Yan Zhuang

**A lifting of the Goulden–Jackson cluster method to the Malvenuto–Reutenauer algebra**
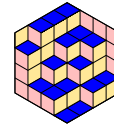
# A lifting of the Goulden–Jackson cluster method to the Malvenuto–Reutenauer algebra

Yan Zhuang

ABSTRACT The Goulden–Jackson cluster method is a powerful tool for counting words by occurrences of prescribed subwords, and was adapted by Elizalde and Noy for counting permutations by occurrences of prescribed consecutive patterns. In this paper, we lift the cluster method for permutations to the Malvenuto–Reutenauer algebra. Upon applying standard homomorphisms, our result specializes to both the cluster method for permutations as well as a $q$-analogue which keeps track of the inversion number statistic. We construct additional homomorphisms using the theory of shuffle-compatibility, leading to further specializations which keep track of various "inverse statistics", including the inverse descent number, inverse peak number, and inverse left peak number. This approach is then used to derive formulas for counting permutations by occurrences of two families of consecutive patterns—monotone patterns and transpositional patterns—refined by these statistics.

## 1. INTRODUCTION

Let $\mathfrak{S}_n$ denote the symmetric group of permutations on the set $[n] \coloneqq \{1, 2, \ldots, n\}$ (where $\mathfrak{S}_0$ consists of the empty permutation), and let $\mathfrak{S} \coloneqq \bigsqcup_{n=0}^{\infty} \mathfrak{S}_n$. We write permutations in one-line notation—that is, $\pi = \pi_1 \pi_2 \cdots \pi_n$—and the $\pi_i$ are called *letters* of $\pi$. The *length* of $\pi$ is the number of letters in $\pi$, so that $\pi$ has length $n$ whenever $\pi \in \mathfrak{S}_n$.

For a sequence of distinct integers $w$, the *standardization* of $w$—denoted $\mathrm{std}(w)$—is defined to be the permutation in $\mathfrak{S}$ obtained by replacing the smallest letter of $w$ with 1, the second smallest with 2, and so on. As an example, we have $\mathrm{std}(73184) = 42153$. Given permutations $\pi \in \mathfrak{S}_n$ and $\sigma \in \mathfrak{S}_m$, we say that $\pi$ *contains* $\sigma$ (as a *consecutive pattern*) if $\mathrm{std}(\pi_i \pi_{i+1} \cdots \pi_{i+m-1}) = \sigma$ for some $i \in [n - m + 1]$, and in this case we call $\pi_i \pi_{i+1} \cdots \pi_{i+m-1}$ an *occurrence* of $\sigma$ (as a consecutive pattern) in $\pi$. For instance, the permutation $315497628$ has three occurrences of the consecutive pattern $213$, namely $315$, $549$, and $628$. On the other hand, $137258469$ has no occurrences of $213$.

Let $\mathrm{occ}_\sigma(\pi)$ denote the number of occurrences of $\sigma$ in $\pi$. If $\mathrm{occ}_\sigma(\pi) = 0$, then we say that $\pi$ *avoids* $\sigma$ (as a consecutive pattern). If $\Gamma \subseteq \mathfrak{S}$, then we let $\mathfrak{S}_n(\Gamma)$ denote the subset of permutations in $\mathfrak{S}_n$ avoiding every permutation in $\Gamma$ as a consecutive pattern. When $\Gamma$ consists of a single permutation $\sigma$, we shall simply write $\mathfrak{S}_n(\sigma)$ as opposed to $\mathfrak{S}_n(\{\sigma\})$. (We use the same convention for other notations involving a set $\Gamma$ of permutations when $\Gamma$ is a singleton.) As observed earlier, we have $137258469 \in \mathfrak{S}_9(213)$.

For the rest of this paper, the notions of occurrence and avoidance of patterns in permutations always refer to consecutive patterns unless otherwise stated.

The study of consecutive patterns in permutations, initiated by Elizalde and Noy [9] in 2003, extends the study of classical patterns in permutations originating in the work of Simion and Schmidt [29]. Consecutive patterns in permutations are analogous to consecutive subwords in words, where repetition of letters is allowed. In the latter realm, the cluster method of Goulden and Jackson [19] provides a very general formula expressing the generating function for words by occurrences of prescribed subwords in terms of a "cluster generating function", which is easier to compute. By setting the variable keeping track of occurrences to zero, this yields a powerful approach for counting words avoiding a prescribed set of subwords. In 2012, Elizalde and Noy [10] adapted the Goulden–Jackson cluster method to the setting of permutations, which they used to obtain differential equations satisfied by $\omega_\sigma(s, x) = (\sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\sigma(\pi)} x^n / n!)^{-1}$ for various families of consecutive patterns $\sigma$, including "monotone patterns", "chain patterns", and "non-overlapping patterns". Solving these differential equations for $\omega_\sigma(s, x)$ then allows one to count permutations by the number of occurrences of $\sigma$.

Over the past decade, Elizalde and Noy's adaptation of the cluster method for permutations has become a standard tool in the study of consecutive patterns; see [2, 3, 6, 7, 8, 22] for a selection of references. One recent development is a $q$-analogue of the cluster method for permutations which also keeps track of the inversion number statistic. This $q$-cluster method, due to Elizalde, was first mentioned in his survey [8] on consecutive patterns, and was applied to monotone patterns and non-overlapping patterns by Crane, DeSalvo, and Elizalde [3] in their study of the Mallows distribution.

To explain the philosophy which guides our work, let us briefly discuss a paper by Josuat-Vergès, Novelli, and Thibon [21], in which the authors study alternating permutations (and their analogues in other Coxeter groups) from the perspective of combinatorial Hopf algebras. Their starting point is André's [1] famous exponential generating function $\sec x + \tan x$ for the number of alternating permutations. The authors note that André's formula has a natural lifting in the Malvenuto–Reutenauer algebra **FQSym**, a Hopf algebra whose basis elements correspond to permutations and whose multiplication encodes "shifted concatenation" of permutations. They then recover André's formula by applying a certain homomorphism $\phi$ to its lifting in **FQSym**, and in their words:

> "Such a proof is not only illuminating, it says much more than the original statement. For example, one can now replace $\phi$ by more complicated morphisms, and obtain generating functions for various statistics on alternating permutations."

A similar approach to permutation enumeration was taken in a series of papers by Gessel and the present author [14, 16, 35, 36], but instead utilizing homomorphisms on noncommutative symmetric functions.

The main result of this present paper is an analogous lifting of the Goulden–Jackson cluster method for permutations to the Malvenuto–Reutenauer algebra. Since the basis elements of the Malvenuto–Reutenauer algebra correspond to permutations, our cluster method in **FQSym** is in a sense the most general cluster method possible for permutations. By applying the same homomorphism $\phi$ used by Josuat-Vergès–Novelli–Thibon to our generalized cluster method, we can recover Elizalde and Noy's cluster method for permutations, and we can use another homomorphism to recover Elizalde's $q$-analogue. We also construct other homomorphisms which lead to new specializations of our cluster method that can be used to count permutations by occurrences of prescribed patterns while keeping track of other permutation statistics.

## 1.1. PERMUTATION STATISTICS.

The permutation statistics that we shall consider are the "inverses" of several classical permutation statistics related to descents and peaks: the *descent number* des, the *major index* maj, the *comajor index* comaj, the *peak number* pk, and the *left peak number* lpk. We define these statistics below.

- We call $i \in [n-1]$ a *descent* of $\pi \in \mathfrak{S}_n$ if $\pi_i > \pi_{i+1}$. Then $\mathrm{des}(\pi)$ is defined to be the number of descents of $\pi$, and $\mathrm{maj}(\pi)$ the sum of all descents of $\pi$. In other words, if $\mathrm{Des}(\pi) := \{\, i \in [n-1] : i \text{ is a descent of } \pi \,\}$—that is, $\mathrm{Des}(\pi)$ is the *descent set* of $\pi$—then

$$\mathrm{des}(\pi) := |\mathrm{Des}(\pi)| \quad \text{and} \quad \mathrm{maj}(\pi) := \sum_{i \in \mathrm{Des}(\pi)} i.$$

  The comajor index comaj is a variant of the major index maj, and is defined by

$$\mathrm{comaj}(\pi) := \sum_{i \in \mathrm{Des}(\pi)} (n-i) = n\mathrm{des}(\pi) - \mathrm{maj}(\pi). \tag{1}$$

- We call $i \in \{2, 3, \ldots, n-1\}$ a *peak* of $\pi \in \mathfrak{S}_n$ if $\pi_{i-1} < \pi_i > \pi_{i+1}$. Then $\mathrm{pk}(\pi)$ is defined to be the number of peaks of $\pi$.
- We call $i \in [n-1]$ a *left peak* of $\pi \in \mathfrak{S}_n$ if $i$ is a peak of $\pi$, or if $i = 1$ and $i$ is a descent of $\pi$. Then $\mathrm{lpk}(\pi)$ is defined to be the number of left peaks of $\pi$.[1]

For example, if $\pi = 72163584$, then we have $\mathrm{Des}(\pi) = \{1, 2, 4, 7\}$, $\mathrm{des}(\pi) = 4$, $\mathrm{maj}(\pi) = 14$, $\mathrm{comaj}(\pi) = 18$, $\mathrm{pk}(\pi) = 2$, and $\mathrm{lpk}(\pi) = 3$. We note that, in the language of consecutive patterns, descents correspond to occurrences of 21 and peaks correspond to occurrences of 132 and 231.

Given a permutation statistic st, we define its *inverse statistic* ist by $\mathrm{ist}(\pi) := \mathrm{st}(\pi^{-1})$. Continuing with the example from above, the inverse of $\pi$ is $\pi^{-1} = 32586417$, so we have $\mathrm{iDes}(\pi) = \{1, 4, 5, 6\}$, $\mathrm{ides}(\pi) = 4$, $\mathrm{imaj}(\pi) = 16$, $\mathrm{icomaj}(\pi) = 16$, $\mathrm{ipk}(\pi) = 1$, and $\mathrm{ilpk}(\pi) = 2$. While st and ist are obviously equidistributed over $\mathfrak{S}_n$, it is worth studying the joint distribution of ist and other permutation statistics over $\mathfrak{S}_n$, or the distribution of ist over restricted sets of permutations (such as pattern avoidance classes). For instance, Garsia and Gessel [11] studied the joint distribution of des, ides, maj, and imaj over $\mathfrak{S}_n$.

Let $\Gamma$ be a set of consecutive patterns and $\mathrm{occ}_\Gamma(\pi)$ the number of occurrences in $\pi$ of patterns in $\Gamma$. In this paper, we will consider the polynomials

$$A_{\Gamma,n}^{(\mathrm{ides},\mathrm{imaj})}(s,t,q) := \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{imaj}(\pi)},$$

$$A_{\Gamma,n}^{(\mathrm{ides},\mathrm{icomaj})}(s,t,q) := \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)},$$

$$A_{\Gamma,n}^{\mathrm{ides}}(s,t) := \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ides}(\pi)+1},$$

$$P_{\Gamma,n}^{\mathrm{ipk}}(s,t) := \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ipk}(\pi)+1}, \text{ and}$$

$$P_{\Gamma,n}^{\mathrm{ilpk}}(s,t) := \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ilpk}(\pi)}$$

where $n \geqslant 1$, and with each of these polynomials defined to be 1 when $n = 0$. These polynomials give the joint distribution of the occurrence statistic $\mathrm{occ}_\Gamma$ along with each of the statistics $(\mathrm{ides}, \mathrm{imaj})$, $(\mathrm{ides}, \mathrm{icomaj})$, ides, ipk, and ilpk. Setting $s = 0$ in any

---

[1]Equivalently, $\mathrm{lpk}(\pi)$ is the number of peaks of the permutation $0\pi$ obtained by prepending 0 to $\pi$.

of these polynomials then gives the distribution of the corresponding statistic over the pattern avoidance class $\mathfrak{S}_n(\Gamma)$. For convenience, let us define $A_{\Gamma,n}^{(\mathrm{ides},\mathrm{imaj})}(t,q) :=$ $A_{\Gamma,n}^{(\mathrm{ides},\mathrm{imaj})}(0,t,q)$ and the polynomials $A_{\Gamma,n}^{(\mathrm{ides},\mathrm{icomaj})}(t,q)$, $A_{\Gamma,n}^{\mathrm{ides}}(t)$, $P_{\Gamma,n}^{\mathrm{ipk}}(t)$, and $P_{\Gamma,n}^{\mathrm{ilpk}}(t)$ analogously.

The reason why we consider the statistics $(\mathrm{ides},\mathrm{imaj})$, $(\mathrm{ides},\mathrm{icomaj})$, ides, ipk, and ilpk is because they are inverses of "shuffle-compatible" statistics. Roughly speaking, a permutation statistic st is shuffle-compatible if the distribution of st over the set of shuffles of two permutations $\pi$ and $\sigma$ depends only on $\mathrm{st}(\pi)$, $\mathrm{st}(\sigma)$, and the lengths of $\pi$ and $\sigma$. (See Section 2.3 for precise definitions.) If st is shuffle-compatible and is a coarsening of the descent set, then st induces a quotient of the algebra QSym of quasisymmetric functions, denoted $\mathcal{A}_{\mathrm{st}}$. By composing the quotient map from QSym to $\mathcal{A}_{\mathrm{st}}$ with the canonical surjection from **FQSym** to QSym, we obtain a homomorphism on **FQSym** which can be used to count permutations by the corresponding inverse statistic. Applying these homomorphisms to our generalized cluster method in **FQSym** yields specializations that refine by the statistics $(\mathrm{ides},\mathrm{icomaj})$, ides, ipk, and ilpk.[2]

1.2. OUTLINE. The structure of this paper is as follows. Section 2 is devoted to background material. We first give a brief expository account of the Goulden–Jackson cluster method, both for words and for permutations. Then, we define quasisymmetric functions and the Malvenuto–Reutenauer algebra, and review some basic symmetries on permutations (reversal, complementation, and reverse-complementation) which will play a role in our work.

The focus of Section 3 is on our main result, the cluster method in Malvenuto–Reutenauer. We prove our generalized cluster method and show how it specializes to Elizalde and Noy's cluster method for permutations as well as its $q$-analogue. In this section, we also use the theory of shuffle-compatibility to construct homomorphisms which we then use to obtain further specializations of our generalized cluster method for the statistics $(\mathrm{ides},\mathrm{icomaj})$, ides, ipk, and ilpk.

In Sections 4 and 5, we apply our general results from Section 3 to produce formulas for the polynomials $A_{\sigma,n}^{\mathrm{ides}}(s,t)$, $P_{\sigma,n}^{\mathrm{ipk}}(s,t)$, and $P_{\sigma,n}^{\mathrm{ilpk}}(s,t)$—and their $s=0$ evaluations—where $\sigma$ is a specific type of consecutive pattern. Section 4 focuses on *monotone patterns*, i.e. the patterns $12\cdots m$ and $m\cdots 21$. Section 5 focuses on the patterns $12\cdots(a-1)(a+1)a(a+2)(a+3)\cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$; these patterns were considered in [10] as a subfamily of "chain patterns", and here we call them *transpositional patterns* because $12\cdots(a-1)(a+1)a(a+2)(a+3)\cdots m$ is precisely the elementary transposition $(a, a+1)$. Most of our formulas involve the Hadamard product operation on formal power series, although some "Hadamard product-free" formulas are obtained for monotone patterns. In the case of monotone patterns, we also give a formula for counting $12\cdots m$-avoiding permutations by inverse descent number and inverse major index.

We conclude this paper in Section 6 with a brief discussion of ongoing work and future directions of research. See [38] for an extended abstract summarizing the results of this paper, as well as [34] for proofs of two observations (Claims 4.6 and 4.9) which are left unproven here.

## 2. Preliminaries

2.1. THE CLUSTER METHOD FOR WORDS. We first introduce the Goulden–Jackson cluster method for words, which we will use to prove our lifting of the cluster method

---

[2]We do not explicitly give a specialization for $(\mathrm{ides},\mathrm{imaj})$, but one can be obtained using the one for $(\mathrm{ides},\mathrm{icomaj})$ and the formula $\mathrm{imaj}(\pi) = n\,\mathrm{ides}(\pi) - \mathrm{icomaj}(\pi)$, which is equivalent to (1).

for permutations to the Malvenuto–Reutenauer algebra. The exposition in this section follows that in [37].

For a finite or countably infinite set $A$, let $A^*$ be the set of all finite sequences of elements of $A$, including the empty sequence. We call $A$ an *alphabet*, the elements of $A$ *letters*, and the elements of $A^*$ *words*. The *length* $|w|$ of a word $w \in A^*$ is the number of letters in $w$. For $v, w \in A^*$, we say that $v$ is a *subword* of $w$ if $w = uvu'$ for some $u, u' \in A^*$, and in this case we also say $w$ *contains* $v$ and that $v$ is an *occurrence* of $w$. The *total algebra* of $A^*$ over $\mathbb{Q}$, denoted $\mathbb{Q}\langle\langle A^*\rangle\rangle$, is the $\mathbb{Q}$-algebra of formal sums of words in $A^*$ where multiplication is the concatenation product.

Given a word $w = w_1 w_2 \cdots w_n \in A^*$ and a set $B \subseteq A^*$, we say that $(i, v)$ is a *marked occurrence* of $v \in B$ in $w$ if

$$v = w_i w_{i+1} \cdots w_{i+|v|-1},$$

that is, $v$ is a subword of $w$ starting at position $i$. Moreover, we say that $(w, T)$ is a *marked word* on $w$ (with respect to $B$) if $w \in A^*$ and $T$ is a set of some marked occurrences in $w$ of words in $B$.

To illustrate, suppose that $A = \{a, b, c\}$ and $B = \{cab, bc\}$. Then

$$(cabcabbca, \{(1, cab), (3, bc), (7, bc)\}) \tag{2}$$

is a marked word on $w = cabcabbca$ with respect to $B$. Informally, we will display a marked word $(w, T)$ as the word $w$ with the marked occurrences in $T$ circled, so that (2) is displayed as

$$c\,a\,b\,c\,a\,b\,b\,c\,a\ .$$

We define the concatenation of two marked words in the obvious way. For example, (2) can be obtained by concatenating $(cabca, \{(1, cab), (3, bc)\})$ and $(bbca, \{(2, bc)\})$, i.e.

$$c\,a\,b\,c\,a \quad \text{and} \quad b\,b\,c\,a\ .$$

A marked word is called a *cluster* if it is not a concatenation of two nonempty marked words. (In particular, we will call a cluster with respect to $B$ a *$B$-cluster*.) So, (2) is not a cluster, but

$$b\,c\,a\,b\,c\,a\,b$$

is a cluster.

For a word $w \in A^*$, let $\mathrm{occ}_B(w)$ be the number of occurrences in $w$ of words in $B$ and let $C_{B,w}$ be the set of all $B$-clusters on $w$. If $c$ is a $B$-cluster, then we let $\mathrm{mk}_B(c)$ be the number of marked occurrences in $c$. Define

$$F_B(s) := \sum_{w \in A^*} w\, s^{\mathrm{occ}_B(w)} \quad \text{and} \quad R_B(s) := \sum_{w \in A^*} w \sum_{c \in C_{B,w}} s^{\mathrm{mk}_B(c)},$$

so that $F_B(s)$ is the generating function for words in $A^*$ by the number of occurrences of words in $B$, and $R_B(s)$ is the generating function for $B$-clusters by the number of marked occurrences. Both $F_B(s)$ and $R_B(s)$ are elements of the formal power series algebra $\mathbb{Q}\langle\langle A^*\rangle\rangle[[s]]$, so the variable $s$ commutes with letters in $A$ (but the letters in $A$ do not commute with each other).

THEOREM 2.1 (Cluster method for words). *Let $A$ be an alphabet and let $B \subseteq A^*$ be a set of words, each of length at least 2. Then*

$$F_B(s) = \left(1 - \sum_{a \in A} a - R_B(s-1)\right)^{-1}.$$

This is a noncommutative version of the original cluster method of Goulden and Jackson, but the proofs are essentially the same; see, e.g. [37, Theorem 1] for details.

2.2. The cluster method for permutations. Next, we describe Elizalde and Noy's [10] adaptation of the cluster method for permutations, as well as its *q*-analogue which refines by the inversion number. The terms *marked occurrence*, *marked permutation*, *concatenation*, and *cluster* are defined for permutations in the analogous way as for words, but with the notion of word containment replaced by permutation containment (in the sense of consecutive patterns). It is worth pointing out that, unlike concatenation of marked words, concatenation of marked permutations is not unique. For instance, both

$$3\,\overparen{2\ 1\ 4}\,6\ 7\,\overparen{8}\,5\ 9 \quad \text{and} \quad 7\,\overparen{5\ 1\ 8}\,3\ 4\,\overparen{6}\,2\ 9$$

are concatenations of

$$3\,\overparen{2\ 1\ 4} \quad \text{and} \quad \overparen{2\ 3\,\overparen{4}\,1\ 5}.$$

However, this does not make a difference in defining clusters for permutations or in adapting the cluster method to the setting of permutations.

Let $\Gamma \subseteq \mathfrak{S}$. Recall that $\mathrm{occ}_\Gamma(\pi)$ is the number of occurrences in $\pi$ of patterns in $\Gamma$, and let $C_{\Gamma,\pi}$ be the set of all $\Gamma$-clusters on $\pi$. If $c$ is a $\Gamma$-cluster, let $\mathrm{mk}_\Gamma(c)$ be the number of marked occurrences in $c$. Define

$$F_\Gamma(s,x) := \sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} \frac{x^n}{n!} \quad \text{and}$$

$$R_\Gamma(s,x) := \sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)} \frac{x^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} r_{\Gamma,n,k} s^k \frac{x^n}{n!}$$

where $r_{\Gamma,n,k}$ is the number of $\Gamma$-clusters of length $n$ with $k$ marked occurrences.

THEOREM 2.2 (Cluster method for permutations). *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$F_\Gamma(s,x) = (1 - x - R_\Gamma(s-1,x))^{-1}.$$

Elizalde and Noy give Theorem 2.2 in the special case where $\Gamma$ consists of a single pattern [10, Theorem 1.1], but in Section 3 we will recover this more general result from our cluster method in the Malvenuto–Reutenauer algebra.

The *n*th *q-factorial* $[n]_q!$ is defined by

$$[n]_q! := (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1})$$

for $n \geqslant 1$ and $[0]_q! := 1$. Later, we will also need the *q-binomial coefficient* defined by

$$\binom{n}{k}_q := \frac{[n]_q!}{[k]_q!\,[n-k]_q!}$$

for all $n \geqslant 0$ and $0 \leqslant k \leqslant n$.

We say that $(i,j) \in [n]^2$ is an *inversion* of $\pi \in \mathfrak{S}_n$ if $i < j$ and $\pi_i > \pi_j$. Let $\mathrm{inv}(\pi)$ denote the number of inversions of $\pi$. Define

$$F_\Gamma(s,q,x) := \sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} s^{\mathrm{occ}_\Gamma(\pi)} q^{\mathrm{inv}(\pi)} \frac{x^n}{[n]_q!} \quad \text{and}$$

$$R_\Gamma(s,q,x) := \sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} q^{\mathrm{inv}(\pi)} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)} \frac{x^n}{[n]_q!} = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} r_{\Gamma,n,k,j} q^j s^k \frac{x^n}{[n]_q!}$$

where $r_{\Gamma,n,k,j}$ is the number of $\Gamma$-clusters of length $n$ with $k$ marked occurrences and whose underlying permutation has $j$ inversions. The next result is [3, Theorem 2.3], but for a set $\Gamma$ of patterns rather than a single pattern $\sigma$.

THEOREM 2.3 ($q$-Cluster method for permutations). *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$F_\Gamma(s,q,x) = (1 - x - R_\Gamma(s-1,q,x))^{-1}.$$

See [27, Corollary 1] for a related result. Like with Theorem 2.2, we will later recover Theorem 2.3 as a specialization of our generalized cluster method.

Let us give one more definition before continuing. Given $\sigma \in \mathfrak{S}_m$, let

$$O_\sigma := \{\, i \in [m-1] : \mathrm{std}(\sigma_{i+1}\sigma_{i+2}\cdots\sigma_m) = \mathrm{std}(\sigma_1\sigma_2\cdots\sigma_{m-i}) \,\}$$

be the *overlap set* of $\sigma$. The notion of overlap set is useful for characterizing $\Gamma$-clusters where $\Gamma$ consists of a single pattern $\sigma$, and we will do this in Sections 4.1 and 5.1.

2.3. QUASISYMMETRIC FUNCTIONS AND SHUFFLE-COMPATIBILITY. A permutation in $\mathfrak{S}_n$ can be characterized as a word in $[n]^*$ of length $n$ consisting of distinct letters. Let $\mathbb{P}$ be the set of positive integers, and let $\mathfrak{P}_n$ denote the set of words in $\mathbb{P}^*$ of length $n$ consisting of distinct letters—not necessarily from 1 to $n$. Also, let $\mathfrak{P} := \bigsqcup_{n=0}^{\infty} \mathfrak{P}_n$. In this section only, we will use the term "permutation" to refer more generally to elements of $\mathfrak{P}$. Observe that any statistic st defined on permutations in $\mathfrak{S}$ can be extended to $\mathfrak{P}$ by letting $\mathrm{st}(\pi) := \mathrm{st}(\mathrm{std}(\pi))$ for $\pi \in \mathfrak{P}$.

Every permutation in $\mathfrak{P}$ can be uniquely decomposed into a sequence of maximal increasing consecutive subsequences, which we call *increasing runs*. Equivalently, an increasing run of $\pi$ is a maximal consecutive subsequence containing no descents. The *descent composition* of $\pi$, denoted $\mathrm{Comp}(\pi)$, is the composition whose parts are the lengths of the increasing runs of $\pi$ in the order that they appear. For instance, the increasing runs of $\pi = 85712643$ are 8, 57, 126, 4, and 3, so the descent composition of $\pi$ is $\mathrm{Comp}(\pi) = (1,2,3,1,1)$. We use the notations $L \vDash n$ and $|L| = n$ to indicate that $L$ is a composition of $n$, so that $L \vDash n$ and $|L| = n$ whenever $L$ is the descent composition of a permutation in $\mathfrak{P}_n$. For a composition $L = (L_1, L_2, \ldots, L_k)$, let $\mathrm{Des}(L) := \{L_1, L_1 + L_2, \ldots, L_1 + \cdots + L_{k-1}\}$. It is easy to see that if $L$ is the descent composition of $\pi$, then $\mathrm{Des}(L)$ is the descent set of $\pi$.

If $\pi \in \mathfrak{P}_m$ and $\sigma \in \mathfrak{P}_n$ are *disjoint*—that is, if they have no letters in common—then we call $\tau \in \mathfrak{P}_{m+n}$ a *shuffle* of $\pi$ and $\sigma$ if both $\pi$ and $\sigma$ are subsequences of $\tau$. The set of shuffles of $\pi$ and $\sigma$ is denoted $S(\pi, \sigma)$. For example, we have

$$S(31, 25) = \{3125, 3215, 3251, 2315, 2351, 2531\}.$$

Let $x_1, x_2, \ldots$ be commuting variables. A formal power series $f \in \mathbb{Q}[[x_1, x_2, \ldots]]$ of bounded degree is called a *quasisymmetric function* if for any positive integers $a_1, a_2, \ldots, a_k$, if $i_1 < i_2 < \cdots < i_k$ and $j_1 < j_2 < \cdots < j_k$ then

$$[x_{i_1}^{a_1} x_{i_2}^{a_2} \cdots x_{i_k}^{a_k}] f = [x_{j_1}^{a_1} x_{j_2}^{a_2} \cdots x_{j_k}^{a_k}] f.$$

Let $\mathrm{QSym}_n$ denote the set of quasisymmetric functions homogeneous of degree $n$. As a vector space, $\mathrm{QSym}_n$ has as a basis the *fundamental quasisymmetric functions* $\{F_L\}_{L \vDash n}$ defined by

$$F_L := \sum_{\substack{i_1 \leqslant i_2 \leqslant \cdots \leqslant i_n \\ i_j < i_{j+1} \text{ if } j \in \mathrm{Des}(L)}} x_{i_1} x_{i_2} \cdots x_{i_n}.$$

If $L \vDash m$ and $K \vDash n$, then

$$F_L F_K = \sum_{\tau \in S(\pi,\sigma)} F_{\mathrm{Comp}(\tau)} \tag{3}$$

where $\pi$ and $\sigma$ are any disjoint permutations satisfying $\mathrm{Comp}(\pi) = L$ and $\mathrm{Comp}(\sigma) = K$. Hence,

$$\mathrm{QSym} \coloneqq \bigoplus_{n=0}^{\infty} \mathrm{QSym}_n$$

is a graded subalgebra of $\mathbb{Q}[[x_1, x_2, \dots]]$; this is the *algebra of quasisymmetric functions* (over $\mathbb{Q}$).

Motivated by Stanley's theory of $P$-partitions, quasisymmetric functions were first defined and studied by Gessel [13] and are now ubiquitous in algebraic combinatorics. References on quasisymmetric functions include [32, Section 7.19], [20, Section 5], and [23].

Let us now return to the product formula (3) for fundamental quasisymmetric functions. In order for (3) to make sense, the multiset $\{\,\mathrm{Comp}(\tau) : \tau \in S(\pi, \sigma)\,\}$ must only depend on the descent compositions of $\pi$ and $\sigma$, or equivalently, $\{\,\mathrm{Des}(\tau) : \tau \in S(\pi, \sigma)\,\}$ only depends on $\mathrm{Des}(\pi)$, $\mathrm{Des}(\sigma)$, and the lengths of $\pi$ and $\sigma$. More generally, a permutation statistic st is called *shuffle-compatible* if for any disjoint permutations $\pi$ and $\sigma$, the multiset $\{\,\mathrm{st}(\tau) : \tau \in S(\pi, \sigma)\,\}$ depends only on $\mathrm{st}(\pi)$, $\mathrm{st}(\sigma)$, and the lengths of $\pi$ and $\sigma$. Therefore, the descent set Des is a shuffle-compatible permutation statistic.

In [17], Gessel and the present author develop a theory of shuffle-compatibility for *descent statistics*: statistics st such that $\mathrm{Comp}(\pi) = \mathrm{Comp}(\sigma)$ implies $\mathrm{st}(\pi) = \mathrm{st}(\sigma)$. The statistics des, maj, comaj, pk, and lpk are all examples of shuffle-compatible descent statistics. If st is a descent statistic and if $L$ is a composition, then we let $\mathrm{st}(L)$ denote the value of st on any permutation with descent composition $L$. Two compositions $L$ and $K$ are called st-*equivalent* if $\mathrm{st}(L) = \mathrm{st}(K)$ and $|L| = |K|$. The following is Theorem 4.3 of [17], and provides a necessary and sufficient condition for a descent statistic to be shuffle-compatible.

THEOREM 2.4. *A descent statistic* st *is shuffle-compatible if and only if there exists a $\mathbb{Q}$-algebra homomorphism* $\phi_{\mathrm{st}} \colon \mathrm{QSym} \to \mathcal{A}_{\mathrm{st}}$, *where $\mathcal{A}_{\mathrm{st}}$ is a $\mathbb{Q}$-algebra with basis $\{u_\alpha\}$ indexed by* st-*equivalence classes $\alpha$ of compositions, such that* $\phi_{\mathrm{st}}(F_L) = u_\alpha$ *whenever $L$ is in the* st-*equivalence class $\alpha$.*

Gessel and the present author call $\mathcal{A}_{\mathrm{st}}$ the *shuffle algebra* of st, because the basis elements $u_\alpha$ can be viewed as encoding the distribution of st over shuffles of permutations. Theorem 2.4 implies that $\mathcal{A}_{\mathrm{st}}$ is isomorphic to a quotient of QSym whenever st is a shuffle-compatible descent statistic. We will not be working with the algebras $\mathcal{A}_{\mathrm{st}}$ themselves, but rather with the homomorphisms $\phi_{\mathrm{st}}$. Note that in the special case of the descent set, $\phi_{\mathrm{Des}}$ is an isomorphism and the basis $\{u_\alpha\}$ of $\mathcal{A}_{\mathrm{Des}}$ corresponds directly to the fundamental basis of QSym.

2.4. THE MALVENUTO–REUTENAUER ALGEBRA. Let $\mathbb{Q}[\mathfrak{S}]$ denote the $\mathbb{Q}$-vector space with basis elements the permutations in $\mathfrak{S}$. The *Malvenuto–Reutenauer algebra*, first defined in [24], is the $\mathbb{Q}$-algebra on $\mathbb{Q}[\mathfrak{S}]$ with the product

$$\pi \cdot \sigma = \sum_{\tau \in C(\pi, \sigma)} \tau$$

where $C(\pi, \sigma)$ is the set of *shifted concatenations* of $\pi$ and $\sigma$. That is, if $\pi \in \mathfrak{S}_m$ and $\sigma \in \mathfrak{S}_n$ then

$$C(\pi, \sigma) \coloneqq \{\, \tau \in \mathfrak{S}_{m+n} : \mathrm{std}(\tau_1 \cdots \tau_m) = \pi \text{ and } \mathrm{std}(\tau_{m+1} \cdots \tau_{m+n}) = \sigma \,\}.$$

Note that the Malvenuto–Reutenauer algebra is graded by the length of the permutation, and that its identity element is the empty permutation.

Rather than using the original construction of the Malvenuto–Reutenauer algebra as given above, we will follow the approach of Duchamp, Hivert, and Thibon [5], who gave another realization of the Malvenuto–Reutenauer algebra as a subalgebra of $\mathbb{Q}\langle\langle A^* \rangle\rangle$ where $A$ consists of the noncommuting variables $X_1, X_2, \ldots$. In order to describe their construction, we must revisit the standardization map std. We extend the map std to all words on the alphabet $\mathbb{P}$ of positive integers using the following rule: if a letter repeats, then they are viewed as increasing from left to right. For example, $\mathrm{std}(145411) = 146523$. We will later use the following fact, which is Proposition 5.3.2 of [20].

PROPOSITION 2.5. *Let $w = w_1 w_2 \cdots w_n$ be a word in $\mathbb{P}^*$ of length $n$, and let $\tau = \tau_1 \tau_2 \cdots \tau_n = \mathrm{std}(w)$. Then $\tau$ is the unique permutation in $\mathfrak{S}_n$ such that, whenever $1 \leqslant i < j \leqslant n$, we have $\tau_i < \tau_j$ if and only if $w_i \leqslant w_j$.*

For the remainder of this section, let $A = \{X_1, X_2, \ldots\}$ where the $X_i$ are noncommuting variables. Given a monomial $\mathbf{X} = X_{i_1} X_{i_2} \cdots X_{i_n}$, define $\mathrm{std}(\mathbf{X}) := \mathrm{std}(i_1 i_2 \cdots i_n)$. Then we associate to each permutation $\pi \in \mathfrak{S}$ an element $\mathbf{G}_\pi \in \mathbb{Q}\langle\langle A^* \rangle\rangle$ defined by

$$\mathbf{G}_\pi := \sum_{\substack{\mathbf{X} \in A^* \\ \mathrm{std}(\mathbf{X}) = \pi}} \mathbf{X}.$$

It can be shown that the $\mathbf{G}_\pi$ are linearly independent and multiply by the rule

$$\mathbf{G}_\pi \mathbf{G}_\sigma = \sum_{\tau \in C(\pi, \sigma)} \mathbf{G}_\tau,$$

so $\{\mathbf{G}_\pi\}_{\pi \in \mathfrak{S}}$ spans a $\mathbb{Q}$-subalgebra of $\mathbb{Q}\langle\langle A^* \rangle\rangle$, called the *algebra of free quasisymmetric functions* and denoted $\mathbf{FQSym}$. Since $\pi \mapsto \mathbf{G}_\pi$ is clearly a $\mathbb{Q}$-algebra isomorphism between $\mathbb{Q}[\mathfrak{S}]$ and $\mathbf{FQSym}$, we will henceforth refer to $\mathbf{FQSym}$ as the Malvenuto–Reutenauer algebra. We use $\mathbf{FQSym}$ instead of $\mathbb{Q}[\mathfrak{S}]$ because, by identifying permutations with elements of $\mathbb{Q}\langle\langle A^* \rangle\rangle$, we can prove our generalized cluster method for permutations using the cluster method for words.[3]

The Malvenuto–Reutenauer algebra $\mathbf{FQSym}$ contains an important subalgebra related to descent sets. Given a composition $L$, let $\mathbf{r}_L$ be the sum of all $\mathbf{G}_\pi$ for which $\pi$ has descent composition $L$; that is, let

$$\mathbf{r}_L := \sum_{\mathrm{Comp}(\pi) = L} \mathbf{G}_\pi.$$

The $\{\mathbf{r}_L\}_{L \vDash n, \, n \geqslant 0}$ is a linearly independent set and spans a $\mathbb{Q}$-subalgebra of $\mathbf{FQSym}$ called the *algebra of noncommutative symmetric functions*, denoted $\mathbf{Sym}$. Noncommutative symmetric functions were introduced in the seminal paper [12] of Gelfand et al., but implicitly appeared earlier in Gessel's Ph.D. thesis [18].

Let $\iota \colon \mathbf{Sym} \to \mathbf{FQSym}$ denote the canonical inclusion map from $\mathbf{Sym}$ to $\mathbf{FQSym}$. There is also a natural surjection $\rho \colon \mathbf{FQSym} \to \mathrm{QSym}$ given by

$$\rho(\mathbf{G}_\pi) := F_{\mathrm{Comp}(\pi^{-1})}. \tag{4}$$

The map $\rho$ explains the name "free quasisymmetric functions", as the elements of $\mathbf{FQSym}$ lift quasisymmetric functions to a noncommutative setting. We will need $\rho$ to define the homomorphisms on $\mathbf{FQSym}$ that we will use to study inverse statistics.

It is worth mentioning that QSym, $\mathbf{FQSym}$, and $\mathbf{Sym}$ are prototypical examples of combinatorial Hopf algebras, but we only need the algebra structure in our work. See

---

[3]It is possible to prove our generalized cluster method in $\mathbb{Q}[\mathfrak{S}]$, and we do this in the extended abstract [38]. While that approach is more direct, our approach here gives a unified treatment of the cluster method for words and the cluster method for permutations.

[20] for a survey on Hopf algebras in combinatorics, including more on the relationship between QSym, **FQSym**, and **Sym**.

2.5. SYMMETRIES ON PERMUTATIONS. Given $\pi \in \mathfrak{S}_n$, we define its *reverse* $\pi^r$ and its *complement* $\pi^c$ by

$$\pi^r := \pi_n \pi_{n-1} \cdots \pi_1 \quad \text{and} \quad \pi^c := (n+1-\pi_1)(n+1-\pi_2)\cdots(n+1-\pi_n),$$

respectively, and its *reverse-complement* $\pi^{rc}$ by $\pi^{rc} := (\pi^r)^c = (\pi^c)^r$.

It is clear that—along with permutation inversion $(\pi \mapsto \pi^{-1})$—reversion, complementation, and reverse-complementation are all involutions on $\mathfrak{S}_n$, and they can be identified with rigid motions in the dihedral group of the square acting on permutation matrices. As such, it is easy to see that $(\pi^{-1})^r = (\pi^c)^{-1}$ and $(\pi^{-1})^c = (\pi^r)^{-1}$—and hence $(\pi^{-1})^{rc} = (\pi^{rc})^{-1}$—for all $\pi \in \mathfrak{S}$.

PROPOSITION 2.6. *For any $\pi \in \mathfrak{S}_n$ with $n \geqslant 1$, we have*

(a) $\operatorname{imaj}(\pi^r) = \binom{n}{2} - \operatorname{imaj}(\pi)$,
(b) $\operatorname{imaj}(\pi^{rc}) = \operatorname{icomaj}(\pi)$,
(c) $\operatorname{ides}(\pi^r) = \operatorname{ides}(\pi^c) = n - 1 - \operatorname{ides}(\pi)$,
(d) $\operatorname{ides}(\pi^{rc}) = \operatorname{ides}(\pi)$, *and*
(e) $\operatorname{ipk}(\pi^c) = \operatorname{ipk}(\pi)$.

*Proof.* Since $\operatorname{maj}(\pi^c) = \binom{n}{2} - \operatorname{maj}(\pi)$ for all $\pi \in \mathfrak{S}_n$ [4, Lemma 2.5], we have

$$\operatorname{imaj}(\pi^r) = \operatorname{maj}((\pi^r)^{-1}) = \operatorname{maj}((\pi^{-1})^c) = \binom{n}{2} - \operatorname{maj}(\pi^{-1}) = \binom{n}{2} - \operatorname{imaj}(\pi),$$

which proves (a). The proofs of (b)–(e) are similar, using the identities $\operatorname{maj}(\pi^{rc}) = \operatorname{comaj}(\pi)$, $\operatorname{des}(\pi^r) = \operatorname{des}(\pi^c) = n - 1 - \operatorname{des}(\pi)$, $\operatorname{des}(\pi^{rc}) = \operatorname{des}(\pi)$, and $\operatorname{pk}(\pi^r) = \operatorname{pk}(\pi)$. $\qquad\square$

Let us define $\Gamma^r := \{ \pi^r : \pi \in \Gamma \}$ for a set $\Gamma \subseteq \mathfrak{S}$ of permutations, and we define $\Gamma^c$ and $\Gamma^{rc}$ in the analogous way. The next proposition tells us that if two sets of patterns $\Gamma$ and $\Delta$ are related by one of these symmetries, then we may be able to compute $A_{\Gamma,n}^{(\operatorname{ides},\operatorname{imaj})}(s,t,q)$, $A_{\Gamma,n}^{\operatorname{ides}}(s,t)$, or $P_{\Gamma,n}^{\operatorname{ipk}}(s,t)$ using the corresponding polynomials for $\Delta$. For example, once we obtain a generating function formula for the polynomials $A_{12\cdots m,n}^{(\operatorname{ides},\operatorname{imaj})}(t,q)$ in Section 4.2, we can use this formula along with Proposition 2.7 (a) to compute the polynomials $A_{m\cdots21,n}^{(\operatorname{ides},\operatorname{imaj})}(t,q)$.

PROPOSITION 2.7. *For any $\pi \in \mathfrak{S}_n$ with $n \geqslant 1$, we have*

(a) $A_{\Gamma^r,n}^{(\operatorname{ides},\operatorname{imaj})}(s,t,q) = t^{n+1} q^{\binom{n}{2}} A_{\Gamma,n}^{(\operatorname{ides},\operatorname{imaj})}(s,t^{-1},q^{-1})$,
(b) $A_{\Gamma^{rc},n}^{(\operatorname{ides},\operatorname{imaj})}(s,t,q) = A_{\Gamma,n}^{(\operatorname{ides},\operatorname{icomaj})}(s,t,q)$,
(c) $A_{\Gamma^{rc},n}^{\operatorname{ides}}(s,t) = A_{\Gamma,n}^{\operatorname{ides}}(s,t)$,
(d) $A_{\Gamma^r,n}^{\operatorname{ides}}(s,t) = A_{\Gamma^c,n}^{\operatorname{ides}}(s,t) = t^{n+1} A_{\Gamma,n}^{\operatorname{ides}}(s,t^{-1})$, *and*
(e) $P_{\Gamma^c,n}^{\operatorname{ipk}}(s,t) = P_{\Gamma,n}^{\operatorname{ipk}}(s,t)$.

*Proof.* Each of these identities follows from algebraic manipulations, Proposition 2.6, and the fact that an occurrence of a pattern $\sigma$ in $\pi$ directly corresponds to an occurrence of $\sigma^r$ (respectively, $\sigma^c$ and $\sigma^{rc}$) in $\pi^r$ (respectively, $\pi^c$ and $\pi^{rc}$). We demonstrate the

proof for (a) and leave the rest to the reader:

$$t^{n+1}q^{\binom{n}{2}}A_{\Gamma,n}^{(\mathrm{ides,imaj})}(s,t^{-1},q^{-1}) = t^{n+1}q^{\binom{n}{2}}\sum_{\pi\in\mathfrak{S}_n}s^{\mathrm{occ}_\Gamma(\pi)}(t^{-1})^{\mathrm{ides}(\pi)+1}(q^{-1})^{\mathrm{imaj}(\pi)}$$

$$= \sum_{\pi\in\mathfrak{S}_n}s^{\mathrm{occ}_\Gamma(\pi)}t^{n+1-(n-1-\mathrm{ides}(\pi^r))-1}q^{\binom{n}{2}-\mathrm{imaj}(\pi)}$$

$$= \sum_{\pi\in\mathfrak{S}_n}s^{\mathrm{occ}_{\Gamma^r}(\pi^r)}t^{\mathrm{ides}(\pi^r)+1}q^{\mathrm{imaj}(\pi^r)}$$

$$= \sum_{\pi\in\mathfrak{S}_n}s^{\mathrm{occ}_{\Gamma^r}(\pi)}t^{\mathrm{ides}(\pi)+1}q^{\mathrm{imaj}(\pi)}$$

$$= A_{\Gamma^r,n}^{(\mathrm{ides,imaj})}(s,t,q). \qquad \square$$

## 3. The cluster method in Malvenuto–Reutenauer

### 3.1. Main result.
Given a set of permutations $\Gamma\subseteq\mathfrak{S}$, define

$$\bar{F}_\Gamma(s) := \sum_{\pi\in\mathfrak{S}}\mathbf{G}_\pi s^{\mathrm{occ}_\Gamma(\pi)} \quad\text{and}\quad \bar{R}_\Gamma(s) := \sum_{\pi\in\mathfrak{S}}\mathbf{G}_\pi\sum_{c\in C_{\Gamma,\pi}}s^{\mathrm{mk}_\Gamma(c)},$$

which are liftings of the exponential generating functions $F_\Gamma(s,x)$ and $R_\Gamma(s,x)$ from Section 2.2.

**Theorem 3.1** (Cluster method in **FQSym**). *Let $\Gamma\subseteq\mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$\bar{F}_\Gamma(s) = \left(1 - \mathbf{G}_1 - \bar{R}_\Gamma(s-1)\right)^{-1}. \tag{5}$$

We will prove Theorem 3.1 using Theorem 2.1, the noncommutative version of the original Goulden–Jackson cluster method for words. Our proof will rely on several preliminary lemmas, which we establish below.

**Lemma 3.2.** *Let $u = u_1u_2\cdots u_n$ and $v = v_1v_2\cdots v_n$ be two words in $\mathbb{P}^*$. If $\mathrm{std}(u) = \mathrm{std}(v)$, then for any $0\leqslant m\leqslant n-1$ and $1\leqslant k\leqslant n-m$, we have $\mathrm{std}(u_ku_{k+1}\cdots u_{k+m}) = \mathrm{std}(v_kv_{k+1}\cdots v_{k+m})$.*

*Proof.* Recall from Proposition 2.5 that if $\tau = \mathrm{std}(u)$, then whenever $1\leqslant i < j\leqslant n$, we have $u_i\leqslant u_j$ if and only if $\tau_i < \tau_j$. Since we are given that $\mathrm{std}(u) = \mathrm{std}(v)$, we have $u_i\leqslant u_j$ if and only if $v_i\leqslant v_j$ for all $1\leqslant i < j\leqslant n$. In particular, we have $u_i\leqslant u_j$ if and only if $v_i\leqslant v_j$ for all $k\leqslant i < j\leqslant k+m$; thus $\mathrm{std}(u_ku_{k+1}\cdots u_{k+m}) = \mathrm{std}(v_kv_{k+1}\cdots v_{k+m})$. $\square$

For the remainder of this section, let $A$ be the set of noncommuting variables $\{X_1, X_2, \dots\}$, let $M(\pi)$ be the set of monomials in these variables whose standardization is $\pi$, and let $B = \bigsqcup_{\sigma\in\Gamma}M(\sigma)$.

**Lemma 3.3.** *If $\mathbf{X}\in M(\pi)$, then $\mathrm{occ}_B(\mathbf{X}) = \mathrm{occ}_\Gamma(\pi)$.*

*Proof.* Write $\pi = \pi_1\pi_2\cdots\pi_n$ and $\mathbf{X} = X_{i_1}X_{i_2}\cdots X_{i_n}$. Since $\mathbf{X}\in M(\pi)$, we have that $\mathrm{std}(i_1i_2\cdots i_n) = \pi$. Suppose that $X_{i_k}X_{i_{k+1}}\cdots X_{i_{k+m}}$ is an occurrence of a word from $B$, so $X_{i_k}X_{i_{k+1}}\cdots X_{i_{k+m}}\in M(\sigma)$ for some $\sigma\in\Gamma$ and thus

$$\mathrm{std}(i_ki_{k+1}\cdots i_{k+m}) = \mathrm{std}(X_{i_k}X_{i_{k+1}}\cdots X_{i_{k+m}}) = \sigma.$$

Since $\mathrm{std}(i_1i_2\cdots i_n) = \pi$ and $\mathrm{std}(i_ki_{k+1}\cdots i_{k+m}) = \sigma$, it follows from Lemma 3.2 that $\mathrm{std}(\pi_k\pi_{k+1}\cdots\pi_{k+m}) = \sigma$. In other words, $\pi_k\pi_{k+1}\cdots\pi_{k+m}$ is an occurrence of $\sigma$ in $\pi$. We can go backward to see that there is a bijection between occurrences of words from $B$ in $\mathbf{X}$ and patterns from $\Gamma$ in $\pi$, which shows $\mathrm{occ}_B(\mathbf{X}) = \mathrm{occ}_\Gamma(\pi)$. $\square$

LEMMA 3.4. *If $\mathbf{X} \in M(\pi)$, then $\sum_{c \in C_{B,\mathbf{x}}} s^{\mathrm{mk}_B(c)} = \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)}$.*

*Proof.* Similar reasoning as above can be used to show that there is a bijection between $B$-clusters on $\mathbf{X}$ and $\Gamma$-clusters on $\pi$ which preserves the number (and positions) of marked occurrences; we omit the details. □

We are now ready to prove Theorem 3.1.

*Proof of Theorem 3.1.* As a consequence of Lemma 3.3, we have that

$$\sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi s^{\mathrm{occ}_\Gamma(\pi)} = \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi s^{\mathrm{occ}_B(\pi)}$$

where $\mathrm{occ}_B(\pi) := \mathrm{occ}_B(\mathbf{X})$ for any $\mathbf{X} \in M(\pi)$. Hence

$$\begin{aligned}
\bar{F}_\Gamma(s) &= \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi s^{\mathrm{occ}_\Gamma(\pi)} \\
&= \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi s^{\mathrm{occ}_B(\pi)} \\
&= \sum_{\mathbf{X} \in A^*} \mathbf{X} s^{\mathrm{occ}_B(\mathbf{X})} \\
&= F_B(s).
\end{aligned} \tag{6}$$

Similarly, Lemma 3.4 implies

$$\sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)} = \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi \sum_{c \in C_{B,\pi}} s^{\mathrm{mk}_B(c)}$$

where $C_{B,\pi} := C_{B,\mathbf{X}}$ for any $\mathbf{X} \in M(\pi)$. Thus

$$\begin{aligned}
\bar{R}_\Gamma(s) &= \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)} \\
&= \sum_{\pi \in \mathfrak{S}} \mathbf{G}_\pi \sum_{c \in C_{B,\pi}} s^{\mathrm{mk}_B(c)} \\
&= \sum_{\mathbf{X} \in A^*} \mathbf{X} \sum_{c \in C_{B,\mathbf{x}}} s^{\mathrm{mk}_B(c)} \\
&= R_B(s).
\end{aligned} \tag{7}$$

Finally, we use Theorem 2.1 along with Equations (6) and (7) to conclude

$$\begin{aligned}
\bar{F}_\Gamma(s) &= F_B(s) \\
&= \left( 1 - \mathbf{G}_1 - R_B(s-1) \right)^{-1} \\
&= \left( 1 - \mathbf{G}_1 - \bar{R}_\Gamma(s-1) \right)^{-1}.
\end{aligned}$$ □

3.2. TWO BASIC HOMOMORPHISMS. We now demonstrate how Elizalde and Noy's cluster method for permutations, as well as Elizalde's $q$-analogue, can be recovered from the cluster method in **FQSym**.

Given $\pi \in \mathfrak{S}_n$, define the maps $\Psi \colon \mathbf{FQSym} \to \mathbb{Q}[[x]]$ and $\Psi_q \colon \mathbf{FQSym} \to \mathbb{Q}[[q,x]]$ by

$$\Psi(\mathbf{G}_\pi) := \frac{x^n}{n!} \qquad \text{and} \qquad \Psi_q(\mathbf{G}_\pi) := q^{\mathrm{inv}(\pi)} \frac{x^n}{[n]_q!}$$

and extending linearly.

PROPOSITION 3.5. *The maps $\Psi$ and $\Psi_q$ are $\mathbb{Q}$-algebra homomorphisms.*

*Proof.* Let $\pi \in \mathfrak{S}_m$ and $\sigma \in \mathfrak{S}_n$. Using the multiplication rule for the $\mathbf{G}_\pi$, the definition of the map $\Psi_q$, and the identity

$$\sum_{\tau \in C(\pi,\sigma)} q^{\mathrm{inv}(\tau)} = q^{\mathrm{inv}(\pi)+\mathrm{inv}(\sigma)} \binom{m+n}{n}_q$$

(see [3, Lemma 2.1]), we obtain

$$\begin{aligned}
\Psi_q(\mathbf{G}_\pi \mathbf{G}_\sigma) &= \sum_{\tau \in C(\pi,\sigma)} \Psi_q(\mathbf{G}_\tau) \\
&= \sum_{\tau \in C(\pi,\sigma)} q^{\mathrm{inv}(\tau)} \frac{x^{m+n}}{[m+n]_q!} \\
&= q^{\mathrm{inv}(\pi)+\mathrm{inv}(\sigma)} \binom{m+n}{n}_q \frac{x^{m+n}}{[m+n]_q!} \\
&= q^{\mathrm{inv}(\pi)} \frac{x^m}{[m]_q!} q^{\mathrm{inv}(\sigma)} \frac{x^n}{[n]_q!} \\
&= \Psi_q(\mathbf{G}_\pi) \Psi_q(\mathbf{G}_\sigma);
\end{aligned}$$

hence, $\Psi_q$ is a homomorphism. Setting $q = 1$ above yields

$$\Psi(\mathbf{G}_\pi \mathbf{G}_\sigma) = \sum_{\tau \in C(\pi,\sigma)} \frac{x^{m+n}}{(m+n)!} = \frac{x^m}{m!} \frac{x^n}{n!} = \Psi(\mathbf{G}_\pi)\Psi(\mathbf{G}_\sigma),$$

which shows that $\Psi$ is a homomorphism as well. □

The homomorphism $\Psi$ is precisely the homomorphism $\phi$ of Josuat-Vergès, Novelli, and Thibon mentioned in the introduction of this paper. It is easy to see that upon applying $\Psi$ to our cluster method in **FQSym** (Theorem 3.1), we recover Elizalde and Noy's cluster method for permutations (Theorem 2.2). Applying $\Psi_q$ instead yields a proof of Elizalde's $q$-cluster method (Theorem 2.3).

We also note that $\Psi$ and $\Psi_q$ are closely related to two homomorphisms, denoted $\Phi$ and $\Phi_q$, which appear in [36]. (The map $\Phi$ also appears in [14, 16, 35].) These two homomorphisms are defined on the algebra **Sym** of noncommutative symmetric functions by the formulas

$$\Phi(\mathbf{h}_n) \coloneqq \frac{x^n}{n!} \quad \text{and} \quad \Phi_q(\mathbf{h}_n) \coloneqq \frac{x^n}{[n]_q!},$$

where $\mathbf{h}_n \coloneqq \sum_{i_1 \leqslant \cdots \leqslant i_n} X_{i_1} X_{i_2} \cdots X_{i_n} = \mathbf{G}_{12 \cdots n}$. In fact, $\Phi$ and $\Phi_q$ are related to our homomorphisms $\Psi$ and $\Psi_q$ by

$$\Phi = \Psi \circ \iota \quad \text{and} \quad \Phi_q = \Psi_q \circ \iota$$

where $\iota$ is the canonical inclusion from **Sym** to **FQSym**.

3.3. A NOTE ON THE HADAMARD PRODUCT. Our next goal is to define a family of homomorphisms on **FQSym** that can be used to produce other specializations of our generalized cluster method. Our starting point is Theorem 2.4, which states that every shuffle-compatible descent statistic st gives rise to a homomorphism $\phi_{\mathrm{st}}$ from the algebra QSym of quasisymmetric functions to the shuffle algebra $\mathcal{A}_{\mathrm{st}}$ of st. Many of these algebras $\mathcal{A}_{\mathrm{st}}$ can be characterized as subalgebras of various formal power series algebras in which the multiplication is the "Hadamard product" in a variable $t$, which we define below.

The operation of *Hadamard product* $*$ on formal power series in $t$ is defined by

$$\Big( \sum_{n=0}^{\infty} a_n t^n \Big) * \Big( \sum_{n=0}^{\infty} b_n t^n \Big) := \sum_{n=0}^{\infty} a_n b_n t^n.$$

In our notation for formal power series algebras, we write $t*$ to indicate that multiplication is the Hadamard product in $t$. For example, $\mathbb{Q}[[t*, x]]$ is the $\mathbb{Q}$-algebra of formal power series in the variables $t$ and $x$, where multiplication is ordinary multiplication in $x$ but is the Hadamard product in $t$. Thus we have

$$\Big( \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} a_{n,k} x^k t^n \Big) * \Big( \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} b_{n,k} x^k t^n \Big) = \sum_{n=0}^{\infty} \Big( \sum_{k=0}^{\infty} a_{n,k} x^k \Big) \Big( \sum_{k=0}^{\infty} b_{n,k} x^k \Big) t^n$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{k} a_{n,j} b_{n,k-j} x^k t^n$$

in $\mathbb{Q}[[t*, x]]$. Note that the identity of $\mathbb{Q}[[t*, x]]$ is $1/(1-t)$.

We write $f^{*\langle n \rangle}$ to mean the $n$-fold Hadamard product of $f$, and $f^{*\langle -1 \rangle}$ the inverse of $f$ with respect to Hadamard product. For example, we have

$$(3t + t^2)^{*\langle n \rangle} = \underbrace{(3t + t^2) * (3t + t^2) * \cdots * (3t + t^2)}_{n \text{ times}} = 3^n t + t^2$$

and

$$\Big( \frac{1}{1-t} - 2xt \Big)^{*\langle -1 \rangle} = \sum_{n=0}^{\infty} (2xt)^{*\langle n \rangle} = \frac{1}{1-t} + 2xt + 4x^2 t + 8x^3 t + \cdots .$$

We will always use the notations $*$, $*\langle n \rangle$, or $*\langle -1 \rangle$ for any expression involving the Hadamard product; all other expressions should be interpreted as using ordinary multiplication.

3.4. HOMOMORPHISMS ARISING FROM SHUFFLE-COMPATIBILITY. Given $\pi \in \mathfrak{S}_n$, let us define

$$\Psi_{(\mathrm{ides, icomaj})} \colon \mathbf{FQSym} \to \mathcal{A}_{(\mathrm{des, comaj})} \subseteq \mathbb{Q}[[t*, q, x]],$$
$$\Psi_{\mathrm{ipk}} \colon \mathbf{FQSym} \to \mathcal{A}_{\mathrm{pk}} \subseteq \mathbb{Q}[[t*, x]], \text{ and}$$
$$\Psi_{\mathrm{ilpk}} \colon \mathbf{FQSym} \to \mathcal{A}_{\mathrm{lpk}} \subseteq \mathbb{Q}[[t*, x]]$$

by

$$\Psi_{(\mathrm{ides, icomaj})}(\mathbf{G}_\pi) := \begin{cases} \dfrac{t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)}}{\prod_{i=0}^{n}(1 - tq^i)} x^n, & \text{if } n \geqslant 1, \\ 1/(1-t), & \text{if } n = 0, \end{cases}$$

$$\Psi_{\mathrm{ipk}}(\mathbf{G}_\pi) := \begin{cases} \dfrac{2^{2\,\mathrm{ipk}(\pi)+1} t^{\mathrm{ipk}(\pi)+1}(1+t)^{n-2\,\mathrm{ipk}(\pi)-1}}{(1-t)^{n+1}} x^n, & \text{if } n \geqslant 1, \\ 1/(1-t), & \text{if } n = 0, \text{ and} \end{cases}$$

$$\Psi_{\mathrm{ilpk}}(\mathbf{G}_\pi) := \frac{2^{2\,\mathrm{ilpk}(\pi)} t^{\mathrm{ilpk}(\pi)}(1+t)^{n-2\,\mathrm{ilpk}(\pi)}}{(1-t)^{n+1}} x^n$$

and extending linearly.

THEOREM 3.6. *The maps* $\Psi_{(\mathrm{ides, icomaj})}$, $\Psi_{\mathrm{ipk}}$, *and* $\Psi_{\mathrm{ilpk}}$ *are* $\mathbb{Q}$*-algebra homomorphisms.*

*Proof.* By Theorems 4.5, 4.8, and 4.10 of [17], the descent statistics (des, comaj), pk, and lpk are all shuffle-compatible and their homomorphisms

$$\phi_{(\mathrm{des,comaj})}\colon \mathrm{QSym} \to \mathcal{A}_{(\mathrm{comaj,des})} \subseteq \mathbb{Q}[[t*, q, x]],$$
$$\phi_{\mathrm{pk}}\colon \mathrm{QSym} \to \mathcal{A}_{\mathrm{pk}} \subseteq \mathbb{Q}[[t*, x]], \text{ and}$$
$$\phi_{\mathrm{lpk}}\colon \mathrm{QSym} \to \mathcal{A}_{\mathrm{lpk}} \subseteq \mathbb{Q}[[t*, x]]$$

(see Theorem 2.4) are defined by

$$\phi_{(\mathrm{des,comaj})}(F_L) := \begin{cases} \dfrac{t^{\mathrm{des}(L)+1} q^{\mathrm{comaj}(L)}}{\prod_{i=0}^{n}(1-tq^i)} x^n, & \text{if } n \geqslant 1, \\ 1/(1-t), & \text{if } n = 0, \end{cases}$$

$$\phi_{\mathrm{pk}}(F_L) := \begin{cases} \dfrac{2^{2\,\mathrm{pk}(L)+1} t^{\mathrm{pk}(L)+1}(1+t)^{n-2\,\mathrm{pk}(L)-1}}{(1-t)^{n+1}} x^n, & \text{if } n \geqslant 1, \\ 1/(1-t), & \text{if } n = 0, \text{ and} \end{cases}$$

$$\phi_{\mathrm{lpk}}(F_L) := \dfrac{2^{2\,\mathrm{lpk}(L)} t^{\mathrm{lpk}(L)}(1+t)^{n-2\,\mathrm{lpk}(L)}}{(1-t)^{n+1}} x^n,$$

where $L \vDash n$. The maps $\Psi_{(\mathrm{ides,icomaj})}$, $\Psi_{\mathrm{ipk}}$, and $\Psi_{\mathrm{ilpk}}$ are simply the result of composing these homomorphisms $\phi_{(\mathrm{des,comaj})}$, $\phi_{\mathrm{pk}}$, and $\phi_{\mathrm{lpk}}$ with the canonical surjection $\rho$ from **FQSym** to QSym—see Equation (4). Since compositions of homomorphisms are homomorphisms, the result follows. $\qquad\square$

3.5. Further specializations of the generalized cluster method. We now use the homomorphisms defined in the previous section to produce further specializations of our generalized cluster method which can be used to relate the polynomials $A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)$, $A_{\Gamma,n}^{\mathrm{ides}}(s,t)$, $P_{\Gamma,n}^{\mathrm{ipk}}(s,t)$, and $P_{\Gamma,n}^{\mathrm{ilpk}}(s,t)$—defined in Section 1.1—to "refined cluster polynomials". These specializations are similar in spirit to Elizalde's $q$-cluster method in that they count permutations by occurrences of prescribed patterns but also keep track of additional statistics.

We begin with (ides, icomaj). Given a set $\Gamma \subseteq \mathfrak{S}$, let

$$R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s,t,q) := \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)},$$

which counts $\Gamma$-clusters of length $k$ by the number of marked occurrences as well as the inverse descent number and inverse comajor index of the underlying permutation.

**Theorem 3.7.** *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$\sum_{n=0}^{\infty} \frac{A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$$

$$= \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)(1-tq)} + \sum_{k=2}^{\infty} R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)} \right)^{*\langle n \rangle}.$$

We note that this formula lives inside the formal power series algebra $\mathbb{Q}[[s, t*, q, x]]$, although the Hadamard product is only present on the right-hand side.

*Proof.* Take Equation (5) from Theorem 3.1, and then apply the homomorphism $\Psi_{(\mathrm{ides,icomaj})}$ to both sides. Observe that

$$\Psi_{(\mathrm{ides,icomaj})}(\bar{F}_\Gamma(s)) = \sum_{\pi \in \mathfrak{S}_n} \Psi_{(\mathrm{ides,icomaj})}(\mathbf{G}_\pi) s^{\mathrm{occ}_\Gamma(\pi)}$$

$$= \frac{1}{1-t} + \sum_{n=1}^{\infty} \sum_{\pi \in \mathfrak{S}_n} \frac{s^{\mathrm{occ}_\Gamma(\pi)} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)}}{\prod_{i=0}^{n}(1-tq^i)} x^n$$

$$= \sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} \frac{A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$$

and

$$\Psi_{(\mathrm{ides,icomaj})}(\bar{R}_\Gamma(s-1)) = \sum_{\pi \in \mathfrak{S}} \Psi_{(\mathrm{ides,icomaj})}(\mathbf{G}_\pi) \sum_{c \in C_{\Gamma,\pi}} (s-1)^{\mathrm{mk}_\Gamma(c)}$$

$$= \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \frac{t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)} x^k}{\prod_{i=0}^{k}(1-tq^i)} \sum_{c \in C_{\Gamma,\pi}} (s-1)^{\mathrm{mk}_\Gamma(c)}$$

$$= \sum_{k=2}^{\infty} R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)},$$

and also $\Psi_{(\mathrm{ides,icomaj})}(1) = 1/(1-t)$ and $\Psi_{(\mathrm{ides,icomaj})}(\mathbf{G}_1) = tx/((1-t)(1-tq))$. Thus, we have

$$\sum_{n=0}^{\infty} \sum_{\pi \in \mathfrak{S}_n} \frac{A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$$

$$= \left( \frac{1}{1-t} - \frac{tx}{(1-t)(1-tq)} - \sum_{k=2}^{\infty} R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)} \right)^{*\langle -1 \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)(1-tq)} + \sum_{k=2}^{\infty} R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)} \right)^{*\langle n \rangle}. \qquad \square$$

Let us give two remarks before proceeding. First, recall the identity

$$\mathrm{comaj}(\pi) = n\,\mathrm{des}(\pi) - \mathrm{maj}(\pi),$$

which is equivalent to

$$\mathrm{imaj}(\pi) = n\,\mathrm{ides}(\pi) - \mathrm{icomaj}(\pi).$$

It follows that

$$A_{\Gamma,n}^{(\mathrm{ides,imaj})}(s,t,q) = q^{-n} A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,tq^n,q^{-1}),$$

so we can compute the polynomials $A_{\Gamma,n}^{(\mathrm{ides,imaj})}(s,t,q)$ from the $A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)$. In other words, having a formula for the polynomials $A_{\Gamma,n}^{(\mathrm{ides,icomaj})}(s,t,q)$ is equivalent to having one for the $A_{\Gamma,n}^{(\mathrm{ides,imaj})}(s,t,q)$.

Furthermore, in using Theorem 3.7 to compute the polynomial $A_{\Gamma,j}^{(\mathrm{ides,icomaj})}(s,t,q)$, one only needs to sum from $n=0$ to $n=j$ on the right-hand side. This is because, by the definition of Hadamard product in $t$, the coefficient of $x^j$ in

$$\left( \frac{tx}{(1-t)(1-tq)} + \sum_{k=2}^{\infty} R_{\Gamma,k}^{(\mathrm{ides,icomaj})}(s-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)} \right)^{*\langle n \rangle}$$

is zero unless $n \leqslant j$. The same is true for the other formulas in this section.

We now specialize Theorem 3.7 to an analogous result solely for the inverse descent number. Let

$$R_{\Gamma,k}^{\mathrm{ides}}(s,t) := \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ides}(\pi)+1} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)}$$

be the refined cluster polynomial for ides.

THEOREM 3.8. *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$\sum_{n=0}^{\infty} \frac{A_{\Gamma,n}^{\mathrm{ides}}(s,t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)^2} + \frac{1}{1-t} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ides}}(s-1,t) z^k \right)^{*\langle n \rangle}$$

*where $z = x/(1-t)$.*

*Proof.* This follows immediately from setting $q = 1$ in Theorem 3.7 and simplifying. $\square$

For the inverse peak statistics ipk and ilpk, let us define

$$R_{\Gamma,k}^{\mathrm{ipk}}(s,t) := \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ipk}(\pi)+1} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)} \quad \text{and} \quad R_{\Gamma,k}^{\mathrm{ilpk}}(s,t) := \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ilpk}(\pi)} \sum_{c \in C_{\Gamma,\pi}} s^{\mathrm{mk}_\Gamma(c)}.$$

Then the following two theorems can be proven in the same way as Theorem 3.7, but using the homomorphisms $\Psi_{\mathrm{ipk}}$ and $\Psi_{\mathrm{ilpk}}$. We outline the steps for Theorem 3.9 but omit the proof of Theorem 3.10.

THEOREM 3.9. *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P_{\Gamma,n}^{\mathrm{ipk}}(s,u) z^n$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} + \frac{1+t}{2(1-t)} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ipk}}(s-1,u) z^k \right)^{*\langle n \rangle}$$

*where $u = 4t/(1+t)^2$ and $z = (1+t)x/(1-t)$.*

*Proof.* We shall apply $\Psi_{\mathrm{ipk}}$ to both sides of (5). Observe that

$$\Psi_{\mathrm{ipk}}(\bar{F}_\Gamma(s)) = \frac{1}{1-t} + \sum_{n=1}^{\infty} \sum_{\pi \in \mathfrak{S}_n} \frac{2^{2\,\mathrm{ipk}(\pi)+1} t^{\mathrm{ipk}(\pi)+1} (1+t)^{n-2\,\mathrm{ipk}(\pi)-1} s^{\mathrm{occ}_\Gamma(\pi)}}{(1-t)^{n+1}} x^n$$

$$= \frac{1}{1-t} + \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{1+t}{1-t} \right)^{n+1} P_{\Gamma,n}^{\mathrm{ipk}}\left( s, \frac{4t}{(1+t)^2} \right) x^n$$

$$= \frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P_{\Gamma,n}^{\mathrm{ipk}}(s,u) z^n$$

and

$$\Psi_{\mathrm{ipk}}(\bar{R}_\Gamma(s-1))$$

$$= \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \frac{2^{2\,\mathrm{ipk}(\pi)+1} t^{\mathrm{ipk}(\pi)+1} (1+t)^{k-2\,\mathrm{ipk}(\pi)-1} x^k}{(1-t)^{k+1}} \sum_{c \in C_{\Gamma,\pi}} (s-1)^{\mathrm{mk}_\Gamma(c)}$$

$$= \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ipk}}\left( s-1, \frac{4t}{(1+t)^2} \right) \frac{(1+t)^{k+1} x^k}{2(1-t)^{k+1}}$$

$$= \frac{1+t}{2(1-t)} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ipk}}(s-1,u) z^k.$$

Also, we have $\Psi_{\mathrm{ipk}}(1) = 1/(1-t)$ and $\Psi_{\mathrm{ipk}}(\mathbf{G}_1) = 2tx/(1-t)^2$. Hence, we obtain

$$
\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P_{\Gamma,n}^{\mathrm{ipk}}(s,u) z^n
$$
$$
= \left( \frac{1}{1-t} - \frac{2tx}{(1-t)^2} - \frac{1+t}{2(1-t)} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ipk}}(s-1,u) z^k \right)^{*\langle -1 \rangle}
$$
$$
= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} + \frac{1+t}{2(1-t)} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ipk}}(s-1,u) z^k \right)^{*\langle n \rangle}. \qquad \square
$$

THEOREM 3.10. *Let $\Gamma \subseteq \mathfrak{S}$ be a set of permutations, each of length at least 2. Then*

$$
\frac{1}{1-t} \sum_{n=0}^{\infty} P_{\Gamma,n}^{\mathrm{ilpk}}(s,u) z^n = \sum_{n=0}^{\infty} \left( \frac{z}{1-t} + \frac{1}{1-t} \sum_{k=2}^{\infty} R_{\Gamma,k}^{\mathrm{ilpk}}(s-1,u) z^k \right)^{*\langle n \rangle}
$$

*where $u = 4t/(1+t)^2$ and $z = (1+t)x/(1-t)$.*

## 4. MONOTONE PATTERNS $12\cdots m$ AND $m\cdots 21$

4.1. CLUSTER GENERATING FUNCTIONS FOR MONOTONE PATTERNS. In this section, we will study the polynomials $A_{\sigma,n}^{(\mathrm{ides,imaj})}(t,q)$, $A_{\sigma,n}^{\mathrm{ides}}(s,t,q)$, $A_{\sigma,n}^{\mathrm{ides}}(t,q)$, $P_{\sigma,n}^{\mathrm{ipk}}(s,t)$, $P_{\sigma,n}^{\mathrm{ipk}}(t)$, $P_{\sigma,n}^{\mathrm{ilpk}}(s,t)$, and $P_{\sigma,n}^{\mathrm{ilpk}}(t)$ for $\sigma = 12\cdots m$ and $\sigma = m\cdots 21$. Our formulas will mostly be for the pattern $\sigma = 12\cdots m$, but we can use these formulas along with Proposition 2.7 to compute most of these polynomials for $\sigma = m\cdots 21$ as well.

We begin with a lemma establishing closed-form generating functions for refined $12\cdots m$-cluster polynomials, which we need in order to apply our results from Section 3.5. Note that, in general, there is no straightforward way to count clusters by inverse statistics. As a matter of fact, the simpler problem of counting clusters (without keeping track of any statistic) is equivalent to counting linear extensions of a certain poset [10], which is itself difficult in general. Yet, counting $\sigma$-clusters by our inverse statistics is essentially trivial when $\sigma$ is a monotone pattern.

LEMMA 4.1. *For all $m \geqslant 2$, we have*

$$
\sum_{k=2}^{\infty} R_{12\cdots m,k}^{(\mathrm{ides,icomaj})}(s,t,q) x^k = \sum_{k=2}^{\infty} R_{12\cdots m,k}^{\mathrm{ides}}(s,t) x^k = \sum_{k=2}^{\infty} R_{12\cdots m,k}^{\mathrm{ipk}}(s,t) x^k
$$
$$
= \frac{stx^m}{1 - s\sum_{l=1}^{m-1} x^l}
$$

*and*

$$
\sum_{k=2}^{\infty} R_{12\cdots m,k}^{\mathrm{ilpk}}(s,t) x^k = \frac{sx^m}{1 - s\sum_{l=1}^{m-1} x^l}.
$$

*Proof.* It is easy to see that there exists a $12\cdots m$-cluster on $\pi$ if and only if $\pi$ is itself of the form $12\cdots n$ where $n \geqslant m$, and that the overlap set of $12\cdots m$ is given by $O_{12\cdots m} = \{1, 2, \ldots, m-1\}$. Hence, we can uniquely generate $12\cdots m$-clusters by first taking the permutation $12\cdots m$, and then repeatedly appending the next $l$ largest integers (for any $1 \leqslant l \leqslant m-1$) in increasing order—each iteration creates an additional marked occurrence of $12\cdots m$. Figure 1 provides an illustration for the
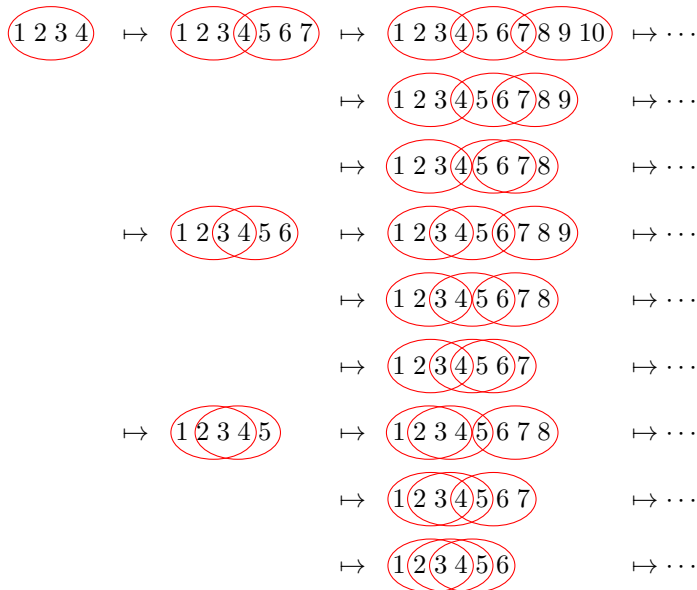
$$\boxed{1\,2\,3\,4} \;\mapsto\; \boxed{1\,2\,3\,\boxed{4}\,5\,6\,7} \;\mapsto\; \boxed{1\,2\,3\,\boxed{4}\,5\,6\,\boxed{7}\,8\,9\,10} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,2\,3\,\boxed{4}\,5\,\boxed{6}\,\boxed{7}\,8\,9} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,2\,3\,\boxed{4}\,5\,6\,7\,8} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,2\,\boxed{3\,4}\,5\,6} \;\mapsto\; \boxed{1\,2\,\boxed{3\,4}\,5\,\boxed{6}\,7\,8\,9} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,2\,\boxed{3\,4}\,5\,\boxed{6}\,7\,8} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,2\,\boxed{3}\,\boxed{4}\,5\,6\,7} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,\boxed{2}\,3\,4\,5} \;\mapsto\; \boxed{1\,\boxed{2}\,3\,4\,\boxed{5}\,6\,7\,8} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,\boxed{2}\,3\,\boxed{4}\,5\,6\,7} \;\mapsto \cdots$$

$$\mapsto\; \boxed{1\,\boxed{2}\,3\,4\,\boxed{5}\,6} \;\mapsto \cdots$$

FIGURE 1. 1234-clusters

case $m = 4$. Thus, we have the formula

$$\sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{12\cdots m, \pi}} s^{\mathrm{mk}_{12\cdots m}(c)} x^k = \frac{sx^m}{1 - s(x + x^2 + \cdots + x^{m-1})}$$

$$= \frac{sx^m}{1 - s\sum_{l=1}^{m-1} x^l} \tag{8}$$

(see also [10, p. 356]).

Moreover, since $12\cdots m$-clusters are themselves monotone increasing, their inverses are also monotone increasing and therefore have no descents, peaks, or left peaks. Using (8), it follows that

$$\sum_{k=2}^{\infty} R_{12\cdots m, k}^{(\mathrm{ides},\mathrm{icomaj})}(s, t, q) x^k = \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{icomaj}(\pi)} \sum_{c \in C_{12\cdots m, \pi}} s^{\mathrm{mk}_{12\cdots m}(c)} x^k$$

$$= t \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{12\cdots m, \pi}} s^{\mathrm{mk}_{12\cdots m}(c)} x^k$$

$$= \frac{stx^m}{1 - s\sum_{l=1}^{m-1} x^l};$$

our formulas for $\sum_{k=2}^{\infty} R_{12\cdots m, k}^{\mathrm{ipk}}(s, t) x^k$ and $\sum_{k=2}^{\infty} R_{12\cdots m, k}^{\mathrm{ilpk}}(s, t) x^k$ are obtained in the same way. Lastly, since our formula for $\sum_{k=2}^{\infty} R_{12\cdots m, k}^{(\mathrm{ides},\mathrm{icomaj})}(s, t, q) x^k$ does not depend on $q$, we have the same formula for $\sum_{k=2}^{\infty} R_{12\cdots m, k}^{\mathrm{ides}}(s, t) x^k$. $\qquad\square$

4.2. MONOTONE PATTERNS, INVERSE DESCENT NUMBER, AND INVERSE MAJOR INDEX. We will now derive generating function formulas for $A_{12\cdots m, n}^{(\mathrm{ides},\mathrm{imaj})}(t, q)$, $A_{12\cdots m, n}^{\mathrm{ides}}(s, t)$, and $A_{12\cdots m, n}^{\mathrm{ides}}(t)$.

THEOREM 4.2. *Let $m \geqslant 2$. We have*

(a) $\displaystyle\sum_{n=0}^{\infty} \frac{A_{12\cdots m,n}^{(\mathrm{ides,imaj})}(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$

$$= \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)(1-tq)} - \sum_{j=1}^{\infty} \left( \frac{tx^{jm}}{\prod_{i=0}^{jm}(1-tq^i)} - \frac{tx^{jm+1}}{\prod_{i=0}^{jm+1}(1-tq^i)} \right) \right)^{*\langle n \rangle}$$

*and*

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{A_{12\cdots m,n}^{(\mathrm{ides,imaj})}(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$

$$= 1 + \sum_{k=1}^{\infty} \left[ \sum_{j=0}^{\infty} \left( \binom{k+jm-1}{k-1}_q x^{jm} - \binom{k+jm}{k-1}_q x^{jm+1} \right) \right]^{-1} t^k.$$

*Proof.* In light of Proposition 2.7 (b) and the fact that $12\cdots m$ is invariant under reverse-complementation, it suffices to prove these formulas with the polynomial $A_{12\cdots m,n}^{(\mathrm{ides,imaj})}(t,q)$ replaced by $A_{12\cdots m,n}^{(\mathrm{ides,icomaj})}(t,q)$.

We first apply Theorem 3.7 to $\Gamma = \{12\cdots m\}$ and set $s = 0$ to obtain

$$\sum_{n=0}^{\infty} \frac{A_{12\cdots m,n}^{(\mathrm{ides,icomaj})}(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n$$

$$= \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)(1-tq)} + \sum_{k=2}^{\infty} R_{12\cdots m,k}^{(\mathrm{ides,icomaj})}(-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)} \right)^{*\langle n \rangle}. \quad (9)$$

By Lemma 4.1, we have

$$\sum_{k=2}^{\infty} R_{12\cdots m,k}^{(\mathrm{ides,icomaj})}(-1,t,q) x^k = \frac{-tx^m}{1+x+x^2+\cdots+x^{m-1}}$$

$$= -\frac{tx^m(1-x)}{1-x^m}$$

$$= -\sum_{j=1}^{\infty}(tx^{jm} - tx^{jm+1})$$

and thus

$$\sum_{k=2}^{\infty} R_{12\cdots m,k}^{(\mathrm{ides,icomaj})}(-1,t,q) \frac{x^k}{\prod_{i=0}^{k}(1-tq^i)}$$

$$= -\sum_{j=1}^{\infty} \left( \frac{tx^{jm}}{\prod_{i=0}^{jm}(1-tq^i)} - \frac{tx^{jm+1}}{\prod_{i=0}^{jm+1}(1-tq^i)} \right). \quad (10)$$

Combining (9) with (10) yields part (a).

Now we prove part (b). Here we begin with (10), and use the identity

$$\frac{1}{\prod_{i=0}^{n}(1-tq^i)} = \sum_{k=0}^{\infty} \binom{n+k}{k}_q t^k$$

[31, p. 68] to arrive at

$$\sum_{n=0}^{\infty} \frac{A_{12\cdots m,n}^{(\mathrm{ides},\mathrm{icomaj})}(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n =$$

$$\sum_{n=0}^{\infty} \left( \sum_{k=0}^{\infty} \left( \binom{k+1}{k}_q x - \sum_{j=1}^{\infty} \binom{k+jm}{k}_q x^{jm} + \sum_{j=1}^{\infty} \binom{k+jm+1}{k}_q x^{jm+1} \right) t^{k+1} \right)^{*\langle n \rangle}.$$

A sequence of algebraic manipulations yields

$$\sum_{n=0}^{\infty} \frac{A_{12\cdots m,n}^{(\mathrm{ides},\mathrm{icomaj})}(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n = \frac{1}{1-t}$$

$$+ \sum_{k=0}^{\infty} \sum_{n=1}^{\infty} \left( \binom{k+1}{k}_q x - \sum_{j=1}^{\infty} \binom{k+jm}{k}_q x^{jm} + \sum_{j=1}^{\infty} \binom{k+jm+1}{k}_q x^{jm+1} \right)^n t^{k+1}$$

$$= 1 + \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \left( \binom{k+1}{k}_q x - \sum_{j=1}^{\infty} \binom{k+jm}{k}_q x^{jm} + \sum_{j=1}^{\infty} \binom{k+jm+1}{k}_q x^{jm+1} \right)^n t^{k+1}$$

$$= 1 + \sum_{k=0}^{\infty} \left( 1 - \binom{k+1}{k}_q x + \sum_{j=1}^{\infty} \binom{k+jm}{k}_q x^{jm} - \sum_{j=1}^{\infty} \binom{k+jm+1}{k}_q x^{jm+1} \right)^{-1} t^{k+1}$$

$$= 1 + \sum_{k=0}^{\infty} \left( \sum_{j=0}^{\infty} \binom{k+jm}{k}_q x^{jm} - \sum_{j=0}^{\infty} \binom{k+jm+1}{k}_q x^{jm+1} \right)^{-1} t^{k+1}$$

$$= 1 + \sum_{k=1}^{\infty} \left[ \sum_{j=0}^{\infty} \left( \binom{k+jm-1}{k-1}_q x^{jm} - \binom{k+jm}{k-1}_q x^{jm+1} \right) \right]^{-1} t^k,$$

thus completing the proof. $\qquad\square$

Let $A_n(t,q) := \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{des}(\pi)+1} q^{\mathrm{maj}(\pi)}$ for $n \geqslant 1$ and $A_0(t,q) := 1$; these are called *q-Eulerian polynomials* and encode the joint distribution of des and maj over $\mathfrak{S}_n$. Observe that

$$\lim_{m \to \infty} A_{12\cdots m,n}^{(\mathrm{ides},\mathrm{imaj})}(t,q) = \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{ides}(\pi)+1} q^{\mathrm{imaj}(\pi)} = \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{des}(\pi)+1} q^{\mathrm{maj}(\pi)} = A_n(t,q);$$

we can exploit this limit to recover from Theorem 4.2 a classical identity involving $q$-Eulerian polynomials. By taking the limit as $m \to \infty$ of both sides of Theorem 4.2 (b), we obtain

$$\sum_{n=0}^{\infty} \frac{A_n(t,q)}{\prod_{i=0}^{n}(1-tq^i)} x^n = 1 + \sum_{k=1}^{\infty} (1-[k]_q x)^{-1} t^k$$

$$= 1 + \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} [k]_q^n x^n t^k$$

$$= 1 + \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} [k]_q^n t^k x^n,$$

and extracting coefficients of $x^n$ yields the famous *Carlitz identity* [26, Corollary 6.1]

$$\frac{A_n(t,q)}{\prod_{i=0}^{n}(1-tq^i)} = \sum_{k=1}^{\infty} [k]_q^n t^k.$$

Next, we have the following formulas for the polynomials $A^{\mathrm{ides}}_{12\cdots m,n}(s,t)$ and $A^{\mathrm{ides}}_{12\cdots m,n}(t)$.

THEOREM 4.3. *Let* $m \geqslant 2$. *Then*

(a) $\displaystyle\sum_{n=0}^{\infty} \frac{A^{\mathrm{ides}}_{12\cdots m,n}(s,t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)^2} + \frac{(s-1)tz^m}{(1-t)(1-(s-1)\sum_{l=1}^{m-1} z^l)} \right)^{*\langle n \rangle},$

(b) $\displaystyle\sum_{n=0}^{\infty} \frac{A^{\mathrm{ides}}_{12\cdots m,n}(t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tz(1-z^{m-1})}{(1-t)(1-z^m)} \right)^{*\langle n \rangle},$

*and*

(c) $\displaystyle\sum_{n=0}^{\infty} \frac{A^{\mathrm{ides}}_{12\cdots m,n}(t)}{(1-t)^{n+1}} x^n =$

$$1 + \sum_{k=1}^{\infty} \left[ \sum_{j=0}^{\infty} \left( \binom{k+jm-1}{k-1} x^{jm} - \binom{k+jm}{k-1} x^{jm+1} \right) \right]^{-1} t^k,$$

*where* $z = x/(1-t)$.

*Proof.* Part (a) follows immediately from Theorem 3.8 and Lemma 4.1, and part (c) is obtained from substituting $q = 1$ into Theorem 4.2 (b). Then taking $s = 0$ in part (a), we have

$$\sum_{n=0}^{\infty} \frac{A^{\mathrm{ides}}_{12\cdots m,n}(t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)^2} - \frac{tz^m}{(1-t)(1+z+z^2+\cdots+z^{m-1})} \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{tz}{1-t} - \frac{tz^m(1-z)}{(1-t)(1-z^m)} \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{tz(1-z^{m-1})}{(1-t)(1-z^m)} \right)^{*\langle n \rangle}$$

which proves part (b). $\square$

We use Theorem 4.3 to compute the first ten polynomials $A^{\mathrm{ides}}_{12\cdots m,n}(t)$ for $m = 3$ and $m = 4$, which are displayed in Tables 1–2. (By Proposition 2.7 (d), the polynomials $A^{\mathrm{ides}}_{m\cdots 21,n}(t)$ are the same as the $A^{\mathrm{ides}}_{12\cdots m,n}(t)$ but with the order of their coefficients reversed.)

Curiously, beginning with $n = 3$, the quadratic coefficient of $A^{\mathrm{ides}}_{123,n}(t)$ alternates between 4 and 5. We state this observation in the following proposition.

PROPOSITION 4.4. *Let* $n \geqslant 3$. *The number of permutations* $\pi$ *in* $\mathfrak{S}_n(123)$ *with* $\mathrm{ides}(\pi) = 1$ *is 4 if* $n$ *is odd, and is 5 if* $n$ *is even.*

Proposition 4.4 can be proven using Theorem 4.3, but we shall instead sketch a combinatorial proof, which is more enlightening. Our proof relies on the notion of reading sequences of permutations. Given a permutation $\pi \in \mathfrak{S}_n$, we read the letters $1, 2, \ldots, n$ in $\pi$ from left-to-right in order, going back to the beginning of $\pi$ when

| $n$ | $A_{123,n}^{\text{ides}}(t)$ |
|---|---|
| 0 | $1$ |
| 1 | $t$ |
| 2 | $t + t^2$ |
| 3 | $4t^2 + t^3$ |
| 4 | $5t^2 + 11t^3 + t^4$ |
| 5 | $4t^2 + 39t^3 + 26t^4 + t^5$ |
| 6 | $5t^2 + 91t^3 + 195t^4 + 57t^5 + t^6$ |
| 7 | $4t^2 + 193t^3 + 904t^4 + 795t^5 + 120t^6 + t^7$ |
| 8 | $5t^2 + 396t^3 + 3420t^4 + 6400t^5 + 2889t^6 + 247t^7 + t^8$ |
| 9 | $4t^2 + 761t^3 + 11610t^4 + 39275t^5 + 37450t^6 + 9774t^7 + 502t^8 + t^9$ |

TABLE 1. Distribution of ides over $\mathfrak{S}_n(123)$

| $n$ | $A_{1234,n}^{\text{ides}}(t)$ |
|---|---|
| 0 | $1$ |
| 1 | $t$ |
| 2 | $t + t^2$ |
| 3 | $t + 4t^2 + t^3$ |
| 4 | $11t^2 + 11t^3 + t^4$ |
| 5 | $18t^2 + 66t^3 + 26t^4 + t^5$ |
| 6 | $28t^2 + 254t^3 + 302t^4 + 57t^5 + t^6$ |
| 7 | $40t^2 + 814t^3 + 2160t^4 + 1191t^5 + 120t^6 + t^7$ |
| 8 | $64t^2 + 2358t^3 + 12030t^4 + 14340t^5 + 4293t^6 + 247t^7 + t^8$ |
| 9 | $96t^2 + 6538t^3 + 57804t^4 + 127250t^5 + 82102t^6 + 14608t^7 + 502t^8 + t^9$ |

TABLE 2. Distribution of ides over $\mathfrak{S}_n(1234)$

necessary; this process realizes $\pi$ as a shuffle of *reading sequences* [31, p. 37]. For example, take $\pi = 748361259$; then the reading sequences of $\pi$ are 12, 3, 45, 6, and 789. It is easy to see that the lengths of reading sequences of $\pi$ are precisely the lengths of the increasing runs of $\pi^{-1}$. Thus, the inverse of a permutation $\pi$ avoids $12\cdots m$ if and only if every reading sequence of $\pi$ has length less than $m$.

*Proof.* The number of permutations $\pi$ in $\mathfrak{S}_n(123)$ with $\text{ides}(\pi) = 1$ is equal to the number of permutations $\pi$ with $\text{des}(\pi) = 1$ whose inverse $\pi^{-1}$ is in $\mathfrak{S}_n(123)$, so it suffices to prove the result for the latter family of permutations. More specifically, we claim that if $n$ is odd, then the permutations $\pi$ in $\mathfrak{S}_n$ with $\text{des}(\pi) = 1$ whose inverse $\pi^{-1}$ is in $\mathfrak{S}_n(123)$ are

- $13\cdots n24\cdots(n-1)$,
- $24\cdots(n-1)13\cdots n$,
- $24\cdots(n-1)n13\cdots(n-2)$, and
- $35\cdots n124\cdots(n-1)$.

Moreover, if $n$ is even, then the permutations $\pi$ in $\mathfrak{S}_n$ with $\text{des}(\pi) = 1$ whose inverse $\pi^{-1}$ is in $\mathfrak{S}_n(123)$ are

- $13\cdots(n-1)24\cdots n$
- $13\cdots(n-1)n24\cdots(n-2)$,

- $24\cdots n13\cdots(n-1)$,
- $35\cdots(n-1)124\cdots n$, and
- $35\cdots(n-1)n124\cdots(n-2)$.

Clearly, all of these permutations have exactly one descent and the lengths of their reading sequences are all less than 3. A careful case analysis shows that these are the only permutations in $\mathfrak{S}_n$ with these properties; we omit the details. $\qquad\square$

4.3. Monotone patterns and inverse peak number. Next, we proceed to the polynomials $P^{\mathrm{ipk}}_{12\cdots m,n}(s,t)$ and $P^{\mathrm{ipk}}_{12\cdots m,n}(t)$, which are equal to the polynomials $P^{\mathrm{ipk}}_{m\cdots 21,n}(s,t)$ and $P^{\mathrm{ipk}}_{m\cdots 21,n}(t)$ by Proposition 2.7 (e).

THEOREM 4.5. *Let* $m \geqslant 2$. *We have*

(a) $\qquad \dfrac{1}{1-t} + \dfrac{1+t}{2(1-t)} \displaystyle\sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(s,u)z^n$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} + \frac{2t(s-1)z^m}{(1-t^2)(1-(s-1)\sum_{l=1}^{m-1} z^l)} \right)^{*\langle n\rangle},$$

(b) $\qquad \dfrac{1}{1-t} + \dfrac{1+t}{2(1-t)} \displaystyle\sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(u)z^n = \sum_{n=0}^{\infty} \left( \frac{2tz(1-z^{m-1})}{(1-t^2)(1-z^m)} \right)^{*\langle n\rangle},$

*and*

(c) $\qquad \dfrac{1}{1-t} + \dfrac{1+t}{2(1-t)} \displaystyle\sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(u)z^n$

$$= 1 + \sum_{k=1}^{\infty} \left[ 1 - 2kx + \sum_{j=1}^{\infty} (c_{m,j,k}x^{jm} - c'_{m,j,k}x^{jm+1}) \right]^{-1} t^k,$$

*where* $u = 4t/(1+t)^2$, $z = (1+t)x/(1-t)$, *and*

$$c_{m,j,k} = 2\sum_{l=1}^{k} \binom{l+jm-1}{l-1}\binom{jm-1}{k-l} \quad \text{and} \quad c'_{m,j,k} = 2\sum_{l=1}^{k} \binom{l+jm}{l-1}\binom{jm}{k-l}.$$

*Proof.* Part (a) follows immediately from Theorem 3.9 and Lemma 4.1. Setting $s = 0$ in part (a), we obtain

$$\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(u)z^n$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} - \frac{2tz^m}{(1-t^2)(1+z+z^2+\cdots+z^{m-1})} \right)^{*\langle n\rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tz}{1-t^2} - \frac{2tz^m(1-z)}{(1-t^2)(1-z^m)} \right)^{*\langle n\rangle} \qquad (11)$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tz(1-z^{m-1})}{(1-t^2)(1-z^m)} \right)^{*\langle n\rangle},$$

thus proving part (b). For part (c), we shall use the well-known identities

$$\frac{1}{(1-t)^{n+1}} = \sum_{k=0}^{\infty} \binom{n+k}{k}t^k \qquad \text{and} \qquad (1+t)^n = \sum_{k=0}^{n} \binom{n}{k}t^k,$$

which imply

$$\frac{2t(1+t)^{jm-1}}{(1-t)^{jm+1}} = \sum_{k=1}^{\infty} c_{m,j,k} t^k \qquad \text{and} \qquad \frac{2t(1+t)^{jm}}{(1-t)^{jm+2}} = \sum_{k=1}^{\infty} c'_{m,j,k} t^k.$$

Then, continuing from (11), we have

$$\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(u) z^n$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} - \frac{2tz^m(1-z)}{(1-t^2)(1-z^m)} \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} - \frac{2t}{1-t^2} \sum_{j=1}^{\infty} (z^{jm} - z^{jm+1}) \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} - \sum_{j=1}^{\infty} \left( \frac{2t(1+t)^{jm-1} x^{jm}}{(1-t)^{jm+1}} - \frac{2t(1+t)^{jm} x^{jm+1}}{(1-t)^{jm+2}} \right) \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \sum_{k=1}^{\infty} 2kxt^k - \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (c_{m,j,k} x^{jm} - c'_{m,j,k} x^{jm+1}) t^k \right)^{*\langle n \rangle}$$

$$= \sum_{n=0}^{\infty} \left( \sum_{k=1}^{\infty} \left( 2kx - \sum_{j=1}^{\infty} (c_{m,j,k} x^{jm} - c'_{m,j,k} x^{jm+1}) \right) t^k \right)^{*\langle n \rangle}$$

$$= \frac{1}{1-t} + \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \left( 2kx - \sum_{j=1}^{\infty} (c_{m,j,k} x^{jm} - c'_{m,j,k} x^{jm+1}) \right)^n t^k$$

$$= 1 + \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \left( 2kx - \sum_{j=1}^{\infty} (c_{m,j,k} x^{jm} - c'_{m,j,k} x^{jm+1}) \right)^n t^k$$

$$= 1 + \sum_{k=1}^{\infty} \left[ 1 - 2kx + \sum_{j=1}^{\infty} (c_{m,j,k} x^{jm} - c'_{m,j,k} x^{jm+1}) \right]^{-1} t^k;$$

this completes the proof. $\qquad \square$

In order to use Theorem 4.5 to compute the polynomials $P^{\mathrm{ipk}}_{12\cdots m,n}(s,t)$ and $P^{\mathrm{ipk}}_{12\cdots m,n}(t)$, one must "invert" the expression $u = 4t/(1+t)^2$. Let us first replace the variable $t$ with $v$, and $u$ with $t$, to obtain $t = 4v/(1+v)^2$. Then, solving $t = 4v/(1+v)^2$ for $v$ yields $v = 2t^{-1}(1 - \sqrt{1-t}) - 1$. Thus, Theorem 4.5 (b) is equivalent to

$$\frac{1}{1-v} + \frac{1+v}{2(1-v)} \sum_{n=1}^{\infty} P^{\mathrm{ipk}}_{12\cdots m,n}(t) z^n = \sum_{n=0}^{\infty} \left( \frac{2tz(1-z^{m-1})}{(1-t^2)(1-z^m)} \right)^{*\langle n \rangle} \Bigg|_{t \mapsto v}$$

where $z = (1+t)x/(1-t)$ and $v = 2t^{-1}(1 - \sqrt{1-t}) - 1$. (Note that substitution does not commute with Hadamard product, so we cannot simply replace $t$ with $v$ inside the Hadamard product.) With some additional algebraic manipulations, we get the

| $n$ | $P_{123,n}^{\mathrm{ipk}}(t)$ | $n$ | $P_{123,n}^{\mathrm{ipk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $8t + 52t^2 + 10t^3$ |
| 1 | $t$ | 6 | $13t + 200t^2 + 136t^3$ |
| 2 | $2t$ | 7 | $21t + 714t^2 + 1170t^3 + 112t^4$ |
| 3 | $3t + 2t^2$ | 8 | $34t + 2468t^2 + 8180t^3 + 2676t^4$ |
| 4 | $5t + 12t^2$ | 9 | $55t + 8348t^2 + 50786t^3 + 37978t^4 + 2210t^5$ |

Table 3. Distribution of ipk over $\mathfrak{S}_n(123)$

| $n$ | $P_{1234,n}^{\mathrm{ipk}}(t)$ | $n$ | $P_{1234,n}^{\mathrm{ipk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $13t + 82t^2 + 16t^3$ |
| 1 | $t$ | 6 | $24t + 364t^2 + 254t^3$ |
| 2 | $2t$ | 7 | $44t + 1502t^2 + 2553t^3 + 248t^4$ |
| 3 | $4t + 2t^2$ | 8 | $81t + 5976t^2 + 20436t^3 + 6840t^4$ |
| 4 | $7t + 16t^2$ | 9 | $149t + 23286t^2 + 146636t^3 + 112192t^4 + 6638t^5$ |

Table 4. Distribution of ipk over $\mathfrak{S}_n(1234)$

formula

$$\sum_{n=1}^{\infty} P_{12\cdots m,n}^{\mathrm{ipk}}(t)x^n = \frac{2(1-v)}{1+v}\sum_{n=0}^{\infty}\left(\frac{2tz(1-z^{m-1})}{(1-t^2)(1-z^m)}\right)^{*\langle n\rangle}\Bigg|_{x\mapsto (1-t)x/(1+t),\, t\mapsto v} - \frac{2}{1+v}$$

where $z$ and $v$ are the same as above; this formula can be used to compute the polynomials $P_{12\cdots m,n}^{\mathrm{ipk}}(t)$. We can carry out a similar process with Theorem 4.5 (a) and (c), as well as with Theorems 4.7, 4.8, 5.3, and 5.4 appearing later in this paper.

Tables 3–4 list the first ten polynomials $P_{12\cdots m,n}^{\mathrm{ipk}}(t)$ for $m = 3$ and $m = 4$.

The linear coefficients of $P_{123,n}^{\mathrm{ipk}}(t)$ are Fibonacci numbers [30, A000045], and those of $P_{1234,n}^{\mathrm{ipk}}(t)$ are tribonacci numbers [30, A000073].[4] In fact, we can make a more general statement relating the linear coefficients of $P_{123,n}^{\mathrm{ipk}}(t)$ to "higher-order" Fibonacci numbers. The *Fibonacci sequence of order $k$* (also called the *$k$-generalized Fibonacci sequence*) $\{f_n^{(k)}\}_{n\geqslant 0}$ is defined by the recursion

$$f_n^{(k)} := f_{n-1}^{(k)} + f_{n-2}^{(k)} + \cdots + f_{n-k}^{(k)}$$

with $f_0^{(k)} := 1$ (and where we treat $f_n^{(k)}$ as 0 for $n < 0$). Hence, the Fibonacci sequence of order two is the usual Fibonacci sequence, and the Fibonacci sequence of order three is the tribonacci sequence. We give three proofs of the following claim in [34].

Claim 4.6. Let $n \geqslant 1$ and $m \geqslant 3$. The number of permutations $\pi$ in $\mathfrak{S}_n(12\cdots m)$ with $\mathrm{ipk}(\pi) = 0$ is equal to the $(m-1)$th order Fibonacci number $f_n^{(m-1)}$.

4.4. Monotone patterns and inverse left peak number. Finally, we produce analogous formulas for the inverse left peak polynomials $P_{12\cdots m,n}^{\mathrm{ilpk}}(s,t)$ and $P_{12\cdots m,n}^{\mathrm{ilpk}}(t)$. We omit the proofs of these formulas, as they follow essentially the same steps as the proof of Theorem 4.5.

---

[4] Note that [30] uses a different indexing for these sequences.

| $n$ | $P_{123,n}^{\text{ilpk}}(t)$ | $n$ | $P_{123,n}^{\text{ilpk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $27t + 43t^2$ |
| 1 | 1 | 6 | $63t + 248t^2 + 38t^3$ |
| 2 | $1 + t$ | 7 | $144t + 1225t^2 + 648t^3$ |
| 3 | $5t$ | 8 | $333t + 5591t^2 + 6882t^3 + 552t^4$ |
| 4 | $12t + 5t^2$ | 9 | $765t + 24304t^2 + 58552t^3 + 15756t^4$ |

TABLE 5. Distribution of ilpk over $\mathfrak{S}_n(123)$

| $n$ | $P_{1234,n}^{\text{ilpk}}(t)$ | $n$ | $P_{1234,n}^{\text{ilpk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $50t + 61t^2$ |
| 1 | 1 | 6 | $138t + 443t^2 + 61t^3$ |
| 2 | $1 + t$ | 7 | $378t + 2659t^2 + 1289t^3$ |
| 3 | $1 + 5t$ | 8 | $1042t + 14501t^2 + 16524t^3 + 1266t^4$ |
| 4 | $18t + 5t^2$ | 9 | $2866t + 74941t^2 + 167780t^3 + 43314t^4$ |

TABLE 6. Distribution of ilpk over $\mathfrak{S}_n(1234)$

THEOREM 4.7. *Let $m \geqslant 2$. We have*

(a) $$\frac{1}{1-t} \sum_{n=0}^{\infty} P_{12\cdots m,n}^{\text{ilpk}}(s,u)z^n = \sum_{n=0}^{\infty} \left( \frac{z}{1-t} + \frac{(s-1)z^m}{(1-t)(1-(s-1)\sum_{l=1}^{m-1} z^l)} \right)^{*\langle n \rangle},$$

(b) $$\frac{1}{1-t} \sum_{n=0}^{\infty} P_{12\cdots m,n}^{\text{ilpk}}(u)z^n = \sum_{n=0}^{\infty} \left( \frac{z(1-z^{m-1})}{(1-t)(1-z^m)} \right)^{*\langle n \rangle},$$

*and*

(c) $$\frac{1}{1-t} \sum_{n=0}^{\infty} P_{12\cdots m,n}^{\text{ilpk}}(u)z^n = \sum_{k=0}^{\infty} \left[ \sum_{j=0}^{\infty} (d_{m,j,k} x^{jm} - d'_{m,j,k} x^{jm+1}) \right]^{-1} t^k,$$

*where $u = 4t/(1+t)^2$, $z = (1+t)x/(1-t)$, and*

$$d_{m,j,k} = \sum_{l=0}^{k} \binom{l+jm}{l} \binom{jm}{k-l} \qquad and \qquad d'_{m,j,k} = \sum_{l=0}^{k} \binom{l+jm+1}{l} \binom{jm+1}{k-l}.$$

In Tables 5–6, we display the first ten polynomials $P_{12\cdots m,n}^{\text{ilpk}}(t)$ for $m = 3$ and $m = 4$.

Unfortunately, we cannot use symmetries to translate Theorem 4.7 into a result about the pattern $m \cdots 21$. However, we can obtain an analogous result for the polynomials $P_{m\cdots 21,n}^{\text{ilpk}}(s,t)$ and $P_{m\cdots 21,n}^{\text{ilpk}}(t)$ separately; this is given below. The proof is omitted but follows the same general line of reasoning as in the previous few theorems. We only note that the underlying permutation of any $m \cdots 21$-cluster has exactly one left peak (rather than having no left peaks as for $12 \cdots m$-clusters), which results in the generating function

$$\sum_{k=2}^{\infty} R_{m\cdots 21,k}^{\text{ilpk}}(s,t)x^k = \frac{stx^m}{1 - s\sum_{l=1}^{m-1} x^l}$$

for the refined cluster polynomials $R_{m\cdots 21,n}^{\text{ilpk}}(s,t)$.

| $n$ | $P^{\text{ilpk}}_{321,n}(t)$ | $n$ | $P^{\text{ilpk}}_{321,n}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $1 + 37t + 32t^2$ |
| 1 | 1 | 6 | $1 + 101t + 222t^2 + 25t^3$ |
| 2 | $1 + t$ | 7 | $1 + 269t + 1251t^2 + 496t^3$ |
| 3 | $1 + 4t$ | 8 | $1 + 710t + 6349t^2 + 5899t^3 + 399t^4$ |
| 4 | $1 + 13t + 3t^2$ | 9 | $1 + 1865t + 30186t^2 + 54825t^3 + 12500t^4$ |

TABLE 7. Distribution of ilpk over $\mathfrak{S}_n(321)$

| $n$ | $P^{\text{ilpk}}_{4321,n}(t)$ | $n$ | $P^{\text{ilpk}}_{4321,n}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $1 + 53t + 57t^2$ |
| 1 | 1 | 6 | $1 + 158t + 428t^2 + 55t^3$ |
| 2 | $1 + t$ | 7 | $1 + 462t + 2668t^2 + 1195t^3$ |
| 3 | $1 + 5t$ | 8 | $1 + 1342t + 15074t^2 + 15765t^3 + 1151t^4$ |
| 4 | $1 + 17t + 5t^2$ | 9 | $1 + 3886t + 80338t^2 + 164337t^3 + 40339t^4$ |

TABLE 8. Distribution of ilpk over $\mathfrak{S}_n(4321)$

THEOREM 4.8. *Let $m \geqslant 2$. We have*

(a)
$$\frac{1}{1-t} \sum_{n=0}^{\infty} P^{\text{ilpk}}_{m\cdots 21,n}(s,u) z^n$$
$$= \sum_{n=0}^{\infty} \left( \frac{z}{1-t} + \frac{4t(s-1)z^m}{(1-t^2)(1+t)(1-(s-1)\sum_{l=1}^{m-1} z^l)} \right)^{*\langle n \rangle},$$

(b)
$$\frac{1}{1-t} \sum_{n=0}^{\infty} P^{\text{ilpk}}_{m\cdots 21,n}(u) z^n = \sum_{n=0}^{\infty} \left( \frac{(1+t)^2 z - 4tz^m - (1-t)^2 z^{m+1}}{(1-t^2)(1+t)(1-z^m)} \right)^{*\langle n \rangle},$$

*and*

(c)
$$\frac{1}{1-t} \sum_{n=0}^{\infty} P^{\text{ilpk}}_{m\cdots 21,n}(u) z^n$$
$$= \frac{1}{1-x} + \sum_{k=1}^{\infty} \left[ 1 - (2k+1)x + \sum_{j=1}^{\infty} (e_{m,j,k} x^{jm} - e'_{m,j,k} x^{jm+1}) \right]^{-1} t^k,$$

*where $u = 4t/(1+t)^2$, $z = (1+t)x/(1-t)$, and*

$$e_{m,j,k} = 4 \sum_{l=1}^{k} \binom{l+jm-1}{l-1} \binom{jm-2}{k-l} \quad \text{and} \quad e'_{m,j,k} = 4 \sum_{l=1}^{k} \binom{l+jm}{l-1} \binom{jm-1}{k-l}.$$

The first ten polynomials $P^{\text{ilpk}}_{m\cdots 21,n}(t)$ for $m = 3$ and $m = 4$ are given in Tables 7–8. The linear coefficients of the $P^{\text{ilpk}}_{321,n}(t)$ match OEIS sequence A080145 [30, A080145], which involves the Fibonacci numbers $f_n := f_n^{(2)}$. We give two proofs of this claim in [34].
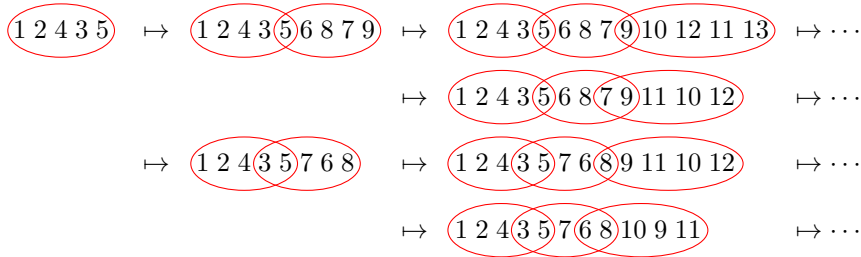
FIGURE 2. 12435-clusters

CLAIM 4.9. Let $n \geqslant 1$. The number of permutations $\pi$ in $\mathfrak{S}_n(321)$ with $\mathrm{ilpk}(\pi) = 1$ is equal to

$$\sum_{i=1}^{n-1} \sum_{j=1}^{i} f_{j-1} f_j = f_{n-1} f_n - \left\lfloor \frac{n+1}{2} \right\rfloor .$$

## 5. TRANSPOSITIONAL PATTERNS $12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$

### 5.1. CLUSTER GENERATING FUNCTIONS FOR TRANSPOSITIONAL PATTERNS.
In this section, we turn our attention to patterns of the form $\sigma = 12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$. These are precisely the elementary transpositions $(a, a+1)$ of $\mathfrak{S}_m$—aside from the transpositions $(1,2)$ and $(m-1, m)$—and form another family of patterns for which it is straightforward to obtain closed-form generating functions for our refined cluster polynomials.

LEMMA 5.1. *Let $\sigma = 12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$. Let $i = \min(a, m-a)$. Then*

$$\sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ides}}(s,t) x^k = \sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ipk}}(s,t) x^k = \frac{st^2 x^m}{1 - st \sum_{l=1}^{i} x^{m-l}}$$

*and*

$$\sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ilpk}}(s,t) x^k = \frac{stx^m}{1 - st \sum_{l=1}^{i} x^{m-l}} .$$

*Proof.* Let us first assume that $m - a \leqslant a$; then the overlap set of $\sigma$ is given by $O_\sigma = \{a, a+1, \ldots, m-1\}$. In this case, we can uniquely generate $\sigma$-clusters by first taking the permutation $\sigma$ and then repeatedly appending the next $l$ largest integers (where $a \leqslant l \leqslant m-1$) in the order compatible with the pattern $\sigma$—each iteration creates an additional marked occurrence of $\sigma$. See Figure 2 for an illustration in the case $m = 5$ and $a = 3$. Thus, we have the formula

$$\sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{\sigma,\pi}} s^{\mathrm{mk}_\sigma(c)} x^k = \frac{sx^m}{1 - s(x^a + x^{a+1} + \cdots + x^{m-1})}$$

when $m - a \leqslant a$ (see also [10, p. 357]).

When $a \leqslant m - a$, we instead have $O_\sigma = \{m - a, m - a + 1, \ldots, m - 1\}$, and we can uniquely generate clusters in essentially the same way as above; the only differences are that (1) $m - a \leqslant l \leqslant m - 1$, and (2) whenever we append $l = m - a$ integers to create a larger cluster, we increase the last entry of the existing cluster by 1 so that
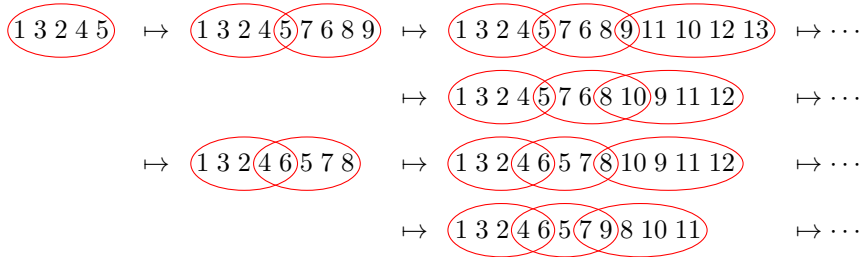
FIGURE 3. 13245-clusters

the new cluster contains an additional occurrence of $\sigma$. See Figure 3 for an illustration in the case $m = 5$ and $a = 2$. Thus, we have the formula

$$\sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{\sigma,\pi}} s^{\mathrm{mk}_\sigma(c)} x^k = \frac{sx^m}{1 - s(x^{m-a} + x^{m-a+1} + \cdots + x^{m-1})}$$

when $a \leqslant m - a$. In either case, we have

$$\sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{\sigma,\pi}} s^{\mathrm{mk}_\sigma(c)} x^k = \frac{sx^m}{1 - s(x^{m-i} + x^{m-i+1} + \cdots + x^{m-1})} = \frac{sx^m}{1 - s\sum_{l=1}^{i} x^{m-l}}$$

where $i = \min(a, m - a)$.

It is clear that the underlying permutation of any $\sigma$-cluster is a product of disjoint elementary transpositions, hence an involution. So, whenever $\pi$ is the underlying permutation of a $\sigma$-cluster, we have $\mathrm{des}(\pi) = \mathrm{ides}(\pi)$, $\mathrm{pk}(\pi) = \mathrm{ipk}(\pi)$, and $\mathrm{lpk}(\pi) = \mathrm{ilpk}(\pi)$. Each marked occurrence of a $\sigma$-cluster $c$ contributes exactly one descent, which is also a peak and a left peak. Hence, we have

$$\sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ides}}(s,t) x^k = \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} t^{\mathrm{ides}(\pi)+1} \sum_{c \in C_{12\cdots m,\pi}} s^{\mathrm{mk}_{12\cdots m}(c)} x^k$$

$$= t \sum_{k=2}^{\infty} \sum_{\pi \in \mathfrak{S}_k} \sum_{c \in C_{12\cdots m,\pi}} (st)^{\mathrm{mk}_{12\cdots m}(c)} x^k$$

$$= \frac{st^2 x^m}{1 - st \sum_{l=1}^{i} x^{m-l}}.$$

Our formulas for $\sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ipk}}(s,t) x^k$ and $\sum_{k=2}^{\infty} R_{\sigma,k}^{\mathrm{ilpk}}(s,t) x^k$ are obtained in the same way. $\qquad\square$

5.2. TRANSPOSITIONAL PATTERNS AND INVERSE DESCENT NUMBER. We now apply our results from Section 3.5 to the patterns $\sigma = 12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$ for arbitrary $m \geqslant 5$ and $2 \leqslant a \leqslant m - 2$. All of these formulas follow immediately from combining either Theorem 3.8, 3.9, or 3.10 with Lemma 5.1, and then setting $s = 0$.

THEOREM 5.2. *Let* $\sigma = 12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$ *where* $m \geqslant 5$ *and* $2 \leqslant a \leqslant m - 2$. *Let* $i = \min(a, m-a)$. *We have*

(a) $\displaystyle \sum_{n=0}^{\infty} \frac{A_{\sigma,n}^{\mathrm{ides}}(s,t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)^2} + \frac{(s-1)t^2 z^m}{(1-t)(1 - (s-1)t \sum_{l=1}^{i} z^{m-l})} \right)^{*\langle n \rangle}$

| $n$ | $A^{\text{ides}}_{13245,n}(t)$ |
|---|---|
| 0 | $1$ |
| 1 | $t$ |
| 2 | $t + t^2$ |
| 3 | $t + 4t^2 + t^3$ |
| 4 | $t + 11t^2 + 11t^3 + t^4$ |
| 5 | $t + 25t^2 + 66t^3 + 26t^4 + t^5$ |
| 6 | $t + 53t^2 + 294t^3 + 302t^4 + 57t^5 + t^6$ |
| 7 | $t + 108t^2 + 1125t^3 + 2368t^4 + 1191t^5 + 120t^6 + t^7$ |
| 8 | $t + 215t^2 + 3934t^3 + 14923t^4 + 15363t^5 + 4293t^6 + 247t^7 + t^8$ |
| 9 | $t + 422t^2 + 12985t^3 + 82066t^4 + 150240t^5 + 86954t^6 + 14608t^7 + 502t^8 + t^9$ |

TABLE 9. Distribution of ides over $\mathfrak{S}_n(13245)$

*and*

(b)
$$\sum_{n=0}^{\infty} \frac{A^{\text{ides}}_{\sigma,n}(t)}{(1-t)^{n+1}} x^n = \sum_{n=0}^{\infty} \left( \frac{tx}{(1-t)^2} - \frac{t^2 z^m}{(1-t)(1 + t\sum_{l=1}^{i} z^{m-l})} \right)^{*\langle n \rangle}$$

*where $z = x/(1-t)$.*

We use Theorem 5.2 to compute the first ten polynomials $A^{\text{ides}}_{13245,n}(t)$; see Table 9. By the symmetry present in Theorem 5.2, the polynomials $A^{\text{ides}}_{13245,n}(t)$ displayed above are the same as the polynomials $A^{\text{ides}}_{12435,n}(t)$ for corresponding $n$.

## 5.3. TRANSPOSITIONAL PATTERNS AND INVERSE PEAK NUMBER.

THEOREM 5.3. *Let $\sigma = 12\cdots(a-1)(a+1)a(a+2)(a+3)\cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$. Let $i = \min(a, m-a)$. Then*

(a)
$$\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P^{\text{ipk}}_{\sigma,n}(s,u) z^n$$
$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} + \frac{(1+t)(s-1)u^2 z^m}{2(1-t)(1-(s-1)u\sum_{l=1}^{i} z^{m-l})} \right)^{*\langle n \rangle}$$

*and*

(b)
$$\frac{1}{1-t} + \frac{1+t}{2(1-t)} \sum_{n=1}^{\infty} P^{\text{ipk}}_{\sigma,n}(u) z^n$$
$$= \sum_{n=0}^{\infty} \left( \frac{2tx}{(1-t)^2} - \frac{(1+t)u^2 z^m}{2(1-t)(1+u\sum_{l=1}^{i} z^{m-l})} \right)^{*\langle n \rangle}$$

*where $u = 4t/(1+t)^2$ and $z = (1+t)x/(1-t)$.*

We give the first ten polynomials $P^{\text{ipk}}_{13245,n}(t)$—which are also the first ten polynomials $P^{\text{ipk}}_{12435,n}(t)$—in Table 10.

| $n$ | $P_{13245,n}^{\text{ipk}}(t)$ | $n$ | $P_{13245,n}^{\text{ipk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $16t + 87t^2 + 16t^3$ |
| 1 | $t$ | 6 | $32t + 408t^2 + 268t^3$ |
| 2 | $2t$ | 7 | $64t + 1776t^2 + 2808t^3 + 266t^4$ |
| 3 | $4t + 2t^2$ | 8 | $128t + 7424t^2 + 23745t^3 + 7680t^4$ |
| 4 | $8t + 16t^2$ | 9 | $256t + 30336t^2 + 178029t^3 + 131542t^4 + 7616t^5$ |

TABLE 10. Distribution of ipk over $\mathfrak{S}_n(13245)$

### 5.4. TRANSPOSITIONAL PATTERNS AND INVERSE LEFT PEAK NUMBER.

THEOREM 5.4. *Let* $\sigma = 12\cdots(a-1)(a+1)a(a+2)(a+3)\cdots m$ *where* $m \geqslant 5$ *and* $2 \leqslant a \leqslant m-2$. *Let* $i = \min(a, m-a)$. *Then*

(a)
$$\frac{1}{1-t}\sum_{n=0}^{\infty} P_{\sigma,n}^{\text{ilpk}}(s,u)z^n = \sum_{n=0}^{\infty}\left(\frac{z}{1-t} + \frac{(s-1)uz^m}{(1-t)(1-(s-1)u\sum_{l=1}^{i} z^{m-l})}\right)^{*\langle n\rangle}$$

*and*

(b)
$$\frac{1}{1-t}\sum_{n=0}^{\infty} P_{\sigma,n}^{\text{ilpk}}(u)z^n = \sum_{n=0}^{\infty}\left(\frac{z}{1-t} - \frac{uz^m}{(1-t)(1+u\sum_{l=1}^{i} z^{m-l})}\right)^{*\langle n\rangle}$$

*where* $u = 4t/(1+t)^2$ *and* $z = (1+t)x/(1-t)$.

Table 11 lists the first ten polynomials $P_{13245,n}^{\text{ilpk}}(t) = P_{12435,n}^{\text{ilpk}}(t)$.

| $n$ | $P_{13245,n}^{\text{ilpk}}(t)$ | $n$ | $P_{13245,n}^{\text{ilpk}}(t)$ |
|---|---|---|---|
| 0 | 1 | 5 | $1 + 57t + 61t^2$ |
| 1 | 1 | 6 | $1 + 173t + 473t^2 + 61t^3$ |
| 2 | $1 + t$ | 7 | $1 + 516t + 3030t^2 + 1367t^3$ |
| 3 | $1 + 5t$ | 8 | $1 + 1528t + 17551t^2 + 18536t^3 + 1361t^4$ |
| 4 | $1 + 18t + 5t^2$ | 9 | $1 + 4511t + 95867t^2 + 198379t^3 + 49021t^4$ |

TABLE 11. Distribution of ilpk over $\mathfrak{S}_n(13245)$

## 6. CONCLUSION

In summary, we have proven a lifting of Elizalde and Noy's adaptation of the Goulden–Jackson cluster method for permutations to the Malvenuto–Reutenauer algebra **FQSym**. By applying two homomorphisms to the cluster method in **FQSym**, we recover both Elizalde and Noy's cluster method and Elizalde's $q$-cluster method as special cases. We have also defined several other homomorphisms, by way of the theory of shuffle-compatibility, which lead to new specializations of our generalized cluster method that keep track of various inverse statistics. Finally, we applied these results to two families of patterns: the monotone patterns $12\cdots m$ and $m\cdots21$, and the transpositional patterns $12\cdots(a-1)(a+1)a(a+2)(a+3)\cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$.

We chose to study monotone patterns as well as the transpositional patterns of the form above because, for these patterns, it is easy to count clusters by the inverse statistics that we consider. In particular, these patterns have two nice properties:

(1) These patterns are *chain patterns*. Elizalde and Noy [10] showed that counting clusters is equivalent to counting linear extensions in a certain poset, and the poset associated with a chain pattern is a chain. This means that if we fix the length of a cluster as well as the positions of the marked occurrences within the cluster, then there is at most one cluster of that length and with that set of positions.

(2) Clusters formed from any one of these patterns are involutions, so counting clusters by ist is the same as counting them by st.

In forthcoming work, joint with Sergi Elizalde and Justin Troyka, we study the transpositional patterns $2134 \cdots m$ and $12 \cdots (m-2)m(m-1)$ for $m \geqslant 3$. Interestingly, the enumeration of $2134 \cdots m$-clusters and $12 \cdots (m-2)m(m-1)$-clusters by ides, ipk, and ilpk turns out to have connections to generalized Stirling permutations [15] and $1/k$-Eulerian polynomials [28]. Although these are not chain patterns and their clusters are not involutions, they are examples of *non-overlapping patterns*: patterns whose overlap set is equal to $\{m-1\}$. Both the non-overlapping condition and the condition of being a chain pattern greatly restrict how clusters can be formed, making them easier to characterize and thus more amenable to study. As such, one direction of future work is to apply our results to other families of non-overlapping patterns and chain patterns.

We also present the following conjecture, which is suggested by computational evidence.

CONJECTURE 6.1. *Let $\sigma$ be $12 \cdots m$ or $m \cdots 21$ where $m \geqslant 3$, or $12 \cdots (a-1)(a+1)a(a+2)(a+3) \cdots m$ where $m \geqslant 5$ and $2 \leqslant a \leqslant m-2$. Then the polynomials $A_{\sigma,n}^{\mathrm{ides}}(t)$, $P_{\sigma,n}^{\mathrm{ipk}}(t)$, and $P_{\sigma,n}^{\mathrm{ilpk}}(t)$ have only real roots for all $n \geqslant 2$.*

In particular, Conjecture 6.1 would imply that—for all patterns $\sigma$ considered in this paper—the polynomials $A_{\sigma,n}^{\mathrm{ides}}(t)$, $P_{\sigma,n}^{\mathrm{ipk}}(t)$, and $P_{\sigma,n}^{\mathrm{ilpk}}(t)$ are unimodal and log-concave, and that the distributions of the statistics ides, ipk, and ilpk over $\mathfrak{S}_n(\sigma)$ converge to a normal distribution as $n \to \infty$. It is worth noting that the Eulerian, peak, and left peak polynomials

$$A_n(t) := \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{des}(\pi)+1} = \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{ides}(\pi)+1},$$

$$P_n^{\mathrm{pk}}(t) := \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{pk}(\pi)+1} = \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{ipk}(\pi)+1}, \text{ and}$$

$$P_n^{\mathrm{lpk}}(t) := \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{lpk}(\pi)} = \sum_{\pi \in \mathfrak{S}_n} t^{\mathrm{ilpk}(\pi)}$$

are all real-rooted (see, e.g. [25, 26, 33]). In light of this fact, one might intuitively expect the polynomials $A_{\sigma,n}^{\mathrm{ides}}(t)$, $P_{\sigma,n}^{\mathrm{ipk}}(t)$, and $P_{\sigma,n}^{\mathrm{ilpk}}(t)$ to be real-rooted as well, since avoiding a single consecutive pattern is not a very restrictive condition (especially when compared to classical pattern avoidance) and therefore might be expected to preserve unimodality or asymptotic normality.

One can use the theory of shuffle-compatibility from [17] to define other homomorphisms on **FQSym** which can be used to count permutations by inverses of shuffle-compatible statistics other than the ones we consider here. For example, the bistatistic (pk, des) is shuffle-compatible, so we can define a homomorphism $\Psi_{(\mathrm{pk},\mathrm{ides})}$ that can be used to produce yet another specialization of our generalized cluster method that simultaneously refines by ipk and ides. Finally, in a different direction, one may apply our homomorphisms to other formulas in **FQSym** which lift classical formulas in permutation enumeration—such as the lifting of André's exponential

generating function considered in [21]—leading to new refinements of these classical formulas by inverse statistics.

## References

[1] Désiré André, *Sur les permutations alternées*, J. Math. Pures Appl. **7** (1881), 167–184.

[2] Nicholas R. Beaton, Andrew R. Conway, and Anthony J. Guttmann, *On consecutive pattern avoiding permutations of length 4, 5 and beyond*, Discrete Math. Theor. Comput. Sci. **19** (2017), no. 2, Paper No. 8, 24 pp.

[3] Harry Crane, Stephen DeSalvo, and Sergi Elizalde, *The probability of avoiding consecutive patterns in the Mallows distribution*, Random Structures Algorithms **53** (2018), no. 3, 417–447.

[4] Theodore Dokos, Tim Dwyer, Bryan P. Johnson, Bruce E. Sagan, and Kimberly Selsor, *Permutation patterns and statistics*, Discrete Math. **312** (2012), no. 18, 2760–2775.

[5] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon, *Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras*, Internat. J. Algebra Comput. **12** (2002), no. 5, 671–717.

[6] Tim Dwyer and Sergi Elizalde, *Wilf equivalence relations for consecutive patterns*, Adv. in Appl. Math. **99** (2018), 134–157.

[7] Sergi Elizalde, *The most and the least avoided consecutive patterns*, Proc. Lond. Math. Soc. (3) **106** (2013), no. 5, 957–979.

[8] ———, *A survey of consecutive patterns in permutations*, in Recent trends in combinatorics, IMA Vol. Math. Appl., vol. 159, Springer, [Cham], 2016, pp. 601–618.

[9] Sergi Elizalde and Marc Noy, *Consecutive patterns in permutations*, Adv. in Appl. Math. **30** (2003), no. 1–2, 110–125.

[10] ———, *Clusters, generating functions and asymptotics for consecutive patterns in permutations*, Adv. in Appl. Math. **49** (2012), no. 3–5, 351–374.

[11] A. M. Garsia and I. Gessel, *Permutation statistics and partitions*, Adv. in Math. **31** (1979), no. 3, 288–305.

[12] Israel M. Gelfand, Daniel Krob, Alain Lascoux, Bernard Leclerc, Vladimir S. Retakh, and Jean-Yves Thibon, *Noncommutative symmetric functions*, Adv. Math. **112** (1995), no. 2, 218–348.

[13] Ira M. Gessel, *Multipartite P-partitions and inner products of skew Schur functions*, Contemp. Math. **34** (1984), 289–317.

[14] ———, *Reciprocals of exponential polynomials and permutation enumeration*, Australas. J. Combin. **74** (2019), 364–370.

[15] ———, *A note on Stirling permutations*, 2020 (originally written in 1978), `https://arxiv.org/abs/2005.04133`.

[16] Ira M. Gessel and Yan Zhuang, *Counting permutations by alternating descents*, Electron. J. Combin. **21** (2014), no. 4, Paper P4.23, 21 pp.

[17] ———, *Shuffle-compatible permutation statistics*, Adv. Math. **332** (2018), 85–141.

[18] Ira Martin Gessel, *Generating Functions and Enumeration of Sequences*, Ph.D. thesis, Massachusetts Institute of Technology, 1977.

[19] I. P. Goulden and D. M. Jackson, *An inversion theorem for cluster decompositions of sequences with distinguished subsequences*, J. London Math. Soc. (2) **20** (1979), no. 3, 567–576.

[20] Darij Grinberg and Victor Reiner, *Hopf Algebras in Combinatorics*, 2020, `https://arxiv.org/abs/1409.8356v7`.

[21] Matthieu Josuat-Vergès, Jean-Christophe Novelli, and Jean-Yves Thibon, *The algebraic combinatorics of snakes*, J. Combin. Theory Ser. A **119** (2012), no. 8, 1613–1638.

[22] Mitchell Lee and Ashwin Sah, *Constraining strong c-Wilf equivalence using cluster poset asymptotics*, Adv. in Appl. Math. **103** (2019), 43–57.

[23] Kurt Luoto, Stefan Mykytiuk, and Stephanie van Willigenburg, *An Introduction to Quasisymmetric Schur Functions. Hopf Algebras, Quasisymmetric Functions, and Young Composition Tableaux.*, Springer, 2013.

[24] Claudia Malvenuto and Christophe Reutenauer, *Duality between quasi-symmetric functions and the Solomon descent algebra*, J. Algebra **177** (1995), no. 3, 967–982.

[25] T. Kyle Petersen, *Enriched P-partitions and peak algebras*, Adv. Math. **209** (2007), no. 2, 561–610.

[26] _____, *Eulerian numbers*, Birkhäuser/Springer, New York, 2015.

[27] Don Rawlings, *The q-exponential generating function for permutations by consecutive patterns and inversions*, J. Combin. Theory Ser. A **114** (2007), no. 1, 184–193.

[28] Carla D. Savage and Gopal Viswanathan, *The 1/k-Eulerian polynomials*, Electron. J. Combin. **19** (2012), no. 1, Paper 9, 21 pp.

[29] Rodica Simion and Frank W. Schmidt, *Restricted permutations*, European J. Combin. **6** (1985), no. 4, 383–406.

[30] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org`.

[31] Richard P. Stanley, *Enumerative Combinatorics, vol. 1*, 2nd ed., Cambridge University Press, 2011.

[32] _____, *Enumerative Combinatorics, vol. 2*, Cambridge University Press, 2001.

[33] John R. Stembridge, *Enriched P-partitions*, Trans. Amer. Math. Soc. **349** (1997), no. 2, 763–788.

[34] Justin M. Troyka and Yan Zhuang, *Fibonacci numbers, consecutive patterns, and inverse peaks*, Adv. in Appl. Math. **141** (2022), Paper No. 102406, 19 pages.

[35] Yan Zhuang, *Counting permutations by runs*, J. Comb. Theory Ser. A **142** (2016), 147–176.

[36] _____, *Eulerian polynomials and descent statistics*, Adv. in Appl. Math. **90** (2017), 86–144.

[37] _____, *A generalized Goulden-Jackson cluster method and lattice path enumeration*, Discrete Math. **341** (2018), no. 2, 358–379.

[38] _____, *Refined consecutive pattern enumeration via a generalized cluster method*, to appear in *Sém. Lothar. Combin.*, 2022.

Yan Zhuang, Davidson College, 405 N. Main St., Davidson, NC 28035 (USA)
*E-mail :* `yazhuang@davidson.edu`