Reduced word enumeration, complexity, and randomization

Cara Monical
Sandia National Laboratories
Albuquerque, NM 87185, U.S.A.
caramonical.math@gmail.com

Benjamin Pankow Alexander Yong*
Department of Mathematics
U. Illinois at Urbana-Champaign
Urbana, IL 61801, USA
{bpankow, ayong}@illinois.edu

Submitted: Mar 7, 2019; Accepted: Apr 25, 2022; Published: Jun 3, 2022

Abstract

A *reduced word* of a permutation w is a minimal length expression of w as a product of simple transpositions. We examine formulas and (randomized) algorithms for their enumeration. In particular, we prove that the *Edelman-Greene statistic*, defined by S. Billey-B. Pawlowski, is typically exponentially large. This implies a result of B. Pawlowski, that it has exponentially growing expectation. Our result is established by a formal run-time complexity analysis of A. Lascoux-M.-P. Schützenberger's *transition algorithm*. The more general problem of Hecke word enumeration, and its closely related question of counting *set-valued standard Young tableaux*, is also investigated. The latter enumeration problem is further motivated by work on *Brill-Noether varieties* due to M. Chan-N. Pflueger and D. Anderson-L. Chen-N. Tarasca. We also state some related problems about counting computational complexity.

Mathematics Subject Classifications: 05E05, 05E14, 68Q25

1 Introduction

1.1 Reduced word combinatorics

Let S_n denote the symmetric group on $\{1, 2, \ldots, n\}$. Each $w \in S_n$ can be expressed as a product of $\ell(w)$ simple transpositions $s_i = (i, i+1)$, where $\ell(w)$ is the number of *inversions* of w, i.e., pairs $i < j$ such that $w(i) > w(j)$. Such an expression $w = s_{i_1} s_{i_2} \cdots s_{i_{\ell(w)}}$ is a *reduced word* for w.

*Supported by NSF grant DMS 1500691 and Simons Collaboration Grant 582242

Let $\mathsf{Red}(w)$ be the set of reduced words for w. R. P. Stanley [36] defined a symmetric function F_w such that

$$\#\mathsf{Red}(w) = \text{the coefficient of } x_1 x_2 \cdots x_{\ell(w)} \text{ in } F_w. \tag{1}$$

In connection to *ibid.*, P. Edelman-C. Greene [13, Section 8] proved that

$$\#\mathsf{Red}(w) = \sum_{\lambda} a_{w,\lambda} f^{\lambda}, \quad \text{where} \tag{2}$$

- f^{λ} is the number of *standard Young tableaux* of shape λ, that is, row and column increasing bijective fillings of the Young diagram of λ using $1, 2, \ldots, |\lambda|$. The *hook-length formula* of J. S. Frame-G. de B. Robinson-R. M. Thrall [16] states

$$f^{\lambda} = \frac{|\lambda|!}{\prod_b h_b}, \tag{3}$$

  where the product is over all boxes $b \in \lambda$ and h_b is the *hooklength* of b, i.e., the number of boxes weakly right and strictly below b.

- $a_{w,\lambda}$ counts *EG tableaux*: row and column increasing fillings T of λ such that reading the entries $(i_1, i_2, \ldots, i_{|\lambda|})$ of T along columns, top to bottom, and right to left, gives a reduced word $s_{i_1} \cdots s_{i_{|\lambda|}}$ for w (cf. [10]).

Let $w_0 = n \; n-1 \; n-2 \; \ldots \; 3 \; 2 \; 1$ be the unique longest length permutation of S_n (hence $\ell(w_0) = \binom{n}{2}$). R. P. Stanley [36] proved that, in this case, (2) is short:

$$\#\mathsf{Red}(w_0) = f^{(n-1,n-2,\ldots,3,2,1)}; \tag{4}$$

hence $\#\mathsf{Red}(w_0)$ is computed by (3).

One measure of the brevity of (2) is the *Edelman-Greene statistic* on S_n,

$$\mathsf{EG}(w) = \sum_{\lambda} a_{w,\lambda};$$

this was introduced by S. Billey-B. Pawlowski [5]. From (4), one sees $\mathsf{EG}(w_0) = 1$. Permutations w such that $\mathsf{EG}(w) = 1$ are *vexillary*. These permutations are characterized by 2143-*pattern avoidance*: there are no indices $i_1 < i_2 < i_3 < i_4$ such that $w(i_1), w(i_2), w(i_3), w(i_4)$ are in the same relative order as 2143. For instance, $w = \underline{54}27\underline{8}31\underline{6}$ is not vexillary; the underlined positions give a 2143 pattern. Each such w has *shape* $\lambda(w)$ (defined in Section 2.2). Extending (4), whenever w is vexillary,

$$\#\mathsf{Red}(w) = f^{\lambda(w)}; \tag{5}$$

see, e.g., [29, Corollary 2.8.2]. Our main result (Theorem 7) is that EG is typically large. This implies a (weak version) of a Theorem of B. Pawlowski [32, Theorem 3.2.7]:

Theorem 1 (Average exponential growth). $\mathbb{E}[\mathsf{EG}] = \Omega(c^n)$*, for some fixed constant* $c > 1$.

1.2 Run-time complexity of transition

Our proof of Theorem 1 uses the *transition algorithm* of A. Lascoux-M. P. Schützenberger [26] (cf. [29, Sections 2.7, 2.8]). This algorithm constructs the *Lascoux-Schützenberger transition tree* $\mathcal{T}(w)$ whose root is w and the leaves $\mathcal{L}(w)$ are labelled with vexillary permutations (with multiplicity). With this,

$$\#\mathsf{Red}(w) = \sum_{v \in \mathcal{L}(w)} f^{\lambda(v)}; \tag{6}$$

see Section 2 for details. Different v may give the same $\lambda(v)$. After combining such terms, (6) is the same as (2); see Lemma 6.

The (practical) efficiency of (extensions/variations of) transition has been mentioned a number of times. S. Billey [3] calls transition "one of the most efficient methods" to compute Schubert polynomials. See also A. Buch [9, Sections 3.4, 3.5] who discusses a quantum cohomology version of transition that is "quite efficient" for computing Gromov-Witten invariants (based on "practical experiments"). Another remark in this vein is found in the abstract to Z. Hamaker-E. Marberg-B. Pawlowski [19] who develop a different variation on the transition algorithm to "efficiently compute the decomposition" of involution Stanley symmetric functions into Schur P-functions. On the other hand, concerning the application of transition to computing the Littlewood-Richardson coefficients [26], A. Garsia [18, p. 52] writes:

> "Curiously, their algorithm (in spite of their claims to the contrary) is hopelessly inefficient as compared with well known methods."

He also refers to transition as "efficient" for a different purpose in his study of $\mathsf{Red}(w)$.

Theorem 1 is actually a reformulation of the following result which may be interpreted as a lower-bound for the typical run-time of transition:

Theorem 2. $\mathbb{E}(\#\mathcal{L}) = \Omega(c^n)$ *for a fixed constant* $c > 1$*. That is the average running time of transition, as an algorithm to compute* $\#\mathsf{Red}(w)$*, is at least exponential in* n*.*

Theorem 7 strengthens Theorem 2 to show that the "typical" running time is exponentially large. To prove Theorem 2 we use that the expected number of occurences of a fixed pattern $\pi \in S_k$ in $w \in S_n$ is $\binom{n}{k}/k!$. Thus for $u = 2143$, this expectation is $O(n^4)$. One shows each step of transition reduces the number of 2143 patterns by $O(n^3)$. Using the graphical description of transition by A. Knutson and the third author [25], a node u of $\mathcal{T}(w)$ has exactly one child u' only if u' has weakly more 2143 patterns than u does. Consequently, $\mathcal{T}(w)$ has $\Omega(n)$ branch points along any root-to-leaf path and thus exponentially many leaves. (In fact, the $c > 1$ from our argument is close to 1.)

Section 6 collects some remarks and questions about the related matter of the *computational* complexity of counting $\#\mathsf{Red}(w)$.

1.3 Hecke words

Section 4 studies the more general problem of counting $\mathsf{Hecke}(w, N)$, the set of *Hecke words* of length N whose *Demazure product* is a given $w \in S_n$. Here, the role of Stanley's symmetric polynomial is played by the *stable Grothendieck polynomial* defined by S. Fomin and A. N. Kirillov [15]. Using work of S. Fomin and C. Greene [14] and of A. Buch, A. Kresch, M. Shimozono, H. Tamvakis and the third author [10], one has two analogues of the results of Edelman-Greene [13]. However, finding useful enumeration *formulas* for Hecke words is a challenge. This is even true for the case that w is vexillary.

As explained by Proposition 19, enumerating Hecke words is closely related to the problem of counting $f^{\lambda,N}$, the number of *set-valued tableaux* [8] that are N-*standard* of shape λ. These are fillings T of the boxes of λ by $1, 2, \ldots, N$, where each entry appears exactly once, and if one chooses precisely one entry from each box of T, one obtains a semistandard tableau. For example, if $N = 8$ and $\lambda = (3, 2)$, one tableau is $\begin{array}{|c|c|c|}\hline 1,2 & 4,5 & 8 \\ \hline 3 & 6,7 \\ \cline{1-2}\end{array}$.

There is no algorithm to compute $f^{\lambda,N}$ that is polynomial-time in the bit-length of the input (λ, N) (see Section 6, after Observation 39).

Problem 3. Does there exist an algorithm to compute $f^{\lambda,N}$ that is polynomial in $|\lambda|$ and N?

Clearly, (3) gives a solution when $N = |\lambda|$. Using a theorem of C. Lenart [27], there exists an $|\lambda|^{O(1)}$ algorithm for any λ and where $N = |\lambda| + k$, if k is *fixed* (Proposition 21).

Recent work of M. Chan-N. Pflueger [11] and D. Anderson-L. Chen-N. Tarasca [2] motivates study of $f^{\lambda,N}$ in terms of *Brill-Noether varieties.* We remark on two manifestly nonnegative formulas for the Euler characteristics of these varieties (Corollary 26).

1.4 Randomization

Section 5 gives three randomized algorithms to estimate $\#\mathsf{Red}(w)$ and/or $\#\mathsf{Hecke}(w, N)$ using *importance sampling.* That is, let S be a finite set. Assign $s \in S$ probability p_s. Let Z be a random variable on S with $Z(s) = 1/p_s$. Then

$$\mathbb{E}(\mathsf{Z}) = \sum_{s \in S} p_s \times \frac{1}{p_s} = \#S.$$

Using this, one can devise simple Monte Carlo algorithms to estimate $\#S$. The idea goes back to at least a 1951 article of H. Kahn-T. E. Harris [21], who furthermore credit J. von Neumann. The application to combinatorial enumeration was popularized through D. Knuth's article [22] which applies it to estimating the number of self-avoiding walks in a grid. An application to approximating the *permanent* was given by L. E. Rasmussen [33]. More recently, J. Blitzstein-P. Diaconis [6] develop an importance sampling algorithm to estimate the number of graphs with a given degree sequence. We are suggesting another avenue of applicability, to core objects of algebraic combinatorics.

2 The Graphical Transition Algorithm

2.1 Preliminaries

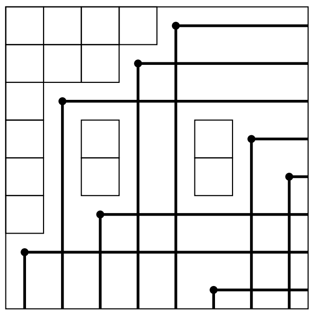
The *graph* $G(w)$ of a permutation $w \in S_n$ is the $n \times n$ grid, with a $\bullet$ placed in position $(i, w(i))$ (in matrix coordinates). The *Rothe diagram* of w is given by

$$D(w) = \{(i,j) : 1 \leqslant i,j \leqslant n,\ j < w(i),\ i < w^{-1}(j)\}.$$

Pictorially, this is described by striking out boxes below and to the right of each $\bullet$ in $G(w)$. $D(w)$ consists of the remaining boxes. If it exists, the connected component involving $(1,1)$ is the *dominant component.* The *essential set* of w consists of the maximally southeast boxes of each connected component of $D(w)$, i.e.,

$$\mathcal{E}ss(w) = \{(i,j) \in D(w) : (i+1,j), (i,j+1) \notin D(w)\}.$$

If it exists, the *accessible box* is the southmost then eastmost essential set box *not in the dominant component.* For example, if $w = 54278316 \in S_8$, $D(w)$ is depicted by:



Also, $\mathcal{E}ss(w) = \{(1,4),(2,3),(5,3),(5,6),(6,1)\}$, and the accessible box is at $(5,6)$. The *Lehmer code* of $w \in S_\infty$, denoted $\mathsf{code}(w)$ is the vector $(c_1, c_2, \ldots, c_L)$ where c_i equals the number of boxes in row i of the Rothe diagram of w. We will assume L is minimum (i.e., $\mathsf{code}(w)$ does not have trailing zeros). By this convention, $\mathsf{code}(id) = ()$.

Fulton's criterion [17, Remark 9.17] states that u is vexillary if and only if there does not exist two essential set boxes where one is strictly northwest of the other. Thus, using the above picture of $D(w)$ we can see that w is not vexillary because of, e.g., $(1,4)$ and $(5,6)$.

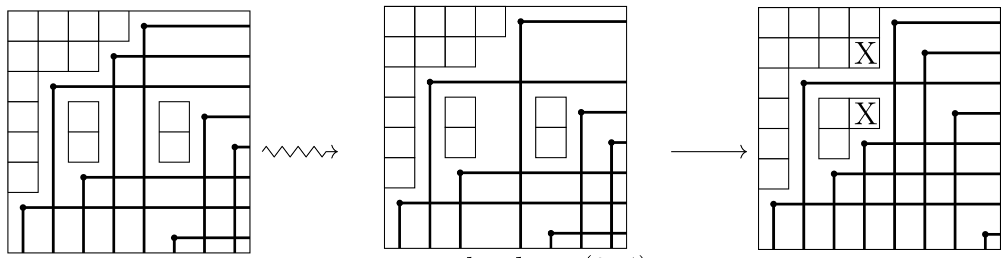
2.2 Description of $\mathcal{T}(w)$

Transition was invented by A. Lascoux-M. P. Schützenberger's [26]; see also the exposition [29, Sections 2.7.3, 2.7.4, 2.8.1]. We use the graphical description given in [25] and its elaboration in [1]. There are some minor choices in describing this *Lascoux-Schützenberger transition tree*, and those of [25, 1] differ slightly from [26, 29].

We describe the graphical version of the transition algorithm to compute $\#\mathsf{Red}(w)$. The root of the tree is labelled by $D(w)$. If w is vexillary, stop. Otherwise, there exists an accessible box. (If not, $D(w)$ consists only of the dominant component and, by Fulton's criterion, w is vexillary, a contradiction.) The *pivots* of $D(w)$ are the maximally southeast $\bullet$'s of $G(w)$, say $b_1, b_2, \ldots, b_t$ that are northwest of the accessible box e.

If w is not vexillary, the children of w are defined as follows. For each $i = 1, 2, \ldots, t$, let R_i be the rectangle defined by b_i and e. Remove b_i and its rays from $G(w)$ to form $G^{(i)}(w)$. Order the boxes $\{v_i\}_{i=1}^m$ in English reading order. Move v_1 strictly north and strictly west to the closest position not occupied by another box of $D(w)$ or a ray from $G^{(i)}(w)$. Now, iterate this procedure with $v_2, v_3, \ldots$. At each step, v_j may move to a position vacated by earlier moves. The result is the diagram $D(w^{(i)})$ of some permutation $w^{(i)}$. These $D(w^{(i)})$'s are the children of $D(w)$. We call the transformation $D(w) \to D(w^{(i)})$ a *marching move.*

Example 4. Continuing our example, the pivots of w are $(1,5), (2,4)$ and $(3,2)$. We now obtain the child corresponding to the pivot $b_2 = (2,4)$:



w remove hook at $(2,4)$ $w^{(2)} = 56274318$

We have indicated by "X" the boxes that have moved. This process constructs one of the three children of w. In Figure 1 we draw the remainder of $\mathcal{T}(w)$. $\square$

If u is vexillary we define $\lambda(u)$ graphically by pushing all boxes of $D(u)$ northwest along the diagonal that it sits until a partition shape is reached; see [24, Section 3.2]. Concluding our running example, from Figure 1 we have

$$\begin{aligned}\#\mathsf{Red}(54278316) &= f^{\lambda(54672318)} + f^{\lambda(56274318)} + f^{\lambda(65342718)} + f^{\lambda(64532718)} \\ &= f^{4,3,3,3,1,1} + f^{4,4,3,2,1,1} + f^{5,4,2,2,1,1} + f^{5,3,3,2,1,1} \\ &= 730158.\end{aligned}$$

This result is a mild variation of [29, Proposition 2.8.1] (cf. [26]) using the marching moves. We make no claim of originality.

Theorem 5 (cf. [26, 29, 25])**.** $\#\mathsf{Red}(w) = \sum_{v \in \mathcal{L}(w)} f^{\lambda(v)}$.

Proof: We follow [1, Section 5.2], which elaborates on the notions from [25] in the case of Schubert polynomials $\mathfrak{S}_w$. We refer to [29, Chapter 2] for background.

Let (r,c) be the accessible box of $w \in S_n$ and set $k = w^{-1}(c)$. Also let $w' = w \cdot (r,k)$. Transition gives this recurrence for the *Schubert polynomials*:

$$\mathfrak{S}_w = x_r \mathfrak{S}_{w'} + \sum_{w''} \mathfrak{S}_{w''}, \tag{7}$$

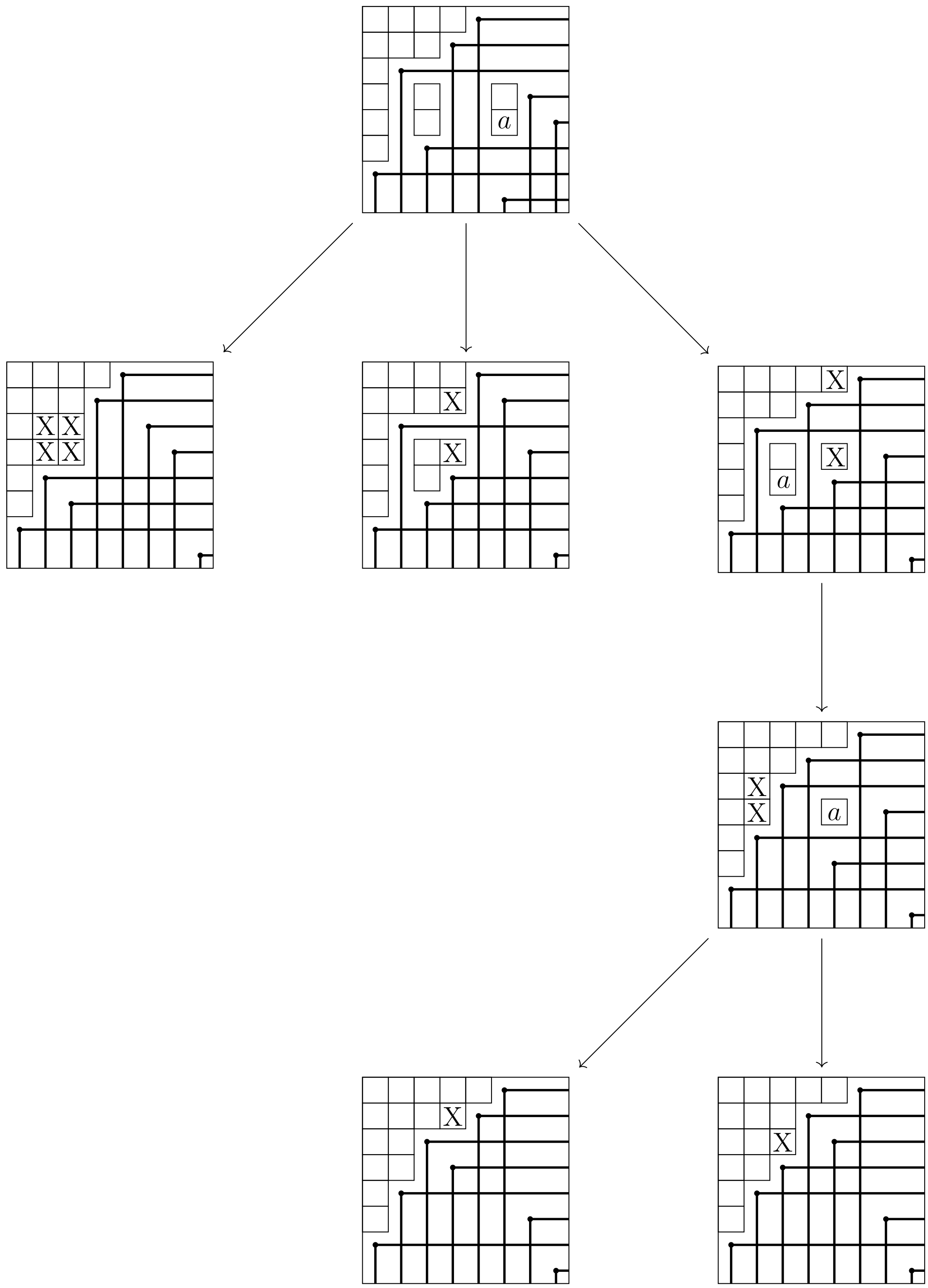


Figure 1: $\mathcal{T}(w)$ for $w = 54278316$. The a indicates the accessible box of each node. The X's describe which boxes of the parent moved. From this tree, we compute $\#\mathsf{Red}(w) = 730158$.

where the summation is over the children w'' of w in $\mathcal{T}(w)$.

Let $1^N \times w \in S_{N+n}$ send $j \mapsto j$ for $1 \leqslant j \leqslant n$ and $j \mapsto w(j-N+1)+N$ for $j \geqslant N$. Then

$$F_w = \lim_{N\to\infty} \mathfrak{S}_{1^N\times w} \in \mathbb{Z}[[x_1, x_2, \dots]].^1$$

Moreover, since $w \in S_n$ then

$$F_w(x_1, x_2, \dots, x_n, 0, 0, \dots) = \mathfrak{S}_{1^n\times w}(x_1, x_2, \dots, x_n, 0, 0, \dots). \tag{8}$$

Now, by repeated application of (7) to $1^n \times w$,

$$\mathfrak{S}_{1^n\times w} = J(x_1, x_2, \dots, x_{2n}) + \sum_{v\in\mathcal{L}(w)} \mathfrak{S}_{1^n\times v}, \tag{9}$$

where $J(x_1, x_2, \dots, x_n, 0, 0, \dots) \equiv 0$.

Hence by setting $x_i = 0$ for $i > n$ in (9) we obtain, using (8) that

$$F_w(x_1, \dots, x_n) = \sum_{v\in\mathcal{L}(w)} F_v(x_1, \dots, x_n). \tag{10}$$

Let $s_\alpha(x_1, \dots, x_n)$ be the Schur polynomial for a shape α. Since $v \in \mathcal{L}(w)$ is vexillary,

$$F_v(x_1, \dots, x_n) = s_{\lambda(v)}(x_1, \dots, x_n);$$

see, e.g., [29, Section 2.8.1]. Hence

$$F_w(x_1, \dots, x_n) = \sum_{v\in\mathcal{L}(w)} s_{\lambda(v)}(x_1, \dots, x_n). \tag{11}$$

We have that $[x_1x_2\cdots x_{\ell(w)}]F_w = \#\mathsf{Red}(w)$ and $[x_1x_2\cdots x_{\ell(w)}]s_{\lambda(v)}(x_1, \dots, x_n) = f^{\lambda(v)}$. Now the result follows from these two facts combined with (11). □

3 Proof of Theorems 1 and 2

3.1 On the distribution of $\mathsf{EG}(w)$

Lemma 6. *For any $w \in S_n$, $\mathsf{EG}(w) = \#\mathcal{L}(w)$.*

Proof. Combining results of [36, 13] gives

$$F_w(x_1, \dots, x_{\ell(w)}) = \sum_\lambda a_{w,\lambda} s_\lambda(x_1, \dots, x_{\ell(w)}) \tag{12}$$

where the sum is over partitions λ of size $\ell(w)$, and $a_{w,\lambda}$ is defined in Section 1.

The Schur polynomials $s_\lambda(x_1, \dots, x_{\ell(w)})$ for $|\lambda| = \ell(w)$ are a basis of the vector space $\Lambda^{(\ell(w))}_{\mathbb{Q}}[x_1, \dots, x_{\ell(w)}]$ of degree $\ell(w)$ symmetric polynomials in $\{x_1, \dots, x_{\ell(w)}\}$. Since (12) and (11) (where $n = \ell(w)$) are linear combinations for the same vector, we are done. □

[1] In the conventions of [36], the limit is an expression for $F_{w^{-1}}$.

In view of Lemma 6, Theorems 1 and 2 are equivalent. It is easy to see that Theorem 1 follows from our main result:

Theorem 7. *Fix $0 < \gamma < \frac{1}{2}$. There exists $\alpha > 0$ such that for n sufficiently large,*

$$\mathbb{P}(\mathsf{EG}(w) \geqslant 2^{\alpha n}) \geqslant 1 - \frac{1}{n^{2\gamma}}.$$

Our goal is therefore to prove Theorem 7. To do so we need some preparatory results. Let $N_{\pi,n}(w)$ be the number of π patterns contained in $w \in S_n$.

Proposition 8. *Suppose in $\mathcal{T}(w)$ that the node u has exactly one child u'. Then*

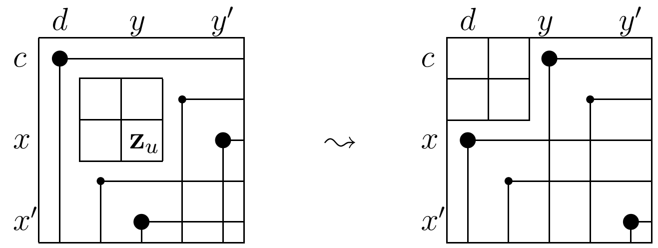
$$N_{2143,n}(u') \geqslant N_{2143,n}(u).$$

Proof of Proposition 8: Let the accessible box $\mathbf{z}_u$ of u be in position (x, y). By definition of $D(u)$, there is a $\bullet$ of $G(u)$ at $C = (x, y')$ for some $y' > y$, and there is a $\bullet$ at $B = (x', y)$ for some $x' > x$. Let b_1 be the unique pivot of $D(u)$, i.e., the southeastmost $\bullet$ that is northwest of $\mathbf{z}_u$ (as in Section 2.2). Suppose b_1 is at position $A = (c, d)$. Thus, $c < x$ and $d < y$.

By definition of the transition algorithm, all $\bullet$'s of $G(u)$ and $G(u')$ are in the same position, except A, B, C in $G(u)$ are respectively replaced by A', B', C' in $G(u')$ where

$$\begin{aligned} A = (c, d) &\mapsto A' = (x, d) \\ B = (x', y) &\mapsto B' = (c, y) \\ C = (x, y') &\mapsto C' = (x', y') \end{aligned}$$

Schematically, the march move looks as follows (we have thickened the moving $\bullet$'s).



Claim 9. *If there are two $\bullet$'s, other than $\{B, C\}$, that are weakly south and weakly east of $\mathbf{z}_u$ then one $\bullet$ must be (strictly) southeast of the other.*

Proof of Claim 9: Suppose not. Then let the two $\bullet$'s be at (q, r) and (m, n) where $q > m$ and $r < n$. Then there is a box $\mathbf{z} \neq \mathbf{z}_u$ of $D(u)$ in position (m, q), which is weakly south and weakly east of $\mathbf{z}_u$. Since $\mathbf{z}_u$ is not in the dominant component of $D(u)$, then $\mathbf{z}$ cannot be in that component either. Therefore, $\mathbf{z}_u$ is not the accessible box of $D(u)$, a contradiction. ☐

Claim 10. *There is no $\bullet$ of $G(u)$ strictly north of row c and strictly between columns d and y. Similarly, there is no $\bullet$ of $G(u)$ strictly west of column d and strictly between rows c and x.*

Proof of Claim 10: We prove only the first sentence of the claim, as the second sentence is analogous. Suppose not; we may assume this $\bullet$ is maximally southeast with the assumed properties. Then $A = (c, d)$ and this $\bullet$ are two pivots for $D(u)$, which implies u has at least two children, contradicting the hypothesis of the Proposition. □

Let $\mathcal{F}_u$ consist of all embedding positions $i_1 < i_2 < i_3 < i_4$ of a 2143-pattern in u. Also, let $\mathcal{F}'_u$ be the subset of $\mathcal{F}_u$ consisting of those $i_1 < i_2 < i_3 < i_4$ such that

$$\{i_1, i_2, i_3, i_4\} \cap \{c, x, x'\} = \emptyset$$

(i.e., the positions do not involve the rows of A, B or C). Let

$$\mathcal{F}''_u = \mathcal{F}_u \setminus \mathcal{F}'_u.$$

Similarly, we define $\mathcal{F}_{u'}$, $\mathcal{F}'_{u'}$ and $\mathcal{F}''_{u'}$ in exactly the same way, except with respect to u'.

Since $\mathcal{F}'_u = \mathcal{F}'_{u'}$, it suffices to establish an injection

$$\psi : \mathcal{F}''_u \hookrightarrow \mathcal{F}''_{u'}.$$

In what follows, we will let $\bullet_i$ refer to the $\bullet$ in the diagram corresponding to the "i" in the 2143 pattern, for $1 \leqslant i \leqslant 4$. In addition, if i_1 is in the row of A we will write "$A = \bullet_2$", *etc.* We define now ψ in cases:

Case 1: ($B = \bullet_1$ or $B = \bullet_2$) The $\bullet_4$ and $\bullet_3$ appear strictly right of column y. This contradicts Claim 9. Hence, no elements of $\mathcal{F}''_u$ fall into this case.

Case 2: ($C = \bullet_1$ or $C = \bullet_2$) $\bullet_4$ and $\bullet_3$ appear strictly southeast of $\mathbf{z}_u$. As in Case 1, this contradicts Claim 9. Again, no elements of $\mathcal{F}''_u$ fall into this case.

Case 3: ($A = \bullet_1$) Let $\bullet_2$ be at position (r, s). Hence $r < c$ and $s > d$. If moreover, $s < y$ we contradict the first sentence of Claim 10. Hence, $s > y$. We must have that $\bullet_2 \notin \{A, B, C\}$ and $\bullet_4$ and $\bullet_3$ are strictly to the right of column y.

Subcase 3a: ($\bullet_4$ and $\bullet_3$ are both strictly south of row x) This contradicts Claim 9.

Subcase 3b: ($\bullet_4$ and $\bullet_3$ are both strictly north of row x) Thus $\{\bullet_3, \bullet_4\} \cap \{A, B, C\} = \emptyset$. The 2143 pattern $[\bullet_2, A, \bullet_4, \bullet_3]$ is destroyed by the marching move, i.e., $[\bullet_2, A', \bullet_4, \bullet_3]$ is not a 2143 pattern in u'. Now, in u' we now have the 2143 pattern $[\bullet_2, B', \bullet_4, \bullet_3]$. Hence we define

$$\psi([\bullet_2, A, \bullet_4, \bullet_3]) := [\bullet_2, B', \bullet_4, \bullet_3].$$

Subcase 3c: ($\bullet_4$ is strictly north of row x and $\bullet_3$ is strictly south of row x). Since $s > y$, $C \neq \bullet_3$. Hence $\{\bullet_3, \bullet_4\} \cap \{A, B, C\} = \emptyset$. The 2143 pattern $[\bullet_2, A, \bullet_4, \bullet_3]$ is destroyed by the marching move. However, in u' we now have the 2143 pattern $[\bullet_2, B', \bullet_4, \bullet_3]$. We again define

$$\psi([\bullet_2, A, \bullet_4, \bullet_3]) := [\bullet_2, B', \bullet_4, \bullet_3].$$

Subcase 3d: ($\bullet_3$ is in row x and $\bullet_4$ is strictly above row x) Then in fact $\bullet_3 = C$ while $\bullet_4 \notin \{A, B, C\}$. In this case,

$$\psi([\bullet_2, A, \bullet_4, C]) := [\bullet_2, B', \bullet_4, C'].$$

Subcase 3e: ($\bullet_4$ is in row x and $\bullet_3$ is strictly south of row x) Thus $\bullet_4 = C$ and $\bullet_3$ is strictly southeast of z_u. This contradicts Claim 9.

Case 4: ($A = \bullet_2$) Let the 1 be at position (r, s). Hence $r > c$ and $s < d$. If $r \leqslant x$ then we contradict the second sentence of Claim 10. Hence $r > x$. We have that $\bullet_4$ and $\bullet_3$ are in rows strictly south of x. Moreover, there must be a box e of $D(u)$ in the row of $\bullet_4$ and the column of $\bullet_3$ that is therefore strictly south of $\mathbf{z}_u$. Since the columns of $\bullet_4$ and $\bullet_3$ are strictly east of column d, the box e is not part of the dominant component of $D(u)$. Hence, $\mathbf{z}_u$ cannot be the accessible box, a contradiction. Thus, no elements of $\mathcal{F}''_u$ are in this case.

Case 5: ($A = \bullet_3$) Hence, in u, $\bullet_2, \bullet_1, \bullet_4$ are strictly north of row c. Thus $\{\bullet_1, \bullet_2, \bullet_4\} \cap \{A, B, C\} = \emptyset$ and $\bullet_2, \bullet_1, \bullet_4$ remain in the same place in u'. Set

$$\psi([\bullet_2, \bullet_1, \bullet_4, A]) := [\bullet_2, \bullet_1, \bullet_4, A'].$$

Case 6: ($A = \bullet_4$) $\bullet_3$ is strictly south of the row of A. If it is also weakly north of x, we contradict the second sentence of Claim 10. Hence $\bullet_3$ is strictly south of x, i.e., the row of e. Now, $\bullet_2, \bullet_1, \bullet_3$ are the same position in u and u' and $\{\bullet_1, \bullet_2, \bullet_3\} \cap \{A, B, C\} = \emptyset$. Here,

$$\psi([\bullet_2, \bullet_1, A, \bullet_3]) := [\bullet_2, \bullet_1, A', \bullet_3].$$

Since the row of A' is x the output is a 2143 pattern in u'.

Case 7: ($B = \bullet_4$) Let $\bullet_3$ be at (r, s). Thus $r > x'$ and $s < y$. There must be a box $e \in D(w)$ in position (x', s). Now, $\bullet_2$ and $\bullet_1$ are in columns strictly left of s and strictly above row r. Hence e cannot be in the dominant component of $D(w)$. Thus, since e is further south than $\mathbf{z}_u$, the latter is not accessible, a contradiction. So, no elements of $\mathcal{F}''_u$ appear in this case.

Case 8: ($B = \bullet_3$) Let $\bullet_4$ be in position (r, s).

Subcase 8a: ($r < c$) Therefore, $\bullet_1$ and $\bullet_2$ are also strictly above row c. Since $\bullet_1, \bullet_2$ and $\bullet_4$ stay in the same place in u and u' and B' is in row c in u'. Moreover, $\{\bullet_1, \bullet_2, \bullet_4\} \cap \{A, B, C\} = \emptyset$. We may define

$$\psi([\bullet_2, \bullet_1, \bullet_4, B]) := [\bullet_2, \bullet_1, \bullet_4, B'].$$

Subcase 8b: ($x < r < x'$) This contradicts Claim 9.

Subcase 8c: ($r = c$) This implies $A = \bullet_4$, which is impossible.

Subcase 8d: ($c < r < x$) We may assume $A \neq \bullet_1, \bullet_2$ since those cases are handled by Case 3 and Case 4. Now, $\bullet_1$ and $\bullet_2$ are strictly west of column y and strictly north of row x. By the assumption that $A = b_1$ is the (unique) pivot, combined with Claim 10, both

$\bullet_1$ and $\bullet_2$ are strictly northwest of A. Thus, $\bullet_1$ and $\bullet_2$ are in the same place in u', and $\{\bullet_1, \bullet_2, \bullet_4\} \cap \{A, B, C\} = \emptyset$. Since A' is in row x, it make sense to let

$$\psi([\bullet_2, \bullet_1, \bullet_4, B]) := [\bullet_2, \bullet_1, \bullet_4, A'].$$

Subcase 8e: $(r = x)$ Hence $C = \bullet_4$. For the same reasons as in Subcase 8d, both $\bullet_1$ and $\bullet_2$ are strictly northwest of A. Thus, $\bullet_1$ and $\bullet_2$ are in the same place in u' and $\{\bullet_1, \bullet_2\} \cap \{A, B, C\} = \emptyset$. In this case set

$$\psi([\bullet_2, \bullet_1, C, B]) := [\bullet_2, \bullet_1, B', A'].$$

Case 9: $(C = \bullet_3)$ Let $\bullet_4$ be in position (r, s). Hence $s > y'$.
Subcase 9a: $(r < c)$ Hence $\bullet_1, \bullet_2$ and $\bullet_4$ remain in the same place in u' and $\{\bullet_1, \bullet_2, \bullet_4\} \cap \{A, B, C\} = \emptyset$. Since C' is further south than C, we may set

$$\psi([\bullet_2, \bullet_1, \bullet_4, C]) := [\bullet_2, \bullet_1, \bullet_4, C'].$$

Subcase 9b: $(c < r < x)$ We may also assume that $A \neq \bullet_1$ and $A \neq \bullet_2$, since those are handled in Case 3 and Case 4, respectively. Thus $\{\bullet_1, \bullet_2, \bullet_4\} \cap \{A, B, C\} = \emptyset$. Here,

$$\psi([\bullet_2, \bullet_1, \bullet_4, C]) := [\bullet_2, \bullet_1, \bullet_4, C'].$$

Subcase 9c: $(r = c)$ Then $A = \bullet_4$, which is impossible.

Case 10: $(C = \bullet_4)$ We may assume that $A \neq \bullet_1, \bullet_2$ (Case 3 and Case 4) and also $B \neq \bullet_3$ (Case 8). Therefore $\{\bullet_1, \bullet_2, \bullet_3\} \cap \{A, B, C\} = \emptyset$. Let $\bullet_3$ be in position (r, s).
Subcase 10a: $(y < s < y')$ This contradicts Claim 9.
Subcase 10b: $(s = y)$ This means $\bullet_3 = B$, a situation we have ruled out/refer to Case 8.
Subcase 10c: $(s < y)$ If moreover $r > x'$ then there exists $e \in D(w)$ in position (x', s), which is therefore strictly south of $\mathbf{z}_u$. Since column s is strictly east of the column of $\bullet_2$, e is not in the dominant component. Hence $\mathbf{z}_u$ is not accessible, a contradiction. Now $r \neq x'$ (since we assumed $B \neq \bullet_3$). Thus, $x < r < x'$ and it follows that $\bullet_3$ is in the same place in u'. By the reasoning of the first paragraph of Subcase 8d, $\bullet_1, \bullet_2$ are strictly northwest of A. Hence $\bullet_1, \bullet_2$ also remain in the same place in u'. Summing up, since B' is in row c, we may define

$$\psi([\bullet_2, \bullet_1, C, \bullet_3]) := [\bullet_2, \bullet_1, B', \bullet_3].$$

ψ is well-defined: The above cases handle each of the possibilities for A, B, C being one of $1, 2, 3, 4$. Our definition of ψ is shown to send an element of $\mathcal{F}''_u$ to an element of $\mathcal{F}''_{u'}$.

We also need that if an element of $\mathcal{F}''_u$ occurs in two cases, ψ sends them to the *same* element of $\mathcal{F}''_{u'}$. By inspection, the only overlapping situations are Subcase 3d$\leftrightarrow$Subcase 9b and Subcase 8d$\leftrightarrow$Case 10. In both these cases we define ψ to be consistent on the overlap.

ψ is an injection: This is by inspection of pairs of subcases where ψ's output was given. By our choice of notation, if $\bullet_i$ appears in the description of the input to ψ, it cannot be

equal to A, B or C and hence in the output, it cannot be equal to A', B' or C' (as $\{A, B, C\}$ and $\{A', B', C'\}$ occupy the same rows). Therefore, if in two cases, some coordinate of the two outputs differ *symbolically*, those outputs cannot be equal. After ruling out these pairs, we are left with a few to check:

Subcase 3b, Subcase 3c: These differ in the fourth coordinate since in the former case, $\bullet_3$ is strictly north of row x and in the latter case, $\bullet_3$ is strictly south of row x.

Case 5 and Subcase 8d: These differ in the third coordinate since in the former case, $\bullet_4$ appears above row c whereas in the latter case, $\bullet_4$ is below row c.

Subcase 9a and Subcase 9b: These differ in the third coordinate for the same reason as the previous pair. □

Lemma 11. *Let $w \in S_n$ and suppose $u \to u'$ in $\mathcal{T}(w)$. Then*

$$N_{2143,n}(u) - N_{2143,n}(u') \leqslant 2n^3 + 3n^2 - n.$$

Proof of Lemma 11: Since $u \to u'$ in $\mathcal{T}(w)$, exactly three positions a, i, j differ between u and u'. We are claiming that

$$N_{2143,n}(u) - N_{2143,n}(u') \leqslant \binom{3}{3} 4 \binom{n}{1} + \binom{3}{2} 6 \binom{n}{2} + \binom{3}{1} 4 \binom{n}{3} = 2n^3 + 3n^2 - n. \quad (13)$$

Let $t_1 < t_2 < t_3 < t_4$ be the indices of a 2143-pattern in u. First suppose $\{t_1, t_2, t_3, t_4\} \cap \{a, i, j\} = \emptyset$. Clearly, $t_1 < t_2 < t_3 < t_4$ are indices of a 2143-pattern in u'. Therefore this case does not contribute to $N_{2143,n}(u) - N_{2143,n}(u')$.

Next assume $\#(\{t_1, t_2, t_3, t_4\} \cap \{a, i, j\}) = 1$. There are $\binom{3}{1}$ choices for which of a, i or j is in $\{t_1, t_2, t_3, t_4\}$. Then there are at most $\binom{n}{3}$ choices for $\{t_1, t_2, t_3, t_4\} \setminus \{a, i, j\}$. Finally there are 4 choices for which k satisfies $t_k \in \{a, i, j\}$. Therefore, this case contributes at most $\binom{3}{1} 4 \binom{n}{3}$ to $N_{2143,n}(u) - N_{2143,n}(u')$, thus explaining the third term of (13).

Similar arguments explain the first and second terms of (13) as the contributions to $N_{2143,n}(u) - N_{2143,n}(u')$ from the cases that

$$\#(\{t_1, t_2, t_3, t_4\} \cap \{a, i, j\}) = 3 \text{ and } \#(\{t_1, t_2, t_3, t_4\} \cap \{a, i, j\}) = 2,$$

respectively. The lemma thus follows. □

The following is known; see work of M. Bona [7] and of S. Janson, B. Nakamura, and D. Zeilberger [20]. The proof being not difficult, we include it for completeness.

Lemma 12. *For any $\pi \in S_k$, the expected number of occurrences of π as a pattern in $w \in S_n$ (selected using the uniform distribution) is $\binom{n}{k}/k!$.*

Proof of Lemma 12: For an increasing sequence $I = \{i_1 < i_2 < \ldots < i_k\}$ (in $[1, n]$), let

$$\mathsf{X}_I(w) = \begin{cases} 1 & \text{if } \pi \text{ is a pattern at the positions of } I; \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

Thus, $N_{\pi,n} = \sum_I \mathsf{X}_I$. There are $\binom{n}{k}(n-k)!$ permutations such that I has pattern π. By linearity of expectation,

$$\mathbb{E}[N_{\pi,n}] = \sum_I \mathbb{E}[\mathsf{X}_I] = \binom{n}{k}^2 (n-k)!/n!,$$

and the lemma follows. □

Lemma 13. *Let $\mathcal{T}$ be a rooted tree with the property that along any path from the root to a leaf there are d nodes with at least two children. Then that tree has at least 2^d leaves.*

Proof of Lemma 13: Arbitrarily left-right order the descendants of the root of $\mathcal{T}$. After pruning, if necessary, we may assume each node as at most two children. Along any path from the root to a leaf, record "S" if a node has one child, and "L" if one steps to the left child and "R" if one goes to the right child. Thus, each leaf is uniquely encoded by an $\{S, L, R\}$ sequence. By hypothesis, each such sequence has at least d from $\{L, R\}$. Also, each of the 2^d-many $\{L, R\}$-sequences must be a subsequence of a unique leaf sequence. Hence there are at least 2^d leaves. □

Proof of Theorem 7: By Chebyshev's inequality, for any $t \in \mathbb{R}_{>0}$, $\mathbb{P}(|N_{\pi,n} - \mu| \geqslant t\sigma) \leqslant 1/t^2$, and hence

$$\mathbb{P}(N_{\pi,n} \geqslant \mu - t\sigma) \geqslant 1 - 1/t^2.$$

For $\pi = 2143$, $\mu = \binom{n}{4}/4!$. Let $t = n^\gamma$ for the fixed choice $0 < \gamma < \frac{1}{2}$. Thus, we obtain

$$\mathbb{P}\left(\frac{N_{2143,n}}{2n^3 + 3n^2 + n} \geqslant \frac{\binom{n}{4}/4! - n^\gamma\sigma}{2n^3 + 3n^2 + n}\right) \geqslant 1 - \frac{1}{n^{2\gamma}}. \tag{15}$$

Define a random variable $Q : S_n \to \mathbb{Z}_{\geqslant 0}$ by

$$Q(w) = \min_{u \in \mathcal{L}(w)} \#\{v \text{ appears in a path from } w \text{ to } u \text{ in } \mathcal{T}(w) : \exists v' \neq v'', v \to v', v \to v''\}.$$

By Proposition 8 and Lemma 11,

$$Q(w) \geqslant \frac{N_{2143,n}(w)}{2n^3 + 3n^2 + n}. \tag{16}$$

Combining (15) and (16) gives

$$\mathbb{P}\left(Q \geqslant \frac{\binom{n}{4}/4! - n^\gamma\sigma}{2n^3 + 3n^2 + n}\right) \geqslant 1 - \frac{1}{n^{2\gamma}}.$$

By [20, Section 4.1], the r-th central moment for $N_{\pi,n}$, i.e., $\mathbb{E}[(N_{\pi,n} - \mathbb{E}(N_{\pi,n}))^r]$, is a polynomial in n of degree $\lfloor r(k - \frac{1}{2})\rfloor$ where, recall, $\pi \in S_k$. Hence $\mathsf{Var}(N_{2143,n}) \in O(n^7)$ and $\sigma \in O(n^{3.5})$. Therefore there exists $\alpha > 0$ such that for n sufficiently large

$$\mathbb{P}(Q \geqslant \alpha n) \geqslant 1 - \frac{1}{n^{2\gamma}}. \tag{17}$$

Finally, by Lemma 6 and Lemma 13,

$$\mathsf{EG}(w) = \#\mathcal{L}(w) \geqslant 2^{Q(w)}. \tag{18}$$

The desired equality holds by (17) and (18) combined. □

3.2 Remarks

M. Bona [7] proves that the sequence of random variables

$$\widetilde{X_n} := \frac{N_{2143,n} - \mathbb{E}[N_{2143,n}]}{\sqrt{\mathsf{Var}(N_{2143,n})}}$$

is asymptotically normal, i.e., X_n converges in distribution to the standard normal variable $N(0,1)$. In particular, this means that for any $\epsilon > 0$, for any $a, b \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, $|\mathbb{P}(\widetilde{X}_n \in [a,b]) - \mathbb{P}(N(0,1) \in [a,b])| < \epsilon$. Thus one could use Bona's theorem to prove a more refined version of Theorem 7. However, this does not affect our basic conclusions, so we opted to state a result/proof that only appeals to Chebyshev's inequality.

Using the relations

$$s_i s_j = s_j s_i \text{ for } |i-j| \geqslant 2, \text{ and } \; s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \tag{19}$$

one can transform between any two reduced words

$$s_{i_1} s_{i_2} \cdots s_{i_\ell} \iff s_{j_1} s_{j_2} \cdots s_{j_\ell} \in \mathsf{Red}(w);$$

see, e.g., [29, Proposition 2.1.6]. Hence, it follows that

$$\{i_1, i_2, \ldots, i_\ell\} = \{j_1, j_2, \ldots, j_\ell\}. \tag{20}$$

Let $\sigma^{(n)} = 214365 \cdots 2n \; 2n-1 \in S_{2n}$. This next fact is well-known to experts (for example, an anonymous referee states one can derive it from [28, Theorem 9]). We give a proof here for sake of completeness and make no claims of originality.

Proposition 14. $a_{\sigma^{(n)},\lambda} = f^\lambda$.

Proof. Fix any partition λ of size $2n-1$. Consider any row and column increasing filling T of λ, using each of the labels $\{1, 3, 5, \ldots, 2n-1\}$ precisely once. Let $\mathcal{A}_\lambda$ be the set of these tableaux. Also, let $\mathcal{B}_\lambda$ be the set of EG tableaux for the coefficient $a_{\sigma^{(n)},\lambda}$. $\mathsf{Red}(\sigma^{(n)})$ consists of all $n!$ rearrangements of the factors of $s_1 s_3 \cdots s_{2n-1}$. Hence, the column reading word of any $T \in \mathcal{A}_\lambda$ gives a reduced word for w. Thus, $\mathcal{A}_\lambda \subseteq \mathcal{B}_\lambda$. By (20), if $S \in \mathcal{B}_\lambda$, it must use each label of $\{1, 3, 5, \ldots, 2n-1\}$ exactly once. Since S must also be row and column increasing, we see $S \in \mathcal{A}_\lambda$. This gives $\mathcal{A}_\lambda = \mathcal{B}_\lambda$.

Given $T \in \mathcal{A}_\lambda (= \mathcal{B}_\lambda)$, let $\phi(T) \in \mathsf{SYT}(\lambda)$ be the standard Young tableau of shape λ obtained by sending label i in T to $\lceil \frac{i}{2} \rceil$. Clearly, $\phi : \mathcal{A}_\lambda \to \mathsf{SYT}(\lambda)$ is a bijection. Hence $a_{\sigma^{(n)},\lambda} = \#\mathcal{A}_\lambda = \#\mathsf{SYT}(\lambda) = f^\lambda$. ☐

Let $\mathsf{inv}(n)$ be the number of involutions of S_n. The following shows that the worst case and average case running time of transition is quite different:

Corollary 15. $\#\mathcal{L}(\sigma^{(n)}) = \mathsf{inv}(n) \sim \left(\frac{n}{e}\right)^{n/2} \frac{e^{\sqrt{n}}}{(4e)^{\frac{1}{4}}}$.

Proof. The equality holds since

$$\#\mathcal{L}(\sigma^{(n)}) = \mathsf{EG}(\sigma^{(n)}) = \sum_{\lambda} f^{\lambda} = \mathsf{inv}(n). \tag{21}$$

The first equality of (21) is Lemma 6, the second is Proposition 14 and the third is textbook (e.g., [37, Corollary 7.13.9]). The asymptotic statement is [23, Section 5.1.4]. □

Conjecture 16. $a_{w,\lambda} \leqslant f^{\lambda}$.

Since the original preprint version of this paper was posted to the arXiv, this conjecture has been proved by G. Orelowitz [31].

4 Counting Hecke words

A sequence $(i_1, i_2, \ldots, i_N)$ is a *Hecke word* for $w \in S_n$ if $s_{i_1} \star s_{i_2} \star \cdots \star s_{i_N} = w$ where $\star$ is the *Demazure product* defined by

$$u \star s_i = \begin{cases} us_i & \text{if } \ell(us_i) = \ell(u) + 1 \\ u & \text{otherwise.} \end{cases}$$

Therefore, $N \geqslant \ell(w)$. Let $\mathsf{Hecke}(w, N)$ denote the set of Hecke words for w of length N.

4.1 Two generalizations of the Edelman-Greene formula (2)

We now give two formulas for computing $\mathsf{Hecke}(w, N)$. Both are known to experts, but we are unaware of any specific place that they appear in the literature.

We need three (stable) Grothendieck polynomial formulas from the literature. For the purposes of this paper, the reader may take these formulas as definitions.

First, S. Fomin-A. N. Kirillov [15] prove the following combinatorial formula for the *stable Grothendieck polynomial* G_w:

$$G_w = \sum_{(\mathbf{i},\mathbf{j})} (-1)^{\ell(w)-|\mathbf{j}|} \mathbf{x}^{\mathbf{j}}, \tag{22}$$

where $\mathbf{i} = (i_1, \ldots, i_N) \in \mathsf{Hecke}(w, N)$, and $\mathbf{j} = (j_1 \leqslant j_2 \leqslant \cdots \leqslant j_N)$ are positive integers satisfying $j_t < j_{t+1}$ whenever $i_t \leqslant i_{t+1}$. This is a formal power series in $x_1, x_2, \ldots$.

Second, S. Fomin-C. Greene [14, Theorem 1.2] states that, up to change of conventions,

$$G_w = \sum_{\lambda} (-1)^{|\lambda|-\ell(w)} b_{w,\lambda} s_{\lambda}, \tag{23}$$

where $b_{w,\lambda}$ be the number of row strictly increasing and column weakly increasing tableaux of shape λ whose top to bottom, right to left, column reading word is a Hecke word for w.

Third, C. Lenart [27] gave an expression for the *symmetric Grothendieck polynomial*:

$$G_{\mu}(x_1, x_2, \ldots, x_t) = \sum_{\lambda} (-1)^{|\lambda|-|\mu|} g_{\mu,\lambda} s_{\lambda}(x_1, \ldots, x_t) \tag{24}$$

where $\mu \subseteq \lambda \subseteq \widehat{\mu}$. Here $\widehat{\mu}$ is the unique maximal partition with t rows obtained by adding at most $i-1$ boxes to row i of μ for $2 \leqslant i \leqslant t$. In addition, $g_{\mu,\lambda}$ counts the number of *Lenart tableaux*, i.e., column and row strict tableaux of shape μ/λ with entries in the i-th row restricted to $1, 2, \ldots, i-1$ for each i.

Since $\mathsf{Hecke}(w, \ell(w)) = \mathsf{Red}(w)$, our first formula (25) below generalizes (2). Our second point is that in contrast with (5), even for vexillary permutations, (25) is not short.

Proposition 17. *There is a manifestly nonnegative combinatorial formula*

$$\#\mathsf{Hecke}(w, N) = \sum_{\lambda, |\lambda|=N} b_{w,\lambda} f^{\lambda}. \tag{25}$$

Let $M \geqslant 1$. There is a vexillary permutation $\pi \in S_{2M}$ with $\ell(\pi) = M^2$ such that

$$\#\{\lambda \in \mathsf{par}(M^2 + M) : b_{\pi,\lambda} > 0\} \geqslant \mathsf{par}(M),$$

where $\mathsf{par}(M)$ is the number of partitions of size M. That is when $w = \pi$ and $N = M^2+M$, (25) has at least $\mathsf{par}(M)$-many terms. Moreover,

$$\sum_{\lambda:|\lambda|=M^2+M} b_{\pi,\lambda} \geqslant \mathsf{inv}(M).$$

Proof. Looking at (22), for any $\mathbf{i} = \mathsf{Hecke}(w, N)$, the sequence $(1, 2, \ldots, N)$ can be used for $\mathbf{j}$. Hence,

$$(-1)^{N-\ell(w)} \#\mathsf{Hecke}(w, N) = [x_1 x_2 \cdots x_N] G_w. \tag{26}$$

Combining (26) and (23) gives (25).

Pick $\mu = M \times M$ and fix $t \geqslant M^2 + M$. Therefore $\widehat{\mu} = t \times M$. Hence, by (24),

$$(-1)^M [x_1 \cdots x_{M^2+M}] G_{M\times M}(x_1, \ldots, x_t) = \sum_{\lambda} g_{M\times M,\lambda} f^{\lambda}. \tag{27}$$

Here the sum is over $\mu \subseteq \lambda \subseteq \widehat{\mu}$ with $|\lambda| = M^2 + M$. Now, each such λ is of the form $(M \times M, \overline{\lambda})$ where $\overline{\lambda} \in \mathsf{par}(M)$ and is contained in $M \times M$. Notice that $g_{M\times M,\lambda} \geqslant f^{\overline{\lambda}}$, since for each such λ we can obtain a Lenart-tableau by filling the $\overline{\lambda}$ part with $1, 2, \ldots, M$ to obtain a standard tableau, in all possible ways. Hence, using [37, Corollary 7.13.9],

$$\sum_{\lambda \in \mathsf{Par}(M^2+M)} g_{M\times M,\lambda} \geqslant \sum_{\overline{\lambda}:|\overline{\lambda}|=M} f^{\overline{\lambda}} = \mathsf{inv}(M).$$

Finally, let $\pi = M+1, M+2, \ldots, 2M, 1, 2, 3 \ldots, M \in S_{2M}$. This is a vexillary permutation π with $\lambda(\pi) = M \times M$. By, e.g., [24, Lemma 5.4],

$$G_\pi(x_1, \ldots, x_t, 0, 0, \ldots,) = G_{M\times M}(x_1, \ldots, x_t, 0, 0, \ldots).$$

Since the Schur polynomials form a basis of the ring of symmetric polynomials, the right-hand sides of (24) and (23) coincide, i.e., $b_{\pi,\lambda} = g_{M\times M,\lambda}$ for every λ. The result follows. ☐

Example 18. Let $w = 31524 = s_4s_2s_3s_1$. Using (25) we obtain

$$\#\mathsf{Hecke}(w,5) = \left(\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 1 & 3 & \\ \cline{1-2} \end{array}, \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 4 & \\ \cline{1-2} \end{array} \right) f^{3,2} + \left(\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} 3 \\ \cline{1-1} \end{array}, \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 1 \\ \cline{1-1} 3 \\ \cline{1-1} \end{array} \right) f^{3,1,1}$$

$$+ \left(\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 3 \\ \cline{1-1} \end{array}, \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 1 & 4 \\ \hline 3 \\ \cline{1-1} \end{array} \right) f^{2,2,1}$$

$$= 2f^{3,2} + 2f^{3,1,1} + 2f^{2,2,1} = 32,$$

which the reader may confirm by direct check. □

For our second formula, we use work of A. Buch, A. Kresch, M. Shimozono, H. Tamvakis and the third author [10] that proves

$$G_w = \sum_{\lambda} (-1)^{\ell(w)-|\lambda|} c_{w,\lambda} G_\lambda \quad \text{where} \quad G_\lambda = \sum_{T} (-1)^{|T|-|\lambda|} \mathbf{x}^T \tag{28}$$

and the latter sum is over all semistandard set-valued tableaux of shape λ [8]. Above, $c_{w,\lambda}$ is the number of row and column strict tableaux of shape λ whose top to bottom, right to left, column reading word is a Hecke word for w.

This next generalization of (2) is also manifestly nonnegative. It specializes in the vexillary case in a tantalizing way.

Proposition 19.

$$\#\mathsf{Hecke}(w,N) = \sum_{\lambda:\ell(w)\leqslant|\lambda|\leqslant N} c_{w,\lambda} f^{\lambda,N}. \tag{29}$$

If w is vexillary, then

$$\#\mathsf{Hecke}(w,N) = f^{\lambda(w),N}.$$

Proof. In view of (26) and (28) we have

$$\begin{aligned} (-1)^{N-\ell(w)} \#\mathsf{Hecke}(w,N) &= [x_1x_2\cdots x_N]G_w \\ &= [x_1x_2\cdots x_N] \sum_{\lambda} (-1)^{\ell(w)-|\lambda|} c_{w,\lambda} G_\lambda \\ &= \sum_{\lambda:|\lambda|\leqslant N} (-1)^{\ell(w)-|\lambda|} c_{w,\lambda} [x_1x_2\cdots x_N] G_\lambda \\ &= \sum_{\lambda:|\lambda|\leqslant N} (-1)^{\ell(w)-|\lambda|} c_{w,\lambda} (-1)^{N-|\lambda|} f^{\lambda,N} \\ &= \sum_{\lambda:|\lambda|\leqslant N} (-1)^{N+\ell(w)} c_{w,\lambda} f^{\lambda,N}, \end{aligned}$$

proving (29). For the second statement, by [34, Lemma 5.4], when w is vexillary then $G_w = G_\lambda$, and the above sequence of equalities simplifies, as desired. □

Example 20. Again let $w = 31524$ as in Example 18. Now applying (29) gives

$$\#\mathsf{Hecke}(w,5) = \begin{pmatrix} \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \end{pmatrix} f^{(2,2),5} + \begin{pmatrix} \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 \\ \cline{1-1} \end{array} \end{pmatrix} f^{(3,1),5} + \begin{pmatrix} \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 4 \\ \cline{1-2} \end{array} \end{pmatrix} f^{(3,2),5}$$
$$= 10 + 17 + 5 = 32,$$

in agreement with Example 18. One can check the $f^{\lambda,N}$ computations either directly, or by using

$$f^{\lambda,N} = [x_1 \cdots x_N] G_\lambda \tag{30}$$

combined with (24). □

Proposition 19 is our central motivation for Problem 40.

Proposition 21. *Fix k. There is an $|\mu|^{O(1)}$ algorithm to compute $f^{\mu,N}$ where $N \leqslant |\mu|+k$.*

Proof. We use (30) combined with (24) and describe the possible Lenart tableaux. First, we look for $\mu \subseteq \lambda \subseteq \overline{\mu}$ where $|\lambda| = |\mu| + k$. Such λ correspond to a choice of k rows $r_1 \leqslant r_2 \leqslant \ldots \leqslant r_k$ to add a box, among $\ell(\mu)+k-1$ choices (rows $2, 3, \ldots, \ell(\mu)+k$). There are $\binom{\ell(\mu)+2k-2}{k} \in |\mu|^{O(1)}$ many ways to do this. For each such choice, it takes constant time to verify that λ is a partition. For those cases, we construct a possible Lenart tableau T by filling row r_i in at most $(r_i - 1)^k$ ways. Since $r_i \leqslant \ell(\mu) + k$, there are $|\mu|^{O(1)}$-many possible row strictly increasing tableaux T. It remains to determine if T is actually a Lenart tableau, which takes constant time (since k is fixed). Finally, to each tableau, we must compute f^λ via (3). This takes $|\lambda|^{O(1)}$-time. Now, since $|\lambda| = |\mu| + k$ and k is fixed, it also takes $|\mu|^{O(1)}$-time. Moreover,

$$\log(f^\lambda) \leqslant \log |\lambda|! \in O(|\mu| \log |\mu|).$$

Hence, summing the at most $\binom{\ell(\mu)+2k-2}{k}$ hook-length calculations, also takes $|\mu|^{O(1)}$-time, as desired. □

Example 22. We elaborate on the proof of Theorem 21 in the case $k = 2$. Here we look for $\mu \subseteq \lambda \subseteq \overline{\mu}$ where $|\lambda| = |\mu| + 2$. Such λ correspond to a choice of two rows $r_1 \leqslant r_2$ to add a box, among $\ell(\mu) + 1$ choices (rows $2, 3, \ldots, \ell(\mu) + 2$). If $r_1 = r_2$, $g_{\mu,\lambda} = \binom{r_2-1}{2}$. Otherwise if $r_1 < r_2$, there are $\binom{\ell(\mu)+1}{2}$ many choices. Assuming λ is a partition, there are two cases. If the two boxes are in different columns $g_{\mu,\lambda} = (r_1 - 1)(r_2 - 1)$. Otherwise, if they are in the same column (and hence $r_2 = r_1 + 1$), then $g_{\mu,\lambda} = \binom{r_2-1}{2}$. Now apply (24).

For $\lambda = \delta_{100} = (100, 99, \ldots, 3, 2, 1)$ and $N = \binom{100}{2}+2$, this procedure exactly computes $f^{\lambda,N} = \#\mathsf{Hecke}(w_0, N) = 3.75\ldots \times 10^{7981}$. □

4.2 Application to Euler characteristics of Brill-Noether varieties (after [2, 11])

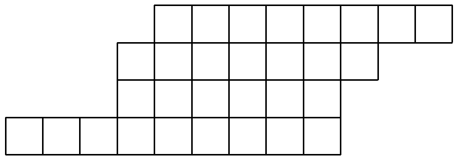
Counting standard set-valued tableaux has been given geometric impetus through work of [2, 11] on *Brill-Noether varieties.* More precisely, following [11, Definition 1.2], let

$g, r, d \in \mathbb{Z}_{\geqslant 0}$. Suppose $\alpha = (\alpha_0 \leqslant \alpha_1 \leqslant \cdots \leqslant \alpha_r)$ and $\beta = (\beta_0 \leqslant \beta_1 \leqslant \cdots \leqslant \beta_r)$ be sequences in $\mathbb{Z}_{\geqslant 0}^{r+1}$. Let $\mathsf{CP} = \mathsf{CP}(g, r, d, \alpha, \beta)$ be the skew Young diagram with boxes

$$\{(x, y) \in \mathbb{Z}^2 \ : \ 0 \leqslant y \leqslant r, -\alpha_y \leqslant x < g - d + r + \beta_{r-y}\}.$$

Example 23. If $g = 45, d = 43, r = 3, \alpha = (0, 1, 1, 4), \beta = (0, 0, 1, 3)$ then

$\mathsf{CP} =$



□

Let $\chi(G_d^{r,\alpha,\beta}(X, p, q))$ be the algebraic Euler characteristic of the Brill-Noether variety $G_d^{r,\alpha,\beta}(X, p, q)$.

Theorem 24 (M. Chan-N. Pfleuger [11])**.** $(-1)^{g-|\mathsf{CP}|}\chi(G_d^{r,\alpha,\beta}(X, p, q)) = f^{\mathsf{CP},g}$.

Given λ/μ, construct a (321-avoiding) permutation $w_{\lambda/\mu}$ by filling all boxes in the same northwest-southeast diagonal with the same entry, starting with 1 on the northeastmost diagonal and increasing consecutively as one moves southwest. Call this filling $T_{\lambda/\mu}$. Let $(r_1, r_2, \ldots, r_{|\lambda/\mu|})$ be the left-to-right, top-to-bottom, row reading word of $T_{\lambda/\mu}$. Define $w_{\lambda/\mu} = s_{r_1} s_{r_2} \cdots s_{r_{|\lambda/\mu|}}$.

Example 25. If $\lambda/\mu = (12, 10, 9, 9)/(4, 3, 3, 0)$ be CP from Example 23. Then

$$T_{\lambda/\mu} = \begin{array}{cccccccccccc} & & & & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ & & & 10 & 9 & 8 & 7 & 6 & 5 & 4 & & \\ & & & 11 & 10 & 9 & 8 & 7 & 6 & & & \\ 15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & & & \end{array}$$

The reading word is

$$(8, 7, 6, 5, 4, 3, 2, 1, 10, 9, 8, 7, 6, 5, 4, 11, 10, 9, 8, 7, 6, 15, 14, 13, 12, 11, 10, 9, 8, 7)$$

and $w_{\lambda/\mu} = 9, 1, 2, 11, 3, 12, 16, 4, 5, 6, 7, 8, 10, 13, 14, 15 \in S_{16}$. □

Earlier, in [2], a *cancellative* combinatorial formula for $\chi(G_d^{r,\alpha,\beta}(X, p, q))$ is given. In [11], another cancellative formula is given, as a signed sum involving counts of (ordinary) skew standard Young tableaux. See [11, Theorem 6.6] (and the discussion of [14] in [11, Section 6]) as well as [2, Theorem A, Theorem C].

Proposition 26. *Let w_{CP} is a 321-avoiding permutation defined as above. Then:*

$$(-1)^{g-|\mathsf{CP}|}\chi(G_d^{r,\alpha,\beta}(X, p, q)) = \sum_{\lambda, |\lambda|=g} b_{w_{\mathsf{CP}},\lambda} f^{\lambda},$$

and

$$(-1)^{g-|\mathsf{CP}|}\chi(G_d^{r,\alpha,\beta}(X, p, q)) = \sum_{\lambda : |CP| \leqslant |\lambda| \leqslant g} c_{w_{\mathsf{CP}},\lambda} f^{\lambda,g}.$$

Proof. By [8, Section 2], $G_{w_{\nu/\mu}} = G_{\nu/\mu} := \sum_T (-1)^{|T|-|\nu/\mu|}\mathbf{x}^T$, where the sum is over semistandard set-valued tableaux of skew shape ν/μ. Therefore if $f^{\nu/\mu,N}$ is the number of standard set-valued tableaux of this shape with N entries then

$$f^{\nu/\mu,N} = \#\mathsf{Hecke}(w_{\nu/\mu}, N). \tag{31}$$

Now combine this with Theorem 24, Proposition 17 and 19. ☐

The second formula expresses $(-1)^{g-|\mathsf{CP}|}\chi(G_d^{r,\alpha,\beta}(X,p,q))$ as a cancellation-free sum of Euler characteristics of other Brill-Noether varieties. *Is there a geometric explanation of this?*

5 Three importance sampling algorithms

5.1 Estimating $\#\mathsf{Red}(w)$

Define a random variable Y_w for $w \in S_n$, as follows:

---

if w is vexillary then
  $\mathsf{Y}_w = f^{\lambda(w)}$
else
  $C = \{w' \text{ is a child of } w \text{ in } \mathcal{T}(w)\}$
  Choose $W' \in C$ uniformly at random
  $\mathsf{Y}_w = \#C \times \mathsf{Y}_{w'}$

Proposition 27. *Let $w \in S_n$. Then $\mathbb{E}(\mathsf{Y}_w) = \#\mathsf{Red}(w)$.*

Proof. We induct on $h = h(w) \geqslant 0$, the height of $\mathcal{T}(w)$, i.e., the maximum length of any path from the root to a leaf. In the base case, $h = 0$, w is vexillary and thus, by (5),

$$\mathbb{E}(\mathsf{Y}_w) = f^{\lambda(w)} = \#\mathsf{Red}(w).$$

Our induction hypothesis is that $\mathbb{E}(\mathsf{Y}_u) = \#\mathsf{Red}(u)$ whenever $h(u) < h(w)$. Now

$$\begin{aligned}
\mathbb{E}(\mathsf{Y}_w) &= \sum_{w'\in C} \mathbb{E}(\mathsf{Y}_w | W' = w')\mathbb{P}(W' = w') \\
&= \frac{1}{\#C} \sum_{w'\in C} \mathbb{E}(\mathsf{Y}_w | W' = w') \\
&= \frac{1}{\#C} \sum_{w'\in C} \mathbb{E}(\#C \times \mathsf{Y}_{w'}) \\
&= \sum_{w'\in C} \mathbb{E}(\mathsf{Y}_{w'}) \\
&= \sum_{w'\in C} \#\mathsf{Red}(w') \quad \text{(induction hypothesis)} \\
&= \#\mathsf{Red}(w).
\end{aligned}$$

The last equality is by construction of the transition algorithm and Theorem 5. $\square$

Example 28. Let $w = 43817625 \in S_8$. We have the following sequence of transition steps

$$43817625 \xrightarrow{3} 53817426 \xrightarrow{1} 53827146 \xrightarrow{3} 63825147 \xrightarrow{2} 63842157 \xrightarrow{2} 73642158.$$

The number of children is indicated at each stage. The final permutation is vexillary, and $f^{\lambda(73642158)} = f^{6,4,2,2,1} = 243243$. Hence one sample is $3 \times 1 \times 3 \times 2 \times 2 \times 243243 = 8756748$. Using sample size 2×10^3 gives an estimate of $2.09(\pm 0.04) \times 10^6$, versus $\#\mathsf{Red}(w) = 2085655$.[2] $\square$

Example 29. Let $w = \sigma^{(n)} = 2143 \cdots 2n\ 2n-1$. When $n = 10$ (so $\sigma^{(n)} \in S_{20}$), using sample size 10^5 gives an estimate of $3.63(\pm 0.02) \times 10^6$, which is close to the exact value $10! = 3628800$. When $n = 30$ ($\sigma^{(n)} \in S_{60}$), using sample size 2×10^6 one estimates $2.18(\pm 0.49) \times 10^{32}$ whereas $30! = 2.65 \ldots \times 10^{32}$. $\square$

Example 30 (Estimating the number of skew standard Young tableaux)**.** We continue Example 25. Let $f^{\lambda/\mu}$ be the number of standard Young tableaux of shape λ/μ. By a result of S. Billey-W. Jockusch-R. P. Stanley [4, Corollary 2.4], $F_{w_{\lambda/\mu}} = s_{\lambda/\mu}$. Taking the coefficient of $x_1 x_2 \cdots x_{|\lambda/\mu|}$ on both sides implies $\#\mathsf{Red}(w_{\lambda/\mu}) = f^{\lambda/\mu}$. One has the textbook determinantal formula

$$f^{\lambda/\mu} = |\lambda/\mu|! \det \left(\frac{1}{(\lambda_i - \mu_j - i + j)!} \right)_{i,j=1}^{t}. \tag{32}$$

So $f^{\lambda/\mu} = 73064598262110 \approx 7.31 \times 10^{13}$. A 10^4 sample size estimate is $7.30(\pm 0.04) \times 10^{13}$. $\square$

5.2 Estimating $\#\mathsf{Hecke}(w, N)$

We propose a different importance sampling algorithm, to compute $\#\mathsf{Hecke}(w, N)$. For $N < \ell(w)$ the random variable $\mathsf{Z}_{w,N}$ is equal to 0 and for $N \geqslant \ell(w)$, it is recursively defined by:

---

if $w = id$ then
    if $N = 0$ then $\mathsf{Z}_{w,N} = 1$ else $\mathsf{Z}_{w,N} = 0$
else
    $D = \{i : w(i) > w(i+1)\}$
    Choose $I \in D$ and $\theta \in \{0, 1\}$ independently and uniformly at random
    if $\theta = 0$ then $\mathsf{Z}_{w,N} = 2\#D \times \mathsf{Z}_{w,N-1}$ else $\mathsf{Z}_{w,N} = 2\#D \times \mathsf{Z}_{ws_I,N-1}$

Proposition 31. *Let $w \in S_n$ and $N \geqslant \ell(w)$. Then $\mathbb{E}(\mathsf{Z}_{w,N}) = \#\mathsf{Hecke}(w, N)$.*

[2] The "($\pm$0.04)" refers to the standard error of the mean. All estimates are based on twelve trials of an indicated sample size. Code is available at https://github.com/ICLUE/reduced-word-enumeration

Proof. First we claim

$$\#\mathsf{Hecke}(w,N)=\begin{cases}1 & \text{if } w=id \text{ and } N=0\\ 0 & \text{if } w=id \text{ and } N>0\\ \sum_{i\in D}(\#\mathsf{Hecke}(ws_i,N-1)+\#\mathsf{Hecke}(w,N-1)) & \text{otherwise.}\end{cases} \tag{33}$$

The unique Hecke word for $w=id$ is the empty word; this explains the first two cases.

Thus assume $w\neq id$ and $N\geqslant \ell(w)$. Suppose that $(i_1,i_2,\dots,i_N)\in \mathsf{Hecke}(w,N)$.

Claim 32. *i_N is the position of a descent of w, i.e., $w(i_N)>w(i_N+1)$.*

Proof of Claim 32: Consider $w':=s_{i_1}\star s_{i_2}\star\cdots\star s_{i_{N-1}}$. Either $\ell(w')=\ell(w)$ or $\ell(w')=\ell(w)-1$. In the former case then if i_N is the position of an ascent of $w'=w$ then $w=w'\star s_{i_N}$ would create a descent at that position, a contradiction. In the latter case, w' had an ascent at position i_N which becomes a descent in $w'\star s_{i_N}=w's_{i_N}$. □

Claim 32 implies the existence of a bijection

$$\mathsf{Hecke}(w,N)\xrightarrow{\sim}\left(\bigcup_{i\in D}\mathsf{Hecke}(ws_i,N-1)\times\{i\}\right)\cup\left(\bigcup_{i\in D}\mathsf{Hecke}(w,N-1)\times\{i\}\right), \tag{34}$$

defined by $(i_1,i_2,\dots,i_{N-1},i_N)\in\mathsf{Hecke}(w,N)\mapsto((i_1,i_2,\dots,i_{N-1}),i_N)$.[3] Therefore, by taking cardinalities on both sides of (34) we obtain the third case of (33).

Returning to proposition itself, we induct on $N\geqslant 0$. The case $N=0$ holds by the first case of (33) and the definition $Z_{w,N}=0$ if $N<\ell(w)$. For $N>0$,

$$\begin{aligned}\mathbb{E}(\mathsf{Z}_{w,N}) &= \sum_{i\in D}\mathbb{E}(\mathsf{Z}_{w,N}|I=i,\theta=0)\mathbb{P}(I=i)\mathbb{P}(\theta=0)\\ &\quad+\sum_{i\in D}\mathbb{E}(\mathsf{Z}_{w,N}|I=i,\theta=1)\mathbb{P}(I=i)\mathbb{P}(\theta=1)\\ &=\sum_{i\in D}\mathbb{E}(2\#D\times\mathsf{Z}_{w,N-1})\frac{1}{\#D}\times\frac{1}{2}+\sum_{i\in D}\mathbb{E}(2\#D\times\mathsf{Z}_{ws_i,N-1})\frac{1}{\#D}\times\frac{1}{2}\\ &=\sum_{i\in D}(\mathbb{E}(\mathsf{Z}_{w,N-1})+\mathbb{E}(\mathsf{Z}_{ws_i,N-1}))\\ &=\sum_{i\in D}(\#\mathsf{Hecke}(w,N-1)+\#\mathsf{Hecke}(ws_i,N-1))\\ &=\#\mathsf{Hecke}(w,N),\end{aligned}$$

where we have applied induction (on N) and the third case of (33). □

Example 33. One can explicitly generate all 2030964 elements of $\mathsf{Hecke}(351624,13)$. A 2000 sample size estimate is $2.04(\pm 0.10)\times 10^6$. □

[3] If $N=\ell(w)$, then $\mathsf{Hecke}(w,N)=\mathsf{Red}(w)$ and $\mathsf{Hecke}(w,N-1)=\emptyset$. In this case, (34) reduces to the bijection $\mathsf{Red}(w)\xrightarrow{\sim}\bigcup_{i\in D}\mathsf{Red}(ws_i)\times\{i\}$.

Example 34. By [34, Corollary 1.3],

$$\#\mathsf{Hecke}\left(w_0, \binom{n}{2}+1\right) = \frac{\binom{n}{2}\left[\binom{n}{2}+1\right]}{n} \times \#\mathsf{Red}(w_0). \tag{35}$$

For $n = 10$, $\#\mathsf{Hecke}(w_0, 46) = 5.65\ldots \times 10^{28}$. Using sample size 10^8, we obtained an estimate of $\approx 4.26(\pm 1.94) \times 10^{28}$. ☐

The Z-algorithm restricts to an algorithm to compute $\#\mathsf{Red}(w)$. However, the Y-algorithm of Subsection 5.1 sometimes has better convergence in this case. This suggests a "hybrid" algorithm. Define $\mathsf{H}_{w,N}$ to be 0 if $N < \ell(w)$. Otherwise,

---

if $N = \ell(w)$ then $\mathsf{H}_{w,N} = Y_w$
else if $w = id$ then
    if $N = 0$ then $\mathsf{H}_{w,N} = 1$ else $\mathsf{H}_{w,N} = 0$
else
    $D = \{i : w(i) > w(i+1)\}$
    Choose $I \in D$ and $\theta \in \{0, 1\}$ independently and uniformly at random
    if $\theta = 0$ then $\mathsf{H}_{w,N} = 2\#D \times \mathsf{H}_{w,N-1}$ else $\mathsf{H}_{w,N} = 2\#D \times \mathsf{H}_{ws_I,N-1}$

Proposition 35. *Let $w \in S_n$. Then $\mathbb{E}[\mathsf{H}_{w,N}] = \#\mathsf{Hecke}(w, N)$.*

We omit the proof, as it is a straightforward modification of the argument for Proposition 31, using Proposition 27.

Example 36. Let $w = 361824795 \in S_9$; hence $\ell(w) = 12$. Using sample size 10^6 with the Z algorithm gives $\#\mathsf{Hecke}(w, 25) \approx 5.98(\pm 0.04)\times 10^{16}$. The estimate from the H algorithm (with the same sample size) is $\#\mathsf{Hecke}(w, 25) \approx 6.02(\pm 0.08)\times 10^{16}$. For Example 34, with 10^8 samples, the H algorithm estimates $\#\mathsf{Hecke}(w_0, 46)$ as $6.09(\pm 4.69) \times 10^{28}$. ☐

Example 37. We use Proposition 21 to compute $\#\mathsf{Hecke}(w_0, \binom{n}{2} + 2)$. When $n = 7$, $\#\mathsf{Hecke}(w_0, 23) = 2.54\ldots \times 10^{12}$. A 10^6 sample size estimate is $2.60(\pm 0.22) \times 10^{12}$. For $n = 10$, $\#\mathsf{Hecke}(w_0, 47) = 6.01\ldots \times 10^{30}$. A 10^8 sample size estimate is $\approx 4.04(\pm 2.17) \times 10^{30}$. ☐

Example 38 (Skew set-valued tableaux)**.** To estimate $f^{\lambda/\mu,N}$ for

$$\lambda/\mu = (12, 10, 9, 9)/(4, 3, 3, 0) \text{ and } N = 45,$$

we use (31) and the Z-algorithm with sample size 10^7 to predict

$$f^{\lambda/\mu,45} = \#\mathsf{Hecke}(w_{\lambda/\mu}, 45) \approx 1.30(\pm 0.03) \times 10^{33}.$$

This is backed by the estimate $1.29(\pm 0.06) \times 10^{33}$ using the H-algorithm with sample size 10^6. We have thus estimated the value of $(-1)^{g-|\mathsf{CP}|}\chi(G_d^{r,\alpha,\beta}(X, p, q))$ for the parameters of Example 23. There are a number of ways to theoretically compute this value ([2], [11], Proposition 26). *What is the exact value?* ☐

6 Remarks and questions about computational complexity

The exponential average run-time of transition (Theorem 2) does not imply computing $\#\mathsf{Red}(w)$ is hard. Suppose one encodes a permutation w by its *Lehmer code* $\mathsf{code}(w) = (c_1, c_2, \ldots, c_L)$. *What is the worst case complexity of computing* $\#\mathsf{Red}(w)$ *given input* $\mathsf{code}(w)$*?*

L. Valiant [38] introduced the complexity class $\#\mathsf{P}$ of problems that count the number of accepting paths of a non-deterministic Turing machine running in polynomial time in the length of the input. Let FP be the class of function problems solvable in polynomial time on a deterministic Turing machine. It is basic theory that $\mathsf{FP} \subseteq \#\mathsf{P}$.

Observation 39 (*cf.* [30, Section 3])**.** $\#\mathsf{Red}(w) \notin \#\mathsf{P}$. *In particular,* $\#\mathsf{Red}(w) \notin \mathsf{FP}$.

Proof. Let θ_n be the vexillary permutation with $\mathsf{code}(\theta_n) = (n, n)$. Then, using (5),

$$\#\mathsf{Red}(\theta_n) = f^{(n,n)} = C_n := \frac{1}{n+1}\binom{2n}{n}. \tag{36}$$

The middle equality is textbook: there is a bijection between standard Young tableaux of shape (n, n) and Dyck paths from $(0, 0)$ to $(2n, 0)$; both are enumerated by the *Catalan number* C_n. Now, $\#\mathsf{Red}(\theta_n)$ is *doubly* exponential in the input length $O(\log n)$. No such problem can be in $\#\mathsf{P}$; see [30, Section 3] which also inspired this observation. ($\#\mathsf{Red}(w) \notin \mathsf{FP}$ is true from this argument for the simple reason that it takes exponential time just to write down the output.) □

By Observation 39's reasoning, (36) shows there is no algorithm to compute $f^{\lambda,N}$ that is polynomial-time in the bit-length of the input (λ, N).

A counting problem $\mathcal{P}$ is $\#\mathsf{P}$-hard if any problem in $\#\mathsf{P}$ has a polynomial-time counting reduction to $\mathcal{P}$. *Is* $\#\mathsf{Red}(w) \in \#\mathsf{P}$-$\mathsf{hard}$*?*

Observation 39 is dependent on the choice of encoding. For example, if one encodes a permutation $w \in S_n$ in the inefficient one-line notation, the input takes $O(n \log n)$ space. Since $\ell(w) \leqslant \binom{n}{2}$ is polynomial in the input length, it follows that $\#\mathsf{Red}(w) \in \#\mathsf{P}$; see [35].

Problem 40. Does there exist an $n^{O(1)}$-algorithm to compute $\#\mathsf{Red}(w)$?

It is easy to see that $\#\mathsf{Red}(u) \leqslant \#\mathsf{Red}(us_i)$ whenever $\ell(us_i) = \ell(u) + 1$. Hence, $\#\mathsf{Red}(w)$ is maximized at $w = w_0$. So, by (4), $\log(\#\mathsf{Red}(w)) \in n^{O(1)}$. Thus, unlike Observation 39, there is no easy negative solution to Problem 40 (and any negative solution implies $\mathsf{FP} \neq \#\mathsf{P}$, which is a famous open problem). Indeed, in the vexillary case (5), the hook-length formula (3) gives a $n^{O(1)}$-algorithm for $\#\mathsf{Red}(w)$.

Acknowledgements

We thank Anshul Adve, David Anderson, Alexander Barvinok, Melody Chan, Yuguo Chen, Anna Chlopecki, Michael Engen, Neil Fan, Sergey Fomin, Sam Hopkins, Allen

Knutson, Tejo Nutalapati, Gidon Orelowitz, Colleen Robichaux, Renming Song, John Stembridge, Anna Weigandt and Harshit Yadav for helpful remarks/discussion. We are especially grateful to Brendan Pawlowski for pointing out Theorem 1 appears as Theorem 3.2.7 of [32], as well as other remarks. We also thank the anonymous referee for their careful reading and comments that improved our presentation. This work is part of ICLUE, the Illinois Combinatorics Lab for Undergraduate Experience.

References

[1] A. Adve, C. Robichaux, and A. Yong, *An efficient algorithm for deciding vanishing of Schubert polynomial coefficients*, Adv. Math., **383** (2021), 107669.

[2] D. Anderson, L. Chen, and N. Tarasca, *K-classes of Brill-Noether loci and a determinantal formula*, International Mathematics Research Notices, 2021, rnab025.

[3] S. Billey, *Transition equations for isotropic flag manifolds*, Selected papers in honor of Adriano Garsia (Taormina, 1994). Discrete Math. 193 (1998), no. 1-3, 69–84.

[4] S. Billey, W. Jockusch and R. P. Stanley, *Some combinatorial properties of Schubert polynomials*, J. Algebraic Combin. **2**(1993), no. 4, 345–374.

[5] S. Billey and B. Pawlowski, *Permutation patterns, Stanley symmetric functions, and generalized Specht modules*, J. Combin. Theory Ser. A 127 (2014), 85–120.

[6] J. Blitzstein and P. Diaconis, *A sequential importance sampling algorithm for generating random graphs with prescribed degrees*, Internet Math. 6 (2010), no. 4, 489?522.

[7] M. Bona, *The copies of any permutation pattern are asymptotically normal*, preprint, 2007. arXiv:0712.2792.

[8] A. Buch, *A Littlewood-Richardson rule for the K-theory of Grassmannians*, Acta Math. 189 (2002), no. 1, 37–78.

[9] A. Buch, *Quantum cohomology of partial flag manifolds*, Trans. Amer. Math. Soc. 357 (2005), 443–458.

[10] A. Buch, A. Kresch, M. Shimozono, H. Tamvakis, and A. Yong, *Stable Grothendieck polynomials and K-theoretic factor sequences*, Math. Ann. 340 (2008), no. 2, 359–382.

[11] M. Chan and N. Pflueger, *Euler characteristics of Brill-Noether varieties*, Trans. Amer. Math. Soc. 374 (2021), no. 3, 1513–1533.

[12] S. Chatterjee and P. Diaconis, *The sample size required in importance sampling*, Ann. Appl. Probab. 28 (2018), no. 2, 1099–1135.

[13] P. Edelman and C. Greene, *Balanced tableaux*, Adv. in Math. 63 (1987), no. 1, 42–99.

[14] S. Fomin and C. Greene, *Noncommutative Schur functions and their applications*, Discrete Math. **193**(1998), 179–200, Selected papers in honor of Adriano Garsia (Taormina, 1994).

[15] S. Fomin and A. N. Kirillov, *Grothendieck polynomials and the Yang-Baxter equation*, Formal power series and algebraic combinatorics/Séries formelles et combinatoire algébrique, 183–189, DIMACS, Piscataway, NJ, s.d..

[16] J. S. Frame, Robinson, G. de B. Robinson and R. M. Thrall, *The hook graphs of the symmetric group*, Can. J. Math. 6 (1954), 316–325.

[17] W. Fulton, *Flags, Schubert polynomials, degeneracy loci, and determinantal formulas*, Duke Math. J. 65 (1992), no. 3, 381–420.

[18] A. Garsia, *The saga of reduced factorizations of elements of the symmetric group*, Publications du Laboratoire de Combinatoire et d'Informatique Mathématique, 29, 2002.

[19] Z. Hamaker, E. Marberg and B. Pawlowski, *Schur P-positivity and Involution Stanley Symmetric Functions*, International Mathematics Research Notices, rnx274, 2017.

[20] S. Janson, B. Nakamura, and D. Zeilberger, *On the asymptotic statistics of the number of occurrences of multiple permutation patterns*, J. Comb. 6 (2015), no. 1-2, 117–143.

[21] H. Kahn and T. E. Harris, *Estimation of particle transmission by random sampling*, National Bureau of Standards applied mathematics series, 12 (1951):27–30.

[22] D. Knuth, *Mathematics and computer science: coping with finiteness*, Science, 194(4271):1235–1242, 1976.

[23] D. Knuth, *The art of computer programming*, Volume 3, 2nd ed, Addison Wesley Longman, 1988.

[24] A. Knutson, E. Miller, and A. Yong, *Gröbner geometry of vertex decompositions and of flagged tableaux*, J. Reine Angew. Math. 630 (2009), 1–31.

[25] A. Knutson and A. Yong, *A formula for K-theory truncation Schubert calculus*, Int. Math. Res. Not. 2004, no. 70, 3741–3756.

[26] A. Lascoux and M. -P. Schützenberger, *Schubert polynomials and the Littlewood-Richardson rule*, Letters in Math. Physics 10 (1985), 111–124.

[27] C. Lenart, *Combinatorial aspects of the K-theory of Grassmannians*, Ann. Comb. 4 (2000), no. 1, 67–82.

[28] D. Little, *Factorization of the Robinson-Schensted-Knuth correspondence*, J. Combin. Theory Ser. A 110 (2005), no. 1, 147–168.

[29] L. Manivel, *Symmetric functions, Schubert polynomials and degeneracy loci*. Translated from the 1998 French original by John R. Swallow. SMF/AMS Texts and Monographs, American Mathematical Society, Providence, 2001.

[30] H. Narayanan, *On the complexity of computing Kostka numbers and Littlewood-Richardson coefficients*, J. Alg. Comb., Vol. 24, N. 3, 2006, 347–354.

[31] G. Orelowitz, *Maximizing the Edelman-Green statistic*, preprint, 2019. arXiv:1908.11455.

[32] B. Pawlowski, *Permutation diagrams in symmetric function theory and Schubert calculus*, PhD thesis, University of Washington, 2014.

[33] L. E. Rasmussen, *Approximating the permanent: a simple approach*, Random Structures Algorithms 5 (1994), no. 2, 349–361.

[34] V. Reiner, B. Tenner, and A. Yong, *Poset edge densities, nearly reduced words, and barely set-valued tableaux*, Journal of Combinatorial Theory, Series A Volume 158, August 2018, 66–125.

[35] M. Samuels, *Word posets, complexity, and Coxeter groups*, preprint, 2011. arXiv:1101.4655.

[36] R. P. Stanley, *On the number of reduced decompositions of elements of coxeter groups*, European Journal of Combinatorics **5** (1984), no. 4, 359 – 372.

[37] R. P. Stanley, *Enumerative combinatorics*, Vol. 2. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999. xii+581 pp.

[38] L. G. Valiant, *The complexity of computing the permanent*, Theoret. Comput. Sci., 8(2):189–201, 1979.