

# MULTINOMIAL RANDOM COMBINATORIAL STRUCTURES AND $r$ -VERSIONS OF STIRLING, EULERIAN AND LAH NUMBERS

ALEXANDER IKSANOV, ZAKHAR KABLUCHKO, ALEXANDER MARYNYCH, VITALI WACHTEL

ABSTRACT. We introduce multinomial and  $r$ -variants of several classic objects of combinatorial probability, such as the random recursive and Hoppe trees, random set partitions and compositions, the Chinese restaurant process, Feller’s coupling, and some others. Just as various classic combinatorial numbers – like Stirling, Eulerian and Lah numbers – emerge as essential ingredients defining the distributions of the mentioned processes, the so-called  $r$ -versions of these numbers appear in exact distributional formulas for the multinomial and  $r$ -counterparts. This approach allows us to offer a concise probabilistic interpretation for various identities involving  $r$ -versions of these combinatorial numbers, which were either unavailable or meaningful only for specific values of the parameter  $r$ . We analyze the derived distributions for fixed-size structures and establish distributional limit theorems as the size tends to infinity. Utilizing the aforementioned generalized Stirling numbers of both kinds, we define and analyze  $(r, s)$ -Lah distributions, which have arisen in the existing literature on combinatorial probability in various contexts.

## CONTENTS

1. Introduction .....	2
1.1. $r$ -Stirling numbers .....	2
1.2. Three families of probability distributions related to the $r$ -Stirling numbers .....	4
1.3. Notation .....	5
2. Multinomial structures .....	6
2.1. Multinomial Stirling numbers .....	6
2.2. Multinomial Ewens permutation .....	7
2.3. Multinomial Chinese restaurant and the colored Feller coupling .....	9
2.4. Multinomial Hoppe forests .....	10
3. $r$ -Ewens permutations, random $r$ -recursive trees .....	15
3.1. $r$ -Ewens permutations .....	15
3.2. Random $r$ -recursive trees and $r$ -Hoppe trees .....	18
4. Random $r$ -partitions .....	20
4.1. Incomplete partitions derived from urn schemes .....	20
4.2. Random uniform $r$ -partitions and random Gibbs $r$ -partitions .....	22
5. Random $r$ -compositions .....	24
5.1. Couplings of random $r$ -compositions and the Dirichlet multinomial distribution ....	26
5.2. Marginal distributions of $r$ -compositions and their limits .....	29
6. The $(r, s)$ -Lah distribution .....	31
6.1. Motivation .....	31

---

2020 *Mathematics Subject Classification*. Primary: 60C05, 11B73; Secondary: 60E05, 60F05.

*Key words and phrases*.  $r$ -Stirling numbers,  $r$ -Lah numbers, central limit theorem, Chinese restaurant process, Hoppe trees, random compositions, random permutations, random recursive trees, random set partitions.

6.2.	The $r$ -Lah numbers .....	31
6.3.	Definition of the $(r, s)$ -Lah distribution .....	32
6.4.	Generating polynomial of the $(r, s)$ -Lah distribution and a recursion for the probability mass function .....	33
6.5.	Stochastic representation in terms of random recursive trees .....	34
6.6.	Stochastic representation in terms of random compositions .....	36
6.7.	Expectation and variance .....	38
7.	Limit theorems for the $(r, s)$ -Lah distribution .....	42
	Acknowledgement .....	50
	References .....	50

## 1. INTRODUCTION

**1.1.  $r$ -Stirling numbers.** The usual Stirling numbers of both kinds play a pervasive role in combinatorics and discrete probability, often appearing prominently in diverse counting problems as an indispensable element. The list of quintessential examples begins with the enumeration of cycles in permutations, records in random samples, and blocks in set partitions. It extends by enumerating alignments and surjections, progressing to the counting of recursive trees with various profile restrictions. This list could have potentially been indefinitely supplemented. We refer to the excellent treatises [10, 22, 46, 47, 50, 64] for comprehensive information on the classic Stirling numbers. A thorough historic account and an extensive bibliography can be found in the books [46] and [47].

As detailed in the aforementioned references, there exist numerous generalizations of the Stirling numbers. In the focus of the present paper are the so-called  $r$ -Stirling numbers which depend on an additional parameter  $r \in \mathbb{R}$ . These numbers have been intensively investigated in the combinatorics literature. Given that the usual Stirling numbers play a central role in combinatorial probability, it is natural to ask whether there exist natural  $r$ -deformations of the classic discrete random structures (such as random permutations and partitions, or random recursive trees) whose distributional properties are described by the  $r$ -Stirling numbers. One of the purposes of the present paper is to construct these  $r$ -deformed structures and explore their properties.

Perhaps, the easiest way to define the  $r$ -Stirling numbers  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  and  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  is by using the finite differences calculus. Let  $\mathbb{C}[r]$  be the vector space of polynomials in the variable  $r$  with coefficients from  $\mathbb{C}$ . Let  $D$  be the differentiation operator and  $\Delta$  the finite difference operator acting on  $\mathbb{C}[r]$  as follows:

$$Df(r) = f'(r), \quad \Delta f(r) = f(r+1) - f(r).$$

Then, using the notation  $r^{n\uparrow} = r(r+1)\cdots(r+n-1)$  for the rising factorial,

$$\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r := \frac{D^k}{k!} r^{n\uparrow}, \quad \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r := \frac{\Delta^k}{k!} r^n, \quad n \in \mathbb{N}, \quad k \in \{0, 1, \dots, n\}. \quad (1.1)$$

For example,  $\left[ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right]_r = r^{n\uparrow}$  and  $\left\{ \begin{smallmatrix} n \\ 0 \end{smallmatrix} \right\}_r = r^n$ . The usual Stirling numbers  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]$  and  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$  are recovered by evaluating the expressions at  $r = 0$ , that is,

$$\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right] = \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_0 = \left( \frac{D^k}{k!} r^{n\uparrow} \right) \Big|_{r=0}, \quad \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_0 = \left( \frac{\Delta^k}{k!} r^n \right) \Big|_{r=0}.$$

In general, the parameter  $r$  may take arbitrary real (or complex) values. If we fix  $n$  and  $k$ , then  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  and  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  are polynomials in  $r$ , as directly follows from (1.1). The above definitions can be naturally extended to arbitrary  $n \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$  and  $k \in \mathbb{Z}$  by putting

$$\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r := \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r := 0 \text{ for } n \in \mathbb{N}_0 \text{ and } k \notin \{0, \dots, n\}, \quad \left[ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right]_r := \left\{ \begin{smallmatrix} 0 \\ 0 \end{smallmatrix} \right\}_r := 1.$$

The first systematic investigation of the  $r$ -Stirling numbers (also called the ‘weighted Stirling numbers’ and ‘non-central Stirling numbers’) is due to Carlitz [6, 7]. Some follow-up papers are [5, 12, 42, 58]. A clear exposition of properties of the  $r$ -Stirling numbers can be found in the paper of Broder [5] (where the term ‘ $r$ -Stirling numbers’ appears for the first time) and in the books of Mező [50, Chapter 8] and Charalambides [10, § 8.5], [11, Chapter 3]. The paper [54], whose notation we follow, contains a nice overview of the identities satisfied by the  $r$ -Stirling numbers. We note in passing that our notation corresponds to Carlitz’s notation  $R(n, k, r) = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  and  $R_1(n, k, r) = \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  and what we denote by  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  and  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  corresponds to  $\left[ \begin{smallmatrix} n+r \\ k+r \end{smallmatrix} \right]_r$  and  $\left\{ \begin{smallmatrix} n+r \\ k+r \end{smallmatrix} \right\}_r$  in [5, 50]. In particular, the second definition in (1.1) can be read from Eq. (3.8) in [6].

While Carlitz’s approach is purely analytic and hinges on manipulations with generating functions, Broder’s starting point is a combinatorial definition, see Eq. (1) and (2) in [5]. Adapting to our notation and assuming that  $r$  is a non-negative integer, this is

**Definition 1.1.** For  $r \in \mathbb{N}_0$ ,  $n \in \mathbb{N}$  and  $k \in \{0, 1, \dots, n\}$ ,  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  is the number of permutations of the set  $\{1, 2, \dots, n+r\}$  having  $k+r$  cycles such that the numbers  $1, 2, \dots, r$  are in different cycles;  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  is the number of partitions of the set  $\{1, 2, \dots, n+r\}$  into  $k+r$  disjoint subsets such that the numbers  $1, 2, \dots, r$  are in distinct blocks.

While being combinatorially appealing Definition 1.1 has a drawback of being only valid for integer  $r$ . In contrast, Carlitz’s method is free of this shortcoming. Unless stated otherwise, in all subsequent considerations we stick to Carlitz’s viewpoint and do not assume that  $r$  is integer.

The properties of the  $r$ -Stirling numbers are parallel to those of the usual Stirling numbers. For example, see Eq. (3.4) and (5.8) in [6] or Eq. (48) and (50) in [5],

$$(x+r)^{n\uparrow} = \sum_{k=0}^n \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r x^k, \quad (x+r)^n = \sum_{k=0}^n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r x^{k\downarrow}, \quad (1.2)$$

where  $x^{k\downarrow} = x(x-1)\cdots(x-k+1)$  denotes the falling factorial. The (vertical) exponential generating functions of the  $r$ -Stirling numbers are given by

$$\sum_{n=k}^{\infty} \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r \frac{x^n}{n!} = \frac{1}{k!} (\log((1-x)^{-1}))^k (1-x)^{-r}, \quad (1.3)$$

$$\sum_{n=k}^{\infty} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r \frac{x^n}{n!} = \frac{1}{k!} (e^x - 1)^k e^{rx}, \quad (1.4)$$

see Eq. (3.9) in [6] and Eq. (36) in [5]. It is known (see, e.g. p. 1661 in [54]) that

$$\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r = \sum_{j=k}^n \left[ \begin{smallmatrix} n \\ j \end{smallmatrix} \right]_r \binom{j}{k} r^{j-k} = \sum_{j=k}^n \binom{n}{j} \left[ \begin{smallmatrix} j \\ k \end{smallmatrix} \right]_r^{(n-j)\uparrow}, \quad (1.5)$$

$$\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r = \sum_{j=k}^n \binom{n}{j} \left\{ \begin{smallmatrix} j \\ k \end{smallmatrix} \right\}_r r^{n-j} = \sum_{j=k}^n \left\{ \begin{smallmatrix} n \\ j \end{smallmatrix} \right\}_r \binom{j}{k} r^{(j-k)\downarrow}. \quad (1.6)$$

These formulas demonstrate once again that, for fixed  $n \in \mathbb{N}_0$ ,  $k \in \{0, \dots, n\}$ , the  $r$ -Stirling numbers  $\left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r$  and  $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_r$  are polynomials in  $r$ .

**1.2. Three families of probability distributions related to the  $r$ -Stirling numbers.** There are various probability distributions related to usual Stirling numbers, see [11, 13, 53, 56, 57]. The most prominent example is the so-called *Stirling distribution of the first kind* (or a *Stirling-Karamata* distribution)  $\text{Stir1}[n; \theta]$  with parameters  $n \in \mathbb{N}$  and  $\theta > 0$ , which is a discrete probability measure on  $\{0, \dots, n\}$  given by

$$\text{Stir1}[n; \theta](\{k\}) = \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right] \frac{\theta^k}{\theta^{n\uparrow}}, \quad k \in \{0, \dots, n\}. \quad (1.7)$$

If  $\theta = 1$ ,  $\text{Stir1}[n; 1]$  is the distribution of the number of cycles in a random permutation picked uniformly at random from the symmetric group of  $n$  elements. More generally,  $\text{Stir1}[n; \theta]$  is the distribution of the number of cycles in a random Ewens permutation with parameter  $\theta$ . This distribution also pops up as the number of occupied tables in the Chinese restaurant process, see Eq. (3.11) in [57], or the root degree in the random Hoppe tree [26]. For a review of the related concept of the Ewens sampling formula and similar topics we refer to [15]. We will discuss these connections later on in a larger generality.

Related to the Stirling numbers of the second kind is the *Stirling-Sibuya* distribution defined by

$$\text{StirSibuya}[n; \theta](\{k\}) = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \frac{\theta^{k\downarrow}}{\theta^n}, \quad k \in \{0, \dots, n\}, \quad n \in \mathbb{N}. \quad (1.8)$$

Here,  $\theta$  is either a positive integer or a real number larger than  $n - 1$ , so that the falling factorial is nonnegative. If  $\theta$  is a positive integer,  $\text{StirSibuya}[n; \theta]$  is the distribution of the number of occupied boxes after allocating  $n$  balls over  $\theta$  urns equiprobably. The fact that both (1.7) and (1.8) do define probability distributions follows from equalities (1.2) with  $r = 0$ . Yet another instance is the *Stirling distribution of the second kind*  $\text{Stir2}[n; \theta]$ . To define it, let  $T_n$  be the Touchard polynomial given by

$$T_n(z) = \sum_{k=0}^n \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} z^k, \quad z \in \mathbb{C}, \quad n \in \mathbb{N}.$$

Then, for a parameter  $\theta > 0$ ,

$$\text{Stir2}[n; \theta](\{k\}) = \frac{1}{T_n(\theta)} \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} \theta^k, \quad k \in \{0, \dots, n\}, \quad n \in \mathbb{N}.$$

This is the distribution of the number of blocks in a Gibbs (uniform, for  $\theta = 1$ ) random partition of  $\{1, \dots, n\}$ . We refer to [11, 13, 37] and references therein for further properties and applications of the aforementioned distributions.

From a purely analytic viewpoint one can introduce the  $r$ -versions of the above distributions by simply using (1.2) and the  $r$ -Touchard polynomials.

**Definition 1.2** ([53, 60, 61]). Let  $n \in \mathbb{N}$ ,  $\theta > 0$  and  $r \geq 0$ . The  $r$ -*Stirling distribution of the first kind*  $r\text{-Stir1}[n; \theta]$  is a discrete probability measure on  $\{0, \dots, n\}$  given by

$$r\text{-Stir1}[n; \theta](\{k\}) = \left[ \begin{smallmatrix} n \\ k \end{smallmatrix} \right]_r \frac{\theta^k}{(\theta + r)^{n\uparrow}}, \quad k \in \{0, \dots, n\}. \quad (1.9)$$

The distribution defined by (1.9) appears in [53], see also [11, Example 2.1] and a survey [60]. Also, in [61] it pops up as the distribution of a random variable  $X_{n+1} - 1$ , where  $X_{n+1}$  is the number of records in a sample of size  $n + 1$  of independent non-identically distributed random variables (Nevzorov's model). Our notation  $r\text{-Stir1}[n; \theta]$  corresponds to  $\text{Str1F}(n, \theta, r)$  in [53] and [61].

For  $r \geq 0$  and  $n \in \mathbb{N}$ , let the  $r$ -Touchard polynomial, also called the  $r$ -Bell polynomial in [50, § 8.3] and [48], be defined by

$$T_{n,r}(z) := \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r z^k, \quad z \in \mathbb{C}.$$

**Definition 1.3.** Let  $n \in \mathbb{N}$ ,  $\theta > 0$  and  $r \geq 0$ . The  $r$ -Stirling distribution of the second kind  $r\text{-Stir2}[n; \theta]$  is a discrete probability measure on  $\{0, \dots, n\}$  given by

$$r\text{-Stir2}[n; \theta](\{k\}) = \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \frac{\theta^k}{T_{n,r}(\theta)}, \quad k \in \{0, \dots, n\}. \quad (1.10)$$

**Definition 1.4** ([53, 60]). Assume that  $n \in \mathbb{N}$ ,  $r \geq 0$  and either  $\theta \in \mathbb{N}$  or  $\theta > n - 1$ . The  $r$ -Stirling–Sibuya distribution  $r\text{-StirSibuya}[n; \theta]$  is a discrete probability measure on  $\{0, \dots, n\}$  given by

$$r\text{-StirSibuya}[n; \theta](\{k\}) = \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \frac{\theta^{k\downarrow}}{(\theta + r)^n}, \quad k \in \{0, \dots, n\}. \quad (1.11)$$

The  $r$ -Stirling–Sibuya distribution  $r\text{-StirSibuya}[n; \theta]$  appears in [53], see Section 5 therein, as the distribution of a position at time  $n$  of a certain time-inhomogeneous random walk on  $\mathbb{N}_0$ ; see also [11, Example 2.6]. An urn model leading to the  $r\text{-StirSibuya}[n; \theta]$  distribution is also discussed in [53] and will be important in our considerations of random  $r$ -partitions in Section 4 below. Our notation  $r\text{-StirSibuya}[n; \theta]$  corresponds to  $\text{Str2F}(n, \theta, r)$  in [53].

Distributions related to further generalizations of Stirling numbers are discussed in [11], [57, Example 3.2.3] and [28].

In this work, we introduce a general concept of multinomial combinatorial structures which we find interesting in their own. Multinomial structures are obtained by taking the classic random structures (such as random permutations, partitions or compositions) and coloring their cycles or blocks into  $d \in \mathbb{N}$  colors. It will become clear that the  $r$ -distributions defined above (and some other) occur in the special case  $d = 2$ .

The rest of the paper is organized as follows. In Section 2 we introduce and investigate the multinomial versions of some classic combinatorial stochastic processes. In Sections 3 we show how the  $r$ -distributions arise as the distributions of some functionals defined on the multinomial structures of Section 2. Section 4 is devoted to the analysis of random incomplete partitions and an accompanying notion of random  $r$ -partitions. A similar analysis of incomplete compositions and random  $r$ -compositions is carried out in Section 5. In Section 6 we discuss yet another class of combinatorial numbers, called  $r$ -Lah numbers, and introduce generalized Lah distributions. Specific occurrences of these distributions have already received some attention in the literature on combinatorial probability across diverse contexts. Various limit theorems for these distributions are discussed in the concluding Section 7.

**1.3. Notation.** Besides the already introduced notation, we also use the following abbreviations and conventions. We put  $[n] := \{1, 2, \dots, n\}$ . For the asymptotic equivalence we use the symbol  $\simeq$ . Thus,  $a_n \simeq b_n$  if, and only if,  $a_n/b_n \rightarrow 1$  as  $n \rightarrow \infty$ . We use the standard notation  $x_+ = \max(x, 0)$

and  $x_- = \max(-x, 0)$  for  $x \in \mathbb{R}$ . The cardinality of a finite set  $A$  is denoted by  $\#A$ . For a formal power series  $f(x) = \sum_{n=0}^{\infty} f_n x^n$ , we stipulate that  $[x^n]f(x) = f_n$ . If  $X$  is a random element taking values in some space  $S$  and  $\mu$  is a probability distribution on  $S$ , we write  $X \sim \mu$  to denote that  $X$  has distribution  $\mu$ . If  $(\mu[\theta])_{\theta \in P}$  is a family of probability measures depending on a parameter  $\theta$  and  $\mathfrak{M}$  is a probability measure on  $P$ , we denote by  $\mu[\mathfrak{M}]$  the mixture of  $(\mu[\theta])_{\theta \in P}$ . Thus,

$$(\mu[\mathfrak{M}])(\cdot) := \int_P \mu[\theta](\cdot) \mathfrak{M}(d\theta). \quad (1.12)$$

In addition to the already introduced probability measures related to the Stirling numbers, many standard distributions show up throughout the text. We denote by  $\text{Bin}[n; p]$  the binomial distribution with parameters  $n$  and  $p$  and by  $\text{Bern}[p]$  the Bernoulli distribution with parameter  $p$ , that is,  $\text{Bern}[p](\{1\}) = p = 1 - \text{Bern}[p](\{0\})$ . Furthermore,  $\text{Poi}[\lambda]$  denotes the Poisson distribution with parameter  $\lambda$  and  $\text{Mult}[n; (p_1, \dots, p_d)]$  denotes the multinomial distribution with parameters  $n$  and  $p_1, \dots, p_d$ . Notation for other distributions will be introduced in the text just before the first appearance. Keeping in mind Eq. (1.12) we ubiquitously use formulas like  $\text{Bin}[\text{Poi}[\lambda]; p] = \text{Poi}[\lambda p]$  involving the mixtures of distributions. With an abuse of notation,  $\text{Bern}[p]$  can denote either the distribution or a random variable with this distribution (implicitly assumed to be defined on some probability space). The same disclaimer also applies to the other distributions.

## 2. MULTINOMIAL STRUCTURES

In this section we introduce  $r$ -analogues of some classic random structures. Examples include random Ewens permutations, random recursive (and Hoppe) trees and the Chinese restaurant process. The  $r$ -versions will be defined as special cases of more general, multinomial versions of these structures.

**2.1. Multinomial Stirling numbers.** The *multinomial Stirling cycle number*  $\left[ \begin{smallmatrix} n \\ k_1, \dots, k_d \end{smallmatrix} \right]$  is defined as the number of permutations of  $[n]$  with  $k_1 + \dots + k_d$  cycles which are colored in  $d$  possible colors  $1, \dots, d$  such that the number of cycles with color  $j$  is  $k_j$ , for all  $j = 1, \dots, d$ . Clearly,

$$\left[ \begin{smallmatrix} n \\ k_1, \dots, k_d \end{smallmatrix} \right] = \left[ \begin{smallmatrix} n \\ k_1 + \dots + k_d \end{smallmatrix} \right] \binom{k_1 + \dots + k_d}{k_1, \dots, k_d}.$$

Similarly, the *multinomial Stirling partition number*  $\left\{ \begin{smallmatrix} n \\ k_1, \dots, k_d \end{smallmatrix} \right\}$  is defined as the number of partitions of  $[n]$  into  $k_1 + \dots + k_d$  nonempty blocks which are colored in  $d$  possible colors such that  $k_j$  is the number of blocks with color  $j$ , for all  $j = 1, \dots, d$ . Plainly,

$$\left\{ \begin{smallmatrix} n \\ k_1, \dots, k_d \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n \\ k_1 + \dots + k_d \end{smallmatrix} \right\} \binom{k_1 + \dots + k_d}{k_1, \dots, k_d}.$$

In both definitions, the parameters  $n \in \mathbb{N}_0$  and  $k_1, \dots, k_d \in \mathbb{N}_0$  satisfy  $k_1 + \dots + k_d \leq n$  with the convention that  $\left[ \begin{smallmatrix} 0 \\ 0, \dots, 0 \end{smallmatrix} \right] = 1$ . Additionally, if  $n \in \mathbb{N}_0$  and  $k_1, \dots, k_d \in \mathbb{Z}$  fail to satisfy the above restrictions (meaning that one of the  $k_j$ 's is negative or their sum is larger than  $n$ ), we define the corresponding multinomial Stirling numbers to be 0. The properties of the multinomial Stirling numbers are reviewed in [4, Section 4]. Here we only need the following result.

**Proposition 2.1.** *With  $x_1, \dots, x_d$  denoting variables,*

$$(x_1 + \dots + x_d)^{n\uparrow} = \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \left[ \begin{matrix} n \\ k_1, \dots, k_d \end{matrix} \right] x_1^{k_1} \dots x_d^{k_d}, \quad (2.1)$$

$$(x_1 + \dots + x_d)^n = \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \left\{ \begin{matrix} n \\ k_1, \dots, k_d \end{matrix} \right\} x_1^{k_1\downarrow} \dots x_d^{k_d\downarrow}. \quad (2.2)$$

*Proof.* To prove (2.1), we use the formula  $x^{n\uparrow} = \sum_{k=0}^n \binom{n}{k} x^k$  with  $x = x_1 + \dots + x_d$  and then the multinomial theorem:

$$\begin{aligned} (x_1 + \dots + x_d)^{n\uparrow} &= \sum_{k=0}^n \binom{n}{k} (x_1 + \dots + x_d)^k = \sum_{k=0}^n \binom{n}{k} \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d = k}} \binom{k}{k_1, \dots, k_d} x_1^{k_1} \dots x_d^{k_d} \\ &= \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \binom{n}{k} \binom{k}{k_1, \dots, k_d} x_1^{k_1} \dots x_d^{k_d} = \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \left[ \begin{matrix} n \\ k_1, \dots, k_d \end{matrix} \right] x_1^{k_1} \dots x_d^{k_d}. \end{aligned}$$

Similarly, to prove (2.2), we use the formula  $x^n = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^{k\downarrow}$  together with the analogue of the multinomial theorem for falling/rising factorials:

$$\begin{aligned} (x_1 + \dots + x_d)^n &= \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} (x_1 + \dots + x_d)^{k\downarrow} = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d = k}} \binom{k}{k_1, \dots, k_d} x_1^{k_1\downarrow} \dots x_d^{k_d\downarrow} \\ &= \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \left\{ \begin{matrix} n \\ k \end{matrix} \right\} \binom{k}{k_1, \dots, k_d} x_1^{k_1\downarrow} \dots x_d^{k_d\downarrow} = \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \left\{ \begin{matrix} n \\ k_1, \dots, k_d \end{matrix} \right\} x_1^{k_1\downarrow} \dots x_d^{k_d\downarrow}. \end{aligned}$$

The proof is complete.  $\square$

**2.2. Multinomial Ewens permutation.** Let  $\mathfrak{S}_n$  be the set of all permutations of the set  $[n]$ . Let  $\text{cycles}(\rho)$  denote the set of cycles of a permutation  $\rho \in \mathfrak{S}_n$  and  $\#\text{cycles}(\rho)$  denote the number of cycles.

A random *Ewens permutation* with parameter  $\theta > 0$  is a random element  $\rho$  of  $\mathfrak{S}_n$  such that for each deterministic permutation  $\rho \in \mathfrak{S}_n$

$$\mathbb{P}[\rho = \rho] = \frac{\theta^{\#\text{cycles}(\rho)}}{\theta^{n\uparrow}},$$

see Example 2.19 and Chapter 5 in [2], and also [57, Section 3.2]. We denote this distribution by  $\text{Ewens}[n; \theta]$ , so that  $\rho \sim \text{Ewens}[n; \theta]$ . Note that the special case  $\theta = 1$  corresponds to a random uniform permutation. It is important for what follows that the Ewens distribution can also be defined for  $\theta = 0$ . In this case it coincides with the uniform distribution on the subset of  $\mathfrak{S}_n$  consisting of permutations with exactly one cycle. The total number of such permutations is  $(n-1)!$ .

The number of cycles of  $\rho$  follows the Stirling distribution of the first kind with parameter  $\theta$ :

$$\mathbb{P}[\#\text{cycles}(\rho) = k] = \left[ \begin{matrix} n \\ k \end{matrix} \right] \frac{\theta^k}{\theta^{n\uparrow}}, \quad k \in \{0, \dots, n\}$$

(in short,  $\#\text{cycles}(\rho) \sim \text{Stir1}[n; \theta]$ ). Indeed,  $\begin{bmatrix} n \\ k \end{bmatrix}$  is the number of permutations of  $[n]$  with  $k$  cycles, and each such a permutation has probability  $\theta^k/\theta^{n\uparrow}$  under the Ewens distribution  $\text{Ewens}[n; \theta]$ .

Now we introduce the multinomial analogue of the Ewens distribution. A *colored permutation* is a permutation of  $[n]$  in which each cycle has one of  $d$  possible colors. The number of colors  $d \in \mathbb{N}$  is fixed once and for all. Formally, a colored permutation is a pair  $(\rho, \phi)$ , where  $\rho \in \mathfrak{S}_n$  is a permutation and  $\phi : \text{cycles}(\rho) \rightarrow \{1, \dots, d\}$  is a map assigning colors to cycles of  $\rho$ . A *multinomial random Ewens permutation* is obtained by taking a Ewens permutation and coloring the cycles independently of each other, with  $p_j$  denoting the probability of assigning color  $j \in \{1, \dots, d\}$  to a cycle. More precisely, we fix parameters  $\theta > 0$  and  $p_1, \dots, p_d \geq 0$  such that  $p_1 + \dots + p_d = 1$ , define  $\theta_j := \theta p_j$  and adopt the following

**Definition 2.2.** A random colored permutation  $(\rho, \phi)$  has a *multinomial Ewens distribution*  $\text{MultiEwens}[n; \theta; (p_1, \dots, p_d)]$  if

$$\mathbb{P}[(\rho, \phi) = (\rho, \phi)] = \frac{\theta^{\#\text{cycles}(\rho)}}{\theta^{n\uparrow}} p_1^{\#\text{cycles}_1(\rho, \phi)} \dots p_d^{\#\text{cycles}_d(\rho, \phi)} = \frac{\theta_1^{\#\text{cycles}_1(\rho, \phi)} \dots \theta_d^{\#\text{cycles}_d(\rho, \phi)}}{\theta^{n\uparrow}} \quad (2.3)$$

for each deterministic colored permutation  $(\rho, \phi)$ . Here,  $\text{cycles}_j(\rho, \phi)$  denotes the set of cycles of color  $j$  in the colored permutation  $(\rho, \phi)$ .

From the definition of multinomial Stirling numbers it follows that the cycle counts of such  $(\rho, \phi)$  are distributed as follows:

$$\mathbb{P}[\#\text{cycles}_1(\rho, \phi) = k_1, \dots, \#\text{cycles}_d(\rho, \phi) = k_d] = \begin{bmatrix} n \\ k_1, \dots, k_d \end{bmatrix} \frac{\theta^{k_1 + \dots + k_d}}{\theta^{n\uparrow}} p_1^{k_1} \dots p_d^{k_d}$$

for all  $k_1, \dots, k_d \in \mathbb{N}_0$  subject to  $k_1 + \dots + k_d \leq n$ . This motivates the following

**Definition 2.3.** A random vector  $(K_1, \dots, K_d)$  with values in  $\mathbb{N}_0^d$  has a *multinomial Stirling distribution* of the first kind  $\text{MultiStir1}[n; \theta; (p_1, \dots, p_d)]$  if

$$\mathbb{P}[K_1 = k_1, \dots, K_d = k_d] = \begin{bmatrix} n \\ k_1, \dots, k_d \end{bmatrix} \frac{\theta^{k_1 + \dots + k_d}}{\theta^{n\uparrow}} p_1^{k_1} \dots p_d^{k_d} = \begin{bmatrix} n \\ k_1, \dots, k_d \end{bmatrix} \frac{\theta_1^{k_1} \dots \theta_d^{k_d}}{\theta^{n\uparrow}}$$

for all  $k_1, \dots, k_d \in \mathbb{N}_0$  satisfying  $k_1 + \dots + k_d \leq n$ , where  $\theta_j = \theta p_j$  for  $j \in \{1, \dots, d\}$ .

Thus, the cycle counts of  $(\rho, \phi) \sim \text{MultiEwens}[n; \theta; (p_1, \dots, p_d)]$  satisfy

$$(\#\text{cycles}_1(\rho, \phi), \dots, \#\text{cycles}_d(\rho, \phi)) \sim \text{MultiStir1}[n; \theta; (p_1, \dots, p_d)].$$

**Proposition 2.4.** The multivariate generating function of the random vector  $(K_1, \dots, K_d)$  that follows  $\text{MultiStir1}[n; \theta; (p_1, \dots, p_d)]$  distribution is given by

$$\mathbb{E}[t_1^{K_1} \dots t_d^{K_d}] = \frac{(\theta_1 t_1 + \dots + \theta_d t_d)^{n\uparrow}}{(\theta_1 + \dots + \theta_d)^{n\uparrow}}, \quad (2.4)$$

where  $\theta_j = \theta p_j$  for  $i \in \{1, \dots, d\}$ . Moreover, if  $I_1, \dots, I_n$  are independent random vectors in  $\mathbb{N}_0^d$  with

$$\mathbb{P}[I_\ell = \mathbf{e}_j] = \frac{\theta_j}{\theta + \ell - 1}, \quad j \in \{1, \dots, d\}, \quad \mathbb{P}[I_\ell = \mathbf{0}] = \frac{\ell - 1}{\theta + \ell - 1}, \quad (2.5)$$

where  $\mathbf{e}_1, \dots, \mathbf{e}_d$  is the standard basis of  $\mathbb{R}^d$  and  $\mathbf{0}$  is the zero vector in  $\mathbb{R}^d$ , then  $(K_1, \dots, K_d)$  has the same distribution as  $I_1 + \dots + I_n$ . In particular,  $K_j \sim (\theta - \theta_j)\text{-Stir1}[n; \theta_j]$ .



*Proof.* By definition of the multinomial Stirling distribution,

$$\mathbb{E}[t_1^{K_1} \cdots t_d^{K_d}] = \frac{1}{\theta^{n\uparrow}} \sum_{\substack{k_1, \dots, k_d \in \mathbb{N}_0 \\ k_1 + \dots + k_d \leq n}} \begin{bmatrix} n \\ k_1, \dots, k_d \end{bmatrix} (\theta_1 t_1)^{k_1} \cdots (\theta_d t_d)^{k_d} = \frac{(\theta_1 t_1 + \cdots + \theta_d t_d)^{n\uparrow}}{\theta^{n\uparrow}},$$

where the second equality follows from (2.1). It remains to note that the generating function of the random vector  $I_1 + \cdots + I_n$  is the same. To identify the distribution of  $K_j$ , observe that (2.4) implies

$$\mathbb{E}[t_j^{K_j}] = \frac{(\theta_j t_j + \theta - \theta_j)^{n\uparrow}}{\theta^{n\uparrow}} \stackrel{(1.2)}{=} \sum_{k=0}^n \begin{bmatrix} n \\ k \end{bmatrix}_{\theta - \theta_j} \frac{\theta_j^k t_j^k}{(\theta_j + \theta - \theta_j)^{n\uparrow}},$$

which proves that  $K_j \sim (\theta - \theta_j)\text{-Stir1}[n; \theta_j]$ , see Definition 1.2.  $\square$

The following aggregation property follows directly from the definition:

**Proposition 2.5.** *If  $(K_1, \dots, K_d) \sim \text{MultiStir1}[n; \theta; (p_1, p_2, \dots, p_d)]$ , then*

$$(K_1 + K_2, K_3, \dots, K_d) \sim \text{MultiStir1}[n; \theta; (p_1 + p_2, p_3, \dots, p_d)].$$

Recalling our convention (1.12) concerning the mixtures of distributions we write

$$\text{MultiStir1}[n; \theta; (p_1, \dots, p_d)] = \text{Mult}[\text{Stir1}[n; \theta]; (p_1, \dots, p_d)]. \quad (2.6)$$

This provides another way to prove Proposition 2.4.

**2.3. Multinomial Chinese restaurant and the colored Feller coupling.** Fix parameters  $\theta_1, \dots, \theta_d \geq 0$  and suppose that  $\theta := \theta_1 + \cdots + \theta_d > 0$ . Customers labeled by  $1, 2, \dots, n$  arrive in discrete time and take seats at round tables. Each table is colored in one of  $d$  colors. Each customer  $\ell$ , after arrival, either creates a new table of color  $j \in \{1, \dots, d\}$  with probability  $\theta_j / (\theta + \ell - 1)$  or takes a seat at some already existing table next to customer  $i$  (in the counterclockwise direction) with probability  $1 / (\theta + \ell - 1)$ , for each customer  $i = 1, \dots, \ell - 1$ . After arrival of  $n$  customers, this process generates a random permutation of  $[n]$ , if we regard tables (read counterclockwise) as cycles.

**Proposition 2.6.** *The colored permutation generated by the multinomial Chinese restaurant process has the distribution*

$$\text{MultiEwens}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)].$$

*In particular, if  $C_{n;j}$  denotes the number of tables (cycles) of color  $j \in \{1, \dots, d\}$ , then*

$$(C_{n;1}, \dots, C_{n;d}) \sim \text{MultiStir1}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)].$$

*First proof.* Consider the multinomial Chinese restaurant process and remove the colors of the tables. This is the usual Chinese restaurant process with parameter  $\theta$ . It is known, see Example 2.19 in [2], that the process generates a random permutation which is distributed according to Ewens $[n; \theta]$ . Now, recall that in the colored version of the process each new table receives color  $j$  with probability  $\theta_j / \theta$ . By definition, this gives the multinomial Ewens distribution.  $\square$

*Second proof.* Another approach hinges on a colored version of the classic Feller coupling for Ewens' permutations, see pp. 95-96 in [2]. Let  $(D_i)_{i \in \mathbb{N}}$ , be independent random variables such that  $D_i$  takes values in the set

$$\{\text{close}_1, \dots, \text{close}_d\} \cup \{2, 3, \dots, i\}$$

(the meaning of ‘close<sub>*j*</sub>’ will be explained a bit later) and has the distribution

$$\mathbb{P}[D_i = \text{close}_j] = \frac{\theta_j}{\theta + i - 1}, \quad j = 1, \dots, d, \quad \mathbb{P}[D_i = k] = \frac{1}{\theta + i - 1}, \quad k = 2, \dots, i, \quad i \in \mathbb{N}.$$

These variables are then used to generate a colored random permutation in the ordered cycle notation as follows. Start the first cycle with ‘(1’ and make a  $(n + d - 1)$ -choice using the variable  $D_n$ . If  $D_n = k$ , then  $k$  is added to the first cycle transforming ‘(1’ to ‘(1  $k$ ’. If  $D_n = \text{close}_j$ , then one closes the first cycle, paints it using the color  $j$  and opens the next cycle. Continuing in this way and using  $D_{n-1}, \dots, D_1$  produces a colored random permutation, say  $(\boldsymbol{\rho}, \boldsymbol{\phi})$ , in the ordered cycle notation. Thus, the value  $D_i = \text{close}_j$  corresponds to closure of the current cycle, painting it using the color  $j$  and starting a new cycle with the smallest unused integer. The total number of cycles in  $(\boldsymbol{\rho}, \boldsymbol{\phi})$  is equal to

$$C_n = \sum_{i=1}^n \mathbf{1}[D_i \in \{\text{close}_1, \dots, \text{close}_d\}]$$

and, since  $\mathbb{P}[D_i \in \{\text{close}_1, \dots, \text{close}_d\}] = \theta/(\theta + i - 1)$ , follows the Stirling distribution of the first kind  $\text{Stir1}[n; \theta]$ . Furthermore, the probability of seeing any fixed (uncolored) permutation  $\rho \in \mathfrak{S}_n$ , is equal to

$$\mathbb{P}[\boldsymbol{\rho} = \rho] = \frac{\theta^{\#\text{cycles}(\rho)}}{\theta^{n\uparrow}}.$$

Thus, erasing the colors in  $(\boldsymbol{\rho}, \boldsymbol{\phi})$  produces Ewens’ permutation. The number of cycles of color  $j$  is

$$C_{n;j} = \sum_{i=1}^n \mathbf{1}[D_i = \text{close}_j], \quad j = 1, \dots, d.$$

Using that, independently of  $i = 1, \dots, n$ ,

$$\mathbb{P}[D_i = \text{close}_j | D_i \in \{\text{close}_1, \dots, \text{close}_d\}] = \frac{\theta_j}{\theta}, \quad j = 1, \dots, d,$$

we conclude that  $(\boldsymbol{\rho}, \boldsymbol{\phi}) \sim \text{MultiEwens}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$  and  $(C_{n;1}, \dots, C_{n;d})$  has the multinomial Stirling distribution  $\text{MultiStir1}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$ .  $\square$

Since the cycles are painted independently, the following is a simple consequence of the marking theorem for Poisson processes in conjunction with Theorem 5.1 in [2].

**Proposition 2.7.** *Let  $(\boldsymbol{\rho}, \boldsymbol{\phi})$  be the multinomial Ewens permutation and let  $C_{n;j;k}$  be the number of cycles of color  $j \in \{1, \dots, d\}$  and length  $k \in \mathbb{N}$ . Then,*

$$(C_{n;j;k} : j \in \{1, \dots, d\}, k \in \mathbb{N}) \xrightarrow[n \rightarrow \infty]{d} (\text{Poi}[\theta_j/k] : j \in \{1, \dots, d\}, k \in \mathbb{N})$$

*in the product topology on  $\mathbb{N}_0^{d \times \mathbb{N}}$ , where the limiting Poisson variables are mutually independent.*

**2.4. Multinomial Hoppe forests.** The classic sequential construction of a *random recursive tree*  $\text{RRT}[n]$  with  $n$  nodes  $1, \dots, n$  and the root labeled by 0 proceeds by recursively attaching node  $\ell \in \{1, \dots, n\}$  to one of the existing nodes  $0, 1, \dots, \ell - 1$  picked uniformly at random and independently of the previous choices; see, e.g., [62]. The *Hoppe tree*  $\text{Hoppe}[n; \theta]$  is defined similarly, see [43], but now the probability of attaching the node  $\ell$  to the root is  $\theta/(\theta + \ell - 1)$ , while the probability of attaching  $\ell$  to any other node  $1, \dots, \ell - 1$  is  $1/(\theta + \ell - 1)$ . Here,  $\theta > 0$  is a parameter representing the ‘weight’ of the root, while the ‘weight’ of any other node is 1. In the special case  $\theta = 1$  one recovers the random recursive tree  $\text{RRT}[n]$ . The name ‘Hoppe tree’ originates from the

corresponding Hoppe's urn model introduced in [25, 26, 27] in the context of the Ewens sampling formula.

We now define a multinomial generalization of Hoppe trees, called *multinomial Hoppe forest* and denoted by  $\text{MultiHoppe}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$ . Fix parameters  $\theta_1, \dots, \theta_d \geq 0$  and suppose that  $\theta := \theta_1 + \dots + \theta_d > 0$ . Start at time 0 with  $d$  roots, labeled  $\text{root}_1, \dots, \text{root}_d$ , which have weights  $\theta_1, \dots, \theta_d$ . Nodes, labeled by  $1, 2, \dots, n$ , arrive in discrete time and are added as children to the roots (with probabilities proportional to the weights  $\theta_1, \dots, \theta_d$ ) or to the nodes already present in the forest (with probability proportional to 1 for each node). More precisely, suppose that the nodes  $1, \dots, \ell - 1$  have already arrived, where  $\ell \in \{1, \dots, n\}$ . Then, the next node, with label  $\ell$ , arrives at time  $\ell$  and is attached to  $\text{root}_j$  with probability  $\theta_j/(\theta + \ell - 1)$ , for each  $j \in \{1, \dots, d\}$ , or to any of the already existing nodes  $1, \dots, \ell - 1$ , with probability  $1/(\theta + \ell - 1)$  for each node. The usual Hoppe trees are recovered by taking  $d = 1$ .

Recall that a random vector  $(X_1, \dots, X_d)$  has the *Dirichlet multinomial distribution*, denoted hereafter by  $\text{DirMult}[n; \theta_1, \dots, \theta_d]$ , if

$$\begin{aligned} \mathbb{P}[(X_1, \dots, X_d) = (k_1, \dots, k_d)] \\ = \frac{\Gamma(n+1)\Gamma(\theta_1 + \dots + \theta_d)}{\Gamma(n + \theta_1 + \dots + \theta_d)} \prod_{i=1}^d \frac{\Gamma(k_i + \theta_i)}{\Gamma(\theta_i)\Gamma(k_i + 1)}, \quad k_i \geq 0, \quad \sum_{i=1}^d k_i = n, \end{aligned} \quad (2.7)$$

see Eq. (35.151) in [35] with the convention that  $\mathbb{P}[X_i = 0] = 1$  if  $\theta_i = 0$ , for some  $i \in \{1, \dots, d\}$ .

Let  $\text{Beta}[\alpha; \beta]$  denote the beta distribution with parameters  $\alpha, \beta > 0$  and having the density  $x \mapsto x^{\alpha-1}(1-x)^{\beta-1}/\text{B}(\alpha, \beta)$  for  $x \in (0, 1)$ . If  $\alpha = 0$  (respectively  $\beta = 0$ ), we stipulate that  $\text{Beta}[\alpha; \beta]$  is the degenerate distribution at 0 (respectively, 1). The following result is standard, see [34, Chapter 4.5.1].

**Proposition 2.8.** *The joint distribution of the sizes of the connected components of  $\text{root}_1, \dots, \text{root}_d$  in  $\text{MultiHoppe}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$  (not counting the roots themselves) is  $\text{DirMult}[n; \theta_1, \dots, \theta_d]$ . In particular, the size of the connected component of  $\text{root}_j$  (not counting  $\text{root}_j$  itself) has the so-called beta-binomial distribution  $\text{Bin}[n; \text{Beta}[\theta_j; \theta - \theta_j]]$ , for all  $j = 1, \dots, d$ .*

**Proposition 2.9.** *For each  $j \in \{1, \dots, d\}$ , let  $D_{n;j}$  be the degree of  $\text{root}_j$  in the random forest  $\text{MultiHoppe}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$ . Then,*

$$(D_{n;1}, \dots, D_{n;d}) \sim \text{MultiStir1}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)].$$

*Proof.* For  $\ell \in \{1, \dots, n\}$  let  $I_\ell$  be a random vector in  $\mathbb{N}_0^d$  with distribution (2.5) and assume that  $I_1, \dots, I_n$  are mutually independent. We interpret the event  $\{I_\ell = e_j\}$  as  $\{\text{the node } \ell \text{ was attached to } \text{root}_j\}$  and the event  $\{I_\ell = 0\}$  as  $\{\ell \text{ was attached to one of the nodes } 1, \dots, \ell - 1\}$ . Then,  $(D_{n;1}, \dots, D_{n;d})$  has the same distribution as  $I_1 + \dots + I_n$ . It remains to apply Proposition 2.4.  $\square$

**2.4.1. Expected profiles of multinomial Hoppe forests.** The expected number of nodes in an  $\text{RRT}[n]$  having distance  $k$  to the root is known to be  $\frac{1}{n!} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix}$ , for all  $k \in \{1, \dots, n\}$ ; see [18, Lemma 6.16] or [62, Theorem 1] (actually both references attribute this formula to [17]). The next theorem generalizes this result by identifying the expected profile of the multinomial Hoppe trees. Interestingly, even for  $d = 1$  and arbitrary  $\theta > 0$ , the next formula contains the  $\theta$ -Stirling numbers.

**Theorem 2.10.** *Consider  $\text{MultiHoppe}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$  and let  $L_{n;j}(k)$  be the number of nodes in the connected component of the root  $j \in \{1, \dots, d\}$  which have distance  $k$  to the root*

*j*. Then,

$$\mathbb{E}[L_{n;j}(k)] = \frac{\theta_j}{\theta^{n\uparrow}} \left[ \begin{matrix} n \\ k \end{matrix} \right]_{\theta}, \quad k \in \{1, \dots, n\}, \quad j \in \{1, \dots, d\}. \quad (2.8)$$

*Proof.* Denote the left-hand side of (2.8) by  $a_j(n, k) := \mathbb{E}[L_{n;j}(k)]$ . For  $n \in \mathbb{N}_0$ , let  $\mathcal{F}_n$  be the  $\sigma$ -algebra generated by the multinomial Hoppe forest, which starts with  $d$  roots, after adding  $n$  new nodes. Then  $a_j(0, k) = 0$  for  $k \in \mathbb{N}$  and, for  $k \geq 2$ ,

$$\begin{aligned} a_j(n, k) &= \mathbb{E}[L_{n;j}(k)] = \mathbb{E}[\mathbb{E}[L_{n;j}(k)] | \mathcal{F}_{n-1}] = \mathbb{E} \left[ L_{n-1;j}(k) + \frac{L_{n-1;j}(k-1)}{n + \theta - 1} \right] \\ &= a_j(n-1, k) + \frac{a_j(n-1, k-1)}{n + \theta - 1}, \quad n \in \mathbb{N}, \quad j = 1, \dots, d, \end{aligned} \quad (2.9)$$

where  $1/(n + \theta - 1)$  is the probability that the node  $n$  joins a particular node at the level  $k - 1$ . For  $k = 1$ , the recursion is slightly different accounting to a special weight of the root  $j$ :

$$a_j(n, 1) = a_j(n-1, 1) + \frac{\theta_j}{n + \theta - 1}, \quad n \in \mathbb{N}, \quad j = 1, \dots, d, \quad (2.10)$$

where  $\theta_j/(n + \theta - 1)$  is the probability of joining the root  $j$ . Introduce the generating functions

$$\phi_{n;j}(z) := \theta_j + \sum_{k=1}^{\infty} a_j(n, k) z^k, \quad n \in \mathbb{N}_0, \quad j = 1, \dots, d.$$

Then (2.9) and (2.10) imply

$$\phi_{0;j}(z) = \theta_j, \quad \phi_{n;j}(z) = \left( 1 + \frac{z}{n + \theta - 1} \right) \phi_{n-1;j}(z), \quad n \in \mathbb{N}.$$

Thus,

$$\phi_{n;j}(z) = \theta_j \prod_{k=0}^{n-1} \left( 1 + \frac{z}{k + \theta} \right) = \theta_j \frac{(\theta + z)^{n\uparrow}}{\theta^{n\uparrow}} = \frac{\theta_j}{\theta^{n\uparrow}} \sum_{k=0}^n \left[ \begin{matrix} n \\ k \end{matrix} \right]_{\theta} z^k, \quad n \in \mathbb{N}_0,$$

and (2.8) follows.  $\square$

*Remark 2.11.* Observe that the expected number of nodes in the  $j$ -th tree (not counting the root) is

$$\sum_{k=1}^n \mathbb{E}[L_{n;j}(k)] = \frac{\theta_j}{\theta^{n\uparrow}} \sum_{k=1}^n \left[ \begin{matrix} n \\ k \end{matrix} \right]_{\theta} \stackrel{(1.2)}{=} \frac{\theta_j}{\theta^{n\uparrow}} ((1 + \theta)^{n\uparrow} - \theta^{n\uparrow}) = n\theta_j/\theta,$$

as it should be. We have used here that  $\left[ \begin{matrix} n \\ 0 \end{matrix} \right]_{\theta} = \theta^{n\uparrow}$ .

**2.4.2. Leaves in multinomial Hoppe forests and generalized Eulerian numbers.** Let  $T_n$  be the number of leaves in a random recursive tree  $\text{RRT}[n]$  with one root and  $n$  other nodes. According to a result of Najock and Heyde [52] (alternatively, see Eq. (6.15) in [18]):

$$\mathbb{P}[T_n = k] = \frac{1}{n!} \left\langle \begin{matrix} n \\ k-1 \end{matrix} \right\rangle, \quad k = 1, \dots, n, \quad n \in \mathbb{N}, \quad (2.11)$$

where the second factor is the standard Eulerian number, see [22, Chapter 6.2] and [55]. For our purposes the following recursive definition of the Eulerian number serves best:

$$\left\langle \begin{matrix} n \\ k \end{matrix} \right\rangle = (n - k) \left\langle \begin{matrix} n-1 \\ k-1 \end{matrix} \right\rangle + (k + 1) \left\langle \begin{matrix} n-1 \\ k \end{matrix} \right\rangle, \quad 0 \leq k < n, \quad n \in \mathbb{N} \quad (2.12)$$

with the boundary conditions  $\langle \binom{n}{0} \rangle = 1$ ,  $n \in \mathbb{N}_0$ ,  $\langle \binom{n}{n} \rangle = 0$ ,  $n \in \mathbb{N}$ . In the following we generalize (2.11) first to Hoppe trees and then to Hoppe forests. Put  $A(r, s) := \langle \binom{r+s+1}{s} \rangle$  and note that (2.12) implies

$$A(r, s) = (r+1)A(r, s-1) + (s+1)A(r-1, s), \quad r, s \in \mathbb{N}_0, \quad r+s > 0,$$

with  $A(0, 0) = 1$  and  $A(r, s) = 0$  if  $r = -1$  or  $s = -1$ . This formula suggests the following generalization of Eulerian numbers, proposed by Carlitz and Scoville [8]. Let  $\alpha, \beta$  be arbitrary real parameters. Put

$$A(r, s|\alpha, \beta) = (r+\beta)A(r, s-1|\alpha, \beta) + (s+\alpha)A(r-1, s|\alpha, \beta), \quad r, s \in \mathbb{N}_0, \quad r+s > 0, \quad (2.13)$$

and  $A(0, 0|\alpha, \beta) = 1$  and  $A(r, s|\alpha, \beta) = 0$  if  $r = -1$  or  $s = -1$ . Note that  $A(r, s) = A(r, s|1, 1)$ . According to Eq. (1.8) in [8],

$$\sum_{r,s=0}^{\infty} A(r, s|\alpha, \beta) \frac{u^r v^s}{(r+s)!} = (1 + uF(u, v))^\alpha (1 + vF(u, v))^\beta,$$

where

$$F(u, v) = \frac{e^u - e^v}{ue^v - ve^u}.$$

It turns out that the generalized Eulerian numbers defined by (2.13) are an indispensable ingredient of the version of (2.11) for Hoppe trees.

**Theorem 2.12** (Leaves in Hoppe trees). *Let  $T_n$  be the number of leaves in a random tree  $\text{Hoppe}[n; \theta]$  with  $\theta > 0$ . Then,*

$$\mathbb{P}[T_n = k] = \frac{A(n-k, k|0, \theta)}{\theta^{n\uparrow}}, \quad k = 1, \dots, n, \quad n \in \mathbb{N}. \quad (2.14)$$

*Proof.* Denote by  $p_n(k) := \mathbb{P}[T_n = k]$  the left-hand side of (2.14) and let  $\mathcal{F}_n$  denote the  $\sigma$ -algebra generated by the Hoppe tree upon adding  $n$  nodes. Note that the root cannot be a leaf if  $n \geq 1$  and it is also convenient not to regard it as a leaf if  $n = 0$ , thus we put  $p_0(0) = 1$ . Then,

$$p_n(k) = \mathbb{E}[\mathbb{P}[T_n = k|\mathcal{F}_{n-1}]] = \frac{k}{n+\theta-1}p_{n-1}(k) + \frac{n-k+\theta}{n+\theta-1}p_{n-1}(k-1), \quad k = 0, \dots, n, \quad n \in \mathbb{N},$$

where we stipulate that  $p_n(k) = 0$ , for  $k > n$ , and  $p_n(-1) = 0$ . Put  $Q_{n;\ell} := \theta^{(n+\ell)\uparrow} p_{n+\ell}(\ell)$  and note that the above recursion transforms into

$$Q_{n;\ell} = \ell Q_{n-1;\ell} + (n+\theta)Q_{n;\ell-1}, \quad n \in \mathbb{N}_0, \quad \ell \in \mathbb{N}_0, \quad n+\ell > 0$$

with the initial values

$$Q_{0,0} = 1, \quad Q_{-1;\ell} = Q_{n;-1} = 0, \quad n, \ell \in \mathbb{N}_0.$$

Thus,  $Q_{n;\ell} = A(n, \ell|0, \theta)$  and (2.14) follows.  $\square$

**Example 2.13.** Note that the RRT result (2.11) is a particular case when  $\theta = 1$ , since  $A(n-k, k|0, 1) = \langle \binom{n}{k-1} \rangle$  which, in turn, is a consequence of  $A(r, s|0, 1) = \langle \binom{r+s}{s-1} \rangle$ .

The generalized Eulerian distribution with probability mass function  $k \mapsto A(k, n-k|\alpha, \beta)/(\alpha + \beta)^{n\uparrow}$  for  $k \in \{0, \dots, n\}$  is investigated in [9, 30, 31] in the context of Morisita's work on mutual repulsive behavior of ant lions [51]. Among other results, Charalambides [9] computes the factorial moments of this distribution and proves a central limit theorem.

*Remark 2.14.* The  $r$ -analogues of Eulerian numbers appearing in [50, § 8.5–8.7] and [49] and denoted there by  $\langle \binom{n}{k} \rangle_r$  satisfy the same recurrence relation, namely,  $\langle \binom{n}{k} \rangle_r = (n - k + r) \langle \binom{n-1}{k-1} \rangle_r + (k + 1) \langle \binom{n-1}{k} \rangle_r$  (see relation (8.30) on p. 218 in [50]) as the numbers  $A(n - k - 1, k + 1 | 0, r + 1)$  (see (2.13)). However, the boundary conditions imposed on these arrays are different; see [50, p. 216]. Related generalizations of Eulerian numbers can be found in [32] and [45].

**Theorem 2.15** (Leaves in Hoppe forests). *Let  $T_{n;j}$  be the number of leaves in the connected component of the root  $j \in \{1, \dots, d\}$  in the multinomial Hoppe forest  $\text{MultiHoppe}[n; \theta; (\theta_1/\theta, \dots, \theta_d/\theta)]$  (if the root  $j$  has no descendants, we stipulate that  $T_{n;j} = 0$ ). Then,*

$$\mathbb{P}[T_{n;j} = k] = \frac{1}{\theta^{n\uparrow}} \sum_{m=k}^n \binom{n}{m} (\theta - \theta_j)^{(n-m)\uparrow} A(m - k, k | 0, \theta_j).$$

*Proof.* The number of nodes in the connected component of  $\text{root}_j$  (not counting the root itself) is  $S_{n;j} \sim \text{Bin}[n; \text{Beta}[\theta_j; \theta - \theta_j]]$ , see Proposition 2.8. Conditionally on  $S_{n;j} = m$ , the component of  $\text{root}_j$  is distributed as  $\text{Hoppe}[m; \theta_j]$ . By the total probability formula and Theorem 2.12, the number of leaves in the connected component of  $\text{root}_j$  satisfies

$$\begin{aligned} \mathbb{P}[T_{n;j} = k] &= \sum_{m=k}^n \mathbb{P}[S_{n;j} = m] \cdot \frac{A(m - k, k | 0, \theta_j)}{\theta_j^{m\uparrow}} \\ &= \sum_{m=k}^n \binom{n}{m} \frac{\Gamma(m + \theta_j) \Gamma(n - m + \theta - \theta_j) \Gamma(\theta)}{\Gamma(n + \theta) \Gamma(\theta - \theta_j) \Gamma(\theta_j)} \cdot \frac{\Gamma(\theta_j)}{\Gamma(m + \theta_j)} A(m - k, k | 0, \theta_j) \\ &= \frac{1}{\theta^{n\uparrow}} \sum_{m=k}^n \binom{n}{m} (\theta - \theta_j)^{(n-m)\uparrow} A(m - k, k | 0, \theta_j), \end{aligned}$$

which proves the claim. □

**Corollary 2.16** (Subtrees of the  $\ell$ -th node in a Hoppe tree). *Consider a  $\text{Hoppe}[n; \theta]$ -tree and let  $\ell \in \{1, \dots, n\}$  be a node. If  $T_n^{(\ell)}$  denotes the number of leaves in the subtree rooted at the node  $\ell$ , then*

$$\mathbb{P}[T_n^{(\ell)} = k] = \frac{1}{(\theta + \ell)^{(n-\ell)\uparrow}} \sum_{m=k}^{n-\ell} \binom{n-\ell}{m} (\ell + \theta - 1)^{(n-\ell-m)\uparrow} \left\langle \binom{m}{k-1} \right\rangle \quad k = 1, \dots, n - \ell$$

and  $\mathbb{P}[T_n^{(\ell)} = 0] = (\ell - 1 + \theta)/(n - 1 + \theta)$ , which is the probability that the subtree only consists of  $\ell$ .

*Proof.* Indeed, after the node  $\ell$  has arrived, we can consider it as  $\text{root}_1$  with weight  $\theta_1 = 1$  and combine the nodes  $1, \dots, \ell - 1$  and the old root to form  $\text{root}_2$  with weight  $\theta_2 = \ell - 1 + \theta$ . The remaining nodes  $\ell + 1, \dots, n$  behave as in the corresponding multinomial Hoppe tree with  $d = 2$ . Applying Theorem 2.15 and recalling that  $A(m - k, k | 0, 1) = \langle \binom{m}{k-1} \rangle$  gives the claim. □

In the case of RRT (that is, for  $\theta = 1$ ) our formula is equivalent to the formula of Mahmoud and Smythe [44, p. 410]. Note, however, that they stipulate a tree consisting of a single root to have one leaf, whereas we declare it to have no leaves. Summarizing, their formula gives the distribution of  $\max(T_{n-1}^{(\ell)}, 1)$ .

### 3. $r$ -EWENS PERMUTATIONS, RANDOM $r$ -RECURSIVE TREES

Now we are going to define the  $r$ -analogues of the random structures introduced above. The general procedure is as follows. Fix parameters  $\tau = \theta_1 \geq 0$  and  $r = \theta_2 \geq 0$  such that  $\theta = \tau + r > 0$ . Consider some multinomial random structure with  $d = 2$  colors having probabilities  $p_1 := \tau/(\tau + r)$  and  $p_2 := r/(\tau + r)$  and discard all elements having color 2. The remaining elements of color 1 form a random structure on some subset  $B \subseteq [n]$  (which is also random and, moreover, can be empty with positive probability). This random structure, defined on a random subset, is the  $r$ -analogue we are searching for.

**3.1.  $r$ -Ewens permutations.** The  $r$ -Ewens distribution  $r$ -Ewens $[n; \tau]$  is informally defined as follows. Consider a random permutation following the Ewens $[n; \tau + r]$  distribution. Independently remove each cycle of this permutation with probability  $r/(\tau + r)$ . The remaining cycles define a random permutation of some subset  $B$  of  $[n]$ , which we call the  $r$ -Ewens permutation and whose distribution is  $r$ -Ewens $[n; \tau]$ . We now give a precise definition.

**Definition 3.1.** An *incomplete permutation* on  $[n]$  is a pair  $(\rho, B)$ , where  $B \subseteq [n]$  is a subset of  $[n]$ , which might be empty, and  $\rho : B \rightarrow B$  is a bijection. The set  $B$  is called the *domain* of  $\rho$  and its elements is colored white. The complement of  $B$  is denoted by  $B_0 := [n] \setminus B$  and its elements are colored red.

The set of incomplete permutations of  $[n]$  is denoted by  $\mathfrak{S}_n^{\text{inc}} = \{(\rho, B) : B \subseteq [n], \rho \in \mathfrak{S}_B\}$ , where  $\mathfrak{S}_B$  is the symmetric group of  $B$ .

**Definition 3.2.** An  *$r$ -Ewens permutation* with parameters  $\tau \geq 0$  and  $r \geq 0$  such that  $\tau + r > 0$  is a random variable  $(\rho, \mathbf{B})$  taking values in  $\mathfrak{S}_n^{\text{inc}}$  with the distribution

$$\mathbb{P}[(\rho, \mathbf{B}) = (\rho, B)] = r^{(n-\#B)\uparrow} \cdot \frac{\tau^{\#\text{cycles}(\rho)}}{(\tau + r)^{n\uparrow}}, \quad (\rho, B) \in \mathfrak{S}_n^{\text{inc}}, \quad (3.1)$$

which is called the  $r$ -Ewens $[n; \tau]$  distribution. If  $\tau = 1$ , then we say that  $(\rho, \mathbf{B})$  is an  *$r$ -uniform random permutation*.

**Proposition 3.3.** Consider a multinomial Chinese restaurant process with  $d = 2$  colors having weights  $\theta_1 = \tau$  and  $\theta_2 = r$ . Let  $(\rho, \mathbf{B})$  be an incomplete permutation whose cycles are the tables of color 1 and whose domain  $\mathbf{B}$  is the set of all customers sitting at tables of color 1. Let  $\mathbf{B}_0 := [n] \setminus \mathbf{B}$  be the set of all customers sitting at tables of color 2. Then,  $(\rho, \mathbf{B}) \sim r$ -Ewens $[n; \tau]$ .

*Proof.* Fix  $(\rho, B) \in \mathfrak{S}_n^{\text{inc}}$  and let  $\Omega_{\rho, B}$  be the set of colored permutations of  $[n]$  (with two colors) such that the elements belonging to cycles with color 2 are precisely the elements of the set  $B_0$  and cycles of color 1 form permutation  $\rho$  of  $B$ . According to Proposition 2.6 we need to check that the sum of the right-hand sides of (2.3) over the set  $\Omega_{\rho, B}$  is equal to the right-hand side of (3.1). Indeed,

$$\sum_{(\rho', \phi) \in \Omega_{\rho, B}} \frac{\tau^{\#\text{cycles}_1(\rho', \phi)} r^{\#\text{cycles}_2(\rho', \phi)}}{(t + r)^{n\uparrow}} = \frac{\tau^{\#\text{cycles}(\rho)}}{(t + r)^{n\uparrow}} r^{(n-\#B)\uparrow} \sum_{(\rho', \phi) \in \Omega_{\rho, B}} \frac{r^{\#\text{cycles}_2(\rho', \phi)}}{r^{(n-\#B)\uparrow}}.$$

The sum is equal to 1, since the summands form a probability distribution Ewens $[n - \#B; r]$  and the summation is taken over all possible cyclic structures on  $[n] \setminus B$ .  $\square$

**Proposition 3.4.** Let  $A_{n,r}(b_0, b_1, \dots, b_k)$  be the event that  $(\boldsymbol{\rho}, \mathbf{B}) \sim r$ -Ewens $[n; \tau]$  has a red set of size  $b_0 \in \mathbb{N}_0$  and  $k$  white cycles of sizes  $b_1, \dots, b_k \in \mathbb{N}$ , where  $b_0 + b_1 + \dots + b_k = n$ . Then,

$$\mathbb{P}[A_{n,r}(b_0, b_1, \dots, b_k)] = \frac{n!}{k!} \frac{\tau^k}{(\tau + r)^{n\uparrow}} \cdot \frac{r^{b_0\uparrow}}{b_0!} \cdot \frac{1}{b_1 \cdots b_k}. \quad (3.2)$$

*Proof.* Equality (3.2) follows immediately from (3.1) upon noticing that  $\binom{n}{b_0, \dots, b_k} \frac{1}{k!}$  counts the number of partitions of  $[n]$  into blocks  $B_0, B_1, \dots, B_k$  of sizes  $b_0 \in \mathbb{N}_0, b_1, \dots, b_k \in \mathbb{N}$ , respectively, whereas  $(b_j - 1)!$  is the number of ways to organize a cycle on  $b_j$  elements,  $j = 1, \dots, k$ .  $\square$

**Proposition 3.5.** If  $(\boldsymbol{\rho}, \mathbf{B})$  is an  $r$ -Ewens $[n; \tau]$ -distributed incomplete permutation, then the number of cycles in  $\boldsymbol{\rho}$  has the  $r$ -Stir1 $[n; \tau]$  distribution, that is,

$$\mathbb{P}[\#\text{ cycles}(\boldsymbol{\rho}) = k] = \left[ \begin{matrix} n \\ k \end{matrix} \right]_r \frac{\tau^k}{(\tau + r)^{n\uparrow}}, \quad k \in \{0, \dots, n\}. \quad (3.3)$$

Before proceeding with a proof of Proposition 3.5 we note that upon multiplying Taylor expansions (1.3) and (1.4), one obtains

$$\left[ \begin{matrix} n \\ k \end{matrix} \right]_r = \frac{n!}{k!} \sum_{(b_0, b_1, \dots, b_k)} \frac{1}{b_1 \cdots b_k} \cdot \frac{r^{b_0\uparrow}}{b_0!}, \quad \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r = \frac{n!}{k!} \sum_{(b_0, b_1, \dots, b_k)} \frac{1}{b_1! \cdots b_k!} \cdot \frac{r^{b_0}}{b_0!}, \quad (3.4)$$

where the sums are taken over all  $(k + 1)$ -tuples  $(b_0, b_1, \dots, b_k)$  such that  $b_0 \in \mathbb{N}_0, b_1, \dots, b_k \in \mathbb{N}$  and  $b_0 + b_1 + \dots + b_k = n$ .

We give three elementary proofs of Proposition 3.5 with the aim of clarifying the probabilistic meaning of the first equality in (3.4) (the second will be explained later on, see Proposition 4.5) and also of the formulas

$$\left[ \begin{matrix} n \\ k \end{matrix} \right]_r = \sum_{j=k}^n \left[ \begin{matrix} n \\ j \end{matrix} \right] \binom{j}{k} r^{j-k} = \sum_{j=k}^n \binom{n}{j} \left[ \begin{matrix} j \\ k \end{matrix} \right]_r^{(n-j)\uparrow}, \quad (3.5)$$

see Eq. (1.5) in the introduction.

*First proof.* Formula (3.3) for the distribution of the number of cycles follows from the first equality in (3.4) upon summation of the right-hand sides of (3.2) over all  $b_0 \in \mathbb{N}_0$  and  $b_1, \dots, b_k \in \mathbb{N}$  satisfying  $b_0 + \dots + b_k = n$ .  $\square$

*Second proof.* We use the first equality in (3.5). Since the total number of cycles of both colors has the Stir1 $[n; \tau + r]$  distribution and the probability for each particular cycle to be of color 1 is  $\tau/(\tau + r)$  independently of the other cycles,

$$\mathbb{P}[\#\text{ cycles}(\boldsymbol{\rho}) = k] = \text{Bin}[\text{Stir1}[n; \tau + r]; \tau/(\tau + r)](\{k\}), \quad k \in \{0, \dots, n\}.$$

Thus, we need to check that

$$r\text{-Stir1}[n; \tau] = \text{Bin}[\text{Stir1}[n; \tau + r]; \tau/(\tau + r)]. \quad (3.6)$$

This equality follows from

$$\begin{aligned} \text{Bin}[\text{Stir1}[n; \tau + r]; \tau/(\tau + r)](\{k\}) &= \sum_{j=k}^n \text{Stir1}[n; \tau + r](\{j\}) \text{Bin}[j; \tau/(\tau + r)](\{k\}) \\ &= \sum_{j=k}^n \left[ \begin{matrix} n \\ j \end{matrix} \right] \frac{(\tau + r)^j}{(\tau + r)^{n\uparrow}} \binom{j}{k} \frac{\tau^k r^{j-k}}{(\tau + r)^j} \stackrel{(3.5)}{=} \left[ \begin{matrix} n \\ k \end{matrix} \right]_r \frac{\tau^k}{(\tau + r)^{n\uparrow}} = r\text{-Stir1}[n; \tau](\{k\}) \end{aligned}$$



for each  $k \in \{0, 1, \dots, n\}$ .  $\square$

*Third proof.* In this proof we exploit the second equality in (3.5). Using definition (3.1) of the  $r$ -Ewens distribution we obtain

$$\mathbb{P}[\#\text{ cycles}(\boldsymbol{\rho}) = k] = \sum_{j=k}^n \mathbb{P}[\#\text{ cycles}(\boldsymbol{\rho}) = k, \#B = j] \stackrel{(3.1)}{=} \sum_{j=k}^n \frac{r^{(n-j)\uparrow} \tau^k}{(\tau+r)^{n\uparrow}} \binom{n}{j} \left[ \begin{matrix} j \\ k \end{matrix} \right], \quad k \in \{0, \dots, n\},$$

where  $\binom{n}{j}$  counts the number of ways to choose  $j$  elements of  $[n]$  and  $\left[ \begin{matrix} j \\ k \end{matrix} \right]$  counts the number of ways to organize  $k$  cycles on the chosen elements. By the second equality in (3.5) the right-hand side of the last centered formula is equal to  $r\text{-Stir1}[n; \tau](\{k\})$ .  $\square$

Similarly to the classic scenario, the number of cycles in the  $r$ -Ewens permutation  $(\boldsymbol{\rho}, \mathbf{B})$  can be represented as the sum of independent indicators. This follows from the factorization

$$\mathbb{E}[s^{\#\text{ cycles}(\boldsymbol{\rho})}] = \frac{1}{(r+\tau)^{n\uparrow}} \sum_{k=0}^n \left[ \begin{matrix} n \\ k \end{matrix} \right]_r (\tau s)^k = \frac{(r+\tau s)^{n\uparrow}}{(r+\tau)^{n\uparrow}} = \prod_{j=1}^n \left( \frac{r+j-1}{r+\tau+j-1} + \frac{\tau s}{r+\tau+j-1} \right).$$

This representation immediately yields a central limit theorem for  $\#\text{ cycles}(\boldsymbol{\rho})$  and, therefore, for the  $r$ -Stirling distribution of the first kind. Interestingly, the limit theorem does not depend on  $r$  at all!

**Corollary 3.6.** *Assume that  $\tau, r \geq 0, \tau+r > 0$  are fixed and let  $(\boldsymbol{\rho}, \mathbf{B})$  be the  $r$ -Ewens permutation. Then,*

$$\frac{\#\text{ cycles}(\boldsymbol{\rho}) - \tau \log n}{\sqrt{\tau \log n}} \xrightarrow[n \rightarrow \infty]{d} \mathbf{N}[0; 1].$$

For the ease of reference we collect various representations for the  $r$ -Stirling distribution of the first kind derived above in a single

**Proposition 3.7.** *Fix  $r \geq 0, \tau > 0$ . The  $r$ -Stirling distribution of the first kind  $r\text{-Stir1}[n; \tau]$  admits the following representations:*

- (i)  $\sum_{j=1}^n \text{Bern}[\tau/(\tau+r+j-1)] \sim r\text{-Stir1}[n; \tau]$ , where the random variables on the left-hand side are independent;
- (ii)  $r\text{-Stir1}[n; \tau] = \text{Bin}[\text{Stir1}[n; \tau+r]; \tau/(\tau+r)]$ ;
- (iii)  $r\text{-Stir1}[n; \tau] = \text{Stir1}[\text{Bin}[n; \text{Beta}[\tau; r]]; \tau]$ , where  $\text{Beta}[\tau; r]$  is the beta-distribution.

*Proof.* Part (i) can be verified by comparing generating functions. Alternatively, it is just the last claim of Proposition 2.4 with  $d = 2, j = 1, \theta_1 = \tau, \theta_2 = r$ . Part (ii) is formula (3.6). For the proof of part (iii) note that, for  $k \in \{0, 1, \dots, n\}$ ,

$$\begin{aligned} \text{Stir1}[\text{Bin}[n; \text{Beta}[\tau; r]]; \tau](\{k\}) &= \sum_{m=0}^n \text{Bin}[n; \text{Beta}[\tau; r]](\{m\}) \text{Stir1}[m; \tau](\{k\}) \\ &= \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(m+\tau)\Gamma(n-m+r)}{\Gamma(n+\tau+r)} \frac{\Gamma(\tau+r)}{\Gamma(\tau)\Gamma(r)} \left[ \begin{matrix} m \\ k \end{matrix} \right] \frac{\tau^k}{\tau^{m\uparrow}} \\ &= \frac{\Gamma(\tau+r)}{\Gamma(n+\tau+r)} \tau^k \sum_{m=0}^n \binom{n}{m} \left[ \begin{matrix} m \\ k \end{matrix} \right]_r^{(n-m)\uparrow} \frac{\Gamma(m+\tau)}{\Gamma(\tau)\tau^{m\uparrow}} \\ &\stackrel{(1.5)}{=} \frac{\tau^k}{(\tau+r)^{n\uparrow}} \left[ \begin{matrix} n \\ k \end{matrix} \right]_r = r\text{-Stir1}[n; \tau](\{k\}). \end{aligned}$$

The proof is complete.  $\square$

*Remark 3.8.* The recent article [29] investigates properties of the Bernoulli trials with unequal harmonic-like success probabilities as in part (i) of Proposition 3.7.

We close the discussion of  $r$ -Ewens permutations by noting a combinatorial construction of the  $r$ -uniform permutations which works for integer  $r \in \mathbb{N}_0$  and establishes connections to Broder's Definition 1.1. Call the elements  $1, 2, \dots, n$  of the set  $[n+r]$  *white* and elements  $n+1, \dots, n+r$  *red*. Let  $\mathfrak{S}_{n,r}$  be the set of all permutations of  $[n+r]$  such that red elements are in different cycles and let  $\mathfrak{d}_{\text{perm}} : \mathfrak{S}_{n,r} \mapsto \mathfrak{S}_n^{\text{inc}}$  be the mapping defined as follows. For  $\sigma \in \mathfrak{S}_{n,r}$ , the image  $\mathfrak{d}_{\text{perm}}(\sigma)$  is an incomplete permutation  $(\rho, B) \in \mathfrak{S}_n^{\text{inc}}$  such that the red set  $B_0 = [n] \setminus B$  is obtained by merging  $r$  cycles of  $\sigma$  containing red elements, removing these red elements from the union, and keeping other cycles unchanged.

**Proposition 3.9.** *Fix  $n \in \mathbb{N}$ ,  $r \in \mathbb{N}_0$ . Let  $\hat{\sigma}_{n,r}$  be a permutation of  $[n+r]$  picked uniformly at random from  $\mathfrak{S}_{n,r}$ . Then  $\mathfrak{d}_{\text{perm}}(\hat{\sigma}_{n,r})$  is the  $r$ -uniform random permutation ( $r$ -Ewens permutation with  $\tau = 1$ ).*

*Proof.* Each  $\sigma \in \mathfrak{S}_{n,r}$  can be constructed from an incomplete permutation  $(\sigma_1, B) \in \mathfrak{S}_n^{\text{inc}}$  of  $[n]$ , by distributing  $n - \#B$  elements of the block  $[n] \setminus B$  among  $r$  cycles of  $\sigma$  containing  $n+1, \dots, n+r$  red elements. Thus,

$$\begin{aligned} \#\mathfrak{S}_{n,r} &= \sum_{j=0}^n \#\{(\sigma_1, B) \in \mathfrak{S}_n^{\text{inc}} : \#B = j\} r^{(n-j)\uparrow} \\ &= \sum_{j=0}^n \binom{n}{j} j! r^{(n-j)\uparrow} = \sum_{j=0}^n \binom{n}{j} 1! j! r^{(n-j)\uparrow} = (1+r)^{n\uparrow}, \end{aligned}$$

where the last equality follows from the binomial theorem for raising factorials. By the same reasoning, for each incomplete permutation  $(\sigma_1, B) \in \mathfrak{S}_n^{\text{inc}}$ ,

$$\#\{\sigma \in \mathfrak{S}_{n,r} : \mathfrak{d}_{\text{perm}}(\sigma) = (\sigma_1, B)\} = r^{(n-\#B)\uparrow}.$$

The proof concludes by taking the ratio.  $\square$

**3.2. Random  $r$ -recursive trees and  $r$ -Hoppe trees.** The construction is very similar to the construction of the  $r$ -Ewens permutations.

**Definition 3.10.** Fix  $n \in \mathbb{N}_0$ . An *incomplete recursive tree* with the root 0 and  $n$  additional nodes is a pair  $(T, B)$ , where  $B \subseteq [n]$  and  $T$  is a recursive tree with the root 0 and  $\#B$  nodes labeled by the elements of  $B$  and colored white. The elements of  $[n] \setminus B$  are called red nodes (these do not belong to the tree).

Denote by  $\mathcal{T}_n$  the set of all incomplete recursive trees with the root 0 and  $n$  additional nodes.

**Definition 3.11.** Fix parameters  $r \geq 0$  and  $\tau > 0$ . A *random  $r$ -Hoppe tree*  $r\text{-Hoppe}[n; \tau]$  with the root 0 and  $n$  additional nodes is a random element  $(\mathbf{T}, \mathbf{B})$  of  $\mathcal{T}_n$  with the distribution

$$\mathbb{P}[(\mathbf{T}, \mathbf{B}) = (T, B)] = \frac{r^{(n-\#B)\uparrow} \tau^{\deg_T 0}}{(r+\tau)^{n\uparrow}}, \quad (T, B) \in \mathcal{T}_n,$$

where  $\deg_T 0$  is the degree of the root 0 in  $T$ . A *random  $r$ -recursive tree* is a special case corresponding to  $\tau = 1$ .

Note that this definition is correct in a sense that the quantities on the right-hand side sum up to one. Indeed,

$$\begin{aligned} \sum_{(T,B) \in \mathcal{T}_n} \frac{r^{(n-\#B)\uparrow} \tau^{\deg 0}}{(r+\tau)^{n\uparrow}} &= \sum_{j=0}^n \sum_{k=0}^j \frac{r^{(n-j)\uparrow} \tau^k}{(r+\tau)^{n\uparrow}} \sum_{\substack{(T,B) \in \mathcal{T}_n \\ \#B=j, \deg 0=k}} 1 = \frac{\sum_{j=0}^n r^{(n-j)\uparrow} \binom{n}{j} \sum_{k=0}^j \binom{j}{k} \tau^k}{(r+\tau)^{n\uparrow}} \\ &= \frac{\sum_{j=0}^n \binom{n}{j} r^{(n-j)\uparrow} \tau^{j\uparrow}}{(r+\tau)^{n\uparrow}} = 1, \end{aligned}$$

where the penultimate equality is a consequence of the fact that there are  $\binom{n}{j}$  ways to pick  $j$  white nodes and  $\binom{j}{k}$  ways to construct a recursive tree with the root 0 and  $j$  additional white nodes such that the  $\deg 0 = k$ . The last equality follows from the binomial-type formula for the rising factorial.

The next proposition provides a sequential construction of the  $r$ -Hoppe trees.

**Proposition 3.12.** *Consider a Hoppe forest  $\text{MultiHoppe}[n; \tau+r; (\tau/(\tau+r), r/(\tau+r))]$  with  $d = 2$  roots  $\text{root}_1 := 0$  and  $\text{root}_2$  having weights  $\theta_1 = \tau > 0$  and  $\theta_2 = r \geq 0$ , and with  $n$  additional nodes. Let  $\mathbf{T}$  be the connected component of  $\text{root}_1$  and  $\mathbf{B}$  the set of labels of the nodes of  $\mathbf{T}$ . Then,  $(\mathbf{T}, \mathbf{B})$  is distributed as a random tree  $r\text{-Hoppe}[n; \tau]$ .*

*Proof.* Let  $(T, B)$  be a fixed incomplete recursive tree. The event  $\{(\mathbf{T}, \mathbf{B}) = (T, B)\}$  occurs if, and only if, the nodes with labels in  $[n] \setminus B$  were attached to the subtree rooted at  $\text{root}_2$  and the nodes of  $B$  formed a particular tree  $T$  rooted at  $\text{root}_1$ . By definition of the Hoppe forest, the probability of the former event is

$$r^{(n-\#B)\uparrow} \prod_{j \in [n] \setminus B} \frac{1}{j-1+\tau+r},$$

whereas the probability of the latter (given the former) is

$$\tau^{\deg_T 0} \prod_{j \in B} \frac{1}{j-1+\tau+r}.$$

Thus,  $\mathbb{P}[(\mathbf{T}, \mathbf{B}) = (T, B)] = r^{(n-\#B)\uparrow} \tau^{\deg_T 0} / (r+\tau)^{n\uparrow}$ . The proof is complete.  $\square$

Observe that if  $r = 0$ , then in the setting of Proposition 3.12 no nodes are attached to  $\text{root}_2$  (that is,  $\mathbb{P}[\#\mathbf{B} = n] = 1$ ) and  $\mathbf{T}$  becomes the usual  $\text{Hoppe}[n; \tau]$  tree. If, additionally,  $\tau = 1$ , we recover  $\text{RRT}[n]$ .

The following is a specialization of the results on multinomial Hoppe forests proved in Section 2.4.

**Proposition 3.13.** *Consider a random  $r$ -Hoppe tree  $(\mathbf{T}, \mathbf{B}) \sim r\text{-Hoppe}[n; \tau]$ .*

- (a) *The root degree of  $\mathbf{T}$  satisfies  $\deg \text{root}_1 \sim r\text{-Stir1}[n; \tau]$ , that is,  $\mathbb{P}[\deg \text{root}_1 = k] = \frac{\tau^k}{(\tau+r)^{n\uparrow}} \binom{n}{k}_r$  for all  $k = 0, \dots, n$ .*
- (b) *The expected number of nodes in  $\mathbf{T}$  at distance  $k$  from the  $\text{root}_1$  is  $\tau \binom{n}{k}_{\tau+r} / (\tau+r)^{n\uparrow}$ .*
- (c) *The number of leaves of  $\mathbf{T}$ , denoted by  $T_{n;1}$  has the following distribution:*

$$\mathbb{P}[T_{n;1} = k] = \frac{1}{(\tau+r)^{n\uparrow}} \sum_{m=k}^n \binom{n}{m} r^{(n-m)\uparrow} A(m-k, k|0, \tau), \quad k = 1, \dots, n.$$

#### 4. RANDOM $r$ -PARTITIONS

Recall that a *partition* of  $[n]$  is a collection  $B_1, B_2, \dots, B_k$  of pairwise disjoint, nonempty subsets  $B_j \subseteq [n]$  called blocks, whose union is  $[n]$ . The number  $k$  of blocks may vary from 1 for the trivial one-block partition, to  $n$  for the finest partition into singletons. Let  $\Pi_n$  be the set of all partitions of  $[n]$ .

A natural way to generate a partition of  $[n]$  is via an urn scheme. Let  $N \in \mathbb{N}$  be a fixed integer. Drop  $n$  balls labeled by the elements of  $[n]$  into  $N$  urns uniformly at random and independently. Define a random partition  $\lambda_{n,N}$  of  $[n]$  by declaring two balls to be in the same block if, and only if, they fall into the same urn. As has already been mentioned in the introduction, the number  $\#\text{blocks}(\lambda_{n,N})$  of blocks in the resulting partition  $\lambda_{n,N}$  has the Stirling–Sibuya distribution  $\text{StirSibuya}[n; N]$  defined by Eq. (1.8):

$$\mathbb{P}[\#\text{blocks}(\lambda_{n,N}) = k] = \binom{n}{k} \frac{N^{k\downarrow}}{N^n} = \text{StirSibuya}[n; N](\{k\}), \quad k \in \{1, \dots, \min(n, N)\}. \quad (4.1)$$

**4.1. Incomplete partitions derived from urn schemes.** Next, we introduce an  $r$ -version of (4.1). To this end, it is necessary to replace partitions by incomplete partitions.

**Definition 4.1.** An *incomplete partition* of  $[n]$  is a collection  $(B_0, \{B_1, B_2, \dots, B_k\})$  of pairwise disjoint subsets  $B_j \subseteq [n]$ , whose union is  $[n]$  and  $B_j \neq \emptyset$  for all  $j \in \{1, \dots, k\}$ . Note that  $B_0$ , called the *red block*, is allowed to be empty. The blocks  $B_1, \dots, B_k$ , called the *white blocks*, are non-empty. Let  $\Pi_n^{\text{inc}}$  be the set of all incomplete partitions of  $[n]$ .

The already constructed  $r$ -versions of Ewens permutations and random recursive trees suggest the following way to generate a random incomplete partition leading to an  $r$ - $\text{StirSibuya}[n; N]$ -distribution. We consider an urn scheme with  $N + 1$  urns  $0, 1, \dots, N$ . The urn 0 is red and has frequency  $r/(N + r) \geq 0$ , whereas urns  $1, 2, \dots, N$  are white and have frequencies  $1/(N + r)$ . Drop  $n$  balls into the urns and define an incomplete partition  $\lambda_{n,N,r}$  of  $[n]$  by declaring the balls in the urn 0 to form the red block  $B_0$  whereas the balls in those urns  $1, \dots, N$  that are non-empty to form white blocks as before.

**Proposition 4.2.** *Let  $N \in \mathbb{N}$  be an integer and  $r \geq 0$ . Then,*

$$\mathbb{P}[\lambda_{n,N,r} = (B_0, \{B_1, \dots, B_k\})] = \frac{r^{\#B_0} N^{k\downarrow}}{(N + r)^n}, \quad (B_0, \{B_1, \dots, B_k\}) \in \Pi_n^{\text{inc}}. \quad (4.2)$$

*Proof.* The formula follows upon noticing that

$$\frac{r^{\#B_0} N^{k\downarrow}}{(N + r)^n} = \left( \frac{r}{N + r} \right)^{\#B_0} \left( \frac{N^{k\downarrow}}{(N + r)^{n - \#B_0}} \right).$$

The first factor on the right-hand side is the probability for balls with labels in  $B_0$  to fall in the urn 0. The second factor is the probability for remaining balls to form a partition  $B_1, \dots, B_k$  of  $[n] \setminus B_0$ .  $\square$

*Remark 4.3.* For  $r = 0$ , the right-hand side of (4.2) vanishes whenever  $\#B_0 \neq 0$ . Thus, neglecting the (empty) red block in  $\lambda_{n,N,0}$  we obtain the usual partition  $\lambda_{n,N}$  of  $[n]$ .

Let  $\text{type}$  be the mapping which sends an incomplete partition  $\pi \in \Pi_n^{\text{inc}}$  to the multiset  $\text{type}(\pi)$  of block sizes of  $\pi$  with a distinguished part  $b_0 \in \mathbb{N}_0$ . Formally, an incomplete partition  $\pi =$

$(B_0, \{B_1, \dots, B_k\})$  has type  $\text{type}(\pi) = (b_0, \{b_1, \dots, b_k\})$ , where  $b_0 \in \mathbb{N}_0$  and  $b_1, \dots, b_k \in \mathbb{N}$  if, and only if,  $\#B_0 = b_0$  and  $\{b_1, \dots, b_k\}$  is the multiset of block sizes of  $B_1, \dots, B_k$ .

From Proposition 4.2 we immediately conclude the following.

**Corollary 4.4.** *Let  $N \in \mathbb{N}$  be an integer and  $r \geq 0$ . Then,*

$$\mathbb{P}[\text{type}(\boldsymbol{\lambda}_{n,N,r}) = (b_0, \{b_1, \dots, b_k\})] = \binom{n}{b_0, \dots, b_k} \frac{1}{k!} \frac{r^{b_0} N^{k\downarrow}}{(N+r)^n} \quad (4.3)$$

for each collection  $(b_0, \{b_1, \dots, b_k\})$  of integers such that  $b_0 \in \mathbb{N}_0$ ,  $b_1, \dots, b_k \in \mathbb{N}$  and  $b_0 + b_1 + \dots + b_k = n$ .

For an incomplete partition  $\pi \in \Pi_n^{\text{inc}}$  let  $\# \text{ blocks}(\pi)$  denote the number of *white* blocks.

**Proposition 4.5.** *Let  $N, n \in \mathbb{N}$  be integers and  $r \geq 0$ . The number of white blocks in  $\boldsymbol{\lambda}_{n,N,r}$  has the  $r$ -StirSibuya $[n; N]$  distribution, that is,*

$$\mathbb{P}[\# \text{ blocks}(\boldsymbol{\lambda}_{n,N,r}) = k] = \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \frac{N^{k\downarrow}}{(N+r)^n}, \quad k \in \{0, 1, \dots, \min(n, N)\}. \quad (4.4)$$

Similarly to Proposition 3.5 we give three simple proofs of this assertion exploiting various representations of  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r$ .

*First proof.* This proof uses the second equality in (3.4). One just needs to sum the right-hand sides of (4.3) over all  $(k+1)$ -tuples  $(b_0, b_1, \dots, b_k)$  such that  $b_0 \in \mathbb{N}_0$ ,  $b_1, \dots, b_k \in \mathbb{N}$  and  $b_0 + b_1 + \dots + b_k = n$ . Formula (4.4) follows then from (3.4).  $\square$

*Second proof.* This proof hinges on an equality in (1.6), namely,  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r = \sum_{j=k}^n \binom{n}{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} r^{n-j}$ . Conditioning on the number of balls that do not fall in the urn 0 and using (4.1), we obtain

$$\begin{aligned} \mathbb{P}[\# \text{ blocks}(\boldsymbol{\lambda}_{n,N,r}) = k] &= \text{StirSibuya}[\text{Bin}[n; N/(N+r)]; N](\{k\}) \\ &= \sum_{j=k}^n \binom{n}{j} \left( \frac{r}{N+r} \right)^{n-j} \left( \frac{N}{N+r} \right)^j \left\{ \begin{matrix} j \\ k \end{matrix} \right\} \frac{N^{k\downarrow}}{N^j} \stackrel{(1.6)}{=} \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \frac{N^{k\downarrow}}{(N+r)^n} \end{aligned}$$

for each  $k \in \{0, 1, \dots, \min(n, N)\}$ .  $\square$

*Third proof.* In this proof we suppose that  $r \in \mathbb{N}$  is integer and exploit the formula  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r = \sum_{j=k}^n \left\{ \begin{matrix} n \\ j \end{matrix} \right\} \binom{j}{k} r^{(j-k)\downarrow}$ , see (1.6). Replace the red urn with frequency  $r/(N+r)$  by  $r$  urns with frequency  $1/(N+r)$ , so that we allocate  $n$  balls in  $N+r$  equiprobable urns ( $N$  white ones and  $r$  red ones). In total, there are  $(N+r)^n$  possible allocations. Then,

$$\begin{aligned} \mathbb{P}[\# \text{ blocks}(\boldsymbol{\lambda}_{n,N,r}) = k] &= \sum_{j=k}^n \mathbb{P}[k \text{ white urns and } j-k \text{ red urns are non-empty}] \\ &= \frac{1}{(N+r)^n} \sum_{j=k}^n \left\{ \begin{matrix} n \\ j \end{matrix} \right\} \binom{j}{k} N^{k\downarrow} r^{(j-k)\downarrow} = \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \frac{N^{k\downarrow}}{(N+r)^n}, \end{aligned}$$

where  $\left\{ \begin{matrix} n \\ j \end{matrix} \right\}$  is the number of ways to decompose  $n$  balls into  $j$  blocks,  $\binom{j}{k}$  is the number of ways to choose  $k$  blocks to be placed into white urns, and, finally,  $N^{k\downarrow}$ , respectively  $r^{(j-k)\downarrow}$ , is the number of ways to choose white, respectively red, urns to be filled.  $\square$

*Remark 4.6.* Proposition 4.5 can also be found in [53, Section 5].

**4.2. Random uniform  $r$ -partitions and random Gibbs  $r$ -partitions.** The cardinality of the set  $\Pi_n$  of all partitions of  $[n]$  is the Bell number  $\mathfrak{B}_n$ . Let  $\boldsymbol{\pi}$  be a random uniform partition of  $[n]$ , that is, a random element with values in  $\Pi_n$  and such that  $\mathbb{P}[\boldsymbol{\pi} = \pi] = 1/\mathfrak{B}_n$  for each partition  $\pi \in \Pi_n$ . Stam [63] discovered a randomization of the number of urns  $N$  in the above scheme which leads to a random partition with the uniform distribution on  $\Pi_n$ . Let  $M$  be a random variable with

$$\mathbb{P}[M = m] = \frac{1}{e \cdot \mathfrak{B}_n} \frac{m^n}{m!}, \quad m \in \mathbb{N}.$$

The numbers on the right-hand side sum up to 1 by the Dobiński formula. Stam showed that the distribution of  $\boldsymbol{\lambda}_{n,M}$  is uniform on  $\Pi_n$ , where  $M$  is assumed independent of everything else. The proof is rather simple. From (4.2) applied with  $r = 0$  we conclude that, for each  $\{B_1, \dots, B_k\} \in \Pi_n$ ,

$$\mathbb{P}[\boldsymbol{\lambda}_{n,M} = \{B_1, \dots, B_k\}] = \frac{1}{e \cdot \mathfrak{B}_n} \sum_{m=k}^{\infty} \frac{m^n}{m!} \frac{m^{k\downarrow}}{m^n} = \frac{1}{e \cdot \mathfrak{B}_n} \sum_{m=k}^{\infty} \frac{1}{(m-k)!} = \frac{1}{\mathfrak{B}_n}.$$

Moreover, Stam showed that the number of empty boxes is Poisson distributed with parameter 1 and independent of the partition.

We now focus on a two-fold generalization of Stam's construction. This leads us to the  $r$ -Stirling distributions of the second kind  $r\text{-Stir2}[n; \theta]$ .

**Definition 4.7.** Fix some  $\theta > 0$  and  $r \geq 0$ . A random variable  $\boldsymbol{\pi}_{\theta,r}$  with values in  $\Pi_n^{\text{inc}}$  is called a *random Gibbs  $r$ -partition* if it has the distribution

$$\mathbb{P}[\boldsymbol{\pi}_{\theta,r} = (B_0, \{B_1, \dots, B_k\})] = \frac{\theta^{k_r \# B_0}}{T_{n,r}(\theta)}, \quad (B_0, \{B_1, \dots, B_k\}) \in \Pi_n^{\text{inc}},$$

where we recall the definition

$$T_{n,r}(\theta) = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \theta^k$$

of the  $r$ -Touchard polynomial. This distribution is denoted by  $r\text{-GibbsPart}[n; \theta]$ . If  $\theta = 1$ , an incomplete partition  $\boldsymbol{\pi}_{1,r}$  is called *uniform  $r$ -partition*. The number  $T_{n,r}(1) =: \mathfrak{B}_{n,r}$  is called the  $n$ -th  $r$ -Bell number, see [50, § 8.3] and [48].

*Remark 4.8.* In the literature, see, for example, [57] and [19], it is more common to use the term ‘Gibbs partition’ in a more general sense when the weight of a block also depends on its size. In Definition 4.7 we use this terminology in a narrow sense, when all white blocks have the same (but arbitrary) weight  $\theta > 0$ .

Note that the above distribution resembles (4.2) with  $N$  replaced by  $\theta$ . The change of the factor  $\theta^{k\downarrow}$  to  $\theta^k$  results in a more sophisticated normalizing factor  $T_{n,r}(\theta)$  in place of  $(\theta + r)^n$  in (4.2).

The correctness of the definition follows from the chain of equalities (there are  $\binom{n}{j}$  choices for a red block  $B_0$  containing  $n - j$  elements and  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$  choices to partition the remaining  $j$  elements into  $k$  white blocks  $B_1, \dots, B_k$ ):

$$\begin{aligned} \sum_{\boldsymbol{\pi} \in \Pi_n^{\text{inc}}} \mathbb{P}[\boldsymbol{\pi}_{\theta,r} = \boldsymbol{\pi}] &= \frac{1}{T_{n,r}(\theta)} \sum_{j=0}^n \sum_{k=0}^j \theta^k r^{n-j} \binom{n}{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} = \frac{1}{T_{n,r}(\theta)} \sum_{k=0}^n \theta^k \sum_{j=k}^n r^{n-j} \binom{n}{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} \\ &\stackrel{(1.6)}{=} \frac{1}{T_{n,r}(\theta)} \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \theta^k = 1. \end{aligned}$$

**Proposition 4.9.** *Let  $\pi_{\theta,r} \sim r$ -GibbsPart[ $n; \theta$ ]. Then*

$$\mathbb{P}[\text{type}(\pi_{\theta,r}) = (b_0, \{b_1, \dots, b_k\})] = \frac{r^{b_0} \theta^k}{T_{n,r}(\theta)} \binom{n}{b_0, \dots, b_k} \frac{1}{k!} = \frac{n!}{k!} \frac{r^{b_0} \theta^k}{b_0! \dots b_k! T_{n,r}(\theta)}. \quad (4.5)$$

Furthermore, the number of white blocks follows the  $r$ -Stirling distribution of the second kind:

$$\mathbb{P}[\#\text{ blocks}(\pi_{\theta,r}) = k] = \frac{\theta^k}{T_{n,r}(\theta)} \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r, \quad k = 0, \dots, n. \quad (4.6)$$

*Proof.* Equality (4.5) is obvious. Formula (4.6) follows from the second equality in (3.4) upon summing the right-hand side of (4.3) over all  $b_0 \in \mathbb{N}_0$  and  $b_1, \dots, b_k \in \mathbb{N}$  satisfying  $b_0 + \dots + b_k = n$ .  $\square$

Next, in the spirit of Stam's construction, we show that  $\pi_{\theta,r}$  can be realized by randomizing  $N$  in the incomplete partition  $\lambda_{n,N,r}$  defined at the beginning of this section. Consider a random variable  $M_{\theta,r}$  with the following distribution:

$$\mathbb{P}[M_{\theta,r} = m] = \frac{1}{e^\theta T_{n,r}(\theta)} \frac{(r+m)^n \theta^m}{m!}, \quad m \in \mathbb{N}_0. \quad (4.7)$$

We first check that this is indeed a probability distribution. The next lemma is an  $r$ -analogue of the Dobiński formula, see [48, Theorem 5.1]. For completeness, we provide a proof.

**Lemma 4.10.** *For each  $\theta > 0$  and  $r \geq 0$ ,*

$$\sum_{m=0}^{\infty} \frac{(r+m)^n \theta^m}{m!} = e^\theta T_{n,r}(\theta).$$

*Proof.* Using the second formula in (1.2) yields

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{(r+m)^n \theta^m}{m!} &= \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r m^{k\downarrow} = \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \sum_{m=0}^{\infty} \frac{\theta^m}{m!} m^{k\downarrow} \\ &= \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \theta^k \sum_{m=k}^{\infty} \frac{\theta^{m-k}}{(m-k)!} = \left( \sum_{k=0}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r \theta^k \right) \left( \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \right) = T_{n,r}(\theta) e^\theta. \end{aligned}$$

$\square$

Here is a version of Stam's result including the claim that the number of empty urns is independent of the incomplete partition and has a Poisson distribution. Naturally,  $M_{\theta,r}$  is taken independent of everything else.

**Proposition 4.11.** *For each  $\theta > 0$  and  $r \geq 0$ ,*

$$(\lambda_{n,M_{\theta,r},r}, M_{\theta,r} - \#\text{ blocks}(\lambda_{n,M_{\theta,r},r})) \sim r\text{-GibbsPart}[n; \theta] \otimes \text{Poi}[\theta].$$

*Proof.* Fix  $v \in \mathbb{N}_0$  and an incomplete partition  $(B_0, \{B_1 \dots, B_k\}) \in \Pi_n^{\text{inc}}$ . Then, by (4.7) and Proposition 4.2,

$$\begin{aligned}
& \mathbb{P}[\boldsymbol{\lambda}_{n, M_{\theta, r}, r} = (B_0, \{B_1 \dots, B_k\}), M_{\theta, r} = k + v] \\
&= \mathbb{P}[M_{\theta, r} = k + v] \cdot \mathbb{P}[\boldsymbol{\lambda}_{n, M_{\theta, r}, r} = (B_0, \{B_1 \dots, B_k\}) | M_{\theta, r} = k + v] \\
&= \mathbb{P}[M_{\theta, r} = k + v] \cdot \mathbb{P}[\boldsymbol{\lambda}_{n, k+v, r} = (B_0, \{B_1 \dots, B_k\})] \\
&= \frac{1}{e^{\theta} T_{n, r}(\theta)} \frac{(r + k + v)^n \theta^{k+v}}{(k + v)!} \cdot \frac{r^{\#B_0} (k + v)^{k \downarrow}}{(k + v + r)^n} \\
&= \frac{\theta^k r^{\#B_0}}{T_{n, r}(\theta)} \cdot e^{-\theta} \frac{\theta^v}{v!} \\
&= r\text{-GibbsPart}[n; \theta]((B_0, \{B_1 \dots, B_k\})) \cdot \text{Poi}[\theta](\{v\}),
\end{aligned}$$

where in the last line we have used Definition 4.7.  $\square$

As we did for  $r$ -permutations, we conclude the discussion of  $r$ -partitions with a remark on Broder's definition of the  $r$ -Stirling numbers of the second kind given in the introduction. If  $r \in \mathbb{N}_0$  is an integer, then the  $r$ -uniform partition can be constructed combinatorially as follows. Consider the set  $[n + r]$  in which elements  $1, 2, \dots, n$  are *white* and  $n + 1, \dots, n + r$  are *red*. Let  $\Pi_{n, r}$  be the set of all partitions of  $[n + r]$  such that red elements are in different blocks and let  $\mathfrak{d}_{\text{part}} : \Pi_{n, r} \mapsto \Pi_n^{\text{inc}}$  be the mapping defined as follows. For  $\pi \in \Pi_{n, r}$ , the image  $\mathfrak{d}_{\text{part}}(\pi)$  is an incomplete partition  $(\widehat{B}_0, \{\widehat{B}_1, \dots, \widehat{B}_k\})$  of  $[n]$  such that the red block  $\widehat{B}_0$  is obtained by merging  $r$  blocks of  $\pi$  containing red elements, removing these red elements from the union, and keeping white blocks  $\widehat{B}_1, \dots, \widehat{B}_k$  unchanged. Let  $\widehat{\pi}_{n, r}$  be a partition of  $[n + r]$  picked uniformly at random from  $\Pi_{n, r}$ .

**Proposition 4.12.** *For each  $n \in \mathbb{N}$  and  $r \in \mathbb{N}_0$ ,*

$$\mathfrak{d}_{\text{part}}(\widehat{\pi}_{n, r}) \sim r\text{-GibbsPart}[n; 1],$$

*that is,  $\mathfrak{d}_{\text{part}}(\widehat{\pi}_{n, r})$  is the uniform  $r$ -partition.*

*Proof.* Each  $\pi \in \Pi_{n, r}$  can be constructed from an incomplete partition  $\pi_1 := (B_0, \{B_1, \dots, B_k\}) \in \Pi_n^{\text{inc}}$  of  $[n]$  by distributing  $b_0 = \#B_0$  elements of the block  $B_0$  among  $r$  blocks of  $\pi$  containing  $r$  red elements (there is an  $r$ -fold choice, for each  $x \in B_0$ ). Thus,

$$\#\Pi_{n, r} = \sum_{j=0}^n \#\{\pi_1 \in \Pi_n^{\text{inc}} : \#B_0 = j\} r^j = \sum_{j=0}^n \binom{n}{j} \mathfrak{B}_{n-j} r^j = \mathfrak{B}_{n, r} = T_{n, r}(1),$$

where the penultimate inequality follows from (1.6). Furthermore, for each incomplete partition  $\pi_1 \in \Pi_n^{\text{inc}}$  with red block  $B_0$  of size  $b_0$ ,

$$\#\{\pi \in \Pi_{n, r} : \mathfrak{d}_{\text{part}}(\pi) = \pi_1\} = r^{b_0}.$$

The proof concludes by taking the ratio.  $\square$

## 5. RANDOM $r$ -COMPOSITIONS

A *composition* of  $n$  into  $k$  summands is a tuple  $(b_1, \dots, b_k)$  with  $b_1, \dots, b_k \in \mathbb{N}$  and  $b_1 + \dots + b_k = n$ . Using stars-and-bars argument, one can check that the number of compositions of  $n$  into  $k$  summands is  $\binom{n-1}{k-1}$ . Picking one of these compositions at random, we obtain a *uniform random composition* of  $n$  into  $k$  summands. In the following we introduce an  $r$ -analogue of this notion.



**Definition 5.1.** An *incomplete composition* of  $n$  is a tuple  $(b_0, b_1, \dots, b_k)$  such that  $b_0 \in \mathbb{N}_0$ ,  $b_1, \dots, b_k \in \mathbb{N}$  and  $b_0 + b_1 + \dots + b_k = n$ . Note that  $b_0$  is allowed to be 0, while  $b_1, \dots, b_k$  are not.

The number of such incomplete compositions is  $\binom{n}{k}$  by stars-and-bars argument. Incomplete compositions are closely related to Weyl chambers of type  $B$ . Consider Weyl chambers of types  $A$  and  $B$  which are polyhedral cones in  $\mathbb{R}^n$  given by  $A^{(n)} := \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \leq \dots \leq x_n\}$  and  $B^{(n)} := \{(x_1, \dots, x_n) \in \mathbb{R}^n : 0 \leq x_1 \leq \dots \leq x_n\}$ . Then, the  $k$ -dimensional faces of the cone  $A^{(n)}$  are in bijective correspondence with compositions of  $n$  into  $k$  summands and have the form

$$\{x_1 = \dots = x_{b_1} \leq x_{b_1+1} = \dots = x_{b_1+b_2} \leq \dots \leq x_{b_1+\dots+b_{k-1}+1} = \dots = x_n\}.$$

Similarly,  $k$ -dimensional faces of  $B^{(n)}$  are in bijective correspondence with the incomplete compositions  $(b_0, b_1, \dots, b_k)$  and have the form

$$\{0 = x_1 = \dots = x_{b_0} \leq x_{b_0+1} = \dots = x_{b_0+b_1} \leq x_{b_0+b_1+1} = \dots \\ = x_{b_0+b_1+b_2} \leq \dots \leq x_{b_0+b_1+\dots+b_{k-1}+1} = \dots = x_n\}.$$

**Definition 5.2.** Fix  $r \geq 0$ . A *random  $r$ -composition* of  $n$  is a random vector  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  such that, for each incomplete composition  $(b_0, b_1, \dots, b_k)$  of  $n$ ,

$$\mathbb{P}[b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k] = \frac{r^{b_0} / b_0!}{\binom{n+r-1}{k+r-1}} = \frac{\binom{b_0+r-1}{b_0}}{\binom{n+r-1}{k+r-1}}. \quad (5.1)$$

**Example 5.3.** For  $r = 0$ , we understand the right-hand side as  $1/\binom{n-1}{k-1}$  for  $b_0 = 0$  and as 0 for  $b_0 \geq 1$ . So,  $b_0^{(n,k)} = 0$  with probability 1 and, discarding the 0-th entry, we recover the uniform distribution on the set of compositions of  $n$ .

**Example 5.4.** For  $r = 1$ , each incomplete composition  $(b_0, b_1, \dots, b_k)$  is equally likely, with the corresponding probability being  $1/\binom{n}{k}$ . Thus, we recover the uniform distribution on the set of all incomplete compositions which are in bijective correspondence with the  $k$ -faces of the Weyl chamber  $B^{(n)}$ . Note in passing that it would be possible to give a similar interpretation for the  $r$ -uniform *partitions* with  $r = 1/2$  by considering uniformly distributed elements of the subspace lattice generated by the type  $B$  reflection arrangement, see [1, § 6.3, 6.7] for a description of the correspondence between (incomplete) partitions and flats in reflection arrangements.

With the help of the next lemma we will show that Definition 5.2 indeed defines a probability distribution.

**Lemma 5.5.** *Let  $r \geq 0$ . Then, for each  $m \in \mathbb{N}$  and  $\ell \in \{0, 1, \dots, m\}$ ,*

$$\sum_{b_0=0}^{m-\ell} \binom{b_0+r-1}{b_0} \binom{m-b_0}{\ell} = \sum_{b_0=0}^{m-\ell} \binom{b_0+r-1}{r-1} \binom{m-b_0}{\ell} = \binom{m+r}{\ell+r}. \quad (5.2)$$

*Proof.* The first equality is trivial and the second one is Eq. (5.26) in [22], provided that  $r \in \mathbb{N}$ . The latter constraint is superfluous (indeed, both sides are polynomials in  $r$ ). For the reader's convenience we give below yet another proof of (5.2) valid for all real  $r \geq 0$ . We use the Taylor expansions

$$\sum_{b_0=0}^{\infty} \binom{b_0+r-1}{b_0} x^{b_0} = (1-x)^{-r} \quad \text{and} \quad \sum_{j=\ell}^{\infty} \binom{j}{\ell} x^j = \frac{x^\ell}{(1-x)^{\ell+1}}.$$

Multiplying these Taylor series and evaluating the coefficient of  $x^m$ , we arrive at the left-hand side of (5.2). On the other hand, the same coefficient can be computed as follows:

$$[x^m] \left( \frac{(1-x)^{-r} x^\ell}{(1-x)^{\ell+1}} \right) = [x^{m-\ell}] (1-x)^{-\ell-r-1} = [x^{m-\ell}] \sum_{j=0}^{\infty} \frac{x^j}{j!} (\ell+r+1)^{j\uparrow} = \frac{(\ell+r+1)^{j\uparrow}}{(m-\ell)!} = \binom{m+r}{\ell+r}.$$

Comparing these results completes the proof.  $\square$

*Remark 5.6.* Now we can show that (5.1) indeed defines a probability distribution on the set of incomplete compositions of  $n$ . The number of incomplete compositions with a given value of  $b_0 \in \{0, \dots, n-k\}$  is  $\binom{n-b_0-1}{k-1}$ . Each such an incomplete composition has probability given by the right-hand side of (5.1). The sum of all such probabilities is

$$\frac{1}{\binom{n+r-1}{k+r-1}} \sum_{b_0=0}^{n-k} \binom{n-b_0-1}{k-1} \binom{b_0+r-1}{b_0} = 1$$

by Lemma 5.5 with  $m = n-1$  and  $\ell = k-1$ .

**5.1. Couplings of random  $r$ -compositions and the Dirichlet multinomial distribution.** In this subsection we aim at constructing an  $r$ -composition via a certain urn scheme, in a way similar to that used for uniform  $r$ -partitions in Section 4. As a result, we provide several couplings of  $r$ -compositions which are consistent either in  $n$  or in  $k$ . It turns out that the random  $r$ -composition can be constructed using an urn scheme with random frequencies.

Formula (5.1) implies that the random vector  $(b_0^{(n,k)}, b_1^{(n,k)} - 1, \dots, b_k^{(n,k)} - 1)$  has the Dirichlet multinomial distribution

$$\text{DirMult}[n-k; r, \underbrace{1, \dots, 1}_{k \text{ times}}], \quad (5.3)$$

see definition (2.7). It is known that the probability measure  $\text{DirMult}[n; \alpha_0, \alpha_1, \dots, \alpha_k]$  can be regarded as a mixture of multinomial distributions. More precisely,

$$\text{DirMult}[n; \alpha_0, \alpha_1, \dots, \alpha_k] = \text{Mult}[n; \text{Dir}[\alpha_0, \alpha_1, \dots, \alpha_k]],$$

where the Dirichlet distribution  $\text{Dir}[\alpha_0, \alpha_1, \dots, \alpha_k]$  is defined by the density

$$(x_0, \dots, x_k) \mapsto \frac{\Gamma(\alpha_0)\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)}{\Gamma(\alpha_0 + \alpha_1 + \cdots + \alpha_k)} x_0^{\alpha_0-1} x_1^{\alpha_1-1} \cdots x_k^{\alpha_k-1}, \quad \sum_{i=0}^k x_i = 1, \quad x_i > 0. \quad (5.4)$$

The above interpretation suggests a coupling of random  $r$ -compositions which is consistent in  $n \geq k$  if  $k$  is fixed. The coupling works as follows. Take a *random* partition  $\mathcal{P}$  of  $[0, 1)$  into  $k+1$  intervals (urns)

$$I_{k,0} := [0, P_0), \quad I_{k,1} := [P_0, P_0 + P_1), \quad \dots \quad I_{k,k} := [P_0 + \cdots + P_{k-1}, 1),$$

where  $(P_0, P_1, \dots, P_k)$  follows  $\text{Dir}[r, 1, \dots, 1]$  distribution. In a role of balls take an independent of  $\mathcal{P}$  sample  $U_1, U_2, \dots, U_n, \dots$  from the uniform distribution on  $[0, 1]$ . Let  $Z_{n,i}$ ,  $i = 0, \dots, k$  be the number of balls among  $U_1, \dots, U_n$  in the urn  $I_{k,i}$ , that is,

$$Z_{n,i} := \#\{1 \leq j \leq n : U_j \in I_{k,i}\}, \quad 0 \leq i \leq k.$$

The following trivially holds true.

**Proposition 5.7.** *Fix  $k \in \mathbb{N}$ . For each  $n \geq k$ , the random vectors  $(Z_{n-k,0}, Z_{n-k,1}, \dots, Z_{n-k,k})$  and  $(b_0^{(n,k)}, b_1^{(n,k)} - 1, \dots, b_k^{(n,k)} - 1)$  have the same Dirichlet multinomial distribution (5.3).*

The aggregation property of the Dirichlet distribution asserts that if

$$(X_0, X_1, \dots, X_k) \sim \text{Dir}[\alpha_0, \alpha_1, \dots, \alpha_k],$$

then

$$(X_0 + X_1, X_2, \dots, X_k) \sim \text{Dir}[\alpha_0 + \alpha_1, \alpha_2, \dots, \alpha_k].$$

This property suggests another coupling of  $r$ -compositions which is consistent in both parameters  $n$  and  $k$  but works only when  $r$  is an integer. This coupling hinges on a simple observation that the collection of gaps between consecutive order statistics from the uniform distribution on  $[0, 1]$  has a symmetric Dirichlet distribution  $\text{Dir}[1, 1, \dots, 1]$ . More precisely, let  $V_1, V_2, \dots$  be a sample from the uniform distribution on  $(0, 1)$ . For each  $M \in \mathbb{N}$ , let

$$0 < V_{M,1} < V_{M,2} < \dots < V_{M,M} < 1$$

be the order statistics of the first  $M$  points of the sample. Then

$$(V_{M,1}, V_{M,2} - V_{M,1}, \dots, V_{M,M} - V_{M,M-1}, 1 - V_{M,M}) \sim \text{Dir}[\underbrace{1, \dots, 1}_{M+1 \text{ times}}].$$

By the aggregation property, if  $r$  is an integer

$$(V_{k+r,r}, V_{k+r,r+1} - V_{k+r,r}, \dots, V_{k+r,k+r} - V_{k+r,k+r-1}, 1 - V_{k+r,k+r}) \sim \text{Dir}[r, \underbrace{1, \dots, 1}_{k+1 \text{ times}}].$$

Thus, a family  $(I_{k,0}, I_{k,1}, \dots, I_{k,k})$ ,  $k \in \mathbb{N}$  of nested partitions of  $[0, 1]$  such that the lengths of  $(I_{k,0}, I_{k,1}, \dots, I_{k,k})$  have the  $\text{Dir}[r, 1, \dots, 1]$  distribution can be sequentially constructed as follows. Start with a sample  $V_1, \dots, V_r$  of size  $r$  from the uniform distribution on  $[0, 1]$ . This sample yields a partition of  $[0, 1]$  into two intervals  $I_{1,0} := [0, V_{r,r})$  and  $I_{1,1} := [V_{r,r}, 1)$ , whose lengths follow the  $\text{Dir}[r, 1]$  distribution. Add a new point  $V_{r+1}$ . If  $V_{r+1}$  lands into  $I_{1,1}$ , then it splits  $I_{1,1}$  into two subintervals  $I_{2,1} := [V_{r,r}, V_{r+1})$  and  $I_{2,2} := [V_{r+1}, 1)$ . In this case, put also  $I_{2,0} := [0, V_{r,r})$ . On the other hand, if  $V_{r+1}$  lands into  $I_{1,0}$ , the splitting is  $I_{2,0} := [0, V_{r+1,r})$  and  $I_{2,1} := [V_{r+1,r}, V_{r,r})$ . In this case, put also  $I_{2,2} := I_{1,1}$ . More generally, if the partition  $(I_{k,0}, I_{k,1}, \dots, I_{k,k})$  has already been constructed, adding a point  $V_{r+k}$  yields a partition  $(I_{k+1,0}, I_{k+1,1}, \dots, I_{k+1,k+1})$  depending on whether  $V_{r+k} \in I_{k,0}$  or  $V_{r+k} \notin I_{k,0}$ . Formally, the construction is given by the equations

$$I_{k+1,0} := [0, V_{r+k,r}), \quad I_{k+1,j} := [V_{r+k,r+j-1}, V_{r+k,r+j}), \quad j = 1, \dots, k, \\ I_{k+1,k+1} = [V_{r+k,r+k}, 1). \quad (5.5)$$

We used the word ‘nested’ to highlight that the partition  $(I_{k+1,0}, I_{k+1,1}, \dots, I_{k+1,k+1})$  is obtained from the partition  $(I_{k,0}, I_{k,1}, \dots, I_{k,k})$  by splitting one of the intervals  $I_{k,j}$  into two subintervals. By taking samples  $(V_i)$  and  $(U_i)$  independent and appealing to Proposition 5.7 we obtain (in the case  $r \in \mathbb{N}$ ) a coupling of  $r$ -compositions with the following properties:

- an  $r$ -partition  $(b_0^{(n,k+1)}, b_1^{(n,k+1)}, \dots, b_{k+1}^{(n,k+1)})$  is derived from an  $r$ -partition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  by splitting one of blocks  $b_j^{(n,k)}$  into two parts;
- an  $r$ -partition  $(b_0^{(n+1,k)}, b_1^{(n+1,k)}, \dots, b_k^{(n+1,k)})$  is derived from  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  by increasing one of blocks by one without changing other blocks.

Proposition 5.7 is a de Finetti-type result, which provides a realization of  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  via a randomized multinomial allocation scheme. Given next is a Markovian construction that gives a realization of the process  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})_{n \geq k}$ , with  $k$  fixed, as consecutive values of a certain Markov chain.

**Proposition 5.8.** Fix  $k \in \mathbb{N}$ . Let  $(X(l))_{l \in \mathbb{N}_0}$  be a Markov chain on  $[0, \infty) \times \mathbb{N}^k$  with the initial state  $X(0) = (r, 1, \dots, 1)$  and transition probabilities

$$\mathbb{P}[X(l+1) = (i_0, i_1, \dots, i_k) + e_j | X(l) = (i_0, i_1, \dots, i_k)] = \frac{i_j}{i_0 + i_1 + \dots + i_k},$$

$$j = 0, \dots, k, \quad l \in \mathbb{N}_0,$$

where  $e_j$  is the  $j$ -th unit vector in  $\mathbb{R}^{k+1}$ . For each fixed  $n \geq k$ ,  $X(n-k) - re_0$  has the same distribution as  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$ .

The Markov chain constructed in Proposition 5.8 is nothing else but a Pólya's urn. Initially, the urn contains  $r$  balls of color 0 and one ball of each of other colors  $1, 2, \dots, k$ . At each step, a ball is drawn uniformly at random from the urn, its color is observed, and the ball is returned to the urn with an additional ball of the same color. The vector  $X(n)$  encodes the colors of balls in the urn after  $n$  steps.

For later needs, we formulate a lemma.

**Lemma 5.9.** Fix  $n, k \in \mathbb{N}$ . The random vector  $(X_0^{(n)}, \dots, X_k^{(n)})$  with the  $\text{DirMult}[n; \alpha_0, \alpha_1, \dots, \alpha_k]$  distribution is negatively associated, that is, for each pair of disjoint subsets  $A_1, A_2$  of  $\{0, 1, \dots, k\}$  and all nondecreasing in each coordinate functions  $f : \mathbb{R}^{|A_1|} \rightarrow \mathbb{R}$  and  $g : \mathbb{R}^{|A_2|} \rightarrow \mathbb{R}$

$$\mathbb{E}[f(X_i^{(n)}, i \in A_1)g(X_j^{(n)}, j \in A_2)] \leq \mathbb{E}[f(X_i^{(n)}, i \in A_1)]\mathbb{E}[g(X_j^{(n)}, j \in A_2)].$$

*Proof.* This claim is given on p. 292 in [33] without proof. We now give a proof. Let  $Y_0, Y_1, \dots, Y_k$  be independent random variables such that, for  $i \in \{0, \dots, k\}$ ,  $Y_i$  has a negative binomial distribution with parameters  $\alpha_i$  and  $\beta \in (0, 1)$ , that is,

$$\mathbb{P}[Y_i = k] = \text{NBin}[\alpha_i; \beta](\{k\}) = \frac{\alpha_i^{k \uparrow}}{k!} \beta^{\alpha_i} (1 - \beta)^k, \quad k \in \mathbb{N}_0.$$

We claim that the distribution of  $(X_0^{(n)}, \dots, X_k^{(n)})$  coincides with the conditional distribution of  $(Y_0, \dots, Y_k)$  given  $\sum_{i=0}^k Y_i = n$ . Indeed, for each  $(b_0, b_1, \dots, b_k) \in \mathbb{N}_0^{k+1}$  summing up to  $n$ ,

$$\begin{aligned} \mathbb{P} \left[ Y_0 = b_0, Y_1 = b_1, \dots, Y_k = b_k \mid \sum_{i=0}^k Y_i = n \right] &= \frac{\mathbb{P}[Y_0 = b_0] \mathbb{P}[Y_1 = b_1] \cdots \mathbb{P}[Y_k = b_k]}{\mathbb{P}[\sum_{i=0}^k Y_i = n]} \\ &= \left( \prod_{i=0}^k \text{NBin}[\alpha_i; \beta](\{b_i\}) \right) / \text{NBin}[\alpha_0 + \dots + \alpha_k; \beta](\{n\}). \end{aligned}$$

After cancellations this becomes the right-hand side of (2.7) with  $d = k + 1$ ,  $k_i = b_i$  and  $\theta_i = \alpha_i$ , for  $i = 0, \dots, k$ . In view of the above representation we may use [33, Theorem 2.6], which states it is sufficient to show that for each subset  $A$  of  $\{0, 1, \dots, k\}$  and all nondecreasing in each coordinate functions  $f : \mathbb{R}^{|A|} \rightarrow \mathbb{R}$  the sequence

$$n \mapsto \mathbb{E} \left[ f(Y_i, i \in A) \mid \sum_{j=0}^k Y_j = n \right] = \mathbb{E}[f(X_i^{(n)}, i \in A)]$$

is nondecreasing. This property is secured by the aforementioned coupling, which enables us to think of  $X_i^{(n)}$  as the occupancy count of the box  $i$  corresponding to  $(P_0, P_1, \dots, P_k)$  having the  $\text{Dir}[\alpha_0, \alpha_1, \dots, \alpha_k]$  distribution. Throwing a new ball does not decrease an occupancy count, that is,  $X_i^{(n)} \leq X_i^{(n+1)}$  a.s.  $\square$

Lemma 5.9 in combination with Proposition 5.7 immediately implies a result to be used later in the proof of Lemma 6.23.

**Corollary 5.10.** *The random  $r$ -composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  is negatively associated.*

## 5.2. Marginal distributions of $r$ -compositions and their limits.

**Proposition 5.11.** *The marginal distributions of the random  $r$ -composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  are given by*

$$\mathbb{P} \left[ b_0^{(n,k)} = b_0 \right] = \frac{\binom{n-b_0-1}{k-1} \binom{b_0+r-1}{b_0}}{\binom{n+r-1}{k+r-1}}, \quad b_0 \in \{0, 1, \dots, n-k\}, \quad (5.6)$$

$$\mathbb{P} \left[ b_j^{(n,k)} = b_j \right] = \frac{\binom{n-b_j+r-1}{k+r-2}}{\binom{n+r-1}{k+r-1}}, \quad b_j \in \{1, \dots, n-k+1\}. \quad (5.7)$$

Moreover, for  $0 < i < j \leq k$ , the bivariate distributions are given by

$$\mathbb{P} \left[ b_i^{(n,k)} = b_i, b_j^{(n,k)} = b_j \right] = \frac{\binom{n+r-b_i-b_j-1}{k+r-3}}{\binom{n+r-1}{k+r-1}}, \quad b_i, b_j \in \{1, \dots, n-k+1\}, \quad b_i + b_j \leq n-k+2. \quad (5.8)$$

*Proof.* For (5.6), just observe that the number of incomplete compositions with fixed  $b_0$  is  $\binom{n-b_0-1}{k-1}$  and the probability of each such an incomplete composition is given by the right-hand side of (5.1). For (5.7), we can take  $j = 1$  without loss of generality. In the case  $k = 1$ , (5.6) secures

$$\mathbb{P} \left[ b_1^{(n,k)} = b_1 \right] = \mathbb{P} \left[ b_0^{(n,k)} = n - b_1 \right] = \frac{\binom{n-b_1+r-1}{n-b_1}}{\binom{n+r-1}{k+r-1}} = \frac{\binom{n-b_1+r-1}{r-1}}{\binom{n+r-1}{k+r-1}}.$$

Assume now that  $k \geq 2$ . Observe that the number of incomplete compositions with fixed  $b_0$  and  $b_1$  is given by  $\binom{n-b_0-b_1-1}{k-2}$ . Since the probability of each such an incomplete composition is given by the right-hand side of (5.1), we obtain

$$\mathbb{P} \left[ b_1^{(n,k)} = b_1 \right] = \sum_{b_0=0}^{n-b_1-k+1} \binom{n-b_0-b_1-1}{k-2} \frac{\binom{b_0+r-1}{b_0}}{\binom{n+r-1}{k+r-1}} = \frac{\binom{n-b_1-1+r}{k-2+r}}{\binom{n+r-1}{k+r-1}},$$

where we evaluated the sum using Lemma 5.5 with  $m := n - b_1 - 1$  and  $\ell := k - 2$ .

Alternatively, one can use a known property of the Dirichlet distribution  $\text{Dir}[\alpha_0, \alpha_1, \dots, \alpha_k]$ , which asserts that its  $i$ -th component has the beta distribution  $\text{Beta}[\alpha_i; \alpha - \alpha_i]$ , where  $\alpha := \sum_{j=0}^k \alpha_j$ . Then, by Proposition 5.7,

$$b_0^{(n,k)} \sim \text{Bin}[n-k; \text{Beta}[r; k]] \quad \text{and} \quad b_j^{(n,k)} \sim 1 + \text{Bin}[n-k; \text{Beta}[1; k+r-1]]. \quad (5.9)$$

Formulas (5.6) and (5.7) follow by simple manipulations with beta-integrals.

For the proof of (5.8), note that by the aggregation property of the Dirichlet distribution in combination with Proposition 5.7,

$$\left( b_0^{(n,k)} + \sum_{\ell \notin \{0, i, j\}} (b_\ell^{(n,k)} - 1), b_i^{(n,k)} - 1, b_j^{(n,k)} - 1 \right) \sim \text{DirMult}[n-k; \text{Dir}[r+k-2, 1, 1]]$$

for each pair of indices  $0 < i < j \leq k$ . Thus,

$$\mathbb{E}[t^{b_i^{(n,k)}-1} s^{b_j^{(n,k)}-1}] = \frac{\Gamma(r+k)}{\Gamma(r+k-2)} \iiint (x+ty+sz)^{n-k} x^{r+k-3} dx dy dz, \quad t, s \in \mathbb{C},$$

where the integration is taken over the simplex  $x + y + z = 1$ ,  $x, y, z \geq 0$ . By expanding  $(x + ty + sz)^{n-k}$  with the aid of the multinomial theorem and equating the coefficients, we conclude that

$$\mathbb{P}\left[b_i^{(n,k)} = b_i, b_j^{(n,k)} = b_j\right] = \frac{\Gamma(r+k)}{\Gamma(r+k-2)} \frac{(n-k)!}{(n-k+2-b_i-b_j)!} \frac{\Gamma(n+r-b_i-b_j)}{\Gamma(n+r)}.$$

This yields (5.8) after elementary manipulations.  $\square$

**Corollary 5.12.** *The conditional distribution of  $(b_1^{(n,k)}, \dots, b_k^{(n,k)})$  given  $b_0^{(n,k)}$  is uniform on the set of all (usual) compositions of  $n - b_0^{(n,k)}$  into  $k$  blocks.*

*Remark 5.13.* For  $r = 0$ ,  $\mathbb{P}[b_j^{(n,k)} = b_j] = \binom{n-b_j-1}{k-2} / \binom{n-1}{k-1}$  for all  $j \in \{1, \dots, k\}$ . Now observe that the formula for the marginal distribution of  $b_j^{(n,k)}$  for general  $r \geq 0$  can be obtained from the  $r = 0$  case if we replace  $n$  and  $k$  by  $n+r$  and  $k+r$ , respectively. This observation has several implications given below.

**Proposition 5.14.** *For  $r \geq 0$ ,*

$$\mathbb{E}[b_j^{(n,k)}] = \frac{n+r}{k+r}, \quad j \in \{1, \dots, k\}, \quad \mathbb{E}[b_0^{(n,k)}] = \frac{r(n-k)}{k+r}. \quad (5.10)$$

*Proof.* For  $r = 0$  it is clear, by exchangeability, that  $\mathbb{E}[b_j^{(n,k)}] = n/k$  for all  $j \in \{1, \dots, k\}$ . According to Remark 5.13, for general  $r \geq 0$ , we have to replace  $n$  and  $k$  by  $n+r$  and  $k+r$ , which gives the first formula in (5.10). The second formula follows from  $b_0^{(n,k)} = n - b_1^{(n,k)} - \dots - b_k^{(n,k)}$ .  $\square$

A strong law of large numbers for the multinomial distribution immediately implies a limit theorem for  $r$ -compositions with  $k$  being fixed.

**Proposition 5.15.** *Let  $k \in \mathbb{N}$  be a fixed integer and  $r \geq 0$ . Then*

$$\frac{1}{n} (b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)}) \xrightarrow[n \rightarrow \infty]{d} \text{Dir}[r, \underbrace{1, \dots, 1}_{k \text{ times}}].$$

**Proposition 5.16.** *Fix some  $r \geq 0$  and let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow \alpha$  for some  $\alpha \in (0, 1]$ . Then, for a random  $r$ -composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$ ,*

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}\left[b_0^{(n,k)} = b_0\right] &= \binom{b_0+r-1}{b_0} (1-\alpha)^{b_0} \alpha^r, \quad b_0 \in \{0, 1, \dots\}, \\ \lim_{n \rightarrow \infty} \mathbb{P}\left[b_j^{(n,k)} = b_1\right] &= \alpha(1-\alpha)^{b_1-1}, \quad b_1 \in \{1, 2, \dots\}, j \in \{1, \dots, k\}. \end{aligned}$$

*That is to say,  $b_0^{(n,k)}$  converges weakly to the negative binomial distribution  $\text{NBin}[r; \alpha]$ , while  $b_j^{(n,k)}$  converges weakly to the geometric distribution  $\text{Geom}[\alpha]$ .*

*Proof.* Follows from Proposition 5.11 after letting  $n \rightarrow \infty$ .  $\square$

**Proposition 5.17.** *Fix some  $r \geq 0$  and let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow 0$ . Then, for a random  $r$ -composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$ ,*

$$\frac{b_0^{(n,k)}}{n/k} \xrightarrow[n, k \rightarrow \infty]{d} \text{Gamma}[r; 1] \quad \text{and} \quad \frac{b_1^{(n,k)}}{n/k} \xrightarrow[n, k \rightarrow \infty]{d} \text{Exp}[1], \quad (5.11)$$

*where  $\text{Gamma}[r; 1]$  is the gamma distribution with parameters  $r$  and 1 and  $\text{Exp}[1]$  is the standard exponential distribution.*

*Proof.* The easiest way to prove this is to use representations (5.9). The first claim in (5.11) follows from the next three facts. First, with  $(\gamma(t))_{t \geq 0}$  being the standard gamma-subordinator,

$$\frac{\gamma(r)}{\gamma(r+k)} \sim \text{Beta}[r; k],$$

see [41, Section 9.1]. Second, for each fixed  $\lambda > 0$ ,

$$\frac{k}{n} \text{Bin}[n-k; \lambda/k] \xrightarrow{\mathbb{P}} \lambda, \quad k, n \rightarrow \infty, \quad k/n \rightarrow 0,$$

as can be checked by using Chebyshev's inequality. Third, by the strong law of large numbers

$$\frac{\gamma(r+k)}{k} \xrightarrow[k \rightarrow \infty]{a.s.} 1.$$

It remains to note that  $\gamma(r) \sim \text{Gamma}[r; 1]$ .

The second claim in (5.11) follows by the same reasoning with the help of the second relation in (5.9). The proof is complete.  $\square$

## 6. THE $(r, s)$ -LAH DISTRIBUTION

**6.1. Motivation.** In this section we investigate the so-called  $(r, s)$ -Lah distribution which is defined in terms of the  $r$ -Stirling numbers of both kinds. Special cases of this distribution appear in [24] in the context of random recursive trees (and the closely related shortest path trees on the complete graph with exponentially distributed weights), and in [36, 38]. In the last two articles, limit theorems for this distribution are applied to obtain asymptotics of the face numbers of random walk convex (and positive) hulls [20, 21, 39, 40]. All these papers investigate the  $(r, s)$ -Lah distribution for certain special values of parameters, for example, the distribution appearing in [24] corresponds to  $(r, s) = (1, 0)$ , the distributions in [36, 39, 40] and [21] correspond to  $(0, 0)$  and  $(1/2, 1/2)$ , respectively, while [38] is concerned with the case  $r = s$ .

Our purpose is to introduce a general distributional family unifying all these special cases. We relate these general  $(r, s)$ -Lah distributions to the  $(r + s)$ -versions of the classic combinatorial probability structures defined in the previous sections. In particular, we derive several stochastic representations of the  $(r, s)$ -Lah distribution in terms of random  $(r + s)$ -compositions and (multinomial) Hoppe trees. In Section 7 we use these representations when proving limit theorems for the  $(r, s)$ -Lah distribution.

**6.2. The  $r$ -Lah numbers.** We start by reviewing some properties of the  $r$ -Lah numbers  $L(n, k)_r$ ; see [54] as well as [3, 14, 59] for more information. For  $n \in \mathbb{N}_0$ ,  $k \in \{0, \dots, n\}$  and  $r \in \mathbb{R}$ , these numbers may be defined [54, Theorem 3.10] by their exponential generating function

$$\sum_{n=k}^{\infty} L(n, k)_r \frac{x^n}{n!} = \frac{1}{k!} \left( \frac{x}{1-x} \right)^k (1-x)^{-2r}. \quad (6.1)$$

We extend this definition by putting  $L(n, k)_r := 0$  for  $n \in \mathbb{N}_0$  and  $k \notin \{0, \dots, n\}$ . The Taylor expansion of the right-hand side of (6.1) gives [54, Theorem 3.7]

$$L(n, k)_r = \binom{n+2r-1}{k+2r-1} \frac{n!}{k!}. \quad (6.2)$$

As a consequence,  $L(n, k)_r$  is a polynomial of  $r$ , for fixed  $n$  and  $k$ . The following formula [54, Theorem 3.2] complements (1.2):

$$(x + r)^{n\uparrow} = \sum_{k=0}^n L(n, k)_r (x - r)^{k\downarrow}. \quad (6.3)$$

Using (1.2) and (6.3) it can be shown [54, Theorem 3.11] that

$$L(n, k)_r = \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_r. \quad (6.4)$$

More generally, from the same reference it is known that

$$L(n, k)_{\frac{r+s}{2}} = \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s. \quad (6.5)$$

The cited paper [54] only considers nonnegative integer  $r$  and requires that  $r$  and  $s$  have the same parity. In fact, since both sides of the identity are polynomials in  $r$  and  $s$ , it holds for arbitrary  $r, s \in \mathbb{R}$ .

From Eq. (5.1) in the definition of  $r$ -compositions we conclude that

$$L(n, k)_r = \frac{n!}{k!} \sum_{\substack{\ell_0 \in \mathbb{N}_0, \ell_1, \dots, \ell_k \in \mathbb{N} \\ \ell_0 + \ell_1 + \dots + \ell_k = n}} \frac{(2r)^{\ell_0\uparrow}}{\ell_0!},$$

which already indicates a connection of the  $r$ -Lah numbers with the random  $2r$ -compositions.

### 6.3. Definition of the $(r, s)$ -Lah distribution.

**Definition 6.1.** An  $(r, s)$ -Lah distribution  $\text{Lah}[n, k]_{r, s}$  with parameters  $n \in \mathbb{N}$ ,  $k \in \{0, \dots, n\}$  and  $r, s \in [0, \infty)$  (where the case  $k = r = s = 0$  is always excluded) is a discrete probability measure defined by

$$\text{Lah}[n, k]_{r, s}(\{j\}) = \frac{1}{L(n, k)_{\frac{r+s}{2}}} \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s, \quad j \in \{k, k+1, \dots, n\}. \quad (6.6)$$

**Definition 6.2.** We call  $(n, k, r, s)$  an *admissible quadruple* of parameters for the  $(r, s)$ -Lah distribution if  $n \in \mathbb{N}$ ,  $k \in \{0, \dots, n\}$ ,  $r, s \in [0, \infty)$  and  $\max\{k, r, s\} > 0$ .

We exclude the case  $k = r = s = 0$  in which  $L(n, 0)_0 = 0$ ,  $n \in \mathbb{N}$ , since  $(-1)! = \infty$ . Equation (6.6) indeed defines a probability distribution due to identity (6.5).

Now we review several particular cases of the  $(r, s)$ -Lah distribution that have already appeared in the literature.

**Example 6.3.** If  $k = 0$  and  $s \geq 0$ , then in view of  $\left\{ \begin{matrix} j \\ 0 \end{matrix} \right\}_s = s^j$ , the distribution takes the form

$$\text{Lah}[n, 0]_{r, s}(\{j\}) = \frac{1}{(r+s)^{n\uparrow}} \begin{bmatrix} n \\ j \end{bmatrix}_r s^j, \quad j \in \{0, 1, \dots, n\}.$$

Thus,  $\text{Lah}[n, 0]_{r, s}$  is the  $r$ -Stirling distribution of the first kind  $r\text{-Stirl}[n; s]$ . In particular, if also  $s = 0$ , then  $\text{Lah}[n, 0]_{r, 0}$  is the degenerate at 0 distribution for each  $r > 0$ .



**Example 6.4.** Let now  $k = 1$  and  $s = 0$ . Note that, for  $j \in \{1, \dots, n\}$ ,

$$\text{Lah}[n, 1]_{r,0}(\{j\}) \stackrel{(6.2)}{=} \frac{1}{n! \binom{n+r-1}{r}} \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ 1 \end{matrix} \right\}_0 \stackrel{(1.1)}{=} \frac{1}{n! \binom{n+r-1}{r}} \begin{bmatrix} n \\ j \end{bmatrix}_r = \frac{n+r}{n} \frac{1}{(r+1)^{n\uparrow}} \begin{bmatrix} n \\ j \end{bmatrix}_r.$$

Thus, if  $Z \sim r\text{-Stir1}[n, 1]$ , then  $\text{Lah}[n, 1]_{r,0}$  is the conditional distribution of  $Z$  given  $\{Z > 0\}$ .

**Example 6.5.** Motivated by [39, 40] the  $(r, s)$ -Lah distribution with  $r = s = 0$  was investigated in [36]. This distribution has a combinatorial interpretation in terms of fragmented permutations. A fragmented permutation of  $n$  elements with  $k$  fragments is a partition of the set  $\{1, \dots, n\}$  into  $k$  nonempty blocks together with the structure of permutation on each block. The number of fragmented permutations with  $k$  blocks is the Lah number  $L(n, k)_0$ . Consider a random fragmented permutation with  $k$  blocks sampled uniformly from the set of all such fragmented permutations. Then, the number of cycles of this fragmented permutation has the  $(r, s)$ -Lah distribution with parameters  $r = s = 0$ .

**Example 6.6.** Consider a complete graph on  $n + 1$  nodes in which the edges have i.i.d. unit exponential weights. The union of shortest paths from some node of this graph to the  $k$  uniformly chosen nodes is a random tree. The number of edges in this tree, denoted by  $H_{n+1}(k)$ , is analyzed in [24]. According to Theorem 2.1 of that paper

$$\mathbb{P}[H_{n+1}(k) = j] = \frac{k!}{n! \binom{n}{k}} \begin{bmatrix} n+1 \\ j+1 \end{bmatrix}_0 \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_0, \quad j \in \{k, k+1, \dots, n\}.$$

Note that  $\begin{bmatrix} n \\ j \end{bmatrix}_1 = \begin{bmatrix} n+1 \\ j+1 \end{bmatrix}_0$ . This means that  $H_{n+1}(k)$  has the same distribution as  $\text{Lah}[n, k]_{1,0}$ . We will return to this example in Section 6.5.

**6.4. Generating polynomial of the  $(r, s)$ -Lah distribution and a recursion for the probability mass function.** The next result identifies the generating polynomial of the  $(r, s)$ -Lah distribution (up to a normalizing factor). It is a generalization of [38, Lemma 3.3 or Lemma 3.6].

**Lemma 6.7.** For any  $n \in \mathbb{N}$ ,  $k \in \{0, \dots, n\}$ ,  $r, s \geq 0$  and  $t \in \mathbb{C}$ ,

$$\sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s t^j = \frac{n!}{k!} [x^n] \left( ((1-x)^{-t} - 1)^k (1-x)^{-(st+r)} \right).$$

*Proof.* We compute the Taylor series using the exponential generating functions (1.3) and (1.4):

$$\begin{aligned} \frac{1}{k!} ((1-x)^{-t} - 1)^k (1-x)^{-(st+r)} &= (1-x)^{-r} \left( \frac{1}{k!} \left( e^{-t \log(1-x)} - 1 \right)^k e^{-st \log(1-x)} \right) \\ &\stackrel{(1.4)}{=} (1-x)^{-r} \sum_{j=k}^{\infty} \left( \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s \frac{(-t \log(1-x))^j}{j!} \right) = \sum_{j=k}^{\infty} \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s t^j \left( \frac{(-\log(1-x))^j}{j!} (1-x)^{-r} \right) \\ &\stackrel{(1.3)}{=} \sum_{j=k}^{\infty} \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s t^j \left( \sum_{n=j}^{\infty} \begin{bmatrix} n \\ j \end{bmatrix}_r \frac{x^n}{n!} \right) = \sum_{j=k}^{\infty} \sum_{n=j}^{\infty} \frac{1}{n!} \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s t^j x^n. \end{aligned}$$

Taking the coefficient of  $x^n$  gives the claimed identity.  $\square$

**Corollary 6.8.** Fix some  $n \in \mathbb{N}$  and  $k \in \{0, \dots, n\}$ . Let  $r, s \in \mathbb{N}_0$  and  $m \in \mathbb{N}_0$  satisfy  $m < (n+r)/(k+s)$ . Then,

$$\sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s (-m)^j = 0.$$

*Proof.* We apply Lemma 6.7 with  $t := -m$ . The degree of the polynomial  $(1-x)^m - 1)^k(1-x)^{sm-r}$  is  $km + sm - r < n$  under our assumption on  $m$ . Hence, the coefficient of  $x^n$  vanishes.  $\square$

**Corollary 6.9.** *Fix some  $n \in \mathbb{N}$ ,  $k \in \{0, \dots, n\}$  and  $r \in \mathbb{R}$ . Let  $m \in \mathbb{N}_0$  satisfy  $m < n/k$ . Then,*

$$\sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_r (-m)^j = 0.$$

*Proof.* If  $r \in \mathbb{N}$  is sufficiently large, then  $m < (n+r)/(k+r)$  and we can apply Corollary 6.8 with  $r = s$ . Since  $\begin{bmatrix} n \\ j \end{bmatrix}_r$  and  $\left\{ \begin{matrix} j \\ k \end{matrix} \right\}_r$  are polynomials in  $r$ , the identity extends to all real  $r$ .  $\square$

Now we derive recurrence relations for the probability mass function and the generating polynomial of  $\text{Lah}[n, k]_{r,s}$ . The triangles of generalized Stirling numbers satisfy the recursions [54, p. 1661]:

$$\begin{bmatrix} n \\ j \end{bmatrix}_r = (n+r-1) \begin{bmatrix} n-1 \\ j \end{bmatrix}_r + \begin{bmatrix} n-1 \\ j-1 \end{bmatrix}_r, \quad \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s = \left\{ \begin{matrix} j-1 \\ k-1 \end{matrix} \right\}_s + (k+s) \left\{ \begin{matrix} j-1 \\ k \end{matrix} \right\}_s.$$

Upon multiplying these identities we conclude that

$$\begin{bmatrix} n \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s = (n+r-1) \begin{bmatrix} n-1 \\ j \end{bmatrix}_r \left\{ \begin{matrix} j \\ k \end{matrix} \right\}_s + \begin{bmatrix} n-1 \\ j-1 \end{bmatrix}_r \left\{ \begin{matrix} j-1 \\ k-1 \end{matrix} \right\}_s + (k+s) \begin{bmatrix} n-1 \\ j-1 \end{bmatrix}_r \left\{ \begin{matrix} j-1 \\ k \end{matrix} \right\}_s. \quad (6.7)$$

After simple manipulations with binomial coefficients we obtain from (6.7) the following trivariate recursion for the probability mass function of the  $(r, s)$ -Lah distribution. Put

$$p_{n,k}(j) = \mathbb{P}[\text{Lah}[n, k]_{r,s} = j].$$

**Proposition 6.10.** *For  $j \in \{k, k+1, \dots, n\}$ , the following trivariate recursive formula holds true*

$$p_{n,k}(j) = \frac{(n+r-1)(n-k)}{n(n+r+s-1)} p_{n-1,k}(j) + \frac{(k+s)(n-k)}{n(n+r+s-1)} p_{n-1,k}(j-1) + \frac{k(k+r+s-1)}{n(n+r+s-1)} p_{n-1,k-1}(j-1). \quad (6.8)$$

Multiplying both sides of (6.8) by  $t^j$  and summing over  $j$  yields a bivariate recursion for the generating functions

$$G_{n,k}(t) := \sum_{j=k}^n p_{n,k}(j) t^j = \mathbb{E}[t^{\text{Lah}[n, k]_{r,s}}],$$

which is,

$$G_{n,k}(t) = G_{n-1,k}(t) \left( \frac{(n+r-1)(n-k)}{n(n+r+s-1)} + \frac{(k+s)(n-k)}{n(n+r+s-1)} t \right) + \frac{k(k+r+s-1)}{n(n+r+s-1)} t G_{n-1,k-1}(t). \quad (6.9)$$

**6.5. Stochastic representation in terms of random recursive trees.** It is known, see [16, Theorem O1] or [62], that in  $\text{RRT}[n]$  both the degree of the root and the distance from the root to node  $n$  are  $\text{Stir}[n; 1]$ -distributed. The next theorem unifies and generalizes both claims.

**Theorem 6.11.** *Consider a random tree  $\text{Hoppe}[n; r+s]$  with root 0 of weight  $r+s$  and nodes  $1, \dots, n$ . Think of the root as being divided into two components with weights  $r$  and  $s$  which means that if some node is attached to the root, it chooses one of these components with probabilities  $r/(r+s)$  and  $s/(r+s)$ . Let  $\mathfrak{A}_{n,k}$  be a random subset of  $\{1, \dots, n\}$  consisting of  $k$  nodes sampled*

uniformly without replacement. Let  $\mathfrak{C}_{n,k}$  be the set of all nodes which are attached directly to the weight  $s$  component of the root. If  $\mathfrak{T}_{n,k}$  denotes the minimal subtree rooted at 0 and containing all nodes from  $\mathfrak{A}_{n,k} \cup \mathfrak{C}_{n,k}$ , then the number of edges in  $\mathfrak{T}_{n,k}$  is distributed according to  $\text{Lah}[n, k]_{r,s}$ .

Before proving the theorem we consider some special cases.

**Corollary 6.12.** *The minimal subtree of a random tree  $\text{Hoppe}[n; r]$  spanned by a random uniform subset of  $k$  nodes (and the root) has  $\text{Lah}[n, k]_{r,0}$  edges.*

*Proof.* Choose  $s = 0$  in Theorem 6.11 (meaning that the set  $\mathfrak{C}_{n,k}$  is empty).  $\square$

In the case  $r = 1$  and  $s = 0$ , Theorem 6.11 recovers Proposition 3.1 in [24]. Note that in [24] the results are stated in terms of the so-called shortest path trees; these can be reduced to the RRT's, as explained in [67, Sections 16.2, 16.3], [66, Sections 1,2] and [68, Section 6.2]; see also [24, 69].

In the case where  $r > 0$  is arbitrary and  $k = 1$ , recalling a characterization of  $\text{Lah}[n, 1]_{r,0}$  given in Example 6.4 we arrive at the following

**Corollary 6.13.** *The distance from the root of a random tree  $\text{Hoppe}[n; r]$  to a node  $V_n$  picked uniformly at random from the set of nodes  $\{1, \dots, n\}$  has the same distribution as  $Z \sim r\text{-Stir1}[n; 1]$  conditioned on  $Z \neq 0$ .*

In the next corollary we provide a simple representation of the *standard* ( $r = s = 0$ ) Lah distribution.

**Corollary 6.14.** *Consider a random recursive tree with root denoted for the time being by 1 and nodes  $2, \dots, n$ . Then, the number of **nodes** in a subtree generated by the root and a uniform random subset of  $k$  nodes from  $\{1, \dots, n\}$  (note that 1 may be in the subset) is distributed according to  $\text{Lah}[n, k]_{0,0}$ .*

*Proof.* We consider a  $\text{Hoppe}[n; r]$ -tree with  $r$  tending to zero and apply Corollary 6.12. Note that node 1 is always attached to 0, while nodes  $2, \dots, n$  are attached to 0 with probability tending to zero as  $r \rightarrow 0$ . Hence, if we remove 0 and the edge between 0 and 1, then with probability tending to one we obtain the random recursive tree described in Corollary 6.14. It remains to apply Corollary 6.12 and observe that the number of nodes in the RRT is the number of edges plus the removed edge between 0 and 1.  $\square$

In the case  $k = 1$ , Corollary 6.14 recovers a known result on the insertion depth in the RRT; see [62, Theorem 2] or [65].

**Example 6.15.** The case  $k = 0$  of Theorem 6.11 is straightforward. In this scenario  $\mathfrak{A}_{n,k}$  is empty and  $\mathfrak{C}_{n,k}$  has the mixed binomial distribution  $\text{Bin}[\text{Stir1}[n; r+s]; s/(r+s)]$ , where  $\text{Stir1}[n; r+s]$  is the degree distribution of the root in  $\text{Hoppe}[n; r+s]$ . By part (ii) of Proposition 3.7  $\text{Bin}[\text{Stir1}[n; r+s]; s/(r+s)] = r\text{-Stir1}[n; s]$  and by Example 6.3  $r\text{-Stir1}[n; s] = \text{Lah}[n, 0]_{r,s}$  in full accordance with the  $k = 0$  case of Theorem 6.11.

*Proof of Theorem 6.11.* Replacing  $n$  by  $n+1$ , we can write recurrence relation (6.9) for  $G_{n,k}(t) = \mathbb{E}[t^{\text{Lah}[n,k]_{r,s}}]$  in the form

$$G_{n+1,k}(t) = \left(1 - \frac{k}{n+1}\right) \left(G_{n,k}(t) \frac{n+r}{n+r+s} + G_{n,k}(t) \frac{st}{n+r+s}\right) \quad (6.10)$$

$$+ t \cdot \frac{k}{n+1} \left(G_{n,k-1}(t) \frac{k-1+r+s}{n+r+s} + G_{n,k}(t) \frac{n-k+1}{n+r+s}\right). \quad (6.11)$$

We show that the generating function of the number of edges in  $\mathfrak{T}_{n,k}$  satisfies the same recurrence relation as  $G_{n,k}$ . Our argument generalizes the one in [24]. Consider the tree Hoppe[ $n + 1; r + s$ ]. With probability  $k/(n + 1)$ , node  $n + 1$  belongs to  $\mathfrak{A}_{n+1,k}$ . Denote this event by  $\Omega$ . Given  $\Omega$ , there are two possibilities:

*Case 1.* If the node  $n + 1$  is attached to a node from the set  $\mathfrak{A}_{n,k-1}$  or to the root with weight  $r + s$ , then the tree  $\mathfrak{T}_{n+1,k}$  differs from  $\mathfrak{T}_{n,k-1}$  by one edge containing the node  $n + 1$ . This additional edge accounts for a factor  $t$  in (6.11). The conditional (given  $\Omega$ ) probability of the event occurring in Case 1 is  $(k - 1 + r + s)/(n + r + s)$ .

*Case 2.* If the node  $n + 1$  is attached to some node  $v$  from the set  $\{1, \dots, n\} \setminus \mathfrak{A}_{n,k}$ , then the tree  $\mathfrak{T}_{n+1,k}$  can be constructed as follows. Construct the minimal tree containing the root and the set  $\mathfrak{A}_{n,k-1} \cup \{v\} \cup \mathfrak{C}_{n,k}$  and add to this tree the edge connecting  $n + 1$  and  $v$ . Then,  $\mathfrak{T}_{n+1,k}$  differs by one edge from the tree having the same law as  $\mathfrak{T}_{n,k}$ , which again gives a factor  $t$ . The conditional (given  $\Omega$ ) probability of the event occurring in Case 2 is  $(n - k + 1)/(n + r + s)$ . Altogether, these considerations explain the term (6.11).

Now, it is possible that the node  $n + 1$  does not belong to  $\mathfrak{A}_{n+1,k}$ , which happens with probability  $\mathbb{P}[\Omega^c] = 1 - k/(n + 1)$ . Given  $\Omega^c$ , there are again two possibilities:

*Case 3.* If the node  $n + 1$  is attached to the weight  $s$  component of the root, then it belongs to  $\mathfrak{C}_{n,k}$  and the tree  $\mathfrak{T}_{n+1,k}$  differs from  $\mathfrak{T}_{n,k}$  by one edge. The conditional probability (given  $\Omega^c$ ) of this event is  $s/(n + r + s)$ .

*Case 4:* If the node  $n + 1$  is attached to any node in  $\{1, \dots, n\}$  or to the weight  $r$  component of the root, then  $\mathfrak{T}_{n+1,k} = \mathfrak{T}_{n,k}$ . The conditional probability (given  $\Omega^c$ ) of this event is  $(n + r)/(n + r + s)$ . Altogether, Cases 3 and 4 explain (6.10).

It remains to note that on the boundary  $k = 0$  the generating function of the number of edges in  $\mathfrak{T}_{n,0}$  also coincides with  $G_{n,0}(t) = \mathbb{E}[t^{\text{Lah}[n,0]_{r,s}}]$  as we have seen in Example 6.15.  $\square$

**6.6. Stochastic representation in terms of random compositions.** The next proposition provides a representation of the  $(r, s)$ -Lah distribution as a sum of conditionally independent random variables. It will be the starting point in our asymptotic analysis of the  $(r, s)$ -Lah distribution.

**Proposition 6.16.** *Let  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  be a random  $(r + s)$ -composition of  $n$ . Further, let  $(Z_{r,s}^{(n,0)}, Z_{r,s}^{(n,1)}, \dots, Z_{r,s}^{(n,k)})$  be a random vector such that, conditionally on the event  $\{b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k\}$ , where  $(b_0, b_1, \dots, b_k)$  is an arbitrary fixed incomplete composition of  $n$ , the random variables*

$$Z_{r,s}^{(n,0)} \sim r\text{-Stir1}[b_0; s], \quad Z_{r,s}^{(n,1)} \sim \text{Stir1}[b_1; 1], \quad \dots, \quad Z_{r,s}^{(n,k)} \sim \text{Stir1}[b_k; 1]$$

are independent, that is,

$$\mathbb{P} \left[ Z_{r,s}^{(n,j)} = \ell \mid b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k \right] = \mathbb{P} \left[ \sum_{m=1}^{b_j} \text{Bern}[1/m] = \ell \right]$$

for all  $j = 1, \dots, k$  and

$$\mathbb{P} \left[ Z_{r,s}^{(n,0)} = \ell \mid b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k \right] = \mathbb{P} \left[ \sum_{m=1}^{b_0} \text{Bern}[s/(r + s + m - 1)] = \ell \right].$$

Here, all Bernoulli random variables are assumed independent. Then,

$$Z_{r,s}^{(n,0)} + Z_{r,s}^{(n,1)} + \cdots + Z_{r,s}^{(n,k)} \sim \text{Lah}[n, k]_{r,s}.$$

**Example 6.17.** Taking  $k = 0$  gives  $\text{Lah}[n, 0]_{r,s} = r\text{-Stir1}[n; s]$ .

**Example 6.18.** For  $r = s = 0$ , Proposition 6.16 reduces to the representation of the Lah distribution given in [36, Proposition 2.3].

**Example 6.19** (Cumulative degree of roots in the multinomial Hoppe tree). Consider a random forest  $\text{MultiHoppe}[n - k; \theta; ((r + s)/\theta, 1/\theta, \dots, 1/\theta)]$  with  $\text{root}_0$  having weight  $r + s$  and  $k$  roots  $\text{root}_1, \dots, \text{root}_k$  having weight 1 each. The total weight of all roots is  $\theta = k + r + s$ . According to Proposition 5.7 and Proposition 2.8, the vector encoding the sizes of connected components (not counting  $\text{root}_0$  but counting  $\text{root}_1, \dots, \text{root}_k$ ) has the same distribution as a random  $(r + s)$ -composition of  $n$ . Let  $D_{n;j}$  be the degree of  $\text{root}_j$ , for  $j = 0, \dots, k$ . Then, Proposition 6.16 implies that

$$D_{n;1} + \cdots + D_{n;k} + \text{Bin}[D_{n;0}; s/(r + s)] \sim \text{Lah}[n, k]_{r,s}.$$

For  $r = 0$  the right-hand side simplifies to  $D_{n;0} + D_{n;1} + \cdots + D_{n;k}$ .

**Example 6.20.** For  $r = s = 1/2$ , the conditional distribution (as in Proposition 6.16) of  $Z_{r,s}^{(n,j)}$  with  $j \in \{1, \dots, k\}$  is the same as  $\sum_{m=1}^{b_j} \text{Bern}[1/m]$ , which is the distribution of the number of records in  $b_j$  i.i.d. observations from a continuous distribution, while the conditional distribution of  $Z_{r,s}^{(n,0)}$  is the same as  $\sum_{m=1}^{b_0} \text{Bern}[1/(2m)]$ . The distribution of the random 1-composition is uniform on the set of all incomplete compositions of  $n$ . Motivated by the analysis of positive hulls of random walks carried out in [21] (see also [20]) the case  $r = s$  has already been investigated in [38].

*Proof of Proposition 6.16.* Recall from Lemma 6.7 and (6.2) that

$$G_{n,k}(t) = \mathbb{E}[t^{\text{Lah}[n,k]_{r,s}}] = \frac{1}{\binom{n+r+s-1}{k+r+s-1}} [x^n] \left( ((1-x)^{-t} - 1)^k (1-x)^{-(st+r)} \right). \quad (6.12)$$

From the definition of the random variables  $Z_{r,s}^{(n,0)}, Z_{r,s}^{(n,1)}, \dots, Z_{r,s}^{(n,k)}$  it follows that

$$\mathbb{E} \left[ t^{Z_{r,s}^{(n,j)}} \mid b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k \right] = \frac{t^{b_j \uparrow}}{b_j!}, \quad j \in \{1, \dots, k\}$$

and

$$\mathbb{E} \left[ t^{Z_{r,s}^{(n,0)}} \mid b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k \right] = \frac{(st+r)^{b_0 \uparrow}}{(r+s)^{b_0 \uparrow}}.$$

By conditional independence, this gives

$$\mathbb{E} \left[ t^{Z_{r,s}^{(n,0)} + Z_{r,s}^{(n,1)} + \cdots + Z_{r,s}^{(n,k)}} \mid b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k \right] = \frac{t^{b_1 \uparrow}}{b_1!} \cdots \frac{t^{b_k \uparrow}}{b_k!} \cdot \frac{(st+r)^{b_0 \uparrow}}{(r+s)^{b_0 \uparrow}}. \quad (6.13)$$

Also, by the definition of the random  $(r + s)$ -composition (Definition 5.2),

$$\mathbb{P}[b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k] = \frac{(r+s)^{b_0 \uparrow} / b_0!}{\binom{n+r+s-1}{k+r+s-1}}. \quad (6.14)$$

Using the formulas

$$(1-x)^{-t} - 1 = \sum_{b=1}^{\infty} \frac{x^b}{b!} t^{b \uparrow} \quad \text{and} \quad (1-x)^{-(st+r)} = \sum_{b_0=0}^{\infty} \frac{x^{b_0}}{b_0!} (st+r)^{b_0 \uparrow},$$

we obtain

$$\begin{aligned} G_{n,k}(t) &= \frac{1}{\binom{n+r+s-1}{k+r+s-1}} \sum_{(b_0, b_1, \dots, b_k)} \frac{t^{b_1 \uparrow}}{b_1!} \cdots \frac{t^{b_k \uparrow} (st+r)^{b_0 \uparrow}}{b_k! b_0!} \\ &= \sum_{(b_0, b_1, \dots, b_k)} \frac{t^{b_1 \uparrow}}{b_1!} \cdots \frac{t^{b_k \uparrow} (st+r)^{b_0 \uparrow}}{b_k! (r+s)^{b_0 \uparrow}} \cdot \frac{(r+s)^{b_0 \uparrow} / b_0!}{\binom{n+r+s-1}{k+r+s-1}}, \end{aligned}$$

where both sums are taken over all incomplete compositions of  $n$ . Recognizing on the right-hand side the conditional expectation of  $t^{Z_{r,s}^{(n,0)} + Z_{r,s}^{(n,1)} + \dots + Z_{r,s}^{(n,k)}}$  given  $\{b_0^{(n,k)} = b_0, b_1^{(n,k)} = b_1, \dots, b_k^{(n,k)} = b_k\}$ , see (6.13), and the probability of this event, see (6.14), we apply the total expectation formula to obtain

$$G_{n,k}(t) = \mathbb{E}[t^{\text{Lah}[n, k]_{r, s}}] = \mathbb{E}[t^{Z_{r,s}^{(n,0)} + Z_{r,s}^{(n,1)} + \dots + Z_{r,s}^{(n,k)}}].$$

The proof is complete.  $\square$

Using Corollary 5.12 and recalling the representation of the standard ( $r = s = 0$ ) Lah distribution given in [36, Proposition 2.3] we obtain the following.

**Corollary 6.21.** *Let  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  be a random  $(r+s)$ -composition and  $(\text{Lah}[m, k]_{0,0})_{m \geq k}$  a sequence of random variables having the standard Lah distributions, which is independent of  $(Z_{r,s}^{(n,0)}, b_0^{(n,k)})$ , where  $Z_{r,s}^{(n,0)}$  is as defined in Proposition 6.16. Then,*

$$\text{Lah}[n, k]_{r, s} \text{ has the same distribution as } Z_{r,s}^{(n,0)} + \text{Lah}[n - b_0^{(n,k)}, k]_{0,0}.$$

Corollary 6.21 is designed to derive limit theorems for  $\text{Lah}[n, k]_{r, s}$  from the known limit theorems for  $\text{Lah}[n, k]_{0,0}$ .

**6.7. Expectation and variance.** We employ the notation

$$L(n, k, r, s) := \mathbb{E}[\text{Lah}[n, k]_{r, s}].$$

The following explicit formula can be found in Theorem 3.1 of [38]

$$L(n, k, 0, 0) = \mathbb{E}[\text{Lah}[n, k]_{0,0}] = \frac{nk}{n - (k-1)} (H_n - H_{k-1}), \quad (6.15)$$

where  $H_n := 1 + 1/2 + \dots + 1/n$  is the  $n$ -th harmonic number. For  $r, s \geq 0$ , put  $H_0^{(r,s)} := 0$  and, for  $n \in \mathbb{N}$ ,  $H_n^{(r,s)} := \sum_{m=1}^n s(r+s+m-1)^{-1}$ . Plainly,  $H_n^{(r,0)} = 0$  for all  $n \in \mathbb{N}_0$ . Corollary 6.21 immediately yields

$$L(n, k, r, s) = \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] + \mathbb{E}[L(n - b_0^{(n,k)}, k, 0, 0)],$$

where  $(b_0^{(n,k)}, \dots, b_k^{(n,k)})$  is a random  $(r+s)$ -composition. However, this formula is of little use for practical purposes. Note that Proposition 6.16 implies that

$$L(n, k, r, s) = \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] + k \mathbb{E}[H_{b_j^{(n,k)}}].$$

We do not derive exact formulas for either  $L(n, k, r, s)$  or the variance, although it can be done relatively straightforwardly by generalizing the proofs given in [38, Theorems 3.1, 3.7]. Instead, our focus is on asymptotic estimates, which are crucial for proving limit theorems for the Lah distributions in subsequent sections.

**Lemma 6.22.** For  $r, s \geq 0$ , the following estimate holds true

$$\left| H_n^{(r,s)} - s \log \left( \frac{r+s+(n-1)_+}{r+s} \right) \right| \leq \frac{s}{r+s}, \quad n \in \mathbb{N}_0. \quad (6.16)$$

*Proof.* For  $n \in \mathbb{N}_0$ ,

$$H_n^{(r,s)} = \sum_{m=1}^n \frac{s}{r+s+m-1} \leq \frac{s}{r+s} + \int_0^{(n-1)_+} \frac{s \, dx}{r+s+x} = \frac{s}{r+s} + s \log \left( \frac{r+s+(n-1)_+}{r+s} \right).$$

This together with an analogous estimate from below completes the proof of (6.16).  $\square$

**Lemma 6.23.** Fix some  $r, s \geq 0$  and let  $(b_0^{(n,k)}, \dots, b_k^{(n,k)})$  be a random  $(r+s)$ -composition. Assume that  $n, k \rightarrow \infty$  such that  $k/n \rightarrow 0$ . Then, for each fixed  $a \geq 1$ ,

$$\mathbb{E} \left[ (H_{b_0^{(n,k)}}^{(r,s)})^a + \sum_{j=1}^k (H_{b_j^{(n,k)}})^a \right] \simeq \mathbb{E} \left[ \sum_{j=1}^k (H_{b_j^{(n,k)}})^a \right] \simeq k (\log(n/k))^a, \quad n, k \rightarrow \infty. \quad (6.17)$$

In particular,

$$L(n, k, r, s) \simeq \mathbb{E} \left[ \sum_{j=1}^k H_{b_j^{(n,k)}} \right] \simeq k \log(n/k), \quad n \rightarrow \infty. \quad (6.18)$$

Furthermore,

$$\text{Var} \left[ \sum_{j=1}^k H_{b_j^{(n,k)}} \right] = O(k), \quad n \rightarrow \infty \quad (6.19)$$

and

$$\text{Var} \left[ H_{b_0^{(n,k)}}^{(r,s)} + \sum_{j=1}^k H_{b_j^{(n,k)}} \right] = O(k), \quad n \rightarrow \infty. \quad (6.20)$$

*Proof.* When dealing with  $H_{b_0^{(n,k)}}^{(r,s)}$  we always assume that  $s > 0$ , for  $H_{b_0^{(n,k)}}^{(r,0)} = 0$  a.s.

For a proof of (6.17) it suffices to check that  $\mathbb{E}[(H_{b_1^{(n,k)}})^a] \simeq (\log(n/k))^a$  and  $\mathbb{E}[(H_{b_0^{(n,k)}}^{(r,s)})^a] = O(\log(n/k))$  as  $n, k \rightarrow \infty$ . In view of (5.11),

$$\frac{\log b_1^{(n,k)}}{\log(n/k)} \xrightarrow{\mathbb{P}} 1, \quad n, k \rightarrow \infty.$$

This in combination with  $|H_n - \log n| \leq C$  for all  $n \in \mathbb{N}$  and a constant  $C > 0$  entails

$$\frac{(H_{b_1^{(n,k)}})^a}{(\log(n/k))^a} \xrightarrow{\mathbb{P}} 1, \quad n, k \rightarrow \infty$$

and thereupon by Fatou's lemma  $\liminf_{n,k \rightarrow \infty} (\mathbb{E}[(H_{b_1^{(n,k)}})^a]/(\log(n/k))^a) \geq 1$ . To prove the converse inequality for the upper limit, note that  $x \mapsto (\log x + C)^a$  is concave on  $[1, \infty)$  for each  $C > a - 1$ . In view of this we take  $C$  sufficiently large and write with the help of Jensen's inequality

$$\begin{aligned} \mathbb{E}[(H_{b_1^{(n,k)}})^a] &\leq \mathbb{E}[(\log b_1^{(n,k)} + C)^a] \leq (\log \mathbb{E}[b_1^{(n,k)}] + C)^a \\ &= (\log \mathbb{E}[(n - b_0^{(n,k)})/k] + C)^a \leq (\log(n/k) + C)^a. \end{aligned}$$

Using (6.16), we conclude that

$$(H_n^{(r,s)})^a \leq 2^{a-1} \left( \left( \frac{s}{r+s} \right)^a + s^a \log^a \left( \frac{r+s+(n-1)_+}{r+s} \right) \right).$$

Further, Jensen's inequality and formula (5.10) with  $r$  replaced by  $r + s$  yields

$$\begin{aligned}
\mathbb{E}[(H_{b_0^{(n,k)}}^{(r,s)})^a] - 2^{a-1} \left(\frac{s}{r+s}\right)^a &\leq 2^{a-1} s^a \mathbb{E} \left[ \left( \log \left( \frac{r+s+(b_0^{(n,k)}-1)_+}{r+s} \right) + C \right)^a \right] \\
&\leq 2^{a-1} s^a \mathbb{E} \left[ \left( \log \left( \frac{r+s+\mathbb{E}[b_0^{(n,k)}]}{r+s} \right) + C \right)^a \right] \\
&= 2^{a-1} s^a \mathbb{E} \left[ \left( \log \left( 1 + \frac{n-k}{k+r+s} \right) + C \right)^a \right]. \tag{6.21}
\end{aligned}$$

The right-hand side is  $O((\log(n/k))^a)$ . The proof of (6.17) is complete.

We now turn to the proof of (6.19) and (6.20). The functions  $x \mapsto H_{[x]}^{(r,s)}$  and  $x \mapsto H_{[x]}$  are nondecreasing on  $[0, \infty)$  and  $[1, \infty)$ , respectively. An application of Corollary 5.10, with  $r$  replaced by  $r + s$ , then yields, for  $j \in \{1, \dots, k\}$ ,

$$\mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)} H_{b_j^{(n,k)}}] \leq \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] \mathbb{E}[H_{b_j^{(n,k)}}]$$

and, for  $i, j \in \{1, \dots, k\}$ ,  $i \neq j$ ,

$$\mathbb{E}[H_{b_i^{(n,k)}} H_{b_j^{(n,k)}}] \leq \mathbb{E}[H_{b_i^{(n,k)}}] \mathbb{E}[H_{b_j^{(n,k)}}].$$

Thus, both (6.19) and (6.20) follow provided we can prove

$$\text{Var} H_{b_0^{(n,k)}}^{(r,s)} = O(1) \quad \text{and} \quad \text{Var} H_{b_1^{(n,k)}} = O(1), \quad n, k \rightarrow \infty. \tag{6.22}$$

To check this, we show that, for each  $p \geq 1$ ,

$$\begin{aligned}
\sup_{n,k \in \mathbb{N}, k \leq n} \mathbb{E} \left| \log(r+s+(b_0^{(n,k)}-1)_+) - \log(n/k) \right|^p &< \infty \quad \text{and} \\
\sup_{n,k \in \mathbb{N}, k \leq n} \mathbb{E} \left| \log b_1^{(n,k)} - \log(n/k) \right|^p &< \infty. \tag{6.23}
\end{aligned}$$

Our argument is a slight adaptation of the proof of formula (5.33) in [36]. We start by proving the second inequality in (6.23). In view of  $|x| = x_+ + x_-$  for  $x \in \mathbb{R}$  and  $(x+y)^p \leq 2^{p-1}(x^p + y^p)$  for  $x, y \geq 0$  it suffices to show that

$$\sup_{n,k \in \mathbb{N}, k \leq n} \mathbb{E}[\log_+((k/n)b_1^{(n,k)})]^p < \infty \quad \text{and} \quad \sup_{n,k \in \mathbb{N}, k \leq n} \mathbb{E}[\log_-((k/n)b_1^{(n,k)})]^p < \infty.$$

The former inequality follows from  $[\log_+ x]^p = O(x)$  as  $x \rightarrow \infty$  and the estimate  $(k/n)\mathbb{E}[b_1^{(n,k)}] = (n - \mathbb{E}[b_0^{(n,k)}])/n \leq 1$ . To prove the latter inequality, we first conclude with the help of (5.7) that

$$\frac{\mathbb{P}[b_1^{(n,k)} = j]}{\mathbb{P}[b_1^{(n,k)} = j+1]} = 1 + \frac{k+r+s-2}{n-j-k+1} > 1$$



for all  $k > (2 - r - s)_+$  and all  $j \in \{1, \dots, n - k\}$ . Thus, the sequence  $j \mapsto \mathbb{P}[b_1^{(n,k)} = j]$  is strictly decreasing. This entails

$$\begin{aligned} \mathbb{E}[\log_-((k/n)b_1^{(n,k)})]^p &= \sum_{j=1}^{\lfloor n/k \rfloor} [\log_-(jk/n)]^p \mathbb{P}[b_1^{(n,k)} = j] \leq \mathbb{P}[b_1^{(n,k)} = 1] \sum_{j=1}^{\lfloor n/k \rfloor} [\log_-(jk/n)]^p \\ &= \frac{k+q-1}{n+q-1} \sum_{j=1}^{\lfloor n/k \rfloor} [\log_-(jk/n)]^p \rightarrow \int_0^1 |\log x|^p dx < \infty, \quad n, k \rightarrow \infty. \end{aligned}$$

The proof of the second inequality in (6.23) is complete.

Turning to the first inequality in (6.23), we conclude that it is enough to show that

$$\begin{aligned} \sup_{n, k \in \mathbb{N}, k \leq n} \mathbb{E}[\log_+((k/n)(r+s+(b_0^{(n,k)}-1)_+))]^p < \infty \quad \text{and} \\ \sup_{n, k \in \mathbb{N}, k \leq n} \mathbb{E}[\log_-((k/n)(r+s+(b_0^{(n,k)}-1)_+))]^p < \infty. \end{aligned}$$

The first of these is secured by

$$\begin{aligned} (k/n)\mathbb{E}[r+s+(b_0^{(n,k)}-1)_+] &\leq r+s+(k/n)\mathbb{E}[b_0^{(n,k)}] \\ &= r+s+(k/n)((r+s)(n-k)/(k+r+s)) \leq 4(r+s). \end{aligned}$$

We have used (5.10) for the equality. As for the second, we write

$$\begin{aligned} \mathbb{E}[\log_-((k/n)(r+s+(b_0^{(n,k)}-1)_+))]^p &= [\log_-((r+s)(k/n))]^p \mathbb{P}[(b_0^{(n,k)}-1)_+ = 0] \\ &\quad + \sum_{j=1}^{\lfloor n/k-(r+s) \rfloor} [\log_-((k/n)(r+s+j))]^p \mathbb{P}[b_0^{(n,k)} = j+1] =: I_1(n, k) + I_2(n, k). \end{aligned}$$

In view of (5.6),  $\mathbb{P}[(b_0^{(n,k)}-1)_+ = 0] \simeq (r+s+1)(k/n)^{r+s}$  as  $n, k \rightarrow \infty$ . This implies that the first summand  $I_1$  vanishes. Now we analyze  $I_2$ . Since

$$\frac{(n-j-1)!(n-j-1-k)!}{(n-j-k)!(n-j-2)!} = \frac{n-j-1}{n-j-k} > 1$$

for all  $k \geq 2$  and all  $j, n \in \mathbb{N}$  satisfying  $n-j-k \geq 1$ , we conclude that the sequence  $j \mapsto \binom{n-j-1}{k-1}$  is decreasing. Thus, for  $j \geq 1$ ,

$$\frac{\binom{n-j-2}{k-1}}{\binom{n+r+s-1}{k+r+s-1}} \leq \frac{\binom{n-3}{k-1}}{\binom{n+r+s-1}{k+r+s-1}} \simeq \left(\frac{k}{n}\right)^{r+s}, \quad n, k \rightarrow \infty.$$

By a standard estimate for the gamma function, there exists a constant  $c_{r,s} > 0$  such that, for all  $j \geq 1$ ,

$$\binom{r+s+j}{1+j} \leq c_{r,s} j^{r+s-1}.$$

Combining fragments together and recalling (5.6) we infer

$$I_2(n, k) = O\left(\frac{k}{n}\right) \sum_{j=1}^{\lfloor n/k-(r+s) \rfloor} [\log_-((k/n)(r+s+j))]^p \left(\frac{kj}{n}\right)^{r+s-1} = O(1), \quad n, k \rightarrow \infty.$$

The last equality is justified by a Riemann approximation

$$\lim_{n,k \rightarrow \infty} \frac{k}{n} \sum_{j=1}^{\lfloor n/k - (r+s) \rfloor} [\log_-((k/n)(r+s+j))]^p \left(\frac{kj}{n}\right)^{r+s-1} = \int_0^1 |\log x|^p x^{r+s-1} dx < \infty.$$

This completes the proof of (6.23).

We only show how to obtain the first relation in (6.22), the argument for the second is analogous. In view of (6.16) and the first inequality in (6.23),

$$|\mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] - s \log(n/k)| = O(1) \quad \text{and} \quad \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)} - s \log(n/k)]^2 = O(1), \quad n, k \rightarrow \infty.$$

The latter entails

$$\mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}]^2 = 2s \log(n/k) \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] - (s \log(n/k))^2 + O(1), \quad n, k \rightarrow \infty,$$

whence

$$\text{Var}[H_{b_0^{(n,k)}}^{(r,s)}] = \mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}]^2 - (\mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}])^2 = -(\mathbb{E}[H_{b_0^{(n,k)}}^{(r,s)}] - s \log(n/k))^2 + O(1) = O(1), \quad n, k \rightarrow \infty.$$

The proof is complete.  $\square$

*Remark 6.24.* A perusal of the proof of (6.17) reveals that, for fixed  $k$ ,

$$\mathbb{E}[H_{b_0^{(n,k)}}] = s \log n + O(1), \quad n \rightarrow \infty,$$

whence

$$L(n, k, r, s) = (k + s) \log n + O(1), \quad n \rightarrow \infty.$$

## 7. LIMIT THEOREMS FOR THE $(r, s)$ -LAH DISTRIBUTION

In this concluding section we prove central limit theorems for the  $(r, s)$ -Lah distribution. Not only do these theorems generalize the special cases  $r = s = 0$  and  $r = s$  investigated in [36] and [38], but they also treat asymptotic regimes not covered in those papers. Let  $N[0; \sigma^2]$  denote a normal random variable with zero mean and variance  $\sigma^2 > 0$ . Throughout this section  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  denotes the random  $(r + s)$ -composition.

**Theorem 7.1** (CLT in the constant  $k$  regime). *Let  $k \in \mathbb{N}_0$  and  $r, s \in [0, \infty)$  be fixed and such that  $\max\{k, s\} > 0$ . Then,*

$$\frac{\text{Lah}[n, k]_{r, s} - (k + s) \log n}{\sqrt{(k + s) \log n}} \xrightarrow[n \rightarrow \infty]{d} N[0; 1].$$

*Proof.* We assume that  $r + s > 0$ . The proof hinges on Corollary 6.21 and a known CLT in the case of constant  $k$  and  $r = s = 0$ , see Theorem 4.2 in [36]. A classic functional limit theorem for independent Bernoulli variables together with Theorem 4.2 in [36] imply

$$\left( \left( \frac{\sum_{m=1}^{\lfloor n^t \rfloor} \text{Bern}[s/(r+s+m-1)] - ts \log n}{\sqrt{s \log n}} \right)_{t \geq 0}, \frac{\text{Lah}[n - b_0^{(n,k)}, k]_{0,0} - k \log(n - b_0^{(n,k)})}{\sqrt{k \log(n - b_0^{(n,k)})}} \right) \xrightarrow[n \rightarrow \infty]{d} ((B(t))_{t \geq 0}, N[0; 1]) \quad (7.1)$$

with the components on both sides being independent and  $(B(t))_{t \geq 0}$  being a standard Brownian motion. According to Proposition 5.15,  $b_0^{(n,k)}/n \xrightarrow[n \rightarrow \infty]{d} \text{Beta}[r+s; k]$ , whence

$$\frac{\log_+ b_0^{(n,k)}}{\log n} \xrightarrow{\mathbb{P}} 1 \quad \text{and} \quad \frac{\log(n - b_0^{(n,k)}) - \log n}{\sqrt{\log n}} \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty. \quad (7.2)$$

Using continuity of the composition operation and Slutsky's lemma we conclude from (7.1) and (7.2) that

$$\left( \frac{Z_{r,s}^{(n,0)} - s \log_+ b_0^{(n,k)}}{\sqrt{\log n}}, \frac{\text{Lah}[n - b_0^{(n,k)}, k]_{0,0} - k \log n}{\sqrt{\log n}} \right) \xrightarrow[n \rightarrow \infty]{d} \left( \sqrt{s}B(1), \sqrt{k}N[0; 1] \right). \quad (7.3)$$

By adding the coordinates and using that, see Proposition 5.15,

$$\frac{\log_+ b_0^{(n,k)} - \log n}{\sqrt{\log n}} \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty,$$

we obtain the desired statement. The proof is complete.  $\square$

Theorem 7.1 was previously known in case  $r = s > 0$  from [38] and in case  $r = s = 0$  from [36]. The next theorem is new even in the standard scenario  $r = s = 0$ .

**Theorem 7.2** (CLT in the intermediate regime). *Fix some  $r, s \geq 0$  and let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow 0$ . Then,*

$$\frac{\text{Lah}[n, k]_{r,s} - L(n, k, r, s)}{\sqrt{k \log(n/k)}} \xrightarrow[n, k \rightarrow \infty]{d} N[0; 1].$$

*Proof of Theorem 7.2.* According to Proposition 6.16, it is sufficient to show that

$$\frac{Z_{r,s}^{(n,0)} + \dots + Z_{r,s}^{(n,k)} - L(n, k, r, s)}{\sqrt{k \log(n/k)}} \xrightarrow[n, k \rightarrow \infty]{d} N[0; 1].$$

For each  $n \in \mathbb{N}$  and  $k \in \mathbb{N}$ , denote by  $\mathcal{G}_{n,k}$  the  $\sigma$ -algebra generated by  $b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)}$ . We first prove that, conditionally on  $\mathcal{G}_{n,k}$ ,

$$\frac{\sum_{j=0}^k (Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}])}{\sqrt{\text{Var}[\sum_{j=0}^k Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]}} \xrightarrow[n, k \rightarrow \infty]{d} N[0; 1]. \quad (7.4)$$

Conditionally on  $\mathcal{G}_{n,k}$ , the left-hand side is a sum of independent centered random variables with finite third absolute moments. In view of this, we will prove the last limit relation by an application of the Berry–Esseen inequality:

$$\begin{aligned} I_{n,k} &:= \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left[ \frac{\sum_{j=0}^k (Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}])}{\sqrt{\text{Var}[\sum_{j=0}^k Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]}} \leq x \right] - \mathbb{P}[N[0; 1] \leq x] \right| \\ &\leq A \frac{\sum_{j=0}^k \mathbb{E}[|Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]|^3 | \mathcal{G}_{n,k}]}{(\text{Var}[\sum_{j=0}^k Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}])^{3/2}} \end{aligned}$$

for a deterministic constant  $A > 0$  which does not depend on  $n$ , nor  $k$ . To complete the proof of (7.4) it is enough to show that

$$\frac{\sum_{j=0}^k \mathbb{E}[|Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]|^3 | \mathcal{G}_{n,k}]}{(\text{Var} [\sum_{j=0}^k Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}])^{3/2}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty. \quad (7.5)$$

As a preparation, we prove that

$$\frac{\text{Var} [\sum_{j=0}^k Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]}{k \log(n/k)} \xrightarrow{\mathbb{P}} 1, \quad n, k \rightarrow \infty. \quad (7.6)$$

In what follows we assume that  $s > 0$ , for otherwise  $Z_{r,s}^{(n,0)} = 0$  a.s. In view of (6.16),

$$\begin{aligned} \text{Var} [Z_{r,s}^{(n,0)} | \mathcal{G}_{n,k}] &= \sum_{m=1}^{b_0^{(n,k)}} \frac{s}{r+s+m-1} \left(1 - \frac{s}{r+s+m-1}\right) \leq H_{b_0^{(n,k)}}^{(r,s)} \\ &\leq s \log \left( \frac{r+s+(b_0^{(n,k)}-1)_+}{r+s} \right) + \frac{s}{r+s}. \end{aligned}$$

As a consequence of (5.11),  $\log(b_0^{(n,k)} - 1)_+ / \log(n/k) \xrightarrow{\mathbb{P}} 1$  as  $n, k \rightarrow \infty$ . This shows that

$$\frac{\text{Var} [Z_{r,s}^{(n,0)} | \mathcal{G}_{n,k}]}{k \log(n/k)} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty.$$

Since, for  $j \in \mathbb{N}$ ,

$$\text{Var} [Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}] = \sum_{m=1}^{b_j^{(n,k)}} \frac{1}{m} \left(1 - \frac{1}{m}\right) \in [H_{b_j^{(n,k)}} - \pi^2/6, H_{b_j^{(n,k)}}],$$

relation (7.6) is equivalent to

$$\frac{\sum_{j=1}^k H_{b_j^{(n,k)}}}{k \log(n/k)} \xrightarrow{\mathbb{P}} 1, \quad n, k \rightarrow \infty.$$

The latter is secured by Lemma 6.23 and Chebyshev's inequality. Thus, for the proof of (7.5) it remains to check

$$\frac{\sum_{j=0}^k \mathbb{E}[|Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]|^3 | \mathcal{G}_{n,k}]}{(k \log(n/k))^{3/2}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty. \quad (7.7)$$

Invoking once again Proposition 6.16 and using Rosenthal's inequality, see Theorem 9.1 on p. 152 in [23], we obtain for an absolute constant  $B > 0$  which does not depend on  $n$  and  $k$

$$\begin{aligned} \mathbb{E}[|Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}]|^3 | \mathcal{G}_{n,k}] &\leq B \left( \left( \sum_{m=1}^{b_j^{(n,k)}} \mathbb{E} [\text{Bern}[m^{-1}] - m^{-1}]^2 \right)^{3/2} \right. \\ &\quad \left. + \sum_{m=1}^{b_j^{(n,k)}} \mathbb{E} [|\text{Bern}[m^{-1}] - m^{-1}|]^3 \right) \leq B \left( (H_{b_j^{(n,k)}})^{3/2} + H_{b_j^{(n,k)}} \right) \quad \text{a.s.} \end{aligned}$$

for  $j \geq 1$  and analogously

$$\mathbb{E}[|Z_{r,s}^{(n,0)} - \mathbb{E}[Z_{r,s}^{(n,0)} | \mathcal{G}_{n,k}]|^3 | \mathcal{G}_{n,k}] \leq B \left( (H_{b_0^{(n,k)}}^{(r,s)})^{3/2} + H_{b_0^{(n,k)}}^{(r,s)} \right) \quad \text{a.s.}$$

Therefore, (7.7) is secured by Eq. (6.17) with  $a = 3/2$  and Markov's inequality.

We have proved that (7.4) holds conditionally on  $\mathcal{G}_{n,k}$ . By Lebesgue's dominated convergence theorem, it also holds unconditionally. Furthermore, in view of (7.6),

$$\frac{\sum_{j=0}^k (Z_{r,s}^{(n,j)} - \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}])}{\sqrt{k \log(n/k)}} \xrightarrow[n,k \rightarrow \infty]{d} \mathbb{N}[0; 1].$$

Observe that

$$L(n, k, r, s) = \mathbb{E} \left[ H_{b_0^{(n,k)}}^{(r,s)} + \sum_{j=1}^k H_{b_j^{(n,k)}} \right], \quad \sum_{j=0}^k \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}] = H_{b_0^{(n,k)}}^{(r,s)} + \sum_{j=1}^k H_{b_j^{(n,k)}}.$$

The proof finishes by an appeal to Eq. (6.20) in Lemma 6.23, which ensures that

$$\frac{\sum_{j=0}^k \mathbb{E}[Z_{r,s}^{(n,j)} | \mathcal{G}_{n,k}] - L(n, k, r, s)}{\sqrt{k \log(n/k)}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty.$$

The proof is complete.  $\square$

**Theorem 7.3** (CLT in the central regime). *Fix some  $r, s \geq 0$  and let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow \alpha$  for some  $\alpha \in (0, 1)$ . Then,*

$$\frac{\text{Lah}[n, k]_{r,s} - L(n, k, 0, 0)}{\sqrt{n}} \xrightarrow[n,k \rightarrow \infty]{d} \mathbb{N} \left[ 0; - \left( \frac{\alpha}{1-\alpha} + \frac{\alpha(\alpha+1) \log \alpha}{(1-\alpha)^2} + \frac{\alpha^2 \log^2 \alpha}{(1-\alpha)^3} \right) \right].$$

*Proof.* Our proof relies on Corollary 6.21 in conjunction with the fact that in the central regime  $b_0^{(n,k)}$  converges in distribution to the negative binomial distribution; see Proposition 5.16.

We first check that

$$\frac{\text{Lah}[n, k]_{0,0} - \text{Lah}[n - b_0^{(n,k)}, k]_{0,0}}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty. \quad (7.8)$$

In view of Proposition 2.4 in [36] (or by using the couplings from Section 5.1) and conditional Markov's inequality it is sufficient to show that

$$\frac{L(n, k, 0, 0) - L(n - b_0^{(n,k)}, k, 0, 0)}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty. \quad (7.9)$$

Indeed, (7.9) implies that (7.8) holds conditionally on  $b_0^{(n,k)}$  and, by the Lebesgue dominated convergence theorem, also unconditionally.

Recall that  $L(n, k, 0, 0)$  is given explicitly in Eq. (6.15). Convergence (7.9) follows immediately from the estimates, with  $a \in \mathbb{N}_0$  fixed,

$$\begin{aligned} 0 \leq L(n, k, 0, 0) - L(n - a, k, 0, 0) &\leq \frac{k}{n - k + 1} (n(H_n - H_{k-1}) - (n - a)(H_{n-a} - H_{k-1})) \\ &\leq \frac{k}{n - k + 1} (2aH_n + n(H_n - H_{n-a})) \\ &\leq \frac{k}{n - k + 1} \left( 2aH_n + n \min \left( H_n, \frac{a}{n - a} \right) \right), \end{aligned}$$

and the fact that  $b_0^{(n,k)}$  converges in distribution. The latter fact also ensures that

$$\frac{Z_{r,s}^{(n,0)}}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty,$$

see Proposition 6.16. Combining this with (7.8) and appealing to Corollary 6.21 shows that the claim of the theorem is equivalent to

$$\frac{\text{Lah}[n, k]_{0,0} - L(n, k, 0, 0)}{\sqrt{n}} \xrightarrow[n, k \rightarrow \infty]{d} \mathbb{N} \left[ 0; - \left( \frac{\alpha}{1-\alpha} + \frac{\alpha(\alpha+1) \log \alpha}{(1-\alpha)^2} + \frac{\alpha^2 \log^2 \alpha}{(1-\alpha)^3} \right) \right].$$

But this is secured by Theorem 5.1 in [36]. The proof is complete.  $\square$

The next result extends Theorem 7.3 to the border case  $\alpha = 1$  in which the limiting variance in Theorem 7.3 vanishes.

**Theorem 7.4.** *Fix some  $r, s \geq 0$  and let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow 1$  and  $n - k \rightarrow \infty$ . Then,*

$$\frac{\text{Lah}[n, k]_{r,s} - L(n, k, r, s)}{\sqrt{n-k}} \xrightarrow[n, k \rightarrow \infty]{d} \mathbb{N} [0; 1/4].$$

Before delving into a proof we briefly explain the idea. In the regime  $k/n \rightarrow 1$ , the number of blocks in the incomplete composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$  is of the same magnitude as  $n$ . As we will see, this means that the variability of the number of blocks of sizes  $> 2$  or of size 1 is negligible in the sense that they only contribute to  $\text{Lah}[n, k]_{r,s} \sim Z_{r,s}^{(n,0)} + Z_{r,s}^{(n,1)} + \dots + Z_{r,s}^{(n,k)}$  through the expectations. The principal contribution to the variance of  $\text{Lah}[n, k]_{r,s}$  comes from the blocks of size 2, and the number of these blocks turns out to be sufficiently close to  $n - k$  in probability. Independently for each block of size 2 the corresponding sum  $Z_{r,s}^{(n,j)}$  takes values 1 or 2 with probability 1/2 resulting in the limiting variance 1/4. We formalize and summarize the above heuristics in a lemma.

**Lemma 7.5.** *For  $m \in \mathbb{N}$ ,  $m \leq n - k + 1$  and a random  $(r + s)$ -composition  $(b_0^{(n,k)}, b_1^{(n,k)}, \dots, b_k^{(n,k)})$ , let  $\rho_m(n, k)$  be the number of its (usual) blocks of size  $m$ , that is,*

$$\rho_m(n, k) = \sum_{j=1}^k \mathbf{1}[b_j^{(n,k)} = m].$$

*Let  $n \rightarrow \infty$ ,  $k \rightarrow \infty$  such that  $k/n \rightarrow 1$  and  $n - k \rightarrow \infty$ . Then*

- (a)  $\text{Var}[\rho_1(n, k)] \simeq n^{-1}(n - k)^2 = o(n - k)$  as  $n, k \rightarrow \infty$ .
- (b)  $\text{Var}[\rho_2(n, k)] \simeq 4n^{-1}(n - k)^2 = o(n - k)$  and  $\rho_2(n, k)/(n - k) \xrightarrow{\mathbb{P}} 1$  as  $n, k \rightarrow \infty$ .
- (c)  $\sum_{m=3}^{n-k+1} H_m \mathbb{E}[\rho_m(n, k)] = O(n^{-1}(n - k)^2) = o(n - k)$  as  $n, k \rightarrow \infty$ .
- (d)  $\mathbb{E} \left[ \sum_{m=3}^{n-k+1} H_m (\rho_m(n, k) - \mathbb{E}[\rho_m(n, k)]) \right]^2 = O(n^{-1}(n - k)^2) = o(n - k)$  as  $n, k \rightarrow \infty$ .

*Proof.* According to (5.7), for  $m \in \mathbb{N}$ ,

$$\mathbb{E}[\rho_m(n, k)] = k \mathbb{P}[b_1^{(n,k)} = m] = k \frac{\binom{n-m+r+s-1}{k+r+s-2}}{\binom{n+r+s-1}{k+r+s-1}}.$$

According to (5.8), for  $1 \leq i < j \leq k$ ,

$$\mathbb{P}[b_i^{(n,k)} = m, b_j^{(n,k)} = m] = \frac{\binom{n+r+s-2m-1}{k+r+s-3}}{\binom{n+r+s-1}{k+r+s-1}}.$$

This yields

$$\begin{aligned}
\text{Var} [\rho_m(n, k)] &= \mathbb{E}[\rho_m(n, k)] + 2 \sum_{1 \leq i < j \leq k} \mathbb{P}[b_i^{(n,k)} = m, b_j^{(n,k)} = m] - (\mathbb{E}[\rho_m(n, k)])^2 \\
&= \mathbb{E}[\rho_m(n, k)] + k(k-1) \mathbb{P}[b_1^{(n,k)} = m, b_2^{(n,k)} = m] - (\mathbb{E}[\rho_m(n, k)])^2 \\
&= k \frac{\binom{n-m+r+s-1}{k+r+s-2}}{\binom{n+r+s-1}{k+r+s-1}} + k(k-1) \frac{\binom{n+r+s-2m-1}{k+r+s-3}}{\binom{n+r+s-1}{k+r+s-1}} - \left( k \frac{\binom{n-m+r+s-1}{k+r+s-2}}{\binom{n+r+s-1}{k+r+s-1}} \right)^2.
\end{aligned}$$

We will make a repeated use of these formulas without further reference and in conjunction with the identities

$$\binom{\alpha}{\beta} = \frac{\alpha}{\beta} \binom{\alpha-1}{\beta-1} \quad \text{and} \quad \binom{\alpha}{\beta} = \frac{\alpha-\beta+1}{\beta} \binom{\alpha}{\beta-1}.$$

(a) Plugging  $m = 1$  yields

$$\begin{aligned}
\text{Var} [\rho_1(n, k)] &= \frac{k(k+r+s-1)}{n+r+s-1} + \frac{k(k-1)(k+r+s-1)(k+r+s-2)}{(n+r+s-1)(n+r+s-2)} - \frac{k^2(k+r+s-1)^2}{(n+r+s-1)^2} \\
&= \frac{k(n-k)(n-k+r+s-1)(k+r+s-1)}{(n+r+s-2)(n+r+s-1)^2} \simeq \frac{(n-k)^2}{n} = o(n-k), \quad n, k \rightarrow \infty.
\end{aligned}$$

(b) Plugging  $m = 2$  yields

$$\begin{aligned}
\text{Var} [\rho_2(n, k)] &= k \frac{(k+r+s-1)(n-k)}{(n+r+s-2)(n+r+s-1)} \\
&\quad + k(k-1) \frac{(k+r+s-1)(k+r+s-2)(n-k-1)(n-k)}{(n+r+s-4)(n+r+s-3)(n+r+s-2)(n+r+s-1)} \\
&\quad - \left( k \frac{(k+r+s-1)(n-k)}{(n+r+s-2)(n+r+s-1)} \right)^2 \\
&\simeq 4n^{-1}(n-k)^2 = o(n-k), \quad n, k \rightarrow \infty.
\end{aligned}$$

Noting that

$$\mathbb{E}[\rho_2(n, k)] = k \frac{(k+r+s-1)(n-k)}{(n+r+s-2)(n+r+s-1)} \sim n-k, \quad n, k \rightarrow \infty,$$

the claim about the convergence in probability follows with the help of Chebyshev's inequality.

(c) For  $m \geq 3$ ,

$$\mathbb{E}[\rho_m(n, k)] = \frac{k(k+r+s-1)}{n+r+s-1} \prod_{i=0}^{m-2} \frac{n-k-i}{n+r+s-2-i} \leq k \left( \frac{n-k}{n+r+s-2} \right)^{m-1}.$$

It remains to note that

$$k \sum_{m=3}^{n-k+1} H_m \left( \frac{n-k}{n+r+s-2} \right)^{m-1} \sim H_3 \frac{(n-k)^2}{n}, \quad n, k \rightarrow \infty,$$

by the Lebesgue dominated convergence theorem.

(d) By the triangle inequality in  $L_2$ , subadditivity of  $x \mapsto x^{1/2}$  on  $[0, \infty)$  and the inequality  $(x+y)^2 \leq 2x^2 + 2y^2$  for  $x, y \geq 0$

$$\begin{aligned} \mathbb{E} \left[ \sum_{m=3}^{n-k+1} H_s(\rho_m(n, k) - \mathbb{E}[\rho_m(n, k)]) \right]^2 &\leq \left( \sum_{m=3}^{n-k+1} H_m(\text{Var}[\rho_m(n, k)])^{1/2} \right)^2 \\ &\leq 2 \left( \sum_{m=3}^{n-k+1} H_m(k(k-1)\mathbb{P}\{b_1^{(n,k)} = m, b_2^{(n,k)} = m\} - (\mathbb{E}[\rho_m(n, k)])^2)^{1/2} \right)^2 \\ &\quad + 2 \left( \sum_{m=3}^{n-k+1} H_m(\mathbb{E}[\rho_m(n, k)])^{1/2} \right)^2. \end{aligned}$$

As in part (c) we infer  $\left( \sum_{m=3}^{n-k+1} H_m(\mathbb{E}[\rho_m(n, k)])^{1/2} \right)^2 = O(n^{-1}(n-k)^2) = o(n-k)$  as  $n, k \rightarrow \infty$ . Further, for  $m \geq 3$

$$\begin{aligned} &k(k-1)\mathbb{P}[b_1^{(n,k)} = m, b_2^{(n,k)} = m] - (\mathbb{E}[\rho_m(n, k)])^2 \\ &= \frac{k(k+r+s-1)}{n+r+s-1} \prod_{j=0}^{m-2} \frac{n-k-j}{n+r+s-2-j} \\ &\quad \times \left( \frac{(k-1)(k+r+s-2)}{n+r+s-2m} \prod_{j=0}^{m-2} \frac{n-k-m+1-j}{n+r+s-m-1-j} - \frac{k(k+r+s-1)}{n+r+s-1} \prod_{j=0}^{m-2} \frac{n-k-j}{n+r+s-2-j} \right) \\ &\leq \frac{k(k+r+s-1)}{n+r+s-1} \left( \frac{n-k}{n+r+s-2} \right)^{m-1} \left( \frac{n-k-m+2}{n+r+s-m} \right)^{m-1} \\ &\quad \times \left( \frac{(k-1)(k+r+s-2)}{n+r+s-2m} - \frac{k(k+r+s-1)}{n+r+s-1} \right) \\ &\leq \frac{k^2(k+r+s-1)^2}{n+r+s-1} \left( \frac{n-k}{n+r+s-2} \right)^{2m-2} \left( \frac{1}{n+r+s-2m} - \frac{1}{n+r+s-1} \right) \\ &= \frac{(2m-1)k^2}{n+r+s-2m} \frac{(k+r+s-1)^2}{(n+r+s-1)^2} \left( \frac{n-k}{n+r+s-2} \right)^{2m-2} \\ &\leq \frac{(2m-1)k^2}{n+r+s-2m} \left( \frac{n-k}{n+r+s-2} \right)^{2m-2}. \end{aligned}$$

Observe that  $n/(n-2m) \rightarrow 1$  as  $n, k \rightarrow \infty$  uniformly in  $m \in \mathbb{N}$ ,  $m \leq n-k+1$ . This in combination with Lebesgue's dominated convergence theorem enables us to conclude that

$$\begin{aligned} &k^2 \left( \sum_{m=3}^{n-k+1} H_m \left( \frac{2m-1}{n+r+s-2m} \right)^{1/2} \left( \frac{n-k}{n+r+s-2} \right)^{m-1} \right)^2 \\ &\quad \sim 5H_3^2 \frac{(n-k)^4}{n^3} = o\left(\frac{(n-k)^2}{n}\right), \quad n, k \rightarrow \infty. \end{aligned}$$

□

*Proof of Theorem 7.4.* We first show that

$$\frac{\sum_{m=1}^{n-k+1} \sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = m] - \mathbb{E} \left[ \sum_{m=1}^{n-k+1} \sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = m] \right]}{2^{-1}\sqrt{n-k}} \xrightarrow[n, k \rightarrow \infty]{d} \text{N}[0; 1]. \quad (7.10)$$



By Lemma 7.5(a) and Chebyshev's inequality,

$$\frac{\sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = 1] - \mathbb{E}\left[\sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = 1]\right]}{\sqrt{n-k}} = \frac{\rho_1(n,k) - \mathbb{E}[\rho_1(n,k)]}{\sqrt{n-k}} \xrightarrow{\mathbb{P}} 0, \quad n, k \rightarrow \infty.$$

The  $m = 2$  term of the sum gives a principal contribution. Indeed, write

$$\begin{aligned} & \frac{\sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = 2] - \mathbb{E}\left[\sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = 2]\right]}{2^{-1}\sqrt{n-k}} \\ &= \frac{\sum_{j=1}^k (2Z_{r,s}^{(n,j)} - 3) \mathbf{1}[b_j^{(n,k)} = 2]}{\sqrt{\rho_2(n,k)}} \sqrt{\frac{\rho_2(n,k)}{n-k}} + 3 \frac{\rho_2(n,k) - \mathbb{E}[\rho_2(n,k)]}{\sqrt{n-k}}. \end{aligned}$$

By Lemma 7.5(b) and Chebyshev's inequality, the second term converges to 0 in probability. Conditionally on  $\mathcal{G}_{n,k}$ ,  $\sum_{j=1}^k (2Z_{r,s}^{(n,j)} - 3) \mathbf{1}[b_j^{(n,k)} = 2]$  has the same distribution as the sum of  $\rho_2(n,k)$  independent random variables taking values  $\pm 1$  with probability  $1/2$ . Hence, by the central limit theorem for random walks the first factor converges in distribution to  $N[0; 1]$  conditionally on  $\mathcal{G}_{n,k}$ , hence also unconditionally. By Lemma 7.5 (b),

$$\frac{\rho_2(n,k)}{n-k} \xrightarrow{\mathbb{P}} 1, \quad n, k \rightarrow \infty.$$

This secures the distributional convergence of the first term and the sum of two terms to  $N[0; 1]$ .

Now we prove that the contributions of the counts  $\rho_m(n,k)$ ,  $3 \leq m \leq n-k+1$  are negligible. Write

$$\begin{aligned} & \sum_{m=3}^{n-k+1} \sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = m] - \mathbb{E}\left[\sum_{m=3}^{n-k+1} \sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = m]\right] \\ &= \left( \sum_{m=3}^{n-k+1} \sum_{j=1}^k Z_{r,s}^{(n,j)} \mathbf{1}[b_j^{(n,k)} = m] - \sum_{m=3}^{n-k+1} H_m \rho_m(n,k) \right) + \left( \sum_{m=3}^{n-k+1} H_m (\rho_m(n,k) - \mathbb{E}[\rho_m(n,k)]) \right) \\ &=: I(n,k) + J(n,k). \end{aligned}$$

We first calculate

$$\mathbb{E}[(I(n,k))^2 | \mathcal{G}_{n,k}] = \sum_{m=3}^{n-k+1} \left( \sum_{\ell=1}^m \ell^{-1} (1 - \ell^{-1}) \right) \rho_m(n,k) \leq \sum_{m=3}^{n-k+1} H_m \rho_m(n,k).$$

Using now Lemma 7.5(c) we obtain  $\mathbb{E}[I(n,k)^2] = o(n-k)$  as  $n, k \rightarrow \infty$ . Finally, by Lemma 7.5(d)  $\mathbb{E}[J(n,k)^2] = o(n-k)$  as  $n, k \rightarrow \infty$ . The proof of (7.10) is complete.

In view of (7.10) and according to Proposition 6.16, it remains to show that  $\mathbb{E}[Z_{r,s}^{(n,0)}] = O(1)$ . But this follows immediately from (6.21).  $\square$

Theorems 7.1, 7.2, 7.3 and 7.4 prove Conjecture 2.4 in [24] in a more general form. Hsien-Kuei Hwang has kindly informed us that the asymptotic normality of the distribution appearing in [24] was proved by him (unpublished notes) using singularity analysis and the saddle point method.

**Acknowledgement.** We are grateful to Hsien-Kuei Hwang for pointing out the article [24]. ZK has been supported by the German Research Foundation under Germany’s Excellence Strategy EXC 2044 - 390685587, Mathematics Münster: Dynamics - Geometry - Structure and by the DFG priority program SPP 2265 Random Geometric Systems. AM was supported by the Alexander von Humboldt Foundation.

#### REFERENCES

- [1] M. Aguiar and S. Mahajan. *Topics in hyperplane arrangements*. American Mathematical Society, Providence, RI, 2017.
- [2] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. European Mathematical Society (EMS), Zürich, 2003.
- [3] H. Belbachir and A. Belkhir. Cross recurrence relations for  $r$ -Lah numbers. *Ars Combinatoria*, 110:199–203, 2013.
- [4] A. Belkhir. The multivariate Lah and Stirling numbers. *J. Integer Seq.*, 23(4):Art. 20.4.5, 13, 2020.
- [5] A. Broder. The  $r$ -Stirling numbers. *Discrete Mathematics*, 49(3):241–259, 1984.
- [6] L. Carlitz. Weighted Stirling numbers of the first and second kind. I. *Fibonacci Quart.*, 18(2):147–162, 1980.
- [7] L. Carlitz. Weighted Stirling numbers of the first and second kind. II. *Fibonacci Quart.*, 18(3):242–257, 1980.
- [8] L. Carlitz and R. Scoville. Generalized Eulerian numbers: combinatorial applications. *J. Reine Angew. Math.*, 265:110–137, 1974.
- [9] Ch. A. Charalambides. On a generalized Eulerian distribution. *Ann. Inst. Statist. Math.*, 43(1):197–206, 1991.
- [10] Ch. A. Charalambides. *Enumerative combinatorics*. Chapman & Hall/CRC, 2002.
- [11] Ch. A. Charalambides. *Combinatorial methods in discrete distributions*. Wiley Ser. Probab. Stat. John Wiley & Sons, 2005.
- [12] Ch. A. Charalambides and M. Koutras. On the differences of the generalized factorials at an arbitrary point and their combinatorial applications. *Discrete Math.*, 47:183–201, 1983.
- [13] Ch. A. Charalambides and J. Singh. A review of the Stirling numbers, their generalizations and statistical applications. *Comm. Statist. Theory Methods*, 17(8):2533–2595, 1988.
- [14] G.-S. Cheon and J.-H. Jung.  $r$ -Whitney numbers of Dowling lattices. *Discrete Mathematics*, 312:2337–2348, 2012.
- [15] H. Crane. The ubiquitous Ewens sampling formula. *Statist. Sci.*, 31(1):1–19, 2016.
- [16] L. Devroye. Applications of the theory of records in the study of random trees. *Acta Inform.*, 26(1-2):123–130, 1988.
- [17] M. Dondajewski and J. Szymański. On the distribution of vertex-degrees in a strata of a random recursive tree. *Bull. Acad. Polon. Sci. Sér. Sci. Math.*, 30(5–6):205–209, 1982.
- [18] M. Drmota. *Random trees: An interplay between combinatorics and probability*. Springer-Verlag, Vienna, 2009.
- [19] A. Gnedin and J. Pitman. Exchangeable Gibbs partitions and Stirling triangles. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 325:83–102, 244–245, 2005.
- [20] T. Godland and Z. Kabluchko. Angle sums of Schläfli orthoschemes. *Discrete and Computational Geometry*, 68:125–164, 2022.

- [21] T. Godland and Z. Kabluchko. Positive hulls of random walks and bridges. *Stochastic Processes and their Applications*, 147:327–362, 2022.
- [22] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company, second edition, 1994.
- [23] A. Gut. *Probability: a graduate course*. Springer, New York, 2005.
- [24] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Size and weight of shortest path trees with exponential link weights. *Combin. Probab. Comput.*, 15(6):903–926, 2006.
- [25] F. Hoppe. Pólya-like urns and the Ewens’ sampling formula. *J. Math. Biol.*, 20(1):91–94, 1984.
- [26] F. Hoppe. Size-biased filtering of Poisson-Dirichlet samples with an application to partition structures in genetics. *J. Appl. Probab.*, 23(4):1008–1012, 1986.
- [27] F. Hoppe. The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.*, 25(2):123–159, 1987.
- [28] T. Huillet. Occupancy problems related to the generalized Stirling numbers. *J. Stat. Phys.*, 191(1):Paper No. 5, 2024.
- [29] T. Huillet and M. Möhle. On Bernoulli trials with unequal harmonic success probabilities. *Metrika*, 2023. to appear.
- [30] K. G. Janardan. Relationship between Morisita’s model for estimating the environmental density and the generalized Eulerian numbers. *Ann. Inst. Statist. Math.*, 40(3):439–450, 1988.
- [31] K. G. Janardan. Some properties of the generalized Eulerian distribution. *J. Statist. Plann. Inference*, 34(2):159–169, 1993.
- [32] S. Janson. Euler-Frobenius numbers and rounding. *Online J. Anal. Comb.*, 8:34, 2013.
- [33] K. Joag-Dev and F. Proschan. Negative association of random variables, with applications. *Ann. Statist.*, 11(1):286–295, 1983.
- [34] N. L. Johnson and S. Kotz. *Urn models and their application: an approach to modern discrete probability theory*. John Wiley & Sons, Inc., New York, 1977.
- [35] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete multivariate distributions*. John Wiley & Sons, Inc., New York, 1997.
- [36] Z. Kabluchko and A. Marynych. Lah distribution: Stirling numbers, records on compositions, and convex hulls of high-dimensional random walks. *Probability Theory and Related Fields*, 184:969–1028, 2022.
- [37] Z. Kabluchko, A. Marynych, and H. Pitters. Mod- $\varphi$  convergence of Stirling distributions and limit theorems for zeros of their generating functions. *J. Math. Anal. Appl.*, 529(1):Paper No. 127571, 27, 2024.
- [38] Z. Kabluchko and D. A. Steigenberger.  $r$ -Lah distribution: properties, limit theorems and an application to compressed sensing. *Adv. in Appl. Math.*, 150:Paper No. 102575, 23, 2023.
- [39] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. Convex hulls of random walks: expected number of faces and face probabilities. *Adv. Math.*, 320:595–629, 2017.
- [40] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. Convex hulls of random walks, hyperplane arrangements, and Weyl chambers. *Geometric and Functional Analysis*, 27(4):880–918, 2017.
- [41] J. F. C. Kingman. *Poisson processes*. The Clarendon Press, Oxford University Press, New York, 1993.
- [42] M. Koutras. Noncentral Stirling numbers and some applications. *Discrete Math.*, 42(1):73–89, 1982.

- [43] K. Leckey and R. Neininger. Asymptotic analysis of Hoppe trees. *J. Appl. Probab.*, 50(1):228–238, 2013.
- [44] H. M. Mahmoud and R. T. Smythe. On the distribution of leaves in rooted subtrees of recursive trees. *Ann. Appl. Probab.*, 1(3):406–418, 1991.
- [45] R. S. Maier. Triangular recurrences, generalized Eulerian numbers, and related number triangles. *Adv. in Appl. Math.*, 146:Paper No. 102485, 62, 2023.
- [46] T. Mansour. *Combinatorics of set partitions*. CRC Press, Boca Raton, FL, 2013.
- [47] T. Mansour and M. Schork. *Commutation relations, normal ordering, and Stirling numbers*. CRC Press, Boca Raton, FL, 2016.
- [48] I. Mező. The  $r$ -Bell numbers. *J. Integer Seq.*, 14(1):Article 11.1.1, 14, 2011.
- [49] I. Mező. Recent developments in the theory of Stirling numbers. *RIMS Kôkyûroku*, 2013:68–80, 2013.
- [50] I. Mező. *Combinatorics and number theory of counting sequences*. CRC Press, 2020.
- [51] M. Morisita. Measuring of habitat value by environmental density method. In G. P. et al. Patil, editor, *Statistical Ecology*, pages 379–401. The Pennsylvania State University Press, 1971.
- [52] D. Najock and C. C. Heyde. On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *J. Appl. Probab.*, 19(3):675–680, 1982.
- [53] K. Nishimura and M. Sibuya. Extended Stirling family of discrete probability distributions. *Comm. Statist. Theory Methods*, 26(7):1727–1744, 1997.
- [54] G. Nyul and G. Rácz. The  $r$ -Lah numbers. *Discrete Mathematics*, 338(10):1660–1666, 2015.
- [55] T. Petersen. *Eulerian numbers*. Birkhäuser/Springer, New York, 2015.
- [56] R. Pinsky. A view from the bridge spanning combinatorics and probability. *Enumer. Comb. Appl.*, 1(3):Paper No. S2S3, 31, 2021.
- [57] J. Pitman. *Combinatorial stochastic processes*. Springer-Verlag, Berlin, 2006.
- [58] R. Shanmugam. On central versus factorial moments. *South African Statist. J.*, 18(2):97–110, 1984.
- [59] M. Shattuck. Generalized  $r$ -Lah numbers. *Proc. Indian Acad. Sci. Math. Sci.*, 126(4):461–478, 2016.
- [60] M. Sibuya. *Stirling family of distributions*. John Wiley & Sons, 2014.
- [61] M. Sibuya and K. Nishimura. Prediction of record-breakings. *Statist. Sinica*, 7(4):893–906, 1997.
- [62] R. T. Smythe and H. M. Mahmoud. A survey of recursive trees. *Teor. ĽmovĽr. Mat. Stat.*, 51:1–29, 1994.
- [63] A. J. Stam. Generation of a random partition of a finite set by an urn model. *J. Combin. Theory Ser. A*, 35(2):231–240, 1983.
- [64] R. Stanley. *Enumerative combinatorics. Vol. 1*. Cambridge University Press, 1997.
- [65] J. Szymański. On the maximum degree and the height of a random recursive tree. In *Random graphs '87 (Poznań, 1987)*, pages 313–324. Wiley, Chichester, 1990.
- [66] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. First-passage percolation on the random graph. *Probab. Engrg. Inform. Sci.*, 15(2):225–237, 2001.
- [67] P. Van Mieghem. *Performance analysis of communications networks and systems*. Cambridge University Press, 2006.
- [68] P. Van Mieghem, G. Hooghiemstra, and R. van der Hofstad. A scaling law for the hopcount in internet. *Delft University of Technology, report*, 2000125, 2000.

- [69] P. Van Mieghem, G. Hooghiemstra, and R. van der Hofstad. Stochastic model for the number of traversed routers in internet. *Proceedings of Passive and Active Measurement (PAM), RIPE NCC (Amsterdam, The Netherlands, 2001)*, 2001.