ISOLATING NODES IN RECURSIVE TREES

MARKUS KUBA AND ALOIS PANHOLZER

ABSTRACT. We consider the number of random cuts that are necessary to isolate the node with label λ, $1 \le \lambda \le n$, in a random recursive tree of size n. At each stage of the edge-removal procedure considered an edge is chosen at random from the tree and cut, separating the tree into two subtrees. The procedure is then continued with the subtree containing the specified label λ, whereas the other subtree is discarded. The procedure stops when the node with label λ is isolated. Using a recursive approach we are able to give asymptotic expansions for all ordinary moments of the random variable $X_{n,\lambda}$, which counts the number of random cuts required to isolate the vertex with label λ in a random size-n recursive tree, for small labels, i. e., $\lambda = l$, and large labels, i. e., $\lambda = n + 1 - l$, with $l \ge 1$ fixed and $n \to \infty$. Moreover, we can characterize the limiting distribution of a scaled variant of $X_{n,\lambda}$, for the instance of large labels.

1. INTRODUCTION

The model of random recursive trees is definitely one of the most popular and most studied random tree models, since it turned out to be of interest for various applications; we mention here the stemma construction in philology [11], for pyramid schemes [3], and the spread of contamination in an organism [10]. For a survey of applications and results of recursive trees see [9].

A rooted labelled tree T of size $|T| = n$ (the size of a tree T is measured as usual by the number of nodes of T), where the nodes are labelled by distinct integers of the set $\{1, 2, \ldots, n\}$ is a recursive tree, if the root is labelled by 1, and for each node v holds that the labels of the vertices on the unique path from the root to v form an increasing sequence. It is well known that there are $T_n = (n-1)!$ different size-n recursive trees. A random recursive tree of size n is then a recursive tree with n nodes chosen with equal probability $\frac{1}{T_n}$ from the space of all such trees. Besides the previous description the construction of random recursive trees can be described via the following simple growth rule: a random tree T of size n is obtained from a random tree T' of size $n-1$ by choosing uniformly at random a node in T' and adjoining the node labelled by n to it.

In [10] Meir and Moon considered the following edge-removal procedure (= cutting-down procedure) for a rooted tree with n vertices. 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 the subtree that does not contain the original root is discarded and the procedure is continued recursively for the remaining subtree until the original root is isolated. In paper [10] the random variable X_n is studied, which counts the number of edges that will be removed from a randomly chosen recursive tree of size n by above edge-removal procedure until the root, i. e., the node labelled by 1, is isolated. This problem was studied in the context of the spread of contamination in an organism, where it is assumed that the first node is the source of all the contamination. By separating nodes from the source by successively removing edges one eventually isolates the source node. It was shown in [10] the following asymptotic equivalent of the expectation: $\mathbb{E}(X_n) \sim \frac{n}{\log n}$. Thus on average $\sim \frac{n}{\log n}$ random edges have to be removed from a random size-n recursive tree before the root node is isolated.

Recently Javanian and Vahidi-Asl [8] have studied a modification of above edge-removal procedure, motivated by considerations concerning the hierarchy of a workforce of a company: at each stage after removing a random edge, the subtree containing the node with the largest label, i. e., label n, is kept and the other subtree is discarded. Thus finally the node labelled by n will be isolated. Again one is interested in a study of the random variable Y_n, which counts the number of edges that will be removed from a randomly chosen recursive tree of size n until node n, i. e., the last recent entry, is isolated. It was shown in [8] the following asymptotic equivalent of the expectation: $\mathbb{E}(Y_n) \sim \frac{n}{2 \log n}$. Therefore on average $\sim \frac{n}{2 \log n}$ random edges have to be removed from a random size-n recursive tree before node n is isolated.

But isolating node 1 and isolating node n in a tree by removing random edges can be considered as special instances of a natural generalization of the edge-removal procedures described above. In order to isolate via random cuttings the node with a specified label λ, with $1 \le \lambda \le n$, in a tree T with nodes labelled by $1, 2, \ldots, n$ we consider the following procedure:

(1) Pick one of the $n-1$ edges of the tree at random and remove it. This separates the tree T into two subtrees $\hat{T}$ and $\tilde{T}$. Let us assume that $\lambda \in \hat{T}$.
(2) Continue the edge-removal procedure recursively for the subtree $\hat{T}$, which contains the node labelled by λ, until node λ is isolated.

In the present paper we are studying this general edge-removal procedure by analyzing the random variable $X_{n,\lambda}$, with $1 \le \lambda \le n$, which counts the number of edge-cuts that are necessary to isolate the node labelled by λ in a random recursive tree of size n. Of course, the margin cases X_n (isolating the root) and Y_n (isolating the largest node) are contained as the special instances $X_{n,1}$ and $X_{n,n}$, respectively. This general edge-removal procedure has a natural interpretation in the model for the spread of a contamination in an organism mentioned above: instead of assuming that the root node is the contamination source we assume that a certain node λ is the contamination source, which one wants to isolate. Figure 1 and Figure 2 give examples of isolating certain nodes via the edge-removal procedure considered here.
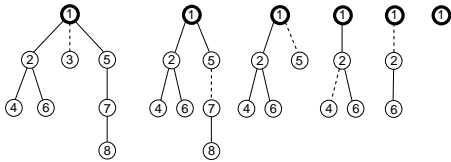

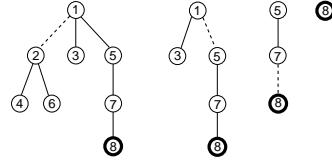Figure 1: Isolating the root in a size-8 recursive tree with 5 cuts.

Figure 2: Isolating the largest node in a size-8 recursive tree with 5 cuts.

We will analyze the random variable $X_{n,\lambda}$ from "both ends", i. e., we are studying $X_{n,\lambda}$ for small labels: $\lambda = l$, with $l \ge 1$ fixed and $n \to \infty$, and for large labels: $\lambda = n+1-l$, with $l \ge 1$ fixed and $n \to \infty$. By using a recursive approach we are able to give asymptotic expansions of the moments of the random variables $X_{n,l}$ and $X_{n,n+1-l}$, for l fixed and $n \to \infty$. For the instance of large labels we can apply the Theorem of Fréchet and Shohat and characterize the limit law of the normalized random variable $\frac{\log n}{n} X_{n,n+1-l}$ by its moments. It turns out that the random variable $X_{n,n+1-l}$, which counts the number of edge cuts necessary for isolating the l-th largest node, is (after scaling with $\frac{\log n}{n}$) asymptotically, for l fixed and $n \to \infty$, uniformly distributed on $[0, 1]$. For the instance of small labels we can show that $\frac{\log n}{n} X_{n,l}$ converge, for l fixed and $n \to \infty$, in probability to 1, but it turns out that a zero-mean and unit-variance normalization of $X_{n,l}$ has (for $s \ge 2$) s-th moments of order $\log^{\frac{s}{2}-1} n$. Thus existence of the limit law (and in the affirmative case a characterization of the limit law) of this normalized random variable cannot be shown by the method of moments. This was already observed for the special instance of isolating the root, i. e., for $X_{n,1}$, in [12].

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. $\Psi(x) := \frac{\Gamma'(x)}{\Gamma(x)}$ denotes the Psi

function (= digamma function), and $H_n := \sum_{k=1}^n \frac{1}{k}$ and $H_n^{(2)} := \sum_{k=1}^n \frac{1}{k^2}$ denote the first and second order harmonic numbers. Note that the Psi function is related to the harmonic numbers by the equation $\Psi(n+1) = H_n - \gamma$, where γ is the Euler-Mascheroni constant. Further we denote by $X \stackrel{(d)}{=} Y$ the equality in distribution of the random variables X and Y, and by $X_n \stackrel{(d)}{\longrightarrow} X$ the weak convergence, i. e., the convergence in distribution, of the sequence of random variables X_n to a random variable X. The symbol $[\![A]\!]$ denotes Iverson's bracket for the predicate A, i. e.$[\![A]\!] = 1$, if A is true and 0 otherwise.

Remarks:
(i) We want to remark that for the problem of isolating the root node of a tree via random cuttings Janson [7] gave an alternative approach by establishing a very useful connection between the number of cuts to isolate the root and the number of records when assigning random values to the edges of the tree. We want to sketch in the following that one can extend the arguments used in [7] to give also a connection between the number of cuts and the number of records for the problem of isolating a specified label λ. We consider a randomly chosen recursive tree of size n and attach to each edge e a random value γ_e, where we assume that the values γ_e are i. i. d. with an arbitrary continuous distribution. For a given label λ, with $1 \leq \lambda \leq n$, we call a value γ_e a record if it is the largest value in the path from the node labelled by λ to the edge e. Then it holds that the number of records is again distributed as $X_{n,\lambda}$.
(ii) Furthermore we want to remark that the cutting-down procedure for isolating the root node of a random recursive tree has also been used to give an alternative representation of the so called Bothausen-Sznitman coalescent (see [5]).

2. Results and outline of the proof

2.1. **Results.** For the s-th moments of the random variables $X_{n,l}$ and $X_{n,n+1-l}$ we get the following asymptotic expansions, for $l \geq 1$ fixed and $n \to \infty$, stated as Theorem 1 and Theorem 2.

Theorem 1. *The s-th moments $\mathbb{E}(X_{n,l}^s)$ of the number of random cuts necessary to isolate node l in a random recursive tree of size n are, for $l, s \geq 1$ fixed and $n \to \infty$, asymptotically given by*

$$\mathbb{E}(X_{n,l}^s) = \frac{n^s}{\log^s n} + \frac{\gamma_{l,s} \, n^s}{\log^{s+1} n} + \mathcal{O}\left(\frac{n^s}{\log^{s+2} n}\right), \tag{1}$$

where the constants $\gamma_{l,s}$ appearing in above expansion are given by

$$\gamma_{l,s} = (s+1)H_s - \frac{1}{2}\left(H_{l-1}^2 - H_{l-1}^{(2)}\right) - H_s H_{l-1} - s\gamma + \sum_{k=1}^{l-1} \frac{H_{s+k}}{k} - \sum_{k=1}^{l-1} \frac{1}{k^2 \binom{s+k}{k}},$$

where $\gamma = 0.57721$ is the Euler-Mascheroni constant.

Theorem 2. *The s-th moments $\mathbb{E}(X_{n,n+1-l}^s)$ of the number of random cuts necessary to isolate node $n+1-l$ in a random recursive tree of size n are, for $l, s \geq 1$ and $n \to \infty$, asymptotically given by*

$$\mathbb{E}(X_{n,n+1-l}^s) = \frac{n^s}{(s+1)\log^s n} + \mathcal{O}\left(\frac{n^s}{\log^{s+1} n}\right). \tag{2}$$

From these asymptotic expansions of the s-th moments we obtain the following results for the limiting behavior of $X_{n,l}$ and $X_{n,n+1-l}$, for l fixed and $n \to \infty$, given as Corollary 1 and Theorem 3.

Corollary 1. *The s-th centered moments $\mathbb{E}\left(\left[X_{n,l} - \mathbb{E}(X_{n,l})\right]^s\right)$ of the number of random cuts necessary to isolate node l in a random recursive tree of size n are, for $l \geq 1$, $s \geq 2$ fixed and $n \to \infty$, asymptotically given by*

$$\mathbb{E}\left(\left[X_{n,l} - \mathbb{E}(X_{n,l})\right]^s\right) = \frac{\delta_{l,s} \, n^s}{\log^{s+1} n} + \mathcal{O}\left(\frac{n^s}{\log^{s+2} n}\right), \tag{3}$$

where the constants $\delta_{l,s}$ appearing in above expansion are given by

$$\delta_{l,s} = \frac{(-1)^s}{s}\left(H_{l+s-1} - H_s\right) + \frac{(-1)^s}{s^2(s-1)} + \frac{(-1)^s}{s^2\binom{l+s-1}{l-1}}.$$

Thus the scaled random variable $\frac{\log n}{n} X_{n,l}$ converges, for $l \geq 1$ fixed and $n \to \infty$, in probability to 1 with convergence of all moments.

Theorem 3. *The limiting distribution of the normalized random variable $\frac{\log n}{n} X_{n,n+1-l}$ is, for $l \geq 1$ fixed and $n \to \infty$, a standard uniform distribution U_1 with support $[0,1)$:*

$$\frac{\log n}{n} X_{n,n+1-l} \xrightarrow{(d)} X, \quad X \overset{(d)}{=} U_1. \tag{4}$$

2.2. **Outline of the proof.** In order to show our results we will basically use a recursive approach, which allows to describe the number of random cuts necessary to isolate label l in a random recursive tree of size n via the corresponding quantities for smaller tree sizes $k < n$ and labels r not larger than l, i. e., $r \leq l$. Such a recursive approach is amenable, since it is well known (see [5, 12]) that random recursive trees satisfy a certain randomness-preservation property, which is stated in Subsection 3.1. Using this property we can easily give a distribution recurrence for $X_{n,l}$ (and X_{n+1-l}), where the behavior of the random variables considered are determined by the splitting probabilities $p_{(n,l),(k,r)}$ (and $\mathfrak{p}_{(n,l),(k,r)}$), which give the probability that when starting with a random size-n recursive tree and removing a random edge the subtree containing node l (node $n+1-l$) is of size k and where furthermore node l (node $n+1-l$) is the r-th smallest (the r-th largest) node in this subtree. Using a bijective argument we can give exact formulæ for these splitting probabilities. They are computed in Subsection 3.2 and given as Lemma 1.

From the distribution recurrences for $X_{n,l}$ and X_{n+1-l} we easily obtain recurrences for the s-th moments $\mathbb{E}\left(X_{n,l}^s\right)$ and $\mathbb{E}\left(X_{n,n+1-l}^s\right)$. In order to treat these recurrences we use a generating functions approach, which allows to translate these recurrences into linear differential equations for suitably introduced generating functions $M_{l,s}(z)$ and $N_{l,s}(z)$. Since we are able to determine the general solutions of the corresponding homogeneous differential equations it is possible to describe the solutions of these differential equations rather "explicitly". To determine the asymptotic growth behavior of $\mathbb{E}\left(X_{n,l}^s\right)$ and $\mathbb{E}\left(X_{n,n+1-l}^s\right)$, and thus essentially of the coefficients of the generating functions $M_{l,s}(z)$ and $N_{l,s}(z)$, we use singularity analysis (see [2]), i. e., we study the growth behavior of the functions in a neighborhood of the dominant singularity, together with certain lemmata for singular differentiation and integration; the corresponding Lemma 2 is stated in Subsection 4.1. Since, for a given pair (l,s), all generating functions $M_{r,j}(z)$ (and $N_{r,j}(z)$), with $r \leq l$, $j \leq s$ and $(r,j) \neq (l,s)$, appear in the inhomogeneous part of the differential equations determining $M_{l,s}(z)$ (and $N_{l,s}(z)$), we are forced to "pump out" the asymptotic expansions of the generating functions $M_{l,s}(z)$ (and $N_{l,s}(z)$) around the dominant singularity via induction on both parameters l and s. The corresponding computations for small labels: labels $l \geq 1$ fixed, are carried out in Section 4, whereas the computations for large labels: labels $n + 1 - l$, with $l \geq 1$ fixed, are given in Section 5.

3. THE RECURSIVE APPROACH

3.1. **Recurrences.** It has been observed in earlier work that random recursive trees satisfy the following "randomness-preservation" property, which will allow a recursive approach for the analysis of the parameter considered.

> Choose a random recursive tree of size n and then one of its $n - 1$ edges uniformly at random. Cutting this edge produces a pair of trees of size k and $n - k$. Then, conditioned on k and after an order preserving relabelling of the subtrees with labels $\{1, \ldots, k\}$ and $\{1, \ldots, n - k\}$, the subtrees themselves are random recursive trees of size k and $n - k$.

An important step for the recursive description of the probabilities $\mathbb{P}\{X_{n,l} = m\}$ is to introduce the splitting probabilities $p_{(n,l),(k,r)}$: they give the probability that when starting with a random size-n recursive tree and removing a random edge the subtree containing node l is of size k and where furthermore node l is the r-th smallest node in this subtree.

When we treat the analogous problem of isolating the node $n+1-l$ it is convenient to introduce the splitting probabilities $\mathfrak{p}_{(n,l),(k,r)}$: they give the probability that when starting with a random size-n recursive tree and removing a random edge the subtree containing node $n+1-l$ is of size k and where furthermore node $n + 1 - l$ is the r-th largest node in this subtree. Of course, these quantities are connected via the trivial relation

$$p_{(n,l),(k,r)} = \mathfrak{p}_{(n,n+1-l),(k,k+1-r)}. \tag{5}$$

From the recursive nature of the problem together with the randomness-preservation property immediately follows the distribution recurrence for the number of random cuts necessary to isolate the l-th smallest node in a random recursive tree of size n given below.

$$\mathbb{P}\{X_{n,l} = m\} = \sum_{r=1}^{l} \sum_{k=r}^{n-1} p_{(n,l),(k,r)} \mathbb{P}\{X_{k,r} = m - 1\}, \quad n \geq 2, \tag{6}$$

with initial value $\mathbb{P}\{X_{1,1} = 0\} = 1$. Furthermore, the distribution recurrence for the number of random cuts necessary to isolate the l-th largest node, i. e., node $n+1-l$, in a random recursive tree of size n is given by:

$$\mathbb{P}\{X_{n,n+1-l} = m\} = \sum_{r=1}^{l} \sum_{k=r}^{n-1} \mathfrak{p}_{(n,l),(k,r)} \mathbb{P}\{X_{k,k+1-r} = m - 1\}, \quad n \geq 2, \tag{7}$$

with initial value $\mathbb{P}\{X_{1,1} = 0\} = 1$.

The splitting probabilities $p_{(n,l),(k,r)}$ appearing in (6) are given by Lemma 1, which also determine the splitting probabilities $\mathfrak{p}_{(n,l),(k,r)}$ appearing in (7) due to equation (5).

3.2. **The splitting probabilities.** We obtain the following explicit formulæ for the splitting probabilities $p_{(n,l),(k,r)}$ appearing in (6).

Lemma 1. *The splitting probabilities $p_{(n,l),(k,r)}$ are, for $1 \leq l \leq n$, $1 \leq r \leq k$, $1 \leq k \leq n - 1$ and $n \geq 2$, given as follows:*

$$p_{(n,l),(k,r)} = \begin{cases} \left[(l-1)\binom{n-l}{n-k} + \binom{n-l+1}{n-k+1} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!}, & r = l, \\ \left[\binom{l-1}{r}\binom{n-l}{k-r} + \binom{l-1}{r-2}\binom{n-l}{k-r} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!}, & r < l. \end{cases}$$

Proof. If we remove an edge e of a size-n recursive tree we split the tree into two subtrees: we denote with T' the subtree containing the original root, i. e., label 1, and with T'' the other subtree, which is rooted at the vertex adjacent to the edge e that was cut. After an order preserving relabelling with labels $\{1, \ldots, |T'|\}$ and $\{1, \ldots, |T''|\}$ both subtrees can be considered as recursive trees. Furthermore we denote with B the arising subtree, which contains the node labelled by l in the original tree; we assume that this subtree has size k, with $1 \leq k \leq n - 1$. We distinguish now the cases $r = l$ and $r < l$.

If $r = l$ then it follows that $B = T'$. We want to determine the number of possibilities of removing an edge e of a recursive tree of size n leading (after an order preserving relabelling) to the pair (T', T'') of subtrees. To do this we count the number of different ways of distributing the labels $\{1, \ldots, n\}$ order preserving to T' and T'' and adjoining the root of T'' to a node of T' (by inserting edge e), such that the resulting tree is a recursive tree. We consider now the node of T' incident with e: if the node of T' incident with e has label j, with $1 \leq j \leq k$, then it follows that the labels of T'' must all be larger than j. For $1 \leq j \leq l$ we can choose $n - k$ of the labels $l + 1, l + 2, \ldots, n$ and distribute them order preserving to T'', whereas the remaining labels are distributed order preserving to T', leading to $\binom{n-l}{n-k}$ possibilities. For $l + 1 \leq j \leq k$ we can choose $n - k$ of the labels $j + 1, j + 2, \ldots, n$ and distribute them order preserving to T'', whereas the remaining labels are distributed order preserving to T', leading to $\binom{n-j}{n-k}$ possibilities. Thus this quantity is independent of the actual choice of T' with

$|T'| = k$ and T'' with $|T''| = n - k$. Since there are $T_k = (k-1)!$ and $T_{n-k} = (n-k-1)!$ different recursive trees of size k and $n-k$, this leads together with the fact that there are $n-1$ ways of selecting an edge e for any of the $T_n = (n-1)!$ recursive trees of size n to the following formula:

$$
\begin{aligned}
p_{(n,l),(k,l)} &= \left[l \binom{n-l}{n-k} + \sum_{j=l+1}^{k} \binom{n-j}{n-k} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!} \\
&= \left[(l-1) \binom{n-l}{n-k} + \binom{n-l+1}{n-k+1} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!},
\end{aligned}
$$

appealing to a well known identity.

If $r < l$ we have to distinguish further between the two cases $B = T'$ and $B = T''$. If $B = T'$ and we distribute the labels $\{1, \dots, n\}$ order preserving to T' and T'' we have the restriction that exactly $l - r$ nodes of the nodes $2, \dots, l-1$ have to be in T''. If $B = T''$ then we have the restriction that exactly $r - 1$ nodes of the nodes $2, \dots, l-1$ have to be in T''. Proceeding the same way as before we obtain eventually the following formula.

$$
\begin{aligned}
p_{(n,l),(k,r)} &= \left[\binom{n-l}{n-k-(l-r)} \sum_{j=1}^{r-1} \binom{l-1-j}{l-r} + \binom{n-l}{k-r} \sum_{j=1}^{l-r} \binom{l-1-j}{r-1} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!} \\
&= \left[\binom{l-1}{r-2} \binom{n-l}{k-r} + \binom{l-1}{r} \binom{n-l}{k-r} \right] \frac{(k-1)!(n-k-1)!}{(n-1)(n-1)!}
\end{aligned}
$$

$\square$

4. Isolating nodes with small labels

4.1. **Singular differentiation and integration.** In order to treat the recurrences for the s-th moments of $X_{n,l}$ and $X_{n,n+1-l}$ that will be obtained in the sequel we use a generating functions approach, which leads "in principle" to exact formulæ for suitably introduced generating functions. To obtain asymptotic information for the s-th moments we will basically use singularity analysis of generating functions, i. e., the transfer lemmata of Flajolet and Odlyzko [2] to "translate" the asymptotic growth behavior of a generating function in the neighborhood of its dominant singularity into the growth behavior of its coefficients. For the functions studied in the present paper the unique dominant singularity is always located at $z = 1$, thus we will specialize the considerations given below to this case. In order to apply singularity analysis it is necessary that the functions involved are analytic for a domain larger than the circle of convergence, namely the functions have to be analytic for indented discs $\Delta := \Delta(\phi, \eta) = \{z : |z| < 1 + \eta, |\operatorname{Arg}(z-1)| > \phi\}$, with $\eta > 0$, $0 < \phi < \frac{\pi}{2}$. Such functions are called Δ-regular (see [1]). We want to point out that the functions considered in this paper are always Δ-regular, since they are generated from Δ-regular functions via basic arithmetical functions and the operations differentiation and integration.

We will require the following O-transfer lemma for a Δ-regular function with a certain growth estimate in a neighborhood of $z = 1$:

$$
f(z) = \mathcal{O}\left(\frac{1}{(1-z)^a \log^b \left(\frac{1}{1-z} \right)} \right), \text{ for } z \to 1 \quad \Longrightarrow \quad [z^n] f(z) = \mathcal{O}\left(\frac{n^{a-1}}{\log^b n} \right). \tag{8}
$$

together with an asymptotic expansion of the coefficients of the following functions:

$$
f(z) = \frac{1}{(1-z)^a \log^b \left(\frac{1}{1-z} \right)} \quad \Longrightarrow \quad [z^n] f(z) = \frac{n^{a-1}}{\Gamma(a) \log^b n} \left(1 + \frac{b \Psi(a)}{\log n} + \mathcal{O}\left(\frac{1}{\log^2 n} \right) \right), \tag{9}
$$

where both formulæ (8) and (9) hold (at least) for $a > 0$ and $b \geq 0$ (see [2]).

However, for a study of the functions appearing we require also lemmata, which describe the asymptotic behavior of the derivative $f'(z)$ and the antiderivative $\int_0^z f(t) dt$ of a Δ-regular function $f(z)$ in the neighborhood of the dominant singularity $z = 1$, supposed that the asymptotic behavior around $z = 1$ of the function $f(z)$ itself is of a certain kind. Such theorems are known as theorems for singular differentiation and integration and can be found, e. g., in [1]. But in the sequel we will

require slightly more general theorems than given there, which are stated in Lemma 2. The proof of this lemma is omitted, since one can essentially "repeat" the arguments used in the proof of the corresponding theorems given in [1].

Lemma 2 (Singular differentiation and integration). *Let $f(z)$ be a Δ-regular function (see [1]), an analytic function in the domain $\Delta := \Delta(\phi, \eta)$,*

$$\Delta(\phi, \eta) = \{z : |z| < 1 + \eta, |\operatorname{Arg}(z - 1)| > \phi\},$$

with $\eta > 0$, $0 < \phi < \frac{\pi}{2}$, satisfying, for $z \to 1$, the expansion

$$f(z) = \mathcal{O}\Big(\frac{1}{(1 - z)^a \log^b \left(\frac{1}{1-z}\right)}\Big),$$

for $a > 1$ and $b \geq 1$. Then $\int_0^z f(t)dt$ and $f'(z)$ are also Δ-regular and they admit, for $z \to 1$, the expansions

$$\int_0^z f(t)dt = \mathcal{O}\Big(\frac{1}{(1 - z)^{a-1} \log^b \left(\frac{1}{1-z}\right)}\Big), \quad \text{and} \quad f'(z) = \mathcal{O}\Big(\frac{1}{(1 - z)^{a+1} \log^b \left(\frac{1}{1-z}\right)}\Big).$$

4.2. **Expectations.** The first step in our proof of Theorem 1 is to show the special case $s = 1$, i. e., asymptotic expansions of the expectations $\mathbb{E}(X_{n,l})$, for l fixed and $n \to \infty$, given as Lemma 3.

Lemma 3. *The expectations $\mathbb{E}(X_{n,l})$ of the number of random cuts necessary to isolate node l in a random recursive tree of size n are, for $l \geq 1$ fixed and $n \to \infty$, asymptotically given by*

$$\mathbb{E}(X_{n,l}) = \frac{n}{\log n} + \Big(4 - H_l - \gamma - \frac{1}{l}\Big)\frac{n}{\log^2 n} + \mathcal{O}\Big(\frac{n}{\log^3 n}\Big).$$

The proof of this lemma will be carried out by a generating functions approach using induction on l. The recurrences for the expectations $\mathbb{E}(X_{n,l})$ are obtained easily from the distribution recurrence (6) and are given by

$$\mathbb{E}(X_{n,l}) = 1 + \sum_{r=1}^{l} \sum_{k=r}^{n-1} p_{(n,l),(k,r)} \mathbb{E}(X_{k,r}), \tag{10}$$

with splitting probabilities given by Lemma 1.

We introduce for $l \geq 1$ the generating functions

$$M_l(z) := \sum_{n \geq l} (n - 1)^{\underline{l-1}} \mathbb{E}(X_{n,l}) z^{n-l}, \tag{11}$$

which allow to translate recurrence (10) by multiplication with $(n - 1)(n - 1)^{\underline{l-1}} z^{n-l}$ and summing up over $n \geq l$ into the following first order linear differential equation for $M_l(z)$, where the functions $M_r(z)$, with $r < l$, are appearing in the inhomogeneous part $R_l(z)$:

$$(1 - z) \log \Big(\frac{1}{1-z}\Big) \frac{d}{dz} M_l(z) + \Big((l - 1) - l \log \Big(\frac{1}{1-z}\Big)\Big) M_l(z) = R_l(z), \tag{12}$$

with inhomogeneous part

$$R_l(z) = \frac{(l - 1)!(l - 1 + z)}{(1 - z)^{l+1}} + \sum_{r=1}^{l-1} \Big[\binom{l - 1}{r} + \binom{l - 1}{r - 2}\Big] \frac{(l - r - 1)!}{(1 - z)^{l-r}} M_r(z),$$

and initial condition $M_l(0) = (l - 1)! \, \mathbb{E}(X_{l,l})$.

The homogeneous differential equation corresponding to (12) has the following general solution, with C an arbitrary constant:

$$M_l^{[h]}(z) = \frac{C}{(1 - z)^l \log^{l-1} \left(\frac{1}{1-z}\right)}.$$

The method of variation of constants leads then to the following particular solution of (12):

$$M_l^{[p]}(z) = \frac{1}{(1 - z)^l \log^{l-1} \left(\frac{1}{1-z}\right)} \int_0^z (1 - t)^{l-1} \log^{l-2} \Big(\frac{1}{1-t}\Big) R_l(t) dt, \tag{13}$$

and it can be shown that this particular solution matches the initial condition and is thus the wanted function, so $M_l(z) = M_l^{[p]}(z)$.

It will suffice to show the following asymptotic expansion of $M_l(z)$ around the dominant singularity $z = 1$, since a direct application of the transfer lemmata (8) and (9) leads then to Lemma 3:

$$M_l(z) = \frac{l!}{(1-z)^{l+1} \log\left(\frac{1}{1-z}\right)} + \frac{(l-1)!(4l-1-2lH_l)}{(1-z)^{l+1} \log^2\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+1} \log^3\left(\frac{1}{1-z}\right)}\right). \quad (14)$$

The proof of the expansion (14) will be done by induction. The case $l = 1$ gives the following solution, which already appeared in [10]:

$$M_1(z) = \frac{1}{1-z} \int_0^z \frac{t}{(1-t)^2 \log\left(\frac{1}{1-t}\right)} dt. \quad (15)$$

Integration by parts together with an application of Lemma 2 for singular integration leads then from (15) to the expansion (14) for the instance $l = 1$.

Now we assume that the functions $M_r(z)$ satisfy for all $r < l$ and a given $l > 1$ the asymptotic expansion (14). Plugging these expansions into the formula for the remainder term $R_l(z)$ given above easily leads to the expansion

$$R_l(z) = \frac{l!}{(1-z)^{l+1}} + \frac{(l-1)!(l^2 + 2l - 1 - 2lH_l)}{(1-z)^{l+1} \log\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+1} \log^2\left(\frac{1}{1-z}\right)}\right),$$

and furthermore to

$$\int_0^z (1-t)^{l-1} \log^{l-2}\left(\frac{1}{1-t}\right) R_l(t) dt = \frac{l! \log^{l-2}\left(\frac{1}{1-z}\right)}{1-z} + \frac{(l-1)!(4l-1-2lH_l) \log^{l-3}\left(\frac{1}{1-z}\right)}{1-z}$$

$$+ \mathcal{O}\left(\frac{\log^{l-4}\left(\frac{1}{1-z}\right)}{1-z}\right).$$

Due to formula (13) for $M_l(z) = M_l^{[p]}(z)$ expansion (14) is also shown for l. Thus (14) and as a consequence Lemma 3 is shown for all $l \geq 1$.

4.3. **Higher moments.** In order to show Theorem 1 for the asymptotic behavior of the moments $\mathbb{E}(X_{n,l}^s)$ we will continue our generating functions approach, where we will now use double induction on both parameters: the label l considered and the order s of the moments. To obtain a recurrence for the s-th moments of $X_{n,l}$ we multiply the distribution recurrence (6) with $m^s = \sum_{j=0}^s \binom{s}{j}(m-1)^j$ and sum up for $m \geq 1$. This leads to the following recurrence valid for $1 \leq l \leq n$ and $n \geq 2$ (with splitting probabilities given by Lemma 1).

$$\mathbb{E}(X_{n,l}^s) = \sum_{j=0}^s \binom{s}{j} \sum_{r=1}^l \sum_{k=r}^{n-1} p_{(n,l),(k,r)} \mathbb{E}(X_{k,r}^j). \quad (16)$$

We proceed as before and introduce for $l \geq 1$ and $s \geq 1$ the generating functions

$$M_{l,s}(z) := \sum_{n \geq l} (n-1)^{\underline{l-1}} \mathbb{E}(X_{n,l}^s) z^{n-l}. \quad (17)$$

Thus it holds $M_{l,1}(z) = M_l(z)$ for the functions $M_l(z)$ introduced in Subsection 4.2. Again we can translate above recurrence (16) into the following first order differential equation for $M_{l,s}(z)$, where the functions $M_{r,j}(z)$, with $r \leq l$, $j \leq s$ and $(r,j) \neq (l,s)$, appear in the inhomogeneous part $R_{l,s}(z)$:

$$(1-z) \log\left(\frac{1}{1-z}\right) \frac{d}{dz} M_{l,s}(z) + \left((l-1) - l \log\left(\frac{1}{1-z}\right)\right) M_{l,s}(z) = R_{l,s}(z), \quad (18)$$

with inhomogeneous part

$$R_{l,s}(z) = \sum_{j=1}^{s-1} \binom{s}{j} \left[\left(z - (1-z) \log\left(\frac{1}{1-z}\right)\right) \frac{d}{dz} M_{l,j}(z) + l \log\left(\frac{1}{1-z}\right) M_{l,j}(z)\right]$$

$$+ \frac{(l-1)!(l-1+z)}{(1-z)^{l+1}} + \sum_{j=1}^{s} \binom{s}{j} \sum_{r=1}^{l-1} \left[\binom{l-1}{r} + \binom{l-1}{r-2} \right] \frac{(l-r-1)!}{(1-z)^{l-r}} M_{r,j}(z),$$

and initial condition $M_{l,s}(0) = (l-1)! \, \mathbb{E}(X_{l,l}^s)$. Since the homogeneous differential equations corresponding to (18) and (12) coincide, we already know the shape of the general solution of (18):

$$M_{l,s}(z) = \frac{C}{(1-z)^l \log^{l-1}\left(\frac{1}{1-z}\right)} + \frac{1}{(1-z)^l \log^{l-1}\left(\frac{1}{1-z}\right)} \int_0^z (1-t)^{l-1} \log^{l-2}\left(\frac{1}{1-t}\right) R_{l,s}(t) dt,$$

with C an arbitrary constant. It turns out that the particular solution obtained for $C = 0$ matches the initial condition and we get thus

$$M_{l,s}(z) = \frac{1}{(1-z)^l \log^{l-1}\left(\frac{1}{1-z}\right)} \int_0^z (1-t)^{l-1} \log^{l-2}\left(\frac{1}{1-t}\right) R_{l,s}(t) dt. \tag{19}$$

Again it suffices to show the following asymptotic expansion around the dominant singularity $z = 1$ for the generating functions $M_{l,s}(z)$, since basic singularity analysis immediately leads from this expansion and (17) to Theorem 1.

$$M_{l,s}(z) = \frac{(l+s-1)!}{(1-z)^{l+s} \log^s\left(\frac{1}{1-z}\right)} + \frac{\alpha_{l,s}}{(1-z)^{l+s} \log^{s+1}\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+s} \log^{s+2}\left(\frac{1}{1-z}\right)}\right), \tag{20}$$

with constants

$$\alpha_{l,s} = (l+s-1)! \left[(s+1)H_s - sH_{l+s-1} - \frac{1}{2}\left(H_{l-1}^2 - H_{l-1}^{(2)}\right) - H_s H_{l-1} + \sum_{k=1}^{l-1} \frac{H_{s+k}}{k} - \sum_{k=1}^{l-1} \frac{1}{k^2 \binom{s+k}{k}} \right].$$

To show expansion (20) for all $l, s \geq 1$ we use induction on both parameters. The case $s = 1$ with arbitrary $l \geq 1$ was already treated in Subsection 4.2, where we computed the following expression, which coincides with (20):

$$M_{l,1}(z) = M_l(z) = \frac{l!}{(1-z)^{l+1} \log\left(\frac{1}{1-z}\right)} + \frac{(l-1)!(4l-1-2lH_l)}{(1-z)^{l+1} \log^2\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+1} \log^3\left(\frac{1}{1-z}\right)}\right).$$

Now we assume that for all pairs $(r, j) < (l, s)$, which means for $r \leq l$, $j \leq s$ and $(r, j) \neq (l, s)$, the functions $M_{r,j}(z)$ have in a neighborhood of the dominant singularity $z = 1$ the asymptotic expansion (20). We want to show that (20) also holds for the pair (l, s), where we may assume $s > 1$, since the case $s = 1$ is already shown. Plugging the expansions of the functions $M_{r,j}(z)$ into the formula for the inhomogeneous part $R_{l,s}(z)$ we obtain the expansion

$$R_{l,s}(z) = \frac{s(l+s-1)!}{(1-z)^{l+s} \log^{s-1}\left(\frac{1}{1-z}\right)} + \frac{\beta_{l,s}}{(1-z)^{l+s} \log^s\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+s} \log^{s+1}\left(\frac{1}{1-z}\right)}\right),$$

with

$$\beta_{l,s} = s(l+s-1)\alpha_{l,s-1} + (l+s-1)! \left[l - 2 - (s+1)(H_{l+s-1} - H_{s+1}) \right]$$
$$+ (l-1)!(s-1)! \left(\binom{l+s-1}{s} - 1 \right) - s(s-1)(l+s-2)!.$$

This further leads by an application of singular integration to the following expansion around $z = 1$:

$$\int_0^z (1-t)^{l-1} \log^{l-2}\left(\frac{1}{1-t}\right) R_{l,s}(t) dt = \frac{(l+s-1)!}{(1-z)^s \log^{s-l+1}\left(\frac{1}{1-z}\right)} + \frac{(l+s-1)!(s-l+1) + \beta_{l,s}}{s(1-z)^s \log^{s-l+2}\left(\frac{1}{1-z}\right)}$$
$$+ \mathcal{O}\left(\frac{1}{(1-z)^s \log^{s-l+3}\left(\frac{1}{1-z}\right)}\right).$$

Using equation (19) leads for $M_{l,s}(z)$ to the expansion

$$M_{l,s}(z) = \frac{(l+s-1)!}{(1-z)^{l+s} \log^s\left(\frac{1}{1-z}\right)} + \frac{(l+s-1)!(s-l+1) + \beta_{l,s}}{s(1-z)^{l+s} \log^{s+1}\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+s} \log^{s+2}\left(\frac{1}{1-z}\right)}\right),$$

and thus to the following recurrence for the coefficients $\alpha_{l,s}$:

$$\alpha_{l,s} = \frac{(l+s-1)!(s-l+1) + \beta_{l,s}}{s}$$

$$= (l+s-1)\alpha_{l,s-1} + \frac{(l+s-1)!\big[s-1-(s+1)(H_{l+s-1}-H_{s+1})\big]}{s}$$

$$+ \frac{(l-1)!(s-1)!\big(\binom{l+s-1}{s}-1\big)}{s} - (s-1)(l+s-2)!,$$

with initial value $\alpha_{l,1} = (l-1)!(4l-1-2lH_l)$. It is not hard to check that the coefficients $\alpha_{l,s}$ defined in equation (20) satisfy this recurrence. Thus expansion (20) and also Theorem 1 are shown for all $l,s \geq 1$.

4.4. **The centered moments.** It remains to prove Corollary 1 for the centered moments of $X_{n,l}$. To show (3) we use the corresponding expansion for the ordinary moments as given by Theorem 1. This leads to

$$\mathbb{E}\Big(\big[X_{n,l} - \mathbb{E}(X_{n,l})\big]^s\Big) = \sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j}\mathbb{E}(X_{n,l}^j)\big(\mathbb{E}(X_{n,l})\big)^{s-j}$$

$$= \left(\sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j}\right)\frac{n^s}{\log^s n} + \frac{\delta_{l,s}\, n^s}{\log^{s+1} n} + \mathcal{O}\Big(\frac{n^s}{\log^{s+2} n}\Big), \tag{21}$$

with constants

$$\delta_{l,s} = \sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j} f_{l,s}(j),$$

where the functions $f_{l,s}(j)$ are given as follows:

$$f_{l,s}(j) = (j+1)H_j - j\gamma - \frac{1}{2}\big(H_{l-1}^2 - H_{l-1}^{(2)}\big) - H_j H_{l-1} + \sum_{k=1}^{l-1}\frac{H_{j+k}}{k} - \sum_{k=1}^{l-1}\frac{1}{k^2\binom{j+k}{k}} + (s-j)\gamma_{l,1}.$$

Since it holds that $\sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j} = 0$, for all $s \geq 1$, the first term of (21) vanishes. To show Corollary 1 it only remains to simplify the expression for the constants $\delta_{l,s}$. One can do this, for instance, by using the calculus of higher order differences (see, e. g., [4]). Below we give two identities that can be shown by this method, which are required to obtain a closed form expression for $\delta_{l,s}$.

$$\sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j}\sum_{k=1}^{l-1}\frac{H_{j+k}}{k} = (-1)^{s-1}\Big(\frac{1}{s^2} - \frac{1}{s^2\binom{l+s-1}{l-1}}\Big),$$

$$\sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j}\sum_{k=1}^{l-1}\frac{1}{k^2\binom{j+k}{k}} = \frac{(-1)^s}{s}\big(H_{l-1} - H_{l+s-1} + H_s\big).$$

This eventually leads to Corollary 1.

5. Isolating nodes with large labels

5.1. **Isolating node n.** Now we are studying the random variable $X_{n,n+1-l}$, for $l \geq 1$ fixed. First we will consider the special case $l=1$, i. e., the instance of isolating the node with largest label n in a size-n recursive tree. We show the following lemma and prove thus the case $l=1$ of Theorem 2.

Lemma 4. *The s-th moments $\mathbb{E}(X_{n,n}^s)$ of the number of cuts necessary to isolate node n in a random recursive tree of size n are, for $s \geq 1$ and $n \to \infty$, asymptotically given by*

$$\mathbb{E}(X_{n,n}^s) = \frac{n^s}{(s+1)\log^s n} + \mathcal{O}\Big(\frac{n^s}{\log^{s+1} n}\Big).$$

An asymptotic equivalent of the expectation $\mathbb{E}(X_{n,n})$ together with a $\mathcal{O}$-bound for the variance of $X_{n,n}$ was already given in [8].

After simplifying the expressions for the splitting probabilities $\mathfrak{p}_{(n,1),(k,1)}$ as computed in Section 3 we can write the distribution recurrence (5) as follows:

$$
\begin{aligned}
\mathbb{P}\{X_{n,n} = m\} &= \sum_{k=1}^{n-1} \mathfrak{p}_{(n,1),(k,1)} \mathbb{P}\{X_{k,k} = m-1\} \\
&= \frac{1}{n-1} \sum_{k=1}^{n-1} \left(\frac{1}{k} + \frac{k-1}{(n-k)(n+1-k)} \right) \mathbb{P}\{X_{k,k} = m-1\}, \quad n \geq 2,
\end{aligned}
\tag{22}
$$

with $\mathbb{P}\{X_{1,1} = 0\} = 1$. For computing the s-th moments of $X_{n,n}$ we multiply (22) with $m^s = \sum_{j=0}^{s} \binom{s}{j}(m-1)^s$ and sum up for $m \geq 1$, which leads to the following recurrence:

$$
\mathbb{E}(X_{n,n}^s) = \frac{1}{n-1} \sum_{j=0}^{s} \binom{s}{j} \sum_{k=1}^{n-1} \left(\frac{1}{k} + \frac{k-1}{(n-k)(n+1-k)} \right) \mathbb{E}(X_{k,k}^j).
\tag{23}
$$

We treat (23) by introducing the generating functions

$$
N_{1,s}(z) := \sum_{n \geq 1} \frac{1}{n} \mathbb{E}(X_{n,n}^s) z^n.
\tag{24}
$$

In the sequel we will obtain the following asymptotic expansion of the generating function $N_{1,s}(z)$, which leads, after applying basic singularity analysis, to Lemma 4.

$$
N_{1,s}(z) = \frac{(s-1)!}{(s+1)(1-z)^s \log^s\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^s \log^{s+1}\left(\frac{1}{1-z}\right)} \right).
\tag{25}
$$

To show (25) we will use induction on s. First we have to consider the case $s = 1$. Plugging $s = 1$ into the recurrence (23) leads after multiplying with $(n-1)z^{n-1}$ and summing up for $n \geq 2$ to the following second order linear differential equation for the generating function $N_{1,1}(z)$:

$$
(1-z) \log\left(\frac{1}{1-z}\right) \frac{d^2}{dz^2} N_{1,1}(z) - \frac{1}{1-z} N_{1,1}(z) = \frac{z}{(1-z)^2},
\tag{26}
$$

with initial conditions $N_{1,1}(0) = 0$ and $\left(\frac{d}{dz} N_{1,1}(z)\right)|_{z=0} = \mathbb{E}(X_{1,1}) = 0$. The solution of the homogeneous differential equation corresponding to (26) is given as follows, with C_1, C_2 arbitrary constants:

$$
N_1^{[h]}(z) = C_1 N_1^{[h_1]}(z) + C_2 N_1^{[h_2]}(z) = C_1 \log\left(\frac{1}{1-z}\right) + C_2 \log\left(\frac{1}{1-z}\right) \int_\alpha^z \frac{dt}{\log^2\left(\frac{1}{1-t}\right)},
\tag{27}
$$

where we may choose any real $0 < \alpha < 1$. Since the Wronski determinant of the two homogeneous solutions equals one,

$$
N_1^{[h_1]}(z) \frac{d}{dz} N_1^{[h_2]}(z) - N_1^{[h_2]}(z) \frac{d}{dz} N_1^{[h_1]}(z) = 1,
\tag{28}
$$

a particular solution of (26) is given by

$$
\begin{aligned}
N_{1,1}^{[p]}(z) &= N_1^{[h_1]}(z) \int_0^z -b_{1,1}(t) N_1^{[h_2]}(t)\, dt + N_1^{[h_2]}(z) \int_0^z b_{1,1}(t) N_1^{[h_1]}(t)\, dt \\
&= \log\left(\frac{1}{1-z}\right) \left(\int_\alpha^z \frac{dt}{\log^2\left(\frac{1}{1-t}\right)} \int_0^z \frac{t}{(1-t)^3}\, dt - \int_0^z \frac{t}{(1-t)^3} \int_\alpha^t \frac{1}{\log^2\left(\frac{1}{1-u}\right)}\, du\, dt \right),
\end{aligned}
\tag{29}
$$

with $b_{1,1}(z) = \frac{z}{(1-z)^3 \log\left(\frac{1}{1-z}\right)}$. Using integration by parts we can simplify the particular solution as follows:

$$
N_{1,1}^{[p]}(z) = \log\left(\frac{1}{1-z}\right) \int_0^z \frac{t^2}{2(1-t)^2 \log^2\left(\frac{1}{1-t}\right)}\, dt.
\tag{30}
$$

It turns out that the particular solution (30) satisfies also the initial conditions of (26) and is thus the required solution, i. e., $N_{1,1}(z) = N_{1,1}^{[p]}(z)$. An application of partial integration and singular integration, Lemma 2, shows then the instance $s = 1$ of expansion (25).

Now we assume that expansion (25) is valid for all values $j < s$, with an arbitrary $s > 1$. To show the expansion also for s we again translate recurrence (23) by multiplying with $(n-1)z^{n-1}$ and summing up for $n \geq 2$ into the following inhomogeneous second order linear differential equation:

$$(1-z)\log\Big(\frac{1}{1-z}\Big)\frac{d^2}{dz^2}N_{1,s}(z) - \frac{1}{1-z}N_{1,s}(z) = R_{1,s}(z), \tag{31}$$

with inhomogeneous part

$$R_{1,s}(z) = \frac{z}{(1-z)^2} + \sum_{j=1}^{s-1}\binom{s}{j}\left[\Big(z-(1-z)\log\Big(\frac{1}{1-z}\Big)\Big)\frac{d^2}{dz^2}N_{1,j}(z) + \frac{N_{1,j}(z)}{1-z}\right],$$

and initial conditions $N_{1,s}(0) = 0$ and $\big(\frac{d}{dz}N_{1,s}(z)\big)|_{z=0} = \mathbb{E}\big(X_{1,1}^s\big) = 0$. Since we have already given by (27) the general solution of the corresponding homogeneous differential equation we can proceed as before and obtain by the method of variation of constants the following particular solution of (31):

$$N_{1,s}^{[p]}(z) = \log\Big(\frac{1}{1-z}\Big)\int_0^z \Big(\int_0^t \frac{R_{1,s}(u)}{1-u}\,du\Big)\frac{dt}{\log^2\big(\frac{1}{1-t}\big)}. \tag{32}$$

It can be checked easily that (32) also satisfies the initial conditions and is thus the solution of (31), i. e., $N_{1,s}(z) = N_{1,s}^{[p]}(z)$.

Plugging the expansions (25) of $N_{1,j}(z)$, for $j < s$, into the formula for $R_{1,s}(z)$ we obtain the following expansion of the inhomogeneous part in a neighborhood of $z = 1$:

$$R_{1,s}(z) = \frac{s!}{(1-z)^{s+1}\log^{s-1}\big(\frac{1}{1-z}\big)} + \mathcal{O}\Big(\frac{1}{(1-z)^{s+1}\log^s\big(\frac{1}{1-z}\big)}\Big).$$

This gives by applying singular integration, Lemma 2,

$$\Big(\int_0^t \frac{R_{1,s}(u)}{1-u}\,du\Big)\frac{1}{\log^2\big(\frac{1}{1-t}\big)} = \frac{s!}{(s+1)(1-t)^{s+1}\log^{s+1}\big(\frac{1}{1-t}\big)} + \mathcal{O}\Big(\frac{1}{(1-t)^{s+1}\log^{s+2}\big(\frac{1}{1-t}\big)}\Big),$$

and eventually leads due to (32), after a further application of Lemma 2, that expansion (25) also holds for s. Thus (25) and as a consequence Lemma 4 are shown for all $s \geq 1$.

5.2. **Expectations for large nodes.** Next we show the following asymptotic expansion of the expectation $\mathbb{E}(X_{n+1-l})$, for l fixed. Lemma 5 gives thus the special case $s = 1$ of Theorem 2.

Lemma 5. *The expectations $\mathbb{E}(X_{n,n+1-l})$ of the number of random cuts necessary to isolate node $n+1-l$ in a random recursive tree of size n are, for $l \geq 1$ fixed and $n \to \infty$, asymptotically given by*

$$\mathbb{E}(X_{n,n+1-l}) = \frac{n}{2\log n} + \mathcal{O}\Big(\frac{n}{\log^2 n}\Big).$$

To show this lemma we study the following recurrence for $\mathbb{E}(X_{n+1-l})$, which is obtained from the distribution recurrence (7) after multiplying with $m = (m-1)+1$ and summing up for $m \geq 1$:

$$\mathbb{E}(X_{n,n+1-l}) = 1 + \sum_{r=1}^l \sum_{k=r}^{n-1} \mathfrak{p}_{(n,l),(k,r)}\mathbb{E}(X_{k,k+1-r}). \tag{33}$$

After simplifying the expressions of the splitting probabilities $\mathfrak{p}_{(n,l),(k,r)}$ as given by (5) and Lemma 1 we obtain

$$\mathfrak{p}_{(n,l),(k,r)} = [\![k \leq n+r-l]\!]\frac{\binom{l-1}{r-1}}{(n-1)(n-1)^{\underline{l-1}}}\big[(k-1)^{\underline{r-2}}(n-k-1)^{\underline{l-r}} + (k-1)^{\underline{r}}(n-k-1)^{\underline{l-r-2}}\big]$$

$$+ [\![k = n+r-l]\!]\binom{l}{r-1}\frac{(l-r-1)!}{(n-1)(n-1)^{\underline{l-r}}}, \quad r \leq l,$$

$$\tag{34}$$

where we use the convention $(j-1)^{\underline{-p}} := (j^{\overline{p}})^{-1}$, $p \in \mathbb{N}$, see e. g. [4].

To treat recurrence (33) we introduce for $l \geq 1$ the following generating functions:

$$N_{l,1}(z) := \sum_{n \geq l} (n-1)^{\underline{l-2}} \mathbb{E}(X_{n,n+1-l}) z^{n+1-l}. \tag{35}$$

Note that this definition also holds for $l = 1$, where it matches the definition of $N_{1,1}(z)$ given by (24). Again, due to an application of basic singularity analysis, it suffices to show that $N_{l,1}(z)$ admits the following expansion in a neighborhood of the dominant singularity $z = 1$, which proves Lemma 5.

$$N_{l,1}(z) = \frac{(l-1)!}{2(1-z)^l \log\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^l \log^2\left(\frac{1}{1-z}\right)}\right). \tag{36}$$

We will show expansion (36) by induction on l. The case $l = 1$ was treated already in Subsection 5.1, where it turned out that (36) holds.

Above recurrence (33) can be translated into the following second order linear differential equation for $N_{l,1}(z)$, where the functions $N_{r,1}(z)$, with $r < l$, are all appearing in the inhomogeneous part $R_{l,1}(z)$. $R_{l,1}(z)$ is now a bit "unpleasant", since one had to consider separately the four cases $r = 1$, $1 < r < l-1$, $r = l-1$ and $r = l$. One obtains

$$(1-z)\log\left(\frac{1}{1-z}\right)\frac{d^2}{dz^2}N_{l,1}(z) + (l-1)\frac{d}{dz}N_{l,1}(z) - \frac{1}{1-z}N_{l,1}(z) = R_{l,1}(z), \tag{37}$$

with inhomogeneous part

$$R_{l,1}(z) = \frac{(l-1)!(l-1+z)}{(1-z)^{l+1}} + \sum_{r=1}^{l-1}\sum_{n \geq l}\left[\sum_{k=r}^{n-1}(n-1)(n-1)^{\underline{l-1}}\mathfrak{p}_{(n,l),(k,r)}\mathbb{E}(X_{k,k+1-r})\right]z^{n-l}$$

$$= \frac{(l-1)!(l-1+z)}{(1-z)^{l+1}} + \sum_{r=1}^{l-1}\binom{l}{r-1}(l-r-1)!\frac{d}{dz}N_{r,1}(z)$$

$$+ \sum_{r=1}^{l-1}\binom{l-1}{r-1}\frac{(l-r)!}{(1-z)^{l-r+1}}N_{r,1}(z) + \sum_{r=1}^{l-2}\binom{l-1}{r-1}(l-r-2)!\left(\frac{1}{(1-z)^{l-r-1}}-1\right)\frac{d^2}{dz^2}N_{r,1}(z)$$

$$+ (l-1)\log\left(\frac{1}{1-z}\right)\frac{d^2}{dz^2}N_{l-1,1}(z),$$

and initial conditions $N_{l,1}(0) = 0$, $\left(\frac{d}{dz}N_{l,1}\right)\big|_{z=0} = (l-1)!\,\mathbb{E}(X_{l,1})$. One gets that the homogeneous differential equation corresponding to (37) has the following general solution, with C_1, C_2 arbitrary constants:

$$N_{l,1}^{[h]}(z) = C_1\left(l-1+\log\left(\frac{1}{1-z}\right)\right) + C_2\left(l-1+\log\left(\frac{1}{1-z}\right)\right)\int_\alpha^z \frac{dt}{\log^{l-1}\left(\frac{1}{1-t}\right)\left[l-1+\log\left(\frac{1}{1-t}\right)\right]^2},$$

where we may choose any real $0 < \alpha < 1$. Applying the method of variation of constants we obtain the following particular solution of (37):

$$N_{l,1}^{[p]}(z) = \left(l-1+\log\left(\frac{1}{1-z}\right)\right)\times \tag{38}$$

$$\times \int_0^z\left(\int_0^t \frac{\log^{l-2}\left(\frac{1}{1-u}\right)\left(l-1+\log\left(\frac{1}{1-u}\right)\right)R_{l,1}(u)}{1-u}du\right)\frac{1}{\log^{l-1}\left(\frac{1}{1-t}\right)\left(l-1+\log\left(\frac{1}{1-z}\right)\right)^2}dt.$$

It turns out that (38) matches the initial conditions of (37) and is thus the required solution, i. e., $N_{l,1}(z) = N_{l,1}^{[p]}(z)$.

Using the explicit description (38) of the generating function $N_{l,1}(z)$ we can show the required expansion (36). We assume that expansion (36) of $N_{r,1}(z)$ holds, for all $r < l$ with an arbitrary $l > 1$. Plugging these expansions into the formula of the inhomogeneous part $R_{l,1}(z)$ given above we obtain, after a heavy use of singular differentiation, the following expansion around $z = 1$:

$$R_{l,1}(z) = \frac{(l+1)!}{2(1-z)^{l+1}} + \mathcal{O}\left(\frac{1}{(1-z)^{l+1}\log\left(\frac{1}{1-z}\right)}\right). \tag{39}$$

Plugging (39) into (38) we obtain after an application of partial integration and singular integration that expansion (36) also holds for $l > 1$. Thus expansion (36) and as a consequence Lemma 5 are shown for all $l \geq 1$.

5.3. **Higher moments for large nodes.** The method applied in Subsection 5.1-5.2 for a study of the s-th moments of $X_{n,n}$ and the expectations of $X_{n,n+1-l}$ can be extended naturally to show Theorem 2, leading to an asymptotic expansion of the s-th moments of the number of random cuts necessary to isolate the l-th largest node in a random size-n recursive tree.

Again we use the distribution recurrence (7) to give a recurrence for the s-th moments of $X_{n,n+1-l}$. After multiplying with $m^s = \sum_{j=0}^{s} \binom{s}{j}(m-1)^j$ and summing up for $m \geq 1$ we obtain:

$$\mathbb{E}\big(X_{n,n+1-l}^s\big) = \sum_{j=0}^{s} \binom{s}{j} \sum_{r=1}^{l} \sum_{k=r}^{n-1} \mathfrak{p}_{(n,l),(k,r)} \mathbb{E}\big(X_{k,k+1-r}^j\big). \tag{40}$$

It is now appropriate to introduce for $l, s \geq 1$ the generating functions

$$N_{l,s}(z) := \sum_{n \geq l}(n-1)^{\underline{l-2}}\, \mathbb{E}\big(X_{n,n+1-l}^s\big) z^{n+1-l}, \tag{41}$$

which are generalizations of $N_{1,s}(z)$ and $N_{l,1}(z)$ as used in Subsection 5.1-5.2. It is again sufficient to show that $N_{l,s}(z)$ admits the following expansion in a neighborhood of the dominant singularity $z = 1$:

$$N_{l,s}(z) = \frac{(l+s-2)!}{(s+1)(1-z)^{l+s-1}\log^s\left(\frac{1}{1-z}\right)} + \mathcal{O}\Big(\frac{1}{(1-z)^{l+s-1}\log^{s+1}\left(\frac{1}{1-z}\right)}\Big), \tag{42}$$

which will be done by induction on both parameters, l and s. Basic singularity analysis leads then directly to Theorem 2. The margin cases $l = 1$, $s \geq 1$ and $l \geq 1$, $s = 1$ are already shown in Subsection 5.1-5.2.

We proceed by translating the recurrence (40) into the following second order linear differential equation for $N_{l,s}(z)$:

$$(1-z)\log\Big(\frac{1}{1-z}\Big)\frac{d^2}{dz^2}N_{l,s}(z) + (l-1)\frac{d}{dz}N_{l,s}(z) - \frac{1}{1-z}N_{l,s}(z) = R_{l,s}(z), \tag{43}$$

with inhomogeneous part

$$R_{l,s}(z) = \frac{(l-1)!(l-1+z)}{(1-z)^{l+1}} + \sum_{j=1}^{s} \binom{s}{j} \sum_{r=1}^{l-1} \binom{l}{r-1}(l-r-1)! \frac{d}{dz}N_{r,j}(z)$$

$$+ \sum_{j=1}^{s} \binom{s}{j} \sum_{r=1}^{l-1} \binom{l-1}{r-1} \frac{(l-r)!}{(1-z)^{l-r+1}} N_{r,j}(z) + \sum_{j=1}^{s} \binom{s}{j}(l-1)\log\Big(\frac{1}{1-z}\Big)\frac{d^2}{dz^2}N_{l-1,j}(z)$$

$$+ \sum_{j=1}^{s} \binom{s}{j} \sum_{r=1}^{l-2} \binom{l-1}{r-1}(l-r-2)!\Big(\frac{1}{(1-z)^{l-r-1}} - 1\Big)\frac{d^2}{dz^2}N_{r,j}(z)$$

$$+ \sum_{j=1}^{s-1} \binom{s}{j}\frac{1}{1-z}N_{l,j}(z) + \sum_{j=1}^{s-1} \binom{s}{j}\Big(z - (1-z)\log\Big(\frac{1}{1-z}\Big)\Big)\frac{d^2}{dz^2}N_{l,j}(z),$$

and initial conditions $N_{l,s}(0) = 0$ and $\big(\frac{d}{dz}N_{l,s}\big)\big|_{z=0} = (l-1)!\,\mathbb{E}(X_{l,1}^s)$.

Since the homogeneous solution corresponding to (42) was already computed in Subsection 5.2 we obtain again a particular solution by applying the method of variation of constants. We get after simplifications

$$N_{l,s}^{[p]}(z) = \Big(l-1+\log\Big(\frac{1}{1-z}\Big)\Big)\times \tag{44}$$

$$\times \int_0^z \Big(\int_0^t \frac{\log^{l-2}\left(\frac{1}{1-u}\right)\big(l-1+\log\left(\frac{1}{1-u}\right)\big)R_{l,s}(u)}{1-u}\,du\Big)\frac{1}{\log^{l-1}\left(\frac{1}{1-t}\right)\big(l-1+\log\left(\frac{1}{1-z}\right)\big)^2}\,dt.$$

Now we assume that for all pairs $(r, j) < (l, s)$, which means for $r \leq l$, $j \leq s$ and $(r, j) \neq (l, s)$, the functions $N_{r,j}(z)$ admit the asymptotic expansion (42) in a neighborhood of the dominant singularity $z = 1$. We want to show that (42) also holds for the pair (l, s), where we may assume $s > 1$ and $l > 1$, since the margin cases are already treated. Plugging the expansions of the functions $N_{r,j}(z)$ into the formula for the inhomogeneous part $R_{l,s}(z)$ and using singular differentiation we obtain that the main contributions in the expansion of $R_{l,s}(z)$ around $z = 1$ are stemming from the terms $\binom{s}{s-1} \frac{d^2}{dz^2} N_{l,s-1}(z)$ and $\binom{s}{s}(l-1) \log\left(\frac{1}{1-z}\right) \frac{d^2}{dz^2} N_{l-1,s}(z)$. This gives

$$R_{l,s}(z) = \frac{\binom{s}{s-1}(l+s-1)!}{s(1-z)^{l+s} \log^{s-1}\left(\frac{1}{1-z}\right)} + \frac{(l-1)(l+s-1)!}{(s+1)(1-z)^{l+s} \log^{s-1}\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+s} \log^s\left(\frac{1}{1-z}\right)}\right)$$

$$= \frac{(l+s)!}{(s+1)(1-z)^{l+s} \log^{s-1}\left(\frac{1}{1-z}\right)} + \mathcal{O}\left(\frac{1}{(1-z)^{l+s} \log^s\left(\frac{1}{1-z}\right)}\right). \tag{45}$$

Plugging (45) into equation (44) we obtain after applying singular integration that expansion (42) also holds for pairs (l, s), with $l > 1$ and $s > 1$. Thus (42) is shown for all $l \geq 1$ and $s \geq 1$, which completes the proof of Theorem 2.

5.4. **The limiting distribution.** From the asymptotic expansion of the s-th moments of $X_{n,n+1-l}$ as given by Theorem 2 we obtain for the s-th moments of the scaled random variable $\tilde{X}_{n,n+1-l} := \frac{\log n}{n} X_{n,n+1-l}$, for $l \geq 1$, $s \geq 0$ and $n \to \infty$:

$$\mathbb{E}\left(\tilde{X}_{n,n+1-l}^s\right) \to \frac{1}{s+1}.$$

If we consider a random variable X, which has a standard uniform distribution U_1 with support $[0, 1)$ then the s-th moments of X are, for $s \geq 0$, given as follows:

$$\mathbb{E}(X^s) = \frac{1}{s+1}.$$

A direct application of the Theorem of Fréchet and Shohat (the second central limit theorem, see, e. g, [6]) proves then Theorem 3.

6. CONCLUSION

Using a generating functions approach in combination with singularity analysis and lemmata for singular differentiation and integration we obtain distribution results for the number of random cuts necessary to isolate large nodes and small nodes in random recursive trees via random cuttings. Although the recurrences given in the present paper could be applied to treat the parameter studied for arbitrary labels l and tree sizes n, it turns out that for general growth rates of l compared to n (e. g., when isolating nodes in the central region $l/n \sim \rho$, with $0 < \rho < 1$) it seems considerably more difficult to attack the problem.

REFERENCES

[1] P. Flajolet, J. A. Fill and N. Kapur, Singularity Analysis, Hadamard Products, and Tree Recurrences, *Journal of Computational and Applied Mathematics* 174, 271–313, 2005.
[2] P. Flajolet and A. Odlyzko, Singularity Analysis of Generating Functions, *SIAM Journal on Discrete Mathematics* 3, 216–240, 1990.
[3] J. L. Gastwirth and P. K. Bhattacharya, Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable, *Operations Research* 32, 527–536, 1984.
[4] R. Graham, D. Knuth and O. Patashnik, *Concrete Mathematics*, Second Edition, Addison-Wesley, Reading, 1994.
[5] C. Goldschmidt and J. B. Martin, Random recursive trees and the Bolthausen-Sznitman coalescent, *Electronic Journal of Probability* 10, paper no. 21, 718–745, 2005.
[6] M. Loève, *Probability Theory I*, 4th Edition, Springer, New York, 1977.
[7] S. Janson, Random cutting and records in deterministic and random trees, *Random Structures and Algorithms*, to appear, 2006.
[8] M. Javanian and M. Q. Vahidi-Asl, Isolating the most recent entry in a random recursive tree by random cuts, *Journal of Applied Mathematics & Computing* 16, 115–123, 2004.

[9] H. Mahmoud and R. Smythe, A Survey of Recursive Trees, *Theoretical Probability and Mathematical Statistics* 51, 1–37, 1995.

[10] A. Meir and J. W. Moon, Cutting down recursive trees, *Mathematical Biosciences* 21, 173–181, 1974.

[11] D. Najock and C. C. Heyde, On the number of terminal vertices in certain random trees with an application to stemma construction in philology, *Journal of Applied Probability* 19, 675–680, 1982.

[12] 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.

MARKUS KUBA, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDNER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA

E-mail address: `markus.kuba+e104@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`