LAH DISTRIBUTION: STIRLING NUMBERS, RECORDS ON COMPOSITIONS, AND CONVEX HULLS OF HIGH-DIMENSIONAL RANDOM WALKS

ZAKHAR KABLUCHKO AND ALEXANDER MARYNYCH

ABSTRACT. Let $\xi_1, \xi_2, \dots$ be a sequence of independent copies of a random vector in $\mathbb{R}^d$ having an absolutely continuous distribution. Consider a random walk $S_i := \xi_1 + \cdots + \xi_i$, and let $C_{n,d} := \operatorname{conv}(0, S_1, S_2, \dots, S_n)$ be the convex hull of the first $n+1$ points it has visited. The polytope $C_{n,d}$ is called k-neighborly if for every indices $0 \le i_0 < \cdots < i_k \le n$ the convex hull of the $k+1$ points $S_{i_0}, \dots, S_{i_k}$ is a k-dimensional face of $C_{n,d}$. We study the probability that $C_{n,d}$ is k-neighborly in various high-dimensional asymptotic regimes, i.e. when n, d, and possibly also k diverge to ∞. There is an explicit formula for the expected number of k-dimensional faces of $C_{n,d}$ which involves Stirling numbers of both kinds. Motivated by this formula, we introduce a distribution, called the Lah distribution, and study its properties. In particular, we provide a combinatorial interpretation of the Lah distribution in terms of random compositions and records, and explicitly compute its factorial moments. Limit theorems which we prove for the Lah distribution imply neighborliness properties of $C_{n,d}$. This yields a new class of random polytopes exhibiting phase transitions parallel to those discovered by Vershik and Sporyshev, Donoho and Tanner for random projections of regular simplices and crosspolytopes.

CONTENTS

2010 *Mathematics Subject Classification.* Primary: 11B73, 60C05; Secondary: 60D05, 52A22, 52A23, 60F05, 60F10, 30C15, 26C10, 05A16, 05A18.

Key words and phrases. Stirling numbers, Lah numbers, Lah distribution, records, random compositions, random walks, random polytopes, convex hulls, neighborliness, f-vectors, mod-Poisson convergence, central limit theorem, large deviations, Lambert W-function, threshold phenomena, intrinsic conic volumes, Weyl chambers.

1. Introduction and summary of main results

1.1. **Introduction.** The aim of the present paper is to introduce and study a family of discrete probability distributions defined in terms of Stirling numbers of both kinds and Lah numbers. Recall, see for example [30, Section 6.1], that the Stirling numbers of the first kind, denoted by $\left[{n \atop k}\right]$, count the number of permutations of n objects with exactly k disjoint cycles, while the Stirling numbers of the second kind, denoted by $\left\{{n \atop k}\right\}$, count the number of ways to partition a set of n elements into k nonempty subsets. Alternatively, Stirling numbers can be defined by the exponential generating functions via the identities

$$\frac{1}{k!}\left(\log\frac{1}{1-x}\right)^k = \sum_{n=k}^{\infty}\left[{n \atop k}\right]\frac{x^n}{n!} \qquad \text{and} \qquad \frac{1}{k!}(\mathrm{e}^x-1)^k = \sum_{n=k}^{\infty}\left\{{n \atop k}\right\}\frac{x^n}{n!}, \qquad k=0,1,2,\dots. \tag{1.1}$$

The *Lah number* $L(n,k)$ can be defined as the number of ways to partition the set $\{1,\dots,n\}$ into k non-empty subsets and to linearly order the elements inside each subset. It is known that $L(n,k)$ is given by

$$L(n,k) = \sum_{j=k}^{n}\left[{n \atop j}\right]\left\{{j \atop k}\right\} = \frac{(n-1)!}{(k-1)!}\binom{n}{k} = \frac{n!}{k!}\binom{n-1}{k-1}, \quad n\in\mathbb{N}, \quad k\in\{1,\dots,n\}. \tag{1.2}$$

These numbers were introduced by Ivo Lah [50] whose name they now bear; see entry A105278 in [63] and [11] for their properties. We can now define the family of distributions we are interested in.

Definition 1.1. A random variable $X = \mathrm{Lah}(n,k)$ has a *Lah distribution* with parameters $n\in\mathbb{N}$ and $k\in\{1,\dots,n\}$ if

$$\mathbb{P}[X=j] = \frac{1}{L(n,k)}\left[{n \atop j}\right]\left\{{j \atop k}\right\}, \qquad j\in\{k,k+1,\dots,n\}. \tag{1.3}$$

Throughout the paper, we agree that $\mathrm{Lah}(n,k)$ denotes some random variable with distribution (1.3).

The special case of the Lah distribution with $k=1$ is well known to be the distribution of the number of cycles in a random uniform permutation of $\{1,\dots,n\}$, or the number of records in a sample of n independent identically distributed (i.i.d.) observations with a continuous distribution function. We shall extend the latter interpretation to arbitrary k, but the original motivation for introducing the Lah distribution comes from the study of threshold phenomena for high-dimensional random polytopes initiated in the pioneering work of Vershik and Sporyshev [68] and continued in a series of works of Donoho and Tanner [15, 16, 18, 19]. As has been suggested by Vershik in his Grassmannian approach to linear programming [66], these threshold phenomena have multiple implications in high-dimensional statistics, signal processing, linear optimization and other fields; for numerous examples we refer the above cited papers as well as [3, 4, 17, 20, 21, 66, 67]. Let us briefly recall the problem studied by Vershik, Sporyshev, Donoho and Tanner. Consider n i.i.d. standard Gaussian points $X_1,\dots,X_n$ in the d-dimensional space, where $n\geq d+1$. Their convex hull $G_{n,d} := \mathrm{conv}(X_1,\dots,X_n)$ is called the Gaussian polytope. Let $f_k(G_{n,d})$ be the number of k-dimensional faces of $G_{n,d}$, for $k\in\{0,\dots,d\}$. Clearly, $f_k(G_{n,d})$ is bounded from above by the number of $(k+1)$-tuples of the points $X_1,\dots,X_n$, that is, by $\binom{n}{k+1}$. If this bound is attained, the polytope $G_{n,d}$ is called *k-neighborly*. Vershik and Sporyshev [68] studied the so-called proportional growth regime in which $d,n,k\to\infty$ in such a way that $k/d\to\rho$ and

$d/n \to \delta$ for some constants $\rho \in (0,1)$ and $\delta \in (0,1)$. They proved the existence of what has been later called a weak threshold, that is, a positive function $\delta \mapsto \rho^{\mathrm{GP}}_{\mathrm{weak}}(\delta)$ such that

$$\lim_{n,d,k\to\infty} \frac{\mathbb{E} f_k(G_{n,d})}{\binom{n}{k+1}} = 1 \qquad \text{provided } \rho < \rho^{\mathrm{GP}}_{\mathrm{weak}}(\delta).$$

Later, Donoho and Tanner [16] proved the existence of what they called a strong threshold, that is, a positive function $\delta \mapsto \rho^{\mathrm{GP}}_{\mathrm{strong}}(\delta)$ such that

$$\lim_{n,d,k\to\infty} \left(\binom{n}{k+1} - \mathbb{E} f_k(G_{n,d}) \right) = 0 \qquad \text{provided } \rho < \rho^{\mathrm{GP}}_{\mathrm{strong}}(\delta).$$

From this relation, they deduced that

$$\lim_{n,d,k\to\infty} \mathbb{P}\left[f_k(G_{n,d}) = \binom{n}{k+1} \right] = 1 \qquad \text{provided } \rho < \rho^{\mathrm{GP}}_{\mathrm{strong}}(\delta).$$

The same conclusions hold for the projection of the regular simplex with n vertices on a random uniform d-dimensional subspace, which has the same expected f-vector as $G_{n,d}$ by a result of Baryshnikov and Vitale [6]. Analogous theory exists, see [15], for random projections of the regular crosspolytope or, equivalently, the symmetric Gaussian polytope defined as $\mathrm{conv}(\pm X_1, \ldots, \pm X_n)$. Going beyond the proportional growth setting, Donoho and Tanner [18] studied also the case when $\delta = 0$. Recently, there has been also interest in the threshold phenomena for random cones as the dimension goes to ∞; see [19, 34, 35, 28].

1.2. **Convex hulls of random walks.** In the present paper we shall be interested in neighborliness properties of a class of random polytopes defined as follows. Let $\xi_1, \xi_2, \ldots$ be i.i.d. random variables with an absolutely continuous distribution on $\mathbb{R}^d$. These assumptions can be weakened; see Section 6.1 below. Consider the d-dimensional random walk $(S_i)_{i=0}^\infty$ defined by $S_i := \xi_1 + \cdots + \xi_i$, $i \in \mathbb{N}$, and $S_0 := 0$. We are interested in the convex hull of the random walk which will be denoted by

$$C_{n,d} := \mathrm{conv}(S_0, S_1, \ldots, S_n) = \{\lambda_0 S_0 + \cdots + \lambda_n S_n \colon \lambda_0, \ldots, \lambda_n \geq 0, \lambda_0 + \cdots + \lambda_n = 1\}, \qquad n \geq d. \tag{1.4}$$

Let $f_k(C_{n,d})$ be the number of k-dimensional faces of the polytope $C_{n,d}$, for $k \in \{0, \ldots, d\}$. The following explicit formula for the expected face numbers of $C_{n,d}$ has been obtained in [42] relying on the methods of [43]:

$$\mathbb{E} f_k(C_{n,d}) = \frac{2 \cdot k!}{n!} \sum_{l=0}^{\infty} \begin{bmatrix} n+1 \\ d-2l \end{bmatrix} \begin{Bmatrix} d-2l \\ k+1 \end{Bmatrix}, \qquad k \in \{0, \ldots, d-1\}. \tag{1.5}$$

In terms of the Lah distribution introduced above, the formula can be stated as follows:

$$\frac{\mathbb{E} f_k(C_{n,d})}{\binom{n+1}{k+1}} = 2\mathbb{P}[\mathrm{Lah}(n+1,k+1) \in \{d, d-2, d-4, \ldots\}]. \tag{1.6}$$

We are interested in the high-dimensional limit when n, d and, possibly, also k, go to ∞ in a coupled manner. Let us argue that limit theorems for the Lah distribution imply threshold phenomena for $C_{n,d}$. Suppose, for example, that in some asymptotic regime we were able to prove a weak law of large numbers of the form

$$\frac{\mathrm{Lah}(n+1,k+1)}{\mathbb{E}\,\mathrm{Lah}(n+1,k+1)} \to 1 \text{ in probability.} \tag{1.7}$$

As we shall see below, the right-hand side of (1.6) does not differ much from the distribution function in the sense that the approximation

$$\frac{\mathbb{E} f_k(C_{n,d})}{\binom{n+1}{k+1}} \approx \mathbb{P}[\mathrm{Lah}(n+1,k+1) \leq d].$$

can be justified. The weak law of large numbers (1.7) then implies that

$$\lim_{n,d\to\infty} \frac{\mathbb{E} f_k(C_{n,d})}{\binom{n+1}{k+1}} = \begin{cases} 1, & \text{if } d > (1+\varepsilon)\mathbb{E}\,\mathrm{Lah}(n+1,k+1) \text{ for some } \varepsilon > 0, \\ 0, & \text{if } d < (1-\varepsilon)\mathbb{E}\,\mathrm{Lah}(n+1,k+1) \text{ for some } \varepsilon > 0, \end{cases}$$

which means that there is a threshold phenomenon if d is near $\mathbb{E}\,\mathrm{Lah}(n+1,k+1)$. In a similar way, a central limit theorem for $\mathrm{Lah}(n+1,k+1)$ would imply a characterization of the limit in the critical window.

1.3. **Summary of results.** Our goal is to investigate the properties of the Lah distribution. In particular, limit theorems for $\mathrm{Lah}(n,k)$ which we shall prove in various asymptotic regimes of n and k yield threshold phenomena for convex hulls of d-dimensional random walks as $d \to \infty$. Our main results can be summarized as follows.

(a) We provide a combinatorial interpretation of Lah distributions $\mathrm{Lah}(n,k)$ in terms of random compositions and records, which also allows us to construct the whole family of random variables $\mathrm{Lah}(n,k)$ simultaneously for all $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$ in a consistent way on a common probability space. This yields stochastic monotonicity of $\mathrm{Lah}(n,k)$ in n and k. This combinatorial construction is a subject of Section 2.

(b) We compute explicitly the expectation, the variance and the factorial moments of the Lah distribution. For example, we show that

$$\mathbb{E}\,\mathrm{Lah}(n,k) = \frac{1}{L(n,k)} \sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \frac{k}{\binom{n-1}{k-1}} \sum_{i=1}^{n-k+1} \frac{1}{i} \binom{n-i}{k-1} = \frac{k}{\binom{n-1}{k-1}} \sum_{i=1}^{n-k+1} \frac{(-1)^{i+1}}{i} \binom{n}{k+i-1}.$$

The aforementioned moment results as well as other basic properties of the Lah distribution are presented in Section 3.

(c) We prove that for fixed $k \in \mathbb{N}$, the random variables $\mathrm{Lah}(n,k)$ converge in the mod-Poisson sense with speed $\lambda_n = k \log n$, which implies several limit theorems including the central limit theorem

$$\frac{\mathrm{Lah}(n,k) - k\log n}{\sqrt{k \log n}} \xrightarrow[n\to\infty]{d} \mathrm{N}(0,1)$$

as well as the precise asymptotics for the large deviations probabilities. This regime of fixed k is analyzed in Section 4.

(d) In the regime when $k = k(n)$ grows linearly with n, that is $k(n) \sim \alpha n$ with $\alpha \in (0,1)$, we prove a central limit theorem and a large deviation principle for $\mathrm{Lah}(n,k)$; see Section 5.

(e) We apply these results to establish the aforementioned threshold phenomena for convex hulls of random walks in various asymptotic regimes of n, d, k in Section 6.

(f) We explain how the Lah distribution is related to the conic intrinsic volume sums of Weyl chambers in Section 7.

2. Combinatorial construction of the Lah distribution

In this section we shall establish connections between Lah distributions and several classical probabilistic and combinatorial models. This connection will allow us to construct the family $(\mathrm{Lah}(n,k))_{n\in\mathbb{N},1\le k\le n}$ in a consistent (simultaneously in n and in k) way on a common probability space. This, in turn, leads to a useful representation of $\mathrm{Lah}(n,k)$ and also establishes some basic qualitative properties of Lah distributions such as, for example, stochastic monotonicity. We start by observing that, for $k=1$, the formula for the Lah distribution takes the form

$$\mathbb{P}[\mathrm{Lah}(n,1) = j] = \frac{1}{n!} \begin{bmatrix} n \\ j \end{bmatrix}, \qquad j \in \{1,\dots,n\}. \tag{2.1}$$

This special case pops up at many places in probability theory, for example as the distribution of the number of cycles in a uniform random permutation of n elements, as the distribution of the number of records in an independent sample of size n from a continuous distribution, or as the distribution of $\sum_{\ell=1}^{n} B_\ell$, where $B_1, B_2, \dots$ are independent Bernoulli variables with distributions $\mathbb{P}[B_\ell = 1] = 1 - \mathbb{P}[B_\ell = 0] = 1/\ell$, $\ell \in \mathbb{N}$.

In order to extend these representations of $\mathrm{Lah}(n,1)$ to other values of k, we need to recall the notion of random compositions.

2.1. **Random compositions and records.** A composition of a positive integer n into k summands (blocks) is a representation of n as a sum $i_1 + i_2 + \cdots + i_k$ of k positive integers in which the order of summands is essential. Thus, $1+3$, $3+1$ and $2+2$ are three different compositions of $n = 4$ into $k = 2$ summands. By the standard "stars-and-bars" argument there are exactly $\binom{n-1}{k-1}$ different compositions of n into k summands. Throughout this paper we let $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ denote a random composition of n into k summands picked uniformly at random, that is, with distribution

$$\mathbb{P}[(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)}) = (i_1, i_2, \dots, i_k)] = \frac{1}{\binom{n-1}{k-1}}, \tag{2.2}$$

for every $i_1, i_2, \dots, i_k \in \mathbb{N}$ summing up to n. The family of random compositions $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ can be defined in a consistent way (simultaneously in $n \in \mathbb{N}$ and $k \in \{1, 2, \dots, n\}$) using a so-called Aldous' construction. We present its simplified version here in a form borrowed from [7], see Section 2.1.3 therein. Start with a chain of length n connecting the labeled vertices $U_1, U_2, \dots, U_n$, see Figure 1 (first row). This chain represents the unique composition of n into a single block and corresponds to time 1 of our construction. At time 2, pick one of $n-1$ edges uniformly at random and remove it. The resulting two connected components, see Figure 1 (second row), induce a uniformly distributed random composition of n into two summands. Proceeding this way and removing at time $k \in \{1, \dots, n\}$ an edge picked uniformly at random among existing $n-k+1$ edges, results in a consistent (in k) family of random compositions given by the sizes of connected components counted from left to right. The number of blocks at time k (or in the k-th row) is k. According to Lemma 2.1 in [7], a composition obtained after removing $k-1$ edges is uniformly distributed on the set of all partitions of n into k summands. Note that this construction is also consistent in n in the following sense. If we start with $n+1$ vertices, construct compositions $(b_1^{(n+1)}, b_2^{(n+1)}, \dots, b_k^{(n+1)})$ for $k = 1, 2, \dots, n+1$, and then remove completely the $(n+1)$-th column and a duplicated row (which necessarily appears upon deleting the $(n+1)$-th column), we obtain a family of uniform random compositions of n into k blocks distributed as $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ for $k = 1, 2, \dots, n$.

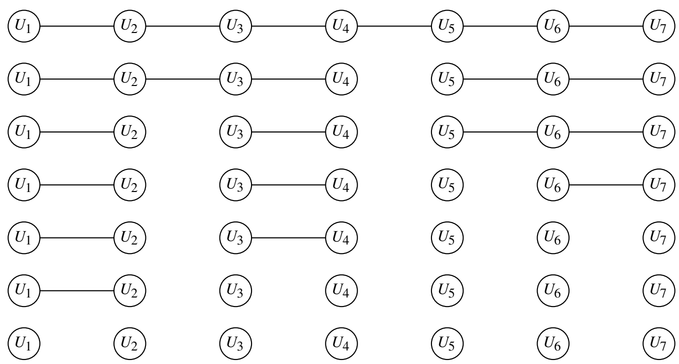

FIGURE 1. Aldous' construction of the consistent family of uniform random compositions. In this example a consistent family of partitions of $n = 7$ is: for $k = 1$, $7 = 7$; for $k = 2$, $7 = 4+3$; for $k = 3$, $7 = 2+2+3$; for $k = 4$, $7 = 2+2+1+2$; for $k = 5$, $7 = 2+2+1+1+1$; for $k = 6$, $7 = 2+1+1+1+1+1$ and, for $k = 7$, $7 = 1+1+1+1+1+1+1$.

So far the labels $U_1, U_2, \dots, U_n$ of vertices in Aldous' construction did not play a role but now we shall exploit them to construct a consistent family $(\text{Lah}(n,k))_{n \in \mathbb{N}, 1 \le k \le n}$. Let $U_1, U_2, \dots, U_n$ be a sample of independent uniformly distributed on $[0,1]$ random variables which is also independent of the above edge-removing process. For fixed $n \in \mathbb{N}$ and $k = 1, \dots, n$ take a composition $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ induced by the k-th row of Aldous' construction. We say that a vertex U_i is a record with respect to the composition $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ if it is a record in the block it occupies, that is, it is larger than all previous elements inside this block counting from left to right. The main result of this section if given by the next proposition.

Proposition 2.1. *The total number $X_{n,k}$ of records with respect to a uniform random composition of n into k summands has the* $\mathrm{Lah}(n,k)$ *distribution.*

The easiest way to prove Proposition 2.1 is via generating functions; but we shall also give a combinatorial proof. Recall that $[x^n]f(x)$ denotes the coefficient of x^n in the Taylor or Laurent expansion of $f(x)$ around 0. The following lemma will be useful on many occasions.

Lemma 2.2. *For all $n,k,j \in \mathbb{N}$ with $k \le j \le n$ we have*

$$\frac{k!}{n!}\begin{bmatrix} n \\ j \end{bmatrix}\begin{Bmatrix} j \\ k \end{Bmatrix} = [t^j][x^n]\left((1-x)^{-t}-1\right)^k.$$

Proof. Using both identities in (1.1) we have

$$\frac{\left((1-x)^{-t}-1\right)^k}{k!} = \frac{\left(\mathrm{e}^{t\log\frac{1}{1-x}}-1\right)^k}{k!} = \sum_{j=k}^{\infty}\left(\log\frac{1}{1-x}\right)^j \begin{Bmatrix} j \\ k \end{Bmatrix}\frac{t^j}{j!} = \sum_{j=k}^{\infty}\sum_{n=j}^{\infty}\frac{1}{n!}\begin{bmatrix} n \\ j \end{bmatrix}\begin{Bmatrix} j \\ k \end{Bmatrix} t^j x^n.$$

The claim follows by equating the coefficients. □

Proof of Proposition 2.1 using generating functions. For $t \in \mathbb{R}$ and $n \in \mathbb{N}$, let $\phi_n(t)$ be the generating function of the number of records in a sample of size n, that is,

$$\phi_n(t) = \mathbb{E}\, t^{\mathrm{Lah}(n,1)} = \sum_{j=1}^{n}\frac{1}{n!}\begin{bmatrix} n \\ j \end{bmatrix} t^j = \frac{t(t+1)\cdots(t+n-1)}{n!},$$

see (2.1). Conditioning on the event $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)}) = (i_1, i_2, \dots, i_k)$ we obtain, for $k \in \mathbb{N}$, $|x|<1$ and $t \in \mathbb{R}$,

$$\sum_{n=k}^{\infty}\binom{n-1}{k-1}\mathbb{E}\, t^{X_{n,k}} x^n = \sum_{n=k}^{\infty}\sum_{\substack{i_1+\dots+i_k=n \\ i_1,\dots,i_k\ge 1}}\prod_{\ell=1}^{k}(\phi_{i_\ell}(t)x^{i_\ell}) = \left(\sum_{i=1}^{\infty}\phi_i(t)x^i\right)^k = ((1-x)^{-t}-1)^k = \sum_{n=k}^{\infty}\binom{n-1}{k-1}\mathbb{E}\, t^{\mathrm{Lah}(n,k)} x^n,$$

where the last equality follows from Lemma 2.2 and equations (1.3) and (1.2). □

Combinatorial proof of Proposition 2.1. Fix $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$. Consider the set $\mathbf{L}_{n,k}$ of all pairs (σ, π), where $\pi = (A_1, \dots, A_k)$ is an ordered partition of the set $\{1, \dots, n\}$ into k non-empty blocks (and the order in which the blocks appear is essential), while $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ is a permutation preserving π meaning that σ only permutes the elements inside the blocks of π but not between the blocks. The total number of pairs (σ, π) in which σ has exactly j cycles is given by $\left[{n \atop j}\right] k! \left\{{j \atop k}\right\}$, which follows from the definition of the Stirling numbers of both kinds. The total number of pairs (σ, π) in $\mathbf{L}_{n,k}$ is $k!L(n,k)$, which either follows from (1.2), or by recalling that the Lah number $L(n,k)$ counts the number of partitions of $\{1, \dots, n\}$ into k blocks and putting linear order on the elements of each block. (The one-line notation of the restriction of σ to each block corresponds to a linear order on that block). Now, let (Σ, Π) be random and uniformly distributed on the finite set $\mathbf{L}_{n,k}$. As we argued above, the number of cycles of Σ has the Lah distribution $\mathrm{Lah}(n,k)$. On the other hand, let us take some deterministic ordered partition $\pi = (I_1, \dots, I_k)$ of $\{1, \dots, n\}$ into k blocks. The number of permutations σ preserving π is $|I_1|! \dots |I_k|!$. The number of ordered partitions of $\{1, \dots, n\}$ with prescribed block sizes $i_1 = |I_1|, \dots, i_k = |I_k|$ is given by $n!/i_1! \dots i_k!$. Hence, the block sizes of Π form a random, uniform composition of n in k summands. Conditionally on Π, the restrictions of Σ to these blocks are independent and uniform random permutations of the elements inside the blocks. Recall now that the number of cycles of a uniform random permutation of i_j elements has the same distribution as the number of records in a uniform sample of size i_j. Hence, the number of cycles of Σ has the same distribution as $X_{n,k}$, and the proof is complete. □

In the sequel we shall frequently use the following representation of the Lah distribution which is an immediate consequence of Proposition 2.1.

Proposition 2.3. *Let* $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$ *denote the uniform random composition of n into k parts, that is, a random composition with distribution* (2.2)*. Moreover, let* $(Z_i^{(j)})_{i,j\in\mathbb{N}}$ *be an array of mutually independent (and independent of* $(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})$*) random variables such that* $Z_n^{(j)} \stackrel{\mathrm{d}}{=} \mathrm{Lah}(n,1)$*, that is, has distribution* (2.1)*, for* $n, j \in \mathbb{N}$*. Then,*

$$\mathrm{Lah}(n,k) \stackrel{\mathrm{d}}{=} \sum_{j=1}^{k} Z^{(j)}_{b_j^{(n)}}, \tag{2.3}$$

where $\stackrel{\mathrm{d}}{=}$ *denotes equality in distribution.*

2.2. **Pólya urn coupling.** The coupling $(X_{n,k})_{n\in\mathbb{N}, 1\le k\le n}$ of the Lah distributions constructed above has the property that by its very definition $X_{n,k} \le X_{n,k+1}$ a.s. However, the monotonicity in n, i.e. the inequality $X_{n,k} \le X_{n+1,k}$, may fail in general. It turns out that, for every fixed $k \in \mathbb{N}$, there is another coupling of the sequence $\mathrm{Lah}(n,k)$, $n \in \{k, k+1, \dots\}$, which is non-decreasing in n. To construct it, let $U_1, U_2, \dots$ be independent random variables with the uniform distribution on $[0,1]$. Consider an urn containing k balls of k different colors and carrying labels $U_1, \dots, U_k$. Suppose that, at some step, there are $n-1$ balls in the urn. Draw one ball from the urn uniformly at random and return it to the urn together with one more ball which has the same color and carries label U_n, and proceed in this way. We say that U_n is a local record if U_n is larger than the labels of all balls which were already in the urn and had the same color as the ball with the label U_n. Let $b_j^{(n)}$ be the number of balls of color j when the number of balls in the urn is n and let $Y_{n,k}$ be the number of local records at this time. Then, $(b_1^{(n)}, \dots, b_k^{(n)})$ has the same distribution as the uniform random composition; see [39, Chapter 40]. Consequently, $Y_{n,k}$ has the Lah distribution $\mathrm{Lah}(n,k)$. Observe that by construction, we have $Y_{n,k} \le Y_{n+1,k}$ for all $n \ge k$.

2.3. **Stochastic monotonicity.** From the above constructions we immediately obtain the following result on stochastic monotonicity. It seems to be a non-trivial task to deduce it from from the definition of the Lah distribution given in (1.3) alone.

Proposition 2.4. *The Lah distributions* $\mathrm{Lah}(n,k)$ *satisfy the following stochastic monotonicity properties: for* $n \in \mathbb{N}$ *and* $k \in \{1, 2, \dots, n-1\}$ *we have*

$$\mathrm{Lah}(n,k) \stackrel{\mathrm{d}}{\le} \mathrm{Lah}(n,k+1), \tag{2.4}$$

and, for $n \in \mathbb{N}$ *and* $k \in \{1, 2, \dots, n\}$*,*

$$\mathrm{Lah}(n,k) \stackrel{\mathrm{d}}{\le} \mathrm{Lah}(n+1,k), \tag{2.5}$$

where for two real-valued random variables X, Y *we write* $X \stackrel{\mathrm{d}}{\le} Y$ *iff* $\mathbb{P}[X \le t] \ge \mathbb{P}[Y \le t]$ *for all* $t \in \mathbb{R}$*.*

Proof. The relations follow from the fact that the couplings constructed in Sections 2.1 and 2.2 satisfy $X_{n,k} \le X_{n,k+1}$ (in the Aldous coupling) and $Y_{n,k} \le Y_{n+1,k}$ (in the Pólya urn coupling), a.s. Both inequalities follow directly from the definitions of the corresponding couplings. □

Corollary 2.5. *For every* $n \in \mathbb{N}$*, the expectation of* $\mathrm{Lah}(n,k)$ *is a nondecreasing function of* $k \in \{1, \dots, n\}$*. For every* $k \in \mathbb{N}$*, the expectation of* $\mathrm{Lah}(n,k)$ *is a nondecreasing function of* $n \in \{k, k+1, \dots\}$*.*

3. Basic properties of the Lah distribution

We start by providing a representation for the generating function of a Lah-distributed random variable $\mathrm{Lah}(n,k)$ which is defined by

$$P_{n,k}(t) := \mathbb{E} t^{\mathrm{Lah}(n,k)} = \frac{1}{L(n,k)} \sum_{j=k}^{n} t^j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix}, \qquad t \in \mathbb{C}.$$

Lemma 3.1. *For all* $n \in \mathbb{N}$*,* $k \in \{1, \dots, n\}$ *and* $t \in \mathbb{C}$ *we have*

$$P_{n,k}(t) = \frac{1}{\binom{n-1}{k-1}} [x^n] \left((1-x)^{-t} - 1 \right)^k \tag{3.1}$$

$$= \frac{1}{\binom{n-1}{k-1}} \sum_{m=1}^{k} (-1)^{k-m} \binom{k}{m} \frac{\Gamma(tm+n)}{\Gamma(tm)n!}. \tag{3.2}$$

Proof. To prove (3.1), multiply the identity from Lemma 2.2 by t^j, sum over $j \in \{k,\dots,n\}$ and divide by $\binom{n-1}{k-1}$. It remains to prove (3.2). Using the binomial formula and (3.1), we obtain

$$P_{n,k}(t) = \frac{1}{\binom{n-1}{k-1}} [x^n] \sum_{m=0}^{k} (-1)^{k-m} \binom{k}{m} (1-x)^{-tm} = \frac{1}{\binom{n-1}{k-1}} \sum_{m=1}^{k} (-1)^{k-m} \binom{k}{m} [x^n](1-x)^{-tm}.$$

We dropped the term with $m=0$ since it vanishes. To complete the proof, recall the Taylor series

$$(1-x)^{-tm} = \sum_{n=0}^{\infty} \frac{tm(tm+1)(tm+2)\dots(tm+n-1)}{n!} x^n = \sum_{n=0}^{\infty} \frac{\Gamma(tm+n)}{\Gamma(tm)n!} x^n.$$

□

3.1. **Expectation and factorial moments of the Lah distribution.**

3.1.1. *Exact formulas for the expectation.* We are going to state exact formulas for the factorial moments of the Lah distribution or, more precisely, for expressions differing from the factorial moments by a missing normalizing factor of $1/L(n,k)$. We begin with the expectation.

Theorem 3.2 (Expectation)**.** *For all $n,k \in \mathbb{N}$ with $n \ge k$ we have*

$$\sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \frac{n!}{(k-1)!} [x^{n-k+1}]((1+x)^n \log(1+x)) \tag{3.3}$$

$$= (-1)^{n-k} \frac{n!}{(k-1)!} [x^{n-k+1}]((1+x)^{-k} \log(1+x)). \tag{3.4}$$

Equivalently, expanding $\log(1+x)$ *and the other terms in Taylor series and multiplying out, we have*

$$\sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \frac{n!}{(k-1)!} \sum_{i=1}^{n-k+1} \frac{1}{i} \binom{n-i}{k-1} = \frac{n!}{(k-1)!} \sum_{i=1}^{n-k+1} \frac{(-1)^{i+1}}{i} \binom{n}{k+i-1}. \tag{3.5}$$

Proof. The starting point of the proof is the formula

$$\frac{k!}{n!} \sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} t^j = [x^n] \left((1-x)^{-t} - 1\right)^k$$

which follows from Lemma 3.1. Since the function $((1-x)^{-t}-1)^k$ is analytic in (x,t) if (x,t) stays in a sufficiently small neighborhood of the point $(0,1)$, we may differentiate it any number of times in x and t and interchange the order of derivatives. Differentiating the above formula in t and putting $t=1$, we obtain

$$\frac{k!}{n!} \sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = -k[x^n] \left(\left(\frac{x}{1-x} \right)^k \frac{\log(1-x)}{x} \right) = -k[x^{n-k+1}]((1-x)^{-k} \log(1-x)).$$

Changing x to $-x$, we obtain (3.3). To prove (3.4), we rewrite (3.3) using the Cauchy formula as

$$\sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = (-1)^{n-k} \frac{n!}{(k-1)!} \frac{1}{2\pi \mathrm{i}} \oint_{\gamma} \frac{\log(1+x)}{(1+x)^k} \frac{\mathrm{d}x}{x^{n-k+2}}, \tag{3.6}$$

where the integration contour γ is a small counterclockwise circle centered at zero. Making the substitution $1+x = \frac{1}{1+y}$, we get

$$\sum_{j=k}^{n} j \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \frac{n!}{(k-1)!} \frac{1}{2\pi \mathrm{i}} \oint_{\gamma'} \frac{(1+y)^n \log(1+y)}{(1+y)^{n-k+2}} \mathrm{d}y \tag{3.7}$$

for some small counterclockwise contour γ' around 0. Using the Cauchy formula one more time, we arrive at (3.4). □

Remark 3.3. Alternatively, the first equality in (3.5) can be derived from (2.3) as follows. First, note that for all $i = 1,\dots,k$ and all $j = 1,\dots,n-k+1$,

$$\mathbb{P}[b_i^{(n)} = j] = \mathbb{P}[b_1^{(n)} = j] = \frac{\binom{n-j-1}{k-2}}{\binom{n-1}{k-1}}. \tag{3.8}$$

Thus, from (2.3), and with H_k denoting the k-th harmonic number, we have

$$\begin{aligned} L(n,k)\,\mathbb{E}\,\mathrm{Lah}(n,k) &= \frac{n!}{k!}\binom{n-1}{k-1}\mathbb{E}\left(\mathbb{E}\left(\sum_{i=1}^{k} Z^{(i)}_{b_i^{(n)}}\Big|(b_1^{(n)}, b_2^{(n)}, \dots, b_k^{(n)})\right)\right) = \frac{n!}{(k-1)!}\binom{n-1}{k-1}\mathbb{E}\,H_{b_1^{(n)}} \\ &= \frac{n!}{(k-1)!}\binom{n-1}{k-1}\sum_{j=1}^{n-k+1} H_j \mathbb{P}[b_1^{(n)} = j] = \frac{n!}{(k-1)!}\sum_{j=1}^{n-k+1} H_j\binom{n-j-1}{k-2} \\ &= \frac{n!}{(k-1)!}\sum_{j=1}^{n-k+1}\sum_{i=1}^{j}\frac{1}{i}\binom{n-j-1}{k-2} = \frac{n!}{(k-1)!}\sum_{i=1}^{n-k+1}\frac{1}{i}\sum_{j=i}^{n-k+1}\binom{n-j-1}{k-2} \\ &= \frac{n!}{(k-1)!}\sum_{i=1}^{n-k+1}\frac{1}{i}\binom{n-i}{k-1}. \end{aligned}$$

Remark 3.4. The Narumi polynomials $s_{\ell;a}(z)$, $\ell \in \mathbb{N}_0$, with parameter $a \in \mathbb{Z}$ are defined by the formula

$$\sum_{\ell=0}^{\infty}\frac{s_{\ell;a}(z)}{\ell!}t^{\ell} = \left(\frac{t}{\log(1+t)}\right)^{a}(1+t)^{z};$$

see [58]. With this notation, Theorem 3.2 takes the form

$$\sum_{j=k}^{n} j\begin{bmatrix} n \\ j\end{bmatrix}\begin{Bmatrix} j \\ k\end{Bmatrix} = (-1)^{n-k}k\binom{n}{k}s_{n-k;-1}(-k) = k\binom{n}{k}s_{n-k;-1}(n).$$

More generally, by taking the p-th derivative of the function $((1-x)^{-t}-1)^k$ at $t=1$ it is possible to express the p-th factorial moment of the Lah distribution through the Narumi polynomials with $a = -p$. For example, for the second factorial moment we get

$$\sum_{j=k}^{n} j(j-1)\begin{bmatrix} n \\ j\end{bmatrix}\begin{Bmatrix} j \\ k\end{Bmatrix} = \frac{(-1)^{n+k}n!s_{n-k;-2}(-k)}{(k-2)!(n-k)!} - \frac{(-1)^{n+k}n!s_{n-k-1;-2}(-k)}{(k-1)!(n-k-1)!}.$$

Expressions for higher factorial moments obtained in this way become more complicated, but we shall present a relatively simple general formula in Theorem 3.5. Note that the Narumi polynomials satisfy the functional equation $s_{\ell;a}(z) = s_{\ell;a}(\ell - a - 1 - z)$ which can be shown by using the Cauchy formula together with the same substitution as the one used to pass from (3.6) to (3.7).

3.1.2. *Exact formula for factorial moments.* The next theorem states a formula for the p-th factorial moment of the Lah distribution, up to a factor of $p!/L(n,k)$.

Theorem 3.5 (Factorial moments)**.** *For all $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$ and $p \in \mathbb{N}$ we have*

$$\sum_{j=k}^{n}\begin{bmatrix} n \\ j\end{bmatrix}\begin{Bmatrix} j \\ k\end{Bmatrix}\binom{j}{p} = n!\sum_{i=1}^{n-k+1}\binom{n-i}{k-1}\sum_{m=1}^{\min\{k,p\}}\frac{\left\{{p \atop m}\right\}}{(k-m)!}\frac{\left[{i+m-1 \atop p}\right]}{(i+m-1)!}.$$

For small values of p, this formula allows to express the p-th factorial moment of the Lah distribution in terms of binomial coefficients and the generalized harmonic numbers

$$H_n^{(m)} = \frac{1}{1^m} + \frac{1}{2^m} + \frac{1}{3^m} + \dots + \frac{1}{n^m}, \qquad H_n := H_n^{(1)}.$$

Indeed, if p is “small”, then the numbers $\left\{ {p \atop m} \right\}$ on the right-hand side are explicit constants, while the numbers $\left[{i \atop p} \right]$ can be expressed in terms of the generalized harmonic numbers, for example

$$\begin{bmatrix} i \\ 1 \end{bmatrix} = (i-1)!, \qquad \begin{bmatrix} i \\ 2 \end{bmatrix} = (i-1)!H_{i-1}, \qquad \begin{bmatrix} i \\ 3 \end{bmatrix} = \frac{1}{2}(i-1)!\left((H_{i-1})^2 - H_{i-1}^{(2)}\right), \qquad \dots .$$

Specifically, for $p=1$ we recover the first formula in (3.5), while for $p=2$, we obtain after some straightforward computations the following expression (which can easily be combined with (3.5) to write down an exact formula for the variance of the Lah distribution).

Corollary 3.6. *For all $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$ we have*

$$\sum_{j=k}^{n} j(j-1) \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \frac{2\cdot n!}{(k-1)!} \sum_{i=1}^{n-k+1} \binom{n-i}{k-1} \left(\frac{H_i \cdot (1+ik)}{i(i+1)} - \frac{1}{i^2} \right).$$

Let us also mention that for $k=1$, the identity of Theorem 3.5 takes the following form: for all $p \in \mathbb{N}$ and $n \in \mathbb{N}$,

$$\frac{1}{n!} \sum_{j=1}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \binom{j}{p} = \sum_{j=1}^{n} \frac{1}{j!} \begin{bmatrix} j \\ p \end{bmatrix}.$$

This equality is well known and, in fact, both sides are equal to $\frac{1}{n!}\left[{n+1 \atop p+1} \right]$; see Entries (6.15) and (6.21) of [30].

Proof of Theorem 3.5. Let $D_x^n\big|_{x=x_0} f(x)$ denote the n-th derivative of a function f evaluated at $x=x_0$. The starting point of the proof is the formula

$$k! \sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} t^j = D_x^n\Big|_{x=0} \left((1-x)^{-t} - 1\right)^k$$

which follows from Lemma 3.1. Taking the p-th derivative at $t=1$ we arrive at

$$\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} j^{\underline{p}} = D_x^n\Big|_{x=0} D_t^p\Big|_{t=1} \frac{\left((1-x)^{-t} - 1\right)^k}{k!},$$

where

$$j^{\underline{p}} := j(j-1)\cdots(j-p+1)$$

denotes the falling factorial. Our next goal is to write the right-hand side as a function of $1-t$. Extracting the factor $1/(1-x)$ and using the binomial formula, we arrive at

$$\begin{aligned}
\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} j^{\underline{p}} &= D_x^n\Big|_{x=0} D_t^p\Big|_{t=1} \left(\frac{1}{(1-x)^k} \frac{\left((1-x)^{1-t} - 1 + x\right)^k}{k!} \right) \\
&= D_x^n\Big|_{x=0} D_t^p\Big|_{t=1} \left(\frac{1}{(1-x)^k} \sum_{m=0}^{k} \frac{1}{k!} \binom{k}{m} ((1-x)^{1-t} - 1)^m x^{k-m} \right) \\
&= D_x^n\Big|_{x=0} D_t^p\Big|_{t=1} \left(\frac{1}{(1-x)^k} \sum_{m=0}^{k} \frac{((1-x)^{1-t} - 1)^m}{m!} \frac{x^{k-m}}{(k-m)!} \right).
\end{aligned}$$

Introducing the variable $s := t-1$, we can write

$$\begin{aligned}
\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} j^{\underline{p}} &= D_x^n\Big|_{x=0} D_s^p\Big|_{s=0} \left(\frac{1}{(1-x)^k} \sum_{m=0}^{k} \frac{((1-x)^{-s} - 1)^m}{m!} \frac{x^{k-m}}{(k-m)!} \right) \\
&= D_x^n\Big|_{x=0} \left(\frac{1}{(1-x)^k} \sum_{m=0}^{k} \frac{x^{k-m}}{(k-m)!} D_s^p\Big|_{s=0} \frac{(\mathrm{e}^{s \log \frac{1}{1-x}} - 1)^m}{m!} \right) \\
&= D_x^n\Big|_{x=0} \left(\frac{1}{(1-x)^k} \sum_{m=0}^{k} \frac{x^{k-m}}{(k-m)!} \left(\log \frac{1}{1-x} \right)^p \begin{Bmatrix} p \\ m \end{Bmatrix} \right),
\end{aligned}$$

where we have used the second relation in (1.1) for the last passage. Interchanging the order of summation, we obtain

$$\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} j^{\underline{p}} = \sum_{m=0}^{k} \begin{Bmatrix} p \\ m \end{Bmatrix} \frac{1}{(k-m)!} D_x^n \Big|_{x=0} \left(\frac{1}{(1-x)^k} x^{k-m} \left(\log \frac{1}{1-x} \right)^p \right)$$
$$= n! \sum_{m=0}^{k} \begin{Bmatrix} p \\ m \end{Bmatrix} \frac{1}{(k-m)!} [x^{n-k+m}] \left(\frac{1}{(1-x)^k} \left(\log \frac{1}{1-x} \right)^p \right).$$

Now we use the formulas

$$\left(\log \frac{1}{1-x} \right)^p = p! \sum_{i=p}^{\infty} \frac{x^i}{i!} \begin{bmatrix} i \\ p \end{bmatrix}, \qquad \frac{1}{(1-x)^k} = \sum_{j=0}^{\infty} x^j \binom{j+k-1}{j}.$$

Multiplying these two series and evaluating the coefficient of x^{n-k+m}, we get

$$\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} j^{\underline{p}} = n!p! \sum_{m=0}^{k} \begin{Bmatrix} p \\ m \end{Bmatrix} \frac{1}{(k-m)!} \sum_{i=p}^{n-k+m} \begin{bmatrix} i \\ p \end{bmatrix} \frac{1}{i!} \binom{n+m-i-1}{k-1}$$
$$= n!p! \sum_{m=0}^{k} \sum_{i=p}^{n-k+m} \begin{Bmatrix} p \\ m \end{Bmatrix} \begin{bmatrix} i \\ p \end{bmatrix} \frac{1}{i!(k-m)!} \binom{n+m-i-1}{k-1}.$$

Observe that the summation range in the first sum can be replaced by $m \in \{1, \dots, \min\{k, p\}\}$ because $\left\{ {p \atop m} \right\} = 0$ for $m = 0$ and $m > p$. After dividing by $p!$ this yields

$$\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} \binom{j}{p} = n! \sum_{m=1}^{\min\{k,p\}} \frac{1}{(k-m)!} \begin{Bmatrix} p \\ m \end{Bmatrix} \sum_{i=p}^{n-k+m} \frac{1}{i!} \begin{bmatrix} i \\ p \end{bmatrix} \binom{n+m-i-1}{k-1}.$$

To complete the proof, introduce the summation index $j = i - m + 1$ and interchange the order of summation. □

3.1.3. *Asymptotics of the expectation.* Based on the exact expression given in Theorem 3.2, we are able to derive the following

Theorem 3.7 (Asymptotics of the expectation)**.** *Let $n \to \infty$ and $k = k(n) \in \{1, \dots, n\}$ be a function of n. Then,*

$$\mathbb{E}\,\mathrm{Lah}(n,k) \sim \begin{cases} k \log(n/k), & \text{if } k = o(n), \\ \frac{\alpha \log(1/\alpha)}{1-\alpha} \cdot n, & \text{if } k \sim \alpha n \text{ for some } \alpha \in (0,1), \\ n, & \text{if } k \sim n. \end{cases} \tag{3.9}$$

We write $a_n \sim b_n$ if $a_n/b_n \to 1$ as $n \to \infty$.

Proof. According to (3.5), see also (1.2), we have

$$\mathbb{E}\,\mathrm{Lah}(n,k) = k \sum_{i=1}^{n-k+1} \frac{1}{i} \frac{\binom{n-i}{k-1}}{\binom{n-1}{k-1}} = k \sum_{i=1}^{n-k+1} \frac{1}{i} \prod_{m=1}^{i-1} \frac{n-k-m+1}{n-m}. \tag{3.10}$$

Now suppose that $k \sim \alpha n$ for some $\alpha \in (0,1)$ or $\alpha = 1$. If $i \in \mathbb{N}$ is fixed, then

$$\lim_{n\to\infty} \frac{1}{i} \prod_{m=1}^{i-1} \frac{n-k-m+1}{n-m} = \frac{(1-\alpha)^{i-1}}{i}.$$

Moreover, we have the bound

$$\max_{m=1,\dots,n-k} \frac{n-k-m+1}{n-m} \le \frac{n-k+1}{n} \underset{n\to\infty}{\longrightarrow} 1-\alpha.$$

It follows that for some sufficiently small $\varepsilon > 0$ and all sufficiently large n we have

$$\frac{1}{i} \prod_{m=1}^{i-1} \frac{n-k-m+1}{n-m} \le \frac{(1-\varepsilon)^{i-1}}{i}, \qquad i \in \{1, \dots, n-k+1\}.$$

Note that the right-hand side is summable in $i \in \mathbb{N}$. Interchanging the limit and the sum by the Lebesgue dominated convergence theorem, we obtain

$$\lim_{n\to\infty} \sum_{i=1}^{n-k+1} \frac{1}{i} \prod_{m=1}^{i-1} \frac{n-k-m+1}{n-m} = \sum_{i=1}^{\infty} \frac{(1-\alpha)^{i-1}}{i} = \begin{cases} \frac{-\log\alpha}{1-\alpha}, & \text{if } \alpha \in (0,1), \\ 1, & \text{if } \alpha = 1. \end{cases}$$

Let us now consider the case $k = o(n)$. Note that we do not require that $k \to \infty$. The idea is to show that in the sum on the right-hand side of (3.10), the summands with $i < n/k$ are approximately equal to $1/i$, whereas the contribution of the remaining summands is $O(1)$. Take some large constant $A > 1$ and let n be sufficiently large in the following. We start with a lower estimate. Recalling (3.10), dropping summands with $i > n/(Ak)$ and using the inequality $\prod_{j=1}^{M}(1-x_j) \geq 1 - \sum_{j=1}^{M} x_j$ which is valid for arbitrary numbers $x_1, \dots, x_M \in [0,1]$ and can be easily established by induction, we get

$$\frac{1}{k}\mathbb{E}\,\mathrm{Lah}(n,k) = \sum_{i=1}^{n-k+1} \frac{1}{i} \prod_{m=1}^{i-1} \left(1 - \frac{k-1}{n-m}\right) \geq \sum_{i=1}^{[n/(Ak)]} \frac{1}{i} \left(1 - \sum_{m=1}^{i-1} \frac{k-1}{n-m}\right).$$

For $1 \leq m < i \leq n/(Ak)$ we have $n - m \geq n - n/(Ak) \geq n/2$ and hence

$$\sum_{m=1}^{i-1} \frac{k-1}{n-m} \leq \sum_{m=1}^{i-1} \frac{k}{n/2} \leq \frac{2ki}{n} \leq \frac{2}{A}.$$

It follows that

$$\frac{1}{k}\mathbb{E}\,\mathrm{Lah}(n,k) \geq \left(1 - \frac{2}{A}\right) \sum_{i=1}^{[n/(Ak)]} \frac{1}{i} = \left(1 - \frac{2}{A}\right) \log\frac{n}{k} - O_A(1).$$

Since $\log(n/k) \to \infty$ and A can be arbitrarily large, we arrive at the lower bound

$$\liminf_{n\to\infty} \frac{\mathbb{E}\,\mathrm{Lah}(n,k)}{k\log(n/k)} \geq 1.$$

To prove the upper bound we shall use (2.3) and an elementary estimate $H_n = \log n + O(1)$, as follows:

$$\mathbb{E}\,\mathrm{Lah}(n,k) = k\,\mathbb{E}\,H_{b_1^{(n)}} = k(\mathbb{E}\log b_1^{(n)} + O(1)) \leq k(\log \mathbb{E}\, b_1^{(n)} + O(1)) = k(\log(n/k) + O(1)),$$

where we have used Jensen's inequality and the fact that $n = \mathbb{E}(b_1^{(n)} + \cdots + b_k^{(n)}) = k\,\mathbb{E}\, b_1^{(n)}$. Therefore,

$$\limsup_{n\to\infty} \frac{\mathbb{E}\,\mathrm{Lah}(n,k)}{k\log(n/k)} \leq 1,$$

which completes the proof. □

Remark 3.8. Let us mention a strange connection of (3.9) to a seemingly unrelated problem studied in [10]. Let L_n be the least common multiple of a random set A_n of integers obtained in the following way: every number from the set $\{1, \dots, n\}$ is included in A_n with probability $\alpha \in (0,1)$, independently from the others. Then, Theorem 1.1 in [10] states that

$$\frac{\log L_n}{n} \underset{n\to\infty}{\longrightarrow} \frac{\alpha\log(1/\alpha)}{1-\alpha} \quad \text{in probability.}$$

The expression on the right-hand side is the same as in (3.9). Moreover, Theorem 1.2 in [10] bears similarity with the $k = o(n)$ case of (3.9). We were not able to explain this coincidence. A central limit theorem for $\log L_n$ has been proved in [1, Corollary 1.5]. The asymptotic variance given in [1, Remark 1.3] does not coincide with the asymptotic variance of the Lah distribution given in Theorem 5.1.

3.2. **Log-concavity and unimodality.** Well-known properties of the Stirling numbers of both kinds yield the following proposition.

Proposition 3.9. *For each $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$, the Lah distribution* $\mathrm{Lah}(n,k)$ *is log-concave, that is*

$$\mathbb{P}[\mathrm{Lah}(n,k)=i]^2 \geq \mathbb{P}[\mathrm{Lah}(n,k)=i-1]\mathbb{P}[\mathrm{Lah}(n,k)=i+1] \qquad \text{for all } i \in \{k,\dots,n\}.$$

Proof. It is well known (see, e.g. [62, Corollary 3.2]) that the Stirling numbers are log-concave, that is

$$\begin{bmatrix} n \\ j \end{bmatrix}^2 \geq \begin{bmatrix} n \\ j+1 \end{bmatrix} \begin{bmatrix} n \\ j-1 \end{bmatrix}.$$

By [62, Theorem 3.3], the sequence $(\left\{{j+1 \atop k}\right\}/\left\{{j \atop k}\right\})_{j=k,k+1,\dots}$ is strictly decreasing for every $k \geq 2$ (and is identically equal to 1 for $k=1$) which means that

$$\begin{Bmatrix} j \\ k \end{Bmatrix}^2 \geq \begin{Bmatrix} j+1 \\ k \end{Bmatrix} \begin{Bmatrix} j-1 \\ k \end{Bmatrix},$$

with a strict inequality for $k \geq 2$. Multiplying these two inequalities, we arrive at the claim. □

Corollary 3.10. *For each $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$, the Lah distribution* $\mathrm{Lah}(n,k)$ *is unimodal. That is, there exists $m_{n,k} \in \{k,\dots,n\}$ such that $i \mapsto \mathbb{P}[\mathrm{Lah}(n,k)=i]$ is nondecreasing for $i \leq m_{n,k}$ and nonincreasing for $i \geq m_{n,k}$.*

3.3. **Zeroes of the generating function.** In the following, we shall prove central limit theorems for the Lah distribution. A natural approach towards this is the Harper method [31], for which one needs to verify that the zeroes of $P_{n,k}(t)$ are real and negative. Numerical simulations, see Figure 2, show that the zeroes are not real except in the special case $k=1$ and suggest the following

Conjecture 3.11. *All complex zeroes of $P_{n,k}$ have negative real parts.*

In fact, this conjecture would also be sufficient to apply Harper's method, see, e.g., [52, Theorem 3.1]. We shall not investigate the properties of zeroes here and mention only one result. In the special case $z=1$ it is well known.

Proposition 3.12. *Let $k \in \mathbb{N}$ and let $n > k$ be integer. Then, for every $z \in \{1,2,\dots,\lfloor \frac{n-1}{k} \rfloor\}$ we have*

$$\sum_{j=k}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} (-z)^j = 0.$$

Proof. It suffices to show that $P_{n,k}(-z)=0$. We use (3.1). Note that $(1-x)^z$ is a polynomial in x of degree z. Hence, $((1-x)^z-1)^k$ is a polynomial in x of degree zk. If $z \leq \lfloor \frac{n-1}{k} \rfloor$, then $zk < n$ and the coefficient of x^n in this polynomial vanishes. Then, (3.1) implies that $P_{n,k}(-z)=0$. □

4. Limit theorems for the Lah distribution: The constant k regime

4.1. **Mod-Poisson convergence and its consequences.** In this section we shall state and prove limit theorems for the random variables $\mathrm{Lah}(n,k)$ in the regime when $k \in \mathbb{N}$ is fixed and $n \to \infty$. The basic tool we shall use is the notion of mod-Poisson convergence introduced by Kowalski and Nikeghbali in [48]; see also [5, 12, 37, 49, 55] for a more general notion of mod-ϕ-convergence and [22] for a monograph treatment of the subject.

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables with values in $\mathbb{R}$ whose Laplace transforms $\mathbb{E}\,\mathrm{e}^{zX_n}$ exist finitely for all $z \in \mathbb{C}$ and $(\lambda_n)_{n\in\mathbb{N}}$ a sequence of positive numbers with $\lim_{n\to\infty}\lambda_n = +\infty$. The sequence $(X_n)_{n\in\mathbb{N}}$ is said to *converge in the mod-Poisson sense* with speed $(\lambda_n)_{n\in\mathbb{N}}$ if

$$\lim_{n\to\infty} \frac{\mathbb{E}\,\mathrm{e}^{zX_n}}{\mathrm{e}^{\lambda_n(\mathrm{e}^z-1)}} = \Psi(z) \tag{4.1}$$

uniformly on compact subsets of some open set $\mathscr{D} \subset \mathbb{C}$ containing the real axis. Here, $\Psi : \mathscr{D} \to \mathbb{C}$ is some analytic function. In the literature, several non-equivalent definitions of mod-Poisson (and, more generally, mod-φ) convergence exist, which differ by the shape of the domain $\mathscr{D}$. The notion we use here is close but not equivalent to the definition used in the book [22], see Definition 1.1.1 therein, where $\mathscr{D}$ is assumed to be a vertical strip containing the imaginary axis. Nevertheless, most important corollaries of the mod-φ convergence continue to hold under the

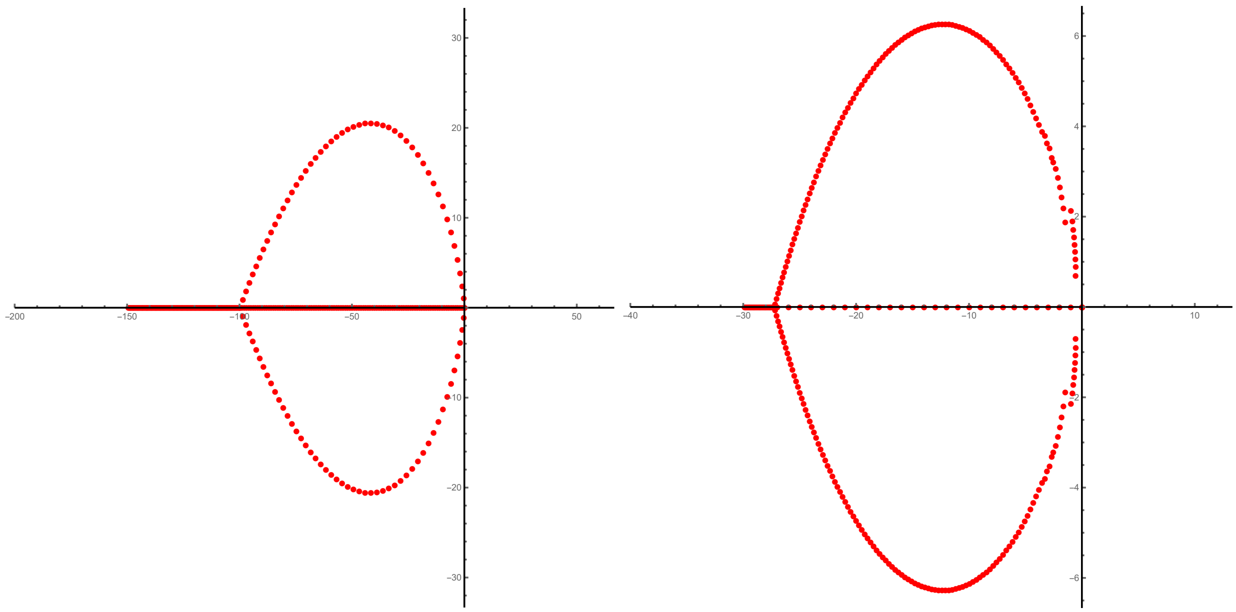

FIGURE 2. Complex zeroes of the polynomial $P_{n,k}$ with $n = 300$ and $k = 2$ (left) and $k = 10$ (right).

assumption that $\mathscr{D} \subset \mathbb{C}$ is an open set containing a segment of the real line, see [41, Remark 2.10]. As we shall see below in Theorem 4.1, definition (4.1) is more suitable for the Lah distribution.

To interpret the above definition, recall that the generating function of the Poisson distributed random variable with parameter λ_n is given by

$$\mathbb{E}\,\mathrm{e}^{z\mathrm{Poi}(\lambda_n)} = \mathrm{e}^{\lambda_n(\mathrm{e}^z-1)},$$

which is the denominator in (4.1). Thus, (4.1) states the heuristic approximation

$$\text{“}X_n \overset{d}{=} \mathrm{Poi}(\lambda_n) + \Xi + o(1)\text{”}, \tag{4.2}$$

where Ξ is a “random variable” with “moment generating function” $\mathbb{E}\,\mathrm{e}^{z\Xi} = \Psi(z)$ that is independent of $\mathrm{Poi}(\lambda_n)$, and $o(1)$ converges to 0 in distribution. Even though usually no random variable Ξ having the required moment generating function $\Psi(z)$ exists, a lot of limit theorems for X_n have the same form as they would do for the sequence of “random variables” $\mathrm{Poi}(\lambda_n) + \Xi$.

The next theorem states that for fixed $k \in \mathbb{N}$, the random variables $(\mathrm{Lah}(n,k))_{n\in\mathbb{N}}$ converge in the mod-Poisson sense with speed $\lambda_n = k \log n$, as $n \to \infty$.

Theorem 4.1 (Mod-Poisson convergence)**.** *Let $k \in \mathbb{N}$ be fixed. Then,*

$$\lim_{n\to\infty} \frac{\mathbb{E}\,\mathrm{e}^{z\,\mathrm{Lah}(n,k)}}{\mathrm{e}^{(k\log n)(\mathrm{e}^z-1)}} = \frac{\Gamma(k)}{\Gamma(k\mathrm{e}^z)} \tag{4.3}$$

for every $z \in \mathscr{D}_{Lah}$, where $\mathscr{D}_{Lah} := \{t \in \mathbb{C} : \cos \mathrm{Im}\, t > 0\} \supset \mathbb{R}$. Moreover, this convergence is uniform as long as z stays in any compact subset K of $\mathscr{D}_{Lah}$, and the speed of convergence is $O(n^{-\varepsilon(K)})$ for some $\varepsilon(K) > 0$.

The proof of Theorem 4.1 is postponed to Section 4.2.

Theorem 4.2 (Central limit theorem)**.** *Let $k \in \mathbb{N}$ be fixed. Then,*

$$\frac{\mathrm{Lah}(n,k) - k\log n}{\sqrt{k\log n}} \underset{n\to\infty}{\overset{d}{\longrightarrow}} N(0,1).$$

Proof. The claim follows from the mod-Poisson convergence (4.3) by [48, Proposition 2.4(2)]. Note that the cited result only requires uniformity of the convergence in a small neighborhood of the origin which is secured by (4.3). □

All subsequent results of this section follow essentially from the corresponding general results on random profiles obtained in [41]. This reference better fits our needs since we have uniform convergence in a horizontal strip rather than a vertical one, preventing us from referring to the standard results on the mod-φ convergence. Note that Assumptions A1–A3 of [41] can be easily verified to hold with

$$\mathbb{L}_n(j) = \mathbb{P}[\mathrm{Lah}(n,k) = j], \quad w_n = k\log n, \quad \beta_\pm = \pm\infty \quad \mathscr{D} = \mathbb{R}\times(-\pi \mathrm{i}/2, +\pi i/2),$$
$$\text{and} \quad \phi(\beta) = \mathrm{e}^\beta - 1, \quad W_\infty(\beta) = \Gamma(k)/\Gamma(k\mathrm{e}^\beta) \quad \text{for} \quad \beta\in\mathscr{D}, \tag{4.4}$$

whereas Assumption A4 will be checked in Remark 4.7.

Theorem 4.3 (Local limit theorem)**.** *For every fixed $k\in\mathbb{N}$ we have*

$$\lim_{n\to\infty}\sqrt{\log n}\sup_{m\in\mathbb{Z}}\left|\mathbb{P}[\mathrm{Lah}(n,k)=m] - \frac{1}{\sqrt{2\pi k\log n}}\exp\left\{-\frac{(m-k\log n)^2}{2k\log n}\right\}\right| = 0.$$

Proof. In view of (4.3) and (4.4) this follows from Theorem 2.7 of [41]. Moreover, a general Edgeworth asymptotic expansion with an arbitrary number of terms could be derived from [41, Theorem 2.1]. □

An illustration of Theorems 4.2 and 4.3 is shown on Figure 3.

Theorem 4.4 (Precise asymptotics of large deviations)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers converging to $x>0$ and such that $kx_n\log n$ is integer for all $n\in\mathbb{N}$. Then,*

$$\mathbb{P}[\mathrm{Lah}(n,k) = kx_n\log n] \sim \frac{n^{-k(x_n\log x_n - x_n+1)}}{\sqrt{2\pi kx\log n}}\frac{\Gamma(k)}{\Gamma(kx)},$$
$$\mathbb{P}[\mathrm{Lah}(n,k) \geq kx_n\log n] \sim \frac{x}{x-1}\frac{n^{-k(x_n\log x_n - x_n+1)}}{\sqrt{2\pi kx\log n}}\frac{\Gamma(k)}{\Gamma(kx)}, \qquad \text{if } x>1,$$
$$\mathbb{P}[\mathrm{Lah}(n,k) \leq kx_n\log n] \sim \frac{1}{1-x}\frac{n^{-k(x_n\log x_n - x_n+1)}}{\sqrt{2\pi kx\log n}}\frac{\Gamma(k)}{\Gamma(kx)}, \qquad \text{if } x<1.$$

Proof. The first claim follows from Theorem 2.8 in [41] applied with $r=0$, K being an arbitrary segment of the real line which contains x, and $\beta_n(kx_n\log n) = \log x_n$. The second claim follows by summation, c.f. [22, Theorem 3.2.2]. The last claim follows from the same results applied to $-\mathrm{Lah}(n,k)$. □

Proposition 4.5 (Location of the mode)**.** *For every fixed $k\in\mathbb{N}$ there is $N_1\in\mathbb{N}$ such that for all integer $n>N_1$, all maximizers of the function $m\mapsto\mathbb{P}[\mathrm{Lah}(n,k)=m]$ are among the following two numbers:*

$$\left\lfloor k\log n - \frac{k\Gamma'(k)}{\Gamma(k)} - \frac{1}{2}\right\rfloor, \quad \left\lceil k\log n - \frac{k\Gamma'(k)}{\Gamma(k)} - \frac{1}{2}\right\rceil.$$

Proof. This follows from Theorem 2.11 of [41]. □

Remark 4.6. A random variable X_n^θ has the Ewens or the Karamata-Stirling distribution with parameters $n\in\mathbb{N}$ and $\theta>0$ if

$$\mathbb{P}[X_n^\theta = j] = \frac{\theta^j}{\theta(\theta+1)\cdots(\theta+n-1)}\begin{bmatrix} n\\ j\end{bmatrix}, \qquad j\in\{1,\dots,n\}.$$

It is well known that X_n^θ has the same distribution as the number of cycles in the Ewens random permutation. Coincidentally, if $\theta = k$ happens to be integer, the sequence $(X_n^\theta)_{n\in\mathbb{N}}$ satisfies the same mod-Poisson convergence as $(\mathrm{Lah}(n,k))_{n\in\mathbb{N}}$; see [40]. Let us mention that the Lah distribution could be generalized by introducing an additional parameter $\theta>0$ (the probability that the random variable $\mathrm{Lah}(n,k;\theta)$ takes the value j is by definition proportional to $\theta^j\left[{n \atop j}\right]\left\{{j \atop k}\right\}$, for $j\in\{k,\dots,n\}$). Most of our results could be generalized to arbitrary $\theta>0$, but since we have no applications for this general setting, we restrict ourselves to the case $\theta=1$.

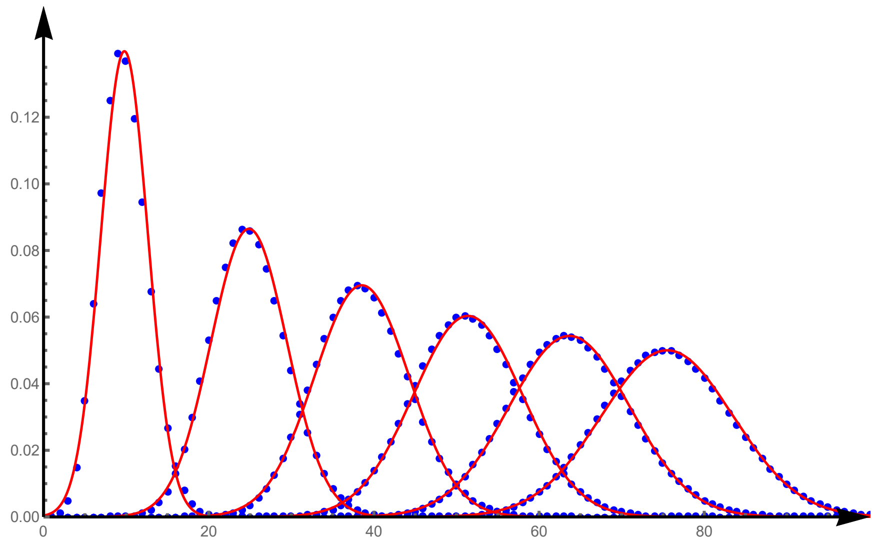

FIGURE 3. The probability mass functions of Lah distributions with $n = 10000$ and $k \in \{1, 3, 5, \dots, 11\}$ (blue dots) together with the approximating normal densities (red curves); see Theorems 4.2 and 4.3. The parameters of the normal densities are the true expectation and variance of the Lah distribution (for the normalization given in Theorems 4.2 and 4.3 the convergence is too slow).

4.2. **Proof of Theorem 4.1.** We need to prove that

$$\lim_{n\to\infty} \frac{P_{n,k}(\mathrm{e}^z)}{n^{k(\mathrm{e}^z-1)}} = \frac{\Gamma(k)}{\Gamma(k\mathrm{e}^z)}.$$

Since the number of summands in (3.2) is fixed, we can consider the asymptotics of each summand separately. For every $m \in \{1, \dots, k\}$, the m-th summand with $t = \mathrm{e}^z$ satisfies

$$(-1)^{k-m}\binom{k}{m}\frac{\Gamma(m\mathrm{e}^z+n)}{\Gamma(m\mathrm{e}^z)n!} = (-1)^{k-m}\binom{k}{m}\frac{n^{m\mathrm{e}^z-1}}{\Gamma(m\mathrm{e}^z)}(1+O(1/n))$$

by the formula $\Gamma(n+\alpha)/\Gamma(n+\beta) = n^{\alpha-\beta}(1+O(1/n))$ as $n \to \infty$, which holds locally uniformly in $\alpha, \beta \in \mathbb{C}$, see, for example, Theorem in [23]. If z stays in a compact subset K of $\mathscr{D}_{Lah} = \{t \in \mathbb{C} : \cos \operatorname{Im} t > 0\}$, then $\operatorname{Re} \mathrm{e}^z > \varepsilon(K) > 0$ stays bounded away from 0 for some sufficiently small $\varepsilon(K) \in (0,1)$. It follows that the summand with $m = k$ dominates in the following sense:

$$P_{n,k}(\mathrm{e}^z) = \frac{1}{\binom{n-1}{k-1}}\left(\frac{n^{k\mathrm{e}^z-1}}{\Gamma(\mathrm{e}^z k)}(1+O(1/n)) + \sum_{m=1}^{k-1}(-1)^{k-m}\binom{k}{m}\frac{n^{m\mathrm{e}^z-1}}{\Gamma(m\mathrm{e}^z)}(1+O(1/n))\right) = \frac{\Gamma(k)}{\Gamma(k\mathrm{e}^z)}n^{k(\mathrm{e}^z-1)}(1+O(n^{-\varepsilon(K)})),$$

which proves the claim. Observe that in the case when $\cos \operatorname{Im} z \le 0$ this argument does not apply. □

Remark 4.7. Assumption A4 of [41] can be verified in a similar way by observing that given $a \in (0, \pi)$ and a compact set $K \subset \mathbb{R}$, for all $\beta \in K$ it holds that

$$n^{-k(\mathrm{e}^\beta-1)}\int_a^\pi |P_{n,k}(\mathrm{e}^{\beta+\mathrm{i}u})|\mathrm{d}u \le Cn^{-k(\mathrm{e}^\beta-1)}\sum_{m=1}^k\int_a^\pi \frac{\left|\Gamma(\mathrm{e}^{\beta+\mathrm{i}u}m+n)\right|}{n!n^{k-1}}\mathrm{d}u \le C\sum_{m=1}^k\int_a^\pi |n^{\mathrm{e}^{\beta+\mathrm{i}u}m-\mathrm{e}^\beta k}|\mathrm{d}u \le Cn^{-\delta}$$

for some constants $C = C(K,k) > 0$ and $\delta = \delta(K,k,a) > 0$.

4.3. **Strong law of large numbers.** For the Pólya urn coupling $(Y_{n,k})_{n=k,k+1,\dots}$ constructed in Section 2.2, where k is fixed, the following strong law of large numbers holds.

Proposition 4.8. *For every fixed $k \in \mathbb{N}$ we have*

$$\frac{Y_{n,k}}{\log n} \xrightarrow[n\to\infty]{a.s.} k.$$

Proof. Differentiating (4.1) and plugging $z=0$, which is legitimate since the convergence is uniform in a neighborhood of the origin, we obtain

$$\lim_{n\to\infty} \frac{\mathbb{E}\,\mathrm{Lah}(n,k)}{k\log n} = \lim_{n\to\infty} \frac{\mathrm{Var}\,\mathrm{Lah}(n,k)}{k\log n} = 1.$$

Thus, by Chebyshev's inequality, for every $\varepsilon > 0$,

$$\mathbb{P}\left[\left|\frac{Y_{n,k}}{\mathbb{E}\,Y_{n,k}} - 1\right| > \varepsilon\right] \le \frac{\mathrm{Var}\,\mathrm{Lah}(n,k)}{\varepsilon^2(\mathbb{E}\,\mathrm{Lah}(n,k))^2} \sim \frac{\varepsilon^{-2}}{k\log n}, \quad n\to\infty.$$

By the Borel-Cantelli lemma,

$$\frac{Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}}{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}} \xrightarrow[n\to\infty]{a.s.} 1.$$

The result now follows from the standard sandwich argument using monotonicity of $(Y_{n,k})_{n=k,k+1,\dots}$. Indeed, for every $m \ge 3$ there exists $n\in\mathbb{N}$ such that $\lfloor \mathrm{e}^{n^2}\rfloor \le m < \lfloor \mathrm{e}^{(n+1)^2}\rfloor$. Therefore,

$$\frac{Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}}{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}} \frac{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}}{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{(n+1)^2}\rfloor,k}} \le \frac{Y_{m,k}}{\mathbb{E}\,Y_{m,k}} \le \frac{Y_{\lfloor \mathrm{e}^{(n+1)^2}\rfloor,k}}{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{(n+1)^2}\rfloor,k}} \frac{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{(n+1)^2}\rfloor,k}}{\mathbb{E}\,Y_{\lfloor \mathrm{e}^{n^2}\rfloor,k}}.$$

Sending $n\to\infty$, finishes the proof. □

5. Limit theorems for the Lah distribution: regimes of growing k

Throughout this section we assume that $n\to\infty$ and $k=k(n)\to\infty$. There are two main regimes: the *central regime* in which

$$\lim_{n\to\infty}\frac{k(n)}{n} = \alpha \qquad \text{for some constant } \alpha\in(0,1), \tag{5.1}$$

and the *intermediate regime*, in which $k(n)=o(n)$. We begin with the central regime.

5.1. Central limit theorem in the central regime.

Theorem 5.1 (CLT in the central regime)**.** *Assume* (5.1). *Then,*

$$\frac{\mathrm{Lah}(n,k)-\mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{n}} \xrightarrow[n\to\infty]{d} N\left(0, -\left(\frac{\alpha}{1-\alpha} + \frac{\alpha(\alpha+1)\log\alpha}{(1-\alpha)^2} + \frac{\alpha^2\log^2\alpha}{(1-\alpha)^3}\right)\right),$$

where $N(m,\sigma^2)$ denotes a normal random varible with mean $m\in\mathbb{R}$ and variance $\sigma^2>0$.

Remark 5.2. Despite of the minus sign in front of the formula for the variance, it is positive. For $\alpha\to 0$ and $\alpha\to 1$ the variance vanishes; the maximal value is attained at $\alpha = 0.23517\dots$. An illustration of Theorem 5.1 is shown on Figure 4.

The proof of Theorem 5.1 relies on the representation (2.3) and a multivariate central limit theorem for the number of blocks of fixed sizes in a uniform random composition $(b_1^{(n)},\dots,b_k^{(n)})$ of n. For $j\in\mathbb{N}$, let us denote by $N_j^{(n)}$ the number of blocks of size j in the composition $(b_1^{(n)},\dots,b_k^{(n)})$. Thus,

$$N_j^{(n)} := \sum_{i=1}^{k} \mathbb{1}_{\{b_i^{(n)}=j\}}.$$

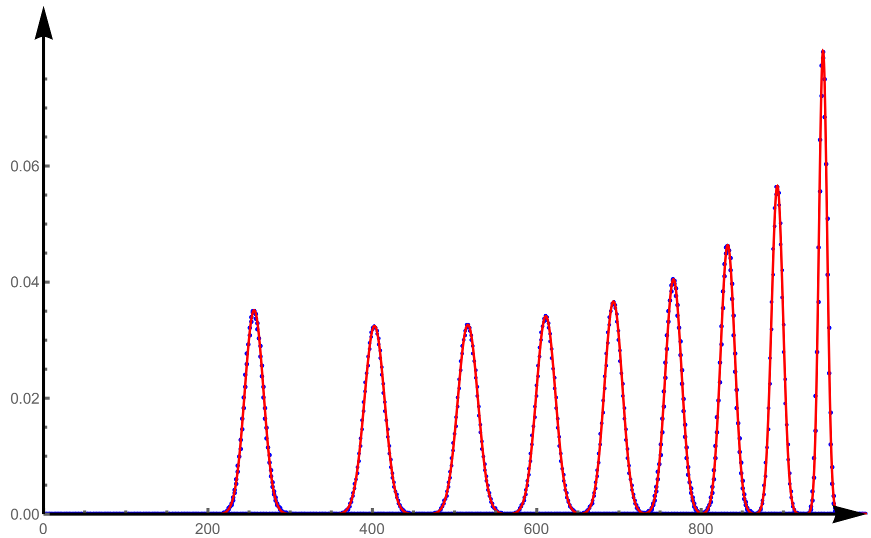

FIGURE 4. The probability mass functions of Lah distributions with $n = 1000$ and $k = \alpha n$ with $\alpha \in \{\frac{1}{10}, \frac{2}{10}, \dots, \frac{9}{10}\}$ (blue dots) together with the approximating normal densities (red curves).

Note that, by (3.8), $\mathbb{E} N_j^{(n)} = k\mathbb{P}[b_1^{(n)} = j] = k\binom{n-1}{k-1}^{-1}\binom{n-j-1}{k-2} \sim k\alpha(1-\alpha)^{j-1}$, where the asymptotic equivalence holds whenever (5.1) is in force. In particular, this implies that under assumption (5.1), the random variables $b_1^{(n)}$ (and, thus $b_i^{(n)}$ for every fixed $i = 1, \dots, k$) converge in distribution, as $n \to \infty$, to a geometric law on $\mathbb{N}$ with success probability α, see [14, Section 4] for much stronger results.

Theorem 5.3 (Central limit theorem for $(N_1^{(n)}, N_2^{(n)}, \dots,))$**.** *Assume* (5.1)*. Then, as $n \to \infty$,*

$$\left(\frac{N_j^{(n)} - k\mathbb{P}[b_1^{(n)} = j]}{\sqrt{k}}\right)_{j\geq 1} \xrightarrow[n\to\infty]{d} (\mathcal{N}_j)_{j\geq 1},$$

in $\mathbb{R}^\infty$ endowed with the product topology, and $(\mathcal{N}_j)_{j\geq 1}$ is a centred Gaussian vector with the covariance

$$\mathrm{Cov}(\mathcal{N}_i, \mathcal{N}_j) = \mathbb{E}\mathcal{N}_i\mathcal{N}_j = p_i \mathbb{1}_{\{i=j\}} - p_i p_j - \frac{p_i p_j}{1-\alpha}(\alpha i - 1)(\alpha j - 1), \quad i, j \in \mathbb{N},$$

where $p_j := \alpha(1-\alpha)^{j-1}$, for $j \in \mathbb{N}$.

Remark 5.4. Let us mention an interpretation of the random vector $(\mathcal{N}_j)_{j\geq 1}$ as a conditional distribution. If $\mathcal{G}_1, \mathcal{G}_2, \dots$ are independent centered Gaussian variables with $\mathrm{Var}\mathcal{G}_j = p_j$, then $(\mathcal{N}_j)_{j\geq 1}$ has the same distribution as $(\mathcal{G}_j)_{j\geq 1}$ conditioned on the event $\{\sum_{j=1}^\infty \mathcal{G}_j = 0, \sum_{j=1}^\infty j\mathcal{G}_j = 0\}$. This can be easily verified using the formulas for the covariance matrix of the conditional Gaussian distribution.

Theorem 5.3 is known and has been rediscovered several times. Its proofs are based on a representation of the distribution of $(b_1^{(n)}, \dots, b_k^{(n)})$ as the law of k independent geometrically distributed random variables conditioned on their sum to be n. More precisely, we have

$$\mathbb{P}[(b_1^{(n)}, \dots, b_k^{(n)}) \in \cdot] = \mathbb{P}[(G_1, \dots, G_k) \in \cdot | G_1 + \dots + G_k = n], \tag{5.2}$$

where $G_1, \dots, G_k$ are independent random variables having the same geometric law on $\mathbb{N}$ with parameter θ. Note that this representation holds for arbitrary $\theta \in (0,1)$ and we are free to choose it as we wish. It is convenient to put $\theta := \theta_n = k/n$, so that the mean of $G_1 + \dots + G_k$ is n. This identifies the random composition as a special case of the generalized allocation scheme introduced by V. F. Kolchin in [46] and much studied thereafter; see, e.g., [47] and [45, Chapter VIII]. In particular, the convergence of one-dimensional distributions in Theorem 5.3 is contained in Theorem 1 of [46]. The full statement of Theorem 5.3 is a special case of the general results of Holst, see [32, Theorem 2] or [33, Theorem 2], but it requires some effort to see this. The proofs by Holst rely on the paper by Le Cam [51] who studied a related question for exponential random variables. A special case of Holst's results, from which Theorem 5.3 follows directly, can be found in the paper by Ivchenko [36, Theorem 4]. The corresponding multidimensional local limit theorem was derived by Trunov [65, Theorem 2.1] who did not rely on [51]. See also [69, Theorem 3] for a weak law of large numbers, and [60, Section 4.4] for a similar result about partitions instead of compositions. To keep the paper self-contained we shall give a sketch of the proof of Theorem 5.3 is the Appendix.

The papers [32, 33, 36, 65] use the sequence $k\theta_n(1-\theta_n)^{j-1} = k\mathbb{P}[G_1 = j]$ to center $N_j^{(n)}$. Let us check that it can be replaced by the sequence $\mathbb{E}N_j^{(n)} = k\mathbb{P}[b_1^{(n)} = j]$ used in Theorem 5.3.

Lemma 5.5. *Assume* (5.1) *and put* $\theta_n := k/n$. *Then, for every fixed* $j \in \mathbb{N}$, *we have*

$$\left| k\mathbb{P}[b_1^{(n)} = j] - k\theta_n(1-\theta_n)^{j-1} \right| = O(1), \quad n \to \infty.$$

Proof. By (3.8), we have

$$\begin{aligned}
\mathbb{P}[b_1^{(n)} = j] &= \frac{\binom{n-j-1}{k-2}}{\binom{n-1}{k-1}} = \frac{k-1}{n-1}\left(1-\frac{k-2}{n-j}\right)\left(1-\frac{k-2}{n-j+1}\right)\cdots\left(1-\frac{k-2}{n-2}\right) \\
&= \left(\frac{k}{n} + O\left(\frac{1}{n}\right)\right)\left(1-\frac{k}{n} + O\left(\frac{1}{n}\right)\right)\left(1-\frac{k}{n} + O\left(\frac{1}{n}\right)\right)\cdots\left(1-\frac{k}{n} + O\left(\frac{1}{n}\right)\right) \\
&= \frac{k}{n}\left(1-\frac{k}{n}\right)^{j-1} + O\left(\frac{1}{n}\right) = \theta_n(1-\theta_n)^{j-1} + O\left(\frac{1}{n}\right),
\end{aligned}$$

which implies the claim after multiplication by k. □

Proof of Theorem 5.1 using Theorem 5.3. Recall that $Z_n^{(j)} \stackrel{\mathrm{d}}{=} \mathrm{Lah}(n,1)$, is an array of mutually independent random variables such that $Z_n^{(j)}$ has distribution (2.1), for $n, j \in \mathbb{N}$. Put

$$\widehat{Z}_n^{(j)} := \sum_{i=1}^{n} Z_j^{(i)}, \quad n, j \in \mathbb{N}.$$

Representation (2.3) is equivalent to the following

$$\mathrm{Lah}(n,k) \stackrel{d}{=} \sum_{j=1}^{n} \widehat{Z}^{(j)}_{N_j^{(n)}}.$$

By Donsker's theorem and using that $\mathbb{E}Z_j^{(i)} = H_j$ and $\mathrm{Var}(Z_j^{(i)}) = H_j - H_j^{(2)}$, for every $j \in \mathbb{N}$,

$$\left(\frac{\widehat{Z}^{(j)}_{\lfloor nt \rfloor} - H_j nt}{\sqrt{n}}\right)_{t\geq 0} \underset{n\to\infty}{\overset{d}{\Longrightarrow}} \left(\sqrt{H_j - H_j^{(2)}}B_j(t)\right)_{t\geq 0}, \tag{5.3}$$

in the Skorokhod space $D[0,\infty)$ endowed with the standard J_1-topology, where $B_1, B_2, \ldots$ are independent standard Brownian motions. Moreover, the convergences in (5.3) hold also mutually for all $j \in \mathbb{N}$ by independence. Combining this with Theorem 5.3 we obtain that, for every $M \in \mathbb{N}$,

$$\begin{aligned}
&\left(\left(\frac{\widehat{Z}^{(j)}_{\lfloor nt \rfloor} - H_j nt}{\sqrt{n}}\right)_{t\geq 0}, \frac{N_j^{(n)}}{n}, \frac{N_j^{(n)} - k\mathbb{P}[b_1^{(n)} = j]}{\sqrt{n}}\right)_{j=1,\ldots,M} \\
&\qquad \underset{n\to\infty}{\overset{d}{\Longrightarrow}} \left(\left(\sqrt{H_j - H_j^{(2)}}B_j(t)\right)_{t\geq 0}, \alpha^2(1-\alpha)^{j-1}, \sqrt{\alpha}\mathscr{N}_j\right)_{j=1,\ldots,M},
\end{aligned} \tag{5.4}$$

in the product topology on $(D[0,\infty)\times\mathbb{R}\times\mathbb{R})^M$, where we have also used that $k\sim\alpha n$ and $\lim_{n\to\infty}\mathbb{P}[b_1^{(n)}=j]=\alpha(1-\alpha)^{j-1}$.

Applying the mapping $D[0,\infty)\times\mathbb{R}\times\mathbb{R}\ni(f(\cdot),x,y)\mapsto f(x)+H_jy\in\mathbb{R}$ which is a.s. continuous at the point given by the right-hand side of (5.4), yields

$$\left(\frac{\widehat{Z}^{(j)}_{N_j^{(n)}}-kH_j\mathbb{P}[b_1^{(n)}=j]}{\sqrt{n}}\right)_{j=1,\dots,M}\overset{d}{\underset{n\to\infty}{\longrightarrow}}\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right)_{j=1,\dots,M}.$$

Summation over $j=1,\dots,M$ gives

$$\frac{\sum_{j=1}^M\widehat{Z}^{(j)}_{N_j^{(n)}}-k\sum_{j=1}^M H_j\mathbb{P}[b_1^{(n)}=j]}{\sqrt{n}}\overset{d}{\underset{n\to\infty}{\longrightarrow}}\sum_{j=1}^M\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right),\tag{5.5}$$

and this relation holds for every fixed $M\in\mathbb{N}$. Note that as $M\to\infty$, the right-hand side of (5.5) converges to

$$\sum_{j=1}^{\infty}\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right)$$

and this series converges almost surely because

$$\sum_{j=1}^{\infty}\left(\sqrt{H_j-H_j^{(2)}}\,\mathbb{E}\,|B_j(\alpha^2(1-\alpha)^{j-1})|+H_j\sqrt{\alpha}\,\mathbb{E}\,|\mathcal{N}_j|\right)<\infty.$$

According to Theorem 3.2 in [8] it remains to prove that for every fixed $\varepsilon>0$,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\left[\left|\sum_{j=M+1}^{n}\widehat{Z}^{(j)}_{N_j^{(n)}}-k\sum_{j=M+1}^{n}H_j\mathbb{P}[b_1^{(n)}=j]\right|>\varepsilon\sqrt{n}\right]=0.\tag{5.6}$$

By Markov's inequality, it is enough to check that

$$\lim_{M\to\infty}\limsup_{n\to\infty}\frac{\sum_{j=M+1}^{n}\mathbb{E}\left|\widehat{Z}^{(j)}_{N_j^{(n)}}-kH_j\mathbb{P}[b_1^{(n)}=j]\right|}{\sqrt{n}}=0.\tag{5.7}$$

In order to prove (5.7) we argue as follows. Using Wald's identity followed by the formula for the conditional variance, we derive, for all $j\in\mathbb{N}$,

$$\begin{aligned}\mathbb{E}\left|\widehat{Z}^{(j)}_{N_j^{(n)}}-kH_j\mathbb{P}[b_1^{(n)}=j]\right|&=\mathbb{E}\left|\widehat{Z}^{(j)}_{N_j^{(n)}}-\mathbb{E}\widehat{Z}^{(j)}_{N_j^{(n)}}\right|\le\left((\mathrm{Var}(\widehat{Z}^{(j)}_{N_j^{(n)}})\right)^{1/2}\\&=\left(\mathbb{E}\,\mathrm{Var}(\widehat{Z}^{(j)}_{N_j^{(n)}}|N_j^{(n)}))+\mathrm{Var}(\mathbb{E}(\widehat{Z}^{(j)}_{N_j^{(n)}}|N_j^{(n)}))\right)^{1/2}\le\left(\mathbb{E}H_{N_j^{(n)}}+\mathrm{Var}(H_{N_j^{(n)}})\right)^{1/2}\\&\le\left(\mathbb{E}H^2_{N_j^{(n)}}\right)^{1/2}\le\mathrm{const}\cdot(\mathbb{E}N_j^{(n)})^{1/2}\le\mathrm{const}\cdot\sqrt{n}\sqrt{\mathbb{P}[b_1^{(n)}=j]},\end{aligned}$$

where 'const' denotes absolute constants whose values are of no importance. It remains to note that

$$\begin{aligned}\lim_{M\to\infty}\limsup_{n\to\infty}\sum_{j=M+1}^{n}\sqrt{\mathbb{P}[b_1^{(n)}=j]}&=\lim_{M\to\infty}\limsup_{n\to\infty}\sum_{j=M+1}^{n-k+1}\sqrt{\frac{\binom{n-j-1}{k-2}}{\binom{n-1}{k-1}}}\\&=\lim_{M\to\infty}\limsup_{n\to\infty}\sum_{j=M+1}^{n-k+1}\sqrt{(k-1)\frac{(n-k-j+2)\cdots(n-k)}{(n-j)\cdots(n-1)}}=0,\end{aligned}$$

as readily follows from the inequality $\alpha n/2\le k\le n$ that holds for large enough n.

Combining (5.5) and (5.6) we obtain

$$\frac{\mathrm{Lah}(n,k)-\mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{n}} \xrightarrow[n\to\infty]{d} \sum_{j=1}^{\infty}\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right).$$

The limiting variable is obviously normal, has zero mean and its variance can be derived as follows:

$$\begin{aligned}
&\mathrm{Var}\left(\sum_{j=1}^{\infty}\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right)\right)\\
&\quad=\mathbb{E}\left(\sum_{j=1}^{\infty}\left(\sqrt{H_j-H_j^{(2)}}B_j(\alpha^2(1-\alpha)^{j-1})+H_j\sqrt{\alpha}\mathcal{N}_j\right)\right)^2\\
&\quad=\sum_{j=1}^{\infty}(H_j-H_j^{(2)})\alpha^2(1-\alpha)^{j-1}+\mathbb{E}\left(\sum_{j=1}^{\infty}H_j\sqrt{\alpha}\mathcal{N}_j\right)^2\\
&\quad=\sum_{j=1}^{\infty}(H_j-H_j^{(2)})\alpha^2(1-\alpha)^{j-1}+\alpha\sum_{i,j=1}^{\infty}H_iH_j\left(p_i\mathbb{1}_{\{i=j\}}-p_ip_j-\frac{p_ip_j}{1-\alpha}(\alpha i-1)(\alpha j-1)\right)\\
&\quad=\alpha\sum_{j=1}^{\infty}(H_j-H_j^{(2)})p_j+\alpha\left(\sum_{j=1}^{\infty}H_j^2p_j-\left(\sum_{j=1}^{\infty}H_jp_j\right)^2-\frac{1}{1-\alpha}\left(\sum_{j=1}^{\infty}H_jp_j(\alpha j-1)\right)^2\right),
\end{aligned}$$

where $p_j=\alpha(1-\alpha)^{j-1}$ for $j\in\mathbb{N}$. We only show how to deal with the sums containing $H_j^{(2)}$ and H_j^2, the other being elementary. We have

$$\begin{aligned}
\sum_{j=1}^{\infty}H_j^2p_j-\sum_{j=1}^{\infty}H_j^{(2)}p_j&=\sum_{j,k,l=1}^{\infty}\frac{1}{kl}\alpha(1-\alpha)^{j-1}\mathbb{1}_{\{k,l\le j\}}-\sum_{j,k=1}^{\infty}\frac{1}{k^2}\alpha(1-\alpha)^{j-1}\mathbb{1}_{\{k\le j\}}\\
&=\sum_{k,l=1}^{\infty}\frac{1}{kl}(1-\alpha)^{\max(k,l)-1}-\sum_{k=1}^{\infty}\frac{1}{k^2}(1-\alpha)^{k-1}\\
&=\sum_{k=1}^{\infty}\frac{1}{k}\left(\sum_{l=1}^{k-1}\frac{(1-\alpha)^{k-1}}{l}+\sum_{l=k}^{\infty}\frac{(1-\alpha)^{l-1}}{l}\right)-\sum_{k=1}^{\infty}\frac{1}{k^2}(1-\alpha)^{k-1}\\
&=\sum_{k=1}^{\infty}\frac{1}{k}H_{k-1}(1-\alpha)^{k-1}+\sum_{k=1}^{\infty}\frac{1}{k}H_k(1-\alpha)^{k-1}-\sum_{k=1}^{\infty}\frac{1}{k^2}(1-\alpha)^{k-1}\\
&=2\sum_{k=1}^{\infty}\frac{H_{k-1}}{k}(1-\alpha)^{k-1}=\frac{2}{1-\alpha}\int_{\alpha}^{1}\left(\sum_{k=1}^{\infty}H_{k-1}(1-t)^{k-1}\right)\mathrm{d}t\\
&=-\frac{2}{1-\alpha}\int_{\alpha}^{1}\frac{\log t}{t}\mathrm{d}t=\frac{\log^2\alpha}{1-\alpha}.
\end{aligned}$$

The proof is complete. □

5.2. **Cumulant generating function in the central regime.** In this and the following section we establish a large deviations principle under assumption (5.1). First we look at the cumulant generating function.

Proposition 5.6. *Assume that* (5.1) *holds. Then, for every fixed $t\in\mathbb{R}$, the limit*

$$\varphi_\alpha(t):=\lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}\,\mathrm{e}^{t\,\mathrm{Lah}(n,k)},\tag{5.8}$$

exists finitely, and, moreover, the function $\varphi_\alpha:\mathbb{R}\to\mathbb{R}$ is given by

$$\varphi_\alpha(t)=2\alpha\log\alpha+(1-\alpha)\log(1-\alpha)+(\alpha-1)\log r(t)+\alpha t-\alpha(\mathrm{e}^t+1)\log(1-r(t)),\tag{5.9}$$

where $r(t) = r(t;\alpha) \in (0,1)$ *is a unique non-zero solution to the equation*

$$(1-r(t))^{e^t+1} - 1 + r(t)(1+\alpha e^t) = 0. \tag{5.10}$$

Remark 5.7. Let us argue that Equation (5.10) has a unique solution $r(t) \in (0,1)$ for every $t \in \mathbb{R}$. For every fixed $\alpha \in (0,1)$ and $t \in \mathbb{R}$, the function $h(r) = h_{\alpha,t}(r) := (1-r)^{e^t+1} - 1 + r(1+\alpha e^t)$, defined for $r \in [0,1]$, is strictly convex since its first derivative

$$h'(r) = -(e^t+1)(1-r)^{e^t} + (1+\alpha e^t), \quad r \in [0,1],$$

strictly increases. Moreover, we have $h(0) = 0$ and $h'(0) < 0$. To complete the proof of the existence and the uniqueness of zero of h in the interval $(0,1)$, observe that $h(1) > 0$. Moreover, it follows from the above that $h'(r(t)) > 0$. By the implicit function theorem we conclude that $r(t)$ is differentiable on $(0,1)$. Clearly, $r(0;\alpha) = 1-\alpha$.

Proof of Proposition 5.6. Fix any $t \in \mathbb{R}$. By Lemma 3.1 we have

$$\mathbb{E}\, e^{t\,\mathrm{Lah}(n,k)} = P_{n,k}(e^t) = \frac{1}{\binom{n-1}{k-1}}[x^n]\left((1-x)^{-e^t} - 1\right)^k.$$

We shall now use the classical saddle-point method to derive the asymptotics of $[x^n]((1-x)^{-e^t}-1)^k$. With this respect Theorem VIII.8 in [24] perfectly fits our needs. Note that

$$[x^n]\left((1-x)^{-e^t} - 1\right)^k = [x^{n-k}]\left(\frac{(1-x)^{-e^t}-1}{x}\right)^k =: [x^{n-k}](B(x))^k,$$

and the function $x \mapsto B(x)$, for every fixed $t \in \mathbb{R}$, is analytic in the interior of the unit disk, has non-negative coefficients and $B(0) \neq 0$. Thus, by Theorem VIII.8 in [24] applied with $A(x) := 1$, $B(x) = x^{-1}((1-x)^{-e^t}-1)$, $N := n-k$, $n := k$, $\lambda = \lambda_n = n/k - 1$, we get[1]

$$[x^n]\left((1-x)^{-e^t} - 1\right)^k = (B(\zeta_n(t)))^k(\zeta_n((t)))^{-n+k-1}(2\pi k\widehat{\zeta}_n(t))^{-1/2}(1+o(1)), \quad n \to \infty, \tag{5.11}$$

where $\zeta_n(t)$ is the unique root of

$$\frac{\zeta_n(t)B'(\zeta_n(t))}{B(\zeta_n(t))} = \lambda_n, \tag{5.12}$$

and

$$\widehat{\zeta}_n(t) := \frac{d}{dz^2}\left(\log B(z) - \lambda_n \log z\right)\Big|_{z=\zeta_n(t)}.$$

Substituting $B(x) = x^{-1}((1-x)^{-e^t}-1)$ into (5.12) and simplifying, we obtain

$$\frac{e^t\zeta_n(t)}{1-\zeta_n(t)-(1-\zeta_n(t))^{e^t+1}} = \frac{n}{k}.$$

Sending $n \to \infty$ we see from the discussion in Remark 5.7 that

$$\lim_{n\to\infty}\zeta_n(t) = r(t),$$

where $r(t)$ is given by (5.10) and also that $\widehat{\zeta}_n(t)$ converges to a finite non-zero constant. Thus, taking logarithms in (5.11), dividing by n and sending $n \to \infty$, yields

$$\lim_{n\to\infty} n^{-1}\log\left([x^n]\left((1-x)^{-e^t}-1\right)^k\right) = \alpha\log B(r(t)) - (1-\alpha)\log r(t).$$

Combining this with

$$\lim_{n\to\infty}\frac{1}{n}\log\binom{n-1}{k-1} = -\alpha\log\alpha - (1-\alpha)\log(1-\alpha), \tag{5.13}$$

which follows from the Stirling formula, we obtain (5.9) after elementary manipulations. □

[1]Note that the claim of Theorem VIII.8 holds locally uniformly in λ.

Remark 5.8 (On linear growth of cumulants). Recall that the ℓ-th cumulant of a random variable X is defined as

$$\kappa_\ell(X) := D_t^\ell\Big|_{t=0}\left(\log \mathbb{E}\,\mathrm{e}^{tX}\right),$$

and we assume that X has finite exponential moments in a neighborhood of 0. Let, as before, $k = k(n) \sim \alpha n$ for some $\alpha \in (0,1)$. Refining the arguments used in the proof of Proposition 5.6, it is possible to show that the cumulants of $\mathrm{Lah}(n,k)$ grow linearly in the sense that

$$\lim_{n\to\infty} \frac{\kappa_\ell(\mathrm{Lah}(n,k))}{n} = \varphi_\alpha^{(\ell)}(0) \qquad \text{for all } \ell \in \mathbb{N}. \tag{5.14}$$

Indeed, this relation can be obtained from (5.8) by differentiating it $\ell \in \mathbb{N}$ times. To justify that the limit and the derivative can be interchanged, it suffices to check that the assertion of Proposition 5.6 continues to hold locally uniformly for complex t in a small neighborhood of $\mathbb{R}$. Then, since all involved functions are analytic, we can interchange D_t^ℓ and the large n limit. The validity of Proposition 5.6 in a small complex neighborhood of $\mathbb{R}$ follows essentially from the stability of the saddle point under small analytic perturbations of the parameter t. Making this argument rigorous (and in particular, checking the local uniformity of convergence) is standard but tedious, and we omit the details.

Taking $\ell = 1$ and $\ell = 2$ in (5.14) yields the formulae $\mathbb{E}\,\mathrm{Lah}(n,k) \sim \varphi_\alpha'(0)n$ and $\mathrm{Var}\,\mathrm{Lah}(n,k) \sim \varphi_\alpha''(0)n$, and the expression for $\varphi_\alpha(t)$ given in (5.9) yields, after lengthy calculations, that

$$\varphi_\alpha'(0) = -\frac{\alpha\log\alpha}{1-\alpha}, \qquad \varphi_\alpha''(0) = -\frac{\alpha}{1-\alpha} - \frac{\alpha(\alpha+1)\log\alpha}{(1-\alpha)^2} - \frac{\alpha^2\log^2\alpha}{(1-\alpha)^3}.$$

Both formulas agree with previously derived results: formula (3.9) for the expectation, and the variance of the limiting normal law in Theorem 5.1. Note that the linear growth of cumulants (5.14) implies the CLT by the method of moments giving an alternative way to prove Theorem 5.1.

5.3. **Large deviations in the central regime.** Recall that we assume (5.1). In the next theorem we shall state a large deviations principle (see [13]) for the Lah distribution which, in particular, implies that, as $n \to \infty$,

$$\mathbb{P}[\mathrm{Lah}(n,k) \le (\beta + o(1))n] = \exp\{-nI_\alpha(\beta) + o(n)\}, \qquad \text{if } \beta \in \left(\alpha, -\frac{\alpha\log\alpha}{1-\alpha}\right), \tag{5.15}$$

$$\mathbb{P}[\mathrm{Lah}(n,k) \ge (\beta + o(1))n] = \exp\{-nI_\alpha(\beta) + o(n)\}, \qquad \text{if } \beta \in \left(-\frac{\alpha\log\alpha}{1-\alpha}, 1\right) \tag{5.16}$$

for a rate function I_α which we shall explicitly identify.

Theorem 5.9 (LDP in the central regime)**.** *Assume* (5.1). *Then, the sequence of random variables* $\frac{1}{n}\mathrm{Lah}(n,k)$*,* $n \in \mathbb{N}$*, satisfies a large deviations principle with a convex rate function* $I_\alpha : [0,1] \to [0,\infty]$ *defined as follows. For* $\beta \in (\alpha, 1)$ *we have*

$$I_\alpha(\beta) = \sup_{t\in\mathbb{R}}(\beta t - \varphi_\alpha(t)) \tag{5.17}$$

$$= -\alpha\log\left(h^{-1}\left(\frac{\beta}{\alpha}\right) - 1\right) + \log\left(1 - h^{-1}(\beta)\right) + \beta\log\left(-\frac{\log h^{-1}(\frac{\beta}{\alpha})}{\log h^{-1}(\beta)}\right) - \alpha\log\alpha - (1-\alpha)\log(1-\alpha), \tag{5.18}$$

where $h^{-1}(\cdot)$ *is the inverse function of*

$$h(x) := \varphi_x'(0) = -\frac{x\log x}{1-x}, \quad x \in (0,1)\cup(1,\infty), \quad h(0) = 0, \quad h(1) := 1. \tag{5.19}$$

For $\beta < \alpha$*, we have* $I_\alpha(\beta) = +\infty$*. Finally, at the boundary points* $\beta = \alpha$ *and* $\beta = 1$ *we have* $I_\alpha(\alpha) = \lim_{\beta\downarrow\alpha} I_\alpha(\beta)$ *and* $I_\alpha(1) = \lim_{\beta\uparrow 1} I_\alpha(\beta)$*, namely*

$$I_\alpha(\alpha) = \log\left(1 - h^{-1}(\alpha)\right) - \alpha\log\left(-\log h^{-1}(\alpha)\right) - \alpha\log\alpha - (1-\alpha)\log(1-\alpha) < +\infty, \tag{5.20}$$

$$I_\alpha(1) = -\alpha\log\left(h^{-1}\left(\frac{1}{\alpha}\right) - 1\right) + \log\log h^{-1}\left(\frac{1}{\alpha}\right) - \alpha\log\alpha - (1-\alpha)\log(1-\alpha) < +\infty. \tag{5.21}$$

Remark 5.10. Using the relation $h(1/x) = h(x)/x$ together with (5.18) one easily verifies that the function $I_\alpha(\beta)$ vanishes at $\beta = h(\alpha) = -\frac{\alpha\log\alpha}{1-\alpha}$, which is not surprising in view of the fact that $\mathbb{E}\,\mathrm{Lah}(n,k) \sim h(\alpha)n$ by Theorem 3.7. A lengthy computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}\beta} I_\alpha(\beta) = \log\left(-\frac{\log h^{-1}(\frac{\beta}{\alpha})}{\log h^{-1}(\beta)}\right), \qquad \beta \in (\alpha, 1).$$

One then verifies that the derivative of $I_\alpha(\beta)$ vanishes at $\beta = h(\alpha)$, is positive for $\beta > h(\alpha)$ and negative for $\beta < h(\alpha)$. Hence, $I_\alpha(\beta) > 0$ for all $\beta \neq h(\alpha)$.

Proof of Theorem 5.9. By Proposition 5.6, $\varphi_\alpha(t) := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\,\mathrm{e}^{t\,\mathrm{Lah}(n,k)}$ exists finitely for every $t \in \mathbb{R}$. Moreover, it is a differentiable function of t; see Remark 5.7. A large deviation principle with a rate function given by (5.17) is then implied by the Gärtner-Ellis theorem; see Theorem 2.3.6 and Exercise 2.3.20 of [13]. Note also that (5.17) implies that I_α is convex.

To prove (5.18), one can use the asymptotics of the Stirling numbers $\left[{n \atop k}\right]$ and $\left\{{n \atop k}\right\}$ in the central regime $k \sim \alpha n$, which is known from the works of Moser and Wyman [56, 57]; see also [59, Sections 3.6,3.7] and [24, Section VIII.8.2]. A review of these and other asymptotic regimes of k can be found in [53] and [54]. For our purposes the most convenient reference is [64]. In particular, when $k \sim \alpha n$ for some $\alpha \in (0,1)$, we have by formulas (14) and (33) in [64], respectively,

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\frac{k!}{n!}\begin{bmatrix} n \\ k \end{bmatrix}\right) = -\log(1 - h^{-1}(\alpha)) + \alpha\log(-\log h^{-1}(\alpha)),$$
$$\lim_{n\to\infty} \frac{1}{n} \log\left(\frac{k!}{n!}\begin{Bmatrix} n \\ k \end{Bmatrix}\right) = \alpha\log(h^{-1}(1/\alpha) - 1) - \log\log h^{-1}(1/\alpha),$$

where h^{-1} is the inverse of the function h defined in (5.19). Note that h has strictly positive derivative and satisfies $\lim_{x\downarrow 0} h(x) = 0$ as well as $\lim_{x\uparrow+\infty} h(x) = +\infty$, which implies that h^{-1} is well-defined and strictly increasing. Note also that $h(1) = 1$ (by continuity), hence $h^{-1}(1) = 1$ and, consequently, $h^{-1}(\alpha) < 1 < h^{-1}(1/\alpha)$. Combining these relations with (5.13), one obtains in the regime when $k \sim \alpha n$ and $j \sim \beta n$ with $0 < \alpha < \beta < 1$ that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}[\mathrm{Lah}(n,k) = j] = \lim_{n\to\infty} \frac{1}{n} \log\left(\frac{1}{\binom{n-1}{k-1}} \cdot \frac{j!}{n!}\begin{bmatrix} n \\ j \end{bmatrix} \cdot \frac{k!}{j!}\begin{Bmatrix} j \\ k \end{Bmatrix}\right)$$
$$= \alpha\log\left(h^{-1}\left(\frac{\beta}{\alpha}\right) - 1\right) - \log\left(1 - h^{-1}(\beta)\right) - \beta\log\left(-\frac{\log h^{-1}(\frac{\beta}{\alpha})}{\log h^{-1}(\beta)}\right) + \alpha\log\alpha + (1-\alpha)\log(1-\alpha). \quad (5.22)$$

Given this, an LDP with a rate function given by (5.18) follows by standard arguments. Namely, by [13, Theorem 4.1.11] it suffices to check that for all $t \in [\alpha, 1]$ we have

$$\inf_{\varepsilon>0} \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\frac{1}{n}\mathrm{Lah}(n,k) \in [t-\varepsilon, t+\varepsilon]\right] \le -I_\alpha(t), \quad (5.23)$$
$$\inf_{\varepsilon>0} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left[\frac{1}{n}\mathrm{Lah}(n,k) \in [t-\varepsilon, t+\varepsilon]\right] \ge -I_\alpha(t), \quad (5.24)$$

where I_α is defined by (5.18), (5.20) and (5.21). For $t \in (\alpha, 1)$ both claims follow immediately from (5.22) together with the union bound. Let us treat the boundary case $t = \alpha$, the case $t = 1$ being similar. One easily checks that the limit of the right-hand side of (5.22) as $\beta \downarrow \alpha$ coincides with $-I_\alpha(\alpha)$ as defined in (5.20). The unimodality together with the union bound and (5.22) with $\beta = \alpha + \varepsilon$ imply (5.23). To prove (5.24) use (5.22) with $\beta = \alpha + \varepsilon/2$ and let $\varepsilon \downarrow 0$. □

Remark 5.11. The function h^{-1} appearing above can be expressed as $h^{-1}(z) = z/v(z)$, for $z > 0$, where $v(z) > 0$ is the solution to the equation $v\mathrm{e}^{-v} = z\mathrm{e}^{-z}$ which is different from $v = z$ (and for $z = 1$, where there is only one solution, we put $v(1) = 1$). Using the standard notation for the branches of the Lambert W-function we have $h^{-1}(z) = -z/W_0(-z\mathrm{e}^{-z})$ for $z \ge 1$ and $h^{-1}(z) = -z/W_{-1}(-z\mathrm{e}^{-z})$ for $0 < z \le 1$.

5.4. **The intermediate regime.** In this section we prove a weak law of large numbers for the Lah distribution in the intermediate regime, that is, when $k = k(n)$ satisfies

$$\lim_{n\to\infty} \frac{k(n)}{n} = 0 \quad \text{but} \quad \lim_{n\to\infty} k(n) = \infty. \tag{5.25}$$

Recall from Theorem 3.7 that $\mathbb{E}\,\mathrm{Lah}(n,k) \sim k\log(n/k)$ in this regime.

Theorem 5.12 (Weak LLN in the intermediate regime)**.** *Under* (5.25) *we have*

$$\lim_{n\to\infty} \frac{\mathrm{Var}\,\mathrm{Lah}(n,k)}{\mathbb{E}\,\mathrm{Lah}(n,k)} = 1. \tag{5.26}$$

Consequently, the following weak law of large numbers holds in L^2 and hence in probability:

$$\frac{\mathrm{Lah}(n,k)}{k\log(n/k)} \underset{n\to\infty}{\longrightarrow} 1.$$

Proof. It is enough to prove (5.26) because the weak LLN follows from (5.26) by Chebyshev's inequality. Recalling the representation (2.3), conditioning on the random uniform composition $(b_1^{(n)},\dots,b_k^{(n)})$ and using the formula of the total variance, we can write

$$\begin{aligned}
\mathrm{Var}\,\mathrm{Lah}(n,k) &= \mathrm{Var}\,\mathbb{E}\left[\sum_{j=1}^k Z^{(j)}_{b_j^{(n)}} \,\Big|\, (b_1^{(n)},\dots,b_k^{(n)})\right] + \mathbb{E}\,\mathrm{Var}\left[\sum_{j=1}^k Z^{(j)}_{b_j^{(n)}} \,\Big|\, (b_1^{(n)},\dots,b_k^{(n)})\right]\\
&= \mathrm{Var}\left[\sum_{j=1}^k H_{b_j^{(n)}}\right] + \mathbb{E}\left[\sum_{j=1}^k \left(H_{b_j^{(n)}} - H^{(2)}_{b_j^{(n)}}\right)\right]\\
&\le \mathrm{Var}\left[\sum_{j=1}^k H_{b_j^{(n)}}\right] + \mathbb{E}\left[\sum_{j=1}^k H_{b_j^{(n)}}\right],
\end{aligned} \tag{5.27}$$

where we recall that $H_N = \sum_{\ell=1}^N \frac{1}{\ell}$ is the N-th harmonic number and $H_N^{(2)} = \sum_{\ell=1}^N \frac{1}{\ell^2}$. Again according to (2.3), the second term on the right-hand side is nothing else but $\mathbb{E}\,\mathrm{Lah}(n,k)$. Note in passing that (5.27) together with the estimate $H_N^{(2)} < \pi^2/6$ yields

$$\mathrm{Var}\,\mathrm{Lah}(n,k) \ge \mathbb{E}\,\mathrm{Lah}(n,k) - \pi^2 k/6 \sim k\log(n/k), \qquad n\to\infty,$$

which proves the lower bound in (5.26). To complete the proof of (5.26), it suffices to check that

$$\mathrm{Var}\,H_{b_1^{(n)}} = O(1), \qquad \text{as } n\to\infty. \tag{5.28}$$

and

$$\mathrm{Cov}\left(H_{b_i^{(n)}}, H_{b_j^{(n)}}\right) \le 0, \qquad \text{for all } 1\le i<j\le k. \tag{5.29}$$

Indeed, then we have

$$\mathrm{Var}\left[\sum_{j=1}^k H_{b_j^{(n)}}\right] = \sum_{j=1}^k \mathrm{Var}\,H_{b_j^{(n)}} + 2\sum_{1\le i<j\le k} \mathrm{Cov}\left(H_{b_i^{(n)}}, H_{b_j^{(n)}}\right) \le \sum_{j=1}^k \mathrm{Var}\,H_{b_j^{(n)}} = O(k).$$

Let us prove (5.28) first. The main step is the following lemma.

Lemma 5.13. *Under assumption* (5.25)*, the random variables $(k/n)b_1^{(n)}$ converge in distribution, as $n\to\infty$, to the standard exponential law Exp(1).*

Proof. Using formula (3.8) and the hockey-stick identity we obtain

$$\mathbb{P}[b_1^{(n)} \ge j] = \frac{\sum_{m=j}^{n-k+1}\binom{n-m-1}{k-2}}{\binom{n-1}{k-1}} = \frac{\sum_{m=k-2}^{n-j-1}\binom{m}{k-2}}{\binom{n-1}{k-1}} = \frac{\binom{n-j}{k-1}}{\binom{n-1}{k-1}}, \qquad j = 1,\dots,n-k+1.$$

For all $x>0$ it follows that

$$\mathbb{P}[(k/n)b_1^{(n)} \geq x] = \mathbb{P}[b_1^{(n)} \geq nx/k] = \mathbb{P}[b_1^{(n)} \geq \lceil nx/k \rceil] = \frac{\binom{n-j_n}{k-1}}{\binom{n-1}{k-1}} = \frac{(n-j_n-k+2)\dots(n-j_n)}{(n-k+1)\dots(n-1)}$$

with $j_n = \lceil nx/k \rceil \sim nx/k$. Taking the logarithm and using that $\log(1-x) = -x + O(x^2)$ as $x \to 0$, we obtain

$$\begin{aligned}\log \mathbb{P}[(k/n)b_1^{(n)} \geq x] &= \sum_{\ell=1}^{k-1} \log\left(1 - \frac{j_n - 1}{n-k+\ell}\right) = -\sum_{\ell=1}^{k-1}\left(\frac{j_n-1}{n-k+\ell} + O\left(\frac{j_n^2}{(n-k)^2}\right)\right)\\ &= -(j_n - 1)\sum_{i=n-k+1}^{n-1} \frac{1}{i} + O\left(\frac{kj_n^2}{(n-k)^2}\right) = x + o(1),\end{aligned}$$

and the proof is complete. □

We can now prove (5.28) as follows. By the Skorokhod representation theorem, we may pass to a different probability space and, after taking the logarithm, write Lemma 5.13 in the form

$$\log b_1^{(n)} - \log(n/k) \xrightarrow[n\to\infty]{a.s.} G, \tag{5.30}$$

where $G := \log \mathrm{Exp}(1)$. This already suggests that $\mathrm{Var} \log b_1^{(n)}$ should be of order $O(1)$. We shall now justify this by a uniform integrability argument. We claim that for every $p \geq 1$,

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left|\log b_1^{(n)} - \log(n/k)\right|^p < \infty. \tag{5.31}$$

To prove this it suffices to check that

$$\sup_{n\in\mathbb{N}} \mathbb{E}\log_+^p((k/n)b_1^{(n)})) < \infty, \qquad \sup_{n\in\mathbb{N}} \mathbb{E}\log_-^p((k/n)b_1^{(n)})) < \infty, \tag{5.32}$$

where for $x>0$ we defined $\log_+(x) = \max(\log x, 0)$ and $\log_-(x) = \max\{-\log x, 0\}$, so that $|\log x| = \log_+ x + \log_- x$. The first claim in (5.32) follows from the estimate $\log_+^p x = O(x)$ together with the identity $\mathbb{E}[(k/n)b_1^{(n)}] = 1$ which holds by exchangeability. To prove the second claim in (5.32), we first observe that $\mathbb{P}[b_1^{(n)} = j]$ is a decreasing function of j, which follows from the explicit formula (3.8). Hence,

$$\mathbb{E}\log_-^p((k/n)b_1^{(n)})) = \sum_{j=1}^{\lfloor n/k \rfloor} \log_-^p(jk/n)\,\mathbb{P}[b_1^{(n)} = j] \leq \mathbb{P}[b_1^{(n)} = 1] \sum_{j=1}^{\lfloor n/k \rfloor} |\log(jk/n)|^p = \frac{k-1}{n-1}\sum_{j=1}^{\lfloor n/k \rfloor} |\log(jk/n)|^p,$$

which is bounded as a Riemann sum for $\int_0^1 |\log x|^p \mathrm{d}x < \infty$. This completes the proof of (5.31).

Relation (5.30), combined with the uniform integrability established in (5.31), implies that

$$\mathbb{E}|\log b_1^{(n)} - \log(n/k)| = O(1), \qquad \mathbb{E}(\log b_1^{(n)} - \log(n/k))^2 = O(1).$$

Using the fact that $|\log b_1^{(n)} - H_{b_1^{(n)}}|$ is bounded by a non-random constant, using the triangle inequality and the inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we get

$$\begin{aligned}&\mathbb{E}\left|H_{b_1^{(n)}} - \log(n/k)\right| \leq \mathbb{E}|\log b_1^{(n)} - \log(n/k)| + O(1) = O(1),\\ &\mathbb{E}\left(H_{b_1^{(n)}} - \log(n/k)\right)^2 \leq 2\,\mathbb{E}\left(H_{b_1^{(n)}} - \log b_1^{(n)}\right)^2 + 2\,\mathbb{E}(\log b_1^{(n)} - \log(n/k))^2 = O(1).\end{aligned}$$

Applying the triangle inequality to the first relation and expanding the square in the second relation yields

$$\mathbb{E}H_{b_1^{(n)}} = \log(n/k) + O(1), \qquad \mathbb{E}H^2_{b_1^{(n)}} = 2\log(n/k)\,\mathbb{E}H_{b_1^{(n)}} - \log^2(n/k) + O(1).$$

It follows that

$$\mathrm{Var}\,H_{b_1^{(n)}} = \mathbb{E}H^2_{b_1^{(n)}} - (\mathbb{E}H_{b_1^{(n)}})^2 = 2\log(n/k)\,\mathbb{E}H_{b_1^{(n)}} - \log^2(n/k) - (\mathbb{E}H_{b_1^{(n)}})^2 + O(1)$$

$$= -(\log(n/k) - \mathbb{E} H_{b_1^{(n)}})^2 + O(1) = O(1).$$

We now proceed to the proof of (5.29). First we need to recall the notion of negative association; see [9]. For $\ell \in \mathbb{N}$ let $\mathscr{M}(\ell)$ denote the set of all real-valued, bounded, Borel functions on $\mathbb{R}^\ell$ that are nondecreasing in each coordinate. A random vector $(X_1,\dots,X_k)$ is called *negatively associated* if for every disjoint sets $I,J \subset \{1,\dots,k\}$ and every functions $f \in \mathscr{M}(|I|)$, $g \in \mathscr{M}(|J|)$ it holds that

$$\operatorname{Cov}(f(X_i : i \in I), g(X_j : j \in J)) \le 0. \tag{5.33}$$

Although the next lemma, claiming the negative association of the uniform random compositions, sounds classical and will be proved by standard methods, we did not find it among the numerous similar examples listed in [9] and [38].

Lemma 5.14. *For every $n \in \mathbb{N}$ and $k \in \{1,\dots,n\}$, the random uniform composition $(b_1^{(n)},\dots,b_k^{(n)})$ is negatively associated.*

Proof. Recall that $(b_1^{(n)},\dots,b_k^{(n)})$ has the same distribution as the vector $(G_1,\dots,G_k)$ of i.i.d. geometric variables with arbitrary parameter $\theta \in (0,1)$ conditioned on the event $\{G_1+\dots+G_k = n\}$, see equation (5.2). According to Theorem 1.23 of [9] or Theorem 2.6 of [38], to prove negative association, it suffices to check that for every set $I \subset \{1,\dots,n\}$ and for every function $f \in \mathscr{M}(I)$ the function

$$n \mapsto \mathbb{E}\left[f(G_i, i \in I) \,\Big|\, \sum_{i \in I} G_i = n \right]$$

is nondecreasing in n. There is no loss of generality in assuming that $I = \{1,\dots,k\}$, so that our task reduces to proving that

$$n \mapsto \mathbb{E}\left[f(b_1^{(n)},\dots,b_k^{(n)}) \right]$$

is nondecreasing in n. To prove this, it suffices to construct the vectors $(b_1^{(n)},\dots,b_k^{(n)})$, with $n = k, k+1, \dots$, on a common probability space in such a way that $b_j^{(n)} \le b_j^{(n+1)}$ for all $j = 1,\dots,k$ and $n \ge k$. But such a coupling using the Pólya urn has already been constructed in Section 2.2. Clearly, in that coupling the number of balls of each color is nondecreasing in time, and the proof is complete. □

To complete the proof of (5.29), and thus of Theorem 5.12, it remains to observe that the function $x \mapsto H_{\lfloor x \rfloor}$, $x \in [1,n]$, is bounded and nondecreasing, which allows us to apply (5.33) with $I = \{i\}$ and $J = \{j\}$. □

Remark 5.15. It is natural to conjecture that in the intermediate regime, all cumulants are asymptotically equivalent to the cumulants of the Poisson distribution with parameter $k \log(n/k)$, meaning that $\kappa_\ell(\mathrm{Lah}(n,k)) \sim k \log(n/k)$ for all $\ell \in \mathbb{N}$. This statement would imply a CLT in the intermediate regime. More generally, we conjecture that

$$\frac{\mathrm{Lah}(n,k) - \mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{\operatorname{Var}\mathrm{Lah}(n,k)}} \xrightarrow[n\to\infty]{d} N(0,1) \;\Leftrightarrow\; \operatorname{Var}\mathrm{Lah}(n,k) \underset{n\to\infty}{\longrightarrow} \infty \;\Leftrightarrow\; n - k(n) \underset{n\to\infty}{\longrightarrow} \infty.$$

Since the existing results are sufficient to obtain a fairly complete picture of the threshold phenomena in the following Section 6, we refrain from studying these more technical questions here.

6. Threshold phenomena for convex hulls of random walks

6.1. **Formula for the expected number of faces.** Let $\xi_1,\dots,\xi_n$ be a collection of possibly dependent random d-dimensional vectors with partial sums

$$S_i = \xi_1 + \dots + \xi_i, \quad 1 \le i \le n, \quad S_0 = 0.$$

The sequence $S_0, S_1, \dots, S_n$ will be referred to as a *random walk*. We impose the following assumptions on the joint distribution of the increments.

(Ex) *Exchangeability:* For every permutation σ of the set $\{1,\dots,n\}$ we have the following distributional equality of joint distributions:

$$(\xi_{\sigma(1)},\dots,\xi_{\sigma(n)}) \stackrel{d}{=} (\xi_1,\dots,\xi_n).$$

(GP) *General position:* For every $1 \le i_1 < \dots < i_d \le n$ the probability that the vectors $S_{i_1}, \dots, S_{i_d}$ are linearly dependent is 0.

For example, it is known from [43, Proposition 2.5] and [42, Example 1.1] that Conditions (Ex) and (GP) are satisfied if $\xi_1, \dots, \xi_n$ are independent identically distributed, and for every hyperplane $H_0 \subset \mathbb{R}^d$ passing through the origin we have $\mathbb{P}[S_i \in H_0] = 0$, for all $1 \le i \le n$. Moreover, the second condition can be replaced by $\mathbb{P}[\xi_1 \in H] = 0$ for every affine hyperplane $H \subset \mathbb{R}^d$.

We are interested in the convex hull of this random walk which is a random polytope $C_{n,d} = \operatorname{conv}(S_0, S_1, \dots, S_n) \subset \mathbb{R}^d$ defined by (1.4). For $\ell \in \{0, \dots, d\}$, the number of ℓ-dimensional faces of the polytope $C_{n,d}$ is denoted by $f_\ell(C_{n,d})$. As has already been mentioned in the introduction, see formula (1.5), the following explicit formula has been obtained in [42].

Theorem 6.1 (Exact formula for expected face numbers)**.** *Let $(S_i)_{i=0}^\infty$ be a random walk in $\mathbb{R}^d$, whose increments satisfy conditions (Ex) and (GP). Then, for $n \ge d$ and all $k \in \{1, \dots, d\}$,*

$$\mathbb{E} f_{k-1}(C_{n,d}) = \frac{2 \cdot (k-1)!}{n!} \sum_{l=0}^\infty \begin{bmatrix} n+1 \\ d-2l \end{bmatrix} \begin{Bmatrix} d-2l \\ k \end{Bmatrix}. \tag{6.1}$$

In the following it will be more convenient to consider the polytope $C_{n-1,d}$ since it is defined as a convex hull of n points. It is known [42, Remark 1.5] that this polytope is simplicial with probability 1, that is, all its facets are simplices. Therefore, the maximal possible number of $(k-1)$-dimensional faces of $C_{n-1,d}$ is $\binom{n}{k}$. If the number of $(k-1)$-faces attains this bound, the polytope $C_{n-1,d}$ is said to be $(k-1)$-neighborly. We shall therefore be interested in the asymptotic behavior of the quantity $f_{k-1}(C_{n-1,d})/\binom{n}{k}$. The formula of Theorem 6.1 can be stated as follows:

$$\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = \frac{2}{L(n,k)} \sum_{l=0}^\infty \begin{bmatrix} n \\ d-2l \end{bmatrix} \begin{Bmatrix} d-2l \\ k \end{Bmatrix} = 2\mathbb{P}[\operatorname{Lah}(n,k) \in \{d, d-2, d-4, \dots\}]. \tag{6.2}$$

It follows from the well-known identities

$$\sum_{j=k}^n (-1)^{n-j} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = 0 \quad (\text{if } n > k), \qquad \sum_{j=k}^n \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = L(n,k)$$

that

$$\mathbb{P}[\operatorname{Lah}(n,k) \text{ takes even value}] = \mathbb{P}[\operatorname{Lah}(n,k) \text{ takes odd value}] = \frac{1}{2} \quad \text{for } n > k.$$

Consequently, we can rewrite (6.2) as follows:

$$1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = 2\mathbb{P}[\operatorname{Lah}(n,k) \in \{d+2, d+4, \dots\}]. \tag{6.3}$$

In the rest of this section we shall use (6.2) and (6.3) to uncover threshold phenomena for convex hulls of random walks. To describe our problem, let us fix some very large dimension d. Let us also take some $k \in \{1, \dots, d\}$ which may be either fixed, or depend on d in some way. We ask whether the number of $(k-1)$-dimensional faces of $C_{n-1,d-1}$ is equal or close to the maximal possible number $\binom{n}{k}$. If n is not much larger than d, we expect $f_{k-1}(C_{n-1,d-1})$ to be close or even equal to $\binom{n}{k}$. On the other hand, if n is sufficiently large, we expect $f_{k-1}(C_{n-1,d-1})/\binom{n}{k}$ to approach 0. Somewhere in between there should be a threshold at which a phase transition occurs. Following Donoho and Tanner [15, 16, 18, 19] we distinguish between weak and strong thresholds, which are statements about $\mathbb{E} f_{k-1}(C_{n-1,d-1})/\binom{n}{k}$ and $\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d-1})$, respectively. As we shall see in the following, the phase transitions occur surprisingly late. For example, for fixed k the weak transition occurs if n is near $\mathrm{e}^{d/k}$.

6.2. **Threshold phenomena for face numbers: the regime of constant k.** We begin by analyzing the case in which k is constant.

Theorem 6.2 (Weak threshold in the constant k regime)**.** *Let $d \to \infty$ and $n = n(d)$ be a function of d such that*

$$\gamma := \lim_{n \to \infty} \frac{\log n(d)}{d} \in [0, +\infty].$$

Then, for every fixed $k\in\mathbb{N}$, we have

$$\lim_{n\to\infty}\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}=\begin{cases}1, & \text{if } \gamma<1/k,\\ 0, & \text{if } \gamma>1/k.\end{cases}$$

Moreover, in the critical case when $\gamma=1/k$, more precisely if $\log n(d)=\frac{1}{k}(d+c\sqrt{d}+o(\sqrt{d}))$ for some fixed $k\in\mathbb{N}$ and some constant $c\in\mathbb{R}$, then

$$\lim_{n\to\infty}\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-c}\mathrm{e}^{-x^2/2}\mathrm{d}x.$$

Proof. Consider first the case $\gamma>1/k$. By (6.2) we have

$$\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\le 2\mathbb{P}[\mathrm{Lah}(n,k)\le d]=2\mathbb{P}[\mathrm{Lah}(n,k)\le kx_n\log n]$$

with $x_n:=d/(k\log n)\to 1/(\gamma k)<1$ as $d\to\infty$. By Theorem 4.4 (or Theorem 4.2), the probability on the right-hand side goes to 0. Let now $\gamma<1/k$. Then, by (6.3),

$$1-\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\le 2\mathbb{P}[\mathrm{Lah}(n,k)\ge d]=2\mathbb{P}[\mathrm{Lah}(n,k)\ge kx_n\log n]$$

with $x_n:=d/(k\log n)\to 1/(\gamma k)>1$. By Theorem 4.4, the probability on the right-hand side goes to 0.

Consider now the critical case, that is, let $\log n=\frac{1}{k}(d+c\sqrt{d}+o(\sqrt{d}))$. If the right-hand side of (6.2) could be replaced by the simpler quantity $\mathbb{P}[\mathrm{Lah}(n,k)\le d]$, the claim could be deduced from the central limit theorem as follows:

$$\begin{aligned}\mathbb{P}[\mathrm{Lah}(n,k)\le d]&=\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)-k\log n}{\sqrt{k\log n}}\le\frac{d-k\log n}{\sqrt{k\log n}}\right]\\&=\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)-k\log n}{\sqrt{k\log n}}\le -c+o(1)\right]\\&\underset{d\to\infty}{\longrightarrow}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-c}\mathrm{e}^{-x^2/2}\mathrm{d}x\end{aligned}\tag{6.4}$$

by Theorem 4.2. Unfortunately, the right-hand of (6.2) runs over $d-\ell$ with even $\ell=0,2,\dots$ (and there is a factor 2 compensating for that), and more efforts are needed to prove the claim. Recall that the Lah distribution is unimodal by Corollary 3.10. If $m_{n,k}$ denotes the largest mode of $\mathrm{Lah}(n,k)$, then by Proposition 4.5 we have

$$m_{n,k}=k\log n+O(1)=d+c\sqrt{d}+o(\sqrt{d}).$$

If $c>0$, then it follows that $d<m_{n,k}$ for sufficiently large n and hence

$$\mathbb{P}[\mathrm{Lah}(n,k)\le d-1]\le 2\mathbb{P}[\mathrm{Lah}(n,k)\in\{d,d-2,d-4,\dots\}]\le\mathbb{P}[\mathrm{Lah}(n,k)\le d+1].$$

Applying the CLT to both sides as explained in (6.4) completes the proof for $c>0$. If $c<0$, we can pass to the complementary events via (6.3) and argue analogously. Finally, the case $c=0$ follows by a sandwich argument. □

The above Theorem 6.2 deals with expected face numbers only. More interesting is to prove that neighborliness holds with high probability rather than only in expectation. The next result is a strong threshold in the terminology of Donoho and Tanner [15, 16, 18, 19].

Theorem 6.3 (Strong threshold in the constant k regime)**.** *Fix $k\in\mathbb{N}$. Let $n=n(d)$ be an integer sequence. If $n(d)\le \mathrm{e}^{d/(k(\mathrm{e}+\varepsilon))}$ for some $\varepsilon>0$ and all sufficiently large d, then*

$$\binom{n}{k}-\mathbb{E} f_{k-1}(C_{n-1,d})=O(n^{-\delta})\tag{6.5}$$

for some $\delta>0$, and the polytope $C_{n-1,d}$ is $(k-1)$-neighborly with probability approaching 1, more precisely

$$\mathbb{P}\left[f_{k-1}(C_{n-1,d})=\binom{n}{k}\right]=1-O(n^{-\delta}).\tag{6.6}$$

If, on the other hand, $n(d) \geq \mathrm{e}^{d/(k(\mathrm{e}-\varepsilon))}$ for some $\varepsilon \in (0,\mathrm{e})$ and all sufficiently large d, then

$$\lim_{n\to\infty}\left(\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d})\right) = +\infty. \tag{6.7}$$

Proof. Let us prove (6.5) and (6.6). Since on the event $f_{k-1}(C_{n-1,d}) \neq \binom{n}{k}$ we even have $\binom{n}{k} - f_{k-1}(C_{n-1,d}) \geq 1$, the following estimate holds:

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) = \mathbb{E}\left[\binom{n}{k} - f_{k-1}(C_{n-1,d})\right] \geq \mathbb{P}\left[f_{k-1}(C_{n-1,d}) \neq \binom{n}{k}\right].$$

It will be convenient to write this inequality in the form

$$\mathbb{P}\left[f_{k-1}(C_{n-1,d}) \neq \binom{n}{k}\right] \leq \binom{n}{k}\left(1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\right). \tag{6.8}$$

The same argumentation as in the proof of Theorem 6.2 yields then

$$\mathbb{P}\left[f_{k-1}(C_{n-1,d}) \neq \binom{n}{k}\right] \leq \binom{n}{k}\left(1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\right) \leq 2n^k \mathbb{P}[\mathrm{Lah}(n,k) \geq k x_n \log n],$$

where $x_n := d/(k\log n) \geq \mathrm{e}+\varepsilon$. Note that the convex function $f(x) := x\log x - x + 1$ satisfies $f(\mathrm{e}) = 1$ and $f(\mathrm{e}+\varepsilon) > 1$. It follows from Theorem 4.4 that

$$\mathbb{P}\left[f_{k-1}(C_{n-1,d}) \neq \binom{n}{k}\right] \leq 2n^k n^{-kf(\mathrm{e}+\varepsilon)+o(1)} = 2n^{k(1-f(\mathrm{e}+\varepsilon))+o(1)} = O(n^{-\delta}),$$

which proves (6.5) and (6.6). To prove (6.7), we assume that $n \geq \mathrm{e}^{d/(k(\mathrm{e}-\varepsilon))}$. Making ε smaller, if necessary, we may assume that $\varepsilon \in (0, \mathrm{e}-1)$. By (6.3), we have

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) = \binom{n}{k}\left(1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\right) = 2\binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k) \in \{d+2, d+4, \dots\}]. \tag{6.9}$$

Let $m_{n,k}$ be the largest mode of $\mathrm{Lah}(n,k)$, then $m_{n,k} = k\log n + O(1)$ by Proposition 4.5. Fix some $\delta \in (0, \mathrm{e}-\varepsilon-1)$. Without loss of generality we may assume that $d > (1+\delta)m_{n,k}$ for all n large enough. Indeed, if $d \leq (1+\delta)m_{n,k}$ along some subsequence, then we may increase d by an even number without destroying the condition $n \geq \mathrm{e}^{d/(k(\mathrm{e}-\varepsilon))}$ and such that d becomes larger than $(1+\delta)m_{n,k}$. Since this operation decreases the right-hand side of (6.9) and we intend to show that it diverges to infinity, we can and do assume that $d > (1+\delta)m_{n,k}$ for all n large enough. Using the unimodality of the Lah distribution, we have

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) \geq \binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k) \geq d+2] \geq cn^k \mathbb{P}[\mathrm{Lah}(n,k) \geq k x_n \log n],$$

where $c > 0$ is sufficiently small and $x_n := d/(k\log n)$ has all its limit points in $[1+\delta, \mathrm{e}-\varepsilon]$. Then, Theorem 4.4 yields (6.7). □

Let us finally mention a conjecture which we verified numerically for all $d \leq 50$, $1 \leq k \leq d$, $n \leq k+100$. Its part (b) is quite surprising in view of (6.3) and Proposition 2.4.

Conjecture 6.4. *Fix $d \in \mathbb{N}$ and $k \in \{1, \dots, d\}$. Then:*

(a) *The function $n \mapsto \mathbb{E} f_{k-1}(C_{n-1,d})$ is increasing for $n \geq k$.*
(b) *The function $n \mapsto \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}$ is decreasing (if $d-k$ is even) and increasing (if $d-k$ is odd), for all $n \geq k$.*

In the setting of Cover-Efron and Schläfli random cones, a function similar to that appearing in (b) is always decreasing, as has been recently shown by Hug and Schneider [35].

6.3. **Threshold phenomena for face numbers: the regime of linearly growing k.** Let us now turn to the proportional growth regime. It has been first studied by Vershik and Sporyshev [68] in the context of random projections of the regular simplex.

Theorem 6.5 (Weak threshold in the linear regime)**.** *Let $d \to \infty$ and $k = k(d)$, $n = n(d)$ be functions of d such that*

$$k \sim \alpha n \qquad \text{and} \qquad d \sim \beta n, \qquad \text{as } d \to \infty,$$

for some constants $0 < \alpha < \beta < 1$. Then,

$$\lim_{n\to\infty} \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = \begin{cases} 1, & \text{if } \beta > -\frac{\alpha \log \alpha}{1-\alpha}, \\ 0, & \text{if } \beta < -\frac{\alpha \log \alpha}{1-\alpha}. \end{cases}$$

In the critical case, more precisely, when $k \sim \alpha n$ and $d = \mathbb{E}\,\mathrm{Lah}(n,k) + c\sqrt{n} + o(\sqrt{n})$ for some $\alpha \in (0,1)$ and $c \in \mathbb{R}$, we have

$$\lim_{n\to\infty} \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c/\sigma(\alpha)} \mathrm{e}^{-x^2/2} \mathrm{d}x,$$

where $\sigma^2(\alpha) := -\frac{\alpha}{1-\alpha} - \frac{\alpha(\alpha+1)\log\alpha}{(1-\alpha)^2} - \frac{\alpha^2 \log^2 \alpha}{(1-\alpha)^3}$ is the variance appearing in Theorem 5.1.

Proof. By (6.2) we have

$$\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = 2\mathbb{P}[\mathrm{Lah}(n,k) \in \{d, d-2, d-4, \ldots\}] \le 2\mathbb{P}[\mathrm{Lah}(n,k) \le d] = 2\mathbb{P}[\mathrm{Lah}(n,k) \le (\beta + o(1))n]. \quad (6.10)$$

If now $\beta < -\frac{\alpha \log \alpha}{1-\alpha}$, then (5.15) is applicable and shows that the probability on the right-hand side converges to 0 exponentially fast.

On the other hand, by (6.3) we have

$$1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}} = 2\mathbb{P}[\mathrm{Lah}(n,k) \in \{d+2, d+4, \ldots\}] \le 2\mathbb{P}[\mathrm{Lah}(n,k) \ge d] = 2\mathbb{P}[\mathrm{Lah}(n,k) \ge (\beta + o(1))n]. \quad (6.11)$$

If $\beta > -\frac{\alpha \log \alpha}{1-\alpha}$, then (5.16) is applicable and shows that the probability on the right-hand side converges to 0 exponentially fast.

Consider finally the critical case. Using the unimodality of the Lah distribution (with the mode satisfying $m_{n,k} = \mathbb{E}\,\mathrm{Lah}(n,k) + o(\sqrt{n})$ by Theorem 5.1) and arguing as in the proof of Theorem 6.2, our task reduces to showing that

$$\lim_{n\to\infty} \mathbb{P}[\mathrm{Lah}(n,k) \le d] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c/\sigma(\alpha)} \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

But this follows by writing

$$\mathbb{P}[\mathrm{Lah}(n,k) \le d] = \mathbb{P}\left[\frac{\mathrm{Lah}(n,k) - \mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{n}} \le \frac{d - \mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{n}}\right] = \mathbb{P}\left[\frac{\mathrm{Lah}(n,k) - \mathbb{E}\,\mathrm{Lah}(n,k)}{\sqrt{n}} \le c + o(1)\right]$$

and applying Theorem 5.1. □

Remark 6.6. Let us restate the above results in the notation consistent with the one used by Vershik and Sporyshev [68] and Donoho and Tanner [16, 18]. Following these papers, define

$$\rho = \lim_{n\to\infty} \frac{k(n)}{d(n)} = \frac{\alpha}{\beta} \in (0,1), \qquad \delta := \lim_{n\to\infty} \frac{d(n)}{n} = \beta \in (0,1). \quad (6.12)$$

Similarly to these papers, we say that a function $\delta \mapsto \rho_{\mathrm{weak}}(\delta)$ defines a *weak threshold* for convex hulls of random walks if

$$\mathbb{E} f_{k-1}(C_{n-1,d}) = (1 - o(1)) \cdot \binom{n}{k}, \qquad \text{provided that } \rho < \rho_{\mathrm{weak}}(\delta), \quad (6.13)$$

$$\mathbb{E} f_{k-1}(C_{n-1,d}) = o(1) \cdot \binom{n}{k}, \qquad \text{provided that } \rho > \rho_{\mathrm{weak}}(\delta). \quad (6.14)$$

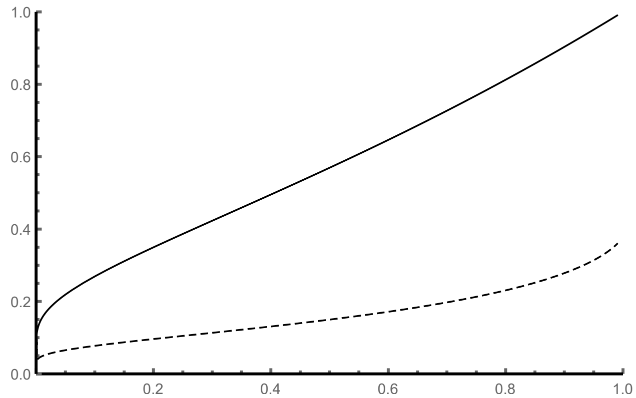

FIGURE 5. Thresholds for convex hulls of random walks from top to bottom: $\rho_{\mathrm{weak}}(\delta)$ (solid) and $\rho_{\mathrm{strong}}(\delta)$ (dashed).

Using Theorem 6.5, we are able to identify the weak threshold explicitly in terms of the Lambert function $W_{-1}(x)$ which is defined as follows: For $-1/\mathrm{e} < x < 0$, the equation $we^w = x$ has two solutions $w = W_0(x)$ and $w = W_{-1}(x)$, the branches of the Lambert function, satisfying $W_{-1}(x) < -1 < W_0(x) < 0$. Then, we claim that the function

$$\rho_{\mathrm{weak}}(\delta) := h^{-1}(\delta)/\delta = -1/W_{-1}(-\delta \mathrm{e}^{-\delta}), \qquad \delta \in (0,1),$$

is the weak threshold; see formula (5.19) and Figure 5 (solid line). Indeed, $\rho < \rho_{\mathrm{weak}}(\delta)$ is equivalent to $\alpha/\beta < h^{-1}(\beta)/\beta$, which is equivalent to $\beta > h(\alpha)$. Knowing this, Theorem 6.5 applies and yields (6.13), while (6.14) follows similarly. Note that $\rho = \rho_{\mathrm{weak}}(\delta)$ is the unique solution to $(1/\rho)\mathrm{e}^{-1/\rho} = \delta \mathrm{e}^{-\delta}$ with $\rho \in (0,1)$, which can be compared to [68, Theorem 1], where a similar characterization of the threshold (involving the Mills ratio function) is given for random projections of the regular simplex. Regarding the behavior of the weak threshold as $\delta \downarrow 0$, it is easy to check that $\rho_{\mathrm{weak}}(\delta) \sim 1/|\log \delta|$, compare [18, Theorems 1.2, 1.4], where a similar asymptotics is stated for weak thresholds of Gaussian polytopes, namely $\rho^{\mathrm{GP}}_{\mathrm{weak}}(\delta) \sim 1/(2|\log \delta|)$ and their symmetric analogues.

Theorem 6.7 (Strong threshold in the linear regime)**.** *Let $d \to \infty$ and $k = k(d)$, $n = n(d)$ be functions of d such that*

$$k \sim \alpha n \qquad \text{and} \qquad d \sim \beta n, \qquad \text{as } d \to \infty,$$

for some constants $0 < \alpha < \beta < 1$. If $I_\alpha(\beta) + \alpha \log \alpha + (1-\alpha)\log(1-\alpha) > 0$, where $I_\alpha(\beta)$ is the rate function from Theorem 5.9, then

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) = O(\mathrm{e}^{-\delta n}) \tag{6.15}$$

for some $\delta > 0$, and the polytope $C_{n-1,d}$ is $(k-1)$-neighborly with probability converging to 1, more precisely,

$$\mathbb{P}\left[f_{k-1}(C_{n-1,d}) = \binom{n}{k}\right] = 1 - O(\mathrm{e}^{-\delta n}). \tag{6.16}$$

If, on the other hand, $I_\alpha(\beta) + \alpha \log \alpha + (1-\alpha)\log(1-\alpha) < 0$, then

$$\lim_{n\to\infty}\left(\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d})\right) = +\infty. \tag{6.17}$$

Proof. Let $I_\alpha(\beta) + \alpha\log\alpha + (1-\alpha)\log(1-\alpha) > 0$, which implies that $I_\alpha(\beta) > 0$ and hence $\beta > -\frac{\alpha\log\alpha}{1-\alpha}$. Using (6.8) and (6.11), we obtain

$$\mathbb{P}\left[f_k(C_{n-1,d}) \neq \binom{n}{k}\right] \le \binom{n}{k}\left(1 - \frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\right) \le 2\binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k) \ge (\beta + o(1))n]. \tag{6.18}$$

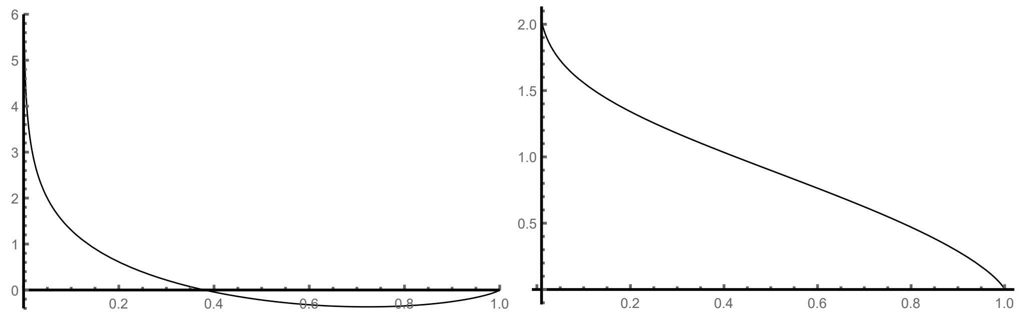

FIGURE 6. Functions needed to define the strong threshold. Left: $\rho \mapsto -\rho \log(h^{-1}(1/\rho)-1) + \log\log h^{-1}(1/\rho)$. Right: $\delta \mapsto -(1/\delta)\log(1-h^{-1}(\delta)) + \log(-\log h^{-1}(\delta))$.

By the Stirling formula,

$$\binom{n}{k} = \mathrm{e}^{-(\alpha\log\alpha+(1-\alpha)\log(1-\alpha))n+o(n)}, \qquad n\to\infty. \tag{6.19}$$

Under the condition $\beta > -\frac{\alpha\log\alpha}{1-\alpha}$ we can apply (5.16) which yields

$$\mathbb{P}[\mathrm{Lah}(n,k) \geq (\beta+o(1))n] = \mathrm{e}^{-I_\alpha(\beta)n+o(n)}. \tag{6.20}$$

Taking everything together, we obtain the claims (6.15) and (6.16).

To prove (6.17), assume that $I_\alpha(\beta)+\alpha\log\alpha+(1-\alpha)\log(1-\alpha)<0$. By (6.3), we have

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) = \binom{n}{k}\left(1-\frac{\mathbb{E} f_{k-1}(C_{n-1,d})}{\binom{n}{k}}\right) = 2\binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k)\in\{d+2,d+4,\dots\}].$$

We may assume that $\beta > -\frac{\alpha\log\alpha}{1-\alpha}$ since otherwise we may increase $d(n)$ by an even sequence which makes the probability on the right side smaller. Under $\beta > -\frac{\alpha\log\alpha}{1-\alpha}$, we can use the unimodality of $\mathrm{Lah}(n,k)$ as in the proof of Theorem 6.2 to estimate

$$\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d}) \geq \binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k)\geq d+2] = \binom{n}{k}\mathbb{P}[\mathrm{Lah}(n,k)\geq(\beta+o(1))n].$$

It follows from (6.19) and (6.20) that the right-hand side is larger than $\mathrm{e}^{\delta n}$, for some $\delta>0$ and all sufficiently large n. □

Remark 6.8. Let us restate the above results in the notation of Donoho and Tanner [16, 18]. We assume (6.12). A function $\delta \mapsto \rho_{\text{strong}}(\delta)$ is said to be a *strong threshold* for convex hulls of random walks if

$$\lim_{n\to\infty}\left(\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d})\right) = 0, \qquad \text{provided that } \rho < \rho_{\text{strong}}(\delta), \tag{6.21}$$

$$\lim_{n\to\infty}\left(\binom{n}{k} - \mathbb{E} f_{k-1}(C_{n-1,d})\right) = +\infty, \qquad \text{provided that } \rho > \rho_{\text{strong}}(\delta). \tag{6.22}$$

Theorem 6.7 yields the following description of the strong threshold: $\rho = \rho_{\text{strong}}(\delta) \in (0,1)$ is the solution of the equation

$$-\rho\log(h^{-1}(1/\rho)-1) + \log\log h^{-1}(1/\rho) = -(1/\delta)\log(1-h^{-1}(\delta)) + \log(-\log h^{-1}(\delta)), \tag{6.23}$$

for $\delta\in(0,1)$; see Figure 5 (dashed line). The plots shown in Figure 6 suggest that the right-hand side, viewed as a function of $\delta\in(0,1)$, decreases from $+\infty$ to 0, whereas the left-hand side, viewed as a function of $\delta\in(0,\rho_*)$ with $\rho_* = 0.3798\dots$, decreases from $+\infty$ to 0, even though we did not verify these claims rigorously. Hence, the solution to the above equation (6.23) exists and is unique. Now, let us prove (6.21) and (6.22). Recalling (5.18) and (6.23) we have

$$\frac{I_\alpha(\beta)+\alpha\log\alpha+(1-\alpha)\log(1-\alpha)}{\beta}=-\rho\log\left(h^{-1}\left(\frac{1}{\rho}\right)-1\right)+\frac{1}{\delta}\log\left(1-h^{-1}(\delta)\right)+\log\left(-\frac{\log h^{-1}(\frac{1}{\rho})}{\log h^{-1}(\delta)}\right).$$

If $\rho<\rho_{\text{strong}}(\delta)$, respectively, $\rho>\rho_{\text{strong}}(\delta)$, then the equality in (6.23) should be replaced by $>$, respectively, $<$, which is equivalent to $I_\alpha(\beta)+\alpha\log\alpha+(1-\alpha)\log(1-\alpha)>0$, respectively, <0. With this at hand, we can apply Theorem 6.7 which yields (6.21), respectively, (6.22).

Remark 6.9. For $\delta=1/2$, that is, when the number of vertices is twice as large as the dimension, the thresholds computed in Remarks 6.6 and 6.8 are $\rho_{\text{weak}}(1/2)=0.5693\ldots$ and $\rho_{\text{strong}}(1/2)=0.1498\ldots$. Let us mention that for the Gaussian polytopes, respectively their symmetric versions, the thresholds are known [18, pp. 6,7] to be

$$\rho^{\text{GP}}_{\text{weak}}(1/2)=0.5581\ldots,\quad \rho^{\text{GP}}_{\text{strong}}(1/2)=0.1335\ldots,\quad \rho^{\pm}_{\text{weak}}(1/2)=0.3848\ldots,\quad \rho^{\pm}_{\text{strong}}(1/2)=0.0894\ldots.$$

Numerically, $\rho_{\text{weak}}(\delta)>\rho^{\text{GP}}_{\text{weak}}(\delta)$ for all $\delta\in(0,1)$, and the difference of these functions is surprisingly close (but not equal to) 0. Thus, convex hulls of random walks are slightly more neighborly than Gaussian polytopes.

6.4. **Threshold phenomena for face numbers: the intermediate regime.** Let us now take some very large dimension d and look at the number of k-dimensional faces, where $k\to\infty$ but $k=o(d)$. The next theorem states that a phase transition occurs if n is near $k\mathrm{e}^{d/k}$.

Theorem 6.10 (Weak threshold in the intermediate regime)**.** *Let $d\to\infty$ and $k=k(d)$ be a function of d such that*

$$\lim_{d\to\infty}k(d)=\infty,\qquad \textit{and}\qquad \lim_{d\to\infty}\frac{k(d)}{d}=0.$$

If an integer sequence $n=n(d)$ is such that $d<n(d)\le k\mathrm{e}^{(1-\varepsilon)d/k}$ for some $\varepsilon>0$ and all sufficiently large d, then

$$\lim_{n\to\infty}\frac{\mathbb{E}f_{k-1}(C_{n-1,d})}{\binom{n}{k}}=1. \tag{6.24}$$

On the other hand, if $n(d)\ge k\mathrm{e}^{(1+\varepsilon)d/k}$ for some $\varepsilon>0$ and sufficiently large d, then

$$\lim_{n\to\infty}\frac{\mathbb{E}f_{k-1}(C_{n-1,d})}{\binom{n}{k}}=0. \tag{6.25}$$

Proof. Let first $n(d)\le k\mathrm{e}^{(1-\varepsilon)d/k}$. Then $k\log(n/k)\le k\log\mathrm{e}^{(1-\varepsilon)d/k}=(1-\varepsilon)d$ for sufficiently large n and hence, recalling (6.3), we can write

$$\begin{aligned}1-\frac{\mathbb{E}f_{k-1}(C_{n-1,d})}{\binom{n}{k}}&=2\mathbb{P}[\mathrm{Lah}(n,k)\in\{d+2,d+4,\ldots\}]\le 2\mathbb{P}[\mathrm{Lah}(n,k)>d]\\&\le 2\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)}{k\log(n/k)}>\frac{d}{k\log(n/k)}\right]\le 2\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)}{k\log(n/k)}>\frac{1}{1-\varepsilon}\right]\underset{n\to\infty}{\longrightarrow}0,\end{aligned}$$

where the last step holds by Theorem 5.12. This proves (6.24). Let now $n(d)\ge k\mathrm{e}^{(1+\varepsilon)d/k}$. Then $k\log(n/k)>k\log\mathrm{e}^{(1+\varepsilon)d/k}=(1+\varepsilon)d$ for sufficiently large n and hence, in view of (6.2) we obtain

$$\begin{aligned}\frac{\mathbb{E}f_{k-1}(C_{n-1,d})}{\binom{n}{k}}&=2\mathbb{P}[\mathrm{Lah}(n,k)\in\{d,d-2,d-4,\ldots\}]\le 2\mathbb{P}[\mathrm{Lah}(n,k)\le d]\\&\le 2\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)}{k\log(n/k)}\le\frac{d}{k\log(n/k)}\right]\le 2\mathbb{P}\left[\frac{\mathrm{Lah}(n,k)}{k\log(n/k)}\le\frac{1}{1+\varepsilon}\right]\underset{n\to\infty}{\longrightarrow}0,\end{aligned}$$

where in the last step we used Theorem 5.12. This proves (6.25). □

In the setting of Gaussian polytopes, the intermediate regime has been studied in [18]. Note that a central limit theorem conjectured in Remark 5.15 would imply a formula for the limit in the critical window.

7. Conic intrinsic volume sums of Weyl chambers

Let us mention an interpretation of the Lah distribution in terms of conic intrinsic volumes. To each convex cone $C \subset \mathbb{R}^n$ it is possible to associate a sequence of quantities $\upsilon_0(C), \dots, \upsilon_n(C)$ which are called conic intrinsic volumes; see [61, Section 6.5] and [2, 3] for their definition and properties. The conic intrinsic volumes form a probability distribution meaning that they are non-negative and sum up to 1. For the Weyl chamber of type A, which is the convex cone defined by

$$A^{(n)} := \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \geq x_2 \geq \dots \geq x_n\},$$

the conic intrinsic volumes are well known to form the $\mathrm{Lah}(n,1)$-distribution meaning that

$$\upsilon_j(A^{(n)}) = \mathbb{P}[\mathrm{Lah}(n,1) = j] = \frac{1}{n!} \begin{bmatrix} n \\ j \end{bmatrix}$$

for all $j = 1, \dots, n$; see, e.g., [42, Theorem 4.2]. To generalize this identity to arbitrary $k = 1, \dots, n$, denote by $\mathscr{F}_k(C)$ the set of k-dimensional faces of a polyhedral cone C, and let $T_F(C)$ be the tangent cone of C at its face F. The next theorem was obtained in [25, Theorem 3.3]; see also [26] for related results.

Theorem 7.1 (Conic intrinsic volume sums of $A^{(n)}$)**.** *For all $j, k \in \{0, \dots, n\}$ we have*

$$\sum_{F \in \mathscr{F}_k(A^{(n)})} \upsilon_j(T_F(A^{(n)})) = \binom{n-1}{k-1} \mathbb{P}[\mathrm{Lah}(n,k) = j] = \frac{k!}{n!} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix}.$$

The proof given in [25] used generating functions. Let us give a combinatorial proof relying on the construction of the Lah distribution given in (2.3).

Proof. By [25, Lemma 3.12], the collection of the tangent cones $T_F(A^{(n)})$, where $F \in \mathscr{F}_k(A^{(n)})$ runs through all k-dimensional faces of $A^{(n)}$ coincides (up to an isometry) with the collection of direct products of the form $A^{(i_1)} \times \dots \times A^{(i_k)}$, where $(i_1, \dots, i_k)$ runs through all compositions of n in k summands. Recalling that $(b_1^{(n)}, \dots, b_k^{(n)})$ denotes a uniform random composition of n in k summands, we can write

$$\sum_{F \in \mathscr{F}_k(A^{(n)})} \upsilon_j(T_F(A^{(n)})) = \binom{n-1}{k-1} \mathbb{E}\, \upsilon_j(A^{(b_1^{(n)})} \times \dots \times A^{(b_k^{(n)})}).$$

Recalling from (2.3) that $(Z_i^{(m)})_{i,m \in \mathbb{N}}$ are independent random variables with $\mathbb{P}[Z_i^{(m)} = \ell] = \frac{1}{i!} \left[\begin{smallmatrix} i \\ \ell \end{smallmatrix}\right] = \upsilon_\ell(A^{(i)})$, and using the formula for the conic intrinsic volumes of direct products, see formula (2.9) in [2], we get

$$\upsilon_j(A^{(i_1)} \times \dots \times A^{(i_k)}) = \sum_{\substack{j_1, \dots, j_k \in \mathbb{N}_0 \\ j_1 + \dots + j_k = j}} \upsilon_{j_1}(A^{(i_1)}) \cdot \ldots \cdot \upsilon_{j_k}(A^{(i_k)}) = \mathbb{P}[Z_{i_1}^{(1)} + \dots + Z_{i_k}^{(k)} = j].$$

Combining everything together, we obtain

$$\sum_{F \in \mathscr{F}_k(A^{(n)})} \upsilon_j(T_F(A^{(n)})) = \binom{n-1}{k-1} \mathbb{P}\left[Z_{b_1^{(n)}}^{(1)} + \dots + Z_{b_k^{(n)}}^{(k)} = j\right] = \binom{n-1}{k-1} \mathbb{P}[\mathrm{Lah}(n,k) = j] = \frac{k!}{n!} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix},$$

where we applied the representation of the Lah distribution given in (2.3). □

In [29] it has been shown that under a minor condition, the conic intrinsic volumes of any sequence of convex cones whose dimension diverges to ∞ satisfy a central limit theorem. One may ask whether there is a natural convex cone $U_{n,k}$ whose conic intrinsic volumes are given by $\upsilon_j(U_{n,k}) = \mathbb{P}[\mathrm{Lah}(n,k) = j]$, for all $j \in \{k, \dots, n\}$. We do not know how to answer this question, but Theorem 7.1 states that $\mathbb{P}[\mathrm{Lah}(n,k) = j]$ is the *expected* j-th conic intrinsic volume of a uniformly selected *random* k-dimensional face of the Weyl chamber $A^{(n)}$; see also [25, Theorem 3.1] and [27, Corollary 2.4] for other examples of this type. Let us also mention that in [2, Lemma 6.5] and [26, Theorem 3.14] (which look similar at a first sight) the Stirling numbers appear in a different order, that is, in the form $\left\{\begin{smallmatrix} n \\ j \end{smallmatrix}\right\} \left[\begin{smallmatrix} j \\ k \end{smallmatrix}\right]$; see [44] and for a review of identities involving this and other types of products.

8. Appendix

8.1. **Proof of Theorem 5.3.** Recall that the distribution of the random uniform composition $(b_1^{(n)},\dots,b_k^{(n)})$ can be represented as

$$\mathbb{P}[(b_1^{(n)},\dots,b_k^{(n)})\in\cdot]=\mathbb{P}[(G_1,\dots,G_k)\in\cdot|G_1+\dots+G_k=n],$$

where $G_1,\dots,G_k$ are independent random variables having the same geometric law on $\mathbb{N}$ with parameter θ. This representation holds for arbitrary $\theta\in(0,1)$ and we are free to choose $\theta:=\theta_n=k/n$. As we demonstrated in Lemma 5.5, it suffices to show that

$$\left(\frac{N_j^{(n)}-k\theta_n(1-\theta_n)^{j-1}}{\sqrt{k}}\right)_{j\ge 1}\xrightarrow[n\to\infty]{d}(\mathscr{N}_j)_{j\ge 1}.$$

By the Cramér–Wold device the last display is equivalent to

$$\frac{\sum_{l=1}^M\beta_l(N_l^{(n)}-k\theta_n(1-\theta_n)^{l-1})}{\sqrt{k}}\xrightarrow[n\to\infty]{d}\sum_{l=1}^M\beta_l\mathscr{N}_l,$$

for arbitrary fixed $M\in\mathbb{N}$ and $\beta_1,\beta_2,\dots,\beta_M\in\mathbb{R}$. Put

$$f_n(x):=\sum_{l=1}^M\beta_l(\mathbb{1}_{\{x=l\}}-\theta_n(1-\theta_n)^{l-1}),\quad x\in\mathbb{N},$$

and, further

$$S_{n,k}:=\sum_{j=1}^k G_j,\quad T_{n,k}:=\sum_{j=1}^k f_n(G_j).$$

The subsequent analysis relies on the following representation

$$\begin{aligned}\mathbb{E}\exp\left(\mathrm{i}tk^{-1/2}\left(\sum_{l=1}^M\beta_l(N_l^{(n)}-k\theta_n(1-\theta_n)^{l-1})\right)\right)&=\mathbb{E}\exp\left(\mathrm{i}tk^{-1/2}\sum_{j=1}^k f_n(b_j^{(n)})\right)\\&\overset{(5.2)}{=}\mathbb{E}\exp\left(\mathrm{i}tk^{-1/2}\sum_{j=1}^k f_n(G_j)\Big|S_{n,k}=n\right)=\mathbb{E}\exp\left(\mathrm{i}tk^{-1/2}T_{n,k}\Big|S_{n,k}=n\right).\end{aligned}$$

Thus, it is enough to prove that, for every fixed $t\in\mathbb{R}$,

$$\lim_{n\to\infty}\mathbb{E}\exp\left(\mathrm{i}tk^{-1/2}T_{n,k}\Big|S_{n,k}=n\right)=\mathbb{E}\exp\left(\mathrm{i}t\sum_{l=1}^M\beta_l\mathscr{N}_l\right).\tag{8.1}$$

According to Theorem 1 in [32] we have

$$\begin{aligned}\mathbb{E}\left(\mathrm{e}^{\mathrm{i}tk^{-1/2}T_{n,k}}\Big|S_{n,k}=n\right)&=\frac{1}{2\pi\mathbb{P}[S_{n,k}=n]}\int_{-\pi}^{\pi}\mathbb{E}\,\mathrm{e}^{\mathrm{i}s(S_{n,k}-n)+\mathrm{i}tk^{-1/2}T_{n,k}}\mathrm{d}s\\&=\frac{1}{2\pi\sqrt{k}\mathbb{P}[S_{n,k}=n]}\int_{-\pi\sqrt{k}}^{\pi\sqrt{k}}\mathbb{E}\,\mathrm{e}^{\mathrm{i}uk^{-1/2}(S_{n,k}-n)+\mathrm{i}tk^{-1/2}T_{n,k}}\mathrm{d}u.\end{aligned}\tag{8.2}$$

Using the Lindeberg–Feller central limit theorem we obtain

$$\left(\frac{S_{n,k}-n}{\sqrt{k}},\frac{T_{n,k}}{\sqrt{k}}\right)\xrightarrow[n\to\infty]{d}(\widetilde{N}_1,\widetilde{N}_2),$$

where $(\widetilde{N}_1,\widetilde{N}_2)$ is a centred Gaussian vector with the following variances and covariance:

$$\begin{aligned}\sigma_1^2&:=\mathrm{Var}\,\widetilde{N}_1=\lim_{n\to\infty}\mathrm{Var}(G_1)=\frac{1-\alpha}{\alpha^2},\\\sigma_2^2&:=\mathrm{Var}\,\widetilde{N}_2=\lim_{n\to\infty}\mathrm{Var}(f_n(G_1))=\sum_{l=1}^M\beta_l^2\alpha(1-\alpha)^{l-1}-\left(\sum_{l=1}^M\beta_l\alpha(1-\alpha)^{l-1}\right)^2,\end{aligned}$$

and

$$r := \mathrm{Cov}(\widetilde{N}_1,\widetilde{N}_2) = \lim_{n\to\infty}\mathrm{Cov}(f_n(G_1),G_1)) = \lim_{n\to\infty}\mathrm{Cov}\left(\sum_{l=1}^{M}\beta_l \mathbb{1}_{\{G_1=l\}}, \sum_{l=1}^{M} l\mathbb{1}_{\{G_1=l\}}\right)$$
$$= \sum_{l=1}^{M}\beta_l l\alpha(1-\alpha)^{l-1} - \alpha^{-1}\sum_{l=1}^{M}\beta_l\alpha(1-\alpha)^{l-1} = \sum_{l=1}^{M}\beta_l(l\alpha-1)(1-\alpha)^{l-1}.$$

Since $S_k - k$ has the negative binomial distribution, direct calculation shows that the limit $\lim_{n\to\infty}\sqrt{k}\mathbb{P}[S_k = n]$ exists and is positive. Thus, by the Lebesgue dominated convergence theorem, we deduce from (8.2)

$$\lim_{n\to\infty}\mathbb{E}\left(\mathrm{e}^{\mathrm{i}tk^{-1/2}T_{n,k}}\middle| S_{n,k}=n\right) = \mathrm{const}\cdot\int_{-\infty}^{\infty}\mathbb{E}\exp\left(\mathrm{i}u\widetilde{N}_1 + \mathrm{i}t\widetilde{N}_2\right)\mathrm{d}u$$
$$= \mathrm{const}\cdot\int_{-\infty}^{\infty}\exp\left(-\frac{u^2\sigma_1^2+t^2\sigma_2^2+2rut}{2}\right)\mathrm{d}u = \exp\left(-\frac{\sigma_1^2\sigma_2^2-r^2}{2\sigma_1^2}t^2\right). \tag{8.3}$$

To ensure applicability of the dominated convergence (which is non-trivial), one can argue as in the paper of Holst [32] who relies on [51]. It remains to note that the right-hand sides of (8.1) and (8.3) coincide as is readily seen by comparing the variances.

ACKNOWLEDGEMENT

ZK acknowledges support 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 National Research Foundation of Ukraine (project 2020.02/0014 "Asymptotic regimes of perturbed random walks: on the edge of modern and classical probability"). ZK is grateful to Thomas Godland for useful discussions.

REFERENCES

[1] G. Alsmeyer, Z. Kabluchko, and A. Marynych. Limit theorems for the least common multiple of a random set of integers. *Trans. Amer. Math. Soc.*, 372(7):4585–4603, 2019.

[2] D. Amelunxen and M. Lotz. Intrinsic volumes of polyhedral cones: A combinatorial perspective. *Discrete Comput. Geom.*, 58(2):371–409, 2017.

[3] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp. Living on the edge: phase transitions in convex programs with random data. *Inf. Inference*, 3(3):224–294, 2014.

[4] P. Baldi and R. Vershynin. A theory of capacity and sparse neural encoding. *Neural Networks, to appear*, 2021.

[5] A. D. Barbour, E. Kowalski, and A. Nikeghbali. Mod-discrete expansions. *Probab. Theory Related Fields*, 158(3-4):859–893, 2014.

[6] Y. M. Baryshnikov and R. A. Vitale. Regular simplices and Gaussian samples. *Discrete Comput. Geom.*, 11(2):141–147, 1994.

[7] N. Berestycki. *Recent progress in coalescent theory*, volume 16 of *Ensaios Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, Rio de Janeiro, 2009.

[8] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 1999.

[9] A. Bulinski and A. Shashkin. *Limit theorems for associated random fields and related systems*, volume 10 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007.

[10] J. Cilleruelo, J. Rué, P. Šarka, and A. Zumalacárregui. The least common multiple of random sets of positive integers. *J. Number Theory*, 144:92–104, 2014.

[11] S. Daboul, J. Mangaldan, M. Z. Spivey, and P. J. Taylor. The Lah numbers and the nth derivative of $e^{1/x}$. *Math. Mag.*, 86(1):39–47, 2013.

[12] F. Delbaen, E. Kowalski, and A. Nikeghbali. Mod-φ convergence. *Int. Math. Res. Not. IMRN*, 2015(11):3445–3485, 2015.

[13] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[14] P. Diaconis and D. Freedman. A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2, suppl.):397–423, 1987.

[15] D. L. Donoho. High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension. *Discrete Comput. Geom.*, 35(4):617–652, 2006.

[16] D. L. Donoho and J. Tanner. Neighborliness of randomly projected simplices in high dimensions. *Proc. Natl. Acad. Sci. USA*, 102(27):9452–9457, 2005.

[17] D. L. Donoho and J. Tanner. Sparse nonnegative solution of underdetermined linear equations by linear programming. *Proc. Natl. Acad. Sci. USA*, 102(27):9446–9451, 2005.

[18] D. L. Donoho and J. Tanner. Counting faces of randomly projected polytopes when the projection radically lowers dimension. *J. Amer. Math. Soc.*, 22(1):1–53, 2009.

[19] D. L. Donoho and J. Tanner. Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete Comput. Geom.*, 43(3):522–541, 2009.

[20] D. L. Donoho and J. Tanner. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 367(1906):4273–4293, 2009. With electronic supplementary materials available online.

[21] D. L. Donoho and J. Tanner. Exponential bounds implying construction of compressed sensing matrices, error-correcting codes, and neighborly polytopes by random sampling. *IEEE Trans. Inform. Theory*, 56(4):2002–2016, 2010.

[22] V. Féray, P.-L. Méliot, and A. Nikeghbali. *Mod-φ convergence: Normality zones and precise deviations*. SpringerBriefs in Probability and Mathematical Statistics. Springer, 2016.

[23] J. L. Fields. The uniform asymptotic expansion of a ratio of two gamma functions. In *Proc. Inter. Conf. Constructive Func. Theory*, pages 171–176, 1970.

[24] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[25] T. Godland and Z. Kabluchko. Angle sums of Schläfli orthoschemes, 2020. Discrete Comput. Geom., to appear. Preprint at arXiv: 2007.02293.

[26] T. Godland and Z. Kabluchko. Projections and angle sums of permutohedra and other polytopes, 2020. Preprint at arXiv: 2009.04186.

[27] T. Godland and Z. Kabluchko. Positive hulls of random walks and bridges, 2021. Preprint at arXiv: 2104.05542.

[28] T. Godland, Z. Kabluchko, and C. Thäle. Random cones in high dimensions I: Donoho-Tanner and Cover-Efron cones, 2020. Preprint at arXiv: 2012.06189.

[29] L. Goldstein, I. Nourdin, and G. Peccati. Gaussian phase transitions and conic intrinsic volumes: Steining the Steiner formula. *Ann. Appl. Probab.*, 27(1):1–47, 2017.

[30] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Publishing Company, Inc., USA, 2nd edition, 1994.

[31] L. H. Harper. Stirling behavior is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[32] L. Holst. Two conditional limit theorems with applications. *Ann. Stat.*, 7(3):551–557, 1979.

[33] L. Holst. A unified approach to limit theorems for urn models. *J. Appl. Probab.*, 16(1):154–162, 1979.

[34] D. Hug and R. Schneider. Threshold phenomena for random cones, 2020. Preprint at arXiv: 2004.11473.

[35] D. Hug and R. Schneider. Another look at threshold phenomena for random cones, 2021. Preprint at arXiv: 2103.11394.

[36] G. I. Ivchenko. On the random coverage of the circle: a discrete model. *Discrete Math. Appl.*, 4(2):147–162, 1994.

[37] J. Jacod, E. Kowalski, and A. Nikeghbali. Mod-Gaussian convergence: new limit theorems in probability and number theory. *Forum Math.*, 23(4):835–873, 2011.

[38] K. Joag-Dev and F. Proschan. Negative association of random variables, with applications. *Ann. Statist.*, 11(1):286–295, 1983.

[39] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Discrete multivariate distributions*. Wiley Series in Probability and Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1997. A Wiley-Interscience

Publication.
[40] Z. Kabluchko, A. Marynych, and H. Sulzbach. Mode and Edgeworth expansion for the Ewens distribution and the Stirling numbers. *J. Integer Seq.*, 19(8):Art. 16.8.8, 17, 2016.
[41] Z. Kabluchko, A. Marynych, and H. Sulzbach. General Edgeworth expansions with applications to profiles of random trees. *Ann. Appl. Probab.*, 27(6):3478–3524, 2017.
[42] 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.
[43] Z. Kabluchko, V. Vysotsky, and D. Zaporozhets. Convex hulls of random walks, hyperplane arrangements, and Weyl chambers. *Geom. Func. Anal.*, 27(4):880–918, 2017.
[44] M. Knežević, V. Krčadinac, and L. Relić. Matrix products of binomial coefficients and unsigned Stirling numbers, 2020. Preprint at arXiv: 2012.15307.
[45] V. F. Kolchin, B. A. Sevastyanov, and V. P. Chistyakov. *Random allocations*. V. H. Winston & Sons, Washington, D.C.; distributed by Halsted Press [John Wiley & Sons], New York-Toronto, Ont.-London, 1978.
[46] V. F. Kolčin. A certain class of limit theorems for conditional distributions. *Litovsk. Mat. Sb.*, 8:53–63, 1968.
[47] V. F. Kolčin. Branching processes, random trees and a generalized particle distribution scheme. *Mat. Zametki*, 21(5):691–705, 1977.
[48] E. Kowalski and A. Nikeghbali. Mod-Poisson convergence in probability and number theory. *Int. Math. Res. Not. IMRN*, 2010(18):3549–3587, 2010.
[49] E. Kowalski and A. Nikeghbali. Mod-Gaussian convergence and the value distribution of $\zeta(\frac{1}{2}+it)$ and related quantities. *J. Lond. Math. Soc. (2)*, 86(1):291–319, 2012.
[50] I. Lah. A new kind of numbers and its application in the actuarial mathematics. *Boletim do Instituto dos Actuários Portugueses*, 9:7–15, 1954.
[51] L. Le Cam. Un théorème sur la division d'un intervalle par des points pris au hasard. *Publ. Inst. Statist. Univ. Paris*, 7(3-4):7–16, 1958.
[52] J. L. Lebowitz, B. Pittel, D. Ruelle, and E. R. Speer. Central limit theorems, Lee-Yang zeros, and graph-counting polynomials. *J. Combin. Theory Ser. A*, 141:147–183, 2016.
[53] G. Louchard. Asymptotics of the Stirling numbers of the first kind revisited: a saddle point approach. *Discrete Math. Theor. Comput. Sci.*, 12(2):167–184, 2010.
[54] G. Louchard. Asymptotics of the Stirling numbers of the second kind revisited. *Appl. Anal. Discrete Math.*, 7(2):193–210, 2013.
[55] P.-L. Méliot and A. Nikeghbali. Mod-Gaussian convergence and its applications for models of statistical mechanics. In *In memoriam Marc Yor—Séminaire de Probabilités XLVII*, volume 2137 of *Lecture Notes in Math.*, pages 369–425. Springer, Cham, 2015.
[56] L. Moser and M. Wyman. Asymptotic development of the Stirling numbers of the first kind. *J. London Math. Soc.*, 33:133–146, 1958.
[57] L. Moser and M. Wyman. Stirling numbers of the second kind. *Duke Math. J.*, 25:29–43, 1958.
[58] S. Narumi. On a power series having only a finite number of algebraico logarithmic singularities on its circle of convergence. *Tôhoku Math. J.*, 30:185–201, 1929.
[59] V. N. Sachkov. *Combinatorial methods in discrete mathematics*, volume 55 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1996.
[60] V. N. Sachkov. *Probabilistic methods in combinatorial analysis*, volume 56 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1997.
[61] R. Schneider and W. Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.
[62] M. Sibuya. Log-concavity of Stirling numbers and unimodality of Stirling distributions. *Ann. Inst. Statist. Math.*, 40(4):693–714, 1988.
[63] N. J. A. Sloane (editor). The On-Line Encyclopedia of Integer Sequences. https://oeis.org.
[64] A. N. Timashev. On asymptotic expansions of Stirling numbers of the first and second kinds. *Discrete Math. Appl.*, 8(5):533–544, 1998.

[65] A. N. Trunov. Limit theorems in the problem of allocation of identical particles among different cells. *Proc. Steklov Inst. Math.*, 177:157–175, 1988.
[66] A. M. Vershik and P. V. Sporyshev. Estimation of the mean number of steps in the simplex method, and problems of asymptotic integral geometry. *Dokl. Akad. Nauk SSSR*, 271(5):1044–1048, 1983.
[67] A. M. Vershik and P. V. Sporyshev. An asymptotic estimate of the average number of steps of the parametric simplex method. *USSR Comput. Math. and Math. Physics*, 26(3):104–113, 1986.
[68] A. M. Vershik and P. V. Sporyshev. Asymptotic behavior of the number of faces of random polyhedra and the neighborliness problem. *Selecta Math. Soviet.*, 11(2):181–201, 1992. Selected translations.
[69] A. M. Vershik and Y. V. Yakubovich. Asymptotics of the uniform measures on simplices and random compositions and partitions. *Func. Anal. Appl.*, 37(4):273–280, 2003.

ZAKHAR KABLUCHKO: INSTITUT FÜR MATHEMATISCHE STOCHASTIK, WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER, ORLÉANS–RING 10, 48149 MÜNSTER, GERMANY
Email address: zakhar.kabluchko@uni-muenster.de

ALEXANDER MARYNYCH: FACULTY OF COMPUTER SCIENCE AND CYBERNETICS, TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV, KYIV 01601, UKRAINE
Email address: marynych@unicyb.kiev.ua