\in E} I$ has Lebesgue measure 0, as required by Lemma 3.4.8. For each n let $t_{n,0} = 0$ and for $1 \leq j \leq c_n$ let $t_{n,j}$ be such that $nt_{n,j}$ is the time the depth-first walk $S_{\mathbb{F}_n}$ finishes visiting the $j-$th tree of $\mathbb{F}_n$. Then almost surely, condition (i) of Lemma 3.4.8 is clearly true and condition (iii) is also true since for each $1 \leq j \leq c_n$, $\zeta_n(t_{n,j}) = \zeta_n(t_{n,j-1}) - \frac{1}{\sigma_n n^{1/2}}$. The definition of Lukasiewicz walk guarantees that the times at which $\frac{S_{\mathbb{F}_n}(tn)}{\sigma_n n^{1/2}}$ 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 3.4.8 is also satisfied. Also note that the vector consisting of decreasingly ranked elements of $\{t_{n,j} - t_{n,j-1}, 1 \leq j \leq c_n\}$ is simply the scaled decreasing ordering of tree component sizes $(|\mathbb{T}_{n,l}|/n)_{1\leq l\leq c_n}$. Hence by Lemma 3.4.8 we know that

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

which immediately implies weak convergence. Lemma 3.4.10 guarantees that this is true for any positive integer j. We also have $h_j(F_\lambda^{br}) \stackrel{d}{=} (|\gamma_l|)_{1\leq l\leq j}$ by definition, and (3.4.4) follows.

To prove (1.3.2) from (3.4.4), we only need to prove that for any $\epsilon > 0$, there exists $I_0 \in \mathbb{N}$ such that $\limsup_{n\to\infty} \mathbf{P}\left\{\sum_{l>I_0} \frac{|\mathbb{T}_{n,l}|}{n} > \epsilon\right\} < \epsilon$. Since by Lemma 3.4.10 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_n be the event that $\sum_{l\leq I_0} \frac{|\mathbb{T}_{n,l}|}{n} < 1-\epsilon$ and A be the event that $\sum_{l\leq I_0} |\gamma_l| < 1-\epsilon$ (which has probability less than $\epsilon/2$ by our choice of I_0). By (3.4.4), we have $|\mathbf{P}\{A_n\} - \mathbf{P}\{A\}| < \epsilon/2$ for n large enough. Therefore

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

as required. □

3.5 Proof of Proposition 3.1.1 and Proposition 3.1.4

We assume that we have the conditions of Theorem 1.3.1 hold. In particular, we have a probability distribution $\mathbf{p}$ on $\mathbb{N}$. Recall that $\sigma = \sigma(\mathbf{p}), \sigma_n = \sigma(\mathbf{p}_n)$. Let $\mathbf{s}_{n,l} = (s_{n,l}^{(i)}, i \geq 0)$ denote the degree sequence of $\mathbb{T}_{n,l}$. Recall that $p_n^{(i)} = s_n^{(i)}/n$ and let $p_{n,l}^{(i)} = s_{n,l}^{(i)}/|\mathbf{s}_{n,l}|$ be the empirical proportion of degree i among all vertices of the $l-$th largest tree $\mathbb{T}_{n,l}$. Note that $p_n^{(i)}$ is deterministic while $p_{n,l}^{(i)}$ is random.

First, we are going to prove Proposition 3.1.1 by using Theorem 3.1.3. To do so, we will have to first show that the assumptions of Theorem 3.1.3 are satisfied in our setting.

Proposition 3.5.1. *Under the assumption of Theorem 1.3.1, for all $l \geq 1$, as $n \to \infty$ we have*

(a) $\mathbf{p}_{n,l} \xrightarrow{p} \mathbf{p}$ *coordinatewise, that is,* $p_{n,l}^{(i)} \xrightarrow{p} p^{(i)}$ *for all* $i \geq 1$.

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

Proof. For (a), we know that by Lemma 3.3.3 and Proposition 3.3.4, for fixed $\epsilon > 0, i, l \in \mathbb{N}$ and n large enough, we have

$$\mathbf{P}\left\{|p_{n,l}^{(i)} - p_n^{(i)}| > \epsilon\right\} \leq 1/n + \mathbf{P}\left\{|\mathbb{T}_{n,l}| \leq n^{1/4}\right\}. \tag{3.5.1}$$

For any $\epsilon' > 0$, there exists $\delta > 0$ such that $\mathbf{P}\{|\gamma_l| < \delta\} < \epsilon'/2$ and by (3.4.4) we can find n_0 such that for all $n \geq n_0$ we have $\mathbf{P}\left\{\frac{|\mathbb{T}_{n,l}|}{n} < \delta\right\} \leq \mathbf{P}\{|\gamma_l| < \delta\} + \epsilon'/2$ and $n^{-3/4} < \delta$. Hence $\mathbf{P}\left\{|\mathbb{T}_{n,l}| \leq n^{1/4}\right\} = \mathbf{P}\left\{\frac{|\mathbb{T}_{n,l}|}{n} \leq n^{-3/4}\right\} \leq \mathbf{P}\left\{\frac{|\mathbb{T}_{n,l}|}{n} \leq \delta\right\} < \epsilon'$. Hence $\mathbf{P}\left\{|\mathbb{T}_{n,l}| \leq n^{1/4}\right\} = o(1)$ as $n \to \infty$. Therefore by (3.5.1) we know that $|p_{n,l}^{(i)} - p_n^{(i)}| \xrightarrow{p} 0$ as $n \to \infty$, which implies (a) since by assumption of Theorem 1.3.1 we have $\mathbf{p}_n$ converges to $\mathbf{p}$ coordinatewise.

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

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

Such ϵ exists by (3.4.4).

Then let M be large enough that $\sigma_{n,>M}^2 := \sum_{i>M} i^2 \frac{s_n^{(i)}}{n} < \epsilon^2$ for all n (such M exists since under the assumption of Theorem 1.3.1 σ_n^2 converges). And let $\sigma_{n,l,>M}^2 =$

$\sum_{i>M} i^2 \frac{s_{n,l}^{(i)}}{|\mathbf{s}_{n,l}|}$ similarly. Note that

$$\sigma_{n,l,>M}^2 \leq \sum_{i>M} i^2 \frac{s_n^{(i)}}{|\mathbb{T}_{n,l}|} = \sigma_{n,>M}^2 \frac{n}{|\mathbb{T}_{n,l}|},$$

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

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

Since $|p_{n,l}^{(i)} - p_n^{(i)}| \to 0$ in probability for all i by part (a), and $\sum_{i>M} i^2 p_n^{(i)} < \epsilon^2 < \epsilon$, and $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p})$ by assumption of Theorem 1.3.1, this yields that

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

which proves part (b). $\square$

Lemma 3.5.2. *Let $\Delta_{n,l}$ be the largest degree of a vertex of $\mathbb{T}_{n,l}$. For any fixed l, we have*

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

Proof. For any $\delta > 0$, we need to prove $\lim_{n\to\infty} \mathbf{P}\left\{\frac{\Delta_{n,l}}{\sqrt{|\mathbb{T}_{n,l}|}} > \delta\right\} = 0$. For any $\epsilon > 0$, by Lemma 3.4.10 we can choose $\epsilon' > 0$ such that $\mathbf{P}\{|\gamma_l| < \epsilon'\} \leq \epsilon/2$. Then choose n_0 such that when $n \geq n_0$ we have

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

This is possible since $\Delta_n = o(n^{1/2})$ by Remark 3.1.1 and $|\mathbb{T}_{n,l}|/n \xrightarrow{d} |\gamma_l|$ by (3.4.4). Therefore

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

hence the claim. □

With Proposition 3.5.1 and Lemma 3.5.2, we are now ready to give the proof of Proposition 3.1.1.

Proof of Proposition 3.1.1. Let $\mathbf{s}_{n,l}$ be the random degree sequence of the $l-$th largest tree in the forest $\mathbb{F}_n$. Then by Proposition 1.3.3, we have

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

By Proposition 3.5.1 and Lemma 3.5.2, we know we can apply Theorem 3.1.3 to $\mathcal{T}_{n,l}$ to conclude that for each fixed $l \le j$,

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

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

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

The result follows by Brownian scaling. □

Finally, we give the proof of Proposition 3.1.4 based on Proposition 3.3.8, with the assumptions of Theorem 1.3.2.

Proof of Proposition 3.1.4. By assumption we have $\sigma_n \to \sigma \in (0, \infty)$ and $s_n^{(1)}/|\mathbf{s}_n| \to p^{(1)} < 1$. 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 3.3.8, so that for all n sufficiently large, if a degree sequence $\mathbf{s}$ satisfies $|\mathbf{s}| = n, \Delta(\mathbf{s}) \leq n^{\frac{1-\epsilon}{2}}, s^{(1)} \leq (1-\epsilon)|\mathbf{s}|$ and $\epsilon \leq \sigma^2(\mathbf{s})/n \leq 1/\epsilon$, then for any $0 < \beta < \beta_0$,

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

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

$$\mathbf{P}\left\{\exists T \in \mathbb{F}(\mathbf{s}_n) : |T| < \beta n, h(T) > \beta^{1/8} n^{1/2}\right\} \leq \rho. \tag{3.5.2}$$

Finally, taking $\beta = (a/\sigma_n)^8$ in (3.5.2), since $\mathcal{T}_{n,l} = \frac{\sigma_n}{2n^{1/2}} \mathbb{T}_{n,l}$ and for all $j > 1/\beta$ we have $|\mathbb{T}_{n,j}| < \beta n$, it follows that for all n sufficiently large,

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

Since $\operatorname{diam}(\mathcal{T}_{n,l}) \leq 2h(\mathcal{T}_{n,l})$, the result now follows easily. $\square$

3.6 Proof of Remark 3.1.1 and Remark 3.1.3

Let's restate Remark 3.1.1 as the following lemma and we want to emphasize that the following lemma applies to the settings of both Chapter 3 and Chapter 4.

Lemma 3.6.1. *Suppose distributions $\mathbf{p}_n$ converges to $\mathbf{p}$ coordinatewise and $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p}) \in (0, \infty)$ and $\frac{c(\mathbf{s}_n)}{n^{1/2}} \to x \in [0, \infty)$, then $\mu(\mathbf{p}_n) \to \mu(\mathbf{p}) = 1$ and $\Delta_n/n^{1/2} \to 0$ as $n \to \infty$.*

Proof. We write $c_n = c(\mathbf{s}_n)$. First, since

$$0 \le \mu(\mathbf{p}) = \sum ip^{(i)} \le \sum i^2 p^{(i)} = \sigma^2(\mathbf{p}) + 1 < \infty,$$

we have $\mu(\mathbf{p}) \in (0, \infty)$. And we can compute the limit of $\mu(\mathbf{p}_n)$ explicitly:

$$\mu(\mathbf{p}_n) = \sum ip_n^{(i)} = \sum i\frac{s_n^{(i)}}{n} = \frac{n - c_n}{n} \to 1$$

by our assumption of the magnitude of c_n.

Next, since $\mathbf{p}_n \to \mathbf{p}$ coordinatewise, for all $M \in \mathbb{N}$ we have

$$\lim_{n\to\infty} |\sum_{i \le M} ip_n^{(i)} - \sum_{i \le M} ip^{(i)}| = 0.$$

It follows that

$$\begin{aligned}
\limsup_{n\to\infty} |\sum ip_n^{(i)} - \sum ip^{(i)}| &= \lim_{M\to\infty} \limsup_{n\to\infty} |\sum_{i \ge M} ip_n^{(i)} - \sum_{i \ge M} ip^{(i)}| \\
&\le \lim_{M\to\infty} \limsup_{n\to\infty} \left(\sum_{i \ge M} ip_n^{(i)} + \sum_{i \ge M} ip^{(i)} \right) \\
&\le \lim_{M\to\infty} \limsup_{n\to\infty} \left(\sum_{i \ge M} i^2 p_n^{(i)} + \sum_{i \ge M} i^2 p^{(i)} \right) \\
&= 0,
\end{aligned}$$

where the final equality holds since $\sigma(\mathbf{p}) < \infty$ and $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p})$. Hence $\mu(\mathbf{p}_n) \to \mu(\mathbf{p})$.

Since $\mathbf{p}_n \to \mathbf{p}$ coordinatewise, it follows that for any integer N,

$$\limsup_{n\to\infty} |\sum_{i \ge N} i^2 p_n^{(i)} - \sum_{i \ge N} i^2 p^{(i)}| = \limsup_{n\to\infty} |\sigma^2(\mathbf{p}_n) - \sigma^2(\mathbf{p})| = 0.$$

Now let $\epsilon > 0$ and let N be large enough that $0 < \sum_{i \geq N} i^2 p^{(i)} < \epsilon$. Then for all n sufficiently large, $0 < \sum_{i \geq N} i^2 p_n^{(i)} < \epsilon$. But $\sum_{i \geq N} i^2 p_n^{(i)} \geq \epsilon \mathbb{1}_{\Delta_n \geq (\epsilon n)^{1/2}}$, so this implies that $\limsup_{n \to \infty} \frac{\Delta_n}{n^{1/2}} \leq \epsilon^{1/2}$. Since $\epsilon > 0$ was arbitrary, the result follows. □

Next we proceed to prove Remark 3.1.3. The following proposition will be useful for our justification of Remark 3.1.3 (see Lemma 2.4 in [45] for a version dealing with Gromov-Hausdorff distance instead of Gromov-Hausdorff-Prokhorov distance):

Proposition 3.6.2 (Proposition 2.9 in [2])**.** *Let f, g be two compactly supported non-negative continuous functions with $f(0) = g(0) = 0$. Then*

$$d_{GHP}(\mathcal{T}_f, \mathcal{T}_g) \leq 6||f - g||_\infty + |\sigma_f - \sigma_g|.$$

Now we prove the following result.

Proposition 3.6.3. *The GH convergence in Theorem 1 in [20] can be strengthened to GHP convergence as in Theorem 3.1.3.*

Proof. Let C_n be the contour function of $\mathbb{T}_n$, define $\hat{C}_n : [0, 1] \to [0, \infty)$ by letting $\hat{C}_n(t) = \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} C_n(2(n-1)t)$, then it is shown in [20] (see Theorem 3 there) that $\hat{C}_n \xrightarrow{d} \mathbf{e}$ in the space $C([0, 1], \mathbb{R})$, equipped with the supremum distance. By Proposition 3.6.2 and Skorokhod's representation theorem, it follows that $\mathcal{T}_{\hat{C}_n} \xrightarrow{d} \mathcal{T}_{\mathbf{e}}$ in the GHP sense.

Next, metrically we may realize $\mathcal{T}_n$ as the subspace of $\mathcal{T}_{\hat{C}_n}$ consisting of the set U of points whose distance from the root is an integer multiple of $\frac{\sigma(\mathbf{p}_n)}{2n^{1/2}}$. With this identification

$$d_H(\mathcal{T}_n, \mathcal{T}_{\hat{C}_n}) = \frac{1}{2} \cdot \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}}.$$

Moreover, the measure $\hat{\mu}_n$ on $\mathcal{T}_{\hat{C}_n}$ is the (normalized) length measure, and the measure μ_n on $\mathcal{T}_n$ is the uniform measure on its points. It follows that

$$d_P(\hat{\mu}_n, \mu_n) \leq \frac{1}{n} + \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}}.$$

To see this, for each $u \in U$ which is not the root of $\mathcal{T}_n$, let e_u be the parent edge of u, which we view as a closed line segment of length $\epsilon = \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}}$ in $\mathcal{T}_{\hat{C}_n}$. For any non-empty set $S \subset U$, we have $\mu_n(S) = |S|/n$. Hence

$$\hat{\mu}_n(S^\epsilon) \geq \frac{|S|-1}{n-1} \geq \mu_n(S) - \frac{1}{n},$$

where the first inequality is because for non-root $u \in S$, we have $e_u \subset S^\epsilon$. On the other hand, let A be a closed set in $\mathcal{T}_{\hat{C}_n}$ and let $l = |\{e \in E(\mathcal{T}_n) : A \cap e \neq \emptyset\}|$. Then A^ϵ contains at least l vertices of $\mathcal{T}_n$ since no cycle exists, so

$$\mu_n(A^\epsilon) \geq \frac{l}{n} = \frac{l}{n-1} - \frac{l}{n(n-1)} \geq \hat{\mu}_n(A) - \frac{1}{n}.$$

Hence $d_{GHP}(\mathcal{T}_n, \mathcal{T}_{\hat{C}_n}) \xrightarrow{d} 0$. By the triangle inequality, it follows that $\mathcal{T}_n \xrightarrow{d} \mathcal{T}_{\mathbf{e}}$ in the GHP sense. $\square$

CHAPTER 4
Supercritical case

4.1 Introduction

The goal of this chapter is to study the asymptotic structure of large random forests with a given degree sequence, in the "supercritical", finite variance regime. That is, let n be the number of nodes of the forest, we consider the case where the number of trees of uniformly random forests is $o(n^{1/2})$. This setting is a natural generalization of [20], where the uniformly random tree is studied (see Theorem 1.3.4). In this "supercritical" setting, the forest typically consists of a single, large tree containing all but a vanishing fraction of the nodes. The scaling limit of this tree is $\mathcal{T}$, the Brownian Continuum Random Tree (BCRT) introduced by Aldous in [6, 7, 8]. The remaining nodes form another random forest, which may be expected to have its own scaling limit (with an appropriate rescaling, which should be different from that of the large tree; In fact, the remaining trees behave like a forest which is *critical*, this behaviour is also observed in the random forest model studied in [44]). The contributions of this chapter confirm that the above picture is correct, and yield a pleasingly straightforward description, which we now provide, for the joint scaling limit of the large tree and the small trees. This chapter is essentially derived from the manuscript of [48]. We now first give the construction of the limit $\mathcal{F}$ in Theorem 1.3.4.

Let $B = (B(t), t \geq 0)$ be a linear Brownian motion. For $t \geq 0$ let $R(t) = B(t) - \inf(B(s), s \leq t)$; the process $R = (R(t), t \geq 0)$ is Brownian motion reflected at its running minimum. Let $Z = \{t \geq 0 : R(t) = 0\}$ be the zero set of R. By definition, this is also the set of times at which B is equal to its running minimum.

Now let $\tau(x) = \inf(t : B(t) \leq -x)$ for $x \geq 0$, and let $Z(x) = Z \cap [0, \tau(x)]$. For $\sigma > 0$, the relative complement $[0, \tau(\frac{1}{\sigma})] \setminus Z(\frac{1}{\sigma})$ is almost surely a countable collection of intervals with distinct lengths, and with total length $\tau(\frac{1}{\sigma})$. List these intervals in decreasing order of length as $((g_i, d_i), i \geq 1)$.

For $i \geq 1$ let $\mathcal{T}_i$ be the continuum random tree coded by B_i, where

$$B_i = (B(g_i + t) - B(g_i), 0 \leq t \leq d_i - g_i) = (R(g_i + t) - R(g_i), 0 \leq t \leq d_i - g_i).$$

Then the scaling limit of the small trees has the law of the sequence $\mathcal{F} = (\mathcal{T}_i^{\downarrow}, i \geq 1)$, which is a decreasing reordering of $(\mathcal{T}_i, i \geq 1)$ according to $(d_i - g_i, i \geq 1)$.

Now we recall Theorem 1.3.4, the main goal of this chapter.

Theorem 1.3.4. *Suppose that there exists a distribution* $\mathbf{p} = (p^{(i)}, i \geq 0)$ *on* $\mathbb{N}_0$ *such that* $\mathbf{p}_n$ *converges to* $\mathbf{p}$ *coordinatewise. Suppose also that* $\sigma(\mathbf{p}_n) \to \sigma := \sigma(\mathbf{p}) \in (0, \infty)$*. If* $c_n := c(\mathbf{s}_n) = o(n^{1/2})$*, then*

$$\left(\mathcal{T}_{n,1}, \hat{\mathcal{F}}_n, \frac{n - |\mathbb{T}_{n,1}|}{c_n^2}\right) \xrightarrow{d} \left(\mathcal{T}, \mathcal{F}, \tau\left(\frac{1}{\sigma}\right)\right)$$

where the first coordinate of the joint convergence is in the GHP sense, the second coordinate is in the sense of coordinatewise GHP convergence, and $\mathcal{T}$ *and* $\mathcal{F}$ *are independent.*

Fix a critical, finite variance offspring distribution ν, and let $\mathcal{F}_n$ be a forest of c_n independent Galton-Watson(ν) trees with offspring distribution ν, conditioned to have total progeny n. It is not hard to check, as in [20], that with high probability the degree sequence of $\mathcal{F}_n$ satisfies the conditions of Theorem 1.3.4, so the distributional convergence of the theorem also applies to $\mathcal{F}_n$. The convergence of the third coordinate, in the Galton-Watson setting, appears as Theorem 1.4.5, and provides a new proof and different perspective on that result; the convergence of the second coordinate strengthens and generalizes and removes a moment assumption from Theorem 1.7 of [23] (part of Theorem 1.4.7).

Outline of the section

In the remainder of this section, we first describe the key ingredients of the proof of our main theorem in Section 4.1.1. Next in Section 4.1.2 we explain how to deduce Theorem 1.3.4 from the results of Section 4.1.1, and outline the remaining sections of the chapter.

4.1.1 Functional convergence and proof of Theorem 1.3.4

Given a degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$ with $|\mathbf{s}| = n$, recall that we let $d(\mathbf{s}) \in \mathbb{Z}^n_{\geq 0}$ be the vector whose entries are weakly increasing and with $s^{(i)}$ entries equal to i, for each $i \geq 0$. Suppose we have a sequence of degree sequences $(\mathbf{s}_n)_{n \in \mathbb{N}}$, with $\mathbf{s}_n = (s_n^{(i)}, i \geq 0), |\mathbf{s}_n| = n, c_n := c(\mathbf{s}_n) = o(n^{1/2})$ and $n^{-1} \cdot \mathbf{s}_n \to \mathbf{p}$ in L^2 for some distribution $\mathbf{p} = (p^{(i)}, i \geq 0)$ with mean 1 and finite variance σ^2 on $\mathbb{N}$. Let $C_{n,1}, \cdots, C_{n,n}$ be a uniformly random permutation of $d(\mathbf{s}_n)$. For $1 \leq k \leq n$, let $X_{n,k} = C_{n,k} - 1$, and set $S_{n,k} = \sum_{j=1}^{k} X_{n,j}$. Our proof makes use of the following functional convergence theorem.

Theorem 4.1.1. *We have the following convergence of processes:*

$$\left(\frac{1}{c_n} S_{n,\lfloor tc_n^2 \rfloor}, t \geq 0\right) \xrightarrow{d} (\sigma B(t), t \geq 0) \tag{4.1.1}$$

where $(B(t), t \geq 0)$ is standard Brownian Motion.

Theorem 4.1.1 will yield a description of the asymptotic behaviour of the sizes of all but the largest tree of $\mathbb{F}_n$.

Corollary 4.1.2. *We have*

$$\left(\frac{|\mathbb{T}_{n,i+1}|}{c_n^2}, i \geq 1\right) \xrightarrow{d} (d_i - g_i, i \geq 1) \text{ in } L^1$$

where $((g_i, d_i), i \geq 1)$ are the excursion intervals of $(R(t), t \leq \tau(\frac{1}{\sigma}))$ in decreasing order of length.

Corollary 4.1.2 is equivalent to the assertions that

$$\frac{1}{c_n^2} \sum_{i \geq 2} |\mathbb{T}_{n,i}| \xrightarrow{d} \tau(\frac{1}{\sigma}) = \sum_{i \geq 1} (d_i - g_i) \tag{4.1.2}$$

and that for any fixed $j \in \mathbb{N}$,

$$\left(\frac{|\mathbb{T}_{n,2}|}{c_n^2}, \frac{|\mathbb{T}_{n,3}|}{c_n^2}, \ldots, \frac{|\mathbb{T}_{n,j}|}{c_n^2}\right) \xrightarrow{d} (d_1 - g_1, d_2 - g_2, \ldots, d_{j-1} - g_{j-1}). \tag{4.1.3}$$

We will prove this corollary in Section 4.3. To describe the limit structure of each tree, we again appeal to Theorem 3.1.3. And we recall it here.

Theorem 3.1.3. *Let $\{\mathbf{s}_n, n \geq 1\}$ be a degree sequence such that $|\mathbf{s}_n| = n \to \infty, \Delta_n := \Delta(\mathbf{s}_n) = o(n^{1/2})$. Suppose that there exists a distribution $\mathbf{p} = (p^{(i)}, i \geq 0)$ on $\mathbb{N}$ with mean 1 such that $\mathbf{p}_n = (s_n^{(i)}/n, i \geq 0)$ converges to $\mathbf{p}$ coordinatewise and such that $\sigma(\mathbf{p}_n) \to \sigma(\mathbf{p}) \in (0, \infty)$. Let $\mathbb{T}_n$ be the random plane tree under $\mathbb{P}_{\mathbf{s}_n}$, the* *uniform measure on the set of plane trees with degree sequence* $\mathbf{s}_n$*. Let* $\mathcal{T}_n$ *denote the measured rooted metric space* $(\mathbb{T}_n, \frac{\sigma(\mathbf{p}_n)}{2n^{1/2}} d_{gr}, \emptyset_n, \mu_n)$ *where* μ_n *denotes the uniform measure putting mass* $\frac{1}{n}$ *on each vertex of* $\mathbb{T}_n$*. Then when* $n \to \infty, \mathcal{T}_n \xrightarrow{d} \mathcal{T}$ *in the Gromov-Hausdorff-Prokhorov sense, where* $\mathcal{T}$ *is BCRT.*

To apply Theorem 3.1.3 to each $\mathbb{T}_{n,i}$, we also need to verify that the assumptions of Theorem 1.3.4 hold. For fixed integers $i \geq 0$ and $l \geq 1$, let

$$p_{n,l}^{(i)} := \frac{|\{v \in \mathbb{T}_{n,l} : k(v) = i\}|}{|\mathbb{T}_{n,l}|} \text{ and } \mathbf{p}_{n,l} = (p_{n,l}^{(i)}, i \geq 0).$$

In Section 4.4 we prove the following assertions:

$$\text{for any fixed } i \geq 0 \text{ and } l \geq 1,\ p_{n,l}^{(i)} - p_n^{(i)} \xrightarrow{p} 0, \text{ as } n \to \infty, \tag{4.1.4}$$

and

$$\text{for any } l \geq 1,\ \sigma^2(\mathbf{p}_{n,l}) - \sigma^2(\mathbf{p}_n) \xrightarrow{p} 0, \text{ as } n \to \infty. \tag{4.1.5}$$

Note that once these two conditions are verified, it follows that for any fixed $l \geq 1$,

$$\max\{i : p_{n,l}^{(i)} \neq 0\} = o_p(|\mathbb{T}_{n,l}|^{1/2}) \text{ as } n \to \infty;$$

as in Lemma 3.6.1.

4.1.2 Proof of Theorem 1.3.4

Now we are ready to give the proof of Theorem 1.3.4, assuming the results of Section 4.1.1.

Proof. It suffices to prove that for any fixed $j \in \mathbb{N}$,

$$\left(\frac{\sigma(\mathbf{p}_n)\mathbb{T}_{n,1}}{2n^{1/2}}, \left(\frac{\sigma(\mathbf{p}_n)\mathbb{T}_{n,l}}{2c_n}, 2 \leq l \leq j \right), \frac{n - |\mathbb{T}_{n,1}|}{c_n^2} \right) \xrightarrow{d} \left(\mathcal{T}, (\mathcal{T}_1^{\downarrow}, \cdots, \mathcal{T}_{j-1}^{\downarrow}), \tau\left(\frac{1}{\sigma}\right) \right).$$

The convergence of the third coordinate is simply (4.1.2). This in particular implies that $\frac{|\mathbb{T}_{n,1}|}{n} \xrightarrow{p} 1$. Since $\mathbf{p}_n \to \mathbf{p}$ in L^2, it straightforwardly follows that with probability $1 - o(1)$, $\mathbb{T}_{n,1}$ satisfies the conditions of Theorem 3.1.3; this yields the convergence of the first coordinate. With (4.1.4) and (4.1.5), we can also apply Theorem 3.1.3 to each $\mathbb{T}_{n,l}$ with $l \geq 2$ and conclude that

$$\frac{\sigma(\mathbf{p}_{n,l})}{2|\mathbb{T}_{n,l}|^{1/2}}\mathbb{T}_{n,l} \xrightarrow{d} \mathcal{T}.$$

Since the trees $(\mathbb{T}_{n,l}, l \geq 1)$ are conditionally independent given their degree sequences, it follows that

$$\left(\frac{\sigma(\mathbf{p}_{n,l})}{2|\mathbb{T}_{n,l}|^{1/2}}\mathbb{T}_{n,l}, 2 \leq l \leq j\right) \xrightarrow{d} \left(\tilde{\mathcal{T}}_{l-1}, 2 \leq l \leq j\right),$$

where $(\tilde{\mathcal{T}}_l)_{l\in\mathbb{N}}$ are independent copies of $\mathcal{T}$. Using (4.1.5) again, together with (4.1.3) and Brownian scaling, the convergence of the second coordinate then follows. $\square$

Outline of the rest of the chapter

In Section 4.2 we describe a combinatorial construction which associates a *marked cyclic forest* with the concatenation of a sequence of first passage lattice bridges, followed by one lattice bridge. This construction is what links Theorem 4.1.1 with random forests. In Section 4.3 we give the proof of Theorem 4.1.1 and Corollary 4.1.2. Finally in Section 4.4 we prove (4.1.4) and (4.1.5) using martingale concentration inequalities.

4.2 Coding marked cyclic forests by lattice paths

We call a sequence of integers $\mathbf{b} = (b(0), b(1), \dots, b(n))$ a $1-$*lattice bridge* if

$$b(0) = 0, b(n) = -1 \text{ and } \forall 0 \le i \le n-1,\ b(i+1) - b(i) \ge -1.$$

If $\mathbf{b}$ is a lattice bridge and $\min_i\{i : b(i) = -1\} = n$, then we call $\mathbf{b}$ a $1-$*first passage lattice bridge*. Given a $1-$lattice bridge $\mathbf{b}$ and a positive integer $k \le n$, we define a $1-$lattice bridge $\mathbf{b}^{(k)}$ as follows. First, for $1 \le i \le n$, let $b(n+i) = b(n) + b(i) = -1 + b(i)$. Then for $0 \le i \le n$, let $\mathbf{b}^{(k)}(i) = b(k+i) - b(k)$. Let $[n] = \{1, \cdots, n\}$. We have the following elementary lemma as a variant of the classical ballot theorem, which is a special case of Lemma 3.2.1.

Lemma 4.2.1 (Lemma 6.1 in [61]). *Fix a $1-$lattice bridge $\mathbf{b} = (b(i), 0 \le i \le n)$, and let $r = r(\mathbf{b}) \in [n]$ be minimal so that $b(r) = \min(b(i), i \le n)$. Then $\mathbf{b}^{(r)}$ is a $1-$first passage lattice bridge, and r is the only such value in $[n]$.*

Lemma 4.2.1 is illustrated by Figure 4–1(a) and Figure 4–1(b). In Figure 4–1(a) we have a $1-$lattice bridge $\mathbf{b} = (0, -1, -1, -2, -1, 1, 0, -1)$. The vertical dashed line shows the position of $\mathbf{b}$ attaining its minimum for the first time, hence the unique position for the cyclic shift to transform $\mathbf{b}$ to a $1-$first passage lattice bridge, as claimed by Lemma 4.2.1. The resulting $1-$first passage lattice bridge, with steps $\mathbf{b}^{(3)}$, is shown in Figure 4–1(b).

Recall that for a plane tree T and a node $v \in v(T)$, we write $k_T(v)$ to denote the degree of v in T. We also write $\mathrm{lex}(T) = (k_T(u_1), \dots, k_T(u_{|T|}))$ where $(u_i, 1 \le i \le |T|) = (u_i(T), 1 \le i \le |T|)$ are nodes of T listed in lexicographic order.

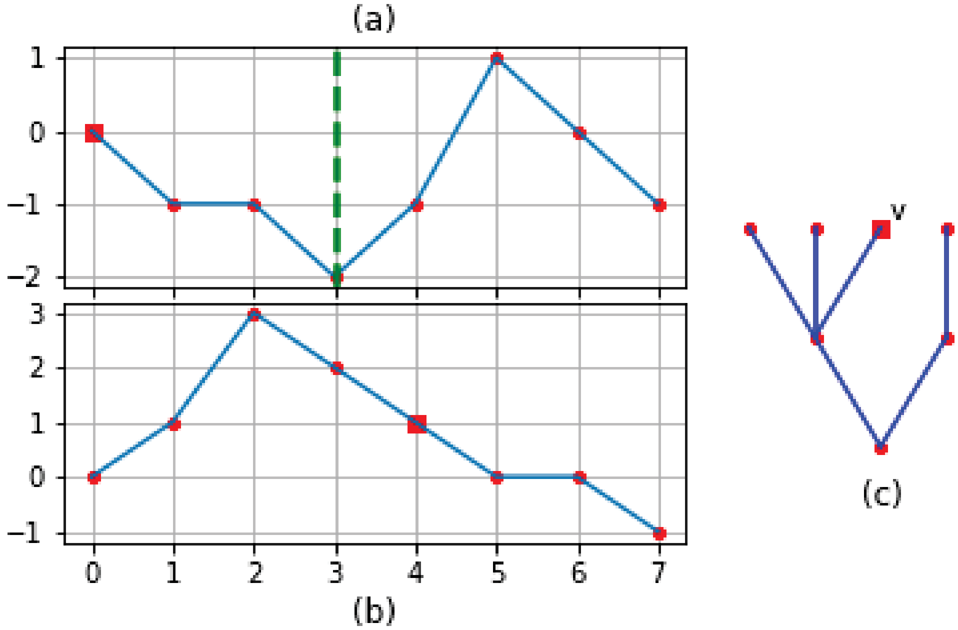

Figure 4–1: (a): a 1−lattice bridge; (b) the corresponding marked 1−first passage lattice bridge; (c) the corresponding marked tree.

For any sequence $\mathbf{c} = (c_1, \cdots, c_n) \in \mathbb{R}^n$, as in Section 3.2, we write $W_{\mathbf{c}}(j) = \sum_{i=1}^{j}(c_i - 1)$ for $j \in [n]$, let $W_{\mathbf{c}}(0) = 0$ and make $W_{\mathbf{c}}$ a continuous function on $[0, n]$ by linear interpolation. A classical bijection between plane trees and 1−first passage lattice bridges associates to a tree T its depth-first walk $(W_{\mathrm{lex}(T)}(i), 0 \leq i \leq n)$, see, e.g. Chapter 6 of [61]. We build on this bijection below.

For a plane tree T and $v \in v(T)$, we call the pair (T, v) a *marked tree* and call v the *mark*. The bijection between 1−first passage lattice bridges and plane trees also leads to a bijection between 1−lattice bridges and marked trees. This bijection, depicted in Figure 4–1, is specified as follows. For a 1−lattice bridge $\mathbf{b}$, let $r = r(\mathbf{b})$ as in Lemma 4.2.1, let $\mathbf{b}' = \mathbf{b}^{(r)}$ be the 1−first passage lattice bridge corresponding to $\mathbf{b}$, and let T be the plane tree with depth-first walk $\mathbf{b}'$. Then the marked node is $v = u_{|T|-r+1}(T)$, the $(|T| - r + 1)$'st node of T in lexicographic order. The mark v is denoted by a red square in Figure 4–1.

Recall that in Section 3.2, a marked forest is a pair (F, v) where F is a plane forest and $v \in v(F)$. We refer v as the mark of (F, v). A *marked cyclic forest* is a marked forest with its mark in its last tree; the name is because we can equivalently view such a forest as having its trees arranged around a cycle.

Fix an integer sequence $W = (W_i : 0 \le i \le n)$ with $W_0 = 0, W_n = -k$, and $W_i - W_{i-1} \ge -1$ for all $1 \le i \le n$. The bijections described above allow us to view W as a marked cyclic forest $(F, v) = (F(W), v(W))$ consisting of $k - 1$ trees and one marked tree, as follows. For integer $b < 0$, let $\tau(b) = \inf\{t \in \mathbb{N} : W_t \le b\}$. For $1 \le j \le k-1$, let T_j be the tree whose depth-first walk is $(W_i - W_{\tau(-(j-1))} : \tau(-(j-1)) \le i \le \tau(-j))$. Let (T_k, v) be the marked tree corresponding to $1-$lattice bridge $(W_i - W_{\tau(-(k-1))} : \tau(-(k-1)) \le i \le n)$. Then $(F(W), v(W)) = ((T_1, \ldots, T_k), v)$. We call W the *coding walk* of the forest, and note that the coding is bijective: W can be recovered from $(F(W), v(W))$ as the concatenation of the $1-$first passage lattice bridges which code $T_1, \ldots, T_{k-1}$ and the $1-$lattice bridge which codes (T_k, v). This bijection is illustrated in Figure 4–2 and Figure 4–3. In Figure 4–2 the whole sequence is decomposed into three segments (divided by vertical dashed lines). The first two segments are $1-$first passage lattice bridges, hence correspond to plane trees T_1, T_2. The last part is a $1-$lattice bridge, hence corresponds to a marked tree (T_3, v) and the node v is again depicted by a square mark. These trees are shown in Figure 4–3.

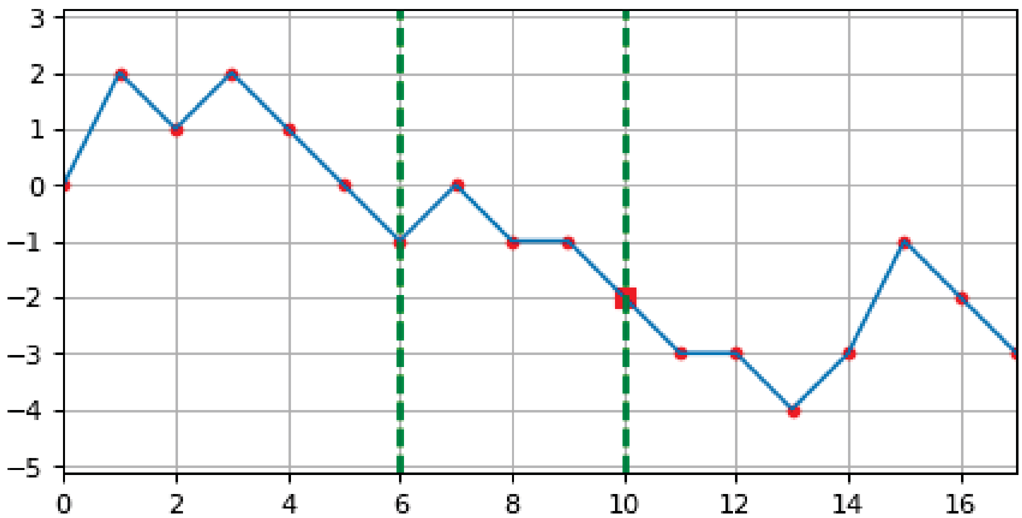

Figure 4–2: A lattice walk $W = (W_i : 0 \leq i \leq 17)$.

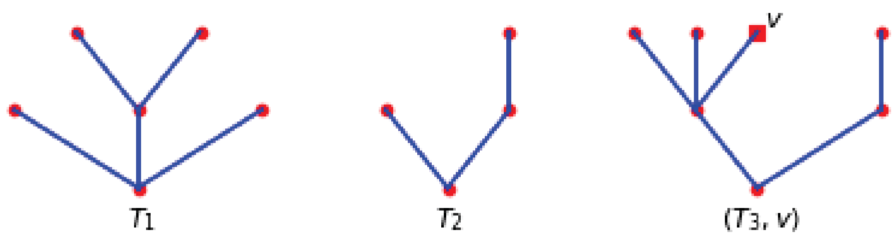

Figure 4–3: The marked forest $(F(W), v(W)) = ((T_1, T_2, T_3), v)$.

Given a degree sequence $\mathbf{s} = (s^{(i)}, i \geq 0)$ with $\sum_{i\geq 0} s^{(i)} = n$, recall that $d(\mathbf{s}) \in \mathbb{Z}^n_{\geq 0}$ is the vector whose entries are weakly increasing and with $s^{(i)}$ entries equal to i, for each $i \geq 0$. Let $\mathrm{D}(\mathbf{s})$ be the set of sequences $\mathbf{d} \in \mathbb{Z}^n_{\geq 0}$ which are permutations of $d(\mathbf{s})$ (there are $n!/(\prod_i s^i!)$ of them). Let $\mathrm{MCF}(\mathbf{s})$ be the set of all marked cyclic forests with degree sequence $\mathbf{s}$. By the correspondence we developed previously, the following is lemma is immediate.

Lemma 4.2.2. *Fix a degree sequence* $\mathbf{s} = (s^{(i)}, i \geq 0)$ *with* $\sum_{i \geq 0} s^{(i)} = n$. *Let* π *be a uniformly random permutation of* $[n]$, *and let* $W = W_{\pi(d(\mathbf{s}))}$. *Then the marked cyclic forest* $(F(W), v(W))$ *coded by* W *is uniformly distributed on* $\mathrm{MCF}(\mathbf{s})$.

In particular, we have the following corollary.

Corollary 4.2.3. *Let* $\mathbf{s} = (s^{(i)}, i \geq 0)$ *with* $\sum_{i \geq 0} s^{(i)} = n$. *Let* (F, v) *be a uniformly random element of* $\mathrm{MCF}(\mathbf{s})$, *and let* M *be the total number of nodes in the non-marked trees of* (F, v). *Let* π *be a uniformly random permutation of* $[n]$ *and let* $S : [0, n] \to \mathbb{R}$, $S(t) = W_{\pi(d(\mathbf{s}))}(t)$. *Then*

$$M \stackrel{d}{=} \inf\{t : S(t) = -c(\mathbf{s}) + 1\}. \tag{4.2.1}$$

We will also need the following easy fact connecting linear forests with marked cyclic forests.

Lemma 4.2.4. *Fix a degree sequence* $\mathbf{s} = (s^{(i)}, i \geq 0)$ *with* $|\mathbf{s}| = n$, *and let* F *be a uniformly random linear forest with degree sequence* $\mathbf{s}$, *and let* (F^*, v) *be the marked cyclic forest obtained from* F *by marking a uniformly random node and applying the requisite cyclic shift of the trees of* F. *Then* (F^*, v) *is a uniformly random element of* $\mathrm{MCF}(\mathbf{s})$.

Proof. Let $\mathrm{F}(\mathbf{s})$ be the set of all plane forests with degree sequence $\mathbf{s}$. The operation of marking a node induces an n-to-$c(\mathbf{s})$ map from $\mathrm{F}(\mathbf{s})$ to $\mathrm{MCF}(\mathbf{s})$, from which the lemma is immediate. □

The preceding lemma allows us to relate the random forest $\mathbb{F}_n$ from Theorem 1.3.4 with the lattice path $S_n = (S_{n,k}, 0 \leq k \leq n)$ from Theorem 4.1.1. Let $(F_n^*, v_n) = ((T_{n,k}, 1 \leq k \leq c_n), v_n)$ be obtained from $\mathbb{F}_n$ by marking a uniformly random node and applying the requisite cyclic shift of the trees of $\mathbb{F}_n$. Then we may couple $\mathbb{F}_n$ and S_n so that $S_n = (S_{n,j}, 0 \le j \le n)$ is the coding walk of (F_n^*, v_n). We work with such a coupling for the remainder of the paper.

4.3 Convergence of the coding processes

The goal of this section is to prove Theorem 4.1.1 and Corollary 4.1.2. To achieve that, we decompose the walk process into two random processes. To be precise, let $d_n := \frac{n^{1/2}}{c_n}$ and fix a sequence $(t_n)_{n\in\mathbb{N}}$, such that $t_n = o(d_n)$ and $t_n = \omega(1)$. This is possible since $d_n \to \infty$ as $n \to \infty$ by our assumption that $c_n = o(n^{1/2})$. We consider the following two processes. Let $(M_{n,k}, k \le n)$ be as follows, $M_{n,0} = 0$ and for $k \ge 1$,

$$M_{n,k} - M_{n,k-1} = X_{n,k}\mathbb{1}_{|X_{n,k}|<t_n}.$$

Similarly, let $(R_{n,k}, k \le n)$ be given by $R_{n,0} = 0$, and for $k \ge 1$,

$$R_{n,k} - R_{n,k-1} = X_{n,k}\mathbb{1}_{|X_{n,k}|\ge t_n}.$$

Then clearly we have $S_{n,k} = M_{n,k} + R_{n,k}$ for all $k \le n$. Define the following quantity:

$$\mu_n^+ := \sum_{j\ge t_n+1}(j-1)\frac{s_n^{(j)}}{n}.$$

Theorem 4.1.1 is an immediate consequence of the following two results:

$$\left(\frac{1}{c_n}(M_{n,\lfloor tc_n^2\rfloor} + \mu_n^+\lfloor tc_n^2\rfloor), t \ge 0\right) \xrightarrow{d} (\sigma B(t), t \ge 0) \tag{4.3.1}$$

and

$$\left(\frac{1}{c_n}(R_{n,\lfloor tc_n^2\rfloor} - \mu_n^+\lfloor tc_n^2\rfloor), t \ge 0\right) \xrightarrow{d} 0, \tag{4.3.2}$$

where 0 denotes a process Z such that $\mathbf{P}\{Z(t)=0,\ \forall t\geq 0\}=1$. For (4.3.1), we are going to use the following theorem from [26].

Theorem 4.3.1 (Theorem 4 in [26])**.** *Suppose an urn U contains n balls, each marked by one or another element of the set S, whose cardinality c is finite. Let H_{Uk} be the distribution of k draws made at random without replacement from U, and M_{Uk} be the distribution of k draws made at random with replacement. Then the two probabilities on S^k satisfy*

$$||H_{Uk}-M_{Uk}||\leq 2ck/n,$$

where $||\cdot||$ denotes the total variation distance.

Proof of (4.3.1). Let $(\tilde{X}_{n,k}, k\leq n)$ be i.i.d. with the law of $X_{n,1}\mathbb{1}_{|X_{n,1}|<t_n}$, set $\tilde{M}_{n,0}=0$ and for $k\geq 1$, let

$$\tilde{M}_{n,k}=\sum_{j=1}^{k}\tilde{X}_{n,j}.$$

Now apply Theorem 4.3.1 with urn U containing n balls, with $s_n^{(j)}$ balls marked by $j-1$ for $0\leq j\leq t_n, j\neq 1$, and $s_n^{(1)}+\sum_{j>t_n}s_n^{(j)}$ balls marked by 0, with $S=\{-1,0,1,\dots,t_n-1\}$, and with $k=k(n)=n/d_n$. This yields that

$$||(X_{n,j}\mathbb{1}_{|X_{n,j}|<t_n}, j\leq k)-(\tilde{X}_{n,j}, j\leq k)||\leq\frac{2t_n(n/d_n)}{n}=2\frac{t_n}{d_n},$$

so for all Borel $B\subset\mathbb{R}^k$,

$$\left|\mathbf{P}\{(M_{n,j}, j\leq k(n))\in B\}-\mathbf{P}\left\{(\tilde{M}_{n,j}, j\leq k(n))\in B\right\}\right|\leq\frac{2t_n}{d_n}.$$

Since $t_n = o(d_n)$ and $k(n) = n/d_n = d_n \cdot c_n^2 = \omega(c_n^2)$, this implies that to establish (4.3.1) it suffices to prove that

$$\left(\frac{1}{c_n} (\tilde{M}_{n,\lfloor tc_n^2 \rfloor} + \mu_n^+ \lfloor tc_n^2 \rfloor), t \geq 0 \right) \xrightarrow{d} (\sigma B(t), t \geq 0) \,. \tag{4.3.3}$$

Note that

$$\mathbf{E}\tilde{X}_{n,1} = \sum_{j \leq t_n} (j-1) \frac{s_n^{(j)}}{n} = \frac{1}{n} \sum_j (j-1) s_n^{(j)} - \sum_{j \geq t_n + 1} (j-1) \frac{s_n^{(j)}}{n} = -\frac{c_n}{n} - \mu_n^+.$$

Define σ_n^- by setting

$$\begin{aligned} (\sigma_n^-)^2 := \mathbf{Var}\left\{ \tilde{X}_{n,1} \right\} &= \mathbf{E}\left[\tilde{X}_{n,1}^2 \right] - \mathbf{E}\left[\tilde{X}_{n,1} \right]^2 \\ &= \sum_{j \leq t_n} (j-1)^2 \frac{s_n^{(j)}}{n} - (-\mu_n^+ - \frac{c_n}{n})^2. \end{aligned} \tag{4.3.4}$$

Applying Donsker's theorem to the process $\left(\tilde{M}_{n,k} + k(\mu_n^+ + \frac{c_n}{n}), k \geq 0 \right)$, we obtain that

$$\left(\frac{1}{a} (\tilde{M}_{n,\lfloor ta^2 \rfloor} + \mu_n^+ \lfloor ta^2 \rfloor) + \frac{c_n \lfloor ta^2 \rfloor}{na}, t \geq 0 \right) \xrightarrow{d} \left(\sigma_n^- B(t), t \geq 0 \right), \tag{4.3.5}$$

as $a \to \infty$.

By our assumption that $n^{-1} \cdot \mathbf{s}_n \to \mathbf{p}$ in L^2 in Theorem 1.3.4, we have $\sum_j (j - 1) \frac{s_n^{(j)}}{n} \to 0$ as $n \to \infty$. Hence for any prescribed $\delta > 0$, we can find L large such that $\sum_{j > L} (j-1) \frac{s_n^{(j)}}{n} < \delta$. Since $t_n \to \infty$, we must have $\mu_n^+ \leq \sum_{j > L} (j-1) \frac{s_n^{(j)}}{n} < \delta$ for n large enough, i.e.

$$\mu_n^+ \to 0 \text{ as } n \to \infty. \tag{4.3.6}$$

Similarly the assumption that $n^{-1} \cdot \mathbf{s}_n \to \mathbf{p}$ in L^2 implies that

$$\sigma_n^2 := \sum_j j(j-1)\frac{s_n^{(j)}}{n} \to \sigma^2 < \infty,$$

so

$$(\sigma_n^+)^2 \to 0 \text{ as } n \to \infty, \tag{4.3.7}$$

where we let $(\sigma_n^+)^2 := \sum_{j \geq t_n+1} j(j-1)\frac{s_n^{(j)}}{n}$. Using (4.3.4),(4.3.6) and (4.3.7), we have

$$\sigma_n^2 - (\sigma_n^-)^2 = (\sigma_n^+)^2 - (\mu_n^+ + \frac{c_n}{n})(1 - \mu_n^+ - \frac{c_n}{n}) \to 0 \text{ as } n \to \infty, \tag{4.3.8}$$

so $\sigma_n^- \to \sigma$ as $n \to \infty$. Taking $a = c_n$ in (4.3.5), then letting $n \to \infty$, now yields that

$$\left(\frac{1}{c_n}(\tilde{M}_{n,\lfloor tc_n^2 \rfloor} + \mu_n^+ \lfloor tc_n^2 \rfloor) + \frac{\lfloor tc_n^2 \rfloor}{n}, t \geq 0\right) \xrightarrow{d} (\sigma B(t), t \geq 0)\,.$$

Since $c_n^2 = o(n)$, (4.3.3) follows. $\square$

To prove (4.3.2), we will again use Proposition 3.3.5.

Proof of (4.3.2). We prove that for all $\epsilon > 0$, we have

$$\limsup_{n \to \infty} \mathbf{P}\left\{\max_{i \leq c_n^2/\epsilon} \left|\frac{R_{n,i} - i\mu_n^+}{c_n}\right| > \epsilon\right\} \leq \epsilon,$$

this immediately implies (4.3.2). Fix n and let $\mathrm{c}_1, \ldots, \mathrm{c}_n$ be such that $|\{1 \leq k \leq n : \mathrm{c}_k = j\}| = s_n^{(j)}$. Let $C_1, \ldots, C_n$ be a uniformly random permutation of $\mathrm{c}_1, \ldots, \mathrm{c}_n$. Fix $t_n \in \mathbb{N}$. Define $(R_i, 0 \leq i \leq n)$ as follows: let $R_0 = 0$, and for $i \geq 0$, let

$$R_{i+1} = \begin{cases} R_i + C_i - 1, & \text{if } C_i \geq t_n + 1; \\ R_i, & \text{if } C_i \leq t_n. \end{cases}$$

For $0 \le i \le n$, let $\mathcal{F}_i = \sigma(C_1, \dots, C_i)$. Since $R_n = n\mu_n^+$ and the process $(R_i, 0 \le i \le n)$ has exchangeable increment,

$$\mathbf{E}\left[R_{i+1} \mid \mathcal{F}_i\right] = R_i + \frac{n\mu_n^+ - R_i}{n-i}. \tag{4.3.9}$$

Now let $K_i = \frac{n\mu_n^+ - R_i}{n-i}$. Then using (4.3.9), we have

$$\mathbf{E}\left[K_{i+1} \mid \mathcal{F}_i\right] = \frac{n\mu_n^+ - R_i}{n-(i+1)} - \frac{n\mu_n^+ - R_i}{(n-i)(n-i+1)} = K_i.$$

Hence K_i is an $\mathcal{F}_i-$martingale.

Since for any $0 \le i \le s$,

$$\frac{n\mu_n^+ - R_i}{n-i} = \mu_n^+ + \frac{i\mu_n^+ - R_i}{n-i},$$

and μ_n^+ is a constant, if we define $\tilde{K}_i = \frac{i\mu_n^+ - R_i}{n-i}$, then $\tilde{K}_i$ is also an $\mathcal{F}_i-$martingale. It follows that for any $\epsilon > 0$,

$$\begin{aligned}
\mathbf{P}\left\{\frac{1}{c_n}\max_{i\le s}|i\mu_n^+ - R_i| > \epsilon\right\} &\le \frac{n^2}{\epsilon^2 c_n^2}\mathbf{E}\left[\left(\max_{i\le s}\frac{|i\mu_n^+ - R_i|}{n-i}\right)^2\right] \\
&\le \frac{4n^2\mathbf{E}\left[(s\mu_n^+ - R_s)^2\right]}{\epsilon^2 c_n^2 (n-s)^2},
\end{aligned} \tag{4.3.10}$$

where in the first line we use Markov's inequality and in the last line we use the L^2 maximal inequality for martingales (see, e.g. Theorem 5.4.3 in [29]).

Since the process $(R_s, 0 \le s \le n)$ has exchangeable increments, we have $\mathbf{E}R_s = s\mu_n^+$. Let $R_s^* = \sum_{i\le s} J_i$ where $J_1, \dots, J_s$ are i.i.d. random variables with $J_1 \stackrel{d}{=} R_1$. Then

Proposition 3.3.5 gives

$$\mathbf{E}\left[R_s^2\right] \leq \mathbf{E}\left[R_s^{*2}\right] = \mathbf{E}\left[(J_1 + \cdots + J_s)^2\right] \;=\; s\mathbf{E}\left[J_1^2\right] + s(s-1)(\mathbf{E}J_1)^2$$
$$= s(\sigma_n^{+2} - \mu_n^+) + s(s-1)\mu_n^{+2}$$

Therefore,

$$\mathbf{E}\left[(s\mu_n^+ - R_s)^2\right] = \mathbf{E}\left[R_s^2\right] - s^2\mu_n^{+2} \leq s(\sigma_n^{+2} - \mu_n^+) - s\mu_n^{+2} \leq s\sigma_n^{+2}.$$

Now take $s = s(n) = c_n^2/\epsilon$ in (4.3.10). For n large this is less than $n/2$, so $(n-s)^2 > n^2/4$ and we obtain

$$\mathbf{P}\left\{\frac{1}{c_n}\max_{i \leq c_n^2/\epsilon} |i\mu_n^+ - R_i| > \epsilon\right\} \leq \frac{16s\sigma_n^{+2}}{\epsilon^2 c_n^2} = \frac{16\sigma_n^{+2}}{\epsilon^3} \leq \epsilon,$$

the last inequality holding for n large since $\sigma_n^+ \to 0$ as $n \to \infty$. This completes the proof. $\square$

Recall that in Section 4.1 we let $\tau(x) = \inf(t : B(t) \leq -x)$ for $x \geq 0$. By (4.2.1) if we let $\tau_n = \sum_{1 \leq i < c_n} |T_{n,i}| = n - |T_{n,c_n}|$ be the total size of non-marked trees of (F_n^*, v_n), then since S_n is the coding process of (F_n^*, v_n) we have

$$\tau_n = \inf\{k : S_{n,k} = -(c_n - 1)\}.$$

From this we immediately get the following corollary of Theorem 4.1.1.

Corollary 4.3.2. *Given the assumptions in Theorem 1.3.4, we have*

$$\frac{\tau_n}{c_n^2} \xrightarrow{d} \tau(\frac{1}{\sigma}), \tag{4.3.11}$$

where $(B(t), t \geq 0)$ is standard Brownian Motion.

Remark 4.3.1. *Note that the right-hand side of (4.3.11) has density*

$$\frac{1}{\sigma\sqrt{2\pi t^3}}\exp\left(-\frac{1}{2t\sigma^2}\right)dt\,;$$

see, e.g., Theorem 6.9 in [62]. This coincides with what we have in Theorem 1.4.5.

The corollary above in fact tells us something about the size of the largest tree $\mathbb{T}_{n,1}$.

Corollary 4.3.3. *For a marked cyclic forest (F,v), let $MT(F,v)$ denoted the marked tree, i.e. the tree of F containing v. Then*

$$\mathbf{P}\left\{MT(F_n^*,v_n)=\mathbb{T}_{n,1}\right\}\to 1$$

as $n\to\infty$.

Proof. It is clear that

$$\begin{aligned}\mathbf{P}\left\{MT(F_n^*,v_n)\neq\mathbb{T}_{n,1}\right\} &\leq \mathbf{P}\left\{|MT(F_n^*,v_n)|<n/2\right\}\\ &= \mathbf{P}\left\{\tau_n>n/2\right\}=\mathbf{P}\left\{\frac{\tau_n}{c_n^2}>\frac{n}{2c_n^2}\right\}\to 0\end{aligned}$$

where in the last line, the first equation is by Lemma 4.2.4 and the final convergence is by Corollary 4.3.2 and the assumption $c_n^2=o(n)$. □

Now we are ready to prove Corollary 4.1.2.

Proof of Corollary 4.1.2. As noted, it suffices to prove (4.1.2) and (4.1.3). Corollary 4.3.2 and Corollary 4.3.3 together imply (4.1.2).

For (4.1.3), first note that by Lemma 4.2.2, the process $S_n=(S_{n,k},0\leq k\leq n)$ has the same law as the coding walk $W(\mathbb{F}_n)$ of $\mathbb{F}_n$. Applying Corollary 4.3.3 then yields that the law of $(|\mathbb{T}_{n,2}|, \ldots, |\mathbb{T}_{n,j}|)$ is asymptotically equivalent to the law of $(d_1^n - g_1^n, \ldots, d_{j-1}^n - g_{j-1}^n)$, the first $j-1$ ranked excursion lengths of S_n above its running minimum before time τ_n. Using this equivalence, (4.1.3) now follows from Theorem 4.1.1 by the Portmanteau Theorem ([57], Theorem 12.6), since the vector $(d_1 - g_1, \ldots, d_{j-1} - g_{j-1})$ has a density. $\square$

4.4 Empirical degree sequences of trees

In this section we aim to prove (4.1.4) and (4.1.5).

For $i \geq 0$ and $x \leq n$, let

$$Q_n^i(x) := |\{1 \leq j \leq x : C_{n,j} = i\}|$$

where $(C_{n,1}, \cdots, C_{n,n})$ is a uniformly random permutation of $d(\mathbf{s}_n)$ and that $S_{n,k} = \sum_{j=1}^{k}(C_{n,j} - 1)$. Let $\mathcal{F}_j = \sigma(C_{n,1}, \ldots, C_{n,j})$. Since $Q_n^i(n) = s_n^{(i)} = np_n^{(i)}$ and the process $(Q_n^i(j), 0 \leq j \leq n)$ has exchangeable increments,

$$\mathbf{E}\left[Q_n^i(j+1) \mid \mathcal{F}_j\right] = Q_n^i(j) + \frac{np_n^{(i)} - Q_n^i(j)}{n-j}.$$

Setting $K_j = \frac{np_n^{(i)} - Q_n^i(j)}{n-j}$ for $0 \leq i \leq n$, then

$$\mathbf{E}\left[K_{j+1} \mid \mathcal{F}_j\right] = \frac{np_n^{(i)} - Q_n^i(j)}{n-(j+1)} - \frac{np_n^{(i)} - Q_n^i(j)}{(n-(j+1))(n-j)} = K_j,$$

so K_j is an $\mathcal{F}_j-$martingale. If we let $\tilde{K}_j = \frac{jp_n^{(i)} - Q_n^i(j)}{n-j}$, then $\tilde{K}_j = K_j - p_n^{(i)}$, so $\tilde{K}_j$ is also an $\mathcal{F}_j-$martingale.

We now use the martingale bound Theorem 2.3.2. We shall apply this theorem to bound the fluctuations of $Q_n^i(s)$.

Proposition 4.4.1. *For any $0 < t < 1$, we have*

$$\mathbf{P}\left\{\exists s > c_n : \left|p_n^{(i)} - \frac{Q_n^i(s)}{s}\right| \geq t\right\} \leq \exp\left(-\frac{3t^2 c_n}{5}\right). \tag{4.4.1}$$

Proof. It is not hard to show that for any $0 \leq j \leq n-2$,

$$\mathbf{Var}\left\{\tilde{K}_{j+1} \mid \mathcal{F}_j\right\} \leq \frac{1}{4} \cdot \frac{1}{(n-(j+1))^2};$$

see Lemma 3.3.1. Thus, for $1 \leq x \leq n-2$,

$$\begin{aligned} V = \sum_{j=0}^{x-1} \mathbf{Var}\left\{\tilde{K}_{j+1} \mid \mathcal{F}_j\right\} &\leq \frac{1}{4} \sum_{j=0}^{x-1} \frac{1}{(n-(j+1))^2} \\ &\leq \frac{1}{4} \int_{n-x-1}^{n-1} \frac{1}{m^2} \mathrm{d}m = \frac{x}{4(n-1)(n-x-1)}. \end{aligned}$$

On the other hand, for $0 \leq j \leq x-1$, if $Q_n^i(j+1) = Q_n^i(j)$, then

$$|\tilde{K}_{j+1} - \tilde{K}_j| = \left|\frac{np_n^{(i)} - Q_n^i(j)}{(n-(j+1))(n-j)}\right| \leq \frac{1}{n-x},$$

while if $Q_n^i(j+1) = Q_n^i(j) + 1$, then

$$|\tilde{K}_{j+1} - \tilde{K}_j| = \left|\frac{np_n^{(i)} - Q_n^i(j)}{(n-(j+1))(n-j)} - \frac{1}{n-(j+1)}\right| \leq \frac{1}{n-x}.$$

Applying Theorem 2.3.2 to both $\{\tilde{K}_j\}_{j=0}^x$ and $\{-\tilde{K}_j\}_{j=0}^x$ with $x = n - c_n$, we have

$$v \leq \frac{1}{2c_n}, \ b \leq \frac{1}{c_n}.$$

Hence, for $t \leq 1$,

$$\mathbf{P}\left\{\max_{0\leq j \leq n-c_n} \left|p_n^{(i)} - \frac{np_n^{(i)} - Q_n^i(j)}{n-j}\right| \geq t\right\} \leq \exp\left(-\frac{t^2}{\frac{1}{c_n} + \frac{2t}{3c_n}}\right) \leq \exp\left(-\frac{3t^2 c_n}{5}\right).$$

Using the exchangeability of $C_{n,1},\ldots,C_{n,n}$, it follows that

$$\begin{aligned}\mathbf{P}\left\{\exists s>c_n : |p_n^{(i)}-\frac{Q_n^i(s)}{s}|\geq t\right\} &= \mathbf{P}\left\{\max_{0\leq j\leq n-c_n}\left|p_n^{(i)}-\frac{np_n^{(i)}-Q_n^i(j)}{n-j}\right|\geq t\right\}\\ &\leq \exp\left(-\frac{3t^2c_n}{5}\right).\end{aligned}$$

□

We next give the proofs of (4.1.4) and (4.1.5). In both proofs we use the coupling between $\mathbb{F}_n$, (F_n^*, v_n) and S_n explained at the end of Section 4.2.

Proof of (4.1.4). Fix $i\geq 0$ and $l\geq 2$. By Corollary 4.3.3, with high probability $\mathbb{T}_{n,1}=T_{n,c_n}$, i.e., $\mathbb{T}_{n,1}$ is the last tree of (F_n^*, v_n), in which case $\mathbb{T}_{n,l}=T_{n,j}$ for some $j<c_n$. Recall that $\tau_n=\sum_{1\leq k<c_n}|T_{n,k}|$.

Let $1\leq j<c_n$, and suppose $|\{v\in T_{n,j}: k(v)=i\}|/|T_{n,j}|\notin[p_n^{(i)}-\delta, p_n^{(i)}+\delta]$. Suppose that $|T_{n,j}|>\delta c_n^2>c_n$ and $\tau_n<c_n^3$. Then there must exist $m>c_n$ and $1\leq u\leq \tau_n-m$ such that

$$\left|\frac{|\{t\in[m]: C_{n,u+t}=i\}|}{m}-p_n^{(i)}\right|>\delta.$$

By union bound and the exchangeability of $(C_{n,1},\ldots,C_{n,n})$, the probability of this is bounded above by $\tau_n\mathbf{P}\left\{\exists m>c_n : \left|\frac{Q_n^i(m)}{m}-p_n^{(i)}\right|>\delta\right\}$. Thus, for $l\geq 2$, for n large enough that $\delta c_n^2>c_n$, we have

$$\begin{aligned}\mathbf{P}\left\{\left|p_{n,l}^{(i)}-p_n^{(i)}\right|>\delta\right\} \leq{}& \mathbf{P}\left\{\tau_n>c_n^3\right\}+\mathbf{P}\left\{|\mathbb{T}_{n,l}|<\delta c_n^2\right\}\\ &+\mathbf{P}\left\{\mathbb{T}_{n,1}\neq T_{n,c_n}\right\}+c_n^3\mathbf{P}\left\{\max_{s>c_n}\left|p_n^{(i)}-\frac{Q_n^i(s)}{s}\right|>\delta\right\}\end{aligned}$$

For any $\epsilon > 0, \mathbf{P}\{\tau_n > c_n^3\} < \epsilon/3$ by Corollary 4.3.2 for n large enough, and $\mathbf{P}\{|\mathbb{T}_{n,l}| < \delta c_n^2\} < \epsilon/3$ by Corollary 4.1.2. The second last probability tends to zero by Corollary 4.3.3. And for the last probability, for n large enough, $\sqrt{\frac{5}{3}} c_n^{-1/3} < \delta$, hence Proposition 4.4.1 gives upper bound $c_n^3 \exp\left(-c_n^{1/3}\right)$, which tends to zero. Thus, $\mathbf{P}\left\{\left|p_{n,l}^{(i)} - p_n^{(i)}\right| > \delta\right\} < \epsilon$ for n large; this proves (4.1.4) for $i \geq 0$ and $l > 1$.

Finally, since $|\mathbb{T}_{n,1}|/n \to 1$, the fact that $\left|p_{n,1}^{(i)} - p_n^{(i)}\right| \to 0$ in probability for each $i \geq 0$ is immediate. □

Proof of (4.1.5). Fix $\epsilon > 0$. By Corollary 4.3.2, we can pick $M > 0$ large enough such that for n large enough,

$$\mathbf{P}\left\{\tau_n > M c_n^2\right\} < \epsilon. \tag{4.4.2}$$

By Corollary 4.3.3, we have that for n large enough,

$$\mathbf{P}\{T_{n,c_n} \neq \mathbb{T}_{n,1}\} < \epsilon\,. \tag{4.4.3}$$

For this $\epsilon > 0$, there exists $\delta > 0$ such that $\mathbf{P}\{g_{l-1} - d_{l-1} \leq \delta\} < \epsilon/2$, so by Corollary 4.1.2, for n large

$$\mathbf{P}\left\{\frac{|\mathbb{T}_{n,l}|}{c_n^2} \leq \delta\right\} < \epsilon. \tag{4.4.4}$$

Next we fix $t > 0$ large enough such that

$$\mathbf{E}\left[C_{n,1}^2 \mathbb{1}_{C_{n,1} \geq t}\right] < \frac{\epsilon^2 \delta}{M} \text{ and } \sum_{i > t} i^2 p_n^{(i)} < \epsilon\,; \tag{4.4.5}$$

this is possible since $\mathbf{p}_n = (p_n^{(i)}, i \geq 0) \to \mathbf{p} = (p^{(i)}, i \geq 0)$ in L^2. For fixed $l \geq 2$ we have

$$\begin{aligned} |\sigma^2(\mathbf{p}_{n,l}) - \sigma^2(\mathbf{p}_n)| &\leq |\sum_{i \leq t} i^2(p_{n,l}^{(i)} - p_n^{(i)})| + \sum_{i>t} i^2 p_n^{(i)} + \sum_{i>t} i^2 p_{n,l}^{(i)} \\ &\leq |\sum_{i \leq t} i^2(p_{n,l}^{(i)} - p_n^{(i)})| + \epsilon + \sum_{i>t} i^2 p_{n,l}^{(i)} \end{aligned} \tag{4.4.6}$$

where we use (4.4.5) in the second line.

Let $L_n = \sum_{j \leq Mc_n^2} C_{n,j}^2 \mathbb{1}_{C_{n,j} \geq t}$. If $T_{n,c_n} = \mathbb{T}_{n,1}$ and $\tau_n \leq Mc_n^2$ then $\sum_{i>t} i^2 p_{n,l}^{(i)} \leq L_n/|\mathbb{T}_{n,l}|$. Hence

$$\begin{aligned} &\mathbf{P}\left\{|\sigma^2(\mathbf{p}_{n,l}) - \sigma^2(\mathbf{p}_n)| \geq 3\epsilon\right\} \\ \leq\ &\mathbf{P}\left\{|\sigma^2(\mathbf{p}_{n,l}) - \sigma^2(\mathbf{p}_n)| \geq 3\epsilon, \tau_n \leq Mc_n^2, T_{n,c_n} = \mathbb{T}_{n,1}, \frac{|\mathbb{T}_{n,l}|}{c_n^2} > \delta\right\} \\ +\ &\mathbf{P}\left\{\tau_n > Mc_n^2\right\} + \mathbf{P}\left\{T_{n,c_n} \neq \mathbb{T}_{n,1}\right\} + \mathbf{P}\left\{\frac{|\mathbb{T}_{n,l}|}{c_n^2} \leq \delta\right\} \\ \leq\ &\mathbf{P}\left\{|\sum_{i \leq t} i^2(p_{n,l}^{(i)} - p_n^{(i)})| \geq \epsilon\right\} + \mathbf{P}\left\{\frac{L_n}{|\mathbb{T}_{n,l}|} > \epsilon, \frac{|\mathbb{T}_{n,l}|}{c_n^2} > \delta\right\} + 3\epsilon \end{aligned} \tag{4.4.7}$$

where we use (4.4.2), (4.4.3), (4.4.4), (4.4.6) and the aforementioned stochastic dominance in the last line.

Since t is fixed, we can use (4.1.4) to conclude that the first summand of (4.4.7) can be made arbitrarily small by taking n large enough. For the second summand, note that by exchangeability and (4.4.5),

$$\mathbf{E}L_n = Mc_n^2 \mathbf{E}\left[C_{n,1}^2 \mathbb{1}_{C_{n,1} \geq t}\right] < c_n^2 \epsilon^2 \delta\,,$$

so

$$\mathbf{P}\left\{\frac{L_n}{|\mathbb{T}_{n,l}|} > \epsilon, \frac{|\mathbb{T}_{n,l}|}{c_n^2} > \delta\right\} \le \mathbf{P}\left\{\frac{L_n}{c_n^2} > \epsilon\delta\right\} \le \frac{\mathbf{E}\left[\frac{L_n}{c_n^2}\right]}{\epsilon\delta} < \epsilon.$$

This completes the proof of (4.1.5) for $l \ge 2$. Again since $|\mathbb{T}_{n,1}|/n \to 1$, (4.1.5) is immediate for $l = 1$ case. $\square$

CHAPTER 5
Conclusions

5.1 Summary of Main Results

In this thesis, we study the scaling limits of a sequence of uniformly random plane forests $\mathbb{F}_n$ with given degree sequences $\mathbf{s}_n$. For the convenience of notation simpleness, we assume that $|\mathbb{F}_n| = |\mathbf{s}_n| = n$. In general we can have degree sequence $\mathbf{s}_\kappa$ depending on some index κ (as in [20]). Then as long as $|\mathbf{s}_\kappa| \to \infty$ as $\kappa \to \infty$, if we replace n by $|\mathbf{s}_\kappa|$, our conclusions still hold as we consider the limit as $\kappa \to \infty$. Under reasonable conditions on $(\mathbf{s}_n)_{n\geq 1}$, we explicitly prove the convergence of $\mathbb{F}_n$ in two different regimes.

In the case that the number of trees c_n satisfies $\frac{c_n}{\sigma_n n^{1/2}} \to \lambda > 0$, we show that the ranked forest $\mathcal{F}_n^{\downarrow}$, viewed as a measured metric space, converges to the real trees coded by the ranked excursions of first passage bridge F_λ^{br} reflected at its minima, in the sense of coordinatewise GHP convergence. With a stronger condition on the maximum degrees Δ_n, the result can be strengthened to the the convergence in the sense of $(\mathbb{L}_\infty, d_{GHP}^\infty)$.

In the case that the number of trees $c_n = o(n^{1/2})$, we confirm the picture that there is a giant tree $\mathbb{T}_{n,1}$ containing all but a negligible fraction of all vertices (hence naturally has scaling limit $\mathcal{T}$). Moreover, for the forest $\hat{\mathcal{F}}_n$ containing all small trees, we give a finer calculation of the limit of size of $\hat{\mathcal{F}}_n$ and under a different scaling, prove the convergence of $\hat{\mathcal{F}}_n$ to a sequence of real trees coded by the excursions of Brownian motion reflected at its minima, run until the local time at zero reaches $1/\sigma$.

5.2 Summary of Methodologies

The first technical aspect is based on using the depth first walk to show the walk convergence and, as a corollary, proving the convergence of the sizes of the tree components. This is achieved by first proving certain combinatorial bijection results, which couple uniformly permuted degree sequences with random forests. Namely, in Chapter 3, the coupling is via a random rotation map; and in Chapter 4, we connect the depth first walk with random forests using the concept of marked cyclic forests. The second component of our work is to show that for each tree component of the forest, the empirical degree sequences behave well, i.e. is close to the "expected" degree sequence governed by the sequence $\mathbf{s}_n/n$. For this part, we use concentration inequalities proved via the martingale difference method from [54].

5.3 Potential Future Works

There are a few possible future research directions following this work. First, it would be nice to get rid of the extra maximum degree condition in Theorem 1.3.2. This condition is used in Proposition 3.3.8, where we proved a technical proposition of bounding the probability of $\mathbb{F}(\mathbf{s})$ containing trees with unusually large height. The second natural direction is to work in the subcritical regime that the number of trees $c_n = \omega(n^{1/2})$. In this regime we would expect no trees of linear order. But the detailed picture of scaling limits is not very clear. In fact, the possible behaviours are more complex in this case. A natural guess is that the largest tree will have size of order $\left(\frac{n}{c_n}\right)^2 \log \frac{c_n^2}{n}$. This guess arises from the following heuristic. The search process for the forest $\mathbb{F}_n$ should look roughly like a random walk with drift $-c_n/n$, run for n steps. At time t, the expected position of such walk is $-tc_n/n$, and the fluctuations of its position have order $\sqrt{t}$. Thus, the random walk typically hits a new local minimum in $O((n/c_n)^2)$ steps. Since new running minima corresponding to the times when tree explorations conclude, this suggests that most trees are not larger than $O((n/c_n)^2)$. The logarithmic factor arises because an interval of length n may be divided into c_n^2/n intervals of length $(n/c_n)^2$, and at some point one expects to see an "exceptional" run of $O(\log \frac{c_n^2}{n})$ consecutive intervals of length $(n/c_n)^2$ on which the walk does not reach a new running minimum. This heuristic is correct in some cases (e.g. Theorem 2.1.3 in [59], see below), but the full story will necessarily be more complicated. To see this, consider for example a binary forest with n nodes and with $c_n \cong n^{5/6}$ trees. If we add a single node of degree $\lfloor n^{2/5} \rfloor$ to such a degree sequence, then we obtain a random forest with $c_n + 1 - \lfloor n^{2/5} \rfloor \cong n^{5/6}$ trees. However, the size of the largest tree will be at least $n^{2/5}$, whereas the heuristic would suggest a largest tree of size $O(n^{1/3} \log n)$.

Finally, we would like to include Pavlov's result in the subcritical regime and an example to showcase that the aforementioned heuristic gives the same result as his result. Let $F(z) = \sum_{k=0}^{\infty} p_k z^k$ be the probability generating functions and recall as in Section 1.4.4 that η denotes the size of the largest tree of $\mathfrak{F}_{N,n}$.

Theorem 5.3.1 (Theorem 2.1.3 in [59])**.** *Let $F'''(1) < \infty, N, n \to \infty$ in such a way that n takes values which are divided by d. Assume $n/N \to \infty, n/N^2 \to 0$, let λ be defined by*

$$\frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{N+n}.$$

Then

$$\mathbf{P}\left\{\beta\eta - \mu \le z\right\} \to e^{-e^{-z}}$$

where $\beta = \beta(\lambda) = -\ln(\lambda/F(\lambda))$ *and* $u = u(\lambda)$ *is chosen so that*

$$N\beta^{1/2}u^{-3/2}e^{-u} = \sqrt{2\pi\sigma^2}.$$

Example 5.3.1 (Poisson case)**.** *Let* n *be number of non-rooted vertices and* m *be number of vertices. Assume* $c_m = m^{5/6}$ *so* $n = m - m^{5/6}$*.*

For simplicity of calculation, suppose $\xi \sim Poisson(1)$ *so* $p_k = \frac{e^{-1}}{k!}$*. Then* $F(z) = e^{z-1}$ *hence* $F'(z) = e^{z-1}$*. So*

$$\lambda = \frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{m} = 1 - m^{-1/6}.$$

Let $\epsilon = m^{-1/6}$*, then*

$$F(\lambda) = F(1-\epsilon) = e^{-\epsilon} \approx 1 - \epsilon + \frac{\epsilon^2}{2}.$$

Then

$$\beta = -\ln(\lambda/F(\lambda)) = -\ln\left(\frac{1-\epsilon}{1-\epsilon+\frac{\epsilon^2}{2}}\right) = -\ln\left(1 - \frac{\epsilon^2/2}{1-\epsilon+\epsilon^2/2}\right) \approx \epsilon^2/2.$$

So $\beta \approx m^{-1/3}$*.*

By the definition of u:

$$m^{5/6} \cdot m^{-1/6} \cdot u^{-3/2}e^{-u} = \sqrt{2\pi}.$$

So $m^{2/3} \approx e^u$*, that is,*

$$u \approx \log m^{2/3} = \log(\frac{c_m^2}{m}).$$

$\mathbf{P}\{\beta\eta - u \leq z\} \to e^{-e^{-z}}$ *implies* $\eta \sim (z+u)\beta^{-1}$, $u\beta^{-1}$ *gives order*

$$\log(\frac{c_m^2}{m}) \cdot m^{1/3} = \log(\frac{c_m^2}{m}) \cdot (\frac{m}{c_m})^2.$$

References

[1] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18:no. 14, 21, 2013.

[2] Romain Abraham, Jean-François Delmas, and Patrick Hoscheit. Exit times for an increasing Lévy tree-valued process. *Probab. Theory Related Fields*, 159(1-2):357–403, 2014.

[3] L. Addario-Berry. Tail bounds for the height and width of a random tree with a given degree sequence. *Random Structures Algorithms*, 41(2):253–261, 2012.

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

[5] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields*, 152(3-4):367–406, 2012.

[6] David Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

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

[8] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[9] David Aldous and Jim Pitman. The standard additive coalescent. *Ann. Probab.*, 26(4):1703–1726, 1998.

[10] David J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.

[11] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York-Heidelberg, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[12] Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978.

[13] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001.

[14] Jean Bertoin, Loï c Chaumont, and Jim Pitman. Path transformations of first passage bridges. *Electron. Comm. Probab.*, 8:155–166, 2003.

[15] Jérémie Bettinelli. Scaling limits for random quadrangulations of positive genus. *Electron. J. Probab.*, 15(52), 2010.

[16] Jérémie Bettinelli. Scaling limit of random planar quadrangulations with a boundary. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(2):432–477, 2015.

[17] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[18] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

[19] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.

[20] Nicolas Broutin and Jean-François Marckert. Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures Algorithms*, 44(3):290–316, 2014.

[21] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[22] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms*, 21(1):76–119, 2002.

[23] I. A. Cheplyukova. The emergence of a giant tree in a random forest. *Discrete Math. Appl*, 8(1):17–33, 1998.

[24] Herman Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statistics*, 23:493–507, 1952.

[25] Nicolas Curien, Bénédicte Haas, and Igor Kortchemski. The CRT is the scaling limit of random dissections. *Random Structures Algorithms*, 47(2):304–327, 2015.

[26] P. Diaconis and D. Freedman. Finite exchangeable sequences. *Ann. Probab.*, 8(4):745–764, 08 1980.

[27] Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.

[28] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.

[29] Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

[30] P. Erdös and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[31] P. Erdös and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 5:17–61, 1960.

[32] Steven N. Evans. *Probability and real trees*, volume 1920 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008.

[33] Pat Fitzsimmons, Jim Pitman, and Marc Yor. Markovian bridges: construction, Palm interpretation, and splicing. In *Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992)*, volume 33 of *Progr. Probab.*, pages 101–134. Birkhäuser Boston, Boston, MA, 1993.

[34] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

[35] Bailey K. Fosdick, Daniel B. Larremore, Joel Nishimura, and Johan Ugander. Configuration random graph models with fixed degree sequences. *arXiv:1608.00607v2 [stat.ME]*.

[36] Bénédicte Haas and Grégory Miermont. Scaling limits of Markov branching trees with applications to Galton-Watson and random unordered trees. *Ann. Probab.*, 40(6):2589–2666, 2012.

[37] T. E. Harris. First passage and recurrence distributions. *Trans. Amer. Math. Soc.*, 73:471–486, 1952.

[38] Theodore E. Harris. *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin; Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

[39] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[40] Svante Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012.

[41] Svante Janson and Sigurdur Örn Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.*, 43(3):1045–1081, 2015.

[42] Thordur Jonsson and Sigurdur Örn Stefánsson. Condensation in nongeneric trees. *J. Stat. Phys.*, 142(2):277–313, 2011.

[43] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.

[44] Tomasz Ł uczak and Boris Pittel. Components of random forests. *Combin. Probab. Comput.*, 1(1):35–52, 1992.

[45] Jean-François Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311, 2005.

[46] J.-F. LeGall. *Random Trees and Spatial Branching Processes*, volume 9 of *Maphysto Lecture Notes Series*. University of Aarhus, 2000.

[47] T. Lei. Scaling limit of random forests with prescribed degree sequences. arXiv:1704.02064 [math.PR], 2017.

[48] T. Lei. Small trees in supercritical random forests. arXiv:1710.02744 [math.PR], 2017.

[49] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[50] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96(6):933–957, 2006.

[51] R Dean Malmgren, Julio M Ottino, and Luís A Nunes Amaral. The role of mentorship in protégé performance. *Nature*, 465(7298):622–626, 2010.

[52] Jean-François Marckert and Grégory Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures Algorithms*, 38(4):467–501, 2011.

[53] James Martin and Dominic Yeo. Critical random forests. *arXiv:1709.07514 [math.PR]*.

[54] Colin McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.

[55] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canad. J. Math.*, 30(5):997–1015, 1978.

[56] Grégory Miermont. Invariance principles for spatial multitype Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6):1128–1161, 2008.

[57] Peter Mörters and Yuval Peres. *Brownian motion*, volume 30 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.

[58] Konstantinos Panagiotou, Benedikt Stufler, and Kerstin Weller. Scaling limits of random graphs from subcritical classes. *Ann. Probab.*, 44(5):3291–3334, 2016.

[59] Y.L. Pavlov. *Random Forests*. VSP, 2000.

[60] Yu. L. Pavlov. Limit distributions for the maximum size of a tree in a random forest. *Diskret. Mat.*, 7(3):19–32, 1995.

[61] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006.

[62] René L. Schilling and Lothar Partzsch. *Brownian motion*. De Gruyter, Berlin, 2012.

[63] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[64] Daniel B Stouffer, Juan Camacho, Wenxin Jiang, and Luís A Nunes Amaral. Evidence for the existence of a robust pattern of prey selection in food webs. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1621):1931–1940, 2007.

[65] Benedikt Stufler. The continuum random tree is the scaling limit of unlabelled unrooted trees. *arXiv:1412.6333v4 [math.PR]*.

[66] Arun Sundararajan. Local network effects and complex network structure. *The BE Journal of Theoretical Economics*, 7(1), 2007.

[67] Michel Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, (81):73–205, 1995.

[68] Michel Talagrand. New concentration inequalities in product spaces. *Invent. Math.*, 126(3):505–563, 1996.

[69] Duncan J Watts and Peter Sheridan Dodds. Influentials, networks, and public opinion formation. *Journal of consumer research*, 34(4):441–458, 2007.

[70] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.