ANALYSIS OF STATISTICS FOR GENERALIZED STIRLING PERMUTATIONS

MARKUS KUBA AND ALOIS PANHOLZER

ABSTRACT. In this work we give a study of generalizations of Stirling permutations, a restricted class of permutations of multisets introduced by Gessel and Stanley [15]. First we give several bijections between such generalized Stirling permutations and various families of increasing trees extending the known correspondences of [20, 21]. Then we consider several permutation statistics of interest for generalized Stirling permutations as the number of left-to-right minima, the number of left-to-right maxima, the number of blocks of specified sizes, the distance between occurrences of elements, and the number of inversions. For all these quantities we give a distributional study, where the established connections to increasing trees turn out be very useful. To obtain the exact and limiting distribution results we use several techniques ranging from generating functions, connections to urn models, martingales and Stein's method.

1. INTRODUCTION

Stirling permutations were defined by Gessel and Stanley [15]. A Stirling permutation is a permutation of the multiset $\{1, 1, 2, 2, \ldots, n, n\}$ such that, for each i, $1 \le i \le n$, the elements occurring between the two occurrences of i are larger than i. E.g., 1122, 1221 and 2211 are Stirling permutations, whereas the permutations 1212 and 2112 of $\{1, 1, 2, 2\}$ aren't. The name of these combinatorial objects is due to relations with the Stirling numbers, see [15] for details.

A straightforward generalization of Stirling permutations is to consider permutations of a more general multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$, with $k_i \in \mathbb{N}$ for $1 \le i \le n$ (here and throughout this work we use in this context $j^l := \underbrace{j, \ldots, j}_{l}$, for $l \ge 1$). We call a permutation of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$ a *generalized Stirling permutation*, if for each i, $1 \le i \le n$, the elements occurring between two occurrences of i are at least i. (Alternatively, one might say that the elements occurring between two consecutive occurrences of i are larger than i.) Such permutations have already been considered previously by Brenti [8, 9]. Probably the most natural generalizations of Stirling permutations are obtained when all labels of the multiset appear with the same multiplicity; as in [21] we thus define the class $\mathcal{Q}_n = \mathcal{Q}_n(k)$ of *k-Stirling permutations* of order n as generalized Stirling permutations of the multiset $\{1^k, 2^k, \ldots, n^k\}$, with an integer $k \ge 1$. Note that $k = 2$ yields exactly Stirling permutations as defined by Gessel and Stanley [15], whereas $k = 1$ gives just ordinary permutations. This class of k-Stirling permutations has been introduced already by Park [32, 33, 34] under the name k-multipermutations.

In the recent papers [7, 20, 21] several classes of generalized Stirling permutations have been considered, where a main focus has been given to a study of the number of ascents, descents and plateaux. We call an element σ_i, with $1 \le i \le s$, an ascent of the generalized Stirling permutation $\sigma = \sigma_1 \sigma_2 \cdots \sigma_s$, if $\sigma_{i-1} < \sigma_i$ (we set here $\sigma_0 := 0$); analogous definitions can be given for descents and plateaux. During the studies [20, 21] close relations between classes of generalized Stirling permutations and families of so-called increasing trees have been established. Increasing trees are rooted labelled trees, where the nodes are labelled in such a way, that each child of a node has a label larger than its parent node. In the first part of this work we reconsider this issue and give three general constructions, which relate generalized Stirling permutations with various kinds of increasing trees. Such connections between classes of generalized Stirling permutations and increasing tree

families turn out to be quite useful, since they also link specific quantities in generalized Stirling permutations with quantities in increasing trees.

The second part of this work is devoted to a study of interesting quantities in k-Stirling permutations. There, besides obtaining exact results, a main focus is given to a description of the limiting behaviour of the parameters. For our study we will throughout this paper always assume that all $Q_n = Q_n(k) := |\mathcal{Q}_n(k)| = \prod_{i=1}^{n-1} \big(ki+1\big)$ (see Section 2) different k-Stirling permutations of order n, i.e., of the multiset $\{1^k, 2^k, \ldots, n^k\}$, are occurring with the same probability; we will thus speak about *random k-Stirling permutations* of order n. In particular, if we introduce a random variable X_n measuring a certain quantity in the class $Q_n(k)$ we use this model of randomness; furthermore, for any particular $\sigma \in \mathcal{Q}_n(k)$, $X_n(\sigma)$ gives the value of the quantity for the k-Stirling permutation σ.

First we consider the number of *left-to-right minima* and the number of *left-to-right maxima* in k-Stirling permutations. For a generalized Stirling permutation $\sigma = \sigma_1 \sigma_2 \cdots \sigma_s$ of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$, with $s = \sum_{i=1}^{n} k_i$, an element σ_i, with $1 \le i \le s$, is a left-to-right minimum if $\sigma_{i'} > \sigma_i$, for all $i' < i$; analogous σ_i is a left-to-right maximum if $\sigma_{i'} < \sigma_i$, for all $i' < i$. Thus we are introducing the random variables L_n and M_n, which will count the number of left-to-right minima and the left-to-right maxima, respectively, in a random k-Stirling permutation of order n. Of course, due to symmetry the number of right-to-left minima (maxima), which can be defined analogously, are equidistributed with the number of left-to-right minima (maxima) and they do not have to be studied separately. Furthermore, for $k = 1$, due to the complementation $j \mapsto n + 1 - j$ the random variables L_n and M_n are equidistributed, which is no more true for $k \ge 2$. The number of left-to-right minima in k-Stirling permutations has already been considered in [32], where a formula for the exponential generating function has been given. Here we add a simple explicit formula for the number of k-Stirling permutations of order n with exactly m left-to-right minima (or alternatively, for the exact probability $\mathbb{P}\{L_n = m\}$) and also establish a central limit theorem for L_n. The parameter M_n is more difficult to treat, but we are able to give exact and asymptotic results for the expected number $\mathbb{E}(M_n)$ of left-to-right maxima in a random k-Stirling permutation of order n. During the analysis we even show refined results, i.e., exact and asymptotic results for the expected number of left-to-right maxima amongst the first i elements $\sigma_1 \sigma_2 \cdots \sigma_i$, with $i \le kn$, of a k-Stirling permutation $\sigma = \sigma_1 \cdots \sigma_{kn}$.

Then we carry out a precise *local study* of the block structure of k-Stirling permutations. A *block* in a generalized Stirling permutation $\sigma = \sigma_1 \cdots \sigma_s$ is a substring $\sigma_a \cdots \sigma_b$, with $\sigma_a = \sigma_b$, that is maximal, i.e., which is not contained in any larger such substring. There is obviously at most one block for every $j \in \{1, 2, \ldots, n\}$, extending from the first occurrence of j to the last one; we say that j forms a block if this substring is indeed a block, i.e., when it is not contained in a string $j' \cdots j'$, for some $j' < j$. It can be shown easily by induction that any generalized Stirling permutation has a unique decomposition as a sequence of its blocks. Janson et al. [21] studied several *global parameters* related to the block structure of k-Stirling permutations; more precisely, they analyzed the (asymptotic) distribution of the total number of blocks as well as the (joint) distribution of the sizes of the largest blocks. To obtain their results the authors used various methods, in particular descriptions via so-called urn models (see [17, 19]); such relations to urn models are also of interest here. In the present work we are interested in the number of blocks of a certain size, i.e., blocks of size $\tilde{\ell}$, with $1 \le \tilde{\ell} \le n$, in a random k-Stirling permutation of order n, where we explicitly allow that $\tilde{\ell}$ may also be a function of n such that $\tilde{\ell} = \tilde{\ell}(n)$; we use here the convention that the size of a block $\sigma_a \cdots \sigma_b$ is given by $b - a + 1$, i.e., by the number of its elements. Since for k-Stirling permutations the size of a block is always a multiple of k we introduce the random variable $X_{n,\ell}$ that counts the number of blocks of size $\tilde{\ell} = k\ell$ in a random k-Stirling permutation of order n. In our studies we obtain for $X_{n,\ell}$ the exact distribution, the moments, and a full characterization of the limiting distributions appearing depending on the growth of $\ell = \ell(n)$. For the analysis the correspondence between k-Stirling permutations and increasing trees as well as generating functions techniques are heavily used. Moreover, we briefly show how to extend our study to obtain the joint distribution of the number of blocks of sizes $k\ell_1, \ldots, k\ell_r$, with $1 \le \ell_1 < \cdots < \ell_r \le n$, for arbitrary but fixed $r \in \mathbb{N}$. Furthermore, we show that our problem is also related to an urn model, which turns out to be exactly solvable by using methods of Flajolet et al. [13].

Then we study the distribution of *distances* in k-Stirling permutations, a parameter which was considered briefly by Janson [20] for the case $k = 2$. For the analysis Janson used the specific connection between distances in 2-Stirling permutations and descendants in plane-oriented increasing trees, which only holds for

$k = 2$, see Section 2 and also [20, 21]. Let $\Delta_{n;j,p}$ measure the distance between the p-th and the $(p + 1)$-th occurrence of j in a random k-Stirling permutation of order n, with $1 \le p \le k - 1$ and $1 \le j \le n$. For a given k-Stirling permutation $\sigma = \sigma_1\sigma_2 \cdots \sigma_{kn}$ of order n let i_p and i_{p+1} denote the indices of the p-th and $(p + 1)$-th occurrence of the number j in a k-Stirling permutation of order n (i.e., $\sigma_{i_p} = j$, $\sigma_{i_{p+1}} = j$, and there are exactly $p - 1$ indices $1 \le i_1 < i_2 < \cdots < i_{p-1} < i_p$ such that $\sigma_{i_q} = j$, $1 \le q \le p - 1$); then the distance between the p-th and $(p + 1)$-th occurrence of j is defined as $i_p - i_{p+1}$. Alternatively $\Delta_{n;j,p}$ counts one plus the number of elements between the p-th and $(p + 1)$-th occurrence of j in a random k-Stirling permutation of order n. In our analysis we will determine the exact distribution of $\Delta_{n;j,p}$, $1 \le p \le k - 1$, and again we are able to give a full characterization of the limiting distributions appearing depending on the growth of $j = j(n)$. Moreover, using relations between k-Stirling permutations and increasing trees, we also obtain distribution results of statistics related to the number of descendants of nodes in $(k+1)$-ary increasing trees; more precisely, using the bijective correspondence between $(k + 1)$-ary increasing trees and k-Stirling permutations described in Section 2 we will show that $\Delta_{n;j,p}$ counts one plus k times the number of descendants of node j contained in the $(p + 1)$-th branch of node j (i.e., the number of nodes in the $(p+1)$-th branch of node j), with $1 \le p \le k - 1$, in a random $(k + 1)$-ary increasing tree of order n. Thus we also obtain a refinement of known results for the number of descendants of the nodes in such random increasing trees, see [23, 28, 36].

The last section of this work is devoted to a distributional study of the number of *inversions* in k-Stirling permutations. An inversion in a generalized Stirling permutation $\sigma = \sigma_1\sigma_2 \cdots \sigma_s$ is defined as a pair of indices (i, i'), $1 \le i < i' \le s$, such that $\sigma_i > \sigma_{i'}$. An enumerative study of the number of inversions in k-Stirling permutations has been carried out already by Park [32]. Here we are interested in the limiting behaviour of this quantitiy and introduce the random variable I_n, which counts the number of inversions in a random k-Stirling permutation of order n. By an application of Stein's method (see, e.g., [2]) we show a central limit theorem for I_n.

We find a study of parameters in generalized Stirling permutations of interest at least for the following two reasons. Firstly, generalized Stirling permutations are a natural restricted class of permutations (avoiding the pattern 212, see, e.g., [6] for permutation patterns) of a multiset and the influence of this restriction to the behaviour of "classical" permutation statistics (as the number of left-to-right maxima, minima and inversions) arises as a natural question. Secondly, classes of generalized Stirling permutations can be considered as encodings of important increasing tree families as recursive trees, plane-oriented recursive trees and d-ary increasing trees, and a study of quantities in generalized Stirling permutations also leads to results for corresponding parameters in such tree families (as it is the case for the number of left-to-right maxima, minima, blocks or the distance between occurrences of elements).

The following example shall illustrate the parameters studied in this work.

Example 1. We consider the 3-Stirling permutation $\sigma = 355777534443112888221666$ of order 8. There are two left-to-right minima, namely $\sigma_1 = 3$ and $\sigma_{13} = 1$, and four left-to-right maxima, namely $\sigma_1 = 3$, $\sigma_2 = 5$, $\sigma_4 = 7$ and $\sigma_{16} = 8$, thus $L_8(\sigma) = 2$ and $M_8(\sigma) = 4$. Furthermore σ can be decomposed into three blocks leading to the block decomposition $[355777534443][112888221][666]$, which consists of single blocks of sizes 3, 9 and 12, respectively. Thus $X_{8,1}(\sigma) = 1$, $X_{8,3}(\sigma) = 1$ and $X_{8,4}(\sigma) = 1$ (recall that $X_{n,\ell}$ counts the number of blocks of size $k\ell$). Moreover, the distance between the first and second occurrence of 3 is 7, whereas the distance between the second and third occurrence of 3 is 4; thus $\Delta_{8;3,1}(\sigma) = 7$ and $\Delta_{8;3,2}(\sigma) = 4$. Finally we denote by q_j the number of inversions obtained by elements with label j together with elements to the right that are smaller than j and collect them in an "inversion table" $(q_1, q_2, \ldots, q_n)$. This leads for σ to $(0, 3, 18, 21, 33, 0, 45, 18)$ and thus to a total number of inversions $I_8(\sigma) = \sum_{i=1}^{8} q_i = 138$.

Throughout this work $\Psi(x) = (\ln \Gamma(x))'$ denotes the Psi-function, $\Psi'(x)$ its derivative, whereas $\Phi(x)$ denotes the distribution function of the standard normal distribution $\mathcal{N}(0, 1)$. Moreover, we use the abbreviation $\begin{bmatrix} n \\ k \end{bmatrix}$ for the signless Stirling numbers of the first kind, $\begin{Bmatrix} n \\ k \end{Bmatrix}$ for the Stirling numbers of the second kind, and $H_n := \sum_{j=1}^{n} \frac{1}{j}$ for the harmonic numbers. The floor and ceiling function returning the largest integer $\le x$ and the smallest integer $\ge x$, respectively, are denoted by $\lfloor x \rfloor$ and $\lceil x \rceil$, respectively. We write $x^{\underline{m}} := x(x - 1) \cdots (x - m + 1)$ for the falling factorials and $\delta_{i,j}$ for the Kronecker delta function. Sometimes it is convenient

to use the notion of multifactorials $n!^{(k)}$, which are, for integers $n \geq 0$, recursively defined as follows:

$$n!^{(k)} := \begin{cases} 1, & \text{for } 0 \leq n < k, \\ n \cdot \big((n-k)!^{(k)}\big), & \text{for } n \geq k. \end{cases}$$

For independent random variables X and Y we write $X \oplus Y$ for the sum of X and Y. For not necessarily mutually independent random variables X and Y we write $X + Y$. Furthermore, $X \overset{(d)}{=} Y$ denotes the equality in distribution of the random variables X and Y, whereas $X_n \overset{(d)}{\longrightarrow} X$ denotes the weak convergence, i.e., the convergence in distribution, of the sequence of random variables X_n to a random variable X; the almost sure convergence is denoted by $X_n \overset{(a.s.)}{\longrightarrow} X$.

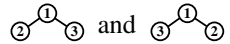## 2. Stirling permutations and increasing trees

2.1. **Generalized Stirling permutations.** We consider the class of generalized Stirling permutations of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$. It is seen easily and already stated in [21] that there are exactly $\prod_{j=1}^{n} \big(1 + \sum_{i=1}^{j-1} k_i\big)$ different generalized Stirling permutations of that kind. This can be shown by induction, since the k_n copies of n have to form a substring and thus each such generalized Stirling permutation can be obtained uniquely by inserting the string n^{k_n} into a generalized Stirling permutation of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, (n-1)^{k_{n-1}}\}$ at one of the $1 + \sum_{i=1}^{n-1} k_i$ possible positions (viz., anywhere, including first or last). Moreover, when considering the insertion process, where we start with a string 1^{k_1} and successively insert at step $j > 1$ the string j^{k_j} randomly at one of the $1 + \sum_{i=1}^{j-1} k_j$ possible positions (or "gaps") into the generated string, we obtain after step n each of the $\prod_{j=1}^{n} \big(1 + \sum_{i=1}^{j-1} k_i\big)$ generalized Stirling permutations of this multiset with the same probability, i.e., we obtain a random generalized Stirling permutation of this multiset.

Example 2. In the following we list all 12 generalized Stirling permutations of the multiset $\{1^2, 2, 3^3\}$: 112333, 113332, 133312, 333112, 121333, 123331, 133321, 333121, 211333, 213331, 233311, 333211.

When specializing to the class of k-Stirling permutations $\mathcal{Q}_n$ we obtain that the number $Q_n = Q_n(k) := |\mathcal{Q}_n(k)|$ of different k-Stirling permutations of order n is given by

$$Q_n(k) = \prod_{i=1}^{n-1} (ki+1) = n! k^n \binom{n-1+\frac{1}{k}}{n} = k^n \frac{\Gamma(n+1/k)}{\Gamma(1/k)} = (k(n-1)+1)!^{(k)}. \tag{1}$$

For $k = 2$ this number is just $Q_n(2) = (2n-1)!!$. Moreover, the insertion process, where one starts with a string 1^k and successively inserts at step $j > 1$ the string j^k randomly at one of the $k(j-1)+1$ possible positions into the generated string, yields after step n a random k-Stirling permutation of order n, i.e., each of the Q_n different k-Stirling permutations of order n is obtained with the same probability $1/Q_n$.

2.2. **Relations to increasing trees.** Increasing trees are rooted labelled trees, where the nodes of a tree with n nodes (i.e., of order n) are labelled with distinct integers from a given label set L in such a way that each child node has a larger label than its parent node. Alternatively one might say that the labels along any path starting at the root node and ending at an arbitrary node v are forming an increasing sequence. Usually we use as label set $L = \{1, 2, \ldots, n\}$ to label the nodes of an increasing tree of order n; if we use another label set L then we express this by saying "increasing trees of L", e.g., sometimes the set $L = \{0, 1, \ldots, n-1\}$ is more convenient. If we want to express that the left-to-right order of the children of the nodes is important we speak about ordered trees, otherwise we use the term unordered trees. E.g., the increasing trees and are considered as two different ordered trees, but as unordered trees they are considered as the same tree.

To formulate the first correspondence with generalized Stirling permutations we consider ordered increasing tree families, where, depending on the labels, the nodes are of different node types. We say that node v is of node type A_d, with $d \geq 1$, if there are exactly d positions, where a child might be attached or not. Thus there are exactly $\binom{d}{r}$ different possibilities that the sequence of $0 \leq r \leq d$ nodes $v_1, v_2, \ldots, v_r$ is attached to v in this left-to-right order.

Theorem 1. *There is a bijective correspondence between the class of generalized Stirling permutations of the multiset* $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$, *with* $k_i \geq 1$ *for* $1 \leq i \leq n$, *and the family of ordered increasing trees, where node 1 is of node type* A_{k_1+1}, *node 2 is of node type* A_{k_2+1}, *..., node* n *is of node type* A_{k_n+1}.

Proof. We introduce a mapping φ from the family of ordered increasing trees of the given node types to the class of generalized Stirling permutations of the given multiset and show that φ is indeed a bijection. We start with an ordered increasing tree T. In order to describe the image of φ we use a depth-first walk of T: the depth-first walk of a rooted (ordered) tree starts at the root, goes first to the leftmost child of the root, explores that branch (recursively, using the same rules), returns to the root, and continues with the next child of the root, until there are no more children left. To do this we think of the increasing tree T, which consists of n nodes with the given node types, as a tree, where all empty places (i.e., each position at a node, where no child is attached) are represented by "external nodes". Hence, any (internal) node of node type A_d always has d children, some of which may be external nodes. Between these d edges going out from a node labelled v, we place $d - 1$ integers v. (External nodes have no children and no labels.) Now we perform the depth-first walk and encode the increasing tree T by the sequence of the labels visited as we go around the tree (one may think of actually going around the tree like drawing the contour). Thus, for each node labelled v of node type A_d, we add label v to the code the $d - 1$ first times we return to node v, but not the first time we arrive there or the last time we return. Of course, the sequence of labels σ we obtain by this depth-first walk is a permutation of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$. Since the nodes in each branch of a node labelled v in the increasing tree T have labels larger than v it is guaranteed that all labels between two occurrences of v are larger than v, which shows that σ is indeed a generalized Stirling permutation. It is apparent that different trees lead to different generalized Stirling permutations, i.e., that φ is injective. To show that φ is also surjective we first describe the following insertion procedure, which allows to generate all increasing trees with the given node types: start with the root node 1, where $k_1 + 1$ external nodes are attached. Then successively insert nodes with label $j > 1$ by substituting in the tree generated so far one of the external nodes by node j and attaching $k_j + 1$ external nodes to j. It can be shown easily by induction that there are exactly $1 + \sum_{i=1}^{j-1} k_i$ external nodes before inserting node j, which furthermore shows that there are exactly $\prod_{j=1}^{n} \left(1 + \sum_{i=1}^{j-1} k_i\right)$ different increasing trees with the given node types. Moreover, when coupling the insertion procedures for increasing trees and generalized Stirling permutations by the rule that inserting the string j^{k_j} at the r_j-th possible position (at the r_j-th gap) into the string generated so far corresponds to a substitution of the r_j-th external node (in left-to-right order) by node j, then we obtain after n steps exactly the tree, whose mapping φ leads to the desired generalized Stirling permutation.

To describe the inverse mapping φ^{-1} we start with a generalized Stirling permutation σ and recursively construct a corresponding increasing tree T as follows. We consider the decomposition of σ into substrings according to the occurrences of the smallest label, i.e., to the k_1 occurrences of 1: $\sigma = \tau_1 1 \tau_2 1 \tau_3 \ldots \tau_{k_1} 1 \tau_{k_1+1}$. The substrings $\tau_1, \tau_2, \ldots, \tau_{k_1+1}$ can itself be considered as generalized Stirling permutations of certain multisets. More precisely, the multiset $S = \{2^{k_2}, \ldots, n^{k_n}\}$ (the original multiset, where all occurrences of the smallest label have been removed) is partitioned into possibly empty multisets $S_1, S_2, \ldots, S_{k_1+1}$, with $S_i \cap S_{i'} = \emptyset$, for $i \neq i'$, such that, for each q, τ_q is a generalized Stirling permutation of the multiset S_q. Now we can recursively apply the (analogous) construction to each of the substrings τ_q and obtain increasing trees $T_1, \ldots, T_{k_1+1}$; if τ_q is the empty string we set T_q as the empty tree (or as an external node, if we like). E.g., if $S_1 = \{j_1^{k_{j_1}}, j_2^{k_{j_2}}, \ldots, j_t^{k_{j_t}}\}$ we obtain an increasing tree T_1 with labels $j_1, j_2, \ldots, j_t$, such that node j_1 is of node type $A_{k_{j_1}+1}$, ..., node j_t is of node type $A_{k_{j_t}+1}$. Finally, the root nodes of each tree $T_1, \ldots, T_{k_1+1}$ are attached at positions $1, \ldots, k_1 + 1$ to the node labelled by 1, which is of node type A_{k_1+1}; if T_q is the empty tree then at position q nothing (or alternatively an external node) will be attached. The resulting tree T is indeed an ordered increasing tree, where each node j is of node type A_{k_j+1}. Theorem 1 is illustrated in Figure 1. $\square$

For the instance of k-Stirling permutations, Theorem 1 has been stated already in [21] and [32]: then all nodes of the corresponding increasing trees are of the same node type A_{k+1} and one gets thus so-called $(k+1)$-ary increasing trees, members of well-known combinatorial tree families. The tree family $\mathcal{T} = \mathcal{T}(k+1)$ of
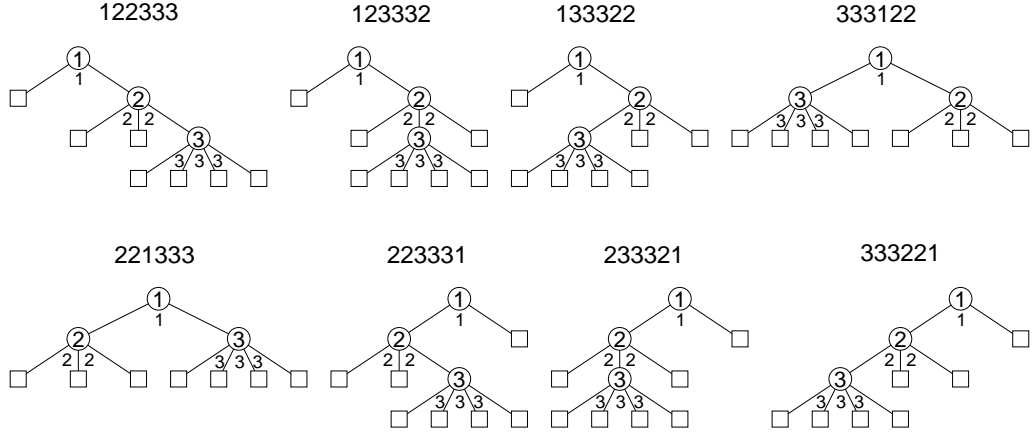
FIGURE 1. All eight generalized Stirling permutations of the multiset $\{1, 2^2, 3^3\}$ and the corresponding increasing trees, where the nodes 1, 2 and 3 are of node types A_2, A_3 and A_4, respectively.

$(k+1)$-ary increasing trees can also be described by the following formal recursive equation:

$$\mathcal{T} = \textcircled{1} \times \underbrace{\left(\{\epsilon\} \mathbin{\dot\cup} \mathcal{T}\right) * \left(\{\epsilon\} \mathbin{\dot\cup} \mathcal{T}\right) * \cdots * \left(\{\epsilon\} \mathbin{\dot\cup} \mathcal{T}\right)}_{k+1 \text{ times}} = \textcircled{1} \times \left(\{\epsilon\} \mathbin{\dot\cup} \mathcal{T}\right)^{k+1}, \tag{2}$$

where $\textcircled{1}$ denotes the node labelled by 1, ϵ the empty tree, $\times$ the cartesian product, $\dot\cup$ the disjoint union, and $*$ the partition product for labelled objects (see, e.g., the books [14, 38]).

Moreover, we denote by $\mathcal{T}_n = \mathcal{T}_n(k+1)$ the family of $(k+1)$-ary increasing trees of order n and by $T_n = T_n(k+1) := |\mathcal{T}_n(k+1)|$ the number of different $(k+1)$-ary increasing trees of order n. When introducing the generating function $T(z) = \sum_{n \geq 1} T_n \frac{z^n}{n!}$, then a direct application of the symbolic method (see, e.g., [14]) yields from (2) the following differential equation:

$$T'(z) = (1 + T(z))^{k+1}, \quad T(0) = 0,$$

with solution

$$T(z) = \frac{1}{(1 - kz)^{\frac{1}{k}}} - 1. \tag{3}$$

Extracting coefficients also shows that $T_n(k+1) = \prod_{i=1}^{n-1}(ki+1)$, $n \geq 1$, which coincides with the number $Q_n(k)$ of k-Stirling permutations of order n.

For the second correspondence we again consider ordered increasing tree families, where, depending on the labels, the nodes are of different node types. We say that node v is of node type B_d, with $d \geq 1$, if there are d positions, such that at each position a (possibly empty) sequence of children might be attached. Equivalently one might think of node v having $d - 1$ separating walls, which can be regarded as a special type of edges and which are separating the sequence of children attached to node v into d parts. Note that node type B_1 simply means that a possibly empty sequence of nodes might be attached to v.

Theorem 2. *There is a bijective correspondence between the class of generalized Stirling permutations of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$, with $k_1 \geq 0$ and $k_i \geq 2$, for $2 \leq i \leq n$, and the family of ordered increasing trees, where node 1 is of node type B_{k_1+1}, node 2 is of node type B_{k_2-1}, node 3 is of node type B_{k_3-1}, ..., node n is of node type B_{k_n-1}.* [1]

---

[1] We allow here $k_1 = 0$, i.e., that there are no occurrences of element 1 in the multiset; on the other hand we have the restriction that each element $2 \leq j \leq n$ appears at least twice in the multiset.
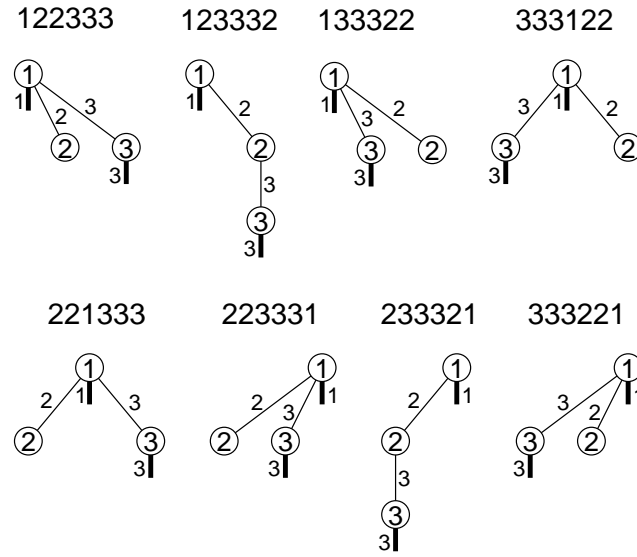
FIGURE 2. All eight generalized Stirling permutations of the multiset $\{1, 2^2, 3^3\}$ and the corresponding increasing trees, where the nodes 1, 2 and 3 are of node types B_2, B_1 and B_2, respectively.

Proof. Again we introduce a mapping ψ from the family of increasing trees with the given node types to the class of generalized Stirling permutations of the given multiset by using a depth-first walk. We start with an ordered increasing tree T and label each auxiliary separation wall of a node labelled v by the label of the node v. Moreover, we label any (proper) edge by the label of the child. Hence, node 1 of node type B_{k_1+1} has k_1 separation walls labelled 1, whereas each node labelled $v \neq 1$ of node type B_d has $d - 1$ separation walls labelled v and additionally an incoming edge labelled v. Now we perform the depth-first walk of T by going around the tree and encode it by the sequence of the labels visited on the edges, under the additional rule that a label on a separation wall only contributes once. Since every proper edge is traversed twice, and every label except 1 occurs on exactly one proper edge, the sequence of labels σ we obtain is a permutation of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$. Note that for each j, $1 \leq j \leq n$, the elements occurring between two occurrences of j are larger than j, since T is an increasing tree and thus we can only visit nodes with higher labels. Hence σ is indeed a generalized Stirling permutation of the given multiset. Finally, in a way analogous to Theorem 1 one shows that ψ is a bijection; we skip this (and also formulating explicitly the inverse mapping ψ^{-1}) here, but rather illustrate Theorem 2 in Figure 2. □

For the particular instance $k_1 = 0$ and $k_j = 2$, for $2 \leq j \leq n$, Theorem 2 gives a bijection between generalized Stirling permutations of the multiset $\{2^2, 3^2, \ldots, n^2\}$ (which are of course isomorphic to 2-Stirling permutations of order $n - 1$) and ordered increasing trees of order n, where all nodes are of node type B_1 (i.e., ordered increasing trees, where at each node a sequence of children might be attached). But such increasing trees are so-called plane-oriented recursive trees and Theorem 2 gives then exactly the bijection between 2-Stirling permutations of order n and plane-oriented recursive trees of order $n + 1$ introduced by Janson [20]. The family $\mathcal{P}$ of plane-oriented recursive trees is one of the most important increasing tree families with many applications, see, e.g., [4]. It can be described by the following formal recursive equation:

$$\mathcal{P} = \textcircled{1} \times \left(\{\epsilon\} \,\dot\cup\, \mathcal{P} \,\dot\cup\, \mathcal{P}^2 \,\dot\cup\, \mathcal{P}^3 \,\dot\cup\, \cdots \right). \tag{4}$$

Additionally the instances $k_1 = k - 2$ and $k_i = k$, for $2 \leq i \leq n$, of Theorem 2 have been described for arbitrary $k \geq 2$ in [21].
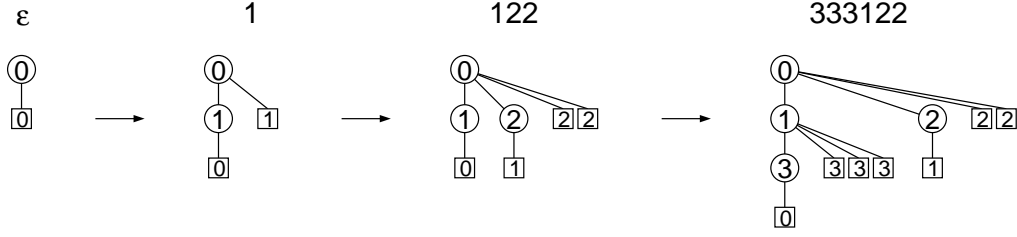
FIGURE 3. The construction of the generalized Stirling permutation 333122 of the multiset $\{1, 2^2, 3^3\}$ and the corresponding construction of an increasing tree with 4 internal and 7 external nodes.

For our last correspondence between generalized Stirling permutations and increasing trees we consider the insertion process for generalized Stirling permutations together with a corresponding one that will generate certain increasing trees. For each generalized Stirling permutation of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$ we obtain a specific increasing tree, which has $n + 1$ internal nodes labelled by distinct integers of $\{0, 1, \ldots, n\}$; moreover, it has $1 + \sum_{i=1}^{n} k_i$ external nodes, where one external node is labelled by 0 and, for $1 \leq j \leq n$, k_j are labelled by j. We note that the tree we obtain is an increasing tree by satisfying that the label of each internal node is larger than the label of its parent node; however, this condition is not necessarily satisfied for the external nodes.

Construction 1.

  Generalized Stirling permutations:

  - Step 0: we start with the empty word ϵ
  - Step j: we insert the string j^{k_j} at the q-th possible position (starting from the left) in the string obtained so far, with $1 \leq q \leq 1 + \sum_{i=1}^{j-1} k_i$

  Increasing trees:

  - Step 0: we start with the root node labelled 0, which has exactly one external node labelled 0
  - Step j: we insert the internal node labelled j together with k_j external nodes labelled j into the tree obtained so far by the following procedure: take the q-th external node x (in left-to-right order; it holds that $1 \leq q \leq 1 + \sum_{i=1}^{j-1} k_i$) in the tree and replace it by the internal node j, but attach the external node x to j as the only child of j; furthermore attach a sequence of k_j external nodes labelled j to the parent of the internal node j as the immediate right brothers of it.

The construction is illustrated in Figure 3. Apparently it holds (by construction) that the corresponding generalized Stirling permutation can be retrieved from the generated increasing tree by collecting the labels of the external nodes during a depth-first walk and removing the preceeding label 0. Furthermore, as a consequence of the construction, it holds that an internal node j does not have internal nodes as child nodes (i.e., if we remove all external nodes then j is a leaf node in the resulting tree) if and only if the first occurrence of j in the corresponding generalized Stirling permutation is an ascent.

We consider now the problem of characterizing the class of increasing trees generated by Construction 1. We treat the important case of k-Stirling permutations separately. First we have a look on the instance $k = 1$, i.e., ordinary permutations. If we start with a permutation σ of the set $\{1, 2, \ldots, n\}$ we obtain, by applying Construction 1, an increasing tree T with $n + 1$ internal nodes labelled by $\{0, 1, \ldots, n\}$; furthermore, T has $n + 1$ external nodes labelled also by distinct integers of $\{0, 1, \ldots, n\}$ and it holds that the rightmost child of each internal node is an external node. We remove now all external nodes of T and consider the resulting tree as an unordered increasing tree T', i.e., the left-to-right order of the children of a node is irrelevant, thus we can always consider them in increasing order, which is exactly the way they are obtained by Construction 1 (for $k = 1$). It is now important to note that the increasing tree T and thus the labels of the external nodes and so the permutation σ can be retrieved from the unordered increasing tree T' simply by applying Construction 1:

now at each step j we know the parent node of j, but since during the procedure there is always exactly one external node attached to each internal node and thus in particular to the parent node of j, we regain the external node where j is inserted and can indeed apply Construction 1. It is now immediate to see that the class of permutations of $\{1, 2, \ldots, n\}$ corresponds bijectively with the family of unordered increasing trees of $\{0, 1, \ldots, n\}$. The family of unordered increasing trees is maybe the most important increasing tree family with applications in many branches and is usually called the family of recursive trees. This bijection between recursive trees and permutations can also be described in an alternative way by using a depth-first walk of the recursive tree (where we assume that the children of each node are arranged in increasing order of their labels): when going around the tree each edge is traversed twice; we encode now the tree by the sequence of labels obtained when encoding only the second traversal of an edge by the label of its incident child node. As a consequence one obtains that node j is exactly a leaf node in the recursive tree if element j is an ascent in the corresponding permutation. We remark that there are many well-known correspondences between recursive trees and permutations; the present one is described, in a different way via the rotation correspondence between binary trees and ordered trees, in [4].

Now we stick to the more general case $k \geq 2$. When starting with a k-Stirling permutation σ we obtain via Construction 1 an increasing tree T with $n + 1$ internal nodes labelled by $\{0, 1, \ldots, n\}$, which additionally has $kn + 1$ external nodes, one is labelled by 0 and, for each $1 \leq j \leq n$, k ones are labelled by j. If we remove all these external nodes from T and just consider the resulting unordered increasing tree, i.e., recursive tree, T', then it is no more possible to reconstruct σ in a unique way. The point is that, although we know for each node the parent node, we do not know which external node attached to the parent node has to be replaced when inserting a particular node via Construction 1. Consider a node v with d children $j_1 < j_2 < \cdots < j_d$: before attaching node j_1 to v there was only one external node attached to v, before attaching node j_2 to v there were k external nodes attached to v, before attaching node j_3 to v there were $2k - 1$ external nodes attached to v, $\ldots$, before attaching node j_d to v there were $(d - 1)(k - 1) + 1$ external nodes attached to v. Thus in order to reconstruct the tree T and so to retrieve the k-Stirling permutation σ, we have to store this information. We do this by introducing the following edge-labelled recursive trees: each r-th child edge gets a label from the set $\{1, 2, \ldots, (k - 1)(r - 1) + 1\}$; here an r-th child edge is an edge $e = (v, w)$, whose corresponding child node w is the r-th (smallest) child amongst all children of the corresponding parent node v. If an edge $e = (v, w)$ has label q then the meaning in Construction 1 is that the child node w is inserted at the q-th external node attached to the parent node v. Thus the following theorem follows.

Theorem 3. *There is a bijective correspondence between the class of k-Stirling permutations of order n and the family of edge-labelled recursive trees of $\{0, 1, \ldots, n\}$, where each r-th child edge gets a label from the set $\{1, 2, \ldots, (k - 1)(r - 1) + 1\}$. Moreover, node j is a leaf node in the edge-labelled recursive tree if and only if the first occurrence of j in the corresponding k-Stirling permutation is an ascent.*

It holds, for each node v in a recursive tree of out-degree d, that there are exactly $\omega_d := \prod_{i=0}^{d-1} \left(1 + (k - 1)i\right)$ possibilities to label the out-going edges of v to obtain a proper edge-labelled recursive tree. Thus if we consider the family $\mathcal{R}$ of weighted recursive trees (i.e., weighted unordered increasing trees), where each node of degree d gets a weight ω_d and the weight of a tree is simply the product of the weights of the nodes in the tree, then $\mathcal{R}$ can be described by the following formal equation (recall that we are considering unordered trees and that 0 is contained in the label set of the nodes):

$$\mathcal{R} = \textcircled{0} \times \left(\omega_0 \cdot \{\epsilon\} \mathbin{\dot\cup} \frac{\omega_1}{1!} \cdot \mathcal{R} \mathbin{\dot\cup} \frac{\omega_2}{2!} \cdot \mathcal{R}^2 \mathbin{\dot\cup} \frac{\omega_3}{3!} \cdot \mathcal{R}^3 \mathbin{\dot\cup} \cdots\right). \tag{5}$$

Furthermore, it holds that the distribution of the number of leaves amongst the trees of order $n + 1$ of $\mathcal{R}$ is equal to the distribution of the number of ascents in the class of k-Stirling permutations of order n. Thus the bijection given here gives an answer to the question raised in [21] concerning a combinatorial meaning of (5).

It is easy to see that analogous considerations can be made also for arbitrary generalized Stirling permutations, where we again introduce edge-labelled recursive trees; but now the size of the set of allowed labels for each r-th child edge $e = (v, w)$ depends on the $r - 1$ brothers of w with a label smaller than w.

Theorem 4. *There is a bijective correspondence between the class of generalized Stirling permutations of the multiset $\{1^{k_1}, 2^{k_2}, \ldots, n^{k_n}\}$ and the family of edge-labelled recursive trees of $\{0, 1, \ldots, n\}$, where each r-th*
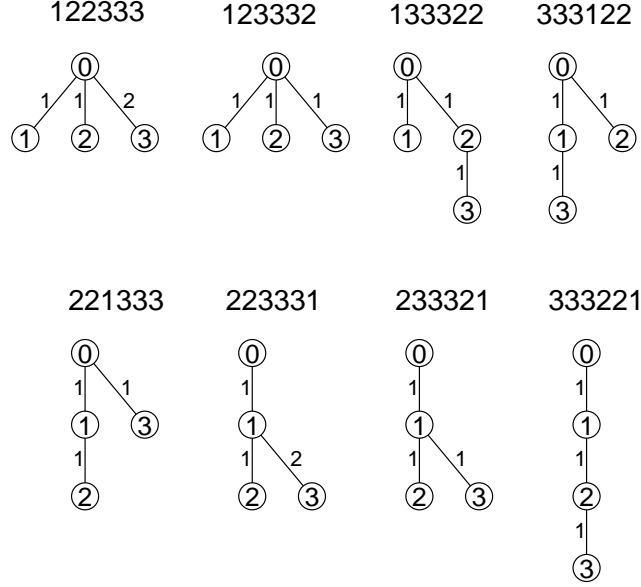
FIGURE 4. All eight generalized Stirling permutations of the multiset $\{1, 2^2, 3^3\}$ and the corresponding edge-labelled recursive trees according to Theorem 4.

child edge $e = (v, w)$, with v the parent node and w the child node, gets a label from the set $\{1, 2, \ldots, 1 + \sum_{i=1}^{r-1} (k_{j_i} - 1)\}$, where $j_1 < j_2 < \cdots < j_{r-1}$ are the labels of the $r - 1$ brothers of w with a label smaller than w. Moreover, node j is a leaf node in the edge-labelled recursive tree if and only if the first occurrence of j in the corresponding generalized Stirling permutation is an ascent.

Theorem 4 is illustrated in Figure 4.

2.3. **Relations between quantities in increasing trees and generalized Stirling permutations.** The bijections stated in Subsection 2.2 often also give a connection between interesting quantities in generalized Stirling permutations and parameters in incrasing trees. In the following we will give such connections, where we restrict ourselves to correspondences between parameters in k-Stirling permutations $\mathcal{Q}_n(k)$ and $(k + 1)$-ary increasing trees $\mathcal{T}_n(k + 1)$ of order n as a consequence of Theorem 1. Sometimes it is here useful to think about $(k + 1)$-ary trees as trees, where external nodes are added to the empty places of the internal nodes.

First we consider the random variable L_n, i.e., the number of left-to-right minima in a random k-Stirling permutation of order n. We introduce now the random variable $\tilde{L}_n$, which counts the length of the path (i.e., the number of edges) from the root to the leftmost external node in a random $(k + 1)$-ary tree of order n. Due to Theorem 1 it holds that $L_n(\sigma) = \tilde{L}_n(T)$, for corresponding $\sigma \in \mathcal{Q}_n(k)$ and $T \in \mathcal{T}_n(k + 1)$.

For M_n, i.e., the r.v. counting the number of left-to-right maxima in a random object of $\mathcal{Q}_n(k)$, there is not such a direct correspondence to a quantity in $\mathcal{T}_n(k + 1)$. However, we give a description by using the so-called inorder traversal code of a $(k + 1)$-ary tree. The inorder traversal code $S(T)$ of a $(k + 1)$-ary tree T is a sequence containing exactly the (internal) nodes of T, which is obtained via the following recursive definition (here $\square$ denotes an external node and ϵ the empty sequence):

$$S(\square) = \epsilon \qquad \text{and} \qquad S(T) = S(T_1) \, r \, S(T_2) \, S(T_3) \, \ldots \, S(T_{k+1}),$$

if the tree T consists of a root node r with the sequence of subtrees $T_1, T_2, \ldots, T_{k+1}$. We introduce now the r.v. $\tilde{M}_n$ counting the number of left-to-right maxima of the inorder traversal code of a random object of $\mathcal{T}_n(k + 1)$. Apparently it holds then $M_n(\sigma) = \tilde{M}_n(T)$ for corresponding objects σ and T.

We turn now to the random variable $X_{n,\ell}$ counting the number of blocks of size $k\ell$ in a random k-Stirling permutation of order n. We say that an internal node v in a $(k+1)$-ary tree is a left-right-node if the path from the root node to v exclusively consists of the leftmost or rightmost possible edge at each node, i.e., contains only edges to the first or the $(k+1)$-th branch, but does not contain "inner" edges leading to the second, third, $\ldots$, k-th branch. It is easily seen that node j is forming a block in $\sigma \in Q_n(k)$ iff j is a left-right-node in the corresponding tree $T \in \mathcal{T}_n(k+1)$. We further say that the internal node w is an inner-descendant of node v in a $(k+1)$-ary tree if w is contained either in the second, the third, $\ldots$, or the k-th branch of v. One can then give the following refinement of the previous statement: element j is forming a block of size $k\ell$ iff node j is a left-right-node that has exactly $\ell - 1$ inner descendants. Thus, if we introduce the r.v. $\tilde{X}_{n,\ell}$, which counts for a random $(k+1)$-ary tree of order n the number of left-right-nodes that have exactly ℓ inner-descendants, it apparently holds $X_{n,\ell}(\sigma) = \tilde{X}_{n,\ell-1}(T)$ for corresponding objects σ and T.

Finally we consider the random variable $\Delta_{n;j,p}$, which counts the distance between the p-th and the $(p+1)$-th occurrence of j in a random k-Stirling permutation of order n. We introduce now the r.v. $D_{n;j,p}$, which counts the number of descendants of node j contained in the p-th branch of node j in a random $(k+1)$-ary increasing tree of order n. It holds then $\Delta_{n;j,p}(\sigma) = 1 + kD_{n;j,p+1}(T)$ for corresponding objects σ and T.

Of course, this also implies the following distributional equations.

Theorem 5. *The random variables L_n, M_n, $X_{n,\ell}$ and $\Delta_{n;j,p}$ for quantities in random k-Stirling permutations of order n and the random variables $\tilde{L}_n$, $\tilde{M}_n$, $\tilde{X}_{n,\ell}$ and $D_{n;j,p}$ (with $1 \le \ell \le n$, $1 \le j \le n$ and $1 \le p \le k-1$) for $(k+1)$-ary increasing trees satisfy the following distributional equations:*

$$L_n \overset{(d)}{=} \tilde{L}_n, \quad M_n \overset{(d)}{=} \tilde{M}_n, \quad X_{n,\ell} \overset{(d)}{=} \tilde{X}_{n,\ell-1}, \quad \Delta_{n;j,p} \overset{(d)}{=} 1 + kD_{n;j,p+1},$$

where the latter ones can be readily extended to joint distributions (with $r \in \mathbb{N}$ and $1 \le \ell_1 < \ell_2 < \cdots < \ell_r \le n$):

$$\left(X_{n,\ell_1}, X_{n,\ell_2}, \ldots, X_{n,\ell_r}\right) \overset{(d)}{=} \left(\tilde{X}_{n,\ell_1-1}, \tilde{X}_{n,\ell_2-1}, \ldots, \tilde{X}_{n,\ell_r-1}\right),$$
$$\left(\Delta_{n;j,1}, \Delta_{n;j,2}, \ldots, \Delta_{n;j,k-1}\right) \overset{(d)}{=} \left(1 + kD_{n;j,2}, 1 + kD_{n;j,3}, \ldots, 1 + kD_{n;j,k}\right).$$

Theorem 5 is illustrated in Figure 5.

3. NUMBER OF LEFT-TO-RIGHT MINIMA

Here we consider the number of left-to-right minima L_n in a random k-Stirling permutation of order n. As pointed out in Subsection 2.3, L_n is equidistributed with $\tilde{L}_n$ counting the distance between the root and the leftmost external node in a random $(k+1)$-ary increasing tree of order n. We use this correspondence and introduce the bivariate generating function

$$L(z,v) = \sum_{n \ge 1} \sum_{m \ge 0} T_n(k+1) \, \mathbb{P}\{\tilde{L}_n = m\} \frac{z^n}{n!} v^m,$$

with $T_n(k+1) = Q_n(k)$ and $T(z)$ given by (1) and (3), respectively. Using the symbolic method (see, e.g., [14]) the formal equation (2) of the family $\mathcal{T}$ of $(k+1)$-ary increasing trees, i.e., the decomposition of a tree with respect to the root node, can be translated into a differential equation for $L(z,v)$. One just has to take into account that the length of the path from the root to the leftmost external node is one plus the length of the corresponding path in the leftmost branch of the root node. This immediately gives the following differential equation (note that the derivative of $L(z,v)$ with respect to z on the left side of the equation reflects the fact that the root is labelled by 1):

$$\frac{\partial}{\partial z} L(z,v) = v\left(1 + L(z,v)\right) \cdot \left(1 + T(z)\right)^k = \frac{v}{1 - kz}\left(1 + L(z,v)\right), \quad L(0,v) = 0.$$

Solving this differential equation leads to the following solution

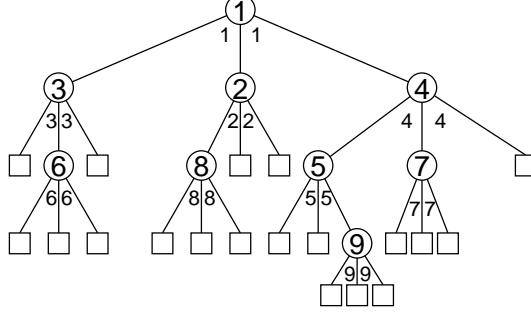$$L(z,v) = \frac{1}{(1 - kz)^{\frac{v}{k}}} - 1, \tag{6}$$

$$366318822155994774$$

FIGURE 5. The 2-Stirling permutation $\sigma = 366318822155994774$ and the corresponding 3-ary increasing tree T. $L_9(\sigma) = \tilde{L}_9(T) = 2$, since 3 and 1 are the only left-to-right minima of σ and the length of the path from the root to the leftmost external node is two. $M_9(\sigma) = \tilde{M}_9(T) = 4$, since $3, 6, 8, 9$ are left-to-right maxima of σ and they are also left-to-right maxima in the inorder-traversal code 361825947 of T. The block decomposition of σ is given by $[3663][188221][55][99][4774]$. Thus the elements $3, 1, 5, 9, 4$ are forming a block of σ, but they are also the left-right nodes of T. There are two blocks of size 4 formed by the elements 3 and 4, respectively, whereas each of the elements 3 and 4 has one inner-descendant. Thus $X_{9,2} = \tilde{X}_{9,1} = 2$. Finally the distance between the two occurrences of 1 is 5, whereas the number of descendants of 1 in the second branch is 2, thus $\Delta_{9;1,1} = 5 = 1 + 2D_{9;1,1}$.

which has been given already in [32]. From (6) we can easily deduce the exact distribution of L_n as well as the limiting distribution, for $n \to \infty$. Using the well-known formula $\frac{1}{(1-z)^v} = \sum_{n \geq 0} \sum_{m \geq 0} \begin{bmatrix} n \\ m \end{bmatrix} \frac{z^n}{n!} v^m$ for the bivariate generating function of the signless Stirling numbers of first kind $\begin{bmatrix} n \\ m \end{bmatrix}$, see, e.g., [14], we immediately obtain an exact formula for the number $L_{n,m} = Q_n(k)\mathbb{P}\{L_n = m\}$ of k-Stirling permutations of order n with exactly m left-to-right minima. It holds, for $n \geq 1$ and $0 \leq m \leq n$:

$$L_{n,m} = n![z^n v^m]L(z,v) = n!k^{n-m}[(kz)^n(\frac{v}{k})^m]\frac{1}{(1-kz)^{\frac{v}{k}}} = k^{n-m}\begin{bmatrix} n \\ m \end{bmatrix},$$

with initial value $L_{0,0} = 1$. To obtain a limiting distribution result for L_n we consider the moment generating function $\mathbb{E}(e^{tL_n}) = \sum_{m \geq 0} \mathbb{P}\{L_n = m\}e^{tm}$, for which we get by using Stirling's formula for the Gamma function:

$$\Gamma(z) = \left(\frac{z}{e}\right)^z \frac{\sqrt{2\pi}}{\sqrt{z}}\left(1 + \mathcal{O}(\frac{1}{z})\right), \tag{7}$$

the following asymptotic expansion, which holds uniformly in a neighbourhood of $t = 0$:

$$\mathbb{E}(e^{tL_n}) = \frac{n!}{Q_n}[z^n]L(z,e^t) = \frac{\binom{n+\frac{e^t}{k}-1}{n}}{\binom{n+\frac{1}{k}-1}{n}} = \frac{\Gamma(\frac{1}{k})}{\Gamma(\frac{e^t}{k})}n^{\frac{e^t-1}{k}} \cdot \left(1 + \mathcal{O}(\frac{1}{n})\right)$$

$$= e^{\left(\frac{e^t-1}{k}\right)\log n + \log\left(\frac{\Gamma(\frac{1}{k})}{\Gamma(\frac{e^t}{k})}\right)} \cdot \left(1 + \mathcal{O}(\frac{1}{n})\right). \tag{8}$$

From expansion (8) a direct application of the so-called quasi power theorem of Hwang [16] shows a central limit theorem for L_n. Additionally exact formulæ for the expectation $\mathbb{E}(L_n)$ and the variance $\mathbb{V}(L_n)$ of L_n can be obtained easily by extracting coefficients from the first and second derivative of $L(z,v)$ with respect to v evaluated at $v = 1$: $\mathbb{E}(L_n) = \frac{[z^n]F_v(z,v)|_{v=1}}{[z^n]T(z)}$ and $\mathbb{V}(L_n) = \frac{[z^n]F_{vv}(z,v)|_{v=1}}{[z^n]T(z)} + \mathbb{E}(L_n) - \left(\mathbb{E}(L_n)\right)^2$. We omit these straightforward computations and collect the results in the following theorem.

Theorem 6. *The number of k-Stirling permutations $L_{n,m} = Q_n(k)\mathbb{P}\{L_n = m\}$ of order n with exactly m left-to-right minima is given by*

$$L_{n,m} = k^{n-m}\begin{bmatrix} n \\ m \end{bmatrix}.$$

Moreover, the number L_n of left-to-right minima in a random k-Stirling permutation of order n is, for $n \to \infty$, asymptotically normal distributed, where the rate of convergence is of order $\mathcal{O}\big(\frac{1}{\sqrt{\log n}}\big)$:

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

and the expectation $\mathbb{E}(L_n)$ and the variance $\mathbb{V}(L_n)$ satisfy

$$\mathbb{E}(L_n) = \frac{1}{k}\Big(\Psi\big(n+\frac{1}{k}\big) - \Psi\big(\frac{1}{k}\big)\Big) = \frac{1}{k}\log n + \mathcal{O}(1),$$

$$\mathbb{V}(L_n) = \frac{1}{k}\Big(\Psi\big(n+\frac{1}{k}\big) - \Psi\big(\frac{1}{k}\big)\Big) + \frac{1}{k^2}\Big(\Psi'\big(n+\frac{1}{k}\big) - \Psi'\big(\frac{1}{k}\big)\Big) = \frac{1}{k}\log n + \mathcal{O}(1).$$

4. The number of left-to-right maxima

A study of the number M_n of left-to-right maxima in a random k-Stirling permutation of order n is more involved than the previous one for L_n. Now the correspondence to $(k+1)$-ary increasing trees and the combinatorial decomposition of a $(k+1)$-ary increasing tree with respect to the root node, i.e., equation (2), does not simplify the problem. Thus we want to treat M_n by considering the insertion process for the random generation of k-Stirling permutations of order n described in Subsection 2.1. In order to do that we have to introduce a refinement of M_n: the random variable $M_{n,i}$, for $0 \leq i \leq kn$, counts the number of left-to-right maxima in the substring $\sigma_1 \sigma_2 \cdots \sigma_i$ of a random k-Stirling permutation $\sigma = \sigma_1 \cdots \sigma_{kn}$ of order n.

We consider now the quantities $T_{n,i,m} = Q_n(k) \cdot \mathbb{P}\{M_{n,i} = m\}$, which give the number of k-Stirling permutations of order n with exactly m left-to-right maxima amongst the first i elements. By distinguishing whether the first occurrence of element n (in the block n^k) is amongst the first i elements of a k-Stirling permutation $\sigma = \sigma_1 \cdots \sigma_{kn}$ (then the number of left-to-right maxima amongst the first i elements of σ is one plus the number of left-to-right maxima amongst the first ℓ elements of σ', if $\sigma_{\ell+1}$ is the first occurrence of n in σ and where σ' denotes the k-Stirling permutation of order $n-1$ obtained after removing the block n^k from σ) or not (then the number of left-to-right maxima amongst the first i elements in σ is equal to the corresponding number in σ'), we obtain the following recurrence for $T_{n,i,m}$:

$$T_{n,i,m} = \big(k(n-1) - i + 1\big)T_{n-1,i,m} + \sum_{\ell=0}^{i-1} T_{n-1,\ell,m-1}, \quad \text{for } n \geq 1 \text{ and } 0 \leq i \leq k(n-1)+1, \quad (9a)$$

$$T_{n,i,m} = \sum_{\ell=0}^{kn-n} T_{n-1,\ell,m-1}, \quad \text{for } n \geq 1 \text{ and } k(n-1)+1 < i \leq kn, \quad (9b)$$

with initial values $T_{0,0,m} = \delta_{0,m}$. Plugging $i = k(n-1)+1$ into (9a) shows, by comparing with recurrence (9b), that $T_{n,i,m} = T_{n,k(n-1)+1,m}$, for $k(n-1)+1 \leq i \leq kn$ and $n \geq 1$; thus we can restrict ourselves to a study of $T_{n,i,m}$, with $0 \leq i \leq k(n-1)+1$. In a probabilistic language this implies that $M_{n,i} = M_{n,k(n-1)+1}$, for $k(n-1)+1 \leq i \leq kn$, and in particular that $M_n = M_{n,kn} = M_{n,k(n-1)+1}$.

By a differencing argument (replacing i by $i-1$ and subtracting the resulting equation from the original one) we obtain from (9a) the following recurrence (for $n \geq 1$ and $1 \leq i \leq k(n-1)+1$):

$$T_{n,i,m} - T_{n,i-1,m} = (kn - k + 1 - i)\big(T_{n-1,i,m} - T_{n-1,i-1,m}\big) + T_{n-1,i-1,m-1} - T_{n-1,i-1,m}. \quad (10)$$

This recurrence seems difficult to treat (e.g., a generating functions approach leads to partial differential equations with unknown boundary values). We continue by introducing the numbers $S_{n,i} := \sum_{m \geq 0} m T_{n,i,m} = Q_n(k) \cdot \mathbb{E}(M_{n,i})$ for a study of the expectation $\mathbb{E}(M_{n,i})$. We note that $\sum_{m \geq 0} T_{n,i,m} = Q_n(k) = (k(n-1)+1)!^{(k)}$, for $0 \leq i \leq kn$. Thus we obtain from (10) the following recurrence for $S_{n,i}$, for $n \geq 1$ and $1 \leq i \leq k(n-1)+1$, with $S_{n,0} = 0$:

$$S_{n,i} - S_{n,i-1} = (kn - k + 1 - i)\big(S_{n-1,i} - S_{n-1,i-1}\big) + (k(n-2)+1)!^{(k)}. \quad (11)$$

This first order linear recurrence (11) for the difference $S_{n,i} - S_{n,i-1}$ can be treated by standard methods and gives the solution:

$$S_{n,i} - S_{n,i-1} = (kn - k + 1 - i)!^{(k)} \sum_{\ell=0}^{n-1-\lceil \frac{i-1}{k} \rceil} \frac{(kn - 2k + 1 - k\ell)!^{(k)}}{(kn - k + 1 - i - k\ell)!^{(k)}}. \tag{12}$$

Iterating (12) leads then to the following exact solution for the numbers $S_{n,i}$, for $0 \le i \le kn - k + 1$:

$$S_{n,i} = \sum_{q=0}^{i} \sum_{\ell=0}^{n-1-\lceil \frac{q-1}{k} \rceil} \frac{(kn - k + 1 - q)!^{(k)} (kn - 2k + 1 - k\ell)!^{(k)}}{(kn - k + 1 - q - k\ell)!^{(k)}}. \tag{13}$$

Using elementary relations between the multifactorials and the ordinary factorials equation (13) can be written as follows:

$$S_{n,i} = (kn - 2k + 1)!^{(k)} \sum_{q=1}^{i} \sum_{\ell=0}^{n-1-\lceil \frac{q-1}{k} \rceil} \frac{\binom{n-1+\frac{1-q}{k}}{\ell}}{\binom{n-2+\frac{1}{k}}{\ell}}. \tag{14}$$

The inner sum can be evaluated, since the summand has a hypergeometric term antidifference, and we obtain

$$S_{n,i} = (kn - 2k + 1)!^{(k)} \sum_{q=1}^{i} \left(\frac{k(n-1)+1}{q} - \frac{\left(k(n-1)+1-k(n-\lceil \frac{q-1}{k} \rceil)\right)\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-2+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}} \right).$$

Simple manipulations lead then to the following expression for $S_{n,i}$, $0 \le i \le kn - k + 1$:

$$S_{n,i} = (kn - k + 1)!^{(k)} H_i - (kn - 2k + 1)!^{(k)} \sum_{q=1}^{i} \frac{\left(k\lceil \frac{q-1}{k} \rceil - k + 1\right)\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-2+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}$$

$$= (kn - k + 1)!^{(k)} H_i - (kn - k + 1)!^{(k)} \sum_{q=1}^{i} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}$$

$$= Q_n(k) \left(H_i - \sum_{q=1}^{i} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}} \right). \tag{15}$$

We will also give an asymptotic estimate of $S_{n,i}$, which holds uniformly for $0 \le i \le kn - k + 1$ (with more effort refined asymptotic expansions can be obtained, but we refrain from giving them here). Since the sum only contains non-negative terms it obviously holds

$$\sum_{q=1}^{i} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}} \le \sum_{q=1}^{kn-k+1} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}.$$

Furthermore we get

$$\frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}} = \frac{\left(k(n-1)+1-q\right) \cdot \left(k(n-2)+1-q\right) \cdots \left(k\lceil \frac{q-1}{k} \rceil + 1 - q\right)}{q\left(k(n-1)+1\right) \cdot \left(k(n-2)+1\right) \cdots \left(k\lceil \frac{q-1}{k} \rceil + 1\right)}$$

$$= \begin{cases} 0, & q \equiv 1 \mod k, \\ \frac{(k(n-1)+1-q)!^{(k)} \left(k(\lceil \frac{q-1}{k} \rceil - 1) + 1\right)!^{(k)}}{q(k(n-1)+1)!^{(k)}}, & \text{otherwise.} \end{cases}$$

Thus we can write

$$\sum_{q=1}^{kn-k+1} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil \frac{q-1}{k} \rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil \frac{q-1}{k} \rceil}} = \sum_{p=1}^{n-1} \sum_{r=1}^{k-1} \frac{(k(n-p-1)+r)!^{(k)} (k(p-1)+1)!^{(k)}}{(kp+1-r)(k(n-1)+1)!^{(k)}}.$$

The trivial estimates $(k(n-p-1)+r)!^{(k)} \leq (k(n-p-1)+k-1)!^{(k)}$ and $kp+1-r \geq kp$, for $1 \leq r \leq k-1$, give thus

$$\sum_{q=1}^{kn-k+1} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil\frac{q-1}{k}\rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil\frac{q-1}{k}\rceil}} \leq \sum_{p=1}^{n-1} \frac{(k(n-p-1)+k-1)!^{(k)}(k(p-1)+1)!^{(k)}}{p(k(n-1)+1)!^{(k)}} = \sum_{p=1}^{n-1} \frac{\binom{n-p-\frac{1}{k}}{n-p}}{p\binom{n-1+\frac{1}{k}}{n-p}}. \tag{16}$$

The sum occurring in (16) can be estimated easily by dissecting into three parts, $p=1$, $2 \leq p \leq \lfloor\frac{n}{2}-\frac{1}{k}\rfloor$ and $\lfloor\frac{n}{2}-\frac{1}{k}\rfloor < p \leq n-1$ and using monotonicity considerations. We omit these straightforward computations and just give the result:

$$\sum_{q=1}^{kn-k+1} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil\frac{q-1}{k}\rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil\frac{q-1}{k}\rceil}} = \mathcal{O}\left(n^{-\frac{2}{k}}\right) + \mathcal{O}\left(n^{-1}\right). \tag{17}$$

Actually the $\mathcal{O}(n^{-1})$ term can be removed, since for $k=1$ the whole sum is actually 0. Thus collecting (15) and (17) we obtain via $S_{n,i} = Q_n(k)\mathbb{E}(M_{n,i})$ the following theorem.

Theorem 7. *The expected number $\mathbb{E}(M_{n,i})$ of left-to-right maxima in the substring $\sigma_1\sigma_2\cdots\sigma_i$ of a random k-Stirling permutation $\sigma = \sigma_1\cdots\sigma_{kn}$ of order n is given by the following expression, where the asymptotic estimate holds for $n \to \infty$ and uniformly for $0 \leq i \leq kn-k+1$:*

$$\mathbb{E}(M_{n,i}) = H_i - \sum_{q=1}^{i} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil\frac{q-1}{k}\rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil\frac{q-1}{k}\rceil}} = H_i + \mathcal{O}(n^{-\frac{2}{k}}).$$

For the number M_n of left-to-right maxima of a random k-Stirling permutation of order n holds $M_n = M_{n,kn} = M_{n,kn-k+1}$, which implies

$$\mathbb{E}(M_n) = H_{kn-k+1} - \sum_{q=1}^{kn-k+1} \frac{\binom{n-1-\frac{q-1}{k}}{n-\lceil\frac{q-1}{k}\rceil}}{q\binom{n-1+\frac{1}{k}}{n-\lceil\frac{q-1}{k}\rceil}} = \log n + \log k + \gamma + \mathcal{O}(n^{-\frac{2}{k}}) + \mathcal{O}(n^{-1}).$$

For the asymptotic expansion of $\mathbb{E}(M_n)$ we additionally used the following well-known asymptotic expansion of the harmonic numbers, where $\gamma = 0.57721\ldots$ denotes the Euler-Masceroni constant:

$$H_n = \log n + \gamma + \mathcal{O}(n^{-1}). \tag{18}$$

Although we do not provide a distributional study of M_n we remark that it is possible to show a central limit theorem for M_n by using the so-called "method of moments" (applied later in Section 5) starting with the distributional recurrence (9). Since the centered moments have to be studied the approach requires a lot of computations, which we do not give here.

To end this section we give a result concerning the number of k-Stirling permutations possessing the maximum possible number of left-to-right maxima.

Theorem 8. *The number $T_{n,kn,n}$ of k-Stirling permutations of order n with exactly n left-to-right maxima is equal to the number of k-ary trees and thus given by the generalized Catalan numbers:*

$$T_{n,kn,n} = \frac{1}{kn-n+1}\binom{kn}{n}.$$

Proof. We consider the representation of k-Stirling permutations as $(k+1)$-ary increasing trees. We first note that node j can only contribute as a left-to-right maximum, if the leftmost branch of j is empty. Thus, since each node shall contribute as a left-to-right maximum, it must hold that the leftmost branch of each node in the $(k+1)$-ary increasing tree is empty. It is easily seen that for each such $(k+1)$-ary tree of order n, where the leftmost branch of each node is empty, there exists exactly one increasing labelling with $\{1, 2, \ldots, n\}$, such that all nodes contribute as a left-to-right maximum. This labelling can be described recursively as follows: the root gets the smallest label and all remaining labels are distributed amongst the branches T_2, T_3 $\ldots$, T_{k+1} of the root, such that all labels of the branch T_r are smaller than all labels of the branch T_s, if $r < s$. Of course, the trees obtained are isomorphic to (unlabelled) k-ary trees. □

5. THE NUMBER OF BLOCKS OF A CERTAIN SIZE

5.1. **Results.** We collect here the exact and limiting distribution results obtained for the random variable $X_{n,\ell}$, which counts the number of blocks of size $k\ell$ in a random k-Stirling permutation of order n. Due to the relation $X_{n,\ell} \stackrel{(d)}{=} \tilde{X}_{n,\ell-1}$, where $\tilde{X}_{n,\ell}$ counts the number of left-right-nodes with exactly ℓ inner descendants in a random $(k+1)$-ary increasing tree of order n, corresponding results also hold for $\tilde{X}_{n,\ell}$.

Theorem 9. *The probability mass function of the random variable $X_{n,\ell}$, which counts the number of blocks of size $k\ell$ in a random k-Stirling permutation of order n, is, for $1 \le \ell \le n$ and $m \ge 0$, given by the following exact formula:*

$$\mathbb{P}\{X_{n,\ell} = m\} = \frac{1}{\binom{n-1+\frac{1}{k}}{n}} \sum_{h=0}^{\lfloor \frac{n-m\ell}{\ell} \rfloor} (-1)^h \frac{\binom{\ell-1-\frac{1}{k}}{\ell-1}^{h+m} \binom{m+h}{h} \binom{n-1-(m+h)\ell+\frac{m+1+h}{k}}{n-(m+h)\ell}}{(k\ell)^{h+m}}.$$

The s-th factorial moments $\mathbb{E}(X_{n,\ell}^{\underline{s}}) = \mathbb{E}\big(X_{n,\ell}(X_{n,\ell}-1)\dots(X_{n,\ell}-(s-1))\big)$ are, for integers $s \ge 0$, given by

$$\mathbb{E}(X_{n,\ell}^{\underline{s}}) = \frac{s!}{(k\ell)^s} \left(\frac{\ell-1-\frac{1}{k}}{\ell-1} \right)^s \cdot \frac{\binom{n-\ell s+\frac{s+1}{k}-1}{n-\ell s}}{\binom{n-1+\frac{1}{k}}{n}}.$$

Theorem 10. *The limiting distribution of the random variable $X_{n,\ell}$ is, depending on the growth of $\ell = \ell(n)$ as $n \to \infty$, characterized as follows, where we define $\lambda_{n,\ell}$ via*

$$\lambda_{n,\ell} := \frac{1}{k\ell} \left(\frac{\ell-1-\frac{1}{k}}{\ell-1} \right) n^{\frac{1}{k}}. \tag{19}$$

- *Range (ℓ, n), such that $\lambda_{n,\ell} \to \infty$ as $n \to \infty$. The normalized random variable $X_{n,\ell}/\lambda_{n,\ell}$ converges in distribution to a random variable ζ, which is characterized by its sequence of integer moments:*

$$\frac{X_{n,\ell}}{\lambda_{n,\ell}} \xrightarrow{(d)} \zeta, \quad \text{with} \quad \mathbb{E}(\zeta^s) = (s+1)! \frac{\Gamma(1+\frac{1}{k})}{\Gamma(1+\frac{s+1}{k})}, \quad s \ge 0.$$

Furthermore, ζ has a density function $f(x)$ that can be written as $f(x) = \Gamma(\frac{1}{k}) x^{-k} g(x^{-k})$, $x > 0$, where g is the density function of a positive $\frac{1}{k}$-stable distribution with Laplace transform $e^{-\lambda^{1/k}}$; thus $f(x)$ is given by the series expansion

$$f(x) = \frac{\Gamma(\frac{1}{k})}{\pi} \sum_{j=1}^{\infty} (-1)^{j-1} \frac{\Gamma(\frac{j}{k}+1) \sin \frac{j\pi}{k}}{j!} x^j, \quad x > 0.$$

- *Range (ℓ, n), such that $\lambda_{n,\ell} \to \lambda \in \mathbb{R}^+ \cup \{0\}$ as $n \to \infty$. The random variable $X_{n,\ell}$ converges in distribution to a random variable ϑ_λ, which is characterized by its sequence of integer moments:*

$$X_{n,\ell} \xrightarrow{(d)} \vartheta_\lambda, \quad \text{with} \quad \mathbb{E}(\vartheta_\lambda^s) = \sum_{q=0}^{s} \begin{Bmatrix} s \\ q \end{Bmatrix} (q+1)! \frac{\Gamma(1+\frac{1}{k})}{\Gamma(1+\frac{q+1}{k})} \lambda^q, \quad s \ge 0.$$

Furthermore, let $\psi_\zeta(t) = \mathbb{E}(e^{t\zeta})$ and $\varphi_{\vartheta_\lambda}(t) = \mathbb{E}(e^{t\vartheta_\lambda})$ denote the moment generating functions of the random variables ζ and ϑ_λ; then both random variables are related as follows:

$$\varphi_{\vartheta_\lambda}(t) = \psi_\zeta\Big(\lambda \cdot (e^t - 1)\Big).$$

Remark 1. Note that the random variable ζ already appeared in the characterization of the limiting distribution of the total number of blocks in random k-Stirling permutations, see [21]. For the instance $\lambda = 1$ the moments of ζ and $\vartheta = \vartheta_1$ are related by a particularly appealing classical pair of combinatorial relations:

$$\mathbb{E}(\vartheta^s) = \sum_{q=0}^{s} \begin{Bmatrix} s \\ q \end{Bmatrix} \mathbb{E}(\zeta^q), \quad \mathbb{E}(\zeta^s) = \sum_{q=0}^{s} (-1)^{s-q} \begin{bmatrix} s \\ q \end{bmatrix} \mathbb{E}(\vartheta^q).$$

Furthermore, we want to point out that a phase change similar to the one presented here occurred in the study of certain "diminishing" urn models, see [24], where a corresponding relation connects a gamma distribution with a geometric distribution.

5.2. **Deriving the explicit formulas.** We use the correspondence between k-Stirling permutations and $(k+1)$-ary increasing trees and study the r.v. $\tilde{X}_{n,\ell}$ by introducing the bivariate generating function

$$T_\ell(z, v) = \sum_{n \geq 1} \sum_{m \geq 0} T_n(k+1) \, \mathbb{P}\{\tilde{X}_{n,\ell} = m\} \frac{z^n}{n!} v^m.$$

We consider the decomposition of a $(k + 1)$-ary increasing tree w.r.t. the root node as given by the formal equation (2) and derive from it a differential equation for $T_\ell(z, v)$. To do this we take into account that, apart from the root node, all left-right-nodes with ℓ inner descendants can be found in the leftmost and rightmost branch of the root, which give the contribution $(1 + T_\ell(z, v))^2 (1 + T(z))^{k-1}$, where $T(z)$ is given by (3). But if the root has exactly ℓ inner descendants then it also contributes to $\tilde{X}_{n,\ell}$, which gives the following correction term: $(v - 1)z^\ell [z^\ell](1 + T(z))^{k-1}$. Thus we obtain the following differential equation:

$$\frac{\partial}{\partial z} T_\ell(z, v) = (1 + T_\ell(z, v))^2 \Big((1 + T(z))^{k-1} - (1 - v)z^\ell \kappa_\ell \Big), \quad T(0, v) = 0,$$

with κ_ℓ given by

$$\kappa_\ell = [z^\ell](1 + T(z))^{k-1} = k^\ell \binom{\ell - \frac{1}{k}}{\ell}.$$

This differential equation can be solved by separation of variables, which leads to the exact solution

$$T_\ell(z, v) = \frac{1}{(1 - kz)^{\frac{1}{k}} + (1 - v)k^\ell \binom{\ell - \frac{1}{k}}{\ell} \frac{z^{\ell+1}}{\ell+1}} - 1. \tag{20}$$

We use now the exact solution (20) to obtain our exact results for the probabilities and the factorial moments of $X_{n,\ell}$. We use that, for $n \geq 1$, it holds:

$$\mathbb{P}\{X_{n,\ell} = m\} = \frac{n!}{Q_n(k)} [z^n v^m] T_{\ell-1}(z, v)$$

$$= \frac{n!}{Q_n(k)} [z^n v^m] \frac{1}{\Big((1 - kz)^{\frac{1}{k}} + k^{\ell-1} \binom{\ell-1-\frac{1}{k}}{\ell-1} \frac{z^\ell}{\ell}\Big) \Big(1 - \frac{vk^{\ell-1} \binom{\ell-1-\frac{1}{k}}{\ell-1} z^\ell}{\ell\big((1-kz)^{\frac{1}{k}} + k^{\ell-1} \binom{\ell-1-\frac{1}{k}}{\ell-1} \frac{z^\ell}{\ell}\big)}\Big)}.$$

Consequently, we obtain

$$\mathbb{P}\{X_{n,\ell} = m\} = \frac{n! k^{m(\ell-1)} \binom{\ell-1-\frac{1}{k}}{\ell-1}^m}{\ell^m Q_n(k)} [z^{n-m\ell}] \frac{1}{\Big((1 - kz)^{\frac{1}{k}} + k^{\ell-1} \binom{\ell-1-\frac{1}{k}}{\ell-1} \frac{z^\ell}{\ell}\Big)^{m+1}}$$

$$= \frac{n! k^{m(\ell-1)} \binom{\ell-1-\frac{1}{k}}{\ell-1}^m}{\ell^m Q_n(k)} \sum_{h=0}^{\lfloor \frac{n-m\ell}{\ell} \rfloor} (-1)^h \frac{k^{h(\ell-1)} \binom{\ell-1-\frac{1}{k}}{\ell-1}^h \binom{m+h}{h} k^{n-(m+h)\ell} \binom{n-1-(m+h)\ell + \frac{m+1+h}{k}}{n-(m+h)\ell}}{\ell^h},$$

which leads, by using the explicit result (1) for $Q_n(k)$, to the formula stated in Theorem 9.

For the factorial moments we proceed as before, but use the relation

$$\mathbb{E}(X_{n,\ell}^{\underline{s}}) = \frac{s! n!}{Q_n(k)} [z^n w^s] T_{\ell-1}(z, 1 + w).$$

One obtains

$$[z^n w^s] T(z, 1 + w) = \frac{k^{n-s} \binom{\ell-1-\frac{1}{k}}{\ell-1}^s}{\ell^s} \binom{n - \ell s + \frac{s+1}{k} - 1}{n - \ell s},$$

which leads to the result given in Theorem 9.

5.3. **Deriving the limiting distributions.** To show the limiting distribution results given in Theorem 10 we use the so-called method of moments, i.e., the Theorem of Fréchet and Shohat (see, e.g., [25]), which works as follows. Let $(X_n)_n$ be a sequence of random variables; if one can show that $\mathbb{E}(X_n^s) \to \mathbb{E}(X^s)$, for each integer $s \geq 0$, where X has to be a random variable that is uniquely defined by its sequence of positive integer moments (guaranteed if the moment sequence satisfies certain growth conditions), then it holds that $X_n \xrightarrow{(d)} X$.

First assume that a sequence (ℓ, n) is given, such that $\lambda_{n,\ell} \to \infty$, with $\lambda_{n,\ell}$ given in (19). It is immediate to see that this implies $\ell = o(n^{\frac{1}{k+1}})$. We consider the explicit formula for the factorial moments $\mathbb{E}(X_{n,\ell}^{\underline{s}})$ as given in Theorem 9 and apply Stirling's formula (7). We obtain, for α fixed, the asymptotic expansion

$$\binom{n+\alpha}{n} = \frac{n^\alpha}{\Gamma(\alpha+1)}\Big(1+\mathcal{O}(\frac{1}{n})\Big),$$

which gives

$$\mathbb{E}(X_{n,\ell}^{\underline{s}}) = \frac{s!}{(k\ell)^s}\binom{\ell-1-\frac{1}{k}}{\ell-1}^s \cdot \frac{(n-\ell s)^{\frac{s+1}{k}-1}\Gamma(\frac{1}{k})}{\Gamma(\frac{s+1}{k})n^{\frac{1}{k}-1}}\Big(1+\mathcal{O}(\frac{1}{n-\ell s})\Big)$$

$$= (s+1)!\frac{\Gamma(1+\frac{1}{k})}{\Gamma(1+\frac{s+1}{k})}\lambda_{n,\ell}^s\Big(1+\mathcal{O}(\frac{\ell}{n})\Big). \tag{21}$$

Using the relation

$$\mathbb{E}(X_{n,\ell}^s) = \mathbb{E}(X_{n,\ell}^{\underline{s}}) + \sum_{q=0}^{s-1}\begin{Bmatrix}s\\q\end{Bmatrix}\mathbb{E}(X_{n,\ell}^q) \tag{22}$$

between the factorial moments and the ordinary moments together with the asymptotic expansion (21) of the factorial moments proves convergence of the (ordinary) moments of the scaled random variable $\frac{X_{n,\ell}}{\lambda_{n,\ell}}$. It can be shown easily by elementary growth estimates that the arising moment sequence determines a distribution uniquely; thus the theorem of Fréchet and Shohat is applicable. Moreover, random variables giving such a sequence of integer moments have been studied previously by Janson [19] (see also Janson et al. [21]) and we can use results obtained there to characterize the density function of the distribution.

Now assume that a sequence (ℓ, n) is given, such that $\lambda_{n,\ell} \to \lambda \in \mathbb{R}^+ \cup \{0\}$, with $\lambda_{n,\ell}$ given in (19). This implies that $\ell \sim c \cdot n^{\frac{1}{k+1}}$, with $c > 0$. Proceeding exactly as before, we obtain the asymptotic expansion

$$\mathbb{E}(X_{n,\ell}^{\underline{s}}) = (s+1)!\frac{\Gamma(1+\frac{1}{k})}{\Gamma(1+\frac{s+1}{k})}\lambda_{n,\ell}^s\Big(1+\mathcal{O}(\frac{\ell}{n})\Big) \sim (s+1)!\frac{\Gamma(1+\frac{1}{k})}{\Gamma(1+\frac{s+1}{k})}\lambda^s.$$

Hence, by using the relation between the factorial moments and the ordinary moments (22), we obtain the moment sequence stated in Theorem 10. In order to show that the arising sequence of moments defines a distribution uniquely, we compute the corresponding moment generating function $\varphi_{\vartheta_\lambda}(t) = \mathbb{E}(e^{t\vartheta_\lambda})$ and show that it is analytic around $t = 0$ (which is a well-known sufficient criterion). We get

$$\varphi_{\vartheta_\lambda}(t) = \sum_{s\geq 0}\mathbb{E}(\vartheta_\lambda^s)\frac{t^s}{s!} = \sum_{s\geq 0}\sum_{q=0}^s\begin{Bmatrix}s\\q\end{Bmatrix}\mathbb{E}(\zeta^q)\lambda^q\frac{t^s}{s!} = \sum_{q\geq 0}\lambda^q\mathbb{E}(\zeta^q)\sum_{s\geq q}\begin{Bmatrix}s\\q\end{Bmatrix}\frac{t^s}{s!}.$$

Using the bivariate generating function identity of the Stirling numbers of the second kind:

$$\sum_{n\geq 0}\sum_{m\geq 0}\begin{Bmatrix}n\\m\end{Bmatrix}\frac{t^n}{n!}u^m = e^{u(e^t-1)} = \sum_{m\geq 0}(e^t-1)^m\frac{u^m}{m!},$$

we obtain further

$$\varphi_{\vartheta_\lambda}(t) = \sum_{q\geq 0}\mathbb{E}(\zeta^q)\frac{\lambda^q(e^t-1)^q}{q!}.$$

Since the latter expression is the moment generating function $\psi_\zeta(u) = \mathbb{E}(e^{u\zeta})$ of the random variable ζ evaluated at $u = \lambda(e^t-1)$, we obtain that $\varphi(t)$ is indeed analytic around $t = 0$, which completes the proof.

5.4. **The joint distribution of the number of blocks of specified sizes.** The generating function approach presented above can be extended to study even the joint distribution of the number of blocks of sizes $k\ell_1, \ldots, k\ell_r$. We find it of interest to note that such an extension is indeed possible and we will state the corresponding results; however, we will only sketch the proof of the results, since the ideas and the actual computations are straightforward extensions of the ones given in Subsections 5.2–5.3.

Here it is appropriate to use a vector notation. We denote with $\mathbf{X_l} = (X_{n,\ell_1}, X_{n,\ell_2}, \ldots, X_{n,\ell_r})$ the random vector counting the joint distribution of the number of blocks of sizes $k\mathbf{l}$, with $\mathbf{l} = (\ell_1, \ldots, \ell_r)$, and with $\tilde{\mathbf{X}}_\mathbf{l} = (\tilde{X}_{n,\ell_1}, \tilde{X}_{n,\ell_2}, \ldots, \tilde{X}_{n,\ell_r})$ the random vector counting the joint distribution of the number of left-right nodes that have exactly $\ell_1, \ldots, \ell_r$ inner descendants, with $1 \le \ell_1 < \cdots < \ell_r \le n$. Note that $\mathbf{X_l} \overset{(d)}{=} \tilde{\mathbf{X}}_{\mathbf{l-1}}$. Furthermore, we use $\mathbf{v} = (v_1, \ldots, v_r)$ and $\mathbf{s} = (s_1, \ldots, s_r)$.

We introduce now the multivariate generating function

$$T_\mathbf{l}(z, \mathbf{v}) = \sum_{n \ge 1} \sum_{\mathbf{m} \ge \mathbf{0}} T_n(k+1) \, \mathbb{P}\{\tilde{\mathbf{X}}_\mathbf{l} = \mathbf{m}\} \frac{z^n}{n!} \mathbf{v^m}.$$

Using the tree decomposition (2) with respect to the root node we obtain the following differential equation for $T_\mathbf{l}(z, \mathbf{v})$:

$$\frac{\partial}{\partial z} T_\mathbf{l}(z, \mathbf{v}) = (1 + T_\mathbf{l}(z, \mathbf{v}))^2 \Big((1 + T(z))^{k-1} - \sum_{i=1}^{r} z^{\ell_i} k^{\ell_i} \binom{\ell_i - \frac{1}{k}}{\ell_i} (1 - v_i) \Big),$$

with initial condition $T_\mathbf{l}(0, \mathbf{v}) = 0$. Separation of variables gives the explicit solution

$$T_\mathbf{l}(z, \mathbf{v}) = \frac{1}{(1 - kz)^{\frac{1}{k}} + \sum_{i=1}^{r} (1 - v_i) k^{\ell_i} \binom{\ell_i - \frac{1}{k}}{\ell_i} \frac{z^{\ell_i+1}}{\ell_i + 1}} - 1.$$

We refrain from deriving the joint distribution of $\mathbf{X_l}$ and proceed directly to the computation of the mixed factorial moments, where we use the identity

$$\mathbb{E}(\mathbf{X_l^s}) := \mathbb{E}\Big(\prod_{q=1}^{r} X_{n,\ell_q}^{s_q} \Big) = \frac{\big(\prod_{q=1}^{r} s_q! \big) n!}{Q_n(k)} [z^n \mathbf{w^s}] T_{\mathbf{l-1}}(z, \mathbf{1} + \mathbf{w}).$$

It is not difficult to extract the coefficients

$$[z^n \mathbf{w^s}] T_{\mathbf{l-1}}(z, \mathbf{1} + \mathbf{w}) = [z^n w_1^{s_1} \ldots w_r^{s_r}] \frac{1}{(1 - kz)^{\frac{1}{k}} - \sum_{i=1}^{r} w_i k^{\ell_i - 1} \binom{\ell_i - 1 - \frac{1}{k}}{\ell_i - 1} \frac{z^{\ell_i}}{\ell_i}} - 1,$$

by using the identity

$$[w_1^{s_1} \ldots w_r^{s_r}] \frac{1}{1 - w_1 \alpha_1 - \cdots - w_r \alpha_r} = \frac{\big(\sum_{q=1}^{r} s_q \big)!}{\prod_{q=1}^{r} (s_q!)} \prod_{q=1}^{r} \alpha_q^{s_q}.$$

We eventually obtain the following theorem, where the limiting distributions are again derived by using the method of moments. The computations occurring are very similar to the univariate case and are therefore omitted.

Theorem 11. *The mixed factorial moments of the random vector $\mathbf{X_l}$, which counts the joint distribution of the number of blocks of sizes $k\mathbf{l}$, with $\mathbf{l} = \big(\ell_1, \ldots, \ell_r \big)$, in a random k-Stirling permutation of order n, are given by*

$$\mathbb{E}(\mathbf{X_l^s}) = \big(\sum_{q=1}^{r} s_q \big)! \left(\prod_{q=1}^{r} \Big(\frac{1}{k\ell_q} \binom{\ell_q - 1 - \frac{1}{k}}{\ell_q - 1} \Big)^{s_q} \right) \frac{\binom{n - \sum_{q=1}^{r} \ell_q s_q + \frac{1 + \sum_{q=1}^{r} s_q}{k} - 1}{n - \sum_{q=1}^{r} \ell_q s_q}}{\binom{n - 1 + \frac{1}{k}}{n}}.$$

The limiting distribution of $\mathbf{X_l}$ is, depending on the growth of $\ell_q = \ell_q(n)$, with $1 \le q \le r$, as $n \to \infty$, characterized as follows, where we define λ_{n,ℓ_q} as before:

$$\lambda_{n,\ell_q} := \frac{1}{k\ell_q} \binom{\ell_q - 1 - \frac{1}{k}}{\ell_q - 1} n^{\frac{1}{k}}.$$

Assume that sequences (ℓ_q, n) are given such that $\lambda_{n,\ell_1}, \ldots, \lambda_{n,\ell_h} \to \infty$, with $1 \leq h \leq r$, and $\lambda_{n,\ell_i} \to \lambda_i \in \mathbb{R}^+$, with $h+1 \leq i \leq r$, as $n \to \infty$. The properly normalized random vector $\hat{\mathbf{X}}_1 = \left(\frac{X_{n,\ell_1}}{\lambda_{n,\ell_1}}, \ldots, \frac{X_{n,\ell_h}}{\lambda_{n,\ell_h}}, X_{n,\ell_{h+1}}, \ldots, X_{n,\ell_r} \right)$ converges in distribution to a limiting random vector ϑ, which is characterized by its sequence of mixed (factorial) moments $\mathbf{s} \geq \mathbf{0}$:

$$\hat{\mathbf{X}}_1 \xrightarrow{(d)} \vartheta = (\vartheta_1, \ldots, \vartheta_r), \quad \mathbb{E}\left(\prod_{j=1}^{h} \vartheta_j^{s_j} \prod_{i=h+1}^{r} \vartheta_i^{s_i} \right) = \left(\sum_{h=1}^{r} s_h \right)! \frac{\Gamma(\frac{1}{k})}{\Gamma\left(\frac{1+\sum_{h=1}^{r} s_h}{k} \right)} \prod_{i=h+1}^{r} \lambda_i^{s_i}.$$

5.5. An alternative approach using a description via Pólya-Eggenberger urn models.

In this section we introduce a Pólya-Eggenberger urn model, which also describes the joint distribution of the random variables $X_{n,1}, \ldots, X_{n,\ell}$. For a detailed description of Pólya-Eggenberger urn models see, e.g., [13, 19, 22, 27]; we just give the basic notions for an urn model with balls of r different colours. At step 0 there are $a_{i,0}$, for $1 \leq i \leq r$, balls of colour i in the urn; we might call this the initial configuration of the urn. The urn evolves at discrete time steps according to the ball addition matrix $A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,r} \\ \alpha_{2,1} & \cdots & \alpha_{2,r} \\ \vdots & \cdots & \vdots \\ \alpha_{r,1} & \cdots & \alpha_{r,r} \end{pmatrix}$: at each step a ball is chosen at random from all of the balls in the urn, the colour the ball inspected and afterwards it is thrown back into the urn. If the ball has colour i then balls are added to or removed from the urn as listed in the i-th row of the ball addition matrix A, i.e., $\alpha_{i,j}$ balls of colour j, for $1 \leq j \leq n$, are added/removed. An urn is called balanced if all the row sums of A are equal, i.e., if the total number of added/removed balls in each draw is constant, independently of the colour observed; this implies that, after each draw, the total number of balls in the urn is deterministic.

For balanced urns in [13] a method is described, which, if applicable, leads to the exact distribution of the number of balls of colour i, $1 \leq i \leq r$, contained in the urn after n draws. It is sufficient to study the numbers $H_n \begin{pmatrix} a_{1,0} & a_1 \\ a_{2,0} & a_2 \\ \vdots & \vdots \\ a_{r,0} & a_r \end{pmatrix}$, which give the number of urn histories of length n (i.e., sequences of n draws), such that the urn contains exactly a_i, for $1 \leq i \leq r$, balls of colour i when starting with $a_{i,0}$, $1 \leq i \leq r$, balls. These numbers are encoded in the generating function

$$H(y_1, \ldots, y_r, z) := \sum_{n, a_1, a_2, \ldots, a_r} H_n \begin{pmatrix} a_{1,0} & a_1 \\ a_{2,0} & a_2 \\ \vdots & \vdots \\ a_{r,0} & a_r \end{pmatrix} \cdot y_1^{a_1} y_2^{a_2} \cdots y_r^{a_r} \frac{z^n}{n!}.$$

Crucial to the approach is a study of the associated ordinary system of differential equations

$$\Sigma : \quad \dot{y}_i = y_i^{\alpha_{i,i}+1} \prod_{1 \leq j \leq r, j \neq i} y_j^{\alpha_{i,j}},$$

where the functions $y_i = y_i(t)$ are assumed to be functions of an independent variable t. Let us assume that the solution of the associated system of differential equations Σ with initial conditions $y_{1,0}, y_{2,0}, \ldots, y_{r,0}$ is given as follows:

$$Y_1\left(t \,\middle|\, \begin{matrix} y_{1,0} \\ y_{2,0} \\ \vdots \\ y_{r,0} \end{matrix} \right), \ldots, Y_r\left(t \,\middle|\, \begin{matrix} y_{1,0} \\ y_{2,0} \\ \vdots \\ y_{r,0} \end{matrix} \right).$$

Then the generating function $H(y_1, \ldots, y_r, z)$ of the urn histories is given as follows:

Theorem 12 (Flajolet, Dumas and Puyhaubert [13]).

$$H(y_1, \ldots, y_r, z) = \left(Y_1\left(z \,\middle|\, \begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{matrix} \right) \right)^{a_{1,0}} \cdot \left(Y_2\left(z \,\middle|\, \begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{matrix} \right) \right)^{a_{2,0}} \cdots \left(Y_r\left(z \,\middle|\, \begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{matrix} \right) \right)^{a_{r,0}}.$$

Of course, the exact joint distribution of the number of balls of colour i, $1 \leq i \leq r$, can be obtained from $H(y_1, \ldots, y_r, z)$ by extracting coefficients.

Now we go back to the problem of describing the joint distribution of the number of blocks of specified sizes by means of an urn model. To do this we introduce the following urn.

Urn I. Consider a balanced urn with balls of $\ell + 2$ colours and let $(Z_{n,0}, \ldots, Z_{n,\ell+1})$ count the number of balls of each color at time n. At each time step draw one ball at random from the urn, observe its color, and add balls according to the ball addition matrix

$$
A = \begin{pmatrix}
1 & k-1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & -(k-1) & 2k-1 & \ddots & \ddots & & \ddots & 0 & 0 \\
0 & 0 & -(2k-1) & 3k-1 & \ddots & & \ddots & 0 & 0 \\
\vdots & \ddots & & \ddots & \ddots & \ddots & & \ddots & \vdots & 0 \\
\vdots & \ddots & & & \ddots & \ddots & \ddots & & \vdots & 0 \\
0 & \ddots & & \ddots & & \ddots & 0 & -((\ell-1)k-1) & \ell k-1 & 0 \\
0 & \ddots & & \ddots & & \ddots & 0 & 0 & -(\ell k-1) & (\ell+1)k-1 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & k
\end{pmatrix}.
$$

The initial configuration of the urn (it is here convenient to start at time 1) is specified by $(Z_{1,0}, \ldots, Z_{1,\ell+1}) = (2, k-1, 0, \ldots, 0)$.

Note that, at each time step, a total number of k balls will be added to the urn. We further remark that we opted to introduce balls of colour $0, 1, \ldots, \ell + 1$; the reason for that will be obvious after the following statement.

Theorem 13. *The random variables $Z_{n,i}$, with $1 \leq i \leq \ell$, described by urn model I are related to the random variables $X_{n,i}$, $1 \leq i \leq \ell$, which count the number of blocks of size ki in a random k-Stirling permutation of order n, as follows:*

$$
Z_{n,i} = (ki - 1)X_{n,i}, \quad 1 \leq i \leq \ell.
$$

Proof. We give the following coupling between the growth process describing k-Stirling permutations and the urn model. Consider a k-Stirling permutation σ. Then the balls of colour i, with $1 \leq i \leq \ell$, will represent the gaps within blocks of size ki in σ. Moreover, balls of colour $\ell + 1$ will represent the gaps within blocks of sizes $\geq k(\ell + 1)$ in σ, whereas balls of colour 0 will represent the gaps between blocks including the gap before and the gap after σ (thus the number of balls of colour 0 is exactly one plus the number of blocks in σ).

Obviously at step 1, i.e., for $\sigma = 1^k$, there are two balls of colour 0 and $k - 1$ balls of colour 1, which matches the initial configuration. At step n a block of size ki, $i \geq 1$, in σ "attracts" the sequence n^k with probability $(ki-1)/((n-1)k+1)$ and gives then a block of size $k(i+1)$. In the corresponding urn model this is reflected by removing $ki - 1$ balls of colour i and adding $k(i + 1) - 1$ balls of colour $i + 1$, for $1 \leq i \leq \ell$; furthermore, if $i \geq \ell + 1$, then one simply adds k balls of colour $\ell + 1$. Moreover, at step n a new block (which is of size k) is created in σ with a probability proportional to one plus the number of blocks of σ; in the urn model this is reflected by adding one ball of colour 0 and $k - 1$ balls of colour 1. $\square$

Remark 2. As pointed out before the r.v. $Z_{n,0}$ gives one plus the total number of blocks, a quantity that has been already studied in [21]. The idea of introducing an additional random variable $Z_{n,\ell+1}$, which is serving as some kind of "superball' collecting all contributions of larger sizes, has been applied in this context, e.g., by Mahmoud and Smythe [28] or Janson [18].

The evolution of this balanced Urn I can be analyzed by using the approach of Flajolet et al. [13]. One has to study the following associated system of differential equations Σ, where $y_i = y_i(t)$:

$$
\begin{aligned}
\dot{y}_0 &= y_0^2 y_1^{k-1}, & y_0(0) &= y_{0,0}, \\
\dot{y}_i &= y_i^{-(ik-1)+1} y_{i+1}^{(i+1)k-1}, \quad \text{for } 1 \leq i \leq \ell, & y_i(0) &= y_{i,0}, \\
\dot{y}_{\ell+1} &= y_{\ell+1}^{k+1}, & y_{\ell+1}(0) &= y_{\ell+1,0}.
\end{aligned}
$$

First, one easily obtains

$$
y_{\ell+1}(t) = \frac{1}{(-kt + C_{\ell+1})^{\frac{1}{k}}}, \quad C_{\ell+1} = \frac{1}{y_{\ell+1,0}^k}.
$$

Then, by induction with respect to i, it is not difficult to prove the general formula

$$
y_{\ell+1-i}(t) = \left(\sum_{q=1}^{i} C_{\ell+1-q} \frac{t^{i-q}}{(i-q)!} \prod_{p=q+1}^{i} ((\ell+1-p)k - 1) + \frac{1}{(-kt + C_{\ell+1})^{\frac{(\ell+1-i)k-1}{k}}} \right)^{\frac{1}{(\ell+1-i)k-1}}, \quad (23)
$$

for $0 \le i \le \ell$, with

$$C_{\ell+1-i} + C_{\ell+1}^{\frac{-(\ell+1-i)k+1}{k}} = y_{\ell+1-i,0}^{(\ell+1-i)k-1}, \quad \text{and thus} \quad C_{\ell+1-i} = y_{\ell+1-i,0}^{(\ell+1-i)k-1} - y_{\ell+1,0}^{(\ell+1-i)k-1}.$$

Consequently, we easily obtain $y_0(t) = 1/(-\int y_1(t)^{k-1} dt + C_0)$ and furthermore

$$y_0(t) = \frac{1}{C_0 - \sum_{q=1}^{\ell} C_{\ell+1-q} \frac{t^{\ell+1-q}}{(\ell+1-q)!} \prod_{p=q+1}^{\ell} \left((\ell+1-p)k-1\right) + (-kt + C_{\ell+1})^{\frac{1}{k}}},$$

with

$$(C_0 + C_{\ell+1}^{\frac{1}{k}})^{-1} = y_{0,0}, \quad \text{and thus} \quad C_0 = y_{0,0}^{-1} - y_{\ell+1,0}^{-1}.$$

After simple manipulations we observe that the formula for $y_{\ell+1-i}(t)$ given in (23) is also valid for $i = \ell + 1$.

Applying Theorem 12 we obtain the following result concerning the generating function of the urn histories of Urn I.

Theorem 14. *The $(\ell + 2)$-variable generating function $H(\mathbf{y}, z) = H(y_0, y_1, \ldots, y_{\ell+1}, z)$ of urn histories of Urn I, with general initial conditions $(Z_{1,0}, \ldots, Z_{1,\ell+1}) = (a_{1,0}, a_{1,1}, \ldots, a_{1,\ell+1})$, is given by*

$$H(\mathbf{y}, z) = \prod_{i=0}^{\ell+1} \left(U_i(\mathbf{y}, z) \right)^{a_{1,i}},$$

where the functions $U(\mathbf{y}, z)$ are given by

$$U_{\ell+1-i}(\mathbf{y}, z) = \left(\sum_{q=1}^{i} \left(y_{\ell+1-q}^{(\ell+1-q)k-1} - y_{\ell+1}^{(\ell+1-q)k-1} \right) \frac{z^{i-q}}{(i-q)!} \prod_{p=q+1}^{i} \left((\ell+1-p)k-1\right) \right.$$
$$\left. + \frac{y_{\ell+1}^{(\ell+1-i)k-1}}{(1 - kzy_{\ell+1}^k)^{\frac{(\ell+1-i)k-1}{k}}} \right)^{\frac{1}{(\ell+1-i)k-1}}, \quad \text{for } 0 \le i \le \ell + 1.$$

Moreover, the generating function concerning the distribution of blocks of sizes $k, 2k \ldots, \ell k$ in random k-Stirling permutations is obtained by the initial conditions $(Z_{1,0}, \ldots, Z_{1,\ell+1}) = (2, k-1, 0, \ldots, 0)$ leading to $H(\mathbf{y}, z) = U_0^2(\mathbf{y}, z) U_1^{k-1}(\mathbf{y}, z)$.

Remark 3. We want to point out that an advantage of this approach is that one obtains a result valid for arbitrary initial conditions (chosen in such a way that the urn model is still tenable, i.e., that the urn never gets "stuck"), whereas the previous approach using the description via $(k+1)$-ary increasing trees is only applicable for the particular initial values $(Z_{1,0}, \ldots, Z_{1,\ell+1}) = (2, k-1, 0, \ldots, 0)$. From Theorem 14 it is not difficult to reprove our results for $X_{n,\ell}$ obtained in Theorems 9–11.

6. DISTANCES IN STIRLING PERMUTATIONS AND DESCENDANTS IN INCREASING TREES

In the following we use the notation $\mathbf{\Delta} = (\Delta_{n;j,1}, \ldots, \Delta_{n;j,k-1})$ for the random vector counting the joint distribution of distances of element j, and $\mathbf{D} = (D_{n;j,1}, \ldots, D_{n;j,k}, D_{n;j,k+1})$ for the random vector counting the joint distribution of descendants of node j. Recall from Theorem 5 that the distances $\Delta_{n;j,p}$ between the p-th and the $(p+1)$-th occurrence of element j in a random k-Stirling permutation of order n and the numbers $D_{n;j,p}$ of descendants of node j in the p-th branch of a random $(k+1)$-ary increasing tree of order n are related by the equation $\Delta_{n;j,p} \overset{(d)}{=} 1 + k \cdot D_{n;j,p+1}, 1 \le p \le k - 1$. Thus it holds that the distribution of the random vector $\mathbf{\Delta}$ is fully characterized by the distribution of the more general random vector $\mathbf{D}$. Hence, in this section we will mainly formulate our results for the random vector $\mathbf{D}$, i.e., give results for $(k+1)$-ary increasing trees, but note, that all the results can be translated immediately into corresponding ones for the random vector $\mathbf{\Delta}$, i.e., for k-Stirling permutations.

In order to describe the evolution of $\mathbf{D}$ and $\mathbf{\Delta}$, respectively, we introduce a purely autistic hyper-Pólya urn model (using the terminology of Flajolet et al. [13]).

Urn II. Consider a balanced urn with balls of $k+2$ colors, and let $(W_{n,0}, W_{n,1}, \ldots, W_{n,k+1})$ count the number of balls of each colour at time n. At each time step, draw one ball at random from the urn, observe its colour, and add k balls of the observed colour to the urn, i.e., the urn evolves according to the ball addition matrix

$$A = (k\delta_{i,j})_{1 \leq i,j \leq k+2} = \begin{pmatrix} k & 0 & 0 & \ldots & 0 \\ 0 & k & 0 & \ddots & 0 \\ 0 & 0 & k & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & k \end{pmatrix}.$$

The initial configuration of the urn is specified by $(W_{1,0}, W_{1,1}, \ldots, W_{1,k+1}) = ((j-1)k, 1, 1, 1, \ldots, 1)$.

Then it holds the following theorem for Urn II.

Theorem 15. *The distribution of the random vector $(W_{n,1}, \ldots, W_{n,k+1})$ is related to $\mathbf{D}$ and $\mathbf{\Delta}$ by the equations*

$$(W_{n,1} - 1, \ldots, W_{n,k+1} - 1) = (k \cdot D_{n;j,1}, \ldots, k \cdot D_{n;j,k+1}),$$
$$(W_{n,2} + 1, \ldots, W_{n,k} + 1) = (\Delta_{n;j,1}, \ldots, \Delta_{n;j,k-1}).$$

Consequently, the distribution of $\mathbf{D}$, (and of $\mathbf{\Delta}$) is exchangeable, invariant under any permutation of the $k+1$ variables (or $k-1$ variables when considering $\mathbf{\Delta}$).

Proof. We give the following coupling between the growth process describing $(k+1)$-ary increasing trees and the urn model. Due to the correspondence between $(k+1)$-ary increasing trees and k-Stirling permutations this also gives a coupling with the growth process of k-Stirling permutations. Consider a $(k+1)$-ary increasing tree T. Then the balls of colour p, for $1 \leq p \leq k+1$, will represent the external nodes of T, which are contained in the p-th branch of node j. Moreover, balls of colour 0 will represent the external nodes of T, which are not contained in the subtree rooted at node j.

The initial configuration of the urn corresponds to a tree T of order j, i.e., directly after inserting node j, which leads to exactly one ball of each colour p, with $1 \leq p \leq k+1$, whereas all remaining $k(j-1)+1$ balls are of colour 0. Furthermore, in a tree T of order $n-1$, each branch of node j "attracts" the newly inserted node n with a probability proportional to the number of external nodes contained in this branch. If node n is inserted into a certain branch of node j then the number of external nodes contained in this branch increases by k. Moreover, the probability that node n is not inserted in the subtree rooted at j is proportional to the number of external nodes not contained in one of the branches of j; by inserting node n the number of such external nodes increases by k in this case. But this is exactly what is reflected by the ball addition matrix A. The relation between $W_{n,p}$ and $D_{n;j,p}$ (and also $\Delta_{n;j,p}$) follows then by taking into account that a branch containing exactly m internal nodes is containing $km+1$ external nodes. $\square$

We obtain then the following exact and asymptotic results for the random vector $\mathbf{D}$.

Proposition 1. *The distribution of the random vector $\mathbf{D}$, which counts the number of descendants of node j according to its branches $1, \ldots, k+1$ in a random $(k+1)$-ary increasing tree of order n, is given as follows:*

$$\mathbb{P}\{\mathbf{D} = \mathbf{m}\} = \frac{\binom{m_1 - 1 + \frac{1}{k}}{m_1} \cdots \binom{m_{k+1} - 1 + \frac{1}{k}}{m_{k+1}} \binom{n - 2 - \sum_{i=1}^{k+1} m_k}{j-2}}{\binom{n-1+\frac{1}{k}}{n-j}}, \quad \text{for } \mathbf{m} \geq \mathbf{0} \text{ and } 1 \leq j \leq n.$$

(For the instance $j = 1$ we obtain $\sum_{i=1}^{k+1} m_k = n - 1$, and use the convention $\binom{-1}{-1} = 1$.)

The mixed binomial moments $\mathbb{E}\left(\prod_{p=1}^{k+1} \binom{D_{n;j,p} + \frac{1}{k} + s_p}{s_p}\right)$ are given by the formula

$$\mathbb{E}\left(\prod_{p=1}^{k+1} \binom{D_{n;j,p} + \frac{1}{k} + s_p}{s_p}\right) = \frac{\binom{n-1+\frac{1}{k}+\sum_{p=1}^{k+1} s_p}{\sum_{p=1}^{k+1} s_p}}{\binom{j-1+\frac{1}{k}+\sum_{p=1}^{k+1} s_p}{\sum_{p=1}^{k+1} s_p}} \prod_{p=1}^{k+1} \binom{s_p - 1 + \frac{1}{k}}{s_p}.$$

Theorem 16. *The limiting distribution of the random vector $\mathbf{D}$ can be characterized according to the growth of $j = j(n)$.*

- *Range arbitrary but fixed $j \geq 1$. The normalized random vector $(\mathbf{D}, n - \sum_{p=1}^{k+1} D_{n;j,p})/n$ converges almost surely to a Dirichlet distributed random vector of length $k+2$ with parameters $(\frac{1}{k}, \ldots, \frac{1}{k}, \frac{j-1}{k})$:*

$$\frac{\mathbf{D}}{n} \xrightarrow{(a.s.)} \mathrm{Dir}_{k+2}\left(\frac{1}{k}, \ldots, \frac{1}{k}, \frac{j-1}{k}\right).$$

- *Range $j \to \infty$, such that $j = o(n)$. The normalized random vector $j\mathbf{D}/n$ converges in distribution to a random vector of i.i.d. gamma distributed random variables, $\gamma_i \overset{(d)}{=} \gamma(\frac{1}{k}, 1)$, for $1 \leq i \leq k+1$:*

$$\frac{j\mathbf{D}}{n} \xrightarrow{(d)} (\gamma_1, \ldots, \gamma_{k+1}).$$

- *Range $j \to \infty$, such that $j = \rho n$, with $0 < \rho < 1$. The random vector $\mathbf{D}$ converges in distribution to a random vector of i.i.d. negative binomial distributed random variables, $\mathrm{NegBin}_i \overset{(d)}{=} \mathrm{NegBin}(\frac{1}{k}, \rho)$, for $1 \leq i \leq k+1$:*

$$\mathbf{D} \xrightarrow{(d)} (\mathrm{NegBin}_1, \ldots, \mathrm{NegBin}_{k+1}).$$

- *Range $j \to \infty$, such that $j = n - l$ with $l = o(n)$. The random vector $\mathbf{D}$ is asymptotically degenerate, $\mathbf{D} \xrightarrow{(d)} \mathbf{0}$.*

Remark 4. The explicit result for the distribution of $\mathbf{D}$ is well known in the literature, but the asymptotic results generalize some of earlier results obtained by the authors [23] concerning descendants in increasing trees. Note that the particular instance $j = 1$ describes the asymptotics of the subtree sizes of the subtrees attached to the root node in a $(k+1)$-ary increasing tree, which has been studied in a more general context by Broutin et al. [10]. Interestingly, the case $j \to \infty$ corresponds to an urn model, where the initial conditions depend on the discrete time, as the time tends to infinity.

Proof. We will give here only a brief outline of the proof of Proposition 1 and Theorem 16, where we focus on the main ideas leading to the results, but skip many of the (straightforward) computations.

The exact distribution of $\mathbf{D}$ given in Proposition 1 could be derived by using generating function techniques based on the formal equation of $(k+1)$-ary increasing trees, or alternatively by studying the associated system of differential equations of Urn II. However, one can even use a more direct approach, which is based on lattice path counting arguments similar to the ones used by Prodinger [36]. Crucial to this approach is to observe that the probability $w = w(\mathbf{m})$ to obtain a certain configuration $\mathbf{m} = (m_1, m_2, \ldots, m_{k+1})$ of descendants (in the branches $1, 2, \ldots, k+1$) of node j is independent of the different histories to reach the configuration $\mathbf{m}$. Hence, we simply have to determine w, and the number $H = H(\mathbf{m})$ of different histories leading to the configuration $\mathbf{m}$ using lattice path counting arguments given by [36]. It is not difficult to show the following formulas, which directly leads to the exact distribution of $\mathbf{D}$:

$$\mathbb{P}\{\mathbf{D} = \mathbf{m}\} = w \cdot H, \quad \text{with} \quad H = \binom{n-j}{m_1, \ldots, m_{k+1}, n-j-\sum_{p=1}^{k+1} m_p},$$

$$w = \frac{\left(\prod_{p=1}^{k+1} \prod_{i_p=0}^{m_p-1} (ki_p + 1)\right) \prod_{i=0}^{n-j-\sum_{p=1}^{k+1} m_p+1} k(i+j-1)}{\prod_{i=j}^{n-1} (ik+1)}.$$

The exact moments of $\mathbf{D}$ could be obtained by using the explicit formula for $\mathbb{P}\{\mathbf{D} = \mathbf{m}\}$ and carrying out the summations required. However, the most convenient way to get the simple expressions for the binomial moments is to use martingales. First we only consider a single variable $W_{n,1}$. Let $\mathcal{F}_n$ denote the σ-field generated by the first $n-j$ draws of the urn model (i.e., after inserting node n in the corresponding growth process of $(k+1)$-ary increasing trees). Moreover denote by $\Theta_n = W_{n,1} - W_{n-1,1} \in \{0, k\}$ the increment at step n (i.e., due to inserting node n). We get

$$\mathbb{E}(W_{n,1} \mid \mathcal{F}_{n-1}) = \mathbb{E}(W_{n-1,1} + \Theta_n \mid \mathcal{F}_{n-1}) = W_{n-1,1} + \mathbb{E}(\Theta_n \mid \mathcal{F}_{n-1}).$$

Since the probability that a ball of colour 1 is chosen at step n is proportional to the number of existing balls of colour 1 (at step $n-1$), we obtain further

$$\mathbb{E}(W_{n,1} \mid \mathcal{F}_{n-1}) = W_{n-1,1} + \frac{kW_{n-1,1}}{k(n-1)+1} = \frac{kn+1}{k(n-1)+1}W_{n-1}, \quad n \geq j+1.$$

Let $\tilde{W}_n = W_{n,1}\binom{n-1+\frac{1}{k}}{n-1}/\binom{n+\frac{1}{k}}{n-1} = W_{n,1}(1+\frac{1}{k})/(n+\frac{1}{k})$, then it holds

$$\mathbb{E}(\tilde{W}_n \mid \mathcal{F}_{n-1}) = \tilde{W}_{n-1}, \quad n \geq j+1.$$

Hence, $\tilde{W}_n$ (and also $\tilde{W}_n/k$) is a positive martingale.

The idea can be extended to study the random variables $\mathcal{W}_{n,\mathbf{s}} = \prod_{p=1}^{k+1} \binom{W_{n,p}/k-1+s_p}{s_p}$, where we obtain due to the dynamics of the urn model:

$$\mathbb{E}(\mathcal{W}_{n,\mathbf{s}} \mid \mathcal{F}_{n-1}) = \mathcal{W}_{n-1,\mathbf{s}} + \sum_{i=1}^{k+1} \frac{W_{n-1,i}}{(n-1)k+1} \prod_{p=1}^{k+1} \binom{\frac{W_{n-1,p}}{k} - 1 + s_p}{s_p - \delta_{p,i}}$$

$$= \mathcal{W}_{n-1,\mathbf{s}} + \mathcal{W}_{n-1,\mathbf{s}} \sum_{i=1}^{k+1} \frac{s_i}{(n-1)k+1} = \mathcal{W}_{n-1,\mathbf{s}} \frac{n-1+\frac{1}{k}+\sum_{i=1}^{k+1}ks_i}{n-1+\frac{1}{k}}.$$

Consequently, we observe that the random variables $\tilde{\mathcal{W}}_{n,\mathbf{s}} := \mathcal{W}_{n,\mathbf{s}}\binom{n-1+\frac{1}{k}}{n-1}/\binom{n-1+\sum_{i=1}^{k+1}s_i+\frac{1}{k}}{n-1}$ are forming martingales with respect to the filtration $\mathcal{F}_n$. The results for the binomial moments stated in Proposition 1 are obtained easily by taking the expected value.

Next we outline the proof of the limiting distribution result given in Theorem 16 for the range j fixed as $n \to \infty$. The martingale description given above shows immediately the almost sure convergence of $\tilde{\mathcal{W}}_{n,\mathbf{s}}$, for j fixed, to limiting random variables, since $\tilde{\mathcal{W}}_{n,\mathbf{s}}$ is a positive martingale. In particular, all the random variables $\tilde{W}_{n,p} := \left(\frac{W_{n,p}}{k}+1\right)(1+\frac{1}{k})/(n+\frac{1}{k})$ converge almost surely to limiting random variables. This also implies that the random vector $(\frac{W_{n,1}}{n}, \ldots, \frac{W_{n,k+1}}{n})$ converges almost surely to a limiting random vector; moreover, using Stirling's formula for the Gamma function (7), one can easily characterize the limiting distribution by its mixed moments. Furthermore, it is not difficult to obtain even a local limit law using the explicit formula for the probability mass function given in Proposition 1.

The remaining cases of the limiting distribution results given in Theorem 16 can be proven easily either by the method of moments or directly by asymptotic considerations for the probability mass function; we skip here these computations. $\square$

7. The number of inversions

The number of inversions in ordinary permutations is a classical parameter of interest. It has been studied by many authors, e.g., by Sachkov [37], Bender [3], Margolius [30], and Louchard and Prodinger [26]. Park [33] has studied the random variable I_n, which counts the number of inversions in a random k-Stirling permutation of order n, from a combinatorial point of view. Here we can add some probabilistic considerations. Park already derived the probability generating function of the total number of inversions I_n using combinatorial methods. Using simple probabilistic arguments we rederive his result, but furthermore we show a central limit theorem for I_n. We first collect the results concerning I_n in the following theorem.

Theorem 17. *The random variable I_n can be decomposed into a sum of mutually independent random variables U_i, where U_i denotes the discrete uniform distribution on the set $\{mk \mid 0 \leq m \leq ik\}$:*

$$I_n = \bigoplus_{i=0}^{n-1} U_i, \quad \text{with } \mathbb{P}\{U_i = km\} = \frac{1}{ik+1} \quad 0 \leq m \leq ik.$$

The expectation and the variance of I_n are given by

$$\mathbb{E}(I_n) = \frac{k^2}{2}\binom{n}{2} \sim \frac{k^2n^2}{4}, \quad \mathbb{V}(I_n) = \frac{k^3\binom{n}{2}}{6}\left(\frac{k(2n-1)}{6}+1\right) \sim \frac{k^4n^3}{36}.$$

The centered and normalized random variable I_n^* converges in distribution to a standard normal distributed random variable, where the speed of convergence is of order $\mathcal{O}\big(1/\sqrt{n}\big)$ with respect to the d_1-metric:

$$I_n^* = \frac{I_n - \mathbb{E}(I_n)}{\sqrt{\mathbb{V}(I_n)}} \xrightarrow{(d)} \mathcal{N}(0,1), \quad d_1\big(\mathcal{L}(I_n^*), \mathcal{N}(0,1)\big) \leq \frac{\kappa(k)}{\sqrt{n}},$$

where $\kappa(k)$ is a constant, which depends only on k, but is independent of n.

Remark 5. It is well known that convergence with respect to the d_1-metric of random variables implies the weak convergence, i.e., the convergence in distribution, see [1, 2]. The speed of convergence is given with respect to the metric d_1, see Barbour, Karoński and Ruciński [2]: for two probability measures P and Q their d_1-distance is defined as $d_1(P, Q) := \sup_{\|h\| \leq 1} |\mathbb{E}(h(X)) - \mathbb{E}(h(Y))|$, where $\|h\| = K \cdot (\|h\|_\infty + \|h'\|_\infty)$, for bounded functions h with bounded derivative and some constant K, with X and Y denoting random variables with distribution P and Q, respectively. We refer the reader to [1] and [2] for a precise definition.

Proof. First we derive the decomposition of the random variable I_n into a sum of independent random variables. To do this we use the description of random k-Stirling permutations via the simple insertion process: at step n the sequence n^k is inserted into a k-Stirling permutation of order $n-1$ at one of the $k(n-1)+1$ places. If n^k is inserted at the m-th place, taken from right to left, this induces exactly km additional inversions. Since the insertion is independent of the structure of the actual k-Stirling permutation, we obtain the stated result.

The expectation and the variance of I_n are obtained using the independence of the random variables U_i; e.g., the expectation is easily derived as follows:

$$\mathbb{E}(U_i) = \sum_{m=0}^{ik} \frac{km}{ik+1} = \frac{ik^2}{2}, \quad \mathbb{E}(S_n) = \sum_{i=0}^{n-1} \mathbb{E}(U_i) = \frac{k^2}{2}\binom{n}{2}.$$

The central limit theorem for I_n can be established via various methods. E.g., one could use directly the method of moments, which also proves the moment convergence, or one can check that Lyapunov's condition in the corresponding theorem for sums of independent random variables is satisfied, or one could use the approach of [26] based on saddle point methods.

However, in order to obtain the stated result concerning the speed of convergence of the centered and normalized r.v. I_n^* to $\mathcal{N}(0,1)$ with respect to the d_1-metric, we used Stein's method [1, 2], which, in the present case, can be applied easily due to the independence of the random variables involved. In the following we will use the notations

$$\hat{U}_i = U_i - \mathbb{E}(U_i), \quad V_i = \frac{\hat{U}_i}{\sqrt{\mathbb{V}(I_n)}}.$$

Since U_i takes values in $\{k\ell \mid 0 \leq \ell \leq ki\}$, the random variable $\hat{U}_i$ is uniformly distributed on the set $\{\frac{ik^2}{2} - k\ell \mid 0 \leq \ell \leq ki\}$. By definition the random variables V_i are mutually independent and satisfy $\mathbb{E}(V_i) = 0$ and $\mathbb{V}(\oplus_{i=0}^{n-1} V_i) = 1$; furthermore it holds $\bigoplus_{i=0}^{n-1} V_i = I_n^*$. To proceed we use the following consequence of [2]:

$$d_1(\mathcal{L}(I_n^*), \mathcal{N}(0,1)) \leq 2 \sum_{i=0}^{n-1} \big(\mathbb{E}(|V_i^3|) + \mathbb{E}(|V_i^2|)\mathbb{E}(|V_i|)\big)$$

$$= \frac{2}{(\mathbb{V}(I_n))^{\frac{3}{2}}} \sum_{i=0}^{n-1} \big(\mathbb{E}(|\hat{U}_i^3|) + \mathbb{E}(|\hat{U}_i^2|)\mathbb{E}(|\hat{U}_i|)\big).$$

It is easy to show that

$$\sum_{i=0}^{n-1} \big(\mathbb{E}(|\hat{U}_i^3|) + \mathbb{E}(|\hat{U}_i^2|)\mathbb{E}(|\hat{U}_i|)\big) = \sum_{i=0}^{n-1} \big(\kappa_1(k)i^3 + \kappa_2(k)i^2\kappa_3(k)i\big) + \mathcal{O}(n^3) = \kappa_4(k)n^4 + \mathcal{O}(n^3),$$

where the $\kappa_j = \kappa_j(k)$ appearing are constants depending only on k, which are not important to our purpose. We get thus

$$d_1\big(\mathcal{L}(I_n^*), \mathcal{N}(0,1)\big) \leq \frac{\kappa(k)}{\sqrt{n}},$$

by using the asymptotic expansion of the variance of I_n. This proves the stated result. $\square$

Acknowledgement

References

[1] A. D. Barbour and G. K. Eagleson, Multiple comparisons and sums of dissociated random variables, *Adv. Appl. Probab.* 17, 147–162, 1985.

[2] A. D. Barbour, M. Karoński and A. Ruciński, A Central Limit Theorem for Decomposable Random Variables with Applications to Random Graphs, *Journal of Combinatorial Theory, Series B* 47, 125–145, 1989.

[3] E. A. Bender, Central and local limit theorems applied to asymptotics enumeration, *Journal of Combinatorial Theory, Series A* 15, 91–111, 1973.

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

[5] B. Bollobas and O. M. Riordan, Mathematical results on scale-free random graphs, in *Handbook of Graphs and Networks*, 1–34, Wiley-VCH, Weinheim, 2003.

[6] M. Bona, *Combinatorics of permutations*, Chapman & Hall, Boca Raton, 2004.

[7] M. Bona, Real zeros and normal distribution for statistics on Stirling permutations defined by Gessel and Stanley, *SIAM J. Discrete Math.* 23, 401–406, 2008.

[8] F. Brenti, Unimodal, log-concave, and Polya frequency sequences in combinatorics, *Memoirs Amer. Math. Soc.* 81, no. 413, 1989.

[9] F. Brenti, Hilbert polynomials in combinatorics, *J. Algebraic Combinatorics* 7, 127–156, 1998.

[10] N. Broutin, L. Devroye, E. McLeish and M. de la Salle, The height of increasing trees, *Random Structures and Algorithms* 32, 494–518, 2008.

[11] D. Y. C. Chan, B. D. Hughes, A. S. Leong and W. J. Reed, Stochastically evolving networks, *Physical Review E* 68, 066124, 2003.

[12] F. Eggenberger and G. Pólya, Über die Statistik verketteter Vorgänge, *Zeitschrift Angew. Math. Mech.* 3, 279–289, 1923.

[13] P. Flajolet, P. Dumas, and V. Puyhaubert, Some exactly solvable models of urn process theory, *Discrete Mathematics and Computer Science*, Proceedings of Fourth Colloquium on Mathematics and Computer Science, Ph. Chassaing Editor., vol. AG, 59–118, 2006.

[14] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, Cambridge, 2009.

[15] I. Gessel and R. P. Stanley, Stirling polynomials, *J. Combin. Theory Ser. A* 24, no. 1, 24–33, 1978.

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

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

[18] S. Janson, Asymptotic degree distribution in random recursive trees, *Random Structures Algorithms* 26, 69–83, 2005.

[19] S. Janson, Limit theorems for triangular urn schemes, *Probability Theory and Related Fields* 134, 417–452, 2006.

[20] S. Janson, Plane recursive trees, Stirling permutations and an urn model, *DMTCS: Proceedings; Fifth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities.* 541-548, 2008.

[21] S. Janson, M. Kuba and A. Panholzer, Generalized Stirling permutations, families of increasing trees and urn models, *Journal of Combinatorial Theory, Series A*, to appear, 2010.

[22] N. L. Johnson and S. Kotz, *Urn models and their application*, Wiley, New York, 1977.

[23] M. Kuba and A. Panholzer, Descendants in increasing trees, *Electronic Journal of Combinatorics* 13, research paper #8, 2006.

[24] M. Kuba and A. Panholzer, Limiting distributions for a class of diminishing urn models, manuscript, 2007.

[25] M. Loève, *Probability Theory I*, 4th Edition, Springer-Verlag, New York, 1977.

[26] G. Louchard and H. Prodinger, Inversions in permutations: A saddle point approach, *Journal of Integers Sequences* 6, article 3.2.8., 19 p., 2003.

[27] H. Mahmoud, Urn models and connections to random trees: a review, *Journal of the Iranian Mathematical Society* 2, 53-114, 2003.

[28] H. Mahmoud and R. Smythe, On the distribution of leaves in rooted subtrees of recursive trees, *Annals of Applied Probability* 1, 406–418, 1991.

[29] H. M. Mahmoud, R. T. Smythe & J. Szymański, On the structure of random plane-oriented recursive trees and their branches, *Random Structures and Algorithms* 4, no. 2, 151–176, 1993.

[30] B. H. Margolius, Permutations with inversions. *Journal of Integer Sequences* 4, 1–13, 2001.

[31] A. Panholzer and H. Prodinger, The level of nodes in increasing trees revisited, *Random Structures and Algorithms*, 31, 203–226, 2007.

[32] S. K. Park, The r-multipermutations, *J. Combin. Theory Ser. A* 67, no. 1, 44–71, 1994.

[33] S. K. Park, Inverse descents of r-multipermutations, *Discrete Math.* 132, no. 1–3, 215–229, 1994.

[34] S. K. Park, P-partitions and q-Stirling numbers, *J. Combin. Theory Ser. A* 68, no. 1, 33–52, 1994.

[35] G. Pólya, Sur quelques points de la théorie des probabilités, *Ann. Inst. Poincaré* 1, 117–161, 1931.

[36] H. Prodinger, Descendants in heap ordered trees, or, A triumph of computer algebra, *Electronic Journal of Combinatorics* 3, paper #29, 1996.

[37] V. N. Sachkov, *Probabilistic Methods in Combinatorial Analysis*, Cambridge University Press, Cambridge, 1997.

[38] J. Vitter and P. Flajolet, Average case analysis of algorithms and data structures, in *Handbook of Theoretical Computer Science*, 431–524, Elsevier, Amsterdam, 1990.

[39] H. Wilf, *Generatingfunctionology*, 2nd ed., Academic Press, Boston, 1994.

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

E-mail address: `kuba@dmg.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`