4-2014

New combinatorial formulations of the shuffle conjecture

Nicholas A. Loehr

Elizabeth Niese
Marshall University, niese@marshall.edu

Recommended Citation

Loehr, N. A., & Niese, E. (2014). New combinatorial formulations of the shuffle conjecture. Advances in Applied Mathematics, 55, 22-47.

New combinatorial formulations of the shuffle conjecture

Nicholas A. Loehr[1]

Virginia Tech Dept. of Mathematics, Blacksburg, VA 24061
U.S. Naval Academy Math Dept., Annapolis, MD 21402

Elizabeth Niese

Department of Mathematics, Marshall University, Huntington, WV 25755

Abstract

The shuffle conjecture (due to Haglund, Haiman, Loehr, Remmel, and Ulyanov) provides a combinatorial formula for the Frobenius series of the diagonal harmonics module DH_n, which is the symmetric function $\nabla(e_n)$. This formula is a sum over all labeled Dyck paths of terms built from combinatorial statistics called area, dinv, and IDes. We provide three new combinatorial formulations of the shuffle conjecture based on other statistics on labeled paths, parking functions, and related objects. Each such reformulation arises by introducing an appropriate new definition of the inverse descent set. Analogous results are proved for the higher-order shuffle conjecture involving $\nabla^m(e_n)$. We also give new versions of some recently proposed combinatorial formulas for $\nabla(C_\alpha)$ and $\nabla(s_{(k,1^{(n-k)})})$, which translate expansions based on the dinv statistic into equivalent expansions based on Haglund's bounce statistic.

Keywords: labeled Dyck paths, parking functions, diagonal harmonics
2000 MSC: 05E05,05E10

---

1. Introduction

F. Bergeron and Garsia [2] introduced the operator ∇ (nabla) as part of a program to study the modified Macdonald polynomials $\tilde{H}_\mu$. The nabla operator is the unique $\mathbb{Q}(q,t)$-linear map on symmetric functions that sends each $\tilde{H}_\mu$ to $q^{n(\mu')}t^{n(\mu)}\tilde{H}_\mu$. Letting e_n denote the elementary symmetric function, Haiman [10] proved that $\nabla(e_n)$ is the Frobenius series of DH_n, the module of diagonal harmonics. More details may be found in [5, 9, 16].

In [6], Haglund, Haiman, Loehr, Remmel, and Ulyanov conjectured an explicit combinatorial formula for the Frobenius series $\nabla(e_n)$ involving a weighted

[1]This work was partially supported by a grant from the Simons Foundation (#244398 to Nicholas Loehr).

sum of labeled Dyck paths; this conjecture is known as the *shuffle conjecture*. This combinatorial formula uses two statistics on labeled Dyck paths called dinv and area, as well as the notion of the inverse descent set (IDes) of a certain reading word. The dinv and area statistics had appeared previously in Haglund, Haiman, and Loehr's conjectured formula for the Hilbert series of DH_n [7].

It turns out that there are four equivalent combinatorial formulas for this Hilbert series. In addition to the formula using area and dinv, Haglund and Loehr [7] gave a formula based on statistics on modified labeled Dyck paths called area$'$ and bounce. They defined a bijection proving the equidistribution of the pair of statistics $(\mathrm{dinv}, \mathrm{area})$ with the pair $(\mathrm{area}', \mathrm{bounce})$. The third formula appears in [13], where Loehr introduced a statistic on parking functions called pmaj computed using a suitable new parking policy. The statistics $(\mathrm{dinv}, \mathrm{area})$ were shown to be equidistributed with $(\mathrm{area}, \mathrm{pmaj})$. These two pairs of statistics were linked via the fourth "fermionic" formula involving a sum over intermediate objects $(\sigma; \vec{u})$, where σ is a permutation of n objects and $\vec{u}$ is a vector of integers satisfying certain conditions determined by σ.

The main goal of this paper is to extend these formulas to provide three new combinatorial reformulations of the shuffle conjecture: one involving area$'$ and bounce on modified labeled Dyck paths, one involving area and pmaj on parking functions, and one involving the intermediate objects $(\sigma; \vec{u})$ appearing in the fermionic formula. In each case, we introduce a new notion of IDes on a suitable reading word and give a bijective proof that the new formulation is equivalent to the original shuffle conjecture. We extend two of these results to give combinatorial reformulations of the higher-order shuffle conjecture for $\nabla^m(e_n)$. We also give new versions of some recently proposed combinatorial formulas for $\nabla(C_\alpha)$ and $\nabla(s_{(k,1^{(n-k)})})$, which translate expansions based on the dinv statistic into equivalent expansions based on Haglund's bounce statistic.

We hope that our new formulations of the shuffle conjecture may lead to further progress in proving the conjecture, since the other statistics have led to progress on related problems. For instance, the pmaj statistic was used in [13] to prove an explicit formula for the $t = 1/q$ specialization of q, t-parking functions, and Armstrong [1] has recently discovered a connection between the statistics area$'$ and bounce and certain hyperplane arrangements.

The paper is organized as follows. §2 reviews the original version of the shuffle conjecture. §3 deduces a version of the shuffle conjecture based on the fermionic formula for the Hilbert series of DH_n. This leads in §4 to a formula based on the statistics pmaj and area for parking functions. §5 describes an expansion based on modified labeled Dyck paths. We use this expansion in §7 to convert dinv-based formulas for $\nabla(C_\alpha)$ (recalled in §6) to equivalent formulas using the bounce statistic for Dyck paths. The last three sections extend some of the preceding results to labeled trapezoidal paths; in particular, we provide two new versions of the higher-order shuffle conjecture for $\nabla^m(e_n)$.

2. Original Formulation of Shuffle Conjecture

We begin by reviewing the combinatorial definitions needed to state the original shuffle conjecture. A *parking function* of order n is a function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $|\{x : f(x) \leq i\}| \geq i$ for $1 \leq i \leq n$. For example, the function $f : \{1, \ldots 6\} \to \{1, \ldots, 6\}$ defined by $f(1) = 5, f(2) = 2, f(3) = 1, f(4) = 2, f(5) = 6, f(6) = 1$ is a parking function. A *Dyck path* of order n is a lattice path from $(0, 0)$ to (n, n) consisting of n unit-length north steps and n unit-length east steps that never go below the line $y = x$. A *labeled Dyck path* is a Dyck path where each north step is labeled with an integer between 1 and n, each used exactly once, so that the labels in each column increase from bottom to top. See Fig. 1 for an example. There is a bijection between labeled Dyck paths of order n and parking functions of order n. This bijection creates a parking function f from a given labeled Dyck path by letting $f(x) = i$ iff the label x appears in column i. For instance, the labeled Dyck path in Fig. 1 corresponds to the parking function f defined above. Let $\mathcal{P}_n$ be the set of labeled Dyck paths of order n. It is well-known [12] that $|\mathcal{P}_n| = (n + 1)^{n-1}$.

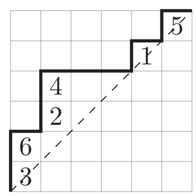


Figure 1: A labeled Dyck path from $(0, 0)$ to $(6, 6)$.

It is often useful to describe a labeled Dyck path P as a pair of vectors $(g(P), p(P))$ where for each i, g_i is the number of complete cells between the path and the diagonal in row i (with the rows numbered 1 to n from bottom to top), and p_i is the label in row i. A pair of vectors (g, p) corresponds to a labeled Dyck path of order n if and only if (a) g and p have length n; (b) $g_1 = 0$; (c) $g_i \geq 0$ for $1 \leq i \leq n$; (d) $g_{i+1} \leq g_i + 1$ for $1 \leq i \leq n - 1$; (e) p is a permutation of $\{1, \ldots, n\}$; and (f) $g_{i+1} = g_i + 1$ implies $p_i < p_{i+1}$. For example, the labeled Dyck path P in Fig. 1 has $g(P) = (0, 1, 1, 2, 0, 0)$ and $p(P) = (3, 6, 2, 4, 1, 5)$.

We now review the relevant statistics on labeled Dyck paths. The *area* of a labeled Dyck path $P = (g, p)$ is

$$\text{area}(P) = \sum_{i=1}^{n} g_i,$$

which is the number of complete lattice squares between the path and the line $y = x$. The *dinv* (diagonal inversion count) of $P = (g, p)$ is defined by

$$\text{dinv}(P) = \sum_{i<j} [\chi(g_i = g_j \text{ and } p_i < p_j) + \chi(g_i = g_j + 1 \text{ and } p_i > p_j)],$$

where, for any logical statement A, $\chi(A)$ is 1 if A is true and 0 if A is false. For example, the path in Fig. 1 has $\text{area}(P) = 4$ and $\text{dinv}(P) = 5$.

The *diagonal reading word* of a labeled Dyck path P, denoted $\text{drw}(P)$, is obtained by scanning down each diagonal in the path, from northeast to southwest, starting with the highest diagonal, and recording the labels in the order they are scanned. Note that $\text{drw}(P) \in S_n$, where S_n is the set of permutations of $\{1, 2, \ldots, n\}$. For any $w = w_1 \cdots w_n \in S_n$, let $\text{Des}(w) = \{i < n : w_i > w_{i+1}\}$ and $\text{IDes}(w) = \text{Des}(w^{-1})$. Note that $\text{IDes}(w)$ consists of all $i < n$ such that $i+1$ appears before i in $w_1 \cdots w_n$. For example, the labeled Dyck path P in Fig. 1 has $\text{drw}(P) = 426513$ and $\text{IDes}(\text{drw}(P)) = \{1, 3, 5\}$.

The shuffle conjecture uses the inverse descent set of the diagonal reading word to index *Gessel's quasisymmetric functions* [4]. Given $n \in \mathbb{N}$ and $D \subseteq \{1, \ldots, n-1\}$, Gessel's quasisymmetric function can be defined by

$$Q_{n,D}(X) = \sum_{\substack{a_1 \leq a_2 \leq \cdots \leq a_n \\ i \in D \Rightarrow a_i < a_{i+1}}} x_{a_1} x_{a_2} \cdots x_{a_n}.$$

Conjecture 1 (Shuffle Conjecture [6]). *For all* $n \geq 1$,

$$\nabla(e_n) = \sum_{P \in \mathcal{P}_n} q^{\text{dinv}(P)} t^{\text{area}(P)} Q_{n,\text{IDes}(\text{drw}(P))}(X).$$

For convenience, we will denote the sum on the right side of the shuffle conjecture by $CF_n(X; q, t)$. On the other hand, we write

$$CH_n(q,t) = \sum_{P \in \mathcal{P}_n} q^{\text{dinv}(P)} t^{\text{area}(P)}$$

to denote the combinatorial formula for the Hilbert series of DH_n conjectured in [7].

3. Fermionic Formulation of Shuffle Conjecture

Our first reformulation of the shuffle conjecture involves the *fermionic formula* expressing $CH_n(q,t)$ as a weighted sum of permutations [7, 13]. For $\sigma \in S_n$, factor $\sigma_1 \cdots \sigma_n$ into ascending runs R_i separated by descents. That is, write $\sigma_1 \cdots \sigma_n = R_k > R_{k-1} > \cdots > R_1 > R_0$ where runs are indexed from right to left, with least index zero. By convention, set $R_{-1} = 0$. Let $1 \leq i \leq n$ and suppose σ_i is in run R_j. Define $w_i(\sigma)$ to be the number of symbols greater than σ_i in R_j plus the number of symbols less than σ_i in R_{j-1}. For example, if $\sigma = 78236154$, then $R_3 = 78$, $R_2 = 236$, $R_1 = 15$, $R_0 = 4$, and $(w_1(\sigma), \ldots, w_8(\sigma)) = (4, 3, 3, 2, 2, 1, 1, 1)$. Define

$$\mathcal{I}_n = \{(\sigma; u_1, \ldots, u_n) : \sigma \in S_n \text{ and } 0 \leq u_i < w_i(\sigma) \text{ for } 1 \leq i \leq n\}.$$

Recall that the *descent set* of a permutation $\sigma \in S_n$ is the set $\text{Des}(\sigma) = \{i < n : \sigma_i > \sigma_{i+1}\}$. The *major index* of σ is $\text{maj}(\sigma) = \sum_{i \in \text{Des}(\sigma)} i$. We define statistics on objects $(\sigma; u) \in \mathcal{I}_n$ as follows. First, let $\mathrm{tstat}(\sigma; u) = \mathrm{maj}(\sigma)$ and $\mathrm{qstat}(\sigma; u) = \sum_{i=1}^n u_i$. Define

$$\begin{aligned} \mathrm{IDes}(\sigma; u) = &\mathrm{IDes}(\sigma) \\ &\cup \{i < n : \text{for some } k,\ \sigma_k = i, \sigma_{k+1} = i+1, \text{ and } u_k > u_{k+1}\}. \end{aligned}$$

For example, $(\sigma; u) = (78236154; 2, 1, 0, 1, 1, 0, 0, 0) \in \mathcal{I}_8$ has $\mathrm{qstat}(\sigma; u) = 5$, $\mathrm{tstat}(\sigma; u) = \mathrm{maj}(78236154) = 14$, and $\mathrm{IDes}(\sigma; u) = \{1, 4, 5, 6, 7\}$.

Setting $[m]_q = 1+q+q^2+\cdots+q^{m-1}$ for all positive integers m, the fermionic formula can be stated as follows.

Theorem 2 ([7, 13])**.** *For all $n \geq 1$,*

$$\sum_{P \in \mathcal{P}_n} q^{\mathrm{dinv}(P)} t^{\mathrm{area}(P)} = \sum_{\sigma \in S_n} t^{\mathrm{maj}(\sigma)} \prod_{k=1}^{n} [w_k(\sigma)]_q.$$

It is immediate from the definitions that

$$\sum_{\sigma \in S_n} t^{\mathrm{maj}(\sigma)} \prod_{k=1}^{n} [w_k(\sigma)]_q = \sum_{(\sigma;u) \in \mathcal{I}_n} t^{\mathrm{tstat}(\sigma;u)} q^{\mathrm{qstat}(\sigma;u)}.$$

Generalizing this formula leads to our first reformulation of the shuffle conjecture:

Theorem 3. *For each positive integer n,*

$$CF_n(X; q, t) = \sum_{(\sigma;u) \in \mathcal{I}_n} q^{\mathrm{qstat}(\sigma;u)} t^{\mathrm{tstat}(\sigma;u)} Q_{n,\mathrm{IDes}(\sigma;u)}(X).$$

We prove this result by analyzing a bijection $\beta : \mathcal{P}_n \to \mathcal{I}_n$ used in [13] to prove Theorem 2. This map has the property that $\mathrm{dinv}(P) = \mathrm{qstat}(\beta(P))$ and $\mathrm{area}(P) = \mathrm{tstat}(\beta(P))$ for all $P \in \mathcal{P}_n$. Given $P \in \mathcal{P}_n$, to find $\beta(P) = (\sigma(P); u(P))$, first write $P = (g, p)$. Let $D = \max_{1 \leq i \leq n} g_i$. For each d from D to 0 (in this order), list the labels p_i such that $g_i = d$ in increasing order. This process defines a permutation σ such that the run R_d consists exactly of the labels p_i such that $g_i = d$. For $1 \leq k \leq n$, define

$$u_k = \sum_{i<j} [\chi(g_i = g_j \text{ and } p_i = \sigma_k < p_j) + \chi(g_i = g_j + 1 \text{ and } p_i = \sigma_k > p_j)].$$

For example, β sends the labeled path P in Fig. 1 to $(\sigma; u)$ where $\sigma = 426135$ and $u = (0, 1, 2, 1, 1, 0)$. Note that $\mathrm{tstat}(\sigma; u) = \mathrm{maj}(426135) = 4 = \mathrm{area}(P)$, $\mathrm{qstat}(\sigma; u) = 5 = \mathrm{dinv}(P)$, and $\mathrm{IDes}(\sigma; u) = \{1, 3, 5\} = \mathrm{IDes}(\mathrm{drw}(P))$.

To compute β^{-1}, we start with $(\sigma; u) \in \mathcal{I}_n$ and reconstruct (g, p) to determine a labeled Dyck path in $\mathcal{P}_n$. First, factor $\sigma = R_s > R_{s-1} > \cdots > R_1 > R_0$ into ascending runs R_i. For $1 \leq k \leq n$, create tiles $\begin{bmatrix} \sigma_k \\ i \end{bmatrix}$ where σ_k is in R_i. For example, given $(\sigma; u) = (78236154; 2, 1, 0, 1, 1, 0, 0, 0) \in \mathcal{I}_8$, the tiles are:

$$\begin{bmatrix} 7 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 8 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 6 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 5 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 4 \\ 0 \end{bmatrix}.$$

Working through the list of tiles from right to left, we insert these tiles into an initially empty sequence. After inserting all the tiles, the top row of the sequence will give $p_1 \cdots p_n$, and the bottom row will give $g_1 \cdots g_n$. For each tile, the number u_k determines a unique insertion position of $\begin{bmatrix} \sigma_k \\ i \end{bmatrix}$ that will cause this tile to contribute u_k to $\mathrm{dinv}(g,p)$. More specifically, there are $w_k(\sigma)$ legal insertion positions for this tile, and the change in dinv caused by inserting this tile in these positions (from right to left) is $0, 1, 2, \ldots, w_k(\sigma) - 1$. In the particular example considered above, the result of inserting the tiles is:

$$\begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 5 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 6 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 7 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 8 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix},$$

which gives $g = (0, 1, 2, 2, 3, 3, 1, 2)$ and $p = (4, 5, 6, 3, 7, 8, 1, 2)$. The resulting labeled Dyck path is drawn in Fig. 2.

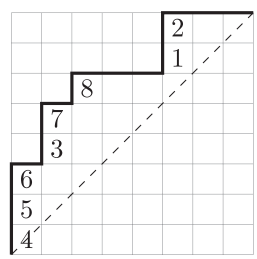


Figure 2: Result of β^{-1}.

Theorem 3 is a consequence of the properties of β and the following lemma.

Lemma 4. *For all* $P \in \mathcal{P}_n$, $\mathrm{IDes}(\mathrm{drw}(P)) = \mathrm{IDes}(\beta(P))$.

Proof. Let $P \in \mathcal{P}_n$, let $\beta(P) = (\sigma; u)$, and suppose $i \in \mathrm{IDes}(\mathrm{drw}(P))$. This can happen in two ways. First, $i + 1$ could be on a higher diagonal than i in P. Then $i+1$ will appear earlier in σ than i, and hence $i \in \mathrm{IDes}(\sigma) \subseteq \mathrm{IDes}(\beta(P))$. Second, $i + 1$ could occur to the northeast of i on the same diagonal of P. In this case, let $i = p_k = \sigma_l$ and $i + 1 = p_j = \sigma_{l+1}$ for some $k < j$ and some l. Note that $g_k = g_j$ since i and $i + 1$ are on the same diagonal. If $g_j = g_m$ and $i + 1 = p_j < p_m$ for some $m > j$, then $g_k = g_m$ and $i = p_k < i + 1 < p_m$. Further, $g_k = g_j$ and $i = p_k < i + 1 = p_j$ where $k < j$. Thus,

$$\sum_{m:\, k<m} \chi(g_k = g_m \text{ and } i = p_k < p_m) > \sum_{m:\, j<m} \chi(g_j = g_m \text{ and } i+1 = p_j < p_m). \tag{3.1}$$

Similarly, if $g_j = g_m + 1$ and $i + 1 = p_j > p_m$ for some $m > j$, then $g_k = g_m + 1$

and $i = p_k > p_m$. So

$$\sum_{m\colon k<m} \chi(g_k = g_m + 1 \text{ and } i = p_k > p_m) \geq \sum_{m\colon j<m} \chi(g_j = g_m + 1 \text{ and } i+1 = p_j > p_m). \quad (3.2)$$

Adding (3.1) and (3.2) shows that $u_l > u_{l+1}$, so $i \in \mathrm{IDes}(\beta(P))$.

On the other hand, suppose $i \in \mathrm{IDes}(\beta(P))$. If $i \in \mathrm{IDes}(\sigma)$, then $i \in R_k$ and $i+1 \in R_j$ for some $j > k$. Then in P, $i+1$ must appear on diagonal j and i on diagonal k. Thus, since $j > k$, $i+1$ appears earlier in $\mathrm{drw}(P)$ than i, and hence $i \in \mathrm{IDes}(\mathrm{drw}(P))$. If instead $i, i+1 \in R_k$, then $i = \sigma_j$ and $i+1 = \sigma_{j+1}$ for some j. Then, since $i \in \mathrm{IDes}(\beta(P))$, $u_j > u_{j+1}$. When constructing P from $(\sigma; u)$, it follows that the tile $\begin{bmatrix} i \\ k \end{bmatrix}$ will be placed to the left of the tile $\begin{bmatrix} i+1 \\ k \end{bmatrix}$, so again $i \in \mathrm{IDes}(\mathrm{drw}(P))$. □

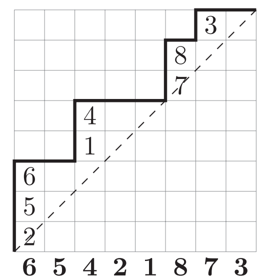
4. Formulation using pmaj and area

Given a parking function f, we recall the definition of $\mathrm{pmaj}(f)$ from [13]. We say that a car i *prefers parking space* j iff $f(i) = j$. For $1 \leq j \leq n$, define $X_j = \{i : f(i) = j\}$. That is, X_j is the set of cars that prefer to park in spot j. In terms of labeled Dyck paths, X_j is the set of labels in column j. Let $T_j = \bigcup_{k=1}^{j} X_k$ be the set of cars that want to park at or before spot j. Then, $|T_j| \geq j$ for $1 \leq j \leq n$ since f is a parking function. We now define the *parking order* used to compute the pmaj statistic. The spots $1, \ldots, n$ will be filled with cars $c_1, \ldots, c_n$ from left to right. The car c_1 that parks in spot 1 is the largest car in the set T_1. The car c_2 that parks in spot 2 is the largest car less than c_1 in the set $T_2 \sim \{c_1\}$, or the largest car in $T_2 \sim \{c_1\}$ if there is no such car. Continuing likewise, the car c_j that parks in spot j is the largest car less than c_{j-1} in $T_j \sim \{c_1, \ldots, c_{j-1}\}$, or the largest car in this set if there is no car smaller than c_{j-1}. We then define a word σ by setting $\sigma_j = c_{n+1-j}$ (i.e., σ is the reverse of $c_1 \ldots c_n$). Then $\mathrm{pmaj}(f) = \mathrm{maj}(\sigma)$. Letting $P = (g,p) \in \mathcal{P}_n$ be the labeled Dyck path representing the parking function f, we set $\mathrm{pmaj}(P) = \mathrm{pmaj}(f)$. For example, the labeled Dyck path P shown in Fig. 3 has parking order 65421873, so $\mathrm{pmaj}(P) = \mathrm{maj}(37812456) = 3$ and $\mathrm{area}(P) = 8$.

We now state our second reformulation of the shuffle conjecture. For $P = (g,p) \in \mathcal{P}_n$, define the *vertical reading word* of P, denoted $\mathrm{vrw}(P)$, to be $p_n p_{n-1} \cdots p_1 \in S_n$. As usual in S_n, we let $\mathrm{IDes}(\mathrm{vrw}(P))$ be the set of $i < n$ with $i+1$ to the left of i in $\mathrm{vrw}(P)$; i.e., $\mathrm{IDes}(\mathrm{vrw}(P))$ is the set of $i < n$ such that $i+1$ is in a higher row of P than i. For the example in Fig. 3, $\mathrm{vrw}(P) = 38741652$ and $\mathrm{IDes}(\mathrm{vrw}(P)) = \{2, 5, 6, 7\}$.

Theorem 5. *For all $n \geq 1$,*

$$CF_n(X; q, t) = \sum_{P \in \mathcal{P}_n} q^{\mathrm{area}(P)} t^{\mathrm{pmaj}(P)} Q_{n, \mathrm{IDes}(\mathrm{vrw}(P))}(X).$$



$\sigma = 37812456$

$u = (1, 1, 0, 2, 3, 0, 1, 0)$

Figure 3: The new parking order and $\gamma(P)$.

To prove this theorem, we need the bijection $\gamma : \mathcal{P}_n \to \mathcal{I}_n$, defined in [13], which has the property that $\operatorname{area}(P) = \operatorname{qstat}(\gamma(P))$ and $\operatorname{pmaj}(P) = \operatorname{tstat}(\gamma(P))$. For $P \in \mathcal{P}_n$ with associated parking function f, we compute $\gamma(P) = (\sigma; u) \in \mathcal{I}_n$ as follows. Let σ be the reverse of the parking order of P. Set $u_k = n + 1 - k - f(\sigma_k)$ for $1 \leq k \leq n$. One can show that $0 \leq u_k < w_k(\sigma)$ using the parking rules. For example, the object P in Fig. 3 maps under γ to $(37812456; 1, 1, 0, 2, 3, 0, 1, 0)$. Observe that this object has $\operatorname{IDes} = \{2, 5, 6, 7\} = \operatorname{IDes}(\operatorname{vrw}(P))$.

To compute $\gamma^{-1}(\sigma; u_1, \ldots, u_n)$, first define $f(\sigma_k) = n + 1 - k - u_k$ for $1 \leq k \leq n$. One can show that f is a parking function; let $\gamma^{-1}(\sigma; u_1, \ldots, u_n)$ be the labeled Dyck path in $\mathcal{P}_n$ representing this parking function. For example, let $(\sigma; u) = (78236154; 2, 1, 0, 1, 1, 0, 0, 0) \in \mathcal{I}_8$. The parking function f and the labeled Dyck path $P = \gamma^{-1}(\sigma; u)$ are shown in Fig. 4. Observe that $\operatorname{IDes}(\operatorname{vrw}(P)) = \{1, 4, 5, 6, 7\} = \operatorname{IDes}(\sigma; u_1, \ldots, u_n)$.

i	$f(i)$
1	3
2	6
3	4
4	1
5	2
6	3
7	6
8	6

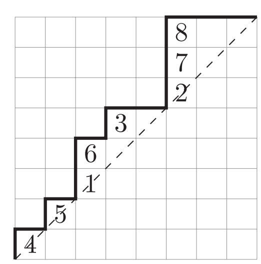


Figure 4: Computation of $\gamma^{-1}(78236154; 2, 1, 0, 1, 1, 0, 0, 0)$.

Theorem 5 follows from properties of the bijection γ, Theorem 3, and the lemma below.

Lemma 6. *For all $n \geq 1$ and all $P \in \mathcal{P}_n$,*

$$\operatorname{IDes}(\gamma(P)) = \operatorname{IDes}(\operatorname{vrw}(P)).$$

Proof. Let $P \in \mathcal{P}_n$, let $\gamma(P) = (\sigma; u)$, and suppose $i \in \mathrm{IDes}(\mathrm{vrw}(P))$. This can happen only if car i prefers a parking spot weakly before the spot preferred by car $i+1$. If car i prefers the same parking spot as car $i+1$ (i.e., i and $i+1$ are in the same column of P), then $i+1$ must park just before i in the parking order. Hence, i and $i+1$ are in the same increasing run in σ, with $i = \sigma_j$ and $i+1 = \sigma_{j+1}$ for some j. Then, since $f(\sigma_j) = f(\sigma_{j+1})$, $u_j = n+1-j-f(\sigma_j) > n+1-(j+1)-f(\sigma_{j+1}) = u_{j+1}$, so $i \in \mathrm{IDes}(\gamma(P))$. If car i prefers an earlier parking spot than $i+1$, then either $i+1$ is parked after i, so $i \in \mathrm{IDes}(\sigma) \subseteq \mathrm{IDes}(\gamma(P))$, or $i+1$ is parked just before i. Once again, in this case, i and $i+1$ are in the same increasing run and $i = \sigma_j$ and $i+1 = \sigma_{j+1}$ for some j. Since $f(\sigma_j) < f(\sigma_{j+1})$, we see as above that $u_j > u_{j+1}$ and $i \in \mathrm{IDes}(\gamma(P))$.

Now, suppose $i \in \mathrm{IDes}(\gamma(P))$. If $i \in \mathrm{IDes}(\sigma)$, then $i+1$ appears earlier than i in σ, which means i parks earlier than $i+1$ in the parking order for P. Now if i starts in a higher row of P than $i+1$, it follows from the parking rules that i must park in a later spot than $i+1$, which is not true here. So $i+1$ is in a higher row of P than i, which means $i \in \mathrm{IDes}(\mathrm{vrw}(P))$ in this case. The other possibility is that i and $i+1$ are in the same increasing run of σ, with $i = \sigma_j$ and $i+1 = \sigma_{j+1}$ for some j and $u_j > u_{j+1}$. In this case, $f(i) = n+1-j-u_j \leq n+1-(j+1)-u_{j+1} = f(i+1)$. Then $i+1$ must appear before i in $\mathrm{vrw}(P)$. Therefore, $\mathrm{IDes}(\gamma(P)) = \mathrm{IDes}(\mathrm{vrw}(P))$. □

5. Formulation using area′ and bounce

In [7] Haglund and Loehr introduce a pair of statistics $(\mathrm{area}', \mathrm{bounce})$ on objects we call *modified labeled Dyck paths*. We review the statistics here as well as the weight-preserving bijection from labeled Dyck paths to modified labeled Dyck paths. To obtain a modified labeled Dyck path from a given Dyck path, label the path with a permutation $w_1 \dots w_n \in S_n$ such that w_i is on the main diagonal in the ith row of the diagram (counting up from the bottom row). The permutation must obey the restriction that for each east step of P followed immediately by a north step, the label w_i due east of the north step must be less than the label w_j due south of the east step. See Fig. 5 for an example. Let the set of all modified labeled Dyck paths of order n be denoted $\mathcal{Q}_n$.

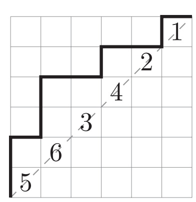


Figure 5: A modified labeled Dyck path.

We now define $\mathrm{bounce}(P)$ for $P \in \mathcal{Q}_n$; this statistic depends only on the underlying Dyck path of P, not the labels. Start by placing a ball in the lower left corner $(0,0)$ of the path. Move the ball up until it touches the beginning of an east step in the path. The ball now "bounces" and heads to the right until it intersects the diagonal, at which point it bounces up again, and the process is repeated until the ball reaches (n,n). This process is illustrated in Fig. 6. Let v_i denote the length of the ith bounce, starting the indexing at $i=0$. Then $\text{bounce}(P) = \sum_{i=0}^{k} iv_i$. Thus, for the object P in Fig. 6, $\text{bounce}(P) = 0\cdot 3 + 1\cdot 3 + 2\cdot 4 + 3\cdot 2 = 17$.

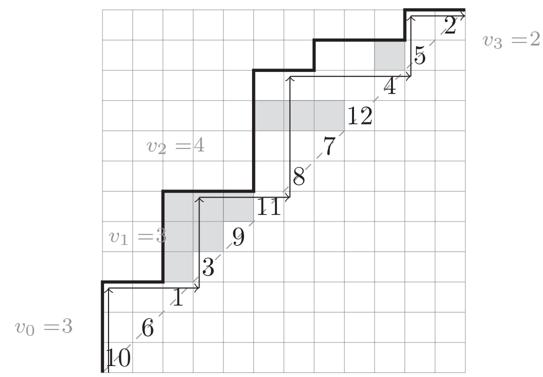


Figure 6: A bounce path (shown by the arrows) and cells contributing to area′ (shaded).

For $P \in \mathcal{Q}_n$, we define $\text{area}'(P)$ to be the number of cells c in the diagram such that c is strictly between the Dyck path and $y=x$, and the label south of c is less than the label east of c. The cells contributing to $\text{area}'(P)$ are shaded in Fig. 5; here, $\text{area}'(P) = 10$.

The reading word for a modified labeled Dyck path $P \in \mathcal{Q}_n$, denoted $\text{rw}(P)$, is the word obtained by reading down the main diagonal starting in the northeast corner. For example, the reading word for the object P in Fig. 6 is $\text{rw}(P) = 2,5,4,12,7,8,11,9,3,1,6,10$ and $\text{IDes}(\text{rw}(P)) = \{1,3,4,6,10,11\}$.

Theorem 7. *For all $n \geq 1$,*

$$CF_n(X;q,t) = \sum_{S\in\mathcal{Q}_n} q^{\text{area}'(S)} t^{\text{bounce}(S)} Q_{n,\text{IDes}(\text{rw}(S))}(X).$$

For the proof, we need the bijection $F : \mathcal{P}_n \to \mathcal{Q}_n$ given in [7], which has the property that $\text{dinv}(P) = \text{area}'(F(P))$ and $\text{area}(P) = \text{bounce}(F(P))$. Given $P \in \mathcal{P}_n$, consider the area vector $g(P) = (g_1, g_2, \ldots, g_n)$. Let $s = \max_i(g_i)$. For each k with $0 \leq k \leq s$, let $w_k = |\{g_i : g_i = k\}|$. We use the w_k's to construct a bounce path that goes north w_0 steps, then east w_0 steps, then north w_1 steps, then east w_1 steps, and so on. The regions above each horizontal bounce and to the left of the subsequent vertical bounce form a sequence of s rectangles, $R_0, R_1, \ldots, R_{s-1}$, as shown in Fig. 7. To determine the portion of the path $F(P)$ in rectangle R_k, use the subword of $g(P)$ obtained by considering only entries k and $k+1$. More specifically, we start at the southwest corner of R_k

and take an east step every time a k occurs in the subword and take a north step every time a $k+1$ occurs in the subword. (The part of the path before R_0 consists of vertical steps from $(0,0)$ to the southwest corner of R_0, and the part of the path after R_{s-1} consists of horizontal steps from the northeast corner of R_{s-1} to (n,n).) This algorithm creates a Dyck path with vertical bounce moves $w_0, w_1, \ldots, w_s$. By definition, the labels of $F(P)$ are obtained by writing the diagonal reading word of P from northeast to southwest along the main diagonal of $F(P)$. We see from this definition that $\mathrm{drw}(P) = \mathrm{rw}(F(P))$. Thus, $\mathrm{IDes}(\mathrm{drw}(P)) = \mathrm{IDes}(\mathrm{rw}(F(P)))$, and Theorem 7 follows.

Figure 7: Example of F; note $g(P) = (0,1,2,1,2,2,1,1)$.

6. Expansion of $\nabla(C_\alpha)$ using dinv

This section reviews some recent progress on the shuffle conjecture involving certain symmetric functions denoted C_α, described in more detail below. In [8], Haglund, Morse, and Zabrocki conjectured combinatorial formulas for $\langle \nabla(C_\alpha), s_{(1^n)} \rangle$ and $\nabla(C_\alpha)$ involving subcollections of Dyck paths determined by the composition α. The formula for the sign character was subsequently proved by Garsia, Hicks, Xin, and Zabrocki [3, 11]. These results imply similar formulas for $\nabla(s_{(a,1^{n-a})})$ which had been conjectured earlier by Loehr and Warrington [15]. All of these identities are based on the statistics dinv and area for labeled and unlabeled Dyck paths. In the next section, we provide new combinatorial versions of these results using the statistics area, area′, and bounce.

First we recall the definition of the symmetric functions C_α. Here, α is a *composition* of a fixed integer n, which is an ordered sequence $(\alpha_1, \alpha_2, \ldots, \alpha_l)$ of positive integers with $\alpha_1 + \alpha_2 + \cdots + \alpha_l = n$. Write $\mathrm{Comp}(n)$ for the set of compositions of n. For each positive integer a, we can define an operator $\mathbb{C}_a$ on symmetric functions by the plethystic formula

$$\mathbb{C}_a(P[X]) = (-1/q)^{a-1} P\left[X - \frac{1-1/q}{z}\right] \Omega[zX]\Big|_{z^a}.$$

In more detail, let ϕ be the unique ring homomorphism on the ring of symmetric functions sending each power-sum p_k to $p_k - (1-q^{-k})/z^k$. Then for any symmetric function P, $\mathbb{C}_a(P)$ is the coefficient of z^a in the formal power series

$$(-1/q)^{a-1}\phi(P)(h_0 + h_1 z + h_2 z^2 + h_3 z^3 + \cdots),$$

where h_k is the complete homogeneous symmetric function. Define

$$C_\alpha = C_\alpha[X;q] = \mathbb{C}_{\alpha_1} \circ \mathbb{C}_{\alpha_2} \circ \cdots \circ C_{\alpha_l}(1),$$

which is a homogeneous symmetric function of degree n with coefficients in $\mathbb{Q}(q)$. For example, one can check that $C_{(n)} = (-1)^{n-1}q^{-(n-1)}s_{(n)}$ and $C_{(a,1)} = (-1)^{a-1}(q^{-a}s_{(a+1)} + q^{-(a-1)}s_{(a,1)})$.

Let $\mathcal{D}_n$ be the set of unlabeled Dyck paths ending at (n,n), which can be identified with the set of vectors $g = (g_1,\ldots,g_n)$ satisfying conditions (b), (c), and (d) in §2. For $g \in \mathcal{D}_n$, define $\operatorname{dinv}(g) = \sum_{i<j} \chi(g_i - g_j \in \{0,1\})$ and $\operatorname{area}(g) = \sum_{i=1}^n g_i$. Let $1 = i_1 < i_2 < \cdots < i_s \leq n$ be the set of all indices i such that $g_i = 0$, and define

$$\operatorname{touch}(g) = (i_2 - i_1, i_3 - i_2, \ldots, i_s - i_{s-1}, n+1-i_s),$$

which is a composition of n. Informally, the indices $i_1, i_2, \ldots$ mark locations where the Dyck path g touches the diagonal $y = x$, and the parts of $\operatorname{touch}(g)$ count the number of rows between consecutive touch points. For example, given

$$g = (0,1,1,2,3,1,2,0,1,2,2,0,1,1,2,3,0,1) \in \mathcal{D}_{18},$$

we have $\operatorname{area}(g) = 23$, $\operatorname{dinv}(g) = 71$, and $\operatorname{touch}(g) = (7,4,5,2)$.

Theorem 8. *[3, 8, 11] For all $n \geq 1$ and all compositions α of n,*

$$\langle \nabla(C_\alpha), s_{(1^n)} \rangle = \sum_{\substack{g \in \mathcal{D}_n: \\ \operatorname{touch}(g)=\alpha}} t^{\operatorname{area}(g)} q^{\operatorname{dinv}(g)}.$$

Conjecture 9. *[8] For all $n \geq 1$ and all compositions α of n,*

$$\nabla(C_\alpha) = \sum_{\substack{P=(g,p) \in \mathcal{P}_n: \\ \operatorname{touch}(g)=\alpha}} t^{\operatorname{area}(P)} q^{\operatorname{dinv}(P)} Q_{n,\operatorname{IDes}(\operatorname{drw}(P))}.$$

Now consider a Schur symmetric function $s_{(k,1^{n-k})}$ indexed by a "hook partition" $(k,1^{n-k})$. Proposition 5.3 of [8] provides the following expansion of this Schur function in terms of the C_α's:

$$s_{(k,1^{n-k})} = (-q)^{k-1} \sum_{\substack{\alpha \in \operatorname{Comp}(n): \\ \alpha_1 \geq k}} C_\alpha. \tag{6.1}$$

For $g \in \mathcal{D}_n$, note that $\alpha = \operatorname{touch}(g)$ satisfies $\alpha_1 \geq k$ iff $g_i > 0$ for $1 < i \leq k$. By this remark and linearity, the previous results for the C_α's yield the following results for Schur hooks.

Theorem 10. *For* $1 \leq k \leq n$*,*

$$(-1)^{k-1}\langle\nabla(s_{(k,1^{n-k})}), s_{(1^n)}\rangle = \sum_{\substack{g\in\mathcal{D}_n:\\ g_i>0 \text{ for } 1<i\leq k}} t^{\mathrm{area}(g)}q^{\mathrm{dinv}(g)+k-1}.$$

Theorem 11. *If Conjecture 9 is true, then for* $1 \leq k \leq n$*,*

$$(-1)^{k-1}\nabla(s_{(k,1^{n-k})}) = \sum_{\substack{P=(g,p)\in\mathcal{P}_n:\\ g_i>0 \text{ for } 1<i\leq k}} t^{\mathrm{area}(P)}q^{\mathrm{dinv}(P)+k-1}Q_{n,\mathrm{IDes}(\mathrm{drw}(P))}.$$

7. Expansion of $\nabla(C_\alpha)$ using bounce

We now derive a new formula for $\langle\nabla(C_\alpha), s_{(1^n)}\rangle$ involving Haglund's bounce statistic. In §5, we defined a set $\mathcal{Q}_n$, a bijection $F : \mathcal{P}_n \to \mathcal{Q}_n$, and statistics area' and bounce on $\mathcal{Q}_n$. By ignoring all the labels, the map F induces a bijection $F : \mathcal{D}_n \to \mathcal{D}_n$ on unlabeled Dyck paths. For $P \in \mathcal{D}_n$, we define $\mathrm{area}(P)$ and $\mathrm{bounce}(P)$ as before (note that the definitions of these statistics do not involve the labels).

To state our new formula, we need to generalize the bounce statistic for Dyck paths as follows. Fix $P \in \mathcal{D}_n$. For each j between 0 and n, we can define a bounce path $B_j(P)$ that starts at (j, j) and moves north and east to (n, n) using the rules explained earlier. Let $v_{0,j}(P), v_{1,j}(P), \ldots$ be the lengths of the vertical moves of this bounce path, and define $b_j(P) = \sum_{i\geq 0} iv_{i,j}(P)$. Note that $b_0(P) = \mathrm{bounce}(P)$ and $v_{0,0}(P)$ is the number of north steps at the beginning of P. For example, consider the Dyck path $P \in \mathcal{D}_{18}$ shown in Fig. 8. (The letters a,b,c,d in that figure will be explained below.) The bounce paths $B_0(P)$, $B_1(P)$, and $B_2(P)$ are shown in the figure. We compute $b_i(P)$ for $0 \leq i \leq 4$ in the following table:

j	$v_{0,j}(P), v_{1,j}(P), \ldots$	$b_j(P)$
0	4,7,5,2	23
1	6,6,4,1	17
2	6,7,2,1	14
3	7,6,2	10
4	7,5,2	9

Note that since $v_{0,0} = 4$, the bounce path starting at $(4, 4)$ is the same as the bounce path starting at $(0, 0)$ with the first bounce removed.

Theorem 12. *For all* $n \geq 1$ *and all compositions* $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ *of* n*,*

$$\langle\nabla(C_\alpha), s_{(1^n)}\rangle = \sum_{\substack{P\in\mathcal{D}_n:\ v_{0,0}(P)=\ell \text{ and}\\ \alpha=(b_0(P)-b_1(P)+1,b_1(P)-b_2(P)+1,\ldots,b_{\ell-1}(P)-b_\ell(P)+1)}} t^{\mathrm{bounce}(P)}q^{\mathrm{area}(P)}.$$

Recall that the bijection $F : \mathcal{D}_n \to \mathcal{D}_n$ has the property that $\mathrm{bounce}(F(g)) = \mathrm{area}(g)$ and $\mathrm{area}(F(g)) = \mathrm{dinv}(g)$ for all $g \in \mathcal{D}_n$. Theorem 12 follows from this observation, Theorem 8, and the next lemma.

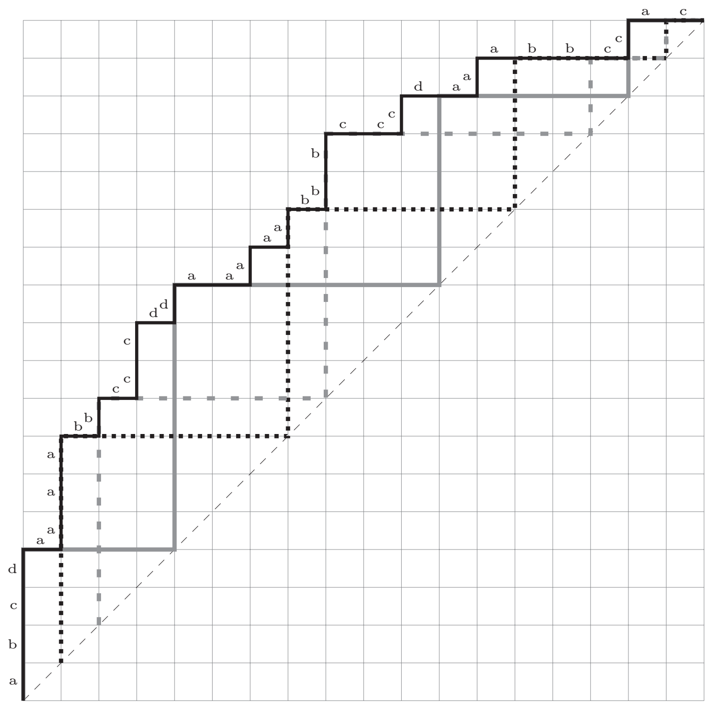


Figure 8: A Dyck path and the bounce paths $B_0(P)$, $B_1(P)$, $B_2(P)$.

Lemma 13. *Given* $g \in \mathcal{D}_n$*, let* $\mathrm{touch}(g) = \alpha = (\alpha_1, \ldots, \alpha_\ell)$ *and* $P = F(g) \in \mathcal{D}_n$*. Then* $\ell = v_{0,0}(P)$ *and*

$$\alpha = (b_0(P) - b_1(P) + 1, b_1(P) - b_2(P) + 1, \ldots, b_{\ell-1}(P) - b_\ell(P) + 1).$$

The proof of the lemma is most easily understood in the context of a specific running example. Let $g \in \mathcal{D}_{18}$ be the Dyck path

$$g = (\underbrace{0,1,1,2,3,1,2}_{\text{`a' block}},\ \underbrace{0,1,2,2}_{\text{`b' block}},\ \underbrace{0,1,1,2,3}_{\text{`c' block}},\ \underbrace{0,1}_{\text{`d' block}}),$$

where we have marked the entries between each successive return to $y = x$ using the letters a,b,c,d. One may check that $F(g)$ is the Dyck path P shown in Fig. 8; the lowercase letter labeling each step records which block of entries in g that step came from. By the calculations given above, we have $\alpha = \mathrm{touch}(g) = (7,4,5,2) = (b_0(P) - b_1(P) + 1, b_1(P) - b_2(P) + 1, b_2(P) - b_3(P) + 1, b_3(P) - b_4(P) + 1)$, illustrating the lemma. Note that $P = F(g)$ begins with 4 north steps, since g contains 4 zeroes. In general, by definition of F, $v_{0,0}(P)$ is the number of zeroes in g, which is also the length ℓ of the composition $\alpha = \mathrm{touch}(g)$.

To explain why $\alpha_i = b_{i-1}(P) - b_i(P) + 1$ in general, we introduce the following notation for $P \in \mathcal{D}_n$. For $0 \leq j \leq n$ and $i \geq -1$, let $v_{-1,j} = j$ and $V_{i,j} =$

$\sum_{k=-1}^{i} v_{i,j}$. Then for $i \geq 0$, the i'th bounce for the bounce path $B_j(P)$ starts on the diagonal at $(V_{i-1,j}, V_{i-1,j})$, moves north until blocked by P at $(V_{i-1,j}, V_{i,j})$, then moves east to the point $(V_{i,j}, V_{i,j})$ on the diagonal. One readily checks that $b_j(P) = \sum_{i \geq 0}(n - V_{i,j})$ for $j \geq 0$. Then $b_{j-1}(P) - b_j(P) + 1 = 1 + \sum_{i \geq 0}(V_{i,j} - V_{i,j-1})$ for $j > 0$.

In our running example, let us explain in detail why $b_0(P) - b_1(P) + 1 = 7 = \alpha_1$. First observe, in Fig. 8, that each northwest corner of $B_1(P)$ is at a vertex of P between an 'a' step coming from the first block of g and a 'b' step coming from the second block of g. To see why this happens, consider the first bounce rectangle for P, which contains the subpath

$$\underbrace{ENNN}_{\text{a}}\underbrace{EN}_{\text{b}}\underbrace{ENN}_{\text{c}}\underbrace{EN}_{\text{d}}. \tag{7.1}$$

Each east step in this rectangle comes from a zero in g. The first east step is followed by three north steps coming from the three 1's in the first block of g; the second east step is followed by one north step coming from the one 1 in the second block of g; and so on. The first north bounce move of $B_0(P)$ is blocked by the first east step in this bounce rectangle. Since $B_1(P)$ starts at $(1,1)$, the first north bounce move of $B_1(P)$ is blocked by the second east step, and thus moves north three steps further compared to $B_0(P)$. In other words, $V_{0,1} - V_{0,0} = 3$ in this example; and in general, $V_{0,1} - V_{0,0}$ will be the number of 1's in the first block of g.

Continuing to the second bounce rectangle, we see the subpath

$$\underbrace{EENEN}_{\text{a}}\underbrace{ENN}_{\text{b}}\underbrace{EEN}_{\text{c}}\underbrace{E}_{\text{d}}. \tag{7.2}$$

Note that the first three east steps in this rectangle came from the three 1's in the first block of g. Since $B_1(P)$ moved north three steps further than $B_0(P)$ in the previous move, $B_1(P)$ will bounce three steps futher east compared to $B_0(P)$ and skip over these three east steps. Thus, the next north move of $B_1(P)$ will be blocked by the fourth east step in the second bounce rectangle, which is the first east step in that rectangle coming from the second block of g. Accordingly, the extra northward distance $V_{1,1} - V_{1,0} = 2$, which is the number of 2's in the first block of g (by definition of F). The same argument applies in general to show that $V_{1,1} - V_{1,0}$ is the number of 2's in the first block of g.

The same analysis applies in the third bounce rectangle and all subsequent bounce rectangles. In general, we find that for $i \geq 0$, $V_{i,1} - V_{i,0}$ is the number of $(i+1)$'s in the first block of g (i.e., between the first two zeroes in g). The first block of g begins with exactly one zero, and hence $b_0(P) - b_1(P) + 1 = 1 + \sum_{i \geq 0}(V_{i,1} - V_{i,0})$ is the number of entries in the first block of g, namely α_1.

A similar argument shows that $b_{j-1}(P) - b_j(P) + 1 = \alpha_j$ for $1 \leq j \leq \ell = v_{0,0}(P)$. For instance, consider $B_1(P)$ and $B_2(P)$ in our running example. The first north bounce of $B_1(P)$ is blocked by the 'b' east step in (7.1) corresponding to the second zero in g, whereas the first north bounce of $B_2(P)$ is blocked by the 'c' east step in (7.2) corresponding to the third zero in g. The number of north steps between these two east steps is the number of 1's in the second block of g (by definition of F) and is also the vertical difference $V_{0,2} - V_{0,1}$. In the second bounce rectangle, $B_2(P)$ skips over the same number of extra east steps (compared to $B_1(P)$) before moving north again. Thus the next east step blocking a north bounce of $B_2(P)$ will correspond to the first 'c' east step in the second bounce rectangle. Since $B_1(P)$ is blocked by the first 'b' east step in this rectangle, we see that $V_{1,2} - V_{1,1}$ is the number of 2's in the second block of g. Continuing similarly, $V_{i,2} - V_{i,1}$ is the number of $(i+1)$'s in the second block of g. Adding over all $i \geq 0$ and remembering that the second block of g contains one initial zero, we see that $b_1(P) - b_2(P) + 1 = \alpha_2$. The argument for larger j is exactly the same, completing the proof of the lemma and Theorem 12.

The above analysis comparing touch(g) to the bounce numbers $b_i(F(g))$ does not change if we add labels to the Dyck paths. Accordingly, we deduce the following result.

Theorem 14. *Fix $n \geq 1$ and a composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ of n. If Conjecture 9 is true, then*

$$\nabla(C_\alpha) = \sum_{\substack{P \in \mathcal{Q}_n:\ v_{0,0}(P)=\ell \text{ and} \\ \alpha=(b_0(P)-b_1(P)+1,\ldots,b_{\ell-1}(P)-b_\ell(P)+1)}} t^{\mathrm{bounce}(P)} q^{\mathrm{area}'(P)} Q_{n,\mathrm{IDes}(\mathrm{rw}(P))}.$$

Using (6.1) and Theorem 12, we also obtain the following formulas for Schur hooks.

Theorem 15. *(a) For $1 \leq k \leq n$,*

$$(-1)^{k-1}\langle \nabla(s_{(k,1^{n-k})}), s_{(1^n)}\rangle = \sum_{\substack{P\in\mathcal{D}_n: \\ b_0(P)-b_1(P)+1\geq k}} t^{\mathrm{bounce}(P)} q^{\mathrm{area}(P)+k-1}.$$

(b) If Conjecture 9 is true, then for $1 \leq k \leq n$,

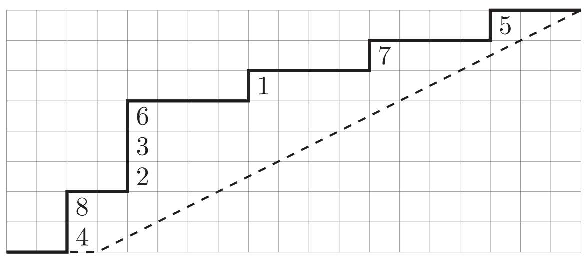
$$(-1)^{k-1}\nabla(s_{(k,1^{n-k})}) = \sum_{\substack{P\in\mathcal{Q}_n: \\ b_0(P)-b_1(P)+1\geq k}} t^{\mathrm{bounce}(P)} q^{\mathrm{area}'(P)+k-1} Q_{n,\mathrm{IDes}(\mathrm{rw}(P))}.$$

8. Trapezoidal Lattice Paths

Our next goal is to generalize the results in §3 and §4 to provide an alternate formulation of the shuffle conjecture for $\nabla^m(e_n)$. At the combinatorial level, we can proceed in somewhat greater generality by considering *labeled trapezoidal lattice paths*. We first recall some relevant definitions from [14].

Let $n, k, m \geq 0$ be integers. Denote the trapezoid connecting the points $(0,0)$,
$(k,0)$,$(0,n)$, and $(mn+k,n)$ by $TZ_{n,k,m}$. A *trapezoidal lattice path* of type (n,k,m) is a lattice path with n vertical steps contained in $TZ_{n,k,m}$. We denote the set of (unlabeled) trapezoidal lattice paths by $\mathcal{T}_{n,k,m}$. A *labeled trapezoidal*

lattice path of type (n, k, m) is a trapezoidal lattice path of type (n, k, m) with the vertical steps labeled by $1, 2, \ldots, n$ (used once each) such that labels increase up columns. Denote the set of labeled trapezoidal lattice paths of type (n, k, m) by $\mathcal{P}_{n,k,m}$. As with labeled Dyck paths, a labeled trapezoidal lattice path P can be described using a pair of vectors $g(P) = (g_1, g_2, \ldots, g_n)$ and $p(P) = (p_1, p_2, \ldots, p_n)$, where g_i is the number of area cells in the ith row from the bottom and p_i is the label of the vertical step in the ith row from the bottom. A pair of vectors (g, p) corresponds to a path in $\mathcal{P}_{n,k,m}$ if and only if (a) $g_1 \in \{0, 1, \ldots, k\}$; (b) $g_i \geq 0$ for $1 \leq i \leq n$; (c) $g_{i+1} \leq g_i + m$ for all $1 \leq i < n$; (d) $p_1, p_2, \ldots, p_n$ is a permutation of $1, 2, \ldots, n$; and (e) for $1 \leq i < n$, if $g_{i+1} = g_i + m$, then $p_i < p_{i+1}$. An example of a labeled trapezoidal lattice path in $\mathcal{P}_{8,3,2}$ and its associated vectors is given in Fig. 9. Note that $\mathcal{P}_{n,0,1}$ is the set of labeled Dyck paths of order n.



$$g(P) = (1, 3, 3, 5, 7, 5, 3, 1)$$
$$p(P) = (4, 8, 2, 3, 6, 1, 7, 5)$$

Figure 9: A labeled trapezoidal lattice path.

The *area* of $P \in \mathcal{P}_{n,k,m}$ is given by

$$\text{area}(P) = \sum_{i=1}^{n} g_i.$$

Thus the area of the path in Fig. 9 is $\text{area}(P) = 28$. The analogue of diagonal inversions in this setting is the statistic

$$\text{dinv}_{m,k}(P) = \sum_{i=1}^{n} (k - g_i)^+ + \sum_{i<j} \sum_{d=0}^{m-1} \chi(A_{i,j,d})$$

where $x^+ = \max\{x, 0\}$, and $A_{i,j,d}$ is the statement

$$(g_i - g_j + d = 0 \text{ and } p_i < p_j) \text{ or}$$
$$(g_i - g_j + d \in \{1, 2, \ldots, m-1\}) \text{ or}$$
$$(g_i - g_j + d = m \text{ and } p_i > p_j)$$

For P in Fig. 9, $\text{dinv}_{2,3}(P) = 14$.

Following [14], we define the *combinatorial Hilbert series* of type (n,k,m) by

$$CH_{n,k,m}(q,t) = \sum_{P\in\mathcal{P}_{n,k,m}} q^{\mathrm{dinv}_{m,k}(P)} t^{\mathrm{area}(P)}.$$

As in the case $m=1$, we modify $CH_{n,k,m}$ to obtain a quasisymmetric polynomial denoted $CF_{n,k,m}$ by introducing reading words. The *diagonal reading word* of $P\in\mathcal{P}_{n,k,m}$, denoted $\mathrm{drw}(P)$, is obtained by reading labels down each diagonal from northeast to southwest, starting on the highest diagonal. More precisely, for each k from $k=\max_i g_i(P)$ down to $k=0$, we scan $p(P)$ from right to left and write all labels p_j with $g_j=k$. For example, the object P in Fig. 9 has $\mathrm{drw}(P)=6,1,3,7,2,8,5,4$ and $\mathrm{IDes}(\mathrm{drw}(P))=\{2,4,5\}$. Define

$$CF_{n,k,m}(X;q,t) = \sum_{P\in\mathcal{P}_{n,k,m}} q^{\mathrm{dinv}_{m,k}(P)} t^{\mathrm{area}(P)} Q_{n,\mathrm{IDes}(\mathrm{drw}(P))}.$$

Conjecture 16 (Higher-Order Shuffle Conjecture [6, §6])**.** *For all $n,m\geq 1$,*

$$\nabla^m(e_n) = CF_{n,0,m}(X;q,t).$$

9. Fermionic Formulation of Higher-Order Shuffle Conjecture

In order to define the objects used in [14] to obtain a fermionic formula for $CH_{n,k,m}$, we start first by considering functions $f:\{1,2,\ldots,n\}\to\{0,1,\ldots,k+m(n-1)\}$, and denote the set of all such functions by $\mathcal{F}_{n,k,m}$. Let $f\in\mathcal{F}_{n,k,m}$ and $T\subseteq\{0,1,\ldots,k+m(n-1)\}$. Define

(i) $f^{-1}(T)=\{x\in\{1,2,\ldots,n\}: f(x)\in T\}$,

(ii) for any $i\in\mathbb{N}$, $f^{-1}_{<i}(T)=\{x: x<i \text{ and } f(x)\in T\}$, and

(iii) for any $i\in\mathbb{N}$, $f^{-1}_{>i}(T)=\{x: x>i \text{ and } f(x)\in T\}$.

When $T=\{a\}$, we write $f^{-1}(T)=f^{-1}(a)$. For each $i\in\{0,1,\ldots,k+m(n-1)\}$, let $W_i=f^{-1}(i)$. Then define the *word* of f as $w(f)=W_0|W_1|\cdots|W_{k+m(n-1)}|$ where the elements of W_i are written in decreasing order, with a bar symbol at the end of each W_i. Note that some W_i might be empty, leading to consecutive bars. We can now define the following statistics for $f\in\mathcal{F}_{n,k,m}$:

(i) $\mathrm{maj}(f)=\sum_{i=1}^{n} f(i)$

(ii) $$\begin{aligned}\mathrm{count}(f,i) = &\chi(f(i)\leq k) + |f^{-1}_{<i}(f(i)-m)| \\ &+ |f^{-1}_{>i}(f(i))| + |f^{-1}(\{f(i)-1,\ldots,f(i)-(m-1)\})|\end{aligned}$$

(iii) $x_0(f)=\sum_{i<j}(m-|f(i)-f(j)|)^+ + \sum_{i=1}^{n}(k-f(i))^+$

(iv) $x_i(f) = -|f^{-1}_{>i}(\{f(i), f(i)-1, \ldots, f(i)-(m-1)\})|$

For example, given $f \in \mathcal{F}_{8,2,3}$ defined by

$$f(1) = 3, f(2) = 0, f(3) = 2, f(4) = 2, f(5) = 3, f(6) = 0, f(7) = 2, f(8) = 4,$$

the word of f is

$$w(f) = 62||743|51|8||||||||||||||||||||.$$

Note that f can be recovered from $w(f)$. Further,

$$\begin{array}{ll}
\mathrm{maj}(f) = 16 & x_0(f) = 44 \\
\mathrm{count}(f,1) = 4 & x_1(f) = -4 \\
\mathrm{count}(f,2) = 2 & x_2(f) = -1 \\
\mathrm{count}(f,3) = 5 & x_3(f) = -3 \\
\mathrm{count}(f,4) = 4 & x_4(f) = -2 \\
\mathrm{count}(f,5) = 4 & x_5(f) = -1 \\
\mathrm{count}(f,6) = 1 & x_6(f) = 0 \\
\mathrm{count}(f,7) = 3 & x_7(f) = 0 \\
\mathrm{count}(f,8) = 5 & x_8(f) = 0.
\end{array}$$

Define

$$\begin{aligned}
R_i(f) &= -x_i(f) = |f^{-1}_{>i}(\{f(i), f(i)-1, \ldots, f(i)-(m-1)\})|, \\
L_i(f) &= x_i(f) + \mathrm{count}(f,i) - 1 \\
&= \chi(f(i) \leq k) - 1 + |f^{-1}_{<i}(\{f(i)-1, \ldots, f(i)-m\})|, \\
\mathcal{I}_{n,k,m} &= \{I = (f;u) : f \in \mathcal{F}_{n,k,m} \text{ and } -R_i(f) \leq u_i \leq L_i(f) \text{ for } 1 \leq i \leq n\}.
\end{aligned}$$

Define the *inverse descent set* for $I \in \mathcal{I}_{n,k,m}$, where $I = (f;u)$ and $w(f) = W_0|W_1|\cdots|W_{k+m(n-1)}|$, by

$$\begin{aligned}
\mathrm{IDes}(f;u) = \{i : i \in W_j \text{ and } i+1 \in W_l \text{ for some } l > j\} \\
\cup \{i : i, i+1 \in W_j \text{ and } u_i \geq u_{i+1}\}.
\end{aligned}$$

Continuing the previous example, consider $I \in \mathcal{I}_{8,2,3}$ defined by
$I = (f; -2, 0, -1, -1, 0, 0, 1, 2)$. Then $\mathrm{IDes}(f;u) = \{2, 3, 4, 6, 7\}$.

Theorem 17 ([14])**.** *For all $n \geq 1, k \geq 0, m \geq 1$,*

$$CH_{n,k,m} = \sum_{f \in \mathcal{F}_{n,k,m}} t^{\mathrm{maj}(f)} q^{x_0(f)} \prod_{i=1}^{n} \sum_{p=-R_i(f)}^{L_i(f)} q^p.$$

For $I = (f;u) \in \mathcal{I}_{n,k,m}$ define $\mathrm{tstat}(I) = \mathrm{maj}(f)$ and $\mathrm{qstat}(I) = x_0(f) + \sum_{i=1}^n u_i$. It follows from the definitions that

$$\sum_{f \in \mathcal{F}_{n,k,m}} t^{\mathrm{maj}(f)} q^{x_0(f)} \prod_{i=1}^{n} \sum_{p=-R_i(f)}^{L_i(f)} q^p = \sum_{I \in \mathcal{I}_{n,k,m}} q^{\mathrm{qstat}(I)} t^{\mathrm{tstat}(I)}.$$

The next theorem gives our fermionic reformulation of the higher-order shuffle conjecture.

Theorem 18. *For all $n \geq 1, k \geq 0, m \geq 1$,*

$$CF_{n,k,m} = \sum_{I \in \mathcal{I}_{n,k,m}} q^{\text{qstat}(I)} t^{\text{tstat}(I)} Q_{n,\text{IDes}(I)}.$$

We will prove Theorem 18 by analyzing the bijection $F : \mathcal{P}_{n,k,m} \to \mathcal{I}_{n,k,m}$ in [14]. The bijection F has the property that given $P \in \mathcal{P}_{n,k,m}$, $\text{area}(P) = \text{tstat}(F(P))$ and $\text{dinv}_{m,k}(P) = \text{qstat}(F(P))$. Let $P \in \mathcal{P}_{n,k,m}$. Find $F(P) = (f, u)$ by first setting $f(p_i) = g_i$ for $1 \leq i \leq n$. To find $u = (u_1, u_2, \ldots, u_n)$, we think of P as a list of tiles of the form

$$\begin{bmatrix} g_1 \\ p_1 \end{bmatrix}, \begin{bmatrix} g_2 \\ p_2 \end{bmatrix}, \ldots, \begin{bmatrix} g_n \\ p_n \end{bmatrix}.$$

Let $w_1 w_2 \cdots w_n$ be $w(f)$ with the vertical bars removed. To obtain u, we remove tiles from the above list in the order

$$\begin{bmatrix} f(w_n) \\ w_n \end{bmatrix}, \begin{bmatrix} f(w_{n-1}) \\ w_{n-1} \end{bmatrix}, \ldots, \begin{bmatrix} f(w_1) \\ w_1 \end{bmatrix}.$$

Just before removing each tile $\begin{bmatrix} f(w_j) \\ w_j \end{bmatrix}$, compute u_{w_j} by first noting that $w_j = p_i$ for some i. Then

$$\begin{aligned} u_{w_j} = u_{p_i} = & \sum_{t:\, i<t} \chi(g_i - g_t \in \{1, \ldots, m\} \text{ and } p_i > p_t) \\ & - \sum_{t:\, t<i} \chi(g_t - g_i \in \{0, -1, \ldots, -(m-1)\} \text{ and } p_t > p_i) \end{aligned}$$

For example, let $n = 8, k = 2, m = 3$, and let $P = (g, p)$ be defined by $g = (1, 4, 3, 5, 8, 5, 2, 3)$ and $p = (2, 8, 1, 5, 6, 7, 4, 3)$. Writing P as a list of tiles gives

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} \begin{bmatrix} 8 \\ 6 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

Note that the function f is defined by

$$\begin{aligned} &f(1) = 3, \quad f(2) = 1, \quad f(3) = 3, \quad f(4) = 2, \\ &f(5) = 5, \quad f(6) = 8, \quad f(7) = 5, \quad f(8) = 4. \end{aligned}$$

Then the word of f without bars is 24318756. Table 1 shows the result of removing each tile in the prescribed order and the computed value of u_i for each i. Thus $F(P) = (f; -1, 0, -1, 0, 1, 0, 1, 3)$. Note that $\text{dinv}_{3,2}(P) = 31 = \text{tstat}(f; u)$ and $\text{area}(P) = 31 = \text{qstat}(f; u)$.

To compute $F^{-1}(f; u)$ given $(f; u) \in \mathcal{I}_{n,k,m}$, first construct tiles of the form $\begin{bmatrix} f(i) \\ i \end{bmatrix}$ for each i, $1 \leq i \leq n$. Let $w_1 w_2 \cdots w_n$ be the word of f with bars removed. Starting with an empty list, insert the tiles in the order

$$\begin{bmatrix} f(w_1) \\ w_1 \end{bmatrix}, \begin{bmatrix} f(w_2) \\ w_2 \end{bmatrix}, \ldots, \begin{bmatrix} f(w_n) \\ w_n \end{bmatrix}.$$

$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}4\\8\end{bmatrix} \begin{bmatrix}3\\1\end{bmatrix} \begin{bmatrix}5\\5\end{bmatrix} \begin{bmatrix}\mathbf{8}\\\mathbf{6}\end{bmatrix} \begin{bmatrix}5\\7\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\3\end{bmatrix}$	$u_6 = 0$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}4\\8\end{bmatrix} \begin{bmatrix}3\\1\end{bmatrix} \begin{bmatrix}\mathbf{5}\\\mathbf{5}\end{bmatrix} \begin{bmatrix}5\\7\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\3\end{bmatrix}$	$u_5 = 1$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}4\\8\end{bmatrix} \begin{bmatrix}3\\1\end{bmatrix} \begin{bmatrix}\mathbf{5}\\\mathbf{7}\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\3\end{bmatrix}$	$u_7 = 1$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}\mathbf{4}\\\mathbf{8}\end{bmatrix} \begin{bmatrix}3\\1\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\3\end{bmatrix}$	$u_8 = 3$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}\mathbf{3}\\\mathbf{1}\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}3\\3\end{bmatrix}$	$u_1 = -1$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}2\\4\end{bmatrix} \begin{bmatrix}\mathbf{3}\\\mathbf{3}\end{bmatrix}$	$u_3 = -1$
$\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}\mathbf{2}\\\mathbf{4}\end{bmatrix}$	$u_4 = 0$
$\begin{bmatrix}\mathbf{1}\\\mathbf{2}\end{bmatrix}$	$u_2 = 0$

Table 1: Removing tiles and computing u.

Insert each tile, one at a time, so that when tile $\begin{bmatrix} f(w_i) \\ w_i \end{bmatrix}$ is inserted and the new partial word is

$$\begin{bmatrix} g_1 \\ p_1 \end{bmatrix}, \ldots, \begin{bmatrix} g_i \\ p_i \end{bmatrix}$$

with $w_i = p_j$ for some j, $1 \le j \le i$,

$$u_{w_i} = \sum_{t\colon j<t} \chi(g_j - g_t \in \{1, \ldots, m\} \text{ and } p_j > p_t) - \sum_{t\colon t<j} \chi(g_t - g_j \in \{0, -1, \ldots, -(m-1)\} \text{ and } p_t > p_j). \quad (9.1)$$

It is proved in [14] that there is a unique insertion position for this tile satisfying (9.1). More specifically, the valid insertion positions for $\begin{bmatrix} f(w_i) \\ w_i \end{bmatrix}$, scanned from right to left, cause the right side of (9.1) to assume the values $-R_{w_i}(f)$, $-R_{w_i}(f)+1, \ldots, L_{w_i}(f)$ in this order.

For example, given $(f; -2, 0, -1, -1, 0, 0, 1, 2) \in \mathcal{I}_{8,2,3}$ with

$$f(1) = 3, f(2) = 0, f(3) = 2, f(4) = 2, f(5) = 3, f(6) = 0, f(7) = 2, f(8) = 4,$$

the word of f with bars removed is 62743518. This yields a set of tiles to be inserted in the order

$$\begin{bmatrix} 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \end{bmatrix}.$$

The result of inserting each tile in the prescribed order is

$$\begin{bmatrix} 0 \\ 2 \end{bmatrix} \begin{bmatrix} 2 \\ 7 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} 4 \\ 8 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 0 \\ 6 \end{bmatrix}.$$

Thus the corresponding labeled trapezoidal lattice path in vector form is $P = (g, p)$ where $g = (0, 2, 2, 3, 4, 3, 2, 0)$ and $p = (2, 7, 3, 1, 8, 5, 4, 6)$. The path can be seen in Fig. 10.

Lemma 19. *Let* $P \in \mathcal{P}_{n,k,m}$. *Then* $\mathrm{IDes}(P) = \mathrm{IDes}(F(P))$.

Proof. Let $P \in \mathcal{P}_{n,k,m}$. Suppose $i \in \mathrm{IDes}(P)$. Recall that this means $i+1$ appears earlier than i in the diagonal reading word of P. Thus $i+1$ either occurs on a higher diagonal than i or $i+1$ and i are on the same diagonal with $i+1$ occuring to the right (and hence also above) i in the labeled trapezoidal lattice path P. We first consider the case where $i+1$ is on a higher diagonal than i. Suppose i is on diagonal j and $i+1$ is on diagonal l with $l > j$. Let $F(P) = (f; u)$. In $w(f)$, $i \in W_j$ and $i{+}1 \in W_l$ with $l > j$. Thus, $i \in \mathrm{IDes}(F(P))$.

Suppose next that i and $i+1$ are both on diagonal d with $i+1$ appearing to the right of i. By definition of the word of f, the tile $\begin{bmatrix} f(i) \\ i \end{bmatrix}$ is removed prior

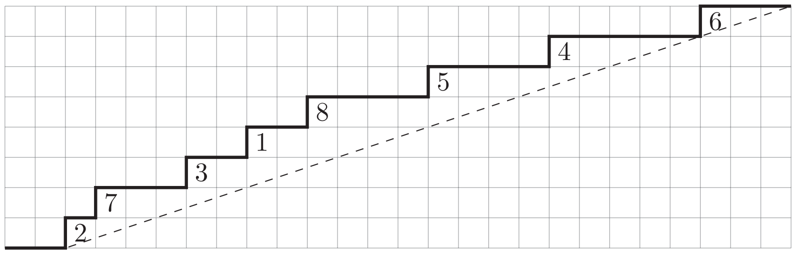


Figure 10: The path corresponding to $(f;u)$.

to the tile $\begin{bmatrix} f(i+1) \\ i+1 \end{bmatrix}$. Prior to the removal of $\begin{bmatrix} f(i) \\ i \end{bmatrix}$, the (partial) list of tiles will be

$$\cdots \begin{bmatrix} f(i) \\ i \end{bmatrix} \cdots \begin{bmatrix} f(j) \\ j \end{bmatrix} \cdots \begin{bmatrix} f(i+1) \\ i+1 \end{bmatrix} \cdots .$$

Note that any tiles to the left of $\begin{bmatrix} f(i) \\ i \end{bmatrix}$ will contribute the same amount to u_i and u_{i+1} since $f(i) = f(i+1)$. Similarly, any tiles to the right of $\begin{bmatrix} f(i+1) \\ i+1 \end{bmatrix}$ will contribute the same amount to u_i and u_{i+1}. Consider the tiles of the form $\begin{bmatrix} f(j) \\ j \end{bmatrix}$ between $\begin{bmatrix} f(i) \\ i \end{bmatrix}$ and $\begin{bmatrix} f(i+1) \\ i+1 \end{bmatrix}$. If $\begin{bmatrix} f(j) \\ j \end{bmatrix}$ contributes to u_i, then $f(i) - f(j) \in \{1, \ldots, m\}$ and $i > j$. Thus, $f(j) - f(i+1) \in \{-1, \ldots, -m\}$ and $i+1 > j$. Thus, $\begin{bmatrix} f(j) \\ j \end{bmatrix}$ cannot contribute to u_{i+1} if it also contributes to u_i. Similarly, if $\begin{bmatrix} f(j) \\ j \end{bmatrix}$ contributes to u_{i+1} (note that this contribution is negative), then $f(j) - f(i+1) \in \{0, -1, \ldots, -(m-1)\}$ and $j > i+1$. It follows that $f(i) - f(j) \in \{0, 1, \ldots, m-1\}$ and $j > i$, so $\begin{bmatrix} f(j) \\ j \end{bmatrix}$ does not also contribute to u_i. Therefore, $u_i \geq u_{i+1}$ and hence $i \in \mathrm{IDes}(F(P))$.

A similar argument shows that if $i \in \mathrm{IDes}(F(P))$, then $i \in \mathrm{IDes}(P)$, and therefore $\mathrm{IDes}(P) = \mathrm{IDes}(F(P))$. □

Theorem 18 follows immediately from the bijection F and Lemma 19.

10. Formulation using $\mathbf{pmaj}_{m,k}$ and area

This section gives one more formulation of the higher-order shuffle conjecture obtained by analyzing a bijection $G : \mathcal{I}_{n,k,m} \to \mathcal{P}_{n,k,m}$ defined in [14]. This bijection has the property that for all $I \in \mathcal{I}_{n,k,m}$, $\mathrm{area}(G(I)) = \mathrm{qstat}(I)$ and $\mathrm{pmaj}_{m,k}(G(I)) = \mathrm{tstat}(I)$ (this equation is used to define $\mathrm{pmaj}_{m,k}$).

As in the special case $k = 0$, $m = 1$ considered earlier, we define the *vertical reading word* of $P = (g, p) \in \mathcal{P}_{n,k,m}$ to be the permutation $\mathrm{vrw}(P) = p_n \cdots p_2 p_1$

obtained by reading the labels in P from top to bottom. Then $\mathrm{IDes}(\mathrm{vrw}(P))$ is the set of $i < n$ such that $i+1$ appears in a higher row than i in P. For instance, the labeled path P in Fig. 9 has $\mathrm{vrw}(P) = 57163284$ and $\mathrm{IDes}(\mathrm{vrw}(P)) = \{2, 4, 6\}$.

We can also identify labeled trapezoidal paths P with certain *generalized parking functions* π. Specifically, we map $P \in \mathcal{P}_{n,k,m}$ to the function $\pi : \{1, 2, \ldots, n\} \to \mathbb{N}$ such that $\pi(x) = i$ for all labels x in column i of P (where the leftmost column is column 1). For example, the labeled path P in Fig. 9 maps to the function π given by

$$\begin{array}{cccc} \pi(1) = 9, & \pi(2) = 5, & \pi(3) = 5, & \pi(4) = 3, \\ \pi(5) = 17, & \pi(6) = 5, & \pi(7) = 13, & \pi(8) = 3. \end{array}$$

Since labels in lower-numbered columns appear in P below labels in higher-numbered columns, and since labels in a given column increase reading up, it follows that for all $i < n$, $i \in \mathrm{IDes}(\mathrm{vrw}(P))$ iff $\pi(i) \leq \pi(i+1)$. Setting $\mathrm{IDes}(\pi) = \{i < n : \pi(i) \leq \pi(i+1)\}$, we therefore have $\mathrm{IDes}(\mathrm{vrw}(P)) = \mathrm{IDes}(\pi)$.

Theorem 20. *For all $n \geq 1$, $k \geq 0$, $m \geq 1$,*

$$CF_{n,k,m} = \sum_{P \in \mathcal{P}_{n,k,m}} q^{\mathrm{area}(P)} t^{\mathrm{pmaj}_{m,k}(P)} Q_{n,\mathrm{IDes}(\mathrm{vrw}(P))}.$$

To prove this, we must recall the definition of the bijection $G : \mathcal{I}_{n,k,m} \to \mathcal{P}_{n,k,m}$ from [14, §5.4]. Given $I = (f; u_1, \ldots, u_n) \in \mathcal{I}_{n,k,m}$, set $W_j = f^{-1}(j)$, $v_j = |W_j|$, $h_j = \chi(j < k) + \sum_{i=0}^{m-1} v_{j-i}$, and $H_j = \sum_{i=0}^{j} h_i$ for all $j \geq 0$; also let $H_{-1} = 0$. For $1 \leq i \leq n$, set $\pi(i) = H_{f(i)-1} + 1 - u_i$. Then $G(I)$ is defined to be the labeled trapezoidal path P associated to the generalized parking function π. It is proved in [14] that G is a bijection with $\mathrm{qstat}(I) = \mathrm{area}(G(I))$ and $\mathrm{tstat}(I) = \mathrm{pmaj}_{m,k}(G(I))$. Thus, Theorem 20 follows from Theorem 18 and the next lemma.

Lemma 21. *For all $I = (f; u_1, \ldots, u_n) \in \mathcal{I}_{n,k,m}$, $\mathrm{IDes}(I) = \mathrm{IDes}(\mathrm{vrw}(G(I)))$.*

Proof. Define π as above. By our earlier remarks relating $\mathrm{IDes}(\mathrm{vrw}(P))$ to $\mathrm{IDes}(\pi)$, we need only prove that for all $i < n$, $i \in \mathrm{IDes}(I)$ iff $\pi(i) \leq \pi(i+1)$. Fix $i < n$, and consider three cases.

Case 1: for some j, i and $i+1$ both belong to W_j, which means $f(i) = j = f(i+1)$. In this case, $\pi(i) \leq \pi(i+1)$ iff $H_{j-1} + 1 - u_i \leq H_{j-1} + 1 - u_{i+1}$, which holds iff $u_i \geq u_{i+1}$, which holds iff $i \in \mathrm{IDes}(I)$ by definition.

Case 2: for some $j < l$, $i \in W_j$ and $i+1 \in W_l$, which means $f(i) = j < l = f(i+1)$. Here $i \in \mathrm{IDes}(I)$, so we must show $\pi(i) \leq \pi(i+1)$, or equivalently $\pi(i+1) - \pi(i) \geq 0$. We compute

$$\begin{aligned} \pi(i+1) - \pi(i) &= (H_{l-1} + 1 - u_{i+1}) - (H_{j-1} + 1 - u_i) \\ &= h_j + \cdots + h_{l-1} + (u_i - u_{i-1}) \\ &\geq h_j + \cdots + h_{l-1} - R_i(f) - L_{i+1}(f) \\ &= h_j + \cdots + h_{l-1} - |f^{-1}_{>i}(\{j, j-1, \ldots, j-(m-1)\})| \\ &+ (1 - \chi(l \leq k)) - |f^{-1}_{<i+1}(\{l-1, \ldots, l-m\})|. \end{aligned}$$

If $j < l-1$, we can discard nonnegative terms to see that

$$\pi(i+1) - \pi(i) \geq [h_j - |f_{>i}^{-1}(\{j, j-1, \ldots, j-(m-1)\})|] + [h_{l-1} - |f_{<i+1}^{-1}(\{l-1, \ldots, l-m\})|]. \quad (10.1)$$

Now, $h_j \geq v_j + \cdots + v_{j-(m-1)} = |f^{-1}(\{j, j-1, \ldots, j-(m-1)\})| \geq |f_{>i}^{-1}(\{j, j-1, \ldots, j-(m-1)\})|$, so the first bracketed term in (10.1) is nonnegative. The second term is nonnegative for similar reasons, so $\pi(i+1) - \pi(i) \geq 0$ when $j < l-1$. On the other hand, if $j = l-1$, our original calculation becomes

$$\pi(i+1) - \pi(i) \geq [1 - \chi(l \leq k)] + [h_j - |f_{>i}^{-1}(\{j, \ldots, j-(m-1)\})| - |f_{<i+1}^{-1}(\{j, \ldots, j-(m-1)\})|].$$

Note $|f_{>i}^{-1}(\{j, \ldots, j-(m-1)\})| + |f_{<i+1}^{-1}(\{j, \ldots, j-(m-1)\})| = |f^{-1}(\{j, \ldots, j-(m-1)\})| = v_j + \cdots + v_{j-(m-1)} \leq h_j$, and $1 - \chi(l \leq k) \geq 0$. So $\pi(i+1) - \pi(i) \geq 0$ when $j = l-1$.

Case 3: for some $j < l$, $i+1 \in W_j$ and $i \in W_l$, which means $f(i) = l > j = f(i+1)$. Here $i \notin \mathrm{IDes}(I)$, so we must show $\pi(i) > \pi(i+1)$, or equivalently $\pi(i) - \pi(i+1) > 0$. Calculating as in case 2, we find

$$\begin{aligned} \pi(i) - \pi(i+1) &= (H_{l-1} + 1 - u_i) - (H_{j-1} + 1 - u_{i+1}) \\ &= h_j + \cdots + h_{l-1} + (u_{i+1} - u_i) \\ &\geq h_j + \cdots + h_{l-1} - R_{i+1}(f) - L_i(f) \\ &> h_j + \cdots + h_{l-1} - \chi(l \leq k) \\ &\quad - |f_{>i+1}^{-1}(\{j, j-1, \ldots, j-(m-1)\})| \\ &\quad - |f_{<i}^{-1}(\{l-1, \ldots, l-m\})|. \end{aligned}$$

If $j < l-1$, we can discard nonnegative terms to get

$$\pi(i) - \pi(i+1) > [h_j - |f_{>i+1}^{-1}(\{j, \ldots, j-(m-1)\})|] + [h_{l-1} - \chi(l-1 < k) - |f_{<i}^{-1}(\{l-1, \ldots, l-m\})|].$$

As in case 2, we see that each bracketed term is nonnegative (for the second term, note $h_{l-1} - \chi(l-1 < k) = v_{l-1} + \cdots + v_{l-m}$). Thus $\pi(i) - \pi(i+1) > 0$ when $j < l-1$. On the other hand, if $j = l-1$, we find that

$$\pi(i) - \pi(i+1) > h_j - \chi(j < k) - |f_{>i+1}^{-1}(\{j, \ldots, j-(m-1)\})| - |f_{<i}^{-1}(\{j, \ldots, j-(m-1)\})|.$$

Since $h_j - \chi(j < k) = v_j + \cdots + v_{j-(m-1)} = |f^{-1}(\{j, \ldots, j-(m-1)\})|$, it follows that $\pi(i) - \pi(i+1) > 0$ when $j = l-1$. $\square$

Remark 22. It is natural to ask if there is a formulation of the higher-order shuffle conjecture using statistics generalizing area′ and bounce (see §5). We leave this as an open problem, noting that it is not evident how to define the analogue of the set $\mathcal{Q}_n$ when $m > 1$.

[1] D. Armstrong. Hyperplane arrangements and diagonal harmonics. *Discrete Math. Theor. Comp. Sci. Proc.*, AO:39–50, 2011.

[2] F. Bergeron and A. Garsia. Science fiction and Macdonald polynomials. *CRM Proceedings and Lecture Notes AMS VI*, 3:363–429, 1999.

[3] A. Garsia, G. Xin, and M. Zabrocki. Hall-Littlewood operators in the theory of parking functions and diagonal harmonics. *Int. Math. Res. Notices*, article ID rnr060 (36 pages), 2011.

[4] I. Gessel. Multipartite p-partitions and inner products of skew Schur functions. *Combinatorics and Algebra, Contemp. Math.*, 34:289–317, 1984.

[5] J. Haglund. *The q,t-Catalan numbers and the space of diagonal harmonics, with an appendix on the combinatorics of Macdonald polynomials.* AMS University Lecture Series, 2008.

[6] J. Haglund, M. Haiman, N. Loehr, J. B. Remmel, and A. Ulyanov. A combinatorial formula for the character of the diagonal coinvariants. *Duke Math. J.*, 126:195–232, 2005.

[7] J. Haglund and N. Loehr. A conjectured combinatorial formula for the Hilbert series for diagonal harmonics. *Discrete Math.*, 298:189–204, 2005.

[8] J. Haglund, J. Morse, and M. Zabrocki. A compositional shuffle conjecture specifying touch points of the Dyck path. *Canad. J. Math.*, 64:822–844, 2012.

[9] M. Haiman. Macdonald polynomials and geometry. In *New perspectives in algebraic combinatorics*, volume 38 of *Math. Res. Sci. Inst. Publ.*, pages 207–254. Cambridge University Press, 1999.

[10] M. Haiman. Vanishing theorems and character formulas for the Hilbert scheme of points in the plane. *Invent. Math.*, 149:371–407, 2002.

[11] A. Hicks. Two parking function bijections: a sharpening of the q,t-Catalan and Schröder theorems. *Int. Math. Res. Notices*, article ID rnr132 (25 pages), 2011.

[12] A. Konheim and B. Weiss. An occupancy discipline and applications. *SIAM J. Appl. Math.*, 14:1266–1274, 1966.

[13] N. Loehr. Combinatorics of q,t-parking functions. *Adv. in Appl. Math.*, 34:408–425, 2005.

[14] N. Loehr and J. Remmel. Conjectured combinatorial models for the Hilbert series of generalized diagonal harmonics modules. *Electron. J. Combin.*, 11:research paper R68; 64 pages (electronic), 2004.

[15] N. Loehr and G. Warrington. Nested quantum Dyck paths and $\nabla(s_\lambda)$. *Int. Math. Res. Notices*, article ID rnm157 (29 pages), 2008.

[16] I. Macdonald. *Symmetric functions and Hall polynomials.* Oxford University Press, second edition, 1995.