EXPLICIT POLYNOMIAL SEQUENCES WITH MAXIMAL SPACES OF PARTIAL DERIVATIVES AND A QUESTION OF K. MULMULEY

FULVIO GESMUNDO AND J.M. LANDSBERG

Abstract. We answer a question of K. Mulmuley: In [5] it was shown that the method of shifted partial derivatives cannot be used to separate the padded permanent from the determinant. Mulmuley asked if this “no-go” result could be extended to a model without padding. We prove this is indeed the case using the iterated matrix multiplication polynomial. We also provide several examples of polynomials with maximal space of partial derivatives, including the complete symmetric polynomials. We apply Koszul flattenings to these polynomials to have the first explicit sequence of polynomials with symmetric border rank lower bounds higher than the bounds attainable via partial derivatives.

1. Introduction

Let $S^d\mathbb{C}^N$ denote the space of homogeneous polynomials of degree d in N variables and let $p \in S^d\mathbb{C}^N$. Let $S^e\mathbb{C}^{N*}$ denote the space of homogenous differential operators of order e with constant coefficients, which acts on $S^d\mathbb{C}^N$ when $e \leq d$. The *e-th partial derivative map* of p (or *e-th flattening of* p) is

$$\begin{aligned} p_{e,d-e} : S^e\mathbb{C}^{N*} &\to Sd-e\mathbb{C}^N \\ D &\mapsto D(p). \end{aligned} \tag{1}$$

We call the image of $p_{e,d-e}$ the e-th *space of partial derivatives* of p; it is straightforward to verify that $\mathrm{rank}(p_{e,d-e}) = \mathrm{rank}(p_{d-e,e})$ and that given $e' \leq e \leq d/2$, if $p_{e,d-e}$ has full rank then $p_{e',d-e'}$ has full rank.

Let $M \geq N$. A polynomial $p \in S^d\mathbb{C}^N$ is a *degeneration* of $q \in S^d\mathbb{C}^M$, if $p \in \overline{GL_M \cdot q} \subset S^d\mathbb{C}^M$, where here we have chosen a linear inclusion $\mathbb{C}^N \subset \mathbb{C}^M$ and we consider p as a polynomial in M variables that just happens to only use N of them. Similarly, p is a *specialization* of q if $p \in \mathrm{End}_M \cdot q \subseteq S^d\mathbb{C}^M$. Notice that if p is a specialization of q then it is a degeneration of q. In complexity theory, one is interested in finding obstructions to specialization of a polynomial p to a polynomial q.

The *method of partial derivatives* and other flattening methods (see, e.g., [4]) exploit semicontinuity of matrix rank to prove that a polynomial is not a degeneration of another. Indeed, if p is a degeneration of q, then $\mathrm{rank}(p_{e,d-e}) \leq \mathrm{rank}(q_{e,d-e})$ for all e, therefore, comparing the ranks of the partial derivatives maps of p and q for various (or all) e, one can prove that p is not a degeneration of q (and thus nor is p a specialization of q).

Key words and phrases. computational complexity, Koszul flattening, shifted partial derivatives, Waring border rank, MSC 68Q15 .
Landsberg supported by NSF grant DMS-1405348.

The method of partial derivatives dates back to Sylvester in 1852 [22], who called the maps (1) *catalecticants*. These maps have been used to obtain lower bounds on the Waring rank, Waring border rank and cactus border rank of polynomials (see, e.g., [3, 12]). The *symmetric* or *Waring rank* of a polynomial $p \in S^d\mathbb{C}^N$ is the smallest r such that $p = \sum_{j=1}^r \ell_j^d$ where $\ell_j \in \mathbb{C}^N$ are linear forms. One writes $\mathbf{R}_S(p) = r$. The *symmetric* or *Waring border rank* of p is the smallest r such that p is a limit of polynomials of Waring rank r, and one writes $\underline{\mathbf{R}}_S(p) = r$. The ranks of the partial derivatives maps give lower bounds for the symmetric border rank of p: $\underline{\mathbf{R}}_S(p) \leq \min_e\{\mathrm{rank}(p_{e,d-e})\}$.

In [9, 11], it was shown that for a general polynomial p all the maps $p_{e,d-e}$ are of maximal rank. When the second author was preparing [15], he asked several experts if they knew of an explicit sequence of polynomials (e.g., in the complexity class **VNP**) with partial derivatives of maximal rank, as the standard references [12] in mathematics and [4] in computer science did not have one. Those asked did not furnish any example, so we wrote down several, see below. One example we found surprised us: the polynomial $(p_{n,2})^k = (x_1^2 + \cdots + x_n^2)^k$, because it is in the complexity class $\mathbf{VP}_e$ of sequences of polynomials admitting polynomial size formulas. It turns out this example had been discovered by Reznick in 1991 [20, Thm. 8.15], and in the same memoir he describes an explicit sequence that essentially dates back to Bierman [2] (the proof, if not the statement appeared in 1903), see below.

Let $p_{n,d} = x_1^d + \cdots + x_n^d$ denote the power sum polynomial of degree d in n variables and $h_{n,d} = \sum_{|\alpha|=d} x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ the complete symmetric polynomial of degree d in n variables.

For the following polynomial sequences, all partial derivatives map have full rank:

- $P_{Bier,n,d} := \sum_{|\alpha|=d}(\alpha_1 x_1 + \cdots + \alpha_n x_n)^d$ where α ranges over all multi-indices $(\alpha_1, ..., \alpha_n)$ of non-negative integers such that $\alpha_1 + \cdots + \alpha_n = d$ i.e., exponents of monomials of degree d in n variables (Bierman [2],Reznick [20]);
- $f_{n,k} := (p_{n,2})^k \in S^{2k}\mathbb{C}^n$, (Reznick [20], a proof is given in §2);
- $\widetilde{f}_{n,k} := p_{n,1} f_{n,k} \in S^{2k+1}\mathbb{C}^n$ (Theorem 7);
- $h_{n,d} \in S^d\mathbb{C}^n$ (Theorem 14).

Note that $h_{n,d} \in \mathbf{VP}_s$, the complexity class determined by the determinant, because the complete symmetric functions can be expressed as a determinant of a matrix whose entries are power sum functions. While it is obvious to the experts, since it is important, we present a proof of the following:

Proposition 1. *Let $n(m), k(m)$ be polynomially bounded functions of m. Then the sequences $\{f_{n,k}\}_m$ and $\{\widetilde{f}_{n,k}\}_m$ are in the algebraic complexity class $\mathbf{VP}_e$ of sequences admitting polynomial size formulas.*

A $\Sigma\Lambda\Sigma$ circuit is a depth three arithmetic circuit whose first and third layers consist of addition gates and whose middle layer consists of powering gates, sending $z \mapsto z^d$ for some d. Our results show this model is quite weak.

The *method of shifted partial derivatives* is a variant of the method of partial derivatives. It was introduced in [13] and exploited in [10] to prove super-polynomial lower complexity bounds for depth four circuits for the permanent (and determinant). In the same paper the authors ask if the method could be used to separate **VP** from **VNP**.

For $p \in S^d\mathbb{C}^N$ the method of shifted partials is based on the study of the following maps (for judiciously chosen e and τ):

$$\begin{aligned} p_{(e,d-e)[\tau]} : S^e\mathbb{C}^{N^*} \otimes S^\tau\mathbb{C}^N &\to S^{d-e+\tau}\mathbb{C}^N \\ D\otimes q &\mapsto qD(p). \end{aligned}$$

Notice that if $\tau = 0$, then $p_{(e,d-e)[\tau]}$ is the partial derivative map defined in (1). Let $\langle \partial^{=e} p\rangle_{=\tau} := p_{(e,d-e)[\tau]}(S^e\mathbb{C}^{N^*} \otimes S^\tau\mathbb{C}^N)$.

Again, semicontinuity of matrix rank guarantees that if p is a degeneration of q, then $\dim\langle \partial^{=e} q\rangle_{=\tau} \geq \dim\langle \partial^{=e} p\rangle_{=\tau}$ and the method of shifted partials can be used to prove that p is not a degeneration of q by showing that $\dim\langle \partial^{=e} p\rangle_{=\tau} > \dim\langle \partial^{=e} q\rangle_{=\tau}$.

There is a geometric interpretation of the image of the shifted partial derivative map: given $p \in S^d\mathbb{C}^n$, let $V(p) \subseteq \mathbb{P}(\mathbb{C}^n)^*$ be the hypersurface of degree d cut out by p. The image of $p_{e,d-e}$ generates an ideal in $S^d\mathbb{C}^n$ that we denote by $\mathcal{J}_e(p)$; it cuts out a subvariety of $V(p)$ that is called the e-th *Jacobian locus of* p. The image of the shifted partials map $p_{(e,d-e)[\tau]}$ is the component of degree $d+\tau$ of $\mathcal{J}_e(p)$. In particular, the study of the ranks of the shifted partials maps of p is equivalent to the study of the growth of the ideals $\mathcal{J}_e(p)$.

Definition 2. Given n, d, the (n,d)-*iterated matrix multiplication* polynomial $\mathrm{IMM}^d_n \in S^d(\mathbb{C}^{dn^2})$ is

$$\mathrm{IMM}^d_n : (X_1, ..., X_d) \mapsto \mathrm{trace}(X_1 \cdots X_d).$$

where $X_\alpha = ((\xi_\alpha)^i_j)_{i,j=1,\dots,n}$ are $n \times n$ matrices of indeterminates.

The (n,d)-*matrix powering* polynomial $\mathrm{Pow}^d_n \in S^d\mathbb{C}^{n^2}$ is

$$\mathrm{Pow}^d_n : X \mapsto \mathrm{trace}(X^d)$$

where $X = (\xi^i_j)_{i,j=1,\dots,n}$ is an $n \times n$ matrix of indeterminates.

By [18], both IMM^d_n and Pow^d_n are $\mathbf{VP}_s$-complete, the same complexity class for which the determinant polynomial $\det_n$ is complete. By the homogenization result of [18], the $\mathbf{VP}_s \neq \mathbf{VNP}$ conjecture can be rephrased by stating that there is no polynomially bounded function $n(m)$ such that the permanent polynomial perm_m is a specialization of $\mathrm{IMM}^m_{n(m)}$ (or of $\mathrm{Pow}^m_{n(m)}$). Note that in the cases of Pow^m_n and IMM^m_n, we are comparing directly with the permanent polynomial rather than its padded version as in the classical setting of the determinant versus permanent conjecture. Note further that Pow^m_n is a specialization of IMM^m_n.

We prove that the method of shifted partials cannot be used to achieve a superpolynomial separation between perm_m and IMM^m_n:

Theorem 3. *If $n > m^6$, then* perm_m *cannot be separated from* IMM^m_n *by the method of shifted partial derivatives. More precisely, given any linear inclusion $\mathbb{C}^{m^2} \subset \mathbb{C}^{mn^2}$ considering* $\mathrm{perm}_m \in S^m\mathbb{C}^{mn^2}$ *as a polynomial that just involves m^2 of the mn^2 variables, then for all choices of e,τ,* $\dim\langle \partial^{=e}(\mathrm{perm}_m \in S^m\mathbb{C}^{mn^2})\rangle_{=\tau} \leq \dim\langle \partial^{=e}\mathrm{IMM}^m_n\rangle_{=\tau}$.

Additional results. We give *a priori* upper bounds for the utility of Koszul flattenings (Proposition 16), another variant of the partial derivatives map. We show that these bounds are sharp for the first Koszul flattenings in low dimensions and degree

(Proposition 21). We obtain explicit (but not sharp) lower bounds for the Koszul flattenings of $\widetilde{f}_{k,n}$, showing one obtains better Waring border rank lower bounds with this method than by the method of partial derivatives (Proposition 19). Ironically, now the simple polynomial $\widetilde{f}_{k,n}$ has the highest Waring border rank lower bound of all explicit polynomials of odd degree.

Related work. Let $e_{n,d} = \sum_{1\le i_1<i_2<\cdots<i_d\le n} x_{i_1}\cdots x_{i_d} \in S^d\mathbb{C}^n$ denote the elementary symmetric polynomial. The complexity of $e_{n,d}$ has been well studied: its symmetric border rank is bounded below by $\binom{n}{\lfloor \frac{d}{2}\rfloor}$ because its space of partial derivatives of order e is spanned by the square free monomials of degree $d-e$. It symmetric rank is bounded above by $\sum_{i=0}^{\lfloor \frac{d}{2}\rfloor}\binom{n}{i}$ when d is odd and there is a similar formula for even d [16]. Its padded version can be computed by a homogeneous depth three ($\Sigma\Pi\Sigma$) circuit of size n^2 [1], and when $\log(n)\le d\le \frac{2n}{3}$, one has the lower bound of $max(\Omega(\frac{n^2}{d}),\Omega(nd))$ from [21] for its depth three circuit size. The lower bounds appear to translate to complete symmetric functions, however the upper bound relies on the generating function for the elementary symmetric functions being a product of linear forms, whereas the complete symmetric functions have generating function $\Pi_{i=1}^n(1-x_it)^{-1}$. The gap between the padded and unpadded depth three circuit complexity may have led researchers to think the results of [5] might fail in a model without padding, motivating Mulmuley's question (although Mulmuley himself anticipated our affirmative answer).

The shifted partial derivative complexity of elementary symmetric polynomials is studied in [6], where strong lower bounds are proved, which in turn give complexity lower bounds for depth four circuits.

Acknowledgements. We thank C. Yan for discussions on the Gessel-Viennot method, B. Reznick and Z. Teitler for historical information, and N. Stein and Mathoverflow for the proof of Lemma 11.

2. Proofs that $f_{k,n}$, $\widetilde{f}_{k,n}$ have maximal partial derivatives

Introduce the notation $q_n := p_{n,2} = x_1^2+\cdots+x_n^2$, $\ell_n := p_{n,1} = x_1+\cdots+x_n$.

We first treat the case of $f_{k,n}$:

Proposition 4. *For every $d\ge 0$, we have $(\mathbb{C}^n)^*\cdot\left(q_nS^d\mathbb{C}^n\right) = S^{d+1}\mathbb{C}^n$.*

Proof. If $d=0$, the statement holds as

$$\frac{\partial}{\partial x_i}q_n = 2x_i.$$

Let $d\ge 1$. Consider a monomial $\mathbf{x}^\alpha$ of degree $d-1$, where α is a multi-index, $|\alpha| = d-1$. For $i\in\{1,...,n\}$, define

$$g_i := \frac{\partial}{\partial x_i}(q_nx_i\mathbf{x}^\alpha) = 2x_i^2\mathbf{x}^\alpha + (\alpha_i+1)q_n\mathbf{x}^\alpha = \mathbf{x}^\alpha(2x_i^2+(\alpha+1)q_n).$$

For every i, $g_i\in(\mathbb{C}^n)^*\cdot\left(q_nS^d\mathbb{C}^n\right)$. Let $\mathbf{g} = (g_1,\dots,g_n)^T$ and $\mathbf{p} = (x_1^2,\dots,x_n^2)^T$ be column vectors. We have

$$\mathbf{g} = \mathbf{x}^\alpha(2\mathrm{Id}+A)\mathbf{p}$$

where A is the $(n+1)\times(n+1)$ rank 1 matrix whose entries in the i-th row are α_i+1 for $i=1,\dots,n$. Notice $A=\mathbf{1}\mathbf{a}^T$, where $\mathbf{1}=(1,\dots,1)^T$ and $\mathbf{a}=(\alpha_1+1,\dots,\alpha_n+1)^T$. In particular by the Sylvester determinant identity (see, e.g., [19, Chap. 1, Prob. 3.1]) $\det(2\mathrm{Id}+A)\neq 0$, so $2\mathrm{Id}+A$ is invertible.

Therefore, $x_i^2\mathbf{x}^\alpha \in (\mathbb{C}^n)^*\cdot\left(q_nS^d\mathbb{C}^n\right)$ for every monomial $\mathbf{x}^\alpha\in S^{d-1}\mathbb{C}^n$. This shows that every non-square free monomial belongs to $(\mathbb{C}^n)^*\cdot(q_nS^d\mathbb{C}^n)$.

Now let $\beta\in S^{d+1}\mathbb{C}^n$ be a square-free monomial; suppose $\beta_1=1$ and let γ be the multi-index with $\gamma_1=0$ and $\gamma_j=\beta_j$ for $j\geq 2$. Then

$$2\mathbf{x}^\beta=\frac{\partial}{\partial x_1}(q_n\mathbf{x}^\gamma).$$

This concludes the proof. □

Remark 5. The same argument as Proposition 4 allows us to prove a more general result. Use q_n to identify $\mathbb{C}^n$ with $(\mathbb{C}^n)^*$ and let W be a subspace of $\mathbb{C}^n$ such that $q_n|_W$ is nondegenerate. Then, via the identification $q_n:\mathbb{C}^n\to\mathbb{C}^{n*}$ induced by q_n, we may consider W^* as a subspace of $(\mathbb{C}^n)^*$. Then

$$W^*\cdot(q_nS^dW)=S^{d+1}W.$$

The argument is the same as above, considering that $q_n=q_n|_W\oplus q_n|_{W^\perp}$, where $W^\perp$ is the orthogonal complement of W in $\mathbb{C}^n$ with respect to q_n.

Theorem 6 (Reznick, [20])**.** *For every n,k and $e\leq k$, we have $(f_{n,k})_{e,2k-e}(S^e(\mathbb{C}^n)^*)=q_n^{k-e}S^e\mathbb{C}^n$. In particular, for all n,k,e, the flattening $(f_{n,k})_{e,2k-e}$ has full rank.*

Proof. We proceed by induction on e. For $e=0$ there is nothing to prove. Let $e\geq 1$. By the induction hypothesis, the $(e-1)$-st flattening surjects onto $q_n^{k-e+1}S^{e-1}\mathbb{C}^n$. It suffices to show that

$$(\mathbb{C}^n)^*\cdot(q_n^{k-e+1}S^{e-1}\mathbb{C}^n)=q_n^{k-e}S^e\mathbb{C}^n.$$

Notice that, for a monomial $\mathbf{x}^\alpha\in S^{e-1}\mathbb{C}^n$,

$$\begin{aligned}\frac{\partial}{\partial x_i}(q_n^{k-e+1}\mathbf{x}^\alpha)&=2(k-e+1)x_iq_n^{k-e}\mathbf{x}^\alpha+q_n^{k-e+1}\frac{\partial\mathbf{x}^\alpha}{\partial x_i}\\&=q_n^{k-e}\left(2(k-e+1)x_i\mathbf{x}^\alpha+q_n\frac{\partial\mathbf{x}^\alpha}{\partial x_i}\right).\end{aligned}$$

Up to rescaling q_n and the differential operators, the term in parenthesis is $\frac{\partial}{\partial x_i}(\mathbf{x}^\alpha q_n)$. These terms span $S^e\mathbb{C}^n$ by Proposition 4. □

Theorem 7. *For every n,k,e, the flattening $(\widetilde{f}_{n,k})_{e,2k+1-e}$ has full rank.*

Proof. Write $\mathbb{C}^n=\langle\ell_n\rangle\oplus W$ as representation of the symmetric group $\mathfrak{S}_n$, where $\langle\ell_n\rangle$ spans a $\mathfrak{S}_n$-invariant and W is isomorphic to the Specht modules $[n-1,1]$. S^2W contains a unique $\mathfrak{S}_n$-invariant up to scale, that we denote by g_n. Notice that $q_n=\frac{1}{n}\ell_n+g_n$ (after possibly rescaling g_n).

Let $e\leq k$. We show that the image of $(\widetilde{f}_{n,k})_{e,2k+1-e}$ has dimension $\binom{e+n-1}{e}$.

The decomposition $\mathbb{C}^n=W\oplus\langle\ell_n\rangle$ is orthogonal with respect to the $\mathfrak{S}_n$-invariant non-degenerate inner-product q_n, and q_n is non-degenerate restricted to each space. Write $\ell_n^*=\frac{1}{n}\left(\frac{\partial}{\partial x_1}+\cdots+\frac{\partial}{\partial x_n}\right)$, so that $(\mathbb{C}^n)^*=\langle\ell_n^*\rangle\oplus W^*$ with $W^*\cdot\ell_n=0$ and $\ell_n^*\cdot W=0$.

As an $\mathfrak{S}_n$-module, $S^e(\langle \ell_n^* \rangle \oplus W^*) = \bigoplus_{j=0}^e \langle (\ell_n^*)^j \rangle \otimes S^{e-j} W^*$, so

$$
\begin{aligned}
S^e(\langle \ell_n^* \rangle \oplus W^*) \cdot (\ell_n q_n^k) &= \left(\bigoplus_{j=0}^{e} \langle (\ell_n^*)^j \rangle \otimes S^{e-j} W^* \right) \cdot (\ell_n q_n^k) \\
&= \sum_{j=0}^{e} \langle (\ell_n^*)^j \rangle (S^{e-j} W^* \cdot (\ell_n q_n^k)) \\
&= \sum_{j=0}^{e} \langle (\ell_n^*)^j (\ell_n q_n^{k-e+j}) \rangle S^{e-j} W.
\end{aligned} \tag{2}
$$

The last equality follows from Remark 5: for every j, $S^{e-j}W^* \cdot q_n^k = q_n^{k-e+j} S^{e-j} W$.

Now, since $\ell_n^* \cdot g_n = 0$, we obtain

$$
\begin{aligned}
(\ell_n^*)^j \cdot (\ell q_n^{k-e+j}) &= (\ell_n^*)^j \cdot \sum_{i=0}^{k-e+j} \binom{k-e+j}{i} \ell_n^{2i+1} g_n^{k-e+j-i} \\
&= \sum_{i=\lceil \frac{j-1}{2} \rceil}^{k-e+j} \binom{k-e+j}{i} \frac{(2i+1)!}{(2i+1-j)!} \ell_n^{2i+1-j} g_n^{k-e+j-i}.
\end{aligned} \tag{3}
$$

In particular, the highest power of ℓ_n in $(\ell_n^*)^j \cdot (\ell_n q_n^{k-e+j})$ is $2(k-e+j)+1-j = 2k-2e+1+j$. This shows that all the summands in the last line of (2) are linearly independent. Therefore, their span has dimension

$$
\begin{aligned}
\textstyle\sum_{j=0}^{e} \dim S^{e-j} W = \sum_{j=0}^{e} \binom{e-j+n-2}{e-j} = \\
= \sum_{j=0}^{e} \binom{n-2+j}{j} = \binom{e+n-1}{e}.
\end{aligned}
$$

□

3. Two auxiliary results

Proof of Proposition 1. Consider the circuits in Figure 1 and Figure 2. Together, they give a formula of size $k(4n-2)+2k-2$ that computes $f_{n,k} = q_n^k$. Adding $n-1$ addition gates to compute ℓ_n and one multiplication gate to multiply ℓ with q_n^k gives a formula of size $k(4n-2)+2k-2+n$ for $\widetilde{f}_{n,k}$.

□

Proposition 8. *If $n = 2m$ (resp. $n = 2m+1$) then the polynomial $f_{n,k}$ is a specialization of the matrix powering polynomial Pow_{m+1}^{2k} (resp. Pow_{m+2}^{2k}). If $n = 2m$ (resp. $n = 2m+1$) then the polynomial $\widetilde{f}_{n,k}$ is a specialization of the iterated matrix multiplication polynomial $\mathrm{IMM}_{m+2}^{2k+1}$ (resp. $\mathrm{IMM}_{m+3}^{2k+1}$).*

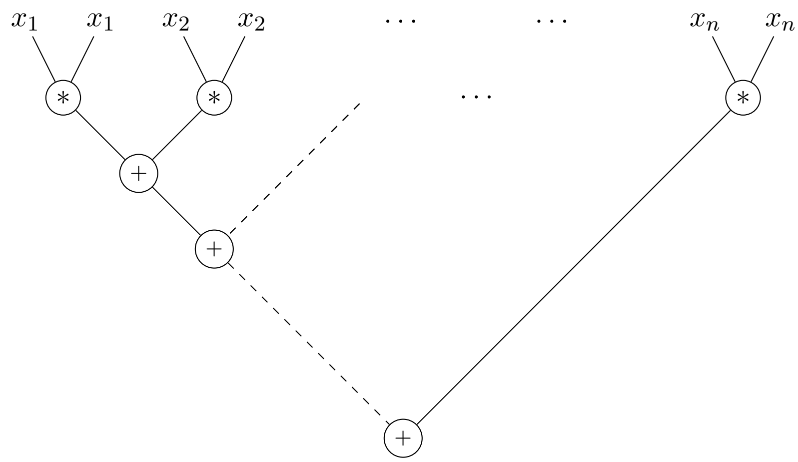

Figure 1. A circuit to compute q_n of size $4n-2$

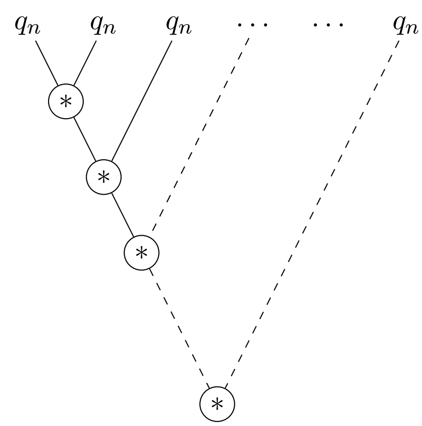

Figure 2. A circuit with input q_n's to compute $f_{n,k} = q_n^k$ of size $2k-2$

Proof. Let $n = 2m+1$ and set $y_j^{\pm} = x_{2j-1} \pm \sqrt{-1}x_{2j}$ for $j = 1, \dots, m$. Consider the specialization of Pow_{m+2}^{2k} to the matrix

$$Q_m = \begin{pmatrix} 0 & y_1^+ & \cdots & y_m^+ & x_n \\ y_1^- & & & & \\ \vdots & & \mathbf{0} & & \\ y_m^- & & & & \\ x_n & & & & \end{pmatrix},$$

of size $m+2$.

We show that $\mathrm{Pow}_{m+2}^{2k}(Q_m) = f_{n,k}$, up to scale. The characteristic polynomial of Q_m is

$$\det(Q_m - t\mathrm{Id}_{m+2}) = (-1)^{m+2}t^m \left(t^2 - \textstyle\sum_j y_j^+ y_j^- - x_n^2\right) = (-1)^{m+2}t^m \left(t^2 - q_n\right).$$

Thus, the nonzero eigenvalues of Q_m (as functions of $x_1, ..., x_n$), are $\pm\sqrt{q_n}$. In particular, $\mathrm{Pow}_{m+2}^d(Q_m) = (\sqrt{q_n})^d + (-\sqrt{q_n})^d$ is 0 if d is odd and it is $2q_n^k = 2f_{n,k}$ if $d = 2k$ is even.

If $n = 2m$ is even, apply the same argument to the matrix obtained from Q_m by removing the last row and the last column.

Similarly, $\widetilde{f}_{n,k}$ is a specialization of $\mathrm{IMM}^{2k+1}_{m+2}$ or $\mathrm{IMM}^{2k+1}_{m+3}$ (depending on the parity of n) by making the first matrix $\ell_n \mathrm{Id}$ and specializing the remaining matrices to the matrix above. $\square$

4. Proof of Theorem 3

Choose a linear inclusion $\mathbb{C}^{m^2} \subset \mathbb{C}^{mn^2}$ and regard $\mathrm{perm}_m \in S^m\mathbb{C}^{mn^2}$. Our goal is to show that for every s, τ

$$\dim\langle \partial^{=s}\mathrm{IMM}^m_n\rangle_{=\tau} \geq \dim\langle \partial^{=s}\mathrm{perm}_m\rangle_{=\tau}.$$

We split the proof into three cases. In the first and in the second case, we degenerate IMM^m_n to $f_{n,k}$ if $m = 2k$ is even and to $\widetilde{f}_{n,k}$ if $m = 2k+1$ is odd. This is possible by Proposition 8. Write $F_{m,n}$ for either $f_{n,k}$ or $\widetilde{f}_{n,k}$ in what follows. Since IMM^m_n degenerates to $F_{m,n}$, we have $\dim\langle \partial^{=s}\mathrm{IMM}^m_n\rangle_{=\tau} \geq \dim\langle \partial^{=s}F_{m,n}\rangle_{=\tau}$.

In the third case, we specialize IMM^m_n to the power sum polynomial of degree m in m^2 variables $y_1^m + \cdots + y_{m^2}^m$ by specializing every argument of IMM^m_n to the diagonal matrix of size $n \times n$ with $y_1, \dots, y_{m^2}$ in the first m^2 diagonal entries and 0 elsewhere.

Case 1: $s \geq \lceil \frac{m}{2} \rceil$. We show that $\dim\langle \partial^{=s}F_{m,n}\rangle_{=\tau} \geq \dim\langle \partial^{=s}\mathrm{perm}_m\rangle_{=\tau}$ when $s \geq \lceil \frac{m}{2} \rceil$. Up to the action of GL_{mn^2}, assume $\mathbb{C}^m \subseteq \mathbb{C}^n \subseteq \mathbb{C}^{mn^2}$, where $\mathbb{C}^m$ is the space spanned by the variables of perm_m and $\mathbb{C}^n$ is the space spanned by the variables of $F_{m,n}$. It will suffice to prove $\dim\langle \partial^{=s}(\mathrm{perm}_m \in S^m\mathbb{C}^n)\rangle_{=\tau} \leq \dim\langle \partial^{=s}F_{m,n} \in S^m\mathbb{C}^n\rangle_{=\tau}$ because the remaining $mn^2 - n$ variables will contribute the same growth to the ideals $\mathcal{J}_s(\mathrm{perm}_m)$ and $\mathcal{J}_s(F_{m,n})$. Since $s \geq \lceil \frac{m}{2} \rceil$, $(F_{m,n})_{s,m-s}$ surjects onto $S^{m-s}\mathbb{C}^n$, and thus, for every shift τ, the shifted partial derivative map surjects onto $S^{m-s+\tau}\mathbb{C}^n$ for all τ. This shows $\langle \partial^{=s}F_{m,n} \in S^m\mathbb{C}^n\rangle_{=\tau} = S^{m-s+\tau}\mathbb{C}^n$ and proves this case.

Case 2: $s < \lceil \frac{m}{2} \rceil$ and $\tau < 2m^3$. Again, it suffices to prove $\dim\langle \partial^{=s}\mathrm{perm}_m \in S^m\mathbb{C}^n\rangle_{=\tau} \leq \dim\langle \partial^{=s}F_{m,n} \in S^m\mathbb{C}^n\rangle_{=\tau}$. Since $s \leq \lceil m/2 \rceil$, the partials of $F_{m,n}$ have image of dimension $\binom{n+s-1}{s}$. By a variant of Macaulay's theorem [5, Cor. 2.4] we have the estimate (that is an equality in the case m is even)

$$\dim\langle \partial^{=s}F_{m,n} \in S^m\mathbb{C}^n\rangle_{=\tau} \geq \binom{n+s+\tau-1}{s+\tau}.$$

We compare this with the crude estimate for perm_m that ignores syzygies of its s-th Jacobian ideal: $\langle \partial^{=s}\mathrm{perm}_m \subset S^m\mathbb{C}^n\rangle_{=0}$ has dimension $\binom{m}{2}^2$ so, ignoring syzygies of $\mathcal{J}_s(\mathrm{perm}_m)$,

$$\binom{m}{s}^2\binom{n+\tau-1}{\tau} \geq \dim\langle \partial^{=s}(\mathrm{perm}_m \subset S^m\mathbb{C}^n)\rangle_{=\tau}.$$

We will conclude that $\binom{n+s+\tau-1}{s+\tau} > \binom{m}{s}^2\binom{n+\tau-1}{\tau}$ in the range we consider. This is equivalent to

$$\frac{(n+s+\tau-1)(n+s+\tau-2)\cdots(n+\tau)}{(\tau+s)(\tau+s-1)\cdots(\tau+1)} > \binom{m}{s}^2. \tag{4}$$

Taking the logarithm of the left hand side of (4), we have

$$\ln\left(\frac{(n+s+\tau-1)(n+s+\tau-2)\cdots(n+\tau)}{(\tau+s)(\tau+s-1)\cdots(\tau+1)}\right) \geq s\ln\left(1+\frac{n-s}{\tau+s}\right).$$

Taking the logarithm of the right hand side of (4), we obtain

$$\ln\binom{m}{s}^2 \leq s\ln(m^2).$$

Therefore (4) holds if

$$\frac{n-s}{\tau+s} > m^2.$$

Using $n \geq m^6$ and $\tau \leq 2m^3$ we conclude because $s \leq m$.

Case 3: $s < \frac{m}{2}$ and $\tau > m^3$. Here set all matrices $(X_1, ..., X_m)$ equal to a matrix that is zero except for the first m^2 entries on the diagonal, call them $y_1, ..., y_{m^2}$. The resulting degeneration of IMM_n^m is $y_1^m + \cdots + y_{m^2}^m$. We compare the shifted partials of both polynomials in m^2 variables. The space of partial derivatives of order s is $\langle y_1^{m-s}, ..., y_{m^2}^{m-s}\rangle$. The image of the τ-th shifted partial map consists of all polynomials in m^2 variables of degree $m-s+\tau$ as soon as $m-s+\tau > m^2(m-s-1)+1$, so that every monomial of degree $m-s+\tau$ is divisible by at least one power of order $m-s$. In particular, the shifted partials derivative map is surjective whenever when $\tau > m^3$.

5. Complete symmetric functions

Recall that $h_{n,d}$ is the complete symmetric function of degree d in n variables:

$$h_{n,d} = \sum_{|\alpha|=d} \mathbf{x}^\alpha,$$

where the summation is over all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_1 + \cdots + \alpha_n = d$.

Proposition 9. *For every monomial $\mathbf{x}^\beta$ with $|\beta| = e \leq d$, we have*

$$\frac{\partial^e}{\partial \mathbf{x}^\beta} h_{n,d} = \beta! \cdot h_{n+e,d-e}(x_1, \ldots, x_n, \mathbf{x}^{(\beta)}),$$

where $\mathbf{x}^{(\beta)} = (\underbrace{x_1, \ldots, x_1}_{\beta_1}, \ldots, \underbrace{x_n, \ldots, x_n}_{\beta_n})$ and $\beta! = \beta_1! \cdots \beta_n!$. In particular the image of the flattening $(h_{n,d})_{e,d-e}$ is the space

$$\left\langle h_{n+e,d-e}(x_1, \ldots, x_n, \mathbf{x}^{(\beta)}) : |\beta| = e \right\rangle \subseteq S^{d-e}\mathbb{C}^n.$$

Proof. We proceed by induction on e. If $e=1$, suppose $\mathbf{x}^\beta = x_n$ and write $h_{n,d} = \sum_{j=0}^d x_n^j h_{n-1,d-j}(x_1,\dots,x_{n-1})$, so that

$$\begin{aligned}\frac{\partial}{\partial x_n} h_{n,d} &= \sum_{j=0}^{d} j \cdot x_n^{j-1} h_{n-1,d-j}(x_1,\dots,x_{n-1}) = \\ &= \sum_{\ell=0}^{d-1} (\ell+1)\cdot x_n^\ell h_{n-1,d-\ell-1}(x_1,\dots,x_{n-1}) = \\ &= \left[\sum_{\ell=0}^{d-1} x_n^\ell h_{n-1,d-\ell-1}(x_1,\dots,x_{n-1})\right] + x_n \left[\sum_{\ell=0}^{d-2} (\ell+1)\cdot x_n^\ell h_{n-1,d-\ell-2}(x_1,\dots,x_{n-1})\right]\end{aligned}$$

where the first summation in the last line contains one term from each summand in the previous line, and the second summation contains the remaining terms (with shifted indices). The first summation adds up to $h_{n,d-1}(x_1,\dots,x_n)$; by repeating this on the second summation we obtain

$$h_{n,d-1}(x_1,\dots,x_n) + x_n h_{n,d-2}(x_1,\dots,x_n) + x_n^2 \left[\sum_{\ell=0}^{d-3} (\ell+1)\cdot x_n^\ell h_{n-1,d-\ell-3}(x_1,\dots,x_{n-1})\right]$$

and iterating this process we obtain $\sum_{j=0}^d x_n^j h_{n,d-1-j}(x_1,\dots,x_n) = h_{n+1,d-1}(x_1,\dots,x_n,x_n)$ proving the base case.

Let $e \geq 1$ and suppose $\beta_1 \geq 1$. Let $\gamma = (\beta_1 - 1, \beta_2, \dots, \beta_n)$. We have

$$\begin{aligned}\frac{\partial^e}{\partial \mathbf{x}^\beta} h_{n,d} &= \frac{\partial}{\partial x_1}\frac{\partial^{e-1}}{\partial \mathbf{x}^\gamma} h_{n,d} = \\ &= \gamma! \cdot \frac{\partial}{\partial x_1} h_{n+e-1,d-e+1}(x_1,\dots,x_n,\mathbf{x}^{(\gamma)}).\end{aligned}$$

By chain rule and by symmetry

$$\begin{aligned}&\frac{\partial}{\partial x_1} h_{n+e-1,d-e+1}(x_1,\dots,x_n,\mathbf{x}^{(\gamma)}) = \\ &= (\gamma_1+1)\frac{\partial}{\partial y_1}\Big|_{(x_1,\dots,x_n,\mathbf{x}^{(\gamma)},x_1)} h_{n+e-1,d-e+1}(y_1,\dots,y_{n+e-1}) = \\ &= \beta_1 h_{n+e,d-e}((x_1,\dots,x_n,\mathbf{x}^{(\gamma)},x_1)),\end{aligned}$$

where we used the case $e=1$ again. Since $\gamma! \cdot \beta_1 = \beta!$, we conclude. □

Proposition 10. *For any choice of multi-indices β,γ with $|\beta| = p$ and $|\gamma| = e$, the coefficient of $\mathbf{x}^\gamma$ in $h_{n+p,e}(x_1,\dots,x_n,\mathbf{x}^{(\beta)})$ is*

$$\prod_{i=1}^n \binom{\beta_i+\gamma_i}{\gamma_i}.$$

Proof. Write $[f]_\gamma$ for the coefficient of $\mathbf{x}^\gamma$ in the polynomial f.

We use induction on p. If $p=0$, then $h_{n+p,e}(x_1,\dots,x_n,\mathbf{x}^{(\beta)}) = h_{n,p}$ and for every γ we have $[h_{n,d}]_\gamma = 1 = \prod_{i=1}^n \binom{\gamma_i}{\gamma_i}$.

Let $p \geq 1$ and suppose $\beta_1 \geq 1$. Write $h_{n+p,e}(\mathbf{y}) = \sum_{j=0}^{e} y_1^j h_{n+p-1,e-j}(y_2, \ldots, y_{n+p})$. Let $\eta_1 = (1, 0, \ldots, 0) \in \mathbb{Z}^n_{\geq 0}$. We have

$$[h_{n+p,e}(x_1, \ldots, x_n, \mathbf{x}^{(\beta)})]_\gamma = \sum_{j=0}^{\gamma_1} [h_{n+p-1,e-j}(x_1, \ldots, x_n, \mathbf{x}^{(\beta-\eta_1)})]_{\gamma - j\eta_1}.$$

Apply the inductive hypothesis to the summands of the right hand side to get

$$\begin{aligned} [h_{n+p,e}(x_1, \ldots, x_n, \mathbf{x}^{(\beta)})]_\gamma &= \sum_{j=0}^{\gamma_1} \left[\binom{\gamma_1 - j + \beta_1 - 1}{\gamma_1 - j} \cdot \prod_{i=2}^{n} \binom{\beta_i + \gamma_i}{\gamma_i} \right] = \\ &= \left(\sum_{j=0}^{\gamma_1} \binom{\beta_1 - 1 + j}{j} \right) \prod_{i=2}^{n} \binom{\beta_i + \gamma_i}{\gamma_i} = \\ &= \prod_{i=1}^{n} \binom{\beta_i + \gamma_i}{\gamma_i} \end{aligned}$$

□

Let $a_1, \ldots, a_N \in \mathbb{Z}_{\geq 0}$ be nonnegative integers and let $G(a_1, \ldots, a_N)$ to be the $N \times N$ symmetric matrix whose (i,j)-th entry is $\binom{a_i + a_j}{a_i}$. The Lindström-Gessel-Viennot Lemma (see [8, §2]) guarantees that $G(a_1, \ldots, a_N)$ is a totally nonnegative matrix (in the sense that every minor is nonnegative), and its rank is equal to the number of distinct a_i's. In particular, $G(a_1, \ldots, a_N)$ is always positive semidefinite and it is positive definite if and only if the a_i's are distinct. Moreover if $a_{i_1} = a_{i_2}$ for some i_1, i_2, then the i_1-th and i_2-th rows are equal.

Given two matrices A, B of the same size, define $A \odot B$ to be the *Hadamard product* of A and B. For vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N \in \mathbb{Z}^m_{\geq 0}$, with $\mathbf{a}_i = (a_{ij})_{j=1,\ldots,m}$ define

$$G(\mathbf{a}_1, \ldots, \mathbf{a}_N) := \bigodot_{i=1}^{n} G(a_{1,i}, \ldots, a_{N,i}).$$

Our goal is to prove that $G(\mathbf{a}_1, \ldots, \mathbf{a}_N)$ is positive definite if the $\mathbf{a}_i$ are distinct.

We will need the following two technical results. Given a matrix A we denote by $A^i_\bullet$ (resp. $A^\bullet_i$) the i-th row (resp. column) of A and by A^I_J the submatrix consisting of rows in the set of indices I and columns in the set of indices J.

Lemma 11. *Let A be symmetric, positive semidefinite. Let $I = \{i_1, \ldots, i_r\}$ be a set of indices such that the r vectors $\{A^\bullet_i\}_{i \in I}$ are linearly independent. Then the principal submatrix A^I_I of A has full rank.*

Proof. Without loss of generality, suppose $I = \{1, \ldots, r\}$ and let $R = A^I_I$. We want to prove that R is full rank, namely that $Ru = 0$ for some $u \in \mathbb{R}^r$ implies $u = 0$. Let $v \in \mathbb{R}^n$ such that $v_i = u_i$ if $i \leq r$ and $v_i = 0$ if $i > r$. Since A is positive semidefinite, write $A = B^T B$. We have

$$0 = u^T R u = v^T A v = v^T B^T B v = \|Bv\|.$$

In particular $Bv = 0$, therefore $Av = 0$; since the first r columns of A are linearly independent, we deduce $v = 0$, so that $u = 0$ and R is nonsingular. □

Lemma 12. *Let A, B be symmetric $N \times N$ matrices such that A is positive definite and B is positive semidefinite with strictly positive diagonal entries. Then $A \odot B$ is positive definite.*

Proof. Given a symmetric matrix C, denote by $C_{(k)}$ the k-th leading principal submatrix of C, namely the $k \times k$ submatrix consisting of the first k rows and k columns of C. C is positive definite if an only if $\det(C_{(k)})$ is positive for every k. Therefore it suffices to show that the leading principal minors of $A \odot B$ are positive.

Let k_0 be the smallest k such that $\det(B_{(k)}) = 0$. From Schur's Product Theorem (see, e.g. [19, Ex. 36.2.1]), the Hadamard product of positive definite matrices is positive definite. For every $k < k_0$, we have that $B_{(k)}$ is positive definite, so $(A \odot B)_{(k)} = A_{(k)} \odot B_{(k)}$ is positive definite as well and in particular it has positive determinant.

If $k \geq k_0$, from [17, Eqn. 1.11], we have

$$\det(A_{(k)} \odot B_{(k)}) + \det(A_{(k)}B_{(k)}) \geq \det(A_{(k)})\textstyle\prod_{i=1}^{k} b_{ii} + \det(B_{(k)})\prod_{i=1}^{k} a_{ii},$$

and so, since $\det(B_k) = 0$ for every $k \geq k_0$,

$$\det((A \odot B)_{(k)}) \geq \det(A_{(k)})\textstyle\prod_{i=1}^{k} b_{ii} > 0$$

proving that $(A \odot B)_{(k)}$ has positive determinant. □

Proposition 13. *Let $\mathbf{a}_1, \dots, \mathbf{a}_N \in \mathbb{Z}^m_{\geq 0}$. Then the rank of $G(\mathbf{a}_1, \dots, \mathbf{a}_N)$ equals the number of distinct m-tuples $\mathbf{a}_1, \dots, \mathbf{a}_N$.*

Proof. We proceed by induction on m. For $m = 1$, the statement follows from the Lindström-Gessel-Viennot Lemma.

If $m \geq 2$, for every $i = 1, \dots, N$, write $\mathbf{a}_i = (\mathbf{a}'_i, a_{i,N})$. Let $A = G(\mathbf{a}'_1, \dots, \mathbf{a}'_N)$ and $B = G(a_{m,1}, \dots, a_{m,N})$. By the induction hypothesis A is positive semidefinite and its rank is equal to the number of distinct $\mathbf{a}'_i$'s. Similarly for B.

Let $C = G(\mathbf{a}_1, \dots, \mathbf{a}_N) = A \odot B$. If two pairs $\mathbf{a}_i = \mathbf{a}_j$, then $A^\bullet_i = A^\bullet_j$ and $B^\bullet_i = B^\bullet_j$ so that the corresponding two columns of C are equal. Conversely, if two columns $C^\bullet_i, C^\bullet_j$ are equal, we show that the corresponding m-tuples $\mathbf{a}_i$ and $\mathbf{a}_j$ are equal. Consider the principal 2×2 submatrix obtained from these two columns:

$$C^{ij}_{ij} = A^{ij}_{ij} \odot B^{ij}_{ij}.$$

If $A^\bullet_i \neq A^\bullet_j$, then they are linearly independent by induction hypothesis, and by Lemma 11 the submatrix A^{ij}_{ij} is positive definite. The submatrix B^{ij}_{ij} is positive semidefinite and has strictly positive diagonal entries, therefore by Lemma 12 C^{ij}_{ij} is positive definite, in contradiction with the assumption. This shows that if $C^\bullet_i = C^\bullet_j$, then $A^\bullet_i = A^\bullet_j$ and therefore $B^\bullet_i = B^\bullet_j$. In particular $\mathbf{a}_i = \mathbf{a}_j$.

Therefore, we may assume that C has distinct columns and our goal is to show that C has full rank. Suppose by contradiction that C does not have full rank and let

$$0 = \alpha_1 C^\bullet_1 + \cdots + \alpha_N C^\bullet_N \tag{5}$$

be a vanishing linear combination of the columns of C.

Up to conjugation by a permutation matrix, suppose there exist $0 = k_0 < k_1 < \cdots < k_r = N$ such that $\mathbf{a}'_i = \mathbf{a}'_j$ if $k_s < i, j \leq k_{s+1}$ and $\mathbf{a}'_i \neq \mathbf{a}'_j$ otherwise. Notice that if the

$\mathbf{a}_i$'s are distinct, then A has full rank and so does C from Lemma 12 because B is positive semidefinite with strictly positive entries. Therefore, suppose $k_1 \geq 2$ and up to reducing to a principal submatrix suppose that $\alpha_{k_1} \neq 0$. Since the first k_1 columns (and rows) of A are equal, the first k_1 columns (and rows) of B are linearly independent, otherwise two of them would be equal, providing that two m-tuples $\mathbf{a}_i$ and $\mathbf{a}_j$ for $i, j \leq k_1$ would be equal. By Lemma 11, the principal submatrix $B^{1,\dots,k_1}_{1,\dots,k_1}$ is positive definite.

The linear combination (5) can be written as

$$0 = A^\bullet_1 \odot (\alpha_1 B^\bullet_1 + \cdots + \alpha_{k_1} B^\bullet_{k_1}) + \textstyle\sum_{i>k_1} \alpha_i (A^\bullet_i \odot B^\bullet_i).$$

Define $\widetilde{A}$ to be the matrix obtained from A by removing the first $k_1 - 1$ rows and columns, that is $\widetilde{A} = G(\mathbf{a}'_{k_1}, \dots, \mathbf{a}'_N)$. $\widetilde{A}$ has the same rank as A. Let $B' = P^T B P$, for

$$P = \begin{pmatrix} 1 & & & \alpha_1 & & & \\ & \ddots & & \vdots & & & \\ & & 1 & \vdots & & & \\ & & & \alpha_k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix};$$

notice that B' is obtained from B by performing row and column operations. In particular B' has the same signature as B; moreover, from the block structure of P, we deduce that the submatrix ${B'}^{1,\dots,k_1}_{1,\dots,k_1}$ has the same signature as $B^{1,\dots,k_1}_{1,\dots,k_1}$, namely it is positive definite. This shows that the k_1-th diagonal entry of B' is strictly positive. Define $\widetilde{B}$ to be the submatrix obtained from B' by removing the first $k_1 - 1$ rows and columns. Define $\widetilde{C} = \widetilde{A} \odot \widetilde{B}$. The linear combination of (5) induces a vanishing linear combination among the columns of $\widetilde{C}$.

By repeating this procedure at most r times, we find a singular $r \times r$ matrix $\widetilde{\widetilde{C}} = \widetilde{\widetilde{A}} \odot \widetilde{\widetilde{B}}$ with $\widetilde{\widetilde{A}}$ positive semidefinite and of full rank (so positive definite) and $\widetilde{\widetilde{B}}$ positive semidefinite with strictly positive diagonal entries. By Lemma 12, we obtain a contradiction. This concludes the proof. □

Using these results, we can finally prove:

Theorem 14. *For every n, d, e, the flattening $(h_{n,d})_{e,d-e} : S^e(\mathbb{C}^n)^* \to S^{d-e}\mathbb{C}^n$ of the complete symmetric function $h_{n,d}$ has full rank.*

Proof. First, we consider the case $d = 2k$ even.

It suffices to prove the result for $e = k$.

$$(h_{n,d})_{k,k}(S^k(\mathbb{C}^n)^*) = \left\langle \frac{\partial^{|\beta|}}{\partial \mathbf{x}^\beta} h_{n,d} : |\beta| = k \right\rangle = \\ = \left\langle h_{n+k,k}(x_1, \dots, x_n, \mathbf{x}^{(\beta)}) : |\beta| = k \right\rangle.$$

Define two column vectors

$$\mathbf{h} = \left(h_{n+k,k}(x_1, \dots, x_n, \mathbf{x}^{(\beta)}) : |\beta| = k\right)^T$$
$$\mathbf{b} = \left(\mathbf{x}^{\beta} : |\beta| = k\right)^T.$$

From Proposition 9 and Proposition 10, we have $\mathbf{h} = A\mathbf{b}$ where the (β, γ)-th entry of A is

$$A_{\beta,\gamma} = \prod_{i=1}^n \binom{\beta_i + \gamma_i}{\gamma_i}$$

namely $A = G(\beta : |\beta| = k)$. Since the multi-indices β are all distinct, by applying Proposition 13 we deduce that A is nonsingular and therefore the entries of $\mathbf{b}$ are linear combinations of the entries of $\mathbf{h}$. This shows that $(h_{n,d})_{k,k}$ is full rank.

Now consider $d = 2k + 1$ odd. It suffices to prove the result for $e = k + 1$. Let $g = \frac{\partial}{\partial x_1} h_{n,d} = h_{n+1,d-1}(x_1, x_1, x_2, \dots, x_n) \in S^{2k}\mathbb{C}^n$. The image of the flattening $g_{k,k} : S^k(\mathbb{C}^n)^* \to S^k\mathbb{C}^n$ is contained in the image of $(h_{n,d})_{k+1,k}$. To conclude, we will show that $g_{k,k}$ is full rank.

Let $h = h_{n+1,d-1}(y_1, \dots, y_{n+1})$. By the result in the case of even degree, we know that $(h)_{k,k}$ has full rank, namely

$$\left\langle h_{n+1+k,k}(y_1, \dots, y_{n+1}, \mathbf{y}^{(\beta)}) : |\beta| = k \right\rangle = S^k\mathbb{C}^{n+1}.$$

In particular, the image of this space under the specialization $(y_1, \dots, y_{n+1}) = (x_1, , \dots, x_n, x_1)$ is $S^k\mathbb{C}^n$.

On the other hand, notice that,

$$h_{n+1+k,k}(y_1, \dots, y_{n+1}, \mathbf{y}^{(\beta)})|_{(y_1,\dots,y_{n+1})=(x_1,x_1,\dots,x_n)} = h_{n+k+1,k}(x_1, \dots, x_n, \mathbf{x}^{(\gamma)})$$

where $\gamma_1 = \beta_1 + \beta_{n+1} + 1$ and $\gamma_i = \beta_i$ for $i = 2, \dots, n$ (indeed $|\gamma| = k + 1$). This shows that the image of $(h)_{k,k}$ is $S^k\mathbb{C}^n$, and therefore $(h_{n,d})_{k+1,k}$ is full rank. $\square$

Theorem 14 gives us the following immediate result

Corollary 15. *For every n, d, we have*

$$\mathbf{R}_S(h_{n,d}) \geq \underline{\mathbf{R}}(h_{n,d}) \geq \binom{n+p-1}{p}.$$

For readers familiar with cactus rank and border rank, by [12, Thm. 5.3D], we obtain the same lower bounds for cactus rank and border rank.

6. Koszul flattenings

We recall the Koszul flattenings introduced in [14]: Tensor the s-th partial derivative map (1) with the identity map on $\Lambda^q\mathbb{C}^N$ for some q, and then compose with the exterior derivative map

$$\delta = \delta_{q,d-s} : \Lambda^q\mathbb{C}^N \otimes S^{d-s}\mathbb{C}^N \to \Lambda^{q+1}\mathbb{C}^N \otimes S^{d-s-1}\mathbb{C}^N$$
$$X \otimes f \mapsto \sum_j \left((x_j \wedge X) \otimes \frac{\partial f}{\partial x_j} \right).$$

Let $p^{\wedge q}_{s,d-s} : \Lambda^q\mathbb{C}^N\otimes S^s\mathbb{C}^{N*} \to \Lambda^{q+1}\mathbb{C}^N\otimes S^{d-s-1}\mathbb{C}^N$ denote the composition of the partial derivative map and the exterior derivative.

Then, by [14, Prop 4.1.1]

$$\underline{\mathbf{R}}_S(p) \geq \frac{\operatorname{rank}(p^{\wedge q}_{s,d-s})}{\operatorname{rank}((\ell^d)^{\wedge q}_{s,d-s})} = \frac{\operatorname{rank}(p^{\wedge q}_{s,d-s})}{\binom{N-1}{q}}. \tag{6}$$

When $p_{s,d-s}$ is of maximal rank for all s, Koszul flattenings can only give a better Waring border rank lower bound than the partial derivative maps when $d = 2k+1$ is odd and $s = k$. For example, if $d = 2k$ is even, then the rank of the Koszul flattening is bounded above by $\dim\left(\Lambda^q\mathbb{C}^N \otimes S^{k-1}\mathbb{C}^N\right)$ so the Waring border rank lower bound from Koszul flattenings is bounded above by $\binom{N+k-2}{k-1}\frac{N}{N-q}$ whereas from flattenings alone, one already gets the larger lower bound of $\binom{N+k-1}{k}$. Similarly, if d is odd but $N \neq 2q+1$, the bound from Koszul flattenings is lower than the one from standard flattenings.

Suppose $d = 2k+1$ and $p_{k,k+1}$ is of maximal rank. Viewed naïvely, one might think Koszul flattenings could prove Waring border rank lower bounds of up to

$$\frac{\dim(\Lambda^q\mathbb{C}^N \otimes S^k\mathbb{C}^{N*})}{\binom{N-1}{q}} = \binom{N+k-1}{k}\frac{N}{N-q}$$

but this is not possible because the exterior derivative map is a GL_N-module map that is not surjective unless $q = N, N-1$. Indeed, we have the decompositions

$$\Lambda^q\mathbb{C}^N\otimes S^{k+1}\mathbb{C}^N = S_{k+1,1^q}\mathbb{C}^N \oplus S_{k+2,1^{q-1}}\mathbb{C}^N$$
$$\Lambda^{q+1}\mathbb{C}^N\otimes S^{k}\mathbb{C}^N = S_{k,1^{q+1}}\mathbb{C}^N \oplus S_{k+1,1^{q}}\mathbb{C}^N$$

so, by Schur's Lemma, $S_{k+2,1^{q-1}}\mathbb{C}^N = \ker(\delta)$ and $\operatorname{image}(\delta) = S_{k+1,1^q}\mathbb{C}^N$.

The following result gives an *a priori* upper bound for the rank of the Koszul flattening, by determining a lower bound for the dimension of $\ker(p^{\wedge q}_{k,k+1})$.

Proposition 16. *Let $p \in S^{2k+1}\mathbb{C}^N$. Then for every q*

$$\operatorname{rank}(p^{\wedge q}_{k,k+1}) \leq \sum_{j=0}^{k}(-1)^j \dim(\Lambda^{q-j}\mathbb{C}^N \otimes S^{k-j}\mathbb{C}^{N*}). \tag{7}$$

Proof. We will prove that

$$\dim\operatorname{image}(p^{\wedge(q-1)}_{e-1,d-e+1}) \leq \dim\ker(p^{\wedge q}_{e,d-e})$$

and conclude that the estimate holds for every q via an induction argument. Indeed, $\operatorname{image}(p^{\wedge(q-1)}_{e-1,d-e+1}) \subseteq \operatorname{image}(\mathrm{Id}_{\Lambda^q\mathbb{C}^N}\otimes p_{e,d-e})$ because the exterior derivative map takes derivatives on the factor $S^{d-e+1}\mathbb{C}^N$. Moreover, since $\delta^2 = 0$, we have $\operatorname{image}(p^{\wedge q-1}_{e-1,d-e+1}) \subseteq \ker(\delta_{q,d-e})$. Therefore, passing to dimensions

$$\dim(\operatorname{image}(p^{\wedge q-1}_{e-1,d-e+1})) \leq \dim(\ker(\delta_{q,d-e}|_{\operatorname{image}(\mathrm{Id}_{\Lambda^q\mathbb{C}^N}\otimes p_{e,d-e})}) \leq \dim\ker(p^{\wedge q}_{e,d-e}).$$

Now,

$$\begin{aligned}\operatorname{rank}(p^{\wedge q}_{k,k+1}) &= \dim(\Lambda^q\mathbb{C}^N \otimes S^k\mathbb{C}^{N*}) - \dim(\ker(p^{\wedge q}_{k,k+1}))\\ &\leq \dim(\Lambda^q\mathbb{C}^N \otimes S^k\mathbb{C}^{N*}) - \operatorname{rank}(p^{q-1}_{k-1,k+2}).\end{aligned}$$

and we conclude by induction. $\square$

Remark 17. This is still not the end of the story: when $N = 2q+1$ with q odd, then the linear map $p^{\wedge q}_{k,k+1} : \Lambda^q\mathbb{C}^N \otimes S^k\mathbb{C}^{N*} \to \Lambda^{q+1}\mathbb{C}^N S^k\mathbb{C}^N \simeq \Lambda^q\mathbb{C}^{N*}S^k\mathbb{C}^N$ was observed in [14], at least in certain cases, to be skew-symmetric. In particular, in this case, if the bound in (7) is odd, it cannot be attained.

Remark 18. Since the border rank bound is obtained by dividing $\mathrm{rank}(p^{\wedge q}_{k,k+1})$ by $\binom{N-1}{q}$, we see asymptotically, the best potential lower bound is obtained when $N = 2q+1$ and there the limit of the method is twice the bounds obtained via flattenings minus lower order terms. This improvement is irrelevant for complexity. It is known more generally that the improvement in best possible lower bounds beyond the best possible bounds of partial derivatives are limited for any determinantal method. This was observed independently by Efremenko, Garg, Oliviera and Wigderson (personal communication), and Galcazka [7] for completely different reasons.

We now show Koszul flattenings can indeed give border rank lower bounds beyond the best lower bound attainable via the method of partial derivatives.

Proposition 19. *The Koszul flattening $(\widetilde{f}_{k,n})^{\wedge q}_{k,k+1}$, when $n > 2$, and all $q < \frac{n}{2}$, has rank at least*

$$\binom{n-1}{q}\left(\binom{n+k-1}{k} + q - 1\right).$$

In particular, it implies $\underline{\mathbf{R}}_S(\widetilde{f}_{k,n}) \geq \binom{n+k-1}{k} + q - 1$, which is greater than the lower bound obtainable by flattenings.

Proof. For fixed n, recall the unique $\mathfrak{S}_n$-invariant $g := g_n \in S^2W$ from the proof of Thm. 7, where W is the subspace of $\mathbb{C}^N$ isomorphic to the Specht module $[n-1,1]$ under the action of $\mathfrak{S}_n$. Let L be the span of $\ell := \ell_n$. For every s, write $p_s := (\ell^*)^{s-1} \cdot (\ell q^{s-1})$, which from Eqn. (3) is a polynomial of degree s with non-zero projection onto L^s. From Eqn. 2, the image of the k-th flattening map of $\widetilde{f}_{n,k}$ is $\mathrm{image}(\widetilde{f}_{n,k})_{k,k+1} = \bigoplus_{s=0}^k p_{s+1}S^{k-s}W \subseteq S^{k+1}\mathbb{C}^n$.

We will give a lower bound for the dimension of the image of $\Lambda^q\mathbb{C}^n \otimes \mathrm{image}(\widetilde{f}_{n,k})_{k,k+1}$ under the exterior derivative map. We have $\Lambda^q\mathbb{C}^n = \Lambda^qW \oplus \left(L \wedge \Lambda^{q-1}W\right)$ as a $\mathfrak{S}_n$-module.

Consider the image of $(L \wedge \Lambda^{q-1}W) \otimes (p_{s+1}S^{k-s}W)$ under the exterior derivative. The $\mathfrak{S}_n$-equivariant projection of this space onto $(L \wedge \Lambda^{q-1}W) \otimes (L^{s+1} \otimes S^{k-s}W)$ commutes with the exterior derivative, which is GL_n-equivariant and therefore $\mathfrak{S}_n$-equivariant. The image of $(L \wedge \Lambda^{q-1}W) \otimes (L^{s+1} \otimes S^{k-s}W)$ under the exterior derivative is (when $s \leq k-1$ and after reordering the factors) the subspace $S_{(k-s,1^{q-1})}W \otimes L \otimes L^{s+1} \subseteq L \wedge \Lambda^qW \otimes L^{s+1}S^{k-s-1}W$.

Now consider the image of $\Lambda^qW \otimes p_{s+1}S^{k-s}W$. Again, consider its projection onto $\Lambda^qW \otimes (L^{s+1} \otimes S^{k-s}W)$. By applying the exterior derivative map, we obtain a subspace of $\left((L \wedge \Lambda^qW) \otimes (L^s \otimes S^{k-s}W)\right) \oplus \left(\Lambda^{q+1}W \otimes (L^{s+1} \otimes S^{k-s-1}W)\right)$; consider its projection to the second summand $\Lambda^{q+1}W \otimes L^{s+1}S^{k-s-1}W$ when $s \leq k-1$. For the same reason as above, the image of this projection is $S_{(k-s,1^q)}W \otimes L^{s+1}$ up to reordering the factors.

Note that $S_{(k-t+1,1^{q-1})}W \oplus S_{(k-t,1^q)}W = S^{k-t}W \otimes \Lambda^q W$ as a $GL(W)$-module. Consider the summands for s ranging from 0 to $k-2$ in the first case and 1 to $k-1$ in the second, we obtain components $S^{k-t}W \otimes \Lambda^q W$ for t from 1 to $k-1$. We obtain a subspace in the image of the Koszul flattening that is isomorphic as a $GL(W)$-module to

$$(\bigoplus_{t=1}^{k-1} S^{k-t}W) \otimes \Lambda^q W.$$

The first factor of the space above has the same dimension as $S^{k-1}\mathbb{C}^n$ minus $\dim(S^0 W) = 1$.

So far we have a contribution to the rank of

$$\binom{n-1}{q}(\binom{n+k-2}{k-1} - 1).$$

Next consider the $s = 0$ contribution that one obtains by applying the exterior derivative to the component with the factor $\Lambda^q W$. The exterior derivative map is $\Lambda^q W \otimes (L \otimes S^k W) \to \Lambda^q W \wedge L \otimes S^k W \oplus \Lambda^q W \otimes (L \otimes S^{k-1} W)$. This projects isomorphically onto the first term in the target and $(\Lambda^q W \wedge L) \otimes S^k W$ does not intersect the image of any other term that we considered so far, so we obtain an additional contribution to the image of dimension $\binom{n-1+k-1}{k}\binom{n-1}{q}$. We now have a contribution of

$$\binom{n-1}{q}(\binom{n+k-2}{k-1} - 1 + \binom{n+k-2}{k}) = \binom{n-1}{q}(\binom{n+k-1}{k} - 1)$$

to the rank.

Finally consider the $s = k$ and $s = k-1$ terms in the term with a factor $L \wedge \Lambda^{q-1}W$: the sources are respectively $L \wedge \Lambda^{q-1} W \otimes \langle p_{k+1} \rangle$ and $(L \wedge \Lambda^{q-1} W) \otimes p_k W$. Consider the projections respectively to $(L \wedge \Lambda^{q-1}W) \otimes \ell^{k-1} S^2 W$ (the second factor is $\ell^{k-1} g$) and $(L \wedge \Lambda^{q-1} W) \otimes \ell^{k-2} S^3 W$ (the second factor is $\ell^{k-2} gW$). Now, applying the exterior derivative, these spaces map injectively to $(L \wedge \Lambda^q W) \otimes L^{k-1} W$ and $(L \wedge \Lambda^q W) \otimes \ell^{k-2} S^2 W$ respectively. These targets do not appear in other terms analyzed above, so we pick up $\binom{n-1}{q-1} = \frac{q}{n}\binom{n-1}{q}$ and $(n-1)\binom{n-1}{q-1} = \frac{q(n-1)}{n}\binom{n-1}{q}$ additional contributions to the rank.

Collecting all the contributions together, we conclude. □

Remark 20. A more careful analysis of the $q = 1$ case shows it also improves the flattening lower bound.

Computer experiments indicate the situation may be significantly better:

Proposition 21. *For $n = 3, \ldots, 6$, $k = 1, \ldots, 6$, let $p \in S^{2k+1}\mathbb{C}^n$ be generic. The Koszul flattening $p^{\wedge 1}_{k,k+1}$ has rank equal to the bound in* (6) *except if $n = 3$ and k is even (in accordance with Remark 17).*

For $n = 3, \ldots, 6$ and $k = 1, \ldots, 6$, let $p = h_{n,2k+1}$ or $p = \widetilde{f}_{k,n}$. The Koszul flattenings of $p^{\wedge 1}_{k,k+1}$ have rank equal to the bound in (6) *if k is odd and one less than the bound in* (6) *if k is even.*

Notice that in the cases $n = 4, 5, 6$ with k even, the Koszul flattenings give a border rank lower bound for $h_{n,2k+1}$ and $\widetilde{f}_{k,n}$ that is one less than the bound for a generic polynomial.

Question 22. *What are the ranks of the Koszul flattenings for $\widetilde{f}_{k,n}$ and $h_{n,2k+1}$ in general?*

References

1. Micheal Ben Or and Richard Cleve, *Computing algebraic formulas using a constant number of registers*, SIAM J. Comput. **21** (1992), no. 21, 54–58.
2. O. Biermann, *Über näherungsweise Cubaturen*, Monatsh. Math. Phys. **14** (1903), no. 1, 211–225. MR 1547094
3. W. Buczyńska and J. Buczyński, *Secant varieties to high degree Veronese reembeddings, catalecticant matrices and smoothable Gorenstein schemes*, J. Algebraic Geom. **23** (2014), no. 1, 63–90. MR 3121848
4. X. Chen, N. Kayal, and A. Wigderson, *Partial derivatives in arithmetic complexity and beyond*, Found. Trends Theor. Comput. Sci. **6** (2010), no. 1-2, front matter, 1–138 (2011). MR 2901512
5. K. Efremenko, J. M. Landsberg, H. Schenck, and J. Weyman, *The method of shifted partial derivatives cannot separate the permanent from the determinant*, ArXiv e-prints, to appear in MCOM (2016).
6. Hervé Fournier, Nutan Limaye, Meena Mahajan, and Srikanth Srinivasan, *The shifted partial derivative complexity of elementary symmetric polynomials*, Mathematical foundations of computer science 2015. Part II, Lecture Notes in Comput. Sci., vol. 9235, Springer, Heidelberg, 2015, pp. 324–335. MR 3419496
7. M. Galcazka, *Vector bundles give equations of cactus varieties*, ArXiv e-prints (2016).
8. I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. in Math. **58** (1985), no. 3, 300–321. MR 815360
9. E. L. Green, *Complete intersections and Gorenstein ideals*, J. Algebra **52** (1978), no. 1, 264–273. MR 0480472
10. A. Gupta, P. Kamath, N. Kayal, and R. Saptharishi, *Approaching the chasm at depth four*, Proceedings of the Conference on Computational Complexity (CCC) (2013).
11. A. Iarrobino and J. Emsalem, *Some zero-dimensional generic singularities; finite algebras having small tangent space*, Compositio Math. **36** (1978), no. 2, 145–188. MR 515043
12. A. Iarrobino and V. Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, vol. 1721, Springer-Verlag, Berlin, 1999, Appendix C by Iarrobino and Steven L. Kleiman. MR MR1735271 (2001d:14056)
13. N. Kayal, *An exponential lower bound for the sum of powers of bounded degree polynomials*, Electronic Colloquium on Computational Complexity (ECCC) **19** (2012), 81.
14. J. M. Landsberg and G. Ottaviani, *Equations for secant varieties of Veronese and other varieties*, Ann. Mat. Pura Appl. (4) **192** (2013), no. 4, 569–606. MR 3081636
15. J.M. Landsberg, *Geometry and complexity theory*, Cambridge University Press, to appear, 2017.
16. Hwangrae Lee, *Power sum decompositions of elementary symmetric polynomials*, Linear Algebra Appl. **492** (2016), 89–97. MR 3440150
17. M.S. Lynn, *On the Schur product of H-matrices and non-negative matrices, and related inequalities*, Mathematical Proceedings of the Cambridge Philosophical Society **60** (1964), no. 3, 425âĂŞ431.
18. N. Nisan, *Lower bounds for non-commutative computation*, Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '91, ACM, 1991, pp. 410–418.
19. V. V. Prasolov, *Problems and theorems in linear algebra*, Translations of Mathematical Monographs, vol. 134, American Mathematical Society, Providence, RI, 1994, Translated from the Russian manuscript by D. A. Leĭtes. MR 1277174
20. B. Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992), no. 463, viii+155. MR 1096187 (93h:11043)
21. Amir Shpilka and Avi Wigderson, *Depth-3 arithmetic circuits over fields of characteristic zero*, Comput. Complexity **10** (2001), no. 1, 1–27. MR 1867306
22. J. J. Sylvester, *On the principles of the calculus of forms*, Cambridge and Dublin Mathematical Journal (1852), 52âĂŞ97.

E-mail address: `jml@math.tamu.edu, fulvio.gesmundo@gmail.com`