The earth mover’s distance as the symmetric difference of Young diagrams

William Q. Erickson

Abstract Classically, the earth mover’s distance is the solution to an optimal transport problem given two input vectors (supply and demand); in the discrete setting, the input vectors are tuples of integers and therefore can be regarded as histograms. In this paper, we express the earth mover’s distance between two histograms as the symmetric difference of a pair of Young diagrams. Using this combinatorial approach, we prove a result (which extends even to the continuous setting) relating the earth mover’s distance to the difference of the histograms’ averages. Generalizing the notion of “difference” to an arbitrary number of histograms, we show that the distribution of these differences (on tuples of histograms) is encoded in the character of a certain virtual representation of the special linear group. We therefore visualize this distribution as a single weight diagram with support in the Type A root lattice.

Keywords Earth mover’s distance · Young diagrams · virtual representations · weight diagrams

Mathematics Subject Classification (2010) 90C27 · 05E10

1 Introduction

The earth mover’s distance (also known as the first Wasserstein distance) is a metric on a probability space. Equivalently, and more historically, it can be regarded as the solution to a specialization of the Hitchcock transportation problem: what is the cheapest way to transport supplies among \( n \) locations in order to meet each location’s demand? Both in a continuous setting (probability distributions on \( \{1, \ldots, n\} \)) and in
a discrete setting (histograms with \( n \) bins), the earth mover’s distance has found an enormous range of modern applications, especially in image retrieval [13], physics [10], and even political science [12].

In the discrete setting, we interpret this statistical metric in a purely combinatorial fashion: our main result (Theorem 1) expresses the earth mover’s distance as the symmetric difference of Young diagrams. As a preview, consider the histograms (2, 0, 2, 4, 0, 0, 1) and (0, 5, 1, 0, 2, 1, 0, 0); we claim that their earth mover’s distance can be calculated as follows. Encode each histogram’s entries as tallies of ascending row-lengths in a Young diagram; then superimpose the two diagrams and count the size of their symmetric difference (i.e., their union minus their intersection):

In this case, the symmetric difference contains 11 boxes, and so the earth mover’s distance is 11. We find this method to be strikingly simple and conceptual, given the classical definition of the earth mover’s distance which we review in Section 1.1. The earth mover’s distance can be naturally generalized to compare an arbitrary number of histograms (see [1] and [2]), and our main result accounts for this full generality as well. Hence from now on, rather than “distance,” we will refer to the earth mover’s coefficient (EMC) of a collection of histograms.

In the second part of the paper, we consider not only the EMC but also the averages of the histograms. We compare these averages via the difference \( D \) of the weighted totals, and we determine (using only Young diagrams) the probability that \( EMC = D \) in the continuous setting. (See Corollary 2.) This difference \( D \) can also be generalized to an arbitrary number \( d \) of histograms, which raises the combinatorial question of counting \( d \)-tuples of histograms with a given \( D \)-value. We give a clean solution to this problem (Theorem 3) by showing that the answers are encoded as weight multiplicities in the character of a certain virtual representation of the Lie algebra \( \mathfrak{sl}(d, \mathbb{C}) \).

1.1 Classical definition of EMC

As context for our alternative approach to the EMC, we present its original definition in the language of transport theory (the “Monge–Kantorovich problem”) and linear programming (the “Hitchcock–Koopmans problem”). We therefore consider the following \( d \)-dimensional discrete transportation problem. This is a generalization of the two-dimensional problem which, in the literature, is named after various combinations of Hitchcock, Koopmans, Monge and Kantorovich (see [2]).

Let the (generalized) “supply–demand vectors” \( \alpha_1, \ldots, \alpha_d \) be \( n \)-tuples of nonnegative integers. Denote by \( \alpha_i(j) \) the \( j \)th entry of \( \alpha_i \), and assume that \( \sum_j \alpha_i(j) = s \) for all \( i = 1, \ldots, d \). (In our context, we will regard the \( \alpha_i \) as histograms with \( s \) data points and \( n \) bins.) Let \( C \) (for “cost”) be the \( d \)-dimensional, \( n \times \cdots \times n \) array
The earth mover’s distance as the symmetric difference of Young diagrams

whose value at each position equals the taxicab distance from the main diagonal. (If $d = 2$, then $C$ is just the $n \times n$ matrix where $C(i, j) =$ $|i - j|$. For a position vector $x \in [n]^d = \{(x_1, \ldots, x_d) \mid 1 \leq x_i \leq n\}$, we let $C(x)$ denote the entry of $C$ in position $x$. The objective is to choose another $n \times \cdots \times n$ integer array $J$, in order to solve the following problem:

Minimize $\sum_{x \in [n]^d} C(x)J(x)$,

subject to $\sum_{x \text{ such that } x_i = j} J(x) = \alpha_i(j)$ for $1 \leq i \leq d$ and $1 \leq j \leq n$,

and $\quad J(x) \geq 0$ for all $x \in [n]^d$.

(Geometrically, the constraint on $J$ states that its hyperplane sums, in the direction perpendicular to the $i^{th}$ standard basis vector, must match the entries in $\alpha_i$. See Figure 1 to visualize this in the case $d = 3$ and $n = 5$.)

![Fig. 1: An illustration of the conditions on $J$ in the transport problem (1) defining the EMC. (Here $d = 3$ and $n = 5$.) Each arrow represents the sum of the entries in the designated plane.](image)

The earth mover’s coefficient (EMC) of the $\alpha_i$ is the minimum value of the objective function $\sum_x C(x)J(x)$. The intuition is this: imagine each $\alpha_i$ as a collection of $s$ pebbles, distributed among $n$ piles located at $1, \ldots, n$ on the number line, with $\alpha_i(j)$ pebbles in the $j^{th}$ pile. If a “move” consists in moving one pebble to a neighboring pile, then the EMC is the fewest number of moves required to equalize all of the $\alpha_i$. (In this paper, we place the piles 1 unit apart because of the motivation from histograms; the most general version of this transport problem involves arbitrary distances between piles, which need not be collinear, in which case the cost array $C$ would look very different.)

Due to a special property of the cost array $C$ (known as the Monge property, which is essentially submodularity), the optimization problem (1) can be solved in $O(d^2 n)$ time. (See [1], [2], and [9] for details behind this greedy algorithm known as the “northwest corner rule.”) Our approach in this paper, however, provides a simple combinatorial method to compute the EMC by counting boxes shared by Young diagrams corresponding to the $\alpha_i$. 


2 Combinatorial background

2.1 Integer compositions

For \( s, n \in \mathbb{Z}_{\geq 0} \), let \( \mathcal{C}(s, n) \) denote the set of (weak) integer compositions of \( s \) into \( n \) parts. In other words, a composition \( \alpha \in \mathcal{C}(s, n) \) is an \( n \)-tuple \( \alpha = (a_0, \ldots, a_{n-1}) \) such that the \( a_i \) are nonnegative integers summing to \( s \). It is well known that \( |\mathcal{C}(s, n)| = \binom{s+n-1}{n-1} \). From this point forward, we use a Greek letter for a composition and an indexed Roman letter for its integer entries; this will avoid confusion in cases when we must also index the compositions themselves as \( \alpha_1, \ldots, \alpha_d \). We will no longer use the \( \alpha_i(j) \) notation from the introduction to denote a specific entry of a composition. (The reader is of course welcome to substitute the statistical term “histogram,” or the transportational term “supply–demand vector,” for the more combinatorial term “composition.”) In light of this shift in perspective, we now regard the earth mover’s coefficient as a function \( \text{EMC} : \mathcal{C}(s, n)^d \rightarrow \mathbb{Z}_{\geq 0} \) which measures, in some sense, the “closeness” of an arbitrary number of compositions. When necessary, we use \( d \) to denote the number of compositions being considered, so that \( \mathcal{C}(s, n)^d \) denotes the Cartesian product of \( d \) copies of \( \mathcal{C}(s, n) \).

2.2 RSK algorithm to compute EMC

As a steppingstone to our main result, we present a streamlined algorithm which improves upon the northwest corner rule mentioned above, by calculating only the support of the optimal array \( J \). This idea was introduced in [3] for the \( d = 2 \) case, and generalized by this author for arbitrary \( d \) in [6]. The following algorithm is, in thin disguise, a highly specialized application of the multivariate Robinson–Schensted–Knuth (RSK) correspondence; see [14], Sections 7.11–14 for its full scope.

Let \( \alpha = (a_0, \ldots, a_{n-1}) \in \mathcal{C}(s, n) \). We define the word of \( \alpha \) to be the sequence \( w(\alpha) \) constructed by writing “0” \( a_0 \) times, then “1” \( a_1 \) times, and so on, ending with “\( n-1 \)” written \( a_{n-1} \) times:

\[
w(\alpha) = (0, \ldots, 0, 1, \ldots, 1, \ldots, n-1, \ldots, n-1).
\]

Hence \( w(\alpha) \) is a weakly increasing sequence of length \( s \), with elements from the set \{0, \ldots, n-1\}. For example, \( w(3, 2, 0, 3, 1) = (0, 0, 0, 1, 1, 3, 3, 3, 4) \). In the pebble-pile imagery from above, a word is the sequence obtained by recording the location of each pebble from left to right.

Given compositions \( \alpha_1, \ldots, \alpha_d \), we form the matrix

\[
M = M(\alpha_1, \ldots, \alpha_d) = \begin{bmatrix} w(\alpha_1) \\ \vdots \\ w(\alpha_d) \end{bmatrix}
\]

whose \( i^{th} \) row is the word \( w(\alpha_i) \). Then the RSK correspondence takes this \( d \times s \) matrix \( M \) to a \( d \)-dimensional, \( n \times \cdots \times n \) array \( J \) as follows:
1. Start with $J$ as the array of all zeros;
2. For each column vector of $M$, interpret that vector as the coordinates of an entry in $J$, and add 1 to that entry.

In other words, the support of $J$ is the set of columns of $M$. As the reader can easily check, the coordinate hyperplane sums in $J$ are given by the $\alpha_i$, and so $J$ satisfies the constraint in the transport problem (1). Moreover, since words are nondecreasing, the support of $J$ is a chain, in the sense that the position vectors of its support are pairwise comparable under the product order on $[n]^d$. This fact, along with the Monge property of $C$, implies that $J$ truly is a solution of the transport problem (1). (See [2] and [6] for details.)

Now that we have an optimal array $J$, by definition we have

$$\text{EMC}(\alpha_1, \ldots, \alpha_d) = \sum_x J(x)C(x).$$

Rather than sum over all array positions $x$, however, we need only sum over the support of $J$, namely, the columns of $M$:

$$\text{EMC}(\alpha_1, \ldots, \alpha_d) = \sum_{j=1}^s C(M_{*j}), \quad (3)$$

where $M_{*j}$ denotes the (transpose of the) $j$th column of $M$. It remains to compute the cost $C(x)$ for a given $d$-tuple $x$. The most convenient way to compute the taxi-cab distance between $x$ and the main diagonal is to consider the “sorted” vector $\mathbf{e}_x$, whose coordinates are those of $x$ rearranged in ascending order; for instance, if $x = (7, 4, 5, 4, 1)$, then $\mathbf{e}_x = (1, 4, 4, 5, 7)$.

**Proposition 1** Let $x \in [n]^d$, with the sorted vector $\mathbf{e}_x = (e_{x_1}, \ldots, e_{x_d})$ as above. Then

$$C(x) = \sum_{i=1}^{\lfloor d/2 \rfloor} \overline{e}_{d+1-i} - \overline{e}_i.$$

The proof is in [6], Section 3. As an example, take $x = (7, 4, 5, 4, 1)$ as above. To compute $C(x)$, we instead inspect $\mathbf{e}_x$ and sum up the pairwise differences, working outside-in:

$$\overline{x} = (1, 4, 5, 7),$$

therefore $C(x) = 6 + 1 = 7$.

We conclude this subsection with a full example computing the EMC using the RSK algorithm.

**Example 1** Set $d = 4$, $n = 5$, and $s = 6$. We aim to compute the EMC of these four compositions in $\mathcal{C}(6, 5)$:

$$\alpha = (4, 1, 1, 0, 0),$$
$$\beta = (3, 0, 0, 0, 3),$$
$$\gamma = (0, 4, 2, 0, 0),$$
$$\delta = (1, 1, 2, 1, 1).$$
We initiate the RSK algorithm outlined above by writing out the words

\[
\begin{align*}
\omega(\alpha) &= (0, 0, 0, 1, 2), \\
\omega(\beta) &= (0, 0, 4, 4, 4), \\
\omega(\gamma) &= (1, 1, 1, 2, 2), \\
\omega(\delta) &= (0, 1, 2, 3, 4),
\end{align*}
\]

which become the rows of the matrix

\[
M = \begin{bmatrix}
0 & 0 & 0 & 1 & 2 \\
0 & 0 & 0 & 4 & 4 \\
1 & 1 & 1 & 2 & 2 \\
0 & 1 & 2 & 3 & 4
\end{bmatrix}.
\]

Using Proposition 1, we compute the cost \( C \) for each of the six columns of \( M \):

\[
\begin{align*}
C(0, 0, 1, 0) &= (1 - 0) + (0 - 0) = 1, \\
C(0, 0, 1, 1) &= (1 - 0) + (1 - 0) = 2, \\
C(0, 0, 1, 2) &= (2 - 0) + (1 - 2) = 3, \\
C(0, 4, 1, 2) &= (4 - 0) + (2 - 1) = 5, \\
C(1, 4, 2, 3) &= (4 - 1) + (3 - 2) = 4, \\
C(2, 4, 2, 4) &= (4 - 2) + (4 - 2) = 4.
\end{align*}
\]

Finally, summing these costs, we conclude

\[
\text{EMC}(\alpha, \beta, \gamma, \delta) = 1 + 2 + 3 + 5 + 4 + 4 = 19.
\]

2.3 Compositions as Young diagrams

The main result of this paper is a reinterpretation of the EMC in terms of the ubiquitous combinatorial objects known as Young diagrams. A Young diagram is a finite set of congruent square cells arranged in left-justified rows, with row lengths weakly decreasing from top to bottom. Given a composition \( \alpha \in \mathcal{C}(s, n) \), we can represent it uniquely as the Young diagram whose ascending row lengths are given by the word \( \omega(\alpha) \). For example:

\[
\alpha = (3, 2, 0, 3, 1) \implies \omega(\alpha) = (0, 0, 0, 1, 1, 3, 3, 4) \sim
\]

since we draw the Young diagram from the bottom up, and 0’s correspond to empty rows. Because of this correspondence, we will often refer to a composition \( \alpha \) as if it is a Young diagram. Specifically, we write \( |\alpha| \) to denote the size (number of cells) of the Young diagram of \( \alpha \). It is clear that if \( \alpha \) is a composition, then

\[
|\alpha| \text{ is the sum of the entries in } \omega(\alpha).
\]
Note that if \( \alpha \in \mathcal{C}(s, n) \), then as a Young diagram, \( \alpha \) has at most \( s \) rows, and each of those row lengths is at most \( n - 1 \). Thus we have a bijective correspondence
\[
\mathcal{C}(s, n) \leftrightarrow \mathcal{Y}(s, n - 1) := \left\{ \text{Young diagrams which fit inside an } s \times (n - 1) \text{ rectangle} \right\}.
\] (5)

By taking conjugates of Young diagrams (i.e., reflection across the main diagonal), we have the additional bijective correspondences:
\[
\begin{align*}
\mathcal{C}(s, n) & \leftrightarrow \mathcal{Y}(s, n - 1) \\
\mathcal{C}(n - 1, s + 1) & \leftrightarrow \mathcal{Y}(n - 1, s)
\end{align*}
\] (6)

This duality between the values \( (s, n) \) and \( (n - 1, s + 1) \) will appear throughout our results (particularly the plots in Sections 4 and 5).

We will briefly refer to a corner in a Young diagram; this is a cell such that the result of its removal is still a Young diagram. For example, in the Young diagram \( \alpha \) depicted above, there are three corners: the last cell in the first row, the last cell in the fourth row, and the cell in the bottom row. We will write \( \text{cor}(\alpha) \) for the number of corners in \( \alpha \), which can easily be computed as the number of distinct row lengths.

2.4 The \( q \)-binomial coefficients

We recall here the \( q \)-analog of the binomial coefficients \( \binom{a}{b} \). These \( q \)-binomial coefficients often appear in one of two guises, both of which we will encounter near the end of this paper; therefore we distinguish the two with the following notation.

On one hand, let \( [a]_q := 1 + q + \cdots + q^{a-1} \). Then with \( [a]_q! := [a]_q[a-1]_q \cdots [1]_q \), we define
\[
\binom{a}{b}_q := \frac{[a]_q!}{[b]_q![a-b]_q!}.
\]
This is actually a polynomial (not just a rational function) in \( q \), which encodes combinatorial information central to this paper (see [15], Proposition 1.7.3):
\[
\text{The coefficient of } q^w \text{ in } \binom{s+n-1}{s}_q \text{ equals the number of Young diagrams in } \mathcal{Y}(s, n - 1) \text{ with size } w.
\] (7)

On the other hand, let \( (a)_q := q^{-(a-1)} + q^{-(a-3)} + \cdots + q^{-3} + q^{-1} \). Then again with \( (a)_q! := (a)_q \cdots (1)_q \), we define
\[
\binom{a}{b}_q := \frac{(a)_q!}{(b)_q!(a-b)_q!}.
\]
This is a Laurent polynomial, invariant under \( q \leftrightarrow q^{-1} \), which (as we will see) arises in the representation theory of the Lie algebra \( \mathfrak{sl}(2, \mathbb{C}) \).

To convert between the two guises of \( q \)-binomial coefficients, we observe that
\[
\binom{s+n-1}{s}_q = q^{-s(a-1)} \binom{s+n-1}{s}_{q^2}.
\] (8)
2.5 Plane partitions

In proving a major result in Section 4, we will use a generalization of Young diagrams known as plane partitions. A plane partition is a Young diagram in which each cell is filled with a nonnegative integer entry, such that the entries weakly decrease along each row and column. A plane partition is visualized three-dimensionally by stacking cubes on a Young diagram such that the height above each cell is its integer entry; for this reason, a plane partition whose underlying Young diagram has at most $x$ rows and $y$ columns, with entries at most $z$, is said to fit inside an $x \times y \times z$ box. The number of such partitions is

$$PP(x, y, z) = \prod_{i=1}^{x} \prod_{j=1}^{y} \prod_{k=1}^{z} \frac{i + j + k - 1}{i + j + k - 2}.$$  \hfill (9)

(See [14], sections 7.20–22, for an expansive treatment of the combinatorics of plane partitions.) In Section 4, we will appeal to the following specialization:

**Lemma 1** In the case $z = 2$, we have

$$PP(x, y, 2) = \frac{(x + 1) \cdot \cdots \cdot (x + y) \cdot (x + 2) \cdots (x + y + 1)}{y!(y + 1)!}$$  \hfill (10)

where the omitted factors increase by 1.

**Proof** Setting $z = 2$, we separate (9) into two main factors (one for $k = 1$ and the other for $k = 2$) and account for the telescoping cancellations:

$$PP(x, y, 2) = \prod_{i,j} \frac{i + j}{i + j - 1} \left( \prod_{i,j} \frac{i + j + 1}{i + j} \right) = \prod_{i,j} \frac{i + j + 1}{i + j - 1}.$$

Now we use induction on $x$ and $y$ to prove that this product is equivalent to (10). In the base case $x = y = 1$, the product above is just the single factor $\frac{1+1}{1+1} = 3$, and likewise the expression in (10) is $\frac{(1+1)(1+2)}{1!2!} = \frac{6}{2} = 3$. (Note that the second string of factors in the numerator of (10) is the empty product 1 in the case $x = y = 1$.)

To induct on $x$, assume that the corollary holds for $x - 1$ and $y$. Then we have

$$PP(x, y, 2) = \prod_{i=1}^{x} \prod_{j=1}^{y} \frac{i + j + 1}{i + j - 1} = \prod_{j=1}^{y} \frac{x + j + 1}{x + j - 1} \cdot PP(x - 1, y, 2) = \frac{(x + y)(x + y + 1)}{(x + 1)} \cdot \frac{(x + y - 1) \cdots (x + 1) \cdot (x + y) \cdots (x + y + 1)}{y!(y + 1)!} = \frac{(x + 1) \cdots (x + y) \cdot (x + 2) \cdots (x + y + 1)}{y!(y + 1)!}$$

as in (10). The induction on $y$ works out similarly. \hfill $\Box$
3 Main result: EMC as symmetric difference of Young diagrams

The symmetric difference of two sets, denoted by the symbol $\triangle$, is

$$A \triangle B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B),$$

the set of elements contained in exactly one of the two sets. In many contexts it is natural to extend this definition associatively to any finite number of sets, but to capture the behavior of the EMC we introduce a different generalization.

**Definition 1** We say the unimodal symmetric difference of sets $A_1, \ldots, A_d$ is

$$\blacktriangle(A_1, \ldots, A_d) := \sum_{k=1}^{d-1} \min\{k, d-k\} \cdot |\{\text{elements contained in exactly } k \text{ of the } A_i\}|.$$

(It would be equivalent to let $k$ range from 0 to $d$.) Note that $\blacktriangle$ returns a number rather than a set. Intuitively, the unimodal symmetric difference counts elements by assigning the weight 1 to those elements contained in exactly one, or all but one, set; then the weight 2 to those elements in exactly two, or all but two, sets; and so forth, finally assigning the greatest weight $[d/2]$ to those elements contained in exactly “half” of the sets. (See Figure 2, where successively higher weights are shown by successively darker shades of grey.) Note that in the $d = 2$ case, $\blacktriangle(A, B)$ agrees with $|A \triangle B|$.

Now, consider each Young diagram in $\mathcal{Y}(s, n)$ as a subset of the full $s \times (n-1)$ rectangle of cells. It then makes sense to say that a given cell is contained within several different Young diagrams; for example, the upper-left cell is contained in every Young diagram except for the empty diagram. Therefore, by identifying compositions with their Young diagrams, we can speak of their unimodal symmetric difference. We now state our main result.

**Theorem 1** On any tuple of compositions in $\mathcal{C}(s, n)$, we have $\text{EMC} = \blacktriangle$.

**Proof** Let $\alpha_1, \ldots, \alpha_d \in \mathcal{C}(s, n)$. By [3], we have $\text{EMC}(\alpha_1, \ldots, \alpha_d) = \sum_{j=1}^{s} C(M_{\alpha_j})$, where $M$ is the matrix defined in [2]. Fix $j$, and put $M_{\alpha_j} = (m_1, \ldots, m_d)$. Since $C(x)$ is symmetric in the entries of $x$, we can assume without loss of generality that $m_1 \leq \cdots \leq m_d$. Then by Proposition [1] we have

$$C(M_{\alpha_j}) = \sum_{i=1}^{[d/2]} m_{d+1-i} - m_i$$

$$= \sum_{i=1}^{[d/2]} \sum_{k=i}^{d-i} m_{k+1} - m_k.$$
For $d = 2$, weights are $0, 1, 0$.  

For $d = 3$, weights are $0, 1, 1, 0$.  

For $d = 4$, weights are $0, 1, 2, 0$.  

For $d = 5$, weights are $0, 1, 2, 2, 0$.  Note that in this example, the intersection of all five sets is empty.  

For $d = 6$, weights are $0, 1, 2, 3, 2, 1, 0$.  

For $d = 8$, weights are $0, 1, 2, 3, 4, 3, 2, 1, 0$.  This image depicts a chain of eight subsets.

Fig. 2: A visualization of the unimodal symmetric difference $\mathbb{A}(A_1, \ldots, A_d)$. Each ellipse is one of the sets $A_i$, and the “weights” $0, 1, 2, \ldots$ are depicted by successively darker shades of grey. Elements contained in $0$ or $d$ sets are not counted (white regions); elements contained in exactly $1$ or $d - 1$ sets are counted once (lightest grey); elements in exactly $2$ or $d - 2$ sets are counted twice (slightly darker grey); and so forth, until elements in the darkest regions are counted $\lfloor d/2 \rfloor$ times.

Now we switch the order of summation, using the fact that $i \leq k \leq d - i$ implies both $i \leq k$ and $i \leq d - k$:

$$C(M_{ij}) = \sum_{k=1}^{d-1} \left( \sum_{i=1}^{\min\{k, d-k\}} m_{k+1} - m_k \right)$$

$$= \sum_{k=1}^{d-1} \min\{k, d-k\} \cdot (m_{k+1} - m_k)$$

$$= \mathbb{A}(\{m_1\}, \ldots, \{m_d\}),$$

where $\{m_i\}$ denotes the set $\{1, \ldots, m_i\}$. Now, recalling the construction of the matrix $M$, we know that $m_k$ is the $j$th element from the word $w(a_k)$, and therefore $m_k$ is the length of the $j$th row (from the bottom) of the Young diagram $a_k$. Hence $\mathbb{A}(\{m_1\}, \ldots, \{m_d\})$ is the result of applying $\mathbb{A}$ to the collection containing the $j$th row (from the bottom) of each of the Young diagrams $a_1, \ldots, a_d$. Letting $j$ range over all $s$ rows and taking the sum, we conclude that $\text{EMC}(a_1, \ldots, a_d) = \mathbb{A}(a_1, \ldots, a_d)$. ☐
Example 2 We return to Example 1 but this time we will calculate the EMC combinatorially, using Theorem 1. We translate the word of each composition into its corresponding Young diagram:

\[
\begin{align*}
\alpha &= \begin{array}{ccc}
\cdot & \cdot & \\
& \\
& \\
\end{array} \\
\beta &= \begin{array}{cccc}
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \\
& \cdot & \cdot \\
\end{array} \\
\gamma &= \begin{array}{ccc}
& \cdot & \\
& \cdot & \cdot \\
& \cdot & \\
\end{array} \\
\delta &= \begin{array}{cccc}
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\end{array}
\end{align*}
\]

Now we superimpose these four diagrams so that they share a common upper-left cell; see the two diagrams below. In the first diagram, we label each cell with the number of diagrams (among \(\alpha, \beta, \gamma, \delta\)) which contain it; then in the second diagram, to each cell labeled \(k\) we assign the weight \(\min(k, 4-k)\), since \(d = 4\), and shade as in Figure 2:

\[
\begin{array}{cccc}
4 & 4 & 2 & 2 \\
4 & 3 & 2 & 1 \\
3 & 2 & 1 & 1 \\
2 & 1 & & \\
2 & & & \\
1 & & & \\
\end{array} ~ \begin{array}{cccc}
0 & 0 & 2 & 2 \\
0 & 1 & 2 & 1 \\
1 & 2 & 1 & 1 \\
2 & 1 & & \\
2 & & & \\
1 & & & \\
\end{array}
\]

The shading is superfluous, of course, since we only need to sum the entries of the second diagram to compute that

\[
\begin{align*}
\text{EMC}(\alpha, \beta, \gamma, \delta) &= \mathbb{1}(\alpha, \beta, \gamma, \delta) \\
&= 2 + 2 + 1 + 2 + 1 + 1 + 2 + 1 + 1 + 2 + 1 + 1 + 2 + 1 + 1 + 2 + 1 + 1 \\
&= 19,
\end{align*}
\]

agreeing with the calculation in Example 1.

We record the following corollary because many applications of the EMC restrict to the \(d = 2\) case, and because this case reduces to the familiar definition of symmetric difference. We appeal constantly to this corollary throughout Section 4.

Corollary 1 For \(\alpha, \beta \in C(s, n)\), we have \(\text{EMC}(\alpha, \beta) = |\alpha \triangle \beta|\).

4 EMC vs. weighted difference: \(d = 2\)

When we regard a composition as a histogram, it makes sense to speak of its average. Comparing the averages of two compositions is quite distinct from taking their EMC. For example, in \(C(10, 3)\), if

\[
\begin{align*}
\alpha &= (5, 0, 5), & \beta &= (0, 10, 0) & \text{and} & \alpha' &= (5, 5, 0), & \beta' &= (0, 5, 5),
\end{align*}
\]

then \(\text{EMC}(\alpha, \beta) = \text{EMC}(\alpha', \beta') = 10\), even though the two pairs are “far apart” in very different ways: as histograms, \(\alpha\) and \(\beta\) have the same average but different
standard deviations, while \( \alpha' \) and \( \beta' \) have different averages but the same standard deviation. In this section, we study how the EMC of two compositions interacts with their averages. We aim to expand upon the treatment in [11] calculating the expected value of the EMC when restricted to composition pairs with identical average. Then in the following section we will generalize to arbitrary \( d \).

To make the notion of “average” precise, we put the most natural weighting on \( \mathcal{C}(s, n) \), namely \( \sum_{i=0}^{n-1} i \cdot a_i \), so that the weighted total \( \mathcal{T} \) of a composition is defined as

\[
\mathcal{T}(\alpha) = \mathcal{T}(a_0, \ldots, a_{n-1}) := \sum_{i=0}^{n-1} i \cdot a_i.
\]  

(11)

In this paper, we will work only with \( \mathcal{T} \), rather than obtaining the actual average by dividing by \( s \). There is no harm in doing this, since all our compositions share a common value of \( s \).

Remark 1 In the case \( n = 5 \), the weighted total is simply the grade-point system familiar to university students, where the five letter grades F, D, C, B, and A are worth 0, 1, 2, 3, and 4 points, respectively. In the context of ranked-ballot voting theory, \( \mathcal{T}(\alpha) \) is the Borda count of a candidate who receives \( a_i \) votes as the \( (n-i) \text{-th} \) choice candidate.

The purpose of the weighted total \( \mathcal{T} \) is to compare compositions to each other, and so we are interested in the signed difference between two weighted totals:

Definition 2 Let \( \alpha, \beta \in \mathcal{C}(s, n) \). We write

\[
\mathcal{D}(\alpha, \beta) := \mathcal{T}(\alpha) - \mathcal{T}(\beta)
\]  

(12)

and call this the weighted difference of the ordered pair \((\alpha, \beta)\).

(This definition will later be subsumed by its generalization in Definition[3].) The weighted total and weighted difference have a simple interpretation in the language of Young diagrams, just as the EMC does. If \( \alpha, \beta \in \mathcal{C}(s, n) \), then

\[
\mathcal{T}(\alpha) = |\alpha|,
\]

\[
\mathcal{D}(\alpha, \beta) = |\alpha| - |\beta|.
\]  

(13)

In words, the weighted total of a composition is the size of its Young diagram. To see this, recall from [4] that \( |\alpha| \) equals the sum of the entries in the word \( w(\alpha) \). But \( w(\alpha) \) contains exactly \( a_i \) copies of \( i \) for each \( i = 0, \ldots, n - 1 \), and so the sum of its entries agrees with the definition of \( \mathcal{T} \) in (11). We are now equipped to study the statistical relationship between EMC and \( \mathcal{D} \) via pairs of Young diagrams.
4.1 Plotting EMC vs. $D$

For fixed $s$ and $n$, we introduce a two-dimensional plot whose vertical axis measures EMC and whose horizontal axis measures $D$. Each ordered pair in $\mathcal{C}(s, n) \times \mathcal{C}(s, n)$ is plotted accordingly, and the number inside each dot indicates the number of ordered pairs at that point. We now take the reader on a brief guided tour of these plots, aided by Figures 3 and 4. (We also include a three-dimensional rendering in Figure 5.) The observations below provide a sort of dictionary between the statistics of histograms and the combinatorics of Young diagram pairs; ultimately, by Corollary 1 and by (13), everything below is a statement about the interplay of two Young diagrams’ symmetric difference with their relative sizes.

- By definition, composition pairs $(\alpha, \beta)$ lying further to the right (higher positive $D$-values) are more “lopsided” (in terms of weighted total) in favor of $\alpha$, and vice versa. But since the plot is symmetric about the vertical axis, from now on we restrict our attention to the right half, i.e., those pairs for which $T(\alpha) \geq T(\beta)$. We therefore refer only to the right half of the plots in the remainder of this list.

- The plot for $\mathcal{C}(s, n)$ is identical to that for $\mathcal{C}(n-1, s+1)$. (Compare, for instance, the first two plots in the second row of Figure 3; this is the only redundancy we have included in the figures.) This duality is not at all obvious a priori from the original definitions of EMC and $D$, but it is immediate when we consider the conjugates of Young diagrams, as in (6): after all, $|\alpha \triangle \beta|$ and $|\alpha|$ and $|\beta|$ are invariant under conjugates, so the EMC-vs.-$D$ plots are as well.

- In general, EMC $\geq D$ since $\alpha \triangle \beta \supseteq \alpha \setminus \beta$.

- The values of both EMC and $D$ range from 0 to $s(n - 1)$. This is obvious when we consider the $s \times (n - 1)$ rectangle containing every diagram in $Y(s, n - 1)$.

- In each plot, the upper-rightmost data point corresponds to the unique pair $(\alpha, \beta)$ where $\alpha$ is the full $s \times (n - 1)$ rectangle and $\beta$ is the empty diagram. Starting from here, moving one unit left on the plot corresponds to either subtracting one cell from $\alpha$ or adding one cell to $\beta$ (in such a way that both remain Young diagrams).

- All composition pairs lie on the integer lattice of points whose coordinates have the same parity. This reflects the fact that adding or subtracting one cell to/from either diagram changes $|\alpha \triangle \beta|$ by exactly 1.

- The “tail” along the diagonal ($y = x$) on the far right reveals that after a certain threshold when $D$ is sufficiently large, EMC and $D$ coincide with each other; in the language of Young diagrams, $|\alpha \triangle \beta| = |\alpha| - |\beta|$ implies $\beta \subseteq \alpha$, so that in any pair on the diagonal, the first diagram contains the second. We claim that this threshold is $D = (s - 1)(n - 2)$.

Proof Starting from the far right at $D = s(n - 1)$ where $\alpha$ is the full rectangle and $\beta$ is empty, the most efficient way to make $\beta$ spill outside $\alpha$ is to subtract the $s$ cells from the rightmost column of $\alpha$ and to add $n - 1$ cells to the top row of $\beta$; at this point, the cell in the upper-right will belong to $\beta$ but not to $\alpha$. Therefore, since moving left from $s(n - 1)$ by $(s + n - 1)$ steps takes us to the line $D = sn - 2s - n + 1$, the “tail” must begin just to the right of this $D$-value, namely $D = sn - 2s - n + 2 = (s - 1)(n - 2)$. $\square$
Fig. 3: Plots created in Mathematica, using the generating function in Section 7. Each numbered dot indicates the number of composition pairs with given EMC and weighted difference.

- The highest concentration occurs relatively close to the origin, on the diagonal. These are pairs in which one Young diagram contains the other, but with a difference of only a “few” cells. More generally, at any point on the diagonal, the composition pairs have the lowest possible EMC given a fixed $D$-value, and also the highest possible $D$-value given a fixed EMC; in other words, if we fix either statistic, then the most “lopsided” possible composition pairs are those on the diagonal. In light of these last two observations, it is natural to ask what proportion of pairs occur along the diagonal, and we answer this question in the next subsection.
A useful statistical goal might be to determine a formula for the distribution of EMC values for a given $D$-value. Combinatorially, the question is this: among all pairs of Young diagrams whose sizes differ by a given number of cells, what is the distribution of the symmetric differences of the pairs? The answer is encoded in the recursive generating function in Section 7, but here we make a first step toward a closed-form expression by attacking the case $D = 0$, i.e., when two diagrams have equal size:

- Clearly the number of pairs with $EMC = 0$ is just $\left| \mathcal{Y}(s, n - 1) \right| = \binom{s+n-1}{s}$, since these are the pairs of identical diagrams. The reader can check that this binomial coefficient is indeed the value at the origin in each plot.
- The next possible EMC value is 2. The question is this: for how many ordered pairs $(\alpha, \beta)$ of equally-sized diagrams in $\mathcal{Y}(s, n - 1)$ is the symmetric difference
Fig. 5: A three-dimensional rendering (of the right half) of the plot for \( s = 10 \) and \( n = 4 \) (dually, for \( s = 3 \) and \( n = 11 \)), seen from four angles. The height above each point is the number inside the dot in the two-dimensional plot.

exactly two cells? For such a pair, consider the union \( \gamma = \alpha \cup \beta \in \mathcal{Y}(s, n - 1) \), and specifically consider the corners of \( \gamma \). Exactly one of these corners must belong to \( \alpha \) alone, and exactly one of the corners (but not the same one!) must belong to \( \beta \) alone. In this way, to each \( \gamma \in \mathcal{Y}(s, n - 1) \) there correspond all possible ordered pairs of diagrams obtained by assigning one corner to each of the diagrams; the number of these pairs is the number of permutations of \( \text{cor}(\gamma) \) objects taken two at a time. We conclude that the value of each plot at the point \((0, 2)\) is given by

\[
\sum_{\gamma \in \mathcal{Y}_s(n-1)} \text{cor}(\gamma) \cdot (\text{cor}(\gamma) - 1).
\]

As an afterthought to our EMC-vs.-\( \mathcal{D} \) plots, we should point out another convenient way to visualize the EMC together with the weighted difference. If we impose a partial order (via inclusion) on the set of Young diagrams, then the result is Young’s lattice. (See [14], Chapter 7; we provide a truncated picture in Figure 6.) Given \( s \) and \( n \), the sublattice \( \mathcal{Y}(s, n - 1) \) is the interval between the empty diagram and the full \( s \times (n - 1) \) rectangle. Young’s lattice is ranked with respect to the size of the diagrams, so that every diagram in the \( i \)th horizontal level has rank \( i \). The least upper
bound (the “join,” written $\alpha \vee \beta$) is the union of the two diagrams, while the greatest lower bound (the “meet” $\alpha \wedge \beta$) is their intersection. Hence, identifying $\alpha, \beta \in \mathcal{C}(s, n)$ with their Young diagrams as usual, we have

\[
\mathcal{D}(\alpha, \beta) = \text{rank}(\alpha) - \text{rank}(\beta),
\]

\[
\text{EMC}(\alpha, \beta) = \text{rank}(\alpha \vee \beta) - \text{rank}(\alpha \wedge \beta).
\]

Fig. 6: Young’s lattice, shown up to rank 5. If compositions $\alpha$ and $\beta$ correspond to the blue and red diagrams respectively, then $\mathcal{D}(\alpha, \beta) = 1$ since $\alpha$ is one row above $\beta$. The darkened arrows converge above at $\alpha \vee \beta$, of rank 5, and converge below at $\alpha \wedge \beta$, of rank 2; therefore $\text{EMC}(\alpha, \beta) = 5 - 2 = 3$.

4.2 Proportion for which $\text{EMC} = \mathcal{D}$ in the continuous setting

As we have observed above, there is a relatively high concentration of pairs along the diagonals $y = |x|$ of our plots. Statistically speaking, at least for the cases we have plotted, in a (perhaps surprisingly) high proportion of composition pairs, the EMC is nothing other than the absolute value of the weighted difference $\mathcal{D}$. We now determine this proportion.

**Theorem 2** The proportion of pairs in $\mathcal{C}(s, n) \times \mathcal{C}(s, n)$ such that $\text{EMC} = |\mathcal{D}|$ is

\[
\frac{2(s + n)}{n(s + 1)} - \frac{(n - 1)!}{(s + 1) \cdots (s + n - 1)}.
\]

**Proof** As we observed above, $\text{EMC}(\alpha, \beta) = |\mathcal{D}(\alpha, \beta)|$ if and only if the diagram of $\alpha$ contains the diagram of $\beta$ or vice versa. But there is a bijective correspondence

\[
\left\{ \text{unordered pairs of Young diagrams from } \mathcal{Y}(s, n-1) \text{ such that one diagram contains the other} \right\} \leftrightarrow \left\{ \text{plane partitions fitting in an } s \times (n-1) \times 2 \text{ box} \right\}.
\]

To see this, consider a pair of diagrams in the left-hand set and superimpose them so that their upper-left cells coincide. Filling any cell common to both diagrams with a “2,” and all the others with a “1,” we obtain a plane partition in the right-hand set. In
the other direction, starting with a plane partition from the right-hand set, we produce a unique pair of Young diagrams: the larger diagram is just the underlying diagram of the plane partition, and the smaller diagram is the collection of cells with the entry \(2\).

Therefore the number of unordered composition pairs such that \(\text{EMC} = |\mathcal{D}|\) is equal to \(\text{PP}(s, n - 1, 2)\). By Lemma 1 this number is

\[
\text{PP}(s, n - 1, 2) = \frac{(s + 1) \cdots (s + n - 1) \cdot (s + 2) \cdots (s + n)}{n!(n - 1)!}.
\]

To find the number of such ordered pairs in \(\mathcal{C}(s, n) \times \mathcal{C}(s, n)\), we multiply by 2 to account for both pairs \((\alpha, \beta)\) and \((\beta, \alpha)\), and then correct for the pairs \((\alpha, \alpha)\) by subtracting \(|\mathcal{C}(s, n)| = \binom{s + n - 1}{n - 1} = \frac{(s + 1) \cdots (s + n - 1)}{(n - 1)!}\). Hence the number of ordered pairs described in the theorem statement is

\[
2 \cdot \text{PP}(s, n - 1, 2) - \frac{(s + 1) \cdots (s + n - 1)}{(n - 1)!}.
\]

Finally, to convert this into a proportion of all ordered pairs, we divide by \(|\mathcal{C}(s, n)|^2\), and after simplifying we obtain

\[
2 \cdot \text{PP}(s, n - 1, 2) = \frac{(s + 1) \cdots (s + n - 1)}{n!(n - 1)!^2},
\]

\[
= \frac{(s + 1) \cdots (s + 2) \cdots (s + n)}{n!(n - 1)!^2} - \frac{(s + 1) \cdots (s + n - 1)}{(n - 1)!^2} = \frac{2(s + 1) \cdots (s + n)}{n(s + 1)} - \frac{(n - 1)!}{(s + 1) \cdots (s + n - 1)}.
\]

**Corollary 2** For fixed \(n\), as \(s \to \infty\), the proportion of elements in \(\mathcal{C}(s, n) \times \mathcal{C}(s, n)\) such that \(\text{EMC} = |\mathcal{D}|\) asymptotically approaches \(2/n\).

**Proof** Take the limit of the proportion in Theorem 2 as \(s \to \infty\).

Suppose that we had defined normalized versions of \(\text{EMC}\) and \(\mathcal{T}\), namely

\[
\widehat{\text{EMC}}(\alpha, \beta) := \frac{\text{EMC}(\alpha, \beta)}{s} \quad \text{and} \quad \widehat{\mathcal{T}}(\alpha) := \frac{\mathcal{T}(\alpha)}{s},
\]

in order to guarantee that both values lie on the interval \([0, n - 1]\). (In this way, \(\widehat{\mathcal{T}}\) is a true weighted average.) Then Corollary 2 is a statement about the continuous setting, i.e., pairs of probability distributions (rather than discrete histograms) on the set \([0, \ldots, n - 1]\). Indeed, both \(\text{EMC}\) and \(\mathcal{T}\) (and therefore \(\mathcal{D}\)) have natural analogs on probability distributions, obtained by letting \(s \to \infty\). Then for a randomly chosen ordered pair \((A, B)\) of probability distributions on \([0, \ldots, n - 1]\), the probability that \(\text{EMC}(A, B) = |\mathcal{D}(A, B)|\) is \(2/n\). This continuous result was proved using only the combinatorics of plane partitions.
5 $D$-distributions as weight diagrams for $\mathfrak{sl}_d$

5.1 Overview

In this section, we generalize the weighted difference $D$ in order to compare $d$ compositions rather than only two. Our $D$-values will no longer be single integers, but elements of the following $\mathbb{Z}$-module:

$$\Xi^d := \left\{ (w_1, \ldots, w_d) \mid w_i \in \mathbb{Z} \right\} / \mathbb{Z}(1, \ldots, 1).$$  \hfill (14)

Since elements of $\Xi^d$ are equivalence classes of $d$-tuples, we write an element using square brackets around any of its representatives. For example, $[5, 2, 1, 3] = [2, -1, -2, 0]$. As in this example, forcing the last entry to be 0 results in a unique way of writing an element as a $(d - 1)$-tuple; we will use this convention below whenever convenient.

We will have cause to use the term “weight” yet again, and in yet another sense: this time, the weights we consider are those from Lie theory (elements of the dual of the Cartan subalgebra). Specifically, we will consider those weights contained in the root lattice $\Lambda_R$ for the Lie algebra $\mathfrak{sl}_d = \mathfrak{sl}(d, \mathbb{C})$. The key observation in (16) will be that

$$\Xi^d \cong \Lambda_R$$

as $\mathbb{Z}$-modules, so that each $D$-value in $\Xi^d$ can be regarded as a weight of $\mathfrak{sl}_d$. (As the reader may have suspected, the results in this section will require some familiarity with representation theory and the Type A root system; we review these details in Section 6.)

In Theorem 3, we prove that the distribution of $D$-values on $\mathcal{C}(s, n)^d$, which is complicated combinatorial data, can be realized as the character of a certain virtual $\mathfrak{sl}_d$-representation; this means that the plot of $D$-values is the weight diagram of this representation. We provide illustrations of these diagrams for $d = 3$.

5.2 Weighted difference for arbitrary $d$

We now generalize the weighted difference $D$ to compare $d$ compositions. We want $D$ to capture just the right amount of information — namely, each of the pairwise differences between weighted totals, but nothing more — and so the most natural definition is the following.

**Definition 3** Given $s, n \geq 1$ and $d \geq 2$, we (re)define the **weighted difference** to be the function $D : \mathcal{C}(s, n)^d \rightarrow \Xi^d$ given by

$$D(\alpha_1, \ldots, \alpha_d) := [\mathcal{T}(\alpha_1), \ldots, \mathcal{T}(\alpha_d)].$$

**Remark 2** As mentioned in the overview, each element of $\Xi^d$ can be represented uniquely by a $d$-tuple whose final coordinate is 0, which can then be truncated into a $(d - 1)$-tuple. With this convention in mind, our new version of $D$ agrees with the old definition in the $d = 2$ case, since $D(\alpha, \beta) = [\mathcal{T}(\alpha), \mathcal{T}(\beta)] = [\mathcal{T}(\alpha) - \mathcal{T}(\beta), 0] \mapsto \mathcal{T}(\alpha) - \mathcal{T}(\beta)$, just as before.
We pose the natural combinatorial question: how many elements in \( \mathcal{C}(s, n)^d \) have a given \( D \)-value? It turns out that the answer can be interpreted as a weight multiplicity in the \( \mathfrak{sl}_d \)-representation described in the following theorem.

**Theorem 3** Let \( s, n \geq 1 \) and \( d \geq 2 \). For \( i = 1, \ldots, d-1 \), put \( x_i = e^{\sigma_i} \), and put \( x_d = e^{-(\sigma_1 + \cdots + \sigma_{d-1})} \), where \( \sigma_i \) is the \( i \)th simple root for \( \mathfrak{sl}_d \). Then the vector space

\[
V = \bigotimes^d \text{Sym}^s(\mathbb{C}^n)
\]

admits a virtual representation of \( \mathfrak{sl}_d \) whose character

\[
\text{char}_{\mathfrak{sl}_d}(V) = \prod_{i=1}^{d} \left[ s + n - 1 \atop s \right]_{x_i}
\]

is the distribution of \( D \)-values on \( \mathcal{C}(s, n)^d \). Specifically, the coefficient of \( x_1^{w_1} \cdots x_d^{w_d} \) equals the number of \( d \)-tuples whose \( D \)-value is \( [w_1, \ldots, w_d] \).

We reserve the proof for Section 6, along with the necessary preliminaries. Note that the relation \( x_d = (x_1 \cdots x_{d-1})^{-1} \) allows us to write the character in terms of just the variables \( x_1, \ldots, x_{d-1} \); then (as the reader may already see, but we will explain in Section 6) the correspondence in the theorem boils down to the following dictionary between \( D \)-values and \( \mathfrak{sl}_d \)-weights:

\[
\begin{array}{c}
\mathcal{C}(s, n)^d \\
\Leftrightarrow \Lambda_R \\
[w_1, \ldots, w_{d-1}, 0] \Leftrightarrow w_1 \sigma_1 + \cdots + w_{d-1} \sigma_{d-1}.
\end{array}
\]

**Example 3** Let \( d = 3 \), with \( s = 1 \) and \( n = 2 \). Then by the character formula in Theorem 3 we have

\[
\text{char}_{\mathfrak{sl}_3}(V) = \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{x_1} \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{x_2} \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{x_3}
\]

\[
= \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{x_1} \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{x_2} \left[ \begin{array}{c}
2 \\
1 \end{array} \right]_{(x_1 x_2)^{-1}}
\]

\[
= x_1 + x_2 + x_1 x_2 + x_1^{-1} + x_2^{-2} + x_1^{-1} x_2^{-1} + 2.
\]

Lining up each term’s exponent vector \( [w_1, w_2] \) with its corresponding \( D \)-value \( [w_1, w_2, 0] \), we conclude that in \( \mathcal{C}(1, 2) \times \mathcal{C}(1, 2) \times \mathcal{C}(1, 2) \), there is one triple having each of the \( D \)-values

\[
[1, 0, 0], \quad [0, 1, 0], \quad [1, 1, 0], \quad [-1, 0, 0], \quad [0, -1, 0], \quad \text{and} \quad [-1, -1, 0],
\]

and two triples for which \( D = [0, 0, 0] \). The reader can easily verify this distribution by hand.
5.3 $\mathcal{D}$-distributions as $\mathfrak{sl}_d$-weight diagrams

We would like to visualize the distribution of $\mathcal{D}$-values for a given $s$ and $n$. Theorem 3 states that such a visualization is nothing other than the weight diagram of $V = \bigotimes^3 \text{Sym}^s(\mathbb{C}^n)$ as a virtual $\mathfrak{sl}_d$-representation.

For $d = 2$, of course, we can obtain the desired diagram simply by taking our planar plots from the previous section and "projecting" (via summation) each vertical strip down onto the horizontal axis; in this way, each integer on the number line is labeled with the total number of composition pairs having that $\mathcal{D}$-value. (Essentially, this is just throwing away the EMC information from the plots.) The root lattice is
generated by the sole simple root $\sigma_1$, pointing in the positive $D$-direction. On one hand, we know that these “column sums” from our old plots count the number of Young diagram pairs which differ in size by a fixed number of cells; on the other hand, Theorem 3 says that these sums are just the weights of $\text{Sym}^s(C^n) \otimes \text{Sym}^t(C^n)$ as an $\mathfrak{sl}_2$-representation.

For $d = 3$, the root lattice is a two-dimensional triangular lattice generated by $\sigma_1$ (pointing due east, at 0 degrees) and $\sigma_2$ (pointing northwest, at 120 degrees). Then (15) tells us where to plot each exponent vector in $\text{char} \mathfrak{sl}_d$, which we label with the corresponding coefficient. From Example 3, we obtain a regular hexagon of dots labeled “1,” along with a dot at the origin labeled “2.” (See the first diagram in Figure 7, which the reader may recognize as the weight diagram of the adjoint representation of $\mathfrak{sl}_3$.) In Figure 7, for $d = 3$, we display the weight diagrams (i.e., $D$-distributions) for several other values of $s$ and $n$. Incidentally, Theorem 3 gives a much more efficient way to create these graphics than does the generating function in Section 7, especially for large values of $n$.

Remark 3 We consider another, perhaps more natural way to think about the plots of $D$-distributions; we remain with the $d = 3$ case, although this approach holds in any dimension.

For any ordered triple $(\alpha, \beta, \gamma)$ of compositions, draw three axes radiating from the origin (as we do in Figure 7): the $\alpha$-axis in the direction of $\sigma_1$, the $\beta$-axis in the direction of $\sigma_2$, and the $\gamma$-axis in the direction of $-(\sigma_1 + \sigma_2)$, i.e., southwest at 240 degrees. (These are the weights corresponding to $x_1, x_2,$ and $x_3$ in the character formula.) Now we can plot a $D$-value $[a, b, c]$, whether or not its final coordinate is 0, by starting at the origin and moving $a$ units parallel to the $\alpha$-axis, $b$ units parallel to the $\beta$-axis, and $c$ units parallel to the $\gamma$-axis. Notice that any multiple of $[1, 1, 1]$ lands on the origin, just as we would expect.

Now the angular position of a $D$-value tells us immediately the ordering of the three weighted totals: for example, at all points strictly between 120 and 180 degrees, we must have $T(\beta) > T(\gamma) > T(\alpha)$. The radial distance is a measure of the “lopsidedness” among the three weighted totals. By the symmetry inherent in our setup, we can get a full picture of the diagram by restricting our attention to just one sector, such as that between 0 and 60 degrees where $T(\alpha) \geq T(\beta) \geq T(\gamma)$. Although we have focused on $D$ at the expense of EMC in this section, we could nevertheless include the EMC frequencies in the $d = 3$ diagrams by plotting them in the third dimension, along strips directly above each $D$-value, just as in our $d = 2$ plots.

From the perspective of representation theory, rather than work with the entire weight diagram of $V$, it is natural to seek the decomposition of $V$ into irreducible subrepresentations of $\mathfrak{sl}_3$. Since finite-dimensional irreducible representations are indexed by their highest weight, we can actually convey the same information from the weight diagram with a much sparser diagram, on which we plot only those highest weights corresponding to the irreducible decomposition of $V$. (See Figure 8, where we show only the region of the root lattice inside the dominant chamber.) As experts will
have anticipated, we programmed these decompositions by multiplying $\text{char}_{\mathfrak{sl}_3^1}(V)$ by the Weyl denominator $(1 - x_1^{-1})(1 - x_2^{-1})(1 - (x_1 x_2)^{-1})$, and then plotting the dominant weights in the resulting expansion. In these plots, the number inside each dot indicates the multiplicity of that irreducible representation inside $V$. Note that some multiplicities are negative (depicted by red dots), reflecting the fact that $V$ is actually a virtual representation of $\mathfrak{sl}_3^1$, i.e., a formal $\mathbb{Z}$-combination of irreducible representations, rather than a direct sum. When $s = 1$, we see that $V$ is isomorphic to the irreducible representation of $\mathfrak{sl}_3^1$ with highest weight $(n - 1) (\sigma_1 + \sigma_2)$. As $s$ or $n$ increases, certain patterns are emerging from the diagrams, but this decomposition will be an interesting subject of further research. The three-dimensional diagrams in the $d = 4$ case will likely be of interest as well.

6 Representation theory and proof of Theorem 3

6.1 Essentials of representation theory

We provide a cursory overview with a view to proving Theorem 3, but this section nonetheless assumes some familiarity with representation theory and the $A_{d-1}$ root system. Specifically, $\text{SL}(d, \mathbb{C})$ is the group of complex $d \times d$ matrices with determinant 1. Its Lie algebra $\mathfrak{sl}_d = \mathfrak{sl}(d, \mathbb{C})$ consists of complex $d \times d$ matrices with trace 0, equipped with the commutator bracket. As usual, $\epsilon_i$ denotes the weight (i.e., a linear functional on the Cartan subalgebra $\mathfrak{h}$ of diagonal matrices in $\mathfrak{sl}_d$) which takes $\text{diag}[h_1, \ldots, h_d]$ to the scalar $h_i \in \mathbb{C}$. The trace-zero condition implies that

$$\epsilon_1 + \cdots + \epsilon_d = 0.$$  

In a slight departure from standard notation, we will denote the $i$th simple root by $\sigma_i := \epsilon_i - \epsilon_{i+1}$ for $i = 1, \ldots, d - 1$. (The only reason we use $\sigma_i$ rather than the usual $\alpha_i$ is because we have been using $\alpha_i$ to denote compositions.) The simple roots generate the root lattice

$$\Lambda_R := \mathbb{Z} \sigma_1 \oplus \cdots \oplus \mathbb{Z} \sigma_{d-1} \subset \mathfrak{h}^*.$$  

Since $\sum_{i=1}^{d-1} \sigma_i = \epsilon_1 - \epsilon_d$, we have

$$\Lambda_R \cong \left\{ \sum_{i=1}^{d-1} w_i \sigma_i + w_d(\epsilon_d - \epsilon_1) \left| w_i \in \mathbb{Z} \right. \right\} \bigg/ \mathbb{Z} \left\{ \sum_{i=1}^{d-1} \sigma_i + (\epsilon_d - \epsilon_1) \right\}$$  

as $\mathbb{Z}$-modules. Comparing this with the definition of $\tilde{\mathbb{Z}}^d$ in (14), we have the $\mathbb{Z}$-module isomorphism

$$\tilde{\mathbb{Z}}^d \cong \Lambda_R,$$

$$[w_1, \ldots, w_d] \mapsto w_1 \sigma_1 + \cdots + w_{d-1} \sigma_{d-1} + w_d (\epsilon_d - \epsilon_1),$$

$$[w_1, \ldots, w_{d-1}, 0] \mapsto \sum_{i=1}^{d-1} w_i \sigma_i.$$  

(16)
Fig. 8: The decomposition of $V = \bigotimes^3 \text{Sym}^s(\mathbb{C}^n)$ as a virtual $\mathfrak{sl}_3$-representation. Each irreducible representation of $\mathfrak{sl}_3$ is plotted (with multiplicity) at its highest weight. Negative multiplicities are colored red.
This isomorphism will be our dictionary between $D$-values on $\mathcal{C}(s, n)^d$ and weights of $\mathfrak{sl}_d$.

A **representation** of a complex Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism $\mathfrak{g} \to \text{End}(W)$ for some complex vector space $W$; we often say that the underlying space $W$ is the representation, with the homomorphism implied, or we say that “$\mathfrak{g}$ acts on $W$.” All representations in this paper are finite-dimensional (with the exception of the supplemental Section 6.3 aimed at experts). Given a representation $W$, its **character** $\text{char}_W$ is a formal sum of $q$-weights $\lambda$, often written in exponential notation, such that the coefficient of $e^{\lambda}$ equals the dimension of the $\lambda$-weight space in $W$. (We omit the definition of a weight space since it is not essential for our result; to supplement our summary here, the reader is referred to [7] or [8], among the many references on representation theory.) The exponential notation ensures that *multiplying* terms formally encodes *adding* weights: $e^{\lambda} e^{\mu} = e^{\lambda + \mu}$. In this way, $\text{char}_W(W \otimes W') = \text{char}_W(W) \cdot \text{char}_W(W')$.

Back in the $\mathfrak{sl}_d$ setting, with the dictionary (16) in mind, we now set

$$x_i = \begin{cases} e^{\alpha_i} & 1 \leq i \leq d-1, \\ e^{\alpha_d - \alpha_i} & i = d, \end{cases}$$

(17)

so that we can write $\mathfrak{sl}_d$-characters in terms of the $x_i$. But by substituting $x_d = (x_1 \cdots x_{d-1})^{-1}$, we can actually write characters in terms of just $x_1, \ldots, x_{d-1}$.

Now we fix $s, n \geq 1$ and $d \geq 2$. The Lie algebra $\mathfrak{gl}_n = \mathfrak{gl}(n, \mathbb{C})$ of complex $n \times n$ matrices acts naturally on the vector space $\mathbb{C}^n$ by matrix multiplication. Therefore $\mathfrak{gl}_n$ also acts (irreducibly) on $\text{Sym}^s(\mathbb{C}^n)$, the $s$th symmetric power of $\mathbb{C}^n$.

Meanwhile, $\mathfrak{sl}_2$ also acts irreducibly on $\mathbb{C}^n \cong \text{Sym}^{n-1}(\mathbb{C}^2)$. The reader can consult Section 2.3.2 of [8] for the explicit action via differential operators on homogeneous polynomials, but for us the important information is the character of this $\mathfrak{sl}_2$-action:

$$\text{char}_{\mathfrak{sl}_2}(\mathbb{C}^n) = q^{-(n-1)} + q^{-(n-3)} + \cdots + q^{-3} + q^{-1} = (n)_q,$$

where $q = e^{\alpha_1}$. Then $\mathfrak{sl}_2$ also acts on $\text{Sym}^s(\mathbb{C}^n)$ with the character

$$\text{char}_{\mathfrak{sl}_2} \text{Sym}^s(\mathbb{C}^n) = \binom{s+n-1}{s}_q,$$

(18)

the second guise of the $q$-binomial coefficient from Section 2.4 (See [8], Theorem 4.1.20.) Because $\mathfrak{gl}_n \cong \text{End}(\mathbb{C}^n)$, and because $\mathbb{C}^n$ is a representation of $\mathfrak{sl}_2$, there naturally exists a **principal embedding** $\mathfrak{sl}_2 \hookrightarrow \mathfrak{gl}_n$, unique up to conjugation. Hence it makes sense to restrict the $\mathfrak{gl}_n$-action on $\text{Sym}^s(\mathbb{C}^n)$ to the principally embedded $\mathfrak{sl}_2$.

Intuitively, we should regard $\text{Sym}^s(\mathbb{C}^n)$ as a linearization of $\mathcal{C}(s, n)$, i.e., a vector space whose basis is in bijective correspondence with $\mathcal{C}(s, n)$: as a $\mathfrak{gl}_n$-representation, its basis consists of one weight vector $\prod_{i=0}^{n-1} e_i^{a_i}$ for each composition $\alpha = (a_0, \ldots, a_{n-1}) \in \mathcal{C}(s, n)$. (Here the $e_i$ are the standard basis vectors of $\mathbb{C}^n$, with indexing shifted down by 1.) Since we want to compare $d$ compositions, we will study the $d$-fold tensor product

$$V = \bigotimes^d \text{Sym}^s(\mathbb{C}^n).$$
6.2 Proof of Theorem 3

The vector space \( V \) is an irreducible representation of \( (\mathfrak{gl}_n)^d = \mathfrak{gl}_n \oplus \cdots \oplus \mathfrak{gl}_n \), with each copy of \( \mathfrak{gl}_n \) acting on the corresponding tensor factor. Restricting to the principally embedded \( \mathfrak{sl}_2 \) in each copy of \( \mathfrak{gl}_n \), we see that \( V \) is also a representation of \( (\mathfrak{sl}_2)^d \). Crucially, each copy of \( \mathfrak{sl}_2 \) embeds into \( \mathfrak{sl}_d \), as follows. For \( i = 1, \ldots, d-1 \), let \( s_i \subset \mathfrak{sl}_d \) denote the image of the \( \mathfrak{sl}_2 \)-embedding given by the Lie algebra homomorphism

\[
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto a \cdot E_{i,i} + b \cdot E_{i,i+1} + c \cdot E_{i+1,i} + d \cdot E_{i+1,i+1},
\]

where \( E_{i,j} \) is the elementary matrix with 1 in position \((i, j)\) and 0’s elsewhere. Similarly, define \( s_d \subset \mathfrak{sl}_d \) as the image of the embedding

\[
\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto -a \cdot E_{1,1} - c \cdot E_{1,d} - b \cdot E_{d,1} - d \cdot E_{d,d}
\]

via the negative transpose in the “four corners.”

In this way, not only does each \( s_i \) inherit (from \( \mathfrak{sl}_2 \)) a natural action on the \( i \)-th tensor factor of \( V \), but this action extends to a (virtual) \( \mathfrak{sl}_d \)-action, as follows:[3] We have

\[
\mathbb{C}^d \cong \text{span}\{e_{i-1}, e_{i+1}\} \oplus \text{span}\{e_1, \ldots, e_{i-2}, e_{i+2}, \ldots, e_d\}
\]

\[
\cong \mathbb{C}^2 \oplus \mathbb{C}^{d-2},
\]

where \( \mathbb{C}^2 \) is the defining representation of \( s_i \cong \mathfrak{sl}_2 \), and \( \mathbb{C}^{d-2} \) is the direct sum of \( d-2 \) copies of the trivial representation of \( \mathfrak{sl}_d \). In other words, we let \( \mathfrak{h} \subset \mathfrak{sl}_d \) act on \( \mathbb{C}^d \) not as in the defining representation, but via

\[
\text{diag}[h_1, \ldots, h_d] \cdot (v_1, \ldots, v_d)^\top = (0, \ldots, 0, h_1v_1, h_{i+1}v_{i+1}, 0, \ldots, 0)^\top.
\]

Hence we are setting \( e_j = 0 \) for all \( j \) except \( i \) and \( i + 1 \), so that the character of this action is \( \text{char}_{\mathfrak{sl}_d}(\mathbb{C}^d) = e^{e_i} + e^{e_{i+1}} + (d-2) \). We then regard \( \mathbb{C}^2 \) as the formal difference \( \mathbb{C}^d - \mathbb{C}^{d-2} \) of \( \mathfrak{sl}_d \)-representations. Recalling that \( e_i = -e_{i+1} \) as \( s_i \)-weights, we have the virtual character

\[
\text{char}_{\mathfrak{sl}_d}(\mathbb{C}^2) = \text{char}_{\mathfrak{sl}_d}(\mathbb{C}^d) - (d-2)(\text{char}_{\mathfrak{sl}_d}(\mathbb{C}))
\]

\[
= e^{e_i} + e^{e_{i+1}} + (d-2) - (d-2)(1)
\]

\[
= e^{e_i} + e^{e_{i+1}}
\]

\[
= e^{e_i} + e^{-e_i}
\]

\[
= q_i + q_i^{-1} = \text{char}_{s_i}(\mathbb{C}^2),
\]

where we write \( q_i = e^{e_i} \). Hence, the \( i \)-th tensor factor \( \text{Sym}^i(\mathbb{C}^n) \cong \text{Sym}^i(\text{Sym}^{n-1}(\mathbb{C}^2)) \) of \( V \) also admits a virtual \( \mathfrak{sl}_d \)-representation, whose character, by [18], is

\[
\text{char}_{\mathfrak{sl}_d}(\text{Sym}^s(\mathbb{C}^n)) = \binom{s+n-1}{s}_{q_i}.
\]

---

[3] In the following calculations, we must treat the special case \( i = d \) differently: namely, replace all instances of \( i + 1 \) by 1, to respect the embedding \( s_d \).
But since $e_i = -e_{i+1}$, we have the substitution $x_i = q_i^2$, with $x_i$ as defined in (17). Thus by (8), we have

$$\text{char}_{\mathfrak{sl}_d} \text{Sym}^n(C^n) = \left( \frac{s + n - 1}{s} \right)_{q_i} = q_i^{-s(n-1)} \left[ \begin{array}{c} s + n - 1 \\ s \end{array} \right]_{x_i}.$$  

Multiplying all these $\mathfrak{sl}_d$-characters for $i = 1, \ldots, d$ in order to obtain the character of the $d$-fold tensor product $V$, we have

$$\text{char}_{\mathfrak{sl}_d}(V) = \prod_{i=1}^{d} q_i^{-s(n-1)} \left[ \begin{array}{c} s + n - 1 \\ s \end{array} \right]_{x_i}.$$  

But by (15), we have $\prod_{i=1}^{d} q_i^{-s(n-1)} = (\prod_{i=1}^{d} e^{e_i})^{-s(n-1)} = (e^{e_1} \cdots e_d)^{-s(n-1)} = (e^{0})^{-s(n-1)} = 1$, leaving us with

$$\text{char}_{\mathfrak{sl}_d}(V) = \prod_{i=1}^{d} \left[ \begin{array}{c} s + n - 1 \\ s \end{array} \right]_{x_i}$$  

which is the main claim of Theorem 3.

It remains to show that the character (19) encodes the distribution of $D$-values on $C(s, n)^d$. By (7), the coefficient of $x_1^{w_1} \cdots x_d^{w_d}$ in the expansion of (19) is the number of $d$-tuples of Young diagrams $(\alpha_1, \ldots, \alpha_d) \in \mathcal{Y}(s, n - 1)^d$ such that $|\alpha_i| = w_i$. But by (15) and Definition 3, this is the number of $d$-tuples $(\alpha_1, \ldots, \alpha_d) \in \mathcal{C}(s, n)^d$ such that $D(\alpha_1, \ldots, \alpha_d) = (w_1, \ldots, w_d)$. Making the substitution $x_d = (x_1 \cdots x_{d-1})^{-1}$ and using the dictionary in (16), we obtain a well-defined exponent vector for each $D$-value.

Remark 4 From this perspective of representation theory, the $(s, n) \leftrightarrow (n-1, s+1)$ duality is a manifestation of Hermite reciprocity: as $\mathfrak{sl}_2$-representations,

$$\text{Sym}^s(C^n) \cong \text{Sym}^s \left( \text{Sym}^{n-1}(C^2) \right)$$  

$$\cong \text{Sym}^{n-1} \left( \text{Sym}^s(C^2) \right)$$  

$$\cong \text{Sym}^{n-1}(C^{s+1}).$$

This can be seen directly from the $q$-binomial coefficient that expresses the $\mathfrak{sl}_2$-character in (18). Since the distribution of $D$-values is completely determined by the $\mathfrak{sl}_2$-action on $V$, which in turn is completely determined by the actions of copies of $\mathfrak{sl}_2$, it is clear why $C(s, n)^d$ must have the same $D$-distribution as $C(n-1, s+1)^d$.

6.3 The first Wallach representation

This subsection offers a deeper perspective into the representation theory of $V$ when $d = 2$. We present here the bare facts, referring the reader to [5], page 343, and [16] for details.
When \( d = 2 \), it turns out that \( V = \text{Sym}^2(\mathbb{C}^n) \otimes \text{Sym}^2(\mathbb{C}^n) \) is the \( 2^{\text{th}} \) graded component of an infinite-dimensional vector space \( W_n \) known as the first Wallach representation of the indefinite unitary group \( U(n,n) \). The underlying space of this representation is

\[
W_n := \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]^{\text{GL}_1} \cong \mathbb{C}[[x]],
\]

the coordinate ring of the determinantal variety \( \mathcal{D}_n \) of \( n \times n \) complex matrices with rank at most 1. In the first line above, the superscript denotes the \( \text{GL}_1 \)-invariants in the polynomial ring, where \( g \in \text{GL}_1 \cong \mathbb{C}^\times \) acts on the polynomial function \( f \) via

\[
(g \cdot f)(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(g^{-1}x_1, \ldots, g^{-1}x_n, g^{-1}y_1, \ldots, g^{-1}y_n).
\]

It follows that \( W_n \) is generated as an algebra by the monomials \( x_1^{a_1} \cdots x_n^{a_n} y_1^{b_1} \cdots y_n^{b_n} \), for \( 1 \leq i, j \leq n \). (This is a special case of the First Fundamental Theorem of Invariant Theory; see [6], Theorem 5.2.1.) Therefore, letting \( \mathbb{C}[\cdot]^s \) denote the subspace of degree-\( s \) homogeneous polynomial functions,

\[
W_n = \bigoplus_{s=0}^{\infty} W_n^s \cong \bigoplus_{s=0}^{\infty} \mathbb{C}[x_1, \ldots, x_n]^s \otimes \mathbb{C}[y_1, \ldots, y_n]^s \cong \bigoplus_{s=0}^{\infty} \text{Sym}^s(\mathbb{C}^n) \otimes \text{Sym}^s(\mathbb{C}^n).
\]

In the context of our paper, given \( s \) and \( n \), we then have \( V \cong W_n^s \). It is clear how the monomial \( x_1^{a_1} \cdots x_n^{a_n} y_1^{b_1} \cdots y_n^{b_n} \) corresponds (via the exponent vectors for the \( x_i \) and \( y_j \)) to a composition pair in \( \mathcal{C}(s,n) \times \mathcal{C}(s,n) \). As we observed earlier in this section, \( V \cong W_n^s \) is naturally a representation of \( \mathfrak{gl}_n \oplus \mathfrak{gl}_n \). The full picture, however, is that all of \( W_n \) is a representation of \( \mathfrak{u}(n,n) \). By restricting the action to the maximal compact subalgebra \( \mathfrak{u}(n,n) \) and then complexifying, we have an action of \( \mathfrak{gl}_n \oplus \mathfrak{gl}_n \), given (up to a central shift) by the Euler operators

\[
\frac{\partial}{\partial x_j} \quad \text{and} \quad \frac{\partial}{\partial y_j}.
\]

Returning to our mental image from the introduction, we can regard each of these operators as acting on one of the two compositions by moving one pebble from pile \( j \) to pile \( i \). But if we consider the full action of \( \mathfrak{u}(n,n) \) (and hence of its complexification \( \mathfrak{gl}_{2n} \)), rather than just the restriction to \( \mathfrak{gl}_n \oplus \mathfrak{gl}_n \), then the graded component \( W_n^s \) is no longer preserved since there are “raising” and “lowering” operators of the forms

\[
\frac{\partial^2}{\partial x_i \partial y_j} \quad \text{and} \quad x_i y_j,
\]

By a “compact Lie algebra” we mean the Lie algebra of a compact Lie group, in this case \( U(n) \times U(n) \).
which decrease or increase the value of $s$ by 1, by removing or adding one pebble from/to each composition. (See [4], Section 2, for the $\mathfrak{g}l_{2n}$-action written out explicitly in terms of these differential operators.)

7 A generating function

In [11], for the $d = 2$ case, a recursive definition is given for a generating function that keeps track of the number of composition pairs with given EMC and given weighted totals. Because of the recursive nature of the definition, we must formally consider composition pairs with unequal numbers of “piles” ($n$ and $m$). As noted in [3], the EMC is still defined on $C(s, n) \times C(s, m)$ since we can append zeros so that all compositions are tuples of length $\max\{n, m\}$. We do not consider the statistical meaning of the EMC in such a scenario; the unequal pile-numbers are just a means to obtain our recursion.

Proposition 2 Formally, let

$$H_{n, m}(q, x, y, t) := \sum_{s=0}^{\infty} \left( \sum_{\alpha \in C(s, n), \beta \in C(s, m)} q^{\text{EMC}(\alpha, \beta)} x^\alpha y^\beta \right) t^s, \quad (20)$$

so that the coefficient of $t^s$ is a polynomial whose coefficients give the number of composition pairs in $C(s, n) \times C(s, m)$ with given EMC and given weighted totals. Then we have the recursion

$$H_{n, m} = H_{n-1, m} + H_{n, m-1} - H_{n-1, m-1} \frac{1 - q}{1 - q C(n, m) x^{n-1} y^{m-1} t} \quad (21)$$

where $H_{1, 1} = \frac{1}{1 - q}$, and $H_{0, m} = H_{n, 0} = 0$.

The proof given in [11] is a specialization of that from [3]. Here, however, we use our main result to give a short proof entirely in terms of Young diagrams.

Proof By Corollary [1] and by [13], we can reinterpret the generating function (20) in terms of Young diagrams:

$$H_{n, m}(q, x, y, t) = \sum_{s=0}^{\infty} \left( \sum_{\alpha \in Y(s, n-1), \beta \in Y(s, m-1)} q^{\alpha \beta} x^{|\alpha|} y^{|\beta|} \right) t^s.$$ 

For the base case $H_{1, 1}$, it is clear that $Y(s, 0)$ contains only the empty diagram, so the coefficient of $t^s$ has only one term; since the empty diagram has size 0, the exponents of $x$ and $y$ in this term are both 0. Moreover, the symmetric difference of two empty diagrams is 0, so the exponent of $q$ is also 0. Hence the coefficient of $t^s$ is $q^0t^0 = 1$, so $H_{1, 1} = 1 + t + t^2 + \cdots = \frac{1}{1 - t}$. In case $n$ or $m$ is 0, we have $Y(s, -1) = \emptyset$, since a Young diagram cannot have a negative number of columns.
Now we consider $H_{n,m}$ in general. For any pair $(\alpha, \beta)$ in the inside sum, the diagram $\alpha$ has a certain number $r_\alpha$ of maximal rows (length $n - 1$) at the top, and likewise $\beta$ has a certain number $r_\beta$ of maximal rows (length $m - 1$) at the top. Hence we can remove $r := \min\{r_\alpha, r_\beta\}$ maximal rows from each diagram. Since $t$ tracks the number of rows in each diagram, each removed row-pair contributes 1 to the exponent of $t$. Note that the symmetric difference of a maximal row from $\alpha$ and a maximal row from $\beta$ is exactly $j_{n-m} = C(1, n-m)$, so each of the removed row-pairs contributes $C(n, m)$ to the exponent of $q$. Moreover, each removed row-pair contributes the size $n - 1$ to the exponent of $x$, and $m - 1$ to the exponent of $y$. Therefore we see that from each term in $H_{nm}$ we can factor out $(q^{C(n,m)}x^{n-1}y^{m-1}t)^r$ for some $r$, leaving us with a pair of diagrams at least one of which contains no maximal rows; that is, at least one of them can fit inside a rectangle with one less column than before. Hence, starting with $H_{nm}$ and factoring out the reciprocal of the denominator in (21), the remaining terms are precisely those appearing in $H_{n-1,m}$ or $H_{n,m-1}$. After correcting for double-counting the terms appearing in both, we have the numerator of (21), and so we are done.

To create the plots in Figures 3–5, we make the substitutions

$$x = z, \quad y = z^{-1}$$

so that the exponent of $z$ in each term encodes $\mathcal{D}(\alpha, \beta)$. Moreover, the generating function $H_{n,m}$ naturally generalizes to an arbitrary number $d$ of compositions, following the methods in [6]; the numerator of $H_{(n_1, \ldots, n_d)}$ simply contains a longer alternating sum to extend the inclusion–exclusion argument to $d$ Young diagrams, and the denominator becomes $1 - q^{C(n_1, \ldots, n_d)}x_1^{n_1-1}\cdots x_d^{n_d-1}t$. Once again, the encoded results can be interpreted either statistically — in terms of the EMC and weighted totals $\mathcal{T}$ of compositions — or combinatorially — in terms of the unimodal symmetric difference $\Delta$ and sizes $|\cdot|$ of Young diagrams.

7.1 Closed-form examples

We first examine the specific case $H_2 := H_{2,2}$, which encodes information about the EMC and weighted totals of composition pairs in $\mathcal{C}(s, 2)$. Either by hand or by software, we use the recursion (21) to find

$$H_2 = \frac{1 - q^2 x y t^2}{(1 - t)(1 - x y t)(1 - q x t)(1 - q y t)}.$$

We interpret this in terms of Young diagram pairs, as above. In the denominator, each of the four factors corresponds to one of the ways to choose a row in each diagram simultaneously: $(1 - t)$ is an empty row in both diagrams, $(1 - x y t)$ is a 1-cell row in both diagrams, $(1 - q x t)$ is a 1-cell row in $\alpha$ only (hence contributing 1 to the symmetric difference), and $(1 - q y t)$ is a 1-cell row in $\beta$ only (again contributing 1 to the symmetric difference). But we cannot choose just any combination of these four row-pairs and still obtain a pair of Young diagrams: specifically, we cannot choose rows from both the third and fourth factors, since it is impossible for a Young diagram...
to have a nonempty row below an empty row. This explains the numerator $1 - q^2 xyt^2$, which effectively kills any term containing the product $(qxt)(qyt)$.

For values of $n$ greater than 2, the closed form of $H_n$ spills over into several lines, or pages, of text; an explicit (as opposed to recursive) formula involves an enormous alternating sum of least common multiples, ranging over the power set of a certain set of monomials (see the partial order we define below, and consider the set of noncomparable pairwise products). As an illustrative example, we take $n = 3$ and inspect the denominator of $H_3$:

$$(1-t)(1-xyt)(1-x^2y^2t)(1-qxt)(1-qxyt)(1-qx^2yt)(1-q^2x^2t)(1-q^2y^2t)$$

Again, each of these factors describes one way to choose a row in each diagram simultaneously, where the rows may now contain 0, 1, or 2 cells; notice that the exponent of $q$ is always the difference between those of $x$ and $y$. But again, we cannot choose just any combination of row-pairs to build a pair of Young diagrams. This time, the “forbidden” combinations seem more daunting to describe; to resolve this, we put a partial order on monomials (ignoring the $q$ and $t$ exponents), such that $x^a y^b \leq x^{a'} y^{b'}$ if and only if $a \leq a'$ and $b \leq b'$. We now see that two given row-pairs can coexist in $\alpha$ and $\beta$ if and only if their monomials are comparable under our partial order. (This is a combinatorial translation of the “matrix straightening” from [3] via modding out by a determinantal ideal.) Hence the “forbidden” combinations of factors are precisely those which contain the product of two noncomparable monomials from the denominator. This necessitates an alternating sum of terms in the numerator, which (despite many cancellations) quickly becomes intractable as $n$ increases, making the recursive definition the only viable computational approach.

The coefficient of $t^2$ in the expansion of $H_3$ is

$$1 + xy + 2x^2y^2 + x^3y^3 + x^4y^4$$
$$+ q(x + y + 2x^2y + 2xy^2 + 2x^3y^2 + 2x^2y^3 + x^4y^3 + x^3y^4)$$
$$+ q^2(2x^2 + x^3y + 2y^2 + 2x^2y^2 + 2x^4y^2 + xy^3 + 2x^2y^4)$$
$$+ q^3(x^3 + x^4y + y^3 + xy^4)$$
$$+ q^4(x^4 + y^4).$$

The reader can verify that these indeed encode all 36 ordered pairs of Young diagrams fitting inside a $2 \times 2$ rectangle, along with the sizes of the diagrams and the symmetric difference between them. The reader can also compare this expansion with the second plot in Figure 3 where $s = 2$ and $n = 3$.

Declarations

Funding: not applicable.
Conflict of interest: the author declares that he has no conflict of interest.
Availability of data and material: not applicable.
Code availability: not applicable.
References

[1] Alok Aggarwal and James K. Park, Sequential searching in multidimensional monotone arrays, IBM T.J. Watson Research Center, Yorktown Heights, NY, 1989. Research Report RC15128.

[2] Wolfgang Bein, Peter Brucker, James K. Park, and Pramod Pathak, A Monge property for the d-dimensional transport problem, Discrete Appl. Math. 58 (1995), no. 2, 97–109.

[3] Rebecca Bourn and Jeb F. Willenbring, Expected value of the one-dimensional earth mover’s distance, Algebr. Stat. 11 (2020), no. 1, 53–78.

[4] Mark Colarusso, William Q. Erickson, and Jeb F. Willenbring, Contingency tables and the generalized Littlewood–Richardson coefficients, Proc. Amer. Math. Soc. (2021). To appear. arXiv:2012.06928.

[5] Thomas J. Enright and Jeb F. Willenbring, Hilbert series, Howe duality, and branching for classical groups, Ann. of Math. 159 (2004), no. 1, 337–375.

[6] William Q. Erickson, A generalization for the expected value of the earth mover’s distance, Algebr. Stat. (2021). To appear. arXiv:2009.12723v2.

[7] William Fulton and Joe Harris, Representation theory: a first course, Graduate Texts in Mathematics: Readings in Mathematics, vol. 129, Springer, 1991.

[8] Roe Goodman and Nolan R. Wallach, Symmetry, representations, and invariants, Graduate Texts in Mathematics, vol. 255, Springer, 2009.

[9] Jeffery Kline, Properties of the d-dimensional earth mover’s problem, Discrete Appl. Math. 265 (2019), 128–141.

[10] Patrick T. Komiske, Eric M. Metodiev, and Jesse Thaler, Metric space of collider events, Phys. Rev. Lett. 123 (2019), 041801.

[11] Jan Kretschmann, Earth mover’s distance between grade distribution data with fixed mean, Master’s Thesis, University of Wisconsin–Milwaukee, 2020.

[12] Noam Lupu, Lucía Selios, and Zach Warner, A new measure of congruence: the earth mover’s distance, Political Analysis 25 (2017), 95–113.

[13] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas, The earth mover’s distance as a metric for image retrieval, Int. J. Comput. Vis. 40 (2000), 99–121.

[14] Richard P. Stanley, Enumerative combinatorics, volume 2, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, 1999.

[15] ———, Enumerative combinatorics, volume 1, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, 2012.

[16] Nolan R. Wallach, The analytic continuation of the discrete series, I and II, Trans. Amer. Math. Soc. 251 (1979), 1–37.