Heights of Random Trees

Céline Kerriou, Department of Mathematics and Statistics

McGill University, Montréal

May, 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Science, Mathematics and Statistics

Abstract

The study of the asymptotic shape, such as the height, width, and size of Bienaymé trees, more commonly called Galton-Watson trees, is well-developed ([2], [3], [12]). However, most bounds on these quantities make restrictive assumptions on the trees' weight sequences and offspring distributions. We prove novel non-asymptotic tail bounds on the height of randomly sampled nodes in trees with given degree statistics. To achieve this, we construct a sampling procedure that generates a random variable with the same law as the height of a random node, and then adapt a Poissonization trick from Camarri and Pitman [7]. Finally, we describe results from joint work with Addario-Berry, Brandenberger, and Hamdan [4]. We show how bounds on the height and width of conditioned Bienaymé trees can be deduced from our previous arguments, resulting in the proofs of conjectures of Janson [11].

Abrégé

Les propriétés asymptotiques d'arbres de Bienaymé, également connus sous le nom d'arbres de Galton-Watson, telles que la taille, la hauteur et la largeur ont été sujet à de nombreuses études ([2], [3], [12]). Désormais, la plupart des bornes établies sur les propriétés nommées ci-dessus nécessitent des conditions restrictives sur la séquence de poids et la loi de reproduction. Nous prouvons de nouvelles bornes exponentielles non asymptotiques sur la hauteur d'un sommet aléatoire dans un arbre combinatoire aléatoire. Notre preuve repose sur la construction d'un processus d'échantillonage qui génère une variable aléatoire avec la même loi que la hauteur d'un sommet aléatoire. Nous utilisons également une technique de poissonisation introduite par Camarri et Pitman [7]. Enfin, nous décrivons les résultats d'un project conjoint avec Addario-Berry, Brandenberger, and Hamdan [4]. Nous démontrons comment des bornes de la hauteur et la largeur d'arbres Bienaymé conditionnés peuvent être déduites à partir des résultats ci-dessus. Nous résolvons ainsi certaines conjectures de Janson [11].

Acknowledgements

I would like to thank Louigi Addario-Berry for his supervision and guidance over the last few years. In my first year as an undergraduate, I sought his advice as academic advisor on whether to enrol in the Honours track at McGill. Thanks to his advice and encouragement, I decided to follow that path. As a supervisor, Louigi taught me how to conduct rigorous research and gave me many opportunities to further develop my research career. Thank you for being such a great mentor.

Further, I would like to thank my parents and sister for their ongoing support, as well as my friends who made my time at McGill so special - Alex, Amy, Adam, Carl, Elaine, Flo, Joanna, Matthieu, Phil, Ralph, Rivka, Saad, Shelomo, Stéphanie, Tom and Zach - to name a few.

Contribution of Authors

Chapter 2 contains well-known results. The proofs of this chapter are based on [13], the author adapted these proofs to forests with fixed degree statistics. The supervisor suggested the approaches to the proofs of Theorems 1.4 and 1.5 from Chapter 3. The supervisor and the author developed the theorems and the proofs together; all writing was done by the author. The application of Theorem 1.4 to prove Theorem 1.2 is the result of a joint project with the supervisor L. Addario-Berry, A. Brandenberger, J. Hamdan and the author [4], to which all four were equal contributors. The remaining chapters are the work of the author.

Table of Contents

Chapter 1

INTRODUCTION

In this thesis we prove novel non-asymptotic tail bounds on the height of a randomly sampled node in a tree with a given degree statistics. These bounds hold for arbitrary degree statistics. The motivation behind studying trees with given degree statistics is to prove bounds on the heights and widths of Bienaymé trees, more commonly called Galton-Watson trees. Moreover, our results allow us to prove a conjecture and solve an open problem from [11]. This thesis focuses on the proof of tail bounds for the height of a random node of a random tree with a given degree statistics and briefly describes how bounds on Bienaymé trees can be deduced.

We begin with some definitions and notation. We define the *Ulam-Harris tree* $\mathcal{U}$ as the infinite rooted tree with root $\emptyset$ and node set

$$V = V_{\mathcal{U}} := \bigcup_{n \geq 0} \mathbb{N}^n,$$

with $\mathbb{N}^0 := \{\emptyset\}$ and where $\mathbb{N}^n$ is the set of all finite strings $i_1 ... i_n$, with $i_1, ..., i_n \in \mathbb{N}$. Further, any node $u = i_1 ... i_{n+1} \in V$ has parent $i_1 ... i_n$ and ordered children $\{ui : i \geq 1\} = \{i_1 ... i_{n+1} i,\, i \geq 1\}$. Next, define the set $\mathfrak{T}$ as the set of rooted ordered subtrees t of $\mathcal{U}$ with the following properties,

$$\emptyset \in V_t, \tag{1.1}$$

$$i_1...i_n \in V_t \implies i_1...i_{n-1} \in V_t, \tag{1.2}$$

$$i_1...i_{n-1}i_n \in V_t \implies i_1...i_{n-1}i \in V_t \text{ for all } 1 \leq i \leq i_n, \tag{1.3}$$

where V_t denotes the node set of t. A rooted tree t is said to be *ordered* if for each node $u \in t$ the children of u are ordered in a sequence. Then, we can note that $\mathfrak{T}$ defines the set of ordered rooted trees, also known as plane trees.

Given some tree $t \in \mathfrak{T}$, write $r(t)$ to denote the root of tree t and note that $r(t) = \emptyset$. If $u \in t$ then we say u is a *node* of t. The *size* of t is the number of nodes in t and is denoted by $|t|$. Define the *height* of a node $u \in t$ as the number of edges on the path from the root $r(t)$ to u and denote it by $|u|$. By this definition, the height of $\emptyset$ is 0. We define the set of *children* of a node $u \in t$ as $c(u) = c_t(u) := \{ui \in V_t : i \geq 1\}$ and denote the size of this set as $|c_t(u)|$. The *degree of a node* $u \in t$, denoted by $d(u) = d_t(u)$, is the number of children of u in t, that is $d_t(u) = |c_t(u)|$. Given a node $u = i_1...i_n \in t$ where $n \geq 0$ and $i_1, ..., i_n \in \mathbb{N}$, we define $u_k := i_1...i_k$ for $0 \leq k \leq n$ to be the $(n-k)^{th}$ ancestor of u, that is, the k^{th} node on the path from $r(t)$ to u. Note that $u_0 := \emptyset$. The *width* of tree t at level $k \geq 0$, denoted by $L_k(t)$, is the number of nodes at height k, thus $L_k(t) = \#\{u \in t : |u| = k\}$. The *width* of t is $\text{wid}(t) := \max(L_k(t) : k \geq 0)$ and the *height* of t is $\text{ht}(t) := \max(|u| : u \in t)$. Define $\mathfrak{T}^* := \{t \in \mathfrak{T} : |t| < \infty\}$ to be the set of all finite plane trees. In this thesis, by tree we refer to a finite plane tree in $\mathfrak{T}^*$. We define the *lexicographic depth-first search* ordering on the nodes of trees $t \in \mathfrak{T}^*$ as the ordering listing the nodes of t in lexicographic order. This lexicographic order has the property that every non-root node $u \in t$ appears in the order after its parent in t.

We now define Bienaymé trees, also known as Galton-Watson trees[1]. They are a class of random ordered rooted trees. Given a random variable ξ with distribution $\mu = (\mu_i, i \geq 0)$ defined on $\mathbb{N}$, a Bienaymé tree is recursively constructed from the root, giving each node a number of children that is an independent copy of ξ. We call μ the *offspring distribution* of such Bienaymé tree. We say a Bienaymé process is *subcritical*, *critical* or *supercritical* if $\mathbf{E}[\xi] < 1$, $\mathbf{E}[\xi] = 1$ or $\mathbf{E}[\xi] > 1$ respectively. Further, when $\mathbf{E}[\xi] \leq 1$ and $\mathbf{P}\{\xi = 1\} < 1$, then the Bienaymé tree is a.s. finite and when $\mathbf{E}[\xi] > 1$ or $\mathbf{P}\{\xi = 1\} = 1$, then the Bienaymé tree is infinite with positive probability. The Bienaymé process induces a measure $\mathbf{P}^\mu$ on random finite trees. Given a tree $t \in \mathfrak{T}^*$,

$$\mathbf{P}^\mu\{t\} = \prod_{u \in t} \mu_{d_t(u)}.$$

Remark, when $\mathbf{E}[\xi] \leq 1$ and $\mathbf{P}\{\xi = 1\} < 1$, then $\mathbf{P}^\mu$ defines a probability measure on $\mathfrak{T}^*$.

Next, we will introduce *simply generated trees*, which are a generalization of Bienaymé trees. We say $\mathbf{w} = (\mathbf{w}_k,\ k \geq 0)$ is a *weight sequence* if it is a sequence of non-negative real numbers with $\mathbf{w}_0 > 0$. Fix a weight sequence $\mathbf{w}$. We define the *weight* of a finite tree $t \in \mathfrak{T}^*$ as

$$\mathbf{w}(t) := \prod_{u \in t} \mathbf{w}_{d_t(u)}.$$

Meir and Moon introduced trees with such weights in [15] and named them *simply generated trees*. For $n \in \mathbb{N}$, we define a probability measure $\mathbf{P}_n^{\mathbf{w}}$ on $\mathfrak{T}^*$ by setting, for $t \in \mathfrak{T}^*$,

$$\mathbf{P}_n^{\mathbf{w}}(t) := \frac{\mathbf{w}(t)}{Z_n} \mathbf{1}_{[|t|=n]}, \tag{1.4}$$

[1]We propose to use the name "Bienaymé tree" instead of "Galton-Watson trees" as Bienaymé [5] was the first to introduce such trees and correctly state the criticality theorem, as can be read in historical surveys by Heyde and Seneta [9] and by Jagers [10].

where Z_n is the normalizing factor defined by $Z_n := Z_n(\mathbf{w}) := \sum_{T\in\mathfrak{T}_n} \mathbf{w}(t)$. Note that Z_n is finite as $|\mathfrak{T}_n|$ is finite and $Z_n > 0$ since $\mathbf{w}_0 > 0$. We now present the correspondence between simply generated trees and conditioned Bienaymé trees, as described by Janson [11, Section 2.3].

Proposition 1.1. *Fix $n \geq 0$ and a weight sequence $\mathbf{w} = (\mathbf{w}_k, k \geq 0)$ with $\sum_{k\geq 0} \mathbf{w}_k = 1$. Then, the law of a Bienaymé tree with offspring distribution $\mathbf{w}$ conditioned on having n nodes is equivalent to the law of a simply generated tree with n nodes and weight sequence $\mathbf{w}$.*

Proof. Let $\mathcal{T}(\mu)$ denote a Bienaymé tree with offspring distribution μ and let $\mathcal{T}_n(\mathbf{w})$ denote a random simply generated tree with n nodes and weight sequence $\mathbf{w}$. Then, considering the Bienaymé tree $\mathcal{T}(\mathbf{w})$, we have for $t \in \mathfrak{T}^*$,

$$\mathbf{P}\{\mathcal{T}(\mathbf{w}) = t\} = \mathbf{P}^{\mathbf{w}}\{t\} = \prod_{u\in t} \mathbf{w}_{d_t(u)} = \mathbf{w}(t).$$

Thus, $Z_n = \sum_{t\in\mathfrak{T}_n} \mathbf{w}(t) = \sum_{t\in\mathfrak{T}_n} \mathbf{P}^{\mathbf{w}}\{t\} = \mathbf{P}\{|\mathcal{T}(\mathbf{w})| = n\}$. It follows that,

$$\begin{aligned}\mathbf{P}\{\mathcal{T}_n(\mathbf{w}) = t\} = \mathbf{P}_n^{\mathbf{w}}(t) = \frac{\mathbf{w}(t)}{Z_n}\mathbf{1}_{[|t|=n]} &= \frac{\mathbf{P}\{\mathcal{T}(\mathbf{w}) = t\}}{\mathbf{P}\{|\mathcal{T}(\mathbf{w})| = n\}}\mathbf{1}_{[|t|=n]} \\ &= \mathbf{P}\{\mathcal{T}(\mathbf{w}) = t \mid |\mathcal{T}(\mathbf{w})| = n\}.\end{aligned}$$

Thus, the law of a Bienaymé tree with offspring distribution $\mathbf{w}$ conditioned on having n nodes is equivalent to the law of a simply generated tree with n nodes and weight sequence $\mathbf{w}$. □

Let $\Phi(z) = \Phi_{\mathbf{w}}(z) = \sum_{k\geq 0} \mathbf{w}_k z^k$ be the generating function of $\mathbf{w}$ and let $\rho = \rho_{\mathbf{w}}$ be the radius of convergence of Φ. For $t > 0$ such that $\Phi(t) < \infty$, let

$$\Psi(t) = \Psi_{\mathbf{w}}(t) = \frac{t\Phi'(t)}{\Phi(t)} = \frac{\sum_{k\geq 0} k\mathbf{w}_k t^k}{\sum_{k\geq 0} \mathbf{w}_k t^k}.$$

If $\Phi(\rho) = \infty$, then define $\Psi(\rho) = \Psi_{\mathbf{w}}(\rho) = \lim_{t\uparrow\rho} \Psi(t)$. By [11, Lemma 3.1(i)], this limit exists. Further, let $\nu = \nu(\mathbf{w}) = \Psi_{\mathbf{w}}(\rho)$ and assume $\nu \leq 1$. Define $\sigma^2 = \rho\Psi'(\rho)$. Janson states the following conjectures and problems [11, Conjectures 21.5, 21.6 and Problems 21.7, 21.8].

Conjecture 1. *Let $\mathbf{w} = (\mathbf{w}_k, k \geq 0)$ be a weight sequence with $\mathbf{w}_0 > 0$ and $\mathbf{w}_k > 0$ for some $k \geq 2$. Let $\mathcal{T}_n$ be a simply generated tree of size n with weight sequence $\mathbf{w}$, whenever $n \geq 0$ satisfies $Z_n(\mathbf{w}) > 0$.*

(1) If $\nu = 1$ and $\sigma^2 = \infty$ then ht$(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} 0$.

(2) If $\nu = 1$ and $\sigma^2 = \infty$ then wid$(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} \infty$.

(3) If $\nu < 1$ then ht$(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} 0$.

(4) If $\nu < 1$ then wid$(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} \infty$.

In the context of joint work with Addario-Berry, Brandenberger, Hamdan [4] we were able to prove results similar to Conjectures (1) and (2) for conditioned Bienaymé trees.

Theorem 1.2. *[4, Theorem 1.2.] Fix a probability distribution μ supported by $\mathbb{N}$ with $|\mu|_1 \leq 1$ and $|\mu|_2 = \infty$. For $n \in \mathbb{N}$, let T_n be a Bienaymé tree with offspring distribution μ conditioned to have size n, and let V_n be a uniformly random node in T_n. Then*

$$\textit{wid}(T_n)/\sqrt{n} \to \infty,\ |V_n|/\sqrt{n} \to 0 \textit{ and } \textit{ht}(T_n)/(\sqrt{n}\log^3 n) \to 0,$$

where the convergence results hold both in probability and in expectation, as $n \to \infty$.

In fact in [4] we also prove (4) and a slight weakening of (3) but these results are not presented in detail in the current work. The study of the asymptotic shape, such as the height, width and size, of simply generated trees and Bienaymé trees is well-developed ([2], [3], [12]). However, most bounds established make restrictive assumptions on weight sequences and offspring distributions. Work by Addario-Berry, Devroye and Janson [3] establishes uniform sub-Gaussian tail bounds for the width and height of critical conditioned Bienaymé trees with finite variance ($|\mu|_1 = 1$, $0 < |\mu|_2 < \infty$),

$$\mathbf{P}\{\text{wid}(T_n) \geq x\} \leq C_1 e^{-c_1 x^2/n}, \text{ and } \mathbf{P}\{\text{ht}(T_n) \geq h\} \leq C_2 e^{-c_2 h^2/n},$$

for $x, h \geq 0$ and some constants C_1, c_1, C_2, c_2. Kortchemski [12] extends these results to hold for critical Bienaymé trees when the offspring distribution μ is in the domain of attraction of a stable law. (We say that a random variable X with distribution μ belongs to the domain of attraction of a stable law of index $\alpha \in (0, 2[$ if $\mathbf{P}\{X \geq k\} = L(k)/k^\alpha$ where $L : \mathbb{R}_+ \to \mathbb{R}_+$ is a slowly varying function, that is, $\lim_{x\to\infty} L(tx)/L(x) = 1$ for $t > 0$.) Theorem 1.2 establishes results on the shape of Bienaymé trees without restrictive assumptions about the offspring distribution μ. This theorem is a consequence of bounds we established on the height of random nodes of *trees with fixed degree statistics*, which we introduce below. The *degree statistics* (singular) of a tree t is the sequence $\mathbf{n}_t = (\mathbf{n}_t(c),\ c \geq 0)$, where $\mathbf{n}_t(c) := |\{u \in t : d_t(u) = c\}|$ is the number of nodes of t with c children. Note that since a tree with n nodes has $n - 1$ edges and every edge is associated to a child in a tree, we have that

$$\sum_{c\geq 0} \mathbf{n}_t(c) = 1 + \sum_{c\geq 0} c\mathbf{n}_t(c).$$

For any sequence $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ define

$$k(\mathbf{n}) := \sum_{c\geq 0} (1 - c)\mathbf{n}(c).$$

A sequence $\mathbf{n} = (\mathbf{n}(c),\ c \geq 0)$ of non-negative integers is the degree statistics of some tree in $\mathfrak{T}^*$ if and only if $k(\mathbf{n}) = 1$. For such sequences, we write $\mathfrak{T}_{\mathbf{n}}$ for the set of finite plane trees with degree statistics $\mathbf{n}$. For $n \geq r \geq 0$, define the degree statistics $bin(n,r) := (n, 0, n-r, 0, ...)$. Then $\mathfrak{T}_{bin(n,1)}$ denotes the set of binary trees with n leaves.

Given sequences $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ and $\mathbf{m} = (\mathbf{m}(c),\, c \geq 0)$ of natural numbers, define $\mathbf{n} + \mathbf{m} := (\mathbf{n}(c) + \mathbf{m}(c),\, c \geq 0)$. Further, write $\|\mathbf{n}\| := \sum_{c\geq 0} \mathbf{n}(c)$ and for $p > 0$ we write $|\mathbf{n}|_p := (\sum_{c\geq 0} c^p \mathbf{n}(c))^{1/p}$. Note that for a tree t, we have $\|\mathbf{n}_t\| = |t|$ and $|\mathbf{n}_t|_1 = |t| - 1$. Thus, $\|\mathbf{n}_t\| = |\mathbf{n}_t|_1 + 1$. We write $n = \|\mathbf{n}\|$ and $n_t = \|\mathbf{n}_t\|$ for readability. For a natural number $n \geq 1$, we write $[n]$ for $\{1, ..., n\}$. For a finite set $\mathcal{S}$, we write $X \in_u \mathcal{S}$ to mean that X is a uniformly random element of set $\mathcal{S}$.

Fix a degree statistics $\mathbf{n}$ with $k(\mathbf{n}) = 1$. We define a probability measure $\mathbf{P_n}$ on $\mathfrak{T}^*$. For $t \in \mathfrak{T}^*$,

$$\mathbf{P_n}(t) := \frac{1}{Z_\mathbf{n}} \mathbf{1}_{[\mathbf{n}_t = \mathbf{n}]}, \tag{1.5}$$

where $Z_\mathbf{n} := |\mathfrak{T}_\mathbf{n}|$. Fix a weight sequence $\mathbf{w} = (\mathbf{w}_k, k \geq 0)$. We define the *weight* of a degree statistics $\mathbf{n}$ as

$$\widetilde{\mathbf{w}}(\mathbf{n}) := \prod_{k\geq 0} \mathbf{w}_k^{\mathbf{n}(k)}.$$

Remark that for any tree $t \in \mathfrak{T}^*$, since $\mathbf{n}_t(c) = |\{u \in t : d_t(u) = c\}|$ for $c \geq 0$,

$$\mathbf{w}(t) = \prod_{u\in t} \mathbf{w}_{d_t(u)} = \prod_{c\geq 0} \prod_{\substack{u\in t \\ d_t(u)=c}} \mathbf{w}_c = \prod_{c\geq 0} \mathbf{w}_c^{\mathbf{n}_t(c)} = \widetilde{\mathbf{w}}(\mathbf{n}_t).$$

The following proposition relates trees with a fixed degree statistics to simply generated trees.

Proposition 1.3. *Fix a degree statistics* $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ *and a weight sequence* $\mathbf{w}$. *Then, the law of a simply generated tree with weight sequence* $\mathbf{w}$ *conditioned on having degree statistics* $\mathbf{n}$ *is equivalent to the law of a tree with fixed degree statistics* $\mathbf{n}$.

Proof. Recall that $n = \sum_{c\geq 0} \mathbf{n}(c)$. Let $\mathcal{T}_\mathbf{n}$ be a random tree chosen according to $\mathbf{P}_\mathbf{n}$ and $\mathcal{T}_n(\mathbf{w})$ be a random tree chosen according to $\mathbf{P}^\mathbf{w}_n$. Then, using (1.4), since $\mathbf{w}(t) = \widetilde{\mathbf{w}}(\mathbf{n}_t)$ and $Z_\mathbf{n} = |\mathfrak{T}_\mathbf{n}|$,

$$\begin{aligned}
\mathbf{P}\left\{\mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}\right\} &= \sum_{t:\mathbf{n}_t=\mathbf{n}} \mathbf{P}\left\{\mathcal{T}_n(\mathbf{w}) = t\right\} = \sum_{t:\mathbf{n}_t=\mathbf{n}} \mathbf{P}^\mathbf{w}_n(t) \\
&= \sum_{t:\mathbf{n}_t=\mathbf{n}} \frac{\mathbf{w}(t)}{Z_n}\mathbf{1}_{[|T|=n]} = \sum_{t:\mathbf{n}_t=\mathbf{n}} \frac{\widetilde{\mathbf{w}}(\mathbf{n}_t)}{Z_n} \\
&= \frac{\widetilde{\mathbf{w}}(\mathbf{n})}{Z_n}|\mathfrak{T}_\mathbf{n}| = \frac{\widetilde{\mathbf{w}}(\mathbf{n})}{Z_n}Z_\mathbf{n}.
\end{aligned}$$

Further, for $t \in \mathfrak{T}^*$,

$$\mathbf{P}\left\{\mathcal{T}_n(\mathbf{w}) = t, \mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}\right\} = \mathbf{P}\left\{\mathcal{T}_n(\mathbf{w}) = t\right\}\mathbf{1}_{[\mathbf{n}_t=\mathbf{n}]} = \frac{\mathbf{w}(t)}{Z_n}\mathbf{1}_{[|t|=n]}\mathbf{1}_{[\mathbf{n}_t=\mathbf{n}]} = \frac{\widetilde{\mathbf{w}}(\mathbf{n})}{Z_n}\mathbf{1}_{[\mathbf{n}_t=\mathbf{n}]}.$$

Combining these two equalities, by (1.5) we conclude that for $t \in \mathfrak{T}^*$,

$$\begin{aligned}
\mathbf{P}\left\{\mathcal{T}_n(\mathbf{w}) = t \mid \mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}\right\} &= \frac{\mathbf{P}\left\{\mathcal{T}_n(\mathbf{w}) = t, \mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}\right\}}{\mathbf{P}\left\{\mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}\right\}} \\
&= \frac{1}{Z_\mathbf{n}}\mathbf{1}_{[\mathbf{n}_t=\mathbf{n}]} = \mathbf{P}_\mathbf{n}(t).
\end{aligned}$$

In words, the random simply generated trees $\mathcal{T}_n(\mathbf{w})$ conditioned on $\mathbf{n}_{\mathcal{T}_n(\mathbf{w})} = \mathbf{n}$ is equivalent to the random tree with fixed degree statistics $\mathcal{T}_\mathbf{n}$. □

The main results of this thesis are contained in the following two theorems.

Theorem 1.4. *Fix a degree statistics $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ with $k(\mathbf{n}) = 1$ and let $(T, V) \in_u \mathfrak{T}^{(1)}_\mathbf{n}$. Then for all $\alpha > 17^{3/2}$,*

$$\mathbf{P}\left\{|V| > \alpha\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right),$$

and if $\mathbf{n}(1) = 0$, *then for all* $\ell \geq 1$,

$$\mathbf{P}\left\{|V| \geq \ell\right\} \leq \exp\left(-\frac{\ell^2}{2|\mathbf{n}|_1}\right).$$

Theorem 1.5. *Let* $m \geq 1$. *Let* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *be a degree statistics with* $\mathbf{n}(0) = m$, $\mathbf{n}(1) = 0$ *and* $k(\mathbf{n}) = 1$. *Further, recall that* $bin(m, 1) := (m, 0, m-1, 0, ...)$. *Let* $(T, V) \in_u \mathfrak{T}^{(1)}_{\mathbf{n}}$ *and* $(T', W) \in_u \mathfrak{T}^{(1)}_{bin(m,1)}$ *be random marked trees. Then,*

$$|V| \preceq_{st} |W|.$$

Theorem 1.4 states a non-asymptotic tail bound on the height of a randomly sampled node in a random tree. Theorem 1.5 states that the height of a random node in a random binary tree with m leaves stochastically dominates the height of a random node in a random tree with m leaves and no degree one nodes. The latter theorem is not used in the proof of Theorem 1.2 but is an interesting result in its own right.

The thesis focuses on the proof of Theorem 1.4 and Theorem 1.5. Chapter 2 describes the correspondence between trees and lattice paths and outlines how this correspondence is used in [3] to prove bounds on the height and width of conditioned Bienaymé trees. Further, Chapter 2 establishes combinatorial identities on the number of trees and forests with a fixed degree statistics using this correspondence. These well-studied identities ([14], [13], [16]) are necessary to prove the theorems from the following chapter. Chapter 3 describes the proofs of Theorem 1.4, Theorem 1.5 and Theorem 1.2. In the conclusion, we present a possible generalisation of Theorem 1.5 and possible future problems around Conjecture 1.

Chapter 2

COMBINATORIAL IDENTITIES

This chapter focuses on the relation between random trees and random lattice walks. The correspondence between random walk paths and trees is well-understood. Harris noted in 1952 [8] that "walks and trees are abstractly identical objects" by encoding trees into paths. This connection has proven to be an important tool in studying properties of trees. The first section of this chapter briefly examines how the correspondence between trees and random walks is used in [3] to prove bounds on the height and width of random trees. The second section focuses on proving known combinatorial identities on the number of trees and forests with a given degree statistics.

2.1 Applications of the correspondence between random walk paths and trees.

The correspondence between random walk paths and trees has been a key tool for studying the shape of random trees ([1], [3], [12]). In [1], Addario-Berry establishes tail bounds for the height and width of a random tree with a given degree statistics $\mathbf{n}$ by encoding trees into lattice paths and using martingale concentration results to bound the maximum height of the respective lattice paths. These bounds are tight when $\sum_{c\geq 0}\mathbf{n}(c)^2 = O(n)$. Similarly, in [12], Kortchemski establishes bounds on the width, height and maximal degree of certain conditioned Bienaymé trees by bounding the height of the encoded lattice paths.

We describe how the correspondence between random walk paths and trees is used in [3] to prove bounds on the height and width of Bienaymé trees. The methods used in [3] are similar to those in [1] and [12]. We begin by defining three orderings on the nodes of a tree t.

- The *breadth first search* (BFS) order list the nodes of t in increasing order of depth, and for nodes of the same depth, in lexicographic order.
- The *lexicographic depth first search* (lex-DFS) order of a tree list the nodes of t in lexicographic order.
- Let t' be the mirror-image of t, then the *reversed depth first search* (rev-DFS) order of t corresponds to the lex-DFS order of t'.

Let $(u_1, ..., u_{|t|})$ be ordered nodes of t. We define a queue process on the nodes of t as follows. Let $Q_t(0) = 0$ and $Q_t(i) = Q_t(i-1) - 1 + d_t(u_i)$ for $i \in [|t|]$. We let $(Q_t^b(i), i \in [|t|])$, $(Q_t^l(i), i \in [|t|])$ and $(Q_t^r(i), i \in [|t|])$ be the respective queue process for the BFS, lex-DFS and rev-DFS order. The lex-DFS queue $(Q_t^l(i), i \in [|t|])$ of a tree t corresponds to the *Łukasiewicz path* $\mathcal{W}(t)$ of t we define in the following section. We begin with stating the bounds on the height and width of conditioned Bienaymé trees established in [3].

Theorem 2.1. *[3, Theorem 1.1., Theorem 1.2.] Fix an offspring distribution μ such that $\mathbf{E}[\mu] = 1$ and $0 < Var(\mu) < \infty$. Let T_n be a Bienaymé tree with offspring distribution μ conditioned to have size n. Then,*

$$\mathbf{P}\{wid(T_n) \geq x\} \leq C_1 \exp\left(-c_1 x^2/n\right),$$

$$\mathbf{P}\{ht(T_n) \geq h\} \leq C_2 \exp\left(-c_2 h^2/n\right),$$

for all $x, h \geq 0$ and $n \geq 1$ and constants C_1, c_1, C_2, c_2.

First, we outline the proof for the bound on the width of conditioned Bienaymé trees. Recall that the width of a tree t at level $k \geq 0$ is defined as $L_k(t) := |\{u \in t : |u| = k\}|$ and the width of a tree t is $\mathrm{wid}(t) := \max(L_k(t) : k \geq 0)$. Remark the following property of the BFS queue process. When exploring a tree t via BFS, we note that after exploring all nodes of a level $k \geq 0$, the queue process then consists precisely of all the nodes at level $k+1$. Therefore, for each $k \geq 0$, let i_k be the step at which we finished exploring all nodes of level k, we have that $L_k(t) = Q_t^b(i_k)$. It follows that for any tree t,

$$\mathrm{wid}(t) = \max_{k \geq 0} L_k(t) \leq \max_{i \geq 0} Q_t^b(i),$$

and further, for $x \geq 0$

$$\mathbf{P}\left\{\mathrm{wid}(t) \geq x\right\} \leq \mathbf{P}\left\{\max_{i \geq 0} Q_t^b(i) \geq x\right\}. \tag{2.1}$$

We now study the BFS queue $(Q_t^b(i), i \in [|t|])$ for Bienaymé trees. Let T be a Bienaymé tree with offspring distribution μ and let $(X_i, i \geq 0)$ be i.i.d. μ distributed random variables. Note that for all $u \in T$, X_i and $d_T(u)$ follow the same law. Therefore, for $i \geq 1$, $Q_T^b(i) = \sum_{j=1}^{i}(d_T(u_j) - 1)$ follows the same law as $W(i) = \sum_{j=1}^{i}(X_j - 1)$. Further, the tree T has size n if and only if $Q_T^b(j) > -1$ for all $0 \leq j < n$ and $Q_T^b(n) = -1$; equivalently $W(j) > -1$ for $0 \leq j < n$ and $W(n) = -1$. We can therefore rewrite the upperbound of (2.1) as follows.

$$\begin{aligned}\mathbf{P}\left\{\mathrm{wid}(T_n) \geq x\right\} &\leq \mathbf{P}\left\{\max_{i \geq 0} Q_{T_n}^b(i) \geq x\right\} \\ &= \mathbf{P}\left\{\max_{i \geq 0} W(i) \geq x \;\middle|\; W(i) > -1, 0 \leq j < n \text{ and } W(n) = -1\right\}.\end{aligned}$$

In the proof of [3, Theorem 1.1], the authors bound the last term using rotation arguments, tail bounds on $W(i)$ and various computations, which we omit the details of.

Next, we describe how bounds on the height of conditioned Bienaymé trees can be established. For simplicity, assume that $\mathbf{P}\{\mu = 1\} = 0$. For $u \in t$, let $l(u)$ and $r(u)$ be the respective index of u when the nodes of t are listed in lex-DFS and rev-DFS. Note that at time $l(u)$, since $\mathbf{P}\{\mu = 1\} = 0$, every ancestor of u has at least two children, at least one of which is not an ancestor of u. Therefore, every ancestor of u contributes at least one to either $Q_T^l(l(u))$ or $Q_T^r(r(u))$. Since u has $|u|$ ancestors, it follows that either $Q_T^l(l(u)) \geq |u|/2$ or $Q_T^r(r(u)) \geq |u|/2$. Then, for $x \geq 1$,

$$\begin{aligned}
\mathbf{P}\{\mathrm{ht}(T) \geq x\} &= \mathbf{P}\left\{\max_{u\in T} |u| \geq x\right\} \\
&\leq \mathbf{P}\left\{\max_{0\leq i\leq n} Q_T^l(i) \geq \lceil x \rceil/2\right\} + \mathbf{P}\left\{\max_{0\leq i\leq n} Q_T^r(i) \geq \lceil x \rceil/2\right\} \\
&= 2\mathbf{P}\left\{\max_{0\leq i\leq n} Q_T^b(i) \geq \lceil x \rceil/2\right\}.
\end{aligned}$$

The last inequality follows from the fact that $(Q_T^l(i), i \in [|T|])$, $(Q_T^r(i), i \in [|T|])$ and $(Q_T^b(i), i \in [|T|])$ all have the same distribution. By similar arguments for the proof of width bounds in [3, Theorem 1.1], bounds on the height of T_n can be deduced.

2.2 Combinatorial identities.

Combinatorial identities around plane trees and forests have been widely studied ([13], [14], [17], [16]). The number of plane trees and forests with a given degree statistics can be found in [17, Exercise 6.2.1]. This chapter derives combinatorial identities on the number of plane forests with given degree statistics by counting the number of encoded lattice paths. In [13], Kortchemski formalizes the correspondence between trees with a fixed number of nodes and lattice walks. We closely follow [13], adapting the propositions and proofs to hold for trees and forests with a fixed degree statistics and to prove the desired identities. In particular, Proposition 2.3 corresponds to [13, Proposition 3.2] and the presentation of the Cyclic Lemma 2.5 follows that given in [13, Section 3.3].

We begin by introducing a slightly modified version of the Ulam Harris Tree $\mathcal{U}$ in order to define forests. Let $\mathcal{U}(j)$ be the infinite rooted subtree of $\mathcal{U}$ consisting of all descendants of node j. For $j \geq 1$, define $\mathfrak{T}(j) := \{jt : t \in \mathfrak{T}\}$ as the set of plane subtrees of $\mathcal{U}(j)$. Further let $\mathfrak{T}^*(j) := \{jt : t \in \mathfrak{T}^*\}$ and $\mathfrak{T}_{\mathbf{n}}(j) := \{jt : t \in \mathfrak{T}_{\mathbf{n}}\}$, for some degree statistics $\mathbf{n}$. We define the *Ulam Harris Forest* $\mathcal{F} := (\mathcal{U}(j), j \geq 1)$ as an infinite forest composed of infinite plane trees. Further, define $\mathfrak{F}^k := \mathfrak{T}^*(1) \times \mathfrak{T}^*(2) \times ... \times \mathfrak{T}^*(k)$ and $\mathfrak{F}^* := \cup_{k \geq 1} \mathfrak{F}^k$ as respectively the set of forests with k finite plane trees and the set of finite forests composed of finite plane trees. We have that $\mathfrak{F}^k$ and $\mathfrak{F}^*$ are both subforests of $\mathcal{F}$. In the sequel, the term forest will refer to a finite forest in $\mathfrak{F}^*$. Given a forest $f \in \mathfrak{F}^k$, we can write $f = (t_1, ..., t_k)$ for some $(t_1, ..., t_k) \in \mathfrak{T}^*(1) \times \mathfrak{T}^*(2) \times ... \times \mathfrak{T}^*(k)$. If $u \in t_j$ for some $j \in [k]$, then we write $u \in f$ and say that u is a node of f. In this case $u = ji_1...i_n$ for some $n \geq 0$ and $i_1, .., i_n \in \mathbb{N}$. The size of a forest f is the number of nodes in f and is denoted by $|f|$. Define the *degree* of a node $u \in f$ as $d(u) := d_f(u) = d_{t_j}(u)$, where $u \in t_j$. By viewing forests as a subset of $\mathcal{F}$, every node u of a forest f is a unique string $u = i_1...i_n$, for some $n \geq 1$ and $i_1, ..., i_n \in \mathbb{N}$.

The lexicographic depth-first search ordering on the nodes of forests $f \in \mathfrak{T}^*$ is the ordering listing the nodes of f in lexicographic order. This order has the property that every non-root node $u \in f = (t_1, ..., t_k)$ appears in the order after its parent in f. Further, for all $1 < j \leq k$, when a node in t_j appears in the order, all nodes in t_i, for $1 \leq i < j$, have already appeared.

We now introduce the *degree statistics* of a forest $f \in \mathfrak{F}^*$. The *degree statistics* of f

is the sequence $\mathbf{n}_f = (\mathbf{n}_f(c),\, c \geq 0)$, where $\mathbf{n}_f(c) := |\{u \in f \,:\, d_f(u) = c\}|$. Note that the definition of the degree statistics in $\mathfrak{F}^*$ is an extension of the respective definition in $\mathfrak{T}^*$. Remark that for integers $n \geq k \geq 1$, any forest f with n nodes and k trees has $n - k$ edges. Further, every edge can be associated to a unique child in f. Thus, recalling that $k(\mathbf{n}_f) := \sum_{c \geq 0}(1 - c)\mathbf{n}_f(c)$, we then have

$$k(\mathbf{n}_f) = \sum_{c \geq 0} \mathbf{n}_f(c) - \sum_{c \geq 0} c\mathbf{n}_f(c) = k.$$

For $k \geq 1$, sequence $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ of non-negative integers is the degree statistics of some forest in $\mathfrak{F}^k$ if and only if $k(\mathbf{n}) = k$. Further, $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ is the degree statistics of some forest in $\mathfrak{F}^*$ if and only if $k(\mathbf{n}) \geq 1$. For such sequences, we write $\mathfrak{F}_{\mathbf{n}}$ for the set of finite forests with degree statistics $\mathbf{n}$.

We define a (n, k)-*lattice path* as a path with starting point $(0, 0)$ and ending point $(n, -k)$, that consists of steps lying in $\{(1, i), i \in \{-1, 0, 1, ..\}\}$. The *length* of a path corresponds to the number of its steps n and we say that the *step size* of $(1, i)$ is i. Given a (n, k)-lattice path, we define its *step sequence* as the sequence $(i_j,\, j \in [n])$ where i_j is the step size of the j^{th} step $(1, i_j)$. Let $f \in \mathfrak{F}^*$, let $u_1, u_2, ..., u_{|f|}$ be the nodes of f, listed in lexicographic order. We can characterize forests by encoding them into lattice paths as follows. We define the *Łukasiewicz path* of a forest f as the vector $\mathcal{W}(f) = (\mathcal{W}_i(f), 0 \leq i \leq |f|)$, where

$$\mathcal{W}_0(f) = 0 \text{ and } \mathcal{W}_{i+1}(f) = \mathcal{W}_i(f) + d_f(u_{i+1}) - 1 \text{ for } 0 \leq i < |f|.$$

Remark that the *Łukasiewicz lattice path* $((i, \mathcal{W}_i(f)),\, 0 \leq i \leq |f|)$ is a $(|f|, k(\mathbf{n}_f))$-lattice path with step sequence $(d_f(u_i) - 1, 1 \leq i \leq |f|)$. We say that $(d_f(u_i) - 1, 1 \leq i \leq |f|)$ is the *step sequence* of f. Further, the step sequence $(d_f(u_i) - 1, 1 \leq i \leq |f|)$ of a forest f satisfies $\#\{1 \leq i \leq |f|,\, d_f(u_i) - 1 = j - 1\} = \mathbf{n}_f(c)$, for all $c \geq 0$.

Now let $\mathbf{x} = (x_j,\, j \geq 0)$ be a finite sequence of integers. We define the vector

$h(\mathbf{x}) = (h(\mathbf{x}, c),\ c \geq 0)$ such that for all $c \geq 0$,

$$h(\mathbf{x}, c) := \#\{j \geq 1 : x_j = c - 1\}. \tag{2.2}$$

Given a sequence $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ of non-negative integers, we define the sets

$$S_{\mathbf{n}} := \{\mathbf{x} = (x_1, ..., x_n) \in \{-1, 0, 1, ...\}^n : \forall c \geq 0,\ h(\mathbf{x}, c) = \mathbf{n}(c)\},$$
$$\bar{S}_{\mathbf{n}} := \{(x_1, ..., x_n) \in S_{\mathbf{n}} : \forall\, 1 \leq j < n,\ x_1 + ... + x_j > -k(\mathbf{n})\}.$$

Note that for $\mathbf{x} \in S_{\mathbf{n}}$, we have $h(\mathbf{x}) = \mathbf{n}$. We can think of $S_{\mathbf{n}}$ as the set of step sequences $\mathbf{x}$ of the corresponding $(n, -k(\mathbf{n}))$-lattice paths with $\mathbf{n}(c)$ steps of size $c - 1$ for each $c \geq 0$. The set $\bar{S}_{\mathbf{n}}$ is the set of step sequences in $S_{\mathbf{n}}$ where the corresponding lattice paths stay above $-k(\mathbf{n})$ until the last step. Remark that the step sequence $(d_f(u_i) - 1, 1 \leq i \leq |f|)$ of a forest f belongs to the set $\bar{S}_{\mathbf{n}_f}$: for $1 \leq j < |f|$, $\sum_{i=1}^{j}(d_f(u_i) - 1) = \sum_{i=1}^{j} d_f(u_i) - j$ and the sum $\sum_{i=1}^{j} d_f(u_i)$ counts the number of children of the nodes $u_1, ..., u_j$. At most $\min(j + 1, k(\mathbf{n}))$ nodes amongst $u_2, ..., u_{j+1}$ are roots of trees in f. Since the nodes are in lexicographic order, it follows that at least $j - k(\mathbf{n})$ nodes amongst $u_2, ..., u_{j+1}$ are children of the nodes $u_1, ..., u_j$. Thus, $\sum_{i=1}^{j} d_f(u_i) \geq j - k(\mathbf{n})$ and so $f \in \bar{S}_{\mathbf{n}_f}$.
The following proposition relates the sets $\mathfrak{F}_{\mathbf{n}}$ and $\bar{S}_{\mathbf{n}}$ by constructing a bijection between forests and sequences in $\bar{S}_{\mathbf{n}}$.

Proposition 2.2. *Let $\mathbf{n}$ be a degree statistics. Define the function $f_{\mathbf{n}}$ as*

$$f_{\mathbf{n}} : \mathfrak{F}_{\mathbf{n}} \to \bar{S}_{\mathbf{n}}$$
$$f \mapsto (d_f(u_i) - 1,\ 1 \leq i \leq n),$$

where $u_1, ..., u_n$ are the nodes in f in lexicographic order. Then $f_{\mathbf{n}}$ is a bijection.

We prove that $f_{\mathbf{n}}$ is a bijection for any degree statistics $\mathbf{n}$ by induction on $k(\mathbf{n})$. We thus begin by stating and proving Proposition 2.2 for degree statistics $\mathbf{n}$ that satisfy $k(\mathbf{n}) = 1$.

Proposition 2.3. *Let $\mathbf{n}$ be a degree statistics satisfying $k(\mathbf{n}) = 1$. Define the function $f_{\mathbf{n}}$ as*

$$\begin{aligned} f_{\mathbf{n}} : \mathfrak{T}_{\mathbf{n}} &\to \bar{S}_{\mathbf{n}} \\ t &\mapsto (d_t(u_i) - 1,\ 1 \leq i \leq n), \end{aligned}$$

where $u_1, ..., u_n$ are the nodes of t in lexicographic order. Then $f_{\mathbf{n}}$ is a bijection.

Since there is an obvious bijection between $\mathfrak{T}_{\mathbf{n}}$ and $\mathfrak{T}_{\mathbf{n}}(j)$ for all $j \geq 1$, Proposition 2.3 implies that the function $f^j_{\mathbf{n}} : \mathfrak{T}_{\mathbf{n}}(j) \to \bar{S}_{\mathbf{n}}$ defined as $f^j_{\mathbf{n}}(jt) := f_{\mathbf{n}}(t)$ for $t \in \mathfrak{T}_{\mathbf{n}}$ is also a bijection.

Proof of Proposition 2.3. We begin by proving the following intermediate result. For any two vectors $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_m)$ we denote the concatenation of $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{xy} := (x_1, ..., x_n, y_1, ..., y_m)$.

Lemma 2.4. *Let $\mathbf{n}$ be some degree statistics and let $\mathbf{x} = (x_1, ..., x_n) \in \bar{S}_{\mathbf{n}}$, then there exists a unique set of $k := k(\mathbf{n})$ vectors $\mathbf{y}^1, ..., \mathbf{y}^k$ such that*

$$\mathbf{x} = \mathbf{y}^1 ... \mathbf{y}^k,$$

where $|\mathbf{y}^i|_1 = -1$ and $\mathbf{y}^i \in \bar{S}_{h(\mathbf{y}^i)}$. Further, for each $i \in [k]$,

$$\mathbf{y}^i := (x_{l_{i-1}+1}, ..., x_{l_i}),$$

where $l_0 := 0$ and $l_i := \min(j \in [n] : x_1 + ... + x_j = -i)$. Remark that for all $i \in [k]$, $h(\mathbf{y}^i)$ is a degree statistics with $k(h(\mathbf{y}^i)) = 1$.

Proof. We start by proving that the vectors $(\mathbf{y}^i = (x_{l_{i-1}+1}, ..., x_{l_i}), i \in [k])$ satisfy the conditions above. A sequence $\mathbf{x} \in \bar{S}_{\mathbf{n}}$ stays above $-k$ until step n, thus $l_k = n$ and $l_i \leq l_{i+1}$ for all $i \in [n-1]$. Therefore, we have that $\mathbf{x} = \mathbf{y}^1 ... \mathbf{y}^k$ and for each $i \in [k]$,

$$|\mathbf{y}^i|_1 = \sum_{j=l_{i-1}+1}^{l_i} x_j = \sum_{j=1}^{l_i} x_j - \sum_{j=1}^{l_{i-1}} x_j = -i + (i-1) = -1.$$

By definition, we know that $h(\mathbf{y}^i)$ for $i \in [k]$ is a sequence of non-negative integers and since $|\mathbf{y}^i|_1 = \sum_{c \geq 0} (c-1) h(\mathbf{y}^i, c) = -1$, we have that $h(\mathbf{y}^i)$ is a degree statistics. By construction of the l_i's, for all $1 < i \leq n$ and $a < l_i$,

$$\sum_{j=l_{i-1}+1}^{a} x_j = \sum_{j=1}^{a} x_j - \sum_{j=1}^{l_{i-1}} x_j > -i - \sum_{j=1}^{l_{i-1}} x_j = -1.$$

Therefore, $\mathbf{y}^i \in \bar{S}_{h(\mathbf{y}^i)}$ for all $i \in [k]$. We now show that our choice of vectors $(\mathbf{y}^i, i \in [k])$ is unique by contradiction. Suppose there exist vectors $(\mathbf{z}^i, i \in [k])$ satisfying the conditions of the lemma. WLOG suppose that $\mathbf{y}^1 = (x_1, ..., x_{l_1}) \neq \mathbf{z}^1 = (x_1, ..., x_j)$ for some $j \neq l_1 \geq 1$. If $j < l_1$, then l_1 is not the smallest time the sum attains -1. If $l_1 < j$, then $x_1 + ... + x_{l_1} = -1$ and $\mathbf{z}^1 \notin \bar{S}_{h(\mathbf{z}^1)}$. $\square$

We now proceed to prove Proposition 2.3. We begin by showing that $f_{\mathbf{n}}$ is injective. Let $t_1, t_2 \in \mathfrak{T}_{\mathbf{n}}$ such that $t_1 \neq t_2$ and let $u_1, ..., u_{|t_1|}$ and $v_1, ..., v_{|t_2|}$ be the respective nodes of t_1 and t_2 listed in lexicographic order. Suppose $|t_1| \neq |t_2|$, then $|f_{\mathbf{n}}(t_1)| \neq |f_{\mathbf{n}}(t_1)|$. Now suppose $|t_1| = |t_2|$, since $t_1 \neq t_2$, there exists at least one $1 \leq i \leq |t_1|$ such that $d_{t_1}(u_i) \neq d_{t_2}(v_i)$ and thus $f_{\mathbf{n}}(t_1) \neq f_{\mathbf{n}}(t_2)$.

We prove by induction on the size n of the degree statistics $\mathbf{n}$ that $f_{\mathbf{n}}$ is surjective. For the base case $n = 1$, we have that $\sum_{c \geq 0} c\mathbf{n}(c) = n - 1 = 0$, hence $\mathbf{n}(0) = 1$ and $\mathbf{n}(c) = 0$ for all $c \geq 1$. Then, $\mathfrak{T}_{\mathbf{n}} = \{\emptyset\}$ and $f_n(\emptyset) = \{(-1)\} = \bar{S}_{\mathbf{n}}$.

For the induction step, let $\mathbf{x} = (k, x_2, ..., x_n) \in \bar{S}_{\mathbf{n}}$ for any $k \geq 1$. Then, $\mathbf{z} := (x_2, ..., x_n) \in \bar{S}_{h(\mathbf{z})}$ and $k(h(\mathbf{z})) = \sum_{c \geq 0} (1-c) h(\mathbf{z}, c) = k + 1$. By the previous lemma we can rewrite $\mathbf{z}$ uniquely as $\mathbf{y}^1 ... \mathbf{y}^{k+1}$, where $h(\mathbf{y}^i)$ is a degree statistics with $|h(\mathbf{y}^i)|_1 < n$ for each $i \in [k+1]$. By the induction hypothesis, $f_{h(\mathbf{y}^i)}$ is a bijection. Thus there exists $t_i \in \mathfrak{T}_{h(\mathbf{y}^i)}$ such that $f_{h(\mathbf{y}^i)}(t_i) = \mathbf{y}^i$. Define the tree t,

$$t := \{\emptyset\} \cup \bigcup_{i=1}^{k+1} t_i,$$

with root $\emptyset$ satisfying $d_t(\emptyset) = k+1$ and with subtrees $t_1, ..., t_{k+1}$, listed in lexicographic order. Then, by construction, $f_{\mathbf{n}}(t) = (k, x_2, ..., x_n) = \mathbf{x}$. □

Proof of Proposition 2.2. We prove the theorem by induction on $k(\mathbf{n})$. Proposition 2.3 handles the base case $k(\mathbf{n}) = 1$. To prove the inductive step, we first introduce some notation. Given degree statistics $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ and $\mathbf{m} = (\mathbf{m}(c),\, c \geq 0)$, we define the set

$$\bar{S}_{\mathbf{n}}\bar{S}_{\mathbf{m}} := \{\mathbf{xy} :\, \mathbf{x} \in \bar{S}_{\mathbf{n}},\, \mathbf{y} \in \bar{S}_{\mathbf{m}}\}.$$

Then, $\bar{S}_{\mathbf{n}}\bar{S}_{\mathbf{m}} \subseteq \bar{S}_{\mathbf{n}+\mathbf{m}}$, where $\mathbf{n}+\mathbf{m} = (\mathbf{n}(c)+\mathbf{m}(c),\, c \geq 0)$. Now let $\mathbf{n}$ be a degree statistics and define $M_{\mathbf{n}} := \{\mathbf{n}_{t_1} :\, (t_1, ..., t_{k(\mathbf{n})}) \in \mathfrak{F}_{\mathbf{n}}\}$ to be the set of all possible degree statistics of the first tree of a forest in $\mathfrak{F}_{\mathbf{n}}$. Thus, we can rewrite $\mathfrak{F}_{\mathbf{n}}$ as the disjoint union of sets,

$$\mathfrak{F}_{\mathbf{n}} = \bigcup_{\mathbf{m} \in M_{\mathbf{n}}} \mathfrak{T}_{\mathbf{m}}(1) \times \mathfrak{F}_{\mathbf{n}-\mathbf{m}}.$$

Let $\mathbf{m} \in M_{\mathbf{n}}$ and $(t, f) \in \mathfrak{T}_{\mathbf{m}}(1) \times \mathfrak{F}_{\mathbf{n}-\mathbf{m}}$. Let $u_1, ..., u_m, u_{m+1}, ..., u_n$ be the nodes of (t, f) in lexicographic order. Then,

$$\begin{aligned} f_{\mathbf{n}}(t,f) &= (d_t(u_1)-1, ..., d_t(u_m)-1, d_f(u_{m+1})-1, ..., d_f(u_{\mathbf{n}})) \\ &= (d_t(u_1)-1, ..., d_t(u_m)-1)(d_f(u_{m+1})-1...d_f(u_{\mathbf{n}})) \\ &= f_{\mathbf{m}}(t) f_{\mathbf{n}-\mathbf{m}}(f). \end{aligned}$$

By the induction hypothesis, since $\mathbf{m}$ and $\mathbf{n}-\mathbf{m}$ are degree statistics satisfying $k(\mathbf{m})=1$ and $k(\mathbf{n}-\mathbf{m})<k$, we have that $f_{\mathbf{m}}$ and $f_{\mathbf{n}-\mathbf{m}}$ are bijections. Thus,

$$f_{\mathbf{n}}(\mathfrak{T}_{\mathbf{m}}(1)\times\mathfrak{F}_{\mathbf{n}-\mathbf{m}})=f_{\mathbf{m}}(\mathfrak{T}_{\mathbf{m}}(1))f_{\mathbf{n}-\mathbf{m}}(\mathfrak{F}_{\mathbf{n}-\mathbf{m}})=\bar{S}_{\mathbf{m}}\bar{S}_{\mathbf{n}-\mathbf{m}}.$$

Since this holds for all $\mathbf{m}\in M_{\mathbf{n}}$,

$$f_{\mathbf{n}}(\mathfrak{F}_{\mathbf{n}})=\bigcup_{\mathbf{m}\in M_{\mathbf{m}}}f_{\mathbf{n}}(\mathfrak{T}_{\mathbf{m}}(1)\times\mathfrak{F}_{\mathbf{n}-\mathbf{m}})=\bigcup_{\mathbf{m}\in M_{\mathbf{n}}}\bar{S}_{\mathbf{m}}\bar{S}_{\mathbf{n}-\mathbf{m}}=\bar{S}_{\mathbf{n}},$$

where the last equality follows from Lemma 2.4. We conclude that $f_{\mathbf{n}}$ is a bijection for all degree statistics $\mathbf{n}$. □

We now study the relationship between $S_{\mathbf{n}}$ and $\bar{S}_{\mathbf{n}}$ via the Cyclic Lemma. Given a vector $\mathbf{z}=(z_1,...,z_r)$, we define $\mathbf{z}^{(i)}$ for $i\geq 0$ to be the permutation of $\mathbf{z}$ generated by a cyclic shift of i:

$$\mathbf{z}^{(i)}=(z_{1+i},...,z_{r+i}),$$

where $z_{j+i}=z_{j+i \mod r}$ for $j\in[r]$. Given a degree statistics $\mathbf{n}$, for $\mathbf{x}\in S_{\mathbf{n}}$ we define the set

$$C_{\mathbf{x}}:=\{i\in\mathbb{Z}/n\mathbb{Z}:\ \mathbf{x}^{(i)}\in\bar{S}_{\mathbf{n}}\}.$$

We write $|C_{\mathbf{x}}|$ to denote the cardinality of $C_{\mathbf{x}}$ and note that for all $i\in\mathbb{Z}$, $|C_{\mathbf{x}}|=|C_{\mathbf{x}^{(i)}}|$.

Proposition 2.5 (Cyclic Lemma). *Let $\mathbf{n}$ be a degree statistics and let $\mathbf{x}=(x_1,...,x_n)\in S_{\mathbf{n}}$. Then,*

$$C_{\mathbf{x}}=\left\{\lambda_1(\mathbf{x}),...,\lambda_{k(\mathbf{n})}(\mathbf{x})\right\},$$

where for $1\leq i\leq k(\mathbf{n})$,

$$\lambda_i(\mathbf{x}):=\min(j\in[n]:\ x_1+...+x_j=m+i-1),$$

where $m = \min(x_1 + ... + x_j : j \in [n])$.

Before proving the Cyclic lemma, we state and prove the following intermediate result.

Lemma 2.6. *Let* $\mathbf{n}$ *be a degree statistics and let* $\mathbf{x} \in S_{\mathbf{n}}$*, then,*

$$|C_{\mathbf{x}}| = k(\mathbf{n}).$$

Proof. Let $\mathbf{n}$ be a degree statistics and let $\mathbf{x} = (x_1, ..., x_n) \in S_{\mathbf{n}}$. For $a \geq 1$, define the sequence $\mathbf{a} := (a, -1, ..., -1)$ composed of a followed by a times -1. Recalling the definition of h from (2.2), note that $h(\mathbf{a}) = (h(\mathbf{a}, c),\ c \geq 0)$ satisfies $k(h(\mathbf{a})) = \sum_{c \geq 0}(1-c)h(\mathbf{a}, c) = 0$ since $h(\mathbf{a}, 0) = a$ and $h(\mathbf{a}, a+1) = 1$. Further, note that since $h(\mathbf{ax}) = h(\mathbf{a}) + h(\mathbf{x})$,

$$k(h(\mathbf{ax})) = \sum_{c \geq 0}(1-c)h(\mathbf{ax}, c) = k(h(\mathbf{a})) + k(h(\mathbf{x})) = k(h(\mathbf{x})) = k(\mathbf{n}). \tag{2.3}$$

It follows that $h(\mathbf{ax})$ is a degree statistics and $\mathbf{ax} \in S_{h(\mathbf{ax})}$. Thus $C_{\mathbf{ax}}$ is defined. We begin by proving that for all $a \geq 1$,

$$|C_{\mathbf{x}}| = |C_{\mathbf{ax}}|. \tag{2.4}$$

Clearly, $0 \in C_{\mathbf{x}}$ if and only if $0 \in C_{\mathbf{ax}}$. Now, for $0 \leq j \leq n-1$, we note that

$$(\mathbf{ax})^{(a+1+j)} = (x_{j+1}, ..., x_n, a, -1, ..., -1, x_1, ..., x_j).$$

We can see that $j \in C_{\mathbf{x}}$ if and only if $a + 1 + j \in C_{\mathbf{ax}}$, for all $0 \leq j \leq n-1$. Now, for $i \in [a]$, the sequence $(\mathbf{ax})^{(i)}$ begins with -1 and can therefore not be in $\bar{S}_{h(\mathbf{ax})}$, it follows that $i \notin C_{\mathbf{ax}}$ for all $i \in [a]$. This shows that (2.4) holds.

We now proceed to show that for all fixed $k \geq 0$ and degree statistics $\mathbf{n}$ with $k(\mathbf{n}) = k$, we have that $|C_{\mathbf{x}}| = k$. Fix $k \geq 0$. We show this result by induction on the size n of $\mathbf{n}$. Since $k(\mathbf{n}) = k$, the base case is given by $n = k$, precisely, when $\mathbf{n} = (k, 0, 0, ...)$. In this case, $\mathbf{x} = (x_1, ..., x_k) = (-1, ..., -1)$, thus for all $i \in [k]$, $i \in C_{\mathbf{x}}$ and $|C_{\mathbf{x}}| = k$. We now prove the induction step. Let $\mathbf{n}$ be a degree statistics with $k(\mathbf{n}) = k$ and let $\mathbf{x} = (x_1, ..., x_n) \in S_{\mathbf{n}}$. By definition of $C_{\mathbf{x}}$, for any cyclic permutation $\mathbf{x}^{(i)}$ of $\mathbf{x}$ we have that $|C_{\mathbf{x}}| = |C_{\mathbf{x}^{(i)}}|$. If $n \geq k$, there exists some $i \in [n]$ such that $x_i \geq 0$. We can assume without loss of generality that $x_1 \geq 0$. Denote by $1 = i_1 < i_2 < ... < i_m$ the indices such that $x_{i_j} \geq 0$. Define $i_{m+1} := n + 1$ and note that for $1 \leq j \leq m$, the number of consecutive -1 following x_{i_j} is given by $i_{j+1} - i_j - 1$. Then we can write,

$$-k(\mathbf{n}) = \sum_{i=1}^{n} x_i = \sum_{j=1}^{m} (x_{i_j} - (i_{j+1} - i_j - 1)).$$

As the sum is negative, there exists some index $J \in [m]$ such that $x_{i_J} \leq i_{J+1} - i_J - 1$. Thus, the sequence $(x_{i_J}, -1, ... - 1)$, containing x_{i_J} times -1 is a substring of $\mathbf{x}$. Let $\mathbf{z}$ be the sequence obtained by removing this substring from $\mathbf{x}$. Note that $k(h(\mathbf{z})) = k(h(\mathbf{x})) = k(\mathbf{n}) = k$ by (2.3). Since $h(\mathbf{z})$ is a degree statistics with $k(h(\mathbf{z})) = k$ and $|h(\mathbf{z})|_1 < n$, by induction we have that $|C_{\mathbf{z}}| = k$. Now, by (2.4) and since $|C_{\mathbf{x}}|$ does not change for cyclic permutations of $\mathbf{x}$, we conclude that $|C_{\mathbf{x}}| = |C_{\mathbf{z}}| = k$. $\square$

Proof of the Cyclic Lemma. Let $\mathbf{n}$ be a degree statistics and let $\mathbf{x} \in S_{\mathbf{n}}$. We now proceed to show that $C_{\mathbf{x}} = \{\lambda_1(\mathbf{x}), ..., \lambda_{k(\mathbf{n})}(\mathbf{x})\}$. For $\lambda = \lambda_1(\mathbf{x}) := \min(j \in [n] : x_1 + ... + x_j)$, we have that for all $j \in [n]$,

$$\sum_{i=1}^{j} x_{i+\lambda} = \sum_{i=1}^{j+\lambda} x_i - \sum_{i=1}^{\lambda} x_i \geq 0 > -1.$$

Therefore $\mathbf{x}^{(\lambda)} = (x_{1+\lambda}, ..., x_{n+\lambda}) \in \bar{S}_{\mathbf{n}}$. Further, by Lemma 2.4, we can rewrite $\mathbf{x}^{(\lambda)}$ as

$$\mathbf{x}^{(\lambda)} = \mathbf{y}^1 ... \mathbf{y}^k,$$

where $k = k(\mathbf{n})$ and $\mathbf{y}^i = (x_{1+\lambda+l_i}, ..., x_{1+\lambda+l_{i+1}})$, with $l_1 = 0$ and $l_i = \min(j \in [n] : x_{1+\lambda} + ... + x_{1+\lambda+j} = -i + 1)$. Since for each $i \in [k]$, $\mathbf{y}^i \in \bar{S}_{h(\mathbf{y}^i)}$, we have that for $0 \leq i \leq k-1$,

$$\mathbf{x}^{(\lambda+l_i)} = \mathbf{y}^i ... \mathbf{y}^{i+k} \in \bar{S}_{\mathbf{n}}.$$

Let $m = \min(x_1 + ... + x_j : j \in [n])$, then for $i \in [k]$,

$$\lambda_i = \lambda + l_i = \min(j \in [n] :\, x_1 + ... + x_j = m + i - 1 : j \in [n]). \qquad \square$$

Corollary 2.7. *For a degree statistics* $\mathbf{n}$*,*

$$|\bar{S}_{\mathbf{n}}| = \frac{k(\mathbf{n})}{n} |S_{\mathbf{n}}|.$$

Proof. The corollary holds since the Cyclic Lemma 2.5 gives that each $\mathbf{x} \in S_{\mathbf{n}}$ generates $k(\mathbf{n})$ vectors in $\bar{S}_{\mathbf{n}}$ and for each $\mathbf{x} \in S_{\mathbf{n}}$, its n cyclic shifts of $\mathbf{x} \in S_{\mathbf{n}}$ generate the same $k(\mathbf{n})$ vectors in $\bar{S}_{\mathbf{n}}$. $\square$

Using these results, we can deduce the following combinatorial identities.

Proposition 2.8. *For any degree statistics* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$*, we have*

$$|\mathfrak{F}_{\mathbf{n}}| = \frac{k(\mathbf{n})}{n} \binom{n}{\mathbf{n}(c),\ c \geq 0}.$$

Proof. Given a degree statistics $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$, by Proposition 2.2 and Corollary 2.7 ,

$$|\mathfrak{F}_{\mathbf{n}}| = |\bar{S}_{\mathbf{n}}| = \frac{k(\mathbf{n})}{n} |S_{\mathbf{n}}|.$$

Remark that by definition, $S_{\mathbf{n}}$ is equivalent to the set of $(n, k(\mathbf{n}))$-lattice paths with $\mathbf{n}(c)$ steps of size $c-1$ for all $c \geq 0$. Thus, $|S_{\mathbf{n}}|$ is the number of such $(n, k(\mathbf{n}))$-lattice paths, precisely,

$$|S_{\mathbf{n}}| = \binom{n}{\mathbf{n}(c),\ c \geq 0}.$$

The proposition follows. □

Proposition 2.9. *For any degree statistics* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *with* $k(\mathbf{n}) = 1$*, we have*

$$|\mathfrak{T}_{\mathbf{n}}| = \frac{1}{n}\binom{n}{\mathbf{n}(c),\ c \geq 0}.$$

Proof. From the definition of $\mathfrak{T}_{\mathbf{n}}(j)$, $|\mathfrak{T}_{\mathbf{n}}(j)| = |\mathfrak{T}_{\mathbf{n}}|$. Further, for degree statistics $\mathbf{n}$ with $k(\mathbf{n}) = 1$ it holds that $|\mathfrak{F}_{\mathbf{n}}| = |\mathfrak{T}_{\mathbf{n}}(1)|$. Applying Proposition 2.8 gives the desired equality. □

Proposition 2.10. *Given* $n \geq 1$*, let bin*$(n, 1) := (n, 0, n-1, 0, 0, ...)$*. Recall that* $\mathfrak{T}_{bin(n,1)}$ *is the set of binary trees with* n *leaves. Then,*

$$|\mathfrak{T}_{bin(n,1)}| = \frac{1}{2n-1}\binom{2n-1}{n}.$$

Proof. This is a straightforward application of Proposition 2.9 with $\mathbf{n} = \text{bin}(n, 1)$. □

Given a degree statistics $\mathbf{n}$, we define a *marked forest* to be a pair (f, u), where $f \in \mathfrak{F}_{\mathbf{n}}$ and $u \in f$ and call the node u the *mark* of f. For $1 \leq i \leq k(\mathbf{n})$, we define the set $\mathfrak{F}_{\mathbf{n}}^{(i)}$ as the set of marked forests, with forests in $\mathfrak{F}_{\mathbf{n}}$ and with a mark in the i^{th} tree,

$$\mathfrak{F}_{\mathbf{n}}^{(i)} := \{((t_1, ..., t_{k(\mathbf{n})}), u) :\ (t_1, ..., t_{k(\mathbf{n})}) \in \mathfrak{F}_{\mathbf{n}},\ u \in t_i\}.$$

In particular, if $\mathbf{n}$ is a degree statistics, then we define a *marked tree* as a pair (t, u), where $t \in \mathfrak{T}_{\mathbf{n}}$ and $u \in t$, and define

$$\mathfrak{T}_{\mathbf{n}}^{(1)} := \{(t, u) :\ t \in \mathfrak{T}_{\mathbf{n}},\ u \in t\}$$

as the set of marked trees with degree statistics $\mathbf{n}$. Similarly, for $j \geq 1$, let $\mathfrak{T}_{\mathbf{n}}^{(1)}(j) := \{(jt, u) : t \in \mathfrak{T}_{\mathbf{n}}, u \in jt\}$ and note that $|\mathfrak{T}_{\mathbf{n}}^{(1)}(j)| = |\mathfrak{T}_{\mathbf{n}}^{(1)}|$. We now deduce combinatorial identities on the number of marked trees and forests.

Proposition 2.11. *For any degree statistics* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *with* $k(\mathbf{n}) = 1$*, the number of marked trees with degree statistics* $\mathbf{n}$ *is given by*

$$\left|\mathfrak{T}_{\mathbf{n}}^{(1)}\right| = \binom{n}{\mathbf{n}(c),\ c \geq 0}.$$

Proof. For all $t \in \mathfrak{T}_{\mathbf{n}}$, we have $|t| = n$ thus $|\{(t, u) : u \in t\}| = n$ and

$$\left|\mathfrak{T}_{\mathbf{n}}^{(1)}\right| = n \cdot |\mathfrak{T}_{\mathbf{n}}| = \binom{n}{\mathbf{n}(c),\ c \geq 0},$$

where the second equality follows by Proposition 2.9. □

Proposition 2.12. *For any degree statistics* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *and* $i \in [k(\mathbf{n})]$*, the number of forests with degree statistics* $\mathbf{n}$ *and a mark in the* i^{th} *tree is given by*

$$\left|\mathfrak{F}_{\mathbf{n}}^{(i)}\right| = \binom{n}{\mathbf{n}(c),\ c \geq 0}.$$

Proof. We begin by noting that for any $1 \leq i < j \leq k := k(\mathbf{n})$, we have $((t_1, ..., t_i, ..., t_j, ..., t_k), u) \in \mathfrak{F}_{\mathbf{n}}^{(i)}$ if and only if $((t_1, ..., t_j, ..., t_i, ..., t_k), u) \in \mathfrak{F}_{\mathbf{n}}^{(j)}$. Therefore $\left|\mathfrak{F}_{\mathbf{n}}^{(i)}\right| = \left|\mathfrak{F}_{\mathbf{n}}^{(j)}\right|$ for all $1 \leq i, j \leq k$ and $\left|\bigcup_{i=1}^{k} \mathfrak{F}_{\mathbf{n}}^{(i)}\right| = k \left|\mathfrak{F}_{\mathbf{n}}^{(1)}\right|$. Now, for $f \in \mathfrak{F}_{\mathbf{n}}$, $|f| = n$ and so $|(f, v) : v \in f| = n$. Thus, every forest $f \in \mathfrak{F}_{\mathbf{n}}$ generates n marked forests in $\bigcup_{i=1}^{k} \mathfrak{F}_{\mathbf{n}}^{(i)}$. Therefore,

$$|\mathfrak{F}_{\mathbf{n}}| = \frac{1}{n} \left|\bigcup_{i=1}^{k} \mathfrak{F}_{\mathbf{n}}^{(i)}\right| = \frac{k}{n} \left|\mathfrak{F}_{\mathbf{n}}^{(1)}\right|.$$

By rearranging we get,

$$\left|\mathfrak{F}_{\mathbf{n}}^{(1)}\right| = \frac{n}{k}\left|\mathfrak{F}_{\mathbf{n}}\right| = \binom{n}{\mathbf{n}(c),\ c \geq 0},$$

where the last equality holds by Proposition 2.8. □

Proposition 2.13. *For $n \geq k \geq 1$ and $1 \leq i \leq k$, the number of binary forests with n leaves, k trees and a mark in the i^{th} tree is given by*

$$\left|\mathfrak{F}_{bin(n,k)}^{(i)}\right| = \binom{2n-k}{n},$$

where bin$(n,k) := (n, 0, n-k, 0, 0, ...)$.

Proof. This is a straightforward application of Proposition 2.12 with $\mathbf{n} = \text{bin}(n,k)$. □

Chapter 3

BOUNDS ON THE HEIGHT OF A RANDOM NODE IN A RANDOM TREE

This chapter focuses on proving novel bounds on the height of random nodes in random trees with a fixed degree statistics. Recall the main theorems of this thesis.

Theorem 1.4. *Fix a degree statistics* $\mathbf{n} = (\mathbf{n}(c),\ c \geq 0)$ *with* $k(\mathbf{n}) = 1$ *and let* $(T, V) \in_u \mathfrak{T}^{(1)}_{\mathbf{n}}$. *Then for all* $\alpha > 17^{3/2}$,

$$\mathbf{P}\left\{|V| > \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right),$$

and if $\mathbf{n}(1) = 0$, *then for all* $\ell \geq 1$,

$$\mathbf{P}\{|V| \geq \ell\} \leq \exp\left(-\frac{\ell^2}{2|\mathbf{n}|_1}\right).$$

Theorem 1.5. *Let* $m \geq 1$. *Let* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *be a degree statistics with* $\mathbf{n}(0) = m$, $\mathbf{n}(1) = 0$ *and* $k(\mathbf{n}) = 1$. *Further, recall that bin*$(m, 1) := (m, 0, m-1, 0, ...)$. *Let*

$(T,V) \in_u \mathfrak{T}^{(1)}_{\mathbf{n}}$ and $(T',W) \in_u \mathfrak{T}^{(1)}_{bin(m,1)}$ be random marked trees. Then,

$$|V| \preceq_{st} |W|.$$

Note that Theorem 1.5 yields the second bound of Theorem 1.4 though this is not exactly how we prove that bound.

3.1 Proof of Theorem 1.5

Theorem 1.5 is essentially a consequence of the following Proposition.

Proposition 3.1. *Under the same assumptions of Theorem 1.5 we have that for all $k \geq 0$,*

$$\mathbf{P}\{|V| = k \mid |V| > k-1\} \geq \mathbf{P}\{|W| = k \mid |W| > k-1\}.$$

In order to prove Theorem 1.5 from Proposition 3.1 we use the following lemma.

Lemma 3.2. *Given two non-negative integer random variables X, Y such that for all $k \in \mathbb{N}$, $\mathbf{P}\{X = k \mid X > k-1\} \geq \mathbf{P}\{Y = k \mid Y > k-1\}$ then $X \preceq_{st} Y$.*

Proof. Define $p_k := \mathbf{P}\{Y = k \mid Y > k-1\}$ and $q_k = \mathbf{P}\{X = k \mid X > k-1\}$ for $k \geq 0$, by assumption $p_k \leq q_k$. Note that $X \preceq_{st} Y$ holds if and only if there exists a coupling $(\hat{X}, \hat{Y})$ of X and Y such that $\mathbf{P}\left\{\hat{Y} \leq k\right\} \leq \mathbf{P}\left\{\hat{X} \leq k\right\}$ for all $k \geq 0$. We thus construct such a coupling. Let $(U_i,\ i \geq 0)$ be a sequence of i.i.d. Uniform$[0,1]$ random variables. Define

$$\hat{X} := \min(i \geq 0 :\ U_i \leq q_i) \text{ and } \hat{Y} := \min(i \geq 0 :\ U_i \leq p_i).$$

By this definition, $\hat{Y} \leq \hat{X}$. We now show that $(\hat{X}, \hat{Y})$ is indeed a coupling of X and Y. For $k = 0$, $\mathbf{P}\left\{\hat{X} = 0\right\} = \mathbf{P}\{X = 0\}$ and for $k \geq 1$,

$$\begin{aligned}\mathbf{P}\left\{\hat{X} = k\right\} &= \mathbf{P}\{\min(i \geq 0 : U_i \leq q_i) = k\} \\ &= (1 - q_0) \cdot (1 - q_1) \cdot \ldots \cdot (1 - q_{k-1}) \cdot q_k \\ &= \mathbf{P}\{X > 0\} \mathbf{P}\{X > 1 \mid X > 0\} \ldots \mathbf{P}\{X = k \mid X > k - 1\} \\ &= \mathbf{P}\{X = k\}.\end{aligned}$$

The same reasoning gives $\mathbf{P}\left\{\hat{Y} = k\right\} = \mathbf{P}\{Y = k\}$ for all $k \geq 0$. □

By applying Proposition 3.1 to Lemma 3.2, we conclude that $|V| \preceq_{st} |W|$, in other words, Theorem 1.5 holds. We now focus on proving Proposition 3.1. The method to prove this proposition consists in decomposing a marked tree into a branch from the root to the marked node, and the forest formed by the trees hanging off the branch. We begin by introducing the notion of a *spine* of a marked tree. Recall that given a node $u = i_1 \ldots i_{|u|} \in t$ where $i_1, \ldots, i_{|u|} \in \mathbb{N}$, we defined $u_k := i_1 \ldots i_k$ for $0 \leq k \leq |u|$, where $u_0 = \emptyset$. Now, given a degree statistics $\mathbf{n} = (\mathbf{n}(c),\ c \geq 0)$ with $k(\mathbf{n}) = 1$, let $(t, v) \in \mathfrak{T}_{\mathbf{n}}^{(1)}$ be a marked tree. We define the *k-spine* of (t, v) as a subtree of t containing all nodes, and its children, that are on the path from the root to the marked node. Formally, the *k-spine* of (t, v) is the subtree

$$\mathrm{S}_k(t, v) := \{\emptyset\} \cup \left\{ \bigcup_{i=0}^{k-1} c_t(v_i) \right\},$$

with $\mathrm{S}_0(t, v) := \emptyset$. By construction, $S_k(t, v)$ contains $k - 1$ internal nodes with degrees $d_t(v_0), \ldots, d_t(v_{k-1})$ and where v_{i-1} is the parent node of v_i for each $1 \leq i \leq k$. Further, we define the *marked k-spine* of (t, v) to be the marked tree

$$\mathrm{S}_k^{\bullet}(t, v) := (\mathrm{S}_k(t, v), v_k).$$

When $k = |v|$, we say $\mathrm{S}(t,v) := \mathrm{S}_{|v|}(t,v)$ is the *spine* of (t,v) and $\mathrm{S}^{\bullet}(t,v) := \mathrm{S}^{\bullet}_{|v|}(t,v) = (\mathrm{S}(t,v),v)$ is the *marked spine* of (t,v).

Let $\hat{t} \in \mathfrak{T}^*$ be some finite tree and let (t,v) be a marked tree such that t is a subtree of $\hat{t}$ and $v \in t$ is not a leaf in $\hat{t}$. For $1 \le b \le d_{\hat{t}}(v)$, we define the marked tree

$$(t,v)_b := (t', vb),$$

where $t' := t \cup \bigcup_{1\le i\le d_{\hat{t}}(v)} \{vi\}$ and vb is the b^{th} child of v. From this definition, we note that for a marked tree (t,v), for $1 \le b \le d_{\hat{t}}(v)$ and $0 \le k \le |v|$,

$$(\mathrm{S}^{\bullet}_k(t,v))_b := (\mathrm{S}_{k+1}(t,v), v_k b).$$

For $(T,V) \in_u \mathfrak{T}^{(1)}_{\mathbf{n}}$, the proposition below gives the probability that $|V| > k$ and $|V| = k$ respectively, given the marked k-spine of (T,V).

Proposition 3.3. *Let $\mathbf{n} = (\mathbf{n}(c), c \ge 0)$ be a degree statistics with $k(\mathbf{n}) = 1$ and let $(T,V) \in_u \mathfrak{T}^{(1)}_{\mathbf{n}}$. Fix $k \ge 0$ and a marked tree (t,v) such that $\mathbf{P}\{S^{\bullet}_k(T,V) = (t,v)\} > 0$, then*

$$\mathbf{P}\{V \ne v \mid S^{\bullet}_k(T,V) = (t,v)\} = \frac{\sum_{c\ge 0} c(\mathbf{n}(c) - \mathbf{n}_t(c))}{n-k}. \tag{3.1}$$

Further,

$$\mathbf{P}\{V = v \mid S^{\bullet}_k(T,V) = (t,v)\} = \frac{\mathbf{n}_t(0)}{n-k}. \tag{3.2}$$

The idea behind these equalities is to decompose a marked tree into its k-spine t, containing $\mathbf{n}_t(0)$ leaves, and a forest composed of $\mathbf{n}_t(0)$ trees hanging off the leaves of t. The number of nodes in this forest is precisely $n-k$ (if we count the leaves of t as roots of the trees in the forest). Given $\mathrm{S}^{\bullet}_k(T,V) = (t,v)$, since $V \in_u T$, the probability that $V = v$ is one over the size of the tree attached to v.

The second equality can then be interpreted as the inverse of the expected size of the subtree above v, given that $\mathrm{S}_k^\bullet(T,V) = (t,v)$. Before proving Proposition 3.3, we prove the following lemmas.

Lemma 3.4. *Let* $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ *be a degree statistics with* $k(\mathbf{n}) = 1$ *and let* $(T,V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$. *Fix* $k \geq 0$ *and a marked tree* (t,v) *such that* $\mathbf{P}\left\{S_k^\bullet(T,V) = (t,v)\right\} > 0$, *then*

$$\mathbf{P}\left\{S_k^\bullet(T,V) = (t,v)\right\} = \left|\mathfrak{F}_{\mathbf{q}_t}^{(1)}\right| \cdot \left|\mathfrak{T}_{\mathbf{n}}^{(1)}\right|^{-1},$$

where $\mathbf{q}_t := (\mathbf{q}_t(c), c \geq 0)$, *with* $\mathbf{q}_t(0) = \mathbf{n}(0)$ *and* $\mathbf{q}_t(c) = \mathbf{n}(c) - \mathbf{n}_t(c)$ *for all* $c \geq 1$.

Proof. Define the set

$$A_k(t,v) := \left\{(\hat{t},\hat{v}) \in \mathfrak{T}_{\mathbf{n}}^{(1)} : \mathrm{S}_k^\bullet(\hat{t},\hat{v}) = (t,v)\right\},$$

as the set of marked trees in $\mathfrak{T}_{\mathbf{n}}^{(1)}$ with k-spine (t,v). Denote by $|A_k(t,v)|$ the size of the set $A_k(t,v)$. Note that since $\mathbf{P}\left\{\mathrm{S}_k^\bullet(T,V) = (t,v)\right\} > 0$, we have $|A_k(t,v)| > 0$. The probability that (T,V) has marked k-spine (t,v) is given by the number of trees in $\mathfrak{T}_{\mathbf{n}}^{(1)}$ with marked k-spine (t,v), divided by the total number of trees in $\mathfrak{T}_{\mathbf{n}}^{(1)}$. Thus,

$$\mathbf{P}\left\{\mathrm{S}_k^\bullet(T,V) = (t,v)\right\} = |A_k(t,v)| \cdot \left|\mathfrak{T}_{\mathbf{n}}^{(1)}\right|^{-1}.$$

We now compute $|A_k(t,v)|$. Define the degree statistics $\mathbf{q}_t := (\mathbf{q}_t(c), c \geq 0)$ as $\mathbf{q}_t(0) := \mathbf{n}(0)$ and $\mathbf{q}_t(c) := \mathbf{n}(c) - \mathbf{n}_t(c)$ for all $c \geq 1$. This is in fact a degree statistics since t is a subtree of $\hat{t}$, thus $\mathbf{q}_t(c) \geq 0$ for all $c \geq 0$ and

$$k(\mathbf{q}_t) = \sum_{c\geq 0}(1-c)\mathbf{q}_t(c) = k(\mathbf{n}) - \sum_{c\geq 1}(1-c)\mathbf{n}_t(c) = k(\mathbf{n}) - k(\mathbf{n}_t) + \mathbf{n}_t(0) = \mathbf{n}_t(0).$$

Consider $\mathfrak{F}_{\mathbf{q}_t}^{(1)}$, the set of forests with degree statistics $\mathbf{q}_t$, with $k(\mathbf{q}_t) = \mathbf{n}_0(t)$ trees and a mark in the first tree. Given a marked forest $(f,u) \in \mathfrak{F}_{\mathbf{q}_t}^{(1)}$, where $f =$

$(t_1, ..., t_{\mathbf{n}_0(t)})$ and $u = 1\hat{u}$, we construct a marked tree in $A_k(t,v)$ as follows. Note that the mark v is a leaf node in t. Let $l_1 := v$ and $l_2, ..., l_{\mathbf{n}_t(0)}$ be the remaining leaves of t in lexicographic order. Now, take the tree t and, for all $1 \leq i \leq \mathbf{n}_t(0)$, respectively attach the tree t_i to the leaf l_i. Call the resulting tree $\hat{t}$ and let $\hat{v} := v\hat{u}$. The degree statistics of $\hat{t}$ is $\mathbf{n}_{\hat{t}} = (\mathbf{n}_{\hat{t}}(c), c \geq 0)$ where $\mathbf{n}_{\hat{t}}(0) = \mathbf{n}(0)$ and $\mathbf{n}_{\hat{t}}(c) = \mathbf{q}_t(c) + \mathbf{n}_t(c) = \mathbf{n}(c)$ for all $c \geq 0$ thus $\mathbf{n}_{\hat{t}} = \mathbf{n}$. Therefore we have that $(\hat{t}, \hat{v}) \in A_k(t,v)$. Further, for any tree in $A_k(t,v)$ we can recover a unique forest in $\mathfrak{F}^{(1)}_{\mathbf{q}_t}$ via the reverse process. We conclude that $|A_k(t,v)| = \left|\mathfrak{F}^{(1)}_{\mathbf{q}}\right|$. ☐

Proof of Proposition 3.3. We now proceed to prove that for a fixed $k \geq 0$ and a marked tree (t,v) with $\mathbf{P}\{\mathrm{S}^\bullet_k(T,V) = (t,v)\} > 0$, we have that

$$\mathbf{P}\{V \neq v \mid \mathrm{S}^\bullet_k(T,V) = (t,v)\} = \frac{\sum_{c\geq 0} c(\mathbf{n}(c) - \mathbf{n}_t(c))}{n-k}.$$

Note that if $V \neq v$ then $d_T(v) > 0$. Hence,

$$\begin{aligned}
&\mathbf{P}\{V \neq v \mid \mathrm{S}^\bullet_k(T,V) = (t,v)\} \\
&= \sum_{b>0} \mathbf{P}\{V \neq v, d_T(v) = b \mid \mathrm{S}^\bullet_k(T,V) = (t,v)\} \\
&= \sum_{b>0} \sum_{c=1}^{b} \mathbf{P}\left\{d_T(v) = b,\, \mathrm{S}^\bullet_{k+1}(T,V) = (t,v)_c \;\middle|\; \mathrm{S}^\bullet_k(T,V) = (t,v)\right\} \\
&= \sum_{b>0} b \cdot \mathbf{P}\left\{d_T(v) = b,\, \mathrm{S}^\bullet_{k+1}(T,V) = (t,v)_1 \;\middle|\; \mathrm{S}^\bullet_k(T,V) = (t,v)\right\},
\end{aligned}$$

where the last equality holds by symmetry. Let $(t', v1) = (t,v)_1$, then note that if $d_T(v) = b$ and $\mathrm{S}^\bullet_{k+1}(T,V) = (t,v)_1$, we have that $\mathbf{n}_{t'}(b) = \mathbf{n}_t(b) + 1$ and $\mathbf{n}_{t'}(c) = \mathbf{n}_t(c)$ for all $c \neq b$. By Lemma 3.4,

$$\mathbf{P}\{V \neq v \mid \mathrm{S}^\bullet_k(T,V) = (t,v)\} = \sum_{b>0} b \left|\mathfrak{F}^{(1)}_{\mathbf{q}_{t'}}\right| \cdot \left|\mathfrak{F}^{(1)}_{\mathbf{q}_t}\right|^{-1}.$$

Note that $\|\mathbf{q}_{t'}\| = \|\mathbf{q}_t\| - 1$, $\mathbf{q}_{t'}(b) = \mathbf{q}_t(b) - 1$ and $\mathbf{q}_{t'}(c) = \mathbf{q}_t(c)$ for all $c \neq b$. Further, since t contains k internal nodes, $\|\mathbf{q}_t\| = \mathbf{n}(0) + \sum_{c\geq 1}(\mathbf{n}(c) - \mathbf{n}_t(c) = n - k$. Therefore, by Proposition 2.12,

$$\begin{aligned}\mathbf{P}\{V \neq v \mid \mathrm{S}_k^\bullet(T,V) = (t,v)\} &= \sum_{b>0} b\binom{\|\mathbf{q}_{t'}\|}{\mathbf{q}_{t'}(c),\, c\geq 0}\binom{\|\mathbf{q}_t\|}{\mathbf{q}_t(c),\, c\geq 0}\\ &= \sum_{b>0} b\frac{\mathbf{q}_t(b)}{\|\mathbf{q}\|} = \sum_{b>0} b\frac{\mathbf{n}(b) - \mathbf{n}_t(b)}{n-k}.\end{aligned}$$

This is precisely the first equality of Proposition 3.3. To show the second equality, note that,

$$\begin{aligned}\mathbf{P}\{V = v \mid \mathrm{S}_k^\bullet(T,V) = (t,v)\} &= 1 - \mathbf{P}\{V \neq v \mid \mathrm{S}_k^\bullet(T,V) = (t,v)\}\\ &= 1 - \frac{\sum_{c>0} c(\mathbf{n}(c) - \mathbf{n}_t(c))}{n-k}\\ &= \frac{n-k-\sum_{c>0} c(\mathbf{n}(c) - \mathbf{n}_t(c))}{n-k}\\ &= \frac{1 + \sum_{c>0}(c-1)\mathbf{n}_t(c)}{n-k},\end{aligned}$$

where the last equality holds since $n = \sum_{c\geq 0} c\mathbf{n}(c) + 1$ and $\sum_{c>0} \mathbf{n}_t(c) = k$, so $n - k - \sum_{c>0} c(\mathbf{n}(c) - \mathbf{n}_t(c)) = 1 - k + \sum_{c>0} c\mathbf{n}_t(c) = 1 + \sum_{c>0}(c-1)\mathbf{n}_t(c)$. Now since $k(\mathbf{n}) = \sum_{c\geq 0}(1-c)\mathbf{n}_t(c) = 1$, it follows that $1 + \sum_{c>0}(c-1)\mathbf{n}_t(c) = \mathbf{n}_t(0)$. We conclude that

$$\mathbf{P}\{V = v \mid \mathrm{S}_k^\bullet(T,V) = (t,v)\} = \frac{n_0(t)}{n-k}. \quad \square$$

We now have all the tools to prove Proposition 3.1.

Proof of Proposition 3.1. Let $m \geq 1$ and let $\mathbf{n} = (\mathbf{n}(c),\, c \geq 0)$ be a degree statistics with $\mathbf{n}(0) = m$, $\mathbf{n}(1) = 0$ and $k(\mathbf{n}) = 1$. Further, let $\text{bin}(m,1) = (m, 0, m-1, 0, \ldots)$. Let $(T,V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$ and $(T',W) \in_u \mathfrak{T}_{\text{bin}(m,1)}^{(1)}$ be random marked trees. In the following equalities we write $\bigcup_{(t,v)}$ and $\sum_{(t,v)}$ to mean that we are iterating through all marked trees (t, v) satisfying $\mathbf{P}\{\mathrm{S}_k^\bullet(T, V) = (t, v)\} > 0$. Similarly, we write $\bigcup_{(t',w)}$ and $\sum_{(t',w)}$ to mean that we are iterating through all marked trees (t', w) satisfying $\mathbf{P}\{\mathrm{S}_k^\bullet(T', W) = (t', w)\} > 0$. Note that we can rewrite the events $\{|V| > k-1\}$ and $\{|W| > k-1\}$ as disjoint unions,

$$\{|V| > k-1\} = \bigcup_{(t,v)} \{\mathrm{S}_k^\bullet(T, V) = (t, v)\} \text{ and } \{|W| > k-1\} = \bigcup_{(t',w)} \{\mathrm{S}_k^\bullet(T', W) = (t', w)\}.$$

Therefore,

$$\begin{aligned}
&\mathbf{P}\left\{|V| = k \mid |V| > k-1\right\} \\
&= \frac{\mathbf{P}\left\{|V| > k-1, |V| = k\right\}}{\mathbf{P}\left\{|V| > k-1\right\}} = \sum_{(t,v)} \frac{\mathbf{P}\left\{\mathrm{S}_k^\bullet(T, V) = (t, v),\, |V| = k\right\}}{\mathbf{P}\left\{|V| > k-1\right\}} \\
&= \sum_{(t,v)} \mathbf{P}\left\{|V| = k \mid \mathrm{S}_k^\bullet(T, V) = (t, v)\right\} \cdot \frac{\mathbf{P}\left\{\mathrm{S}_k^\bullet(T, V) = (t, v)\right\}}{\mathbf{P}\left\{|V| > k-1\right\}} \\
&= \sum_{(t,v)} \mathbf{P}\left\{|V| = k \mid \mathrm{S}_k^\bullet(T, V) = (t, v)\right\} \cdot \mathbf{P}\left\{\mathrm{S}_k^\bullet(T, V) = (t, v) \mid |V| > k-1\right\} \\
&= \sum_{(t,v)} \frac{\mathbf{n}_t(0)}{n-k} \cdot \mathbf{P}\left\{\mathrm{S}_k^\bullet(T, V) = (t, v) \mid |V| > k-1\right\},
\end{aligned}$$

where the last equality holds by Proposition 3.3. Remark that, since $\mathbf{n}(1) = 0$, $n = \sum_{c \geq 0} c\mathbf{n}(c) + 1 \geq 2 \sum_{c \geq 2} \mathbf{n}(c) + 1 = 2n - 2m + 1$. Thus, $n \leq 2m - 1$. Further, since t has k internal nodes and by assumption $\mathbf{n}$ satisfies $\mathbf{n}(1) = 0$, we have that each internal node has at least degree 2. Thus t contains at least $k+1$ leaves, $\mathbf{n}_t(0) \geq k+1$. It follows that $\mathbf{n}_t(0)/(n-k) \geq (k+1)/(2n-1-k)$ and,

$$\begin{aligned}
\mathbf{P}\left\{|V| = k \mid |V| > k-1\right\} &\geq \frac{k+1}{2m-1-k} \sum_{(t,v)} \mathbf{P}\left\{\mathrm{S}_k^\bullet(T, V) = (t, v) \mid |V| > k-1\right\} \\
&= \frac{k+1}{2m-1-k}. \qquad (3.3)
\end{aligned}$$

Similarly, for $k \geq 0$,

$$\begin{aligned}
&\mathbf{P}\left\{|W| = k \mid |W| > k-1\right\} \\
&= \sum_{(t',w)} \mathbf{P}\left\{W = w \mid \mathrm{S}_k^\bullet(T',W) = (t',w)\right\} \mathbf{P}\left\{\mathrm{S}_k^\bullet(T',W) = (t',w) \mid |W| > k-1\right\} \\
&= \frac{k+1}{2m-k-1} \sum_{(t',w)} \mathbf{P}\left\{\mathrm{S}_k^\bullet(T',W) = (t',w) \mid |W| > k-1\right\} = \frac{k+1}{2m-k-1}, \qquad (3.4)
\end{aligned}$$

where the last equality holds by applying Proposition 3.3 with the degree statistics bin$(m,1)$. Combining (3.3) and (3.4) gives us the desired result,. For all $k \geq 0$,

$$\mathbf{P}\left\{|V| = k \mid |V| > k-1\right\} \geq \mathbf{P}\left\{|W| = k \mid |W| > k-1\right\}. \qquad \square$$

3.2 Proof of Theorem 1.4

Theorem 1.4. *Fix a degree statistics* $\mathbf{n} = (\mathbf{n}(c),\ c \geq 0)$ *with* $k(\mathbf{n}) = 1$ *and let* $(T,V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$. *Then for all* $\alpha > 17^{3/2}$,

$$\mathbf{P}\left\{|V| > \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3} \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right),$$

and if $\mathbf{n}(1) = 0$, *then for all* $\ell \geq 1$,

$$\mathbf{P}\left\{|V| \geq \ell\right\} \leq \exp\left(-\frac{\ell^2}{2|\mathbf{n}|_1}\right).$$

We begin with defining a *size biasing*. Let $\mathbf{n} = (\mathbf{n}(c),\ c \geq 0)$ be a degree statistics. A random vector $D = (D_1, \ldots, D_n)$ is a *size biasing* of $\mathbf{n}$ if for any degree sequence $d = (d_1, \ldots, d_n)$ satisfying $|\{i \in [n] : d_i = c\}| = \mathbf{n}(c)$ for all $c \geq 0$, the following holds. For each $k \in [n]$,

$$\mathbf{P}\left\{ D_k = d_k \mid (D_1, ..., D_{k-1}) = (d_1, ..., d_{k-1}) \right\} = \frac{d_k(\mathbf{n}(d_k) - w((d_1, ..., d_{k-1}), d_k))}{|\mathbf{n}|_1 - d_1 - .... - d_{k-1}}, \tag{3.5}$$

for all $k \in [n]$, where $w((d_1, ..., d_k), c) := |\{1 \le i \le k : d_i = c\}|$ for all $c \ge 0$. When $d_1 + ... + d_{k-1} = |\mathbf{n}|_1$, we consider the fraction to be equal to 1. By this definition, we have that the last $\mathbf{n}(0) \ge 1$ entries of D are equal to 0, further $D_n = 0$.

Our method for proving Theorem 1.4 is a sampling procedure that, for $(T, V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$, generates a random variable with the same law as $|V|$. We then derive tail bounds on this new random variable to deduce the desired bounds on $|V|$.

Proposition 3.5. *For a degree statistics $\mathbf{n}$ with $k(\mathbf{n}) = 1$, let $(T, V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$ and let $D = (D_1, ..., D_n)$ be a size biasing of $\mathbf{n}$. Let $(U_1, ..., U_n)$ be i.i.d. Uniform$[0, 1]$ random variables independent of D. For $i \in [n]$, define*

$$A_i := \begin{cases} 1 & \textit{if } U_i \le \frac{1 + \sum_{j=1}^{i-1}(D_j - 1)}{n+1-i}, \\ 0 & \textit{otherwise.} \end{cases}$$

Let $M = \inf(i \ge 1 : A_i = 1)$. Then $|V| \overset{d}{=} M - 1$.

Proposition 3.6. *For a degree statistics $\mathbf{n} = (\mathbf{n}(c),\, c \ge 0)$ with $k(\mathbf{n}) = 1$, let $D = (D_0, ..., D_{n-1})$ be a size biasing of $\mathbf{n}$. Let $(U_1, ..., U_n)$ be i.i.d. Uniform$[0, 1]$ random variables independent of D. For $i \in [n]$, define*

$$B_i := \begin{cases} 1 & \textit{if } U_i \le \frac{\sum_{j=1}^{i-1}(D_j - 1)}{n-i}, \\ 0 & \textit{otherwise.} \end{cases}$$

Let $\sigma = \inf(i \geq 1 : B_i = 1)$*. Then for* $\alpha \geq 17^{3/2}$*,*

$$\mathbf{P}\left\{\sigma > \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right),$$

and if $\mathbf{n}(1) = 0$*, then for all* $\ell \geq 1$*,*

$$\mathbf{P}\{\sigma \geq \ell\} \leq \exp\left(-\frac{(l-1)^2}{2|\mathbf{n}|_1}\right).$$

Theorem 1.4 is a consequence of these two propositions.

Proof of Theorem 1.4. Remark that for a size biasing $D = (D_1, ..., D_n)$ of $\mathbf{n}$, we have that for all $i \in [n]$, $\sum_{j=1}^{i-1} D_j \leq \sum_{j=1}^{n} D_j = n-1$. Thus,

$$\frac{1 + \sum_{j=1}^{i-1}(D_j - 1)}{\sum_{j=1}^{i-1}(D_j - 1)} = 1 + \frac{1}{\sum_{j=1}^{i-1}(D_j - 1)} \geq 1 + \frac{1}{n-i} = \frac{n+1-i}{n-i}.$$

Equivalently, we have that

$$\frac{1 + \sum_{j=1}^{i-1}(D_j - 1)}{n+1-i} \geq \frac{\sum_{j=1}^{i-1}(D_j - 1)}{n-i}.$$

Remark that in Proposition 3.5, $A_i = 1$ if and only if $U_i \leq \frac{1+\sum_{j=1}^{i-1}(D_j-1)}{n+1-i}$ and in Proposition 3.6, $B_i = 1$ if and only if $U_i \leq \frac{\sum_{j=1}^{i-1}(D_j-1)}{n-i}$. Therefore $\sigma = \inf(i \geq 1 : B_i = 1)$ stochastically dominates $M = \inf(i \geq 1 : A_i = 1)$. It follows that for $x \geq 0$,

$$\mathbf{P}\{|V| \geq x\} = \mathbf{P}\{M \geq x+1\} \leq \mathbf{P}\{\sigma \geq x+1\} \leq \mathbf{P}\{\sigma \geq x\}.$$

Then, by the first inequality of Proposition 3.6, we have that for all $\alpha \geq 17^{2/3}$,

$$\begin{aligned}\mathbf{P}\left\{|V| > \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} &\leq \mathbf{P}\left\{M > \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right\} \\ &\leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right).\end{aligned}$$

Finally, to prove the second inequality of Theorem 1.4, we apply the second inequality of Proposition 3.6 and get that for degree statistics with $\mathbf{n} = (\mathbf{n}(c), c \geq 0)$ with $k(\mathbf{n}) = 1$ and $\mathbf{n}(1) = 0$, for all $\ell \geq 1$,

$$\mathbf{P}\left\{|V| \geq \ell\right\} \leq \mathbf{P}\left\{\sigma \geq \ell + 1\right\} \leq \exp\left(-\frac{l^2}{2|\mathbf{n}|_1}\right). \qquad \square$$

The remainder of this chapter focuses on proving Propositions 3.5 and 3.6.

3.2.1 Proof of Proposition 3.5

Proposition 3.5 is a consequence of the following result. Let $(T, V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$ and $d(V_i) = d_T(V_i)$ for all $i \geq 0$. Fix $k \geq 0$. For $\mathbf{d} = (d_1, ..., d_k) \in \mathbb{N}^k$,

$$\mathbf{P}\left\{(d(V_0), ..., d(V_{k-1}) = \mathbf{d},\ |V| > k - 1\right\} = \mathbf{P}\left\{(D_1, ..., D_k) = \mathbf{d},\ M > k\right\}. \tag{3.6}$$

From this, we can easily see that $\mathbf{P}\left\{|V| > k\right\} = \mathbf{P}\left\{M > k + 1\right\}$ for $k \geq 0$,

$$\begin{aligned}
\mathbf{P}\left\{|V| > k - 1\right\} &= \sum_{\mathbf{d}=(d_1,...,d_k)} \mathbf{P}\left\{|V| > k - 1,\ (d(V_0), ..., d(V_{k-1})) = \mathbf{d}\right\} \\
&= \sum_{\mathbf{d}=(d_1,...,d_k)} \mathbf{P}\left\{M > k,\ (D_1, ..., D_k) = \mathbf{d}\right\} = \mathbf{P}\left\{M > k\right\}.
\end{aligned}$$

Thus $|V| \overset{\mathrm{d}}{=} M - 1$, Proposition 3.5 holds. We now prove (3.6) and begin by studying $\mathbf{P}\left\{(D_1, ..., D_k) = \mathbf{d},\ M > k\right\}$. Since $D = (D_1, ..., D_n)$ is a size biasing, (3.5) gives us that for $\mathbf{d} = (d_1, ..., d_k)$ satisfying $w(\mathbf{d}, c) \leq \mathbf{n}(c)$ for all $c \geq 0$,

$$\begin{aligned}
\mathbf{P}\left\{(D_1, ..., D_k) = \mathbf{d}\right\} &= \prod_{i=1}^{k} \mathbf{P}\left\{D_i = d_i \mid (D_1, ..., D_{i-1}) = (d_1, ..., d_{i-1})\right\} \\
&= \prod_{i=1}^{k} \frac{d_i(\mathbf{n}(d_i) - w((d_1, ..., d_{i-1}), d_i)}{n - 1 - d_1 - ... - d_{i-1}}.
\end{aligned} \tag{3.7}$$

Further, given $(D_1, ..., D_k) = \mathbf{d}$, we have that $> k$ if and only if for each $i \in [k]$,

$$U_i > \frac{1 + \sum_{j=1}^{i-1}(d_j - 1)}{n + 1 - i}.$$

Therefore,

$$\begin{aligned}\mathbf{P}\left\{M > k \mid (D_1, ..., D_k) = \mathbf{d}\right\} &= \prod_{i=1}^{k} \mathbf{P}\left\{U_i > \frac{1 + \sum_{j=1}^{i-1}(d_j - 1)}{n + 1 - i}\right\} \\ &= \prod_{i=1}^{k} \frac{n - 1 - d_1 - ... - d_{i-1}}{n + 1 - i},\end{aligned}$$

where the last equality holds since $n+1-i-(1+\sum_{j=1}^{i-1}(d_j-1)) = n-1-d_1-...-d_{i-1}$. We then have that

$$\mathbf{P}\left\{(D_1, ..., D_k) = \mathbf{d},\ M > k\right\} = \prod_{i=1}^{k} \frac{d_i(\mathbf{n}(d_i) - w((d_1, ..., d_{i-1}), d_i)}{n + 1 - i}. \tag{3.8}$$

Next, we study the probability $\mathbf{P}\left\{(d(V_0), .., d(V_{k-1})) = \mathbf{d},\ |V| > k - 1\right\}$. Below, we write $\sum_{\substack{(t,v) \\ d_t(v_{i-1}) = d_i,\, i \in [k]}}$ to mean that we are summing over marked trees (t, v) with non zero degrees $(d_t(v_i), 0 \leq i \leq k - 1)$ and satisfying $\mathbf{P}\left\{\mathrm{S}_k^{\bullet}(T, V) = (t, v)\right\} > 0$. There are precisely $d_1...d_k$ such marked trees. By Lemma 3.4, Propositions 2.11 and 2.12, we have that

$$\begin{aligned}&\mathbf{P}\left\{(d(V_0), .., d(V_{k-1})) = \mathbf{d},\ |V| > k - 1\right\} \\ &= \sum_{\substack{(t,v) \\ d_t(v_{i-1}) = d_i,\, i \in [k]}} \mathbf{P}\left\{\mathrm{S}_k^{\bullet}(T, V) = (t, v)\right\} \\ &= \sum_{\substack{(t,v) \\ d_t(v_{i-1}) = d_i,\, i \in [k]}} \left|\mathfrak{F}_{\mathbf{q}_t}^{(1)}\right| \left|\mathfrak{T}_{\mathbf{n}}^{(1)}\right|^{-1} \\ &= \sum_{\substack{(t,v) \\ d_t(v_{i-1}) = d_i,\, i \in [k]}} \binom{\|\mathbf{q}_t\|}{\mathbf{q}_t(c),\ c \geq 0} \binom{n}{\mathbf{n}(c),\ c \geq 0}^{-1},\end{aligned}$$

where $\mathbf{q}_t = (\mathbf{q}_t(c),\, c \geq 0)$, with $\mathbf{q}_t(0) = \mathbf{n}(0)$ and $\mathbf{q}_t(c) = \mathbf{n}(c) - \mathbf{n}_t(c)$ for $c \geq 1$. Note that $w(\mathbf{d}, 0) = 0$ since $\mathbf{d}$ are the degrees of the internal nodes of t. Further, we are summing over marked trees that satisfy $\sum_{c\geq 1} \mathbf{n}_t(c) = k$ and $\mathbf{n}_t(c) = w(\mathbf{d}, c)$ for $c \geq 1$. From this we have that $\|\mathbf{q}_t\| = n - k$, so

$$\begin{aligned}&\mathbf{P}\left\{(d(V_0), .., d(V_{k-1})) = \mathbf{d},\, |V| > k-1\right\}\\ &= d_1 ... d_k \frac{(n-k)!}{\prod_{c\geq 0}(\mathbf{n}(c) - w(\mathbf{d}, c))!} \frac{\prod_{c\geq 0} \mathbf{n}(c)!}{n!}.\end{aligned}$$

Suppose the following equality holds,

$$\prod_{i=1}^{k}(\mathbf{n}(d_i) - w((d_1, ..., d_{i-1}), d_i)) = \prod_{c\geq 0} \frac{\mathbf{n}(c)!}{(\mathbf{n}(c) - w(\mathbf{d}, c))!}, \tag{3.9}$$

then by (3.8) and since $(n-k)!/n! = \prod_{i=1}^{k} 1/(n+1-i)$,

$$\begin{aligned}\mathbf{P}\left\{(d(V_0), .., d(V_{k-1})) = \mathbf{d},\, |V| > k-1\right\} &= \prod_{i=1}^{k} \frac{d_i(\mathbf{n}(d_i) - w((d_1, ..., d_{i-1}), d_i)}{n+1-i}\\ &= \mathbf{P}\left\{(D_1, ..., D_k) = \mathbf{d},\, M > k\right\}.\end{aligned}$$

Therefore, we are left to prove (3.9). For $b \geq 0$, define $(i_j(b), j \in [w(\mathbf{d}, b)])$ as strictly increasing indices satisfying $d_{i_j(b)} = b$. Note that $w((d_1, ..., d_{i_j(b)-1}), b) = j - 1$ for $j \in [w(\mathbf{d}, b)]$. We can conclude that

$$\begin{aligned}\prod_{i=1}^{k}(\mathbf{n}(d_i) - w((d_1, ..., d_{i-1}), d_i)) &= \prod_{b\geq 0} \prod_{j=1}^{w(\mathbf{d},b)} (\mathbf{n}(b) - w((d_1, ..., d_{i_j(b)-1}), b))\\ &= \prod_{b\geq 0} \prod_{j=1}^{w(\mathbf{d},b)} (\mathbf{n}(b) - j + 1)\\ &= \prod_{b\geq 0} \frac{\mathbf{n}(b)!}{(\mathbf{n}(b) - w(\mathbf{d}, b))!}.\end{aligned}$$

3.2.2 Proof of Proposition 3.6.

For this proof, we construct a random variable $\tilde{\sigma}$ with the same distribution as σ and prove the desired probability bounds on $\tilde{\sigma}$. The key tool is to use a Poisson embedding as described in [7, Section 4.1]. We define a homogenous Poisson process $\mathbf{N}$ on $[0,\infty) \times [0,1]$ of rate 1 per unit area, with points $\{(S_j, U_j),\, j \geq 1\}$ where $0 < S_1 < S_2 < ....$ We treat the S_j's as arrival times for the points U_j. The $(S_j,\, j \geq 1)$ are a Poisson process on $[0,\infty)$ of rate 1 per unit length and the $(U_j,\, j \geq 1)$ are i.i.d. Uniform$[0,1]$ random variables, independent of $(S_j,\, j \geq 1)$. Let $(d_1, ..., d_n)$ be a sequence satisfying $\{1 \leq i \leq n : d_i = c\} = \mathbf{n}(c)$ for each $c \geq 0$. Let $l_1 := 0$ and for $1 \leq i \leq n$, let $l_{i+1} := l_i + d_i/(n-1)$ and $r_i = l_i + \max(0, (d_i - 1)/(n-1)) \leq l_{i+1}$. Define the intervals $I_i := [l_i, l_{i+1})$ and $I_i^- := [l_i, r_i)$. Note that the disjoint intervals $(I_i, i \in [n])$ partition the interval $[0,1)$. Further, $I_i^- \subseteq I_i$ for each $i \in [n]$.

Define, for $\ell \geq 1$, the index of the interval containing the point U_ℓ as $J(\ell)$, thus $U_\ell \in I_{J(\ell)}$. Now, define the sequence of indices $(M(\ell),\, \ell \geq 1)$ satisfying

$$M(\ell+1) = \inf\left(j > M(\ell) :\, U_j \notin \bigcup_{k=1}^{\ell} I_{J(M(k))}\right),$$

and $M(1) = 1$. Then $(M(\ell),\, \ell \geq 1)$ is the sequence of indices such that the points $(U_{M(\ell)},\, \ell \geq 1)$ fall into distinct intervals $(I_i,\, i \in [n])$. Now, let $(j(1), .., j(\ell))$ be a vector of distinct elements of $\{1, ..., n\}$ and let $(m(1), ..., m(\ell))$ be an increasing sequence of integers with $m(1) = 1$. Then

$$\begin{aligned}
&\mathbf{P}\left\{ J(m(\ell)) = j(l),\, M(\ell) = m(\ell) \mid J(m(k)) = j(k),\, M(k) = m(k),\, \forall\, k \in [\ell-1]\right\} \\
&= \mathbf{P}\left\{ U_{m(\ell)} \in I_{j(\ell)},\, U_j \in \bigcup_{k=1}^{\ell-1} I_{j(k)}\, \forall\, j \in \{m(\ell-1)+1, ..., m(\ell)-1\}\right\} \\
&= \frac{d_{j(\ell)}}{n-1} \cdot \left(\frac{\sum_{k=1}^{\ell-1} d_{j(k)}}{n-1}\right)^{m(\ell)-m(\ell-1)-1}.
\end{aligned}$$

Now, if $\sum_{k=1}^{\ell-1} d_{j(k)} < n-1$, we have that

$$\begin{aligned}
&\mathbf{P}\left\{ J(M(\ell)) = j(\ell) \mid J(m(k)) = j(k),\, M(k) = m(k),\, \forall\, k \in [\ell-1] \right\} \\
&= \sum_{a\geq 1} \mathbf{P}\left\{ J(M(\ell)) = j(\ell),\, M(\ell) = m(\ell-1)+a \mid J(m(k)) = j(k),\, M(k) = m(k),\, \forall\, k \in [\ell-1] \right\} \\
&= \sum_{a\geq 1} \frac{d_{j(\ell)}}{n-1} \cdot \left(\frac{\sum_{k=1}^{\ell-1} d_{j(k)}}{n-1} \right)^{a-1} = \frac{d_{j(\ell)}}{n-1-d_{j(1)}-\ldots-d_{j(\ell-1)}}.
\end{aligned}$$

Therefore,

$$\mathbf{P}\left\{ J(M(l)) = j(\ell) \mid J(M(k)) = j(k),\, \forall\, k \in [\ell-1] \right\} = \frac{d_{j(\ell)}}{n-1-d_{j(1)}-\ldots-d_{j(\ell-1)}}.$$

Define the sequence $(D(1), \ldots, D(n))$ by

$$D(\ell) = \begin{cases} d_{J(M(\ell))} & \text{if } 1 \leq \ell \leq n - \mathbf{n}(0) \\ 0 & \text{otherwise.} \end{cases}$$

Then, remark that

$$\begin{aligned}
&\mathbf{P}\left\{ D(\ell) = d \mid (D(1), \ldots, D(\ell-1)) = (d_1, \ldots, d_{\ell-1}) \right\} \\
&= \mathbf{P}\left\{ d_{J(M(\ell))} = d \;\middle|\; (d_{J(M(1))}, \ldots, d_{J(M(\ell-1))}) = (d_1, \ldots, d_{\ell-1}) \right\} \\
&= \sum_{i\in[n]} \mathbf{P}\left\{ J(M(\ell)) = i,\, d_i = d \mid (d_{J(M(1))}, \ldots, d_{J(M(\ell-1))}) = (d_1, \ldots, d_{\ell-1}) \right\}.
\end{aligned}$$

Recall $w((d_1, \ldots, d_{\ell-1}), d) = |\{i \in [\ell-1] : d_i = d\}|$. Suppose $(d_{J(M(1))}, \ldots, d_{J(M(\ell-1))}) = (d_1, \ldots, d_{\ell-1})$, then the number of indices $i \in [n] \setminus \{J(M(k),\, k \in [\ell-1]\}$ satisfying $d_i = d$ is given by $\mathbf{n}(d) - w((d_1, \ldots, d_{\ell-1}), d)$. Thus,

$$\mathbf{P}\left\{ D(\ell) = d \mid (D(1), \ldots, D(\ell-1)) = (d_1, \ldots, d_{\ell-1}) \right\} = \frac{d(\mathbf{n}(d) - w((d_1, \ldots, d_{\ell-1}), d)}{n-1-d_1-\ldots-d_{\ell-1}},$$

showing that the sequence $(D(1),...,D(n))$ is is a size-biasing of $\mathbf{n}$. Next, for $\ell \geq 1$, define $C_\ell = \bigcup_{k=1}^{\ell}[r_{J(k)}, l_{J(k)})$. Then, for each $k \leq l$, $U_k \in I_{J(k)}$ and C_ℓ contains the subinterval $[r_{J(k)}, l_{J(k)}) \subset I_{J(k)}$ of size $1/(n-1)$ subinterval of size $1/(n-1)$. Define

$$\tilde{\tau} = \inf\left(\ell \geq 1 : U_\ell \in \bigcup_{k=1}^{\ell-1} I_{J(k)} \setminus C_{\ell-1}\right) = \inf\left(\ell \geq 1 : U_\ell \in \bigcup_{k=1}^{\ell-1}[l_{J(k)}, r_{J(k)})\right).$$

Recall that the indices $(M(\ell), \ell \geq 1)$ are defined as the first indices of the sequence $(U_j,\, j \geq 1)$ that all fall into distinct intervals of $(I_i,\, i \in [n])$. Therefore, $\tilde{\tau} \notin \{M(\ell),\, \ell \geq 1\}$ and $C_{M(\ell)} = \cup_{k=1}^{\ell}[r_{J(M(k))}, l_{J(M(k))})$. For $U_{M(k)} \in I_{J(M(k))}$, we can think of $[r_{J(M(k))}, l_{J(M(k))}]$ as a censored subinterval of $I_{J(M(k))}$ of size $d_{J(M(k))}/(n-1)$. Then, $C_{M(\ell)}$ is a union of censored subintervals and $\tilde{\tau}$ is the first time a point U_ℓ falls into a previously seen interval in $\bigcup_{k=1}^{\ell-1} I_{J(k)}$ but not in the censored region $C_{\ell-1}$.

Now, fix $(j(1),...,j(\ell))$ a sequence of distinct elements of $[n]$ and $(m(1),...,m(\ell))$ an increasing sequence of integers with $m(1) = 1$. Assume that $J(m(k)) = j(k)$ and $M(k) = m(k)$ for all $k \in [\ell-1]$, and $\tilde{\tau} > m(l-1)$. Then $\tilde{\tau} \leq M(\ell)$ holds if and only if the first point among $(U_m,\, m > m(\ell-1))$ that doesn't fall into $C_{m(\ell-1)} = \cup_{k=1}^{\ell-1}[r_{j(k)}, l_{j(k)})$, falls into

$$\bigcup_{k=1}^{\ell-1} I_{J(M(k))} \setminus C_{m(\ell-1)} = \bigcup_{k=1}^{\ell-1} I_{j(k)} \setminus C_{m(\ell-1)} = \bigcup_{k=1}^{\ell-1}[l_{j(k)}, r_{j(k)}).$$

Therefore, for U a Uniform$[0,1]$ random variable,

$$\begin{aligned}
&\mathbf{P}\left\{\tilde{\tau} < M(\ell) \mid \tilde{\tau} > M(\ell-1),\, J(m(k)) = j(k),\, M(k) = m(k)\, \forall\, 1 \leq k \leq \ell-1\right\} \\
&= \mathbf{P}\left\{ U \in \bigcup_{k=1}^{\ell-1}[l_{j(k)}, r_{j(k)}) \;\middle|\; U \notin \bigcup_{k=1}^{\ell-1}[r_{j(k)}, l_{j(k)+1}) \right\} \\
&= \frac{\sum_{k=1}^{\ell-1}(d_{j(k)}-1)}{n-1}\left(1 - \frac{\ell-1}{n-1}\right)^{-1} = \frac{\sum_{k=1}^{\ell-1}(d_{j(k)}-1)}{n}.
\end{aligned}$$

Note that $\sum_{k=1}^{\ell-1}(d_{j(k)}-1)/n$ only depends on the values $(d_{j(1)},...,d_{j(\ell-1)})$ and that $D(k)=d_{J(M(k))}$ for all $k\geq 1$. Therefore,

$$\begin{aligned}
&\mathbf{P}\left\{\tilde{\tau}\leq M(\ell)\mid\ D(1),...,D(\ell-1),\ \tilde{\tau}>M(\ell-1)\right\}\\
&=\mathbf{P}\left\{\tilde{\tau}<M(\ell)\mid\ D(1),...,D(\ell-1),\ \tilde{\tau}>M(\ell-1)\right\}\\
&=\frac{\sum_{k=1}^{\ell-1}(D(k)-1)}{n-\ell}. \qquad (3.10)
\end{aligned}$$

Next, define $\tilde{\sigma}:=\sup(k\geq 1:\tilde{\tau}>M(k))$. Note that $\tilde{\sigma}\geq l$ if and only if $\tilde{\tau}>M(\ell)$. Thus,

$$\begin{aligned}
\mathbf{P}\left\{\tilde{\sigma}=\ell\mid\ D(1),...,D(\ell-1),\ \tilde{\sigma}>\ell-1\right\}&=\mathbf{P}\left\{\tilde{\tau}\leq M(\ell)\mid\ D(1),...,D(\ell-1),\ \tilde{\tau}>M(\ell-1)\right\}\\
&=\frac{\sum_{k=1}^{\ell-1}(D(k)-1)}{n-\ell}.
\end{aligned}$$

Recall in the statement of Proposition 3.6, we defined $\sigma:=\inf(i\geq 1:B_i=1)$, with

$$B_i:=\begin{cases}1 & \text{if } U_i\leq\frac{\sum_{j=1}^{i-1}(D_j-1)}{n-i},\\ 0 & \text{otherwise,}\end{cases}$$

where $(U_j,\ j\geq 1)$ are i.i.d. Uniform$[0,1]$ random variables and $(D_1,...,D_n)$ is a size-biasing of $\mathbf{n}$. Therefore, for a Uniform$[0,1]$ random variable U,

$$\begin{aligned}
&\mathbf{P}\left\{\sigma=\ell\mid\ (D_1,...,D_{\ell-1}),\ \sigma>\ell-1\right\}\\
&=\mathbf{P}\left\{U\leq\frac{\sum_{j=1}^{\ell-1}(D_j-1)}{n-\ell}\ \middle|\ (D_1,...,D_{\ell-1}),\ \sigma>\ell-1\right\}\\
&=\frac{\sum_{j=1}^{\ell-1}(D_j-1)}{n-\ell}.
\end{aligned}$$

We have that $\sigma \overset{\mathrm{d}}{=} \tilde{\sigma}$ and therefore,

$$\mathbf{P}\{\sigma \geq \ell\} = \mathbf{P}\{\tilde{\sigma} \geq \ell\} = \mathbf{P}\{\tilde{\tau} > M(\ell)\} \leq \mathbf{P}\{\tilde{\tau} > \ell\}.$$

From this, we can deduce the second bound of Proposition 3.6. Note that if $\mathbf{n}(1) = 0$, then for all $d_i \neq 0$ we have that $d_i \geq 2$ and therefore, $|[l_i, r_i)| = (d_i - 1)/(n-1) \geq 1/(n-1)$. Further, if $\tilde{\tau} > M(\ell - 1)$, then $D(k) \geq 2$ for all $1 \leq k \leq \ell - 1$. Then, by (3.10), for all $\ell \geq 2$,

$$\mathbf{P}\{\tilde{\tau} \leq M(\ell) \mid \tilde{\tau} > M(\ell - 1)\} \geq \frac{\ell - 1}{n - \ell}.$$

From the above inequality and the fact that $\mathbf{P}\{\tilde{\tau} > M(1)\} = 1$ as $M(1) = 1$, it follows that

$$\begin{aligned}
\mathbf{P}\{\sigma \geq \ell\} = \mathbf{P}\{\tilde{\tau} > M(\ell)\} &= \prod_{k=2}^{\ell} \mathbf{P}\{\tilde{\tau} > M(k) \mid \tilde{\tau} > M(k-1)\} \\
&= \prod_{k=2}^{\ell} (1 - \mathbf{P}\{\tilde{\tau} \leq M(k) \mid \tilde{\tau} > M(k-1)\}) \\
&\leq \prod_{k=2}^{\ell} \left(1 - \frac{k-1}{n-k}\right) \leq \prod_{k=1}^{\ell-1} \left(1 - \frac{k}{n-k}\right).
\end{aligned}$$

We show by induction that for all $\ell \geq 1$,

$$\mathbf{P}\{\sigma \geq \ell\} \leq \exp\left(-\frac{(\ell-1)^2}{2(n-1)}\right),$$

which is precisely the second inequality in Proposition 3.6. The base case $\ell = 1$ is trivial. By the induction hypothesis and the fact that $1 - x \leq e^{-x}$, we have that

$$\begin{aligned}
\prod_{k=1}^{\ell} \left(1 - \frac{k}{n-k}\right) &= \left(1 - \frac{\ell}{n-\ell}\right) \prod_{k=1}^{\ell-1} \left(1 - \frac{k}{n-k}\right) \\
&\leq \left(1 - \frac{\ell}{n-\ell}\right) \exp\left(-\frac{(\ell-1)^2}{2(n-1)}\right) \leq \exp\left(-\left(\frac{\ell}{n-\ell} + \frac{(\ell-1)^2}{2(n-1)}\right)\right).
\end{aligned}$$

We can conclude the proof by noting that, since $\ell/(n-\ell) \geq \ell/(n-1)$,

$$\frac{\ell}{n-\ell} + \frac{(\ell-1)^2}{2(n-1)} \geq \frac{\ell^2+1}{2(n-1)} \geq \frac{\ell^2}{2(n-1)}.$$

Next, we prove the first inequality of Proposition 3.6 by using the following proposition.

Proposition 3.7. *Let* $\mathrm{v} = \sum_{i:d_i \geq 2} d_i^2/(n-1)$. *Then for all* $\alpha \geq 17^{3/2}$,

$$\mathbf{P}\left\{\tilde{\tau} > \alpha \frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right\} \leq \exp\left(-\frac{1}{3}\left(\frac{\alpha^{2/3}(n-1)}{\mathrm{v}}\right)^{1/2}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right).$$

The first inequality of Proposition 3.6 follows directly from the above proposition since $\mathbf{P}\{\sigma \geq x\} \leq \mathbf{P}\{\tilde{\tau} > x\}$ for $x \geq 1$, $\sum_{i:d_i \geq 2} d_i^2 = |\mathbf{n}|_2^2 - \mathbf{n}(1)$ and $|\mathbf{n}|_1 = n-1$, therefore $\frac{n-1}{\mathrm{v}} = \frac{|\mathbf{n}|_1^2}{|\mathbf{n}|_2^2 - \mathbf{n}(1)}$. To prove Proposition 3.7 we will use the earlier defined Poisson process $\mathbf{N}$. For $t \geq 0$, let $\mathbf{N}(t) := |\mathbf{N} \cap \{[0,t] \times [0,1]\}|$ be the number of points in the Poisson process $\mathbf{N}$ seen up to time t. Further, for $1 \leq i \leq n$, let $\mathbf{N}_i(t) := |\mathbf{N} \cap \{[0,t] \times [l_i, r_i)\}|$ be the number of points seen in the interval $[l_i, r_i)$ up to time t. Remark that if for some $1 \leq i \leq n$ we have that $\mathbf{N}_i(t) \geq 2$, then by construction $\tilde{\tau} \leq t$. Define the stopping time

$$T := \inf\left(t \geq 0 : \max_{i \in [n]} \mathbf{N}_i(t) \geq 2\right).$$

Note that $\tilde{\tau} \leq \mathbf{N}(T)$. Further, for any $h \in \mathbb{N}$, if $\mathbf{N}(t) \leq h$ and $T \leq t$ then $\tilde{\tau} \leq h$, hence

$$\mathbf{P}\{\tilde{\tau} > h\} \leq \inf_{t \geq 0} \left(\mathbf{P}\{\mathbf{N}(t) > h\} + \mathbf{P}\{T > t\}\right). \tag{3.11}$$

To bound $\mathbf{P}\{\mathbf{N}(t) > h\}$ we use the following Poisson tail bound. For $h \geq t$,

$$\mathbf{P}\{Poisson(t) > h\} \leq \exp\left(-t\left(\frac{h}{t}\log\left(\frac{h}{t}\right) - \frac{h}{t} + 1\right)\right). \tag{3.12}$$

The proof of (3.12) can be found in [6, Page 23]. Next we proceed to bound $\mathbf{P}\{T > t\}$. Remark that $(N_i(t),\ 1 \leq i \leq n)$ are independent random variables and $N_i(t)$ is Poisson$(t(r_i - l_i))$-distributed. Thus,

$$\mathbf{P}\{T > t\} = \mathbf{P}\{\mathbf{N}(t) < \tilde{\tau}\} = \prod_{\substack{i \in [n] \\ d_i \geq 2}} \mathbf{P}\{N_i(t) \leq 1\} = \prod_{i : d_i \geq 2} (1 + t(r_i - l_i))e^{-t(r_i - l_i)}.$$

Define $\mathbf{p} = (p_i, 1 \leq i \leq n)$, where $p_i := d_i/2(n-1)$. Recall that when $d_i \leq 1$ we have $r_i - l_i = 0$ and for $d_i \geq 2$, $r_i - l_i = (d_i - 1)/(n-1) \geq p_i$. Since the function $f(x) = (1+x)e^{-x}$ is decreasing for $x \geq 0$, we have that

$$\mathbf{P}\{T > t\} \leq \prod_{i : d_i \geq 2} (1 + p_i t)e^{-p_i t}. \tag{3.13}$$

Next, we bound the RHS using the following Lemma, based on Lemma 9 from [7].

Lemma 3.8. *Let* $\mathbf{d} = (d_1, ..., d_n)$ *and let* $g(t, \mathbf{d}) = \prod_{i : d_i \geq 2}(1 + p_i t)e^{-p_i t}$*, with* $p_i = d_i/2(n-1)$*. Define* $p_{\max} = \max_{i \in [n]} p_i$ *and* $d_{\max} = \max_{i \in [n]} d_i$*. For all* $0 \leq t < 1/p_{\max} = 2(n-1)/d_{\max}$*,*

$$\log g(t, \mathbf{d}) = \sum_{k \geq 2} \frac{(-1)^{k+1}}{k} \sum_{i : d_i \geq 2} \left(\frac{d_i t}{2(n-1)} \right)^k.$$

Further,

$$\left| \log g(t, \mathbf{d}) + \sum_{i : d_i \geq 2} \frac{(d_i t)^2}{8(n-1)^2} \right| \leq \frac{d_{\max} t}{6(n-1) - 3 d_{\max} t} \cdot \sum_{i : d_i \geq 2} \frac{(d_i t)^2}{4(n-1)^2}.$$

Proof. Recall that the expansion of $\log(1+x)$ around 0 for $|x| < 1$ is given by

$$\log(1+x) = \sum_{k \geq 1} (-1)^{k+1} \frac{x^k}{k}. \tag{3.14}$$

Note that for $0 \le t < 1/p_{\max}$,

$$\sum_{k \ge 2} \frac{1}{k} \sum_{i:d_i \ge 2} (p_i t)^k \le n \sum_{k \ge 2} \frac{(p_{\max} t)^k}{k} < \infty.$$

For $0 \le p_{\max} t < 1$, by (3.14) and Tonelli's theorem, we have that

$$\begin{aligned}\log(g(t, \mathbf{d})) &= \log\left(\prod_{i:d_i \ge 2} (1 + p_i t) e^{-p_i t}\right) = \sum_{i:d_i \ge 2} (\log(1 + p_i t) - p_i t) \\ &= \sum_{i:d_i \ge 2} \sum_{k \ge 1} \frac{(-1)^{k+1}}{k} (p_i t)^k - \sum_{i:d_i \ge 2} p_i t \\ &= \sum_{k \ge 2} \frac{(-1)^{k+1}}{k} \sum_{i:d_i \ge 2} (p_i t)^k.\end{aligned}$$

This proves the equality in Lemma 3.8. Now, note that for $k \ge 2$,

$$\sum_{i:d_i \ge 2} p_i^k \le p_{\max}^{k-2} \sum_{i:d_i \ge 2} p_i^2. \tag{3.15}$$

Therefore,

$$\sum_{k \ge 3} \sum_{i:d_i \ge 2} \frac{1}{k} (p_i t)^k \le \left(\sum_{i:d_i \ge 2} (p_i t)^2\right) \sum_{k \ge 3} \frac{(p_{\max} t)^{k-2}}{k} \le \left(\sum_{i:d_i \ge 2} (p_i t)^2\right) \frac{p_{\max} t}{3(1 - p_{\max} t)}.$$

Since

$$\log g(t, \mathbf{d}) = \sum_{k \ge 3} \frac{(-1)^{k+1}}{k} \sum_{i:d_i \ge 2} (p_i t)^k - \frac{1}{2} \sum_{i:d_i \ge 2} (p_i t)^2,$$

we can then conclude that

$$\left| \log g(t, \mathbf{d}) + \sum_{i:d_i \ge 2} \frac{(p_i t)^2}{2} \right| \le \sum_{k \ge 3} \sum_{i:d_i \ge 2} \frac{1}{k} (p_i t)^k \le \left(\sum_{i:d_i \ge 2} (p_i t)^2\right) \frac{p_{\max} t}{3(1 - p_{\max} t)}.$$

Replacing $p_i = d_i / 2(n-1)$ and $p_{\max} = d_{\max} / 2(n-1)$ yields the desired inequality.

□

Corollary 3.9. *Let* $\mathrm{v} = \sum_{i:d_i \geq 2} d_i^2/(n-1)$. *For all* $0 \leq t \leq (n-1)/d_{\max}$,

$$g(t, \mathbf{d}) \leq \exp\left(\frac{-\mathrm{v}t^2}{24(n-1)}\right).$$

Proof. Note that $(d_i t)^2/(n-1)^2 = \mathrm{v}t^2/(n-1)$, and for $t \leq (n-1)/d_{\max}$,

$$\frac{d_{\max} t}{6(n-1) - 3d_{\max} t} \leq \frac{n-1}{3(n-1)} = \frac{1}{3}.$$

Thus, by inequality in Lemma 3.8,

$$\begin{aligned}\log g(t, \mathbf{d}) &\leq \frac{d_{\max} t}{6(n-1) - 3d_{\max} t} \cdot \sum_{i:d_i \geq 2} \frac{(d_i t)^2}{4(n-1)^2} - \sum_{i:d_i \geq 2} \frac{(d_i t)^2}{8(n-1)^2} \\ &\leq \frac{\mathrm{v}t^2}{12(n-1)} - \frac{\mathrm{v}t^2}{8(n-1)} = -\frac{\mathrm{v}t^2}{24(n-1)}. \quad \square\end{aligned}$$

Proof of Proposition 3.7. We begin by noting that when $d_{\max} = 1$, then $\mathrm{v} = 0$ and $\tilde{\tau} \overset{\text{a.s.}}{=} \infty$. Thus, $\mathbf{P}\{\tilde{\tau} > \infty\} = 0$ and the non-negative upper bound in Proposition 3.7 holds. For the rest of the proof we assume that $d_{\max} \geq 2$. Let $\mathrm{v} = \sum_{i:d_i \geq 2} d_i^2/(n-1)$. By (3.11) and (3.13), we have that

$$\mathbf{P}\{\tilde{\tau} > h\} \leq \inf_{t \geq 0} \left(\mathbf{P}\{\mathbf{N}(t) > h\} + \mathbf{P}\{T > t\}\right) \leq \inf_{t \geq 0} \left(\mathbf{P}\{\mathbf{N}(t) > h\} + g(t, \mathbf{d})\right).$$

Further, by (3.12) and Corollary 3.9,

$$\mathbf{P}\{\tilde{\tau} > h\} \leq \inf_{t \geq 0} \left(\exp\left(-t\left(\frac{h}{t}\log\left(\frac{h}{t}\right) - \frac{h}{t} + 1\right)\right) + \exp\left(\frac{-\mathrm{v}t^2}{24(n-1)}\right)\right).$$

Fix $C \geq 2$. If $d_{\max} \leq \left(\sum_{i:d_i \geq 2} d_i^2\right)^{1/2} C^{-1} = \frac{((n-1)\mathrm{v})^{1/2}}{C}$, then for $t = \frac{C(n-1)^{1/2}}{\mathrm{v}^{1/2}}$,

$$\exp\left(\frac{-\mathrm{v}t^2}{24(n-1)}\right) = \exp\left(\frac{-C^2}{24}\right).$$

Further, for $h = 2t = \frac{2C(n-1)^{1/2}}{\mathrm{v}^{1/2}}$, we note that

$$t\left(\frac{h}{t}\log\left(\frac{h}{t}\right) - \frac{h}{t} + 1\right) = t\left(2\log(2) - 1\right) > \frac{t}{3}.$$

Therefore, we have that for fixed $C \geq 2$ and $d_{\max} \leq \left(\sum_{i:d_i\geq 2} d_i^2\right)^{1/2} C^{-1} = \frac{((n-1)\mathrm{v})^{1/2}}{C}$,

$$\mathbf{P}\left\{\tilde{\tau} > \frac{2C(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right\} \leq \exp\left(-\frac{C}{3}\frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right) + \exp\left(\frac{-C^2}{24}\right). \tag{3.16}$$

Next, suppose that $d_{\max} > \left(\sum_{i:d_i\geq 2} d_i^2\right)^{1/2} C^{-1} = \frac{((n-1)\mathrm{v})^{1/2}}{C}$. Let $m = \arg\max_{i\in[n]} d_i$, so $d_m = d_{\max} \geq 2$. Recall that if for some $K > 0$ and $i \in [n]$ we have that $N_i(K) \geq 2$, then $\tilde{\tau} \leq K$. Therefore, for $K > 0$,

$$\begin{aligned}
\mathbf{P}\{\tilde{\tau} > K\} &\leq \mathbf{P}\{\tilde{\tau} > \lfloor K \rfloor\} \leq \mathbf{P}\{\forall\, i \in [n],\ N_i(\lfloor K \rfloor) \leq 1\} \\
&\leq \mathbf{P}\{N_m(\lfloor K \rfloor) \leq 1\} = \mathbf{P}\{|\forall\, k \in [\lfloor K \rfloor] :\ U_k \in [l_m, r_m)| \leq 1\} \\
&= \mathbf{P}\{\mathrm{Bin}(\lfloor K \rfloor, r_m - l_m) \in \{0,1\}\} \\
&= (1 - (r_m - l_m))^{\lfloor K \rfloor - 1}(1 - (r_m - l_m) + \lfloor K \rfloor (r_m - l_m)).
\end{aligned}$$

Now since $d_{\max} \geq 2$, we have that $r_m - l_m = (d_{\max} - 1)/(n-1) \geq d_{\max}/2(n-1)$. Further, $\frac{(n-1)^{1/2}}{v^{1/2}} = (n-1)\left(\sum_{i:d_i\geq 2} d_i^2\right)^{-1/2} \geq 1$. Thus, for $K \geq 2C\frac{(n-1)^{1/2}}{v^{1/2}} \geq 4$, using the fact that $1 - x \leq e^{-x}$ and $\lfloor K \rfloor - 1 > K/2$, it follows that

$$\begin{aligned}
\mathbf{P}\{\tilde{\tau} > K\} &\leq \left(1 - \frac{d_{\max}}{2(n-1)}\right)^{\lfloor K \rfloor - 1}\left(1 + (\lfloor K \rfloor - 1)\frac{d_{\max}}{n-1}\right) \\
&\leq \exp\left(-K\frac{d_{\max}}{4(n-1)}\right)\left(1 + (\lfloor K \rfloor - 1)\frac{d_{\max}}{n-1}\right).
\end{aligned}$$

Next, using that $\frac{(\lfloor K \rfloor - 1)d_{\max}}{n-1} > \frac{K d_{\max}}{2(n-1)} \geq 2$, we have that $1 + \frac{(\lfloor K \rfloor - 1)d_{\max}}{n-1} < \frac{2K d_{\max}}{n-1}$. Thus,

$$\mathbf{P}\{\tilde{\tau} > K\} \leq 2K\frac{d_{\max}}{n-1}\exp\left(-K\frac{d_{\max}}{4(n-1)}\right)$$

Now, let $K = \frac{xC(n-1)^{1/2}}{v^{1/2}}$ where $x \geq 4$. We suppose $d_{\max} > \frac{((n-1)\mathrm{v})^{1/2}}{C}$, thus $Kd_{\max}/(n-1) \geq x$. The function $2xe^{-x/4}$ is decreasing for $x \geq 4$, thus

$$\mathbf{P}\left\{\tilde{\tau} > xC\frac{(n-1)^{1/2}}{v^{1/2}}\right\} \leq 2xe^{-x/4}. \tag{3.17}$$

Finally, we combine the inequalities (3.16) and (3.17) such that we get a bound independent of the value of $d_{\max}$. Fix $\alpha \geq 17^{3/2}$ and let $x = \alpha^{2/3} \geq 4$ and $C = \alpha^{1/3} > 2$. Then, $2C \leq \alpha$ and $xC = \alpha$. In any case, one bound of (3.16) or (3.17) applies, thus,

$$\begin{aligned}
\mathbf{P}\left\{\tilde{\tau} > \alpha\frac{(n-1)^{1/2}}{v^{1/2}}\right\} &= \mathbf{P}\left\{\tilde{\tau} > xC\frac{(n-1)^{1/2}}{v^{1/2}}\right\} \\
&\leq \exp\left(-\frac{C}{3}\frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right) + \exp\left(-\frac{C^2}{24}\right) + 2xe^{-x/4} \\
&= \exp\left(-\frac{\alpha^{1/3}}{3}\frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right) + \exp\left(-\frac{\alpha^{2/3}}{24}\right) + 2\alpha^{2/3}\exp\left(-\frac{\alpha^{2/3}}{4}\right).
\end{aligned}$$

Remark that for $y \geq 17$, we have that $e^{-y/24} + 2ye^{-y/4} \leq 2e^{-y/24}$. We can thus conclude that

$$\mathbf{P}\left\{\tilde{\tau} > \alpha\frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{(n-1)^{1/2}}{\mathrm{v}^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right). \square$$

3.3 Proof of Theorem 1.2

Below we briefly outline the proof of Theorem 1.2 using Theorem 1.4. Recall Theorem 1.2.

Theorem 1.2. *[4, Theorem 1.2.] Fix a probability distribution μ supported by $\mathbb{N}$ with $|\mu|_1 \leq 1$ and $|\mu|_2 = \infty$. For $n \in \mathbb{N}$, let T_n be a Bienaymé tree with offspring distribution μ conditioned to have size n, and let V_n be a uniformly random node in T_n. Then*

$$\mathit{wid}(T_n)/\sqrt{n} \to \infty,\ |V_n|/\sqrt{n} \to 0 \text{ and } \mathit{ht}(T_n)/(\sqrt{n}\log^3 n) \to 0,$$

where the convergence results hold both in probability and in expectation, as $n \to \infty$.

The main fact we use is that Bienaymé trees conditioned to have a fixed size and degree statistics, are uniformly random trees with those degree statistics. Therefore, for degree statistics $\mathbf{n}$ satisfying $\mathbf{P}\{\mathbf{n}_{T_n} = \mathbf{n}\} > 0$, by conditioning that $\mathbf{n}_{T_n} = \mathbf{n}$, we can apply Theorem 1.4 to V_n. That is, for $\alpha \geq 17^{3/2}$, we have

$$\mathbf{P}\left\{|V_n| \geq \alpha \frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}} \;\middle|\; \mathbf{n}_{T_n} = \mathbf{n}\right\} \leq \exp\left(-\frac{\alpha^{1/3}}{3}\frac{|\mathbf{n}|_1}{(|\mathbf{n}|_2^2 - \mathbf{n}(1))^{1/2}}\right) + 2\exp\left(-\frac{\alpha^{2/3}}{24}\right). \tag{3.18}$$

We further use the following proposition from [4, Proposition 2.3].

Proposition 3.10. *Fix a probability distribution μ supported by $\mathbb{N}$ with $|\mu|_1 \leq 1$ and $|\mu|_2 = \infty$. For $n \in \mathbb{N}$, let T_n be a Bienaymé tree with offspring distribution μ conditioned to have size n, and let $\mathbf{n}_{T_n}$ be the degree statistics of T_n. Then for any $C > 0$, there exists $c = c(C) > 0$ such that $\mathbf{P}\{|\mathbf{n}_{T_n}|_2^2 < C|\mathbf{n}_{T_n}|_1\} < e^{-cn}$ for all n sufficiently large.*

Fix $C > 0$, then by Proposition 3.10, for $h \geq 0$,

$$\begin{aligned}\mathbf{P}\{|V_n| \geq h\} &= \mathbf{P}\left\{|V_n| \geq h,\ |\mathbf{n}_{T_n}|_2^2 < C|\mathbf{n}_{T_n}|_1\right\} + \mathbf{P}\left\{|V_n| \geq h,\ |\mathbf{n}_{T_n}|_2^2 \geq C|\mathbf{n}_{T_n}|_1\right\} \\ &\leq \mathbf{P}\left\{|\mathbf{n}_{T_n}|_2^2 < C|\mathbf{n}_{T_n}|_1\right\} + \sup_{\mathbf{n}:|\mathbf{n}|_2^2 \geq C|\mathbf{n}|_1} \mathbf{P}\{|V_n| \geq h \mid \mathbf{n}_{T_n} = \mathbf{n}\} \\ &\leq e^{-cn} + \sup_{\mathbf{n}:|\mathbf{n}|_2^2 \geq C|\mathbf{n}|_1} \mathbf{P}\{|V_n| \geq h \mid \mathbf{n}_{T_n} = \mathbf{n}\}\end{aligned}$$

where $c > 0$ is a constant depending on C. Let $\varepsilon > 0$. By choosing an appropriate $C = C(\varepsilon) = 1 + \varepsilon^4$, we obtain that

$$\mathbf{P}\left\{|V_n| \geq 6^3\varepsilon^2(n-1)^{1/2}\log^3 n\right\} = O\left(\frac{1}{n^2}\right), \tag{3.19}$$

we omit the computational details. Through similar reasoning and a few more computations, it can be shown that $\mathbf{E}[|V_n|] = o(n^{1/2})$. The convergence results on the height of T_n rely on the following inequality. If we suppose that the height of T_n is greater than some positive integer h, then T_n must contain at least one node with height at least h. Then, since V_n is a uniformly random node of T_n, $\mathbf{P}\{|V_n| \geq h \mid \text{ht}(T_n) \geq h\} \geq 1/n$. Using the fact that $|V_n| \geq h$ implies $\text{ht}(T_n) \geq h$, we then have

$$\mathbf{P}\{\text{ht}(T_n) \geq h\} = \frac{\mathbf{P}\{\text{ht}(T_n) \geq h\}\mathbf{P}\{|V_n| \geq h\}}{\mathbf{P}\{|V_n| \geq h, \text{ht}(T_n) \geq h\}} \leq n\mathbf{P}\{|V_n| \geq h\}.$$

By combining (3.19) with the above inequality, it follows that

$$\mathbf{P}\left\{\text{ht}(T_n) \geq 6^3\varepsilon^2(n-1)^{1/2}\log^3 n\right\} = O\left(\frac{1}{n}\right). \tag{3.20}$$

Since this holds for any $\varepsilon > 0$, we get that $\text{ht}(T_n)/((n-1)^{1/2}\log^3 n) \to 0$ in probability. To show convergence in expectation, we note that $\text{ht}(T_n) \leq n-1$ and so for any positive real h,

$$\mathbf{E}[\text{ht}(T_n)] = \mathbf{E}\left[\text{ht}(T_n)\mathbf{1}_{[\text{ht}(T_n)<h]}\right] + \mathbf{E}\left[\text{ht}(T_n)\mathbf{1}_{[\text{ht}(T_n)\geq h]}\right] \leq h + (n-1)\mathbf{P}\{\text{ht}(T_n) \geq h\}.$$

Applying this inequality to $h = 6^3\varepsilon^2(n-1)^{1/2}\log^3 n$, combining it with (3.20) and the fact that (3.20) holds for any $\varepsilon > 0$, we have that $\text{ht}(T_n)/((n-1)^{1/2}\log^3 n) \to 0$ in expectation.

Lastly, we prove the convergence results on the width of T_n. Since V_n is a uniformly random node of T_n, if for some positive integer h, at least half of the nodes of T_n have height h, then $|V_n| \geq h$ with probability at least $1/2$. Now, fix $\varepsilon > 0$ and let n be sufficiently large such that $\mathbf{E}[|V_n|] \leq \varepsilon^2 n^{1/2}$. Then by the above and by applying Markov's inequality,

$$\frac{1}{2}\mathbf{P}\left\{|\{u \in T_n : |u| \geq \varepsilon n^{1/2}\}| \geq \frac{n}{2}\right\} \leq \mathbf{P}\left\{|V_n| \geq \varepsilon n^{1/2}\right\} \leq \frac{\mathbf{E}[|V_n|]}{\varepsilon n^{1/2}} \leq \varepsilon. \tag{3.21}$$

Now, assume that at most half of the nodes of T_n have height greater that $\varepsilon n^{1/2}$. Then there are over $n/2$ nodes within the first $\varepsilon n^{1/2}$ levels of T_n. In this case, the width of T_n is at least $\mathrm{wid}(T_n) \geq (n/2)/(\varepsilon n^{1/2}) = n^{1/2}/(2\varepsilon)$. From this and (3.21), it follows that

$$\mathbf{P}\left\{\mathrm{wid}(T_n) \geq \frac{n^{1/2}}{(2\varepsilon)}\right\} \geq \mathbf{P}\left\{|\{u \in T_n : |u| \geq \varepsilon n^{1/2}\}| < \frac{n}{2}\right\} \geq 1 - 2\varepsilon.$$

Since this holds for arbitrary $\varepsilon > 0$, we can conclude that $\mathrm{wid}(T_n)/n^{1/2} \to \infty$ in probability and in expectation. This concludes the proof of Theorem 1.2.

Chapter 4

CONCLUSION

In this chapter, we briefly describe open problems and future research directions. We begin by stating a possible generalisation of Theorem 1.5. Let $\mathbf{n}, \mathbf{m}$ be degree statistics with $k(\mathbf{n}) = k(\mathbf{m}) = 1$. We define the partial order $\mathbf{n} \preceq \mathbf{m}$ by the following covering relation. We have $\mathbf{n} \preceq \mathbf{m}$ if there exists $a, b \geq 1$ such that for all $c \geq 0$,

$$\mathbf{m}(c) = \mathbf{n}(c) - \mathbf{1}_{[c=a+b-1]} + \mathbf{1}_{[c=a]} + \mathbf{1}_{[c=b]}. \tag{4.1}$$

A possible extension of Theorem 1.5 is described in the following problem.

Problem 4.1. *Let $\mathbf{n}, \mathbf{m}$ be degree statistics with $k(\mathbf{n}) = k(\mathbf{m}) = 1$ and such that there exists $c_1, c_2 > 2$ such that $\mathbf{n}(c_1) > 0$ and $\mathbf{n}(c_2) > 0$. Let $(T, V) \in_u \mathfrak{T}_{\mathbf{n}}^{(1)}$ and $(T', W) \in_u \mathfrak{T}_{\mathbf{m}}^{(1)}$. Does $\mathbf{n} \preceq \mathbf{m}$ imply $|V| \preceq_{st} |W|$?*

To see that this statement is in fact a generalisation of Theorem 1.5, it suffices to show that $\mathbf{n} \preceq \mathbf{m}$, where $\mathbf{n}$ and $\mathbf{m}$ are defined as in the statement of Theorem 1.5. We can prove $\mathbf{n} \preceq \mathbf{m}$ by induction on the maximal degree $d_{\max} = \max(c \geq 0 :$

$\mathbf{n}(c) > 0)$ of $\mathbf{n}$. Define $\mathbf{n}' = (\mathbf{n}'(c),\ c \geq 0)$ such that

$$\mathbf{n}'(c) = \begin{cases} \mathbf{n}(2) + 1 & \text{if } c = 2 \\ \mathbf{n}(d_{\max} - 1) + 1 & \text{if } c = d_{\max} - 1 \\ \mathbf{n}(d_{\max}) - 1 & \text{if } c = d_{\max} \\ \mathbf{n}(c) & \text{otherwise.} \end{cases}$$

It can be easily verified that $\mathbf{n} \preceq \mathbf{n}'$, by taking $a = d_{\max} - 1$ and $b = 2$. Then, by the induction hypothesis, it follows that $\mathbf{n} \preceq \mathbf{n}' \preceq \mathbf{m}$.

Next, we mention future work around Theorem 1.2. Recall the conjectures and problems stated by Janson on simply generated trees [11, Conjectures 21.5, 21.6 and Problems 21.7, 21.8].

Conjecture 1. *Let* $\mathbf{w} = (\mathbf{w}_k, k \geq 0)$ *be a weight sequence with* $\mathbf{w}_0 > 0$ *and* $\mathbf{w}_k > 0$ *for some* $k \geq 2$. *Let* $\mathcal{T}_n$ *be a simply generated tree of size* n *with weight sequence* $\mathbf{w}$, *whenever* $n \geq 0$ *satisfies* $Z_n(\mathbf{w}) > 0$.

(1) If $\nu = 1$ *and* $\sigma^2 = \infty$ *then* $ht(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} 0$.

(2) If $\nu = 1$ *and* $\sigma^2 = \infty$ *then* $wid(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} \infty$.

(3) If $\nu < 1$ *then* $ht(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} 0$.

(4) If $\nu < 1$ *then* $wid(\mathcal{T}_n)/\sqrt{n} \xrightarrow{p} \infty$.

In the joint project with Addario-Berry, Brandenberger and Hamdan [4], we describe how Theorem 1.4 can be applied to prove (2) and (4) and to prove a modification of (1) and (3), with an additional $\log^3 n$ term. It would be desirable to see whether the convergence results on the height of $|V_n|$ in Theorem 1.2 can be stated without the additional $\log^3 n$ factor, similarly for our modified version of (3). Janson suggests that this would be possible in [11].

To conclude, the above work describes new universal bounds on a random node of a random tree and further bounds on the height of conditioned Bienaymé trees. Our method for proving such bounds is to define a sampling procedure that generates a random variable with the same law as the height of a random node in a random tree with fixed degree statistics. We further adapt a Poissonization trick from Camarri and Pitman [7] to trees with fixed degree statistics. While it is notable that our bounds solve and almost solve conjectures from Janson [11, Conjectures 21.5, 21.6 and Problems 21.7, 21.8], it would be desirable to further improve our bounds on the heights of random trees.

Bibliography

[1] 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. doi: https://doi.org/10.1002/rsa.20438. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/rsa.20438. 10, 11

[2] L. Addario-Berry. Most trees are short and fat. *Probability Theory and Related Fields*, 173(1):1–26, 2019. doi: 10.1007/s00440-018-0829-x. URL https://doi.org/10.1007/s00440-018-0829-x. i, ii, 5

[3] L. Addario-Berry, L. Devroye, and S. Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees. *The Annals of Probability*, 41(2):1072 – 1087, 2013. doi: 10.1214/12-AOP758. URL https://doi.org/10.1214/12-AOP758. i, ii, 5, 6, 9, 10, 11, 13

[4] L. Addario-Berry, A. Brandenberger, J. Hamdan, and C. Kerriou. Universal height and width bounds for random trees, 2021. URL https://arxiv.org/abs/2105.03195. i, ii, iv, 5, 51, 52, 56

[5] I.J. Bienaymé. De la loi de multiplication et de la durée des familles. In *L'Institut, journal universel des sciences et des sociétś savantes en France et à l'étranger. 1re Section. - Sciences mathématiques, physiques et naturelles. Société philomatique de Paris, extrait des procès-verbaux des séances pendant l'année 1845*, pages 37–39, RUE du four-Saint-Germain, 47, 1845. Imprimerie

de Gossoin. URL http://www.philomathique.paris/index.php/archives-1791-1917. 3

[6] S. Boucheron, G. Lugosi, and O. Bousquet. *Concentration Inequalities*, pages 208–240. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-540-28650-9. doi: 10.1007/978-3-540-28650-9_9. URL https://doi.org/10.1007/978-3-540-28650-9_9. 47

[7] M. Camarri and J. Pitman. Limit Distributions and Random Trees Derived from the Birthday Problem with Unequal Probabilities. *Electronic Journal of Probability*, 5(none):1 – 18, 2000. doi: 10.1214/EJP.v5-58. URL https://doi.org/10.1214/EJP.v5-58. i, ii, 41, 47, 57

[8] T. E. Harris. First passage and recurrence distributions. *Transactions of the American Mathematical Society*, 73(3):471–486, 1952. ISSN 00029947. URL http://www.jstor.org/stable/1990803. 10

[9] C. C. Heyde and E. Seneta. Studies in the History of Probability and Statistics. XXXI. The simple branching process, a turning point test and a fundamental inequality: A historical note on I. J. Bienaymé. *Biometrika*, 59(3): 680–683, 12 1972. ISSN 0006-3444. doi: 10.1093/biomet/59.3.680. URL https://doi.org/10.1093/biomet/59.3.680. 3

[10] P. Jagers. Some notes on the history of branching processes, from my perspective. 2009. URL https://www.cim.pt/magazines/bulletin/14/article/118/pdf. 3

[11] S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probability Surveys*, 9(none):103 – 252, 2012. doi: 10.1214/11-PS188. URL https://doi.org/10.1214/11-PS188. i, ii, 1, 4, 5, 56, 57

[12] I. Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees. *Probability Theory and Related Fields*, 168(1):1–40, 2017. doi: 10.1007/s00440-016-0704-6. URL `https://doi.org/10.1007/s00440-016-0704-6`. i, ii, 5, 6, 10, 11

[13] I. Kortchemski. Condensation in random trees. Random Trees and Graphs Summer School, July 2019. URL `https://igor-kortchemski.perso.math.cnrs.fr/graal.pdf`. iv, 9, 13, 14

[14] J.-F. Le Gall. Random trees and applications. *Probability Surveys*, 2(none):245 – 311, 2005. doi: 10.1214/154957805100000140. URL `https://doi.org/10.1214/154957805100000140`. 9, 13

[15] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30(5):997–1015, 1978. doi: 10.4153/CJM-1978-085-0. 3

[16] J. Moon. Counting labelled trees. 1970. URL `https://www.math.ucla.edu/~pak/hidden/papers/Moon-counting_labelled_trees.pdf`. 9, 13

[17] J. Pitman. *Random walks and random forests*, pages 121–141. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-34266-3. doi: 10.1007/3-540-34266-4_7. URL `https://doi.org/10.1007/3-540-34266-4_7`. 13