On the Complexity of Partial Derivatives*

Ignacio Garcia-Marco[1], Pascal Koiran[2], Timothée Pecatte[3], and Stéphan Thomassé[4]

1 **LIF and CNRS, Aix-Marseille Université, Marsaille, France**
`iggarcia@ull.es`
2 **LIP,† Ecole Normale Supérieure de Lyon, Université de Lyon, Lyon, France**
`Pascal.Koiran@ens-lyon.fr`
3 **LIP, Ecole Normale Supérieure de Lyon, Université de Lyon, Lyon, France**
`Timothee.Pecatte@ens-lyon.fr`
4 **LIP, Ecole Normale Supérieure de Lyon, Université de Lyon, Lyon, France**
`Stephan.Thomasse@ens-lyon.fr`

Abstract

The method of partial derivatives is one of the most successful lower bound methods for arithmetic circuits. It uses as a complexity measure the dimension of the span of the partial derivatives of a polynomial. In this paper, we consider this complexity measure as a computational problem: for an input polynomial given as the sum of its nonzero monomials, what is the complexity of computing the dimension of its space of partial derivatives?

We show that this problem is $\sharp$P-hard and we ask whether it belongs to $\sharp$P. We analyze the "trace method", recently used in combinatorics and in algebraic complexity to lower bound the rank of certain matrices. We show that this method provides a polynomial-time computable lower bound on the dimension of the span of partial derivatives, and from this method we derive closed-form lower bounds. We leave as an open problem the existence of an approximation algorithm with reasonable performance guarantees.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes, F.2 Analysis of algorithms and problem complexity

Keywords and phrases counting complexity, simplicial complex, lower bounds, arithmetic circuits

Digital Object Identifier 10.4230/LIPIcs.STACS.2017.37

1 Introduction

Circuit lower bounds against a class of circuits $\mathcal{C}$ are often obtained by defining an appropriate complexity measure which is small for small circuits of $\mathcal{C}$ but is high for some explicit " hard function." For arithmetic circuits, one of the most successful complexity measures is based on partial derivatives. Sums of powers of linear forms provide the simplest model where the method of partial derivatives can be presented (see for instance Chapter 10 of the survey by Chen, Kayal and Wigderson [2]). In this model, a homogeneous polynomial $f(x_1, \ldots, x_n)$ of degree d is given by an expression of the form:

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{r} l_i(x_1, \ldots, x_n)^d \tag{1}$$

* Work supported by ANR project CompA (code ANR–13–BS02–0001–01).
† UMR 5668 ENS Lyon, CNRS, UCBL, INRIA.

© Ignacio Garcia-Marco, Pascal Koiran, Timothée Pecatte, and Stéphan Thomassé;
licensed under Creative Commons License CC-BY
34th Symposium on Theoretical Aspects of Computer Science (STACS 2017).
Editors: Heribert Vollmer and Brigitte Vallée; Article No. 37; pp. 37:1–37:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

where the l_i's are linear functions. The smallest possible r is often called the Waring rank of f in the algebra litterature. One takes as complexity measure $\dim \partial^{=k} f$, where $\partial^{=k} f$ denotes the linear space of polynomials spanned by the partial derivatives of f of order k. For any $k \leq d$, the derivatives of order k of a d-th power of a linear form $l(x_1, \ldots, x_n)$ are constant multiples of l^{d-k}. Therefore, by linearity of derivatives we have for any $k \leq d$ the lower bound $r \geq \dim \partial^{=k} f$ on the Waring rank of f.

The method of partial derivatives was introduced in the complexity theory litterature by Nisan and Wigderson [14], where lower bounds were given for more powerful models than (1) such as e.g. depth 3 arithmetic circuits. In such a circuit, the d-th powers in (1) are replaced by products of d affine functions. We then have [14] the lower bound $r \geq (\dim \partial^* f)/2^d$, where r denotes as in (1) the fan-in of the circuit's output gate and $\partial^* f$ denotes the space spanned by partial derivatives of all order. More recently, a number of new lower bound results were obtained using a refinement of the method of partial derivatives. These new results are based on "shifted partial derivatives" (see the continuously updated online survey maintained by Saptharishi [17] for an extensive list of references), but we will stick to "unshifted" derivatives in this paper.

Partial derivatives can also be used for *upper bound* results: see in particular Theorem 5 in [9] for an algorithm that constructs a representation in the Waring model (1) of a polynomial given by a black box. To learn more on the complexity of circuit reconstruction for various classes of arithmetic circuits one may consult Chapter 5 of the survey by Shpilka and Yehudayoff [18].

Our contributions

In this paper we consider the dimension of the set of partial derivatives as a computational problem and provide the first results (that we are aware of) on its complexity. This is quite a natural problem since, as explained above, the knowledge of this dimension for an input polynomial f provides estimates on the circuit size of f for several classes of arithmetic circuits. We assume that the input polynomial f is given in the sparse representation (also called "expanded representation"), i.e., as the sum of its nonzero monomials. We show in Section 4 that computing $\dim \partial^* f$ is hard for Valiant's [20] counting class $\sharp\mathsf{P}$. This remains true even if f is multilinear, homogeneous and has only 0/1 coefficients. The precise complexity of this problem remains open, in particular we do not know whether computing $\dim \partial^* f$ is in $\sharp\mathsf{P}$.

As an intermediate step toward our $\sharp\mathsf{P}$-hardness result, we obtain a result of independent interest for a problem of topological origin: computing the number of faces in an abstract simplicial complex. Our $\sharp\mathsf{P}$-hardness proof for this problem proceeds by reduction from counting the number of independent sets in a graph, a well-known $\sharp\mathsf{P}$-complete problem [15]. It is inspired by the recent proof [16] that computing the Euler characteristic of abstract simplicial complexes is $\sharp\mathsf{P}$-complete.

Since the $\sharp\mathsf{P}$-hardness result rules out an efficient algorithm for the exact computation of $\dim \partial^* f$, it is of interest to obtain efficiently computable upper and lower bounds for this quantity and for $\dim \partial^{=k} f$. Upper bounds are easily obtained from the linearity of derivatives. In Section 2 we give a lower bound that is based on the consideration of a single "extremal" monomial of f. In particular, for a multilinear homogeneous polynomial of degree d with s monomials we have $\binom{d}{k} \leq \dim \partial^{=k} f \leq s\binom{d}{k}$ for every k. In Section 3 we provide lower bounds that take all monomials of f into account. Depending on the choice of the input polynomial, these lower bounds may be better or worse than the lower bound of Section 2. The lower bounds of Section 3 are based on the "trace method." This method was recently used in [10, 11] to lower bound the dimension of *shifted* partial derivatives of a specific "hard" polynomial, the so-called Nisan-Wigderson polynomial. In [10] this method is attributed to Noga Alon [1].

In a nutshell, the principle of the trace method is as follows. Suppose that we want to lower bound the rank of a matrix M. In this paper, M will be the matrix of partial derivatives of a polynomial $f(x_1,\dots,x_n)$. From M, we construct the symmetric matrix $B = M^T.M$. We have $\mathrm{rank}(M) \geq \mathrm{rank}(B)$, with equality if the ranks are computed over the field of real numbers. In the trace method, we replace $\mathrm{rank}(M) = \mathrm{rank}(B)$ by the "proxy rank" $\mathrm{Tr}(B)^2/\mathrm{Tr}(B^2)$. This is legitimate due to the the inequality

$$\mathrm{rank}(B) \geq \mathrm{Tr}(B)^2/\mathrm{Tr}(B^2), \tag{2}$$

which follows from the Cauchy-Schwarz inequality applied to the eigenvalues of B. It is often easier to lower bound the proxy rank than to lower bound the rank directly. In Section 3 we will see that the proxy rank can be computed in polynomial time. This is not self-evident because B may be of size exponential in the number n of variables of f. By contrast, as explained above computing $\mathrm{rank}(B)$ over the field of real numbers is $\sharp$P-hard.

Organization of the paper

In Section 2 we set up the notation for the rest of the paper, and give some elementary estimates. In particular, Theorem 1 provides a lower bound that relies on the consideration of a single extremal monomial of f. Section 3 is devoted to the trace method. We use this method to derive closed-form lower bounds on the dimension of the space of partial derivatives, and compare them to the lower bound from Theorem 1. In Section 3.2 we show that the "proxy rank" $\mathrm{Tr}(B)^2/\mathrm{Tr}(B^2)$ is computable in polynomial time. In Section 3.3 we show that the trace method behaves very poorly on elementary symmetric polynomials: for certain settings of parameters, the matrix of partial derivatives has full rank but the trace method can only show that its rank is larger than 1. Finally, we show in Section 4 that it is $\sharp$P-hard to compute $\dim \partial^* f$ and to compute the number of faces in an abstract simplical complex.

Open problems

Here are three of the main problems that are left open by this work.

1. Give a nontrivial upper bound on the complexity of computing $\dim \partial^* f$ and $\dim \partial^{=k} f$. In particular, are these two problems in $\sharp$P?
2. Give an efficient algorithm that approximates $\dim \partial^* f$ or $\dim \partial^{=k} f$, and comes with a reasonable performance guarantee (or show that such an algorithm does not exist). The proxy rank $\mathrm{Tr}(B)^2/\mathrm{Tr}(B^2)$ is efficiently computable, but certainly does not fit the bill due to its poor performance on symmetric polynomials. For counting the number of independent sets in a graph (the starting point of our reductions), there is already a significant amount of work on approximation algorithms [13, 4] and hardness of approximation [13, 3].
3. We recalled at the beginning of the introduction that partial derivatives are useful as a complexity measure to prove lower bounds against several classes of arithmetic circuits. We saw that computing this measure is hard, but is it hard to compute the Waring rank of a homogeneous polynomial f given in expanded form, or to compute the size of the smallest (homogeneous) depth 3 circuit for f? The former problem is conjectured to be NP-hard already for polynomials of degree 3: see Conjecture 13.2 in [8] which is formulated in the language of symmetric tensors.

2 Elementary bounds

We use the notation $\partial_\beta f$ for partial derivatives of a polynomial $f(x_1,\ldots,x_n)$. Here β is a n-tuple of integers, and β_i is the number of times that we differentiate f with respect to x_i. We denote by $\partial^{=k} f$ the linear space spanned by the partial derivatives of f of order k, and by $\partial^* f$ the space spanned by partial derivatives of all order. For $\alpha \in \{0,1\}^n$, we denote by x^α the multilinear monomial $x_1^{\alpha_1}.\cdots.x_n^{\alpha_n}$. More generally, if α is a n-tuple of integers, x^α denotes the monomial $x_1^{\alpha_1}.\cdots.x_n^{\alpha_n}/(\alpha_1!\cdots\alpha_n!)$. These monomials form a basis of the space $\mathbb{R}[x_1,\ldots,x_n]$ of real polynomials in n variables, which we refer to as the "scaled monomial basis." Dividing by the constant $\alpha_1!\cdots\alpha_n!$ is convenient since differentiation takes the simple form: $\partial_\beta x^\alpha = x^{\alpha-\beta}$. We agree that $x^{\alpha-\beta}=0$ if one of the components of $\alpha-\beta$ is negative.

For a monomial $f = x^\alpha$, $\dim \partial^* x^\alpha = \prod_{i=1}^n(\alpha_i+1)$. One can compute $\dim \partial^{=k} x^\alpha$ by dynamic programming thanks to the recurrence relation:

$$\dim \partial^{=k} x^\alpha = \sum_{j=0}^{\alpha_1} \dim \partial^{=k-j}(x_2^{\alpha_2}.\cdots.x_n^{\alpha_n}).$$

It takes altogether $O((\deg f)^2)$ additions to compute the $\deg(f)+1$ numbers $\dim \partial^{=k} f$ for $k=0,\ldots,\deg(f)$. Equivalently, $\dim \partial^{=k} x^\alpha$ can be computed as the coefficient of t^k in the polynomial

$$(1+t+\ldots+t^{\alpha_1}).(1+t+\ldots+t^{\alpha_2}).\cdots.(1+t+\ldots+t^{\alpha_n}).$$

For a polynomial with more than one monomial, one can obtain simple upper bounds thanks to the linearity of derivatives since $\dim \partial^*(f+g) \le \dim \partial^* f + \dim \partial^* g$ and $\dim \partial^{=k}(f+g) \le \dim \partial^{=k} f + \dim \partial^{=k} g$. Lower bounding the dimension of the space of partial derivatives is slightly less immediate.

▶ **Theorem 1.** *For any polynomial f there is a monomial m in f such that $\dim \partial^{=k} f \ge \dim \partial^{=k} m$ for every k. In particular, if all monomials in f contain at least r variables then $\dim \partial^{=k} f \ge \binom{r}{k}$ for every k.*

Proof. The second claim clearly follows from the first claim. Let n be the number of variables in f. In order to find the monomial m, we fix a total order $\le$ on n-tuple of integers which is compatible with addition, for instance the lexicographic order (different orders may lead to different m's). We will use $\le$ to order monomials as well as tuples β in partial derivatives such as $\partial_\beta f$. We will also use the partial order $\subseteq$ defined by: $\beta \subseteq \alpha$ iff $\beta_i \le \alpha_i$ for all $i=1,\ldots,n$. Let $m = x^\alpha$ be the smallest monomial for $\le$ with a nonzero coefficient in f.

To complete the proof of the theorem, we just need to show that the partial derivatives $\partial_\beta f$ where $\beta \subseteq \alpha$ are linearly independent. The dimension of the space spanned by these partial derivatives is equal to the rank of a certain matrix M. The rows of M are indexed by the n-tuples β such that $\beta \subseteq \alpha$, and row β contains the coordinates of $\partial_\beta f$ in the scaled monomial basis (x^γ). If $f = \sum_\gamma a_\gamma x^\gamma$, we therefore have $M_{\beta,\gamma-\beta} = a_\gamma$. Let us order the rows and columns of M according to $\le$. We have seen that M contains a nonzero coefficient in row β and column $\alpha-\beta$. This coefficient is strictly to the left of any nonzero coefficient in any row above β. Indeed, we have $\alpha-\beta < \alpha'-\beta'$ if $\alpha' \ge \alpha$ and $\beta' < \beta$. Our matrix is therefore in row echelon form, and does not contain any identically zero row. It is therefore of full row rank. ◀

▶ Remark. Recall that the Newton polytope of f is the convex hull of the n-tuples of exponents of monomials of f. By changing the order $\le$ in the proof of Theorem 1 we can take for m any vertex of the Newton polytope.

Theorem 1 lower bounds $\dim \partial^{=k} f$ by the same dimension computed for a suitable monomial of f. This is of course tight if f has a single monomial. We note that adding more monomials does not necessarily increase $\dim \partial^{=k} f$. For instance, the polynomial $f = \prod_{i=1}^{d} \sum_{j=1}^{q} x_{ij}$ has q^d monomials but $\dim \partial^{=k} f$ remains equal to $\binom{d}{k}$ for any q.

▶ **Corollary 2.** *For a multilinear homogeneous polynomial of degree d with s monomials we have $\binom{d}{k} \leq \dim \partial^{=k} f \leq s\binom{d}{k}$ for every k.*

Proof. The upper bound follows from the linearity of derivatives, and the lower bound from Theorem 1. ◀

3 The trace method

The lower bound on $\dim \partial^{=k} f$ in Theorem 1 takes a single monomial of f into account. In this section we give a more "global" result which takes all monomials into account. We will in fact lower bound the dimension of a subspace of $\partial^{=k} f$, spanned by partial derivatives of the form $\partial_I f$ where $I \in \{0,1\}^n$. In other words, we will differentiate at most once with respect to any variable.[1] We can of course view I as a subset of $[n]$ rather than as a vector in $\{0,1\}^n$.

We form a matrix M of partial derivatives as in the proof of Theorem 1. The rows of M are indexed by subsets of $[n]$ of size k, and row I contains the expansion of $\partial_I f$ in the basis (x^J). If $f = \sum_j a_J x^J$, we have seen in Section 2 that $M_{I,J} = a_{I+J}$. In order to lower bound the rank of M, we will apply the following lemma to the symmetric matrix $B = M^T.M$.

▶ **Lemma 3.** *For any real symmetric matrix $B \neq 0$ we have*

$$\mathrm{rank}(B) \geq \frac{(\mathrm{Tr} B)^2}{\mathrm{Tr}(B^2)}.$$

Lemma 3 is easily obtained by applying the Cauchy-Schwarz inequality to the vector of nonzero eingenvalues of B. Note that $B = M^T.M$ has same rank as M since we have: $x^T B x = 0 \Leftrightarrow Mx = 0$ for any vector x.

We first consider the case of polynomials with 0/1 coefficients, for which we have the following lower bound.

▶ **Theorem 4.** *For f a real polynomial with 0/1 coefficients we have*

$$\dim \partial^{=k} f \geq \frac{\sum_{P \in \mathcal{M}} \binom{\sup(P)}{k}}{|\mathcal{M}|^2} \tag{3}$$

where $\mathcal{M}$ denotes the set of monomials occuring in f, and $\sup(P)$ the number of distinct variables occuring in monomial P.

The right-hand side of (3) is sandwiched between $\binom{\mathrm{supmin}}{k}/|\mathcal{M}|$ and $\binom{\mathrm{supmax}}{k}/|\mathcal{M}|$, where supmin and supmax denote respectively the minimum and maximum number of variables occuring in a monomial of f. Theorem 1 provides a better lower bound (by a factor of $|\mathcal{M}|$)

[1] One could lift this restriction and derive similar results for the "full" matrix of k-th order derivatives, i.e., for the case where several differentiations with respect to the same variable are allowed. This would have the effect of replacing the binomial coefficients $\binom{\sup(P)}{k}$ in the lower bounds of the present section by $\dim \partial^{=k} P$. Here P denotes a monomial of f; we have explained at the beginning of Section 2 how to compute $\dim \partial^{=k} P$. We will stick here to a single differentiation for the sake of notational simplicity.

when all the monomials of f have supports of same size. Theorem 4 becomes interesting when all but a few monomials in f have large support. Indeed, the presence of a few monomials of small support can ruin the lower bound of Theorem 1.

▶ **Example 5.** Let $f(x_1,\dots,x_n) = x_1.x_2.\cdots.x_n + \sum_{i=1}^n x_i^n$. The Newton polytope of f is an n-simplex whose vertices correspond to the monomials $x_1^n,\dots,x_n^n$. The point corresponding to the monomial $x_1.x_2.\cdots.x_n$ is the barycenter of this simplex, and in particular it is not a vertex of the Newton polytope. As a result, by Remark 2 the lower bound method of Theorem 1 can only show that $\dim \partial^{=k} f \geq 1$. Theorem 4 shows the better lower bound:

$$\dim \partial^{=k} f \geq \frac{\binom{n}{k}+n}{(n+1)^2}.$$

It is not hard to check by a direct calculation that for this example, the correct value of $\dim \partial^{=k} f$ is:

- 1 for $k \in \{0,n\}$;
- n for $k \in \{1,n-1\}$;
- $\binom{n}{k}+n$ for $2 \leq k \leq n-2$.

Let us now proceeed with the proof of Theorem 4. In view of Lemma 3, we need a lower bound on $\mathrm{Tr}(B)$ and an upper bound on $\mathrm{Tr}(B^2)$.

▶ **Lemma 6.** $\mathrm{Tr}(B) = \sum_{P\in\mathcal{M}} \binom{\sup(P)}{k}$.

Proof. By definition of B, $\mathrm{Tr}(B) = \sum_J (M^T.M)_{J,J} = \sum_{I,J} M_{I,J}^2$. Since $M_{I,J} \in \{0,1\}$, this is nothing but the number of nonzero entries in M. Monomial P contributes $\binom{\sup(P)}{k}$ such entries and they are all distinct. ◀

▶ **Lemma 7.** $\mathrm{Tr}(B^2) \leq |\mathcal{M}|^2 \sum_{P\in\mathcal{M}} \binom{\sup(P)}{k}$.

Proof. Since B is symmetric, $\mathrm{Tr}(B^2) = \sum_{K,L}(B_{K,L})^2$. By definition of B, $B_{K,L} = \sum_I M_{I,K}.M_{I,L}$. Therefore

$$\mathrm{Tr}(B^2) = \sum_{I,J,K,L} M_{I,K}.M_{I,L}.M_{J,K}.M_{J,L}.$$

In this formula, I,J range over row indices (subsets of $[n]$ of size k), and K,L range over column indices. Hence $\mathrm{Tr}(B^2)$ is equal to the number of quadruples (I,J,K,L) such that all 4 entries $M_{I,K}$, $M_{I,L}$, $M_{J,K}$, $M_{J,L}$ are nonzero. Let us say that a quadruple is *valid* if this condition is satisfied. A quadruple is valid if and only if the 4 coefficients a_{I+K}, a_{I+L}, a_{J+K}, a_{J+L} are nonzero. This implies that there are at most $|\mathcal{M}|^2 \sum_{P\in\mathcal{M}} \binom{\sup(P)}{k}$ valid quadruples. Let us indeed denote by P,Q,R the 3 n-tuples $I+K$, $I+L$, $J+K$. For every fixed P we have at most $\binom{\sup(P)}{k}$ choices for I since I is contained in the support of P, and at most $|\mathcal{M}|^2$ choices for the pair (Q,R). The result follows since the quadruple (I,J,K,L) is completely determined by the choices of P,Q,R and I: we must have $K=P-I$, $L=Q-I$, $J=R-K$. ◀

Theorem 4 follows immediately from Lemmas 3 to 7.

3.1 Extension to real coefficients

In this section we generalize Theorem 4 to polynomials with real coefficients. Theorem 8 could itself be generalized to polynomials with complex coefficients by working with a Hermitian matrix in Lemma 3 rather than with a symmetric matrix.

▶ **Theorem 8.** *For any real polynomial f we have*

$$\dim \partial^{=k} f \geq \frac{\sum_{P \in \mathcal{M}} \binom{\sup(P)}{k} a_P^2}{|\mathcal{M}| \sum_{P \in \mathcal{M}} a_P^2} \tag{4}$$

where $\mathcal{M}$ denotes the set of monomials occuring in f.

To make sense of the lower bound in this theorem, it is helpful to look at a couple of special cases. If the coefficients a_P all have the same absolute value, e.g., $|a_P| = 1$ for all $P \in \mathcal{M}$, the right-hand side of (4) reduces to $\sum_{P \in \mathcal{M}} \binom{\sup(P)}{k} / |\mathcal{M}|^2$. This is exactly the lower bound in Theorem 4 (but now the coefficients of f may be in $\{-1, 0, 1\}$ rather than $\{0, 1\}$).

Our lower bound becomes weaker when the vectors $(\binom{\sup(P)}{k})_{P \in \mathcal{M}}$ and $(a_P^2)_{P \in \mathcal{M}}$ are approximately orthogonal. This can happen when the monomials with large support have small coefficients. In this case, as should be expected, Lemma 3 is effectively unable to detect the presence of monomials of large support. A probabilistic analysis shows that this bad behavior is atypical. Consider for instance the following semirandom model: we first choose a set $\mathcal{M}$ of monomials in some arbitrary (worst case) way, and then the a_P are drawn independently at random from some common probability distribution such that $\Pr[a_P = 0] = 0$.

▶ **Corollary 9.** *Let $L(f) = \dfrac{\sum_{P \in \mathcal{M}} \binom{\sup(P)}{k} a_P^2}{|\mathcal{M}| \sum_{P \in \mathcal{M}} a_P^2}$ be the lower bound on the right-hand side of (4).*

In the semirandom model described above, the expectation of $L(f)$ is:

$$E[L(f)] = \sum_{P \in \mathcal{M}} \binom{\sup(P)}{k} / |\mathcal{M}|^2.$$

We omit the (simple) proof due to lack of space. Note that we obtain for $E[L(f)]$ the lower bound from the case where $|a_P| = 1$ for all $P \in \mathcal{M}$.

The remainder of this section is devoted to the proof of Theorem 8. We follow the proof of Theorem 4. In particular, we still differentiate at most once with respect to each variable, we define the same matrix M of partial derivatives and the symmetric matrix $B = M^T.M$. We again have $\dim \partial^{=k} f \geq \mathrm{rank}(M) = \mathrm{rank}(B)$; hence Theorem 8 follows from Lemma 3 and from the next two lemmas.

▶ **Lemma 10.** $\mathrm{Tr}(B) = \sum_{P \in \mathcal{M}} \binom{\sup(P)}{k} a_P^2$.

Proof. By the proof of Lemma 6, $\mathrm{Tr}(B)$ is equal to the sum of squared entries of M; and we have $M_{I,P-I} = a_P$ for each set I of size k contained in the support of P. ◀

▶ **Lemma 11.** $\mathrm{Tr}(B^2) \leq |\mathcal{M}| \mathrm{Tr}(B) \left(\sum_{R \in \mathcal{M}} a_R^2 \right)$.

Proof. By the proof of Lemma 7,

$$\mathrm{Tr}(B^2) = \sum_{(I,J,K,L)} a_{I+K}.a_{I+L}.a_{J+K}.a_{J+L}. \tag{5}$$

Here I, J range over row indices of M while K, L range over column indices. As in the proof of Lemma 7, we call such a quadruple "valid" if the coefficients $a_{I+K}, a_{I+L}, a_{J+K}, a_{J+L}$ are all nonzero. Let $\mathcal{V}$ be the set of valid quadruples. Trying to mimic the proof of Lemma 7, we will write

$$\mathrm{Tr}(B^2) = \sum_{(P,Q,R,I)\in\mathcal{U}} a_P.a_Q.a_R.a_{Q+R-P} \tag{6}$$

where $\mathcal{U}$ is the set of quadruples (P, Q, R, I) such that:

1. P, Q, R and $Q + R - P$ belong to $\mathcal{M}$ (the set of monomials of f) and I is a row index of M, i.e., $I \in \{0,1\}^n$ and $|I| = k$.
2. $I \leq P$ and $I \leq Q$.
3. There exists a (unique) row index J such that $P - I = R - J$.

Equation (6) follows from (5) due to the following one-to-one correspondence between quadruples of $\mathcal{U}$ and $\mathcal{V}$:

(i) Given a quadruple $(I, J, K, L) \in \mathcal{V}$, set $P = I + K$, $Q = I + L$, $R = J + K$. The quadruple (P, Q, R, I) is in $\mathcal{U}$ since $Q + R - P = J + L$ and $P - I = R - J = K$.

(ii) A quadruple $(P, Q, R, I) \in \mathcal{U}$ has a unique preimage $(I, J, K, L) \in \mathcal{V}$, which is obtained as follows. A preimage must satisfy $K = P - I$, $L = Q - I$, $J = R - K$. This defines a quadruple (I, J, K, L) such that $P - I = R - J$, so J must be a row index by condition 3 in the definition of $\mathcal{U}$: $J \in \{0,1\}^n$ and $|J| = k$. It follows that $(I, J, K, L) \in \mathcal{V}$ and that this quadruple is indeed a preimage of (P, Q, R, I).

Since $2a_P.a_Q.a_R.a_{Q+R-P} \leq (a_P.a_R)^2 + (a_Q.a_{Q+R-P})^2$, it follows from (6) that

$$2\mathrm{Tr}(B^2) \leq \sum_{(P,Q,R,I)\in\mathcal{U}} a_P^2.a_R^2 + \sum_{(P,Q,R,I)\in\mathcal{U}} a_Q^2.a_{Q+R-P}^2. \tag{7}$$

The first sum is upper bounded by

$$|\mathcal{M}|.\left(\sum_{P\in\mathcal{M}} \binom{\sup(P)}{k} a_P^2\right).\left(\sum_{R\in\mathcal{M}} a_R^2\right) \tag{8}$$

since there are at most $|\mathcal{M}|$ choices for Q and we must have $I \leq P$ for a quadruple in $\mathcal{U}$. This is equal to

$$|\mathcal{M}|\mathrm{Tr}(B)\left(\sum_{R\in\mathcal{M}} a_R^2\right)$$

by Lemma 10. Likewise, the second sum in (7) is upper bounded by

$$\sum_{P,Q,R\in\mathcal{M}} \binom{\sup(Q)}{k} a_Q^2 a_{Q+R-P}^2$$

since we have $I \leq Q$ for a quadruple in $\mathcal{U}$. For any fixed $Q \in \mathcal{M}$, $\sum_{P,R\in\mathcal{M}} a_{Q+R-P}^2 \leq |\mathcal{M}|.\sum_{S\in M} a_S^2$ since each term a_S^2 on the right-hand side can appear at most $|\mathcal{M}|$ times on the left-hand side. We conclude that the second sum in (7) admits the same upper bound (8) as the first sum, and the lemma is proved. ◀

3.2 Polynomial-time computable lower bounds

The lower bound

$$L(f) = \frac{\sum_{P\in\mathcal{M}} \binom{\sup(P)}{k} a_P^2}{|\mathcal{M}| \sum_{P\in\mathcal{M}} a_P^2}$$

in Theorem 8 is clearly computable in polynomial time from f and k. Recall that we have obtained this lower bound by constructing a symmetric matrix B such that $\dim \partial^{=k} f \geq \mathrm{Tr}(B)^2/\mathrm{Tr}(B^2) \geq L(f)$. The quantity $\mathrm{Tr}(B)^2/\mathrm{Tr}(B^2)$ is therefore a better lower bound on $\dim \partial^{=k} f$ than $L(f)$. Like $L(f)$, it turns out to be computable in polynomial time. This is not self-evident because B may be of exponential size (which is the source of the $\sharp$P-hardness result in the next section).

▶ **Theorem 12.** *There is an algorithm which, given f and k, computes the lower bound $\mathrm{Tr}(B)^2/\mathrm{Tr}(B^2)$ on $\dim \partial^{=k} f$ in polynomial time (as in the rest of this paper, we assume that f is given as the sum of its nonzero monomials).*

Proof. We build on the proof of Theorem 8. Lemma 10 shows that $\mathrm{Tr}(B)$ can be computed in polynomial time, so it remains to do the same for $\mathrm{Tr}(B^2)$. In the proof of Lemma 11, we have defined a set of quadruples $\mathcal{U}$ such that

$$\mathrm{Tr}(B^2) = \sum_{(P,Q,R,I)\in\mathcal{U}} a_P.a_Q.a_R.a_{Q+R-P}.$$

This can be rewritten as:

$$\mathrm{Tr}(B^2) = \sum_{(P,Q,R)\in\mathcal{M}} N(P,Q,R).a_P.a_Q.a_R.a_{Q+R-P}$$

where we denote by $N(P,Q,R)$ the number of row indices I such that $(P,Q,R,I)\in\mathcal{U}$. It therefore remains to show that $N(P,Q,R)$ can be computed in polynomial time. Toward this goal we make two observations.

(i) Condition 2 in the definition of $\mathcal{U}$ means that $I \leq \min(P,Q)$, where the n-tuple $\min(P,Q)$ is the coordinatewise minimum of P and Q.

(ii) The equality $P - I = R - J$ in condition 3 is equivalent to $P - R = I - J$, hence $P - R \in \{-1,0,1\}^n$ since $I, J \in \{0,1\}^n$. Moreover, since I and J each have k nonzero coordinates, $P - R$ must contain the same number of 1's and -1's. By observation (i), the positions of 1's must be positive in $\min(P,Q)$.

We can therefore compute $N(P,Q,R)$ as follows.

1. If $Q + R - P$ is not a monomial of f, $N(P,Q,R) = 0$.
2. If $P - R$ is not in $\{-1,0,1\}^n$, $N(P,Q,R) = 0$.
3. If $P - R$ does not contain the same number of 1's and -1's, $N(P,Q,R) = 0$.
4. If some of the positions of 1's in $P - R$ contain a 0 in $\min(P,Q)$, $N(P,Q,R) = 0$.
5. Let $\mathrm{ones}(P,R)$ be the number of 1's in $P - R$ and $\mathrm{zeros}(P,Q,R)$ the number of 0's in $P - R$ such that we have a positive entry in $\min(P,Q)$ at the same position. Then $N(P,Q,R) = \binom{\mathrm{zeros}(P,Q,R)}{k-\mathrm{ones}(P,R)}$: from the relation $P - R = I - J$, the positions with a 1 in $P - R$ must contain a 1 in I. So it remains to choose the remaining $k - \mathrm{ones}(P,R)$ nonzero positions of I. We can only choose them among the positions that are positive in $\min(P,Q)$ (by (i)) and contain a 0 in $P - R$. ◀

3.3 Elementary symmetric polynomials

A natural question is whether the inequality $\mathrm{rank}(B) \geq \frac{(\mathrm{Tr} B)^2}{\mathrm{Tr}(B^2)}$ in Lemma 3 is tight when $B = M^T.M$ and M comes from a partial derivatives matrix. It is well known that this inequality is in general far from tight, since it is obtained by means of the Cauchy-Schwarz inequality. However in our case, due to the particular shape of the matrix B, it is not a priori clear whether a large gap can exist between $\mathrm{rank}(B)$ and $\frac{(\mathrm{Tr} B)^2}{\mathrm{Tr}(B^2)}$. In the following, we show that arbitrarily large gaps can indeed be achieved. Our source of examples are the elementary symmetric polynomials $Sym_{d,n}(x_1, \ldots, x_n) = \sum_{|I|=d} x^I$.

Here the vector I of exponents belongs to $\{0,1\}^n$, and x^I denotes as usual the multilinear monomial $x_1^{i_1} \cdot \cdots \cdot x_n^{i_n}$.

More precisely, we will show the following.

▶ **Proposition 13.** *For any fixed positive integers $d, k < d$, the family of polynomials $f_n = Sym_{d,n}$ has the following property: if we consider $u_n = \operatorname{rank}(B_n) = \dim \partial^{=k} f_n$ the sequence of dimensions of partial derivatives, and $v_n = \frac{(\operatorname{Tr} B_n)^2}{\operatorname{Tr}(B_n^2)}$ the sequence of lower bounds for the dimension, we have that $v_n \to 1$, whereas B_n is of full rank and, hence, $u_n \to +\infty$.*

Note that since d is fixed, the polynomial f_n is sparse: it contains only $n^{O(1)}$ monomials.

Proof. The matrix M of partial derivatives of f_n has only 0/1-coefficients and the coefficient $M_{I,J}$ is non-zero iff $I \cap J = \emptyset$ (with $|I| = k, |J| = d - k$). This matrix is commonly known as the *disjointness matrix* and has proved useful in communication complexity [12] and of course in algebraic complexity [14] for the study of elementary symmetric polynomials.[2] In particular, by [6], we have that M is of full rank, i.e., that $u_n = \dim \partial^{=k} f_n = \min\{\binom{n}{k}, \binom{n}{d-k}\}$. This directly implies that $u_n \to +\infty$.
We already know that $v_n \geq 1$ for all n, so we only need to compute an upper bound on v_n that tends to 1 to obtain $v_n \to 1$. To do so, we first compute the coefficients of the matrix $B = M^T.M$:

$$B_{I,J} = \sum_{|K|=d-k} M_{I,K} M_{J,K} = \sum_{\substack{|K|=d-k \\ K\cap I = K \cap J = \emptyset}} 1 = \binom{n - |I \cup J|}{d-k}$$

Notice that the value of a diagonal entry $B_{I,I} = \binom{n-k}{d-k}$ is independent of I, hence we can easily compute the trace of B: $\operatorname{Tr}(B) = \binom{n-k}{d-k}\binom{n}{k}$. A diagonal entry of B^2 is of the form $(B^2)_{I,I} = \sum_{|J|=k} (B_{I,J})^2$. In order to obtain an upper bound on v_n, it is enough to lower bound $\operatorname{Tr}(B^2)$. Since all the terms are non-negative, we will consider the following subsum:

$$(B^2)_{I,I} \geq \sum_{\substack{|J|=k \\ I \cap J = \emptyset}} (B_{I,J})^2 = \sum_{\substack{|J|=k \\ I \cap J = \emptyset}} \binom{n-2k}{d-k}^2 = \binom{n-2k}{d-k}^2 \binom{n-k}{k}.$$

Hence $Tr(B^2) = \sum_{|I|=k} (B^2)_{I,I} \geq \binom{n-2k}{d-k}^2 \binom{n-k}{k}\binom{n}{k}$. Finally, we obtain the following upper bound

$$v_n \leq \frac{\binom{n-k}{d-k}^2 \binom{n}{k}^2}{\binom{n-2k}{d-k}^2 \binom{n-k}{k}\binom{n}{k}} \xrightarrow[n\to\infty]{} 1 \tag{9}$$

◀

This proves that for constant k, d, the gap can be as large as we want, but one can ask whether such large gaps can also be achieved when k and d are increasing functions of n. Let us consider the case where k and d are proportional to n, i.e., $k = \alpha n$ and $d = \beta n$ for some constants $\alpha, \beta < 1$. Now, it is no longer true that $v_n \to 1$. For example, for $\alpha = 0.2$ and

[2] Variations on this matrix also proved useful for the analysis of the *shifted* partial derivatives of symmetric polynomials [5].

$\beta = 0.4$ we have that $v_n \to \infty$. However, we can still prove that $\frac{v_n}{u_n} \to 0$ for certain values of α and β. In the following proposition, to make sure that $k = \alpha n$ and $d = \beta n$ are always integers we set $k = k'm, d = d'm$ and $n = n'm$ where m is a new parameter and k', d', n' are constants (so $\alpha = k'/n'$ and $\beta = d'/n'$).

▶ **Proposition 14.** *For any positive integers k', d', n' such that $k' < d' < n'/2$, the family of polynomials $f_m = Sym_{d'm,n'm}$ has the following property: if we consider $u_m = \dim \partial^{=k'm} f_m$ and $v_m = \frac{(\mathrm{Tr} B_m)^2}{\mathrm{Tr}(B_m^2)}$, we have $\frac{v_m}{u_m} \to 0$.*

The proof is omitted due to lack of space.

4 $\sharp$P-hardness result for the space of partial derivatives

In this section it is convenient to work with the space $\partial^+ f$ spanned by partial derivatives of f of order r where $1 \leq r \leq \deg(f) - 1$. We will work with homogeneous polynomials, and for those polynomials we have $\dim \partial^+ f = \dim \partial^* f - 2$.

▶ **Theorem 15.** *It is $\sharp$P-hard to compute $\dim \partial^* f$ for an input polynomial f given in expanded form (i.e., written as a sum of monomials). This result remains true for multilinear homogeneous polynomials with coefficients in $\{0, 1\}$.*

We proceed by reduction from the problem of counting the number of independent sets in a graph, and use as an intermediate step a problem of topological origin. Recall that an (abstract) simplicial complex is a family Δ of subsets of a finite set S such that for every F in Δ, all the nonempty subsets of F are also in Δ. The elements of Δ are also called *faces* of the simplicial complex. We denote by $|\Delta|$ the number of faces of Δ, and more generally by $|X|$ the cardinality of any finite set X. The *dimension* of a face $X \in \Delta$ is $|X| - 1$. The dimension of Δ is the maximal dimension of its faces. If every face of Δ belongs to a face of dimension $\dim(\Delta)$, the simplicial complex is said to be *pure*.

The simplicial complex generated by a family $F_1, \ldots, F_m$ of subsets of S is the smallest simplicial complex containing all of the F_i as faces. This is simply the family of nomempty subsets $Y \subseteq S$ such that $Y \subseteq F_i$ for some i.

▶ **Theorem 16.** *The following problem is $\sharp$P-complete: given a family $F_1, \ldots, F_m$ of subsets of $[n] = \{1, \ldots, n\}$, compute the number of faces of the simplical complex Δ that it generates.*

This result remains true if Δ is pure, i.e., if $F_1, \ldots, F_m$ have the same cardinality.

We deduce Theorem 15 from Theorem 16. Let Δ be the pure simplicial complex generated by a family $F_1, \ldots, F_m$ of subsets of $[n]$, with $|F_i| = d$ for all i. We associate to each F_i the monomial $m_i = \prod_{j \in F_i} X_j$, and to Δ the polynomial $f(X_1, \ldots, X_n, Y_1, \ldots, Y_m) = \sum_{i=1}^m Y_i.m_i(X_1, \ldots, X_n)$. This is a multilinear homogeneous polynomial of degree $d+1$ in $m+n$ variables. Theorem 15 is an immediate consequence of Theorem 16 and of the following lemma.

▶ **Lemma 17.** *A basis of the linear space spanned by $\partial^+ f$ consists of the following set of $2|\Delta|$ polynomials:*

(i) *The $|\Delta|$ monomials of the form $\prod_{j \in F} X_j$, where F is a face of Δ.*

(ii) *The $|\Delta|$ polynomials of the form $\partial f / \partial F$, where F is a face of Δ (we denote by $\partial f / \partial F$ the polynomial obtained from f by differentiating with respect to all variables X_j with $j \in F$).*

In particular, $\dim(\partial^+ f) = 2|\Delta|$.

Proof. We first note that a polynomial (ii) belongs to $\partial^+ f$ by definition. A polynomial in (i) also belongs to $\partial^+ f$ since it can be obtained by picking a maximal face F_i containing F, differentiating with respect to Y_i, and then with respect to all variables X_j where $j \in F_i \setminus F$.

Conversely, any partial derivative which is not identically 0 is of the form (i) if we have differentiated f with respect to exactly one Y_i, or of the form (ii) if we have not differentiated f with respect to any of the variables Y_i. It therefore remains to show that the polynomials in our purported basis are linearly independent.

The monomials in (i) are linearly independent since they are pairwise distinct. To show that the polynomials in (ii) are linearly independent, consider a linear combination

$$g = \sum_{j=1}^{|\Delta|} \alpha_j \frac{\partial f}{\partial G_j},$$

where $G_1, \ldots, G_{|\Delta|}$ are the faces of Δ. By construction of f,

$$g = \sum_{i=1}^{m} Y_i . \left(\sum_{j=1}^{|\Delta|} \alpha_j \frac{\partial m_i}{\partial G_j} \right). \tag{10}$$

Assume that some coefficient α_j, for instance α_1, is different from 0. The face G_1 belongs to some maximal face of Δ, for instance to F_1. We claim that $\sum_{j=1}^{|\Delta|} \alpha_j \frac{\partial m_1}{\partial G_j} \neq 0$. Indeed, the faces G_j which are not included in F_1 contribute nothing to this sum, and the faces that are included in F_1 contribute pairwise distinct monomials. It follows from (10) that $g \neq 0$, and that the polynomials in (ii) are indeed linearly independent.

To complete the proof of the lemma, it remains to note that the spaces spanned by (i) and (ii) are in direct sum. Indeed, the first space is included in $\mathbb{Q}[X_1, \ldots, X_n]$ while the second is included in $\sum_{i=1}^{m} Y_i \mathbb{Q}[X_1, \ldots, X_n]$. ◀

4.1 Proof of Theorem 16

Let $G = (V, E)$ be a graph with vertex set $V = [n]$ and $m = |E|$ edges. We associate to G the simplicial complex Δ generated by the complements of the edges of G, i.e., by the sets $V \setminus \{u, v\}$ where $uv \in E$. This is a pure simplicial complex of dimension $n - 2$. The faces of Δ are the complements of the dependent sets of G. Hence $|\Delta| = 2^n - \mathrm{Ind}(G)$, where $\mathrm{Ind}(G)$ denotes the number of independents sets in G. Computing the number of independent sets of a graph is a well-known $\sharp$P-complete problem. It was shown to be $\sharp$P-complete even for bipartite graphs [15], for planar bipartite graphs of degree at most four [19] and for 3-regular graphs [7]. It follows that computing $|\Delta|$ is $\sharp$P-hard, and membership in $\sharp$P is immediate from the definition. This completes the proof of Theorem 16.

We note that it is easy to shortcircuit Theorem 16 and construct the polynomial f in the proof of Theorem 15 directly from G: we have

$$f = \sum_{uv \in E} Y_{uv} . \prod_{w \notin \{u,v\}} X_w$$

and $\dim \partial^+ f = 2(2^n - \mathrm{Ind}(G))$. Since the maximal faces of Δ have $n - 2$ elements, we have the following refinement of Theorem 15.

▶ **Corollary 18.** *It is $\sharp$P-hard to compute $\dim \partial^* f$ for a multilinear homogenous polynomial f of degree $n - 1$ with coefficients in $\{0, 1\}$, m monomials and $n + m$ variables with $m \leq \binom{n}{2}$.*

References

1 Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability and Computing*, 18(1-2):3–15, 2009.

2 Xi Chen, Neeraj Kayal, and Avi Wigderson. Partial derivatives in arithmetic complexity and beyond. *Found. Trends Theor. Comput. Sci.*, 6(1-2):front matter, 1–138 (2011), 2010.

3 Martin Dyer, Alan Frieze, and Mark Jerrum. On counting independent sets in sparse graphs. *SIAM J. Comput.*, 31(5):1527–1541, 2002.

4 Martin Dyer and Catherine Greenhill. On Markov chains for independent sets. *J. Algorithms*, 35(1):17–49, 2000.

5 Hervé Fournier, Nutan Limaye, Meena Mahajan, and Srikanth Srinivasan. The shifted partial derivative complexity of elementary symmetric polynomials. In *Mathematical foundations of computer science 2015. Part II*, volume 9235 of *Lecture Notes in Comput. Sci.*, pages 324–335. Springer, Heidelberg, 2015.

6 D. H. Gottlieb. A certain class of incidence matrices. *Proc. Amer. Math. Soc.*, 17:1233–1237, 1966.

7 Catherine Greenhill. The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Comput. Complexity*, 9(1):52–72, 2000.

8 Christopher J. Hillar and Lek-Heng Lim. Most tensor problems are NP-hard. *J. ACM*, 60(6):Art. 45, 39, 2013.

9 Neeraj Kayal. Affine projections of polynomials. In *STOC'12 – Proceedings of the 2012 ACM Symposium on Theory of Computing*, pages 643–661. ACM, New York, 2012.

10 Neeraj Kayal, Nutan Limaye, Saha Chiranjib, and Sudarshan Srinivasan. An exponential lower bound for homogeneous depth four arithmetic formulas. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 61–70. IEEE, 2004.

11 Neeraj Kayal and Chandan Saha. Lower bounds for depth three arithmetic circuits with small bottom fanin. In *30th Conference on Computational Complexity*, volume 33 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 158–182. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2015.

12 Eyal Kushilevitz and Noam Nisan. *Communication complexity.* Cambridge University Press, Cambridge, 1997.

13 Michael Luby and Eric Vigoda. Approximately counting up to four. In *Proceedings of the 29th Annual ACM Symposium on Theory of Computing*, pages 682–687. ACM, 1997.

14 Noam Nisan and Avi Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Comput. Complexity*, 6(3):217–234, 1996/97.

15 J. Scott Provan and Michael O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Comput.*, 12(4):777–788, 1983.

16 Bjarke Hammersholt Roune and Eduardo Sáenz-de Cabezón. Complexity and algorithms for Euler characteristic of simplicial complexes. *J. Symbolic Comput.*, 50:170–196, 2013.

17 Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. URL: `https://github.com/dasarpmar/lowerbounds-survey/releases`.

18 Amir Shpilka and Amir Yehudayoff. Arithmetic circuits: a survey of recent results and open questions. *Found. Trends Theor. Comput. Sci.*, 5(3-4):207–388, 2010.

19 Salil P. Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31(2):398–427 (electronic), 2001.

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