ecajournal.haifa.ac.il

Enumerative Combinatorics and Applications

ECA **4:1** (2024) Article #S2R7
https://doi.org/10.54550/ECA2024V4S1R7

The Sum of Width-one Tensors

William Q. Erickson[†] and Jan Kretschmann[‡*]

[†] *Department of Mathematics, Baylor University*
Email: will_erickson@baylor.edu

[‡] *Department of Mathematical Sciences, University of Wisconsin–Milwaukee*
Email: kretsc23@uwm.edu

Received: June 15, 2023, **Accepted**: August 15, 2023, **Published**: August 25, 2023

Abstract: This paper generalizes a recent result concerning the sum of width-one matrices; in the present work, we consider width-one tensors of arbitrary dimensions. A tensor is said to be width-one if, when visualized as an array, its nonzero entries lie along a path consisting of steps in the positive directions of the standard coordinate vectors. We prove two formulas to compute the sum of all width-one tensors with fixed dimensions and fixed sum of (nonnegative integer) components. The first formula is obtained by converting width-one tensors into tuples of one-row semistandard Young tableaux (thereby inverting the northwest corner rule from optimal transport theory). The second formula, which extracts coefficients from products of multiset Eulerian polynomials, is derived via Stanley–Reisner theory, making use of the EL-shelling of the order complex on the standard basis of tensors.

Keywords: Multiset Eulerian polynomials; Optimal transport; Stanley–Reisner rings; Stanley decompositions
2020 Mathematics Subject Classification: Primary 05E45; Secondary 13F55, 90C27

1. Introduction

Let $\mathcal{T}_{\mathbf{n},s}$ be the set of all $\mathbf{n} = (n_1, n_2, \ldots, n_d)$-dimensional tensors (equivalently, hypermatrices) with nonnegative integer entries summing to s, such that the nonzero entries lie on a single path consisting of steps in the positive directions of the standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_d$. For example, Figure 1 shows a typical element of $\mathcal{T}_{(3,3,3),s}$, where the nonzero entries lie on the lattice points along the marked path.

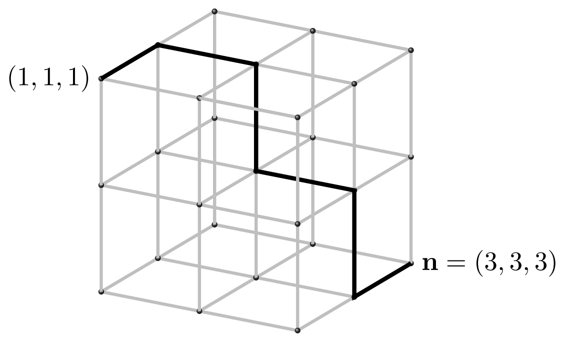

Figure 1: Visualization of the support of an element of $\mathcal{T}_{(3,3,3),s}$.

As shown in [8], computing the sum of such matrices in $d = 2$ dimensions has useful applications in optimal transport, drastically simplifying the problem (first solved recursively in [6]) of computing the expected value of the earth mover's distance (EMD) between two compositions. The methods of [6] were generalized in [9] to find a recursion for the expected value of the generalized EMD between an arbitrary number d of compositions. (For computational treatments of the d-dimensional transport problem, see [2] and [10], for example.) The relationship between the present paper and [9] can be regarded as the d-dimensional analogue of the relationship

*The second author would like to acknowledge partial support through the UWM Advancing Research and Creativity Award.

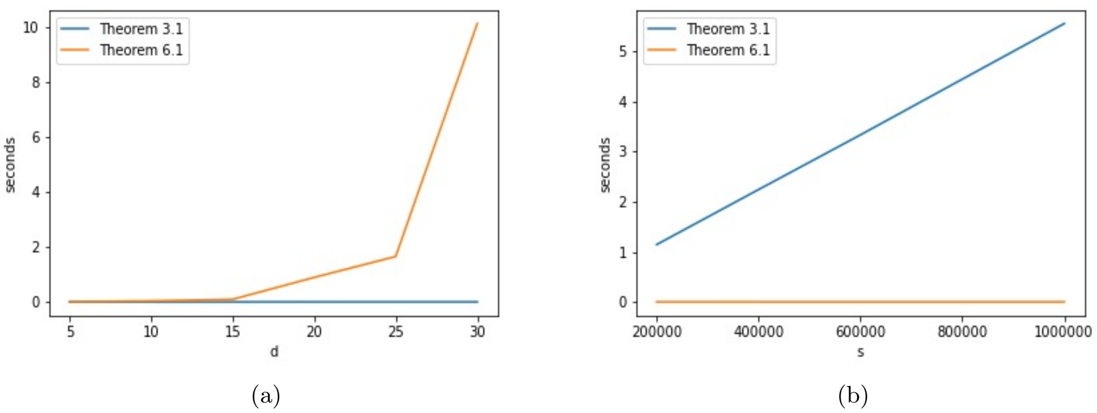

Figure 2: Comparison of computing time with respect to the parameters d and s. In 2a, we fix $s = 5$ and compare the runtime (in seconds) of both approaches for varying d. For arbitrary d, we measure the time it takes to compute the entry at $\mathbf{x} = (\lfloor \frac{d}{2} \rfloor, \dots, \lfloor \frac{d}{2} \rfloor)$ in the d-dimensional hypercube with $\mathbf{n} = (d, \dots, d)$. In 2b, we fix $d = 4$ and let s vary.

between [8] and [6]: in particular, we give two explicit formulas for the sum $\Sigma_{\mathbf{n},s}$ of all tensors in $\mathcal{T}_{\mathbf{n},s}$. These formulas, in turn, can be used to obtain non-recursive formulas for the expected value of the d-fold EMD (easily seen by adapting the argument in [8, §6]).

This connection between width-one tensors and optimal transport theory arises as follows (see [2] for details). In the d-dimensional analogue of the classical transportation problem, the inputs are "supply–demand" vectors $\mathbf{v}_1, \dots, \mathbf{v}_d$ and a "cost" tensor C with dimensions $\mathbf{n}$. Each input vector $\mathbf{v}_i$ has length n_i, and its entries are nonnegative integers summing to s; in combinatorial language, $\mathbf{v}_i$ is a composition of s into n_i parts. The objective is to find a tensor T, with coordinate hyperplane sums prescribed by the $\mathbf{v}_i$, which minimizes the Hadamard product of C and T. This minimum value is said to be the EMD between the input vectors. It turns out that when C has a certain "Monge property" (see [2, Def. 2.1]), there is a greedy algorithm called the northwest corner rule, which outputs the optimal solution T in the form of a width-one tensor. In fact, for fixed s, the northwest corner rule yields a bijection between the set of all possible d-tuples of input vectors and the set $\mathcal{T}_{\mathbf{n},s}$. Therefore, the sum $\Sigma_{\mathbf{n},s}$ can be used to obtain the expected value of the EMD.

In Section 3 we give our first formula for $\Sigma_{\mathbf{n},s}$. We set up a bijection between $\mathcal{T}_{\mathbf{n},s}$ and tuples of one-row semistandard tableaux (essentially the inverse of the northwest corner rule mentioned above), which allows us to write down a formula for our desired sum $\Sigma_{\mathbf{n},s}$ in terms of binomial coefficients (Theorem 3.1). In Sections 4 and 5, we reapproach the problem through the lens of Stanley–Reisner theory. We describe the order complex on the standard basis of the $\mathbf{n}$-dimensional tensors, and we use a special case of an EL-shelling to find the corresponding h-polynomial. This h-polynomial turns out to be a multiset Eulerian polynomial, as we show in Section 5; these polynomials, studied by MacMahon and many others since, enumerate the descents in multiset permutations. We conclude by presenting a second explicit formula for $\Sigma_{\mathbf{n},s}$ using techniques from Stanley–Reisner theory (Theorem 6.1).

Similar to the two formulas for matrices in [8], the two formulas in this paper behave in opposite ways with regard to computing time. Although this issue falls outside the focus of the paper, nonetheless it is not hard to verify that Theorem 3.1 outperforms Theorem 6.1 as s increases for fixed $\mathbf{n}$; the opposite is true, however, as the dimension d or the parameters n_i increase for fixed s. Therefore, as shown in Figure 2, the user who wishes to compute $\Sigma_{\mathbf{n},s}$ should choose between the two theorems according to the sizes of $\mathbf{n}$ and s: roughly speaking, Theorem 6.1 is preferable for low values of d and $|\mathbf{n}|$, while Theorem 3.1 is more efficient for low values of s.

2. Notation and statement of the problem

Throughout the paper, a boldface letter denotes an element of $(\mathbb{Z}_{>0})^d$; in particular, $\mathbf{n} := (n_1, \dots, n_d)$ and $\mathbf{x} := (x_1, \dots, x_d)$. We write $|\mathbf{x}| := \sum_i x_i$, as well as $\min(\mathbf{x}) := \min_i\{x_i\}$ and $\max(\mathbf{x}) := \max_i\{x_i\}$. Let $\mathbf{1} := (1, \dots, 1)$. As usual, we write $[x] := \{1, \dots, x\}$. Define the poset

$$\Pi_{\mathbf{x}} := [x_1] \times \cdots \times [x_d] = \{(a_1, \dots, a_d) : 1 \le a_i \le x_i \text{ for all } i\}, \tag{1}$$

equipped with the product order, so that $\mathbf{a} \le \mathbf{b} \iff a_i \le b_i$ for each $i = 1, \dots, d$. Because our main problem addresses $n_1 \times \cdots \times n_d$ tensors, we will always be working inside $\Pi_{\mathbf{n}}$, but it will be useful to consider the subposets $\Pi_{\mathbf{x}}$, which are the lower-order ideals generated by each $\mathbf{x} \in \Pi_{\mathbf{n}}$.

Recall that a *chain* is a totally ordered subset of $\Pi_{\mathbf{x}}$, and an *antichain* is a subset whose elements are pairwise incomparable. The *width* of a subset $S \subseteq \Pi_{\mathbf{x}}$ is the size of the largest antichain contained in S. Equivalently, by Dilworth's theorem, the width of S equals the minimum number of chains into which S can be partitioned. In particular, S has width 1 if and only if S is a chain.

In this paper, we consider certain tensors of order d. Equivalently, the reader may prefer to consider d-dimensional arrays (also called hypermatrices). Taking the real numbers $\mathbb{R}$ as our ground field, we let

$$\mathcal{T}_{\mathbf{n}} := \mathbb{R}^{n_1} \otimes \cdots \otimes \mathbb{R}^{n_d}$$

denote the space of order-d tensors with dimensions $\mathbf{n}$. Upon fixing the standard basis $\{e_1, \ldots e_{n_i}\}$ for each factor $\mathbb{R}^{n_i}$, every tensor $T \in \mathcal{T}_{\mathbf{n}}$ can be written uniquely in the form

$$T = \sum_{\mathbf{x} \in \Pi_{\mathbf{n}}} T[\mathbf{x}]\, e_{x_1} \otimes \cdots \otimes e_{x_d}, \tag{2}$$

where the scalars $T[\mathbf{x}] \in \mathbb{R}$ are called the *components* of T. Hence T is described completely by its components $T[\mathbf{x}]$, which can be regarded as the entries in an $n_1 \times \cdots \times n_d$ array. The *elementary tensor* $E_{\mathbf{x}}$ is given by

$$E_{\mathbf{x}}[\mathbf{y}] = \delta_{\mathbf{xy}}, \tag{3}$$

where δ is the Kronecker delta. Hence $E_{\mathbf{x}}$ can be regarded as an array with 1 in position $\mathbf{x}$ and 0's elsewhere.

The *support* of a tensor T is the set

$$\operatorname{supp}(T) := \{\mathbf{x} \in \Pi_{\mathbf{n}} : T[\mathbf{x}] \neq 0\}.$$

We say that T is a *width-one* tensor if $\operatorname{supp}(T)$ has width 1 as a subposet of $\Pi_{\mathbf{n}}$. In this paper, we restrict our attention to those width-one tensors whose components are nonnegative integers summing to some positive integer s. We denote this set by

$$\mathcal{T}_{\mathbf{n},s} := \left\{ T \in \mathcal{T}_{\mathbf{n}} : \begin{array}{l} T \text{ is width-one}, \\ T[\mathbf{x}] \in \mathbb{Z}_{\geq 0} \text{ for all } \mathbf{x} \in \Pi_{\mathbf{n}}, \\ \sum_{\mathbf{x}} T[\mathbf{x}] = s. \end{array} \right\}. \tag{4}$$

The main problem of this paper is to write down an explicit formula for the sum of all tensors in $\mathcal{T}_{\mathbf{n},s}$, under the usual componentwise addition. We denote this sum by

$$\Sigma_{\mathbf{n},s} := \sum_{T \in \mathcal{T}_{\mathbf{n},s}} T. \tag{5}$$

3. Main result, first version

We generalize the argument in [8, §3]. It follows from (2) and (3) that each tensor $T \in \mathcal{T}_{\mathbf{n}}$ can be written as a unique sum of elementary tensors

$$T = \sum_{\mathbf{x} \in \Pi_{\mathbf{n}}} T[\mathbf{x}] E_{\mathbf{x}} = \sum_{\mathbf{x} \in \operatorname{supp}(T)} T[\mathbf{x}] E_{\mathbf{x}}.$$

Now suppose that $T \in \mathcal{T}_{\mathbf{n},s}$, as defined in (4). Then by definition, the $\mathbf{x}$'s appearing in the right-hand sum above form a chain in $\Pi_{\mathbf{n}}$. Writing out all of these $\mathbf{x}$'s as column vectors in ascending order, with multiplicities $T[\mathbf{x}]$, we obtain a $d \times s$ matrix. Note that the entries in each row of this matrix are weakly increasing, and the entries in the ith row are elements in $[n_i]$. We will write each row as a row of boxes. (We do this primarily to evoke a one-row semistandard Young tableau, which is the same thing as a weakly increasing sequence.) We thus obtain d rows with s boxes each:

$$\begin{array}{ccccccc}
\boxed{}\boxed{}\boxed{} & \cdots & \boxed{x_1} & \cdots & \boxed{}\boxed{}\boxed{} & & \leftarrow \text{weakly increasing in } [n_1] \\
\boxed{}\boxed{}\boxed{} & \cdots & \boxed{x_2} & \cdots & \boxed{}\boxed{}\boxed{} & & \leftarrow \text{weakly increasing in } [n_2] \\
\vdots & & \vdots & & \vdots & & \vdots \\
\boxed{}\boxed{}\boxed{} & \cdots & \boxed{x_d} & \cdots & \boxed{}\boxed{}\boxed{} & & \leftarrow \text{weakly increasing in } [n_d]
\end{array}$$

This procedure is invertible: given d rows of length s, with entries in the ith row weakly increasing in $[n_i]$, we recover the associated width-one tensor in $\mathcal{T}_{\mathbf{n},s}$ by summing the s elementary tensors $E_{\mathbf{x}}$, for each column $\mathbf{x}$. In our picture above, we have filled in the entries $\mathbf{x} = (x_1, \ldots, x_d)$ for one typical column; note that this column contributes 1 to the component $T[\mathbf{x}]$. It follows that we have a bijection between $\mathcal{T}_{\mathbf{n},s}$ and the set of d-tuples of weakly increasing sequences in $[n_1], \ldots, [n_d]$.

It will be a useful fact that the number of weakly increasing sequences of length ℓ, taken from the set $[p]$, equals the number of (weak) integer compositions of ℓ into p parts, which is well known to be

$$\binom{\ell + p - 1}{\ell} = \binom{\ell + p - 1}{p - 1}. \tag{6}$$

Theorem 3.1. *Let $\Sigma_{\mathbf{n},s}$ be the sum of all tensors in $\mathcal{T}_{\mathbf{n},s}$. For each $\mathbf{x} \in \Pi_{\mathbf{n}}$, we have*

$$\Sigma_{\mathbf{n},s}[\mathbf{x}] = \sum_{j=1}^{s} \prod_{i=1}^{d} \binom{x_i + j - 2}{j - 1} \binom{n_i - x_i + s - j}{s - j}.$$

Proof. For each $T \in \mathcal{T}_{\mathbf{n},s}$, consider its corresponding d-tuple of weakly increasing rows, as described above. Recall that each column $\mathbf{x}$ contributes 1 to the component $T[\mathbf{x}]$. Hence the component $\Sigma_{\mathbf{n},s}[\mathbf{x}]$ equals the number of occurrences of the column $\mathbf{x}$, counted in all possible d-tuples.

First, for fixed j such that $1 \leq j \leq s$, we will find the number of d-tuples whose jth column is $\mathbf{x}$. In other words, we seek the number of d-tuples such that the jth entry in row i is x_i, for each $i = 1, \ldots, d$. For each i, this implies that the $j - 1$ entries to the left of x_i lie in the set $[x_i]$, and that the $s - j$ entries to the right of x_i lie in the set $\{x_i, x_i + 1, \ldots, n_i\}$, which contains $n_i - x_i + 1$ elements. Hence by (6), the number of ways to fill each row i such that the jth entry is x_i equals

$$\binom{(j-1) + x_i - 1}{j - 1} \binom{(s - j) + (n_i - x_i + 1) - 1}{s - j},$$

which simplifies to the expression in the theorem. Taking the product over all rows $i = 1, \ldots, d$, we obtain the number of d-tuples whose jth column equals $\mathbf{x}$, as desired.

Finally, to obtain the number of times $\mathbf{x}$ occurs as any column in a d-tuple, we sum over all columns $j = 1, \ldots, s$. $\square$

Remark 3.2. As mentioned in the introduction, the formula in Theorem 3.1 generalizes the two-dimensional formula in our previous paper [8, Thm. 3.2]. This may not be obvious at first glance, since the two-dimensional version was expressed as a hypergeometric series. Upon setting $d = 2$ in Theorem 3.1 above, one recovers the sum (where each summand is the product of four binomial coefficients) displayed in the proof in [8], immediately before its simplification via hypergeometric identities. (Note, however, that the parameter s in the present paper was denoted by d in [8].)

4. Stanley–Reisner theory

This section, along with the following section, sets out the theory required to prove our second formula for $\Sigma_{\mathbf{n},s}$, which we do in Section 6.

4.1 Abstract simplicial complexes

Our exposition below is standard, citing from [8, §4] almost exactly; further details can be found in [14, Ch. 12]. Given a finite set V, an *(abstract) simplicial complex* on V is a collection Δ of subsets of V such that

- $\{v\} \in \Delta$ for all $v \in V$;
- if $X \in \Delta$ and $Y \subseteq X$, then $Y \in \Delta$.

Elements of V are called *vertices*, and V is called the *vertex set*. The elements of Δ are called *faces*, and the *dimension* of a face is one less than its cardinality. The dimension of Δ is the maximum of the dimensions of its faces.

Let Δ be a nonempty simplicial complex of dimension $a - 1$. We denote by f_i the number of faces of dimension i. In particular, $f_{-1} = 1$ since $\emptyset \in \Delta$. The *f-vector* is the sequence $(f_0, \ldots, f_{a-1})$. It will be more convenient for us to work instead with the *h-vector* $(h_0, \ldots, h_a)$, where the numbers h_k are defined as follows:

$$h_\Delta(t) := \sum_{k=0}^{a} h_k t^k = \sum_{i=0}^{a} f_{i-1} t^i (1 - t)^{a-i}.$$

We call $h_\Delta(t)$ the *h-polynomial* of Δ.

A maximal face in Δ (with respect to inclusion) is called a *facet*. We say that Δ is *pure* if every facet has the same dimension. A pure simplicial complex is said to be *shellable* if there exists an ordering $F_1, \ldots, F_r$ of its facets with the following property: for all $i = 1, \ldots, r$, the power set of F_i has a unique minimal element $\mathrm{R}(F_i)$ not belonging to the subcomplex generated by $F_1, \ldots, F_{i-1}$. Such an ordering is called a *shelling*, and $\mathrm{R}(F_i)$ is called the *restriction* of the facet F_i. Crucial to our method is the following combinatorial description of the h-polynomial: if $F_1, \ldots, F_r$ is a shelling of Δ, then we have

$$h_\Delta(t) = \sum_{i=1}^{r} t^{\#\mathrm{R}(F_i)}. \tag{7}$$

In other words, h_k counts the number of facets whose restrictions have size k:

$$h_k = \#\{i : \#\mathrm{R}(F_i) = k\}.$$

4.2 Lexicographic shellings of posets

Let Π be a finite bounded poset, with $\lessdot$ denoting the covering relation. Let $\mathcal{E} := \{(a, b) : a \lessdot b\}$ be the set of edges of the Hasse diagram of Π. A *labeling* of Π is a function $\lambda : \mathcal{E} \to \mathbb{Z}_{>0}$ assigning a positive integer to each edge of the Hasse diagram. Each labeling λ induces a lexicographic ordering on the set of saturated chains $a_1 \lessdot a_2 \lessdot \cdots \lessdot a_\ell$, via the lexicographic order on the *label sequences* $(\lambda(a_1, a_2), \ldots, \lambda(a_{\ell-1}, a_\ell))$. Following Björner and Wachs [4], we define a special kind of labeling known as an edge-lexicographical (EL) labeling:

Definition 4.1 ([4, Def. 2.1])**.** We say that λ is an *EL-labeling* of Π if, for all $a < c$ in Π, there exists a unique saturated chain $a \lessdot b_1 \lessdot \cdots \lessdot b_\ell \lessdot c$ such that

$$\lambda(a, b_1) \le \lambda(b_1, b_2) \le \cdots \le \lambda(b_\ell, c), \tag{8}$$

and this chain lexicographically precedes all other saturated chains $a \lessdot \cdots \lessdot c$. A chain with the property (8) is called an *ascending chain* with respect to λ.

The *order complex* $\Delta(\Pi)$ is the simplicial complex whose faces are the chains in Π; hence the facets of $\Delta(\Pi)$ are the maximal chains in Π. An EL-labeling of Π induces a shelling order on the facets of $\Delta(\Pi)$, via the lexicographic order on the maximal chains [3, Thm. 2.3]. Note that an EL-labeling does not guarantee that the maximal chains are totally ordered; nevertheless, arbitrarily breaking ties results in a shelling order.

Let λ be an EL-labeling of Π, and F a facet of $\Delta(\Pi)$. An element $b \in F$ is said to be a *descent* of F (with respect to λ) if F contains $a \lessdot b \lessdot c$ such that $\lambda(a, b) > \lambda(b, c)$. With respect to any shelling induced by λ, the restriction of each facet is precisely its set of descents:

$$\mathrm{R}(F) = \{b : b \text{ is a descent of } F\}. \tag{9}$$

(See [5, Thm. 5.8].) We will also use the term *descent* in the context of label sequences, in the obvious sense: namely, i is a descent of the label sequence $(\lambda_1, \ldots, \lambda_\ell)$ if $\lambda_i > \lambda_{i+1}$. In this way, the number of descents in a facet equals the number of descents in its label sequence.

4.3 The Stanley–Reisner ring

Let Δ be a simplicial complex on the vertex set V. Let K be a field, and consider the polynomial ring $K[V] := K[z_v : v \in V]$, where we regard the elements of V as indeterminates. Given a subset $U \subseteq V$, we will use the shorthand

$$z_U := \prod_{v \in U} z_v \qquad \text{and} \qquad K[U] := K[v : v \in U].$$

Let I_Δ be the ideal of $K[V]$ generated by all monomials z_U such that $U \notin \Delta$. Such a U is called a *nonface* of Δ, and it is easy to see that I_Δ is actually generated by those nonfaces which are minimal (i.e., which contain no proper nonface). The *support* of a monomial $m \in K[V]$ is the set $\{v \in V : z_v \text{ divides } m\}$. A K-basis for I_Δ is given by the monomials whose support is not contained in Δ.

The quotient $K[\Delta] := K[V]/I_\Delta$ is called the *Stanley–Reisner* ring of Δ. It is clear that $K[\Delta]$ has a K-basis consisting of the monomials whose support is a face of Δ (where we identify these monomials with their images in the quotient ring).

Each shelling of Δ induces a *Stanley decomposition* of the Stanley–Reisner ring:

$$K[\Delta] = \bigoplus_F K[F]\, z_{\mathrm{R}(F)}, \tag{10}$$

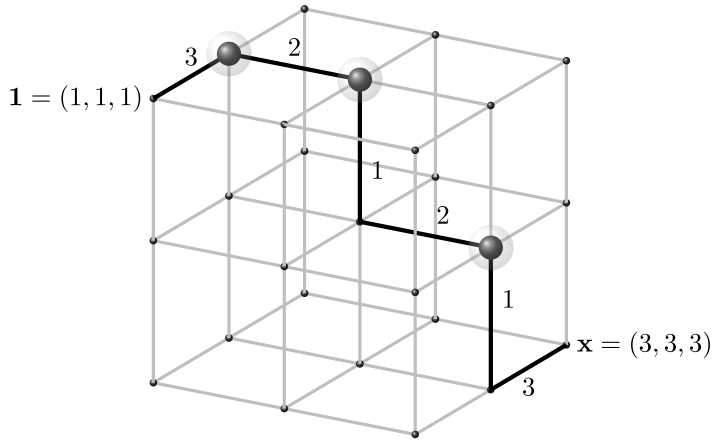

Figure 3: Visualization of a facet F in the order complex $\Delta_{\mathbf{x}}$, where $\mathbf{x} = (3, 3, 3)$. Starting in the upper-left at $\mathbf{1}$, we imagine the coordinate vector $\mathbf{e}_1$ pointing downward, $\mathbf{e}_2$ pointing to the right, and $\mathbf{e}_3$ pointing away from the viewer. With respect to the labeling λ in (13), the label sequence $(3, 2, 1, 2, 1, 3)$ of F is shown in the figure. The descents are indicated by the three large dots; by (9), these are the elements of $\mathrm{R}(F)$.

where the direct sum ranges over the facets F, and their restrictions $\mathrm{R}(F)$ are determined by the shelling. Crucially, each monomial in $K[\Delta]$ lies in exactly one summand of (10). Since I_Δ is generated by homogeneous polynomials (in fact, by monomials), the quotient $K[\Delta]$ inherits from $K[V]$ the natural grading by degree. Writing $K[\Delta]_s$ to denote the graded component consisting of homogeneous polynomials of degree s, we can restrict (10) to a decomposition of each component:

$$K[\Delta]_s = \bigoplus_k \bigoplus_{\substack{F: \\ \#\mathrm{R}(F)=k}} K[F]_{s-k}\, z_{\mathrm{R}(F)}, \tag{11}$$

where k ranges from 0 to the size of the largest restriction $\mathrm{R}(F)$.

4.4 Application to the problem

In this final subsection, we apply the general theory above to the poset $\Pi_{\mathbf{x}}$ defined in (1). We write $\Delta_{\mathbf{x}} := \Delta(\Pi_{\mathbf{x}})$ for its order complex. The facets of $\Delta_{\mathbf{x}}$ are the maximal chains $\mathbf{1} \lessdot \cdots \lessdot \mathbf{x}$. Thus for any facet F of $\Delta_{\mathbf{x}}$, we have

$$\#F = |\mathbf{x}| - d + 1, \tag{12}$$

so $\Delta_{\mathbf{x}}$ is indeed pure.

Let $\mathbf{e}_i$ denote the vector whose ith coordinate is 1, with 0's elsewhere. If $\mathbf{a} \lessdot \mathbf{b}$, then we have $\mathbf{b} = \mathbf{a} + \mathbf{e}_i$ for some $1 \leq i \leq d$. We define the following labeling on $\Pi_{\mathbf{x}}$:

$$\lambda(\mathbf{a}, \mathbf{b}) = i \iff \mathbf{b} = \mathbf{a} + \mathbf{e}_i. \tag{13}$$

For example, if $\mathbf{a} = (3, 6, 4, 1)$ and $\mathbf{b} = (3, 7, 4, 1)$, then $\lambda(\mathbf{a}, \mathbf{b}) = 2$. See Figure 3 for a visualization in the case where $\mathbf{x} = (3, 3, 3)$.

It is easy to see that λ is an EL-labeling: for $\mathbf{a} < \mathbf{b}$, the unique ascending chain $\mathbf{a} \lessdot \cdots \lessdot \mathbf{b}$ with respect to λ is obtained from $\mathbf{a}$ by first adding $\mathbf{e}_1$ a total of $b_1 - a_1$ times, then adding $\mathbf{e}_2$ a total of $b_2 - a_2$ times, etc., and finally adding $\mathbf{e}_d$ a total of $b_d - a_d$ times. Moreover, this chain precedes any other maximal chain between $\mathbf{a}$ and $\mathbf{b}$. Being an EL-labeling, λ induces a unique shelling of $\Delta_{\mathbf{x}}$, since the lexicographical order (in this case) gives a total ordering of the facets.

We now turn to our main problem: writing down a formula for each component of $\Sigma_{\mathbf{n},s}$, which we recall from (5) is the sum of all tensors in $\mathcal{T}_{\mathbf{n},s}$. To this end, note that a K-basis for $K[\Delta_{\mathbf{n}}]$ is given by the monomials whose support is a chain in $\Pi_{\mathbf{n}}$. Restricting to the degree-s component, we observe the bijection

$$\begin{aligned} \mathcal{T}_{\mathbf{n},s} &\longleftrightarrow K\text{-basis of } K[\Delta_{\mathbf{n}}]_s, \\ T &\longleftrightarrow \prod_{\mathbf{x} \in \Pi_{\mathbf{n}}} z_{\mathbf{x}}^{T[\mathbf{x}]}. \end{aligned}$$

Under this correspondence, adding tensors corresponds to multiplying monomials. Therefore, letting m range over the monomials, we have

$$\prod_{m \in K[\Delta_{\mathbf{n}}]_s} m = \prod_{\mathbf{x} \in \Pi_{\mathbf{n}}} z_{\mathbf{x}}^{\Sigma_{\mathbf{n},s}[\mathbf{x}]}. \tag{14}$$

Hence our main problem is equivalent to finding the exponent of each indeterminate $z_{\mathbf{x}}$ in the product of monomials on the left-hand side of (14). We will do this in Section 6. Before that, however, we must explain and exploit the fact that the h-polynomial of $\Delta_{\mathbf{x}}$ is actually a well-known object called the *multiset Eulerian polynomial*.

5. Multiset Eulerian polynomials

Let $\mathbf{p} = (p_1, \ldots, p_d)$. A *multipermutation* of the multiset $\{1^{p_1}, 2^{p_2}, \ldots, d^{p_d}\}$ is a word $\pi = \pi_1 \cdots \pi_{|\mathbf{p}|}$ in which i appears exactly p_i times, for each $1 \le i \le d$. Let $\mathfrak{S}_{\mathbf{p}}$ be the set of all such multipermutations. A *descent* of a multipermutation π is an index i such that $\pi_i > \pi_{i+1}$. Let $\operatorname{des}(\pi)$ denote the number of descents of π. Then the *multiset Eulerian polynomial* $A_{\mathbf{p}}(t)$ is defined to be

$$A_{\mathbf{p}}(t) := \sum_{\pi \in \mathfrak{S}_{\mathbf{p}}} t^{\operatorname{des}(\pi)}.$$

(The special case $A_{\mathbf{1}}(t)$ is just the dth Eulerian polynomial $A_d(t)$, i.e., the descent-generating function over the symmetric group $\mathfrak{S}_d$. See, for example, [13, p. 22], although the convention there is to multiply through by t.) The multiset Eulerian polynomial occurs as the numerator of the following generating function, due to MacMahon [11, p. 211]:

$$\frac{A_{\mathbf{p}}(t)}{(1-t)^{|\mathbf{p}|+1}} = \sum_{\ell=0}^{\infty} \prod_{i=1}^{d} \binom{p_i + \ell}{\ell} t^{\ell}. \tag{15}$$

From (15) we can obtain an explicit expression for the coefficients in $A_{\mathbf{p}}(t)$; see also [7, 1] for combinatorial proofs of the formula below. (These coefficients are known as "Simon Newcomb numbers.") Writing $[t^k]$ for the coefficient of t^k, we have

$$[t^k] A_{\mathbf{p}}(t) = \#\Big\{\pi \in \mathfrak{S}_{\mathbf{p}} : \operatorname{des}(\pi) = k\Big\} = \sum_{\ell=0}^{k} (-1)^{\ell} \binom{|\mathbf{p}|+1}{\ell} \prod_{i=1}^{d} \binom{p_i + k - \ell}{k - \ell}. \tag{16}$$

It is shown in [7, Lemma 2] that the maximum number of descents, i.e., the degree of $A_{\mathbf{p}}(t)$, equals

$$|\mathbf{p}| - \max(\mathbf{p}). \tag{17}$$

For example, if $\mathbf{p} = (3,2,4)$, then $A_{\mathbf{p}}(t) = 24t^5 + 260t^4 + 580t^3 + 345t^2 + 50t + 1$. This is computed directly via (16), and we verify (17) by observing that the degree is indeed $|\mathbf{p}| - \max(\mathbf{p}) = 9 - 4 = 5$.

Lemma 5.1. *The h-polynomial $h_{\mathbf{x}}(t)$ of $\Delta_{\mathbf{x}}$ is the multiset Eulerian polynomial $A_{\mathbf{x}-\mathbf{1}}(t)$.*

Proof. Recall the labeling λ in (13), which induces a shelling of $\Delta_{\mathbf{x}}$. With respect to λ, the label sequence of each facet $\mathbf{1} \lessdot \cdots \lessdot \mathbf{x}$ contains $x_i - 1$ copies of the label i, for each $i = 1, \ldots, d$; conversely, each possible permutation of these labels (where the copies of each i are indistinguishable from each other) is the label sequence of a unique facet. Hence we have a bijection between the set of facets of $\Delta_{\mathbf{x}}$ and the set $\mathfrak{S}_{\mathbf{x}-\mathbf{1}}$. By (9), the size of $\mathrm{R}(F)$ equals the number of descents in F, which in turn equals the number of descents in the label sequence of F. Therefore, comparing (7) and (16), it is clear that $h_{\mathbf{x}}(t) = A_{\mathbf{x}-\mathbf{1}}(t)$. $\square$

Remark 5.2. Combining (7) with the Stanley decomposition (10), it is easy to see that the Hilbert series of a Stanley–Reisner ring $K[\Delta]$ is given by

$$\frac{h_{\Delta}(t)}{(1-t)^{\#F}},$$

where F is any facet of Δ (since Δ is assumed to be pure). Thus by Lemma 5.1 and (12), and by MacMahon's expansion (15), our particular ring $K[\Delta_{\mathbf{n}}]$ has the Hilbert series

$$\frac{A_{\mathbf{n}-\mathbf{1}}(t)}{(1-t)^{|\mathbf{n}|-d+1}} = \sum_{\ell=0}^{\infty} \prod_{i=1}^{d} \binom{n_i + \ell - 1}{\ell} t^{\ell}.$$

This is also the Hilbert series of the coordinate ring of the set of simple (also called pure, or decomposable) tensors in $\mathcal{T}_{\mathbf{n}}$, which is isomorphic to $K[\Delta_{\mathbf{n}}]$. See [12, Thm. 5], where this ring is also viewed as the toric ring defined by the Segre embedding of $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_d}$.

Combining (17) with Lemma 5.1, we note that the degree of $h_{\mathbf{x}}(t) = A_{\mathbf{x}-\mathbf{1}}(t)$ equals

$$|\mathbf{x} - \mathbf{1}| - \max(\mathbf{x} - \mathbf{1}) = |\mathbf{x}| - \max(\mathbf{x}) - d + 1. \tag{18}$$

Equivalently, (18) is the maximum size of $\mathrm{R}(F)$, taken over all facets F of $\Delta_{\mathbf{x}}$.

6. Main result, second version

Theorem 6.1. *For each $\mathbf{x} \in \Pi_{\mathbf{n}}$, we have*

$$\Sigma_{\mathbf{n},s}[\mathbf{x}] = \sum_{k=0}^{\min\{\omega(\mathbf{n},\mathbf{x}),\, s-1\}} \binom{|\mathbf{n}| - d + s - k}{s-k-1} \cdot [t^k]\, A_{\mathbf{x}-\mathbf{1}}(t) A_{\mathbf{n}-\mathbf{x}}(t),$$

where $\omega(\mathbf{n},\mathbf{x}) := |\mathbf{n}| - \max(\mathbf{x}) - \max(\mathbf{n}-\mathbf{x}) - d + 1$.

We record two key lemmas before giving the proof of Theorem 6.1:

Lemma 6.2. *Let $\mathbf{x} \in \Pi_{\mathbf{n}}$, and let $F \ni \mathbf{x}$ be a facet of $\Delta_{\mathbf{n}}$. Then $\binom{|\mathbf{n}|-d+s-k}{s-k-1}$ equals the exponent of $z_{\mathbf{x}}$ in the product of all monomials in*

$$K[F]_{s-k-1} z_{\mathrm{R}(F)\cup\{\mathbf{x}\}}. \tag{19}$$

Proof. It suffices to show that

$$\binom{|\mathbf{n}| - d + s - k}{s-k-1} = \Big(\#\text{ monomials in (19)}\Big)\Big(\text{average exponent of } z_{\mathbf{x}} \text{ in each monomial}\Big).$$

The number of monomials in (19) equals the number of monomials in $K[F]_{s-k-1}$. This number, in turn, equals the number of weak compositions of the degree $s-k-1$ into $\#F$ many parts. Thus, recalling from (12) that $\#F = |\mathbf{n}| - d + 1$ for any facet F of $\Delta_{\mathbf{n}}$, and using the elementary formula (6), we have

$$\#\text{ monomials in (19)} = \binom{s-k-1+|\mathbf{n}|-d}{|\mathbf{n}|-d}. \tag{20}$$

The average exponent of $z_{\mathbf{x}}$, taken over all the monomials in $K[F]_{s-k-1}$, equals the degree $s-k-1$ divided by the number of variables $\#F$. Adding 1 to this average to account for the factor of $z_{\mathbf{x}}$ present in $z_{\mathrm{R}(F)\cup\{\mathbf{x}\}}$ in (19), we obtain

$$\text{average exponent of } z_{\mathbf{x}} \text{ in each monomial} = 1 + \frac{s-k-1}{|\mathbf{n}|-d+1} = \frac{|\mathbf{n}|-d+s-k}{|\mathbf{n}|-d+1}. \tag{21}$$

Multiplying the expressions in (20) and (21), we obtain

$$\binom{s-k-1+|\mathbf{n}|-d}{|\mathbf{n}|-d} \cdot \frac{|\mathbf{n}|-d+s-k}{|\mathbf{n}|-d+1} = \binom{|\mathbf{n}|-d+s-k}{|\mathbf{n}|-d+1} = \binom{|\mathbf{n}|-d+s-k}{s-k-1}. \qquad \square$$

Lemma 6.3. *For $\mathbf{x} \in \Pi_{\mathbf{n}}$, the coefficient*

$$[t^k]\, A_{\mathbf{x}-\mathbf{1}}(t) A_{\mathbf{n}-\mathbf{x}}(t)$$

equals the number of facets $F \ni \mathbf{x}$ of $\Delta_{\mathbf{n}}$, such that $\#\big(\mathrm{R}(F) \setminus \{\mathbf{x}\}\big) = k$.

Proof. Every facet $F \ni \mathbf{x}$ of $\Delta_{\mathbf{n}}$ can be written uniquely as the union of two saturated chains

$$F' : \mathbf{1} \lessdot \cdots \lessdot \mathbf{x} \qquad \text{and} \qquad F'' : \mathbf{x} \lessdot \cdots \lessdot \mathbf{n},$$

which intersect only at $\mathbf{x}$. Then F' can be any facet of $\Delta_{\mathbf{x}}$, viewed as a subposet of $\Delta_{\mathbf{n}}$. Likewise, F'' can be any facet of $\Delta_{\mathbf{n}+\mathbf{x}-\mathbf{1}}$, viewed as a subposet of $\Delta_{\mathbf{n}}$ after translating coordinates by $\mathbf{x}-\mathbf{1}$. Therefore $\#(\mathrm{R}(F) \setminus \{\mathbf{x}\}) = \mathrm{R}(F') + \mathrm{R}(F'')$. Thus by (7), we have

$$\begin{aligned} h_{\mathbf{x}}(t) h_{\mathbf{n}-\mathbf{x}+\mathbf{1}}(t) &= \left(\sum_{F'} t^{\#\mathrm{R}(F')}\right)\left(\sum_{F''} t^{\#\mathrm{R}(F'')}\right) \\ &= \sum_{F',F''} t^{\#\mathrm{R}(F')+\#\mathrm{R}(F'')} \\ &= \sum_{F\ni\mathbf{x}} t^{\#(\mathrm{R}(F)\setminus\{\mathbf{x}\})}, \end{aligned}$$

where the sums range over facets F, F', and F'' of $\Delta_{\mathbf{n}}$, $\Delta_{\mathbf{x}}$, and $\Delta_{\mathbf{n}-\mathbf{x}+\mathbf{1}}$, respectively. By Lemma 5.1, we can rewrite $h_{\mathbf{x}}(t) h_{\mathbf{n}-\mathbf{x}+\mathbf{1}}(t)$ as $A_{\mathbf{x}-\mathbf{1}}(t) A_{\mathbf{n}-\mathbf{x}}(t)$. $\square$

Proof of Theorem 6.1. By (14), we know that $\Sigma_{\mathbf{n},s}[\mathbf{x}]$ equals the exponent of $z_{\mathbf{x}}$ in the product of all monomials in the graded component $K[\Delta_{\mathbf{n}}]_s$, which by (11) has the decomposition

$$K[\Delta_{\mathbf{n}}]_s = \bigoplus_k \bigoplus_{\substack{F: \\ \#\mathrm{R}(F)=k}} K[F]_{s-k}\, z_{\mathrm{R}(F)}, \tag{22}$$

where the F's in the inside sum are facets of $\Delta_{\mathbf{n}}$. The outside sum in (22) ranges from $k = 0$ to $k = \min\{|\mathbf{n}| - \max(\mathbf{n}) - d + 1,\, s\}$; this follows from (18) and from the fact that the degree $s - k$ must be nonnegative. Obviously, the only monomials contributing to the exponent of $z_{\mathbf{x}}$ are those divisible by $z_{\mathbf{x}}$; hence we may ignore all summands in (22) such that $\mathbf{x} \notin F$. If $\mathbf{x} \in F$, then the subspace of $K[F]_{s-k}$ spanned by the monomials divisible by $z_{\mathbf{x}}$ is

$$K[F]_{s-k-1} z_{\mathbf{x}}.$$

Then since $\mathbf{x}$ may or may not lie in $\mathrm{R}(F)$, the subspace of $K[F]_{s-k} z_{\mathrm{R}(F)}$ spanned by monomials divisible by $z_{\mathbf{x}}$ is

$$K[F]_{s-k-1} z_{\mathbf{x}} z_{\mathrm{R}(F)\setminus\{\mathbf{x}\}} = K[F]_{s-k-1} z_{\mathrm{R}(F)\cup\{\mathbf{x}\}}.$$

Combining this with (22), we conclude that $\Sigma_{\mathbf{n},s}[\mathbf{x}]$ equals the exponent of $z_{\mathbf{x}}$ in the product of all monomials in

$$\bigoplus_k \bigoplus_{\substack{F \ni \mathbf{x}: \\ \#(\mathrm{R}(F)\setminus\{\mathbf{x}\})=k}} K[F]_{s-k-1} z_{\mathrm{R}(F)\cup\{\mathbf{x}\}}. \tag{23}$$

Now applying Lemma 6.2 and Lemma 6.3 to (23), we see that the desired exponent of $z_{\mathbf{x}}$ equals

$$\begin{aligned}\Sigma_{\mathbf{n},s}[\mathbf{x}] &= \sum_k \sum_{\substack{F \ni \mathbf{x}: \\ \#(\mathrm{R}(F)\setminus\{\mathbf{x}\})=k}} \binom{|\mathbf{n}| - d + s - k}{s - k - 1} \\ &= \sum_k \binom{|\mathbf{n}| - d + s - k}{s - k - 1} \cdot [t^k]\, A_{\mathbf{x}-\mathbf{1}}(t) A_{\mathbf{n}-\mathbf{x}}(t),\end{aligned}$$

where the nonzero summands are those for which k is less than or equal to both $s - 1$ (otherwise the binomial coefficient is zero) and the degree of $A_{\mathbf{x}-\mathbf{1}}(t)A_{\mathbf{n}-\mathbf{x}}(t)$. This degree is easily computed to be $\omega(\mathbf{n}, \mathbf{x})$, using (18). ☐

Remark 6.4. It is not obvious that the formula in Theorem 6.1 specializes to the two-dimensional formula in our previous paper [8, Thm. 5.1]. This is because in the case $d = 2$, it was straightforward to write down explicitly the coefficient of t^k in the product of two multiset Eulerian polynomials since these polynomials could be expressed without signs (see [8, Lemma 4.2]).

Acknowledgments

The authors would like to thank the anonymous referee for the particularly helpful suggestions that improved this paper.

References

[1] M. Abramson, *A simple solution of Simon Newcomb's problem*, J. Combin. Theory Ser. A 18 (1975), 223–225.
[2] W. W. Bein, P. Brucker, J. K. Park, and P. K. Pathak, *A Monge property for the d-dimensional transport problem*, Discrete Appl. Math. 58:2 (1995), 97–109.
[3] A. Björner, *Shellable and Cohen-Macaulay partially ordered sets*, Trans. Amer. Math. Soc. 260:1 (1980), 159–183.
[4] A. Björner and M. Wachs, *On lexicographically shellable posets*, Trans. Amer. Math. Soc. 277:1 (1983), 323–341.
[5] A. Björner and M. Wachs, *Shellable nonpure complexes and posets. I*, Trans. Amer. Math. Soc. 348:4 (1996), 1299–1327.
[6] R. Bourn and J. F. Willenbring, *Expected value of the one-dimensional earth mover's distance*, Algebr. Stat. 11:1 (2020), 53–78.
[7] J. F. Dillon and D. P. Roselle, *Simon Newcomb's problem*, SIAM J. Appl. Math. 17 (1969), 1086–1093.
[8] W. Q. Erickson and J. Kretschmann, *The sum of all width-one matrices*, European J. Combin., to appear, arXiv:2303.08924.
[9] W. Q. Erickson, *A generalization for the expected value of the earth mover's distance*, Algebr. Stat. 12:2 (2021), 139–166.
[10] J. Kline, *Properties of the d-dimensional earth mover's problem*, Discrete Appl. Math. 265 (2019), 128–141.
[11] P. A. MacMahon, *Combinatory analysis*, Chelsea Publishing Co., New York, 1960.
[12] M. Morales, *Segre embeddings, Hilbert series and Newcomb's problem*, arXiv:1306.6910, 2014.
[13] R. P. Stanley, *Enumerative combinatorics. Volume 1*, Second Edition, Cambridge Studies in Advanced Mathematics, Volume 49, Cambridge University Press, Cambridge, 2012.
[14] R. P. Stanley, *Algebraic combinatorics: walks, trees, tableaux, and more*, Undergraduate Texts in Mathematics, Springer, Cham, 2018.