ISOLATING A LEAF IN ROOTED TREES VIA RANDOM CUTTINGS

MARKUS KUBA AND ALOIS PANHOLZER

Abstract. We consider a recursive procedure for destroying rooted trees and isolating a leaf by removing a random edge and keeping the subtree, which does not contain the original root. For two tree families, the simply generated tree families and increasing tree families, we study here the number of random cuts that are necessary to isolate a leaf. We can show limiting distribution results of this parameter for simply generated trees and certain increasing trees.

1. Introduction

We consider the following edge-removal procedure in a size n rooted tree for isolating a leaf. Pick one of the $n-1$ edges of the tree at random and remove it. This separates the tree into a pair of rooted trees; the tree containing the root of the original tree retains its root while the tree not containing the root of the original tree is rooted at the vertex adjacent to the edge that was cut. Now we discard the subtree containing the original root and continue this procedure in the other subtree, until we end at a size 1 subtree, which contains a leaf. We are going to study for several tree families under the random tree model a random variable Z_n, which counts here the number of edges that will be removed from a randomly chosen tree of size n by this edge-removal procedure until a leaf is isolated. Since all the analyzed tree families can be considered as weighted trees, this means that for starting the edge-removal procedure we choose a tree of size n with probability proportional to its weight. We can give limiting distribution results of Z_n for general simply generated tree families and some classes of so called *increasing tree families*, which contain recursive trees as a special instance (see Subsection 2.2). Surprisingly the multiple zeta function and it's finite counterpart show up in the limit distribution for certain increasing trees.

This edge-removal procedure is the counterpart of another edge-removal procedure, which was already studied in more detail. In the latter procedure the subtree containing the original root of the tree is kept, while the other subtree is discarded (thus it can be seen as the conjugate version of the procedure studied in the present paper) and then the procedure is continued recursively on the subtree containing the root until the original root is isolated.

Best to our knowledge the procedure studied here and the results are new, but for the sake of completeness we collect in the following some known results for the conjugate procedure, which isolates the root. Meir and Moon [17, 18] considered this edge-removal procedure (= cutting-down procedure or root-isolation procedure) on a rooted tree with n vertices. In papers [17, 18] a random variable X_n was studied. This variable counts the number of edges that will be removed from a randomly chosen tree of size n until the root is isolated for the two important tree families *unordered labelled trees* (= Cayley trees) and *recursive trees*. As the underlying model of randomness it is always assumed the random tree model. For both tree families they obtained exact and asymptotic formulæ for the expectation $\mathbb{E}(X_n)$ and also asymptotic formulæ resp. bounds for the second moment $\mathbb{E}(X_n^2)$. Limiting distribution results of X_n for some classes of so called *simply generated tree families*, which contain Cayley trees as a special instance (see Subsection 2.1), are given in [24]. The problem for so called non-crossing trees was considered in [21]. The results are extended to general simply generated tree families in [12]: it turns out that for these tree families (after a suitable normalization) X_n is asymptotically Rayleigh distributed. Finally the limiting distribution of X_n has been characterized for recursive trees in [5], where it turns our that (again after a suitable normalization) X_n converges to a spectrally negative stable distribution.

Date: April 22, 2007.

This work was supported by the Austrian Science Foundation FWF, grant S9608-N13.

We want to mention that also the following two-sided variant of the edge-removal procedure was considered in recent papers: after removing the randomly chosen edge one continues the procedure recursively on both of the obtained subtrees. Of course, when starting with a tree of size n, this two-sided variant leads to n isolated nodes after $n-1$ cuts, but one was interested in the total costs when isolating all nodes in the tree, if one assumes that the cost incurred for selecting an edge and splitting the tree is given by a toll function t_n. For toll functions $t_n = n^{\alpha}$ with $\alpha > 0$, asymptotic results for all moments are obtained in [22] and limiting distribution results for some classes of simply generated tree families are given in [7]. For Cayley trees this procedure is equivalent to a probabilistic model involved in the Union–Find (or equivalence-finding) algorithm, which was analyzed first by Knuth and Schönhage [14].

The random variable Z_n for the leaf-isolation procedure can structurally be viewed as a certain kind of depth-measure and thus the analysis presented can be considered as efforts leading to a deeper understanding of the shape of random trees. We want to describe briefly two common depth-measures for random rooted tree families and compare the asymptotic behaviour with Z_n, namely the depth of a random node in a random rooted tree and the climbing depth of a random rooted tree. The depth of a node v in a rooted tree is given by the distance (measured by the number of edges) between the root node and node v. The depth of a random node in a random rooted tree is a fundamental shape parameter of a random tree family and has been studied extensively, see, e.g., [15, 16, 19] for limiting distribution results for tree families considered in the present paper. The climbing depth of a rooted tree is given by the number of steps to reach a leaf in T, when starting at the root node and choosing at every step one of the out-going edges of the present node at random and proceeding along this edge to a neighbouring node. The climbing depth of a random tree has been studied also for several families of random trees, see, e.g., [23] and references therein for limiting distribution results for tree families considered here. In connection with a digital data structure called Tries, the climbing depth also appears in the analysis of certain randomized leader election algorithms, see [8, 13, 25]. A comparison of the main behaviour of the different depth-measures mentioned for the tree families considered in this paper is given in Table 1, where it is apparent that the behaviour of Z_n is in general different from the other depth parameters.

Another aspect of this study is to obtain further results for a certain kind of divide-and-conquer recurrences (8), which, for different kind of transition probabilities $q_{n,k}$, occur frequently, e.g., in the analysis of some algorithms and in the context of models for the coalescence of particles, and has been considered by several authors, see, e.g., [6, 20, 26].

Table 1. A comparison of the limiting behaviour of different depth-measures for several random tree families stating the order of growth of the expectation and the limiting distribution.

Tree family	Leaf isolation	Root isolation	Climbing depth	Depth
Simply generated trees	Poisson	Rayleigh	Negative binom.	Rayleigh
	$\mathbb{E} = \Theta(1)$	$\mathbb{E} = \Theta(\sqrt{n})$	$\mathbb{E} = \Theta(1)$	$\mathbb{E} = \Theta(\sqrt{n})$
Recursive trees	Discrete	Stable law	Normal	Normal
	$\mathbb{E} \sim \frac{\pi^2}{6}$	$\mathbb{E} \sim \frac{n}{\log n}$	$\mathbb{E} \sim \log \log n$	$\mathbb{E} \sim \log n$
Binary increasing trees	Discrete	unknown	Normal	Normal
	$\mathbb{E} \sim 2$		$\mathbb{E} \sim \log n$	$\mathbb{E} \sim 2 \log n$
Heap ordered trees	Discrete	unknown	Discrete	Normal
	$\mathbb{E} \sim 2 \log 2$		$\mathbb{E} \sim e - 1$	$\mathbb{E} \sim \frac{1}{2} \log n$

As main results of our studies we show for all tree families considered a discrete limit law (and a bounded mean $\mathbb{E}(Z_n)$) of the random variable Z_n. It is not surprising that the mean is bounded, since one may expect that only few steps are required to reach the final unit-size tree and thus to isolate a leaf. But it is maybe more surprising that the limiting distributions obtained can be characterized in a simple way or even turn out to be elementary ones. E.g., when considering simply generated trees we obtain for the leaf-isolation procedure a Poisson limit law. One might like to compare this with the climbing depth of simply generated trees, which converges in law to a negative binomial distribution (see [23] and references therein). It would be nice to have also a (heuristic) structural interpretation of this result, which the authors want to formulate here as an open question.

Basically, to obtain our limiting distribution results for Z_n we treat the recurrences appearing for the probabilities $\mathbb{P}\{Z_n = m\}$ via bivariate generating functions. This leads to exact solvable differential equations. Extracting coefficients of the solutions appearing asymptotically is performed via singularity analysis (see [10]), a complex-analytic technique that relates asymptotics of sequences to the local behaviour of their generating functions in a neighbourhood of the dominant singularities.

Throughout this paper we use the abbreviations $x^{\underline{l}} := x(x-1)\cdots(x-l+1)$ and $x^{\overline{l}} := x(x+1)\cdots(x+l-1)$ for the falling and rising factorials, respectively. Moreover, we use the abbreviations D_x for the differential operator with respect to x, and E_x for the evaluation operator at $x = 1$.

2. Preliminaries

2.1. **Simply generated trees.** Simply generated trees were introduced in [19] and they include several important tree families as special instances, e.g., binary trees, unordered labelled trees, and ordered trees (= planted plane trees). Moreover, they are strongly related to Galton-Watson branching processes, since it is well known (see [1]), that random simply generated trees are essentially the same as conditioned Galton-Watson trees obtained as the family tree of a Galton-Watson process conditioned on the given total size.

A class $\mathcal{T}$ of simply generated trees can be defined in the following way. A sequence of non-negative real numbers $(\varphi_k)_{k\ge 0}$ with $\varphi_0 = 1$ (φ_k can be seen as the multiplicative weight of a node with out-degree k) is used to define the weight $w(T)$ of any ordered tree T by $w(T) := \prod_v \varphi_{d(v)}$, where v ranges over all vertices of T and $d(v)$ is the out-degree (the number of children) of v. In order to avoid degenerate cases we always assume that there exists a $k \ge 2$ such that $\varphi_k > 0$. The family $\mathcal{T}$ consists then of all trees T with $w(T) \neq 0$ together with their weights $w(T)$. It follows further that for a given degree-weight sequence $(\varphi_k)_{k\ge 0}$ the generating function $T(z) := \sum_{n\ge 1} T_n z^n$ of the quantity total weights $T_n := \sum_{|T|=n} w(T)$, where $|T|$ denotes the size of the tree T, satisfies the functional equation

$$T(z) = z\varphi\big(T(z)\big), \tag{1}$$

where the degree-weight generating function $\varphi(t)$ is given by $\varphi(t) = \sum_{k\ge 0} \varphi_k t^k$. We remark that some authors use the relaxed condition $\varphi_0 > 0$ in the above definition of simply generated tree families. However, this case can be reduced to the condition $\varphi_0 = 1$ by considering the weights $\tilde{\varphi}_k := \frac{\varphi_k}{\varphi_0}$, the degree-weight generating function $\tilde{\varphi}(t) := \frac{\varphi(t)}{\varphi_0}$, and using the substitution $w := \varphi_0 z$ in the generating functions.

The asymptotic behaviour of $T(z)$ as solution of (1) is discussed in detail in [9] and we collect some of their results concerning $T(z)$ and the growth of its coefficients T_n, where we have to make only few restrictions on $\varphi(t)$. We will suppose that $\varphi(t)$ has a positive radius of convergence $R > 0$ and assume that there exists a minimal positive solution $\tau < R$ of the equation $t\varphi'(t) = \varphi(t)$.

Defining the period $p := \gcd\{k : \varphi_k > 0\}$, it follows that equation (1) has exactly p solutions of smallest modulus given by $\tau_j = \omega^j \tau$, for $0 \le j \le p-1$, where ω is a primitive p-th root of unity. This leads to p dominant singularities of $T(z)$ at $z = \rho_j$ with $\rho_j = \omega^j \rho$ and $\rho = \frac{\tau}{\varphi(\tau)} = \frac{1}{\varphi'(\tau)}$ ($T(z)$ is analytic for $|z| \le \rho$ except at $z = \rho_j$).

The local expansion around the singularity $z = \rho_j$ is given by the following equation (where κ_j denotes a certain constant), which holds uniformly in a complex neighborhood of ρ_j: $|z - \rho_j| \le \sigma$, with a certain $\sigma > 0$:

$$T(z) = \tau_j - \omega^j \sqrt{\frac{2\varphi(\tau)}{\varphi''(\tau)}} \sqrt{1 - \frac{z}{\rho_j}} + \kappa_j\big(1 - \frac{z}{\rho_j}\big) + \mathcal{O}\Big(\big(1 - \frac{z}{\rho_j}\big)^{\frac{3}{2}}\Big). \tag{2}$$

By applying singularity analysis one obtains the asymptotic expansion

$$T_n = p\sqrt{\frac{\varphi(\tau)}{2\pi\varphi''(\tau)}} \rho^{-n} n^{-\frac{3}{2}} \big(1 + \mathcal{O}(n^{-1})\big), \tag{3}$$

provided that $n \equiv 1 \pmod{p}$. (For $n \not\equiv 1 \pmod{p}$ $T_n = 0$ always holds.)

We want to mention further that it is often advantageous to describe a simply generated tree family $\mathcal{T}$ by the formal recursive equation

$$\mathcal{T} = \bigcirc \times \Big(\{\epsilon\} \mathbin{\dot\cup} \varphi_1 \cdot \mathcal{T} \mathbin{\dot\cup} \varphi_2 \cdot \mathcal{T} \times \mathcal{T} \mathbin{\dot\cup} \varphi_3 \cdot \mathcal{T} \times \mathcal{T} \times \mathcal{T} \mathbin{\dot\cup} \cdots \Big) = \bigcirc \times \varphi(\mathcal{T}), \tag{4}$$

with $\bigcirc$ a node, $\dot\cup$ the disjoint union, $\times$ the cartesian product, and $\varphi(\mathcal{T})$ the substituted structure (see, e.g., [27]).

2.2. **Increasing trees.** Increasing trees are labelled trees where the nodes of a tree of size n are labelled by distinct integers of the set $\{1, \dots, n\}$ in such a way that each sequence of labels along any branch starting at the root is increasing. As the underlying tree model we use the simply generated trees but, additionally, they are equipped with increasing labellings. We will thus speak about *simple families of increasing trees.* A thorough study of families (= varieties) of increasing trees was conducted in [3].

A class $\mathcal{T}$ of a simple family of increasing trees can thus be defined in analogy to the definition of simply generated tree families in the following way. A sequence of non-negative numbers $(\varphi_k)_{k \ge 0}$ with $\varphi_0 = 1$ is used to define the weight $w(T)$ of any ordered tree T by $w(T) = \prod_v \varphi_{d(v)}$, where v ranges over all vertices of T and $d(v)$ is the out-degree of v (again, we always assume that there exists a $k \ge 2$ with $\varphi_k > 0$). Furthermore, $\mathcal{L}(T)$ denotes the set of different increasing labellings of the tree T with distinct integers $\{1, 2, \dots, |T|\}$, where $|T|$ denotes the size of tree T, and $L(T) := \big|\mathcal{L}(T)\big|$ denotes its cardinality. Then the family $\mathcal{T}$ consists of all trees T together with their weights $w(T)$ and the set of increasing labellings $\mathcal{L}(T)$.

For a given degree-weight sequence $(\varphi_k)_{k \ge 0}$ with a degree-weight generating function $\varphi(t) := \sum_{k \ge 0} \varphi_k t^k$, we define now the total weights by $T_n := \sum_{|T|=n} w(T) \cdot L(T)$. It follows then that the exponential generating function $T(z) := \sum_{n \ge 1} T_n \frac{z^n}{n!}$ satisfies the *autonomous* first order differential equation

$$T'(z) = \varphi\big(T(z)\big), \quad T(0) = 0. \tag{5}$$

Again it is sometimes advantageous to describe an increasing tree family $\mathcal{T}$ by the formal recursive equation

$$\mathcal{T} = ① \times \Big(\{\epsilon\} \mathbin{\dot\cup} \varphi_1 \cdot \mathcal{T} \mathbin{\dot\cup} \varphi_2 \cdot \mathcal{T} * \mathcal{T} \mathbin{\dot\cup} \varphi_3 \cdot \mathcal{T} * \mathcal{T} * \mathcal{T} \mathbin{\dot\cup} \cdots \Big) = ① \times \varphi(\mathcal{T}), \tag{6}$$

where additionally $*$ denotes the partition product for labelled objects.

Three specific increasing tree families are of particular interest:

- *Recursive trees* are the family of non-plane increasing trees such that all node degrees are allowed. The degree-weight generating function is $\varphi(t) = \exp(t)$. Solving (5) gives $T(z) = \log\left(\frac{1}{1-z}\right)$ and thus $T_n = (n-1)!$, for $n \ge 1$. For a survey of applications and results on random recursive trees see [16].
- *Heap ordered trees* (also called plane recursive trees) are the family of plane increasing trees such that all node degrees are allowed. The degree-weight generating function is $\varphi(t) = \frac{1}{1-t}$. Equation (5) leads here to $T(z) = 1 - \sqrt{1-2z}$ and thus to $T_n = \frac{(n-1)!}{2^{n-1}}\binom{2n-2}{n-1}$, for $n \ge 1$. See also [16] for a survey on heap ordered trees.
- *Binary increasing trees* (also called tournament trees) have the degree-weight generating function $\varphi(t) = (1+t)^2$. This model is of special importance, since it is isomorphic to the model of *binary search trees* (see [3] and the references therein for binary increasing trees and, e.g., [15] for binary search trees). Thus it must follow $T(z) = \frac{z}{1-z}$ and $T_n = n!$, for $n \ge 1$.

Driven from the inspection that all these important increasing tree families satisfy the equation $\frac{T_{n+1}}{T_n} = c_1 n + c_2$, with fixed constants c_1, c_2, for all $n \ge 1$, we will consider such trees in more detail. Throughout this paper we will call increasing tree families satisfying this equation *very simple increasing tree families*, since it turns out from the characterization given below that the defining degree-weight generating functions $\varphi(t)$ are the same as obtained in [24].

We will give now an exact answer to the question, which degree-weight generating functions are actually defining very simple increasing tree families.

Lemma 1. *The total weights T_n of trees of size n in an increasing tree family satisfy for all $n \in \mathbb{N}$ the equation*

$$\frac{T_{n+1}}{T_n} = c_1 n + c_2, \tag{7}$$

if and only if the degree-weight generating function $\varphi(t) = \sum_{k\ge 0} \varphi_k t^k$ is given by one of the following three formulæ.

$$\begin{aligned}
&\textbf{\textit{Case A}}: \quad \varphi(t) = e^{c_1 t}, \ \ \textit{for } c_1 > 0, \ \ \textit{(leading to } c_2 = 0 \textit{)},\\
&\textbf{\textit{Case B}}: \quad \varphi(t) = (1 + c_2 t)^d, \ \ \textit{for } c_2 > 0, \ d := \frac{c_1}{c_2} + 1 \in \{2,3,4,\dots\},\\
&\textbf{\textit{Case C}}: \quad \varphi(t) = \frac{1}{(1+c_2 t)^{-\frac{c_1}{c_2}-1}}, \ \ \textit{for } 0 < -c_2 < c_1.
\end{aligned}$$

2.3. **The recursive approach.** We will study the random variable Z_n for the tree families considered by treating the recurrence

$$\mathbb{P}\{Z_n = m\} = \sum_{k=1}^{n-1} q_{n,k}\, \mathbb{P}\{Z_k = m-1\}, \quad \text{for } n \ge 2, \ m \ge 1, \tag{8}$$

with initial values $\mathbb{P}\{Z_1 = 0\} = 1$ and $\mathbb{P}\{Z_n = 0\} = 0$, for $n \ge 2$. Here the transition probabilities $q_{n,k}$ are given as follows: $q_{n,k}$ denotes the probability that by choosing a random tree of size n from the given tree family and removing a random edge the resulting subtree, which does not contain the original root of the tree, is of size k.

An analogous approach, with transition probabilities $q_{n,n-k}$, was used in [24] to study for simply generated tree families the random variable X_n, i.e., the number of cuts to isolate the root of the tree. There one had to make a strong assumption on the tree family in order to justify this recursive approach: it was necessary that randomness is preserved by cutting off a random edge, which means that starting with a random tree of size n and removing a random edge, the remaining subtree of size k containing the root is actually a *random* tree of size k in this tree family. It turned out that exactly those tree families with $\varphi(t)$ given by Lemma 1 have this property and could be treated with the recursive approach. In [24] such tree families are called very simple tree families.

For the random variable Z_n studied in the present paper things are easier. For the tree families considered it always holds that randomness is preserved by cutting off a random edge: after removing a random edge from a random tree of size n, the subtree that does not contain the original root is always a random tree of this tree family. This follows immediately from the formal recursive equations (4) and (6).

3. Results

We state here our findings for simply generated tree families and very simple increasing tree families with $\varphi(t)$ satisfying the assumptions made in Subsection 2.1 and Subsection 2.2, respectively. The proofs of these results are given in Section 4 and Section 5.

Theorem 1. *For simply generated tree families with degree-weight generating function $\varphi(t)$, with period p and τ the minimal positive solution of the equation $t\varphi'(t) = \varphi(t)$, the random variable Z_n, which counts the number of random cuts that are required to isolate a leaf from a randomly chosen tree of size n with the edge-removal procedure considered, converges in distribution, for $n \to \infty$ with $n \equiv 1 \pmod p$, to a shifted Poisson distributed random variable Z, which has the distribution*

$$\mathbb{P}\{Z = m\} = \frac{m\lambda^{m-1}}{m!}\, e^{-\lambda}, \quad \textit{for } m \ge 0,$$

with parameter $\lambda := \log\big(\varphi(\tau)\big)$.

Moreover, the r-th factorial moments $\mathbb{E}\big(Z_n^{\underline{r}}\big)$ have the asymptotic expansion

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = \lambda^{r-1}(\lambda + r) + \mathcal{O}(n^{-1}).$$

In particular, we get for the expectation $\mathbb{E}(Z_n)$ and the variance $\mathbb{V}(Z_n)$:

$$\mathbb{E}(Z_n) = \lambda + 1 + \mathcal{O}(n^{-1}), \quad \textit{and} \quad \mathbb{V}(Z_n) = \lambda + \mathcal{O}(n^{-1}).$$

Theorem 2. *Suppose a very simple increasing tree family with degree-weight generating function* $\varphi(t)$ *is given and thus* $\frac{T_{n+1}}{T_n} = c_1 n + c_2$ *is satisfied, for all* $n \geq 1$*, with certain constants* c_1*,* c_2*. Let* Z_n *be the random variable, which counts the number of random cuts that are required to isolate a leaf from a randomly chosen tree of size* n *with the edge-removal procedure considered. Then* Z_n *converges, for* $n \to \infty$*, in distribution to a discrete random variable* Z*. The probabilities* $\mathbb{P}\{Z = m\}$ *are for* $m \geq 0$ *given as the coefficients of the probability generating function* $p(v) := \sum_{m \geq 0} \mathbb{P}\{Z = m\} v^m$ *as given below:*

$$p(v) = \prod_{k=1}^{\infty} \Big(1 + \frac{(c_1 + c_2)(v-1)}{k(c_1 k + c_2)}\Big). \tag{9}$$

Moreover, the r*-th factorial moments* $\mathbb{E}\big(Z_n^{\underline{r}}\big)$ *are given by the following exact formula:*

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = r!(c_1 + c_2)^r \sum_{k_1=1}^{n-1} \frac{1}{k_1(c_1 k_1 + c_2)} \sum_{k_2=k_1+1}^{n-1} \frac{1}{k_2(c_1 k_2 + c_2)} \cdots \sum_{k_r=k_{r-1}+1}^{n-1} \frac{1}{k_r(c_1 k_r + c_2)}.$$

From Theorem 2 one gets the following corollaries, which contain results for particular increasing tree families.

Corollary 1. *Using the notation of Theorem 2 the following closed formulæ for the probabilities* $\mathbb{P}\{Z = m\}$ *hold. For recursive trees:*

$$\mathbb{P}\{Z = m\} = (-1)^m \sum_{k \geq m} (-1)^k \frac{\pi^{2k} \binom{k}{m}}{(2k+1)!}, \tag{10}$$

which leads to the first few values $\mathbb{P}\{Z = 1\} = 1/2$*,* $\mathbb{P}\{Z = 2\} = 3/8$*,* $\mathbb{P}\{Z = 3\} = 5/16 - \pi^2/48$*.*
For binary increasing trees:

$$\mathbb{P}\{Z = m\} = \sum_{k \geq m+1} (-1)^k \frac{\pi^{2k}}{2^{2k}(2k)!} \sum_{j=m+1}^{k} \binom{k}{j} \binom{j-1}{m} (-1)^j 8^j, \tag{11}$$

which gives in particular $\mathbb{P}\{Z = 1\} = 1/3$*.*

Note that the values given for $\mathbb{P}\{Z = 1\}$ are just as expected, since the average number of leaves in recursive trees is $\sim \frac{n}{2}$ for recursive trees and $\sim \frac{n}{3}$ for binary increasing trees (see, e.g., [3]).

Corollary 2. *Using the notation of Theorem 2 the following closed formulæ for the* r*-th factorial moments of* Z_n *resp.* Z *hold for the instance* $c_2 = 0$*. In the context of the multiple zeta functions*

$$\zeta(a_1, \dots, a_l) := \sum_{1 \leq n_1 < n_2 < \dots < n_l} \frac{1}{n_1^{a_1} n_2^{a_2} \dots n_l^{a_l}}, \quad \zeta_N(a_1, \dots, a_l) := \sum_{1 \leq n_1 < n_2 < \dots < n_l \leq N} \frac{1}{n_1^{a_1} n_2^{a_2} \dots n_l^{a_l}}, \tag{12}$$

the factorial moments $\mathbb{E}\big(Z_n^{\underline{r}}\big)$ *can be expressed for* $c_2 = 0$ *as follows:*

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = r! \zeta_{n-1}(2, \dots, 2). \tag{13}$$

Furthermore we obtain for $c_2 = 0$ *the following expression for the factorial moments* $\mathbb{E}\big(Z^{\underline{r}}\big)$*:*

$$\mathbb{E}\big(Z^{\underline{r}}\big) = r! \zeta(2, \dots, 2) = r! \frac{\pi^{2r}}{(2r+1)!}. \tag{14}$$

We remark that our computations of $\mathbb{E}\big(Z^{\underline{r}}\big)$ give thus a further proof of the identity $\zeta(\underbrace{2, \dots, 2}_{r \text{ times}}) = \frac{\pi^{2r}}{(2r+1)!}$, which was shown first in [11].

In Table 2 and 3 we collect some results of the limiting distribution of Z_n for a few important simply generated tree families and very simple increasing tree families, respectively.

We further remark that, at least in principle, this recursive approach can also be used to obtain results for further tree families, in particular for increasing tree families, which are not *very simple* and thus do not fall into one of the classes characterized by Lemma 1. For showing limiting distribution results via the approach presented a detailled description of the transition probabilities $q_{n,k}$ as defined in Subsection 2.3

TABLE 2. Limiting distribution results of Z_n for some important simply generated tree families.

Tree family	Degree-weight generating function $\varphi(t)$	$Z_n \to Z$, shifted Poisson distributed with parameter λ, $\mathbb{E}(Z_n) \sim \lambda + 1$, $\mathbb{V}(Z_n) \sim \lambda$
Cayley trees	$\varphi(t) = e^t$	$\lambda = 1$
d-ary trees	$\varphi(t) = (1+t)^d, \ d \geq 2$	$\lambda = d \log\Big(\frac{d}{d-1}\Big)$
Ordered trees	$\varphi(t) = \frac{1}{1-t}$	$\lambda = \log(2)$
Motzkin trees	$\varphi(t) = 1 + t + t^2$	$\lambda = \log(3)$
Strict binary trees	$\varphi(t) = 1 + t^2$	$\lambda = \log(2)$

TABLE 3. Limiting distribution results of Z_n for some important very simple increasing tree families. $H_n := \sum_{k \geq 1} \frac{1}{k}$ resp. $H_n^{(2)} := \sum_{k \geq 1} \frac{1}{k^2}$ denote the first and second order harmonic numbers.

Tree family	$\varphi(t)$	$\frac{T_{n+1}}{T_n}$	$Z_n \to Z$, with $p(v) = \sum_{m \geq 0} \mathbb{P}\{Z = m\} v^m$	$\mathbb{E}(Z_n)$
Recursive trees	e^t	n	$p(v) = \frac{\sin(\pi\sqrt{1-v})}{\pi\sqrt{1-v}}$	$H_{n-1}^{(2)}$ $\sim \frac{\pi^2}{6} \approx 1.6449$
Binary increasing trees	$(1+t)^2$	$n+1$	$p(v) = \frac{\cos(\frac{\pi}{2}\sqrt{9-8v})}{2\pi(v-1)}$	$2 - \frac{2}{n} \sim 2$
Heap ordered trees	$\frac{1}{1-t}$	$2n-1$	$p(v) = \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{3-\sqrt{9-8v}}{4})\Gamma(\frac{3+\sqrt{9-8v}}{4})}$	$2(H_{2n-2} - H_{n-1})$ $\sim 2\log 2 \approx 1.3863$

would be required, which seems in general quite involved to obtain. On the other hand less detailled information on the transition probabilites $q_{n,k}$ like suitable bounds may lead at least to partial information on Z_n like establishing bounds for the expectation $\mathbb{E}(Z_n)$. As one example in this direction we consider so called polynomial increasing tree families, i.e., increasing trees where the degree-weight generating function $\varphi(t) = 1 + \varphi_1 t + \cdots + \varphi_d t^d$ is a polynomial of degree $d \geq 2$ (thus satisfying $\varphi_d \neq 0$). Using results obtained in [3] for the analytic behaviour of $T(z)$ of polynomial increasing tree families one can show the following bound on the the transition probabilities: $q_{n,k} \leq M k^{-2+\frac{1}{d-1}}$, with M a certain constant independent of n and k. This uniform bound on the transition probabilities is sufficient to show that for every polynomial increasing tree family with $d \geq 3$ there exists a constant $\tilde{M}$ such that the expected number $\mathbb{E}(Z_n)$ of cuts to isolate a leaf is bounded by $\tilde{M}$.

4. SIMPLY GENERATED TREE FAMILIES

4.1. **The transition probabilities.** The required transition probabilities $q_{n,k}$ as defined in Subsection 2.3 were already computed in [24] by a generating functions approach, which is here also sketched. We can define the value $q_{n,k}$ equivalently as the probability that the number of descendants of a node (where the node itself is counted) that was chosen at random from one of the $n-1$ non-root nodes in a random tree of size n is k. We require also the auxiliary value $\tilde{q}_{n,k}$, which denote the probability that the number of descendants of a randomly chosen node in a random tree of size n is k.

Introducing the generating functions

$$G(z,u) = \sum_{n\ge 1}\sum_{k\ge 0} nT_n\, \tilde{q}_{n,k}\, z^n u^k \qquad \text{and} \qquad H(z,u) = \sum_{n\ge 1}\sum_{k\ge 0} (n-1)T_n\, q_{n,k}\, z^n u^k,$$

we can translate the formal equation (4) into the system of equations

$$G(z,u) = T(zu) + z\varphi'\big(T(z)\big)G(z,u), \qquad H(z,u) = z\varphi'\big(T(z)\big)G(z,u),$$

which imply

$$H(z,u) = T(zu)F(z), \quad \text{with} \quad F(z) := \frac{1}{1 - z\varphi'\big(T(z)\big)} - 1. \tag{15}$$

Extracting coefficients from (15) gives

$$F_n := [z^n]F(z) = \begin{cases} [T^n]\big(\varphi(T)\big)^n, & \text{for } n \ge 1, \\ 0, & \text{for } n = 0. \end{cases} \tag{16}$$

Thus the required transition probabilities $q_{n,k}$ are, for $1 \le k \le n-1$, given as follows, with F_n defined by equation (16):

$$q_{n,k} = \frac{[z^n u^k]H(z,u)}{(n-1)T_n} = \frac{T_k F_{n-k}}{(n-1)T_n}. \tag{17}$$

4.2. **Solving the recurrence.** Using (8) we have to study the recurrence

$$\mathbb{P}\{Z_n = m\} = \sum_{k=1}^{n-1} \frac{T_k F_{n-k}}{(n-1)T_n} \mathbb{P}\{Z_k = m-1\}, \tag{18}$$

with $\mathbb{P}\{Z_1 = 0\} = 1$ and $\mathbb{P}\{Z_n = 0\} = 0$, for $n \ge 2$.

We will perform a generating functions approach using the bivariate generating function

$$M(z,v) := \sum_{n\ge 1}\sum_{m\ge 0} T_n \mathbb{P}\{Z_n = m\} z^n v^m.$$

Multiplying (18) with $(n-1)T_n z^n v^m$ and summing up for $n \ge 2$ and $m \ge 1$ leads to the following first order linear differential equation

$$z\frac{\partial}{\partial z}M(z,v) - M(z,v) = vF(z)M(z,v),$$

with initial conditions $M(0,v) = 0$ and $\big(\frac{\partial}{\partial z}M(z,v)\big)|_{z=0} = T_1 = 1$, and the function $F(z)$ given by (15). Solving this differential equation leads to the solution

$$M(z,v) = z\exp\Big(v\int_0^z \frac{F(t)}{t}dt\Big). \tag{19}$$

Then by using $T'(z) = \frac{\varphi(T(z))}{1 - z\varphi'(T(z))}$, which follows from the functional equation (1), and a change of variables, we obtain from (19) the following explicit formula for $M(z,v)$:

$$M(z,v) = z\exp\Big(v\log\varphi\big(T(z)\big)\Big) = z\Big(\varphi\big(T(z)\big)\Big)^v. \tag{20}$$

4.3. **Characterizing the limiting distribution.** Extracting coefficients from (20) immediately gives

$$[v^m]M(z,v) = z\frac{\big(\log\varphi(T(z))\big)^m}{m!}. \tag{21}$$

Before continuing we remark that $\varphi(T(z)) \ne 0$ inside the disk of convergence: if there would exist such a z_0 with $\varphi(T(z_0)) = 0$ then it would also hold that $T(z_0) = 0$ due to equation (1). If we plug in $T(z_0) = 0$ into (1) we obtain further $0 = T(z_0) = z_0\varphi(T(z_0)) = z_0\varphi(0) = z_0\varphi_0$. Since $z_0 \ne 0$ ($\varphi(T(0)) = \varphi(0) = \varphi_0 \ne 0$) this gives a contradiction.

In our asymptotic study of the coefficients $[z^n v^m]M(z,v)$ (and thus of the probabilities $\mathbb{P}\{Z_n = m\}$) via singularity analysis, which is given below, we will only carry out the instance that the degree-weight generating function $\varphi(t)$ is aperiodic, i.e., $p = 1$. But for functions $\varphi(t)$ with period $p > 1$ the proof is fully analogous: then we have to consider the contributions of all p dominant singularities, which must be added. This shows Theorem 1 also for $p > 1$.

Using the singular expansion (2) of $T(z)$ we obtain from (21) for $m \geq 1$ the following local expansion around the dominant singularity $z = \rho$ (with a certain constant $\tilde{\kappa}$), which holds uniformly in a complex neighborhood of $z = \rho$: $|z - \rho| \leq \sigma$, with a certain $\sigma > 0$:

$$\begin{aligned}[v^m]M(z,v) &= \sum_{n\geq 1} T_n \mathbb{P}\{Z_n = m\} z^n \\ &= \frac{\rho}{m!}\Big(\log\varphi(\tau)\Big)^m - \frac{1}{\varphi(\tau)(m-1)!}\Big(\log\varphi(\tau)\Big)^{m-1}\sqrt{\frac{2\varphi(\tau)}{\varphi''(\tau)}}\sqrt{1-\frac{z}{\rho}} + \tilde{\kappa}\big(1-\frac{z}{\rho}\big) + \mathcal{O}\Big(\big(1-\frac{z}{\rho}\big)^{\frac{3}{2}}\Big).\end{aligned} \tag{22}$$

Applying singularity analysis to (22) gives then

$$[z^n v^m]M(z,v) = T_n \mathbb{P}\{Z_n = m\} = \frac{1}{\varphi(\tau)(m-1)!}\Big(\log\varphi(\tau)\Big)^{m-1}\sqrt{\frac{\varphi(\tau)}{2\pi\varphi''(\tau)}}\,\rho^{-n}n^{-\frac{3}{2}}\left(1+\mathcal{O}(n^{-1})\right),$$

and, together with (3), for $m \geq 1$ the asymptotic expansion

$$\mathbb{P}\{Z_n = m\} = \frac{1}{\varphi(\tau)}\frac{\big(\log\varphi(\tau)\big)^{m-1}}{(m-1)!}\big(1+\mathcal{O}(n^{-1})\big). \tag{23}$$

Thus the probabilities $\mathbb{P}\{Z_n = m\}$ converge for all $m \geq 1$ to the probabilities $\mathbb{P}\{Z = m\}$ of a shifted Poisson distributed random variable Z. This shows the first part of Theorem 1.

4.4. **Computing the moments.** From the generating function $M(z,v)$ as given by (20) we can also compute easily the r-th factorial moments $\mathbb{E}\big(Z^{\underline{r}}\big)$.

Evaluating the r-th derivative with respect to v of $M(z,v)$ at $v = 1$ gives

$$E_v D_v^r M(z,v) = E_v\Big[z\Big(\log\varphi\big(T(z)\big)\Big)^r e^{v\log\varphi(T(z))}\Big] = T(z)\Big(\log\varphi\big(T(z)\big)\Big)^r. \tag{24}$$

We further get by using (2) the following local expansion around $z = \rho$ (with a certain constant $\hat{\kappa}$), which again holds uniformly in a complex neighborhood of $z = \rho$:

$$\begin{aligned}T(z)\Big(\log\varphi(T(z))\Big)^r &= \tau\Big(\log\varphi(\tau)\Big)^r - \Big(\log\varphi(\tau)\Big)^{r-1}\big(r+\log\varphi(\tau)\big)\sqrt{\frac{2\varphi(\tau)}{\varphi''(\tau)}}\sqrt{1-\frac{z}{\rho}} \\ &\quad + \hat{\kappa}\big(1-\frac{z}{\rho}\big) + \mathcal{O}\Big(\big(1-\frac{z}{\rho}\big)^{\frac{3}{2}}\Big).\end{aligned} \tag{25}$$

Singularity analysis leads then from (24) and (25) to the asymptotic expansion

$$[z^n]E_v D_v^r M(z,v) = \Big(\log\varphi(\tau)\Big)^{r-1}\big(r+\log\varphi(\tau)\big)\sqrt{\frac{\varphi(\tau)}{2\pi\varphi''(\tau)}}\,\rho^{-n}n^{-\frac{3}{2}}\left(1+\mathcal{O}(n^{-1})\right),$$

and thus by using (3) to the following result, which completes the proof of Theorem 1:

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = \frac{[z^n]E_v D_v^r M(z,v)}{T_n} = \big(\log\varphi(\tau)\big)^{r-1}\big(r+\log\varphi(\tau)\big)\big(1+\mathcal{O}(n^{-1})\big).$$

5. Very simple increasing tree families

5.1. Characterization of very simple increasing tree families.

Proof of Lemma 1. We will show here Lemma 1, which characterizes increasing tree families that satisfy the equation $\frac{T_{n+1}}{T_n} = c_1 n + c_2$, with arbitrary but fixed constants c_1, c_2, for all $n \geq 1$.

We remark that due to the demand $T_n > 0$, for all $n \geq 1$, we get the a priori restrictions: $c_1 \geq 0$ and $c_2 > -c_1$ (otherwise there would exist $n \geq 1$ such that $\frac{T_{n+1}}{T_n} = c_1 n + c_2 < 0$).

• We consider the case $c_1 \neq 0$ and $c_2 \neq 0$ and get for T_n:

$$T_n = \prod_{k=1}^{n-1} (c_1 k + c_2) = c_1^{n-1} \prod_{k=1}^{n-1} \big(\frac{c_2}{c_1} + k\big) = \frac{c_1^n}{c_2} \big(\frac{c_2}{c_1} + n - 1\big)^{\underline{n}} = \frac{(-c_1)^n n!}{c_2} \binom{-\frac{c_2}{c_1}}{n},$$

and further

$$T(z) = \sum_{n \geq 1} T_n \frac{z^n}{n!} = \frac{1}{c_2} \sum_{n \geq 1} \binom{-\frac{c_2}{c_1}}{n} (-c_1 z)^n = \frac{1}{c_2} \Big(\frac{1}{(1 - c_1 z)^{\frac{c_2}{c_1}}} - 1 \Big). \tag{26}$$

In order to decide which values of c_1, c_2 are indeed possible choices we have to compute the corresponding degree-weight generating functions and check, whether they are admissible ($\varphi_k \geq 0$ for all $k \geq 0$). Differentiating (26) gives

$$T'(z) = \frac{1}{(1 - c_1 z)^{\frac{c_2}{c_1} + 1}} = \Big(1 + c_2 T(z)\Big)^{\frac{c_1}{c_2} + 1}. \tag{27}$$

Since $T'(z) = \varphi\big(T(z)\big)$ this further gives

$$\varphi(t) = \sum_{n \geq 0} \varphi_n t^n = \Big(1 + c_2 t\Big)^{\frac{c_1}{c_2} + 1} \quad \text{and} \quad \varphi_n = \binom{\frac{c_1}{c_2} + 1}{n} c_2^n. \tag{28}$$

By considering (28) we can now check, whether the conditions $\varphi_n \geq 0$, for all $n \geq 0$, are satisfied.

(*i*) We consider first the case $c_2 > 0$: if $1 + \frac{c_1}{c_2} \notin \mathbb{N}$, then it follows that there exists $n \in \mathbb{N}$ such that $\binom{1 + \frac{c_1}{c_2}}{n} < 0$ and, since $c_1 > 0$, thus that $\varphi_n < 0$. Therefore we get that this case is not admissible. But if $1 + \frac{c_1}{c_2} =: d \in \mathbb{N}$, then it follows that $\binom{\frac{c_1}{c_2} + 1}{n} = 0$, for all $n > d$ and thus that $\varphi_n > 0$, for all $0 \leq n \leq d$ and $\varphi_n = 0$, for all $n > d$. Such degree-weight generating functions are admissible and are covered by Case B in Lemma 1.

(*ii*) We have to consider also the case $c_2 < 0$: since $c_1 + c_2 > 0$ it follows that $\frac{c_1}{c_2} < -1$ resp. $n - \frac{c_1}{c_2} - 2 > n - 1$ and thus that

$$\varphi_n = \binom{\frac{c_1}{c_2} + 1}{n} (-1)^n \big(-c_2\big)^n = \binom{n - \frac{c_1}{c_2} - 2}{n} \big(-c_2\big)^n > 0,$$

for all $n \geq 0$. Therefore such degree-weight generating functions are also admissible and are covered by Case C in Lemma 1.

• By similar considerations one can treat the case $c_2 = 0$ (and $c_1 > 0$), which eventually gives the degree-weight generating function

$$\varphi(t) = \sum_{n \geq 0} \varphi_n t^n = e^{c_1 t}. \tag{29}$$

Since $c_1 > 0$, we obtain from (29) that $\varphi_n > 0$, for all $n \geq 0$, and thus that all degree-weight generating functions (29) are admissible. They are covered by Case A in Lemma 1.

• The remaining case is $c_1 = 0$ (and thus $c_2 > 0$), which leads to the degree-weight generating function $\varphi(t) = 1 + c_2 t$. This degenerate case (all trees are "chains") is excluded from our further considerations due to the demand that there exists a $k \geq 2$ with $\varphi_k > 0$. ☐

5.2. **The transition probabilities.** First we show for general increasing tree families an expression for the transition probabilities $q_{n,k}$ as defined in Subsection 2.3. We can do this analogously to Subsection 4.1 for simply generated tree families: we use the interpretation of the value $q_{n,k}$ as the probability that the number of descendants of a node that was chosen at random from one of the $n-1$ non-root nodes in a random tree of size n is k, and define the auxiliary value $\tilde{q}_{n,k}$ as the probability that the number of descendants of a randomly chosen node in a random tree of size n is k.

Introducing the generating functions $G(z,u) = \sum_{n \geq 1} \sum_{k \geq 0} n T_n \tilde{q}_{n,k} \frac{z^n}{n!} u^k$ and $H(z,u) = \sum_{n \geq 1} \sum_{k \geq 0} (n-1) T_n q_{n,k} \frac{z^n}{n!} u^k$ we obtain from the formal equation (6) the following system of differential equations:

$$\frac{\partial}{\partial z} G(z,u) = u \varphi\big(T(zu)\big) + \varphi'\big(T(z)\big) G(z,u), \quad \frac{\partial}{\partial z} H(z,u) = \varphi'\big(T(z)\big) G(z,u),$$

with initial conditions $G(0,u) = H(0,u) = 0$. This leads to the following explicit solution for $H(z,u)$:

$$H(z,u) = \varphi\big(T(z)\big) \int_0^z \frac{T(tu)\varphi'\big(T(t)\big)}{\varphi\big(T(t)\big)} dt, \tag{30}$$

which immediately gives an expression for the transition probabilities:

$$q_{n,k} = \frac{n!}{(n-1)T_n}[z^n u^k]H(z,u) = \frac{n! T_k}{(n-1)T_n k!}[z^n]\varphi\big(T(z)\big) \int_0^z \frac{t^k \varphi'\big(T(t)\big)}{\varphi\big(T(t)\big)} dt. \tag{31}$$

From (31) one can obtain explicit formulæ for the probabilities $q_{n,k}$ for the subclass of very simple increasing tree families by specializing $\varphi(t)$ as given in Lemma 1. Since this is a routine task we only state the result:

$$q_{n,k} = \frac{(c_1+c_2)(c_1 n + c_2)}{(n-1)(c_1(k+1)+c_2)(c_1 k + c_2)}. \tag{32}$$

5.3. **Solving the recurrence.** Using (8) we obtain therefore for $n \geq 2$ and $m \geq 1$ the recurrence

$$\mathbb{P}\{Z_n = m\} = \sum_{k=1}^{n-1} \frac{(c_1+c_2)(c_1 n + c_2)}{(n-1)(c_1(k+1)+c_2)(c_1 k + c_2)} \mathbb{P}\{Z_k = m-1\}, \tag{33}$$

with $\mathbb{P}\{Z_1 = 0\} = 1$ and $\mathbb{P}\{Z_n = 0\} = 0$, for $n \geq 2$. Introducing the generating function

$$M(z,v) := \sum_{n\geq 1}\sum_{m\geq 0} \mathbb{P}\{Z_n = m\} \frac{z^{n-1}}{(c_1(n+1)+c_2)(c_1 n + c_2)} v^m,$$

recurrence (33) leads to the following homogeneous second order linear differential equation:

$$z(1-z)\frac{\partial^2}{\partial z^2}M(z,v) + \frac{3c_1+c_2}{c_1}(1-z)\frac{\partial}{\partial z}M(z,v) - \frac{(c_1+c_2)v}{c_1}M(z,v) = 0, \tag{34}$$

with initial conditions $M(0,v) = \frac{1}{(2c_1+c_2)(c_1+c_2)}$ and $\frac{\partial}{\partial z}M(z,v)\big|_{z=0} = \frac{v}{(3c_1+c_2)(2c_1+c_2)}$. One can show that this hypergeometric differential equation has the solution

$$M(z,v) = \frac{1}{(2c_1+c_2)(c_1+c_2)} {}_2F_1\left(\left. \begin{matrix} \frac{2c_1+c_2-\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1}, \frac{2c_1+c_2+\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1} \\ \frac{3c_1+c_2}{c_1} \end{matrix} \right| z \right), \tag{35}$$

where ${}_2F_1\left(\left. \begin{matrix} a, b \\ c \end{matrix} \right| z \right) := \sum_{n\geq 0} \frac{a^{\overline{n}} b^{\overline{n}}}{c^{\overline{n}}} \frac{z^n}{n!}$ denotes the Gauss hypergeometric series.

5.4. **Characterizing the limiting distribution.** To obtain a limiting distribution result we will apply the following instance of the z to $1-z$ transformation (see, e.g., [2]) with $m \in \{1,2,3,\dots\}$:

$$\begin{aligned} {}_2F_1\left(\left. \begin{matrix} a, b \\ a+b+m \end{matrix} \right| z \right) &= \frac{\Gamma(m)\Gamma(a+b+m)}{\Gamma(a+m)\Gamma(b+m)} \sum_{n=0}^{m-1} \frac{a^{\overline{n}} b^{\overline{n}}}{n!(1-m)^{\overline{n}}}(1-z)^n \\ &- \frac{\Gamma(a+b+m)}{\Gamma(a)\Gamma(b)}(z-1)^m \sum_{n=0}^{\infty} \frac{(a+m)^{\overline{n}}(b+m)^{\overline{n}}}{n!(n+m)!}(1-z)^n \\ &\times \Big(\log(1-z) - \Psi(n+1) - \Psi(n+m+1) + \Psi(a+n+m) + \Psi(b+n+m) \Big), \end{aligned}$$

and get from equation (35) the following local expansion of $M(z,v)$ around the dominant singularity $z = 1$ in a complex neighbourhood of $v = 1$ (with certain functions $C_0(v)$, $C_1(v)$, and $C_2(v)$), which holds uniformly in a complex neighborhood of $z = 1$:

$$\begin{aligned} M(z,v) = \frac{1}{(2c_1+c_2)(c_1+c_2)} &\frac{\Gamma(3+\frac{c_2}{c_1})}{\Gamma(\frac{2c_1+c_2-\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1})\Gamma(\frac{2c_1+c_2+\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1})} \\ &\times (z-1)\log\frac{1}{1-z} + C_0(v) + C_1(v)(1-z) + C_2(v)(1-z)^2 + \mathcal{O}\Big((1-z)^2 \log\frac{1}{1-z}\Big). \end{aligned}$$

Singularity analysis of generating functions [10] shows then that the moment generating function (= Laplace transform) $\mathbb{E}(e^{Z_n s}) := \sum_{m\geq 0} \mathbb{P}\{Z_n = m\} e^{ms} = p_n(e^s)$ of Z_n converges in a neighbourhood of

$s = 0$ to the moment generating function $\mathbb{E}(e^{Zs}) = p(e^s)$ of a discrete random variable Z with probability generating function $p(v) := \sum_{m\ge 0} \mathbb{P}\{Z = m\} v^s$ given by

$$p(v) = \frac{\Gamma(1+\frac{c_2}{c_1})}{\Gamma(\frac{2c_1+c_2-\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1})\Gamma(\frac{2c_1+c_2+\sqrt{(2c_1+c_2)^2-4c_1(c_1+c_2)v}}{2c_1})}. \tag{36}$$

We remark that by using the reflection law of the Gamma function $\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin(\pi x)}$, one can further simplify the formula of $p(v)$ for binary increasing trees and recursive trees; see Table 3 and Corollary 1. Furthermore, we can give a representation of $p(v)$ as an infinite product, where we simply use repeatedly the functional equation $\Gamma(x) = \frac{\Gamma(x+1)}{x}$ for the Gamma function expressions in (36):

$$p(v) = \prod_{k=1}^{\infty} \Big(1 + \frac{(c_1+c_2)(v-1)}{k(c_1k+c_2)}\Big). \tag{37}$$

By an application of the continuity theorem for the Laplace transform (see, e.g., [4]) we obtain from equation (37) immediately the first part of Theorem 2.

5.5. **Computing the moments.** From the explicit formula (35) for the generating function $M(z,v)$ one can also compute exact expressions for the r-th factorial moments $\mathbb{E}\big(Z_n^{\underline{r}}\big)$ of Z_n. We first give an exact formula for the probability generating function $p_n(v)$, which can be obtained from $M(z,v)$ easily:

$$p_n(v) = (c_1(n+1)+c_2)(c_1n+c_2)[z^{n-1}]M(z,v) = \prod_{k=1}^{n-1} \Big(1 + \frac{(c_1+c_2)(v-1)}{k(c_1k+c_2)}\Big). \tag{38}$$

Evaluating the r-th derivative of $p_n(v)$ as given by (38) at $v = 1$ leads then to:

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = E_v D_v^r p_n(v) = r!(c_1+c_2)^r \sum_{k_1=1}^{n-1} \frac{1}{k_1(c_1k_1+c_2)} \sum_{k_2=k_1+1}^{n-1} \frac{1}{k_2(c_1k_2+c_2)} \cdots \sum_{k_r=k_{r-1}+1}^{n-1} \frac{1}{k_r(c_1k_r+c_2)},$$

which shows also the second part of Theorem 2. If $c_2 = 0$ we get by using (12) the following result for the r-th factorial moments of Z_n (and also Z by considering the limit $n \to \infty$):

$$\mathbb{E}\big(Z_n^{\underline{r}}\big) = r!\zeta_{n-1}(2,\dots,2) \quad \text{and} \quad \mathbb{E}\big(Z^{\underline{r}}\big) = r!\zeta(2,\dots,2).$$

Furthermore using the probability generating function $p(v)$ for $c_1 = 1$ and $c_2 = 0$ given in Table 3:

$$p(v) = \lim_{n\to\infty} p_n(v) = \frac{\sin(\pi\sqrt{1-v})}{\pi\sqrt{(1-v)}} = \sum_{k\ge 0} \frac{\pi^{2k}(-1)^k(1-v)^k}{(2k+1)!},$$

we obtain the following formula for the r-th factorial moment of Z finishing the proof of Corollary 2:

$$\mathbb{E}(Z^{\underline{r}}) = E_v D_v^r p(v) = \frac{r!\pi^{2r}}{(2r+1)!}.$$

ACKNOWLEDGEMENT

We thank the anonymous referees for many valuable comments, which helped to improve the presentation of the contents.

REFERENCES

[1] D. Aldous, The continuum random tree II: an overview, *London Mathematical Society Lecture Note Series* 167, 23–70, 1991.

[2] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, ninth edition, Dover, New York, 1970.

[3] F. Bergeron, P. Flajolet and B. Salvy, Varieties of increasing trees, *Lecture Notes in Computer Science* 581, 24–48, 1992.

[4] K. L. Chung, *A course in probability theory*, Academic Press, San Diego, 1974.

[5] M. Drmota, A. Iksanov, M. Möhle and U. Rösler, A limiting distribution for the number of cuts needed to isolate the root of a random recursive tree, preprint, 2006.

[6] M. Drmota, A. Iksanov, M. Möhle and U. Rösler, Asymptotic results about the total branch length of the Bolthausen-Sznitman coalescent, *Stochastic Processes and their Applications*, to appear, 2007.

[7] J. A. Fill, N. Kapur and A. Panholzer, Destruction of very simple trees, *Algorithmica* 46, 345–366, 2006.

[8] J. A. Fill, H. Mahmoud and W. Szpankowski, On the distribution for the duration of a randomized leader election algorithm, *The Annals of Applied Probability* 6, 1260–1283, 1996.
[9] P. Flajolet and A. Odlyzko, The average height of binary trees and other simple trees, *Journal of Computer and System Sciences* 25, 171–213, 1982.
[10] P. Flajolet and A. Odlyzko, Singularity analysis of generating functions, *SIAM Journal on Discrete Mathematics* 3, 216–240, 1990.
[11] M. Hoffman, Multiple harmonic series, *Pacific Journal of Mathematics* 152, 275–290, 1992.
[12] S. Janson, Random cutting and records in deterministic and random trees, *Random Structures & Algorithms* 29, 139–179, 2006.
[13] S. Janson and W. Szpankowski, Analysis of an asymmetric leader election algorithm, *Electronic Journal of Combinatorics* 4, #R17 (16 pp.), 1997.
[14] D. Knuth and A. Schönhage, The expected linearity of a simple equivalence algorithm, *Theoretical Computer Science* 6, 281–315, 1978.
[15] H. Mahmoud, *Evolution of random search trees*, Wiley, New York, 1992.
[16] H. Mahmoud and R. Smythe, A survey of recursive trees, *Theoretical Probability and Mathematical Statistics* 51, 1–37, 1995.
[17] A. Meir and J. W. Moon, Cutting down random trees, *Journal of the Australian Mathematical Society* 11, 313–324, 1970.
[18] A. Meir and J. W. Moon, Cutting down recursive trees, *Mathematical Biosciences* 21, 173–181, 1974.
[19] A. Meir and J. W. Moon, On the altitude of nodes in random trees, *Canadian Journal of Mathematics* 30, 997–1015, 1978.
[20] R. Neininger and L. Rüschendorf, On the contraction method with degenerate limit equation, *The Annals of Probability* 32, 2838–2856, 2004.
[21] A. Panholzer, Non-crossing trees revisited: cutting down and spanning subtrees, in *Discrete Random Walks*, C. Banderier and C. Krattenthaler (eds.), *Discrete Mathematics and Theoretical Computer Science*, Proceedings AC, 265–276, 2003.
[22] A. Panholzer, Destruction of recursive trees, *Proceedings of the "Third Colloquium on Mathematics and Computer Science"*, M. Drmota et al (eds.), Birkhäuser, Basel, 267–280, 2004.
[23] A. Panholzer, The climbing depth of random trees, *Random Structures & Algorithms* 26, 84–109, 2005.
[24] A. Panholzer, Cutting down very simple trees, *Quaestiones Mathematicae* 29, 211–227, 2006.
[25] H. Prodinger, How to select a loser, *Discrete Mathematics* 120, 149–159, 1993.
[26] U. Rösler, On the analysis of stochastic divide and conquer algorithms, *Algorithmica* 29, 238–261, 2001.
[27] J. Vitter and P. Flajolet, *Average case analysis of algorithms and data structures*, in *Handbook of Theoretical Computer Science*, 431–524, Elsevier, Amsterdam, 1990.

Markus Kuba, Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstr. 8-10/104, 1040 Wien, Austria
E-mail address: `Markus.Kuba@tuwien.ac.at`

Alois Panholzer, Institut für Diskrete Mathematik und Geometrie, Technische Universität Wien, Wiedner Hauptstr. 8-10/104, 1040 Wien, Austria
E-mail address: `Alois.Panholzer@tuwien.ac.at`