AVERAGE CASE-ANALYSIS OF PRIORITY TREES: A STRUCTURE FOR PRIORITY QUEUE ADMINISTRATION

ALOIS PANHOLZER AND HELMUT PRODINGER

Abstract. Priority trees (p-trees) are a certain variety of binary trees of size n constructed from permutations of the numbers $1, \ldots, n$. In this paper we analyse several parameters depending on n (the size) and j (a number between 1 and n), such as the length of the left path (connecting the root and the leftmost leaf), the height of node j (= distance from the root), the number of left edges on the path from the root to the node j, the number of descendants of node j, the number of key comparisons when inserting an element between j and $j+1$, the number of key comparisons when cutting the p-trees into two p-trees, the number of nodes with 0, 1 or 2 children. Methodologically, recursions are set up according to a fundamental decomposition of the family $\mathcal{A}$ of p-trees (using auxiliary quantities $\mathcal{B}$ and $\mathcal{C}$); using generating functions, they lead to systems of differential equations that can be solved explicitly with some efforts. The quantities of interest can then be identified as coefficients in the explicit forms of the generating functions.

1. Introduction

Priority trees (or "p-trees" for short) are a data structure to implement *priority queues*. There exist applications of priority queues e. g. in operating systems like job scheduling or resource management and in discrete event simulation models. Each element in a priority queue has a fixed associated key value which determines its priority. The lower the key value of an element in the queue is, the higher is its priority. Such a queue must support the two basic operations of inserting an element with an arbitrary given priority (`Insert`) and of removing the element with the highest priority (`Delete`).

A p-tree is either empty or it consists of a sequence of nodes with non-increasing keys, the so called "left path," such that to each node on the left path except the last one, there is associated a possibly empty p-tree, the "right subtree." If z is a node on the left path with key l and x being its left successor with key k, then all nodes y_i of the right path associated with z have key values s_i ranked between k and l, or more precisely $k \le s_i < l$. The element with the smallest key value (i. e. with the highest priority) is the terminal node of the left path and is called the "left leaf."

To insert a new element p into a p-tree T the following recursive algorithm `Insert` can be applied:

- If T is empty or the root of T has a key not greater than p, then let p be the new root and T its left subtree.
- Otherwise search down the left path of T for the first node x that has a key that is not greater than p.

This research was supported by the Austrian Research Society (FWF) under the project number P12599-MAT. Additional new results about priority trees may be found in [14].

Date: December 19, 1997.

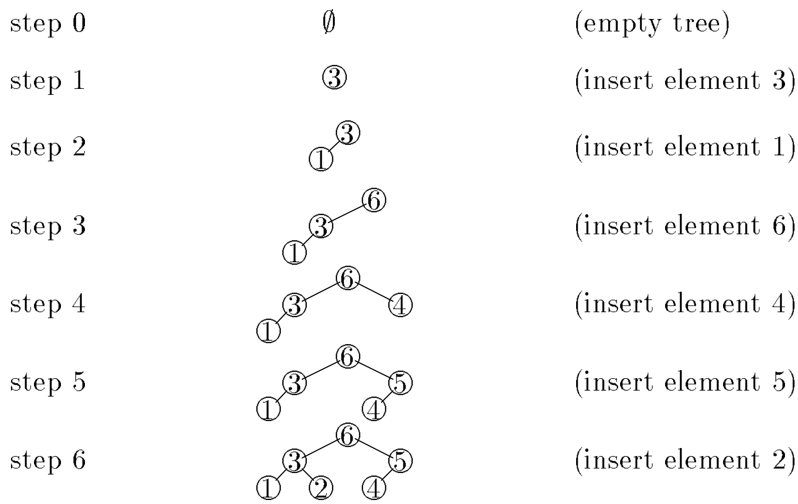

FIGURE 1. Generation of a p-tree.

- If no such node exists, then append p to the left path as the new left leaf.
- Otherwise denote the predecessor of x by z, so that the key of p is ranked between the keys of x and z. In this case the algorithm `Insert` (of node p) will be applied recursively to the right subtree of z.

As an example how the algorithm `Insert` works we show in Figure 1 the generation of the p-tree from the permutation 3 1 6 4 5 2.

2. THE MATHEMATICAL METHODS

For the mathematical analysis of this data structure we consider the model that all $n!$ permutations of the numbers $1, \ldots, n$ generating a p-tree of size n are equally likely.

We will not decompose the p-tree into the left path and the right subtrees and treat the components separately as it was done in [10], but rather work with 3 families of combinatorial objects $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. The family $\mathcal{A}$ are the ordinary p-trees, that are generated from random permutations by starting with the empty tree, the family $\mathcal{B}$ are p-trees that are generated by starting with one additional element "$+\infty$" and the family $\mathcal{C}$ are p-trees that are generated by starting with the extra element "$-\infty$."

For the following we assume that these additional elements are not counted for the size of an object, which means that the size of an object is the number of nodes, without the elements $-\infty$ and $+\infty$. Again all $n!$ permutations from 1 to n are assumed to be equally likely to generate an object of size n from the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$.

For the forthcoming analysis the following decompositions of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ w. r. t. the first element k of a random permutation of the numbers from $1, 2, \cdots, n$ are essential as is shown in Figures 2, 3 and 4.

The parameters of the p-trees that we are going to analyse in this paper satisfy recursions which are consequences of the decompositions of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. Using generating functions, these recurrences translate into a system of differential equations. Although in

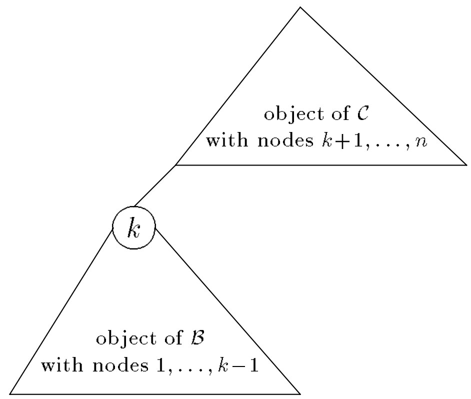

FIGURE 2. Decomposition of the family $\mathcal{A}$

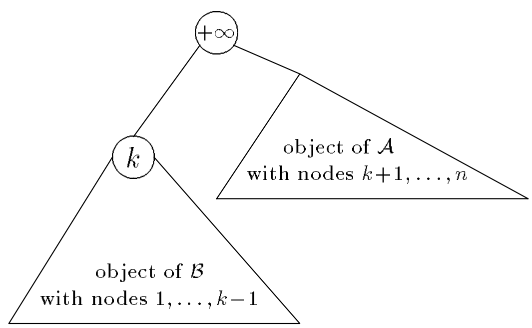

FIGURE 3. Decomposition of the family $\mathcal{B}$

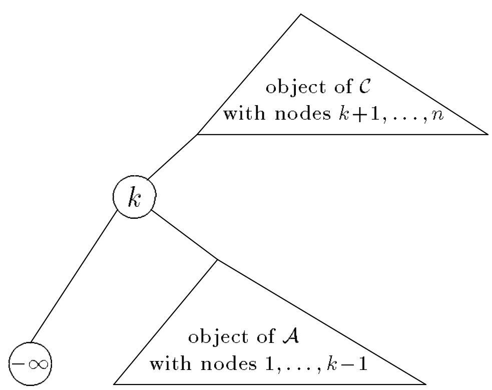

FIGURE 4. Decomposition of the family $\mathcal{C}$

most cases these differential equations for the distributions seem to be not tractable, we can at least compute the expectations for the considered parameters.

Here is an example

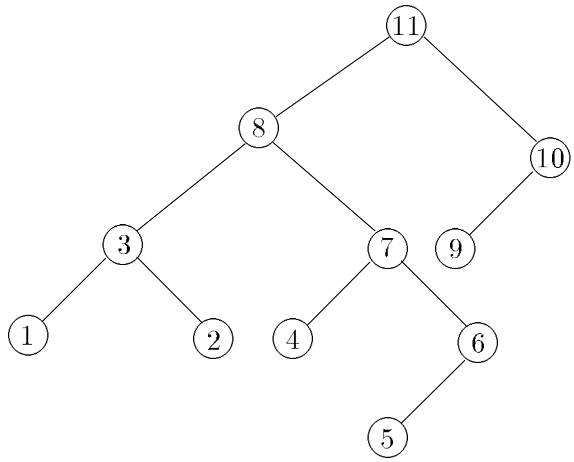

FIGURE 5. A p-tree constructed from the permutation 8 11 9 3 4 1 7 10 6 2 5. The lenght of the left path is 4, the height of node 7 is 3, the number of descendants of node 7 is 4, the number of left edges leading to node 4 is 2, the number of nodes with 0,1,2 successors is 5,2, 4. The number of comparisons to insert a node 8.5 is 4.

3. The length of the left path

In order to use the p-trees as a data structure for priority queue administration it is necessary to be able to remove on demand the element with the highest priority (i. e. with the smallest key) from the queue. In a p-tree this element, is the leaf of the left path starting at the root.

In the following we consider the distributions of the height of the element with highest priority in p-trees. Therefore we have to consider the distributions of the length of the left paths in in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. That means, we have to count the number of nodes on the direct path from the root to the left leaf. To obtain easier recurrences it is advantageous to count the element $+\infty$ for the length of the left path in the family $\mathcal{B}$ but not the element $-\infty$ in the family $\mathcal{C}$.

Decomposing the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the first element k of a random permutation with n elements as described above, we obtain the following recurrences for the probabilities $A_{n,m}$, $B_{n,m}$ and $C_{n,m}$, that an object of these families with size n has a left path of length m:

$$A_{n,m} = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{m} C_{n-k,i}\, B_{k-1,m-i} \quad \text{for } n \geq 1\,, \tag{1a}$$

$$B_{n,m} = \frac{1}{n} \sum_{k=1}^{n} B_{k-1,m-1} \quad \text{for } n \geq 1\,, \tag{1b}$$

$$C_{n,m} = \frac{1}{n} \sum_{k=1}^{n} C_{n-k,m-1} \quad \text{for } n \geq 1\,, \tag{1c}$$

$$A_{0,m} = \delta_{0,m}\,, \quad B_{0,m} = \delta_{1,m}\,, \ \text{ and } \ C_{0,m} = \delta_{0,m}\,. \tag{1d}$$

With the bivariate generating functions

$$A(z,v)=\sum_{n\ge 0}\sum_{m\ge 0}A_{n,m}z^nv^m,\ B(z,v)=\sum_{n\ge 0}\sum_{m\ge 0}B_{n,m}z^nv^m,\ \text{and}\ C(z,v)=\sum_{n\ge 0}\sum_{m\ge 0}C_{n,m}z^nv^m$$

we get from (1) by multiplying with $nz^{n-1}v^m$ and summing up over all $m\ge 0$ and $n\ge 1$ the following system of differential equations

$$\frac{\partial}{\partial z}A(z,v)=B(z,v)\,C(z,v)\,, \tag{2a}$$

$$\frac{\partial}{\partial z}B(z,v)=\frac{v}{1-z}B(z,v)\,, \tag{2b}$$

$$\frac{\partial}{\partial z}C(z,v)=\frac{v}{1-z}C(z,v)\,, \tag{2c}$$

with initial values $A(0,v)=1$, $B(0,v)=v$, and $C(0,v)=1$.

From (2b) we obtain

$$\frac{d\,(B(z,v))}{B(z,v)}=d\,(\log B(z,v))=\frac{v}{1-z}dz\,, \tag{3}$$

which leads through integration and adjusting to the initial values to the solution

$$B(z,v)=\frac{v}{(1-z)^v}\,. \tag{4}$$

Analogously we get from (2c) for $C(z,v)$ the solution

$$C(z,v)=\frac{1}{(1-z)^v}\,. \tag{5}$$

With (2a) these expressions for $B(z,v)$ and $C(z,v)$ lead to the following equations for $A(z,v)$

$$\frac{\partial}{\partial z}A(z,v)=\frac{v}{(1-z)^{2v}}\,, \tag{6}$$

from which we get by integration the solution

$$A(z,v)=\frac{1}{1-2v}\left(1-v-\frac{v}{(1-z)^{2v-1}}\right)\,. \tag{7}$$

In the following we denote the signless Stirling numbers of the first kind by $\left[{n \atop m}\right]$. They are given by the relation (see e. g. [4])

$$\sum_{n\ge 0}\sum_{m=0}^{n}\left[{n \atop m}\right]\frac{z^n}{n!}v^m=\frac{1}{(1-z)^v}\,. \tag{8}$$

With this equation the coefficients of $B(z,v)$ and $C(z,v)$ are obtained immediately whereas for the coefficients of $A(z,v)$ we get

$$\begin{aligned}A_{n,m}&=\frac{1}{n}[z^{n-1}v^m]\frac{\partial}{\partial z}A(z,v)=\frac{1}{n}[z^{n-1}v^m]\frac{v}{(1-z)^{2v}}\\&=\frac{2^{m-1}}{n}[z^{n-1}v^{m-1}]\frac{1}{(1-z)^v}=\frac{2^{m-1}\left[{n-1 \atop m-1}\right]}{n!}\,.\end{aligned}$$

This leads to the following

Theorem 3.1. *The probabilities $A_{n,m}$, $B_{n,m}$ and $C_{n,m}$ of the length of the left paths in objects of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ of size n, i. e. the probabilities, that an object of size n has a left path length m, are given by*

$$A_{n,m} = \frac{2^{m-1}\left[{n-1 \atop m-1}\right]}{n!} \quad \text{for } n \geq 1\,, \tag{9a}$$

$$B_{n,m} = \frac{\left[{n \atop m-1}\right]}{n!} \quad \text{for } n \geq 1\,, \tag{9b}$$

$$C_{n,m} = \frac{\left[{n \atop m}\right]}{n!} \quad \text{for } n \geq 1\,. \tag{9c}$$

To get the expectations A_n, B_n and C_n of the length of the left paths in objects of size n in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ we define

$$A(z) = \frac{\partial}{\partial v} A(z,v)\Big|_{v=1}\,, \quad B(z) = \frac{\partial}{\partial v} B(z,v)\Big|_{v=1} \text{ and } C(z) = \frac{\partial}{\partial v} C(z,v)\Big|_{v=1}\,. \tag{10}$$

Then these expectations are given by $A_n = [z^n]A(z)$, $B_n = [z^n]B(z)$ and $C_n = [z^n]C(z)$.

Differentiating the equations (4), (5) and (6) with respect to v and evaluating at $v = 1$ leads to

$$\frac{\partial}{\partial z} A(z) = \frac{\partial}{\partial v}\left(\frac{\partial}{\partial z} A(z,v)\right)\Big|_{v=1} = \frac{2}{(1-z)^2}\log\left(\frac{1}{1-z}\right) + \frac{1}{(1-z)^2}\,, \tag{11a}$$

$$B(z) = \frac{\partial}{\partial v} B(z,v)\Big|_{v=1} = \frac{1}{1-z} + \frac{1}{1-z}\log\left(\frac{1}{1-z}\right)\,, \tag{11b}$$

$$C(z) = \frac{\partial}{\partial v} C(z,v)\Big|_{v=1} = \frac{1}{1-z}\log\left(\frac{1}{1-z}\right)\,. \tag{11c}$$

To extract the coefficients of such expressions we need the formulas (see [9])

$$[z^n]\,\frac{1}{(1-z)^{\alpha+1}}\log\left(\frac{1}{1-z}\right) = (H_{n+\alpha} - H_\alpha)\binom{n+\alpha}{n}\,, \tag{12a}$$

$$[z^n]\,\frac{1}{(1-z)^{\alpha+1}}\log^2\left(\frac{1}{1-z}\right) = \left((H_{n+\alpha} - H_\alpha)^2 - \left(H_{n+\alpha}^{(2)} - H_\alpha^{(2)}\right)\right)\binom{n+\alpha}{n}\,, \tag{12b}$$

where $H_n = \sum_{k=1}^{n} \frac{1}{k}$ respectively $H_n^{(m)} = \sum_{k=1}^{n} \frac{1}{k^m}$ denote the *harmonic numbers.*

Then with (12a) we obtain from (11)

Theorem 3.2. *The expectations A_n, B_n and C_n for the lengths of the left paths in objects of size n in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are given by*

$$A_n = 2H_n - 1 \ \text{ for } \ n \geq 1\,,\ A_0 = 0\,, \tag{13a}$$

$$B_n = H_n + 1 \ \text{ for } \ n \geq 1\,,\ B_0 = 1\,, \tag{13b}$$

$$C_n = H_n \ \text{ for } \ n \geq 1\,,\ C_0 = 0\,. \tag{13c}$$

In the following we denote by $\breve{A}_n$ the second factorial moment of the length of the left path in p-trees of size n. Furthermore we define $\breve{A}(z) = \frac{\partial^2}{\partial v^2} A(z,v)\Big|_{v=1}$. Then the auxiliary quantity $\breve{A}_n$ is given by $\breve{A}_n = [z^n]\breve{A}(z)$.

Differentiating (6) two times with respect to v and evaluating at $v = 1$ leads to

$$\frac{\partial}{\partial z}\breve{A}(z) = \frac{4}{(1-z)^2}\log^2\left(\frac{1}{1-z}\right) + \frac{4}{(1-z)^2}\log\left(\frac{1}{1-z}\right) . \tag{14}$$

With (12) we get the coefficients $\breve{A}_n$ and so we have

Lemma 3.3. *The second factorial moment $\breve{A}_n$ of the length of the left path in p-trees of size n is given by*

$$\breve{A}_n = 4H_n^2 - 4H_n - 4H_n^{(2)} + 4 \ \text{ for } \ n \geq 1\,, \ \breve{A}_0 = 0\,. \tag{15}$$

The variance $\hat{A}_n$ of the length of the left path in p-trees of size n is given by the relation

$$\hat{A}_n = \breve{A}_n + A_n - A_n^2$$

and therefore we get

Theorem 3.4. *The variance $\hat{A}_n$ of the length of the left path in p-trees of size n is given by*

$$\hat{A}_n = 2H_n - 4H_n^{(2)} + 2 \ \text{ for } \ n \geq 1\,, \ \hat{A}_0 = 0\,. \tag{16}$$

The asymptotic behaviour of the lengths of the left paths in p-trees is described in the following theorem.

Theorem 3.5. *The sequence of random variables*

$$\Omega_n^* = \frac{\Omega_n - \mu_n}{\sigma_n}$$

converges weakly to the standard normal (Gaussian) distribution. Here the random variables Ω_n denote the distributions of the lengths of the left paths in p-trees of size n with expectation μ_n and variance σ_n^2. That means we have

$$\mathbb{P}\left(a < \Omega_n^* < b\right) \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{t^2}{2}}\,dt \tag{17}$$

and further

$$\mu_n = 2\log n + \mathcal{O}(1) \ \text{ and } \ \sigma_n^2 = 2\log n + \mathcal{O}(1)$$

for $n \to \infty$.

Proof. With the asymptotic expansions of the harmonic numbers

$$H_n = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \mathcal{O}\left(\frac{1}{n^4}\right) , \tag{18a}$$

$$H_n^{(2)} = \frac{\pi^2}{6} - \frac{1}{n} + \frac{1}{2n^2} + \mathcal{O}\left(\frac{1}{n^3}\right) \tag{18b}$$

we get from the exact formulas (13a) and (16) immediately

$$\mu_n = A_n = 2\log n + \mathcal{O}(1) \ \text{ and } \ \sigma_n^2 = \hat{A}_n = 2\log n + \mathcal{O}(1)\,.$$

To prove the theorem we use a method that is described in [6]. We start with the sequence of the characteristic functions $\phi_{\Omega_n^*}(t) = \mathbb{E}(e^{i\Omega_n^* t})$ of the Ω_n^* and show, that this sequence converges pointwise to the characteristic function $e^{-\frac{t^2}{2}}$ of the standard normal distribution. In other words, we show, that

$$\phi_{\Omega_n^*}(t) \to e^{-\frac{t^2}{2}} \tag{19}$$

for every $t \in \mathbb{R}$ and $n \to \infty$. By the continuity theorem for characteristic functions of Paul Lévy (see e. g. [3]), we can then deduce from (19) the weak convergence of the distribution functions (17).

From the bivariate generating function $A(z, v)$ of the probabilities $A_{n,m}$, that the length of the left path in a p-tree of size n is equal m as given by (7), we obtain with

$$A_n(v) = [z^n]A(z, v)$$

the probability generating function of the Ω_n. First we obtain

$$A_n(v) = \frac{1}{n}[z^{n-1}]\frac{\partial}{\partial z}A(z, v) = \frac{1}{n}[z^{n-1}]\frac{v}{(1-z)^{2v}} = \frac{v}{2v-1}\binom{n+2v-2}{n}.$$

With the asymptotic expansion for fixed α

$$\binom{n+\alpha}{n} = \frac{(n+\alpha)\cdot(n+\alpha-1)\cdots(n+1)}{\Gamma(\alpha+1)} = \frac{n^\alpha}{\Gamma(\alpha+1)}\cdot\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right),$$

we get

$$A_n(v) = \frac{v}{2v-1}\binom{n+2v-2}{n} = \frac{v\, n^{2v-2}}{\Gamma(2v)}\cdot\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right) \tag{20}$$

for fixed v and $n \to \infty$.

For the characteristic functions $\phi_{\Omega_n^*}(t)$ of the Ω_n^* we have then

$$\begin{aligned}\phi_{\Omega_n^*}(t) &= \mathbb{E}(e^{i\Omega_n^* t}) = \sum_m \mathbb{P}(\Omega_n^* = m)\cdot e^{imt} = \sum_m \mathbb{P}(\Omega_n = \mu_n + \sigma_n m)\cdot e^{imt} \\ &= \sum_{m\ge 0}\mathbb{P}(\Omega_n = m)\cdot e^{it\frac{m-\mu_n}{\sigma_n}} = \sum_{m\ge 0} A_{n,m}e^{it\frac{m-\mu_n}{\sigma_n}} = e^{-\frac{i\mu_n t}{\sigma_n}}A_n(e^{\frac{it}{\sigma_n}}).\end{aligned}$$

With (20) we get the expansion

$$\begin{aligned}A_n(e^{\frac{it}{\sigma_n}}) &= \frac{e^{i\frac{t}{\sigma_n}}}{\Gamma(2e^{i\frac{t}{\sigma_n}})}\exp\left(2\log n(e^{i\frac{t}{\sigma_n}}-1)\right)\cdot\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right) \\ &= \exp\left(2\log n(e^{i\frac{t}{\sigma_n}}-1)\right)\cdot(1+o(1)) = \exp\left(2\log n\frac{it}{\sigma_n}-\frac{t^2}{2\sigma_n^2}\right)\cdot(1+o(1)) \\ &= \exp\left(2\log n\frac{it}{\sigma_n}-\frac{t^2}{2}\right)\cdot(1+o(1)).\end{aligned}$$

This leads to

$$\phi_{\Omega_n^*}(t) = e^{-\frac{i\mu_n t}{\sigma_n}}A_n(e^{\frac{it}{\sigma_n}}) = \exp\left(-2\log n\frac{it}{\sigma_n}\right)\exp\left(2\log n\frac{it}{\sigma_n}-\frac{t^2}{2}\right)\cdot(1+o(1))$$

$$= e^{-\frac{t^2}{2}} \cdot (1 + o(1)) \to e^{-\frac{t^2}{2}}$$

for a fixed t and $n \to \infty$. With the continuity theorem of Lévy the theorem is completely proved. □

4. The height of the nodes

In the following we consider the distributions of the height of the nodes with key j of objects of size n in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. To be precise, we say that the height of a node j is the number of nodes between the root and the node j.

With $A_{n,j,m}$, $B_{n,j,m}$ and $C_{n,j,m}$ we denote the probabilities, that the node with key j in an corresponding object of size n has height m. To get a recurrence for these quantities, we use the already treated distribution of the length of the left path in objects of the family $\mathcal{C}$. To avoid confusion, from now on the probability, that the length of the left path in an object of size n in the family $\mathcal{C}$ is equal m, is denoted by $p_{n,m}$. With (5) and (9c) we have for $p_{n,m}$, respectively its generating function $p(z,v) = \sum_{n\geq 0}\sum_{m\geq 0} p_{n,m} z^n v^m$,

$$p_{n,m} = \frac{\left[{n \atop m}\right]}{n!} \quad \text{and} \quad p(z,v) = \frac{1}{(1-z)^v}\,. \tag{21}$$

With these remarks we get by means of decomposing the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the first element k of a random permutation following recurrences for $1 \leq j \leq n$

$$A_{n,j,m} = \frac{1}{n}\left(\sum_{k=1}^{j-1} C_{n-k,j-k,m} + p_{n-j,m-1} + \sum_{k=j+1}^{n}\sum_{i=0}^{m} p_{n-k,i}\, B_{k-1,j,m-i}\right), \tag{22a}$$

$$B_{n,j,m} = \frac{1}{n}\left(\sum_{k=1}^{j-1} A_{n-k,j-k,m-1} + \delta_{m,2} + \sum_{k=j+1}^{n} B_{k-1,j,m-1}\right), \tag{22b}$$

$$C_{n,j,m} = \frac{1}{n}\left(\sum_{k=1}^{j-1} C_{n-k,j-k,m} + p_{n-j,m-1} + \sum_{k=j+1}^{n}\sum_{i=0}^{m-1} p_{n-k,i}\, A_{k-1,j,m-1-i}\right). \tag{22c}$$

Introducing the trivariate generating functions

$$A(z,u,v) = \sum_{j\geq 1}\sum_{n\geq j}\sum_{m\geq 0} A_{n,j,m} z^n u^j v^m\,,\ B(z,u,v) = \sum_{j\geq 1}\sum_{n\geq j}\sum_{m\geq 0} B_{n,j,m} z^n u^j v^m \ \text{ and}$$

$$C(z,u,v) = \sum_{j\geq 1}\sum_{n\geq j}\sum_{m\geq 0} C_{n,j,m} z^n u^j v^m\,,$$

we get from (22) by multiplying with $nz^{n-1}u^j v^m$ and summing up over all $n \geq j \geq 1$ and $m \geq 0$ the following system of linear differential equations

$$\frac{\partial}{\partial z}A(z,u,v) = \frac{u}{1-uz}C(z,u,v) + \frac{vu}{1-uz}p(z,v) + p(z,v)\,B(z,u,v)\,, \tag{23a}$$

$$\frac{\partial}{\partial z}B(z,u,v) = \frac{vu}{1-uz}A(z,u,v) + \frac{v^2u}{(1-z)(1-uz)} + \frac{v}{1-z}B(z,u,v)\,, \tag{23b}$$

$$\frac{\partial}{\partial z} C(z,u,v) = \frac{u}{1-uz} C(z,u,v) + \frac{vu}{1-uz} p(z,v) + v\, p(z,v)\, A(z,u,v)\,. \tag{23c}$$

This system of differential equations seems to be not tractable, but it is helpful to get factorial moments of the distribution of the height of nodes in p-trees, in particular the *expectation*.

Then the ordinary generating functions

$$A(z,u) = \sum_{n\geq 1} \sum_{j=1}^{n} A_{n,j} z^n u^j\,, \quad B(z,u) = \sum_{n\geq 1} \sum_{j=1}^{n} B_{n,j} z^n u^j\,, \quad C(z,u) = \sum_{n\geq 1} \sum_{j=1}^{n} C_{n,j} z^n u^j$$

of the expectations $A_{n,j}$, $B_{n,j}$ and $C_{n,j}$ of the heights of the node j in objects of size n in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ fulfill the relations

$$A(z,u) = \frac{\partial}{\partial v} A(z,u,v)\Big|_{v=1}\,, \quad B(z,u) = \frac{\partial}{\partial v} B(z,u,v)\Big|_{v=1}\,, \quad C(z,u) = \frac{\partial}{\partial v} C(z,u,v)\Big|_{v=1}\,.$$

Differentiating (23) with respect to v and evaluating at $v=1$ leads to the following system of linear differential equation for $A(z,u)$, $B(z,u)$ and $C(z,u)$

$$\frac{\partial}{\partial z} A(z,u) = \frac{u}{1-uz} C(z,u) + \frac{1}{1-z} B(z,u) + f_1(z,u)\,, \tag{24a}$$

$$\frac{\partial}{\partial z} B(z,u) = \frac{u}{1-uz} A(z,u) + \frac{1}{1-z} B(z,u) + f_2(z,u)\,, \tag{24b}$$

$$\frac{\partial}{\partial z} C(z,u) = \frac{u}{1-uz} C(z,u) + \frac{1}{1-z} A(z,u) + f_3(z,u) \tag{24c}$$

with

$$\begin{aligned} f_1(z,u) &= \frac{u}{(1-z)(1-uz)} + \frac{u}{(1-z)(1-uz)} \log\left(\frac{1}{1-z}\right) \\ &\quad + \frac{uz}{(1-z)^2(1-uz)} \log\left(\frac{1}{1-z}\right)\,, \\ f_2(z,u) &= \frac{u^2 z}{(1-z)(1-uz)^2} + \frac{2u}{(1-z)(1-uz)} + \frac{uz}{(1-z)^2(1-uz)}\,, \\ f_3(z,u) &= \frac{u}{(1-z)(1-uz)} + \frac{u}{(1-z)(1-uz)} \log\left(\frac{1}{1-z}\right) + \frac{uz}{(1-z)^2(1-uz)} \\ &\quad + \frac{uz}{(1-z)^2(1-uz)} \log\left(\frac{1}{1-z}\right)\,. \end{aligned}$$

In the following this system of differential equations will be converted to a single differential equation for $A(z,u)$. To do this, we introduce two functions $\widetilde{B}(z,u)$ and $\widetilde{C}(z,u)$ which are given by the relations $B(z,u) = \frac{\widetilde{B}(z,u)}{1-z}$ and $C(z,u) = \frac{\widetilde{C}(z,u)}{1-uz}$.

Then we get from (24)

$$\frac{\partial}{\partial z} A(z,u) = \frac{u}{(1-uz)^2} \widetilde{C}(z,u) + \frac{1}{(1-z)^2} \widetilde{B}(z,u) + f_1(z,u)\,, \tag{25a}$$

$$\frac{1}{1-z}\frac{\partial}{\partial z}\widetilde{B}(z,u) = \frac{u}{1-uz}A(z,u) + f_2(z,u)\,, \tag{25b}$$

$$\frac{1}{1-uz}\frac{\partial}{\partial z}\widetilde{C}(z,u) = \frac{1}{1-z}A(z,u) + f_3(z,u)\,. \tag{25c}$$

To eliminate $\widetilde{B}(z,u)$, we extract it from (25a), differentiate with respect to z and multiply by $\frac{1}{1-z}$. This leads to

$$\widetilde{B}(z,u) = (1-z)^2\frac{\partial}{\partial z}A(z,u) - \frac{u(1-z)^2}{(1-uz)^2}\widetilde{C}(z,u) - (1-z)^2 f_1(z,u)\,, \tag{26a}$$

$$\begin{aligned}\frac{1}{1-z}\frac{\partial}{\partial z}\widetilde{B}(z,u) = {} & -2\frac{\partial}{\partial z}A(z,u) + (1-z)\frac{\partial^2}{\partial z^2}A(z,u) - \frac{u(1-z)}{(1-uz)^2}\frac{\partial}{\partial z}\widetilde{C}(z,u) \\ & - \frac{2u(u-1)}{(1-uz)^3}\widetilde{C}(z,u) - \frac{1}{1-z}\frac{\partial}{\partial z}\left((1-z)^2 f_1(z,u)\right)\,.\end{aligned} \tag{26b}$$

Inserting (26b) in (25) leads to the following system of differential equations for $A(z,u)$ and $\widetilde{C}(z,u)$

$$\begin{aligned}\frac{u}{1-uz}A(z,u) + 2\frac{\partial}{\partial z}A(z,u) - (1-z)\frac{\partial^2}{\partial z^2}A(z,u) + \frac{u(1-z)}{(1-uz)^2}\frac{\partial}{\partial z}\widetilde{C}(z,u) \\ + \frac{2u(u-1)}{(1-uz)^3}\widetilde{C}(z,u) = -f_2(z,u) - \frac{1}{1-z}\frac{\partial}{\partial z}\left((1-z)^2 f_1(z,u)\right)\,,\end{aligned} \tag{27a}$$

$$\frac{1}{1-uz}\frac{\partial}{\partial z}\widetilde{C}(z,u) = \frac{1}{1-z}A(z,u) + f_3(z,u)\,. \tag{27b}$$

To eliminate also $\widetilde{C}(z,u)$, we insert equation (27b) in (27a) and solve for $\widetilde{C}(z,u)$. This leads to

$$\begin{aligned}\widetilde{C}(z,u) = {} & -\frac{(1-uz)^2}{u-1}A(z,u) - \frac{(1-uz)^3}{u(u-1)}\frac{\partial}{\partial z}A(z,u) + \frac{(1-z)(1-uz)^3}{2u(u-1)}\frac{\partial^2}{\partial z^2}A(z,u) \\ & - \frac{(1-uz)^3}{2u(u-1)}f_2(z,u) - \frac{(1-uz)^3}{2u(u-1)(1-z)}\frac{\partial}{\partial z}\left((1-z)^2 f_1(z,u)\right) \\ & - \frac{(1-uz)^2(1-z)}{2(u-1)}f_3(z,u)\,.\end{aligned} \tag{28}$$

Differentiating this equation with respect to z and multiplying by $\frac{1}{1-uz}$ gives a differential equation for $A(z,u)$ after insertion in (27a).

$$\begin{aligned}& \frac{2uz-u-1}{(1-z)(1-u)}A(z,u) - 2\frac{1-uz}{1-u}\frac{\partial}{\partial z}A(z,u) - \\ & - \frac{3}{2}\frac{(1-uz)(2uz-u-1)}{u(1-u)}\frac{\partial^2}{\partial z^2}A(z,u) - \frac{1}{2}\frac{(1-uz)^2(1-z)}{u(1-u)}\frac{\partial^3}{\partial z^3}A(z,u) \\ & = f_3(z,u) + \frac{1}{1-uz}\frac{\partial}{\partial z}\left[\frac{(1-uz)^3}{2u(u-1)}f_2(z,u) + \right. \\ & \left. + \frac{(1-uz)^3}{2u(u-1)(1-z)}\frac{\partial}{\partial z}\left((1-z)^2 f_1(z,u)\right) + \frac{(1-uz)^2(1-z)}{2(u-1)}f_3(z,u)\right]\,.\end{aligned} \tag{29}$$

Resubstituting $f_1(z,u)$, $f_2(z,u)$ and $f_3(z,u)$ and taking the initial values $A_{0,0}=0$, $A_{1,1}=1$, $A_{2,1}=2$ and $A_{2,2}=1$ into account leads to the following linear differential equation of order 3 for $A(z,u)$ with the given initial conditions:

$$\begin{aligned}
&\frac{2uz-u-1}{(1-z)(1-u)}A(z,u) - 2\frac{1-uz}{1-u}\frac{\partial}{\partial z}A(z,u) - \frac{3}{2}\frac{(1-uz)(2uz-u-1)}{u(1-u)}\frac{\partial^2}{\partial z^2}A(z,u) \\
&\quad - \frac{1}{2}\frac{(1-uz)^2(1-z)}{u(1-u)}\frac{\partial^3}{\partial z^3}A(z,u) = \frac{uz+u-2}{(1-z)^3(1-u)}, \qquad (30)\\
&A(0,u)=0, \quad \frac{\partial}{\partial z}A(z,u)\bigg|_{z=0} = u, \quad \frac{\partial^2}{\partial z^2}A(z,u)\bigg|_{z=0} = 2\left(2u+u^2\right).
\end{aligned}$$

Now we have to find 3 linearly independent solutions of the corresponding homogeneous differential equation

$$\begin{aligned}
&\frac{2uz-u-1}{(1-z)(1-u)}A_h(z,u) - 2\frac{1-uz}{1-u}\frac{\partial}{\partial z}A_h(z,u) - \\
&\quad - \frac{3}{2}\frac{(1-uz)(2uz-u-1)}{u(1-u)}\frac{\partial^2}{\partial z^2}A_h(z,u) - \frac{1}{2}\frac{(1-uz)^2(1-z)}{u(1-u)}\frac{\partial^3}{\partial z^3}A_h(z,u) = 0,
\end{aligned} \qquad (31)$$

then we can apply the *variation of the parameter method* [11] in order to solve (30).

From the homogeneous system of linear differential equations corrresponding to system (24)

$$\frac{\partial}{\partial z}A_h(z,u) = \frac{u}{1-uz}C_h(z,u) + \frac{1}{1-z}B_h(z,u), \qquad (32a)$$

$$\frac{\partial}{\partial z}B_h(z,u) = \frac{u}{1-uz}A_h(z,u) + \frac{1}{1-z}B_h(z,u), \qquad (32b)$$

$$\frac{\partial}{\partial z}C_h(z,u) = \frac{u}{1-uz}C_h(z,u) + \frac{1}{1-z}A_h(z,u) \qquad (32c)$$

it is easy to see that there exists a one dimensional solution space with $A_h(z,u) = B_h(z,u) = C_h(z,u)$. The corresponding differential equation

$$\frac{\partial}{\partial z}A_h(z,u) = \frac{u}{1-uz}A_h(z,u) + \frac{1}{1-z}A_h(z,u) \qquad (33)$$

has the solution

$$A_h(z,u) = \frac{k_1(u)}{(1-z)(1-uz)}. \qquad (34)$$

To reduce the order of the differential equation (31), we substitute

$$A(z,u) = \frac{E(z,u)}{(1-z)(1-uz)} \quad \text{and} \quad F(z,u) = \frac{\partial}{\partial z}E(z,u)$$

with functions $E(z,u)$ and $F(z,u)$ (method of d'Alembert [11]). This leads to the linear differential equation of order 2

$$\frac{1}{(1-z)(1-u)}F(z,u) - \frac{1}{2}\frac{1-uz}{(1-u)u}\frac{\partial^2}{\partial z^2}F(z,u) = 0. \qquad (35)$$

Using the substitution $z = 1 + \frac{1-u}{u}t$ and the notation $\widetilde{F}(t,u) = F(1+\frac{1-u}{u}t, u)$, the differential equation (12) transforms into the *hypergeometric equation*

$$t(1-t)\frac{\partial^2}{\partial t^2}F(t,u) + 2F(t,u) = 0\,. \tag{36}$$

Now we recall [2], that for all $c \neq 1, 2, \cdots$ a one dimensional solution space of the hypergeometric differential equation

$$t(1-t)f''(t) + (c - (1+a+b)t)f'(t) - abf(t) = 0 \tag{37}$$

is in a neighborhood of $t = 0$ given by

$$k(t)t^{1-c}\,{}_2F_1\left(\begin{matrix} a-c+1, b-c+1 \\ 2-c \end{matrix}\middle| t\right)\,, \tag{38}$$

where ${}_2F_1$ denotes a *hypergeometric function* [8].

In our instance the parameters are $c = 0$, $a = -2$ and $b = 1$, thus the above hypergeometric function degenerates to a polynomial, and we get one solution of (36) as

$$\widetilde{F}(t,u) = t - t^2\,. \tag{39}$$

Resubstituting leads to a one dimensional solution space

$$F(z,u) = k_2(u)(1-z)(1-uz) \tag{40}$$

of (35).

Introducing new functions $G(z,u)$ and $H(z,u)$ by

$$F(z,u) = (1-z)(1-uz)G(z,u) \text{ and } H(z,u) = \frac{\partial}{\partial z}G(z,u)\,, \tag{41}$$

we get from (35) a linear differential equation of order one:

$$-\frac{(2uz-u-1)(1-uz)}{(1-u)u}H(z,u) - \frac{1}{2}\frac{(1-z)(1-uz)^2}{(1-u)u}\frac{\partial}{\partial z}H(z,u) = 0\,. \tag{42}$$

This equation has the solution

$$H(z,u) = \frac{k_3(u)}{(1-z)^2(1-uz)^2}\,, \tag{43}$$

with a function $k_3(u)$, depending only on u.

Adding the three linearly independent solutions (34), (40) and (43) we get by resubstituting the complete solution of the homogeneous differential equation (31)

$$\begin{aligned} A_h(z,u) = {} & \frac{c_1(u)}{(1-z)(1-uz)} + \frac{c_3(u)}{(1-z)(1-uz)}\int_0^z (1-t)(1-ut)dt \\ & + \frac{c_3(u)}{(1-z)(1-uz)}\int_0^z \left[(1-x)(1-ux)\int_0^x \frac{1}{(1-t)^2(1-ut)^2}\,dt\right]dx \end{aligned} \tag{44}$$

or, integrated,

$$\begin{aligned}A_h(z,u) = \frac{c_1(u)}{(1-z)(1-uz)} + c_2(u)\frac{z(6-3uz-3z+2uz^2)}{(1-z)(1-uz)} \\ + c_3(u)\Bigg[-\frac{6(1-uz)(2uz-3u+1)}{1-z}\log\left(\frac{1}{1-uz}\right) \\ + \frac{6(1-z)(2uz-3+u)u^2}{1-uz}\log\left(\frac{1}{1-z}\right) \\ - \frac{(u-1)(12u^2z^2-12u(u+1)z-5u^2+4u-5)}{(1-z)(1-uz)}\Bigg]\,.\end{aligned} \tag{45}$$

Now the variation of the parameter method leads after insertion of the initial conditions to the following solution of the generating function $A(z,u)$ of the expectations of the height of node j in a p-tree of size n:

$$\begin{aligned}A(z,u) = \frac{1}{3}(2u^2z-u^2-u)\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt \\ + \frac{1}{3}\frac{(1-u)^2}{1-z}\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt \\ + \frac{1}{3}\frac{u}{(1-z)(1-uz)}\log^2\left(\frac{1}{1-z}\right) \\ + \frac{1}{3}\frac{u(2u^2z-u^2+2uz-2u+1)}{(1-u)^3}\log\left(\frac{1}{1-z}\right) \\ + \frac{7}{3}\frac{u}{(1-u)(1-z)}\log\left(\frac{1}{1-z}\right) - \frac{1}{3}\frac{1}{(1-z)^2}\log\left(\frac{1}{1-z}\right) \\ - \frac{1}{3}\frac{(6u-1)u}{(1-u)(1-uz)}\log\left(\frac{1}{1-z}\right) \\ - \frac{1}{3}\frac{u(u+1)(2uz-u-1)}{(1-u)^3}\log\left(\frac{1}{1-uz}\right) - \frac{1}{3}\frac{u+1}{(1-u)(1-z)}\log\left(\frac{1}{1-uz}\right) \\ - \frac{4}{3}\frac{u^2z}{(1-u)^2} + \frac{u}{(1-u)(1-uz)} - \frac{1}{3}\frac{2u+1}{(1-u)(1-z)} + \frac{1}{3}\frac{1}{(1-z)^2}\,.\end{aligned} \tag{46}$$

To extract the coefficients from equation (46), we use apart from the formulas (12) also

$$[z^nu^j]\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt = \frac{(j+1)(n-j)}{n}(H_{n-j}-1)\,, \tag{47a}$$

$$\begin{aligned}[z^nu^j]\,\frac{1}{1-z}\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt \\ = (j+1)\left((n-j+1)H_{n-j}-2(n-j)\right) - j(j+1)\left(\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} - (H_n-H_j)\right)\,.\end{aligned} \tag{47b}$$

Theorem 4.1. *The expectation $A_{n,j}$ of the height of the node with key j in a p-tree of size n is for $1 \le j \le n$ given by*

$$\begin{aligned} A_{n,j} = &-\frac{2}{3}\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} + \frac{1}{3}(H_n - H_j) + \frac{-\frac{2}{3}l(l+1)H_l + \frac{4}{3}l^2}{n-1} + \frac{\frac{2}{3}(l+1)^2 H_l - \frac{4}{3}l^2 - l}{n} \\ &+ \frac{1}{3}(H_{l+1}^2 - H_{l+1}^{(2)}) + \frac{5}{3}H_{l+1} - 1 + \frac{1}{3(l+1)} \quad \text{with } l := n-j\,. \end{aligned} \tag{48}$$

Now we evaluate our findings asymptotically, under the assumptions j fixed; $j \sim \rho n$, with $0 < \rho < 1$; $n - j$ fixed. In order to handle the sum $\sum_{k=j+1}^{n}\frac{H_{k-j}}{k}$ asymptotically we state the following relations

$$\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} + \sum_{k=1}^{j-1}\frac{H_{n-k}}{k} = \frac{1}{2}\left(H_n^2 - H_n^{(2)}\right) - \frac{1}{2}\left(H_{j-1}^2 + H_{j-1}^{(2)}\right) + H_n H_{j-1}\,, \tag{49a}$$

$$\sum_{k=1}^{j}\frac{H_{n-k}}{k} - \sum_{k=1}^{j}\frac{H_k}{n-k} = H_j H_{n-j-1}\,, \tag{49b}$$

which can be proved by means of generating functions.

For a fixed ratio $\rho = \frac{j}{n}$ with $0 < \rho < 1$ we use for $n \to \infty$ the expansion

$$\sum_{k=1}^{j}\frac{H_{n-k}}{k} = \log^2 n + (\log\rho + 2\gamma)\log n + \gamma^2 + \gamma\log\rho - \operatorname{dilog}(1-\rho) + o(1)\,, \tag{50}$$

which can be proved by means of Euler's summation formula after applying the asymptotics for the harmonic numbers (it can also be found in [12]). We recall the definition of the *dilogarithm* $\operatorname{dilog} z$ (see e. g. [1])

$$\operatorname{dilog} z = \int_1^z \frac{\log t}{1-t}\,dt\,. \tag{51}$$

Combining (18), (49) and (50) we get the following lemma.

Lemma 4.2. *The sum $\sum_{k=j+1}^{n}\frac{H_{k-j}}{k}$ has the following asymptotic equivalents:*

$$\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} = \frac{1}{2}\log^2 n + \gamma\log n - \frac{\pi^2}{12} + \frac{\gamma^2}{2} - \frac{H_j^2}{2} + \frac{H_j}{j} - \frac{H_j^{(2)}}{2} + \mathcal{O}\left(\frac{1}{n}\right) \tag{52a}$$

for fixed j,

$$\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} = \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } l = n-j\,, \tag{52b}$$

$$\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} = -\log\rho\log n + \mathcal{O}(1) \;\text{ for } j = \rho n \text{ and } 0 < \rho < 1\,. \tag{52c}$$

With the relations (18) and (52) we get from (48) the following corollary.

Corollary 4.3. *The expectation $A_{n,j}$ of the height of the node with key j in a p-tree of size n is for $n \to \infty$ asymptotically given by*

$$A_{n,j} = 2\log n + \mathcal{O}(1) \ \text{ for fixed } j\,, \tag{53a}$$

$$A_{n,j} = \frac{1}{3}H_{l+1}^2 + \frac{5}{3}H_{l+1} - \frac{1}{3}H_{l+1}^{(2)} - 1 + \frac{1}{3(l+1)} + \mathcal{O}\left(\frac{1}{n}\right) \ \text{ for fixed } l = n-j\,, \tag{53b}$$

$$A_{n,j} = \frac{1}{3}\log^2 n + \left(-\frac{2}{3}\rho^2 + \frac{2}{3}\rho + \frac{2}{3}\gamma + \frac{5}{3} + \frac{2}{3}\log\rho + \frac{2}{3}\log(1-\rho)\right)\log n + \mathcal{O}(1) \tag{53c}$$
$$\text{for } j = \rho n \text{ and } 0 < \rho < 1\,,$$

Also of interest are the expectations A_n of the height of a random node in a p-tree of size n. Because of

$$A_n = \frac{1}{n}\sum_{j=1}^{n} A_{n,j}$$

these parameters are easily obtained by summing up (48). With basic summation formulas for the harmonic numbers and the relation

$$\sum_{j=1}^{n}\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} = \sum_{k=1}^{n}\frac{1}{k}\sum_{j=1}^{k-1}H_{k-j} = \sum_{k=1}^{n}\frac{1}{k}(kH_k - k) = \sum_{k=1}^{n}H_k - n = (n+1)H_n - 2n\,, \tag{54}$$

we get

Corollary 4.4. *The expectation A_n of the height of a randomly chosen node in a p-tree of size n is given by*

$$A_n = \frac{n+1}{3n}H_n^2 + \frac{2(2n+5)}{9n}H_n - \frac{n+1}{3n}H_n^{(2)} - \frac{13n+1}{27n} \ \text{ for } n \geq 2 \ \text{ and } \ A_1 = 1 \tag{55}$$

or asymptotically by

$$A_n = \frac{1}{3}\log^2 n + \left(\frac{2}{3}\gamma + \frac{4}{9}\right)\log n + \mathcal{O}(1) \ \text{ for } n \to \infty\,. \tag{56}$$

5. The number of left edges to a node

As opposed to the previous section we consider here the *edges* from the root to a specified node, but only the leftsided edges. In this way we get an impression of the *leftist* shape of the tree.

Here we denote by $A_{n,j}^{[L]}$, $B_{n,j}^{[L]}$ and $C_{n,j}^{[L]}$ the expectations of the number of leftsided edges from the root to the node with key j in objects of size n in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$.

To obtain recurrences for these parameters, we use the already computed expectation of the length of the left path of objects in the family $\mathcal{C}$ of size n. As was shown in section 3, this expectation is H_n.

Decomposing the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the first element k of a random permutation, we get for $1 \leq j \leq n$ the following recurrences

$$A_{n,j}^{[L]} = \frac{1}{n}\left(\sum_{k=1}^{j-1} C_{n-k,j-k}^{[L]} + H_{n-j} + \sum_{k=j+1}^{n}\left(B_{k-1,j}^{[L]} + H_{n-k}\right)\right)\,, \tag{57a}$$

$$B_{n,j}^{[L]} = \frac{1}{n}\left(\sum_{k=1}^{j-1} A_{n-k,j-k}^{[L]} + 1 + \sum_{k=j+1}^{n}\left(B_{k-1,j}^{[L]} + 1\right)\right), \tag{57b}$$

$$C_{n,j}^{[L]} = \frac{1}{n}\left(\sum_{k=1}^{j-1} C_{n-k,j-k}^{[L]} + H_{n-j} + \sum_{k=j+1}^{n}\left(A_{k-1,j}^{[L]} + H_{n-k}\right)\right), \tag{57c}$$

with initial values $A_{0,0}^{[L]} = 0$, $B_{0,0}^{[L]} = 0$ and $C_{0,0}^{[L]} = 0$.

With the bivariate generating functions $A^{[L]}(z,u) = \sum_{j\ge 1}\sum_{n\ge j} A_{n,j}^{[L]} z^n u^j$ etc. the above recurrences (multiplied by $nz^{n-1}u^j$ and summed up over all $n \ge j \ge 1$) lead to the following system of differential equations

$$\frac{\partial}{\partial z}A^{[L]}(z,u) = \frac{u}{1-uz}C^{[L]}(z,u) + \frac{1}{1-z}B^{[L]}(z,u) + f_1(z,u), \tag{58a}$$

$$\frac{\partial}{\partial z}B^{[L]}(z,u) = \frac{u}{1-uz}A^{[L]}(z,u) + \frac{1}{1-z}B^{[L]}(z,u) + f_2(z,u), \tag{58b}$$

$$\frac{\partial}{\partial z}C^{[L]}(z,u) = \frac{u}{1-uz}C^{[L]}(z,u) + \frac{1}{1-z}A^{[L]}(z,u) + f_3(z,u), \tag{58c}$$

with

$$f_1(z,u) = \frac{u}{(1-z)(1-uz)}\log\left(\frac{1}{1-z}\right) + \frac{uz}{(1-z)^2(1-uz)}\log\left(\frac{1}{1-z}\right),$$

$$f_2(z,u) = \frac{u}{(1-z)(1-uz)} + \frac{uz}{(1-z)^2(1-uz)},$$

$$f_3(z,u) = \frac{u}{(1-z)(1-uz)}\log\left(\frac{1}{1-z}\right) + \frac{uz}{(1-z)^2(1-uz)}\log\left(\frac{1}{1-z}\right).$$

This system of differential equations can be solved analogously to (24). This leads to the bivariate generating function $A^{[L]}(z,u)$ of the expectations $A_{n,j}^{[L]}$ of the number of leftsided edges from the root to the node j:

$$\begin{aligned} A^{[L]}(z,u) = {} & \left(\frac{1}{3}u(2uz-u-1) + \frac{1}{3}\frac{(u-1)^2}{1-z}\right)\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt \\ & + \left(-\frac{1}{3}\frac{u^2}{(1-uz)(1-u)} + \frac{1}{3}\frac{u}{(1-z)(1-u)}\right)\log^2\left(\frac{1}{1-z}\right) \\ & + \left[\frac{2}{3}\frac{(u+1)(2uz-u-1)u}{(1-u)^3} + \frac{7}{3}\frac{u}{(1-z)(1-u)} - \frac{1}{3}\frac{1}{(1-z)^2} - \right. \\ & \qquad \left. - \frac{1}{3}\frac{(5u-2)u}{(1-uz)(1-u)}\right]\log\left(\frac{1}{1-z}\right) \\ & + \left(-\frac{2}{3}\frac{(u+1)(2uz-u-1)u}{(1-u)^3} - \frac{2}{3}\frac{u+1}{(1-z)(1-u)}\right)\log\left(\frac{1}{1-uz}\right) \\ & + \frac{8}{3}\frac{u^2 z}{(1-u)^2} - \frac{1}{3}\frac{4u+1}{(1-z)(1-u)} + \frac{1}{3}\frac{1}{(1-z)^2} + \frac{5}{3}\frac{u}{(1-uz)(1-u)}. \end{aligned} \tag{59}$$

Extracting the coefficients using (12) and (47) we obtain

Theorem 5.1. *The expectation $A_{n,j}^{[L]}$ of the number of leftsided edges from the root to the node with key j in a p-tree of size n is for $1 \le j \le n$ given by*

$$\begin{aligned} A_{n,j}^{[L]} &= -\frac{2}{3}\sum_{k=j+1}^{n}\frac{H_{k-j}}{k} + \frac{1}{3}H_{n-j}^2 + H_n - H_j - \frac{1}{3}H_{n-j}^{(2)} \\ &+ \left(1 + \frac{2}{3}\frac{(j-1)^2}{n} - \frac{2}{3}\frac{(j-1)(j-2)}{n-1} + \frac{2}{3}\frac{1}{n-j+1}\right)H_{n-j} - \frac{1}{3}\frac{(3j-1)(2j-1)}{n} \\ &+ 2\frac{(j-1)^2}{n-1} + \frac{5}{3}\frac{1}{n-j+1} + \frac{1}{3j} - \frac{4}{3}. \end{aligned} \tag{60}$$

The expectation $A_{n,j}^{[R]}$ of the number of rightsided edges from the root to the node j in a p-tree of size n is easy to obtain using the relation

$$A_{n,j}^{[R]} = A_{n,j} - 1 - A_{n,j}^{[L]}. \tag{61}$$

Theorem 5.2. *The expectation $A_{n,j}^{[R]}$ of the number of rightsided edges from the root to the node with key j in a p-tree of size n is for $n \ge j \ge 1$ given by*

$$A_{n,j}^{[R]} = \frac{2}{3}H_{n-j} - \frac{2}{3}H_n + \frac{2}{3}H_j + \frac{1}{3}\frac{2j^2-2j+1}{n} - \frac{2}{3}\frac{(j-1)^2}{n-1} + \frac{1}{3}\frac{1}{n-j+1} - \frac{1}{3j} - \frac{1}{3}. \tag{62}$$

With the relations (18) and (52) we get the following asymptotic corollaries.

Corollary 5.3. *The expectation $A_{n,j}^{[L]}$ of the number of leftsided edges from the root to the node with key j in a p-tree of size n is for $n \to \infty$ asymptotically given by*

$$A_{n,j}^{[L]} = 2\log n + \mathcal{O}(1) \quad \text{for fixed } j\,, \tag{63a}$$

$$A_{n,j}^{[L]} = \frac{1}{3}H_l^2 + H_l - \frac{5}{3} - \frac{1}{3}H_l^{(2)} + \frac{2}{3}\frac{H_l}{l+1} + \frac{5}{3}\frac{1}{l+1} + \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } l = n-j\,, \tag{63b}$$

$$A_{n,j}^{[L]} = \frac{1}{3}\log^2 n + \left(-\frac{2}{3}\rho^2 + \frac{2}{3}\rho + \frac{2}{3}\gamma + 1 + \frac{2}{3}\log\rho + \frac{2}{3}\log(1-\rho)\right)\log n + \mathcal{O}(1) \tag{63c}$$

for $j = \rho n$ and $0 < \rho < 1$.

Corollary 5.4. *The expectation $A_{n,j}^{[R]}$ of the number of rightsided edges from the root to the node with key j in a p-tree of size n is for $n \to \infty$ asymptotically given by*

$$A_{n,j}^{[R]} = \frac{2}{3}H_j - \frac{1}{3} - \frac{1}{3j} + \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } j\,, \tag{64a}$$

$$A_{n,j}^{[R]} = \frac{2}{3}H_l - \frac{1}{3} + \frac{1}{3}\frac{1}{l+1} + \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } l = n-j\,, \tag{64b}$$

$$A_{n,j}^{[R]} = \frac{2}{3}\log n - \frac{2}{3}\rho^2 + \frac{2}{3}\rho - \frac{1}{3} + \frac{2}{3}\gamma + \frac{2}{3}\log\rho + \frac{2}{3}\log(1-\rho) + \mathcal{O}\left(\frac{1}{n}\right) \tag{64c}$$

for $j = \rho n$ and $0 < \rho < 1$.

6. The number of descendants

In the following we consider the number of descendants of particular nodes in objects of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. The number of descendants of a node x is here defined as the size of the subtree rooted at node x.

To get recurrences for the expectations $A_{n,j}$, $B_{n,j}$ and $C_{n,j}$ of the number of descendants of the node with key j of objects in the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ of size n, we introduce auxiliary parameters $q_{n,j}$, $C_{n,j}^{[Q]}$ and $C_{n,j}^{[\neg Q]}$.

Here $q_{n,j}$ is the probability, that the node with key j in an object of $\mathcal{C}$ of size n is positioned on the left path, that means, that node j lies on the path from the root to the left leaf. This is equivalent to the condition, that the node j is a left-to-right maximum in a random permutation of size n, and so we get

$$q_{n,j} = \frac{1}{n-j+1}\,.$$

With $C_{n,j}^{[Q]}$ we denote the conditional expectation of the number of desendants of node j in an object of $\mathcal{C}$ with size n, under the condition, that the node j lies on the left path.

Consequently $C_{n,j}^{[\neg Q]}$ denotes the conditional expectation of the number of descendants of node j in an object of $\mathcal{C}$ with size n under the condition, that the node j does *not* lie on the left path.

Decomposing the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the first element k of a random permutation leads for $n \ge j \ge 1$ to the following recurrences

$$A_{n,j} = \frac{1}{n}\left[\sum_{k=1}^{j-1}\left(q_{n-k,j-k}\left(C_{n-k,j-k}^{[Q]}+k\right)+(1-q_{n-k,j-k})\,C_{n-k,j-k}^{[\neg Q]}\right)+j+\sum_{k=j+1}^{n}B_{k-1,j}\right], \tag{65a}$$

$$B_{n,j} = \frac{1}{n}\left[\sum_{k=1}^{j-1}A_{n-k,j-k}+j+\sum_{k=j+1}^{n}B_{k-1,j}\right], \tag{65b}$$

$$C_{n,j} = \frac{1}{n}\left[\sum_{k=1}^{j-1}\left(q_{n-k,j-k}\left(C_{n-k,j-k}^{[Q]}+k\right)+(1-q_{n-k,j-k})\,C_{n-k,j-k}^{[\neg Q]}\right)+j+\sum_{k=j+1}^{n}A_{k-1,j}\right]. \tag{65c}$$

With the trivial equivalence

$$C_{n,j} = q_{n,j}C_{n,j}^{[Q]} + (1-q_{n,j})\,C_{n,j}^{[\neg Q]}\,,$$

the system (65) leads for $n \ge j \ge 1$ to the recurrences

$$A_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j-1}\left(C_{n-k,j-k}+\frac{k}{n-j+1}\right)+j+\sum_{k=j+1}^{n}B_{k-1,j}\right), \tag{66a}$$

$$B_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j-1} A_{n-k,j-k} + j + \sum_{k=j+1}^{n} B_{k-1,j}\right), \tag{66b}$$

$$C_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j-1}\left(C_{n-k,j-k} + \frac{k}{n-j+1}\right) + j + \sum_{k=j+1}^{n} A_{k-1,j}\right), \tag{66c}$$

with initial values $A_{0,0} = 0$, $B_{0,0} = 0$ and $C_{0,0} = 0$.

Introducing bivariate generating functions $A(z,u) = \sum_{j\ge 1}\sum_{n\ge j} A_{n,j} z^n u^j$ etc. we get from the system of recurrences (66) by multiplication with $nz^{n-1}u^j$ and summing up over all $n \ge j \ge 1$ the following system of differential equations

$$\frac{\partial}{\partial z}A(z,u) = \frac{u}{1-uz}C(z,u) + \frac{1}{1-z}B(z,u) + f_1(z,u), \tag{67a}$$

$$\frac{\partial}{\partial z}B(z,u) = \frac{u}{1-uz}A(z,u) + \frac{1}{1-z}B(z,u) + f_2(z,u), \tag{67b}$$

$$\frac{\partial}{\partial z}C(z,u) = \frac{u}{1-uz}C(z,u) + \frac{1}{1-z}A(z,u) + f_3(z,u) \tag{67c}$$

with

$$f_1(z,u) = \frac{u^2}{(1-uz)^3}\log\left(\frac{1}{1-z}\right) + \frac{u}{(1-z)(1-uz)^2},$$

$$f_2(z,u) = \frac{u}{(1-z)(1-uz)^2},$$

$$f_3(z,u) = \frac{u^2}{(1-uz)^3}\log\left(\frac{1}{1-z}\right) + \frac{u}{(1-z)(1-uz)^2}.$$

This system can be solved analogously to (24), and we get for $A(z,u)$

$$\begin{aligned}
A(z,u) = \Bigg[&-\frac{1}{3}uz + \frac{1}{6}u + \frac{1}{6} + \frac{1}{2}\frac{(u-1)^2}{u(1-z)} \\
&- \frac{2}{3}\frac{(u-1)^2}{1-uz}\Bigg]\int_0^z \frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt \\
&+ \left(-\frac{1}{6}\frac{(2uz-u-1)u}{(1-u)^3} + \frac{1}{2}\frac{1}{(1-z)(1-u)} - \frac{2}{3}\frac{u}{(1-uz)(1-u)}\right)\log^2\left(\frac{1}{1-z}\right) \\
&+ \left(+\frac{1}{3}\frac{(2uz-u-1)u}{(1-u)^3} + \frac{1}{3}\frac{u}{(1-uz)(1-u)}\right)\log\left(\frac{1}{1-z}\right)\log\left(\frac{1}{1-uz}\right) \\
&+ \left(-\frac{1}{6}\frac{(2uz-u-1)u}{(1-u)^3} - \frac{1}{6}\frac{1}{(1-z)(1-u)}\right)\log^2\left(\frac{1}{1-uz}\right) \\
&+ \Bigg[-\frac{2}{3}\frac{u}{(1-u)^2} + \frac{1}{2}\frac{1}{(1-z)(1-u)} - \frac{1}{2}\frac{1}{(1-z)^2u} - \frac{1}{2}\frac{u}{(1-uz)(1-u)} \\
&\quad + \frac{5}{3}\frac{u}{(1-uz)^2}\Bigg]\log\left(\frac{1}{1-z}\right) \\
&+ \left(\frac{2}{3}\frac{u}{(1-u)^2} - \frac{2}{3}\frac{1}{(1-z)(1-u)} + \frac{2}{3}\frac{u}{(1-uz)(1-u)}\right)\log\left(\frac{1}{1-uz}\right)
\end{aligned}$$

$$+\frac{1}{6}\frac{2uz-u-1}{1-u}+\frac{1}{6}\frac{u^2-3}{(1-z)(1-u)u}+\frac{1}{2}\frac{1}{(1-z)^2u}+\frac{2}{3}\frac{1}{(1-uz)(1-u)}\,. \tag{68}$$

To extract the coefficients of this expression, we need apart from (12) and (47) the following equations

$$\begin{aligned}&[z^nu^j]\frac{1}{(1-u)^3}\log\left(\frac{1}{1-z}\right)\log\left(\frac{1}{1-uz}\right)\\ &\quad= -\frac{1}{2}j+\frac{1}{2}\frac{j^2+3j+2}{n}H_j+\frac{1}{2}\frac{n^2-2jn-3n+j^2+3j+2}{n}\left(H_{n-1}-H_{n-j-1}\right)\,,\end{aligned} \tag{69a}$$

$$[z^nu^j]\frac{1}{(1-u)(1-uz)}\log\left(\frac{1}{1-z}\right)\log\left(\frac{1}{1-uz}\right)=\sum_{k=1}^{j}\frac{H_k}{n-k}\,. \tag{69b}$$

Then we get from (68) for the desired expectations

$$\begin{aligned}A_{n,j}&=\frac{1}{3}\sum_{k=1}^{j-1}\frac{H_k}{n-k}-\sum_{k=j+1}^{n}\frac{H_{k-j}}{k}-\frac{1}{6}H_n^2-\frac{2}{3}H_{n-j}^2+\frac{4}{3}H_{n-j}H_n-\frac{1}{6}H_j^2\\ &\quad+\frac{1}{6}\frac{j-6}{j}H_n-\frac{1}{6}\frac{j-6}{j}H_{n-j}-\frac{1}{6}\frac{n^2j-6n^2+2nj^2-nj+6n-2j^3}{n(n-1)j}H_j\\ &\quad+\frac{1}{6}H_n^{(2)}-\frac{2}{3}H_{n-j}^{(2)}+\frac{1}{6}H_j^{(2)}+\frac{1}{6}\frac{n^3+6n^2j-3nj^2-3nj-n+2j^3-3j^2+j}{n(n-1)(n-j+1)}\,,\end{aligned} \tag{70}$$

which can be simplified by use of equations (49). This leads to

Theorem 6.1. *The expectation $A_{n,j}$ of the number of descendants of the node with key j in a p-tree of size n is for $1\le j\le n$ given by*

$$\begin{aligned}A_{n,j}&=\frac{4}{3}\sum_{k=1}^{j}\frac{H_{n-k}}{k}-\frac{2}{3}H_{n-j}^2+\frac{4}{3}H_{n-j}H_n-\frac{1}{3}H_{n-j}H_j-\frac{2}{3}H_n^2-H_nH_j+\frac{1}{3}H_j^2\\ &\quad-\frac{1}{6}H_{n-j}+\frac{1}{6}H_n+\left(-\frac{1}{6}+\frac{1}{3}\frac{j(j-1)}{n-1}-\frac{j^2}{3n}\right)H_j-\frac{2}{3}H_{n-j}^{(2)}+\frac{2}{3}H_n^{(2)}+\frac{2}{3}H_j^{(2)}\\ &\quad+\frac{j}{n-j+1}-\frac{1}{3}\frac{j(j-1)}{n-1}+\frac{1}{6}\frac{j(2j-1)}{n}+\frac{1}{6}\,.\end{aligned} \tag{71}$$

To evaluate our findings asymptotically, under the assumptions j fixed; $j\sim\rho n$, with $0<\rho<1$; $n-j$, we state here the following lemma, that handles the sum $\sum_{k=1}^{j}\frac{H_{n-k}}{k}$ asymptotically. It can be obtained easily from (52) by use of (49).

Lemma 6.2. *The sum $\sum_{k=1}^{j}\frac{H_{n-k}}{k}$ has the following asymptotic equivalents:*

$$\sum_{k=1}^{j}\frac{H_{n-k}}{k}=H_j\log n+\gamma H_j+\mathcal{O}\left(\frac{1}{n}\right)\quad\text{for fixed } j\,, \tag{72a}$$

$$\sum_{k=1}^{j}\frac{H_{n-k}}{k}=\log^2 n+2\gamma\log n+\gamma^2-\frac{\pi^2}{6}+\mathcal{O}\left(\frac{1}{n}\right)\quad\text{for fixed } l=n-j\,, \tag{72b}$$

$$\sum_{k=1}^{j} \frac{H_{n-k}}{k} = \log^2(n) + (\log\rho + 2\gamma)\log n + \gamma^2 + \gamma\log\rho - \operatorname{dilog}(1-\rho) + o(1) \quad \text{for } j = \rho n \text{ and } 0 < \rho < 1\,. \tag{72c}$$

Now we obtain from (72a) by use of the previous lemma

Corollary 6.3. *The expectation $A_{n,j}$ of the number of descendants with key j in a p-tree of size n is for $n \to \infty$ asymptotically given by*

$$A_{n,j} = \frac{1}{3}H_j^2 - \frac{1}{6}H_j + \frac{2}{3}H_j^{(2)} + \frac{1}{6} + \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } j\,, \tag{73a}$$

$$A_{n,j} = \frac{n}{l+1} + H_l \log n + \mathcal{O}(1) \quad \text{for fixed } l = n-j\,, \tag{73b}$$

$$A_{n,j} = \frac{1}{3}\log^2 n + \left(\frac{1}{3}\rho^2 - \frac{1}{3}\rho - \frac{1}{6} + \frac{2}{3}\gamma + \frac{2}{3}\log\rho - \frac{1}{3}\log(1-\rho)\right)\log n + \mathcal{O}(1) \quad \text{for } j = \rho n \text{ and } 0 < \rho < 1\,. \tag{73c}$$

7. The number of key comparisons when inserting an element

To insert a new element into a p-tree we need the recursive algorithm `Insert` as described in section 1. Here we are interested in the average number of key comparisons that are made, when a particular new element is inserted into a random p-tree.

In the following we denote by $A_{n,j}^{[I]}$, $B_{n,j}^{[I]}$ and $C_{n,j}^{[I]}$ the expectations of the number of key comparisons, that are made when inserting the element with key $j + \frac{1}{2}$ with $j \in \mathbb{Z}$ and $0 \le j \le n$ into an object of size n of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$.

Decomposing the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the first element k of a random permutation leads for $n \ge 1$ and $0 \le j \le n$ to the following recurrences

$$A_{n,j}^{[I]} = \frac{1}{n}\left(\sum_{k=1}^{j} C_{n-k,j-k}^{[I]} + \sum_{k=j+1}^{n}\left(B_{k-1,j}^{[I]} + H_{n-k}\right)\right), \tag{74a}$$

$$B_{n,j}^{[I]} = \frac{1}{n}\left(\sum_{k=1}^{j}\left(A_{n-k,j-k}^{[I]} + 2\right) + \sum_{k=j+1}^{n}\left(B_{k-1,j}^{[I]} + 1\right)\right), \tag{74b}$$

$$C_{n,j}^{[I]} = \frac{1}{n}\left(\sum_{k=1}^{j} C_{n-k,j-k}^{[I]} + \sum_{k=j+1}^{n}\left(A_{k-1,j}^{[I]} + H_{n-k} + 2\right)\right), \tag{74c}$$

with the initial values $A_{0,0}^{[I]} = 0$, $B_{0,0}^{[I]} = 1$ and $C_{0,0}^{[I]} = 1$.

Multiplying (74) by $nz^{n-1}u^j$ and summing up leads by use of the bivariate generating functions

$$A^{[I]}(z,u) = \sum_{j\ge 0}\sum_{n\ge j} A_{n,j}^{[I]} z^n u^j\,, \quad B^{[I]}(z,u) = \sum_{j\ge 0}\sum_{n\ge j} B_{n,j}^{[I]} z^n u^j\,, \quad C^{[I]}(z,u) = \sum_{j\ge 0}\sum_{n\ge j} C_{n,j}^{[I]} z^n u^j$$

to the following system of differential equations

$$\frac{\partial}{\partial z}A^{[I]}(z,u)=\frac{u}{1-uz}C^{[I]}(z,u)+\frac{1}{1-z}B^{[I]}(z,u)+f_1(z,u)\,, \tag{75a}$$

$$\frac{\partial}{\partial z}B^{[I]}(z,u)=\frac{u}{1-uz}A^{[I]}(z,u)+\frac{1}{1-z}B^{[I]}(z,u)+f_2(z,u)\,, \tag{75b}$$

$$\frac{\partial}{\partial z}C^{[I]}(z,u)=\frac{u}{1-uz}C^{[I]}(z,u)+\frac{1}{1-z}A^{[I]}(z,u)+f_3(z,u) \tag{75c}$$

with

$$f_1(z,u)=\frac{1}{(1-z)^2(1-uz)}\log\left(\frac{1}{1-z}\right)\,,$$

$$f_2(z,u)=\frac{2u}{(1-z)(1-uz)^2}+\frac{1}{(1-z)^2(1-uz)}\,,$$

$$f_3(z,u)=\frac{1}{(1-z)^2(1-uz)}\log\left(\frac{1}{1-z}\right)+\frac{2}{(1-z)^2(1-uz)}\,.$$

Analogously to the previous sections we find as solution of (75) for $A^{[I]}(z,u)$:

$$\begin{aligned}A^{[I]}(z,u)=&\left(\frac{2}{3}uz-\frac{1}{3}u-\frac{1}{3}+\frac{1}{3}\frac{(u-1)^2}{(1-z)u}\right)\int_0^z\frac{1}{(1-t)^2(1-ut)^2}\log\left(\frac{1}{1-t}\right)dt\\&+\left(-\frac{1}{3}\frac{u}{(1-uz)(1-u)}+\frac{1}{3}\frac{1}{(1-z)(1-u)}\right)\log^2\left(\frac{1}{1-z}\right)\\&+\left[\frac{1}{3}\frac{(u+1)(2uz-u-1)}{(1-u)^2}+\frac{7}{3}\frac{1}{(1-z)(1-u)}-\frac{1}{3}\frac{1}{(1-z)^2u}\right.\\&\qquad\left.-\frac{1}{3}\frac{u^2+7u-1}{(1-uz)(1-u)}\right]\log\left(\frac{1}{1-z}\right)\\&+\left(-\frac{1}{3}\frac{(u+1)(2uz-u-1)}{(1-u)^2}-\frac{1}{3}\frac{u+1}{(1-z)u}\right)\log\left(\frac{1}{1-uz}\right)\\&+\frac{4}{3}\frac{uz}{(1-u)}+\frac{1}{3}\frac{2u^2-2u-1}{(1-z)(1-u)u}+\frac{1}{3}\frac{1}{(1-z)^2u}-\frac{1}{3}\frac{2u-3}{(1-uz)(1-u)}\,.\end{aligned} \tag{76}$$

Extracting the coefficients leads to

Theorem 7.1. *The expectation $A^{[I]}_{n,j}$ of the number of key comparisons, that are made by inserting a new element with key $j+\frac{1}{2}$ with the procedure* `Insert` *in a p-tree of size n is given by*

$$\begin{aligned}A^{[I]}_{n,j}=&-\frac{2}{3}\sum_{k=j+1}^{n}\frac{H_{k-j}}{k}+\frac{1}{3}H^2_{n-j}+\left(-\frac{1}{3}-\frac{2}{3j}\right)H_n+\left(\frac{1}{3}+\frac{2}{3j}\right)H_j\\&+\left(\frac{7}{3}+\frac{2}{3j}+\frac{2}{3}\frac{j^2}{n}-\frac{2}{3}\frac{j(j-1)}{n-1}\right)H_{n-j}-\frac{1}{3}H^{(2)}_{n-j}-\frac{1}{3}\frac{j(2j+5)}{n}+\frac{2}{3}\frac{j^2+j-1}{n-1}\\&+\frac{1}{3}\frac{1}{n-j+1}+\frac{1}{3}\frac{1}{n-j}-\frac{1}{3}\frac{1}{j+1}-\frac{1}{3j}\quad\text{for } 1\le j\le n-1\,,\\A^{[I]}_{n,n}=&\,1\ \text{ for } n\ge 1\,,\quad A^{[I]}_{0,0}=0\,,\quad A^{[I]}_{n,0}=2H_n-1\ \text{ for } n\ge 1\,.\end{aligned} \tag{77}$$

We also state the following asymptotic equivalents

Corollary 7.2. *The expectation $A_{n,j}^{[I]}$ of the number of key comparisons, that are made when inserting a new element with key $j+\frac{1}{2}$ with the procedure* `Insert` *in a p-tree of size n is for $n \to \infty$ asymptotically given by*

$$A_{n,j}^{[I]} = 2\log n + \mathcal{O}(1) \ \text{ for fixed } j\,, \tag{78a}$$

$$A_{n,j}^{[I]} = \frac{1}{3}H_l^2 + \frac{7}{3}H_l - \frac{1}{3} - \frac{1}{3}H_l^{(2)} + \frac{1}{3l} + \frac{1}{3(l+1)} + \mathcal{O}\left(\frac{1}{n}\right) \ \text{ for fixed } l = n-j\,, \tag{78b}$$

$$A_{n,j}^{[I]} = \frac{1}{3}\log^2 n + \left(-\frac{2}{3}\rho^2 + \frac{2}{3}\rho + \frac{2}{3}\gamma + \frac{7}{3} + \frac{2}{3}\log\rho + \frac{2}{3}\log(1-\rho)\right) + \mathcal{O}(1) \tag{78c}$$

for $j = \rho n$ and $0 < \rho < 1$.

Furthermore we get for the arithmetic mean $A_n^{[I]} = \frac{1}{n+1}\sum_{j=0}^{n} A_{n,j}^{[I]}$ of the number of key comparisons when inserting a new element in a p-tree of size n by summing up (78a)

Corollary 7.3. *The expectation $A_n^{[I]}$ of the number of key comparisons, that are made when inserting a randomly chosen element from the set $\left\{\frac{1}{2},\frac{3}{2},\dots,\frac{n+1}{2}\right\}$ with the algorithm* `Insert` *in a p-tree of size n is given by*

$$A_n^{[I]} = \frac{1}{3}H_{n+1}^2 + \frac{10}{9}H_{n+1} - \frac{1}{3}H_{n+1}^{(2)} - \frac{28}{27} \ \text{ for } n \geq 2\,, \ \ A_1^{[I]} = 1\,. \tag{79}$$

This result can be found already in [10].

8. Analysis of a procedure to split p-trees

Beside the algorithms `Insert` and `Delete` to insert resp. remove an element of a p-tree, we consider here the algorithm `Cut` to split a p-tree at an arbitrary place into two p-trees. This procedure was introduced in [13]. Thereby a tree T will be split into two p-trees according to a given key s in such a way that one tree T_1 contains all nodes from T with keys less or equal s, whereas the remaining tree $T - T_1$ contains all nodes with keys greater than s. The trees obtained in this way might also be empty.

In more detail, the recursive algorithm `Cut` reads as follows. (We use the abbreviation $T(x)$ for the subtree with node x as root, when x is a node of the tree T.)

1. (a) If T is empty or the root of T has a smaller key than s, then set $T_1 := T$, set $T := \emptyset$ and terminate.
   (b) Otherwise set $T_1 := \emptyset$ and go to step 2.
2. Move along the left path of T until the first node y with key less or equal s is found.
   (a) If no such node is found, terminate.
   (b) Otherwise we denote by z the predecessor of y. Then it is known that s lies between the keys of y and z.
      (i) If the right subtree of z is empty, then set $T_1 := T(y)$ and $T := T - T(y)$.
      (ii) Otherwise attach the right subtree of z between z and the left son of z and set the right subtree of the node z empty. Search the left leaf of $T(z)$ and connect the node y to it. Finally repeat step 2 starting from the node z.

In [13] it was stated, that the number of key comparisons splitting a randomly generated p-tree according to a key s in this algorithm is the same as the number of key comparisons inserting the element s in the tree. This is not fully correct, although the difference is marginal.

To analyse this parameter, we denote by $A^{[C]}_{n,j}$, $B^{[C]}_{n,j}$ and $C^{[C]}_{n,j}$ the expectations of the number of key comparisons, that are made by cutting an object of size n of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ according to the key $j+\frac{1}{2}$ with $j \in \mathbb{Z}$ and $0 \le j \le n$.

To get recurrences we decompose the object families w. r. t. the first element k of a random permutation. When the tree is cut at the position $k+\frac{1}{2}$, that means that $j=k$, then k is compared with the key a second time, whenever the node with key k has a predecessor with nonempty right subtree. In the following recurrence this fact leads in opposition to the recurrence (74) for the insertion algorithm to the additional functions $r_1(n,j)$, $r_2(n,j)$ and $r_3(n,j)$.

We have for $n \ge 1$ and $n \ge j \ge 0$

$$A^{[C]}_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j} C^{[C]}_{n-k,j-k} + \sum_{k=j+1}^{n}\left(B^{[C]}_{k-1,j} + H_{n-k}\right) + r_1(n,j)\right), \tag{80a}$$

$$B^{[C]}_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j}\left(A^{[C]}_{n-k,j-k} + 2\right) + \sum_{k=j+1}^{n}\left(B^{[C]}_{k-1,j} + 1\right) + r_2(n,j)\right), \tag{80b}$$

$$C^{[C]}_{n,j} = \frac{1}{n}\left(\sum_{k=1}^{j} C^{[C]}_{n-k,j-k} + \sum_{k=j+1}^{n}\left(A^{[C]}_{k-1,j} + H_{n-k} + 2\right) + r_3(n,j)\right), \tag{80c}$$

with

$$r_1(n,j) = \begin{cases} 1 - \frac{1}{n-j} & \text{for } 1 \le j \le n-1\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

$$r_2(n,j) = \begin{cases} 1 & \text{for } 1 \le j \le n-1\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

$$r_3(n,j) = \begin{cases} 1 - \frac{1}{n-j} & \text{for } 1 \le j \le n-1\,, \\ 0 & \text{otherwise} \end{cases}$$

and the initial values $A^{[C]}_{0,0} = 0$, $B^{[C]}_{0,0} = 1$ and $C^{[C]}_{0,0} = 1$.

Introducing the usual generating functions $A^{[C]}(z,u) = \sum_{j\ge 0}\sum_{n\ge j} A^{[C]}_{n,j} z^n u^j$ etc. these recurrences lead (by multiplying with $nz^{n-1}u^j$ and summing up) to the following system of differential equations,

$$\frac{\partial}{\partial z} A^{[C]}(z,u) = \frac{u}{1-uz} C^{[C]}(z,u) + \frac{1}{1-z} B^{[C]}(z,u) + f_1(z,u)\,, \tag{81a}$$

$$\frac{\partial}{\partial z} B^{[C]}(z,u) = \frac{u}{1-uz} A^{[C]}(z,u) + \frac{1}{1-z} B^{[C]}(z,u) + f_2(z,u)\,, \tag{81b}$$

$$\frac{\partial}{\partial z} C^{[C]}(z,u) = \frac{u}{1-uz} C^{[C]}(z,u) + \frac{1}{1-z} A^{[C]}(z,u) + f_3(z,u) \tag{81c}$$

with

$$\begin{aligned}
f_1(z,u) &= \frac{1}{(1-z)^2(1-uz)} \log\left(\frac{1}{1-z}\right) + \frac{uz}{(1-z)(1-uz)} - \frac{u}{1-uz} \log\left(\frac{1}{1-z}\right), \\
f_2(z,u) &= \frac{2u}{(1-z)(1-uz)^2} + \frac{1}{(1-z)^2(1-uz)} + \frac{uz}{(1-z)(1-uz)}, \\
f_3(z,u) &= \frac{1}{(1-z)^2(1-uz)} \log\left(\frac{1}{1-z}\right) + \frac{2}{(1-z)^2(1-uz)} + \frac{uz}{(1-z)(1-uz)} \\
&\quad - \frac{u}{1-uz} \log\left(\frac{1}{1-z}\right).
\end{aligned}$$

We get the following solution for the generating function

$$\begin{aligned}
A^{[C]}(z,u) &= \left(\frac{2}{3}uz - \frac{1}{3}u - \frac{1}{3} + \frac{1}{3}\frac{(u-1)^2}{(1-z)u}\right) \int_0^z \frac{1}{(1-t)^2(1-ut)^2} \log\left(\frac{1}{1-t}\right) dt \\
&\quad + \left(-\frac{1}{3}\frac{u}{(1-uz)(1-u)} + \frac{1}{3}\frac{1}{(1-z)(1-u)}\right) \log^2\left(\frac{1}{1-z}\right) \\
&\quad + \left(\frac{1}{3}\frac{(2uz-u-1)u}{(1-u)^2} + \frac{7}{3}\frac{1}{(1-z)(1-u)} - \frac{1}{3}\frac{1}{(1-z)^2 u} \right. \\
&\qquad \left. - \frac{1}{3}\frac{(u+6)u}{(1-uz)(1-u)}\right) \log\left(\frac{1}{1-z}\right) \\
&\quad + \left(-\frac{1}{3}\frac{(2uz-u-1)u}{(1-u)^2} - \frac{1}{3}\frac{1}{(1-z)}\right) \log\left(\frac{1}{1-uz}\right) \\
&\quad + \frac{2}{3}\frac{uz-u+1}{1-u} + \frac{2}{3}\frac{1}{1-uz} - \frac{1}{3}\frac{4u+1}{(1-z)u} + \frac{1}{3}\frac{1}{(1-z)^2 u}
\end{aligned} \tag{82}$$

and, by extracting the coefficients, the following theorem.

Theorem 8.1. *The expectation $A_{n,j}^{[I]}$ of the number of key comparisons, that are made by cutting a p-tree of size n with the algorithm* `Cut` *w. r. t. to an element with key $j+\frac{1}{2}$ is given by*

$$\begin{aligned}
A_{n,j}^{[C]} &= -\frac{2}{3}\sum_{k=j+1}^{n} \frac{H_{k-j}}{k} + \frac{1}{3}H_{n-j}^2 + \left(-\frac{1}{3} - \frac{2}{3j}\right) H_n + \left(\frac{1}{3} + \frac{2}{3j}\right) H_j \\
&\quad + \left(\frac{7}{3} + \frac{2}{3j} + \frac{2}{3}\frac{j^2}{n} - \frac{2}{3}\frac{j(j-1)}{n-1}\right) H_{n-j} - \frac{1}{3}H_{n-j}^{(2)} - \frac{2j^2+3j-1}{3n} \\
&\quad + \frac{2}{3}\frac{(j-1)(j+1)}{n-1} + \frac{1}{3}\frac{1}{n-j+1} - \frac{1}{3j} + \frac{1}{3} \quad \text{for } 1 \le j \le n-1, \\
A_{n,n}^{[C]} &= 1 \quad \text{for } n \ge 1, \quad A_{0,0}^{[C]} = 0, \; A_{n,0}^{[C]} = 2H_n - 1 \quad \text{for } n \ge 1.
\end{aligned} \tag{83}$$

We have the following asymptotic equivalents

Corollary 8.2. *The expectation $A_{n,j}^{[I]}$ of the number of key comparisons, that are made by cutting a p-tree of size n with the algorithm* `Cut` *w. r. t. to an element with key $j+\frac{1}{2}$ is for*

$n \to \infty$ asymptotically given by

$$A_{n,j}^{[C]} = 2\log n + \mathcal{O}(1) \quad \text{for fixed } j\,, \tag{84a}$$

$$A_{n,j}^{[C]} = \frac{1}{3}H_l^2 + \frac{7}{3}H_l - \frac{1}{3}H_l^{(2)} + \frac{1}{3(l+1)} + \mathcal{O}\left(\frac{1}{n}\right) \quad \text{for fixed } l = n-j\,, \tag{84b}$$

$$A_{n,j}^{[C]} = \frac{1}{3}\log^2 n + \left(-\frac{2}{3}\rho^2 + \frac{2}{3}\rho + \frac{2}{3}\gamma + \frac{7}{3} + \frac{2}{3}\log\rho + \frac{2}{3}\log(1-\rho)\right) + \mathcal{O}(1) \tag{84c}$$
for $j = \rho n$ and $0 < \rho < 1$.

Furthermore we state

Corollary 8.3. *The expectation $A_n^{[C]}$ of the number of key comparisons, that are made by cutting a p-tree of size n with the algorithm* `Cut` *w. r. t. to a randomly chosen element from the set $\left\{\frac{1}{2}, \frac{3}{2}, \dots, \frac{n+1}{2}\right\}$ is given by*

$$A_n^{[C]} = \frac{1}{3}H_{n+1}^2 + \frac{10}{9}H_{n+1} - \frac{1}{3}H_{n+1}^{(2)} - \frac{19n+46}{27(n+1)} \quad \text{for } n \ge 2\,, \quad A_1^{[C]} = 1\,. \tag{85}$$

9. The number of nodes with 0, 1 and 2 children

In this section we consider the number of nodes in a randomly generated p-tree with 0, 1 or 2 internal nodes as successors. First we state the following nonobvious relation between the p-trees and binary search trees.

Theorem 9.1. *The distribution of the number of nodes with 0, 1 or 2 children in a p-tree of size n is equal to the distribution of the number of nodes with 0, 1 or 2 children in random binary search trees of size n.*

Proof. At first we remark, that to prove the theorem it is enough to show, that the number of leaves in p-trees and binary search trees are identically distributed. If we denote in an arbitrary binary tree of size n with n_0, n_1 resp. n_2 the number of nodes with 0, 1 and 2 children, then we have by counting the nodes and the edges in the tree the following relations

$$n_0 + n_1 + n_2 = n \quad \text{and} \quad n_1 + 2n_2 = n - 1\,.$$

Therefore n_1 and n_2 are obtained from n_0 via

$$n_1 = n + 1 - 2n_0 \quad \text{and} \quad n_2 = n_0 - 1\,.$$

Now we show, that the number of leaves in p-trees and binary search trees of the same size are identically distributed. We denote here by $F_{n,m}$ the number of binary search trees of size n with exactly m leaves. They fulfill for $n \ge 2$ the following recurrence

$$F_{n,m} = \frac{1}{n}\sum_{k=1}^{n}\sum_{l=1}^{m} F_{k-1,l} \cdot F_{n-k,m-l}\,, \tag{86}$$

with initial values $F_{0,0} = 1$ and $F_{1,1} = 1$.

Introducing the bivariate generating function $F(z,v)$ of the $F_{n,m}$

$$F(z,v) = \sum_{n\ge 0}\sum_{m\ge 0} F_{n,m} z^n v^m\,,$$

the above recurrence (86) leads (by multiplying with $nz^{n-1}v^m$ and summing up) to the linear differential equation

$$\frac{\partial}{\partial z} F(z,v) = F^2(z,v) + v - 1\,, \quad F(0,v) = 1\,. \tag{87}$$

Now it suffices to show, that the generating function $A(z,v)$ of the number $A_{n,m}$ of p-trees of size n with exactly m leaves fulfill the same differential equation.

Decomposing the objects of the families $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ w. r. t. the first element k of a random permutation with n elements, the probabilities $A_{n,m}$, $B_{n,m}$ resp. $C_{n,m}$, that an object of size n of the corresponding family has exactly m leaves leads for $n \geq 1$ to the following recurrence

$$A_{n,m} = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{m} C_{n-k,i} \cdot B_{k-1,m-i}\,, \tag{88a}$$

$$B_{n,m} = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{m} A_{n-k,i} \cdot B_{k-1,m-i}\,, \tag{88b}$$

$$C_{n,m} = \frac{1}{n} \sum_{k=1}^{n} \sum_{i=0}^{m} C_{n-k,i} \cdot A_{k-1,m-i} \tag{88c}$$

with initial values $A_{0,m} = \delta_{0,m}$, $B_{0,m} = \delta_{1,m}$ and $C_{0,m} = \delta_{0,m}$.

Introducing the usual generating functions $A(z,v) = \sum_{n\geq 0}\sum_{m\geq 0} A_{n,m} z^n v^m$ etc. the above recurrences lead by multiplying with $nz^{n-1}v^m$ and summing up to the following system of differential equations

$$\frac{\partial}{\partial z} A(z,v) = B(z,v)\, C(z,v)\,, \tag{89a}$$

$$\frac{\partial}{\partial z} B(z,v) = A(z,v)\, B(z,v)\,, \tag{89b}$$

$$\frac{\partial}{\partial z} C(z,v) = A(z,v)\, C(z,v)\,, \tag{89c}$$

with the initial values $A(0,v) = 1$, $B(0,v) = v$ and $C(0,v) = 1$.

Differentiating equation (89a) with respect to z leads to

$$\frac{\partial^2}{\partial z^2} A(z,v) = \frac{\partial}{\partial z} B(z,v)\, C(z,v) + B(z,v)\, \frac{\partial}{\partial z} C(z,v)\,. \tag{90}$$

Now we substitute $\frac{\partial}{\partial z} B(z,v)$ and $\frac{\partial}{\partial z} C(z,v)$ in (90) by the corresponding right sides of the equations (89b) and (89c). Then we get

$$\frac{\partial^2}{\partial z^2} A(z,v) = 2A(z,v)\, B(z,v)\, C(z,v)\,. \tag{91}$$

Substituting the factor $B(z,v)\,C(z,v)$ in this equation by $\frac{\partial}{\partial z}A(z,v)$ according to (89a), we get the following nonlinear differential equation of order 2 for $A(z,v)$:

$$\frac{\partial^2}{\partial z^2}A(z,v) = 2\,A(z,v)\,\frac{\partial}{\partial z}A(z,v)\,,$$
$$A(0,v) = 1\,, \quad \frac{\partial}{\partial z}A(z,v)\bigg|_{z=0} = v\,. \tag{92}$$

This differential equation is of the type $A'' = f(A, A')$, that means the variable z does not appear explicitly. To handle such an equation, we define a function $q(A,v)$ of A and v, with $\frac{\partial}{\partial z}A(z,v) = q(A,v)$ (see e. g. [15]). We get then from (92) a differential equation for the function q by the following procedure:

$$\frac{dq}{dA} = \frac{dq}{dz}\,\frac{dz}{dA} = \frac{dq}{dz}\,\frac{1}{\frac{dA}{dz}} = \frac{\partial^2}{\partial z^2}A\,\frac{1}{\frac{\partial}{\partial z}A} = 2\,A\,. \tag{93}$$

This differential equation has the solution

$$q(A,v) = A(z,v)^2 + c_1(v) \tag{94}$$

with a function $c_1(v)$.

With $q(A,v) = \frac{\partial}{\partial z}A(z,v)$ and fitting to the initial conditions we get the same first order differential equation for $A(z,v)$ as we obtained for $F(z,v)$

$$\frac{\partial}{\partial z}A(z,v) = A(z,v)^2 - (1-v)\,, \quad A(0,v) = 1\,, \tag{95}$$

and this suffices to prove the theorem.

Of course this differential equation can be solved, and we get

$$A(z,v) = \sqrt{1-v} - \frac{2\,v\sqrt{1-v}}{v - \left(1+\sqrt{1-v}\right)^2 e^{-2\sqrt{1-v}\,z}}\,. \tag{96}$$

□

Therefore we can apply theorems that were proved by Devroye in [5] for the distribution of the nodes with 0, 1 and 2 children in binary search trees to the instance of p-trees. Denoting the random variables for the number of nodes with 0, 1 resp. 2 children in p-trees of size n with $I_n^{(0)}$, $I_n^{(1)}$ and $I_n^{(2)}$, we get

Theorem 9.2. *For the expectations $\mathbb{E}\left(I_n^{(i)}\right)$ and the variances $\mathbb{V}\left(I_n^{(i)}\right)$ with $0 \le i \le 2$, we have the following asymptotic expansions for $n \to \infty$*

$$\mathbb{E}\left(I_n^{(0)}\right) \sim \frac{n}{3}\,,\ \mathbb{E}\left(I_n^{(1)}\right) \sim \frac{n}{3}\,,\ \mathbb{E}\left(I_n^{(2)}\right) \sim \frac{n}{3}\,,$$
$$\mathbb{V}\left(I_n^{(0)}\right) \sim \frac{2}{45}n\,,\ \mathbb{V}\left(I_n^{(1)}\right) \sim \frac{8}{45}n\,,\ \mathbb{V}\left(I_n^{(2)}\right) \sim \frac{2}{45}n\,.$$

Furthermore the sequence of random variables

$$\frac{I_n^{(0)} - \frac{n}{3}}{\sqrt{\frac{2}{45}n}}\,,\ \frac{I_n^{(1)} - \frac{n}{3}}{\sqrt{\frac{8}{45}n}}\ \text{and}\ \frac{I_n^{(2)} - \frac{n}{3}}{\sqrt{\frac{2}{45}n}}$$

is for $n \to \infty$ weakly convergent to the standard (Gaussian) normal distribution $\mathcal{N}(0,1)$.

References

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions.* Dover Publications, Inc., New York, ninth edition, 1970.

[2] L. Bieberbach. *Theorie der gewöhnlichen Differentialgleichungen.* Springer Verlag, Berlin, 1965.

[3] K. L. Chung. *A course in Probability Theory.* Academic Press, San Diego, 1974.

[4] L. Comtet. *Advanced Combinatorics.* Reidel, Dordrecht, 1974.

[5] L. Devroye. Limit Laws for Local Counters in Random Binary Search Trees. *Random Structures and Algorithms*, 2:303–315, 1991.

[6] Ph. Flajolet and M. Soria. General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, 114:159–180, 1993.

[7] G. H. Gonnet and R. Baeza-Yates. *Handbook of Algorithms and Data Structures.* Addison-Wesley, Reading, Massachusetts, second edition, 1991.

[8] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, Reading, Massachusetts, second edition, 1994.

[9] D. H. Greene and D. E. Knuth. *Mathematics for the Analysis of Algorithms.* Birkhäuser, Boston, second edition, 1982.

[10] A. Jonassen and O.-J. Dahl. Analysis of an Algorithm for Priority Queue Administration. *BIT*, 15:409–422, 1975.

[11] E. Kamke. *Differentialgleichungen - Lösungsmethoden und Lösungen*, volume 1. Teubner, Stuttgart, tenth edition, 1983.

[12] P. Kirschenhofer and H. Prodinger. Comparisons in Hoare's Find Algorithm. to appear in *Combinatorics, Probability & Computing*, 1996.

[13] O. Nevalainen and J. Teuhola. The Efficiency of two indexed Priority Queue Algorithms. *BIT*, 18:320–333, 1978.

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

[15] H. Wenzel and P. Meinhold. *Gewöhnliche Differentialgleichungen.* Teubner, Stuttgart, 1994.

Institut für Algebra und Diskrete Mathematik, Technical University of Vienna, Wiedner Hauptstrasse 8–10, A-1040 Vienna, Austria.

E-mail address: `Helmut.Prodinger@tuwien.ac.at`

E-mail address: `e9125354@fbma.tuwien.ac.at (Alois Panholzer)`