A COMBINATORIAL APPROACH TO THE ANALYSIS OF BUCKET RECURSIVE TREES AND VARIANTS

MARKUS KUBA AND ALOIS PANHOLZER

ABSTRACT. In this work we provide a combinatorial analysis of bucket recursive trees, which have been introduced previously as a natural generalization of the growth model of recursive trees. Our analysis is based on the description of bucket recursive trees as a special instance of so called bucket increasing trees, which is a family of combinatorial objects introduced in this paper. Using this combinatorial description we obtain exact and limiting distribution results for the parameters depth of a specified element, descendants of a specified element and degree of a specified element.

1. INTRODUCTION

Recursive trees are one of the most natural combinatorial tree models with applications in several fields, e.g., it has been introduced as a model for the spread of epidemics, for pyramid schemes, for the family trees of preserved copies of ancient texts and furthermore it is related to the Bolthausen-Sznitman coalescence model (see, e.g., [5, 10]). A recursive tree with n nodes is an unordered rooted tree, where the nodes are labelled by distinct integers from $\{1, 2, \ldots, n\}$ in such a way that the sequence of labels lying on the unique path from the root node to any node in the tree are always forming an increasing sequence. This implies that the root node is always labelled by 1. Due to this description recursive trees are falling into the combinatorial class of increasing tree families, see, e.g., [1]. It is well known (and easy to show by induction) that there are $(n-1)!$ different recursive trees with n nodes. It is of particular interest in applications to assume the random recursive tree model and to speak about a random recursive tree with n nodes, which means that one of the $(n-1)!$ possible recursive trees with n nodes is chosen with equal probability, i.e., the probability that a particular tree with n nodes is chosen is always $1/(n-1)!$.

The usefulness of this tree model relies at least in parts on the fact that there also exists a probabilistic description of random recursive trees via a simple stochastic growth rule: in order to get a random recursive tree T' with $n+1$ nodes one can choose a random recursive tree T with n nodes and choose uniformly at random one of the n nodes $v \in T$ as a parent node and attach the node $n+1$ to v. Starting with node 1 this leads after $n-1$ insertion steps (inserting successively the labels $2, 3, \ldots, n$) to a random recursive tree with n nodes and easily explains that there are $(n-1)!$ different recursive trees with n nodes.

An interesting and natural generalization of random recursive trees has been introduced in [9], which are called (random) bucket recursive trees. In this model the nodes of a bucket recursive tree are buckets, which can contain up to a fixed integer amount of $b \geq 1$ elements (= labels). A (probabilistic) description of random bucket recursive trees is given by a generalization of the stochastic growth rule for ordinary random recursive trees (which are the special instance $b = 1$), where a tree grows by progressive attraction of increasing integer labels: when inserting element $n+1$ into an existing bucket recursive tree containing n elements (i.e., containing the labels $\{1, 2, \ldots, n\}$) all n existing elements in the tree compete to attract the element $n+1$, where all existing elements have equal chance to recruit the new element. If the element winning this competition is contained in a node with less than b elements (an unsaturated bucket or node), element $n+1$ is added to this node, otherwise if the winning element is contained in a node with already b elements (a saturated bucket or node), element $n+1$ is attached to this node as a new bucket containing only the element $n+1$. Starting with a single bucket as root node containing only element 1 leads after $n-1$ insertion steps, where the labels $2, 3, \ldots, n$ are successively inserted according to this growth rule, to a so called random bucket recursive tree

2000 *Mathematics Subject Classification.* 05C05, 60F05.

Key words and phrases. bucket recursive trees, node-depth, number of descendants, node-degree, limiting distribution.

This work was supported by the Austrian Science Foundation FWF, grant S9608-N13. The second author was also supported by the European FP6-NEST-Adventure Program, contract n° 028875.

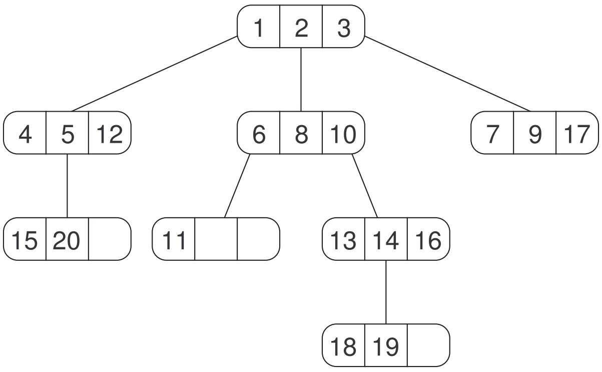

FIGURE 1. A bucket recursive tree of size $n = 20$ with maximal bucket size $b = 3$. The element $j = 8$ has depth 1, 8 descendants and out-degree 2.

with n elements and maximal bucket size b. Of course, the above growth rule for inserting the element $n+1$ could also be formulated by saying that, for an existing bucket recursive tree T with n elements, the probability that a certain node $v \in T$ attracts the new element $n+1$ is proportional to the number of elements contained in v, let us say k with $1 \le k \le b$, and is thus given by $\frac{k}{b}$. As the authors of [9] mention this growth rule for random bucket recursive trees could model a variety of possible recruiting situations, as, e.g, for a business in the service sector. Different bucketing strategies are naturally used in data structures in computer science, as, e.g., for the construction of m-ary search trees (see, e.g., [2]).

The aim of this paper is to give also a combinatorial description of bucket recursive trees generalizing the one for ordinary recursive trees. We do this by generalizing a class of weighted tree families, so called simple families of increasing trees, to a class of bucket trees, which we call families of bucket increasing trees. Bucket recursive trees will then turn out to be a special instance of a bucket increasing tree family. The gain of the combinatorial description provided here is that the natural combinatorial decomposition of a bucket recursive tree into a root bucket and its subtrees will lead to a recursive description of several important tree parameters in random bucket recursive trees. Often this combinatorial decomposition can be translated "almost automatically" into certain equations (here mainly differential equations) for suitable generating functions. Thus besides probabilistic techniques, as a description via Pólya-Eggenberger urn models or embedding into continuous time branching processes (see, e.g., [8]), which rely on the stochastic growth rule of random bucket recursive trees and turn out to be very powerful for a variety of parameters (like "extremal parameters" as the so called height of the tree, see [9]), one is able to apply also techniques of analytic combinatorics (see, e.g., [4]), which itself turn out to be powerful for a variety of parameters.

We illustrate the usefulness of this combinatorial description for a detailed study of some important "local parameters" for random bucket recursive trees. In particular we are interested in the effect of bucketing on "label-based parameters" and we are going to answer the corresponding questions for the random variables "depth" of element j (i.e., the number of edges from the root node to the node containing element j) denoted by $D_{n,j}$, the number of "descendants" of element j (i.e., the total number of elements with a label $\ge j$ contained in the subtree rooted with the node containing element j) denoted by $Y_{n,j}$, and the "out-degree" of element j (i.e., the out-degree of the node containing element j) denoted by $X_{n,j}$ in a random bucket recursive tree with n elements. Since the depth of node j in a random bucket recursive tree with n elements is independent of n, which is a consequence of the description via a stochastic growth rule, we may restrict ourselves to a study of the depth of the largest element n in a random bucket recursive tree with n elements and thus to the r.v. $D_n := D_{n,n}$. However, for all the parameters mentioned and all fixed maximal-bucket sizes b, we are able to give a complete characterization of the limiting distribution behaviour and the phase changes appearing for all regions $j = j(n)$, where the label $1 \le j \le n$ is possibly growing with the total number n of inserted elements. An example of a bucket recursive tree and the parameters considered is given in Figure 1.

We remark that the effect of bucketing on some "global parameters", in particular on the distribution of the r.v. $X_n^{[k]}$, which counts the number of nodes containing a certain number $1 \le k \le b$ of elements in a random bucket recursive tree with n elements, has been considered and described in [2, 9]. For this parameter it turns out that up to a maximal bucket-size $b \le 26$ the random vector $(X_n^{[1]}, \dots, X_n^{[b]})$ satisfies (after suitable normalization) a multivariate normal limit law, but for $b \ge 27$ the behaviour changes and an oscillating behaviour of the variances $\mathbb{V}(X_n^{[k]})$ appears.

The plan of the paper is as follows. In Section 2 we give the combinatorial description of bucket recursive trees and in Section 3 we give limiting distribution results for the parameters depth, number of descendants and node-degree of a specified element, which are all obtained by using this combinatorial description of bucket recursive trees. The proof of these results is given in Sections 4-6.

With $X \stackrel{(d)}{=} Y$ we denote the equality in distribution of two r.v. X and Y and we write $X_n \xrightarrow{(d)} X$ for the weak convergence (i.e., convergence in distribution) of a sequence of r.v. X_n to a r.v. X. We denote by $H_n := \sum_{k=1}^n \frac{1}{k}$ the harmonic numbers and by $H_n^{(r)} := \sum_{k=1}^n \frac{1}{k^r}$ the r-th order harmonic numbers. Furthermore, we use the abbreviation $H_{n+\alpha} - H_\alpha := \sum_{k=1}^n \frac{1}{k+\alpha}$ for the continuation of the harmonic numbers for a complex $\alpha \in \mathbb{C} \setminus \{-1, -2, -3, \dots\}$. Moreover, the signless Stirling numbers of first kind are denoted by $\left[{n \atop m}\right]$ and the Stirling numbers of second kind are denoted by $\left\{{n \atop m}\right\}$. With $x^{\underline{k}} := x(x-1)\cdots(x-k+1)$ and $x^{\overline{k}} := x(x+1)\cdots(x+k-1)$ we denote the falling and rising factorials, respectively.

2. Combinatorial description of bucket recursive trees

2.1. **Bucket increasing tree families.** Our basic objects are rooted ordered trees (the order of the subtrees of a node is of relevance), where the nodes are "buckets" with an integer capacity c, with $1 \le c \le b$ for a given maximal integer bucket-size $b \ge 1$ and the additional restriction, that all internal nodes (i.e., non-leaves) in the tree must be saturated, while the leaves might be either saturated or unsaturated. We always call a node v with capacity $c(v) = b$ "saturated" and otherwise "unsaturated". A tree defined in this way is called a bucket ordered tree with maximal bucket-size b. It will be convenient to define for bucket ordered trees the size $|T|$ of a tree T via $|T| = \sum_v c(v)$, where $c(v)$ ranges over all vertices of T. An increasing labelling of a bucket ordered tree T is then a labelling of T, where the labels $\{1, 2, \dots, |T|\}$ are distributed amongst the nodes of T, such that the following conditions are satisfied: (i) every node v contains exactly $c(v)$ labels, (ii) the labels within a node are arranged in increasing order, (iii) each sequence of labels along any path starting at the root is increasing.

Then a class $\mathcal{T}$ of a family of bucket increasing trees with maximal bucket-size b can be defined in the following way. A sequence of non-negative numbers $(\varphi_k)_{k\ge0}$ with $\varphi_0 > 0$ and a sequence of non-negative numbers $\psi_1, \psi_2, \dots, \psi_{b-1}$ is used to define the weight $w(T)$ of any bucket ordered tree T by $w(T) := \prod_v w(v)$, where $w(v)$ ranges over all vertices of T. The weight $w(v)$ of a node v is given as follows, where $d(v)$ denotes the out-degree (i.e., the number of children) of node v:

$$w(v) = \begin{cases} \varphi_{d(v)}, & \text{if } c(v) = b, \\ \psi_{c(v)}, & \text{if } c(v) < b. \end{cases}$$

Thus for saturated nodes the weight is dependent on the out-degree and described by the sequence φ_k, whereas for unsaturated nodes the weight is dependent on the capacity and described by the sequence ψ_k.

Furthermore, $\mathcal{L}(T)$ denotes the set of different increasing labellings of the tree T with distinct integers $\{1, 2, \dots, |T|\}$, where $L(T) := \big|\mathcal{L}(T)\big|$ denotes its cardinality. Then the family $\mathcal{T}$ consists of all trees T together with their weights $w(T)$ and the set of increasing labellings $\mathcal{L}(T)$.

For a given degree-weight sequence $(\varphi_k)_{k\ge0}$ with a degree-weight generating function $\varphi(t) := \sum_{k\ge0} \varphi_k t^k$ and a bucket-weight sequence $\psi_1, \dots, \psi_{b-1}$, we define now the total weights by $T_n := \sum_{|T|=n} w(T) \cdot L(T)$.

It is then not difficult to show that the exponential generating function $T(z) := \sum_{n\ge1} T_n \frac{z^n}{n!}$ of the total weights T_n is characterized by the following differential equation of order b:

$$\frac{d^b}{dz^b} T(z) = \varphi\big(T(z)\big), \qquad (1)$$
$$T(0) = 0, \quad T^{(k)}(0) = \psi_k, \quad \text{for } 1 \le k \le b-1.$$

This could be done by setting up a recurrence for the total weights T_n:

$$T_n = \sum_{r\ge0} \varphi_r \sum_{\substack{k_1+\dots+k_r=n-b \\ k_1,\dots,k_r\ge1}} T_{k_1} \dots T_{k_r} \binom{n-b}{k_1, k_2, \dots, k_r}, \quad \text{for } n \ge b, \qquad (2)$$

and treat it by introducing the exponential generating function $T(z)$.

However it is advantageous for such enumeration problems to describe a family of increasing trees $\mathcal{T}$ by the following formal recursive equation:

$$\begin{aligned}\mathcal{T} &= \psi_1 \cdot \boxed{1\quad\quad} \,\dot\cup\, \psi_2 \cdot \boxed{1\,|\,2\quad\quad} \,\dot\cup\, \cdots \,\dot\cup\, \psi_{b-1} \cdot \boxed{1\,|\,2\,|\cdots|\,b\text{-}1\,} \,\dot\cup \\ &\qquad \varphi_0 \cdot \boxed{1\,|\,2\,|\cdots|\,b} \,\dot\cup\, \varphi_1 \cdot \boxed{1\,|\,2\,|\cdots|\,b} \times \mathcal{T} \,\dot\cup\, \varphi_2 \cdot \boxed{1\,|\,2\,|\cdots|\,b} \times \mathcal{T} * \mathcal{T} \,\dot\cup\, \varphi_3 \cdot \boxed{1\,|\,2\,|\cdots|\,b} \times \mathcal{T} * \mathcal{T} * \mathcal{T} \,\dot\cup\, \cdots \\ &= \psi_1 \cdot \boxed{1\quad\quad} \,\dot\cup\, \psi_2 \cdot \boxed{1\,|\,2\quad\quad} \,\dot\cup\, \cdots \,\dot\cup\, \psi_{b-1} \cdot \boxed{1\,|\,2\,|\cdots|\,b\text{-}1\,} \,\dot\cup\, \boxed{1\,|\,2\,|\cdots|\,b} \times \varphi\big(\mathcal{T}\big),\end{aligned} \tag{3}$$

where $\boxed{1\,|\,2\,|\cdots|\,k\,}$ denotes a bucket of capacity k labelled by $1, 2, \dots, k$, $\times$ the cartesian product, $*$ the partition product for labelled objects, and $\varphi(\mathcal{T})$ the substituted structure (see, e.g., [13]). Then the differential equation (1) follows immediately by translating equation (3), but this formal description will turn out to be useful in particular when considering certain parameters in bucket increasing trees; see Sections 4-6.

2.2. **Description of bucket recursive trees as a bucket increasing tree family.** In the following we will show that bucket recursive trees can be considered as a certain bucket increasing tree family. We claim that the family of bucket recursive trees can be modeled, e.g., by using the following degree-weight and bucket-weight sequences (the choice of the sequences leading to bucket recursive trees is not unique):

$$\varphi_k = \frac{(b-1)!b^k}{k!}, \quad \text{for } k \ge 0, \qquad \psi_k = (k-1)!, \quad \text{for } 1 \le k \le b-1.$$

To show that this choice of sequences is actually a model for bucket recursive trees we have to show that this combinatorial family $\mathcal{T}$ of bucket increasing trees has the same stochastic growth rule as bucket recursive trees, namely: given an arbitrary bucket increasing tree $T \in \mathcal{T}$ of size $|T| = n$, then the probability that a new element $n+1$ is attracted by a node $v \in T$ with capacity $c(v) = k$ must be given by $\frac{k}{n}$.

We use now the notation $T \to T'$ to denote that T' is obtained from T with $|T| = n$ by incorporating element $n+1$, i.e., either by attaching element $n+1$ to a saturated node $v \in T$ at one of the $d(v)+1$ possible positions (recall that bucket increasing trees are per definition ordered trees and thus the order of the subtrees is of relevance) by creating a new bucket of capacity 1 containing element $n+1$ or by adding element $n+1$ to an unsaturated node $v \in T$ by increasing the capacity of v by 1. If we want to express that node $v \in T$ has attracted the element $n+1$ leading from T to T' we use the notation $T \xrightarrow{v} T'$. If there exists a stochastic growth rule for a bucket increasing tree family $\mathcal{T}$, then it must hold that for a given tree $T \in \mathcal{T}$ of size $|T| = n$ and a given node $v \in T$ the probability $p_T(v)$, which gives the probability that element $n+1$ is attracted by node $v \in T$ is given as follows:

$$p_T(v) = \frac{\sum_{T' \in \mathcal{T} : T \xrightarrow{v} T'} w(T')}{\sum_{\tilde{T} \in \mathcal{T} : T \to \tilde{T}} w(\tilde{T})} = \frac{\sum_{T' \in \mathcal{T} : T \xrightarrow{v} T'} \frac{w(T')}{w(T)}}{\sum_{\tilde{T} \in \mathcal{T} : T \to \tilde{T}} \frac{w(\tilde{T})}{w(T)}}. \tag{4}$$

For a certain tree $\tilde{T}$ with $T \xrightarrow{u} \tilde{T}$ and $u \in T$ the quotient of the weight of the trees $\tilde{T}$ and T is by the definition of bucket increasing trees given as follows, where we define for simplicity $\psi_b := \varphi_0$:

$$\frac{w(\tilde{T})}{w(T)} = \begin{cases} \psi_1 \frac{\varphi_{k+1}}{\varphi_k}, & \text{for } c(u) = b \quad \text{and} \quad d(u) = k, \\ \frac{\psi_{k+1}}{\psi_k}, & \text{for } c(u) = k < b. \end{cases}$$

For a given tree $T \in \mathcal{T}$ we define by $m_k := |\{u \in T : c(u) = k < b\}|$ the number of unsaturated nodes of T with capacity $k < b$ and by $n_k := |\{u \in T : c(u) = b \text{ and } d(u) = k\}|$ the number of saturated nodes of T with out-degree $k \ge 0$. It holds then

$$n = \sum_{u \in T} c(u) = \sum_{k=1}^{b-1} k m_k + b \sum_{k \ge 0} n_k$$

and (where we use that there are $k+1$ possibilities of attaching a new node to a saturated node $u \in T$ with out-degree $d(u) = k$):

$$\sum_{\tilde{T} \in \mathcal{T} : T \to \tilde{T}} \frac{w(\tilde{T})}{w(T)} = \sum_{k=1}^{b-1} m_k \frac{\psi_{k+1}}{\psi_k} + \sum_{k \ge 0} n_k (k+1) \psi_1 \frac{\varphi_{k+1}}{\varphi_k}.$$

Thus if one chooses the weights $\psi_k = (k-1)!$ and $\varphi_k = \frac{(b-1)!b^k}{k!}$ we obtain further

$$\sum_{\tilde{T} \in \mathcal{T}: T \to \tilde{T}} \frac{w(\tilde{T})}{w(T)} = \sum_{k=1}^{b-1} k m_k + \psi_1 \sum_{k \geq 0} n_k (k+1) \frac{b}{k+1} = \sum_{k=1}^{b-1} k m_k + b \sum_{k \geq 0} n_k = n.$$

Furthermore by choosing these weights φ_k and ψ_k we get

$$\sum_{T' \in \mathcal{T}: T \xrightarrow{v} T'} \frac{w(T')}{w(T)} = \begin{cases} (k+1)\psi_1 \frac{\varphi_{k+1}}{\varphi_k} = b, & \text{for } c(v) = b \quad \text{and} \quad d(v) = k, \\ \frac{\psi_{k+1}}{\psi_k} = k, & \text{for } c(v) = k < b, \end{cases}$$

and thus

$$\sum_{T' \in \mathcal{T}: T \xrightarrow{v} T'} \frac{w(T')}{w(T)} = k, \quad \text{for } c(v) = k.$$

Therefore we have shown that by choosing the weight sequences $\psi_k = (k-1)!$ and $\varphi_k = \frac{(b-1)!b^k}{k!}$ the probability $p_T(v)$ that in a bucket increasing tree T of size $|T| = n$ the node v with capacity $c(v) = k$ attracts element $n+1$ is always given by $\frac{k}{n}$, which coincides with the stochastic growth rule for bucket recursive trees.

We obtain then from equation (1) that the exponential generating function $T(z) := \sum_{n \geq 1} T_n \frac{z^n}{n!}$ of the total-weight T_n of bucket recursive trees of size n satisfies the differential equation

$$\frac{d^b}{dz^b} T(z) = (b-1)! e^{bT(z)}, \tag{5}$$

with initial conditions $T(0) = 0$ and $\frac{d^k}{dz^k} T(z)\Big|_{z=0} = (k-1)!$, for $1 \leq k \leq b-1$. The solution of this equation is given by

$$T(z) = \log \frac{1}{1-z} = \sum_{n \geq 1} (n-1)! \frac{z^n}{n!}. \tag{6}$$

Hence the total weight of all size-n bucket recursive trees is given by $T_n = (n-1)!$.

We remark that we have introduced here the more general combinatorial objects "bucket increasing trees" to describe bucket recursive trees by using specific weight sequences $(\varphi_k)_{k \geq 0}$ and $\psi_1, \ldots, \psi_{b-1}$ for the following reasons: (i) the combinatorial decompositions used in Sections 4-6 hold for arbitrary weight sequences and thus for general bucket increasing trees and seem to be more transparent for them. (ii) it seems to be interesting (and it is planned by the authors) to study the effect of bucketing also for other increasing tree families, as, e.g., for growth models with a "preferential attachment rule" like generalized plane-oriented recursive trees.

3. Results for label-based parameters

Here we give our main results for the exact and asymptotic behaviour of the parameters depth of element n, the number descendants of element j and the out-degree of element j in a random bucket recursive tree of size n (and fixed maximal bucket-size b). In the formulation of the theorems there will appear numbers λ_i, with $1 \leq i \leq b$, which are given by the roots of the equation

$$\lambda^{\overline{b}} - b! = \lambda(\lambda+1)\cdots(\lambda+b-1) - b! = 0.$$

To formulate our limiting distribution results we use the notation $\mathcal{N}(0,1)$ for a standard normal distributed r.v. and $\Phi(x)$ for its distribution function. Furthermore we use the notation $\gamma(a,b)$ and $\beta(a,b)$ for a Gamma and Beta distributed r.v. with parameters a and b, respectively, and $\mathrm{NegBin}(m,p)$ for a negative binomial distributed r.v. with parameters m and p.

3.1. **Results for the depth of the largest element.**

Theorem 1. *The random variable D_n, which denotes the depth of the node that contains element n in a random bucket recursive tree of size n with maximal bucket size b, is asymptotically normal distributed with rate of convergence $\mathcal{O}\big(\frac{1}{\sqrt{\log n}}\big)$:*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\Big\{\frac{D_n-\mathbb{E}(D_n)}{\sqrt{\mathbb{V}(D_n)}}\le x\Big\}-\Phi(x)\right|=\mathcal{O}\Big(\frac{1}{\sqrt{\log n}}\Big).$$

Moreover, the expectation $\mathbb{E}(D_n)$ and the variance $\mathbb{V}(D_n)$ of D_n have the following asymptotic expansions:

$$\mathbb{E}(D_n)=\frac{1}{H_b}\log n+\mathcal{O}(1),\qquad \mathbb{V}(D_n)=\frac{H_b^{(2)}}{H_b^3}\log n+\mathcal{O}(1).$$

3.2. **Results for the number of descendants of a specified element.**

Theorem 2. *The exact distribution of the random variable $Y_{n,j}$, which denotes the number of descendants of element j in a random bucket recursive tree of size n with maximal bucket size b, is for $2\le j\le n$ and $1\le m\le n+1-j$ given as follows:*

$$\mathbb{P}\{Y_{n,j}=m\}=\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{\lambda_i+j-2}{j-1}\binom{\ell+m-1}{\ell}\binom{n-m-\ell-1}{j-\ell-2}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{n-1}{j-1}}.$$

Furthermore, it holds $\mathbb{P}\{Y_{n,1}=n\}=1$.

Theorem 3. *The limiting distribution behaviour of the random variable $Y_{n,j}$ is, for $n\to\infty$ and depending on the growth of j, characterized as follows:*

- *The region for $j\ge 2$ fixed. The normalized r.v. $\frac{Y_{n,j}}{n}$ converges in distribution to a r.v. Y_j: $\frac{Y_{n,j}}{n}\xrightarrow{(d)}Y_j$, where Y_j has density $f_j(x)$:*

$$f_j(x)=\sum_{\ell=0}^{b-1}x^\ell(1-x)^{j-\ell-2}(j-1)\binom{j-2}{\ell}\sum_{i=1}^{b}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{\lambda_i+j-2}{j-1}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})},\quad \textit{for }\ 0<x<1.$$

  Thus Y_j is given as a beta distributed random variable: $Y_j\stackrel{(d)}{=}\beta(K_j,j-K_j)$, where the first parameter is given by the random variable $K_j\in\{0,1,\dots,b-1\}$, which is distributed as follows:

$$\mathbb{P}\{K_j=\ell\}=\sum_{i=1}^{b}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{\lambda_i+j-2}{j-1}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})},\quad \textit{for }\ 0\le\ell\le b-1.$$

- *The region for j small: $j\to\infty$ such that $j=o(n)$. The normalized r.v. $\frac{j}{n}Y_{n,j}$ converges in distribution to a r.v. Y: $\frac{j}{n}Y_{n,j}\xrightarrow{(d)}Y$, where Y has density $f(x)$:*

$$f(x)=\sum_{\ell=0}^{b-1}e^{-x}x^\ell\frac{1}{(\ell+1)!H_b}.$$

  Thus Y is given as a gamma distributed random variable: $Y\stackrel{(d)}{=}\gamma(K,1)$, where the first parameter is given by a Zipf distributed random variable $K\in\{1,\dots,b\}$: $\mathbb{P}\{K=i\}=\frac{1}{iH_b}$.
- *The central region for j: $j\to\infty$ such that $j\sim\rho n$, with $0<\rho<1$. The r.v. $Y_{n,j}$ converges in distribution to a discrete r.v. Y_ρ: $Y_{n,j}\xrightarrow{(d)}Y_\rho$, where the probability mass function of Y_ρ is given by*

$$\mathbb{P}\{Y_\rho=m\}=\sum_{\ell=0}^{b-1}\frac{\binom{\ell+m-}{\ell}}{(\ell+1)H_b}\rho^{\ell+1}(1-\rho)^{m-1},\quad \textit{for }\ m=1,2,\dots$$

  Thus $Y_\rho-1$ is given as a negative binomial-distributed random variables, $Y_\rho-1\stackrel{(d)}{=}\mathrm{NegBin}(K,\rho)$, where the first parameter is given by a Zipf distributed random variable $K\in\{1,\dots,b\}$: $\mathbb{P}\{K=i\}=\frac{1}{iH_b}$

- *The region for j large: $j \to \infty$ such that $n - j = o(n)$. The r.v. $Y_{n,j}$ converges to a random variable $\tilde{Y}$, which has all its mass concentrated at 1: $Y_{n,j} \xrightarrow{(d)} \tilde{Y}$, with $\mathbb{P}\{\tilde{Y} = 1\} = 1$.*

3.3. **Results for the node-degree of a specified element.**

Theorem 4. *The limiting distribution behaviour of the random variable $X_{n,j}$, which denotes the out-degree of element j in a random bucket recursive tree of size n with maximal bucket size b, is, for $n \to \infty$ and depending on the growth of j, characterized as follows:*

- *The region for j small: $j = o(n)$. The centered and normalized r.v. $X^*_{n,j}$ is asymptotically Gaussian distributed:*
$$X^*_{n,j} := \frac{X_{n,j} - b(\log n - \log j)}{\sqrt{b(\log n - \log j)}} \xrightarrow{(d)} \mathcal{N}(0,1).$$
- *The central region for j: $j \to \infty$ such that $j \sim \rho n$, with $0 < \rho < 1$. The r.v. $X_{n,j}$ converges in distribution to a discrete r.v. X_ρ: $X_{n,j} \xrightarrow{(d)} X_\rho$, where the probability generating function $p_\rho(v) := \mathbb{E}\big(v^{X_\rho}\big)$ is given by*

$$p_\rho(v) = e^{-b(v-1)\log \rho} \sum_{\ell=0}^{b-1} \frac{\binom{b}{b-\ell-1}}{bH_b\binom{bv-1}{b-\ell-1}} + \sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{b}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)}{bH_b} \rho^{\ell+1}(1-\rho)^{b-1-\ell}.$$

- *The region for j large: $j \to \infty$ such that $n - j = o(n)$. The r.v. $X_{n,j}$ converges to a random variable $\tilde{X}$, which has all its mass concentrated at 0: $X_{n,j} \xrightarrow{(d)} \tilde{X}$, with $\mathbb{P}\{\tilde{X} = 0\} = 1$.*

4. DEPTH OF THE LARGEST ELEMENT

We consider now the random variable D_n, which denotes the depth of element n, i.e., the number of edges lying on the path from the root node to the node that contains element n, in a random bucket recursive tree of size n, i.e., containing n elements. The maximal bucket size is always denoted by b.

In order to study D_n for bucket recursive trees we consider first the corresponding random variable D_n in a bucket increasing tree family with arbitrary weight sequences φ_k and ψ_k. To do this we introduce the bivariate generating function
$$N(z,v) := \sum_{n\geq 1}\sum_{m\geq 0} \mathbb{P}\{D_n = m\} T_n \frac{z^{n-1}}{(n-1)!} v^m. \tag{7}$$
To establish a functional equation for $N(z,v)$ from the formal recursive equation (1) it is convenient to think of specifically bicolored bucket increasing trees, where the elements contained in the nodes are colored as follows: element n in a size-n tree is colored *red* and all elements with a label smaller than n are colored *black*. We are thus interested in the depth of the red element. We consider now a specific bicolored bucket increasing tree T of size n and we assume that the root of T has out-degree $r \geq 1$ and the red element is not captured in the root (thus $n > b$). Then the red element is located in one of the r subtrees of the root node, let us assume it is in the first subtree. Let us consider now these r subtrees: after an order preserving relabelling each of the subtrees $S_1, \dots, S_r$ is itself a bucket increasing tree. The first subtree is again a bicolored tree containing n_1 black elements and one red element, whereas the $n_2, \dots, n_r$ elements in the subtrees $S_2, \dots, S_r$ are all colored black. Since the labels of the $n_1 + n_2 + \cdots + n_r$ black elements are distributed over the black elements in $S_1, \dots, S_r$, each specific r-tuple $S_1, \dots, S_r$ of colored increasing trees appears exactly $\binom{n_1+n_2+\cdots+n_r}{n_1,n_2,\dots,n_r}$ times when starting from all possible bicolored trees of size n. Thus a proper description of this combinatorial decomposition is obtained when introducing univariate and bivariate generating functions, which are exponential in the variable z that marks the black elements. For the bivariate case additionally the variable v counts the depth of the red element.

Since the total weight of bicolored bucket increasing trees with $n-1$ black elements (and thus size n), where the depth of the red element is m, is given by $\mathbb{P}\{D_n = m\}T_n$ their bivariate generating function is exactly given by $N(z,v)$ defined in (7). Of course, the total weight of bucket increasing trees with n elements, where all elements are colored black, is T_n leading to the exponential generating function $T(z)$. Thus the decomposition described above with $r-1$ unicolored trees and one bicolored tree yields to the function $T(z)^{r-1}N(z,v)$. The fact that the depth of the red element in the tree is one more than the depth of the red element in the subtree leads to a factor v. Since the red element can be in the first, second, …, r-th subtree, we additionally get a factor r. Furthermore, according to (1), the event that the root has out-degree r leads to a factor φ_r. Summing over $r \ge 1$ leads to $\sum_{r\ge1} v\varphi_r T(z)^{r-1}N(z,v) = v\varphi'(T(z))N(z,v)$.

Since the elements labelled by $1, 2, \dots, b$ contained in the root node are all colored black (fixing b elements in a labelled object, i.e., the construction $\mathcal{B} = \{1\} \times \{2\} \times \cdots \times \{b\} \times \mathcal{A}$, leads to b differentiations for the corresponding exponential generating functions: $\frac{d^b}{dz^b} = A(z)$), equation (1) leads now to the following differential equation of order b for $N(z,v)$:

$$\frac{\partial^b}{\partial z^b} N(z,v) = v\varphi'(T(z))N(z,v). \tag{8}$$

The case that the element colored red is contained in the root of the tree corresponds of course to the initial conditions, but does not appear (explicitly) in the differential equation itself. The initial conditions of the differential equation (8) are given as follows:

$$\frac{\partial^\ell}{\partial z^\ell} N(z,v)\bigg|_{v=0} = \sum_{m\ge0} \mathbb{P}\{D_{\ell+1} = m\} T_{\ell+1} v^m = T_{\ell+1} = \psi_{\ell+1}, \quad \text{for } 0 \le \ell \le b-1. \tag{9}$$

Now we can specify the sequences $\varphi_k = \frac{(b-1)!b^k}{k!}$ and $\psi_k = (k-1)!$ in above equations and obtain then for bucket recursive trees the following differential equation together with the initial conditions for the bivariate generating function $N(z,v)$:

$$\frac{\partial^b}{\partial z^b} N(z,v) = \frac{vb!}{(1-z)^b} N(z,v), \qquad \frac{\partial^\ell}{\partial z^\ell} N(z,v)\bigg|_{z=0} = \ell!, \quad \text{for } 0 \le \ell \le b-1. \tag{10}$$

This homogeneous differential equation is of Cauchy-Euler-type and can be solved by plugging in $N(z,v) = \frac{1}{(1-z)^{\lambda(v)}}$ with unspecified $\lambda(v)$ into equation (10). This leads then to the indicial equation

$$\lambda(v)^{\overline{b}} - vb! = 0 \qquad \text{or equivalently} \qquad \binom{\lambda(v)+b-1}{b} - v = 0. \tag{11}$$

For our further analysis we require the behaviour of the solutions $\lambda(v)$ in a complex neighbourhood of $v = 1$. For $v = 1$ the corresponding indicial equation $\binom{\lambda+b-1}{b} - 1 = 0$, where we set $\lambda := \lambda(1)$, has been studied in [9] in the context of eigenvalues of a replacement matrix associated to bucket recursive trees. They have shown that all solutions $\lambda_1, \lambda_2, \dots, \lambda_b$ are simple and when arranging the solutions in descending order of real parts it holds

$$1 = \lambda_1 > \Re(\lambda_2) \ge \Re(\lambda_3) \ge \cdots \ge \Re(\lambda_b).$$

An application of the implicit function theorem shows then (see, e.g., [11] for the corresponding treatment of another algebraic equation) that all roots $\lambda_1(v), \lambda_2(v), \dots, \lambda_b(v)$ of (11) are simple in a complex neighbourhood of $v = 1$, i.e., for $|v-1| \le \eta$ with a certain $\eta > 0$, and that the $\lambda_i(v)$ are analytic as functions of v. Since $\lambda_i = \lambda_i(1)$ in above arrangement of the solutions in descending order, it further holds that $\Re(\lambda_1(v)) > \Re(\lambda_i(v))$, for all $2 \le i \le b$, in a complex neighbourhood of $v = 1$. From these considerations follows that the general solution of the differential equation (10) is given by

$$N(z,v) = \sum_{i=1}^{b} \frac{\beta_i(v)}{(1-z)^{\lambda_i(v)}}, \tag{12}$$

with certain functions $\beta_i(v)$, which are specified by the initial conditions. Plugging in the initial conditions given by (10) into the general solution (12) leads then to the following system of b linear equations for the b unknown functions $\beta_i(v)$, $1 \le i \le b$:

$$\sum_{i=1}^{b} \lambda_i(v)^{\overline{\ell}} \beta_i(v) = \ell!, \quad \text{for } 0 \le \ell \le b-1.$$

It can be seen easily by applying Cramer's rule, which expresses the $\beta_i(v)$ as a quotient of determinants involving the solutions of the indicial equation $\lambda_i(v)$ (where the denominator can be transferred into the Vandermonde-determinant), that the functions $\beta_i(v)$ are in a neighbourhood of $v = 1$ analytic functions of v. Moreover, since $N(z,1) = T'(z) = \frac{1}{1-z}$, which follows from the definition, one obtains $\beta_1(1) = 1$ yielding that $\Re(\beta_1(v)) > 0$ in a complex neighbourhood of $v = 1$. We just remark (without showing here details) that by a precise study of the linear system of equations determining $\beta_i(v)$ (analogous to computations carried out in Section 5) one can obtain the following explicit formulæ for the functions $\beta_i(v)$, $1 \le i \le b$:

$$\beta_i(v) = \frac{1-v}{v(H_{\lambda_i(v)+b-1} - H_{\lambda_i(v)-1})(1-\lambda_i(v))}.$$

In order to get an asymptotic expansion of the coefficients of z^n in $N(z,v)$, which holds uniformly in a complex neighbourhood of $v=1$, we can simply apply singularity analysis [3] to the representation (12). This immediately leads to the expansion

$$[z^n]N(z,v) = \frac{\beta_1(v)}{\Gamma(\lambda_1(v))} n^{\lambda_1(v)-1} \cdot \Big(1 + \mathcal{O}\big(n^{\lambda_2 - 1 + \epsilon}\big) + \mathcal{O}\big(n^{-1}\big)\Big),$$

which holds uniformly for $|v-1| \le \eta$, for certain constants $\eta, \epsilon > 0$; recall that $\lambda_2 = \lambda_2(1)$ is a root of the indicial equation (11) for $v=1$ with second largest real part. Thus we obtain the following expansion of the moment generating function $\mathbb{E}(e^{D_n s})$ of the random variable D_n:

$$\mathbb{E}(e^{D_n s}) = [z^{n-1}]N(z,e^s) = e^{U(s)\log n + V(s)} \cdot \big(1 + \mathcal{O}(n^{\lambda_2-1+\epsilon}) + \mathcal{O}(n^{-1})\big), \tag{13}$$

with

$$U(s) = \lambda_1(e^s), \quad \text{and} \quad V(s) = \log\big(\beta_1(e^s)\big) - \log\big(\Gamma(\lambda_1(e^s))\big), \tag{14}$$

which holds uniformly in a complex neighbourhood of $s=0$. A direct application of the so-called quasi-power theorem (see [7]) leads then from (13) to the central limit theorem stated in Theorem 1 together with the following asymptotic expansions of the expectation and the variance of D_n:

$$\mathbb{E}(D_n) = U'(0)\log n + \mathcal{O}(1), \qquad \mathbb{V}(D_n) = U''(0)\log n + \mathcal{O}(1).$$

From (14) we immediately get that $U'(0) = \lambda_1'(1)$ and $U''(0) = \lambda_1''(1) + \lambda_1'(1)$. To compute these values one differentiates the indicial equation (11) w.r.t. v once or twice and evaluates at $v=1$, where one takes into account that $\lambda_1(1) = 1$. One obtains then

$$\lambda_1'(v) = \frac{b!}{\lambda_1(v)^{\bar{b}} \sum_{k=0}^{b-1} \frac{1}{\lambda_1(v)+k}}, \quad \lambda_1''(v) = -\frac{2(\lambda_1'(v))^2 \sum_{0\le i<j\le b-1} \frac{1}{(\lambda_1(v)+i)(\lambda_1(v)+j)}}{\sum_{k=0}^{b-1} \frac{1}{\lambda_1(v)+k}},$$

and thus after some easy manipulations with harmonic numbers:

$$\lambda'(1) = \frac{1}{H_b}, \quad \lambda''(1) = \frac{H_b^{(2)}}{H_b^3} - \frac{1}{H_b}.$$

This completes the proof of Theorem 1.

5. Number of descendants of a specified element

5.1. **The generating functions approach.** We consider now the random variable $Y_{n,j}$, which denotes the number of descendants of element j, i.e., the total number of elements with a label $\ge j$ contained in the subtree rooted with the node containing element j, in a random bucket recursive tree (with maximal bucket size b) of size n.

In order to study $Y_{n,j}$ for bucket recursive trees we consider first the corresponding random variable $Y_{n,j}$ in a bucket increasing tree family with arbitrary weight sequences φ_k and ψ_k. To do this we introduce the trivariate generating function

$$N(z,u,v) := \sum_{k\ge0}\sum_{j\ge1}\sum_{m\ge0} \mathbb{P}\{Y_{k+j,j} = m\} T_{k+j} \frac{z^{j-1}}{(j-1)!}\frac{u^k}{k!} v^m. \tag{15}$$

To establish a functional equation for $N(z,u,v)$ from the formal recursive equation (1) it is now convenient to think of specifically tricolored bucket increasing trees, where the coloring is as follows: exactly one element is colored *red*, all elements with a label smaller than the red element are colored *black*, and all elements with a label larger than the red element are colored *white*. We are then interested in the number of descendants of the red element, i.e., the number of black elements in the subtree rooted with the node containing the red element. Let us consider such a tricolored bucket increasing tree T and assume that the out-degree of the root node of T is $r \ge 1$.

We further assume that the red element of T is not contained in the root node. Then the red element is located in one of the r subtrees of the root of T; let us assume that it is in the r-th subtree. Let us now consider these r subtrees. After order preserving relabellings, each subtree $S_1, \dots, S_r$ is an bucket increasing tree by itself. The first subtree is again a tricolored bucket increasing tree with one red, j_1 black and k_1 white elements, whereas the remaining $r-1$ subtrees are only bicolored in such a way that the elements with the j_i smallest labels (with $2 \le i \le r$ and $0 \le j_i \le |S_i|$) are colored black and the remaining k_i elements in the subtrees are colored white. Then such a specific r-tuple $S_1, \dots, S_r$ of colored bucket increasing trees appears exactly $\binom{j_1+\cdots+j_r}{j_1,\dots,j_r}\binom{k_1+\cdots+k_r}{k_1,\dots,k_r}$ times, where the labels of the $j_1+\cdots+j_r$ black elements and the $k_1+\cdots+k_r$ white elements are distributed over the black and white elements in $S_1, \dots, S_r$ in an order preserving fashion.

Of course, this corresponds to a tricolored bucket increasing tree T of size $|T| = j+k+1$ with $j = j_1+\cdots+j_r$ black elements and $k = k_1+\cdots+k_r$ white elements.

We introduce now generating functions, which are exponential in both variables z and u, where z marks the black elements and u marks the white elements, $f(z,u) = \sum_{j,k\ge 0} f_{j,k}\frac{z^j u^k}{j!k!}$ for sequences $f_{j,k}$ and $f(z,u,v) = \sum_{j,k,m\ge 0} f_{j,k,m}\frac{z^j u^k}{j!k!}v^m$ for sequences $f_{j,k,m}$, where v counts the number of descendants of the red element.

With this setting, the total weight of all suitably tricolored bucket increasing trees with j black and k white elements, where the number of descendants of the red element is exactly m, is given by $\mathbb{P}\{Y_{j+k+1,j+1} = m\}T_{j+k+1}$, and thus its generating function is given by

$$\sum_{k\ge 0}\sum_{j\ge 1}\sum_{m\ge 0}\mathbb{P}\{Y_{k+j,j} = m\}T_{k+j}\frac{z^{j-1}u^k}{(j-1)!k!}v^m = N(z,u,v),$$

whereas the total weight of suitably bicolored bucket increasing trees with j black and k white elements is T_{k+j} and its generating function is given by

$$\sum_{k\ge 0}\sum_{j\ge 0}T_{k+j}\frac{z^j u^k}{j!k!} = T(z+u).$$

The $r-1$ bicolored trees and the tricolored bucket tree lead then to the expression $T(z+u)^{r-1}N(z,u,v)$. Since the red element can be in the first, second, ..., r-th subtree, we additionally get a factor r. Furthermore, the event that the root has out-degree r leads to a factor φ_r. Summing over all $r \ge 1$ leads thus to $\sum_{r\ge 1} r\varphi_r T(z+u)^{r-1}N(z,u,v) = \varphi'(T(z+u))N(z,u,v)$. Since the elements labelled by $1,2,\dots,b$ contained in the root node are all colored black (which again means that b elements in a labelled object are fixed), equation (1) leads thus to the following differential equation of order b for $N(z,u,v)$:

$$\frac{\partial^b}{\partial z^b}N(z,u,v) = \varphi'(T(z+u))N(z,u,v). \tag{16}$$

The cases, where the red element is contained in the root of the tree do not appear explicitly in the differential equation itself, but will be described by the initial conditions. Since $\mathbb{P}\{Y_{n,j} = n+1-j\} = 1$, for $1 \le j \le b$ (if element j is contained in the root node then all elements with a label $\ge j$ are descendants of j), we obtain the following initial conditions, for $0 \le \ell \le b-1$:

$$\begin{aligned}\frac{\partial^\ell}{\partial z^\ell}N(z,u,v)\Big|_{z=0} &= \sum_{k\ge 0}\sum_{m\ge 0}\mathbb{P}\{Y_{k+\ell+1,\ell+1} = m\}T_{k+\ell+1}\frac{u^k}{k!}v^m = \sum_{k\ge 0}T_{k+\ell+1}\frac{u^k}{k!}v^{k+1} \\ &= v\sum_{n\ge \ell+1}T_n\frac{(uv)^{n-\ell-1}}{(n-\ell-1)!} = vT^{(\ell+1)}(uv).\end{aligned} \tag{17}$$

Now we can specify the sequences $\varphi_k = \frac{(b-1)!b^k}{k!}$ and $\psi_k = (k-1)!$ in equations (16) and (17) and obtain then for bucket recursive trees the following differential equation together with the initial conditions for the trivariate generating function $N(z,u,v)$:

$$\frac{\partial^b}{\partial z^b}N(z,u,v) = \frac{b!}{(1-z-u)^b}N(z,u,v), \quad \frac{\partial^\ell}{\partial z^\ell}N(z,u,v)\Big|_{z=0} = \frac{v\ell!}{(1-uv)^{\ell+1}}, \quad \text{for } 0 \le \ell \le b-1. \tag{18}$$

5.2. **The exact distribution.** In order to obtain the exact distribution of the r.v. $Y_{n,j}$ we will give the exact solution of the homogeneous differential equation (18), which is again of Cauchy-Euler-type. Plugging in $N(z,u,v) = \frac{1}{(1-z-u)^{\lambda}}$ with unspecified λ into equation (18) leads to the indicial equation

$$\lambda^{\overline{b}} - b! = 0 \quad \text{or equivalently} \quad \binom{\lambda+b-1}{b} - 1 = 0. \tag{19}$$

As mentioned in Section 4 this equation has been studied in [9], where it has been shown that all solutions $\lambda_1, \lambda_2, \dots, \lambda_b$ are simple and when arranging them in descending order of real parts it holds $1 = \lambda_1 > \Re(\lambda_2) \geq \Re(\lambda_3) \geq \dots \geq \Re(\lambda_b)$. Thus the general solution of (18) is given by

$$N(z,u,v) = \sum_{i=1}^{b} \frac{\beta_i(u,v)}{(1-z-u)^{\lambda_i}}, \tag{20}$$

with certain functions $\beta_i(u,v)$, which are specified by the initial conditions as given in (18). When plugging in the initial conditions into (20) this leads to the following system of linear equations for the unknown functions $\beta_i(u,v)$, $1 \leq i \leq b$:

$$\sum_{i=1}^{b} \frac{\lambda_i^{\overline{\ell}} \beta_i(u,v)}{(1-u)^{\lambda_i+\ell}} = \frac{v\ell!}{(1-uv)^{\ell+1}}, \quad \text{for } 0 \leq \ell \leq b-1.$$

Using the abbreviations

$$\gamma_i := \gamma_i(u,v) := \frac{\beta_i(u,v)}{(1-u)^{\lambda_i}}, \quad \text{and} \quad s_\ell := s_\ell(u,v) := \frac{(1-u)^{\ell} v}{(1-uv)^{\ell+1}}, \tag{21}$$

we obtain the following system of linear equations for the unknown γ_i, $1 \leq i \leq b$:

$$\sum_{i=1}^{b} \binom{\lambda_i+\ell-1}{\ell} \gamma_i = s_\ell, \quad \text{for } 0 \leq \ell \leq b-1. \tag{22}$$

To get explicit solutions for the γ_i we apply Cramer's rule to (22) and write the solutions γ_i, $1 \leq i \leq b$, as a quotient of determinants:

$$\gamma_i = \begin{vmatrix} \lambda_1^{\overline{0}} & \dots & \lambda_{i-1}^{\overline{0}} & 0!s_0 & \lambda_{i+1}^{\overline{0}} & \dots & \lambda_b^{\overline{0}} \\ \lambda_1^{\overline{1}} & \dots & \lambda_{i-1}^{\overline{1}} & 1!s_1 & \lambda_{i+1}^{\overline{1}} & \dots & \lambda_b^{\overline{1}} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \lambda_1^{\overline{b-1}} & \dots & \lambda_{i-1}^{\overline{b-1}} & (b-1)!s_{b-1} & \lambda_{i+1}^{\overline{b-1}} & \dots & \lambda_b^{\overline{b-1}} \end{vmatrix} \cdot \begin{vmatrix} \lambda_1^{\overline{0}} & \dots & \lambda_i^{\overline{0}} & \dots & \lambda_b^{\overline{0}} \\ \lambda_1^{\overline{1}} & \dots & \lambda_i^{\overline{1}} & \dots & \lambda_b^{\overline{1}} \\ \vdots & & \vdots & & \vdots \\ \lambda_1^{\overline{b-1}} & \dots & \lambda_i^{\overline{b-1}} & \dots & \lambda_b^{\overline{b-1}} \end{vmatrix}^{-1}. \tag{23}$$

Using elementary relations between the Stirling numbers and the factorials (see, e.g., [6]) and the abbreviation

$$c_\ell := \sum_{k=0}^{\ell} \left\{ \begin{matrix} \ell \\ k \end{matrix} \right\} (-1)^{\ell-k} k! s_k, \quad \text{for } 0 \leq \ell \leq b-1, \tag{24}$$

we obtain from (23) after elementary transformations the following representation of the solutions γ_i, $1 \leq i \leq b$:

$$\gamma_i = Q_i^{[1]} \cdot Q_i^{[2]},$$

where $Q_i^{[1]}$, $Q_i^{[2]}$ are the following quotients of determinants:

$$Q_i^{[1]} := \begin{vmatrix} \lambda_1^{0} & \dots & \lambda_{i-1}^{0} & \lambda_{i+1}^{0} & \dots & \lambda_b^{0} \\ \lambda_1^{1} & \dots & \lambda_{i-1}^{1} & \lambda_{i+1}^{1} & \dots & \lambda_b^{1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_1^{b-2} & \dots & \lambda_{i-1}^{b-2} & \lambda_{i+1}^{b-2} & \dots & \lambda_b^{b-2} \end{vmatrix} \cdot \begin{vmatrix} \lambda_1^{0} & \dots & \lambda_i^{0} & \dots & \lambda_b^{0} \\ \lambda_1^{1} & \dots & \lambda_i^{1} & \dots & \lambda_b^{1} \\ \vdots & & \vdots & & \vdots \\ \lambda_1^{b-1} & \dots & \lambda_i^{b-1} & \dots & \lambda_b^{b-1} \end{vmatrix}^{-1}$$

$$Q_i^{[2]} := \begin{vmatrix} \lambda_1^{0} & \dots & \lambda_{i-1}^{0} & c_0 & \lambda_{i+1}^{0} & \dots & \lambda_b^{0} \\ \lambda_1^{1} & \dots & \lambda_{i-1}^{1} & c_1 & \lambda_{i+1}^{1} & \dots & \lambda_b^{1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \lambda_1^{b-1} & \dots & \lambda_{i-1}^{b-1} & c_{b-1} & \lambda_{i+1}^{b-1} & \dots & \lambda_b^{b-1} \end{vmatrix} \cdot \begin{vmatrix} \lambda_1^{0} & \dots & \lambda_{i-1}^{0} & \lambda_{i+1}^{0} & \dots & \lambda_b^{0} \\ \lambda_1^{1} & \dots & \lambda_{i-1}^{1} & \lambda_{i+1}^{1} & \dots & \lambda_b^{1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_1^{b-2} & \dots & \lambda_{i-1}^{b-2} & \lambda_{i+1}^{b-2} & \dots & \lambda_b^{b-2} \end{vmatrix}^{-1}.$$

Since $Q_i^{[1]}$ is a quotient of Vandermonde-determinants it is evaluated easily:

$$Q_i^{[1]} = \frac{\prod_{1\le p<q\le b, i\neq p, i\neq q}(\lambda_q-\lambda_p)}{\prod_{1\le p<q\le b}(\lambda_q-\lambda_p)} = \frac{(-1)^{b-i}}{\prod_{1\le p\le b, i\neq p}(\lambda_i-\lambda_p)}.$$

When expanding the i-th column in the numerator of $Q_i^{[2]}$ we obtain

$$Q_i^{[2]} = \sum_{\ell=0}^{b-1} c_\ell (-1)^{\ell+1+i} q_{i,\ell},$$

with

$$q_{i,\ell} := \begin{vmatrix} \lambda_1^0 & \dots & \lambda_{i-1}^0 & \lambda_{i+1}^0 & \dots & \lambda_b^0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_1^{\ell-1} & \dots & \lambda_{i-1}^{\ell-1} & \lambda_{i+1}^{\ell-1} & \dots & \lambda_b^{\ell-1} \\ \lambda_1^{\ell+1} & \dots & \lambda_{i-1}^{\ell+1} & \lambda_{i+1}^{\ell+1} & \dots & \lambda_b^{\ell+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_1^{b-1} & \dots & \lambda_{i-1}^{b-1} & \lambda_{i+1}^{b-1} & \dots & \lambda_b^{b-1} \end{vmatrix} \cdot \begin{vmatrix} \lambda_1^0 & \dots & \lambda_{i-1}^0 & \lambda_{i+1}^0 & \dots & \lambda_b^0 \\ \lambda_1^1 & \dots & \lambda_{i-1}^1 & \lambda_{i+1}^1 & \dots & \lambda_b^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \lambda_1^{b-2} & \dots & \lambda_{i-1}^{b-2} & \lambda_{i+1}^{b-2} & \dots & \lambda_b^{b-2} \end{vmatrix}^{-1}.$$

These considerations lead to the following representation of γ_i:

$$\gamma_i = \frac{1}{\prod_{1\le p\le b, i\neq p}(\lambda_i-\lambda_p)} \sum_{\ell=0}^{b-1} c_\ell (-1)^{b-1-\ell} q_{i,\ell}, \quad \text{for } 1\le i\le b. \tag{25}$$

Next we are going to simplify the expressions appearing in (25). The quotient of determinants $q_{i,\ell}$ has the following representation (this can be obtained, e.g., when writing $q_{i,\ell}$ as a Schur-function and applying the Jacobi-Trudi-identity, see, e.g., [12]):

$$q_{i,\ell} = e_{b-1-\ell}(\lambda_1,\dots,\lambda_{i-1},\lambda_{i+1},\dots,\lambda_b), \quad \text{for } 0\le \ell\le b-1, \tag{26}$$

where $e_r(x_1,\dots,x_n)$, denotes the r-th elementary symmetric polynomial with variables $x_1,\dots,x_n$ which is defined as $e_0=1$, and $e_r=\sum_{i_1<i_2<\dots<i_r} x_{i_1}x_{i_2}\cdots x_{i_r}$, for integers $r\ge 1$.

Furthermore by using the factorization $\lambda^{\overline{b}} - b! = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_b)$ we obtain the identity

$$(\lambda-\lambda_i)\sum_{\ell=0}^{b-1}(-1)^{b-1-\ell} e_{b-1-\ell}(\lambda_1,\dots,\lambda_{i-1},\lambda_{i+1},\dots,\lambda_b)\lambda^\ell = \lambda^{\overline{b}} - b!.$$

This leads to the following evaluation of the elementary symmetric polynomials appearing in (26):

$$\begin{aligned}(-1)^{b-1-\ell} e_{b-1-\ell}(\lambda_1,\dots,\lambda_{i-1},\lambda_{i+1},\dots,\lambda_b) &= [\lambda^\ell]\frac{\lambda^{\overline{b}}-b!}{\lambda-\lambda_i} = [\lambda^\ell]\Big(\frac{\lambda^{\overline{b}}-\lambda_i^{\overline{b}}}{\lambda-\lambda_i} + \frac{\lambda_i^{\overline{b}}-b!}{\lambda-\lambda_i}\Big) = [\lambda^\ell]\frac{\lambda^{\overline{b}}-\lambda_i^{\overline{b}}}{\lambda-\lambda_i} \\ &= [\lambda^\ell]\frac{\sum_{k=1}^b {b \brack k}(\lambda^k-\lambda_i^k)}{\lambda-\lambda_i} = [\lambda^\ell]\sum_{\ell=0}^{b-1}\lambda^\ell \sum_{k=\ell+1}^{b}{b \brack k}\lambda_i^{k-1-\ell} = \sum_{k=\ell+1}^{b}{b \brack k}\lambda_i^{k-1-\ell},\end{aligned}$$

and thus to the formula

$$q_{i,\ell} = (-1)^{b-1-\ell}\sum_{k=\ell+1}^{b}{b \brack k}\lambda_i^{k-1-\ell}. \tag{27}$$

Furthermore when considering the derivative of the indicial polynomial $P(\lambda) := \lambda^{\overline{b}} - b! = \prod_{p=1}^b(\lambda-\lambda_p)$ w.r.t. λ and evaluating at λ_i we obtain the identity

$$P'(\lambda_i) = \lambda_i^{\overline{b}}\sum_{k=0}^{b-1}\frac{1}{\lambda_i+k} = b!(H_{\lambda_i+b-1}-H_{\lambda_i-1}) = \prod_{1\le p\le b, i\neq p}(\lambda_i-\lambda_p). \tag{28}$$

Plugging in equations (24), (27) and (28) into (25) we obtain after easy manipulations the following formula for γ_i:

$$\gamma_i = \frac{1}{b!(H_{\lambda_i+b-1}-H_{\lambda_i-1})}\sum_{r=0}^{b-1} r! s_r A_{b,r}(\lambda_i), \tag{29}$$

where the function $A_{b,r}(x)$ is defined as follows:

$$A_{b,r}(x) := \sum_{\ell=r}^{b-1} \left\{ \begin{matrix} \ell \\ r \end{matrix} \right\} (-1)^{\ell-r} \sum_{k=\ell+1}^{b} \begin{bmatrix} b \\ k \end{bmatrix} x^{k-1-\ell}.$$

By applying basic identities for Stirling numbers, which can be found, e.g., in [6], one obtains that the function $A_{b,r}(x)$ satisfies the recurrence

$$A_{b,r}(x) = (b-1+x)A_{b-1,r}(x), \quad \text{for } b-1 > r,$$

with initial value $A_{r+1,r}(x) = 1$. Thus when iterating this equation we get the following simple expression for $A_{b,r}(x)$:

$$A_{b,r}(x) = (b-1+x)^{\underline{b-1-r}}. \tag{30}$$

Combining (29) and (30) we obtain thus the following exact formulæ for the unknown functions γ_i:

$$\gamma_i = \sum_{r=0}^{b-1} s_r \frac{\binom{\lambda_i+b-1}{b-r-1}}{\binom{b}{r}(b-r)(H_{\lambda_i+b-1} - H_{\lambda_i-1})}, \quad \text{for } 1 \le i \le b. \tag{31}$$

Together with (20) and (21) we obtain then the following exact expression for the trivariate generating function $N(z,u,v)$:

$$N(z,u,v) = \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \frac{\eta_{i,\ell}\,(1-u)^{\lambda_i+\ell} v}{(1-uv)^{\ell+1}(1-z-u)^{\lambda_i}}, \tag{32}$$

with constants

$$\eta_{i,\ell} := \frac{\binom{\lambda_i+b-1}{b-\ell-1}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1} - H_{\lambda_i-1})}, \tag{33}$$

where $\lambda_1, \dots, \lambda_b$ are the roots of the indicial equation $\lambda^{\overline{b}} - b! = 0$ arranged in decreasing order of their real parts. We remark that due to $\lambda_1 = 1$ we obtain in particular $\eta_{1,\ell} = \frac{1}{(\ell+1)H_b}$.

Thus, using the definition (15), we immediately obtain by extracting coefficients from (32) an exact formula for the probability that the number $Y_{n,j}$ of descendants of element j in a bucket recursive tree of size n is equal to m:

$$\begin{aligned}
\mathbb{P}\{Y_{n,j} = m\} &= \frac{(j-1)!(n-j)!}{T_n} [z^{j-1}u^{n-j}v^m] N(z,u,v) \\
&= \frac{1}{\binom{n-1}{j-1}} \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \eta_{i,\ell} [z^{j-1}u^{n-j}v^m] \frac{(1-u)^{\lambda_i+\ell} v}{(1-uv)^{\ell+1}(1-z-u)^{\lambda_i}} \\
&= \frac{1}{\binom{n-1}{j-1}} \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \eta_{i,\ell} \binom{j-1+\lambda_i-1}{j-1} [u^{n-j}v^m] \frac{v}{(1-u)^{j-1-\ell}(1-uv)^{\ell+1}} \\
&= \frac{1}{\binom{n-1}{j-1}} \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \eta_{i,\ell} \binom{j-1+\lambda_i-1}{j-1} \binom{m-1+\ell}{\ell} [u^{n-j-m+1}] \frac{1}{(1-u)^{j-1-\ell}} \\
&= \frac{1}{\binom{n-1}{j-1}} \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \eta_{i,\ell} \binom{j-1+\lambda_i-1}{j-1} \binom{n-1+\ell}{\ell} \binom{n-m-\ell-1}{j-2-\ell} \\
&= \sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{\lambda_i+j-2}{j-1}\binom{\ell+m-1}{\ell}\binom{n-m-\ell-1}{j-\ell-2}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{n-1}{j-1}}, \quad \text{for } j \ge 2 \quad \text{and} \quad 1 \le m \le n+1-j.
\end{aligned} \tag{34}$$

Of course, it also holds $\mathbb{P}\{Y_{n,1} = n\} = 1$ and this completes the proof of Theorem 2.

5.3. **Limiting distribution results.** An advantage of the approach presented leading to the exact distribution of the r.v. $Y_{n,j}$ under consideration is that by using these exact results the asymptotic behaviour of $Y_{n,j}$ can be described in a quite precise manner, where $j = j(n)$, with $1 \le j \le n$, is possibly growing in n. Of course, the asymptotic behaviour of $Y_{n,j}$ is dependant on the "growth function" $j(n)$ and leads to four regions, where different limiting distributions are occuring. Since the asymptotic results are essentially following from (34) by applying Stirling's asymptotic formula for the factorials:

$$n! = n^n e^{-n}\sqrt{2\pi n}\big(1 + \mathcal{O}(n^{-1})\big), \tag{35}$$

we will not carry out here every step of these straightforward computations.

The region for j fixed. Stirling's formula (35) leads for fixed $j \ge 2$ from the exact formula (34) immediately to the following asymptotic evaluation:

$$\mathbb{P}\{Y_{n,j} = m\} = \frac{1}{n}\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\eta_{i,\ell}\binom{\lambda_i + j - 2}{j-1}(j-1)\binom{j-2}{\ell}\big(\frac{m}{n}\big)^{\ell}\big(1-\frac{m}{n}\big)^{j-\ell-2}$$
$$\times\Big(1 + \mathcal{O}\big(\frac{1}{m}\big) + \mathcal{O}\big(\frac{1}{n-m}\big)\Big).$$

Thus, setting $x := \frac{m}{n}$, we obtain for fixed $j \ge 2$ the local expansion

$$\frac{\mathbb{P}\{x \le \frac{Y_{n,j}}{n} < x + \frac{1}{n}\}}{\frac{1}{n}} = f_j(x)\Big(1 + \mathcal{O}\big(\frac{1}{xn}\big) + \mathcal{O}\big(\frac{1}{(1-x)n}\big)\Big),$$

with

$$f_j(x) := \sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\eta_{i,\ell}(j-1)\binom{\lambda_i + j - 2}{j-1}\binom{j-2}{\ell}x^{\ell}(1-x)^{j-\ell-2}$$
$$= \sum_{\ell=0}^{b-1}x^{\ell}(1-x)^{j-\ell-2}(j-1)\binom{j-2}{\ell}\sum_{i=1}^{b}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{\lambda_i+j-2}{j-1}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1} - H_{\lambda_i-1})}.$$

This implies that one obtains for $n^{-\frac{1}{2}} \le x \le 1 - n^{-\frac{1}{2}}$ the uniform local approximation

$$\frac{\mathbb{P}\{x \le \frac{Y_{n,j}}{n} < x + \frac{1}{n}\}}{\frac{1}{n}} = f_j(x)\big(1 + \mathcal{O}(n^{-\frac{1}{2}})\big),$$

which also shows for the region j fixed the corresponding limiting distribution result in Theorem 3.

The region for j small: $j \to \infty$ such that $j = o(n)$. For this region Stirling's formula (35) gives the asymptotic expansion

$$\mathbb{P}\{Y_{n,j} = m\} = \sum_{\ell=0}^{b-1}\frac{\binom{\ell+m-1}{\ell}\binom{n-m-\ell-1}{j-\ell-1}}{\binom{n-1}{j-1}}\eta_{1,\ell}\big(1 + \mathcal{O}(j^{-1}) + \mathcal{O}(j^{\Re\lambda_2 - 1})\big)$$
$$= \sum_{\ell=0}^{b-1}\frac{\binom{\ell+m-1}{\ell}\binom{n-m-\ell-1}{j-\ell-2}}{\binom{n-1}{j-1}(\ell+1)H_b}\big(1 + \mathcal{O}(j^{-1}) + \mathcal{O}(j^{\Re\lambda_2 - 1})\big)$$
$$= \frac{j}{n}\sum_{\ell=0}^{b-1}\frac{(n-m)!(n-j)!}{n!(n-m-j)!}\big(\frac{mj}{n}\big)^{\ell}\frac{1}{(l+1)!H_b}$$
$$\times\big(1 + \mathcal{O}(j^{-1}) + \mathcal{O}(j^{\Re\lambda_2 - 1}) + \mathcal{O}(m^{-1}) + \mathcal{O}(mn^{-1}) + \mathcal{O}(jn^{-1})\big)$$
$$= \frac{j}{n}\sum_{\ell=0}^{b-1}e^{-\frac{jm}{n}}\big(\frac{jm}{n}\big)^{\ell}\frac{1}{(\ell+1)!H_b}$$
$$\times\big(1 + \mathcal{O}(j^{-1}) + \mathcal{O}(j^{\Re\lambda_2 - 1}) + \mathcal{O}(m^{-1}) + \mathcal{O}(mn^{-1}) + \mathcal{O}(jn^{-1}) + \mathcal{O}(jm^2n^{-2}) + \mathcal{O}(j^2mn^{-2})\big).$$

Setting $x := \frac{jm}{n}$ we obtain for $j \to \infty$ with $j = o(n)$ the local expansion

$$\frac{\mathbb{P}\{x \le \frac{j}{n} Y_{n,j} < x + \frac{j}{n}\}}{\frac{j}{n}} = f(x) \\ \times \Big(1 + \mathcal{O}\big(j^{-1}\big) + \mathcal{O}\big(j^{\Re\lambda_2 - 1}\big) + \mathcal{O}\big(\frac{j}{n}\big) + \mathcal{O}\big(\frac{x}{j}\big) + \mathcal{O}\big(\frac{x^2}{j}\big) + \mathcal{O}\big(\frac{jx}{n}\big) + \mathcal{O}\big(\frac{j}{xn}\big)\Big),$$

with

$$f(x) := \sum_{\ell=0}^{b-1} e^{-x} x^{\ell} \frac{1}{(\ell+1)! H_b}.$$

For $\sqrt{\frac{j}{n}} \le x \le \min\big(j^{\frac{1}{4}}, \sqrt{\frac{n}{j}}\big)$ this gives the uniform local approximation

$$\frac{\mathbb{P}\{x \le \frac{j}{n} Y_{n,j} < x + \frac{j}{n}\}}{\frac{j}{n}} = f(x)\Big(1 + \mathcal{O}\big(j^{-\frac{1}{2}}\big) + \mathcal{O}\big(j^{\Re\lambda_2 - 1}\big) + \mathcal{O}\big(\sqrt{\frac{j}{n}}\big)\Big),$$

which leads for the region $j \to \infty$ such that $j = o(n)$ to the corresponding limiting distribution result in Theorem 3.

The central region for j: $j \to \infty$ such that $j \sim \rho n$, with $0 < \rho < 1$. For $\epsilon \le \frac{j}{n} \le 1 - \epsilon$, with an arbitrary $\epsilon > 0$, we obtain with (35) the asymptotic expansion

$$\mathbb{P}\{Y_{n,j} = m\} = \sum_{\ell=0}^{b-1} \frac{\binom{\ell+m-1}{\ell}\binom{n-m-\ell-1}{j-\ell-2}}{\binom{n-1}{j-1}(\ell+1)H_b}\big(1 + \mathcal{O}(n^{-1}) + \mathcal{O}(n^{\Re\lambda_2 - 1})\big),$$

which leads for every $m \ge 1$ fixed to the following local approximation:

$$\mathbb{P}\{Y_{n,j} = m\} = \sum_{\ell=0}^{b-1} \frac{\binom{\ell+m-1}{\ell}}{(\ell+1)H_b}\big(\frac{j}{n}\big)^{\ell+1}\big(1 - \frac{j}{n}\big)^{m-1}\big(1 + \mathcal{O}(n^{-1})\big).$$

Thus, for $\frac{j}{n} \sim \rho$ with $0 \le \rho < 1$, one obtains that for every $m \ge 1$:

$$\mathbb{P}\{Y_{n,j} = m\} \to \sum_{\ell=0}^{b-1} \frac{\binom{\ell+m-1}{\ell}}{(\ell+1)H_b}\rho^{\ell+1}(1-\rho)^{m-1},$$

which shows the discrete limit law for this region presented in Theorem 3.

The region for j large: $j \to \infty$ such that $n - j = o(n)$. For $n - j = o(n)$ equation (35) leads to the asymptotic expansion

$$\mathbb{P}\{Y_{n,j} = 1\} = \sum_{\ell=0}^{b-1} \frac{\binom{n-\ell-2}{j-\ell-2}}{\binom{n-1}{j-1}(\ell+1)H_b}\big(1 + \mathcal{O}(n^{-1}) + \mathcal{O}(n^{\Re\lambda_2 - 1})\big).$$

Since we further obtain for this region the expansion

$$\frac{\binom{n-\ell-2}{j-\ell-2}}{\binom{n-1}{j-1}} = \big(\frac{j}{n}\big)^{\ell+1}\big(1 + \mathcal{O}(n^{-1})\big) = 1 + \mathcal{O}\big(\frac{n-j}{n}\big) = 1 + o(1),$$

we have shown that for the region $j \to \infty$ such that $j = o(n)$ it holds

$$\mathbb{P}\{Y_{n,j} = 1\} \to 1,$$

and this proves the degenerate limit law in the corresponding part of Theorem 3.

6. Node-degree of a specified element

6.1. **The generating functions approach.** Now we consider the random variable $X_{n,j}$, which denotes the out-degree of element j, i.e., the out-degree of the node containing element j, in a random bucket recursive tree (with maximal bucket size b) of size n. Again, in order to study $X_{n,j}$ for bucket recursive trees we consider first the corresponding random variable $X_{n,j}$ in a bucket increasing tree family with arbitrary weight sequences φ_k and ψ_k and introduce the trivariate generating function

$$N(z,u,v) := \sum_{k\ge 0}\sum_{j\ge 1}\sum_{m\ge 0} \mathbb{P}\{X_{k+j,j}=m\} T_{k+j} \frac{z^{j-1}}{(j-1)!}\frac{u^k}{k!}v^m. \tag{36}$$

It can be verified easily that the arguments in Subsection 5.1 for the r.v. $Y_{n,j}$ leading to the differential equation (16) for the corresponding generating function (15) also work for $X_{n,j}$ and the generating function (36). Thus the trivariate generating function defined by (36) also satisfies the differential equation (16) (but, of course, with different initial conditions):

$$\frac{\partial^b}{\partial z^b} N(z,u,v) = \varphi'(T(z+u)) N(z,u,v). \tag{37}$$

We remark that one could also argue that $N(z,u,v)$ defined by (36) has to satisfy (37), since $X_{n,j}$ and $Y_{n,j}$ satisfy, apart from different initial values, the same recurrence, which is obtained from the natural decomposition (3) of bucket increasing trees.

As we will see, in order to obtain the initial conditions for the generating function $N(z,u,v)$ we have to study the degree distribution of the root of a random bucket increasing tree with n elements. Let R_n denote the random variable counting the out-degree of the root and $R(u,v)$ the following bivariate generating function:

$$R(u,v) := \sum_{n\ge 1}\sum_{m\ge 0} \mathbb{P}\{R_n=m\} T_n \frac{u^n}{n!} v^m. \tag{38}$$

By using the combinatorial decomposition (3) of bucket increasing trees one easily obtains that $R(u,v)$ satisfies the following differential equation:

$$\frac{\partial^b}{\partial u^b} R(u,v) = \sum_{k\ge 0} \varphi_k v^k (T(u))^k = \varphi(vT(u)), \tag{39}$$

with initial conditions

$$R(0,v)=0, \quad \text{and} \quad \frac{\partial^\ell}{\partial u^\ell} R(u,v)\Big|_{u=0} = \sum_{m\ge 0} \mathbb{P}\{R_\ell = m\} T_\ell v^m = T_\ell, \quad \text{for } 1\le \ell \le b-1.$$

We further use that $R_n \overset{(d)}{=} X_{n,j}$, for $1\le j\le b$ (elements $1,2,\dots,b$ are all contained in the root node), which gives the following description of the initial conditions corresponding to (37):

$$\begin{aligned}\frac{\partial^\ell}{\partial z^\ell} N(z,u,v)\Big|_{z=0} &= \sum_{k\ge 0}\sum_{m\ge 0} \mathbb{P}\{X_{k+\ell+1,\ell+1}=m\} T_{k+\ell+1}\frac{u^k}{k!}v^m \\ &= \sum_{k\ge 0}\sum_{m\ge 0} \mathbb{P}\{R_{k+\ell+1}=m\} T_{k+\ell+1}\frac{u^k}{k!}v^m = \frac{\partial^{\ell+1}}{\partial u^{\ell+1}} R(u,v), \quad \text{for } 0\le \ell\le b-1.\end{aligned}$$

Now we specify our findings for the instance of bucket recursive trees and obtain that $N(z,u,v)$ satisfies the following differential equation together with the initial conditions:

$$\frac{\partial^b}{\partial z^b} N(z,u,v) = \frac{b!}{(1-z-u)^b} N(z,u,v), \quad \frac{\partial^\ell}{\partial z^\ell} N(z,u,v)\Big|_{z=0} = \frac{\partial^{\ell+1}}{\partial u^{\ell+1}} R(u,v), \quad \text{for } 0\le \ell\le b-1. \tag{40}$$

Moreover, the function $R(u,v)$ satisfies the differential equation

$$\frac{\partial^b}{\partial u^b} R(u,v) = \frac{(b-1)!}{(1-u)^{bv}}, \tag{41}$$

with initial conditions

$$R(0,v)=0, \quad \text{and} \quad \frac{\partial^\ell}{\partial u^\ell} R(u,v)\Big|_{u=0} = (\ell-1)!, \quad \text{for } 1 \le \ell \le b-1. \tag{42}$$

Of course, equation (41) can be solved by integration and leads after adapting to the initial conditions (42) to the following explicit solution:

$$R(u,v) = \frac{1}{b\binom{bv-1}{b}(1-u)^{bv-b}} + \frac{1}{b}\sum_{k=1}^{b-1}\binom{b}{k}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)u^k - \frac{1}{b\binom{bv-1}{b}}. \tag{43}$$

6.2. **The exact distribution.** Since the differential equation (40) coincides apart from the initial conditions with equation (18) we can proceed as in Subsection 5.2 to obtain an exact solution of $N(z,u,v)$. We obtain thus

$$N(z,u,v) = \sum_{i=1}^{b} \frac{\beta_i(u,v)}{(1-z-u)^{\lambda_i}} \tag{44}$$

for the general solution of equation (40) (with unspecified functions $\beta_i(u,v)$). Adapting to the initial conditions leads to the following system of linear equations for the functions $\beta_i(u,v)$, for $1 \le i \le b$:

$$\sum_{i=1}^{b}\binom{\lambda_i+\ell-1}{\ell}\gamma_i(u,v) = s_\ell(u,v), \quad \text{for } 0 \le \ell \le b-1, \tag{45}$$

where we used the abbreviations

$$\gamma_i(u,v) := \frac{\beta_i(u,v)}{(1-u)^{\lambda_i}}, \quad \text{and} \quad s_\ell(u,v) := \frac{(1-u)^\ell}{\ell!}\frac{\partial^{\ell+1}}{\partial u^{\ell+1}}R(u,v).$$

This system of linear equations (45) has been solved in Subsection 5.2 leading to the solutions

$$\gamma_i(u,v) = \sum_{\ell=0}^{b-1} s_\ell(u,v) \frac{\binom{\lambda_i+b-1}{b-\ell-1}}{\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})}, \quad \text{for } 1 \le i \le b.$$

Therefore we obtain the following solution of $N(z,u,v)$ defined by (36):

$$\begin{aligned} N(z,u,v) &= \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}}{\ell!\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})}\frac{(1-u)^{\lambda_i+\ell}}{(1-z-u)^{\lambda_i}}\frac{\partial^{\ell+1}}{\partial u^{\ell+1}}R(u,v) \\ &= \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}}{\ell!\binom{b}{\ell}(b-\ell)(H_{\lambda_i+b-1}-H_{\lambda_i-1})}\frac{(1-u)^{\lambda_i+\ell}}{(1-z-u)^{\lambda_i}} \\ &\quad \times \left(\frac{\binom{b}{\ell+1}(\ell+1)!}{b\binom{bv-1}{b-\ell-1}(1-u)^{bv-b+\ell+1}} + \frac{\binom{b}{\ell+1}(\ell+1)!}{b}\sum_{k=\ell+1}^{b-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)u^{k-\ell-1}\right). \end{aligned} \tag{46}$$

Extracting coefficients from (46) leads thus directly to the following exact solution of the probability generating function $p_{n,j}(v) := \sum_{m\ge 0}\mathbb{P}\{X_{n,j}=m\}v^m = \frac{1}{\binom{n-1}{j-1}}[z^{j-1}u^{n-j}]N(z,u,v)$ of the out-degree $X_{n,j}$ of element j in a random bucket recursive tree of size n:

$$\begin{aligned} p_{n,j}(v) &= \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n+bv-b-1}{n-j}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{n-1}{j-1}\binom{bv-1}{b-\ell-1}} \\ &\quad + \sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{n-1}{j-1}}. \end{aligned} \tag{47}$$

6.3. **Limiting distribution results.** Again, since we have a detailed description of the behaviour of the r.v. $X_{n,j}$ (now via the probability generating function $p_{n,j}(v)$) we are also able to give a quite detailed description of the limiting behaviour of $X_{n,j}$ for all regions $1 \le j \le n$ depending on the growth of $j = j(n)$. Essentially we also only require Stirling's formula (35) for the factorials together with asymptotic equivalents and bounds for the harmonic numbers H_n and $H_n^{(2)}$ of first and second order. Since the asymptotic considerations required to prove our limiting distribution results are essentially straightforward, but nevertheless lengthy when figured out in detail, we will here only sketch these computations.

The region for j small: $j = o(n)$. To obtain a limiting distribution result for this region we first compute exact formulæ for the expectation $\mathbb{E}(X_{n,j}) = p'_{n,j}(1)$ and the variance $\mathbb{V}(X_{n,j}) = p''_{n,j}(1) + p'_{n,j}(1) - (p'_{n,j}(1))^2$. They are given as follows:

$$\begin{aligned}
\mathbb{E}(X_{n,j}) &= b(H_{n-1} - H_{j-1}) - \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-\ell-1}}(H_{b-1} - H_\ell) \\
&\quad + \sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-k}\binom{n-1}{j-1}}(H_{b-1} - H_{k-1}), \\
\mathbb{V}(X_{n,j}) &= b(H_{n-1} - H_{j-1}) - b^2(H_{n-1}^{(2)} - H_{j-1}^{(2)}) \\
&\quad - 2b(H_{n-1} - H_{j-1})\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-k}\binom{n-1}{j-1}}(H_{b-1} - H_{k-1}) \\
&\quad + \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-\ell-1}}\big(b((H_{b-1} - H_\ell)^2 + (H_{b-1}^{(2)} - H_\ell^{(2)})) - (H_{b-1} - H_\ell)\big) \\
&\quad - \sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-k}\binom{n-1}{j-1}} \\
&\qquad \times \big(b((H_{b-1} - H_{k-1})^2 + (H_{b-1}^{(2)} - H_{k-1}^{(2)})) - (H_{b-1} + H_{k-1})\big) \\
&\quad - \Bigg(\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-k}\binom{n-1}{j-1}}(H_{b-1} - H_{k-1}) \\
&\qquad - \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-\ell-1}}(H_{b-1} - H_\ell)\Bigg)^2.
\end{aligned}$$

By using (35) and the asymptotic expansions

$$H_n = \log n + \gamma + \mathcal{O}(n^{-1}), \quad \text{and} \quad H_n^{(2)} = \frac{\pi^2}{6} + \mathcal{O}(n^{-1}), \tag{48}$$

one easily obtains the expansions

$$\mathbb{E}(X_{n,j}) = b(\log n - \log j) + \mathcal{O}(1), \quad \text{and} \quad \mathbb{V}(X_{n,j}) = b(\log n - \log j) + \mathcal{O}(1),$$

where the bound on the remainder term holds uniformly for all $1 \le j \le n$ and $n \ge 1$. In other words there exist constants c_b and d_b independent of j and n, such that $|\mathbb{E}(X_{n,j}) - b(\log n - \log j)| \le c_b$ and $|\mathbb{V}(X_{n,j}) - b(\log n - \log j)| \le d_b$, for all $1 \le j \le n$.

We use now the abbreviations

$$\mu_{n,j} := b(\log n - \log j) \quad \text{and} \quad \sigma_{n,j} := \sqrt{b(\log n - \log j)}, \tag{49}$$

and consider the normalized r.v.

$$X_{n,j}^* := \frac{X_{n,j} - \mu_{n,j}}{\sigma_{n,j}} \tag{50}$$

and the moment generating function

$$\mathbb{E}\big(e^{sX_{n,j}^*}\big) = e^{-\frac{\mu_{n,j}}{\sigma_{n,j}}s}\,\mathbb{E}\big(e^{\frac{X_{n,j}}{\sigma_{n,j}}s}\big) = e^{-\sigma_{n,j}s}p_{n,j}(e^{\frac{s}{\sigma_{n,j}}}). \tag{51}$$

We will consider the two summands of the probability generating function $p_{n,j}(v)$ as given by (47) separately and use thus the abbreviations:

$$p_{n,j}^{[1]}(v) := \sum_{i=1}^{b}\sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}\binom{n+bv-b-1}{n-j}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{n-1}{j-1}\binom{bv-1}{b-\ell-1}},$$
$$p_{n,j}^{[2]}(v) := \sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}}-\frac{1}{\binom{bv-1}{b-k}}\right)\binom{j-2+\lambda_i}{j-1}\binom{n-k-1}{n-j-k+\ell+1}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{n-1}{j-1}}.$$

We assume now that $j=o(n)$ and consider $e^{-\sigma_{n,j}s}p_{n,j}^{[2]}(e^{\frac{s}{\sigma_{n,j}}})$ for a real s fixed. Using the asymptotic expansions

$$\binom{j-2+\lambda_i}{j-1} = \mathcal{O}(1) \quad \frac{\binom{n-k-1}{n-j-k+\ell+1}}{\binom{n-1}{j-1}} = \mathcal{O}\big(\frac{j}{n}\big), \quad \text{and} \quad \frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}} = \mathcal{O}\big(\frac{1}{\sigma_{n,j}}\big),$$

we obtain that

$$e^{-\sigma_{n,j}s}p_{n,j}^{[2]}(e^{\frac{s}{\sigma_{n,j}}}) = \mathcal{O}\Big(\sqrt{\frac{j}{n}}\Big), \tag{52}$$

which will turn out to be negligible compared to $e^{-\sigma_{n,j}s}p_{n,j}^{[1]}(e^{\frac{s}{\sigma_{n,j}}})$.

We consider now the contribution of $p_{n,j}^{[1]}(v)$ and split the considered range $j=o(n)$ into the regions $j>\log n$ and $j\le \log n$. We first assume that $j>\log n$. By a direct application of Stirling's formula (35) we obtain then the following expansion, which holds uniformly around $v=1$:

$$p_{n,j}^{[1]}(v) = K(v)e^{b(v-1)(\log n - \log j)}\big(1+\mathcal{O}((\log n)^{\Re\lambda_2-1})+\mathcal{O}((\log n)^{-1})\big),$$

with

$$K(v) = \frac{v-\frac{1}{\binom{bv-1}{b-1}}}{bH_b(v-1)}.$$

Since $K(v)=1+\mathcal{O}(v-1)$ we obtain for $j=o(n)$ such that $j>\log n$ and for every s fixed the asymptotic expansion

$$e^{-\sigma_{n,j}s}p_{n,j}^{[1]}(e^{\frac{s}{\sigma_{n,j}}}) = e^{\frac{s^2}{2}}\big(1+\mathcal{O}\big((\log n)^{\Re\lambda_2-1}\big)+\mathcal{O}\big(\sigma_{n,j}^{-1}\big)\big). \tag{53}$$

Second we assume that $j\le\log n$. Using a Taylor-series expansion around $v=e^{\frac{s}{\sigma_{n,j}}}=1$ we obtain for this region and s fixed the expansions

$$\binom{be^{\frac{s}{\sigma_{n,j}}}-1}{b-\ell-1} = \binom{b-1}{b-\ell-1}\Big(1+\mathcal{O}\big(\frac{1}{\sqrt{\log n}}\big)\Big), \qquad \binom{j+be^{\frac{s}{\sigma_{n,j}}}-b-1}{j-1} = 1+\mathcal{O}\big(\frac{\log\log n}{\sqrt{\log n}}\big),$$
$$\binom{n+be^{\frac{s}{\sigma_{n,j}}}-b-1}{n-1} = e^{b(e^{\frac{t}{\sigma_{n,j}}}-1)\log n}\Big(1+\mathcal{O}\big(\frac{1}{\sqrt{\log n}}\big)\Big).$$

Using them we obtain the expansion

$$e^{-\sigma_{n,j}s}p_{n,j}^{[1]}(e^{\frac{s}{\sigma_{n,j}}}) = e^{-\sigma_{n,j}s}e^{b(e^{\frac{s}{\sigma_{n,j}}}-1)\log n}\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{b-1}{b-\ell-1}}\big(1+\mathcal{O}\big(\frac{\log\log n}{\sqrt{\log n}}\big)\big).$$

Since we have for $j\le\log n$ and s fixed the expansions

$$\sigma_{n,j} = \sqrt{b}\sqrt{\log n}\Big(1+\mathcal{O}\big(\frac{\log\log n}{\log n}\big)\Big), \quad \text{and} \quad e^{\frac{s}{\sigma_{n,j}}}-1 = \frac{s}{\sigma_{n,j}}+\frac{s^2}{2\sigma_{n,j}^2}+\mathcal{O}\big(\frac{1}{(\log n)^{\frac{3}{2}}}\big),$$

we further get

$$e^{-\sigma_{n,j}s}p_{n,j}^{[1]}(e^{\frac{s}{\sigma_{n,j}}}) = e^{\frac{s^2}{2}}\sum_{i=1}^{b}\sum_{\ell=0}^{b-1}\frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{b(H_{\lambda_i+b-1}-H_{\lambda_i-1})\binom{b-1}{b-\ell-1}}\Big(1+\mathcal{O}\big(\frac{\log\log n}{\sqrt{\log n}}\big)\Big). \tag{54}$$

But due to the binomial identity

$$\sum_{\ell=0}^{b-1} \frac{\binom{\lambda+b-1}{b-\ell-1}}{\binom{b-1}{b-\ell-1}} = \begin{cases} 0, & \text{for } \lambda \neq 1, \\ bH_b, & \text{for } \lambda = 1, \end{cases}$$

the double sum appearing in (54) evaluates to 1:

$$\sum_{i=1}^{b} \sum_{\ell=0}^{b-1} \frac{\binom{\lambda_i+b-1}{b-\ell-1}\binom{j-2+\lambda_i}{j-1}}{b(H_{\lambda_i+b-1} - H_{\lambda_i-1})\binom{b-1}{b-\ell-1}} = 1.$$

Therefore we obtain for $j = o(n)$ such that $j \leq \log n$ and for every s fixed the asymptotic expansion

$$e^{-\sigma_{n,j} s} p_{n,j}^{[1]}(e^{\frac{s}{\sigma_{n,j}}}) = e^{\frac{s^2}{2}} \Big(1 + \mathcal{O}\big(\frac{\log\log n}{\sqrt{\log n}}\big)\Big). \tag{55}$$

Combining the previous results (52), (53) and (55) we obtain that for the whole region $j = o(n)$ the moment generating function $\mathbb{E}(e^{X_{n,j}^* s}) = e^{-\sigma_{n,j} s} p_{n,j}(e^{\frac{s}{\sigma_{n,j}}})$ of the r.v. $X_{n,j}^*$ converges pointwise for every s fixed to $e^{\frac{s^2}{2}}$, which is the moment generating function of a standard normal distributed random variable. This suffices to show that $X_{n,j}^* = \frac{X_{n,j}-\mu_{n,j}}{\sigma_{n,j}}$ converges in distribution to a standard normal distributed r.v. and proves the corresponding part of Theorem 4.

The central region for j: $j \to \infty$ such that $j \sim \rho n$, with $0 < \rho < 1$. We assume now that $\epsilon \leq \frac{j}{n} \leq 1 - \epsilon$, with $\epsilon > 0$, and we assume further that v is in a (complex) neighbourhood of 1. Then we obtain the following asymptotic expansions:

$$\binom{j-2+\lambda_i}{j-1} = j^{\lambda_i - 1}\big(1 + \mathcal{O}(j^{-1})\big) = \begin{cases} 1 + \mathcal{O}(n^{-1}), & i = 1, \\ \mathcal{O}(n^{\Re\lambda_i - 1}), & 2 \leq i \leq b, \end{cases}$$

$$\frac{\binom{n+bv-b-1}{n-j}}{\binom{n-1}{j-1}} = \frac{1}{\big(\frac{j}{n}\big)^{bv-b}}\big(1 + \mathcal{O}(n^{-1})\big) = e^{-b(v-1)\log\frac{j}{n}}\big(1 + \mathcal{O}(n^{-1})\big),$$

$$\frac{\binom{n-k-1}{n-j-k+\ell+1}}{\binom{n-1}{j-1}} = \big(\frac{j}{n}\big)^{\ell+1}\big(1 - \frac{j}{n}\big)^{k-\ell-1}\big(1 + \mathcal{O}(n^{-1})\big).$$

Using them we get from (47) the following asymptotic expansion of $p_{n,j}(v)$, which holds uniformly for $\epsilon \leq \frac{j}{n} \leq 1 - \epsilon$ in a complex neighbourhood of $v = 1$:

$$p_{n,j}(v) = \left(e^{-b(v-1)\log\frac{j}{n}} \sum_{\ell=0}^{b-1} \frac{\binom{b}{b-\ell-1}}{bH_b\binom{bv-1}{b-\ell-1}} + \sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{b}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)}{bH_b} \big(\frac{j}{n}\big)^{\ell+1}\big(1 - \frac{j}{n}\big)^{k-\ell-1} \right) \times \big(1 + \mathcal{O}(n^{-1}) + \mathcal{O}(n^{\Re\lambda_2 - 1})\big).$$

This shows that, for $j \sim \rho n$ with $0 < \rho < 1$, the probability generating function $p_{n,j}(v)$ converges uniformly in a complex neighbourhood of $v = 1$ to a function $p_\rho(v)$ given as follows:

$$p_\rho(v) := e^{-b(v-1)\log\rho} \sum_{\ell=0}^{b-1} \frac{\binom{b}{b-\ell-1}}{bH_b\binom{bv-1}{b-\ell-1}} + \sum_{\ell=0}^{b-1}\sum_{k=\ell+1}^{b-1} \frac{\binom{b}{b-\ell-1}\binom{b-\ell-1}{k-\ell-1}\left(\frac{1}{\binom{b-1}{b-k}} - \frac{1}{\binom{bv-1}{b-k}}\right)}{bH_b} \rho^{\ell+1}(1-\rho)^{b-1-\ell}. \tag{56}$$

Since this also shows that for this region the moment generating function $\mathbb{E}(e^{X_{n,j}s}) = p_{n,j}(e^s)$ converges pointwise in a real neighbourhood of $s = 0$ to the moment generating function $\mathbb{E}(e^{X_\rho s}) = p_\rho(e^s)$ of a r.v. X_ρ, we obtain that, for $j \sim \rho n$ with $0 < \rho < 1$, $X_{n,j}$ converges in distribution to a discrete r.v. X_ρ with probability generating function $p_\rho(v)$. Thus the corresponding part in Theorem 4 is proven.

The region for j large: $j \to \infty$ such that $n-j = o(n)$. For the region $n-j=o(n)$ we obtain the following asmyptotic expansions:

$$\frac{\binom{n-b-1}{n-j}}{\binom{n-1}{j-1}} = 1 + o(1), \quad \text{and} \quad \frac{\binom{n-k-1}{n-j-k+\ell+1}}{\binom{n-1}{j-1}} = \begin{cases} o(1), & \text{for } k > \ell+1, \\ 1+o(1), & \text{for } k = \ell+1. \end{cases}$$

Using these expansions leads then, for $j \to \infty$ such that $n-j=o(n)$, to

$$\mathbb{P}\{X_{n,j} = 0\} = p_{n,j}(0) = 1 + o(1). \tag{57}$$

Thus, for this region, $X_{n,j}$ converges in distribution to a degenerate r.v. $\tilde{X}$, with $\mathbb{P}\{\tilde{X} = 0\} = 1$, as stated in the corresponding part of Theorem 4.

REFERENCES

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

[2] H.-H. Chern and H.-K. Hwang, Phase changes in random m-ary search trees and generalized quicksort, *Random Structures & Algorithms* 19, 316–358, 2001.

[3] P. Flajolet and A. Odlyzko, Singularity Analysis of Generating Functions, *SIAM Journal on Discrete Mathematics* 3, 216–240, 1990.

[4] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, to appear.

[5] C. Goldschmidt and J. B. Martin, Random recursive trees and the Bolthausen-Sznitman coalescent, *Electronic Journal of Probability* 10, 718–745, 2005.

[6] R. Graham, D. Knuth and O. Patashnik, *Concrete Mathematics*, Second Edition, Addison-Wesley, Reading, 1994.

[7] H.-K. Hwang, On Convergence Rates in the Central Limit Theorems for Combinatorial Structures, *European Journal of Combinatorics* 19, 329–343, 1998.

[8] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns, *Stochastic Processes and their Applications* 110, 177–245, 2004.

[9] H. Mahmoud and R. Smythe, Probabilistic analysis of bucket recursive trees, *Theoretical Computer Science* 144, 221–249, 1995.

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

[11] A. Panholzer, Distribution of the Steiner Distance in Generalized M-ary Search Trees, *Combinatorics, Probability and Computing* 13, 717–733, 2004.

[12] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Campbridge University Press, Cambridge, 1999.

[13] J. Vitter and P. Flajolet, *Average Case Analysis of Algorithms and Data Structures*, in *Handbook of Theoretical Computer Science*, 431–524, Elsevier, Amsterdam, 1990.

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

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

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

E-mail address: `Alois.Panholzer@tuwien.ac.at`