On the Number of Descendants and Ascendants in Random Search Trees*

Conrado Martínez[†] Alois Panholzer[‡] Helmut Prodinger[‡]

February 10, 1998

Abstract

We consider here the probabilistic analysis of the number of descendants and the number of ascendants of a given internal node in a random search tree. The performance of several important algorithms on search trees is closely related to these quantities. For instance, the cost of a successful search is proportional to the number of ascendants of the sought element. On the other hand, the probabilistic behavior of the number of descendants is relevant for the analysis of paged data structures and for the analysis of the performance of quicksort, when recursive calls are not made on small subfiles. We also consider the number of ascendants and descendants of a random node in a random search tree, i.e., the *grand averages* of the quantities mentioned above. We address these questions for standard binary search trees and for locally balanced search trees. These search trees were introduced by Poblete and Munro and are binary search trees such that each subtree of size 3 is balanced; in other words, binary search trees where there are not two adjacent internal nodes with only one son each.

In this work, we follow a purely combinatorial approach, extensively using generating functions, and derive exact and asymptotic expressions for the probability distribution and moments of some of the considered quantities, finding several new results as well as alternative derivations for already known results.

1. Introduction

Binary search trees are among the most important and commonly used data structures, their applications spanning a wide range of the areas of Computer Science. Standard binary search trees (BSTs, for short) are still the subject of active research, see for instance the recent articles [2, 27]. Deepening our knowledge about binary search trees is interesting in its own; moreover, most of this knowledge can be translated and applied to other data structures such as heap ordered trees, k-d-trees [32], and to important algorithms like quicksort and Hoare's Find algorithm for selection (also known as quickselect) [11, 12, 29, 30].

We assume that the reader is already familiar with binary search trees and the basic algorithms to manipulate them [19, 30, 8]. Height and weight-balanced versions of the binary search trees, like AVL and red-black trees [1, 10], have been proposed and find many useful applications, since all of them guarantee good worst-case performance of both searches and updates.

*This research was partly done while the third author was visiting the CRM (Centre de Recerca Matemàtica, Institut d'Estudis Catalans). The first author was supported by the ESPRIT Long Term Research Project ALCOM IT (contract no. 20244). The second author was supported by the FWF Project 12599-MAT. All 3 authors are supported by the Project 16/98 of Acciones Integrades 1998/99.

[†]Departament de Llenguatges i Sistemes Informàtics. Polytechnical University of Catalonia. Pau Gargallo 5, E-08028 Barcelona, Spain. E-mail: `Conrado.Martinez@lsi.upc.es`

[‡]Institut für Algebra und Diskrete Mathematik. Technical University of Vienna. Wiedner Hauptstraße 8–10, A-1040 Vienna, Austria. E-mail: `e9125354@fbma.tuwien.ac.at (Alois Panholzer)`, E-mail: `Helmut.Prodinger@tuwien.ac.at`

Locally balanced search trees (LBSTs) were introduced by Bell [4] and Walker and Wood [33], and throughly analyzed by Poblete and Munro in [26]. LBSTs have been proposed as an alternative to more complex balancing schemes for search trees. In these search trees, only local rebalancing is made; after each insertion, local rebalancing is applied to ensure that all subtrees of size 3 in the tree are complete[1]. The basic idea of the heuristic is that the construction of poorly balanced trees becomes less likely. A similar idea, namely, selecting a sample of 3 elements and taking the median of the sample as the pivot element for partitioning in algorithms like quicksort and quickselect has been shown to yield significant improvements in theory and practice [29, 16].

Random search trees, either random BSTs or random LBSTs, are search trees built by performing n random insertions into an initially empty tree [19, 23]. An insertion of a new element into a search tree of size k is said to be random, if the new element falls with equal probability into any of the $k+1$ intervals defined by the k keys already present in the tree (equivalently, the new element replaces any of the $k+1$ external nodes in the tree with equal probability). Random search trees can also be defined as the result of the insertion of the elements of a random permutation of $\{1, \ldots, n\}$ into an initially empty tree.

Ascendants and descendants of the j^{th} internal node of a random search tree of size n are denoted $A_{n,j}$ and $D_{n,j}$, respectively. Besides the two afore mentioned random variables, we also consider other random variables: the number of descendants D_n and the number of ascendants A_n of a randomly chosen internal node in a random search tree of size n. This corresponds to averaging $D_{n,j}$ and $A_{n,j}$ over j. We remark, that all the distributions, as well as the expectations $\mathbb{E}[X]$ and probabilities $\mathbb{P}[X]$ are induced by the creation process of the random search trees (BSTs resp. LBSTs). The number of descendants and the number of ascendants in random BSTs have been investigated in several previous works ([3, 5, 22, 21, 20]). The number of ascendants of a random node in a random LBST has been studied in [26, 25].

We define the number of descendants $D_{n,j}$ as the size of the subtree rooted at the j^{th} node, so we count the j^{th} node as a descendant of itself. The number of ascendants $A_{n,j}$ is the number of internal nodes in the path from the root of the tree to the j^{th} node, both included. It is worth mentioning the following symmetry property (which is very easy to prove) for the random variables we are going to consider.[2]

Proposition 1.1 *For any $n > 0$ and any $1 \le j \le n$,*

$$\begin{aligned} D_{n,j} &= D_{n,n+1-j}, \\ A_{n,j} &= A_{n,n+1-j}. \end{aligned}$$

The performance of a successful search is obviously proportional to the number of ascendants of the sought internal node. The next proposition states this relation, as well as other interesting relationships that hold for both random BSTs and random LBSTs.

Proposition 1.2 *Consider a random search tree of size n and let*

$S_{n,j}$ = # *of comparisons in a successful search for the j^{th} element,*

S_n = # *of comparisons in a successful search for a randomly chosen element,*

U_n = # *of comparisons in a unsuccessful search for a randomly chosen external node,*

$P_{n,j}$ = *depth of the j^{th} element,*

[1]The generalization of the local rebalancing heuristic to subtree sizes larger than 3 is straightforward.

[2]We remark, that here and in the sequel equalities between random variables are equalities in distribution, which is often denoted by $\stackrel{\mathrm{d}}{=}$.

$$I_n = \sum_{1 \le j \le n} P_{n,j} = \textit{internal path length},$$

Then,

$$\begin{aligned} S_{n,j} &= P_{n,j} + 1 = A_{n,j}, \\ S_n &= A_n, \\ \mathbb{E}[U_n] &= \frac{n}{n+1}\left(1 + \mathbb{E}[A_n]\right), \\ \mathbb{E}[I_n] &= n\left(\mathbb{E}[A_n] - 1\right), \\ \mathbb{E}[A_n] &= \mathbb{E}[D_n]. \end{aligned}$$

There is also a close relationship between the performance of quickselect [11, 18, 16] and the number of ascendants.

Proposition 1.3 *Let $F_{n,j}$ be the number of recursive calls made by quickselect to select the j^{th} element out of n elements. Then*

$$F_{n,j} = A_{n,j}.$$

If we consider $A_{n,j}$ in random BSTs, then this corresponds to the selection of the pivots at random in each phase of quickselect. If we consider $A_{n,j}$ in random LBSTs, then the proposition applies for the variant of quickselect that uses the median of a random sample of three elements as the pivot in each partitioning phase.

The study of the number of descendants has applications in the context of paged trees (see for instance [19, 13]). A paged binary search tree with *page capacity* b stores all its subtrees of size $\le b$ (possibly empty) in pages; typically, the pages reside in secondary memory and the elements within a page are not organized as search trees (see Figure 1: the pagination of the search tree at the left is indicated using dashed lines; a more "realistic" representation of the same tree appears at its right).

Let $\mathcal{P}_n^{(b)}$ be the number of pages in a random search tree of size n with page capacity b. It is obvious that $\mathcal{P}_n^{(b)} = \mathcal{I}_n^{(b)} + 1$, where $\mathcal{I}_n^{(b)}$ is the number of internal nodes that are the root of a subtree that contains more than b items. In other words, in a paged search tree, we have external nodes (*pages*) that may contain up to b keys; if $\mathcal{P}_n^{(b)}$ is the number of external nodes or pages in a paged search tree, then $\mathcal{I}_n^{(b)} = \mathcal{P}_n^{(b)} - 1$ is the number of internal nodes in the tree, and these internal nodes are in one-to-one correspondance with the internal nodes with $> b$ descendants in the non-paged search tree.

Proposition 1.4 *For all n, and for any constant $b \ge 1$,*

$$\mathbb{E}\left[\mathcal{P}_n^{(b)}\right] = n\,\mathbb{P}[D_n > b] + 1.$$

Proof. Let δ_j be the indicator random variable for the predicate "the j^{th} element has more than b descendants". Then $\mathcal{I}_n^{(b)} = \sum_{1 \le j \le n} \delta_j$. The proposition follows taking expectations in both sides of this equation, because of the linearity of expectations and $\mathbb{E}[\delta_j] = \mathbb{P}[D_{n,j} > b]$. ☐

Results about the probabilistic behavior of the number of descendants are also useful in the analysis of the performance of quicksort if recursive calls are not made on small subfiles (say, of size $\le b$).

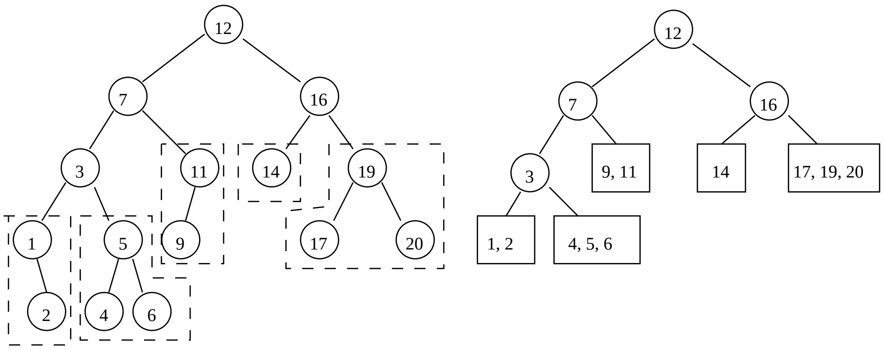

Figure 1: A paged binary search tree with page capacity $b = 3$

Proposition 1.5 *Let $C_n^{(b)}$ and $R_n^{(b)}$ be the number of comparisons*[3] *and the number of partitions made by quicksort to sort n elements, when the recursion halts on subfiles of size $\leq b$. Notice that standard quicksort corresponds to the case where $b = 1$. Then*

$$\mathbb{E}\left[R_n^{(b)}\right] = n\,\mathbb{P}\left[D_n > b\right],$$
$$\mathbb{E}\left[C_n^{(b)}\right] = n\left(\mathbb{E}\left[D_n\right] - 1\right) - n \sum_{1\leq m\leq b} (m-1)\,\mathbb{P}\left[D_n = m\right].$$

The strategy for the selection of pivots is related with the type of random search trees that we consider: for BSTs, we have selection of pivots at random; for LBSTs, we have that the pivots are the medians of random samples of three elements.

Proof. It is well known that we can associate to each particular execution of quicksort a binary search tree: the root contains the pivot element of the first stage, and the left and right subtrees are recursively built for the elements smaller and larger than the pivot, respectively. Each internal node in the search tree corresponds to a recursive call to quicksort. We will make a partitioning of a given subfile if and only if the subfile contains $> b$ elements, i.e. the corresponding internal node has $> b$ descendants, and the claim in the proposition follows.
On the other hand, let ϵ_j be the number of comparisons made between the j^{th} element and other elements, during the partition where the j^{th} element was selected as a pivot. Clearly, if $D_{n,j} \leq b$ then $\epsilon_j = 0$, since no recursive call will be made that chooses the j^{th} element as a pivot. On the other hand, if $D_{n,j} > b$, the j^{th} element will be compared with each of its descendants (except itself) in the associated search tree. Hence, $\mathbb{E}\left[\epsilon_j\right] = \sum_{m=b+1}^{n} (m-1)\,\mathbb{P}\left[D_{n,j} = m\right]$. We need only to sum over j to get the desired result. □

The structure of the paper is as follows. We start with an overview of some basic facts about generating functions and, in particular, about probability generating functions (Section 2).

In Section 3 we develop the main steps of our approach, taking the analysis of the number of descendants in random BSTs as a first introductory example. We provide here alternative derivations to the results of Lent [20], finding the probability that the j^{th} node in a random BST of size n has m descendants (Theorem 3.1). We also find exact and asymptotic values for all ordinary moments, including the expected value and variance (Theorem 3.2). Then we analyze the number of descendants

[3] We only count those made during the partitioning phases.

	BST		LBST	
	Of a given node	Of a random node	Of a given node	Of a random node
Ascendants	Average [3], variance*	Probability, moments, limit distribution [22, 5, 21, 17]	Average [16], variance*	Average, variance [26]* Higher order moments, PGF, limit distribution*
Descendants	Probability, moments [20]*	Probability, moments [20]*	PGF, average, variance*	Probability, moments*

Table 1: Summary of previous works and the results of this paper.

of a random node, obtaining the probability that $D_n = m$, as well as the moments of D_n (Theorems 3.3 and 3.2).

The remaining sections are devoted to the analysis of the number of ascendants and descendants in random LBSTs. In Section 5 we formally define LBSTs and give an equivalent characterization of the model of randomness which is more suitable to our purposes.

Among our new results, in Section 6 we derive an explicit form for the generating function of the probability distribution of $D_{n,j}$ (Theorem 6.1) and closed formulæ for the average (Theorem 6.2) and the second factorial moment (Theorem 6.3). Moreover, we find the probability distribution of D_n (Theorem 6.4) and all its moments (Theorem 6.5).

In Section 7, we compute $\mathbb{E}[A_{n,j}]$, the average number of ascendants of the j^{th} node in a random LBST of size n (Theorem 7.1). We are also able to compute the PGF of A_n, the number of ascendants of a random node (Theorem 7.4), as well as all its moments (Theorems 7.5 and 7.6), thus extending the results of Poblete and Munro [26].

The results of previous works and the new results in this paper are summarized in Table 1. Entries corresponding to new results in this paper and to alternative derivations of previous results are marked by '*'.

2. Mathematical Preliminaries

We start recalling the definition of *generating function*, for the reader's convenience. Given a sequence $\{a_n\}_{n\geq 0}$ its generating function $A(z)$ is the formal power series

$$A(z) = \sum_{n\geq 0} a_n z^n.$$

As usual, $[z^n]A(z)$ denotes the coefficient of z^n in $A(z)$ (the n^{th} coefficient of $A(z)$). Excellent sources of information about generating functions and their applications to combinatorics and the analysis of algorithms are [34, 32, 31, 19].

We make extensive use in this paper of probability generating functions (PGFs) as well as multivariate generating functions whose coefficients are PGFs themselves. We define them in turn. Given a discrete random variable X, its probability generating function $X(z)$ is

$$X(z) = \sum_{m} \mathbb{P}[X = m]\, z^m.$$

If we assume further that $X \geq 0$ and let $p_m = \mathbb{P}[X = m]$, the PGF of the random variable X is nothing but the ordinary generating function of the sequence $\{p_m\}_{m\geq 0}$. We list now a few important, although elementary, properties of PGFs.

Proposition 2.1 *For any discrete random variable X, its probability generating function $X(z)$ satisfies:*

1. $X(1) = 1$.
2. $X'(1) = \left.\frac{dX}{dz}\right|_{z=1} = \mathbb{E}[X]$.
3. $X^{(s)}(1) = \left.\frac{d^s X}{dz^s}\right|_{z=1} = \mathbb{E}[X^{\underline{s}}]$, *where $X^{\underline{s}}$ denotes the s^{th} falling factorial of X, that is, $X^{\underline{s}} = X(X-1)\dots(X-s+1)$. The quantity $\mathbb{E}[X^{\underline{s}}]$ is customarily called the s^{th} factorial moment of the random variable X. Ordinary and central moments may be recovered from factorial moments quite easily. For instance, if $\mu = \mathbb{E}[X]$, the variance of X is given by*
$$\mathbb{V}[X] = \mathbb{E}\left[(X-\mu)^2\right] = \mathbb{E}\left[X^{\underline{2}}\right] + \mathbb{E}[X] - \mathbb{E}[X]^2.$$

Since we will mostly deal with families of random variables, with two (n and j) or one (n) index, we will systematically work with multivariate generating functions of these families. For instance, if we were interested in the family $\{X_{n,j}\}_{1\le j\le n}$, we would introduce a generating function $\mathsf{X}(z,u,v)$ in three variables, such that the coefficient of $z^n u^j v^m$ in $\mathsf{X}(z,u,v)$ is the probability that $X_{n,j}$ is m. Thus

$$\mathsf{X}(z,u,v) = \sum_{n,j,m} \mathbb{P}[X_{n,j} = m]\, z^n u^j v^m, \tag{1}$$

where the indices of summation n, j and m run in the appropriate ranges (or we assume that $\mathbb{P}[X_{n,j} = m]$ is 0 whenever $n < 1$, $j < 1$, $j > n$ or $m < 0$). Notice that, by definition, $[z^n u^j]\mathsf{X}(z,u,v)$ is the PGF of the random variable $X_{n,j}$, and $[z^n u^j v^m]\mathsf{X}(z,u,v) = \mathbb{P}[X_{n,j} = m]$.

For technical reasons that will be clearer later, we will also use sometimes the derivative w.r.t. z of such a multivariate generating function. We will introduce then

$$\begin{aligned}\mathsf{X}_z(z,u,v) &= \frac{\partial}{\partial z}\sum_{n,j,m} \mathbb{P}[X_{n,j} = m]\, z^n u^j v^m \\ &= \sum_{n,j,m} n\,\mathbb{P}[X_{n,j} = m]\, z^{n-1} u^j v^m\end{aligned}$$

rather than the more natural definition given in Equation (1). This means that once we were able to extract coefficients from such a generating function, let us say the coefficient of $z^{n-1}u^j v^m$, we must divide by n to obtain $\mathbb{P}[X_{n,j} = m]$.

Furthermore, we are also interested in investigating all the moments of the random variables: mean, variance, and higher order moments. We differentiate s times the generating function $\mathsf{X}(z,u,v)$ with respect to v and let $v = 1$, to get the generating function for the s^{th} factorial moments, i.e.

$$\mathcal{X}^{(s)}(z,u) = \left.\frac{\partial^s \mathsf{X}(z,u,v)}{\partial v^s}\right|_{v=1}, \qquad s \ge 1. \tag{2}$$

Recall that $[z^n u^j]\mathcal{X}^{(s)}(z,u) = \mathbb{E}\left[X_{n,j}^{\underline{s}}\right]$.

Grand averages correspond to the situation where the *rank* —the parameter j in $X_{n,j}$— is random itself. More precisely, let $X_n \equiv X_{n,Z_n}$, where Z_n is a uniformly distributed random variable in $\{1,\dots,n\}$. Then X_n is the grand average of the random variables $X_{n,1},\dots,X_{n,n}$. It follows that

$$\mathbb{P}[X_n = m] = \frac{1}{n}\sum_{1\le j\le n} \mathbb{P}[X_{n,j} = m]. \tag{3}$$

We remark that $X_n \neq \frac{1}{n}(X_{n,1} + \cdots + X_{n,n})$, even if the $X_{n,j}$'s are independent.

Unless we are dealing with a differentiated version of the generating function $\mathsf{X}(z,u,v)$, we have

$$\overline{\mathsf{X}}(z,v) = \mathsf{X}(z,1,v) = \sum_{n,m} z^n v^m \sum_{1\le j\le n} \mathbb{P}\left[X_{n,j} = m\right]. \tag{4}$$

Thus the coefficient $[z^n v^m]\overline{\mathsf{X}}(z,v)$, divided by n, is the probability that X_n is m. In the case that $\mathsf{X}(z,u,v)$ were a differentiated generating function, then we should divide the coefficient $[z^{n-1}v^m]\overline{\mathsf{X}}(z,v)$ by n^2. Finally, computing the derivatives of $\mathsf{X}(z,v)$ w.r.t. v and setting $v=1$ yields the generating functions for the factorial moments of the grand average X_n.

The main steps of the systematic procedure that we will follow are thus:

1. Set up a recurrence for $\mathbb{P}\left[X_{n,j} = m\right]$;
2. Translate the recurrence to a functional equation over the corresponding generating function $\mathsf{X}(z,u,v)$;
3. Solve the functional equation;
4. Extract the coefficients of $\mathsf{X}(z,u,v)$;
5. Repeatedly differentiate $\mathsf{X}(z,u,v)$ w.r.t. v and set $v=1$; extract the coefficients to get the factorial moments of $X_{n,j}$;
6. Set $\overline{\mathsf{X}}(z,v) = \mathsf{X}(z,1,v)$ and repeat steps 4 and 5 for $\overline{\mathsf{X}}(z,v)$.

In practice, the procedure might fail for several reasons. Typically, because we are not able to solve the equation at step 3 or to extract the coefficients of a given generating function. Although we have (almost) not used them in this paper, the reader should be aware of the existing powerful techniques to extract asymptotic information about the coefficients of a generating function if we know its behaviour near its singularities or in some case, even if we only know the functional equation satisfied by the generating function [32, 6]. Also, if we are not able to solve and get an explicit form for $\mathsf{X}(z,u,v)$, we can still differentiate w.r.t. to v or set $u=1$ and try to solve the (easier) resulting differential equations, to get information about the moments or the grand average.

The functional equations that arise in our study are linear partial differential equations of the first (BSTs) and of the second (LBSTs) order. The former can be solved, in principle, by quadrature through the variation of constant —actually, functions in u and v— method. For the second order differential equations, the theory of hypergeometric differential equations comes into play [15].

Nowadays, most of the necessary mathematical knowledge is embodied into modern computer algebra systems. In our case, MAPLE needed little or no assistance to solve the differential equations that we had.

The last step, that of extracting coefficients in exact form, was, at large, the least systematic and mechanical one. A great deal of combinatorial identities, inspired guessing and patience was needed. Standard MAPLE tools like the function interp or the GFUN package [28] proved also to be useful. However, once the solution is obtained, it is just a matter of minutes to check its correctness. It is quite difficult to provide a detailed and ordered description of the methods that we used to extract coefficients from generating functions. As a result, the paper contains only some hints here and there, while some claims are just stated without further explanation.

3. The number of descendants in random BSTs

The number of the descendants $D_{n,j}$ of the j^{th} node of a BST of size n is recursively computed as the number of descendants in the left subtree of the j^{th} node, plus the number of descendants in its right subtree, plus one (to count the j^{th} node itself). The probability that $D_{n,j} = m$ is computed conditioning on the events "the rank of the root is k", that means the root is the k^{th} node of a search tree. Recall that, for a random BST of size n, the rank of the root is k with probability $1/n$, for $k = 1, \dots, n$. Using the recursive definition of $D_{n,j}$ we have

$$\begin{aligned}\mathbb{P}[D_{n,j} = m] &= \sum_{k=1}^{n} \mathbb{P}\left[D_{n,j} = m \,|\, \text{the root is the } k^{\text{th}} \text{ element}\right] \\ &\qquad \times \mathbb{P}\left[\text{the root is the } k^{\text{th}} \text{ element}\right] \\ &= \frac{1}{n}[\![m = n]\!] + \frac{1}{n}\sum_{k=1}^{j-1} \mathbb{P}[D_{n-k,j-k} = m] + \frac{1}{n}\sum_{k=j+1}^{n} \mathbb{P}[D_{k-1,j} = m], \end{aligned} \tag{5}$$

where $[\![P]\!]$ is 1 if P is true and 0 otherwise [9].

This recursion translates nicely into a functional equation over the generating function for the family of random variables $\{D_{n,j}\}$. Solving the functional equation and extracting coefficients of the generating function, we get the following theorem, which was already found by Lent [20] using probabilistic techniques.

Theorem 3.1 *The probability that the j^{th} internal node of a random binary search tree of size n has m descendants is, assuming that $j \le n + 1 - j$,*

$$\mathbb{P}[D_{n,j} = m] = \begin{cases} \dfrac{2}{(m+1)(m+2)} & \text{for } 1 \le m < j, \\ \dfrac{1}{(m+1)(m+2)}\left(1 + \frac{2j}{m}\right) & \text{for } j \le m < n+1-j, \\ \dfrac{2(n+1)}{m(m+1)(m+2)} & \text{for } n+1-j \le m < n, \\ \dfrac{1}{n} & \text{for } m = n. \end{cases}$$

For the cases where $j > n + 1 - j$ we can use the symmetry on j and $n + 1 - j$ (Proposition 1.1) to compute the corresponding probabilities.

Also, the distribution function for $D_{n,j}$ is

$$\mathbb{P}[D_{n,j} \le m] = \begin{cases} \dfrac{m}{m+2} & \text{for } 1 \le m < j, \\ \dfrac{m+1}{m+2} - \dfrac{j}{(m+1)(m+2)} & \text{for } j \le m < n+1-j, \\ \dfrac{m^2 + 3m + 1 - n}{(m+1)(m+2)} & \text{for } n+1-j \le m < n, \\ 1 & \text{for } m = n. \end{cases}$$

Proof. We start defining the generating function

$$\mathsf{D}(z, u, v) = \sum_{1 \le j, m \le n} \mathbb{P}[D_{n,j} = m]\, z^n u^j v^m.$$

Multiplying both sides of (5) by $nz^{n-1}u^j v^m$ and summing for all $n \geq 1$, $1 \leq j \leq n$ and $m \geq 1$, yields

$$\frac{\partial \mathsf{D}}{\partial z} = \frac{u\mathsf{D}}{1-uz} + \frac{\mathsf{D}}{1-z} + \frac{uv}{(1-vz)(1-uvz)},$$
$$\mathsf{D}(0,u,v) = 0. \tag{6}$$

The solution to the differential equation above is relatively simple

$$\begin{aligned}\mathsf{D}(z,u,v) = {} & \frac{uz}{v(1-z)(1-uz)} \\ & - \frac{u(1-v)(v-u)}{(1-z)(1-uz)v^2(1-u)} \log\frac{1}{1-vz} \\ & - \frac{(1-v)(1-uv)}{(1-z)(1-uz)v^2(1-u)} \log\frac{1}{1-uvz}.\end{aligned} \tag{7}$$

The statement of the theorem follows after extracting the coefficient $[z^n u^j v^m]\mathsf{D}(z,u,v)$. □

The explicit and simple form of the trivariate generating function in Theorem 3.1 allows us to compute all the moments *explicitly*. It is convenient to deal with a sort of shifted factorial moments; the ordinary moments can be computed by linear combinations of the shifted factorial ones.

Theorem 3.2 *Let $d_{n,j}^{(s)} = \mathbb{E}\left[(D_{n,j}+2)^{\underline{s}}\right]$ and $d_{n,j} = d_{n,j}^{(1)}$, where $D_{n,j}$ denotes the number of descendants of the j^{th} internal node in a random binary search tree of size n. For all $n > 0$ and all $1 \leq j \leq n$,*

1. $d_{n,j} = H_j + H_{n+1-j} + 1$.
2. $d_{n,j}^{(2)} = 2(n+1)H_n - 2jH_j - 2(n+1-j)H_{n+1-j} + 2(n+2)$.
3. *For all $s \geq 3$,*

$$d_{n,j}^{(s)} = \frac{s}{s-2}(n+1)^{\underline{s-1}} - \frac{s}{(s-1)(s-2)}\left[j^{\underline{s-1}} + (n+1-j)^{\underline{s-1}}\right].$$

Proof. We begin by introducing

$$\mathcal{D}^{(s)}(z,u) = \left.\frac{\partial^s(v^2\mathsf{D}(z,u,v))}{\partial v^s}\right|_{v=1},$$

and hence its coefficients are

$$d_{n,j}^{(s)} = [z^n u^j]\mathcal{D}^{(s)}(z,u) = \mathbb{E}\left[(D_{n,j}+2)^{\underline{s}}\right].$$

The shifted moments are particularly easy to obtain, since the coefficients of $\mathcal{D}^{(s)}(z,u)$ that we seek are linear combinations of the coefficients of the next generating functions:

$$\begin{aligned}\left.\frac{\partial^s}{\partial v^s}\log\frac{1}{1-vz}\right|_{v=1} &= (s-1)!\left(\frac{z}{1-z}\right)^s, \\ \left.\frac{\partial^s}{\partial v^s}v\log\frac{1}{1-vz}\right|_{v=1} &= (s-1)!\left(\frac{z}{1-z}\right)^s + s(s-2)!\left(\frac{z}{1-z}\right)^{s-1}, \\ \left.\frac{\partial^s}{\partial v^s}v^2\log\frac{1}{1-vz}\right|_{v=1} &= (s-1)!\left(\frac{z}{1-z}\right)^s + 2s(s-2)!\left(\frac{z}{1-z}\right)^{s-1} + s(s-1)(s-3)!\left(\frac{z}{1-z}\right)^{s-2},\end{aligned}$$

$$\frac{\partial^s}{\partial v^s}\log\frac{1}{1-uvz}\Big|_{v=1} = (s-1)!\left(\frac{uz}{1-uz}\right)^s,$$
$$\frac{\partial^s}{\partial v^s}v\log\frac{1}{1-uvz}\Big|_{v=1} = (s-1)!\left(\frac{uz}{1-uz}\right)^s + s(s-2)!\left(\frac{uz}{1-z}\right)^{s-1},$$
$$\frac{\partial^s}{\partial v^s}v^2\log\frac{1}{1-uvz}\Big|_{v=1} = (s-1)!\left(\frac{uz}{1-uz}\right)^s + 2s(s-2)!\left(\frac{uz}{1-uz}\right)^{s-1} + s(s-1)(s-3)!\left(\frac{uz}{1-uz}\right)^{s-2}.$$

We might additionally observe that for all $n \geq 0$ and $1 \leq j \leq n$

$$[z^n u^j]\frac{1}{(1-z)^{s+1}(1-uz)(1-u)} = \binom{s+n+1}{s+1} - \binom{s+n-j}{s+1},$$
$$[z^n u^j]\frac{1}{(1-z)(1-uz)^{s+1}(1-u)} = \binom{s+j+1}{s+1}, \quad \text{and}$$
$$[z^n u^j]\frac{1}{(1-z)^2(1-uz)^2} = (j+1)(n+1-j).$$

Theorem 3.2 is an immediate consequence of the formulæ above. □

Corollary 3.1 *The expected value and variance of $D_{n,j}$ are, respectively,*

$$\mathbb{E}[D_{n,j}] = H_j + H_{n+1-j} - 1,$$
$$\mathbb{V}[D_{n,j}] = 2(n+1)H_n - (2j+1)H_j - (2n-2j+3)H_{n+1-j} + 2(n+2) - H_j^2 - H_{n+1-j}^2 - 2H_jH_{n+1-j}.$$

Furthermore, for $j = \alpha n$, with $0 < \alpha < 1$, we have

$$\mathbb{E}[D_{n,\alpha n}] = 2\log n + \log\alpha + \log(1-\alpha) + 2\gamma - 1 + o(1),$$
$$\mathbb{V}[D_{n,\alpha n}] = 2n\Big(1 - \alpha\log\alpha - (1-\alpha)\log(1-\alpha)\Big) + \mathcal{O}(\log^2 n),$$

where $\gamma = 0.5772156649\ldots$ is Euler's constant.

To recover higher order ordinary moments, we only need to express the ordinary powers as linear combinations of the shifted falling factorials with coefficients $\lambda_{s,k}$. Thus

$$x^s = \sum_{k=0}^{s}\lambda_{s,k}(x+2)^{\underline{k}}.$$

It is easy to show that

$$\lambda_{s,k} = \sum_{i=k}^{s}\left\{ i \atop k \right\}\binom{s}{i}(-2)^{s-i},$$

where $\left\{ i \atop k \right\}$ denote Stirling numbers of the second kind. The coefficients $\lambda_{s,k}$ satisfy a recursion that is similar to that of the Stirling numbers

$$\lambda_{s+1,k} = \lambda_{s,k-1} + (k-2)\lambda_{s,k},$$

and $\lambda_{s,0} = (-2)^s$.

Let us consider now D_n, the number of descendants of a random node in a random BST of size n. The following two theorems give closed formulæ for the probability that D_n is m and for the shifted factorial moments of D_n, i.e. for $d_n^{(s)} = \mathbb{E}[(D_n+2)^{\underline{s}}]$.

Theorem 3.3 *The probability that a randomly chosen internal node in a random binary search tree of size n has m descendants is given by*

$$\mathbb{P}[D_n = m] = \begin{cases} \dfrac{2(n+1)}{n(m+1)(m+2)} & \text{for } 1 \le m < n, \\ \dfrac{1}{n} & \text{for } m = n. \end{cases}$$

Proof. Plug $u = 1$ into the solution of (7) to get

$$\overline{\mathsf{D}}(z,v) = \mathsf{D}(z,1,v) = -\frac{2(1-v)}{v^2(1-z)^2}\log\frac{1}{1-vz} - \frac{z(zv - v^2 + 2v - 2)}{v(1-vz)(1-z)^2}. \tag{8}$$

The coefficient of $[z^n v^m]\overline{\mathsf{D}}(z,v)$, divided by n, is the sought probability. □

Theorem 3.4 *The s^{th} shifted factorial moment $d_n^{(s)} = \mathbb{E}[(D_n+2)^{\underline{s}}]$ of the number of descendants of a random node in a random binary search tree of size n is given by*

1. $d_n = d_n^{(1)} = 2(1 + \frac{1}{n})H_n - 1$.
2. $d_n^{(2)} = 3(n+1)$.
3. *For all $s \geq 3$,*
$$d_n^{(s)} = \frac{1}{n}\left((n+2)^{\underline{s}} + \frac{2}{s-1}(n+1)^{\underline{s}}\right) \sim \frac{s+1}{s-1}n^{s-1}.$$

Proof. Repeated differentiation of the generating function $v^2\overline{\mathsf{D}}(z,v)$ w.r.t. v and setting $v = 1$, gives us the generating functions of the shifted factorial moments. Their coefficients are extracted much in the same way as in Theorem 3.2. □

A few comments concerning the last theorem are in order now. Observe that for $s \geq 3$

$$\frac{1}{n}\sum_{j=1}^{n} d_{n,j}^{(s)} = \frac{(n+1)^{\underline{s-1}}}{s-1}\left(s + 1 + \frac{2}{n}\right).$$

Asymptotically, this quantity is

$$\sim \frac{s+1}{s-1}n^{s-1},$$

one of the observations in the work of Lent [20]. The coincidence in asymptotic behavior with $d_n^{(s)}$ is remarkable; recall that in general

$$\mathbb{E}[D_n^s] \neq \mathbb{E}\left[\left(\frac{1}{n}\sum_{1\le j\le n} D_{n,j}\right)^s\right],$$

except when $s = 1$ and the same observation holds for the shifted factorial moments we were dealing with.

Last, but not least, we can obtain the following corollaries, from Propositions 1.4 and 1.5 and the theorems in this section. These results can already be found in [19], although there is a slight difference in $\mathbb{E}\left[C_n^{(b)}\right]$, because $n+1$ comparisons per partition are counted there, while we count $n-1$ comparison per partition.

Corollary 3.2 *The expected number of pages in a random binary search tree of size n with page capacity b is*

$$\mathbb{E}\left[\mathcal{P}_n^{(b)}\right] = 2\frac{n+1}{b+2}.$$

The filling ratio *for binary search trees is thus*

$$\gamma_b = \frac{n/b}{\mathbb{E}\left[\mathcal{P}_n^{(b)}\right]} \sim \frac{1}{2}.$$

Corollary 3.3 *The expected number of recursive calls to sort a random permutation of size n, when the recursion stops in subfiles of size $\leq b$ is*

$$\mathbb{E}\left[R_n^{(b)}\right] = \frac{2n-b}{b+2}.$$

Also, the expected number of comparisons to sort a random permutation of size n, when the recursion stops in subfiles of size $\leq b$ is

$$\mathbb{E}\left[C_n^{(b)}\right] = 2(n+1)\left(H_n - H_{b+1}\right) + n + 5 - \frac{6(n+1)}{b+2}.$$

4. The number of ascendants in random BSTs

Considering the element k of the root of a BST, we obtain for the number of ascendants $A_{n,j}$ of the j^{th} node of a BST of size n the following recursion:

$$\mathbb{P}\left[A_{n,j} = m\right] = \frac{1}{n}[\![m=1]\!] + \frac{1}{n}\sum_{k=1}^{j-1}\mathbb{P}\left[A_{n-k,j-k} = m-1\right] + \frac{1}{n}\sum_{k=j+1}^{n}\mathbb{P}\left[A_{k-1,j} = m-1\right]. \tag{9}$$

Introducing the generating function for the family of random variables $\{A_{n,j}\}$

$$\mathsf{A}(z,u,v) = \sum_{1\leq j,m\leq n}\mathbb{P}\left[A_{n,j} = m\right] z^n u^j v^m,$$

this recursion translates by multiplying both sides by $nz^{n-1}u^jv^m$ and summing for all $n \geq 1$, $1 \leq j \leq n$ and $m \geq 1$ into the following differential equation:

$$\frac{\partial \mathsf{A}}{\partial z} = \frac{v}{1-z}\mathsf{A} + \frac{uv}{1-uz}\mathsf{A} + \frac{uv}{(1-z)(1-uz)}$$

with the initial condition $\mathsf{A}(0,u,v) = 0$. This differential equation has the following solution

$$\mathsf{A}(z,u,v) = \frac{uv}{(1-z)^v(1-uz)^v}\int_0^z (1-t)^{v-1}(1-ut)^{v-1}dt. \tag{10}$$

Starting with this generating function, it is easy to get the following theorems. At first we obtain an old result from [3]:

Theorem 4.1 *The expected number of ascendants $a_{n,j} = \mathbb{E}\left[A_{n,j}\right]$ of the j^{th} node in a random binary search tree of size n is*

$$a_{n,j} = H_j + H_{n+1-j} - 1.$$

Proof. Starting with (10), taking derivatives w.r.t. v and setting $v = 1$, we get the generating function $\mathcal{A}(z, u)$, whose coefficients are the expected values $a_{n,j} = \mathbb{E}[A_{n,j}]$. It is given by

$$\mathcal{A}(z,u) = \frac{u}{(1-z)(1-uz)} \log\frac{1}{1-z} + \frac{1}{(1-z)(1-uz)} \log\frac{1}{1-uz} - \frac{uz}{(1-z)(1-uz)}.$$

It is easy to extract the coefficients of this expression, which leads immediately to the stated theorem. □

Theorem 4.2 *The second factorial moment* $a^{(2)}_{n,j} = \mathbb{E}\left[(A_{n,j})^{\underline{2}}\right]$ *of the number of ascendants of the* j^{th} *node in a random binary search tree of size* n *is*

$$\begin{aligned} a^{(2)}_{n,j} &= \frac{2\,(n+1)}{(n+1-j)\,j} H_n + H_j^2 + 2 H_j H_{n+1-j} + \frac{2\left(-nj-n+j^2-j-1\right)}{(n+1-j)\,j} H_j + H^2_{n+1-j} \\ &\quad + \frac{2\left(-nj-n+j^2-j-1\right)}{(n+1-j)\,j} H_{n+1-j} - H^{(2)}_j - H^{(2)}_{n+1-j} - \frac{2(-2\,nj+2\,j^2-2\,j-1)}{(n+1-j)\,j}. \end{aligned} \tag{11}$$

Proof. Differentiating equation (10) two times w.r.t. v and setting $v = 1$ gives the generating function $\mathcal{A}^{(2)}(z, u)$ of the second factorial moments $a^{(2)}_{n,j}$ of the number of ascendants:

$$\begin{aligned} \mathcal{A}^{(2)}(z,u) &= -\frac{2zu}{(1-uz)\,(1-z)} \log\frac{1}{1-uz} - \frac{2zu}{(1-uz)\,(1-z)} \log\frac{1}{1-z} \\ &\quad - \frac{2\,(uz-u-1)}{(1-uz)\,(1-z)} \log\frac{1}{1-z} \log\frac{1}{1-uz} + \frac{u}{(1-uz)\,(1-z)} \log^2\frac{1}{1-z} \\ &\quad + \frac{1}{(1-uz)\,(1-z)} \log^2\frac{1}{1-uz} + \frac{2u}{(1-uz)\,(1-z)} \int_0^z \log\frac{1}{1-t} \log\frac{1}{1-ut} dt. \end{aligned}$$

Extracting the coefficients leads to the given theorem. Since one expression in $\mathcal{A}^{(2)}(z, u)$ turns out to be a bit messier, we sketch how to extract the coefficients of it. First we get the following sum

$$\begin{aligned} [z^n u^j] \frac{1}{(1-z)(1-uz)} \int_0^z \log\frac{1}{1-t} \log\frac{1}{1-ut} dt &= \sum_{k=0}^{j} \sum_{l=0}^{n-j+k} [z^l u^k] \int_0^z \log\frac{1}{1-t} \log\frac{1}{1-ut} dt \\ &= \sum_{k=1}^{j} \sum_{l=k+2}^{n-j+k} \frac{1}{lk(l-k-1)}, \end{aligned}$$

which can be simplified to

$$\begin{aligned} \sum_{k=1}^{j} \sum_{l=k+2}^{n-j+k} \frac{1}{lk(l-k-1)} &= \sum_{k=1}^{j} \frac{1}{k} \sum_{l=1}^{n-j-1} \frac{1}{l(l+k+1)} = \sum_{k=1}^{j} \frac{1}{k(k+1)} \sum_{l=1}^{n-j-1} \left(\frac{1}{l} - \frac{1}{l+k+1}\right) \\ &= \sum_{k=1}^{j} \frac{1}{k(k+1)} \left(H_{n-j-1} + H_{k+1} - H_{n-j+k}\right) = \sum_{k=1}^{j} \left(\frac{1}{k} - \frac{1}{k+1}\right) \left(H_{n-j-1} + H_{k+1} - H_{n-j+k}\right) \\ &= H_{n-j-1} \sum_{k=1}^{j} \left(\frac{1}{k} - \frac{1}{k+1}\right) + \sum_{k=1}^{j} \left(\frac{H_k}{k} - \frac{H_{k+1}}{k+1}\right) + \sum_{k=1}^{j} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &\quad - \sum_{k=1}^{j} \left(\frac{H_{n-j+k}}{k} - \frac{H_{n-j+k+1}}{k+1}\right) - \frac{1}{n-j} \sum_{k=1}^{j} \left(\frac{1}{k+1} - \frac{1}{n-j+k+1}\right). \end{aligned}$$

The sums telescope and we finally get

$$[z^n u^j]\frac{1}{(1-z)(1-uz)}\int_0^z \log\frac{1}{1-t}\log\frac{1}{1-ut}dt = \frac{n+1}{(j+1)(n-j)}\left(H_{n+1}-H_{j+1}-H_{n+1-j}\right) + \frac{2jn^2-4nj^2+2j^3+n^2-jn+2n-2j+1}{(n-j)(j+1)(n+1-j)}.$$

□

The next theorem gives the variance, which is now easy to obtain.

Theorem 4.3 *The variance $\mathbb{V}[A_{n,j}]$ of the number of ascendants of the j^{th} node in a random search tree of size n is*

$$\begin{aligned}\mathbb{V}[A_{n,j}] = {} & \frac{2\,(n+1)}{(n+1-j)\,j}H_n + \frac{(nj-2\,n-j^2+j-2)}{(n+1-j)\,j}H_j + \frac{(nj-2\,n-j^2+j-2)}{(n+1-j)\,j}H_{n+1-j} \\ & - H_j^{(2)} - H_{n+1-j}^{(2)} + \frac{2(nj-j^2+j+1)}{(n+1-j)\,j}.\end{aligned} \tag{12}$$

□

5. Locally balanced binary search trees

One approach to avoid drastically unbalanced binary search trees is the introduction of strict balance constraints like in AVLs or red-black trees [1, 10]. Such schemes guarantee logarithmic performance of searches and updates in the worst-case, but they have additional space requirements and are more difficult to implement than standard BSTs. As an alternative, several authors [4, 33, 26] have suggested the use of a simple heuristic that makes the construction of poorly balanced trees much less likely than with the use of the standard algorithms. Furthermore, the heuristic was shown to yield significant savings in the expected search time.

The basic idea is really simple: whenever a son is appended to a node that itself is a single son (its "brother" is an external node), a rotation of the three nodes is performed to place the median of the three elements as the root of the subtree and the other two elements as sons (see Figure 2). Since no other kind of rebalancing operation is ever made, Poblete and Munro refer to this technique as a *fringe* heuristic. We will call the binary search trees constructed in this way *locally balanced binary search trees* (LBST, for short).

Poblete and Munro [26] and Poblete [25] carry on the analysis of this heuristic and some generalizations by means of bottom-up or fringe techniques: they basically study the number of nodes that are at level k and which are the root of a subtree of size 1 or 2.

As we have already mentioned in the introduction, the standard model for random LBSTs states that a random LBST of size n is the result of n random insertions into an initially empty tree. Equivalently, a random LBST of size n is the result of inserting the elements of a random permutation of $\{1,\ldots,n\}$ into an initially empty tree. Here, we show that a recursive, top-down definition of the randomness model is also possible. This characterization of the model of randomness is more amenable to the kind of algebraic manipulations that we want to carry on; as we will see, the recurrence relations for the analyzed quantities translate to equations over generating functions in a natural way, almost automatically.

Definition 5.1 *1. A random binary search tree of size $s \leq 2$ is also a random locally balanced search tree. Recall that a BST of size 2 is random if the smallest (resp. largest) key is the root with probability 1/2.*

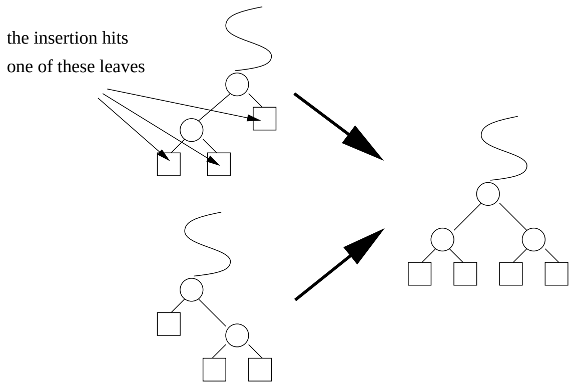

Figure 2: The fringe heuristic

2. *A binary search tree T of size $n \geq 3$, with left and right subtrees T_1 and T_2, is a random LBST if and only if, both T_1 and T_2 are random independent LBSTs, and*

$$\pi_{n,k} = \mathbb{P}\left[\, |T_1| = k-1 \,\middle|\, |T| = n \,\right] = \frac{(k-1)(n-k)}{\binom{n}{3}}, \qquad \textit{for all } 1 \leq k \leq n,$$

The reader should have noticed that the only difference between this definition and that for random BSTs relies on the *splitting probabilities* $\pi_{n,k}$. In the case of BSTs, each element of the random permutation has the same probability (namely, $1/n$) of being the first element and hence of becoming the root. In the case of LBSTs, when $n \geq 3$, the probability that the k^{th} element is one the first three elements of the permutation and is the median of these three elements is

$$\frac{1}{n} \times \frac{k-1}{n-1} \times \frac{n-k}{n-2} \times 3! = \pi_{n,k}.$$

Indeed, the left hand side of the equation above give us the probability that the k^{th} element is the first, times the probability that it is followed by a smaller element, times the probability that the two elements are followed by a larger element. For any permutation of such three elements, we have that the k^{th} element is among the first three elements and it is their median. Now, under these conditions the k^{th} element will be the root of the LBST (after the insertion of the first three elements, with rebalancing if necessary). The insertion of the fourth, fifth, etc. elements will not affect the root of the LBST. The principle applies recursively to the subsequences of elements smaller and greater than the selected element and the definition follows.

This argument also justifies the deep connection between LBSTs, quicksort and quickselect (see Propositions 1.3 and 1.5), when we consider the variants that select the median of 3 elements taken at random as the pivot of each partitioning phase.

6. The number of descendants in random LBSTs

As in Section 3, let $D_{n,j}$ denote the number of descendants of the j^{th} node, but now in a random LBST of size n. The recursion for $\mathbb{P}[D_{n,j} = m]$ is almost the same as for random BSTs, the only difference being the splitting probability $\pi_{n,k}$, the probability that the root of the LBST is the k^{th}

element. Thus,

$$\mathbb{P}[D_{n,j}=m]=\Bigg[\sum_{1\le k<j}\pi_{n,k}\,\mathbb{P}[D_{n-k,j-k}=m]+\pi_{n,j}[\![m=n]\!]$$
$$+\sum_{j<k\le n}\pi_{n,k}\,\mathbb{P}[D_{k-1,j}=m]\Bigg]. \tag{13}$$

Theorem 6.1 *Let*

$$\mathsf{D}_z(z,u,v)=\frac{\partial}{\partial z}\sum_{n,j,m}\mathbb{P}[D_{n,j}=m]\,z^n u^j v^m=\sum_{n,j,m}n\,\mathbb{P}[D_{n,j}=m]\,z^{n-1}u^j v^m.$$

Then,

$$\begin{aligned}\mathsf{D}_z(z,u,v)=&\frac{A_0(z,u,v)}{v(1-z)^2(1-uz)^2(1-uv)^2(1-v)^2(v-u)^2(1-u)^2}\\&+\frac{A_1(z,u,v)}{(1-uz)^2(1-u)(1-v)^3(1-uv)^3}\log\frac{1}{1-z}\\&+\frac{A_2(z,u,v)}{(1-z)^2(u-v)^3(1-u)(1-v)^3}\log\frac{1}{1-uz}\\&+\frac{A_3(z,u,v)}{v^2(1-z)^2(1-uz)^2(v-u)^3(1-u)^3(1-v)^3}\log\frac{1}{1-vz}\\&+\frac{A_4(z,u,v)}{v^2(1-z)^2(1-uz)^2(1-u)^3(1-v)^3(1-uv)^3}\log\frac{1}{1-uvz},\end{aligned}$$

where each of the $A_i(z,u,v)$'s is a complicated polynomial in z, u and v. They are listed in full in the appendix.

Proof. We multiply the recursion (13) by $\binom{n}{3}$ and $z^{n-3}u^jv^m$, sum up over all $n\ge 1$ and $1\le j,m\le n$ to get the following differential equation:

$$\frac{1}{6}\frac{\partial^2\mathsf{D}_z}{\partial z^2}=\frac{u^2v^3}{(1-vz)^2(1-uvz)^2}+\frac{u^2}{(1-uz)^2}\mathsf{D}_z+\frac{1}{(1-z)^2}\mathsf{D}_z, \tag{14}$$

where the initial conditions are $\mathsf{D}_z(0,u,v)=uv$ and $\frac{\partial}{\partial z}\mathsf{D}_z(0,u,v)=uv(1+u)(1+v)$. We use the partial derivative w.r.t. z to define $\mathsf{D}_z(z,u,v)$ because the differential equation just given, which translates the recurrence for $\mathbb{P}[D_{n,j}=m]$, is then of the second order. Had we introduced the generating function $\mathsf{D}_z(z,u,v)$ in the standard manner, we would have had a third order differential equation, with no appearance of the function itself, only the first and third derivatives.

The differential equation (14) is solvable: its explicit form (abridged) is the one given in the statement of the theorem. □

From the explicit form of $\mathsf{D}_z(z,u,v)$ given in Theorem 6.1 we can, in principle, compute exact expressions for $\mathbb{P}[D_{n,j}=m]$ and all moments. However, the task is daunting, and we will content ourselves computing the expected value and the second factorial moment in the next two theorems.

Theorem 6.2 *The expected number of descendants $d_{n,j} = \mathbb{E}[D_{n,j}]$ of the j^{th} node in a random LBST of size n is, when $5 \le j \le n-4$*

$$\begin{aligned} d_{n,j} = & -\frac{12}{7}H_n + \frac{12}{7}H_j + \frac{12}{7}H_{n+1-j} \\ & - \frac{6}{7j} - \frac{6}{7(n+1-j)} + \frac{79}{70} \\ & - \frac{3(3j-5)}{7n} + \frac{6(j-1)^2}{7n^{\underline{2}}} + \frac{2(2j-3)(j-1)^{\underline{2}}}{7n^{\underline{3}}} \\ & + \frac{3(j-2)(j-1)^{\underline{3}}}{7n^{\underline{4}}} - \frac{3(2j-5)(j-1)^{\underline{4}}}{7n^{\underline{5}}} + \frac{2(j-3)(j-1)^{\underline{5}}}{7n^{\underline{6}}}. \end{aligned}$$

The remaining cases when $j \le 4$ (or when $j > n-4$, by symmetry) appear in the appendix.

Proof. Taking the first derivative with respect to v, and setting $v = 1$ we get[4]

$$\begin{aligned} \left.\frac{\partial \mathsf{D}_z}{\partial v}\right|_{v=1} = & \frac{B_0(z,u)}{70(1-uz)^2(1-u)^7(1-z)^2} \\ & + \frac{B_1(z,u)}{7(1-u)^7(1-uz)^2}\log\frac{1}{1-z} \\ & + \frac{B_2(z,u)}{7(1-z)^2(1-u)^7}\log\frac{1}{1-uz}, \end{aligned}$$

where the $B_i(z,u)$'s are polynomials in z and u. Their explicit value can be found in the appendix at the end of this paper.
In order to get to the coefficients, we use formulæ such as

$$\begin{aligned} [z^n u^j]\frac{1}{(1-u)^4(1-uz)^2}\log\frac{1}{1-z} = & (n+1)\binom{j-n+3}{3}\Big(H_n - H_{n-1-j}\Big) \\ & + \frac{(j+1)n^3}{6} - \frac{5(j+1)(j+2)n^2}{12} \\ & + \frac{(j+1)(11j^2+40j+30)n}{36} - \frac{3j-10}{12}\binom{j+3}{3}, \end{aligned}$$

$$\begin{aligned} [z^n u^j]\frac{1}{(1-u)^5(1-uz)^2}\log\frac{1}{1-z} = & (n+1)\binom{j-n+4}{4}\Big(H_n - H_{n-1-j}\Big) \\ & - \frac{(j+1)n^4}{24} + \frac{(7j+18)(j+1)n^3}{48} \\ & - \frac{(j+1)(13j^2+65j+75)n^2}{72} \\ & + \frac{(j+1)(25j^3+173j^2+348j+180)n}{288} \\ & - \frac{12j-65}{60}\binom{j+4}{4}, \end{aligned}$$

[4] It turns out that Maple gets stuck doing the work in the obvious way, i.e. take the derivative, then take the limit when $v \to 1$. But we can produce the differential equations satisfied by the generating functions for the factorial moments from the differential equation (14) and solve them. Also, the problem can be fixed by computing a `series` expansion of the derivatives around $v = 1$.

and similar ones that are not too hard to obtain. To retrieve the final answer, we have also to take into account that we need to shift the coefficients in z^n by 1 and multiply by $\frac{1}{n}$, because we were considering $\frac{\partial}{\partial z} \sum_{n,j,m} \mathbb{P}\left[D_{n,j} = m\right] z^n u^j v^m$. Putting everything together, the theorem follows. $\square$

Theorem 6.3 *The second factorial moment of the number of descendants $d^{(2)}_{n,j} = \mathbb{E}\left[D^{\underline{2}}_{n,j}\right]$ of the j^{th} node in a random LBST of size n is, when $5 \le j \le n-4$,*

$$\begin{aligned} d^{(2)}_{n,j} = {} & \left(\frac{36n}{5} - \frac{12}{35}\right) H_n + \left(\frac{36j}{5} - \frac{36n}{5} - \frac{48}{7}\right) H_{n+1-j} + \left(\frac{12}{35} - \frac{36j}{5}\right) H_j \\ & - \frac{132}{35j} - \frac{132}{35(n+1-j)} + \frac{3489}{175} - \frac{33j}{5} + \left(\frac{66}{7} - \frac{429j}{35} + \frac{33j^2}{5}\right) \frac{1}{n} \\ & + \frac{132(j-1)^2}{35n^{\underline{2}}} + \frac{44(2j-3)(j-1)^{\underline{2}}}{35n^{\underline{3}}} + \frac{66(j-2)(j-1)^{\underline{3}}}{35n^{\underline{4}}} \\ & - \frac{66(2j-5)(j-1)^{\underline{4}}}{35n^{\underline{5}}} + \frac{44(j-3)(j-1)^{\underline{5}}}{35n^{\underline{6}}}. \end{aligned}$$

The formulæ for the second factorial moment in the special cases (when $j \le 4$ or $j > n-4$) are collected in a table in the appendix.

Proof. The second factorial moment $d^{(2)}_{n,j}$ is the coefficient of $z^{n-1}u^j$ times $\frac{1}{n}$ in

$$\begin{aligned} \left.\frac{\partial^2 \mathsf{D}_z}{\partial v^2}\right|_{v=1} = {} & \frac{C_0(z,u)}{35(1-z)^2(1-u)^7(1-uz)^2} \\ & + \frac{C_1(z,u)}{35(1-z)^2(1-u)^7(1-uz)^2} \log\frac{1}{1-z} \\ & + \frac{C_2(z,u)}{35(1-z)^2(1-u)^7(1-uz)^2} \log\frac{1}{1-uz} \end{aligned}$$

where the $C_i(z,u)$'s are polynomials in z and u. They have been listed in the appendix. Using techniques similar to the ones in the proof of Theorem 6.2, we extract the coefficients and obtain the stated result. $\square$

As in Section 3, we shift now our attention to the number of descendants of a random node in a random LBST of size n. We start giving an explicit expression for the probability distribution of D_n.

Theorem 6.4 *The probability that a random node in a random LBST of size n has m descendants is*

$$\mathbb{P}\left[D_n = m\right] = \frac{12}{7} \frac{(n+2+m)(n-1-m)}{n^2(m+1)(m+2)} - \frac{12}{7} \frac{m^{\underline{5}}}{n n^{\underline{6}}} + \frac{12}{7n^2},$$

for $5 \le m < n$. The probability that a random node has n descendants is $\mathbb{P}\left[D_n = n\right] = \frac{1}{n}$. Furthermore, the probability that a random node in a random LBST of size n has no children is

$$\mathbb{P}\left[D_n = 1\right] = \frac{6}{7} \frac{1}{n^2} \binom{n+1}{2} = \frac{3}{7}\left(1 + \frac{1}{n}\right).$$

In the appendix, a table collects the general result for $5 \le m < n$ as well as the special cases where $m < 5$ or $m = n$.

Proof. If we consider the explicit form for $\mathsf{D}_z(z,u,v)$ given in Theorem 6.1 and average w.r.t. j, i.e. we plug $u=1$ there, we get

$$\begin{aligned}\overline{\mathsf{D}}_z(z,v) &= \frac{v}{(1-vz)^2} + \frac{12}{7(1-v)(1-z)} - \frac{24}{7v(1-z)^2} + \frac{2(v^2-6v+12)}{7v(1-z)^3} \\ &\quad + \frac{v}{7(1-v)^5}\Big[-15(1-v)^3 + 20v(1-v)^2(1-z) - 30v^2(1-v)(1-z)^2 \\ &\qquad + 60v^3(1-z)^3 + (1-7v+23v^2-57v^3-22v^4+2v^5)(1-z)^4\Big] \\ &\qquad - \frac{60}{7}\frac{v^5(1-z)^4}{(1-v)^6}\log\frac{1}{1-z} + \Big(\frac{60}{7}\frac{v^5(1-z)^4}{(1-v)^6} - \frac{24}{7}\frac{1-v}{v^2(1-z)^3}\Big)\log\frac{1}{1-vz}. \qquad (15)\end{aligned}$$

Alternatively, we can write down the differential equation for $\overline{\mathsf{D}}_z(z,v) = \mathsf{D}_z(z,1,v)$ and solve it. The differential equation is

$$\frac{1}{6}\frac{\partial^2\overline{\mathsf{D}}_z}{\partial z^2} = \frac{v^3}{(1-vz)^4} + 2\frac{\overline{\mathsf{D}}_z}{(1-z)^2},$$

where the initial conditions are $\overline{\mathsf{D}}_z(0,v) = v$, and $\frac{\partial}{\partial z}\overline{\mathsf{D}}_z(0,v) = 2v(1+v)$. The reader may readily check that the explicit form given in Equation (15) is a solution to the differential equation above. The purely rational term in Equation (15), i.e. the one that is not multiplied by any logarithmic function, although more complicated than the others, has the very pleasant feature that "almost" all coefficients are $\frac{12}{7}$. On the other hand,

$$[z^n v^m]\frac{1}{(1-z)^3}\log\frac{1}{1-vz} = \frac{1}{m}\binom{n-m+2}{2},$$

and thus

$$-[z^n v^m]\frac{24}{7}\frac{1-v}{v^2}\frac{1}{(1-z)^3}\log\frac{1}{1-vz} = \frac{12}{7}\frac{(n+3+m)(n-m)}{(m+2)(m+1)}.$$

This is the main contribution in the coefficient $z^n v^m$ of $\overline{\mathsf{D}}_z(z,v)$, the remaining contributions being small. Indeed,

$$\frac{60}{7}\frac{v^5(1-z)^4}{(1-v)^6}\log\frac{1}{1-vz}$$

produces no coefficients at all, since $m \le n$. And the remaining contribution comes from

$$-[z^n v^m]\frac{60}{7}\frac{v^5(1-z)^4}{(1-v)^6}\log\frac{1}{1-z} = -\frac{12}{7}\frac{m^{\underline{5}}}{n^{\underline{5}}}.$$

The general part of the theorem follows from the considerations made above. The special cases, when $m<5$ or $m=n$ have to be dealt with separately. In particular, to get the probability that a random node in a random LBST of size n has no children (the special case $m=1$) we compute

$$\left.\frac{\partial\overline{\mathsf{D}}_z(z,v)}{\partial v}\right|_{v=0} = \frac{6}{7}\frac{1}{(1-z)^3} + \frac{1}{7}(1-z)^4,$$

extract the coefficient of z^{n-1} in the GF above and divide by n^2, yielding $\mathbb{P}[D_n=1] \sim 3/7$. Also, for evident reasons, $\mathbb{P}[D_n=n] = 1/n$, since only the root has n descendants and we choose it with probability $1/n$. □

Finally, the moments of D_n can be computed after differentiation of $\overline{\mathsf{D}}_z(z,v)$, whose explicit form was given in the proof above. We state now the following result.

Theorem 6.5 *Let $d_n^{(s)} = \mathbb{E}\left[(D_n+2)^{\underline{s}}\right]$, i.e., $d_n^{(s)}$ is the shifted s^{th} factorial moment of the number of descendants of a random node in a random locally balanced binary search tree of size n. Furthermore, let $d_n = d_n^{(1)}$. Then*

1. $d_n = \frac{12}{7}\left(1+\frac{1}{n}\right)H_n - \frac{1}{49}\left(26-\frac{9}{n}\right)$, *for* $n \geq 6$.
2. $d_n^{(2)} = \dfrac{5(n+1)(7n+2)}{14n}$, *for* $n \geq 6$.
3. $d_n^{(3)} = \dfrac{(n+1)(10n^2+5n+6)}{6n}$, *for* $n \geq 6$.
4. *For all* $n \geq s+7$ *and all* $s \geq 4$,
$$d_n^{(s)} = \frac{A(s,n)\,(n+1)^{\underline{s+1}}}{(s+6)^{\underline{6}}\,(n+2-s)^{\underline{2}}\,n\,n^{\underline{6}}\,(s-1)},$$
where
$$\begin{aligned}A(s,n) &= (s+5)^{\underline{5}}(s+3)(s+2)n^7\\ &\quad-(s+4)^{\underline{4}}(s+2)\Big(13s^2+128s+195\Big)n^6\\ &\quad+(s+3)^{\underline{3}}\Big(67s^4+1082s^3+6125s^2+11326s+6600\Big)n^5\\ &\quad-5(s+2)^{\underline{2}}\Big(35s^5+643s^4+4459s^3+15317s^2+15906s+3960\Big)n^4\\ &\quad+4(s+1)\Big(61s^6+1159s^5+8157s^4+24383s^3+60116s^2-9276s-31680\Big)n^3\\ &\quad-4\Big(43s^7+794s^6+5176s^5+10190s^4-80183s^3+29336s^2-220956s-77040\Big)n^2\\ &\quad+48\Big(s^7+17s^6+97s^5+215s^4+1894s^3-39832s^2+41208s-25200\Big)n\\ &\quad-1036800s^2+3110400s-2073600.\end{aligned}$$

Corollary 6.1 *For any $n \geq 6$ and for $j = \alpha n$, with $0 < \alpha < 1$, we have*
$$\begin{aligned}\mathbb{E}\left[D_{n,\alpha n}\right] &= \frac{12}{7}\log n + \mathcal{O}(1),\\ \mathbb{V}\left[D_{n,\alpha n}\right] &= -\frac{3}{5}n\Big(11\alpha(1-\alpha)+12\alpha\log\alpha+12(1-\alpha)\log(1-\alpha)\Big)+\mathcal{O}(\log^2 n).\end{aligned}$$

As in Section 3, several interesting corollaries may be deduced from the results in this section and Propositions 1.4 and 1.5.

Corollary 6.2 *The expected number of pages in a random locally balanced search tree of size n with page capacity $b \geq 2$ is*
$$\mathbb{E}\left[\mathcal{P}_n^{(b)}\right] \sim \frac{12}{7}\frac{n}{b+2}.$$
The filling ratio *for locally balanced search trees is thus*
$$\gamma_b = \frac{n/b}{\mathbb{E}\left[\mathcal{P}_n^{(b)}\right]} \sim \frac{7}{12} = 0.58333\ldots$$

Corollary 6.3 *The expected number of recursive calls to sort a random permutation of size n, when the recursion stops at subfiles of size $\le b$ and the pivots are selected as the median of samples of three elements, is*

$$\mathbb{E}\left[R_n^{(b)}\right] = \mathbb{E}\left[\mathcal{P}_n^{(b)}\right] - 1 \sim \frac{12}{7}\frac{n}{b+2}.$$

Also of interest is the expectation $C_{n,b} := \mathbb{E}\left[C_n^{(b)}\right]$ of the number of comparisons to sort a random permutation of size n with quicksort, where the pivots are selected as the median of samples of three elements (for subfiles of length $n \ge 3$) and the recursion stops at subfiles of size $\le b$. To get these expectations we don't use Proposition 1.5. We take another approach and start with the following recursion for $C_{n,b}$:

$$C_{n,b} = n - 1 + \sum_{k=1}^{n} \pi_{n,k}\left(C_{k-1,b} + C_{n-k,b}\right) \quad \text{for } n > b \ge 0 \text{ and } n \ge 3\,, \tag{16}$$

with initial values $C_{2,0} = 1$, $C_{2,1} = 1$ and $C_{n,b} = 0$ otherwise. (With these initial values we take care of the one additional comparison, sorting a subfile of length 2, when the pages are smaller than 2.)

To solve this recurrence, we introduce the bivariate generating function $\mathsf{C}_z(z,v) = \sum_{n>b\ge 0} C_{n,b} n z^{n-1} v^b$. Multiplying both sides of equation (16) by $n(n-1)(n-2)z^{n-3}v^b$ and summing up over all $n > b \ge 0$ leads to the following differential equation

$$\begin{aligned}\frac{\partial^2}{\partial z^2}\mathsf{C}_z(z,v) = {}& \frac{12}{(1-z)^2}\mathsf{C}_z(z,v) \\ &+ \frac{12(z^6v^4+z^5v^4+z^5v^3-15\,z^4v^3+10\,z^3v^3+10\,z^3v^2-5\,z^2v^3+5\,z^2v^2-5\,z^2v+zv^3-4\,zv^2-4\,zv+z+v^2+v+1)}{(1-z)^5\,(1-zv)^5},\end{aligned} \tag{17}$$

with initial conditions $\mathsf{C}_z(0,v) = 0$ and $\frac{\partial}{\partial z}\mathsf{C}_z(0,v) = 2(1+v)$.

This differential equation is of Eulerian type, and can be solved easily. We get then

$$\begin{aligned}\mathsf{C}_z(z,v) = {}& \left(\frac{120}{7}\,\frac{(1-z)^4\,(v+2)\,v^5}{(1-v)^8} + \frac{24}{7}\,\frac{1}{(1-v)\,(1-z)^3}\right)\log\frac{1}{1-z} \\ &+ \left(-\frac{120}{7}\,\frac{(1-z)^4\,(v+2)\,v^5}{(1-v)^8} + \frac{24}{7}\,\frac{2\,v-3}{(1-v)\,(1-z)^3\,v^2}\right)\log\frac{1}{1-zv} \\ &- 12\,\frac{v}{(1-v)^3\,(1-zv)} + 2\,\frac{v}{(1-zv)^2\,(1-v)} - 2\,\frac{v}{(1-zv)^3\,(1-v)} \\ &- \frac{2}{49}\,\frac{89\,v-252}{(1-v)\,(1-z)^3\,v} - \frac{2}{7}\,\frac{7\,v^2-31\,v+36}{(1-z)^2\,(1-v)^2\,v} - \frac{12}{7}\,\frac{2\,v-3}{(1-v)^3\,(1-z)} \\ &+ \frac{2}{49}\,\frac{R(z,v)}{(1-v)^7}\,.\end{aligned} \tag{18}$$

with

$$\begin{aligned}R(z,v) = {}& 40\,z^4v^6 + 929\,z^4v^5 + 327\,z^4v^4 - 23\,z^4v^3 - 23\,z^4v^2 + 12\,z^4v - 2\,z^4 - 160\,z^3v^6 - 3296\,z^3v^5 \\ &- 468\,z^3v^4 + 92\,z^3v^3 + 92\,z^3v^2 - 48\,z^3v + 8\,z^3 + 240\,z^2v^6 + 4104\,z^2v^5 - 768\,z^2v^4 + 282\,z^2v^3 \\ &- 138\,z^2v^2 + 72\,z^2v - 12\,z^2 - 160\,zv^6 - 1896\,zv^5 + 1632\,zv^4 - 1168\,zv^3 + 372\,zv^2 - 48\,zv \\ &+ 8\,z + 40\,v^6 + 54\,v^5 - 618\,v^4 + 1132\,v^3 - 828\,v^2 + 222\,v - 2\,.\end{aligned}$$

Extracting the coefficients, we get with $\mathbb{E}\left[C_n^{(b)}\right] = C_{n,b} = \frac{1}{n}[z^{n-1}v^b]\mathsf{C}_z(z,v)$ the required expectations. This leads to

Theorem 6.6 *The expected number of comparisons to sort a random permutation of size n, when the recursion stops in subfiles of size $\leq b$ and the pivots are selected as the median of samples of three elements, is for $n > b \geq 0$ and $n \geq 6$ given as*

$$\mathbb{E}\left[C_n^{(b)}\right] = \frac{12}{7}(n+1)H_n - \frac{12}{7}(n+1)H_{b+1} + \frac{37n}{49} + \frac{219}{49} - \frac{36(n+1)}{7(b+2)} + \frac{4(3b-1)(b+1)^{\underline{6}}}{49n^{\underline{6}}}.$$

7. The number of ascendants of a given node in a LBST

As in the case of the number of ascendants in a random BST, computing the probability that the j^{th} node in a random LBST has m ascendants turns out to be an extremely difficult problem.

However, the recursive definition can easily be translated to a differential equation for the corresponding generating function $\mathsf{A}_z(z,u,v)$. Because of the same technical reason discussed in Section 6, the function $\mathsf{A}_z(z,u,v)$ is actually the derivative w.r.t. z of the generating function such that the coefficient of $z^n u^j$ is the PGF of $A_{n,j}$. The recurrence for $A_{n,j}$

$$A_{n,j} = \sum_{k=1}^{j-1} \pi_{n,k}(A_{n-k,j-k}+1) + \pi_{n,j} + \sum_{k=j+1}^{n} \pi_{n,k}(A_{k-1,j}+1) \quad \text{for } n \geq 3\,,$$

with initial values $A_{0,j} = 0$, $A_{1,1} = 1$, $A_{2,1} = \frac{3}{2}$, $A_{2,2} = \frac{3}{2}$ and $A_{n,j} = 0$ otherwise, translates into the second-order differential equation

$$\frac{1}{6}\frac{\partial^2 \mathsf{A}_z}{\partial z^2} = \frac{v}{(1-z)^2}\mathsf{A}_z + \frac{u^2 v}{(1-uz)^2}\mathsf{A}_z + \frac{u^2 v}{(1-z)^2(1-uz)^2}, \tag{19}$$

and the initial values are $\mathsf{A}_z(0,u,v) = uv$ and $\frac{\partial}{\partial z}\mathsf{A}_z(0,u,v) = uv(1+v)(1+u)$. This differential equation is the starting point for our next theorems.

Theorem 7.1 *The expected number of ascendants $a_{n,j} = \mathbb{E}[A_{n,j}]$ of the j^{th} node in a random locally balanced search tree of size n is*

$$\begin{aligned} a_{n,j} &= \frac{24}{35}H_n + \frac{18}{35}H_j + \frac{18}{35}H_{n+1-j} \\ &+ \frac{12}{35j} + \frac{12}{35(n+1-j)} - \frac{279}{175} - \frac{6}{7n} \\ &+ \frac{18j}{35n} - \frac{12(j-1)^2}{35n^{\underline{2}}} - \frac{4(2j-3)(j-1)^{\underline{2}}}{35n^{\underline{3}}} \\ &- \frac{6(j-2)(j-1)^{\underline{3}}}{35n^{\underline{4}}} + \frac{6(2j-5)(j-1)^{\underline{4}}}{35n^{\underline{5}}} - \frac{4(j-3)(j-1)^{\underline{5}}}{35n^{\underline{6}}}, \end{aligned}$$

for $5 \leq j \leq n-4$. In the appendix we give alse the cases $j = 1,2,3,4$. The cases where $j > n-4$ follow from the special cases with $j \leq 4$ and the symmetry in j and $n+1-j$ of $a_{n,j}$.

Proof. Although it is in principle possible to solve the differential equation (19), it is sufficient for our purpose to take derivatives w.r.t. v and setting $v = 1$, to get the differential equation for $\mathcal{A}_z(z,u)$, the generating function whose coefficients are the expected values $a_{n,j} = \mathbb{E}[A_{n,j}]$. It is

$$\frac{1}{6}\frac{\partial^2 \mathcal{A}_z}{\partial z^2} - \left(\frac{1}{(1-z)^2} + \frac{u^2}{(1-uz)^2}\right)\mathcal{A}_z = \frac{u}{1-u}\left(\frac{1}{(1-z)^4} - \frac{u^3}{(1-uz)^4}\right), \tag{20}$$

and the initial conditions are now $\mathcal{A}_z(0,u) = u$ and $\frac{\partial}{\partial z}\mathcal{A}_z(0,u) = 3u(1+u)$.

The solution of the differential equation (20) yields the explicit form

$$\begin{aligned}\mathcal{A}_z(z,u) &= \frac{D_0(z,u)}{(1-z)^2(1-u)^7(1-uz)^2}\\ &+ \frac{D_1(z,u)}{(1-z)^2(1-u)^7(1-uz)^2}\log\frac{1}{1-z}\\ &+ \frac{D_2(z,u)}{(1-z)^2(1-u)^7(1-uz)^2}\log\frac{1}{1-uz},\end{aligned}$$

where the polynomials $D_i(z,u)$ can be found in the appendix.
Once we have the explicit form for $\mathcal{A}_z$, extracting the coefficients is just a matter of patience and careful computations. A possible shortcut is to expand each of the three main parts of $\mathcal{A}_z$ as power series in z and u, and spot a pattern in the shape of the coefficients. The inspired guesses can be readily checked and proved by induction. For instance, the coefficient of $z^n u^j$ in the purely rational term of $\mathcal{A}_z$ is

$$-\frac{69}{175}n + \frac{18}{35}j - \frac{9}{175},$$

whenever $5 \le j \le n-4$; the remaining values of j are special cases that we have to consider separately. Similarly, the coefficient of $z^n u^j$ in the second term —the one that contains $\log(1/(1-uz))$ as a factor— is

$$\frac{18H_j(n+1)}{35} - \frac{18j}{35} + \frac{12n}{35j} - \frac{12}{35} + \frac{12}{35j}.$$

In the same vein, an explicit formula for the coefficient of the first term can be obtained. Finally, we collect everything, consider the coefficient $z^{n-1}u^j$ and divide by n, since A_z is a derivative w.r.t. z. □

The differential equation (20) is exactly the same as the one for the number of passes in quickselect with median-of-three (see Proposition 1.3). The only difference between the expected number of passes in quickselect, as given in the work by Kirschenhofer et al. [16], and the number of ascendants in LBSTs relies on the initial conditions. The reason is that in the mentioned paper only one recursive call is counted if we want to select some element in a file of size ≤ 2, while the average number of ascendants of the j^{th} node in a random LBST of size $n \le 2$ is $3/2$ (for $j=1$ and $j=2$). Then $a_{n,j}$ and the expected number of passes to select the j^{th} element out of n differ in the constant term, exactly by $1/7$.

In a similar way, when differentiating the differential equation (19) two times w.r.t. v and setting $v=1$, we get the differential equation for $\mathcal{A}_z^{(2)}(z,u)$, the generating function whose coefficients are the second factorial moments $a_{n,j}^{(2)}$ of the number of ascendants. We get then

Theorem 7.2 *The second factorial moment $a_{n,j}^{(2)} = \mathbb{E}\left[(A_{n,j})^{\underline{2}}\right]$ of the number of ascendants of the j^{th} node in a random locally balanced search tree of size n is*

$$\begin{aligned}a_{n,j}^{(2)} = &\frac{576}{245}H_nH_{n-j} + \frac{576}{245}H_nH_j - \frac{288}{245}H_n^2 + \frac{324}{1225}H_{n-j}^2 - \frac{1368}{1225}H_{n-j}H_j + \frac{324}{1225}H_j^2 + \\ &+ \left(\frac{72}{42875}\frac{227j+525}{j} + \frac{792}{245}\frac{1}{n+1-j} - \right.\\ &- \frac{2}{1225}\frac{2j^6-6j^5-55j^4+120j^3+953j^2-1014j+540}{n} +\end{aligned}$$

$$
\begin{aligned}
&+\frac{4}{245}\frac{(j-1)^2(j^4-4j^3-11j^2+30j+72)}{n-1}-\\
&-\frac{4}{245}\frac{(j-1)(j-2)(2j^4-12j^3+13j^2+15j-36)}{n-2}+\\
&+\frac{8}{245}\frac{(j-1)^2(j-2)^2(j-3)^2}{n-3}-\\
&-\frac{2}{245}\frac{(j-1)(j-2)(j-3)(j-4)(2j^2-10j+15)}{n-4}+\\
&+\frac{4}{1225}\frac{(j-1)(j-2)(j-3)^2(j-4)(j-5)}{n-5}\Bigg)H_n+\\
+&\Bigg(-\frac{36}{42875}\frac{5697j+70}{j}-\frac{288}{245}\frac{1}{n+1-j}+\\
&+\frac{1}{6125}\frac{20j^6-102j^5-445j^4+2040j^3+8165j^2+402j+360}{n}-\\
&-\frac{2}{1225}\frac{(j-1)(10j^5-71j^4+14j^3+557j^2-42j-1224)}{n-1}+\\
&+\frac{2}{1225}\frac{(j-1)(j-2)(20j^4-162j^3+319j^2+129j-612)}{n-2}-\\
&-\frac{4}{1225}\frac{(j-1)(j-2)(j-3)(10j^3-81j^2+194j-102)}{n-3}+\\
&+\frac{1}{1225}\frac{(j-1)(j-2)(j-3)(j-4)(20j^2-142j+255)}{n-4}-\\
&-\frac{2}{6125}\frac{(j-1)(j-2)(j-3)(j-4)(j-5)(10j-51)}{n-5}\Bigg)H_{n-j}+\\
+&\Bigg(-\frac{36}{42875}\frac{2267j+2030}{j}-\frac{288}{245}\frac{1}{n+1-j}+\\
&+\frac{1}{6125}\frac{20j^6-18j^5-655j^4+360j^3+10895j^2-20682j+10440}{n}-\\
&-\frac{2}{1225}\frac{(j-1)(10j^5-29j^4-154j^3+263j^2+882j-216)}{n-1}+\\
&+\frac{2}{1225}\frac{(j-1)(j-2)(20j^4-78j^3-59j^2+171j-108)}{n-2}-\\
&-\frac{4}{1225}\frac{(j-1)(j-2)(j-3)(10j^3-39j^2+26j-18)}{n-3}+\\
&+\frac{1}{1225}\frac{(j-1)(j-2)(j-3)(j-4)(20j^2-58j+45)}{n-4}-\\
&-\frac{2}{6125}\frac{(j-1)(j-2)(j-3)(j-4)(j-5)(10j-9)}{n-5}\Bigg)H_j+\\
+&\frac{288}{245}H_n^{(2)}-\frac{3204}{1225}H_{n-j}^{(2)}-\frac{3204}{1225}H_j^{(2)}+\\
+&\frac{12}{1225}\frac{1}{j+4}-\frac{24}{175}\frac{1}{j+3}+\frac{792}{1225}\frac{1}{j+2}-\frac{2088}{1225}\frac{1}{j+1}-\\
-&\frac{53148}{42875}\frac{1}{j}-\frac{1044}{875}\frac{1}{j-1}+\frac{396}{875}\frac{1}{j-2}-\frac{12}{125}\frac{1}{j-3}+\frac{6}{875}\frac{1}{j-4}+\frac{144}{175}\frac{1}{j^2}+
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{6}{875}\frac{1}{n-3-j} - \frac{12}{125}\frac{1}{n-2-j} + \frac{396}{875}\frac{1}{n-1-j} - \frac{1044}{875}\frac{1}{n-j} - \\
&- \frac{24}{8575}\frac{1123j-420}{(n+1-j)j} - \frac{2088}{1225}\frac{1}{n+2-j} + \frac{792}{1225}\frac{1}{n+3-j} - \frac{24}{175}\frac{1}{n+4-j} + \\
&+ \frac{12}{1225}\frac{1}{n+5-j} - \frac{792}{245}\frac{1}{(n+1-j)^2} + \frac{2297696}{214375} - \\
&- \frac{1}{643125}(5522j^7 - 8166j^6 - 167395j^5 + 142845j^4 + 3249683j^3 - 1044579j^2 - \\
&\qquad - 2445930j + 1096200)\frac{1}{nj} + \\
&+ \frac{2}{128625}(2761j^7 - 12366j^6 - 25787j^5 + 116883j^4 + 176746j^3 - 361269j^2 + \\
&\qquad + 120672j + 22680)\frac{1}{(n-1)j} - \\
&- \frac{2}{128625}(5522j^7 - 41298j^6 + 82913j^5 + 11427j^4 - 226090j^3 + 256698j^2 - \\
&\qquad - 49482j - 22680)\frac{1}{(n-2)j} + \\
&+ \frac{4}{128625}(j-1)(j-3)(2761j^5 - 17888j^4 + 38198j^3 - 28798j^2 - \\
&\qquad - 5718j + 3780)\frac{1}{(n-3)j} - \\
&- \frac{1}{128625}(5522j^7 - 74430j^6 + 407885j^5 - 1158375j^4 + 1755053j^3 - 1233465j^2 + \\
&\qquad + 169920j + 113400)\frac{1}{(n-4)j} + \\
&+ \frac{2}{643125}(2761j^7 - 45498j^6 + 299185j^5 - 999255j^4 + 1757614j^3 - 1456887j^2 + \\
&\qquad + 313560j + 113400)\frac{1}{(n-5)j} + \\
&+ \frac{2}{1225}(2j^6 - 6j^5 - 55j^4 + 120j^3 + 953j^2 - 1014j + 540)\frac{1}{n^2} - \\
&- \frac{4}{245}(j-1)^2(j^4 - 4j^3 - 11j^2 + 30j + 72)\frac{1}{(n-1)^2} + \\
&+ \frac{4}{245}(j-1)(j-2)(2j^4 - 12j^3 + 13j^2 + 15j - 36)\frac{1}{(n-2)^2} - \\
&- \frac{8}{245}(j-1)^2(j-2)^2(j-3)^2)\frac{1}{(n-3)^2} + \\
&+ \frac{2}{245}(j-1)(j-2)(j-3)(j-4)(2j^2 - 10j + 15)\frac{1}{(n-4)^2} - \\
&- \frac{4}{1225}\frac{(j-1)(j-2)(j-3)^2(j-4)(j-5)}{(n-5)^2}
\end{aligned}
$$

for $n \geq 8$ and $5 \leq j \leq n-4$. In the appendix we give the remaining cases $j = 1, 2, 3, 4$. The cases where $j > n-4$ follow from the special cases with $j \leq 4$ and the symmetry in j and $n+1-j$ of $a_{n,j}$.

The next theorem gives the variance, which is now easy to obtain.

Theorem 7.3 *The variance $\mathbb{V}[A_{n,j}]$ of the number of ascendants of the j^{th} node in a random locally balanced search tree of size n is*

$$
\begin{aligned}
\mathbb{V}[A_{n,j}] = {} & \frac{288}{175} H_n H_j + \frac{288}{175} H_n H_{n+1-j} - \frac{288}{175} H_n^2 - \frac{288}{175} H_j H_{n+1-j} + \\
& + \left(\frac{139488}{42875} + \frac{72}{175}\frac{1}{j} + \frac{72}{175}\frac{1}{n+1-j} - \frac{72}{175}\frac{6j-5}{n} + \frac{288}{175}\frac{(j-1)^2}{n^{\underline{2}}} + \right. \\
& \quad + \frac{96}{175}\frac{(2j-3)(j-1)^{\underline{2}}}{n^{\underline{3}}} + \frac{144}{175}\frac{(j-2)(j-1)^{\underline{3}}}{n^{\underline{4}}} - \frac{144}{175}\frac{(2j-5)(j-1)^{\underline{4}}}{n^{\underline{5}}} + \\
& \quad \left. + \frac{96}{175}\frac{(j-3)(j-1)^{\underline{5}}}{n^{\underline{6}}} \right) H_n + \\
& + \left(-\frac{112734}{42875} - \frac{72}{175}\frac{1}{j} - \frac{72}{35}\frac{1}{n+1-j} + \frac{72}{35}\frac{2j-1}{n} - \frac{144}{175}\frac{(j-1)^{\underline{2}}}{n^{\underline{2}}} - \right. \\
& \quad \left. - \frac{96}{175}\frac{(j-1)^{\underline{3}}}{n^{\underline{3}}} - \frac{72}{175}\frac{(j-1)^{\underline{4}}}{n^{\underline{4}}} + \frac{144}{175}\frac{(j-1)^{\underline{5}}}{n^{\underline{5}}} - \frac{48}{175}\frac{(j-1)^{\underline{6}}}{n^{\underline{6}}} \right) H_{n+1-j} + \\
& + \left(\frac{10746}{42875} - \frac{72}{35}\frac{1}{j} - \frac{72}{175}\frac{1}{n+1-j} - \frac{72}{175}\frac{4j-5}{n} - \frac{144}{175}\frac{j^{\underline{2}}}{n^{\underline{2}}} - \frac{96}{175}\frac{j^{\underline{3}}}{n^{\underline{3}}} - \right. \\
& \quad \left. - \frac{72}{175}\frac{j^{\underline{4}}}{n^{\underline{4}}} + \frac{144}{175}\frac{j^{\underline{5}}}{n^{\underline{5}}} - \frac{48}{175}\frac{j^{\underline{6}}}{n^{\underline{6}}} \right) H_j + \\
& + \frac{288}{245} H_n^{(2)} - \frac{3204}{1225} H_{n+1-j}^{(2)} - \frac{3204}{1225} H_j^{(2)} + \\
& + \frac{1}{926100000}\frac{1}{j^{\underline{5}}} (49336 n^2 j^{17} - 49336 n j^{18} - 1865416 n^2 j^{16} + 1671908 n j^{17} + \\
& \quad + 242844 j^{18} + 32784304 n^2 j^{15} - 25494056 n j^{16} - 9398508 j^{17} - 355430740 n^2 j^{14} + \\
& \quad + 227842328 n j^{15} + 169528380 j^{16} + 2660193907 n^2 j^{13} - 1283644187 n j^{14} - \\
& \quad - 1892353176 j^{15} - 14573920102 n^2 j^{12} + 4328697506 n j^{13} + 14637396963 j^{14} + \\
& \quad + 60447522743 n^2 j^{11} - 4676604687 n j^{12} - 83250254661 j^{13} - \\
& \quad - 193475395850 n^2 j^{10} - 36170564954 n j^{11} + 360438321705 j^{12} + \\
& \quad + 482054276453 n^2 j^9 + 246940887347 n j^{10} - 1212564043197 j^{11} - \\
& \quad - 935042746298 n^2 j^8 - 863949222686 n j^9 + 3203516805087 j^{10} + \\
& \quad + 1401970954277 n^2 j^7 + 2046716285867 n j^8 - 6664776231879 j^9 - \\
& \quad - 1606665630530 n^2 j^6 - 3484586512526 n j^7 + 10876953083775 j^8 + \\
& \quad + 1399475657604 n^2 j^5 + 4312068202636 n j^6 - 13811250629703 j^7 - \\
& \quad - 951315080184 n^2 j^4 - 3893967885528 n j^5 + 13559528011686 j^6 + \\
& \quad + 548876398176 n^2 j^3 + 2697126473376 n j^4 - 10418235359052 j^5 - \\
& \quad - 276724874880 n^2 j^2 - 1671052102848 n j^3 + 6626911103640 j^4 + \\
& \quad + 88189171200 n^2 j + 957718391040 n j^2 - 3803838011424 j^3 + 2612736000 n^2 - \\
& \quad - 274471027200 n j + 2026211137920 j^2 - 108428544000 n - 1064459750400 j + \\
& \quad + 555206400000) \frac{1}{(n+1-j)(n-4)(n-5)} - \\
& - \frac{1}{926100000}\frac{1}{j^{\underline{2}}} (3836 j^{14} - 31892 j^{13} - 92960 j^{12} + 1268876 j^{11} + 1190343 j^{10} -
\end{aligned}
$$

$$
\begin{aligned}
&- 26514831j^9 + 11837170j^8 + 245828918j^7 - 345582349j^6 - 868029827j^5 + \\
&+ 4697373240j^4 - 4702798044j^3 - 3343999680j^2 + 4329547200j - \\
&- 1360800000)\frac{1}{n} - \\
- &\frac{1}{9261000}\frac{1}{j(j-2)}(455j^{14} - 7630j^{13} + 46900j^{12} - 96320j^{11} - 262290j^{10} + \\
&+ 1720740j^9 - 2168948j^8 - 5125976j^7 + 18165755j^6 - 13883206j^5 - \\
&- 41023200j^4 + 119810088j^3 - 98394480j^2 + 16137792j + 10886400)\frac{1}{n-1} - \\
- &\frac{1}{2315250}\frac{1}{j(j-3)}(140j^{14} - 4200j^{13} + 54040j^{12} - 389550j^{11} + 1721685j^{10} - \\
&- 4737915j^9 + 7600984j^8 - 4804158j^7 - 6353236j^6 + 20120466j^5 - \\
&- 28218987j^4 + 24854661j^3 - 8887482j^2 - 6263568j + 4082400)\frac{1}{n-2} + \\
+ &\frac{1}{1157625}\frac{(j-1)(j-3)}{j(j-4)}(70j^{12} - 1050j^{11} + 3500j^{10} + 33810j^9 - 404670j^8 + \\
&+ 2022930j^7 - 6035383j^6 + 11464746j^5 - 13898405j^4 + 11040600j^3 - \\
&- 7054452j^2 + 4642704j - 907200)\frac{1}{n-3} + \\
+ &\frac{12}{1225}\frac{1}{n+5-j} - \frac{24}{175}\frac{1}{n+4-j} + \frac{792}{1225}\frac{1}{n+3-j} - \frac{2088}{1225}\frac{1}{n+2-j} - \\
- &\frac{1044}{875}\frac{1}{n-j} + \frac{396}{875}\frac{1}{n-1-j} - \frac{12}{125}\frac{1}{n-2-j} + \frac{6}{875}\frac{1}{n-3-j} + \\
+ &\frac{864}{1225}\frac{1}{(n+1-j)^2} + \\
+ &\frac{2}{214375}(705517j^{10} - 194925j^9 - 21064710j^8 + 5812050j^7 + 189569541j^6 - \\
&- 52136385j^5 - 550488940j^4 + 146201100j^3 + 318001392j^2 + 12130560j + \\
&+ 43545600)\frac{1}{j\,(j+4)^{\underline{9}}} - \\
- &\frac{1}{4410000}(4j^{12} - 24j^{11} - 184j^{10} + 1140j^9 + 5397j^8 - 28692j^7 - 95542j^6 + \\
&+ 392100j^5 + 1140049j^4 - 2969484j^3 - 7205724j^2 + 8760960j - 3369600)\frac{1}{n^2} - \\
- &\frac{1}{44100}(j-1)^2(j^4 - 4j^3 - 11j^2 + 30j + 72) \\
&(j^6 - 6j^5 - 2j^4 + 48j^3 + j^2 - 114j + 792)\frac{1}{(n-1)^2} - \\
- &\frac{1}{44100}(j-1)(j-2)(2j^6 - 18j^5 + 53j^4 - 48j^3 - 55j^2 + 138j - 792) \\
&(2j^4 - 12j^3 + 13j^2 + 15j - 36)\frac{1}{(n-2)^2} - \\
- &\frac{1}{11025}(j-1)^2(j-2)^2(j-3)^2(j^6 - 12j^5 + 58j^4 - 144j^3 + 193j^2 - 132j + 396)\frac{1}{(n-3)^2} -
\end{aligned}
$$

$$- \frac{1}{176400}(j-1)^{\underline{4}}(2j^2-10j+15)(2j^6-30j^5+185j^4-600j^3+1073j^2-990j-1080)\frac{1}{(n-4)^2} -$$
$$- \frac{1}{1102500}(j-3)(j-1)^{\underline{5}}(j^6-18j^5+130j^4-480j^3+949j^2-942j+3960)\frac{1}{(n-5)^2}$$

for $n \geq 8$ and $5 \leq j \leq n-4$. The remaining cases for $j = 1, 2, 3, 4$ can also be found in the appendix.

□

In [26] the authors considered the expectation and variance of A_n in random LBSTs. To be more precise, they stated the problem in terms of unsuccessful search costs. Here, we are able to reproduce their results and extend them to higher order moments. Since we deal with ascendants of internal nodes, our results can be naturally stated in terms of successful search costs, and then translated to unsuccessful costs using Proposition 1.2. where $a_n = \mathbb{E}[A_n]$ is the expected number of ascendants of a random node in a random tree with n nodes.

Theorem 7.4 *Let*

$$\overline{\mathsf{A}}_z(z,v) = \frac{\partial}{\partial z}\sum_{n,m} \mathbb{P}[A_n = m]\; z^n v^m .$$

Then

$$\begin{aligned}\overline{\mathsf{A}}_z(z,v) &= \frac{v}{(1-2v)(1-z)^2} \\ &\quad - \frac{v^2}{(1-2v)\Delta}\Big((\Delta+4v+3)(1-z)^{-(\Delta-1)/2} + (\Delta-4v-3)(1-z)^{(\Delta+1)/2}\Big),\end{aligned}$$

where $\Delta = \sqrt{1+48v}$.

Proof. The differential equation to be solved (from Equation (19), plugging $u = 1$) is

$$\frac{1}{6}\frac{\partial^2 \overline{\mathsf{A}}_z}{\partial z^2} = \frac{2v}{(1-z)^2}\overline{\mathsf{A}}_z + \frac{v}{(1-z)^4},$$

where $\overline{\mathsf{A}}_z(z,v) = \mathsf{A}_z(z,1,v)$ and the initial conditions are $\overline{\mathsf{A}}_z(0,v) = v$ and $\frac{\partial}{\partial z}\overline{\mathsf{A}}_z(0,v) = 2v(1+v)$. Recall that $\overline{\mathsf{A}}_z(z,v) = \mathsf{A}_z(z,1,v)$. The solution of the differential equation above is the explicit form given in the theorem. □

Extracting coefficients in exact form from there is quite difficult. However, as Philippe Flajolet kindly pointed to us, asymptotic information and most notably, the limiting probability distribution can be established [7, 14]. In this case, it follows that A_n converges in distribution (converges in law) to a Gaussian distribution, i.e.

$$\mathbb{P}\left[\frac{A_n - \frac{12}{7}\log n}{\sqrt{\frac{300}{343}\log n}} < x\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right).$$

This result follows from the asymptotic estimation for the average and the variance of A_n and the fact that $\overline{\mathsf{A}}_z(z,v)$ is essentially a quasi-power of $[z^n]\overline{\mathsf{A}}_z(z,v)$ in a neighborhood of $v = 1$, i.e.

$$[z^n]\overline{\mathsf{A}}_z(z,v) = c\cdot n^{(\Delta-3)/2)}\left(1 + \mathcal{O}(1/n)\right),$$

and the error term is uniformly bounded.

The next step in our programme is to differentiate $\overline{\mathsf{A}}_z$ as many times as needed w.r.t. v and set $v = 1$, in order to get the generating functions for factorial moments.

Theorem 7.5 *The expected number of ascendants $a_n = \mathbb{E}[A_n]$ of a random node in a random LBST of size n, when $n \geq 6$, is*

$$a_n = \frac{12}{7}\Big(1 + \frac{1}{n}\Big) H_n - \frac{1}{49}\Big(124 - \frac{9}{n}\Big).$$

Proof. Let, as usual,

$$\overline{\mathcal{A}}^{(s)}(z) = \left.\frac{\partial^s \overline{\mathsf{A}}_z}{\partial v^s}\right|_{v=1}.$$

To avoid cluttering the notation, we also let $\overline{\mathcal{A}}_z(z) = \overline{\mathcal{A}}^{(1)}(z)$. Here is the generating function for the expectations

$$\overline{\mathcal{A}}_z(z) = \frac{24}{7}\frac{1}{(1-z)^3}\log\frac{1}{1-z} + \frac{4}{49}(1-z)^{-3} + (1-z)^{-2} - \frac{4}{49}(1-z)^4.$$

Then we extract the $(n-1)^{\text{th}}$ coefficient and divide by n^2 to get the expected value of A_n; recall that since we are averaging w.r.t. j and $\overline{\mathsf{A}}_z$ is already a partial derivative w.r.t z, we have in fact

$$\mathbb{E}[A_n] = \frac{1}{n^2}[z^{n-1}]\overline{\mathcal{A}}_z(z).$$

□

Theorem 7.6 *The variance of the number of ascendants, A_n, of a random node in a random LBST with n nodes, or equivalently, the variance of the successful search cost for a random element in a LBST of size n is, when $n \geq 6$,*

$$\begin{aligned}\mathbb{V}[A_n] &= \frac{1}{343}\Big(300 + \frac{2100}{n} - \frac{216}{n^2}\Big)H_n \\ &\quad - \frac{144}{49}\Big(1+\frac{1}{n}\Big)\Big(\frac{H_n^2}{n} + H_n^{(2)}\Big) + \frac{1}{2401}\Big(10758 + \frac{2431}{n} - \frac{81}{n^2}\Big) + \frac{2304}{343n\, n^{\underline{6}}}.\end{aligned}$$

Proof. Analogously to what we did in the proof of the previous theorem, we compute the second derivative of $\overline{\mathsf{A}}_z(z, v)$, and let $v = 1$. Then

$$\begin{aligned}\overline{\mathcal{A}}^{(2)}(z) &= \frac{288}{49}\frac{1}{(1-z)^3}\log^2\frac{1}{1-z} + \Big(-\frac{480}{343}\frac{1}{(1-z)^3} + \frac{96}{343}(1-z)^4\Big)\log\frac{1}{1-z} \\ &\quad + \frac{9988}{2401}\frac{1}{(1-z)^3} - 4\frac{1}{(1-z)^2} - \frac{384}{2401}(1-z)^4.\end{aligned}$$

Extracting the coefficients is not as easy as before, but it is also doable, yielding the second factorial moment:

$$\begin{aligned}\mathbb{E}\big[A_n^{\underline{2}}\big] &= \frac{144}{49}(1+\frac{1}{n})\Big(H_n^2 - H_n^{(2)}\Big) - \frac{1}{343}\Big(3264 + \frac{1248}{n}\Big)H_n \\ &\quad + \frac{1}{2401}\Big(32210 - \frac{242}{n}\Big) + \frac{2304}{343n\, n^{\underline{6}}}\end{aligned}$$

From here, the remaining computations are just mechanical. □

For higher order moments, i.e. $s > 2$, the procedure applies but the computations get messier. If we do only consider the main order term in $a_n^{(s)} = \mathbb{E}\big[A_n^{\underline{s}}\big]$, then the result is much easier.

Theorem 7.7 *The s^{th} factorial moment of the number of ascendants, A_n, of a random node in a random LBST with n nodes, or equivalently, the s^{th} factorial moment of the successful search cost for a random element in a LBST of size n is, when $n \geq 6$,*

$$a_n^{(s)} = \left(\frac{12}{7}\right)^s \log^s n + \mathcal{O}(\log^{s-1} n).$$

Acknowledgements

We thank Philippe Flajolet for useful comments and suggestions. We also wish to thank the authors of the computer algebra system Maple who, although they might not know, greatly contributed to make this paper possible.

References

[1] G.M. Adel'son-Vel'skii and E.M. Landis. An algorithm for the organization of information. *Dokladi Akademia Nauk SSSR*, 146(2):263–266, 1962. English translation in Soviet Math. Doklay 3 ,1962, 1259-1263.

[2] C.R. Aragon and R.G. Seidel. Randomized search trees. In *Proc. of the 30th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 540–545, 1989.

[3] S.R. Arora and W.T. Dent. Randomized binary search technique. *Comm. ACM*, 12(2):77–80, 1969.

[4] C.J. Bell. *An Investigation into the Principles of the Classification and Analysis of Data on an Automatic Digital Computer.* PhD thesis, Leeds University, 1965.

[5] G.G. Brown and B.O. Shubert. On random binary trees. *Mathematics of Operations Research*, 9(1):43–65, 1984.

[6] Ph. Flajolet and A.M. Odlyzko. Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, 3(2):216–240, May 1990.

[7] Ph. Flajolet and M. Soria. General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, 114, 1993.

[8] G.H. Gonnet and R. Baeza-Yates. *Handbook of Algorithms and Data Structures - In Pascal and C.* Addison-Wesley, 2nd edition, 1991.

[9] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, 1989.

[10] L.J. Guibas and R. Sedgewick. A dichromatic framework for balanced trees. In *Proc. of the 19th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 8–21, 1978.

[11] C.A.R. Hoare. Find (Algorithm 65). *Comm. ACM*, 4:321–322, 1961.

[12] C.A.R. Hoare. Quicksort. *Comput. J.*, 5:10–15, 1962.

[13] M. Hoshi and Ph. Flajolet. Page usage in a quadtree index. *BIT*, 32(3):384–402, 1992.

[14] H.-K. Hwang. *Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques.* PhD thesis, Ecole Polytechnique, 1994.

[15] E. Kamke. *Differentialgleichungen: Lösungsmethoden und Lösungen.* Teubner, Stuttgart, 1977.

[16] Peter Kirschenhofer, Conrado Martínez, und Helmut Prodinger. Analysis of Hoare's Find Algorithm with Median-of-three partition. *Random Structures & Algorithms*, 10:143–156, 1997.

[17] P. Kirschenhofer and H. Prodinger. Comparisons in Hoare's Find algorithm. *Combinatorics, Probability and Computing*, 1996. To appear.

[18] D.E. Knuth. Mathematical analysis of algorithms. In *Proc. of the 1971 IFIP Congress*, pages 19–27, Amsterdam, 1972. North-Holland.

[19] D.E. Knuth. *The Art of Computer Programming: Sorting and Searching*, volume 3. Addison-Wesley, 1973.

[20] J. Lent. *Probabilistic analysis of some searching and sorting algorithms.* PhD thesis, George Washington University, 1996.

[21] G. Louchard. Exact and asymptotic distributions in digital and binary search trees. *Theoretical Informatics and Applications*, 21(4):479–496, 1987.

[22] W.C. Lynch. More combinatorial properties of certain trees. *Comput. J.*, 7:299–302, 1965.

[23] H.M. Mahmoud. *Evolution of Random Search Trees.* Wiley Interscience, 1992.

[24] A. Panholzer. *Untersuchungen zur durchschnittlichen Gestalt gewisser Baumfamilien. Mit besonderer Berücksichtigung von Anwendungen in der Informatik.* PhD thesis, Technische Universität Wien, 1997.

[25] P.V. Poblete. The analysis of heuristics for search trees. *Acta Informatica*, 30:233–248, 1993.

[26] P.V. Poblete and J.I. Munro. The analysis of a fringe heuristic for binary search trees. *J. Algorithms*, 6:336–350, 1985.

[27] S. Roura and C. Martínez. Randomization of search trees by subtree size. In J. Díaz and M. Serna, editors, *Proc. of the 4th European Symposium on Algorithms (ESA)*, volume 1136 of *LNCS*, pages 91–106. Springer, 1996.

[28] B. Salvy and P. Zimmermann. Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software*, 20(2):163–177, 1994.

[29] R. Sedgewick. *Quicksort.* Garland, New York, 1978.

[30] R. Sedgewick. *Algorithms in C.* Addison-Wesley, 3rd edition, 1997.

[31] R. Sedgewick and Ph. Flajolet. *An Introduction to the Analysis of Algorithms.* Addison-Wesley, 1996.

[32] J.S. Vitter and Ph. Flajolet. Average-case analysis of algorithms and data structures. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, chapter 9. North-Holland, 1990.

[33] A. Walker and D. Wood. Locally balanced binary trees. *Comput. J.*, 19(4):322–325, 1976.

[34] H. Wilf. *Generatingfunctionology.* Academic Press, 1990.

APPENDIX TO THE ARTICLE "ON THE NUMBER OF DESCENDANTS AND ASCENDANTS IN RANDOM SEARCH TREES" BY C. MARTÍNEZ, A. PANHOLZER, AND H. PRODINGER

APPENDIX A. POLYNOMIALS

We collect here various polynomials in two and three variables that appear as factors or terms in the explicit form of some the generating functions studied in this paper.

A.1. The polynomials $A_i(z, u, v)$

▷ $$A_0(z,u,v) = \frac{1}{70}\Big(3410v^7u^6z^6+7356u^4v^4z+900v^8u^8z^7-1316u^9v^7z^5-1148u^9v^8z^6-880v^7u^8z^7-
210u^9v^8z^4-5874v^7u^5z^5+4886v^7u^8z^4+4080v^4u^6z-2564v^8u^8z^6-4648v^5u^3z+140u^{10}v^8z^7+3822u^5v^8z^5-
2378u^3v^3z-680u^9v^7z^7+7356v^4u^5z-120vu^2z-880v^7u^7z^7+1388v^7u^9z^6-490v^9u^5z+1042v^4u^2z+
3484u^7v^7z^6+120u^5z^2+60zu^3-120z^2u^4-140u^6v^7-140u^4v^8+70u^5v^8+140u^6v^6+70v^2u^3-140u^5v^4-
140v^6u^4+140u^5v^3-140u^4v^4-140v^4u^3-140v^2u^4+60uv^2z+10780v^7u^6z^2-130v^8u^8z^8-3920u^6v^8z^2+
4080u^3v^4z-5874u^8v^7z^5-1378v^5u^2z-4648v^5u^6z-9220v^5u^5z-9220v^5u^4z+12362v^7u^5z^4+60v^7u^8z^8-
80v^8u^9z^8+130u^9v^7z^8-45u^{10}v^8z^8+3410u^8v^7z^6+3822u^8v^8z^5-8606v^7u^7z^5-70v^9u^9z^4-84v^8u^{10}z^6+
1512v^9u^7z^6+868v^9u^8z^6+6986v^8z^5u^6+1106u^9v^8z^5+868v^9u^6z^6-400v^9u^7z^7+900v^8u^7z^7-2564v^8u^6z^6-
2590v^9u^7z^5+70v^9u^8z^8-3560v^8u^7z^6-826v^9u^8z^5-2590v^9u^6z^5-826v^9u^5z^5-2800v^9u^5z^3-400v^9u^8z^7-
5v^9u^{10}z^8-100v^9u^9z^7+1806v^5u^6z^6-2370v^5u^5z^5-520u^3v^2z^2+210uv^6z^4-14104v^4u^5z^2+56u^{10}v^7z^5-
220v^5z^7u^6-1612u^4v^2z^2+110v^3u^2z^2-1612u^6v^2z^2-1246v^6u^9z^6+8400u^6v^5z^4+140u^5v^7+140v^3u^3-
140v^6u^3+140v^4u^2+70v^4u-140v^3u^2-140u^5v^5-140v^5u-140v^5u^3-140u^5v^6+140u^6v^4+140v^7u^3+
560v^5u^4-140v^3u^6+70v^2u^5+60u^6z+140v^6u^2-140v^7u^2+70uv^6+70v^8u^3-140v^5u^7+70u^7v^4+
42v^9u^9z^5-126v^8u^{10}z^5+14v^9u^{10}z^5+84v^9u^9z^6+30v^9u^9z^8-28v^9u^{10}z^6-4062v^3u^4z-4062v^3u^5z+
7732u^5v^6z+770u^6v^8z+7732v^6u^4z+11760v^7u^5z^2-4620v^7u^5z-140v^9u^7z^2-4620v^7u^4z+6986v^8u^7z^5+
1890v^8u^5z+1534v^2u^4z-8606v^7u^6z^5+1960v^9u^5z^2-2800v^9u^6z^3+9240v^8u^6z^3-1260v^8u^7z^2-
15232u^6v^7z^3+12362v^7u^7z^4-3038v^8u^8z^4-7000v^8u^5z^2+2170v^9u^5z^4+3780v^9u^6z^4-7490v^8u^5z^4+
70v^9u^6z-580u^2v^3z+360u^2v^2z-15232v^7u^5z^3-700v^9u^7z^3-180vzu^3+746v^2zu^3+122v^4zu-180v^3uz+
9240v^8u^5z^3+980v^8u^8z^3+210v^9u^8z^4+2170v^9u^7z^4+13888v^7u^6z^4-7490v^8u^7z^4-10724v^8u^6z^4+
840v^9u^6z^2-9464v^7u^7z^3-2520v^7u^6z+4252v^6u^6z-1148u^5v^8z^6+2800u^3v^7z^2+126v^7u^2z^5-3038u^4v^8z^4+
1106u^4v^8z^5+826u^2v^6z-2520v^7u^3z+10v^8u^5z^8+500u^6v^8z^7-680v^7u^6z^7+140v^8u^5z^7+20v^9u^5z^7-
28v^9u^4z^6-210uv^6z-1316v^7u^4z^5-100v^7u^5z^7+25v^7u^6z^8+112v^7u^4z^6-126v^8u^3z^5+56v^7u^3z^5+
42v^9u^4z^5+14v^9u^3z^5-5v^9u^6z^8-84v^8u^4z^6-84v^7u^3z^6-210v^7u^2z^4-210v^8u^3z^4+56v^8u^3z^6-28v^8u^2z^5-
45v^8u^6z^8+1388v^7u^5z^6-70v^9u^3z^4-180u^6vz-140u^2v^8z+70uv^7z+280u^2v^8z^2-70u^2v^7z+140z^4u^2v^8+
126v^5uz+140v^7u^2z^3-2378u^6v^3z+448v^7u^3z^4+746u^6v^2z-300u^5vz+4886v^7u^4z^4-100v^9u^6z^7+
84v^9u^5z^6+140v^9u^3z^3+210v^9u^4z^4+980v^8u^3z^3-1624v^7u^3z^3+1534u^5v^2z-300u^4vz-700v^9u^4z^3-
1260u^3v^8z^2-140u^3v^9z^2+4252v^6u^3z+20v^7u^4z^7+4340v^8u^4z^3-9464v^7u^4z^3+872v^5u^8z^4+176v^4u^8z^4+
7824v^5u^7z^4-100v^7u^{10}z^7+6554u^7v^6z^5+3998u^8v^6z^5+14940u^4v^5z^2+1520v^3u^6z^4+2246v^4u^7z^5-
9220u^4v^4z^2-46u^8v^4z^5+320v^4z^7u^7+1280v^6u^9z^5+380v^6u^9z^7-2804u^5v^3z^3+1806v^5u^8z^6+540u^5vz^2+
120v^6u^{10}z^7-526v^5u^9z^6-286v^4u^8z^6-140v^7uz^2+840u^4v^9z^2+770u^3v^8z-13520u^6v^5z^3+14940v^5u^6z^2+
252u^2v^6z^2+420v^6uz^2+7880v^4u^5z^3+9428v^6u^7z^3-1344u^9v^6z^3+412u^8v^5z^3+2940v^6u^8z^3-5014v^6u^8z^4-
2416v^4u^7z^4+1302v^6u^{10}z^4+2246v^4z^5u^6-1834v^6z^6u^6+460v^3u^5z^4-5200v^4u^6z^4-70v^7u^7z+2800v^7u^7z^2+
1890v^8u^4z+448v^7u^9z^4-210v^7u^{10}z^4-2416v^4u^5z^4-3920v^8u^4z^2+336v^6u^9z^4-1876v^5u^9z^4+1560v^5u^7z^6-
490v^9u^4z-8008v^5u^7z^3+2948u^7v^4z^3+540u^5z^3v^2+150u^4z^2v+70v^9u^3z-5964v^6u^3z^2-2804u^6v^3z^3+
7880u^6v^4z^3+4856v^3u^4z^2+4336v^5u^3z^2+6552u^5v^3z^2+252v^6u^8z^2+4336v^5u^7z^2-1378v^5u^7z-
12914v^6u^4z^2-12914v^6u^6z^2+17336v^5u^5z^2+1042u^7v^4z-3634u^7v^4z^2-580u^7v^3z+1658u^7v^3z^2+
360u^7v^2z+196u^8v^5z^2-19028u^5v^6z^2-9220u^6v^4z^2+4856u^6v^3z^2-2096u^5v^2z^2-4366v^5z^5u^6+
432v^6u^{10}z^6+3998v^6z^5u^5+1812v^5z^5u^4+500v^8u^9z^7+112u^{10}v^7z^6-220u^9v^5z^7+308v^4z^6u^5-100v^4z^7u^6+
872v^5u^4z^4-1246v^6z^6u^5+6554v^6z^5u^6+17116v^6u^6z^3+1280v^6z^5u^4+7824v^5u^5z^4-700v^4z^5u^4-
46v^4z^5u^5-286v^4z^6u^6-1344v^6u^{10}z^5-9772u^7v^6z^4-1834u^8v^6z^6-546v^6z^5u^2+460u^5v^5z^7-70v^7uz^4-$$

$952v^5u^4z^6-4366u^7v^5z^5+120v^6u^5z^7-2370u^8v^5z^5-1344u^3v^6z^5-300u^4v^6z^7+1302u^2v^6z^4+588u^3v^6z^6+$
$1148u^3v^5z^5+432u^4v^6z^6+10780v^7u^4z^2-13724u^6v^6z^4-2648u^7v^6z^6-526v^5z^6u^5-5964v^6u^7z^2+$
$1812u^9v^5z^5-34v^4z^2u^2-1396v^4u^7z^6+320v^4u^8z^7+4340v^8u^7z^3-3634v^4u^3z^2+196v^5u^2z^2+2940u^3v^6z^3+$
$412v^5u^3z^3-1344v^6u^2z^3-8008u^4v^5z^3-13520u^5v^5z^3-88v^4u^3z^3-856v^3u^4z^3+9428v^6u^4z^3-1624v^7u^8z^3+$
$17116v^6u^5z^3+1658u^3v^3z^2-500v^5u^8z^7+140v^7u^9z^3-420v^5z^2u+140v^7uz^3-280u^2v^8z^3-420v^6uz^3+$
$336u^3v^6z^4-1876v^5u^3z^4-5014v^6u^4z^4-910v^5z^4u^2+896v^5z^3u^2+2948v^4u^4z^3-9772v^6u^5z^4+826v^6u^7z+$
$980v^4u^3z^4+176v^4u^4z^4+126v^5u^8z+122u^8v^4z-180u^8v^3z+60u^8v^2z-420u^9v^5z^2+420v^6u^9z^2+$
$210v^6u^{11}z^4+70v^7u^8z-70v^7u^{11}z^4-420u^{10}v^6z^3-120u^7vz-280v^8u^9z^3+140v^9u^8z^3-28v^8u^{11}z^5+$
$20v^9u^{10}z^7+56v^8u^{11}z^6+25u^{10}v^7z^8+10u^{11}v^8z^8-40u^{11}v^8z^7+140u^{10}v^8z^4+70u^7v^6-5v^7u^4z^8+$
$20u^6z^7v^2+14v^7u^{12}z^5+20v^7u^{12}z^7+15v^6u^{12}z^8-5v^7u^{12}z^8-20v^4u^{12}z^7+28v^4u^{12}z^6+60v^5u^{12}z^7+$
$5v^4u^{12}z^8-15v^5u^{12}z^8-42v^6u^{12}z^5+84v^6u^{12}z^6-28v^7u^{12}z^6-60v^6u^{12}z^7+42v^5u^{12}z^5+140v^4u^9z^2+$
$420v^5u^{10}z^3-210v^5u^{11}z^4-14v^4u^{12}z^5-84v^5u^{12}z^6-140v^4u^{10}z^3+70v^4u^{11}z^4+40u^{11}v^3z^7-10v^3z^8u^{11}+$
$5v^2z^8u^{10}-56u^{11}v^3z^6+280u^9v^3z^3-140u^{10}v^3z^4+70z^4u^9v^2+28u^{10}z^6v^2-20u^{10}z^7v^2-14u^{10}z^5v^2+$
$28u^{11}v^3z^5-210v^6u^8z+240u^7vz^2-120u^8z^2v^2-140z^3u^8v^2+140v^7u^{10}z^3-140v^7u^9z^2-60u^5vz^3+$
$150u^6vz^2+240u^3vz^2-520u^7z^2v^2+98z^5u^8v^2+260z^3u^7v^2+70z^4u^8v^2+14z^5u^7v^2-84z^6u^8v^2-$
$210z^4u^7v^2+28u^9z^6v^2+20u^9z^7v^2-98u^9z^5v^2+238u^{10}v^3z^5-70u^8v^3z^4+154u^9v^3z^5-140u^{10}v^3z^6+$
$80u^8v^3z^3-350u^9v^3z^4-856u^7v^3z^3+540u^6v^2z^3+460u^7v^3z^4+20u^6v^2z^4+70u^8v^3z^5-28u^9v^3z^6-$
$550v^3z^5u^7+14v^2z^5u^6+28v^3z^6u^8+56v^2z^6u^7+20u^{10}v^3z^7+20v^3z^7u^9+5v^3z^8u^{10}-10v^2z^8u^9+$
$98v^2z^5u^5-84v^2z^6u^6-210v^2z^4u^5+10v^2z^8u^8-550v^3u^6z^5+392v^3u^7z^6+70v^3z^5u^5+28v^3z^6u^6-$
$80v^3u^8z^7-10v^3u^9z^8-80v^3z^7u^7+30v^3z^8u^8-88v^4u^8z^3+308v^4u^9z^6+392v^4u^{10}z^6-100v^4u^9z^7-$
$700v^4u^9z^5-644v^4u^{10}z^5-100v^4u^{10}z^7-700v^4u^9z^3+490v^4u^{10}z^4+980v^4u^9z^4+5v^4z^8u^{10}-34v^4u^8z^2+$
$1148v^5u^{10}z^5+896v^5u^9z^3-910v^5u^{10}z^4-952v^5u^{10}z^6-182v^4u^{11}z^5-364v^5u^{11}z^6-100v^4u^{11}z^7+$
$434v^5u^{11}z^5+196v^4u^{11}z^6+140v^5u^{11}z^7+20v^4u^{11}z^8+460v^5u^{10}z^7-20v^5u^{11}z^8-95v^5z^8u^{10}+20v^4z^8u^9+$
$520v^6u^8z^7-100v^4z^8u^8+120v^5u^9z^8-500v^5z^7u^7+20v^5z^8u^8-75v^6z^8u^{10}-546v^6u^{11}z^5+588v^6u^{11}z^6-$
$84v^7u^{11}z^6-300v^6u^{11}z^7+126v^7u^{11}z^5+20v^7u^{11}z^7+60v^6u^{11}z^8-20v^6z^8u^9+520v^6z^7u^7-80v^6z^8u^8+$
$14v^7uz^5-182v^4u^2z^5-644u^3v^4z^5-364v^5u^3z^6+196v^4u^3z^6+420uv^5z^3+490v^4u^2z^4-14v^4z^5u+70v^4uz^4-$
$84v^5u^2z^6-700u^2v^4z^3-140uv^4z^3+140uv^4z^2+42v^5z^5u-40v^8u^4z^7+140u^4v^5z^7-42v^6z^5u+434u^2v^5z^5+$
$84u^2v^6z^6+28v^4u^2z^6-20v^4u^3z^7+60v^5u^3z^7-120u^2z^2v^2-140u^3z^3v^2-14z^5u^3v^2+70z^4u^3v^2+28u^4z^6v^2+$
$28u^5z^6v^2-98u^4z^5v^2-20u^5z^7v^2+260u^4z^3v^2+70u^4z^4v^2+20v^7u^3z^7-28u^2v^7z^6-210uv^5z^4-350u^3v^3z^4+$
$154u^4v^3z^5-28u^5v^3z^6+80u^3v^3z^3-70u^4v^3z^4+280u^2v^3z^3-140u^2v^3z^4+238u^3v^3z^5-140u^4v^3z^6+$
$28v^3z^5u^2-56v^3z^6u^3+20u^6v^3z^7-10u^7v^3z^8+5u^6v^2z^8+20u^5v^3z^7+5u^6v^3z^8+40v^3z^7u^4-10v^3z^8u^5-$
$10u^7z^8v^2+392v^4u^4z^6-100v^4u^4z^7-100v^4u^5z^7+20v^4z^8u^5+120v^5u^7z^8+5v^4u^6z^8-95v^5u^6z^8-$
$20v^5z^8u^5+5v^4z^8u^4+20v^4u^7z^8-60v^6u^3z^7-15v^5u^4z^8+60v^6z^8u^5+380v^6u^6z^7+130v^7u^7z^8-75v^6u^6z^8-$
$20v^6u^7z^8+15v^6z^8u^4+30v^9u^7z^8-80v^8u^7z^8-120u^6z^2+110u^8z^2v^3-60z^3u^6v+280v^8u^8z^2-140v^8u^7z\Big)$

▷ $A_1(z,u,v)=\dfrac{6}{35}(1-z)^3\big(5z^3u^3+5z^2u^3-20u^2z^2+3zu^3-16u^2z+28uz+u^3-14-6u^2+14u\big)(2uv-u-1)u^2v^4$

▷ $A_2(z,u,v)=-\dfrac{6}{35}v^4(u-2v+1)(1-uz)^3\big(-5z^3u^3+20z^2u^3-5u^2z^2-28zu^3+16u^2z-3uz+14u^3-14u^2+6u-1\big)$

▷ $A_3(z,u,v)=\dfrac{1}{35}\Big(-672v^7u^6z^6-3444u^4v^4z+720v^8u^8z^7-600v^7u^8z^7+1848v^7u^5z^5-2520v^4u^6z-$
$1008v^8u^8z^6+1260v^5u^3z+504u^5v^8z^5-6888v^4u^5z-240v^7u^7z^7+2016u^7v^7z^6+630u^4v^8+4704u^5v^4+$
$5040v^6u^4-3564u^5v^3+5166u^4v^4-2394u^4v^3+1848v^4u^3+468v^2u^4-180v^8u^8z^8-252u^3v^4z-420u^8v^7z^5+$
$1512v^5u^6z+6552v^5u^5z+6048v^5u^4z-1680v^7u^5z^4+150v^7u^8z^8+840u^8v^7z^6+504u^8v^8z^5-3528v^7u^7z^5+$
$84v^{10}z^5u^7+1680v^9u^7z^6+336v^9u^8z^6+6552v^8z^5u^6-168v^{10}u^7z^6+2520v^9u^6z^6-720v^9u^7z^7+1080v^8u^7z^7-$
$2016v^8u^6z^6+840v^{10}u^6z^5-1848v^9u^7z^5+60v^9u^8z^8-3528v^8u^7z^6-168v^9u^8z^5-4620v^9u^6z^5-3864v^9u^5z^5-$

$$
5040v^9u^5z^3 - 240v^9u^8z^7 + 60u^6 - 420u^5v^7 - 378v^3u^3 + 3780v^6u^3 + 126v^4u^2 - 4032u^5v^5 - 3654v^5u^3 -
3150u^4v^7 + 2520u^5v^6 + 1050u^6v^4 - 630u^6v^5 - 2310v^7u^3 - 630v^5u^2 - 6300v^5u^4 + 1062v^2u^6 - 1332v^3u^6 +
1512v^2u^5 + 630u^2v^8 + 1260v^6u^2 - 1260v^7u^2 - 270u^5v - 420u^6v + 1260v^8u^3 + 126v^5u^7 - 252u^7v^4 - 162u^7v^2 +
756v^3u^4z + 4032v^3u^5z - 5040u^5v^6z - 5040v^6u^4z + 4200v^7u^5z^2 + 840v^7u^5z + 3360v^7u^4z + 4788v^8u^7z^5 -
1512v^7u^6z^5 + 1680v^9u^5z^2 - 1680v^9u^6z^3 + 5040v^8u^6z^3 - 4200u^6v^7z^3 + 2100v^7u^7z^4 - 5040v^8u^5z^2 +
6300v^9u^5z^4 + 4200v^9u^6z^4 - 5040v^8u^5z^4 - 1680v^7u^5z^3 + 7560v^8u^5z^3 + 840v^9u^7z^4 + 5040v^7u^6z^4 -
2520v^8u^7z^4 - 420v^{10}u^6z^4 - 8820v^8u^6z^4 + 1008u^5v^8z^6 + 2520u^4v^8z^4 - 2016u^4v^8z^5 + 2520v^7u^3z +
360v^7u^6z^7 - 360v^8u^5z^7 + 840v^{10}u^4z^5 - 168v^9u^3z^6 - 336v^9u^4z^6 + 252v^7u^4z^5 - 252v^8u^3z^5 - 168v^{10}u^4z^6 -
30v^9u^5z^8 + 84v^{10}u^3z^5 - 336v^9u^4z^5 + 672v^9u^3z^5 + 120v^9u^4z^7 + 84v^9u^2z^5 + 30v^9u^6z^8 + 504v^8u^4z^6 +
1260v^8u^3z^4 + 90v^8u^6z^8 - 504v^7u^5z^6 - 420v^9u^2z^4 - 840v^9u^3z^4 + 540u^6vz + 3096u^6v^3z - 2088u^6v^2z -
1260v^7u^4z^4 - 600v^9u^6z^7 + 1008v^9u^5z^6 - 420v^{10}z^4u^3 - 1680v^{10}u^4z^4 - 672v^{10}z^6u^5 + 840v^{10}u^3z^3 +
2520v^9u^4z^4 + 1512v^{10}u^5z^5 - 2520v^8u^3z^3 + 840v^9u^2z^3 - 840v^9u^2z^2 + 420v^9u^2z + 120v^{10}z^7u^5 - 936u^5v^2z -
4200v^9u^4z^3 + 2520u^3v^8z^2 + 840u^3v^9z^2 - 2520v^6u^3z + 2520v^7u^4z^3 + 3360u^4v^9z^2 - 2520u^3v^8z + 1260v^8u^4z -
2520v^8u^4z^2 - 840v^9u^4z - 420v^9u^3z - 252v^5u^7z + 420u^7v^4z - 432u^7v^3z - 36u^7v^2z - 2520v^7u^4z^2 +
342u^7v^3 - 210v^9u^3 - 30u^8 + 30u^7 + 120u^7v^{10}z^7 - 840u^4v^{10}z^2 + 420u^3v^{10}z + 840u^5v^{10}z^3 + 1680u^4v^{10}z^3 -
840u^3v^{10}z^2 + 360u^8v^2z + 300u^7vz - 1680u^5v^{10}z^4 - 252u^8v^3z - 240u^8vz - 210u^2v^9 - 672u^6v^{10}z^6 -
30u^6v^{10}z^8 - 30u^7v^{10}z^8 + 240u^6v^{10}z^7 + 84u^8v^4z - 30u^7v + 120u^8v - 90v^7u^7z^8 + 120v^9u^7z^8 - 90v^8u^7z^8 -
120u^7z + 60u^8z + 126u^8v^3 - 180u^8v^2 - 42u^8v^4\Big)
$$

$$
\triangleright \quad A_4(z,u,v) = \frac{1}{35}\Big(30 - 840v^7u^6z^6 + 6888u^4v^4z - 1080v^8u^8z^7 - 252u^9v^7z^5 - 1008u^9v^8z^6 +
240v^7u^8z^7 - 1260u^9v^8z^4 + 420v^7u^5z^5 + 1260v^7u^8z^4 + 252v^4u^6z + 2016v^8u^8z^6 - 1512v^5u^3z + 360u^{10}v^8z^7 -
504u^5v^8z^5 - 3096u^3v^3z - 360u^9v^7z^7 + 3444v^4u^5z - 300vu^2z + 600v^7u^7z^7 + 504v^7u^9z^6 + 840v^9u^5z -
420v^4u^2z - 2016u^7v^7z^6 + 420u^2v + 30uv - 60uz + 1260u^6v^7 - 630u^4v^8 - 630u^6v^8 - 1260u^5v^8 -
1260u^6v^6 + 210v^9u^5 - 1512v^2u^3 - 1848u^5v^4 - 5040v^6u^4 + 378u^5v^3 - 5166u^4v^4 + 2394u^4v^3 - 4704v^4u^3 -
468v^2u^4 - 360uv^2z + 240uvz + 2520v^7u^6z^2 + 180v^8u^8z^8 + 2520u^6v^8z^2 + 2520u^3v^4z - 1848u^8v^7z^5 +
252v^5u^2z - 1260v^5u^6z - 6048v^5u^5z - 6552v^5u^4z - 420v^9u^7z - 2100v^7u^5z^4 - 150v^7u^8z^8 + 90v^8u^9z^8 +
90u^9v^7z^8 - 90u^{10}v^8z^8 + 672u^8v^7z^6 - 504u^8v^8z^5 + 1512v^7u^7z^5 + 1680v^{10}u^8z^4 + 840v^9u^9z^4 - 840v^{10}z^5u^7 -
504v^8u^{10}z^6 - 1680v^9u^7z^6 - 2520v^9u^8z^6 - 4788v^8z^5u^6 + 2016u^9v^8z^5 - 240v^{10}u^9z^7 + 672v^{10}u^8z^6 +
168v^{10}u^7z^6 + 420v^{10}u^9z^4 - 840v^{10}u^9z^5 - 336v^9u^6z^6 + 240v^9u^7z^7 + 420v^9u^{10}z^4 - 720v^8u^7z^7 + 1008v^8u^6z^6 -
84v^{10}u^6z^5 + 4620v^9u^7z^5 - 60v^9u^8z^8 + 3528v^8u^7z^6 - 120v^{10}u^8z^7 + 3864v^9u^8z^5 - 840v^{10}u^8z^3 + 840v^{10}u^7z^2 -
420v^{10}u^6z + 1848v^9u^6z^5 + 168v^9u^5z^5 + 30v^{10}u^9z^8 - 1512v^{10}u^8z^5 + 840v^9u^8z^2 + 1680v^9u^5z^3 + 720v^9u^8z^7 -
840v^9u^9z^3 + 30v^9u^{11}z^8 - 30v^9u^{10}z^8 + 30v^{10}u^{10}z^8 + 600v^9u^9z^7 + 180v^2 - 120v + 42v^4 - 126v^3 + 120u^2z -
60u^2 - 30u + 2310u^5v^7 - 342v^3u + 3564v^3u^3 + 162v^2u - 1062u^2v^2 - 2520v^6u^3 + 210v^9u^6 + 270u^3v -
1050v^4u^2 + 252v^4u + 1332v^3u^2 + 3654u^5v^5 - 126v^5u + 4032v^5u^3 + 3150u^4v^7 - 3780u^5v^6 - 126u^6v^4 +
630u^6v^5 + 420v^7u^3 + 630v^5u^2 + 6300v^5u^4 + 672v^{10}u^9z^6 + 336v^9u^9z^5 - 84v^{10}u^{10}z^5 - 120v^9u^{11}z^7 +
252v^8u^{10}z^5 - 672v^9u^{10}z^5 + 168v^9u^{11}z^6 - 84v^9u^{11}z^5 - 1008v^9u^9z^6 - 120v^{10}u^{10}z^7 - 120v^9u^9z^8 + 168v^{10}u^{10}z^6 +
336v^9u^{10}z^6 - 4032v^3u^4z - 756v^3u^5z - 1680v^{10}u^7z^3 + 840v^{10}u^6z^2 + 5040u^5v^6z + 2520u^6v^8z + 5040v^6u^4z -
4200v^7u^5z^2 - 3360v^7u^5z - 840v^9u^7z^2 - 840v^7u^4z - 6552v^8u^7z^5 - 1260v^8u^5z + 936v^2u^4z + 3528v^7u^6z^5 -
1680v^9u^5z^2 - 840v^{10}u^6z^3 + 5040v^9u^6z^3 - 7560v^8u^6z^3 - 2520v^8u^7z^2 + 1680u^6v^7z^3 + 1680v^7u^7z^4 -
2520v^8u^8z^4 + 5040v^8u^5z^2 - 840v^9u^5z^4 - 4200v^9u^6z^4 + 2520v^8u^5z^4 + 420v^9u^6z + 432u^2v^3z + 36u^2v^2z +
4200v^7u^5z^3 + 4200v^9u^7z^3 - 540vzu^3 + 2088v^2zu^3 - 84v^4zu + 252v^3uz - 5040v^8u^5z^3 + 2520v^8u^8z^3 -
2520v^9u^8z^4 - 6300v^9u^7z^4 + 1680v^{10}u^7z^4 - 5040v^7u^6z^4 + 5040v^8u^7z^4 + 420v^{10}u^6z^4 + 8820v^8u^6z^4 -
3360v^9u^6z^2 - 2520v^7u^7z^3 - 2520v^7u^6z + 2520v^6u^6z\Big)
$$

A.2. The polynomials $B_i(z,u)$

$$
\triangleright \quad B_0(z,u) = 10uz - 120u^2z + 1140u^2z^2 - 8190u^5z^2 - 4236z^5u^6 - 1980z^3u^5 - 680z^4u^6 - 980u^4z -
605z^4u^4 - 4826z^5u^5 + 4354z^6u^6 - 5020z^3u^4 + 8585z^4u^5 + 500zu^3 + 2980z^3u^3 - 4140z^2u^3 + 8790z^2u^4 -
1470u^6 - 490u^2 + 1470u^3 - 2450u^4 + 70u + 980u^6z + 2450u^5 + 8190u^6z^2 + 1040u^9z^7 + 960u^{10}z^7 -
$$

$1040u^7z^7+518u^7z^6+490u^7-70u^8+4236u^8z^5+605u^9z^4-518u^8z^6-4354u^9z^6-55u^8z^8-8585u^8z^4+4826u^9z^5+112u^{10}z^6-500u^7z+55u^9z^8-390u^{10}z^8+490u^{10}z^4+5020u^8z^3-8790u^7z^2+4140u^8z^2-2980u^9z^3+150u^{11}z^8+680u^7z^4-1694u^{10}z^5-440u^{11}z^7+196u^{11}z^6+462u^{11}z^5+1980u^7z^3-112u^5z^6-140uz^2+14u^2z^5+1694u^4z^5+40u^5z^8-960u^6z^7+440u^5z^7-5u^4z^8-28u^2z^6+140u^3z^6+20u^3z^7-70uz^4+14z^5u-140u^4z^7-196u^4z^6-150u^6z^8-462u^3z^5+350u^2z^4-490u^3z^4+140uz^3-980u^2z^3-20u^{13}z^7+140u^{10}z^2-140u^{11}z^3+70u^{12}z^4+5u^{13}z^8+28u^{13}z^6-14u^{13}z^5-40u^{12}z^8+140u^{12}z^7-140u^{12}z^6-350u^{11}z^4-14u^{12}z^5-10u^9z+120u^8z+390u^7z^8-1140u^9z^2+980u^{10}z^3$

▷ $B_1(z,u) = -6(1+u)u^5(1-z)^3\Big(5z^3u^3+5z^2u^3-20u^2z^2+3u^3z-16u^2z+28uz+u^3-6u^2+14u-14\Big)$

▷ $B_2(z,u) = 6(1+u)\Big(5z^3u^3-20z^2u^3+5u^2z^2-16u^2z+3uz+28u^3z-6u-14u^3+14u^2+1\Big)(1-uz)^3$

A.3. The polynomials $C_i(z,u)$

▷ $C_0(z,u) = 950uz-6360u^2z+5120u^2z^2+10850u^5z^2-58440z^5u^6-39420z^3u^5+32210z^4u^6-22540u^4z+275z^4u^4-38218z^5u^5+38066z^6u^6-32540z^3u^4+69865z^4u^5+17260zu^3+17660z^3u^3-16280z^2u^3+29070z^2u^4+22540u^6z-10850u^6z^2+13960u^9z^7+7320u^{10}z^7-13960u^7z^7+21574u^7z^6+58440u^8z^5-275u^9z^4-21574u^8z^6-38066u^9z^6+1600u^8z^8-69865u^8z^4+38218u^9z^5-1540u^{10}z^6-17260u^7z-1600u^9z^8-2715u^{10}z^8+3710u^{10}z^4+32540u^8z^3-29070u^7z^2+16280u^8z^2-17660u^9z^3+885u^{11}z^8-32210u^7z^4-9352u^{10}z^5-2680u^{11}z^7+1484u^{11}z^6+2310u^{11}z^5+39420u^7z^3+1540u^5z^6-700uz^2+28u^2z^5+9352u^4z^5+215u^5z^8-7320u^6z^7+2680u^5z^7-25u^4z^8-140u^2z^6+784u^3z^6+100u^3z^7-350uz^4+70z^5u-760u^4z^7-1484u^4z^6-885u^6z^8-2310u^3z^5+1960u^2z^4-3710u^3z^4+700uz^3-5320u^2z^3-100u^{13}z^7+700u^{10}z^2-700u^{11}z^3+350u^{12}z^4+25u^{13}z^8+140u^{13}z^6-70u^{13}z^5-215u^{12}z^8+760u^{12}z^7-784u^{12}z^6-1960u^{11}z^4-28u^{12}z^5-950u^9z+6360u^8z+2715u^7z^8-5120u^9z^2+5320u^{10}z^3$

▷ $C_1(z,u) = 18480u^5z^2-36960u^6z^3-18480z^5u^6-18480z^3u^5+36960z^4u^6+3024u^4z-1848z^5u^5+3696z^6u^6+9240z^4u^5-504zu^3-1260u^6+252u^2-1260u^3+2268u^4+6384u^6z-16800u^5z+588u^5+18480u^6z^2-2640u^9z^7-2640u^7z^7+14784u^7z^6+1212u^7-600u^8+120u^9-18480u^8z^5+14784u^8z^6+3696u^9z^6+660u^8z^8+9240u^8z^4-1848u^9z^5-3864u^7z+660u^9z^8+36960u^7z^4-5280u^8z^7-18480u^7z^3-33264u^7z^5-240u^9z+1440u^8z$

▷ $C_2(z,u) = -120+240uz-1440u^2z-18480u^5z^2+36960u^6z^3+18480z^5u^6+18480z^3u^5-36960z^4u^6-6384u^4z+1848z^5u^5-3696z^6u^6-9240z^4u^5+3864zu^3+1260u^6-1212u^2+1260u^3-588u^4+600u-3024u^6z+16800u^5z-2268u^5-18480u^6z^2+2640u^9z^7+2640u^7z^7-14784u^7z^6-252u^7+18480u^8z^5-14784u^8z^6-3696u^9z^6-660u^8z^8-9240u^8z^4+1848u^9z^5+504u^7z-660u^9z^8-36960u^7z^4+5280u^8z^7+18480u^7z^3+33264u^7z^5$

A.4. The polynomials $D_i(z,u)$

▷ $D_0(z,u) = -\frac{u}{350}\Big(-1960uz^3-19000u^3z^3+19050u^3z^2-84uz^5+370z+840uz^4+4844u^3z^5+400u^3z^4+2520u^9z^4+1960u^9z^3+140u^9z^2-8900u^8z^3-1650u^8z^2+336u^2z^6-644u^2z^5-2520u^2z^4+8900u^2z^3-7650u^2z^2-28uz^6+2450u-1008u^3z^6+7650u^7z^2+1650uz^2+20u^2z^7-350-7350u^2+7350u^5-2450u^6-140z^2-70z^4+140z^3+14z^5-2340uz+5900u^2z-6860u^3z+6860u^5z-5900u^6z+2340u^7z-370u^8z-20790u^4z^2+5840u^4z^3+22630u^4z^4-15546u^4z^5+20790u^5z^2-17040u^5z^4-1528u^5z^5-19050u^6z^2-5840u^6z^3+17040u^6z^4-400u^8z^4+15546u^8z^5-4844u^9z^5+12250u^3-12250u^4+350u^7+19000u^7z^3-22630u^7z^4+1\Big)$

▷ $D_1(z,u) = \frac{-6u}{35}\Big(-7+42u-98u^2-98u^5+82u^6+14uz-98u^2z+294u^3z-630u^4z+434u^5z-238u^6z+74u^7z-10u^8z+280u^4z^2-280u^4z^3+140u^4z^4-28u^4z^5+280u^5z^2-560u^5z^3+560u^5z^4-280u^5z^5-280u^6z^3+560u^6z^4-504u^6z^5-28u^8z^5+98u^3+28u^4-32u^7+5u^8+140u^7z^4-280u^7z^5+224u^7z^6+224u^6z^6+56u^8z^6-80u^7z^7-40u^8z^7+10u^8z^8+56u^5z^6-40u^6z^7+10u^7z^8\Big)$

▷ $D_2(z,u) = -\frac{1}{35}\Big(30-192u+492u^2+588u^5-588u^6-60uz+444u^2z-1428u^3z+2604u^4z-3780u^5z+1764u^6z-588u^7z+84u^8z+1680u^5z^2-1680u^5z^3+840u^5z^4-168u^5z^5+1680u^6z^2-3360u^6z^3+3360u^6z^4-1680u^6z^5+840u^8z^4-1680u^8z^5-168u^9z^5-588u^3+168u^4+252u^7-42u^8-1680u^7z^3+3360u^7z^4-3024u^7z^5+1344u^7z^6+336u^6z^6+1344u^8z^6+336u^9z^6-240u^7z^7-240u^9z^7-480u^8z^7+60u^8z^8+60u^9z^8\Big)$

Appendix B. Tables

The next Tables collect the full statements of the theorems, considering the general cases as they appear in the main text of the paper as well as the special cases not listed there.

j	$d_{n,j}$
1	$\frac{6}{5}, \qquad \text{for } n \geq 5$
	$d_{1,1} = 1,\ d_{2,1} = \frac{3}{2},\ d_{3,1} = 1,\ d_{4,1} = \frac{5}{4}$
2	$\frac{18}{5} - \frac{3}{n}, \qquad \text{for } n \geq 6$
	$d_{2,2} = \frac{3}{2},\ d_{3,2} = 3,\ d_{4,2} = \frac{11}{4},\ d_{5,2} = 3$
3	$\frac{136}{35} - \frac{6}{n}, \qquad \text{for } n \geq 7$
	$d_{3,3} = 1,\ d_{4,3} = \frac{11}{4},\ d_{5,3} = \frac{13}{5},\ d_{6,3} = \frac{29}{10}$
4	$\frac{9}{2} - \frac{9}{n}, \qquad \text{for } n \geq 8$
	$d_{4,4} = \frac{5}{4},\ d_{5,4} = 3,\ d_{6,4} = \frac{29}{10},\ d_{7,4} = \frac{113}{35}$
$5 \leq j \leq n-4$	$-\frac{12}{7}H_n + \frac{12}{7}H_j + \frac{12}{7}H_{n+1-j}$ $-\frac{6}{7j} - \frac{6}{7(n+1-j)} + \frac{79}{70}$ $-\frac{3(3j-5)}{7n} + \frac{6(j-1)^2}{7n^{\underline{2}}} + \frac{2(2j-3)(j-1)^{\underline{2}}}{7n^{\underline{3}}}$ $+\frac{3(j-2)(j-1)^{\underline{3}}}{7n^{\underline{4}}} - \frac{3(2j-5)(j-1)^{\underline{4}}}{7n^{\underline{5}}} + \frac{2(j-3)(j-1)^{\underline{5}}}{7n^{\underline{6}}}$

Table 2: Average number of descendants of the j^{th} node in a random LBST of size n.

j	$d_{n,j}^{(2)}$
1	$\frac{2}{5}$ for $n \geq 5$
	$d_{1,1}^{(2)} = 0,\ d_{2,1}^{(2)} = 1,\ d_{3,1}^{(2)} = 0,\ d_{4,1}^{(2)} = \frac{1}{2}$
2	$\frac{36}{5}H_n - \frac{206}{25}$ for $n \geq 6$
	$d_{2,2}^{(2)} = 1,\ d_{3,2}^{(2)} = 6,\ d_{4,2}^{(2)} = \frac{13}{2},\ d_{5,2}^{(2)} = \frac{41}{5}$
3	$\frac{72}{5}H_n - \frac{4534}{175} + \frac{6}{n}$ for $n \geq 7$
	$d_{3,3}^{(2)} = 0,\ d_{4,3}^{(2)} = \frac{13}{2},\ d_{5,3}^{(2)} = 8,\ d_{6,3}^{(2)} = \frac{52}{5}$
4	$\frac{108}{5}H_n - \frac{7886}{175} - \frac{36}{5(n-1)} + \frac{126}{5n}$ for $n \geq 8$
	$d_{4,4}^{(2)} = \frac{1}{2},\ d_{5,4}^{(2)} = \frac{41}{5},\ d_{6,4}^{(2)} = \frac{52}{5},\ d_{7,4}^{(2)} = \frac{468}{35}$
$5 \leq j \leq n-4$	$\left(\frac{36n}{5} - \frac{12}{35}\right)H_n + \left(\frac{36j}{5} - \frac{36n}{5} - \frac{48}{7}\right)H_{n+1-j} + \left(\frac{12}{35} - \frac{36j}{5}\right)H_j - \frac{132}{35j} - \frac{132}{35(n+1-j)} + \frac{3489}{175} - \frac{33j}{5} + \left(\frac{66}{7} - \frac{429j}{35} + \frac{33j^2}{5}\right)\frac{1}{n} + \frac{132(j-1)^2}{35n^{\underline{2}}} + \frac{44(2j-3)(j-1)^{\underline{2}}}{35n^{\underline{3}}} + \frac{66(j-2)(j-1)^{\underline{3}}}{35n^{\underline{4}}} - \frac{66(2j-5)(j-1)^{\underline{4}}}{35n^{\underline{5}}} + \frac{44(j-3)(j-1)^{\underline{5}}}{35n^{\underline{6}}}$

Table 3: Second factorial moment of the number of descendants of the j^{th} node in a random LBST of size n.

m	$p_{n,m} = \mathbb{P}[D_n = m]$
1	$\frac{3}{7}\left(1+\frac{1}{n}\right)$ for $n \geq 6$
	$p_{1,1} = 1,\ p_{2,1} = \frac{1}{4},\ p_{3,1} = \frac{2}{9},\ p_{4,1} = \frac{1}{8},\ p_{5,1} = \frac{13}{125}$
2	$\frac{1}{7}\left(1+\frac{1}{n}\right),$ for $n \geq 6$
	$p_{2,2} = \frac{1}{4},\ p_{3,2} = 0,\ p_{4,2} = \frac{1}{16},\ p_{5,2} = \frac{4}{125}$
3	$\frac{3}{35}\left(1+\frac{1}{n}\right),$ for $n \geq 6$
	$p_{3,3} = \frac{1}{9},\ p_{4,3} = 0,\ p_{5,3} = \frac{3}{5}$
4	$\frac{2}{35}\left(1+\frac{1}{n}\right)$ for $n \geq 6$
	$p_{4,4} = \frac{1}{16},\ p_{5,4} = 0$
$5 \leq m < n$	$\frac{12}{7}\frac{(n+2+m)(n-1-m)}{n^2(m+2)(m+1)} - \frac{12}{7}\frac{m^{\underline{5}}}{n^2 n^{\underline{5}}} + \frac{12}{7n^2}$
n	$\frac{1}{n}$

Table 4: Probability that a random node in a random LBST of size n has m descendants.

j	$a_{n,j}$
1	$\frac{6}{5}H_n - \frac{6}{25}$, for $n \geq 5$
	$a_{1,1} = 1,\ a_{2,1} = \frac{3}{2},\ a_{3,1} = 2,\ a_{4,1} = \frac{9}{4}$
2	$\frac{6}{5}H_n - \frac{21}{25}$, for $n \geq 6$
	$a_{2,2} = \frac{3}{2},\ a_{3,2} = 1,\ a_{4,2} = \frac{7}{4},\ a_{5,2} = \frac{19}{10}$
3	$\frac{6}{5}H_n - \frac{87}{175} - \frac{6}{5}\frac{1}{n^{\underline{2}}}$, for $n \geq 7$
	$a_{3,3} = 2,\ a_{4,3} = \frac{7}{4},\ a_{5,3} = \frac{11}{5},\ a_{6,3} = \frac{12}{5}$
4	$\frac{6}{5}H_n - \frac{11}{25} - \frac{18}{5}\frac{1}{n^{\underline{2}}} - \frac{12}{5}\frac{1}{n^{\underline{3}}}$, for $n \geq 8$
	$a_{4,4} = \frac{9}{4},\ a_{5,4} = \frac{19}{10},\ a_{6,4} = \frac{12}{5},\ a_{7,4} = \frac{18}{7}$
$5 \leq j \leq n-4$	$\frac{24}{35}H_n + \frac{18}{35}H_j + \frac{18}{35}H_{n+1-j}$ $+\frac{12}{35j} + \frac{12}{35(n+1-j)} - \frac{279}{175} - \frac{6}{7n}$ $+\frac{18j}{35n} - \frac{12(j-1)^2}{35n^{\underline{2}}} - \frac{4(2j-3)(j-1)^{\underline{2}}}{35n^{\underline{3}}}$ $-\frac{6(j-2)(j-1)^{\underline{3}}}{35n^{\underline{4}}} + \frac{6(2j-5)(j-1)^{\underline{4}}}{35n^{\underline{5}}} - \frac{4(j-3)(j-1)^{\underline{5}}}{35n^{\underline{6}}}$

Table 5: Average number of ascendants of the j^{th} node in a random LBST of size n.

j	$a_{n,j}^{(2)}$
1	$\frac{36}{25}H_n^2 - \frac{144}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{766}{625} + \frac{3}{125}\frac{1}{n} - \frac{12}{125}\frac{1}{n-1} + \frac{18}{125}\frac{1}{n-2} - \frac{12}{125}\frac{1}{n-3} + \frac{3}{125}\frac{1}{n-4}$ for $n \geq 5$
	$a_{1,1}^{(2)} = 0\ ,\ a_{2,1}^{(2)} = 1\ ,\ a_{3,1}^{(2)} = 2\ ,\ a_{4,1}^{(2)} = 3$
2	$\frac{36}{25}H_n^2 - \frac{324}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{1726}{625} - \frac{27}{125}\frac{1}{n} + \frac{108}{125}\frac{1}{n-1} - \frac{162}{125}\frac{1}{n-2} + \frac{108}{125}\frac{1}{n-3} - \frac{27}{125}\frac{1}{n-4}$ for $n \geq 6$
	$a_{2,2}^{(2)} = 1\ ,\ a_{3,2}^{(2)} = 0\ ,\ a_{4,2}^{(2)} = 2\ ,\ a_{5,2}^{(2)} = \frac{11}{5}$
3	$\frac{36}{25}H_n^2 - \left(\frac{1548}{875} + \frac{72}{25}\frac{1}{n(n-1)}\right)H_n - \frac{36}{25}H_n^{(2)} + \frac{77094}{30625} + \frac{2169}{875}\frac{1}{n} - \frac{2892}{875}\frac{1}{n-1} + \frac{1476}{875}\frac{1}{n-2} - \frac{1044}{875}\frac{1}{n-3} + \frac{321}{875}\frac{1}{n-4} - \frac{36}{875}\frac{1}{n-5} + \frac{6}{875}\frac{1}{n-6} - \frac{72}{25}\frac{1}{n^2} + \frac{72}{25}\frac{1}{(n-1)^2}$ for $n \geq 7$
	$a_{3,3}^{(2)} = 2\ ,\ a_{4,3}^{(2)} = 2\ ,\ a_{5,3}^{(2)} = \frac{18}{5}\ ,\ a_{6,3}^{(2)} = \frac{22}{5}$
4	$\frac{36}{25}H_n^2 + \left(-\frac{204}{125} + \frac{144}{25}\frac{1}{n} - \frac{72}{25}\frac{1}{n-1} - \frac{72}{25}\frac{1}{n-2}\right)H_n - \frac{36}{25}H_n^{(2)} + \frac{11077}{4375} + \frac{5517}{875}\frac{1}{n} - \frac{786}{125}\frac{1}{n-1} - \frac{246}{125}\frac{1}{n-2} + \frac{378}{125}\frac{1}{n-3} - \frac{177}{125}\frac{1}{n-4} + \frac{54}{125}\frac{1}{n-5} - \frac{12}{125}\frac{1}{n-6} + \frac{6}{875}\frac{1}{n-7} - \frac{144}{25}\frac{1}{n^2} + \frac{72}{25}\frac{1}{(n-1)^2} + \frac{72}{25}\frac{1}{(n-2)^2}$ for $n \geq 8$,
	$a_{4,4}^{(2)} = 3\ ,\ a_{5,4}^{(2)} = \frac{11}{5}\ ,\ a_{6,4}^{(2)} = \frac{22}{5}\ ,\ a_{7,4}^{(2)} = \frac{36}{7}$

Table 6: Second factorial moment of the number of ascendants of the j^{th} node in a random LBST of size n.

j	$\mathbb{V}[A_{n,j}]$
1	$\frac{78}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{116}{125} + \frac{3}{125}\frac{1}{n} - \frac{12}{125}\frac{1}{n-1} + \frac{18}{125}\frac{1}{n-2} - \frac{12}{125}\frac{1}{n-3} + \frac{3}{125}\frac{1}{n-4}$ for $n \geq 5$
	$\mathbb{V}[A_{1,1}] = 0$, $\mathbb{V}[A_{2,1}] = \frac{1}{4}$, $\mathbb{V}[A_{3,1}] = 0$, $\mathbb{V}[A_{4,1}] = \frac{3}{16}$
2	$\frac{78}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{152}{125} - \frac{27}{125}\frac{1}{n} + \frac{108}{125}\frac{1}{n-1} - \frac{162}{125}\frac{1}{n-2} + \frac{108}{125}\frac{1}{n-3} - \frac{27}{125}\frac{1}{n-4}$ for $n \geq 6$
	$\mathbb{V}[A_{2,2}] = \frac{1}{4}$, $\mathbb{V}[A_{2,2}] = 0$, $\mathbb{V}[A_{4,2}] = \frac{11}{16}$, $\mathbb{V}[A_{5,2}] = \frac{49}{100}$
3	$\frac{78}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{2172}{1225} + \frac{249}{125}\frac{1}{n} - \frac{2466}{875}\frac{1}{n-1} + \frac{1476}{875}\frac{1}{n-2} - \frac{1044}{875}\frac{1}{n-3} + \frac{321}{875}\frac{1}{n-4} - \frac{36}{875}\frac{1}{n-5} + \frac{6}{875}\frac{1}{n-6} - \frac{108}{25}\frac{1}{n^2} + \frac{36}{25}\frac{1}{(n-1)^2}$ for $n \geq 7$
	$\mathbb{V}[A_{3,3}] = 0$, $\mathbb{V}[A_{4,3}] = \frac{11}{16}$, $\mathbb{V}[A_{5,3}] = \frac{24}{25}$, $\mathbb{V}[A_{6,3}] = \frac{26}{25}$
4	$\frac{78}{125}H_n - \frac{36}{25}H_n^{(2)} + \frac{1661}{875} + \frac{381}{175}\frac{1}{n} - \frac{12}{125}\frac{1}{n-1} - \frac{528}{125}\frac{1}{n-2} + \frac{378}{125}\frac{1}{n-3} - \frac{177}{125}\frac{1}{n-4} + \frac{54}{125}\frac{1}{n-5} - \frac{12}{125}\frac{1}{n-6} + \frac{6}{875}\frac{1}{n-7} - \frac{288}{25}\frac{1}{n^2} + \frac{36}{25}\frac{1}{(n-1)^2} + \frac{36}{25}\frac{1}{(n-2)^2}$ for $n \geq 8$
	$\mathbb{V}[A_{4,4}] = \frac{3}{16}$, $\mathbb{V}[A_{5,4}] = \frac{49}{100}$, $\mathbb{V}[A_{6,4}] = \frac{26}{25}$, $\mathbb{V}[A_{7,4}] = \frac{54}{49}$

Table 7: Variance of the number of ascendants of the j^{th} node in a random LBST of size n.