On multisets, interpolated multiple zeta values and limit laws

Markus Kuba

Department of Applied Mathematics and Physics
FH Technikum Wien
Höchstädtplatz 6, 1200-Vienna, Austria

`kuba@technikum-wien.at`

Abstract

In this work we discuss a parameter σ on weighted k-element multisets of $[n] = \{1, \dots, n\}$. The sums of weighted k-multisets are related to k-subsets, k-multisets, as well as special instances of truncated interpolated multiple zeta values. We study properties of this parameter using symbolic combinatorics. We (re)derive and extend certain identities for $\zeta_n^t(\{m\}_k)$. Moreover, we introduce random variables on the k-element multisets and derive their distributions, as well as limit laws for k or n tending to infinity.

Mathematics Subject Classifications: 60C05, 11M32

1 Introduction

The multiple zeta values are defined as

$$\zeta(i_1, \dots, i_k) = \sum_{\ell_1 > \cdots > \ell_k \geqslant 1} \frac{1}{\ell_1^{i_1} \cdots \ell_k^{i_k}},$$

with admissible indices $(i_1, \dots, i_k)$ satisfying $i_1 \geqslant 2$, $i_j \geqslant 1$ for $2 \leqslant j \leqslant k$, see Hoffman [12] or Zagier [39]. Their truncated counterpart, sometimes called multiple harmonic sum, is given by

$$\zeta_n(i_1, \dots, i_k) = \sum_{n \geqslant \ell_1 > \cdots > \ell_k \geqslant 1} \frac{1}{\ell_1^{i_1} \cdots \ell_k^{i_k}}.$$

We refer to $i_1 + \cdots + i_k$ as the weight of this multiple zeta value, and k as its depth. For a comprehensive overview as well as a great many pointers to the literature we refer to

the survey of Zudilin [40]. An important variant of the (truncated) multiple zeta values are the so-called multiple zeta star values, where equality is allowed:

$$\zeta^\star(i_1,\ldots,i_k) = \sum_{\ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} \frac{1}{\ell_1^{i_1} \cdots \ell_k^{i_k}},$$

their truncated counterpart $\zeta_n^\star(i_1,\ldots,i_k)$ is defined analogous. On the combinatorial side, it is well known that two special cases, namely the series $\zeta_n(\{1\}_k)$ and $\zeta_n^\star(\{1\}_k)$ occur in a multitude of different places. See for example [28] and the references therein for different representation of the two series. The value $\zeta_n(\{1\}_k)$ is closely related to the Stirling numbers of the first kind, and $\zeta_n^\star(\{1\}_k)$ is related to alternating binomial sums. Here and throughout this work $\{m\}_k$ means m repeated k times. For non-truncated series ζ and $\zeta^\star$, Yamamoto [38] introduced a generalization of both versions called interpolated multiple zeta values. Noting that,

$$\zeta^\star(i_1,\ldots,i_k) = \sum_{\circ = \text{``,'' or ``+''}} \zeta(i_1 \circ i_2 \cdots \circ i_k),$$

let the parameter σ denote the number of plus in the expression $i_1 \circ i_2 \cdots \circ i_k$. Yamamoto defines

$$\zeta^t(i_1,\ldots,i_k) = \sum_{\circ = \text{``,'' or ``+''}} t^\sigma \zeta(i_1 \circ i_2 \cdots \circ i_k). \tag{1}$$

Thus, the series $\zeta^t(i_1,\ldots,i_k)$ interpolates between multiple zeta values, case $t = 0$, and multiple zeta star values, case $t = 1$. It turned out that the interpolated series satisfies many identities generalizing or unifying earlier result for multiple zeta and zeta star values, see for example Yamamoto [38], Hoffman and Ihara [16] or Hoffman [13, 15]. In particular, a so-called quasi-shuffle product [13, 15, 16, 38], sometimes also called stuffle product or harmonic shuffle product, can be defined for the interpolated zeta values. Hoffman and Ihara used an algebra framework, which leads amongst others to expressions for interpolated multiple zeta values $\zeta^t(\{m\}_k)$ in terms of Bell polynomials and ordinary single argument zeta values, $m \geqslant 1$. We discuss and (re)derive in this work the results for $\zeta_n^t(\{m\}_k)$ and $\zeta^t(\{m\}_k)$ using symbolic combinatorics. Moreover, we relate and extend certain multiple harmonic sums identities [33, 34] to $\zeta_n^t(\{m\}_k)$.

We study in particular the interpolated truncated multiple zeta values: we are interested in the truncated series

$$\zeta_n^t(\{1\}_k) = \sum_{\circ = \text{``,'' or ``+''}} t^\sigma \zeta_n(\underbrace{1 \circ 1 \circ \cdots \circ 1}_{k}) \tag{2}$$

and their properties. Note that $\zeta_n^t(\{1\}_k)$ interpolates between $\zeta_n(\{1\}_k)$ and $\zeta_n^\star(\{1\}_k)$, $t = 0$ and $t = 1$, respectively.

On the other hand, classical combinatorial problems are the enumeration of k-element multisets, short k-multisets, of $[n] = \{1,\ldots,n\}$, leading to the multiset coefficients, $\binom{n+k-1}{k}$ and the enumeration of k-element sets, short k-subsets, of $[n] = \{1,\ldots,n\}$ leading to the binomial coefficients, $\binom{n}{k}$.

The goal of this work is to unify these two topics, namely interpolated (truncated) multiple zeta values and the enumeration of k-subsets and k-multisets. We study k-multisets of $[n]$ and introduce a weighted enumeration of k-multisets of $[n]$ in terms of a given weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$. Our weighted enumeration also takes into account an additional parameter $\sigma = \sigma_{n,k}$. It is defined similarly to Yamamoto [38]: σ counts the number of equalities in the nested sum representation, corresponding to the number of plus signs in (1) or (2). The weighted enumeration also has an algebraic interpretation, connecting elementary symmetric functions e_k and complete symmetric functions h_k; moreover power sum $p_k s$ also appear in an alternative representation.

In the next section we introduce our weighted enumeration of k-multisets, given an arbitrary sequence $(a_j)_{j \in \mathbb{N}}$ of positive reals, $a_j > 0$. Closely related ideas have been considered by Vignat and Takhare [37], who considered (non-truncated) generalized multiple zeta values and infinity products, with $(a_j)_{j \in \mathbb{N}}$ given by the zeros of certain special functions. Moreover, Bachmann very briefly discussed related ideas at the end of his article [1]. Also related are the partition zeta functions introduced by Schneider [35].

We study two aspects of the weighted enumerations: first, its relation to k-sets and k-multisets and to the interpolated (truncated) multiple zeta function, leading to our main result in Theorems 13, and second, the probabilistic aspects of σ. While the algebraic and number theoretic aspects of interpolated multiple zeta values have been studied in the literature, no studies so far have been conducted on its stochastic properties concerning structures of large size and its distribution, to the best of the author's knowledge. Questions such as "what is the expected value of the parameter σ" or "what is the limit law of σ" have not been considered before.

Introducing suitably defined random variables $S_{n,k}$, we study the distribution of σ, as well as limit laws for k or n tending to infinity. In particular, we show that the parameter σ in k-multisets follows a hypergeometric distribution and derive Poisson and Gaussian limit laws. For $\zeta_n^\star(\{1\}_k)$ we show that also different limit laws occur. In a later section we introduce refinements of the parameter σ an also discuss related random variables. It turns out that the random variables $S_{n,k}$, its refinements and its limit laws are closely related to many quantities studied earlier in combinatorial probability theory: maxima in hypercubes [2, 22], unsuccessfull search in binary search tree [6], descendants in increasing trees [24], edge-weighted increasing trees [25, 29], distances in increasing trees [5, 6], leaf-isolation procedures in random trees [26], as well as asymptotics of the Poisson distribution [6, 21].

Throughout this work use the notation $H_n = \zeta_n(1)$ for the nth harmonic number, $H_n^{(j)} = \zeta_n(j)$ for the nth generalized harmonic number of order j. We denote with $\mathcal{N}(0,1)$ a standard normal distributed random variable, with $\mathrm{Poisson}(\lambda)$ a Poisson distributed random variable with parameter λ, and with $\mathrm{Hypergeo}(N, K, n)$ a hypergeometric distributed random variable with parameters N, K, n. For the reader's convenience a few basic facts about the probability distributions, appearing later in our analysis, are collected in the appendix of this work. We denote with $x^{\underline{k}}$ the kth falling factorial, $x^{\underline{k}} = x(x-1) \ldots (x - (k-1))$, $k \in \mathbb{N}$, with $x^{\underline{0}} = 1$. Furthermore, we use the abbreviation

E_x for the evaluation operator at $x = 1$. Moreover, we denote by $X \overset{(d)}{=} Y$ the equality in distribution of the random variables X and Y, and by $X_n \overset{(d)}{\longrightarrow} X$ convergence in distribution of the sequence of random variables X_n to a random variable X. We denote with $\mathcal{P}(k)$ the set of ordered partitions of the integer k, often also called compositions of the integer k, and $\mathcal{L}(\mathbf{p})$ the length of a partition $\mathbf{p}$, defined as the number of its summands. E.g., for $k = 3$ we have $\mathcal{P}_O(3) = \{(1,1,1),(2,1),(1,2),(3)\}$ and

$$\mathcal{P}_O(4) = \{(1,1,1,1),(1,1,2),(1,2,1),(2,1,1),(1,3),(3,1),(4)\}.$$

Concerning length of partitions, we have $\mathcal{L}((1,1,1,1)) = 4$ and $\mathcal{L}((4)) = 1$.

2 Interpolated weighted Multisets and parameter sigma

2.1 Parameter sigma

We consider k-multisets of $[n] = \{1, \ldots, n\}$. Strictly speaking, we do not directly consider multisets. Instead, we induce an increasing order on the distinct elements in the multisets and consider the ordered sequences. For example, the unordered multiset $\{1, 1, 4, 2, 2, 6, 2\}$ corresponds to the increasing sequence $(1, 1, 2, 2, 2, 4, 6)$. Let $\mathcal{M}_{n,k}$ denote the set

$$\mathcal{M}_{n,k} = \{\vec{\ell} \in \mathbb{N}^k \colon 1 \leqslant \ell_k \leqslant \ldots \leqslant \ell_2 \leqslant \ell_1 \leqslant n\}, \tag{3}$$

such that each $\vec{\ell}$ corresponds to a k-multiset of $[n]$. In a slight abuse of notation, we denote with $|\vec{\ell}| = k$ the length of $\vec{\ell}$, in other words the cardinality of the corresponding multiset. The cardinality of $\mathcal{M}_{n,k}$ is given by the multiset coefficients, counting the number of k-multisets of $[n]$:

$$|\mathcal{M}_{n,k}| = \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{\ell_1} \cdots \sum_{\ell_k=1}^{\ell_{k-1}} 1 = \binom{n+k-1}{k}.$$

Closely related is the set

$$\mathcal{R}_{n,k} = \{\vec{\ell} \in \mathbb{N}^k \colon 1 \leqslant \ell_k < \cdots < \ell_2 < \ell_1 \leqslant n\}, \tag{4}$$

whose cardinality is given by binomial coefficients, counting the number of k-sets of $[n]$:

$$|\mathcal{R}_{n,k}| = \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{\ell_1-1} \cdots \sum_{\ell_k=1}^{\ell_{k-1}-1} 1 = \binom{n}{k}.$$

In the following we introduce the parameter $\sigma = \sigma_{n,k}$ for elements $\vec{\ell} \in \mathcal{M}_{n,k}$.

Definition 1. Let $\vec{\ell} \in \mathcal{M}_{n,k}$ (3). Then $\sigma = \sigma_{n,k}(\vec{\ell})$ is defined as the number of elements of $\vec{\ell}$ equal to the preceding one in $\vec{\ell}$, or in other words

$$\sigma(\vec{\ell}) = \sigma(\ell_1, \ldots, \ell_k) = |\{1 \leqslant j \leqslant k-1 \colon \ell_j = \ell_{j+1}\}|.$$

Example 2. Given $\vec{\ell} = (1,1,2,2,2,4,6) \in \mathcal{M}_{n,7}$ then $\sigma(\vec{\ell}) = 3$. Given $\vec{\ell} = (5,5,5,5) \in \mathcal{M}_{n,3}$ then $\sigma(\vec{\ell}) = 3$.

Example 3. The parameter σ relates $\mathcal{M}_{n,k}$ and $\mathcal{R}_{n,k}$, as defined in (3), (4), respectively:

$$\mathcal{R}_{n,k} = \{\vec{\ell} \in \mathcal{M}_{n,k} \colon \sigma(\vec{\ell}) = 0\}.$$

Next we define multiplicative weights for $\vec{\ell} \in \mathcal{M}_{n,k}$ and the weighted enumerations of $\mathcal{M}_{n,k}$ (3).

Definition 4 (Interpolated weighted multiset sums)**.** Given an the weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$ of positive real numbers a_j. Then, the weight $w(\vec{\ell}) = w(\vec{\ell}, \boldsymbol{a})$ of $\vec{\ell} = (\ell_1, \ldots, \ell_k) \in \mathcal{M}_{n,k}$ is defined in terms of $\boldsymbol{a}$ as $w(\vec{\ell}) = \prod_{j=1}^{k} a_{\ell_j}$ and $\theta_{n;k}(t) = \theta_{n;k}(t; \boldsymbol{a})$ is defined as

$$\theta_{n;k}(t) = \sum_{\vec{\ell} \in \mathcal{M}_{n,k}} w(\vec{\ell}) t^{\sigma(\vec{\ell})} = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k} t^{\sigma(\vec{\ell})}. \tag{5}$$

Note that we use the convention $\theta_{n;0}(t) = 1$ in the boundary case $k = 0$.

Remark 5. The construction is also well defined for

$$\mathcal{M}_{\infty,k} = \{\vec{\ell} \in \mathbb{N}^k \colon 1 \leqslant \ell_k \leqslant \ldots \leqslant \ell_2 \leqslant \ell_1 < \infty\}, \tag{6}$$

as long as the weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$ is chosen in such a way that $\theta_{\infty;k}(t) = \sum_{\vec{\ell} \in \mathcal{M}_{\infty,k}} w(\vec{\ell}) t^{\sigma(\vec{\ell})}$ converges. This construction is very closely related to the generalized multiple zeta values of Vignat and Wakhare [37].

Remark 6 (Interpolated complete symmetric polynomials)**.** Regarding the weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$ as variables, the values $\theta_{n;k}(t)$ can also be interpreted in an algebraic way. This is no surprise, due to the well known relationship between the multiple zeta values and quasi-symmetric functions [13, 15, 16, 37, 39]. They relate complete elementary symmetric polynomials h_k and elementary symmetric polynomials e_k: $\theta_{n;k}(t) = \theta_{n;k}(t; \boldsymbol{a})$ at $t = 0$ or $t = 1$ is given by

$$\theta_{n;k}(0; \boldsymbol{a}) = e_k(a_1, \ldots, a_n) = \sum_{n \geqslant \ell_1 > \cdots > \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k},$$

and

$$\theta_{n;k}(1; \boldsymbol{a}) = h_k(a_1, \ldots, a_n) = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k}.$$

We will see later that the power sums p_k on n variables,

$$p_k(a_1, \ldots, a_n) = A_n(k) = \sum_{m=1}^{n} a_m^k, \tag{7}$$

also appear in an alternative representation of $\theta_{n;k}(t)$.

Example 7 (Interpolated k-sets and multisets)**.** Given the weight sequence $\boldsymbol{a} = (1)_{j \in \mathbb{N}}$ we obtain

$$\theta_{n;k}(t) = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} t^{\sigma(\vec{\ell})},$$

interpolating between k-sets and multisets of $[n]$. The special values $\theta_{n;k}(0)$ and $\theta_{n;k}(1)$ enumerate k-sets (4) and multisets (3) of $[n]$:

$$\theta_{n;k}(0) = |\mathcal{R}_{n,k}| = \binom{n}{k}, \quad \theta_{n;k}(1) = |\mathcal{M}_{n,k}| = \binom{n+k-1}{k}.$$

Example 8. Given the weight sequence $\boldsymbol{a} = (j)_{j \in \mathbb{N}}$ we obtain

$$\theta_{n;k}(t) = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} \ell_1 \cdots \ell_k \cdot t^{\sigma(\vec{\ell})}.$$

The special values are given in terms of the Stirling numbers of the first and second kind [10]:

$$\theta_{n;k}(0) = \begin{bmatrix} n+1 \\ n+1-k \end{bmatrix} = n! \cdot \zeta_n(\{1\}_{n+k}), \quad \theta_{n;k}(1) = \begin{Bmatrix} n+k \\ n \end{Bmatrix}.$$

Example 9 (Interpolated truncated multiple zeta values)**.** Given the weight sequence $\boldsymbol{a} = (\frac{1}{j^m})_{j \in \mathbb{N}}$, $m > 0$ we obtain a special instance of the truncated interpolated multiple zeta values,

$$\theta_{n;k}(t) = \zeta_n^t(\{m\}_k) = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} \frac{t^{\sigma(\vec{\ell})}}{\ell_1^m \cdots \ell_k^m},$$

such that $\theta_{n;k}(0) = \zeta_n(\{m\}_k)$ and $\theta_{n;k}(1) = \zeta_n^\star(\{m\}_k)$.

2.2 Ordered partitions and Interpolated multiple zeta values

Let $\vec{\ell} \in \mathcal{M}_{n,k}$ (3). We can associate an ordered partition $\mathbf{p} \in \mathcal{P}_O(k)$ to $\vec{\ell}$ as follows: there exist integers $1 \leqslant s \leqslant k$ and $n \geqslant j_1 > j_2 > \cdots > j_s \geqslant 1$ with multiplicities $r_1, \ldots, r_s \in \mathbb{N}$ with $\sum_{i=1}^s r_i = k$ such that $\vec{\ell} = \vec{j} = (j_s^{r_s}, \ldots, j_1^{r_1})$. Here, we use that shorthand notation $j_i^{r_i}$ for j_i appearing exactly r_i times. We refer to the ordered partition $\mathbf{p} = \mathbf{p}(\vec{j}) = (r_1, \ldots, r_s) \in \mathcal{P}_O(k)$ as the shape of $\vec{j}$. The shape, as a map, is a surjection from $\mathcal{M}_{n,k}$ to $\mathcal{P}_O(k)$. Note that

$$\sigma(\vec{j}) = \sum_{i=1}^s (r_i - 1) = k - s = k - \mathcal{L}(\mathbf{p}).$$

We may write

$$\theta_{n;k}(t) = \sum_{\vec{\ell} \in \mathcal{M}_{n,k}} a_{\ell_1} \cdot \cdots \cdot a_{\ell_k} t^{\sigma(\vec{\ell})} = \sum_{\mathbf{p} \in \mathcal{P}_O(k)} \sum_{\substack{\vec{\ell} \in \mathcal{M}_{n,k} \\ \text{shape}(\vec{\ell}) = \mathbf{p}}} a_{\ell_1} \cdot \cdots \cdot a_{\ell_k} t^{\sigma(\vec{\ell})}$$

Thus, for $\boldsymbol{a} = (\frac{1}{j^m})_{j \in \mathbb{N}}$ we get

$$
\begin{aligned}
\theta_{n;k}(t) &= \sum_{\substack{\mathbf{p} \in \mathcal{P}_O(k)}} \sum_{\substack{\vec{j} \in \mathcal{M}_{n,k} \\ \text{shape}(\vec{j})=(r_1,\ldots,r_s)=\mathbf{p}}} \frac{1}{j_1^{m \cdot r_1}} \cdots \cdot \frac{1}{j_s^{m \cdot r_s}} t^{\sigma(\vec{j})} \\
&= \sum_{\substack{\mathbf{p} \in \mathcal{P}_O(k)}} \sum_{\substack{n \geqslant j_1 > j_2 > \cdots > j_s \geqslant 1 \\ \text{shape}(\vec{j})=(r_1,\ldots,r_s)=\mathbf{p}}} \frac{1}{j_1^{m \cdot r_1}} \cdots \cdot \frac{1}{j_s^{m \cdot r_s}} t^{\sigma(\vec{j})} \\
&= \sum_{\mathbf{p} \in \mathcal{P}_O(k)} \zeta_n(m \cdot r_1, \ldots, m \cdot r_s) t^{k-s} = \sum_{\mathbf{p} \in \mathcal{P}_O(k)} \zeta_n(m \cdot \mathbf{p}) t^{k - \mathcal{L}(p)}.
\end{aligned}
$$

3 Generating functions and Bell polynomials

We use symbolic combinatorics to obtain the generating function

$$
\Theta_n(z,t) = \sum_{k \geqslant 0} \theta_{n;k}(t) z^k.
$$

This leads then to several expressions for $\theta_{n;k}(t)$, collected in our main result, Theorem 13, generalizing several results in the literature [3, 4, 16, 33, 28]. To the best of the author's knowledge, our general results have not appeared before in literature, although related ideas have been developed earlier or in parallel, as mentioned in the introduction. Let $\mathcal{M}_n$ denote all multisets of $[n]$, $\mathcal{M}_n = \bigcup_{k=1}^{\infty} \mathcal{M}_{n,k}$, with $\mathcal{M}_{n,k}$ as stated in (3). In order to gain more insight into $\theta_{n;k}(t)$ we use symbolic combinatorics and study the generating function

$$
\begin{aligned}
\Theta_n(z,t) &= \sum_{\vec{\ell} \in \mathcal{M}_n} w(\vec{\ell}) t^{\sigma(\vec{\ell})} z^{|\vec{\ell}|} = \sum_{k \geqslant 0} \theta_{n;k}(t) z^k \\
&= \sum_{k \geqslant 0} z^k \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k} t^{\sigma(\vec{\ell})}.
\end{aligned}
$$

Theorem 10 (Generating functions and power sums). *The generating function* $\Theta_n(z,t)$ *is given by*

$$
\Theta_n(z,t) = \prod_{m=1}^{n} \left(1 + \frac{a_m z}{1 - a_m z t} \right), \tag{8}
$$

as well as

$$
\Theta_n(z,t) = \exp\left(\sum_{j=1}^{\infty} \frac{z^j}{j} A_n(j) \big(t^j - (t-1)^j \big) \right), \tag{9}
$$

with $A_n(j)$ *as in* (7). *Moreover, assume that* $t \neq 0$ *and that the* $a_1, \ldots, a_n > 0$ *are all distinct. Then,*

$$
\Theta_n(z,t) = \frac{(t-1)^n}{t^n} + (-1)^n \sum_{j=1}^{n} \left(\frac{\prod_{m=1}^{n} \left(\frac{1-t}{a_j t} + \frac{1}{a_m} \right)}{\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n} \left(\frac{1}{a_j} - \frac{1}{a_\ell} \right)} \right) \cdot \frac{1}{zt - \frac{1}{a_k}}. \tag{10}
$$

Remark 11. Very similar results also hold for $n \to \infty$ and infinite multisets $\mathcal{M}_{\infty,k}$ (6). In special cases it is possible to derive the complete generating function $T(x, z, t) = \sum_{n \geqslant 0} \Theta_n(z, t) x^n = \sum_{n \geqslant 0} \sum_{k \geqslant 0} \theta_{n;k}(t) x^n z^k$. For $(a_j) = (1)_{j \in \mathbb{N}}$ we get

$$T(x, z, t) = \frac{1 - zt}{(1 - zt)(1 - x) - zx}.$$

For $(a_j) = (\frac{1}{j})$ we get an ordinary hypergeometric function

$$T(x, z, t) = {}_2F_1(1, 1 - z(t-1), 1 - zt; x).$$

Similar, but more involved, partial fraction decompositions exist when the a_m are not all distinct.

Before we state the proof of Theorem 10 we collect a result on partial fraction decomposition.

Lemma 12 (Partial fraction decomposition)**.** *Let $p_n(z)$ and $q_n(z)$ denote monic polynomials of degree n with distinct zeros given by $-\alpha_k$ and β_k, $1 \leqslant k \leqslant n$, respectively, such that $\{-\alpha_1, \ldots, -\alpha_n\} \cap \{\beta_1, \ldots, \beta_n\} = \emptyset$. The rational function $p_n(z)/q_n(z)$ has the partial fraction decomposition*

$$\frac{p_n(z)}{q_n(z)} = \prod_{k=1}^{n} \frac{z + \alpha_k}{z - \beta_k} = 1 + \sum_{j=1}^{n} \left(\frac{p_n(\beta_j)}{\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n} (\beta_j - \beta_\ell)} \right) \cdot \frac{1}{z - \beta_j}.$$

Proof of Theorem 10. First, we derive (8) using symbolic constructions from analytic combinatorics, see Flajolet and Sedgewick [8]. Let $\mathcal{Z}_m = \{m\}$ be a combinatorial class of size one, $1 \leqslant m \leqslant n$. Due to the sequence construction we can describe the class of multisets $\mathcal{B}_m$ of $\mathcal{Z}_m$ as follows

$$\mathcal{B}_m = \mathrm{SEQ}(\mathcal{Z}_m) = \{\epsilon\} + \mathcal{Z}_m + \mathcal{Z}_m \times \mathcal{Z}_m + \mathcal{Z}_m \times \mathcal{Z}_m \times \mathcal{Z}_m + \ldots;$$

Thus, the generating function

$$B_m(z) = \sum_{\beta \in \mathcal{B}_m} w(\beta) t^{\sigma(\beta)} z^{|\beta|} = 1 + \sum_{\epsilon \neq \beta \in \mathcal{B}_m} w(\beta) t^{|\beta|-1} z^{|\beta|}$$

is given by

$$B_m(z) = 1 + \sum_{j=1}^{\infty} t^{j-1} a_m^j z^j = 1 + \frac{a_m z}{1 - a_m t z}.$$

All multisets $\mathcal{M}_n = \bigcup_{k=1}^{\infty} \mathcal{M}_{n,k}$ of $[n]$ can be combinatorially described by

$$\mathcal{M}_n = \mathcal{B}_1 \times \mathcal{B}_2 \times \cdots \times \mathcal{B}_n.$$

Hence, the generating function $\Theta_n(z, t)$ is given by

$$\Theta_n(z, t) = \prod_{m=1}^{n} B_m(z) = \prod_{m=1}^{n} \left(1 + \frac{a_m z}{1 - a_m t z} \right).$$

Next, we use the $\exp - \log$ representation to obtain the expression (9) for $\Theta_n(z,t)$.

$$\Theta_n(z,t) = \prod_{m=1}^{n} \Big(1 + \frac{a_m z}{1 - a_m z t}\Big) = \prod_{m=1}^{n} \exp\Big(\log\Big(1 + \frac{a_m z}{1 - a_m z t}\Big)\Big)$$

$$= \exp\Big(\sum_{m=1}^{n} \log\big(1 - a_m z(t-1)\big) - \sum_{m=1}^{n} \log\big(1 - a_m z t\big)\Big).$$

We use (7) and expand the two logarithm functions:

$$\Theta_n(z,t) = \exp\Big(\sum_{m=1}^{n}\sum_{j=1}^{\infty} \frac{a_m^j z^j}{j}\big(t^j - (t-1)^j\big)\Big) = \exp\Big(\sum_{j=1}^{\infty} \frac{z^j}{j} A_n(j)\big(t^j - (t-1)^j\big)\Big).$$

Finally, we turn to (10). We assume that $t \neq 0$ and that the $a_1, \ldots, a_n > 0$ are all distinct. Then, we can write $\Theta_n(z,t)$ as follows:

$$\Theta_n(z,t) = \frac{(t-1)^n}{t^n} \cdot \prod_{m=1}^{n} \frac{z + \frac{1}{a_m(1-t)}}{z - \frac{1}{a_m t}}.$$

Consequently, applying partial fraction decomposition (see Lemma 12) with $\alpha_k = \frac{1}{a_k(1-t)}$ and $\beta_k = \frac{1}{a_k t}$ gives

$$\Theta_n(z,t) = \frac{(t-1)^n}{t^n}\Big(1 + \sum_{j=1}^{n} \Big(\frac{\prod_{m=1}^{n}\big(\frac{1}{a_j t} + \frac{1}{a_m(1-t)}\big)}{\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n}\big(\frac{1}{a_j t} - \frac{1}{a_\ell t}\big)}\Big) \cdot \frac{1}{z - \frac{1}{a_j t}}\Big).$$

Multiplying the sum with $\frac{(t-1)^n}{t^n}$ directly gives the stated expression (10). $\qquad\square$

As a consequence of the theorem before, we obtain alternative expressions for the values $\theta_{n;k}(t)$ (5). We use the complete Bell polynomials $B_n(x_1, \ldots, x_n)$, which are defined via the identity

$$\exp\Big(\sum_{\ell \geqslant 1} \frac{z^\ell}{\ell!} x_\ell\Big) = \sum_{j \geqslant 0} \frac{B_j(x_1, \ldots, x_j)}{j!} z^j. \tag{11}$$

Theorem 13. *The values $\theta_{n;k}(t)$ (5) can be expressed in various ways:*

- *in terms of values $\theta_{n;k}(0)$ and $\theta_{n;k}(1)$:*

$$\theta_{n;k}(t) = \sum_{j=0}^{k} t^j \theta_{n,j}(1) \cdot (1-t)^{k-j} \theta_{n,k-j}(0). \tag{12}$$

- *Bell polynomials and the power sums $A_n(j)$ (7):*

$$\theta_{n;k}(t) = \frac{1}{k!} B_k\big(0! A_n(1)\big(t^1 - (t-1)^1\big), \ldots, (k-1)! A_n(k)\big(t^k - (t-1)^k\big)\big)$$

$$= \sum_{m_1 + 2m_2 + \cdots = k} \frac{1}{m_1! m_2! \ldots} \Big(\frac{A_n(1)(t^1 - (t-1)^1)}{1}\Big)^{m_1} \Big(\frac{A_n(2)(t^2 - (t-1)^2)}{2}\Big)^{m_2} \ldots. \tag{13}$$

- *Determinantal expression: let $\alpha_n(j;t) = A_n(j)(t^j - (t-1)^j)$, it holds*

$$\theta_{n;k}(t) = \frac{1}{k!} \begin{vmatrix} \alpha_n(1;t) & -1 & 0 & \dots & 0 \\ \alpha_n(2;t) & \alpha_n(1;t) & -2 & \dots & 0 \\ \dots & \dots & \dots & \vdots & \dots \\ \alpha_n(k-1;t) & \alpha_n(k-2;t) & \alpha_n(k-3;t) & \dots & -(k-1) \\ \alpha_n(k;t) & \alpha_n(k-1;t) & \alpha_n(k-2;t) & \dots & \alpha_n(1;t) \end{vmatrix}. \quad (14)$$

- *Combinatorial sum: assume that $t \neq 0$ and that the values $a_1, \dots, a_n > 0$ are all distinct. Then,*

$$\theta_{n;k}(t) = (-1)^{n-1} \sum_{j=1}^{n} \left(\frac{\prod_{m=1}^{n} \left(\frac{1-t}{a_j t} + \frac{1}{a_m} \right)}{\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n} \left(\frac{1}{a_j} - \frac{1}{a_\ell} \right)} \right) \cdot a_j^{k+1} t^k. \quad (15)$$

Example 14. For $t = \frac{1}{2}$ we solely sum over the odd indices $m_1, m_3, \dots$, since only $m_2 = m_4 = \dots = 0$ lead to a positive contribution. We obtain

$$\theta_{n;k}\left(\frac{1}{2}\right) = \frac{1}{k!} B_k\left(0! A_n(1) \cdot 2 \cdot \frac{1}{2^1}, 0, 2! A_n(1) \cdot 2 \cdot \frac{1}{2^3} \dots, (k-1)! A_n(k) \cdot \left(\frac{1}{2^k} - \left(-\frac{1}{2}\right)^k\right) \right)$$

$$= \sum_{m_1 + 3m_3 + \dots = k} \frac{2^{m_1 + m_3 + \dots}}{2^{1 \cdot m_1 + 3 \cdot m_3 + \dots} m_1! m_3! \dots} \left(\frac{A_n(1)}{1} \right)^{m_1} \left(\frac{A_n(3)}{3} \right)^{m_3} \dots.$$

A direct byproduct of this result is an expression for $\zeta_n^t(\{m^k\})$. We state the formula for $\zeta_n^t(\{1^k\})$, generalizing the already known results for $t = 0$, $\zeta_n(\{1\}_k)$ and $t = 1$, $\zeta_n^\star(\{1\}_k)$; see for example [28] and the references therein.

Example 15. Given the weight sequence $\boldsymbol{a} = \left(\frac{1}{j}\right)_{j \in \mathbb{N}}$ we obtain for $\theta_{n;k}(t) = \zeta_n^t(\{1\}_k)$ the results

$$\zeta_n^t(\{1\}_k) = \sum_{j=0}^{k} t^j \zeta_n^\star(\{1\}_j) \cdot (1-t)^{k-j} \zeta_n(\{1\}_{k-j}).$$

as well as

$$\zeta_n^t(\{1\}_k) = \frac{1}{k!} B_k \left(0! H_n^{(1)} \left(t^1 - (t-1)^1 \right), \dots, (k-1)! H_n^{(k)} \left(t^k - (t-1)^k \right) \right)$$

$$= \sum_{m_1 + 2m_2 + \dots = k} \frac{1}{m_1! m_2! \dots} \left(\frac{H_n^{(1)} \left(t^1 - (t-1)^1 \right)}{1} \right)^{m_1} \left(\frac{H_n^{(2)} \left(t^2 - (t-1)^2 \right)}{2} \right)^{m_2} \dots.$$

Moreover, we obtain for $t \neq 0$ the identity

$$\zeta_n^t(\{1\}_k) = t^k \cdot \sum_{j=1}^{n} \binom{n + \frac{1-t}{t} \cdot j}{n} \binom{n}{j} \frac{(-1)^{j-1}}{j^k}.$$

For $t = \frac{1}{2}$ we obtain as a special case a truncated analog of the formula of Hoffman and Ihara [16, eqn. (41)]:

$$\zeta_n^{\frac{1}{2}}(\{1\}_k) = \frac{1}{2^k k!}\left(B_k(0! \cdot 2H_n^{(1)}, 0, 2! \cdot 2H_n^{(3)}, 0, \dots)\right.$$

$$= \sum_{m_1+3m_3+5m_5\cdots=k} \frac{2^{m_1+m_3+m_5+\cdots}}{2^k m_1! m_3! \dots}\left(\frac{H_n^{(1)}}{1}\right)^{m_1}\left(\frac{H_n^{(3)}}{3}\right)^{m_3}\cdots.$$

The combinatorial sum extends a result of Prodinger for $t = \frac{1}{2}$ [33] (see also [32, 34, 36]), originally stated solely in terms of the combinatorial sum and the expression in terms of Bell polynomials (11):

$$\zeta_n^{\frac{1}{2}}(\{1\}_k) = \frac{1}{2^k}\sum_{j=1}^{n}\binom{n+j}{j}\binom{n}{j}(-1)^{j-1}\frac{1}{j^k}.$$

Finally, we mention that for $t = 1$ the combinatorial sum turns into a well-known and often rediscovered formula [3, 4, 9, 28]:

$$\zeta_n^{\star}(\{1\}_k) = \sum_{j=1}^{n}\binom{n}{j}(-1)^{j-1}\frac{1}{j^k}.$$

Proof of Theorem 13. First, we turn to the expression (12) for $\theta_{n,k}(t)$. From Theorem 10 and (8) we obtain

$$\Theta_n(z,t) = \prod_{m=1}^{n}\left(1 + \frac{a_m z}{1 - a_m z t}\right) = \prod_{m=1}^{n}\frac{1 + a_m z(1-t)}{1 - a_m z t}$$

$$= \left(\sum_{j\geqslant 0}(zt)^j \theta_{n,j}(1)\right)\cdot\left(\sum_{j\geqslant 0}z^j(1-t)^j\theta_{n,j}(0)\right).$$

Thus, extraction of coefficients gives the stated result:

$$\theta_{n,k}(t) = [z^k]\left(\sum_{j\geqslant 0}(zt)^j\theta_{n,j}(1)\right)\cdot\left(\sum_{j\geqslant 0}z^j(1-t)^j\theta_{n,j}(0)\right)$$

$$= \sum_{j=0}^{k}t^j\theta_{n,j}(1)\cdot(1-t)^{k-j}\theta_{n,k-j}(0).$$

For (13) we use (9) of Theorem 10. The definition of the Bell polynomials (11) and extraction of coefficients, $\theta_{n,k}(t) = [z^k]\Theta_n(z,t)$, then directly gives the stated results, with the Bell polynomials (11) evaluated at $x_\ell = (\ell-1)!A_n(\ell)\left(t^\ell - (t-1)^\ell\right)$. Following Hoffman [14], the determinant (14) can be obtained using the theory of symmetric functions.

Let $x_1, x_2, \ldots$ denote variables of degree one. As in Remark 6 let h_k denote the complete symmetric functions and p_k the power sums. There exists polynomials Q_k such that

$$h_k = Q_k(p_1, \ldots, p_k).$$

Let $H(z)$ and $P(z)$ denote the generating functions of h_k and p_k:

$$H(z) = \sum_{k=0}^{\infty} h_k z^k = \prod_{i \geqslant 1} \frac{1}{1 - z x_i}, \quad P(z) = \sum_{k=1}^{\infty} p_k z^k = \sum_{i \geqslant 1} \frac{x_i}{1 - z x_i}.$$

Then,

$$H(z) = \exp\left(\int_0^z P(v) dv\right) = \exp\left(-\sum_{i \geqslant 1} \ln(1 - z x_i)\right) = \exp\left(\sum_{j \geqslant 1} z^j \cdot \frac{p_j}{j}\right).$$

This gives the Bell polynomial type expression for the Q_k:

$$Q_k(p_1, \ldots, p_k) = \sum_{m_1 + 2m_2 + \cdots = k} \frac{1}{m_1! m_2! \ldots} \left(\frac{p_1}{1}\right)^{m_1} \left(\frac{p_2}{2}\right)^{m_2} \cdots.$$

On the other hand, MacDonald [30] gives the determinant

$$Q_k(y_1, \ldots, y_k) = \frac{1}{k!} \begin{vmatrix} y_1 & -1 & 0 & \ldots & 0 \\ y_2 & y_1 & -2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \vdots & \ldots \\ y_{k-1} & y_{k-2} & y_{k-3} & \ldots & -(k-1) \\ y_k & y_{k-1} & y_{k-2} & \ldots & y_1 \end{vmatrix},$$

which proves the stated result. Finally, the combinatorial expression (15) is readily obtained by extraction of coefficients from the partial fraction decomposition (10) of $\Theta_n(z, t)$ in Theorem 10. $\square$

4 Further generalizations

4.1 Refinements of the parameter σ

Definition 16 (Refined parameter sigma). Let $\vec{\ell} \in \mathcal{M}_{n,k}$ (3). Then $\sigma^{(i)}(\vec{\ell}) = \sigma_{n,k}^{(i)}(\vec{\ell})$ is defined as the number equal signs in $\vec{\ell}$ stemming from numbers $i \in \mathbb{N}$. Then, as functions acting from $\mathcal{M}_{n,k}$ to $\mathbb{N}_0$:

$$\sigma = \sum_{i=1}^{n} \sigma^{(i)}.$$

Example 17. Given $\vec{\ell} = (1, 1, 2, 2, 2, 4, 6) \in \mathcal{M}_{n,7}$ then $\sigma^{(1)} = 1$, $\sigma^{(2)} = 2$ and $\sigma^{(i)} = 0$, $i > 2$.

In the following we use the vector notation $\vec{t}^{\vec{\sigma}(\vec{\ell})} = t_1^{\sigma_1(\vec{\ell})} \ldots t_n^{\sigma_n(\vec{\ell})}$.

Definition 18 (Multi-interpolated weighted multiset sums). Given $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$, let $\theta_{n;k}(\vec{t}) = \theta_{n;k}(\vec{t}; \boldsymbol{a})$ be defined as

$$\theta_{n;k}(\vec{t}) = \sum_{\vec{\ell} \in \mathcal{M}_{n,k}} w(\vec{\ell}) \vec{t}^{\vec{\sigma}(\vec{\ell})} := \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k} t_1^{\sigma^{(1)}(\vec{\ell})} \ldots t_n^{\sigma^{(n)}(\vec{\ell})}.$$

The generating function $\Theta_n(z, \vec{t}) = \sum_{k \geqslant 0} \theta_{n;k}(\vec{t}) z^k = \sum_{\ell \in \mathcal{M}_n} w(\vec{\ell}) \vec{t}^{\vec{\sigma}(\vec{t})} z^{|\vec{\ell}|}$ can readily be obtained using the symbolic methods.

Theorem 19 (Generating functions and power sums). *The generating function $\Theta_n(z, \vec{t})$ is given by*

$$\Theta_n(z, \vec{t}) = \prod_{m=1}^{n} \left(1 + \frac{a_m z}{1 - a_m z t_m} \right) = \exp \left(\sum_{j=1}^{\infty} \frac{z^j}{j} \sum_{m=1}^{n} a_m^j \left(t_m^j - (t_m - 1)^j \right) \right).$$

Moreover, assume that for $1 \leqslant m \leqslant n$ we have $a_m \cdot t_m > 0$ and distinct. Then,

$$\Theta_n(z, \vec{t}) = \prod_{m=1}^{n} \frac{t_m - 1}{t_m} + (-1)^n \sum_{j=1}^{n} \left(\frac{\prod_{m=1}^{n} \left(\frac{1 - t_m}{a_j t_j} + \frac{1}{a_m} \right)}{\prod_{\substack{\ell=1 \\ \ell \neq j}}^{n} \left(\frac{t_\ell}{a_j t_j} - \frac{1}{a_\ell} \right)} \cdot \frac{1}{z t_j - \frac{1}{a_j}} \right).$$

The proof of the theorem is very similar to the proof of Theorem 13 and therefore omitted.

Example 20 (Even-odd interpolated truncated multiple zeta functions). For even indices m let $t_m = t_E$, whereas for odd indices let $t_m = t_O$. Then,

$$\theta_{n;k}(t_E, t_O) = \sum_{\vec{\ell} \in \mathcal{M}_{n,k}} w(\vec{\ell}) t^{\sigma^E(\vec{\ell})} t^{\sigma^O(\vec{\ell})}$$

The generating function is

$$\Theta_n(z, t_E, t_O) = \exp \left(\sum_{j=1}^{\infty} \frac{z^j}{j} \left[\left(t_O^j - (t_O - 1)^j \right) \sum_{\substack{m=1 \\ m \text{ odd}}}^{n} a_m^j + \left(t_E^j - (t_E - 1)^j \right) \sum_{\substack{m=1 \\ m \text{ even}}}^{n} a_m^j \right] \right).$$

We consider for example $a_m = \frac{1}{m}$, and assume $n = 2N$. Then,

$$\sum_{\substack{m=1 \\ m \text{even}}}^{n} a_m^j = \frac{1}{2^j} \sum_{m=1}^{N} \frac{1}{m^j} = \frac{1}{2^j} H_N^{(j)}$$

and

$$\sum_{\substack{m=1 \\ m \text{ odd}}}^{n} a_m^j = \sum_{m=1}^{N} \frac{1}{(2m-1)^j} = H_{2N}^{(j)} - \frac{1}{2^j} H_N^{(j)}.$$

Thus,

$$\Theta_{2N}(z, t_E, t_O) = \exp \left(\sum_{j=1}^{\infty} \frac{z^j}{j} \left[\left(t_O^j - (t_O - 1)^j \right) \left(H_{2N}^{(j)} - \frac{1}{2^j} H_N^{(j)} \right) + \left(t_E^j - (t_E - 1)^j \right) \frac{1}{2^j} H_N^{(j)} \right] \right).$$

4.2 Sums of multisets: partitions

Yet another direction would be to introduce a new parameter $p = p_{n,k}$ on $\mathcal{M}_{n,k}$ (3), with p the one-norm of $\vec{\ell}$. It measures the numbers $p(\vec{\ell})$ the partitions $\vec{\ell} \in \mathcal{M}_{n,k}$ induce: $p(\vec{\ell}) = \|\vec{\ell}\|_1 = \sum_{i=1}^{k} \ell_k$. Then,

$$\theta_{n;k}(t,q) = \sum_{\vec{\ell} \in \mathcal{M}_{n,k}} w(\vec{\ell}) q^{p(\vec{\ell})} t^{\sigma(\vec{\ell})} = \sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} a_{\ell_1} a_{\ell_2} \ldots a_{\ell_k} q^{\sum_{i=1}^{k} \ell_k} t^{\sigma(\vec{\ell})}.$$

Note that this can alternatively be achieved by a change of weights $a_m \mapsto a_m \cdot q^m$. Then, the generating function $\Theta_n(z,t,q) = \sum_{k \geqslant 0} \theta_{n;k}(t,q)$ satisfies

$$\Theta_n(z,t,q) = \prod_{m=1}^{n} \left(1 + \frac{a_m q^m z}{1 - a_m q^m zt} \right).$$

It is then possible to study the distribution of p, or the joint distribution of p and σ on $\mathcal{M}_{n,k}$.

Example 21 (Partitions - distributions). Setting $t = 1$ we the generating function $\Theta_n(z,1,q) = \sum_{\vec{\ell} \in \mathcal{M}_n} w(\vec{\ell}) q^{p(\vec{\ell})} z^{|\vec{\ell}|}$ simplifies to $\prod_{m=1}^{n} \left(\frac{1}{1 - a_m q^m z} \right)$.

Example 22 (Partitions and infinite multisets). Moreover, setting $a_m = z = t = 1$ and letting $n \to \infty$ in $\Theta_n(z,t,q)$ leads us directly to the generating function of the partition function:

$$\Theta_\infty(1,1,q) = \sum_{\vec{\ell} \in \mathcal{M}_\infty} q^{p(\vec{\ell})} = \prod_{m=1}^{\infty} \frac{1}{1 - q^m}.$$

5 Distributions and limit laws

Before, we treated the variable t as a parameter, mainly between zero and one, interpolating between weighted k-subsets and k-multisets or truncated multiple zeta values and their star counterparts. Here, we are interested in the distribution of σ and introduce a random variable $S_{n,k}$. In the following we denote with $[z^j]$ the extraction of coefficients operator, $[z^j] f(x) = [z^j] \sum_{k=0}^{\infty} f_k \cdot z^k = f_j$, $j \in \mathbb{N}_0$.

Definition 23 (Random variable $\boldsymbol{S_{n,k}}$). Given the weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$ with positive real weights a_j. The random variable $S_{n,k}$ counts the number of equal signs in an element $\vec{\ell} \in \mathcal{M}_{n,k}$ (3). Its probability mass function is defined in terms of $\theta_{n;k}(t)$ (5):

$$\mathbb{P}\{S_{n,k} = j\} := \frac{[t^j] \theta_{n;k}(t)}{\theta_{n;k}(1)}, \quad 0 \leqslant j \leqslant k - 1;$$

equivalently, the probability generating function $\mathbb{E}(t^{S_{n,k}})$ is given by

$$\mathbb{E}(t^{S_{n,k}}) = \frac{\theta_{n;k}(t)}{\theta_{n;k}(1)}.$$

Note the the actual support of $S_{n,k}$ is the range $\max\{0, k-n\} \leqslant j \leqslant k-1$, since for $k \geqslant n$ at most n elements out of a total of k can be unequal.

The expected value

$$\mathbb{E}(S_{n,k}) = \frac{\sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} \sigma(\vec{\ell}) w(\vec{\ell})}{\sum_{n \geqslant \ell_1 \geqslant \cdots \geqslant \ell_k \geqslant 1} w(\vec{\ell})}$$

can be obtained by extraction of coefficients:

$$\mathbb{E}(S_{n,k}) = \frac{\theta'_{n;k}(1)}{\theta_{n;k}(1)} = \frac{[z^k] E_t \frac{\partial}{\partial t} \Theta_n(z,t)}{\theta_{n;k}(1)}. \tag{16}$$

Here E_t denotes the evaluation operator at $t = 1$. The variance is determined via the second factorial moment $\mathbb{E}(S_{n,k}^2)$,

$$\mathbb{V}(S_{n,k}) = \mathbb{E}(S_{n,k}^{\underline{2}}) + \mathbb{E}(S_{n,k}) - \mathbb{E}(S_{n,k})^2, \tag{17}$$

with

$$\mathbb{E}(S_{n,k}^{\underline{2}}) = \frac{\theta''_{n;k}(1)}{\theta_{n;k}(1)} = \frac{[z^k] E_t \frac{\partial^2}{\partial t^2} \Theta_n(z,t)}{\theta_{n;k}(1)}.$$

5.1 Interpolated k-element multisets

We choose the weight sequence $\boldsymbol{a} = (1)_{j \in \mathbb{N}}$ and obtain the distribution of $S_{n,k}$

Theorem 24 (Distribution - k-multisets). *The probability mass function of the random variable $S_{n,k}$ in random k multisets of $[n]$, is given by*

$$\mathbb{P}\{S_{n,k} = j\} = \frac{\binom{n}{k-j}\binom{k-1}{j}}{\binom{n+k-1}{k}}, \quad 0 \leqslant j \leqslant k-1.$$

Thus, $S_{n,k} \stackrel{(d)}{=} Hypergeo(n + k - 1, k - 1, k)$ follows a hypergeometric distribution. Its expected value and variance are given by

$$\mathbb{E}(S_{n,k}) = \frac{k(k-1)}{n+k-1}, \quad \mathbb{V}(S_{n,k}) = \frac{k(k-1)n(n-1)}{(n+k-1)^2(n+k-2)}.$$

Moreover, the factorial moments are given by

$$\mathbb{E}(S_{n,k}^{\underline{s}}) = \frac{k^{\underline{s}} \cdot (k-1)^{\underline{s}}}{(n+k-1)^{\underline{s}}}, \quad s \geqslant 1.$$

Proof. The probability mass function is readily obtained by extraction of coefficients. First, we note that

$$\Theta_n(z,t) = \prod_{m=1}^{n} \left(1 + \frac{z}{1-zt}\right) = \left(\frac{1-zt+z}{1-zt}\right)^n.$$

Then,

$$\mathbb{P}\{S_{n,k} = j\} = \frac{[t^j]\theta_{n,k}(t)}{\theta_{n,k}(1)} = \frac{[z^k t^j]\Theta_{n,k}(z,t)}{\binom{n+k-1}{k}} = \frac{1}{\binom{n+k-1}{k}}[z^{k-j}(zt)^j]\left(\frac{1-zt+z}{1-zt}\right)^n$$

$$= \frac{\binom{n}{k-j}}{\binom{n+k-1}{k}}[(zt)^j]\frac{1}{(1-zt)^{k-j}} = \frac{\binom{n}{k-j}\cdot\binom{k-1}{j}}{\binom{n+k-1}{k}}.$$

Alternatively, observe that j equal signs are distributed amongst $k-1$ places, leading to $\binom{k-1}{j}$, and that the remaining $k-j$ sums give a factor

$$\sum_{n \geqslant \ell_1 > \cdots > \ell_{k-j} \geqslant 1} 1 = |\mathcal{R}_{n,k-j}| = \binom{n}{k-j}.$$

The expected value, the variance and the factorial moments of the hypergeometric distribution are well known, or directly obtained using the Vandermonde identity. $\qquad\square$

Theorem 25 (Limit laws - $\boldsymbol{k}$-multisets). *The limit laws for $\max\{n,k\} \to \infty$ are given by three different distributions:*

1. *Degenerate case: for $n \to \infty$ and $k = o(\sqrt{n})$: $S_{n,k} \to 0$; similarly, for $k \to \infty$ and $n = o(\sqrt{k})$: $S_{n,k} - k + n \to 0$.*

2. *Poisson range: for $k \to \infty$ and $n \sim c\sqrt{k}$: $S_{n,k} - k + n \to Poisson(c^2)$; similarly, for $k, n \to \infty$ and $k \sim c \cdot \sqrt{n}$ with $c > 0$: $S_{n,k} \to Poisson(c^2)$.*

3. *Normal range: for $k, n \to \infty$ and $\sqrt{n} \ll k \ll n^2$:*

$$\frac{S_{n,k} - \mathbb{E}(S_{n,k})}{\sqrt{\mathbb{V}(S_{n,k})}} \to \mathcal{N}(0,1).$$

Proof. First we note that the expected value of the random variable $S_{n,k}$, counting the number of equal signs in random k multisets of $[n]$, has the following behavior:

$$\mathbb{E}(S_{n,k}) \sim \begin{cases} \frac{k^2}{n}, & n \to \infty \text{ and } k = o(\sqrt{n}), \\ c, & k, n \to \infty \text{ and } k \sim c \cdot \sqrt{n}, \\ k - n + c^2, & k \to \infty \text{ and } n \sim c\sqrt{k}, \\ k - n + \frac{n^2}{k}, & k \to \infty \text{ and } n = o(\sqrt{k}), \end{cases}$$

The Poisson limit law for $k \sim c \cdot \sqrt{n}$ follows directly from the asymptotic expansion of the factorial moments:

$$\mathbb{E}(S_{n,k}^{\underline{s}}) = \frac{k^{\underline{s}} \cdot (k-1)^{\underline{s}}}{(n+k-1)^{\underline{s}}} \sim (c^2)^s.$$

Thus, by the method of moments the (factorial) moments $S_{n,k}$ converge to the (factorial) moments of a Poisson distributed random variable with parameter $\lambda = c^2$. In the range

$k \to \infty$ and $n \sim c\sqrt{k}$ the Poisson limit law is indicated by the asymptotic expansions of $\mathbb{E}(S_{n,k})$ and $\mathbb{V}(S_{n,k})$. We obtain elementarily

$$\mathbb{P}\{S_{n,k} = k - n + \ell\} = \frac{\binom{n}{n-\ell}\binom{k-1}{k-n+\ell}}{\binom{n+k-1}{k}}.$$

By our assumption $n \sim c\sqrt{k}$ so $\binom{n}{n-\ell} \sim (nc)^{\ell}/\ell!$. We use the asymptotic expansion of the factorials,

$$n! \sim \frac{n^n}{e^n}\sqrt{2\pi n}$$

as well as a precise expansion of the terms

$$(n+k-1)^{n+k-1}, \quad (k-n+\ell)^{k-n+\ell},$$

using the $\exp - \log$ representation. This gives the desired result

$$\mathbb{P}\{S_{n,k} = k - n + \ell\} \sim \frac{n^{2\ell}(c^2)^{\ell}}{\ell!k^{\ell}}e^{-c^2}, \quad \ell \geqslant 0.$$

The normal limit laws for the hypergeometric distribution are classical and well known, see Nicholson [31] or Feller [7] and the discussion in the introduction; the proofs are omitted. $\qquad\square$

5.2 Interpolated truncated multiple zeta values

We consider the weight sequence $\boldsymbol{a} = (1/j)_{j\in\mathbb{N}}$ and the distribution of $S_{n,k}$, defined by the probability generating function

$$\mathbb{E}(t^{S_{n,k}}) = \frac{\theta_{n,k}(t)}{\theta_{n,k}(1)} = \frac{\zeta_n^t(\{1\}_k)}{\zeta_n^\star(\{1\}_k)}.$$

For example, the boundary values are given by

$$\mathbb{P}\{S_{n,k} = 0\} = \frac{\zeta_n(\{1\}_k)}{\zeta_n^\star(\{1\}_k)}, \quad \mathbb{P}\{S_{n,k} = k - 1\} = \frac{H_n^{(k)}}{\zeta_n^\star(\{1\}_k)}.$$

We derive the expected value

$$\mathbb{E}(S_{n,k}) = \frac{1}{\zeta_n^\star(\{1\}_k)}\sum_{n\geqslant\ell_1\geqslant\cdots\geqslant\ell_k\geqslant 1}\frac{\sigma(\vec{\ell})}{\ell_1^m\cdots\ell_k^m}$$

using (16).

Theorem 26. *The expected value is given by the following exact expression:*

$$\mathbb{E}(S_{n,k}) = \frac{1}{\zeta_n^\star(\{1\}_k)}\sum_{\ell=0}^{k-2}H_n^{(\ell+2)}\zeta_n^\star(\{1\}_{k-\ell-2}).$$

Note that the variance can be obtained using the second factorial moment (17), but it is more involved. The limit laws for $S_{n,k}$ are technically much more involved compared to the case $\boldsymbol{a} = (1)_{j \in \mathbb{N}}$. The cases of either n fixed and $k \to \infty$ or $n \to \infty$ and k fixed are not too difficult to analyze. In contrast, for both $n, k \to \infty$ the asymptotic expansions of the denominator $\zeta_n^\star(\{1\}_k)$ require a more precise analytic combinatorial analysis; see the works of Hwang [18, 20, 22] and Bai et al. [2]. The case of n and $k \to \infty$ indicates that a normal limit should appear. Note that the limit law, given by the sum of independent Bernoulli random variables, also appeared in the analysis of algorithms literature; amongst others, in the the unsuccessful search in binary search trees [6], the degree of the root in increasing trees [6], distances in increasing trees [5, 6] or edge-weights in increasing trees [25, 29].

Proposition 27. *For $k \to \infty$ and n fixed,*

$$k - S_{n,k} \xrightarrow{(d)} D_n = B_1 \oplus B_2 \oplus \cdots \oplus B_n,$$

where $B_j = Be(\frac{1}{j})$ denote independent Bernoulli-distributed random variables; hence the probability mass function of D is given by

$$\mathbb{P}\{D = \ell\} = \frac{\zeta_{n-1}(\{1\}_{\ell-1})}{n}, \quad 1 \leqslant \ell \leqslant n.$$

For $n \to \infty$ and k fixed the random variable $S_{n,k}$ degenerates: $S_{n,k} \xrightarrow{(d)} 0$.

Remark 28. The growth range of the degenerate case can be extended and made more precise using the results of Hwang [19] for $\zeta_n(\{1\}_k)$, stated in terms of the unsigned Stirling numbers of the first kind, and his results [22] (see also Bai et al. [2]) for $\zeta_n^\star(\{1\}_k)$, related to the expected number of maxima in hypercubes.

Remark 29. It is well known that the sum of the limit law of D_n, normalized and centralized, for $n \to \infty$ is asymptotically normal:

$$\frac{D_n - \log n}{\sqrt{\log n}} \xrightarrow{(d)} \mathcal{N}(0, 1).$$

See for example Dobrow and Smythe [6] for an approach using Poisson approximation or Hwang [21]. Thus, we expect that the random variable $S_{n,k}$ is asymptotically normal at least in a certain range of n, k both tending to infinity. We observe that $k - S_{n,k} \sim D_n$, or $S_{n,k} \sim k - D_n$, where D_n is concentrated around $\log n$. Thus $k - \log n$ should govern the behavior of $S_{n,k}$ and also its total mass $\zeta_n^\star(\{1\}_k)$. Note further, that even the asymptotic expansions of $\zeta_n^\star(\{1\}_k)$ for $k, n \to \infty$ derived earlier by Hwang [18, 20, 22] are more involved.

Proof. (Expected value). Our starting point is the generating function

$$\Theta_n(z, t) = \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m}}{1 - \frac{zt}{m}} \right).$$

Differentiating with respect to t gives

$$\frac{\partial}{\partial t}\Theta_n(z,t) = \Theta_n(z,t) \cdot \sum_{j=1}^{n} \frac{z^2}{j^2(1-\frac{zt}{j})^2} \cdot \frac{1}{1+\frac{\frac{z}{j}}{1-\frac{zt}{j}}}$$

$$= \Theta_n(z,t) \cdot \sum_{j=1}^{n} \frac{z^2}{j^2(1-\frac{zt}{j})(1-\frac{zt}{j}+\frac{z}{j})}.$$

Evaluation at $t = 1$ gives

$$E_t\frac{\partial}{\partial t}\Theta_n(z,t) = \Theta_n(z,1) \cdot \sum_{j=1}^{n} \frac{z^2}{j^2(1-\frac{z}{j})}.$$

Extraction of coefficients leads to the stated result:

$$\mathbb{E}(S_{n,k}) = \frac{[z^k]E_t\frac{\partial}{\partial t}\Theta_n(z,t)}{[z^k]\Theta_n(z,1)} = \frac{\sum_{j=1}^{n}[z^{k-2}]\frac{\Theta_n(z,1)}{1-\frac{z}{j}}}{\zeta_n^\star(\{1\}_k)},$$

noting that $[z^\ell]\Theta_n(z,1) = \zeta_n^\star(\{1\}_\ell)$. $\qquad\qquad\square$

Proof. (Limit laws). For $n \to \infty$ and k fixed we study

$$\mathbb{P}\{S_{n,k}=0\} = \frac{\zeta_n(\{1\}_k)}{\zeta_n^\star(\{1\}_k)}.$$

We use the representation given in Corollary 15 for $t = 0$ and $t = 1$. We require the asymptotic expansion of the harmonic numbers H_n for $n \to \infty$:

$$H_n = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \mathcal{O}\Big(\frac{1}{n^4}\Big).$$

Asymptotically, the summand $m_1 = k$ dominates and gives for both $t = 0$ and $t = 1$ the same value $\frac{1}{k!}\log_n^k$. Consequently, $\mathbb{P}\{S_{n,k}=0\} \sim \frac{\log^k n}{\log^k n} = 1$ for large n.

For $k \to \infty$ and n fixed we use singularity analysis [9] and derive the asymptotic equivalent of $\mathbb{P}\{S_{n,k}=k-j\}$, $1 \leqslant j \leqslant n$. We have

$$\mathbb{P}\{S_{n,k}=k-j\} = \frac{1}{\zeta_n^\star(\{1\}_k)}[z^k t^{k-j}]\Theta_n(z,t) = \frac{1}{\zeta_n^\star(\{1\}_k)}[z^k t^{k-j}]\prod_{m=1}^{n}\Big(1+\frac{\frac{z}{m}}{1-\frac{zt}{m}}\Big).$$

Since $z^k t^{k-j} = u^{k-j}z^j$ for $u = zt$ we get

$$\mathbb{P}\{S_{n,k}=k-j\} = \frac{1}{\zeta_n^\star(\{1\}_k)}[z^j u^{k-j}]\prod_{m=1}^{n}\Big(1+\frac{\frac{z}{m}}{1-\frac{u}{m}}\Big).$$

The product has a dominant singularity is at $u = 1$ and can be written as

$$\prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m}}{1 - \frac{u}{m}}\right) = \frac{1}{1-u} \cdot R_n(u, z),$$

with $R_n(u, z)$ analytic inside a circle of radius 2. Consequently, for $k \to \infty$

$$[u^{k-j}] \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m}}{1 - \frac{u}{m}}\right) \sim R_n(1, z) = n \cdot \prod_{m=1}^{n} \left(1 + \frac{z-1}{m}\right).$$

It remains to derive the asymptotic expansion of $\zeta_n^\star(\{1\}_k)$. We can use again singularity analysis as before; alternatively, the well known [9, 2, 28] binomial sum representation directly gives us the desired result:

$$\zeta_n^\star(\{1\}_k) = \sum_{\ell=1}^{n} \binom{n}{\ell} \frac{(-1)^{\ell+1}}{\ell^k} \sim n.$$

Finally, combining our results gives

$$\mathbb{P}\{S_{n,k} = k - j\} \sim [z^j] \prod_{m=1}^{n} \left(1 + \frac{t-1}{m}\right), \quad 1 \leqslant j \leqslant n.$$

The product is exactly the probability generating function of the independent Bernoulli random variables with success probability $\frac{1}{m}$. Extraction of coefficients directly leads to the stated result, using the representation of $\zeta_n(\{1\}_k)$ given in [28]. $\qquad \square$

5.3 Interpolated truncated multiple zeta values - only twos

We consider the weight sequence $\boldsymbol{a} = (1/j^2)_{j \in \mathbb{N}}$ and the distribution of $S_{n,k}$, defined by the probability generating function

$$\mathbb{E}(t^{S_{n,k}}) = \frac{\theta_{n,k}(t)}{\theta_{n,k}(1)} = \frac{\zeta_n^t(\{2\}_k)}{\zeta_n^\star(\{2\}_k)}.$$

For example, the boundary values are given by

$$\mathbb{P}\{S_{n,k} = 0\} = \frac{\zeta_n(\{2\}_k)}{\zeta_n^\star(\{2\}_k)}, \quad \mathbb{P}\{S_{n,k} = k - 1\} = \frac{H_n^{(2k)}}{\zeta_n^\star(\{2\}_k)}.$$

We obtain the following limit laws.

Theorem 30. *For $\max\{n, k\} \to \infty$ the limit laws for $S_{n,k}$ are given by three different distributions:*

1. For $n \to \infty$ and k fixed: $S_{n,k} \xrightarrow{(d)} S_{\infty,k}$, with

$$\mathbb{P}\{S_{\infty,k} = j\} = \frac{1}{\zeta^\star(\{2\}_k)} \cdot \sum_{\substack{\mathbf{p} \in \mathcal{P}_O(k) \\ \mathcal{L}(\mathbf{p}) = k-j}} \zeta(2 \cdot \mathbf{p}), \quad 0 \leqslant j \leqslant k - 1.$$

2. *For $k \to \infty$ and n fixed: $k - S_{n,k} \xrightarrow{(d)} D_n$. The random variable D_n is given by the sum of independent Bernoulli random variables:*

$$D_n = B_1 \oplus B_2 \oplus \cdots \oplus B_n,$$

where $B_j = Be(\frac{1}{j^2})$ denote independent Bernoulli-distributed random variables. Moreover, $\mathbb{E}(D_n^s) = s!\zeta_{n-1}(\{2\}_s)$, $s \geqslant 1$.

3. *For $k, n \to \infty$: $k - S_{n,k} \xrightarrow{(d)} D$;*

$$D = B_1 \oplus B_2 \oplus \cdots = \bigoplus_{m=1}^{\infty} B_m.$$

Here $B_j = Be(\frac{1}{j^2})$, $j \geqslant 1$, denote independent Bernoulli-distributed random variables. Moreover, $\mathbb{E}(D^s) = s!\zeta(\{2\}_s)$, $s \geqslant 1$.

Remark 31. The random variable D is exactly the limit law of the number of cuts in a recursive tree to isolate a leaf [26]; modified weight sequences lead to other families of increasing trees. Moreover, D_n is closely related the the distribution of the number of cuts in a tree of size n.

Proof. Similar to (2) we have,

$$\zeta_n^t(\{2\}_k) = \sum_{\circ = ``,"\text{or} ``+"} t^\sigma \zeta_n(\underbrace{2 \circ 2 \circ \cdots \circ 2}_{k}) = \sum_{\mathbf{p} \in \mathcal{P}_O(k)} t^{k - \mathcal{L}(\mathbf{p})} \zeta_n(2 \cdot \mathbf{p}).$$

By taking the limit $n \to \infty$, k being fixed, and extraction of coefficients we directly obtain the stated result.

For $k \to \infty$ and arbitrary n we use singularity analysis [9] and derive the asymptotic equivalent of $\mathbb{P}\{S_{n,k} = k - j\}$, $1 \leqslant j$. We have

$$\mathbb{P}\{S_{n,k} = k - j\} = \frac{1}{\zeta_n^\star(\{2\}_k)} [z^k t^{k-j}] \Theta_n(z, t) = \frac{1}{\zeta_n^\star(\{2\}_k)} [z^k t^{k-j}] \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m^2}}{1 - \frac{zt}{m^2}}\right).$$

Setting as before $u = zt$, we get

$$\mathbb{P}\{S_{n,k} = k - j\} = \frac{1}{\zeta_n^\star(\{2\}_k)} [z^j u^{k-j}] \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m^2}}{1 - \frac{u}{m^2}}\right).$$

The product has a dominant singularity is at $u = 1$ and can be written as

$$\prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m^2}}{1 - \frac{u}{m^2}}\right) = \frac{1}{1 - u} \cdot R_n(u, z),$$

with $R_n(u, z)$ analytic inside a circle of radius 2. Consequently, for $k \to \infty$

$$[u^{k-j}] \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m^2}}{1 - \frac{u}{m^2}}\right) \sim R_n(1, z) = \frac{2n}{n+1} \cdot \prod_{m=1}^{n} \left(1 + \frac{z-1}{m^2}\right).$$

For the asymptotic expansion of $\zeta_n^\star(\{2\}_k)$ we can use again singularity analysis and obtain $\zeta_n^\star(\{2\}_k) \sim \frac{2n}{n+1}$. Finally, combining our results gives

$$\mathbb{P}\{S_{n,k} = k - j\} \sim [z^j] \prod_{m=1}^{n} \left(1 + \frac{t-1}{m^2}\right), \quad 1 \leqslant j \leqslant n.$$

The product is exactly the probability generating function of n independent Bernoulli random variables each with success probability $\frac{1}{m^2}$ for $1 \leqslant m \leqslant n$. $\square$

5.4 Refined decomposition of sigma, random vectors and marginals

Similarly to the refinements of the parameter σ into $\sigma^{(i)}$, $1 \leqslant i \leqslant n$, the random variable $S_{n,k}$ can be decomposed into $S_{n,k}^{(i)}$: $S_{n,k} = \sum_{i=1}^{n} S_{n,k}^{(i)}$. Given the weight sequence $\boldsymbol{a} = (a_j)_{j \in \mathbb{N}}$, the joint distribution of the random vector $\mathbf{S}_{n,k} = (S_{n,k}^{(1)}, \ldots, S_{n,k}^{(n)})$ is determined by

$$\mathbb{P}\{\mathbf{S}_{n,k} = \vec{j}\} = [\vec{t}^{\vec{j}}] \frac{\theta_{n;k}(\vec{t})}{\theta_{n;k}(\vec{1})}.$$

Another possibility is to distinguish only between odd and even indices, such that $S_{n,k} = S_{n,k}^{(O)} + S_{n,k}^{(E)}$, and to study the bivariate probability generating function

$$\mathbb{E}(t_O^{S_{n,k}^{(O)}} t_E^{S_{n,k}^{(E)}}) = \frac{\theta_{n;k}(t_O, t_E)}{\theta_{n;k}(1, 1)}.$$

For truncated zeta values with the weight sequence $\boldsymbol{a} = (1)_{j \in \mathbb{N}}$, the random vector $\mathbf{S}_{n,k}$ has the probability generating function

$$\mathbb{E}(\mathbf{t}^{\mathbf{S}_{n,k}}) = \frac{1}{\binom{n+k-1}{k}} [z^k] \prod_{m=1}^{n} \left(1 + \frac{z}{1 - zt_m}\right).$$

Example 32 (Truncated zeta values - random vector). For truncated zeta values with weight sequence $\boldsymbol{a} = (1/j)_{j \in \mathbb{N}}$, the random vector $\mathbf{S}_{n,k}$ has the probability generating function

$$\mathbb{E}(\mathbf{t}^{\mathbf{S}_{n,k}}) = \frac{1}{\zeta^\star(\{1\}_k)} [z^k] \prod_{m=1}^{n} \left(1 + \frac{\frac{z}{m}}{1 - \frac{zt_m}{m}}\right).$$

Example 33 (Truncated zeta values - marginals). In contrast to the multiset case, the marginals $S_{n,k}^{(i)}$ are not exchangeable anymore. The probability generating functions are given by

$$\mathbb{E}(t^{S_{n,k}^{(i)}}) = \frac{1}{\zeta^\star(\{1\}_k)} [z^k] (1 - \frac{z}{i}) \left(1 + \frac{\frac{z}{i}}{1 - \frac{zt}{i}}\right) \cdot \prod_{m=1}^{n} \left(\frac{1}{1 - \frac{z}{m}}\right).$$

It is expected that the marginals are asymptotically independent, at least in some growth range of k and n.

Theorem 34 (Marginals - interpolated multisets)**.** *The marginals $S_{n,k}^{(i)}$ are identically distributed, but not independent; the sequence $\mathbf{S}_{n,k}$ is exchangeable. The probability mass function is given by*

$$\mathbb{P}\{S_{n,k}^{(i)} = j\} = \begin{cases} \dfrac{\binom{n-2+k}{k} + \binom{n-3+k}{k}}{\binom{n+k-1}{k}}, & j = 0, \\ \dfrac{\binom{n-3+k-j}{k-j-1}}{\binom{n+k-1}{k}}, & 1 \leqslant j \leqslant k-1. \end{cases}$$

The factorial moments, $s \geqslant 1$, are given by

$$\mathbb{E}(S_{n,k}^{\underline{s}}) = s! \cdot \frac{k^{\underline{s+1}}}{(n+k-1)(n+s-1)^{\underline{s}}}.$$

Theorem 35. *We obtain the following limit laws, depending on the growth of k and n:*

- *for $k \to \infty$ and n fixed the normalized random variable $\frac{S_{n,k}}{k}$ converges to a Beta-distributed random variable $\frac{S_{n,k}}{k} \xrightarrow{(d)} B(1, n-1)$.*

- *for $k/n \to \infty$ and $k, n \to \infty$ the normalized random variable $\frac{n}{k} S_{n,k}$ converges to a standard exponentially distributed random variable, $\frac{n}{k} S_{n,k} \xrightarrow{(d)} Exp(1)$.*

- *for $k/n \to c > 0$ and $k, n \to \infty$ the random variable $S_{n,k}$ converges to a (modified) geometric distribution G:*

$$\mathbb{P}\{G = j\} = \frac{c^{j+1}}{(1+c)^{j+2}}, \quad j \geqslant 1 \quad , \mathbb{P}\{G = 0\} = \frac{1}{1+c} + \frac{c}{(1+c)^2}.$$

- *for $k/n \to 0$ and $n \to \infty$ the random variable $S_{n,k}$ degenerates, $S_{n,k} \xrightarrow{(d)} 0$.*

Remark 36. The random variable $S_{n,k}^{(i)}$ is closely related to the number of descendants $D_{n+k,k}$ of node labelled k in a recursive trees of size $n + k$ [24]. Thus, $S_{n,k}^{(i)}$ can be described in terms of a Polya-urn model and the asymptotics for $k \to \infty$ and n fixed can be refined to almost sure convergence; moreover a classical result is available in terms of a martingale tail sum [11].

Proof of Theorem 34. Let $1 \leqslant i \leqslant n$. The probability generating function of the marginal distribution $S_{n,k}^{(i)}$ is given by

$$\mathbb{E}(t^{S_{n,k}^{(i)}}) = \frac{1}{\binom{n+k-1}{k}} [z^k] \frac{1}{(1-z)^{n-1}} \cdot \left(1 + \frac{z}{1-zt}\right).$$

By the structure of the probability generating function, the random variables are identically distributed and the random vector is exchangable. Thus, we can directly obtain the expected value from $\mathbb{E}(S_{n,k})$.

$$\mathbb{E}(S_{n,k}^{(i)}) = \frac{1}{n}\mathbb{E}(S_{n,k}) = \frac{k(k-1)}{n(n+k-1)}.$$

Additionally, the probability mass function is obtained readily by extraction of coefficients. All factorial moments, $s \geqslant 1$, can be obtained in a straightforward way:

$$\mathbb{E}(S_{n,k}^{\underline{s}}) = [z^k]E_t\frac{\partial^s}{\partial t^s}\mathbb{E}(t^{S_{n,k}^{(i)}}) = \frac{s!}{\binom{n+k-1}{k}}[z^k]\frac{z^{s+1}}{(1-z)^{n+s}}$$

$$= \frac{s!\binom{n+k-2}{k-s-1}}{\binom{n+k-1}{k}} = s!\cdot\frac{k^{\underline{s+1}}}{(n+k-1)(n+s-1)^{\underline{s}}}. \qquad \square$$

Proof of Theorem 35. We use the method of moments and derive asymptotic expansions of the factorial moments. For n fixed and $k \to \infty$ we get

$$\mathbb{E}(S_{n,k}^{\underline{s}}) = s!\cdot\frac{k^{\underline{s+1}}}{(n+k-1)(n+s-1)^{\underline{s}}} \sim \frac{s!}{(n+s-1)^{\underline{s}}}\cdot k^s.$$

Consequently, the power moments of $S_{n,k}/k$ are asymptotically given by $\frac{s!}{(n+s-1)^{\underline{s}}}$, which proves the Beta limiting distributions.

For $n \to \infty$ we observe that

$$\mathbb{E}(S_{n,k}^{\underline{s}}) \sim s!\cdot\lambda_{n,k}^s\cdot\frac{k}{n+k}, \quad \text{with } \lambda_{n,k} = \frac{k}{n}.$$

Thus, the factorial moments are almost of mixed Poisson type [27] with standard exponential mixing distribution; the additional factor $\frac{k}{n+k}$ can be explained by the definition of the parameter σ, which influences the discrete limit case. This directly leads to the stated limit laws using Lemma 2 of [27]. Alternatively, the discrete limit for $k/n \to c$ can be directly obtained as follows:

$$\mathbb{P}\{S_{n,k}^{(i)} = j\} = \frac{\binom{n-3+k-j}{k-j-1}}{\binom{n+k-1}{k}} = \frac{k^{\underline{j+1}}\cdot(n-1)}{(n+k-1)^{\underline{j+2}}} \sim \frac{c^{j+1}}{(1+c)^{j+2}}, \quad j \geqslant 1.$$

The remaining case $j = 0$ is treated in a similar way. For $k/n \to 0$ we observe that

$$\mathbb{P}\{S_{n,k}^{(i)} = 0\} = \frac{n-1}{n+k-1} + \frac{k(n-1)}{(n+k-1)^{\underline{2}}} \sim 1. \qquad \square$$

5.5 Sum theorem for interpolated multiple zeta values

Yamamoto [38] established, amongst many other things, the sum theorem for interpolated multiple zeta values:

$$\sum_{\substack{k_1 \geqslant 2, k_i \geqslant 1 \\ \sum_{\ell=1}^{n} k_\ell = k}} \zeta^t(k_1, \ldots, k_n) = \zeta(k)\cdot\sum_{j=0}^{n-1}\binom{k-1}{j}t^j(1-t)^{n-1-j}, \tag{18}$$

with $k > n$. Note that the case $t = 1$ gives the sum theorem for multiple zeta star values

$$\sum_{\substack{k_1 \geqslant 2, k_i \geqslant 1 \\ \sum_{\ell=1}^{n} k_\ell = k}} \zeta^\star(k_1, \ldots, k_n) = \zeta(k) \cdot \binom{k-1}{n-1},$$

whereas $t = 0$ gives the ordinary sum theorem.

Using (18), we can study the distribution of the parameter σ on sums of interpolated multiple zeta values with the same depth n and weight k. Let $\mathcal{S}_{n,k}$ denote the random variable with probability generating function

$$\mathbb{E}(t^{\mathcal{S}_{n,k}}) = \frac{\zeta(k) \cdot \sum_{j=0}^{n-1} \binom{k-1}{j} t^j (1-t)^{n-1-j}}{\zeta(k) \cdot \binom{k-1}{n-1}} = \frac{\sum_{j=0}^{n-1} \binom{k-1}{j} t^j (1-t)^{n-1-j}}{\binom{k-1}{n-1}}.$$

Let $B_{j,n}(t)$ denote the Bernstein polynomials

$$B_{j,n}(t) = \binom{n}{j} t^j (1-t)^{n-j}, \quad 0 \leqslant j \leqslant n.$$

We obtain the following result.

Theorem 37. *The probability generating function $\mathbb{E}(t^{\mathcal{S}_{n,k}})$ is given by a Bernstein form of degree $n-1$ with Bezier coefficients $\beta_j = \frac{1}{\binom{k-1-j}{k-n}}$*

$$\mathbb{E}(t^{\mathcal{S}_{n,k}}) = \sum_{j=0}^{n-1} \beta_j \cdot B_{j,n-1}(t).$$

Let $\mathcal{R}_{n,k} = n - \mathcal{S}_{n,k}$. The probability mass function $\mathbb{P}\{\mathcal{R}_{n,k} = i\}$, is given by

$$\mathbb{P}\{\mathcal{R}_{n,k} = i\} = \frac{\binom{k-1-i}{j-2}}{\binom{k-1}{j-1}}, \quad 1 \leqslant i \leqslant n,$$

setting with $n = k - j + 1$ and $2 \leqslant j \leqslant k$. Consequently, $\mathcal{R}_{n,k}$, with $n = k - j + 1$, has the same distribution as the random variable $D_{n,j}$, counting the number of descendants of node j in a random recursive tree of size n: $\mathcal{R}_{n,k} \overset{(d)}{=} D_{n,j}$.

As a byproduct of our identification of $\mathcal{R}_{n,k}$, we get the following limit laws from [24].

Corollary 38. *The limiting distribution behaviour of the random variable $\mathcal{R}_{n,k} = n - \mathcal{S}_{n,k}$, with $n = k - j + 1$ and $2 \leqslant j \leqslant k$, is, for $k \to \infty$ and depending on the growth of $j = j(k)$, characterized as follows.*

- *for j fixed, $\frac{\mathcal{R}_{n,k}}{k} \overset{(d)}{\longrightarrow} \beta(1, j-1)$*

- *for small j: $j \to \infty$ such that $j = o(k)$: the normalized random variable $\frac{j}{n} \mathcal{R}_{n,k}$ is asymptotically Exponential distributed,*

- *for j: $j \to \infty$ such that $j \sim \rho k$, with $0 < \rho < 1$. The shifted random variable $\mathcal{R}_{n,k} - 1$ is asymptotically negative binomial-distributed, $\mathcal{R}_{n,k} - 1 \xrightarrow{(d)} NegBin(1, \rho)$,*

- *for large j: $j \to \infty$ such that $\ell := n - j = o(n)$: $\mathcal{R}_{n,k} \to 1$, i.e. $\lim_{k \to \infty} \mathbb{P}\{\mathcal{R}_{n,k} = 1\} = 1$.*

Proof. We note that by definition, the summands are weighted Bernstein polynomials:

$$\mathbb{E}(t^{\mathcal{S}_{n,k}}) = \frac{1}{\binom{k-1}{n-1}} \sum_{j=0}^{n-1} \frac{\binom{k-1}{j}}{\binom{n-1}{j}} B_{j,n-1}(t).$$

Simplification gives the stated expression for the Bezier coefficients. The probability mass function is obtained by extraction of coefficients:

$$\mathbb{P}\{\mathcal{S}_{n,k} = i\} = [t^i]\mathbb{E}(t^{\mathcal{S}_{n,k}}) = \frac{1}{\binom{k-1}{n-1}} \sum_{j=0}^{i} \binom{k-1}{j} [t^{i-j}](1-t)^{n-1-j}$$

$$= \frac{1}{\binom{k-1}{n-1}} \sum_{j=0}^{i} \binom{k-1}{j} \binom{n-1-j}{i-j} (-1)^{i-j}.$$

Using an identity for binomial coefficients [10] then readily gives

$$\mathbb{P}\{\mathcal{S}_{n,k} = i\} = \frac{\binom{k-n+i-1}{k-n-1}}{\binom{k-1}{n-1}}, \quad 0 \leqslant i \leqslant n-1,$$

and thus we get the corresponding result for $\mathcal{R}_{n,k} = n - \mathcal{S}_{n,k}$. We observe that for $n = k - j + 1$ the probability mass functions and thus the distribution of $\mathcal{R}_{n,k}$ and the random variable $D_{n,j}$ (case $c_2 = 0$) in [24] coincide and thus we can directly transfer the limit laws for $D_{n,j}$. $\square$

6 Summary and Outlook

We introduced a parameter σ on weighted k-element multisets and studied properties of it using symbolic combinatorics. This allows to prove several relations for truncated interpolated multiple zeta values $\zeta_n^t(\{m\}_k)$ (as well as reproving identities for truncated multiple zeta values $\zeta^t(\{m\}_k)$). Introducing refined enumeration leads to new refinements of previous identities. Interpreting the parameter σ as a random variable $S_{n,k}$ leads to several different limit laws, depending on the considered weight sequences and the growth of n and k. Interestingly, the limits laws are closely related to a great many results in combinatorial probability and analytic combinatorics.

It is of interest to complete the analysis of $S_{n,k}$ in the case $\boldsymbol{a} = (1/j)$; we will report on this elsewhere. Similar to the random variable $S_{n,k}$, defined in terms of the parameter $\sigma_{n,k}$, one may also study the distribution of the parameter $p = p_{n,k}$ or the joint distribution of p and σ. Moreover, it is certainly of interest to study the distribution of $S_{n,k}$ for other interesting sequences, compare with Vignat and Wakhare [37] or Hoffman and Mező [17].

Acknowledgements

References

[1] H. Bachmann. Interpolated Schur multiple zeta values. *Journal of the Australian Mathematical Society* 104, 289–307, 2018.

[2] Z.-D. Bai, L. Devroye, H. -K. Hwang, T. -H. Tsai. Maxima in hypercubes. *Random Structures and Algorithms* **27**, 290–309, 2005.

[3] N. Batir. On some combinatorial identities and harmonic sums. *International Journal of Number Theory* **13**, 1695–1709, 2017.

[4] K. Dilcher. Some q-series identities related to divisor functions. *Discrete Mathematics* **145**, 83–93, 1995.

[5] R. Dobrow. On the distribution of distances in recursive trees. *Journal of Applied Probability* 33, 749-757, 1996.

[6] R. Dobrow and R. Smythe. Poisson approximations for functionals of random trees. *Random Structures & Algorithms* 9, 79-92, 1996.

[7] W. Feller. *Intoduction to probability theory and its applications.* John Wiley and Sons, Third Edition, 1968.

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

[9] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: finite differences and Rice's integrals. *Theoretical Computer Science* **144**, 101–124, 1995.

[10] R. L. Graham, D. E. Knuth and O. Patashnik. *Concrete Mathematics.* Addison-Wesley, 1994.

[11] P. Hall, C. C. and Heyde. *Martingale limit theory and its application. Probability and Mathematical Statistics.* Academic Press, New York-London, 1980.

[12] M. E. Hoffman. Multiple harmonic series. *Pacific Journal of Mathematics* **152**, 275-290, 1992.

[13] M. E. Hoffman. Quasi-shuffle products. *Journal of Algebraic Combinatorics* 11, 49–68, 2000.

[14] M. E. Hoffman. Harmonic-number summation identities, symmetric functions, and multiple zeta values. *Ramanujan Journal* **42**, 501–526, 2017.

[15] M. E. Hoffman. Quasi-shuffle algebras and applications. In *Algebraic Combinatorics, Resurgence, Moulds and Applications (CARMA)*, vol. 2, F. Chapoton et al. (eds.), European Mathematical Society Publishing House, Berlin, 327–348, 2020.

[16] M. E. Hoffman and K. Ihara. Quasi-shuffle products revisited. *Journal of Algebra* **481**, 293–326, 2017.

[17] I. Mező and M. E. Hoffman. Zeros of the digamma function and its Barnes G-function analogue. *Integral Transforms and Special Functions* 28, 846–858, 2017.

[18] H.-K. Hwang. Théorèmes limites pour les structures combinatoires et les fonctions arithmétiques. Thèse, Ecole polytechnique, 1994.

[19] H.-K. Hwang. Asymptotic expansions for the Stirling numbers of the first kind. *Journal of Combinatorial Theory, Series A*, 71:2, 343–351, 1995.

[20] H.-K. Hwang. A Poisson ∗ geometric convolution law for the number of components in unlabelled combinatorial structures. *Combinatorics, Probability and Computing*, 7:1, 89–110, 1998.

[21] H.-K. Hwang. Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. *Advances in Applied Probability*, 31, 448-491, 1999.

[22] H.-K. Hwang. Phase changes in random recursive structures and algorithms. *Proceedings of the Workshop on Probability with Applications to Finance and Insurance*, edited by T. L. Lai, H. Yang and S. P. Yung, World Scientific, 82–97, 2004.

[23] K. Ihara, J. Kajikawa, Y. Ohno and J. Okuda. Multiple zeta values vs. multiple zeta-star values. *Journal of Algebra* 332, 187–208, 2011.

[24] M. Kuba and A. Panholzer. Descendants in increasing trees. *Electronic Journal of Combinatorics* 13 (1), Paper 8, 2006.

[25] M. Kuba and A. Panholzer. On edge-weighted recursive trees and inversions in random permutations. *Discrete Mathematics*. Volume 308, Issue 4, 529-540, 2008.

[26] M. Kuba and A. Panholzer. Isolating a leaf in rooted trees via random cuttings. *Annals of Combinatorics* 12, 81–99, 2008.

[27] M. Kuba and A. Panholzer. On moment sequences and mixed Poisson distribution. *Probability Surveys* 13, 89–155, 2016.

[28] M. Kuba and A. Panholzer. A Note on harmonic number identities, Stirling series and multiple zeta values. *International Journal of Number Theory* 15 (7), 1323-1348, 2019.

[29] M. Kuba and S. Wagner. Deterministic edge-weights in increasing tree families, *Combinatorics, Probability and Computing*. Volume 19, issue 01, pp. 99-119, 2010.

[30] I. G. MacDonald. *Symmetric functions and Hall polynomials*. 2nd edition, Clarendon Press, Oxford, 1995.

[31] W. L. Nicholson. On the normal approximation to the hypergeometric distribution. *The Annals of Mathematical Statistics* 27(2), 1956.

[32] H. Prodinger. Human proofs of identities by Osburn and Schneider. *Integers*, 8, article A10, 8 pages (electronic), 2008.

[33] H. Prodinger. Identities involving harmonic numbers that are of interest for physicists. *Utilitas Mathematica* 83, 291–299, 2010.

[34] H. Prodinger and R. Tauraso. New multiple harmonic sum identities. *Electronic Journal of Combinatorics* 21 (2), 2014.

[35] R. Schneider. Partition zeta functions. *Research in Number Theory* 2, art. 9 (17pp), 2016.

[36] J. A. Vermaseren. Harmonic sums, Mellin transforms and integrals. *International Journal of Modern Physics*, A14 (13), 2037–2076, 1999.

[37] C. Vignat and T. Wakhare. Multiple zeta values for classical special functions. *Ramanujan Journal* 51, 519–551, 2020.

[38] S. Yamamoto. Interpolation of multiple zeta and zeta-star values. *Journal of Algebra* **385**, 102–114, 2013.

[39] D. Zagier. Values of zeta functions and their applications. *First European Congress of Mathematics* (Paris, 1992), vol. II (A. Joseph et al., eds.), (*Progress in Mathematics*, vol. 120) Birkhäuser, Boston 1994, pp. 497–512.

[40] W. Zudilin. Algebraic relations for multiple zeta values. *Russian Mathematical Surveys* 58:1 1–29, 2003.

7 Appendix: Auxiliary results about probability distributions

A beta-distributed random variable $Z \stackrel{(d)}{=} \beta(\alpha, \beta)$ with parameters $\alpha, \beta > 0$ has a probability density function given by $f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, where $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ denotes the Beta-function. The (power) moments of Z are given by

$$\mathbb{E}(Z^s) = \frac{\prod_{j=0}^{s-1}(\alpha + j)}{\prod_{j=0}^{s-1}(\alpha + \beta + j)} = \frac{(\alpha + s - 1)^{\underline{s}}}{(\alpha + \beta + s - 1)^{\underline{s}}}, \quad s \geqslant 1.$$

The beta-distribution is uniquely determined by the sequence of its moments. In this work we will discuss a beta-distributed random variable $Z \stackrel{(d)}{=} \beta(1, n-1)$, with moments given by $\mathbb{E}(Z^s) = \frac{s!}{(n+s-1)^{\underline{s}}}$, for $s \geqslant 1$.

An exponentially distributed random variable $Z \stackrel{(d)}{=} \mathrm{Exp}(1)$ with parameter one has density $f(x) = e^{-x}$, $x \geqslant 0$ and power moments $\mathbb{E}(Z^s) = s!$.

A Bernoulli distributed random variable $B \stackrel{(d)}{=} \mathrm{Be}(p)$ with parameter $p \in [0, 1]$ has probability mass function determined by $\mathbb{P}\{B = 1\} = p$, $\mathbb{P}\{B = 0\} = 1 - p$. The sum $Z_n = \bigoplus_{j=1}^{n} B_j$ of n independent Bernoulli distributed random variables $B_j = \mathrm{Be}(p_j)$ has probability generating function

$$\mathbb{E}(v^{Z_n}) = \prod_{j=1}^{n} \mathbb{E}(v^{B_j}) = \prod_{j=1}^{n} \big(1 + (v - 1)p_j\big).$$

The factorial moments $\mathbb{E}(Z_n^{\underline{s}})$ of Z_n are given by

$$\mathbb{E}(Z_n^{\underline{s}}) = E_v \frac{\partial^s}{\partial v^s} \mathbb{E}(v^{Z_n}) = s! \cdot \sum_{n \geqslant \ell_1 > \cdots > \ell_s \geqslant 1} p_{\ell_1} p_{\ell_2} \ldots p_{\ell_s}.$$

Under the assumption that the probabilities p_j tend to zero fast enough, we may define $Z = Z_\infty$ as the sum of infinitely many Bernoulli random variable: $Z = \bigoplus_{j=1}^\infty B_j$, with factorial moments formally given by $s!$ times a multiple series:

$$\mathbb{E}(Z^{\underline{s}}) = s! \cdot \sum_{\ell_1 > \cdots > \ell_s \geqslant 1} p_{\ell_1} p_{\ell_2} \cdots p_{\ell_s}.$$

A Poisson distributed random variable $Z \overset{(d)}{=} \mathrm{Poisson}(\lambda)$, $\lambda > 0$, has probability mass function and factorial moments given by

$$\mathbb{P}\{Z = j\} = \frac{\lambda^j}{j!} \cdot e^\lambda,\ j \geqslant 0, \quad \mathbb{E}(Z^{\underline{s}}) = \lambda^s,\ s \geqslant 1.$$

A hypergeometric distributed random variable $Z \overset{(d)}{=} \mathrm{Hypergeo}(N, K, n)$ with parameters N, K, n has probability mass function and factorial moments given by

$$\mathbb{P}\{Z = j\} = \frac{\binom{K}{j}\binom{N-K}{n-j}}{\binom{N}{n}}, 0 \leqslant j \leqslant n \quad \mathbb{E}(Z^{\underline{s}}) = \frac{K^{\underline{s}} n^{\underline{s}}}{N^{\underline{s}}},\ s \geqslant 1.$$

In particular,
$$\mathbb{E}(Z) = n \cdot \frac{K}{N}, \quad \mathbb{V}(Z) = n \cdot \frac{K(N-K)(N-n)}{N^2(N-1)}.$$

Moreover, the following normal limit law can be deduced (see Nicholson [31] or Feller [7]).

Lemma 39. *Let $Z \overset{(d)}{=} Hypergeo(N, K, n)$ denote a hypergeometric distributed random variable. Under the assumption $\min\{K, N\} \to \infty$ and n such that $\mathbb{E}(Z), \mathbb{V}(Z) \to \infty$, the random variable Z, centered and normalized, is asymptotically standard normal distributed:*
$$\frac{Z - \mathbb{E}(Z)}{\sqrt{\mathbb{V}(Z)}} \xrightarrow{(d)} \mathcal{N}(0, 1).$$