La Matematica (2023) 2:743–771
https://doi.org/10.1007/s44007-023-00064-z

ORIGINAL RESEARCH ARTICLE

The Distribution of the Number of Automorphisms of Random Trees

Christoffer Olsson[1] · **Stephan Wagner**[1]

Received: 17 November 2022 / Revised: 26 July 2023 / Accepted: 14 August 2023 /
Published online: 13 September 2023

Abstract
We study the size of the automorphism group of two different types of random trees: Galton–Watson trees and rooted Pólya trees. In both cases, we prove that it asymptotically follows a log-normal distribution and provide asymptotic formulas for the mean and variance of the logarithm of the size of the automorphism group. While the proof for Galton–Watson trees mainly relies on probabilistic arguments and a general result on additive tree functionals, generating functions are used in the case of rooted Pólya trees. We also show how to extend the results to some classes of unrooted trees.

Keywords Random tree · Galton–Watson tree · Pólya tree · Unrooted tree · Automorphism group · Moments · Central limit theorem

1 Introduction

The automorphism group is a fundamental object associated with a graph as it encodes information about its symmetries. Furthermore, counting mathematical objects up to symmetry is a classical subject in combinatorics which naturally relates to the automorphism group. An example is the case of graphs, where the number of different labelings of a graph G of size n is given by $\frac{n!}{|\operatorname{Aut} G|}$. In this paper we study properties of the automorphism groups associated with random trees, in particular Galton–Watson trees and Pólya trees. We show that the size of the automorphism group follows a log-normal distribution with parameters depending on tree type. The size of the automorphism

An earlier version of this paper has previously been published as the extended abstract *Automorphisms of random trees* [18].

✉ Christoffer Olsson
christoffer.olsson@math.uu.se

Stephan Wagner
stephan.wagner@math.uu.se

[1] Department of Mathematics, Uppsala Universitet, Uppsala, Sweden

group has previously been studied in special cases of Galton–Watson trees: binary trees (expected values and limiting distribution: [4]), labeled trees (limiting distribution: [8] and expected value: [23]), binary and ternary trees (expected values: [15] and [16]). It has also been studied for some other types of trees than those considered here: specifically, random recursive trees (expected value: [14]), and d-ary increasing trees (limiting distribution and moments: [19]). We are primarily studying rooted trees but for some classes of trees we can extend the results to the unrooted case. The book [5] is a general reference to this introduction and the different types of random trees discussed in this paper.

Recall now that a *Galton–Watson tree* is a growth model where we start with one vertex, the root, and the number of children it has is given by a (discrete) random variable ξ, supported on some subset of the non-negative integers that includes at least 0 and some number greater than 1. The tree grows by letting each of the vertices have children of their own according to the offspring distribution ξ, independently of all other vertices. Different distributions for ξ give rise to different types of Galton–Watson trees. We are especially interested in the case of *critical* Galton–Watson trees, for which $\mathbb{E}\xi = 1$, as well as *conditioned* Galton–Watson trees where we condition on the size of the tree, i.e., we pick one of all possible Galton–Watson trees on n vertices at random. A related notion is that of the *size-biased* Galton–Watson tree, which has two different types of vertices. The *normal* vertices have the same offspring distribution ξ as before, while the *special* vertices get offspring according to the size-biased distribution $\hat{\xi}$ defined by $\mathbb{P}(\hat{\xi} = k) = k\mathbb{P}(\xi = k)$. We start the growth process with the root being special, and for each special vertex we choose exactly one of its children, uniformly at random, to be special as well. This means that the size-biased Galton–Watson tree has an infinite *spine* of special vertices, with non-biased unconditioned Galton–Watson trees attached to it.

Conditioned Galton–Watson trees are closely connected to, and a special case of, *simply generated families* of trees (or *simple trees*) which are defined in terms of generating functions. For a sequence of non-negative numbers $\{w_k\}$ define

$$\Phi(x) = \sum_{k \geq 0} w_k x^k$$

to be its *weight generating function*. Then the generating function for the class of trees associated with $\{w_k\}$,

$$T(x) = \sum_{T \in \mathcal{T}} w(T) x^{|T|}$$

is defined by the functional equation

$$T(x) = x\Phi(T(x)). \tag{1}$$

The number $w(T)$ is called the *weight* of the tree T. Under the (mild) assumption that there exists a positive τ within the radius of convergence of $\Phi(z)$ such that

$$\Phi(\tau) = \tau\Phi'(\tau) < \infty,$$

we can find $\rho = \frac{\tau}{\Phi(\tau)}$ such that $T(x)$ has the singular expansion

$$T(x) = \tau - c_1\sqrt{1 - \frac{x}{\rho}} + \sum_{k\geq 2}(-1)^k c_k \left(1 - \frac{x}{\rho}\right)^{\frac{k}{2}}, \tag{2}$$

for constants c_k that can be calculated. Through the process of singularity analysis, this implies that the total weight of all trees of size n is asymptotic to

$$Cn^{-3/2}\rho^{-n}.$$

We take the probability of picking a given tree S of size n to be

$$\frac{w(S)}{\sum_{|T|=n} w(T)}. \tag{3}$$

We can see Galton–Watson trees and simple trees as two sides of the same coin, one being probabilistic and the other being combinatorial, where Galton–Watson trees correspond to simply generated trees with weights w_k adding up to 1. In this context, the numbers w_k correspond to the probability of a vertex having k children, $w(T)$ is the probability of obtaining T through the Galton–Watson growth process and (3) is the probability when we condition on the size of the tree. In fact, if we can find a τ as above, we can always assume that our trees, whether they are conditioned Galton–Watson or simply generated ones, are critical Galton–Watson trees as long as we can perform slight modifications (which will not affect the probabilities of individual trees) to the offspring distribution. Then, the critical Galton–Watson trees are those simple trees having their dominant singularity at $\rho = 1$, so that the discussion above indicates that the probability of an (unconditional) Galton–Watson tree having size n decays like $Cn^{-3/2}$. Examples of Galton–Watson (and simply generated) trees are plane trees, labeled trees, d-ary trees, etc.

Pólya trees are unordered, unlabeled trees which can be either rooted or unrooted. Rooted Pólya trees have many properties similar to Galton–Watson trees, but they do not satisfy the definition so we will need other methods to deal with them. They can be characterized by their generating function $P(x) = \sum_{T\in\mathcal{P}} x^{|T|}$, which satisfies

$$P(x) = x\exp\left(\sum_{k=1}^{\infty}\frac{P(x^k)}{k}\right). \tag{4}$$

The number of such trees of size n is asymptotic to $An^{-3/2}\rho_p^{-n}$, where $\rho_p = 0.33832\ldots$ is the dominant singularity of $P(x)$ and A is a constant. For this singularity, we have $P(\rho_p) = 1$.

A classical result gives a bijection between Pólya trees and the union of unrooted unlabeled trees together with pairs of distinct Pólya trees. The bijection translates into the functional equation

$$U(x) = P(x) - \frac{1}{2}P(x)^2 + \frac{1}{2}P(x^2) \tag{5}$$

that describes the generating function for unrooted trees $U(x)$ in terms of $P(x)$. The number of unrooted Pólya trees of size n is asymptotic to $Bn^{-5/2}\rho_p^{-n}$ for a constant B.

We use $\mathcal{T}$ to denote Galton–Watson trees, $\mathcal{T}_n$ to denote conditioned Galton–Watson trees on n vertices and $\hat{\mathcal{T}}$ to denote size-biased trees. Similarly, we use T, T_n and $\hat{T}$ to denote specific realizations of the respective trees. Furthermore, we will use $\mathcal{P}$ and $\mathcal{P}_n$ to denote rooted Pólya trees as well as Pólya trees of size n, respectively and, sometimes, $\mathcal{U}$ and $\mathcal{U}_n$ in the case of unrooted trees. We let $\deg(T)$ denote the degree of the root of T and $\mathrm{mult}(B)$ be the number of occurrences of a particular tree B as root branches of some other tree (the *root branches* of a rooted tree are the subtrees obtained as components when the root is removed). Note that the isomorphism classes of Galton–Waton trees are rooted Pólya trees. In addition to using $w(T)$ for the weight of a simple tree, we will use $W(B)$ to denote the weight of the entire isomorphism class B.

1.1 Results

In this paper, we will show asymptotic normality of $\log|\operatorname{Aut}\mathcal{T}_n|$, for various classes of random trees. This implies asymptotic log-normality of $|\operatorname{Aut}\mathcal{T}_n|$. We prove the following theorem on the automorphism group of Galton–Watson trees.

Theorem 1 *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of size n with offspring distribution ξ, where $\mathbb{E}\xi = 1$, $0 < \operatorname{Var}\xi < \infty$ and $\mathbb{E}\xi^5 < \infty$. Then there exist constants μ and $\sigma^2 \geq 0$, depending on $\mathcal{T}$, such that*

$$\frac{\log|\operatorname{Aut}\mathcal{T}_n| - \mu n}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0,\sigma^2).$$

The condition on $\mathbb{E}\xi^5$ is needed for technical purposes and is valid for combinatorially significant examples such as labeled trees, plane trees and d-ary trees. The exponent 5 is probably not best possible, but required to apply the general result on additive functionals that our proof is based on.

The mean constant μ and even more so the variance constant σ^2 do not seem easy to compute numerically in general. We show how to derive the numerical values for some classes of trees, namely labeled trees as well as general Galton–Watson trees with bounded degrees. Numerical estimates for some types of trees can be found in Table 1.

Note that it is unclear what an unrooted version of a Galton–Watson tree is in general so we cannot expect an unrooted version of Theorem 1, but in the case of

Table 1 Numerical estimates of the mean and variance constants for some types of trees

Class of tree	μ	σ^2
Labeled trees	0.0522901	0.0394984
Full binary trees	0.0939359	0.0252103
Pruned binary trees	0.0145850	0.0084835
Pólya trees	0.1373423	0.1967696

labeled trees, the result for rooted trees translates to the case of unrooted trees as well.

Theorem 2 *Let $\mathcal{T}_n$ be a uniformly random unrooted labeled tree of size n. Then, $\mathbb{E}(\log|\operatorname{Aut}\mathcal{T}_n|) = \mu n + O(1)$ and $\operatorname{Var}(\log|\operatorname{Aut}\mathcal{T}_n|) = \sigma^2 n + O(1)$, with $\mu = 0.0522901\ldots$ and $\sigma^2 = 0.0394984\ldots$. Furthermore, we have*

$$\frac{\log|\operatorname{Aut}\mathcal{T}_n| - \mu n}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0, \sigma^2).$$

We can also prove asymptotic log-normality for the size of the automorphism group of Pólya trees.

Theorem 3 *Let $\mathcal{P}_n$ be a uniformly random Pólya tree of size n, rooted or unrooted. Then, $\mathbb{E}(\log|\operatorname{Aut}\mathcal{P}_n|) = \mu n + O(1)$ and $\operatorname{Var}(\log|\operatorname{Aut}\mathcal{P}_n|) = \sigma^2 n + O(1)$, with $\mu = 0.1373423\ldots$ and $\sigma^2 = 0.1967696\ldots$. Furthermore, we have*

$$\frac{\log|\operatorname{Aut}\mathcal{P}_n| - \mu n}{\sqrt{n}} \xrightarrow{d} \mathrm{N}(0, \sigma^2).$$

The proofs of Theorem 1 and Theorem 3 rely at their cores on the same idea of approximating the additive functionals by simpler ones, but they are fairly different at a glance. We give some preliminary results in Sect. 2. We then prove Theorem 1 in Sect. 3 and Theorem 3 for rooted trees in Sect. 4. The results for unrooted trees are proved in Sect. 5.

2 Preliminaries

For any rooted tree T, we have a recursive formula for the size of its automorphism group. Let $T_1, T_2, \ldots, T_k$ be its root branches up to isomorphism, having multiplicities $m_1, m_2, \ldots, m_k$, respectively. Then we have

$$|\operatorname{Aut} T| = \prod_{i=1}^{k} m_i! |\operatorname{Aut} T_i|^{m_i}, \tag{6}$$

derived from the fact that the automorphism group of a rooted tree is obtained from symmetric groups by iterated direct and wreath products (see [3], Proposition 1.15). In

other words, the tree is invariant under the automorphisms of each of the root branches as well as under permutation of isomorphic branches. By taking logarithms, we find that

$$\log|\operatorname{Aut} T| = \sum_{i=1}^{k} \log(m_i!) + \sum_{i=1}^{k} m_i \log|\operatorname{Aut} T_i|.$$

This means that $\log|\operatorname{Aut} T|$ is an *additive functional* of the tree (see [5, Sect. 3.2] or any of the references below) with *toll function* $f(T) = \sum_{i=1}^{k} \log(m_i!)$.

Limit theorems for additive functionals have been proven for various classes of random trees under different conditions, see [6, 7, 10, 19, 20, 22]. In the case of Galton–Watson trees, we will specifically make use of a general result on *almost local* additive functionals due to Ralaivaosaona, Šileikis and the second author [20], which is in turn based on earlier work by Janson [10]. Intuitively, "almost local" means that looking at the first M levels of the tree gives us substantial (albeit not perfect) information about the value of the toll function at the root. We will let $\mathcal{T}^{(M)}$ denote the restriction of a Galton–Watson tree to its first M levels, where the root is at level 0, with similar definitions for the other classes of trees. The theorem we will use is the following.

Theorem 4 ([20]) *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of size n with offspring distribution ξ, with $\mathbb{E}\xi = 1$ and $0 < \sigma^2 := \operatorname{Var}\xi < \infty$. Assume further that $\mathbb{E}\xi^{2\alpha+1} < \infty$ for some integer $\alpha \geq 0$. Consider a functional F of finite rooted ordered trees with the property that*

$$f(T) = O(\deg(T)^{\alpha}),$$

where f is the toll function associated with the functional.

Furthermore, assume that there exists a sequence $(p_M)_{M\geq 1}$ of positive numbers with $p_M \to 0$ as $M \to \infty$, such that

- *for every integer $M \geq 1$,*

$$\mathbb{E}\left| f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) | \hat{\mathcal{T}}^{(M)} \right) \right| \leq p_M,$$

  for all $N \geq M$,
- *there is a sequence of positive integers $(M_n)_{n\geq 1}$ such that for large enough n,*

$$\mathbb{E}|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M_n)})| \leq p_{M_n}.$$

If $a_n = n^{-1/2}(n^{\max\{\alpha,1\}} p_{M_n} + M_n^2)$ satisfies

$$\lim_{n\to\infty} a_n = 0, \text{ and } \sum_{n=1}^{\infty} \frac{a_n}{n} < \infty,$$

then

$$\frac{F(\mathcal{T}_n) - \mu n}{\sqrt{n}} \xrightarrow{d} N(0, \gamma^2),$$

where $\mu = \mathbb{E} f(\mathcal{T})$ *and* $0 \leq \gamma^2 < \infty$.

The proof shows that the result still holds if we replace $(F(\mathcal{T}_n) - \mu n)/\sqrt{n}$ by $(F(\mathcal{T}_n) - \mathbb{E}F(\mathcal{T}_n))/\sqrt{n}$. We remark that $\gamma = 0$ means that $\frac{F(\mathcal{T}_n)-\mu n}{\sqrt{n}}$ converges in distribution (thus also in probability) to 0. However, this case does not occur in any of the examples considered here.

To prove the result for Pólya trees we will instead rely on generating functions. We can define the generating function of $F(\mathcal{P}_n) = \log |\operatorname{Aut} \mathcal{P}_n|$ to be

$$P(x, t) = \sum_{T \in \mathcal{P}} e^{t \log |\operatorname{Aut} T|} x^{|T|} = \sum_{T \in \mathcal{P}} |\operatorname{Aut} T|^t x^{|T|}. \tag{7}$$

Note that $P(x, 0) = P(x)$. We can now derive a functional equation analogous to (4) as follows. We have the symbolic decomposition

$$\mathcal{P} = \bullet \times \bigotimes_{T \in \mathcal{P}} (\emptyset \uplus \{T\} \uplus \{T, T\} \uplus \cdots),$$

reflecting the fact that a Pólya tree consists of a tree and a multiset of branches. Taking automorphisms into account, this translates to

$$P(x, t) = x \prod_{T \in \mathcal{P}} \left(\sum_{n=0}^{\infty} x^{n|T|} n!^t |\operatorname{Aut} T|^{nt} \right),$$

by general principles for generating functions. We can manipulate this as follows:

$$\begin{aligned} P(x, t) &= x \exp\left(\sum_{T \in \mathcal{P}} \log \left(\sum_{n=0}^{\infty} x^{n|T|} n!^t |\operatorname{Aut} T|^{nt} \right) \right) \\ &= x \exp\left(\sum_{T \in \mathcal{P}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left(\sum_{n=1}^{\infty} x^{n|T|} n!^t |\operatorname{Aut} T|^{nt} \right)^k \right). \end{aligned}$$

The sum in the exponent can be rewritten as

$$\sum_{T \in \mathcal{P}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{\lambda_1 + \lambda_2 + \cdots = k} \binom{k}{\lambda_1, \lambda_2, \ldots} \prod_{n=1}^{\infty} \left(x^{n|T|} n!^t |\operatorname{Aut} T|^{nt} \right)^{\lambda_n}.$$

We now write integer partitions as sequences $\lambda = (\lambda_1, \lambda_2, \ldots)$, where λ_i is the number of i's in the partition. The total number of summands is denoted by $|\lambda| = \lambda_1 + \lambda_2 + \cdots$,

and we write $\lambda \vdash j$ to denote that λ is a partition of j, i.e. $j = \lambda_1 + 2\lambda_2 + 3\lambda_3 + \cdots$. Further manipulations give

$$\begin{aligned}
&\sum_{T\in\mathcal{P}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\sum_{j=1}^{\infty}\sum_{\substack{\lambda_1+\lambda_2+\cdots=k\\ \lambda_1+2\lambda_2+\cdots=j}}\binom{k}{\lambda_1,\lambda_2,\ldots}x^{j|T|}|\operatorname{Aut}T|^{jt}\prod_{n=1}^{\infty}n!^{\lambda_n t}\\
&=\sum_{j=1}^{\infty}\sum_{\lambda\vdash j}\frac{(-1)^{|\lambda|-1}}{|\lambda|}\binom{|\lambda|}{\lambda_1,\lambda_2,\ldots}\left(\prod_{n=1}^{\infty}n!^{\lambda_n t}\right)\sum_{T\in\mathcal{P}}x^{j|T|}|\operatorname{Aut}T|^{jt}\\
&=\sum_{j=1}^{\infty}\sum_{\lambda\vdash j}\frac{(-1)^{|\lambda|-1}}{|\lambda|}\binom{|\lambda|}{\lambda_1,\lambda_2,\ldots}\left(\prod_{n=1}^{\infty}n!^{\lambda_n t}\right)P(x^j,jt).
\end{aligned}$$

For convenience, we can define

$$c(j,t)=j\sum_{\lambda\vdash j}\frac{(-1)^{|\lambda|-1}}{|\lambda|}\binom{|\lambda|}{\lambda_1,\lambda_2,\ldots}\left(\prod_{n=1}^{\infty}n!^{\lambda_n t}\right),$$

and arrive at the functional equation

$$P(x,t)=x\exp\left(P(x,t)+\sum_{j=2}^{\infty}\frac{c(j,t)}{j}P(x^j,jt)\right). \tag{8}$$

Note that $c(j,0)=1$, so that we recover the functional equation (4) if we set $t=0$.

3 The Automorphism Group of Galton–Watson Trees

As indicated in the previous section, we will show that $\log|\operatorname{Aut}\mathcal{T}_n|$ is in fact an almost local additive functional. This will let us apply Theorem 4 to prove that it converges in distribution to a normal random variable.

3.1 Galton–Watson Trees Isomorphic Up to a Certain Level

In applying Theorem 4, we are led to consider the probability that two Galton–Watson trees are of height $\geq M$ and isomorphic. We use $\mathcal{C}$ to denote the set of isomorphism classes of Galton–Watson trees as well as $\mathcal{C}^M$ to denote the set of isomorphism classes of trees of height M (i.e., trees that have $M+1$ generations). The definitions extend to conditioned Galton–Watson trees as $\mathcal{C}_n$ and $\mathcal{C}_n^M$, respectively. We start with the following lemma.

Lemma 1 *There exists some constant* $0 < c < 1$ *such that*

$$\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \leq c^M,$$

uniformly for all isomorphism classes $C \in \mathcal{C}^M$.

Proof We say that a level L of a tree T agrees with C if it has the correct number of vertices and the offsprings $\xi_1, \xi_2, \ldots, \xi_l$ agree with the offsprings of the same level in C, up to permutation. Let $L_1, L_2, \ldots$ denote the levels of the Galton–Watson tree $\mathcal{T}$. Then the probability is bounded by

$$\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \leq \prod_{i=0}^{M-1} \mathbb{P}(L_i \text{ agrees with } C | L_1, L_2, \ldots, L_{i-1}), \tag{9}$$

where we note that, by truncation, the M-th level will always agree with C, as long as the previous ones do. We can bound each factor in (9) by the probability of the level having the correct number of leaves, conditioned on the previous levels. This random variable follows a binomial distribution with probability $p = \mathbb{P}(\xi = 0)$. It is therefore sufficient to prove a bound $0 < c < 1$ (uniform in both l and k) on the probability that a binomial variable $X_l \sim \text{Bin}(l, p)$ takes a specific value k.

We can in fact bound X_l in terms of p, since if we write X_l as a sum of Bernoulli variables, $X_l = Y_1 + Y_2 + \cdots + Y_l$, we have

$$\begin{aligned}\mathbb{P}(Y_1 + Y_2 + \cdots + Y_l = k) &= \sum_{r=0}^{1} \mathbb{P}(Y_1 + Y_2 + \cdots + Y_{l-1} = k - r)\mathbb{P}(Y_l = r) \\ &\leq \sum_{r=0}^{1} \mathbb{P}(Y_1 + Y_2 + \cdots + Y_{l-1} = k - r) \max_{y \in \{0,1\}} \mathbb{P}(Y_l = y) \leq \max\{p, 1-p\}.\end{aligned}$$

We can thus take $c = \max\{p, 1-p\}$ as a uniform bound for all levels, and now (9) gives the result. □

We now see that for two independent trees $\mathcal{T}_1, \mathcal{T}_2$ we have

$$\begin{aligned}\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ iso. and of height} \geq M) &= \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)^2 \\ &\leq \max_{C \in \mathcal{C}^M} \left\{ \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \right\} \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \\ &= \max_{C \in \mathcal{C}^M} \{\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)\}. \end{aligned} \tag{10}$$

Combining this with Lemma 1, we get the following corollary.

Corollary 1 *Let $\mathcal{T}_1, \mathcal{T}_2$ be two independent Galton–Watson trees. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ isomorphic and of height} \geq M) \leq c^M.$$

In fact, the argument in (10) also works when one of the trees is the size-biased tree $\hat{\mathcal{T}}$, which lets us bound the probability that a Galton–Watson tree and the size-biased tree are isomorphic up to level M in terms of the maximum probability that the Galton–Watson tree belongs to a specific isomorphism class. This gives another corollary, which we will need later on.

Corollary 2 *Let $\mathcal{T}$ be a Galton–Watson tree and $\hat{\mathcal{T}}$ be the size-biased tree, assumed to be independent of $\mathcal{T}$. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ isomorphic and of height} \geq M) \leq c^M.$$

We can obtain similar bounds on the probability that two conditioned Galton–Watson trees are isomorphic up to level M. We start by extending Lemma 1 to the conditioned case.

Lemma 2 *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of size n. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_n^{(M)} \text{ belongs to } C) = O\left(n^{\frac{5}{2}} c^M\right),$$

uniformly for all isomorphism classes $C \in \mathcal{C}_n^M$.

Proof Order the offsprings $\xi_1, \xi_2, \ldots$ of T_n in breadth-first order and consider the sums

$$S_m = \sum_{i=1}^{m} (\xi_i - 1) \quad \text{for } 1 \leq m \leq n,$$

where we also define $S_0 = 0$. In each step, $1 \leq i \leq m$, we are deleting 1 for the current vertex while adding the number of children it has. For a conditioned Galton–Watson tree of size n, we necessarily have $S_m > -1$ for $1 \leq m < n$, and $S_n = -1$, since we are adding 1 for all vertices except the root, but deleting 1 for all vertices including the root. This is a well-known construction, see e.g. [2]. Using it, we can formulate the probability we seek to bound in the following way.

$$\mathbb{P}(T_n^{(M)} \text{ belongs to } C) = \frac{\mathbb{P}(\{T' \text{ belongs to } C\} \cap \{S_1, S_2, \ldots, S_{n-1} > -1, S_n = -1\})}{\mathbb{P}(S_1, S_2, \ldots, S_{n-1} > -1, S_n = -1)},$$

where T' is a Galton–Watson tree with offsprings $\xi_1, \xi_2, \ldots, \xi_k$, and k is the number of vertices of each tree in C excluding the last level (since we truncate at level M, the

number of children the vertices on this level have is of no interest to us). Since the trees in C are isomorphic, they will all have the same number of vertices.

Let l_M be the number of vertices at the last level of each tree in C (again, equal due to isomorphism). Then we have

$$\sum_{i=1}^{n}(\xi_i - 1) = \sum_{i=1}^{k}(\xi_i - 1) + \sum_{i=k+1}^{n}(\xi_i - 1) = l_M - 1 + \sum_{i=k+1}^{n}(\xi_i - 1).$$

By the conditions set on S_m, we draw the conclusion that

$$S'_m := \sum_{i=k+1}^{k+m}(\xi_i - 1) > -l_M \quad \text{for } 1 \le m < n-k,$$
$$S'_{n-k} := \sum_{i=k+1}^{n}(\xi_i - 1) = -l_M.$$

By independence, we now have

$$\frac{\mathbb{P}(\{T' \text{ belongs to } C\} \cap \{S_1, S_2, \ldots, S_{n-1} > -1, S_n = -1\})}{\mathbb{P}(S_1, S_2, \ldots, S_{n-1} > -1, S_n = -1)}$$
$$= \frac{\mathbb{P}(T' \text{ belongs to } C)\mathbb{P}(S'_1, S'_2, \ldots, S'_{n-k-1} > -l_M, S'_{n-k} = -l_M)}{\mathbb{P}(S_1, S_2, \ldots, S_{n-1} > -1, S_n = -1)},$$

and using the cycle lemma we find that this equals

$$\frac{\frac{l_M}{n-k}\mathbb{P}(S'_{n-k} = -l_M)}{\frac{1}{n}\mathbb{P}(S_n = -1)}\mathbb{P}(T' \text{ belongs to } C).$$

The probability $\mathbb{P}(S'_{n-k} = -l_M)$ is bounded by 1, and S_n satisfies a local limit theorem. If we also bound $l_M \le n$ as well as $n - k \ge 1$ (k is the number of vertices up to level $M-1$, and by definition there must be at least one vertex at level M) and use Lemma 1 (note that $\mathcal{C}_{n,M}$ is a subset of $\mathcal{C}^M$), we arrive at

$$\mathbb{P}(T_n^{(M)} \text{ belongs to } C) = O\left(n^{\frac{5}{2}}c^M\right),$$

which is what we wanted to prove. □

Furthermore, using calculations similar to (10), we obtain the following corollary.

Corollary 3 *Let $\mathcal{T}_{n_1}, \mathcal{T}_{n_2}$ be two independent conditioned Galton–Watson trees. There exists some constant $0 < c < 1$ such that*

$$\mathbb{P}(\mathcal{T}_{n_1}^{(M)}, \mathcal{T}_{n_2}^{(M)} \textit{ isomorphic and of height} \ge M) = O\left(n^{\frac{5}{2}}c^M\right),$$

where we can take $n = \min\{n_1, n_2\}$.

We are now ready to apply the central limit theorem for additive functionals.

3.2 Applying the CLT for Almost Local Additive Functionals

By Stirling's approximation, we can bound $f(T) \leq \log \deg(T)! = O(\deg(T)^{1+\epsilon})$ for any $\epsilon > 0$, so that the functional satisfies the degree condition of Theorem 4 with $\alpha = 2$. For the expectations, there are two conditions to check, one for the size-biased Galton–Watson tree and one for the conditioned Galton–Watson tree, and in each case the difference inside the expectation can only be non-zero if (at least) two branches are isomorphic up to level M but non-isomorphic when we take all levels into account. We can therefore reduce the problem to studying trees that are isomorphic up to the M-th level.

We note that if l root branches are isomorphic up to level M, this contributes at most $\log(l!) \leq \binom{l}{2}$ to the difference inside the expectation. Therefore, the contribution of a random tree can be bounded by the sum of indicators

$$\sum_{T_i, T_j \text{ root branches}} I(T_i^{(M)}, T_j^{(M)} \text{ isomorphic and of height} \geq M),$$

where we sum over distinct branches. We can thus bound the expectation $\mathbb{E}|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})|$ for the conditioned Galton–Watson tree by

$$\mathbb{E}\left(\sum_{\substack{\mathcal{T}_i, \mathcal{T}_j \\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M) \right).$$

This can, in turn, be bounded by

$$\begin{aligned}
&\sum_{k \geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \sum_{n_1, n_2} \mathbb{P}(|\mathcal{T}_i| = n_1 | \deg(\mathcal{T}_n) = k) \mathbb{P}(|\mathcal{T}_j| = n_2 | \deg(\mathcal{T}_n) = k) \\
&\quad \cdot \binom{k}{2} \mathbb{E}\left(I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ iso. with height} \geq M) \,\middle|\, |\mathcal{T}_i| = n_1, |\mathcal{T}_j| = n_2 \right) \\
&= O\left(\sum_{k \geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \binom{k}{2} n^{\frac{5}{2}} c^M \right) = O\left(n^{\frac{5}{2}} c^M \sum_{k \geq 2} k \mathbb{P}(\xi = k) \binom{k}{2} \right)
\end{aligned}$$

where we use the law of total expectation and the fact that $\mathbb{P}(\deg(\mathcal{T}_n) = k) \leq c_0 k \mathbb{P}(\xi = k)$ for all k and n, where c_0 is constant [9, (2.7)]. By assumptions on the moments of the offspring distribution, this expression is $O(n^{\frac{5}{2}} c^M)$.

The difference $|f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}(f(\hat{\mathcal{T}}^{(N)})|\hat{\mathcal{T}}^{(M)})|$ must also be zero unless some branches are isomorphic up to level M, and reasoning similar to above lets us rewrite

its expectation in the following way.

$$\sum_{k\geq 2} kP(\xi = k)\Bigg(\mathbb{E}\Big(\sum_{\substack{\mathcal{T}_i,\mathcal{T}_j \text{ non-special}\\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ iso. with height} \geq M)\Big)$$
$$+\mathbb{E}\Big(\sum_{\substack{\mathcal{T} \text{ non-special root branch}\\ \hat{\mathcal{T}} \text{ special root branch}}} I(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ iso. with height} \geq M)\Big)\Bigg),$$

which is equal to

$$\sum_{k\geq 3} kP(\xi = k)\binom{k-1}{2}\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ iso. and of height} \geq M)$$
$$+\sum_{k\geq 2} kP(\xi = k)(k-1)\mathbb{P}(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ iso. and of height} \geq M) = O(c^M)$$

by Corollaries 1 and 2 (the constant c is the same for both of these corollaries since they both rely on Lemma 1) as well as assumptions on moments of the offspring distribution.

We now set $p_M = Kc_1^M$, for $c < c_1 < 1$ and some suitable constant K, as well as $M_n = A \log n$, for some positive constant A that is large enough to make $n^{5/2}c^{M_n} \leq c_1^{M_n}$ for all n and $A \log c_1 < -3/2$. Then, the expectations mentioned in Theorem 4 are bounded by p_M and p_{M_n}, respectively. Furthermore, the sequence a_n goes to 0 and satisfies $\sum a_n/n < \infty$. Thus, we can apply Theorem 4 to show that $\log|\operatorname{Aut}\mathcal{T}_n|$ is asymptotically normal, which completes the proof of Theorem 1.

3.3 Mean and Variance for Some Classes of Trees

In general, calculating the mean and variance constants for Galton–Watson trees seems to be a difficult feat, but we show how to do it in the special cases of labeled trees as well as Galton–Watson trees with bounded degrees. In both cases we view the trees as simply generated and rely on generating functions but otherwise the methods for the two cases are different. We stress that the calculations do not rely on Theorem 4 so we do not need to assume that the trees are critical.

3.3.1 Galton–Watson Trees with Bounded Degrees

We now restrict our attention to the case of Galton–Watson trees with degrees restricted to lie in a finite set D. In other words, the degrees are bounded above by some constant. By general principles of generating functions, we know that we can calculate the mean by studying the first derivative of

$$T(x,t) = \sum_{T\in\mathcal{T}} |\operatorname{Aut} T|^t x^{|T|}$$

with respect to t. Likewise, we can find the variance by studying the second derivative. Using the fact that $\log|\operatorname{Aut} T|$ is an additive functional, we start with the following expression, which was derived for general additive functionals in [22]

$$T_t(x,0) = \frac{xT_x(x,0)}{T(x,0)} H(x). \tag{11}$$

Here $H(x) = \sum w(T) f(T) x^{|T|}$, with $f(T)$ being the toll function of the additive functional. We already know the singular expansion for $T(x) = T(x,0)$ from (2) and we can differentiate it termwise to obtain a singular expansion for $T_x(x)$. Thus, it is enough to study $H(x)$. We manipulate the function in the following way.

$$\begin{aligned} H(x) &= \sum_T w(T) x^{|T|} \left(\sum_{i=1}^{k} \log(m_i!) \right) \\ &= \sum_T w(T) x^{|T|} \left(\sum_{B \in \mathcal{B}_I(T)} \log(\operatorname{mult}(B)!) \right) \\ &= \sum_B \sum_{m=1}^{\infty} \log(m!) \sum_{\substack{T : B \ m\text{-fold branch} \\ \text{of } T \text{ up to iso.}}} w(T) x^{|T|}. \end{aligned}$$

Note here that $\sum_B$ is a sum over isomorphism classes B (i.e., rooted Pólya trees).

Using the fact that B occurs exactly m times in T, we can rewrite the innermost sum as

$$\begin{aligned} & x \sum_{k=m}^{\infty} w_k \binom{k}{m} (W(B) x^{|B|})^m \left(\sum_{T \neq B} w(T) x^{|T|} \right)^{k-m} \\ &= x \sum_{k=m}^{\infty} w_k (W(B) x^{|B|})^m \binom{k}{m} \left(T(x,0) - W(B) x^{|B|} \right)^{k-m}. \end{aligned}$$

This gives that $H(x)$ is equal to

$$\begin{aligned} & x \sum_B \sum_{m=1}^{\infty} \frac{\log(m!)}{m!} (W(B) x^{|B|})^m \sum_{k=m}^{\infty} w_k \frac{k!}{(k-m)!} \left(T(x,0) - W(B) x^{|B|} \right)^{k-m} \\ &= x \sum_B \sum_{m=1}^{\infty} \frac{\log(m!)}{m!} (W(B) x^{|B|})^m \Phi^{(m)}(T(x,0) - W(B) x^{|B|}), \end{aligned}$$

due to Taylor's theorem. As the degrees are bounded, there is some k_0 such that $w_k = 0$ for $k \geq k_0$. Thus, Φ is a polynomial, and the inner sum (which is actually finite, as $\Phi^{(m)}(t)$ is eventually 0) is a polynomial in $W(B) x^{|B|}$ and $T(x,0)$. It follows that $H(x)$

can be expressed in the form

$$H(x) = x \sum_{m=2}^{M} \sum_{B} \left(W(B) x^{|B|} \right)^m P_m(T(x,0)),$$

for some polynomial P_m. Note here that the sum starts at $m = 2$ because $\log(1!) = 0$. Let us now consider the sum over B:

$$\sum_{B} \left(W(B) x^{|B|} \right)^m.$$

In [17] it was shown that the probability

$$p_n = \frac{\sum_{|B|=n} W(B)^2}{\left(\sum_{|B|=n} W(B) \right)^2},$$

that two Galton–Watson trees with bounded degrees are isomorphic decays exponentially in n. Let $t_n = [x^n]T(x,0) = \sum_{|B|=n} W(B)$ be the total weight of all trees with n vertices. We have, for every $m \geq 2$,

$$\begin{aligned} \sum_{B} \left(W(B)|x|^{|B|} \right)^m &= \sum_{n \geq 1} \sum_{|B|=n} W(B)^m |x|^{nm} \\ &\leq \sum_{n \geq 1} \left(\sum_{|B|=n} W(B) \right)^{m-2} \sum_{|B|=n} W(B)^2 |x|^{nm} \\ &= \sum_{n \geq 1} p_n (t_n |x|^n)^m. \end{aligned}$$

As p_n decays exponentially, this shows that the sum $\sum_B \left(W(B)|x|^{|B|} \right)^m$ has greater radius of convergence than $T(x,0)$, so it represents an analytic function in a disk around 0 that contains the dominant singularity ρ of $T(x,0)$ in its interior.

Thus, we can write $H(x) = G(x, T(x,0))$, where $G(x,t)$ is a polynomial in t whose coefficients are functions of x that are analytic in a larger region than $T(x,0)$. Furthermore, the singular expansion for $T(x,0)$ carries over to a singular expansion for $H(x)$ around the dominant singularity ρ. Applying this to (11), we obtain a singular expansion for T_t. By the method of singularity analysis we find that the mean has the form $\mu n + O(1)$ for a constant μ given by

$$\mu = \frac{G(\rho, \tau)}{\tau}.$$

For given classes of Galton–Watson trees with bounded degrees, we can estimate the constant μ numerically by truncating the series

$$G(x,t) = x \sum_{B} \sum_{m=1}^{\infty} \frac{\log(m!)}{m!} (W(B)x^{|B|})^m \Phi^{(m)}(t - W(B)x^{|B|})$$

and approximating its value at $(x,t) = (\rho, \tau)$.

To find the variance we need to study $T_{tt}(x,0)$, and, again by [22], we have (slightly modified to account for the change of variables $u \to e^t$)

$$\begin{aligned} T_{tt}(x,0) = {} & \frac{\Phi''(T(x,0))}{\Phi(T(x,0))} x T_x(x,0) T_t(x,0)^2 \\ & + \frac{x T_x(x,0)}{T(x,0)} \sum_{T \in \mathcal{T}} f(T)(2F(T) - f(T)) w(T) x^{|T|}, \end{aligned}$$

It is enough to show that the final sum is well behaved. We note that

$$2F(T) - f(T) = 2 \sum_{B \in \mathcal{B}_I(T)} \operatorname{mult}(B) F(B) + f(T),$$

so that we need to study the sums

$$\sum_{T \in \mathcal{T}} w(T) x^{|T|} f(T)^2,$$

and

$$\sum_{T \in \mathcal{T}} w(T) x^{|T|} f(T) \sum_{B \in \mathcal{B}_I(T)} \operatorname{mult}(B) F(B).$$

These expressions can be shown to have a larger region of convergence than $T_{tt}(x,0)$ by arguments akin to those we used for $H(x)$. This means that we can get a singular expansion of $T_{tt}(x,0)$, and an application of singularity analysis together with division by $[x^n]T(x,0)$ gives us the asymptotic behavior of $\mathbb{E}(\log(|\operatorname{Aut} T|)^2)$. The variance is then obtained by subtracting $\mathbb{E}(\log(|\operatorname{Aut} T|))^2$. We find that it has the form $\sigma^2 n + O(1)$ for some constant σ^2 that can be calculated. Note, in particular, that there is cancellation of terms of order n^2, a common feature for these types of additive functionals due to Theorem 4 where both the mean and variance is of order n.

We can, for example, estimate the moments for full binary trees (where every internal vertex has two children) and pruned binary trees (where every internal vertex has a left child, a right child, or both). Full binary trees have mean constant $\mu \approx 0.0939359$ and variance constant $\sigma^2 \approx 0.0252103$, and in the case of pruned binary trees we get $\mu \approx 0.0145850$ and $\sigma^2 \approx 0.0084835$. Both of these classes are closely related to the phylogenetic trees studied in [4], and the mean constants above agree with the one for phylogenetic trees after translating between the models.

3.3.2 Labeled Trees

We now show how the constants μ and σ^2 in Theorem 1 can be computed for labeled trees with fairly good accuracy. To this end, we use the functional equation (8). Note that we can rewrite it in terms of an analogously defined exponential generating function for rooted labeled trees. Set

$$R(x,t) = \sum_{T\in\mathcal{R}} |\operatorname{Aut} T|^t \frac{x^{|T|}}{|T|!},$$

the sum now being over the set $\mathcal{R}$ of all rooted labeled trees. Since the number of distinct ways to label a Pólya tree T is $|T|!/|\operatorname{Aut} T|$, we have the relation

$$R(x,t) = P(x,t-1),$$

so the functional equation for Pólya trees immediately translates to a functional equation for labeled trees:

$$R(x,t) = x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,t-1)}{j}R(x^j, jt-j+1)\right). \tag{12}$$

When $t = 0$, one verifies easily (compare the calculations below for the derivative with respect to t) that $c(j,-1) = 0$ for $j > 1$ and $c(1,-1) = 1$, so the functional equation reduces to $R(x,0) = x\exp(R(x,0))$ as expected.

In order to determine the desired moments, we need to consider the derivatives with respect to t. To this end, let $x_1, x_2, \ldots$ be auxiliary variables, and note first that

$$\begin{aligned}\sum_{j\ge 0} y^j \sum_{\lambda\vdash j}\prod_{k\ge 1}\frac{x_k^{\lambda_k}}{\lambda_k!k!^{\lambda_k}} &= \prod_{k\ge 1}\sum_{\lambda_k\ge 0}\frac{x_k^{\lambda_k}y^{k\lambda_k}}{\lambda_k!k!^{\lambda_k}}\\ &= \prod_{k\ge 1}\exp\Big(\frac{x_k y^k}{k!}\Big) = \exp\Big(\sum_{k\ge 1}\frac{x_k y^k}{k!}\Big).\end{aligned}$$

Differentiating with respect to x_m and plugging in $x_1 = x_2 = \cdots = x$ yields

$$\sum_{j\ge 0} y^j \sum_{\lambda\vdash j} x^{|\lambda|-1}\lambda_m \prod_{k\ge 1}\frac{1}{\lambda_k!k!^{\lambda_k}} = \frac{y^m}{m!}\exp\Big(\sum_{k\ge 1}\frac{xy^k}{k!}\Big) = \frac{y^m}{m!}\exp(x(e^y-1)).$$

Consequently,

$$\sum_{\substack{\lambda\vdash j\\ |\lambda|=r}}\lambda_m\prod_{k\ge 1}\frac{1}{\lambda_k!k!^{\lambda_k}} = [x^{r-1}y^j]\frac{y^m}{m!}\exp(x(e^y-1)) = [y^{j-m}]\frac{(e^y-1)^{r-1}}{(r-1)!m!}.$$

By definition, we have

$$\frac{d}{dt}\frac{c(j,t)}{j} = \sum_{\lambda\vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|}\binom{|\lambda|}{\lambda_1,\lambda_2,\ldots}\left(\prod_{n=1}^{\infty} n!^{\lambda_n t}\right)\sum_{m=1}^{\infty}\lambda_m \log(m!),$$

which therefore becomes

$$\begin{aligned}
&\frac{d}{dt}\frac{c(j,t)}{j}\Big|_{t=-1} \\
&= \sum_{r=1}^{\infty}\sum_{m=1}^{\infty}(-1)^{r-1}(r-1)!\sum_{\substack{\lambda\vdash j\\ |\lambda|=r}}\lambda_m\log(m!)\prod_{k\geq 1}\frac{1}{\lambda_k!k!^{\lambda_k}} \\
&= \sum_{r=1}^{\infty}\sum_{m=1}^{\infty}(-1)^{r-1}(r-1)!\log(m!)[y^{j-m}]\frac{(e^y-1)^{r-1}}{(r-1)!m!} \\
&= \sum_{m=1}^{\infty}\frac{\log(m!)}{m!}[y^{j-m}]e^{-y} = \sum_{m=1}^{j}\frac{\log(m!)}{m!}\frac{(-1)^{j-m}}{(j-m)!} \\
&= \frac{1}{j!}\sum_{m=1}^{j}(-1)^{j-m}\binom{j}{m}\log(m!) = \frac{1}{j!}\sum_{m=1}^{j}(-1)^{j-m}\binom{j-1}{m-1}\log(m).
\end{aligned}$$

Let us write $d(j)$ for this expression. Differentiating (12) with respect to t and setting $t = 0$, we get

$$\begin{aligned}
R_t(x,0) &= x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,-1)}{j}R(x^j,1-j)\right) \\
&\quad\times\sum_{j=1}^{\infty}\left(c(j,-1)R_t(x^j,1-j) + \frac{d}{dt}\frac{c(j,t)}{j}\Big|_{t=-1}R(x^j,1-j)\right) \\
&= R(x,0)\Big(R_t(x,0) + \sum_{j=1}^{\infty}d(j)R(x^j,1-j)\Big).
\end{aligned}$$

This can be solved for $R_t(x,0)$:

$$R_t(x,0) = \frac{R(x,0)}{1-R(x,0)}\sum_{j=2}^{\infty}d(j)R(x^j,1-j).$$

Here, we are using the fact that $d(1) = 0$. Now note that $d(j)$ rapidly goes to 0 due to the factor $j!$ in the denominator and that the functions $R(x^j,1-j)$ are all analytic in a larger region than $R(x,0)$. Therefore, we can directly apply singularity analysis,

based on the well-known singular expansion

$$R(x,0) = 1 - \sqrt{2(1-ex)} + \cdots$$

of $R(x,0)$ at its singularity $\frac{1}{e}$, which yields

$$R_t(x,0) \sim \frac{1}{\sqrt{2(1-ex)}} \sum_{j=2}^{\infty} d(j) R(e^{-j}, 1-j).$$

The infinite series converges rapidly, allowing for a fairly accurate numerical computation. The mean constant μ in this special case is found to be $\mu = 0.0522901\ldots$, and similar calculations for the second derivative yield the variance constant $\sigma^2 = 0.0394984\ldots$.

4 The Automorphism Group of Pólya Trees

Since Theorem 4 is not available for Pólya trees, we want to prove asymptotic normality by using generating functions and singularity analysis. Recall that we defined the bivariate generating function $P(x,t) = \sum_{T\in\mathcal{P}} e^{t \log |\operatorname{Aut} T|} x^{|T|}$. We now let $\mathcal{B}(T)$ denote the set of root branches of a particular tree, and $\mathcal{B}_I(T)$ denote the set of unique root branches up to isomorphism. Observe that for Pólya trees there is exactly one tree in every isomorphism class so it will not be necessary to introduce separate notation for such classes.

By considering only the terms corresponding to the star on n vertices, for each n, we obtain

$$\sum_n (n-1)!^t x^n.$$

This is not analytic for any choice of $t > 0$ and, thus, neither is the original generating function. This is the main obstacle in proving asymptotic normality. To circumvent this problem, we will introduce a cut-off, ignoring the contribution of highly symmetric vertices. This is similar to the proof of Theorem 4 in [20], but there the cut-off is in terms of the size of the tree instead of symmetric vertices. We can then use the following approximation result to extend the result from the cut-off random variables to the full additive functional.

Lemma 3 *Let $(X_n)_{n\geq 1}$ and $(W_{n,N})_{n,N\geq 1}$ be sequences of centered random variables. If we have*

1. $W_{n,N} \xrightarrow[n]{d} W_N$ *and* $W_N \xrightarrow{d} W$ *for some random variables* $W, W_1, W_2, \ldots$, *and*
2. $\operatorname{Var}(X_n - W_{n,N}) \underset{N}{\longrightarrow} 0$ *uniformly in* n,

then $X_n \xrightarrow{d} W$.

This result follows e.g. from [12, Theorem 4.28]. We will apply Lemma 3 to variables X_n defined by

$$\frac{\log|\operatorname{Aut}\mathcal{P}_n| - \mathbb{E}(\log|\operatorname{Aut}\mathcal{P}_n|)}{\sqrt{n}},$$

and $W_{n,N}$ being the, similarly normalized, random variable for the additive functional $F^{\leq N}(T)$, defined by having the toll function:

$$f^{\leq N}(T) = \sum_{B\in\mathcal{B}_I(T)} I(\operatorname{mult}(B) \leq N)\log(\operatorname{mult}(B)!).$$

We note that $F(T) - F^{\leq N}(T) = F^{>N}(T)$ for an additive functional defined by

$$f^{>N}(T) = \sum_{B\in\mathcal{B}_I(T)} I(\operatorname{mult}(B) > N)\log(\operatorname{mult}(B)!),$$

so that we will, in fact, be interested in $\operatorname{Var}(F^{>N}(T_n))$ for the second condition of Lemma 3. By straightforward modifications of (7), we can define the generating functions

$$P^{\leq N}(x,t) = \sum_{T\in\mathcal{P}} e^{tF^{\leq N}(T)}x^{|T|}$$

and

$$P^{>N}(x,t) = \sum_{T\in\mathcal{P}} e^{tF^{>N}(T)}x^{|T|}$$

for the corresponding cut-off functionals.

4.1 Mean and Variance

We can now derive moments for the additive functionals F, $F^{\leq N}$, $F^{>N}$ with the help of generating functions and singularity analysis. The calculations are essentially the same in all cases so, to simplify the exposition, we perform them only for F and indicate in the end how the results differ.

Due to general principles of generating functions, studying the mean and variance corresponds to studying $P_t(x,0)$ and $P_{tt}(x,0)$. According to calculations for general additive functionals from [22], we can write

$$P_t(x,0) = xP_x(x,0)\frac{\sum_T f(T)x^{|T|} + P(x,0)\sum_{k\geq 2}P_t(x^k,0)}{P(x,0)(1+\sum_{k\geq 2}x^kP_x(x^k,0))}, \tag{13}$$

and

$$P_{tt}(x,0) = \frac{xP_x(x,0)}{P(x,0)(1+\sum_{k\geq 2} x^k P_x(x^k,0))}\Bigg(P(x,0)\Big(\sum_{k\geq 1} P_t(x^k,0)\Big)^2 + P(x,0)\sum_{k\geq 2} kP_{tt}(x^k,0) + \sum_T x^{|T|} f(T)(2F(T)-f(T))\Bigg), \quad (14)$$

for the first and second derivative. To perform singularity analysis, we must first find singular expansions for these expressions. To this end, we study the sums involved in them separately.

Recall that $\rho_p = 0.33832\ldots$ is the dominant singularity of $P(x) = P(x,0)$. Using the facts that $\rho_p < 1$ so that $\rho_p^m < \rho_p$ for $m \geq 2$ and that $\log|\operatorname{Aut} T| = O(|T|\log|T|)$, we see that the derivatives involving higher powers of x are analytic in a larger region than $P(x,0)$. Now, note that we can rewrite

$$2F(T) - f(T) = 2\sum_{B\in\mathcal{B}(T)} F(B) + f(T),$$

so that it is enough to study $\sum x^{|T|} f(T) \sum F(B)$ and $\sum x^{|T|} f(T)^2$, as well as $\sum x^{|T|} f(T)$. We will now show that we can factor each of these expressions as $P(x,0)$ times some function that is analytic in a larger radius than ρ_p. For the sum in the expression for the mean, we have

$$\begin{aligned}
&\sum_T x^{|T|} f(T) \\
&= \sum_T x^{|T|} \sum_{B\in\mathcal{B}_I(T)} \log(\operatorname{mult}(B)!) = \sum_{B\in\mathcal{P}} \sum_{m=1}^{\infty} \log(m!) \sum_{T:\operatorname{mult}(B)=m} x^{|T|} \\
&= \sum_{B\in\mathcal{P}} \sum_{m=1}^{\infty} \log(m!) x^{m|B|} (P(x,0) - x^{|B|}P(x,0)) \\
&= P(x) \sum_{B\in\mathcal{P}} \sum_{m=1}^{\infty} \log(m!) x^{m|B|}(1-x^{|B|}) = P(x) \sum_{B\in\mathcal{P}} \sum_{m=2}^{\infty} \log(m) x^{m|B|},
\end{aligned}$$

where we note that $P(x,0) - x^{|B|}P(x,0)$ equals the generating function for Pólya trees without B as a root branch. By taking absolute values, we can now bound

$$\sum_{B\in\mathcal{P}} \sum_{m=2}^{\infty} \log(m)|x|^{m|B|} = O\Big(\sum_{B\in\mathcal{P}} |x|^{2|B|}\Big),$$

as long as $|x| < 1$. The extra power of 2 means that the sum converges for $|x| < \sqrt{\rho_p} < 1$, so by the Weierstrass M-test, we have analyticity in a larger region than for the original generating function $P(x)$.

For the sum involving $\sum F(B)$, we have

$$\sum_T x^{|T|}\left(\sum_{B\in\mathcal{B}_I(T)}\log(\mathrm{mult}(B)!)\right)\left(\sum_{B\in\mathcal{B}_I(T)}\mathrm{mult}(B)F(B)\right)$$
$$=\sum_{B\in\mathcal{P}}F(B)\sum_{m=1}^{\infty}m\log(m!)\sum_{T:\mathrm{mult}(B)=m}x^{|T|}$$
$$+\sum_{\substack{B_1,B_2\in\mathcal{P}:\,m_1,m_2\geq 1\\ B_1\neq B_2}}\log(m_1!)m_2F(B_2)\sum_{\substack{T:\mathrm{mult}(B_1)=m_1\\ \mathrm{mult}(B_2)=m_2}}x^{|T|}.$$

Using the fact that $\sum_B F(B)x^{m|B|}=P_t(x^m,0)$ and performing calculations similar to above, the first sum can be seen to be

$$P(x,0)\sum_{m=2}^{\infty}\log(m!m^{m-1})P_t(x^m,0),$$

where the sum is analytic in a larger region than the original function. To deal with the other sum, we first rewrite

$$\sum_{\substack{T:\mathrm{mult}(B_1)=m_1\\ \mathrm{mult}(B_2)=m_2}}x^{|T|}=P(x,0)x^{m_1|B_1|}(1-x^{|B_1|})x^{m_2|B_2|}(1-x^{|B_2|}).$$

Then, we note that

$$\sum_{\substack{B_1:\\ B_1\neq B_2}}F(B_1)\sum_{m_1=1}^{\infty}m_1x^{m_1|B_1|}(1-x^{|B_1|})$$
$$=\sum_{m_1=1}^{\infty}\sum_{\substack{B_1:\\ B_1\neq B_2}}F(B_1)x^{m_1|B_1|}=\sum_{j=1}^{\infty}P_t(x^j,0)-\sum_{j=1}^{\infty}F(B_2)x^{j|B_2|}.$$

These observations let us rewrite the larger sum as

$$P(x,0)\left(\sum_{j=1}^{\infty}P_t(x^j,0)\right)\sum_{m=1}^{\infty}\log(m!)x^{m|B|}(1-x^{|B|})$$
$$-P(x,0)\sum_B F(B)\sum_{m=1}^{\infty}\log(m!)x^{m|B|}(1-x^{|B|})\sum_{j=1}^{\infty}x^{j|B|}.$$

The first of these two sums can now be dealt with using calculations identical to those performed earlier, and further simplifications for the second sum allow us to rewrite

the whole expression as $P(x,0)$ multiplied by

$$\left(\left(P_t(x,0)+\sum_{m=2}^{\infty}P_t(x^m,0)\right)\sum_{B}\sum_{m=2}^{\infty}\log(m)x^{m|B|}-\sum_{m=2}^{\infty}\log(m!)P_t(x^{m+1},0)\right).$$

The sum $\sum x^{|T|}f(T)^2$ can be dealt with using similar techniques and we conclude that we can rewrite (13) and (14) as

$$\begin{aligned}P_t(x,0)&=xP_x(x,0)\frac{H(x)+\sum_{k\geq 2}P_t(x^k,0)}{(1+\sum_{k\geq 2}x^kP_x(x^k,0))},\\ P_{tt}(x,0)&=\frac{xP_x(x,0)}{(1+\sum_{k\geq 2}x^kP_x(x^k,0))}\Bigg(\Bigg(P_t(x,0)+\sum_{k\geq 2}P_t(x^k,0)\Bigg)^2\\ &\quad+\sum_{k\geq 2}kP_{tt}(x^k,0)+2(P_t(x,0)H(x)+K(x))+L(x)\Bigg),\end{aligned}\tag{15}$$

for functions $H(x)$, $K(x)$ and $L(x)$ that are analytic in a larger region than $P(x,0)$.

We can now deduce singular expansions for $P_t(x,0)$ and $P_{tt}(x,0)$ by using the well known expansion of $P_x(x,0)$. The process of singularity analysis then gives asymptotic expressions for the first two raw moments after division by (the asymptotic expansion of) $[x^n]P(x,0)$. The variance is obtained by considering $\mathbb{E}(\log(|\operatorname{Aut}T|)^2)-\mathbb{E}(\log(|\operatorname{Aut}T|))^2$, where cancellation of terms of order n^2 gives that both the variance and the mean is of order n. Numerical computations yield $\mu=0.1373423\ldots$ and $\sigma^2=0.1967696\ldots$.

If we instead consider $F^{\leq N}(T)$ or $F^{>N}(T)$, the extra indicator function introduced in the expression will carry trough the calculations and affect the indices in the sums. In the sums with index m above, we will sum up to $m=N$ in the first case and sum from $m=N+1$ to infinity in the second. In the case of $F^{>N}(T)$, the corresponding analytic functions $H^{>N}(x)$, $K^{>N}(x)$ and $L^{>N}(x)$ will converge to zero within their region of convergence when $N\to\infty$.

4.2 Asymptotic Normality for $\log|\operatorname{Aut}\mathcal{P}_n|$

If we introduce a cut-off and study $F^{\leq N}$ instead of $\log|\operatorname{Aut}T|$, we can perform calculations completely analogous to the ones we did for (8) to obtain the functional equation

$$P^{\leq N}(x,t)=x\exp\left(P^{\leq N}(x,t)+\sum_{j=2}^{\infty}\frac{c_N(j,t)}{j}P^{\leq N}(x^j,jt)\right),\tag{16}$$

where we define

$$c_N(j,t) = j \sum_{\lambda \vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \left(\prod_{n=1}^{N} n!^{\lambda_n t} \right).$$

Except for the root, every vertex in the tree occurs as the child of some other vertex. This implies that it contributes to exactly one of the terms

$$I(\mathrm{mult}(B) \leq N) \log(\mathrm{mult}(B)!),$$

in the expansion of $F^{\leq N}(T)$. Thus, as a crude upper bound, each of the n vertices contributes at most $\log N!$ to the total value of the additive functional. Therefore, we see that $F^{\leq N}(T) = O(n)$ and, if we restrict to $|t| < \delta$ for some suitable $\delta > 0$,

$$G(x, y, t) := x \exp\left(y + \sum_{j=2}^{\infty} \frac{c_N(j,t)}{j} P^{\leq N}(x^j, jt) \right)$$

is analytic in a region containing $x = \rho_p$, $y = \tau$. Theorem 2.23 in [5] now gives asymptotic normality for $F^{\leq N}(T)$, i.e. $W_N \sim \mathrm{N}(0, \sigma_N^2)$ for some constant σ_N^2.

Note that

$$\mathrm{Var}(X_n - W_{n,N}) = \frac{\mathrm{Var}(F(\mathcal{P}_n) - F^{\leq N}(\mathcal{P}_n))}{n}.$$

Since $F(T) - F^{\leq N}(T) = F^{>N}(T)$, we want to show that $\mathrm{Var}(F^{>N}(T_n))/n \to 0$ when $N \to \infty$ which leads us to study $P_{tt}^{>N}(x,t)$. The reasoning from the last section shows that coefficients in Taylor expansions of $H^{>N}(x)$, $K^{>N}(x)$ and $L^{>N}(x)$ around $x = \rho_p$ go to zero as $N \to \infty$. By dominated convergence, the same is true for the expressions

$$\sum_{k \geq 2} P_t^{>N}(x^k, 0) \text{ and } \sum_{k \geq 2} k P_{tt}^{>N}(x^k, 0),$$

since all terms of $P_t^{>N}$ and $P_{tt}^{>N}$ involve powers of $F^{>N}(T)$ and these go to zero for any fixed tree as $N \to \infty$. By studying (15) (except with $P_{tt}^{>N}(x,t)$ instead of $P_{tt}(x,t)$) we see that all the coefficients in the singular expansion of $P_{tt}^{>N}(x,t)$ depend on these quantities.

We can derive a singular expansion for $P_{tt}^{>N}(x,t)$ in a way analogous to the one indicated for $P_{tt}(x,t)$ in the last subsection. This expansion must be of the type

$$a_N \left(1 - \frac{x}{\rho_p}\right)^{-3/2} + b_N \left(1 - \frac{x}{\rho_p}\right)^{-1} + c_N \left(1 - \frac{x}{\rho_p}\right)^{-1/2} + O_N(1),$$

where each coefficient, as well as the error, goes to zero with N by the reasoning above. If we now perform singularity analysis and divide by $[x^n]P(x,0)$, we get the

second raw moment for this modified functional. We can derive a singular expansion for $P_t^{>N}(x,t)$ (of leading order $-\frac{1}{2}$) which also has the property that all coefficients go to zero with N. Then, subtracting $\mathbb{E}(F^{>N}(\mathcal{P}_n))^2$ and observing cancellation of n^2-terms, we obtain

$$\operatorname{Var}(F^{>N}(\mathcal{P}_n)) = \gamma_N^2 n + O_N(1).$$

Here, γ_N goes to 0 as $N \to \infty$ since it relies on the coefficients of the singular expansions of $P_t^{>N}(x,t)$ and $P_{tt}^{>N}(x,t)$. We then divide by n to get

$$\operatorname{Var}(X_n - W_{n,N}) = \gamma_N^2 + O_N\left(\frac{1}{n}\right).$$

Note that the O-term converges to zero as $N \to \infty$ and that this convergence is uniform in n. This implies that the variance of $X_n - W_{n,N}$ goes to zero uniformly for all n so that Lemma 3 applies. Thus, we can conclude asymptotic normality for $\log|\operatorname{Aut}\mathcal{P}_n|$ from the asymptotic normality of $F^{\leq N}(\mathcal{P}_n)$ and finish the proof.

5 Automorphisms of Unrooted Trees

We show how to extend our results to unrooted versions of labeled trees and Pólya trees. Even though it is not clear what an unrooted version of a Galton–Watson tree is in general, some special cases can be dealt with using methods similar to the ones below, e.g. labeled unrooted binary trees.

5.1 Unrooted Labeled Trees

We can define unrooted labeled trees on the same probability space as rooted trees by taking a rooted tree and unrooting it. As there are exactly n unique ways of rooting any labeled tree, this gives the uniform probability measure on unrooted trees, assuming that we started with the uniform measure on rooted trees.

Now let T be a rooted tree of size n and T_v be the tree rooted at the vertex v. Note that $\operatorname{Aut}T_v$ is the stabilizer of v in $\operatorname{Aut}T$. Thus, we have

$$1 \leq \frac{|\operatorname{Aut}T|}{|\operatorname{Aut}T_v|} = |\text{Orbit of } v| \leq |T|,$$

due to the orbit-stabilizer theorem [1, Lemma 6.1]. Taking logarithms and normalizing, we find that

$$0 \leq \frac{\log|\operatorname{Aut}T| - \mu n}{\sqrt{n}} - \frac{\log|\operatorname{Aut}T_v| - \mu n}{\sqrt{n}} \leq \frac{\log n}{\sqrt{n}}, \tag{17}$$

with μ being the mean constant for rooted labeled trees from Theorem 1. If we let $X_n = \frac{\log|\operatorname{Aut}T| - \mu n}{\sqrt{n}}$ and likewise $Y_n = \frac{\log|\operatorname{Aut}T_v| - \mu n}{\sqrt{n}}$ for rooted trees, then we see

that we have almost sure convergence of $X_n - Y_n$ to 0, and thus also convergence in probability. Slutsky's theorem together with the result for rooted trees now lets us conclude that

$$X_n = X_n - Y_n + Y_n \xrightarrow{d} N(0, \sigma^2),$$

with σ^2 also coming from the theorem for rooted trees.

5.2 Unrooted Pólya Trees

We derive an analog of equation (5) that takes the size of the automorphism group into account. Let us first recall that a centroid of a tree is a vertex with the property that none of the components obtained by removing it contains more than half of the vertices. It is a classical result going back to Jordan [11] (see also e.g. [13, Ex. 6.21a]) that every tree has either a unique centroid (which we then call a central vertex) or two centroids, connected by an edge (called a central edge). Centroid vertices are also characterized by the property that the sum of the distances to all other vertices is minimized.

A central edge that connects two isomorphic trees will be called a symmetry line and the term "central edge" will be reserved for edges between centroid vertices that are not symmetry lines. The difference between the automorphisms of rooted trees compared to unrooted trees is that in the latter case any automorphism must preserve edges but not necessarily the root. We now have a bijection between Pólya trees $\mathcal{P}$ and the union of unrooted trees $\mathcal{U}$ and pairs of Pólya trees $P_1, P_2 \in \mathcal{P}$ with $P_1 \neq P_2$. Observe that for a rooted tree, there are four cases:

1. The root is a central vertex. There is a bijection from such trees to unrooted trees with a central vertex and, furthermore, any automorphism must preserve a central vertex so that the two trees have the same group of automorphisms.
2. The root is one endpoint of a symmetry line. We have a bijection between trees with a symmetry line where one of its endpoints is the root and unrooted trees with a symmetry line. We simply root the tree at one of the endpoints and note that we get the same rooted Pólya tree no matter which endpoint we choose. Any automorphism must preserve the central edge, but due to symmetry any automorphism of the rooted tree corresponds to two automorphisms of the unrooted version since we can map the endpoints of the symmetry line into each other.
3. The root is one endpoint of a central edge. First note that we have a bijection between unrooted trees with a central edge $\mathcal{U}_{ce}$ and pairs of rooted trees $\mathcal{P}^p_{ce}$ that, if joined by an edge at the roots, result in a tree with that edge as central. We now have a bijection between the union $\mathcal{U}_{ce} \cup \mathcal{P}^p_{ce}$ and rooted trees with a central edge where one of the two endpoints is the root. This can be seen by, in the former case, choosing one of the vertices of the central edge as the root. This means that we have two cases to consider: a given rooted tree corresponds to an unrooted tree with a central edge or it corresponds to a pair of rooted trees. In the first case, we note that any automorphism of an unrooted tree must preserve the central edge, and as it is not a symmetry line this implies that it must fix the root. For rooted trees in bijection with pairs of rooted trees we note that the two trees must be different but

have the same size implying that no additional symmetry can occur when joining them. In both cases, the size of the automorphism group of the rooted tree is the same size as its counterpart.

4. The root satisfies none of the above. Then one root branch contains strictly more than half of the vertices and the tree decomposes into an unordered pair of rooted trees, i.e., the large branch and the rest of the tree (including the root). As the trees have different sizes this makes the decomposition unique and the size of the automorphism group of the original tree is simply the product of the groups of the two subtrees.

Let $U_c(x,t)$ be the generating function for unrooted trees with a central vertex and $U_e(x,t)$ be the generating function for unrooted trees with a central edge or symmetry line. Combining the observations from above, and translating it to the level of generating functions, we find that

$$P(x,t) = U_c(x,t) + U_e(x,t) - 2^t P(x^2,2t) + P(x^2,2t) + \frac{1}{2}P(x,t)^2 - \frac{1}{2}P(x^2,2t)$$

where the two middle terms involving $P(x^2,2t)$ are correction terms corresponding to point 2. above and the last two terms count unordered pairs of distinct rooted trees corresponding to point 3. By noting that $U(x,t) = U_c(x,t)+U_e(x,t)$ and rearranging we get

$$U(x,t) = P(x,t) - \frac{1}{2}P(x,t)^2 + \left(2^t - \frac{1}{2}\right)P(x^2,2t)$$

which is enough to obtain moments for $\log|\operatorname{Aut} T|$ and calculations show that the mean and variance constants are the same as for rooted trees.

To extend the results for rooted trees to a full central limit theorem we use the far-reaching result in [21, Theorem 1.3]. This theorem shows that the random unrooted tree U_n on n vertices is close to a tree T_n obtained by identifying the roots of a rooted Pólya tree $\mathcal{P}_{K_n}$ (of random size K_n) and a tree B_n of stochastically bounded size $|B_n| = n - K_n + 1 = O_P(1)$. To be precise, the total variation distance between U_n and T_n is $O(e^{-cn})$ for a constant $c > 0$.

In other words, an unrooted tree essentially consists of a large rooted Pólya tree and something small. Thus, we have

$$P\left(\frac{\log|\operatorname{Aut} U_n| - \mu n}{\sqrt{n}} \le a\right) = P\left(\frac{\log|\operatorname{Aut} T_n| - \mu n}{\sqrt{n}} \le a\right) + O(e^{-cn}). \quad (18)$$

Moreover, if $|B_n| \le M$ for some fixed M, then we have $\log|\operatorname{Aut} T_n| = \log|\operatorname{Aut} \mathcal{P}_{K_n}| + O(\log n)$ by the same argument that gave us (17), and consequently

$$\begin{aligned}\frac{\log|\operatorname{Aut} T_n| - \mu n}{\sqrt{n}} &= \frac{\log|\operatorname{Aut} \mathcal{P}_{K_n}| - \mu n}{\sqrt{n}} + O\Big(\frac{\log n}{\sqrt{n}}\Big)\\ &= \frac{\log|\operatorname{Aut} \mathcal{P}_{K_n}| - \mu K_n}{\sqrt{K_n}} + O\Big(\frac{\log n}{\sqrt{n}}\Big).\end{aligned}$$

Since $|B_n| = n - K_n + 1$ is stochastically bounded, we see that

$$\frac{\log|\operatorname{Aut} T_n| - \mu n}{\sqrt{n}} - \frac{\log|\operatorname{Aut} \mathcal{P}_{K_n}| - \mu K_n}{\sqrt{K_n}} \xrightarrow{p} 0.$$

So an application of Slutsky's theorem in combination with (18) and the results for rooted Pólya trees proves the central limit theorem for the size of the automorphism group in the case of unrooted trees.

Acknowledgements This work was supported by the Knut and Alice Wallenberg Foundation. The authors would like to thank Benedikt Stufler for helpful suggestions relating to Section 5.

Funding Open access funding provided by Uppsala University.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Aigner, M.: A Course in Enumeration. Graduate Texts in Mathematics, vol. 238. Springer, Berlin (2007)
2. Aldous, D.: Brownian excursions, critical random graphs and the multiplicative coalescent. Ann. Probab. **25**(2), 812–854 (1997). https://doi.org/10.1214/aop/1024404421
3. Babai, L.: Automorphism groups, isomorphism, reconstruction. In: Handbook of Combinatorics, vol. 2, pp. 1447–1540. Elsevier, Amsterdam (1995)
4. Bóna, M., Flajolet, P.: Isomorphism and symmetries in random phylogenetic trees. J. Appl. Probab. **46**(4), 1005–1019 (2009). https://doi.org/10.1239/jap/1261670685
5. Drmota, M.: Random Trees. Springer, Vienna (2009). https://doi.org/10.1007/978-3-211-75357-6
6. Holmgren, C., Janson, S.: Limit laws for functions of fringe trees for binary search trees and random recursive trees. Electron. J. Probab. 20, no. 4, 51 pp. (2015). https://doi.org/10.1214/EJP.v20-3627
7. Holmgren, C., Janson, S., Šileikis, M.: Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. Electron. J. Comb. 24(2), Paper No. 2.51, 49 pp. (2017). https://doi.org/10.37236/6374
8. Isaev, M., Southwell, A., Zhukovskii, M.: Distribution of tree parameters by martingale approach. Comb. Probab. Comput. **31**(5), 737–764 (2022). https://doi.org/10.1017/s0963548321000523
9. Janson, S.: Random cutting and records in deterministic and random trees. Random Struct. Algorithms **29**(2), 139–179 (2006). https://doi.org/10.1002/rsa.20086
10. Janson, S.: Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. Random Struct. Algorithms **48**(1), 57–101 (2016). https://doi.org/10.1002/rsa.20568
11. Jordan, C.: Sur les assemblages de lignes. J. Reine Angew. Math. **70**, 185–190 (1869). https://doi.org/10.1515/crll.1869.70.185
12. Kallenberg, O.: Foundations of Modern Probability. Springer, New York (1997)

13. Lovász, L.: Combinatorial Problems and Exercises, 2nd edn. AMS Chelsea Publishing, Providence, RI (2007). https://doi.org/10.1090/chel/361
14. Matthews, D.: Automorphisms of random recursive trees. PhD thesis, University of Southhampton (2017). https://eprints.soton.ac.uk/415900
15. McKeon, K.A.: The expected number of symmetries in locally-restricted trees. I. In: Graph Theory, Combinatorics, and Applications. vol. 2 (Kalamazoo, MI, 1988). Wiley-Interscience Publisher, pp. 849–860. Wiley, New York (1991)
16. McKeon, K.A.: The expected number of symmetries in locally restricted trees. II. Discret. Appl. Math. **66**(3), 245–253 (1996). https://doi.org/10.1016/0166-218X(94)00164-9
17. Olsson, C.: The probability of random trees being isomorphic. arXiv (2022). https://doi.org/10.48550/ARXIV.2205.06625
18. Olsson, C., Wagner, S.: Automorphisms of Random Trees. In: Ward, M.D. (ed.) 33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022). Leibniz International Proceedings in Informatics (LIPIcs), vol. 225, pp. 16:1–16:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2022). https://doi.org/10.4230/LIPIcs.AofA.2022.16
19. Ralaivaosaona, D., Wagner, S.: A central limit theorem for additive functionals of increasing trees. Comb. Probab. Comput. **28**(4), 618–637 (2019). https://doi.org/10.1017/s0963548318000585
20. Ralaivaosaona, D., Šileikis, M., Wagner, S.: A central limit theorem for almost local additive tree functionals. Algorithmica **82**(3), 642–679 (2020). https://doi.org/10.1007/s00453-019-00622-4
21. Stufler, B.: The continuum random tree is the scaling limit of unlabeled unrooted trees. Random Struct. Algorithms **55**(2), 496–528 (2019). https://doi.org/10.1002/rsa.20833
22. Wagner, S.: Central limit theorems for additive tree parameters with small toll functions. Comb. Probab. Comput. **24**(1), 329–353 (2015). https://doi.org/10.1017/S0963548314000443
23. Yu, L.: Automorphisms of random trees. PhD thesis, Drexel University (2012). https://idea.library.drexel.edu/islandora/object/idea%3A4004/