Automorphisms of Random Trees

Christoffer Olsson
Department of Mathematics, Uppsala University, Sweden

Stephan Wagner
Department of Mathematics, Uppsala University, Sweden

Abstract

We study the size of the automorphism group of two different types of random trees: Galton–Watson trees and Pólya trees. In both cases, we prove that it asymptotically follows a log-normal distribution. 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 Pólya trees.

2012 ACM Subject Classification Mathematics of computing → Random graphs; Mathematics of computing → Generating functions

Keywords and phrases random tree, Galton–Watson tree, Pólya tree, automorphism group, central limit theorem

Digital Object Identifier 10.4230/LIPIcs.AofA.2022.16

Funding *Christoffer Olsson*: Supported by the Knut and Alice Wallenberg Foundation.
Stephan Wagner: Supported by the Knut and Alice Wallenberg Foundation.

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 order 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 group has previously been studied in special cases of Galton–Watson trees: binary trees (expected values and limiting distribution: [2]), labeled trees (limiting distribution: [6] and expected value: [16]), binary and ternary trees (expected values: [11] and [12]). It has also been studied for some other types of trees than those considered here: specifically, random recursive trees (expected value: [10]), and d-ary increasing trees (limiting distribution and moments: [13]).

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

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

derived from the fact that the automorphism group of a rooted tree is obtained from symmetric groups by iterated direct and wreath products (see [1], 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 $\log |\operatorname{Aut} T|$ is an *additive functional* of the tree, which is a real-valued function $F(T)$ that satisfies

$$F(T) = f(T) + \sum_{i=1}^{r} F(T_i),$$

33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022).
Editor: Mark Daniel Ward; Article No. 16; pp. 16:1–16:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

where we sum over the r root branches and $f(T)$ is another function called the *toll* of the additive functional. In our case, the toll function is $\sum \log(m_i!)$. We note that we can rewrite the definition as

$$F(T) = f(T) + \sum_{i=1}^{k} m_i F(T_i),$$

where we are now summing over unique root branches $T_1, T_2, \ldots, T_k$ with multiplicities $m_1, m_2, \ldots, m_k$. We will show asymptotic normality of $\log|\operatorname{Aut} \mathcal{T}_n|$, which implies asymptotic log-normality of $|\operatorname{Aut} \mathcal{T}_n|$, where $\mathcal{T}_n$ denotes a random tree on n vertices. Limit theorems for additive functionals have been proven for various classes of random trees under different conditions, see [4, 5, 8, 13–15]. We will specifically make use of a general result due to Ralaivaosaona, Šileikis and the second author [14] that is based on earlier work by Janson [8].

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 zero and some number greater than one. 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 (indeed, we can see them as two sides of the same coin, one being probabilistic and the other being combinatorial, see [3, Section 1.2.7]). Examples of Galton–Watson (and simply generated) trees are plane trees, labeled trees, d-ary trees, etc. The book [3] gives a general introduction to different types of random trees.

Pólya trees are rooted, unordered, unlabeled trees. They are not Galton–Watson trees, even though they have many similar properties, and cannot be interpreted as growth processes so we will need other methods to deal with them. The trees can be characterized by their generating function $P(x) = \sum_{T \in \mathcal{P}} x^{|T|}$, which satisfies

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

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 Pólya trees and Pólya trees of size n, respectively.

1.1 Results

In this paper, 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 order 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. In the appendix, we show how to derive the numerical values $\mu = 0.0522901\ldots$ and $\sigma^2 = 0.0394984\ldots$ in the special case of labeled trees, where a connection to Pólya trees can be exploited.

We can also prove the following, similar result, for the class of Pólya trees.

▶ **Theorem 2.** *Let $\mathcal{P}_n$ be a Pólya tree of order n. 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*

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

Even though both proofs rely, at their cores, on the same idea of approximating the additive functionals by simpler ones, they are fairly different at a glance. We prove Theorem 1 in Section 2 and Theorem 2 in Section 3.

2 The automorphism group of Galton–Watson trees

To prove asymptotic normality of $\log|\operatorname{Aut}\mathcal{T}_n|$, we will show that it is in fact an almost local additive functional, as defined in [14]. 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 3** ([14]). *Let $\mathcal{T}_n$ be a conditioned Galton–Watson tree of order 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,

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

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 $(F(\mathcal{T}_n) - \mathbb{E}\,F(\mathcal{T}_n))/\sqrt{n}$.

2.1 Galton–Watson trees isomorphic up to a certain level

In applying Theorem 3, 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 4.** *There exists some constant* $0 < c < 1$ *such that*

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

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 right 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

$$\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \leq \prod_{i=0}^{M-1} \mathbb{P}(L_i \text{ agrees with } C | L_1, L_2, \ldots, L_{i-1}), \tag{2}$$

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 (2) 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 $X_l \sim \mathrm{Bin}(l,p)$ 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 $X_l = Y_1 + Y_2 + \ldots + Y_l$ we have

$$\begin{aligned}\mathbb{P}(Y_1 + Y_2 + \ldots + Y_l = k) &= \sum_r \mathbb{P}(Y_1 + Y_2 + \ldots + Y_{l-1} = r)\,\mathbb{P}(Y_l = k - r)\\ &\leq \sum_r \mathbb{P}(Y_1 + Y_2 + \ldots + Y_{l-1} = r) \max_y \mathbb{P}(Y_l = y) = \max_y \mathbb{P}(Y_l = y) = \max\{p, 1-p\}.\end{aligned}$$

We can thus take $c = \max\{p, 1-p\}$ as a uniform bound for all levels, and now (2) gives the result. ◀

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

$$\mathbb{P}(\mathcal{T}_1^{(M)}, \mathcal{T}_2^{(M)} \text{ isomorphic and of height} \geq M) = \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)^2 \tag{3}$$

$$\leq \max_{C \in \mathcal{C}^M} \{\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)\} \sum_{C \in \mathcal{C}^M} \mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C) \tag{4}$$

$$= \max_{C \in \mathcal{C}^M} \{\mathbb{P}(\mathcal{T}^{(M)} \text{ belongs to } C)\}. \tag{5}$$

Combining this with Lemma 4, we get the following corollary.

▶ **Corollary 5.** *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 (3) 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 6.** *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 4 to the conditioned case.

▶ **Lemma 7.** *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 $C \in \mathcal{C}_n^M$.

The proof uses breadth-first exploration and the cycle lemma, a standard trick in the field, and is deferred to the appendix. Furthermore, using calculations similar to (3), we obtain the following corollary.

▶ **Corollary 8.** *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*

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

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

2.2 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 3 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 for the conditioned Galton–Watson tree in the following way.

$$\begin{aligned}
\mathbb{E}\,|f(\mathcal{T}_n) - f(\mathcal{T}_n^{(M)})| &\leq \mathbb{E}\left(\sum_{\substack{\mathcal{T}_i, \mathcal{T}_j \\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M) \right) \\
&= \sum_{k \geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \sum_{n_1, n_2 \geq 1} \mathbb{P}(|\mathcal{T}_i| = n_1 | \deg(\mathcal{T}_n) = k)\, \mathbb{P}(|\mathcal{T}_j| = n_2 | \deg(\mathcal{T}_n) = k) \\
&\qquad \cdot \binom{k}{2} \mathbb{E}\left(I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M) \middle| |\mathcal{T}_i| = n_1, |\mathcal{T}_j| = n_2 \right) \\
&= O\left(\sum_{k \geq 2} \mathbb{P}(\deg(\mathcal{T}_n) = k) \binom{k}{2} n^{\frac{5}{2}} c^M \right) = O\left(n^{\frac{5}{2}} c^M \sum_{k \geq 2} k\, \mathbb{P}(\xi = k) \binom{k}{2} \right) = O(n^{\frac{5}{2}} c^M)
\end{aligned}$$

where we use the law of total expectation, the fact that $\mathbb{P}(\deg(\mathcal{T}_n) = k) \leq ck\,\mathbb{P}(\xi = k)$ for all k and n, where c is constant [7, (2.7)], and the assumption on moments of the offspring distribution.

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 bound its expectation in the following way.

$$\begin{aligned}
&\mathbb{E}\left| f(\hat{\mathcal{T}}^{(M)}) - \mathbb{E}\left(f(\hat{\mathcal{T}}^{(N)}) | \hat{\mathcal{T}}^{(M)} \right) \right| \\
&= \sum_{k \geq 2} kP(\xi = k) \cdot \Bigg(\mathbb{E}\Big(\sum_{\substack{\mathcal{T}_i, \mathcal{T}_j \text{ non-special} \\ \text{root branches}}} I(\mathcal{T}_i^{(M)}, \mathcal{T}_j^{(M)} \text{ are iso. with height} \geq M) \Big) \\
&\qquad\qquad + \mathbb{E}\Big(\sum_{\substack{\mathcal{T} \text{ non-special root branch} \\ \hat{\mathcal{T}} \text{ special root branch}}} I(\mathcal{T}^{(M)}, \hat{\mathcal{T}}^{(M)} \text{ are iso. with height} \geq M) \Big) \Bigg).
\end{aligned}$$

Furthermore, this is equal to

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

by Corollaries 5 and 6 (the constant c is the same for both of these corollaries since they both rely on Lemma 4) 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 $n^{5/2}c^{M_n} \le c_1^{M_n}$ for all n and $A \log c_1 < -3/2$. Then, the expectations mentioned in Theorem 3 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 3 to show that $\log |\operatorname{Aut} \mathcal{T}_n|$ is asymptotically normal, which completes the proof of Theorem 1.

3 The automorphism group of Pólya trees

Since Theorem 3 is not available for Pólya trees, we want to prove asymptotic normality by using generating functions and singularity analysis. Thus, we define the generating function of $F(\mathcal{P}_n) = \log |\operatorname{Aut} \mathcal{P}_n|$ to be

$$P(x,t) = \sum_{T \in \mathcal{P}} e^{t \log |\operatorname{Aut} T|} x^{|T|}. \tag{6}$$

Note that $P(x, 0) = P(x)$, as defined in the introduction. We will use $\rho = 0.33832\dots$ to denote the dominant singularity of the generating function and recall that $P(\rho) = 1$ (see [3, Remark 3.9]). 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. Furthermore, we let $\operatorname{mult}(B)$ be the number of occurrences as root branches of a particular tree B. 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, in [14], of Theorem 3, 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 9.** *Let $(X_n)_{n\ge1}$ and $(W_{n,N})_{n,N\ge1}$ be sequences of centered random variables. If we have*

1. $W_{n,N} \xrightarrow{d}_n W_N$ *and* $W_N \xrightarrow{d} W$ *for some random variables* $W, W_1, W_2, \dots$, *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 [9], Theorem 4.28. We will apply Lemma 9 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 $F^{\le N}(T)$, defined by having the toll function:

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

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

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

so that we will, in fact, be interested in $\mathrm{Var}(F^{>N}(T_n))$ for the second condition of Lemma 9. By straightforward modifications of (6), we can define generating functions $P^{\le N}(x,t)$ and $P^{>N}(x,t)$ for the corresponding cut-off functionals.

3.1 Mean and variance

We can now derive moments for the additive functionals $F, F^{\le N}, F^{>N}$ 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 [15], we can write

$$P_t(x,0) = xP_x(x,0) \frac{\sum_T f(T) x^{|T|} + P(x,0) \sum_{k \ge 2} P_t(x^k,0)}{P(x,0)(1 + \sum_{k \ge 2} x^k P_x(x^k,0))}, \tag{7}$$

$$P_{tt}(x,0) = \frac{xP_x(x,0)}{P(x,0)(1 + \sum_{k \ge 2} x^k P_x(x^k,0))} \Bigg(P(x,0) \Big(P_t(x,0) + \sum_{k \ge 2} \frac{P_t(x^k,0)}{k} \Big)^2 \tag{8}$$

$$+ P(x,0) \sum_{k \ge 2} \frac{P_{tt}(x^k,0)}{k} + \sum_T x^{|T|} f(T)(2F(T) - f(T)) \Bigg), \tag{9}$$

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.

The derivatives involving higher powers of x are analytic in a larger region than $P(x,0)$, since $\rho < 1$ so that $\rho^m < \rho$ for $m \ge 2$. Now, note that we can rewrite

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

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

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

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. For real positive x with $x \le 1 - \epsilon$ for fixed $\epsilon > 0$, we can bound

$$\sum_B \sum_{m=2}^{\infty} \log(m) x^{m|B|} = O(\sum_B x^{2|B|}).$$

The extra power of 2 means that the sum converges for $x < \sqrt{\rho}$, 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

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

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

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

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 let us rewrite the whole expression as

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

The sum $\sum x^{|T|} f(T)^2$ can be dealt with using similar techniques and we conclude that we can rewrite (7) as

$$
\begin{aligned}
P_t(x,0) &= xP_x(x,0)\frac{H(x)+\sum_{k\geq 2} P_t(x^k,0)}{(1+\sum_{k\geq 2} x^k P_x(x^k,0))},\\
P_{tt}(x,0) &= \frac{xP_x(x,0)}{(1+\sum_{k\geq 2} x^k P_x(x^k,0))}\Bigg(\Big(P_t(x,0)+\sum_{k\geq 2}\frac{P_t(x^k,0)}{k}\Big)^2\\
&\qquad + \sum_{k\geq 2}\frac{P_{tt}(x^k,0)}{k} + 2(P_t(x,0)H(x)+K(x))+L(x)\Bigg),
\end{aligned} \tag{10}
$$

for functions $H(x)$, $K(x)$ and $L(x)$ that are analytic in a larger region than $P(x,0)$. This puts us in a situation where we can perform singularity analysis to find the moments. 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 particular, for $F^{>N}(T)$, the corresponding analytic functions $H^{>N}(x)$,$K^{>N}(x)$ and $L^{>N}(x)$ will converge to zero within their region of convergence, if we let $N \to \infty$.

3.2 Asymptotic normality for $\log|\operatorname{Aut}\mathcal{P}_n|$

For Pólya trees we have the symbolic decomposition

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

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. For the case of the cut-off functional $F^{\leq N}(T)$, which is the case we will be interested in, we have

$$
P^{\leq N}(x,t) = x\prod_{T\in\mathcal{P}}\left(\sum_{n=0}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right).
$$

We can manipulate this as follows:

$$
\begin{aligned}
P^{\leq N}(x,t) &= x\exp\left(\sum_{T\in\mathcal{P}}\log\left(\sum_{n=0}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right)\right)\\
&= x\exp\left(\sum_{T\in\mathcal{P}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\left(\sum_{n=1}^{\infty} x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\right)^k\right)\\
&= x\exp\Big(\sum_{T\in\mathcal{P}}\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\sum_{\substack{\lambda_1+\lambda_2\\+\ldots=k}}\binom{k}{\lambda_1,\lambda_2,\ldots}\prod_{n=1}^{\infty}\big(x^{n|T|} n!^{tI(n\leq N)} e^{ntF^{\leq N}(T)}\big)^{\lambda_n}\Big).
\end{aligned} \tag{11}
$$

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 $|\lambda| = \lambda_1 + \lambda_2 + \ldots$, and we write $\lambda \vdash j$ to denote that λ is a partition of j, i.e. $j = \lambda_1 + 2\lambda_2 + 3\lambda_3 + \ldots$. We can now rearrange the terms in the exponent of (11) to get

$$x \exp\left(\sum_{T\in\mathcal{P}} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum_{j=1}^{\infty} \sum_{\substack{\lambda_1+\lambda_2+\ldots=k \\ \lambda_1+2\lambda_2+\ldots=j}} \binom{k}{\lambda_1, \lambda_2, \ldots} x^{j|T|} e^{jtF^{\le N}(T)} \prod_{n=1}^{\infty} n!^{\lambda_n t I(n\le N)}\right)$$

$$= x \exp\left(\sum_{j=1}^{\infty} \sum_{\lambda\vdash j} \frac{(-1)^{|\lambda|-1}}{|\lambda|} \binom{|\lambda|}{\lambda_1, \lambda_2, \ldots} \left(\prod_{n=1}^{N} n!^{\lambda_n t}\right) \sum_{T\in\mathcal{P}} x^{j|T|} e^{jtF^{\le N}(T)}\right)$$

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

For convenience, we can define

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

and arrive at the functional equation

$$P^{\le N}(x,t) = x \exp\left(P^{\le N}(x,t) + \sum_{j=2}^{\infty} \frac{c_N(j,t)}{j} P^{\le N}(x^j, jt)\right). \tag{12}$$

Note that $c_N(j, 0) = 1$, so that we recover the functional equation (1) from the introduction if we set $t = 0$. We can make completely analogous calculations for $P(x,t)$ to obtain a functional equation for the original functional $\log|\operatorname{Aut}\mathcal{P}_n|$, as well, but recall that $P(x,t)$ is not analytic for $t > 0$.

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 $F^{\le N}(T) = O(n)$ and, if we restrict to $|t| < \delta$ for some suitable $\delta > 0$,

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

is analytic in a region containing $x = \rho$, $y = \tau$. Theorem 2.23 in [3] now gives asymptotic normality for $F^{\le N}(T)$, i.e. $W_N \sim \mathrm{N}(0, \sigma_N^2)$ for some constant σ_N^2.

Note that

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

Since $F(T) - F^{\le N}(T) = F^{>N}(T)$, we want to show that $\operatorname{Var}(F^{>N}(T_n))/n \to 0$ 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 $x = \rho$ go to zero as $N \to \infty$. By dominated convergence, the same is true for the expressions

$$\sum \frac{P_t(x^k, 0)}{k} \text{ and } \sum \frac{P_{tt}(x^k, 0)}{k},$$

since all terms of P_t and P_{tt} involve powers of $F^{>N}(T)$ and this goes to zero for any fixed tree as $N \to \infty$. By studying (10) (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. Therefore, the expansion must be of the type

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

where each coefficient, as well as the error, goes to zero with N.

Performing singularity analysis, where we also subtract $\mathbb{E}(F^{>N}(\mathcal{P}_n))^2$ to get the variance, and dividing by n, gives us that

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

for some constant γ_N that goes to 0 as $N \to \infty$. Moreover, the O-term is uniform in N. This implies that the variance of $X_n - W_{n,N}$ goes to zero, uniformly in n so that the approximation lemma 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.

References

1 László Babai. Automorphism groups, isomorphism, reconstruction. In *Handbook of combinatorics, Vol. 1, 2*, pages 1447–1540. Elsevier Sci. B. V., Amsterdam, 1995.

2 Miklós Bóna and Philippe Flajolet. Isomorphism and symmetries in random phylogenetic trees. *J. Appl. Probab.*, 46(4):1005–1019, 2009. doi:10.1239/jap/1261670685.

3 Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability. doi:10.1007/978-3-211-75357-6.

4 Cecilia Holmgren and Svante Janson. Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Probab.*, 20:no. 4, 51, 2015. doi:10.1214/EJP.v20-3627.

5 Cecilia Holmgren, Svante Janson, and Matas Šileikis. Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *Electron. J. Combin.*, 24(2):Paper No. 2.51, 49, 2017. doi:10.37236/6374.

6 Mikhail Isaev, Angus Southwell, and Maksim Zhukovskii. Distribution of tree parameters by martingale approach, 2019. arXiv:1912.09838.

7 Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures & Algorithms*, 29(2):139–179, 2006. doi:10.1002/rsa.20086.

8 Svante Janson. Asymptotic normality of fringe subtrees and additive functionals in conditioned Galton-Watson trees. *Random Structures Algorithms*, 48(1):57–101, 2016. doi:10.1002/rsa.20568.

9 Olav Kallenberg. *Foundations of modern probability*. Springer, New York, 1997.

10 David Matthews. *Automorphisms of Random Recursive Trees*. PhD thesis, University of Southhampton, 2017. URL: https://eprints.soton.ac.uk/415900.

11 Kathleen A. McKeon. The expected number of symmetries in locally-restricted trees. I. In *Graph theory, combinatorics, and applications. Vol. 2 (Kalamazoo, MI, 1988)*, Wiley-Intersci. Publ., pages 849–860. Wiley, New York, 1991.

12 Kathleen A. McKeon. The expected number of symmetries in locally restricted trees. II. *Discrete Appl. Math.*, 66(3):245–253, 1996. doi:10.1016/0166-218X(94)00164-9.

13 Dimbinaina Ralaivaosaona and Stephan Wagner. A central limit theorem for additive functionals of increasing trees. *Combin. Probab. Comput.*, 28(4):618–637, 2019. doi:10.1017/s0963548318000585.

14 Dimbinaina Ralaivaosaona, Matas Šileikis, and Stephan Wagner. A central limit theorem for almost local additive tree functionals. *Algorithmica*, 82(3):642–679, 2020. doi:10.1007/s00453-019-00622-4.
15 Stephan Wagner. Central limit theorems for additive tree parameters with small toll functions. *Combinatorics, Probability and Computing*, 24(1):329–353, 2015. doi:10.1017/S0963548314000443.
16 Le Yu. *Automorphisms of random trees.* PhD thesis, Drexel University, 2012. URL: https://researchdiscovery.drexel.edu/esploro/outputs/doctoral/Automorphisms-of-random-trees/991014632289304721.

A Omitted proofs and calculations

A.1 Proof of Lemma 7

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

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

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

$$\begin{aligned} S_m &> -1 \quad \text{for } 1 \leq m < n, \\ S_n &= -1, \end{aligned}$$

since we are adding 1 for all vertices except the root, but deleting 1 for all vertices including the root. Using this, we can formulate the probability we seek to bound in the following way.

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

where T' is a Galton–Watson tree with offsprings $\xi_1, \xi_2, \dots, \xi_k$, 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 on this level 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

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

By independence, we now have

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

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 4 (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.

A.2 Mean and variance for labeled trees

In this appendix, we 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 generating function approach from Section 3. Recall that the bivariate generating function

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

satisfies (letting $N\to\infty$ in (12))

$$P(x,t)=x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,t)}{j}P(x^j,jt)\right)$$

with

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

We can rewrite this 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:

$$R(x,t)=x\exp\left(\sum_{j=1}^{\infty}\frac{c(j,t-1)}{j}R(x^j,jt-j+1)\right). \tag{13}$$

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, note first that

$$\sum_{j\geq 0} y^j \sum_{\lambda\vdash j} \prod_{k\geq 1} \frac{x_k^{\lambda_k}}{\lambda_k! k!^{\lambda_k}} = \prod_{k\geq 1} \sum_{\lambda_k\geq 0} \frac{x_k^{\lambda_k} y^{k\lambda_k}}{\lambda_k! k!^{\lambda_k}} = \prod_{k\geq 1} \exp\left(\frac{x_k y^k}{k!}\right) = \exp\left(\sum_{k\geq 1} \frac{x_k y^k}{k!}\right).$$

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

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

Consequently,

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

By definition, we have

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

which therefore becomes

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

Let us write $d(j)$ for this expression. Differentiating (13) with respect to t and setting $t = 0$, we get

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

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 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 $\sigma^2 = 0.0394984\ldots$.