Séminaire Lotharingien de Combinatoire **86B** (2022)
Article #16, 12 pp.

Proceedings of the 34^{th} Conference on Formal Power Series and Algebraic Combinatorics (Bangalore)

Troupes, Cumulants, and Stack-Sorting

Colin Defant*[1]

[1]*Department of Mathematics, Princeton University, Princeton, NJ*

Abstract. In several cases, a sequence of free cumulants that counts certain binary plane trees corresponds to a sequence of classical cumulants that counts the decreasing versions of the same trees. Using two new operations on binary plane trees that we call *insertion* and *decomposition*, we prove that this surprising phenomenon holds for families of trees that we call *troupes*. The proof relies on two new formulas, each of which is given as a sum over objects called *valid hook configurations*. The first of these formulas provides detailed information about the preimages of a permutation under the postorder traversal that lie in a given troupe; the second is a new combinatorial formula that converts from a sequence of free cumulants to the corresponding sequence of classical cumulants. The unexpected connection between troupes and cumulants provides a powerful new tool for analyzing the stack-sorting map s (which is defined via the postorder traversal) that hinges on free probability theory. We give numerous applications of this method. For example, we show that if $\sigma \in S_{n-1}$ is chosen uniformly at random and des denotes the descent statistic, then the expected value of $\mathrm{des}(s(\sigma)) + 1$ is

$$\left(3 - \sum_{j=0}^{n} \frac{1}{j!}\right) n.$$

Furthermore, the variance of $\mathrm{des}(s(\sigma)) + 1$ is asymptotically $(2 + 2e - e^2)n$. We obtain similar results concerning the expected number of descents of postorder readings of decreasing binary plane trees of various types. We also obtain improved estimates for $|s(S_n)|$ and an improved lower bound for the degree of noninvertibility of $s\colon S_n \to S_n$.

Keywords: troupes, cumulants, stack-sorting, free probability, postorder traversal

1 Introduction

Cumulants are fundamental combinatorial tools in noncommutative probability theory that encode different notions of independence. There has been a great deal of recent work aimed at finding combinatorial formulas that convert from cumulants of one type to cumulants of another (see [1, 5, 8] and the references therein). In this extended abstract of the article [5], we provide a new formula, which we call the VHC Cumulant Formula,

*cdefant@princeton.edu. The author was supported by a Fannie and John Hertz Foundation Fellowship and an NSF Graduate Research Fellowship.

that converts from free cumulants to classical cumulants via a sum over objects called *valid hook configurations*. The VHC Cumulant Formula appears to be the first combinatorial cumulant conversion formula with applications beyond the specific combinatorial objects involved in the formula. This is due to the surprising fact that valid hook configurations feature prominently in a different formula concerning postorder traversals of decreasing binary plane trees; we call this other formula the Refined Tree Fertility Formula. Postorder traversals are very closely related to West's stack-sorting map, a specific well-studied combinatorial operator on the set S_n of permutations of $[n] = \{1, \ldots, n\}$. We will also introduce *troupes*, which are families of binary plane trees that are closed under two new operations called *insertion* and *decomposition*, as well as the notion of an *insertion-additive* tree statistic. Many classical families of rooted plane trees found in the literature are troupes, and several interesting statistics are insertion-additive.

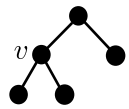
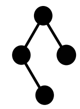
The article [5] has three main aspects. The first, which concerns troupes and cumulants, explains how troupes encode the relationship between (univariate) free and classical cumulants. Indeed, we will show that if $\mathbf{T}$ is a troupe, then a sequence of free cumulants that counts trees in $\mathbf{T}$ according to some insertion-additive tree statistics corresponds to a sequence of classical cumulants that counts decreasing labeled versions of the trees in $\mathbf{T}$ according to the same statistics. The proof of this theorem relies on combining the VHC Cumulant Formula with the Refined Tree Fertility Formula, even though the statement makes no reference to postorder traversals. The second aspect concerns cumulants and stack-sorting. We will outline several instances where tools from combinatorial free probability yield deep facts about the stack-sorting map and, more generally, postorder traversals. Finally, the third aspect, which concerns stack-sorting and troupes, explains how troupes provide a very broad framework for generalizing results about the stack-sorting map. We will not discuss this aspect of the paper here.

The fact that valid hook configurations appear naturally in both the Refined Tree Fertility Formula and the VHC Cumulant Formula is extremely mysterious and is responsible for a great deal of unexpected structure underlying the stack-sorting map.

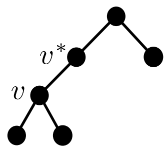
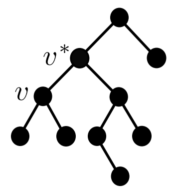
2 Troupes and Tree Traversals

A *binary plane tree* is a rooted tree in which each vertex has at most 2 children and each child is designated as either a left or a right child. Let BPT denote the set of binary plane trees. For $\mathbf{T} \subseteq \mathsf{BPT}$, we let $\mathbf{T}_n$ denote the set of trees in $\mathbf{T}$ that have n vertices.

Let T_1 and T_2 be nonempty binary plane trees, and let v be a vertex of T_1. Let us replace v with two vertices that are connected by a left edge. This produces a new tree T_1^* with one more vertex than T_1. We call the lower endpoint of the new left edge v, identifying it with the original vertex v. We denote the upper endpoint of the new left edge by v^*. The *insertion* of T_2 into T_1 at v, denoted $\nabla_v(T_1, T_2)$, is the tree formed by attaching T_2 as the right subtree of v^* in T_1^*. For instance, if

$T_1 =$  and $T_2 =$  ,

where v is as indicated, then we have

$T_1^* =$  and $\nabla_v(T_1, T_2) =$  .

One can reverse the above procedure. Let T be a binary plane tree, and suppose v^* is a vertex in T with 2 children. Let v be the left child of v^* in T, and let T_2 be the right subtree of v^* in T. Let T_1^* be the tree obtained by deleting T_2 from T, and let T_1 be the tree obtained from T_1^* by contracting the edge connecting v and v^* into a single vertex. We call this contracted vertex v, identifying it with the original v. We say the pair (T_1, T_2) is the *decomposition* of T at v^* and write $\Delta_{v^*}(T) = (T_1, T_2)$.

We say a collection $\mathbf{T} \subseteq \mathsf{BPT}$ is *insertion-closed* if for all nonempty trees $T_1, T_2 \in \mathbf{T}$ and every vertex v of T_1, the tree $\nabla_v(T_1, T_2)$ is in $\mathbf{T}$. We say $\mathbf{T}$ is *decomposition-closed* if for every $T \in \mathbf{T}$ and every vertex v^* of T that has 2 children, the pair $\Delta_{v^*}(T)$ is in $\mathbf{T} \times \mathbf{T}$. A *troupe* is a set of binary plane trees that is insertion-closed and decomposition-closed. The article [5] shows that there are uncountably many troupes and characterizes troupes in terms of their *branch generators*, which are essentially their "indecomposable" elements.

Remark 1. There is actually a more general definition of a troupe in which one is allowed to color the vertices of the trees. This more general setup allows one to give further examples of the interactions between troupes and cumulants. Here, we have limited our definition of a troupe to uncolored trees for the sake of brevity.

Let us say a binary plane tree is *full* if every vertex has either 0 or 2 children. A *Motzkin tree* is a binary plane tree in which every vertex that has a right child also has a left child. Let FBPT and Mot denote the set of full binary plane trees and the set of Motzkin trees, respectively. Then BPT, FBPT, and Mot are all examples of troupes.

Let X be a finite set of positive integers. A *decreasing binary plane tree* on X is a binary plane tree whose vertices are bijectively labeled with the elements of X so that every nonroot vertex has a label that is smaller than the label of its parent. If $\mathcal{T}$ is a decreasing binary plane tree, then the *skeleton* of $\mathcal{T}$, denoted $\mathrm{skel}(\mathcal{T})$, is the binary plane tree obtained by removing the labels from $\mathcal{T}$. Given a set $\mathbf{T}$ of binary plane trees, we let $\mathsf{D}\mathbf{T}$ denote the set of decreasing binary plane trees $\mathcal{T}$ such that $\mathrm{skel}(\mathcal{T}) \in \mathbf{T}$. Thus, DBPT is the set of all decreasing binary plane trees.

A *tree statistic* is a function $f\colon \mathsf{BPT} \to \mathbb{C}$. If we are given a tree statistic $f\colon \mathsf{BPT} \to \mathbb{C}$, we define a function $\ddot{f}\colon \mathsf{DBPT} \to \mathbb{C}$ by $\ddot{f}(\mathcal{T}) = f(\mathrm{skel}(\mathcal{T}))$. We say a tree statistic f is *insertion-additive* if

$$f(\nabla_v(T_1, T_2)) = f(T_1) + f(T_2)$$

for all nonempty binary plane trees T_1 and T_2 and all vertices v in T_1. Two examples of insertion-additive tree statistics are the functions $T \mapsto \mathrm{right}(T) + 1$ and $T \mapsto \mathrm{prol}(T) + 1$, where $\mathrm{right}(T)$ is the number of right edges of T and $\mathrm{prol}(T)$ is the number of vertices in T that have two children (sometimes called *prolific* vertices).

A *permutation* of X is an ordering of the elements of X, which we write as a word in one-line notation. Let S_n be the set of permutations of $[n]$. A *descent* of a permutation $\pi = \pi_1 \cdots \pi_n$ is an index $i \in [n-1]$ such that $\pi_i > \pi_{i+1}$. A *peak* of π is an index $i \in \{2, \ldots, n-1\}$ such that $\pi_{i-1} < \pi_i > \pi_{i+1}$. We write $\mathrm{des}(\pi)$ and $\mathrm{peak}(\pi)$ for the number of descents of π and the number of peaks of π, respectively.

The *in-order traversal* $\mathcal{I}$ and the *postorder traversal* $\mathcal{P}$ are two functions that send decreasing binary plane trees on X to permutations of X. If $\mathcal{T}$ is the empty tree, then $\mathcal{I}(\mathcal{T})$ and $\mathcal{P}(\mathcal{T})$ are both just the empty permutation. Now suppose $\mathcal{T}$ is a nonempty decreasing binary plane tree on X. Let $\mathcal{T}_L$ and $\mathcal{T}_R$ be the (possibly empty) left and right subtrees of the root of $\mathcal{T}$, respectively. Let $m = \max(X)$ be the label of the root of $\mathcal{T}$. The in-order and postorder traversals are defined recursively by

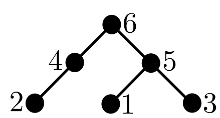
$$\mathcal{I}(\mathcal{T}) = \mathcal{I}(\mathcal{T}_L)\, m\, \mathcal{I}(\mathcal{T}_R) \quad \text{and} \quad \mathcal{P}(\mathcal{T}) = \mathcal{P}(\mathcal{T}_L)\, \mathcal{P}(\mathcal{T}_R)\, m.$$

It is known that $\mathcal{I}$ is a bijection from the set of decreasing binary plane trees on X to the set of permutations of X. Therefore, given a permutation π, we let $\mathcal{I}^{-1}(\pi)$ denote the unique decreasing binary plane tree whose in-order traversal is π.

The stack-sorting map is a function $s\colon S_n \to S_n$ that West introduced in his dissertation [12] as a deterministic version of Knuth's stack-sorting machine [9]. It has received vigorous attention over the last three decades (see [2, 4, 5] and the references therein). There are various ways to define s, one of which makes use of tree traversals. Namely,

$$s = \mathcal{P} \circ \mathcal{I}^{-1}. \tag{2.1}$$

For example,

$246153 \xrightarrow{\mathcal{I}^{-1}}$  $\xrightarrow{\mathcal{P}} 241356,$

so $s(246153) = 241356$.

3 Valid Hook Configurations

The *plot* of a permutation $\pi = \pi_1 \cdots \pi_n$ is the diagram showing the points (i, π_i) for all $i \in [n]$. A *hook* of π is a rotated L shape connecting two points (i, π_i) and (j, π_j) with

$i < j$ and $\pi_i < \pi_j$, as in Figure 1. The point (i, π_i) is the *southwest endpoint* of the hook, and (j, π_j) is the *northeast endpoint* of the hook. For example, Figure 1 shows the plot of the permutation $\pi = 426315789$. The hook shown in this figure has southwest endpoint $(3, 6)$ and northeast endpoint is $(8, 8)$.

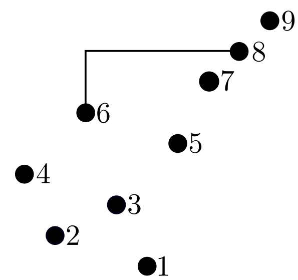


Figure 1: The plot of 426315789 along with a single hook. We have labeled the points in the plot with their heights.

Let π be a permutation with descents $d_1 < \cdots < d_k$. A *valid hook configuration* of π is a tuple $\mathcal{H} = (H_1, \ldots, H_k)$ of hooks of π that satisfy the following properties:

1. For each $i \in [k]$, the southwest endpoint of H_i is (d_i, π_{d_i}).
2. No point in the plot of π lies directly above a hook in $\mathcal{H}$.
3. No two hooks in $\mathcal{H}$ intersect or overlap each other unless the northeast endpoint of one is the southwest endpoint of the other.

Let $\text{VHC}(\pi)$ denote the set of valid hook configurations of π. We make the convention that a valid hook configuration includes its underlying permutation as part of its identity so that $\text{VHC}(\pi) \cap \text{VHC}(\pi') = \emptyset$ when $\pi \neq \pi'$. Given a set S of permutations, let $\text{VHC}(S) = \bigcup_{\pi \in S} \text{VHC}(\pi)$. If π is monotonically increasing, then $\text{VHC}(\pi)$ contains a single element: the empty valid hook configuration of π, which has no hooks.

Fix $\pi = \pi_1 \cdots \pi_n$ with $\text{des}(\pi) = k$. Each valid hook configuration $\mathcal{H} = (H_1, \ldots, H_k) \in \text{VHC}(\pi)$ induces a coloring of the plot of π. To begin this coloring, draw a sky over the entire diagram, and color the sky blue. Assign arbitrary distinct colors other than blue to the hooks $H_1, \ldots, H_k$. In order to decide how to color a point (i, π_i), imagine that this point looks directly upward. If it sees a hook when looking upward, it receives the same color as the hook that it sees. If it does not see a hook, it must see the sky, so it receives the color blue. However, if (i, π_i) is the southwest endpoint of a hook, then it must look around (on the left side of) the vertical part of that hook. We also require that the northeast endpoint of H_i must receive the same color as H_i. Figure 2 shows the coloring of the plot of a permutation induced by a valid hook configuration.

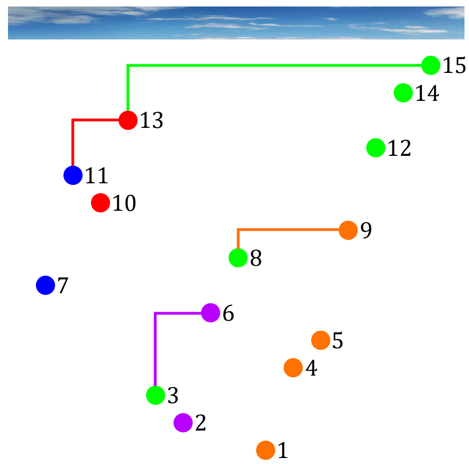


Figure 2: A valid hook configuration and its induced coloring.

Consider the coloring of the plot of a permutation $\pi \in S_{n-1}$ induced by a valid hook configuration $\mathcal{H} \in \mathrm{VHC}(\pi)$. We obtain a partition $|\mathcal{H}$ of the set $[n]$ by declaring that two numbers lie in the same block of the partition if and only if the points with those heights have the same color; here, we think of the sky as a blue point with height n.[1] For example, if $\mathcal{H}$ is the valid hook configuration whose coloring appears in Figure 2, then

$$|\mathcal{H} = \{\{1,4,5,9\},\{2,6\},\{3,8,12,14,15\},\{7,11,16\},\{10,13\}\}.$$

4 Cumulants

Let $\mathbb{K}$ be a field. Let $\Pi(X)$ denote the collection of all set partitions of a totally ordered finite set X. We let $\Pi(n) = \Pi([n])$. We say two distinct blocks B, B' of a set partition $\rho \in \Pi(X)$ form a *crossing* if there exist $i, j \in B$ and $i', j' \in B'$ such that either $i < i' < j < j'$ or $i > i' > j > j'$. A partition is *noncrossing* if no two of its blocks form a crossing. Let $\mathrm{NC}(X)$ be the set of noncrossing partitions in $\Pi(X)$, and let $\mathrm{NC}(n) = \mathrm{NC}([n])$.

A *noncommutative probability space* over $\mathbb{K}$ is a pair $(\mathcal{A}, \varphi)$, where $\mathcal{A}$ is a unital associative algebra and $\varphi \colon \mathcal{A} \to \mathbb{K}$ is a unital linear functional (meaning $\varphi(1_{\mathcal{A}}) = 1_{\mathbb{K}}$). Given $a_1, \ldots, a_n \in \mathcal{A}$ and $B = \{b_1 < \cdots < b_r\} \subseteq [n]$, let $a_B = (a_{b_1}, \ldots, a_{b_r})$. One of the goals of noncommutative probability theory is to understand the *joint moments*

$$m_n(a_1, \ldots, a_n) = \varphi(a_1 \cdots a_n).$$

[1] The notation $|\mathcal{H}$ comes from the fact that the partition is obtained by considering the vertical coordinates of the points. There is a different set partition $\underline{\mathcal{H}}$ discussed in [5], which is obtained by considering the horizontal coordinates.

The *classical cumulants* are the elements $c_n(a_1, \ldots, a_n)$ of $\mathbb{K}$ that satisfy the formula

$$m_n(a_1, \ldots, a_n) = \sum_{\rho \in \Pi(n)} \prod_{B \in \rho} c_{|B|}(a_B). \tag{4.1}$$

The *free cumulants*, originally introduced by Speicher [11], are the elements $\kappa_n(a_1, \ldots, a_n)$ of $\mathbb{K}$ that satisfy the formula

$$m_n(a_1, \ldots, a_n) = \sum_{\eta \in \mathrm{NC}(n)} \prod_{B \in \eta} \kappa_{|B|}(a_B). \tag{4.2}$$

The preceding paragraphs describe moments, classical cumulants, and free cumulants that are *multivariate* in the sense that they involve several (possibly) distinct elements $a_1, a_2, \ldots$ of $\mathcal{A}$. In many applications, it will suffice to consider the *univariate* case in which the elements $a_1, a_2, \ldots$ are all equal. In this case, we drop the notation expressing the dependence on $a_1, a_2, \ldots$ and simply write m_n, c_n, and κ_n. In fact, we will rarely need to refer to the noncommutative probability space $(\mathcal{A}, \varphi)$.

We can view an arbitrary sequence $(\kappa_n)_{n\geq 1}$ of elements of $\mathbb{K}$ as a sequence of free cumulants and use (4.2) to compute the corresponding moment sequence $(m_n)_{n\geq 1}$. We can then invert (4.1) (using Möbius inversion on the partition lattice) in order to obtain a sequence $(c_n)_{n\geq 1}$. For example, one can show that $c_4 = \kappa_4 - \kappa_2^2$. Thus every sequence of free cumulants corresponds to a unique sequence of classical cumulants. Similarly, every sequence of classical cumulants corresponds to a unique sequence of free cumulants. Several recent articles (see [1, 8, 5] and the references therein) have focused on finding combinatorial formulas that convert from one sequence of cumulants to another (there are also other types of cumulants that we will not discuss). In the next section, we will state the VHC Cumulant Formula, which expresses classical cumulants in terms of free cumulants via a sum over valid hook configurations.

One can also use known techniques to convert from the ordinary generating function $\sum_{n\geq 1} \kappa_n z^n$ to the exponential generating function $\sum_{n\geq 1} c_n z^n/n!$. To do so, one combines Voiculescu's R-transform with an inverse Laplace transform and the Exponential Formula (see [5] for more details). Performing these computations explicitly is often infeasible because the expressions become too unwieldy. However, occasionally, the generating functions are nice enough that the computations can be performed. In these cases, we can sometimes combine the generating function techniques with the VHC Cumulant Formula and the Refined Tree Fertility Formula to prove new difficult facts about the stack-sorting map and postorder traversals.

5 The Two Main Formulas

In this section, we state the two main formulas that are given by sums over valid hook configurations; the proofs can be found in [5]. The first formula is the Refined Tree Fertility Formula; the word "fertility" originates from West's thesis [12], where the *fertility* of a permutation π is defined to be the number of preimages of π under the stack-sorting map s.

Theorem 1 (Refined Tree Fertility Formula [5]). *Let* $\mathbf{T}$ *be a troupe, and let* $f_1, \ldots, f_r$ *be insertion-additive tree statistics. For every permutation* π*, we have*

$$\sum_{\mathcal{T}\in\mathcal{P}^{-1}(\pi)\cap\mathsf{D}\mathbf{T}} x_1^{\ddot{f}_1(\mathcal{T})}\cdots x_r^{\ddot{f}_r(\mathcal{T})} = \sum_{\mathcal{H}\in\mathrm{VHC}(\pi)} \prod_{B\in|\mathcal{H}} \sum_{T\in\mathbf{T}_{|B|-1}} x_1^{f_1(T)}\cdots x_r^{f_r(T)}.$$

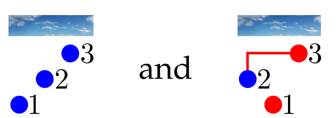
To gain intuition about the preceding theorem, let us restrict to the case in which $r = 0$. Under this specialization, the Refined Tree Fertility Formula tells us that the number of decreasing binary plane trees in $\mathsf{D}\mathbf{T}$ with postorder traversal π is given by $\sum_{\mathcal{H}\in\mathrm{VHC}(\pi)}\prod_{B\in|\mathcal{H}}|\mathbf{T}_{|B|-1}|$. To be even more concrete, we can consider the case in which $r = 0$ and $\mathbf{T} = \mathsf{BPT}$. In this case, it follows from (2.1) that $|\mathcal{P}^{-1}(\pi)\cap\mathsf{DBPT}| = |s^{-1}(\pi)|$ is the fertility of π. Furthermore, $|\mathsf{BPT}_m| = C_m = \frac{1}{m+1}\binom{2m}{m}$ is the m-th Catalan number. Thus, the Refined Tree Fertility Formula expresses the fertility of an arbitrary permutation π as a sum of products of Catalan numbers, where the sum ranges over the valid hook configurations of π. The author has used this formula to uncover several new facts about the stack-sorting map (see [4, 5] and the references therein).

Our second main formula is the VHC Cumulant Formula.

Theorem 2 (VHC Cumulant Formula [5]). *If* $(\kappa_n)_{n\geq 1}$ *is a sequence of free cumulants, then the corresponding classical cumulants are given by*

$$-c_n = \sum_{\mathcal{H}\in\mathrm{VHC}(S_{n-1})} \prod_{B\in|\mathcal{H}} (-\kappa_{|B|}).$$

For example, the two elements of VHC(S_3), drawn with their induced colorings, are



The associated set partitions are $\{\{1,2,3,4\}\}$ and $\{\{1,3\},\{2,4\}\}$, so $-c_4 = -\kappa_4 + (-\kappa_2)^2 = -\kappa_4 + \kappa_2^2$.

The VHC Cumulant Formula actually extends to the multivariate setting (see [5]). However, the univariate version of the formula is sufficient for all of our applications.

6 Troupes and Cumulants

Suppose we define a sequence $(\kappa_n)_{n\geq 1}$ of free cumulants by $\kappa_n = -C_{n-1}$, where $C_m = \frac{1}{m+1}\binom{2m}{m}$ is the m-th Catalan number. Then the corresponding sequence of classical cumulants $(c_n)_{n\geq 1}$ is given by $c_n = -(n-1)!$. Indeed, this is equivalent to the fact that the sequences $((-1)^{n-1}C_{n-1})_{n\geq 1}$ and $((-1)^{n-1}(n-1)!)_{n\geq 1}$ give the Möbius invariants of noncrossing partition lattices and partition lattices, respectively. On the other hand, C_{n-1} is the number of binary plane trees with $n-1$ vertices, while $(n-1)!$ is the number of decreasing binary plane trees with $n-1$ vertices. The main result of this section shows that this is no coincidence.

In what follows, our free and classical cumulants belong to the field $\mathbb{K}=\mathbb{C}(x_1,\ldots,x_r)$ of rational functions in the variables $x_1,\ldots,x_r$. Given a set $\mathbf{T}$ of binary plane trees, let us write $\overline{\mathsf{D}}\mathbf{T}_m$ for the set of decreasing binary plane trees in $\mathsf{D}\mathbf{T}$ on the set $[m]$.

Theorem 3 ([5]). *Let* $\mathbf{T}$ *be a troupe. Let* $f_1,\ldots,f_r$ *be insertion-additive tree statistics, and let* $x_1,\ldots,x_r$ *be variables. If* $(\kappa_n)_{n\geq 1}$ *is the sequence of free cumulants defined by*

$$\kappa_n=-\sum_{T\in\mathbf{T}_{n-1}}x_1^{f_1(T)}\cdots x_r^{f_r(T)},$$

then the corresponding sequence $(c_n)_{n\geq 1}$ *of classical cumulants is given by*

$$c_n=-\sum_{\mathcal{T}\in\overline{\mathsf{D}}\mathbf{T}_{n-1}}x_1^{\ddot{f}_1(\mathcal{T})}\cdots x_r^{\ddot{f}_r(\mathcal{T})}.$$

Proof. By combining the Refined Tree Fertility Formula with the VHC Cumulant Formula, we find that

$$-c_n=\sum_{\mathcal{H}\in\mathrm{VHC}(S_{n-1})}\prod_{B\in|\mathcal{H}}(-\kappa_{|B|})=\sum_{\pi\in S_{n-1}}\sum_{\mathcal{H}\in\mathrm{VHC}(\pi)}\prod_{B\in|\mathcal{H}}(-\kappa_{|B|})$$

$$=\sum_{\pi\in S_{n-1}}\sum_{\mathcal{T}\in\mathcal{P}^{-1}(\pi)\cap\mathsf{D}\mathbf{T}}x_1^{\ddot{f}_1(\mathcal{T})}\cdots x_r^{\ddot{f}_r(\mathcal{T})}=\sum_{\mathcal{T}\in\mathcal{P}^{-1}(S_{n-1})\cap\mathsf{D}\mathbf{T}}x_1^{\ddot{f}_1(\mathcal{T})}\cdots x_r^{\ddot{f}_r(\mathcal{T})}.$$

The desired result follows since $\mathcal{P}^{-1}(S_{n-1})\cap\mathsf{D}\mathbf{T}=\overline{\mathsf{D}}\mathbf{T}_{n-1}$. □

Remark 2. Although the statement of Theorem 3 is quite simple, its proof relies on the Refined Tree Fertility Formula and the VHC Cumulant Formula, which are quite nontrivial to prove. Furthermore, both of these formulas rely on the carefully-defined notions of valid hook configurations and their induced set partitions.

7 Applications to Stack-Sorting

7.1 Descents After Stack-Sorting

Recall that s denotes the stack-sorting map. Viewing s as a sorting operator, it is natural to consider $\mathrm{des}(s(\sigma))$ as a measure of how "far" $s(\sigma)$ is from the identity permutation. It is known that if $\sigma\in S_{n-1}$, then $0\leq\mathrm{des}(s(\sigma))\leq\frac{n-2}{2}$; moreover, each of these bounds is attained by some choice of σ. What happens if we choose $\sigma \in S_{n-1}$ uniformly at random? Then $\mathrm{des}(s(\sigma))$ is a random variable, and we can ask about its expected value. It is not at all clear how one would gain any nontrivial information about this expected value using elementary methods. However, tools from free probability allow us to compute it exactly, and as we will see below, it has a shockingly simple form! In fact, the article [5] outlines an algorithmic procedure for computing the higher moments of this random variable as well. These results are derived from the following theorem.

Theorem 4 ([5])**.** *Let* $F_x(z) = \frac{1}{2}(-x - x^2 z) + x\sqrt{1 - 4z + 2xz + x^2 z^2}$. *Then*

$$\sum_{n\geq 1}\left(\sum_{\sigma\in S_{n-1}} x^{\mathrm{des}(s(\sigma))+1}\right)\frac{z^n}{n!} = -\log(1 + \mathcal{L}^{-1}\{F_x(1/t)/t\}(z)),$$

where $\mathcal{L}^{-1}$ *denotes the inverse Laplace transform with respect to the variable* t.

Proof sketch. The main idea is to define free cumulants κ_n over the field $\mathbb{C}(x)$ by $\kappa_n = -xC_{n-1}$. Combining the VHC Cumulant Formula with the Refined Tree Fertility Formula (in the special case where $\mathbf{T} = \mathsf{BPT}$ and $r = 0$), one can show that the corresponding classical cumulants are given by $c_n = -\sum_{\sigma\in S_{n-1}} x^{\mathrm{des}(s(\sigma))+1}$. The descent statistic comes into play here because if $\mathcal{H} \in \mathrm{VHC}(\pi)$, then the number of blocks of the partition $|\mathcal{H}$ is $\mathrm{des}(\pi) + 1$. Indeed, this follows from the definitions of valid hook configurations and the partitions $|\mathcal{H}$. One can then use the generating function techniques mentioned at the end of Section 4 to show that $\sum_{n\geq 1} c_n z^n/n! = \log(1 + \mathcal{L}^{-1}\{F_x(1/t)/t\}(z))$. □

The article [5] explains how to derive the next theorem (and more) from Theorem 4.

Theorem 5. *If* $\sigma \in S_{n-1}$ *is chosen uniformly at random, then*

$$\mathbb{E}[\mathrm{des}(s(\sigma)) + 1] = \left(3 - \sum_{j=0}^{n}\frac{1}{j!}\right) n \sim (3 - e)n.$$

Furthermore, $\mathrm{Var}[\mathrm{des}(s(\sigma)) + 1] \sim (2 + 2e - e^2)n$.

The techniques used to derive the preceding results are not limited to the stack-sorting map; the same methods generalize readily to allow us to understand postorder readings of trees arising from several other troupes. For brevity, we will state results for just two other troupes, and we will focus on just the expected values of the relevant random variables. Recall the troupe FBPT of full binary plane trees and the troupe Mot of Motzkin trees from Section 2. Let us also recall the notation $\overline{\mathsf{D}}\mathbf{T}_m$ from Section 6.

Theorem 6. *If* $n \geq 2$ *is even and* $\mathcal{T} \in \overline{\mathsf{D}}\mathsf{FBPT}_{n-1}$ *is chosen uniformly at random, then*

$$\mathbb{E}[\mathrm{des}(\mathcal{P}(\mathcal{T})) + 1] = \left(1 - \frac{E_n}{nE_{n-1}}\right) n \sim \left(1 - \frac{2}{\pi}\right) n,$$

where E_m *denotes the* m*-th Euler number.*

Theorem 7. *If $\mathcal{T} \in \overline{\mathsf{D}}\mathsf{Mot}_{n-1}$ is chosen uniformly at random, then*

$$\mathbb{E}[\mathrm{des}(\mathcal{P}(\mathcal{T}))+1] \sim \left(1 - \frac{3\sqrt{3}}{2\pi}\left(e^{\frac{\pi}{3\sqrt{3}}} - 1\right)\right) n.$$

7.2 Uniquely Sorted Permutations

For our next application, we consider *uniquely sorted permutations*, which are permutations with exactly 1 preimage under s. These permutations were introduced by Engen, Miller, and the current author in [6]. In that article, it was shown that there are no uniquely sorted permutations of even size and that the number of uniquely sorted permutations of size $2k-1$ is the k-th term in *Lassalle's sequence*. This is a fascinating sequence introduced by Lassalle in [10], who settled a conjecture of Zeilberger's by showing that its terms are positive and increasing. Lassalle asked for a combinatorial interpretation of this sequence, which uniquely sorted permutations provide (Josuat-Vergès gave a different interpretation in [8]). It turns out that the terms in Lassalle's sequence are the absolute values of the (nonzero) classical cumulants of the *standard semicircular distribution*, a probability distribution that is fundamental in free probability theory. Thus, the combination of the VHC Cumulant Formula and the Refined Tree Fertility Formula neatly explains why uniquely sorted permutations are counted by Lassalle's sequence. The details are given in [5].

7.3 Sorted Permutations

Bousquet-Mélou [3] defined a permutation to be *sorted* if it is in the image of the stack-sorting map. She found a recurrence relation that provides a polynomial-time algorithm for counting sorted permutations; however, the asymptotic behavior of the resulting sequence is still not well understood. The following asymptotic bounds were derived in [5]. The proof of the upper bound of 0.75260 relies heavily on the Refined Tree Fertility Formula, the VHC Cumulant Formula, and the generating function techniques mentioned at the end of Section 4.

Theorem 8. *The limit* $\lim_{n\to\infty}\left(\frac{|s(S_n)|}{n!}\right)^{1/n}$ *exists and lies in the interval* $[0.68631, 0.75260]$.

7.4 Degree of Noninvertibility

Given a finite set X and a function $f\colon X \to X$, Propp and the author [7] defined the *degree of noninvertibility*

$$\deg(f\colon X \to X) = \frac{1}{|X|}\sum_{x\in X}|f^{-1}(x)|^2$$

as a measure of how far the function f is from being invertible. They showed that the limit $\lim_{n\to\infty} \deg(s\colon S_n \to S_n)^{1/n}$ exists and lies in the interval $[1.12462, 4]$, and they conjectured that it actually lies in the interval $(1.68, 1.73)$. In [5], the author combined the Refined Tree Fertility Formula with the VHC Cumulant Formula in order to apply tools from free probability (namely, the generating function techniques mentioned at the end of Section 4) to obtain an improved lower bound of 1.62924 for the limit.

References

[1] O. Arizmendi, T. Hasebe, F. Lehner, and C. Vargas. "Relations between cumulants in noncommutative probability". *Adv. Math.* **282** (2015), pp. 56–92. DOI.

[2] M. Bóna. "A survey of stack-sorting disciplines". *Electron. J. Combin.* **9** (2003). DOI.

[3] M. Bousquet-Mélou. "Sorted and/or sortable permutations". *Discrete Math.* **225**.1 (2000), pp. 25–50. DOI.

[4] C. Defant. "Counting 3-stack-sortable permutations". *J. Combin. Theory Ser. A* **172** (2020). DOI.

[5] C. Defant. "Troupes, cumulants, and stack-sorting". *Adv. Math.* **399** (2022). DOI.

[6] C. Defant, M. Engen, and J. A. Miller. "Stack-sorting, set partitions, and Lassalle's sequence". *J. Combin. Theory Ser. A* **175** (2020). DOI.

[7] C. Defant and J. Propp. "Quantifying noninvertibility in discrete dynamical systems". *Electron. J. Combin.* **27** (2020). DOI.

[8] M. Josuat-Vergès. "Cumulants of the q-semicircular law, Tutte polynomials, and heaps". *Canad. J. Math.* **65** (2013), pp. 863–878. DOI.

[9] D. E. Knuth. *The Art of Computer Programming. Fundamental Algorithms*. Vol. 1. Addison-Wesley, 1968.

[10] M. Lassalle. "Two integer sequences related to Catalan numbers". *J. Combin. Theory Ser. A* **119**.4 (2012), pp. 923–935. DOI.

[11] R. Speicher. "Multiplicative functions on the lattice of non-crossing partitions and free convolution". *Math. Ann.* **298** (1994), pp. 611–628. DOI.

[12] J. West. "Permutations with restricted subsequences and stack-sortable permutations". PhD thesis. MIT, 1990.