ON SEMIRING COMPLEXITY OF SCHUR POLYNOMIALS

SERGEY FOMIN, DIMA GRIGORIEV, DORIAN NOGNENG, AND ÉRIC SCHOST

Abstract. Semiring complexity is the version of arithmetic circuit complexity that allows only two operations: addition and multiplication. We show that semiring complexity of a Schur polynomial $s_\lambda(x_1, \dots, x_k)$ labeled by a partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots)$ is bounded by $O(\log(\lambda_1))$ provided the number of variables k is fixed.

1. Introduction and main results

Let $f(x_1, \dots, x_k)$ be a polynomial with nonnegative integer coefficients. As such, f can be computed using addition and multiplication only—without subtraction or division. To be more precise, one can build an *arithmetic circuit* wherein

- each gate performs an operation of addition or multiplication;
- the inputs are $x_1, \dots, x_k$, possibly along with some positive integer scalars;
- the sole output is $f(x_1, \dots, x_k)$.

The *semiring complexity* (or $\{+, \times\}$-*complexity*) of f is the smallest size of (i.e., the smallest number of gates in) such an arithmetic circuit. This notion is illustrated in Figure 1. For additional details, see [4, Section 2] and references therein.

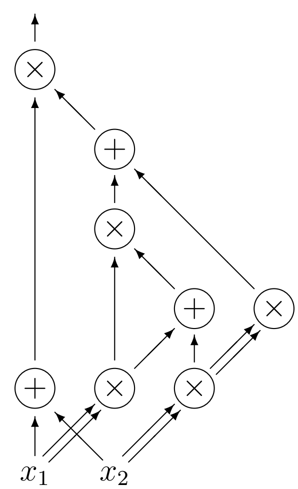

Figure 1. The smallest $\{+, \times\}$-circuit computing the polynomial $f(x_1, x_2) = h_5(x_1, x_2) = x_1^5 + x_1^4 x_2 + x_1^3 x_2^2 + x_1^2 x_2^3 + x_1 x_2^4 + x_2^5$. This circuit is based on the formula $h_5(x_1, x_2) = (x_1 + x_2)(x_1^2(x_1^2 + x_2^2) + x_2^4)$.

Date: August 15, 2016.

Key words and phrases. Semiring complexity, Schur function, Young tableaux.

2010 Mathematics Subject Classification Primary 68Q25, Secondary 05E05.

Partially supported by NSF grant DMS-1361789 (S. F.), RSF grant 16-11-10075 (D. G.), and the NSERC (É. S.). D. G. thanks MCCME Moscow and the Max-Planck Institut für Mathematik for their hospitality and inspiring atmosphere.

This paper is devoted to the problem of determining semiring complexity of symmetric polynomials. More specifically, we focus our attention on Schur functions, an important class of symmetric polynomials which play prominent roles in several branches of mathematics; see, e.g., [8, Chapter I] and [13, Chapter 7].

Let $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq 0)$ be an integer partition. The *Schur function* (or *Schur polynomial*)) $s_\lambda(x_1, \ldots, x_k)$ is a symmetric polynomial of degree $|\lambda| = \sum_i \lambda_i$ in the variables $x_1, \ldots, x_k$ which can be defined in many different ways. One remarkable feature of Schur polynomials that makes them an exciting object of study in algebraic complexity theory is that the classical formulas defining them fall into two categories. On the one hand, there are determinantal expressions (e.g., the Jacobi-Trudi formula or the bialternant formula) which provide efficient ways to compute Schur functions in an unrestricted setting, i.e., when all arithmetic operations are allowed. On the other hand, Schur functions are generating functions for semistandard Young tableaux. This description represents them as polynomials with manifestly positive coefficients; so they can be computed using addition and multiplication only. We note however that the naïve approach based on these monomial expansions yields algorithms whose (semiring) complexity is very high—and indeed very far from the optimum.

Our main result is the following. (We use the notation $\lambda' = (\lambda'_1 \geq \lambda'_2 \geq \cdots)$ for the partition conjugate to λ.)

Theorem 1.1. *The semiring complexity of a Schur polynomial* $s_\lambda(x_1, \ldots, x_k)$ *labeled by partition* $\lambda = (\lambda_1 \geq \cdots \geq \lambda_\ell)$ *is at most* $O(\log(\lambda_1) k^5 2^{k\ell} \ell^d)$ *where* $d = \max_j \lambda'_j (k - \lambda'_j)$.

Since $\ell \leq k$ (or else $s_\lambda(x_1, \ldots, x_k) = 0$) and $d \leq k^2/4$, we obtain:

Corollary 1.2. *The semiring complexity of* $s_\lambda(x_1, \ldots, x_k)$ *is bounded from above by* $k^{k^2(\frac{1}{4}+o(1))} O(\log(\lambda_1))$. *If the number of variables* k *is fixed, then this complexity is* $O(\log(\lambda_1))$.

Remark 1.3. The problem of designing efficient algorithms employing addition and multiplication arises naturally in the context of numerical computation, as these algorithms possess valuable stability properties. Motivated by such considerations, J. Demmel and P. Koev [3] designed $\{+, \times\}$-algorithms for computing Schur polynomials using a dynamic programming approach. In the notation of Theorem 1.1, their [3, Proposition 5.3] asserts that the semiring complexity of $s_\lambda(x_1, \ldots, x_k)$ is bounded from above by $O(e^{5.2\sqrt{|\lambda|}} \ell k)$. When k is fixed, and the shape λ grows, this bound is much larger than the one is Corollary 1.2. On the other hand, in the regime where λ is fixed, and the number of variables k grows, the complexity of the Demmel-Koev algorithm is linear in k whereas the bound in Theorem 1.1 is exponential in k. It would be interesting to find a common generalization of these results.

We prove Theorem 1.1 in two stages. At the first stage (see Section 3), we treat a special case where partition λ has only one (nonzero) part. More explicitly, we obtain the following result.

Let $h_n(x_1, \ldots, x_k)$ denote the *complete homogeneous symmetric polynomial*, i.e., the sum of all monomials of degree n in the variables $x_1, \ldots, x_k$:

$$h_n(x_1, \ldots, x_k) = \sum_{1 \le i_1 \le \cdots \le i_n \le k} x_{i_1} \cdots x_{i_n} \,. \tag{1.1}$$

(See an example in Figure 1.)

Theorem 1.4. *The semiring complexity of $h_n(x_1, \ldots, x_k)$ is $O(k^2 \log(n))$.*

Our proof of Theorem 1.1, presented in Section 6, relies on three main ingredients:

- Theorem 1.4;
- a formula expressing a multichain-generating function of a *shellable poset* in terms of complete homogeneous polynomials, see Section 4; and
- a formula representing a Schur polynomial as a multichain generating function (or more precisely an iterated sum thereof), see Section 5.

2. Related problems

The general problem of determining the semiring complexity of a Schur polynomial is open. In particular, the following tantalizing problem remains out of reach.

Problem 2.1 ([4, Problem 3.2])**.** Can the semiring complexity of $s_\lambda(x_1, \ldots, x_k)$ be bounded by a polynomial in k and $n = |\lambda|$?

Remark 2.2. A general method for obtaining lower bounds on semiring complexity was suggested in 1976 by C. Schnorr [11]. Schnorr's bound only depends on the *support* of a polynomial, i.e., on the set of monomials that contribute with a positive coefficient. Schnorr's argument was further refined by E. Shamir and M. Snir [12]; powerful applications were given in [6]. As mentioned in [4, Remark 3.3], Schnorr-type lower bounds are useless in the case of Schur functions since computing a Schur function is difficult not because of its support but because of the complexity of its coefficients (the Kostka numbers). The problem of computing an individual Kostka number is known to be #**P**-complete (H. Narayanan [9]) whereas the support of a Schur function is very easy to determine.

Remark 2.3. The paper [4] by the first two authors (with G. Koshevoy) investigated the concept of semiring complexity alongside other similar computational models involving restricted sets of arithmetic operations. In brief, the results obtained in [4, 6, 16] demonstrate that that adjoining subtraction and/or division to the two-element set $\{+, \times\}$ of allowed arithmetic operations can, in some cases, dramatically decrease computational complexity. (By contrast, removing division from $\{+, -, \times, \div\}$ comes at merely polynomial cost, as shown by V. Strassen's [14].) We refer the reader to [4] for the discussion of these issues.

Remark 2.4. In the unrestricted model, one can compute $s_\lambda(x_1, \ldots, x_k)$ in time polynomial in k and $\log(\lambda_1)$, via the bialternant formula, see, e.g., [13, Section 7.15], and using repeated squaring to compute the powers of variables appearing in the relevant determinants.

One important complexity model studied in [4] is *subtraction-free complexity*, which allows the operations of addition, multiplication, and division. It turns out that subtraction-free complexity of a Schur function is indeed polynomial.

Theorem 2.5 ([7, Section 6], [2, Section 4], [4, Theorem 3.1])**.** *Subtraction-free complexity of a Schur polynomial $s_\lambda(x_1,\dots,x_k)$ is at most $O(n^3)$ where $n=k+\lambda_1$.*

We emphasize that the algorithms in [2, 4, 7] utilize division in essential ways, so they do not bring us any closer to the resolution of Problem 2.1.

Since subtraction-free complexity is bounded from above by semiring complexity, Theorem 1.1 implies that the subtraction-free complexity of a particular Schur polynomial $s_\lambda(x_1,\dots,x_k)$ can be much smaller (for small k) than the upper bound of Theorem 2.5.

Problem 2.6. Find a natural upper bound on subtraction-free complexity of a Schur polynomial that simultaneously strengthens Theorems 1.1 and 2.5.

Remark 2.7. The paper [5] by G. Koshevoy and the second author provided an exponential lower bound on $\{+,\times\}$-complexity of a monomial symmetric function.

3. Semiring complexity of complete homogeneous polynomials

In this section, we prove Theorem 1.4. We use the notation

$$h_m = h_m(x_1,\dots,x_k) = \sum_{1\le i_1\le\cdots\le i_m\le k} x_{i_1}\cdots x_{i_m}\,, \tag{3.1}$$

$$\tilde h_m = h_m(x_1^2,\dots,x_k^2), \tag{3.2}$$

$$e_m = e_m(x_1,\dots,x_k) = \sum_{1\le i_1<\cdots<i_m\le k} x_{i_1}\cdots x_{i_m}\,. \tag{3.3}$$

Lemma 3.1. *One can compute $h_{n-k+1},\dots,h_n$ starting from $\tilde h_{\lfloor\frac{n}{2}\rfloor-k+1},\dots,\tilde h_{\lfloor\frac{n}{2}\rfloor}$ and $e_1,\dots,e_k$, using $O(k^2)$ additions and multiplications.*

Proof. The key algebraic observation is that

$$\sum_{m\ge0} h_m\,t^m = \prod_{i=1}^k (1-x_it)^{-1} = \prod_{i=1}^k (1+x_it)\prod_{i=1}^k(1-x_i^2t^2)^{-1} = \sum_{a=0}^k e_at^a\sum_{b\ge0}\tilde h_bt^{2b},$$

and consequently

$$h_m = \sum_{m-k\le 2b\le m} e_{m-2b}\,\tilde h_b\,. \tag{3.4}$$

For $n-k+1\le m\le n$, the indices b appearing on the right-hand side of (3.4) satisfy $b\le\lfloor\frac{m}{2}\rfloor\le\lfloor\frac{n}{2}\rfloor$ and $b\ge\lceil\frac{m-k}{2}\rceil\ge\lceil\frac{n-2k+1}{2}\rceil=\lfloor\frac{n}{2}\rfloor-k+1$. Thus we can use (3.4) to compute these h_m; this takes $O(k)$ operations for each of the k values of m. □

Lemma 3.2. *One can compute $e_1,\dots,e_k$ using $O(k^2)$ additions and multiplications.*

Proof. The required algorithm is obtained by iterating the Pascal-type recurrence

$$e_m(x_1,\dots,x_j) = x_je_{m-1}(x_1,\dots,x_{j-1}) + e_m(x_1,\dots,x_{j-1}).$$ □

We note that in the unrestricted model, the complexity of computing $e_1, \ldots, e_k$ is of the order $k \log(k)$, see [15].

Proof of Theorem 1.4. Let us denote by $T(n)$ the semiring complexity of computing $h_{n-k+1}, \ldots, h_n$. Lemmas 3.1–3.2 imply that $T(n) \le T(\lfloor \frac{n}{2} \rfloor) + O(k^2)$. (Squaring the variables $x_1, \ldots, x_k$, which is needed to compute the $\tilde{h}_b$'s, takes linear time.) We conclude that $T(n) = O(k^2 \log(n))$, as desired. □

4. Partially ordered sets: generating functions and shellings

Definition 4.1 (*Poset, chain complex*)**.** Let P be a finite graded partially ordered set (*poset*) with a unique minimal element $\hat{0}$ and a unique maximal element $\hat{1}$. A linearly ordered subset

$$\{p_1 < \cdots < p_m\} \subset P$$

is called a *chain* (of size m). The *chain complex* of P is the simplicial complex on the ground set P whose simplices are the chains in P. Since P is graded, all maximal (by inclusion) chains in P have the same cardinality, so the chain complex is *pure*.

Definition 4.2 (*Shelling*)**.** A linear ordering

$$C_1, \ldots, C_f \tag{4.1}$$

of the set of all maximal chains in P is called a *shelling* (of the chain complex) if for every $j \in \{1, \ldots, f\}$, the subcomplex $\bigcup_{i=1}^{j-1} C_i$ of the chain complex (i.e., the simplical complex whose maximal simplices are $C_1, \ldots, C_{j-1}$) intersects C_j at a union of codimension 1 faces of C_j. We denote by $C_j^* \subset C_j$ the complement (inside C_j) of the intersection of these codimension 1 faces. Equivalently, C_j^* is the (unique) smallest face of C_j not contained in $\bigcup_{i<j} C_i$. Put differently, C_j^* consists of the elements of C_j which can be exchanged with another element to form one of the maximal chains preceding C_j in the shelling:

$$C_j^* = \{x \in C_j \mid \exists y \in P \ \exists i < j \ \ C_j - \{x\} \cup \{y\} = C_i\}. \tag{4.2}$$

Definition 4.3 (*Multichain, support*)**.** A "weakly increasing" sequence $M = (p_1 \le \cdots \le p_m)$ consisting of elements $p_i \in P$ is called a *multichain* of size m; we write $|M| = m$. The elements of P which appear in M (with nonzero multiplicity) form the *support* of M, denoted by $\mathrm{supp}(M)$. The support of any multichain is a chain.

Let us associate a formal variable z_p with each element $p \in P$. For a multiset M of elements in P, we denote by $\mathbf{z}^M$ the corresponding monomial: $\mathbf{z}^M = \prod_{p \in M} z_p$.

Lemma 4.4. *If P has a shelling as above, then the generating function for its multichains of size m is given by*

$$\sum_{\substack{\text{multichain } M \\ |M|=m}} \mathbf{z}^M = \sum_{j=1}^{f} \sum_{\substack{C_j^* \subset \mathrm{supp}(M) \subset C_j \\ |M|=m}} \mathbf{z}^M = \sum_{j=1}^{f} \mathbf{z}^{C_j^*} h_{m-|C_j^*|}((z_p)_{p \in C_j}), \tag{4.3}$$

where we use the notation (4.1)–(4.2)*.*

Proof. For any chain C, there is a unique maximal chain C_j with $C_j^* \subset C \subset C_j$. (The chains C satisfying the latter condition are precisely the ones contained in C_j but in none of the preceding chains C_i $(i < j)$.) Categorizing the multichains M by their support, and applying the above observation to $C = \mathrm{supp}(M)$, we establish (4.3). □

In Section 5, we will relate Schur polynomials to a special case of the above construction involving a class of (shellable) posets $\mathbf{P}_{h,k}$ described in Definition 4.5 below. These posets have been extensively studied in algebraic combinatorics, due to the role they play in representation theory and the classical Schubert Calculus. In particular, $\mathbf{P}_{h,k}$ describes the attachment of Schubert cells in the Grassmann manifold $\mathrm{Gr}(h, k)$.

Definition 4.5 (*Posets* $\mathbf{P}_{h,k}$)**.** Let h and k be positive integers, with $h \le k$. We denote by $\mathbf{P}_{h,k}$ the poset whose elements are column vectors (or simply *columns*) of *height* h whose entries lie in the set $\{1, \ldots, k\}$ and strictly increase downwards:

$$t = \begin{bmatrix} t_1 \\ \vdots \\ t_h \end{bmatrix} \in \mathbb{Z}^h, \quad 1 \le t_1 < \cdots < t_h \le k; \tag{4.4}$$

these are partially ordered component-wise: $\begin{bmatrix} t_1 \\ \vdots \\ t_h \end{bmatrix} \le \begin{bmatrix} t'_1 \\ \vdots \\ t'_h \end{bmatrix}$ if and only if $\begin{cases} t_1 \le t'_1 \\ \vdots \\ t_h \le t'_h \end{cases}$.

Let us make a few simple but useful observations.

Lemma 4.6.

(1) *The cardinality of* $\mathbf{P}_{h,k}$ *is* $\binom{k}{h}$.

(2) *The columns* $\hat{0} = \begin{bmatrix} 1 \\ \vdots \\ h \end{bmatrix}$ *and* $\hat{1} = \begin{bmatrix} k-h+1 \\ \vdots \\ k \end{bmatrix}$ *are the unique minimal and maximal elements of* $\mathbf{P}_{h,k}$*, respectively.*

(3) *The poset* $\mathbf{P}_{h,k}$ *is graded, with the rank function given by*

$$\mathrm{rk}\Big(\begin{bmatrix} t_1 \\ \vdots \\ t_h \end{bmatrix}\Big) = t_1 + \cdots + t_h - \frac{h(h+1)}{2}.$$

(4) *Each maximal chain in* $\mathbf{P}_{h,k}$ *has cardinality* $h(k-h)+1$.

(5) *The poset* $\mathbf{P}_{h,k}$ *is canonically isomorphic to the poset of integer partitions (partially ordered component-wise) having at most* h *parts all of which are* $\le k-h$*. The isomorphism is defined by*

$$\begin{bmatrix} t_1 \\ \vdots \\ t_h \end{bmatrix} \mapsto (t_h - h, \ldots, t_1 - 1).$$

Equivalently, $\mathbf{P}_{h,k}$ is canonically isomorphic to the poset of Young diagrams fitting inside the $h \times (k-h)$ box, ordered by inclusion.

Example 4.7. Let $h = 2$ and $k = 5$. The poset $\mathbf{P}_{2,5}$ consists of $\binom{5}{2} = 10$ elements of the form $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$, with $1 \le a < b \le 5$. These are in bijection with partitions $\mu = (\mu_1, \mu_2) = (b-2, a-1)$ satisfying $3 \ge \mu_1 \ge \mu_2 \ge 0$ (equivalently, Young diagrams fitting inside the 2×3 rectangle). There are 5 maximal chains in $\mathbf{P}_{2,5}$, corresponding to the 5 standard tableaux (cf. Definition 5.1) of this rectangular shape:

maximal chains	standard tableaux
$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}4\\5\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}135\\246\end{smallmatrix}\right]$
$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}4\\5\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}134\\256\end{smallmatrix}\right]$
$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}4\\5\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}125\\346\end{smallmatrix}\right]$
$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}4\\5\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}124\\356\end{smallmatrix}\right]$
$\left[\begin{smallmatrix}1\\2\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\3\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\4\end{smallmatrix}\right] < \left[\begin{smallmatrix}1\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}2\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}3\\5\end{smallmatrix}\right] < \left[\begin{smallmatrix}4\\5\end{smallmatrix}\right]$	$\left[\begin{smallmatrix}123\\456\end{smallmatrix}\right]$

We will later need the following crude estimate.

Lemma 4.8. *The number of maximal chains in* $\mathbf{P}_{h,k}$ *is bounded by* $h^{h(k-h)}$.

Proof. The maximal chains in $\mathbf{P}_{h,k}$ correspond to the standard Young tableaux (see, e.g., Definition 5.1 below) of rectangular shape $h \times (k-h)$. Such a standard tableau can be constructed by consecutively placing the entries $1, \dots, h(k-h)$, in this order. Each time, we have at most h choices (at most one per row). □

Definition 4.9 (*Intervals* $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$*, and lexicographic ordering of maximal chains*)**.** For $\mathbf{a}, \mathbf{b} \in \mathbf{P}_{h,k}$ satisfying $\mathbf{a} \le \mathbf{b}$, we denote by $[\mathbf{a}, \mathbf{b}] = \mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$ the corresponding (order-theoretic) interval:

$$\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}] = \{\mathbf{c} \in \mathbf{P}_{h,k} \mid \mathbf{a} \le \mathbf{c} \le \mathbf{b}\}.$$

In the special case $\mathbf{P}_{h,k}[\hat{0}, \hat{1}] = \mathbf{P}_{h,k}$, we recover the entire poset $\mathbf{P}_{h,k}$.

The *lexicographic ordering* on the set of maximal chains in $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$ is the linear order defined as follows. Let

$$C = \left(\mathbf{a} = \begin{bmatrix} a_{11} \\ \vdots \\ a_{h1} \end{bmatrix} < \cdots < \begin{bmatrix} a_{1N} \\ \vdots \\ a_{hN} \end{bmatrix} = \mathbf{b}\right),$$

$$C' = \left(\mathbf{a} = \begin{bmatrix} a'_{11} \\ \vdots \\ a'_{h1} \end{bmatrix} < \cdots < \begin{bmatrix} a'_{1N} \\ \vdots \\ a'_{hN} \end{bmatrix} = \mathbf{b}\right)$$

be two maximal chains in $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$. Let j indicate the leftmost position where these two chains differ, that is, the smallest index for which there exists i with $a_{ij} \neq a'_{ij}$. Furthermore, let i be the largest value (corresponding to the lowermost location) for which this inequality occurs (for the minimal choice of j). Then $C < C'$ in the lexicographic ordering if and only if $a_{ij} < a'_{ij}$.

Example 4.10. In Example 4.7, the maximal chains in $\mathbf{P}_{2,5} = \mathbf{P}_{2,5}[\hat{0}, \hat{1}]$ are listed in the lexicographic order, top down.

The following result is a special case (for Lie type A) of a well known result about shellability of intervals in Bruhat order quotients for finite Coxeter groups, see [1, 10].

Lemma 4.11. *The lexicographic ordering of the maximal chains in* $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$ *provides a shelling of this poset.*

Combining Lemmas 4.4 and 4.11 enables us to express a generating function for multichains in $\mathbf{P}_{h,k}[\mathbf{a},\mathbf{b}]$ in terms of complete homogeneous symmetric functions. These expressions, reformulated in terms of a certain kind of *tableaux*, will be used in Section 5 to obtain efficient $\{+,\times\}$-algorithms for computing Schur functions.

5. Schur polynomials as multichain generating functions

Let us recall the combinatorial definition of a Schur polynomial $s_\lambda(x_1,\dots,x_k)$ labeled by an integer partition $\lambda=(\lambda_1\geq\cdots\geq\lambda_\ell\geq 0)$. Note that we allow trailing zeroes at the end of λ.

We assume that $\ell\leq k$. This condition does not restrict the generality, since $\lambda_\ell>0$ and $\ell>k$ imply $s_\lambda(x_1,\dots,x_k)=0$.

We use the notation $n=|\lambda|=\lambda_1+\cdots+\lambda_\ell$ for the size of the partition λ.

Definition 5.1 (*Tableaux, Schur functions*)**.** A *semistandard (Young) tableau* T of shape $\lambda=|T|$ is an array of integers

$$T=(t_{i,j}\mid 1\leq i\leq\ell,\ 1\leq j\leq\lambda_i)$$

satisfying $t_{i,j}<t_{i+1,j}$ and $t_{i,j}\leq t_{i,j+1}$ whenever these inequalities make sense. A tableau T is called *standard* if each of the numbers $1,\dots,n$ appears exactly once among the n tableau entries $t_{i,j}$. We denote by $\mathbf{x}^T$ the monomial associated with T:

$$\mathbf{x}^T=\prod_{i,j}x_{t_{i,j}}.$$

The *Schur function* (or *Schur polynomial*) $s_\lambda(x_1,\dots,x_k)$ is the generating function for semistandard tableaux of shape λ and entries in $\{1,\dots,k\}$:

$$s_\lambda(x_1,\dots,x_k)=\sum_{|T|=\lambda}\mathbf{x}^T.$$

By construction, $s_\lambda(x_1,\dots,x_k)$ is a homogeneous polynomial of degree n in the variables $x_1,\dots,x_k$, with positive integer coefficients. It is well known (see, e.g., [13, Chapter 7]) that $s_\lambda(x_1,\dots,x_k)$ is symmetric with respect to permutations of the variables.

Example 5.2. Let $\ell=2$ and $\lambda=(r,r)$. A semistandard tableau of shape λ is a $2\times r$ matrix $T=(t_{i,j})$ with positive integer entries which weakly increase left-to-right in each row, and strictly increase top-down in each column. The corresponding Schur polynomial is given by $s_{(r,r)}(x_1,\dots,x_k)=\sum_T\prod_i\prod_j x_{t_{i,j}}$ where the sum is over all such tableaux with entries $\leq k$. For example, if $r=2$ and $k=3$, then we get 6 different tableaux, and the answer is $s_{(2,2)}(x_1,x_2,x_3)=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2+x_1^2x_2x_3+x_1x_2^2x_3+x_1x_2x_3^2$.

Our next goal is to restate Definition 5.1 using the language of multichain generating functions introduced in Section 4.

The connection between Schur functions and the posets $\mathbf{P}_{h,k}$ comes from the observation that the multichains of size m in $\mathbf{P}_{h,k}$ are in bijection with semistandard tableaux of rectangular shape $h\times m$ and entries $\leq k$. We next extend this correspondence to arbitrary shapes. This will require some preparation.

Definition 5.3 (*Dissecting Young diagrams into rectangular shapes*)**.** Let $\lambda = (\lambda_1 \geq \cdots \geq \lambda_\ell)$ be an integer partition. As usual, we denote by λ' the conjugate partition, i.e., the partition whose parts are the column lengths of (the shape of) λ. We then denote by $\tilde{\lambda}'_1 > \cdots > \tilde{\lambda}'_s$ the integers, listed in the decreasing order, which appear as parts of λ'. In other words, $\tilde{\lambda}'_1, \ldots, \tilde{\lambda}'_s$ are all the (different) heights of columns in the Young diagram of λ. We denote by $\tilde{\lambda} = (\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_\ell)$ the partition conjugate to $\tilde{\lambda}' = (\tilde{\lambda}'_1, \ldots, \tilde{\lambda}'_s)$. To rephrase, the shape $\tilde{\lambda}$ is obtained from λ by keeping one column of each height, and striking out the rest.

We can now dissect the Young diagram λ by vertical cuts into s rectangular shapes of sizes $h \times (\lambda_h - \lambda_{h+1})$ where h runs over the set of parts of $\tilde{\lambda}'$ (equivalently, the distinct column lengths of λ). To simplify notation for the sake of future arguments, we denote $h_j = \tilde{\lambda}'_j$ and $m_j = \lambda_{h_j} - \lambda_{h_j+1} - 1$, so that λ gets dissected into rectangles of sizes $h_j \times (m_j + 1)$, for $j = 1, \ldots, s$.

Example 5.4. Let $\lambda = (6, 6, 4, 1, 1)$, $\ell = 5$. Then $\lambda' = (5, 3, 3, 3, 2, 2)$, $\tilde{\lambda}' = (5, 3, 2)$, $\tilde{\lambda} = (3, 3, 2, 1, 1)$, $s = 3$. The shape λ can be dissected by vertical cuts into three rectangles of sizes 5×1, 3×3, and 2×2, respectively. In this example, we have $h_1 = 5, h_2 = 3, h_3 = 2, m_1 = 0, m_2 = 2, m_3 = 1$.

Definition 5.5 (*Pruning of tableaux*)**.** Let T be a semistandard tableau of shape λ. The *pruning* of T is the semistandard tableau $\tilde{T}$ of shape $\tilde{\lambda}$ obtained from T by selecting the rightmost column of each height (and removing all columns of that height located to the left of it). We denote by $\mathbf{a}_1, \ldots, \mathbf{a}_s$ the columns of $\tilde{T}$, listed left to right. (These columns have heights $h_1, \ldots, h_s$, respectively.) We denote by $\bar{\mathbf{a}}_j$ the column of height h_{j+1} obtained from $\mathbf{a}_j$ by removing the $h_j - h_{j+1}$ bottom entries.

We furthermore denote by $T_1, \ldots, T_s$ the semistandard tableaux of rectangular shapes $h_1 \times m_1, \ldots, h_s \times m_s$ obtained by dissecting T by the vertical cuts described in Definition 5.3, and then removing the rightmost column from each of the resulting tableaux. (If $m_j = 0$, then T_j is empty.) Thus T is obtained by interlacing the rectangular tableaux T_j with the columns of the pruning: $T = [T_1|\mathbf{a}_1|T_2|\mathbf{a}_2|\cdots|T_s|\mathbf{a}_s]$.

Example 5.6. Continuing with Example 5.4, let $T = \begin{bmatrix} 1&1&2&2&2&4 \\ 2&2&3&3&3&5 \\ 4&5&6&6 \\ 5 \\ 6 \end{bmatrix}$. Then

$$\tilde{T} = \begin{bmatrix} 1&2&4 \\ 2&3&5 \\ 4&6 \\ 5 \\ 6 \end{bmatrix}, \quad T_1 = \varnothing, \quad \mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 5 \\ 6 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1&2 \\ 2&3 \\ 5&6 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}, \quad T_3 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Consider the set of semistandard tableaux T of a given shape λ, with entries $\leq k$, and with a given pruning $\tilde{T} = [\mathbf{a}_1|\cdots|\mathbf{a}_s]$. Note that once $\tilde{T}$ and λ have been fixed, each tableau T_j, for $1 \leq j \leq s$, can be chosen independently of the others, as long as it satisfies the following restrictions:

- T_j is a semistandard tableau of rectangular shape $h_j \times m_j$, with entries $\leq k$; as such, it can be viewed as a multichain of size m_j in the poset $\mathbf{P}_{h_j,k}$;
- every column $\mathbf{a}$ in T_j (i.e., every element of this multichain) satisfies the inequalities $\bar{\mathbf{a}}_{j-1} \leq \mathbf{a} \leq \mathbf{a}_j$, with respect to the partial order in $\mathbf{P}_{h_j,k}$.

(We set $\bar{\mathbf{a}}_0 = \hat{0} = \begin{bmatrix} 1 \\ \vdots \\ \ell \end{bmatrix}$ by convention, so that the lower bound is redundant for $j = 1$.) This gives a bijection between the set of tableaux under consideration and a Cartesian product of sets of multichains in the posets $\mathbf{P}_{h_j,k}$:

$$\left\{ \begin{array}{c} \text{semistandard tableaux } T \\ \text{of shape } \lambda, \text{ with entries } \le k, \\ \text{with pruning } \tilde{T} = [\mathbf{a}_1|\cdots|\mathbf{a}_s] \end{array} \right\} \longleftrightarrow \prod_{j=1}^{s} \left\{ \begin{array}{c} \text{multichains of size } m_j \\ \text{in } \mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j] \end{array} \right\}$$

Identifying multichains in $\mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j]$ with semistandard tableaux of rectangular shape, and passing to generating functions, we obtain the following result.

Lemma 5.7. *With the notation as above, we have*

$$s_\lambda(x_1, \ldots, x_k) = \sum_{\tilde{T}} \mathbf{x}^{\tilde{T}} \prod_{j=1}^{s} \sum_{T_j} \mathbf{x}^{T_j}, \tag{5.1}$$

where

- $\tilde{T} = [\mathbf{a}_1|\cdots|\mathbf{a}_s]$ *runs over semistandard tableaux of shape* $\tilde{\lambda}$*, with entries* $\le k$*;*
- T_j *runs over semistandard tableaux of rectangular shape* $h_j \times m_j$ *whose columns form a multichain in* $\mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j]$.

Since the interval $\mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j]$ is (lexicographically) shellable, the sums $\sum_{T_j} \mathbf{x}^{T_j}$ appearing in (5.1) can be computed using formula (4.3):

Lemma 5.8. *We have*

$$\sum_{T} \mathbf{x}^{T} = \sum_{Q} \mathbf{x}^{Q^*} h_{m-|Q^*|}(\mathbf{x}^{\mathbf{c}_1}, \ldots, \mathbf{x}^{\mathbf{c}_N}), \tag{5.2}$$

where

- T *runs over semistandard tableaux of rectangular shape* $h \times m$ *whose columns form a multichain in* $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$*;*
- $Q = [\mathbf{c}_1|\cdots|\mathbf{c}_N]$ *runs over the maximal chains in* $\mathbf{P}_{h,k}[\mathbf{a}, \mathbf{b}]$*, and* Q^* *is a subchain of* Q *obtained from* Q *as in* (4.2)*.*

Let us restate the latter condition in concrete terms. For each pair of consecutive columns $\mathbf{c}_j$ and $\mathbf{c}_{j+1}$, we have $\mathbf{c}_{j+1} = \mathbf{c}_j + \mathbf{e}_{i_j}$ for some $i_j \in \{1, \ldots, h\}$, where $\mathbf{e}_i$ denotes the column whose ith component is equal to 1, and all others are equal to 0. The chain/tableau Q^* is formed by the subset of columns $\mathbf{c}_j$ for which $i_{j-1} > i_j$ and moreover $\mathbf{c}_{j-1} + \mathbf{e}_j \in \mathbf{P}_{h,k}$ (so that replacing $\mathbf{c}_j$ by $\mathbf{c}_{j-1} + \mathbf{e}_j$ transforms Q into a lexicographically smaller maximal chain).

Example 5.9. Let $h = 2$, $k = 5$, $\mathbf{a} = \hat{0}$, $\mathbf{b} = \hat{1}$, cf. Example 4.7. Then (5.2) becomes

$$\begin{aligned} s_{(m,m)}(x_1, \ldots, x_5) = {} & h_m(x_1x_2, x_1x_3, x_2x_3, x_2x_4, x_3x_4, x_3x_5, x_4x_5) \\ & + x_2x_5\, h_{m-1}(x_1x_2, x_1x_3, x_2x_3, x_2x_4, x_2x_5, x_3x_5, x_4x_5) \\ & + x_1x_4\, h_{m-1}(x_1x_2, x_1x_3, x_1x_4, x_2x_4, x_3x_4, x_3x_5, x_4x_5) \\ & + x_1x_4 \cdot x_2x_5\, h_{m-2}(x_1x_2, x_1x_3, x_1x_4, x_2x_4, x_2x_5, x_3x_5, x_4x_5) \\ & + x_1x_5\, h_{m-1}(x_1x_2, x_1x_3, x_1x_4, x_1x_5, x_2x_5, x_3x_5, x_4x_5) \end{aligned}$$

6. Proof of the main theorem

Combining (5.1) and (5.2), we obtain:

Corollary 6.1. *The Schur polynomial* $s_\lambda(x_1, \ldots, x_k)$ *is given by*

$$s_\lambda(x_1, \ldots, x_k) = \sum_{|\tilde{T}|=\tilde{\lambda}} \mathbf{x}^{\tilde{T}} \prod_{j=1}^{s} \sum_{Q} \mathbf{x}^{Q^*} h_{m_j - |Q^*|}(\mathbf{x}^{\mathbf{c}_1}, \ldots, \mathbf{x}^{\mathbf{c}_N}), \tag{6.1}$$

where

- $\tilde{\lambda}$, s, $h_1, \ldots, h_s$, *and* $m_1, \ldots, m_s$ *are described in Definition 5.3;*
- $\tilde{T} = [\mathbf{a}_1 | \cdots | \mathbf{a}_s]$ *runs over semistandard tableaux of shape* $\tilde{\lambda}$*, with entries* $\le k$*;*
- $Q = [\mathbf{c}_1 | \cdots | \mathbf{c}_N]$ *runs over the maximal chains in* $\mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j]$.

To prove Theorem 1.1, we analyze the (semiring) complexity of computing a Schur polynomial $s_\lambda(x_1, \ldots, x_k)$ using the formula (6.1) together with Theorem 1.4.

We begin by computing the monomials $\mathbf{x}^{\mathbf{c}}$, for all columns $\mathbf{c}$ of height h_j with entries $\le k$, for each $j \le s$. This can be done using $\le \ell \sum_{j \le s} \binom{k}{h_j}$ multiplications. (Note that $s \le \ell$.)

Recall that the Young diagram $\tilde{\lambda}$ has s columns, of heights $h_1, \ldots, h_s$. Hence the number of tableaux $\tilde{T}$ appearing in (6.1) is bounded by $\prod_{j \le s} \binom{k}{h_j}$.

Each monomial $\mathbf{x}^{\tilde{T}}$ can be computed by $s-1$ multiplications (given all the $\mathbf{x}^{\mathbf{c}_i}$).

The number of maximal chains in $\mathbf{P}_{h_j,k}[\bar{\mathbf{a}}_{j-1}, \mathbf{a}_j]$ is at most $h_j^{h_j(k-h_j)}$, by Lemma 4.8. Each of these chains has length $N \le h(k-h)+1$. Since $|Q^*| \le |Q| = N$, we can compute $\mathbf{x}^{Q^*}$ in time $\le h(k-h)$. Also, $m_j - |Q^*| \le \lambda_1$. Theorem 1.4 now implies that we can compute $\mathbf{x}^{Q^*} h_{m_j-|Q^*|}(\mathbf{x}^{\mathbf{c}_1}, \ldots, \mathbf{x}^{\mathbf{c}_N})$ in time $O(h^2(k-h)^2 \log(\lambda_1))$. Putting everything together, we obtain the following upper bound on the semiring complexity of $s_\lambda(x_1, \ldots, x_k)$:

$$\ell \sum_{j \le s} \binom{k}{h_j} + \prod_{j \le s} \binom{k}{h_j} \cdot (2s + \sum_{j \le s} h_j^{h_j(k-h_j)} (O(h_j^2 (k-h_j)^2 \log(\lambda_1)))).$$

This can be replaced by $O(\log(\lambda_1)) s \ell^2 k^2 2^{ks} \ell^d$ where

$$d = \max_j h_j(k-h_j) = \max_j \lambda'_j(k - \lambda'_j),$$

and then by $O(\log(\lambda_1)) k^5 2^{k\ell} \ell^d$. □

References

[1] A. Björner and M. Wachs, Bruhat order of Coxeter groups and shellability, *Adv. in Math.* **43** (1982), 87–100.

[2] C. Chan, V. Drensky, A. Edelman, R. Kan, and P. Koev, On computing Schur functions and series thereof, preprint, 2008.

[3] J. Demmel and P. Koev, Accurate and efficient evaluation of Schur and Jack functions, *Math. Comp.* **75** (2006), 223–239.

[4] S. Fomin, D. Grigoriev, and G. Koshevoy, Subtraction-free complexity, cluster transformations, and spanning trees, *Found. Comput. Math.* **16** (2016), 1–31.

[5] D. Grigoriev and G. Koshevoy, Complexity of tropical Schur polynomials, *J. Symbolic Comput.* **74** (2016), 46–54.
[6] M. Jerrum and M. Snir, Some exact complexity results for straight-line computations over semirings, *J. Assoc. Comput. Mach.* **29** (1982), 874–897.
[7] P. Koev, Accurate computations with totally nonnegative matrices, *SIAM J. Matrix Anal. Appl.* **29** (2007), 731–751.
[8] I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Mathematical Monographs, 1999.
[9] H. Narayanan, On the complexity of computing Kostka numbers and Littlewood-Richardson coefficients, *J. Algebraic Combin.* **24**, (2006), 347–354.
[10] R. A. Proctor, Classical Bruhat orders and lexicographic shellability, *J. Algebra* **77** (1982), 104–126.
[11] C. P. Schnorr, A lower bound on the number of additions in monotone computations, *Theor. Comput. Sci.* **2**, (1976), 305–315.
[12] E. Shamir and M. Snir, Lower bounds on the number of multiplications and the number of additions in monotone computations, *Technical Report RC-6757*, IBM, 1977.
[13] R. P. Stanley, *Enumerative combinatorics, vol. 2*, Cambridge University Press, 1999.
[14] V. Strassen, Vermeidung von Divisionen, *J. Reine Angew. Math.* **264** (1973), 184-202.
[15] V. Strassen, Die Berechnungskomplexität von elementarsymmetrischen Funktionen und von Interpolationskoeffizienten, *Numer. Math.* **20** (1972/73), 238–251.
[16] L. G. Valiant, Negation can be exponentially powerful, *Theor. Comput. Sci.* **12**, (1980), 303–314.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
E-mail address: fomin@umich.edu
URL: http://www.math.lsa.umich.edu/~fomin/

CNRS, Mathématiques, Université de Lille, Villeneuve d'Ascq, 59655, France
E-mail address: Dmitry.Grigoryev@math.univ-lille1.fr
URL: http://en.wikipedia.org/wiki/Dima_Grigoriev

LIX, École Polytechnique, 91128 Palaiseau Cedex, France
E-mail address: dorian.nogneng@lix.polytechnique.fr
URL: http://www.lix.polytechnique.fr/~dorian.nogneng/

Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada N2L 3G1
E-mail address: eschost@uwaterloo.ca
URL: https://cs.uwaterloo.ca/~eschost/