"Resistant" Polynomials and Stronger Lower Bounds for Depth-Three Arithmetical Formulas

Maurice J. Jansen and Kenneth W.Regan*
University at Buffalo (SUNY)

Abstract. We derive quadratic lower bounds on the $*$-complexity of sum-of-products-of-sums ($\Sigma\Pi\Sigma$) formulas for classes of polynomials f that have too few partial derivatives for the techniques of Shpilka and Wigderson [10, 9]. This involves a notion of "resistance" which connotes full-degree behavior of f under any projection to an affine space of sufficiently high dimension. They also show stronger lower bounds over the reals than the complex numbers or over arbitrary fields. Separately, by applying a special form of the Baur-Strassen Derivative Lemma tailored to $\Sigma\Pi\Sigma$ formulas, we obtain sharper bounds on $+, *$-complexity than those shown for $*$-complexity by Shpilka and Wigderson [10], most notably for the lowest-degree cases of the polynomials they consider.

Keywords. Computational complexity, arithmetical circuits, lower bounds, constant depth formulas, partial derivatives.

1 Introduction

In contrast to the presence of exponential size lower bounds on constant-depth Boolean circuits for majority and related functions [3, 13, 7], and depth-3 arithmetical circuits over finite fields [5, 6], Shpilka and Wigderson [10] observed that *over fields of characteristic zero* (which are infinite), super-quadratic lower bounds are not known even for constant-depth *formulas*. Indeed they are unknown for unbounded fan-in, depth 3 formulas that are sums of products of affine linear functions, which they call $\Sigma\Pi\Sigma$ formulas. These formulas have notable *upper-bound power* because they can carry out forms of Lagrange interpolation. As they ascribed to M. Ben-Or, $\Sigma\Pi\Sigma$ formulas can compute the elementary symmetric polynomials S_n^k (defined as the sum of all degree-k monomials in n variables, and analogous to majority and threshold-k Boolean functions) in size $O(n^2)$ independent of k. Thus $\Sigma\Pi\Sigma$ formulas present a substantial challenge for lower bounds, as well as being a nice small-scale model to study.

Shpilka and Wigderson defined the *multiplicative size* of an arithmetical (circuit or) formula ϕ to be the total fan-in to multiplication gates. We denote this by $\ell^*(\phi)$, and write $\ell(\phi)$ for the total fan-in to all gates, i.e. $+$ gates as well. The best known lower bound for general arithmetical circuits has remained for thirty years the $\Omega(n \log n)$ lower bound on ℓ^* by the "Degree Method" of Strassen [11] (see also [1, 2]). However, this comes nowhere near the exponential lower bounds conjectured by Valiant [12] for the permanent and expected by many for

* Part of this work by both authors was supported by NSF Grant CCR-9821040. Corresponding author: Dept. of CSE at UB, 201 Bell Hall, Buffalo, NY 14260-2000; (716) 645-3180 x114, fax 645-3464; regan@cse.buffalo.edu.

other NP-hard arithmetical functions. For polynomials f of total degree $n^{O(1)}$, the method is not even capable of $\Omega(n^{1+\epsilon})$ circuit lower bounds, not for any $\epsilon > 0$. Hence it is notable that [10] achieved better lower bounds on $\ell_3^*(f)$, where the subscript-3 refers to $\Sigma\Pi\Sigma$ formulas. These were $\Omega(n^2)$ for $f = S_n^k$ when $k = \Theta(n)$, $n^{2-\epsilon_k}$ for S_n^k with small values of k, and $\Omega(N^2/\operatorname{polylog}(N))$ for the determinant, with $N = n^2$. However, $\Omega(n^2)$ is the best this can do for $\Sigma\Pi\Sigma$ formulas. Shpilka [9] got past this only in some further-restricted cases, and also considered a depth-2 model consisting of an arbitrary symmetric function of sums. This barrier provides another reason to study the $\Sigma\Pi\Sigma$ model, in order to understand the obstacles and what might be needed to surpass them.

The techniques in [8, 10, 9] all depend on the set of dth-order partial derivatives of f being large. This condition fails for functions such as $f(x_1, \ldots, x_n) = x_1^n + \ldots + x_n^n$, which has only n dth-order partials for any d. We refine the analysis to show the sufficiency of f behaving like a degree-r polynomial on any affine subspace A of sufficiently high dimension (for this f, $r = n$ and any affine line suffices). Our technical condition is that for every polynomial g of total degree at most $r - 1$ and every such A, *there exists* a d-th order partial of $f - g$ that is non-constant on A. This enables us to prove an absolutely sharp n^2 bound on $\ell_3^*(f)$ for this f computed over the real or rational numbers, and a lower bound of $n^2/2$ over any field of characteristic zero. Note the absence of "O, Ω" notation. We prove similar tight bounds for sums of powered monomial blocks, powers of inner-products, and functions depending on ℓ_p-norm distance from the origin, and also replicate the bounds of [10, 9] for symmetric polynomials. Even in the last case, we give an example where our simple existential condition may work deeper than the main question highlighted in [9] on the maximum dimension of subspaces A on which S_n^k vanishes.

In Section 5 we prove lower bounds on $+, *$ complexity $\ell_3(f)$ that are significantly higher (but still sub-quadratic) than those given for $\ell_3^*(f)$ in [10] when the degree r of f is small. This is done intuitively by exploiting a closed-form application of the Baur-Strassen "Derivative Lemma" [1] to $\Sigma\Pi\Sigma$ formulas, showing that f and all of its n first partial derivatives can be computed with only a constant-factor increase in ℓ and ℓ^* over $\Sigma\Pi\Sigma$ formulas for f.

2 Preliminaries

A $\Sigma\Pi\Sigma$-formula is an arithmetic formula consisting of four consecutive layers: a layer of inputs, next a layer of addition gates, then a layer of multiplication gates, and finally the output sum gate. The gates have unbounded fan-in from the previous layer (only), and individual wires may carry arbitrary constants from the underlying field. Given a $\Sigma\Pi\Sigma$-formula we can write $p = \sum_{i=1}^{s} M_i$, where $M_i = \Pi_{j=1}^{d_i} l_{i,j}$, and $l_{i,j} = c_{i,j,1}x_1 + c_{i,j,2}x_2 + \ldots + c_{i,j,n}x_n + c_{i,j,0}$. Here d_i is the in-degree of the ith multiplication gate, and $c_{i,j,k}$ is nonzero iff there is a wire from x_k to the addition gate computing $l_{i,j}$.

Let $X = (x_1, \ldots, x_n)$ be an n-tuple of variables. For any affine linear subspace $A \subset F^n$, we can always find a set of variables $B \subset X$, and affine linear forms

l_b in the variables $X \setminus B$, for each $b \in B$, such that A is the set of solutions of $\{x_b = l_b : b \in B\}$. This representation is not unique. The set B is called a *base* of A. The size $|B|$ always equals the co-dimension of A. In the following, we always assume some base B of A to be fixed. Any of our numerical "progress measures" used to prove lower bounds will not depend on the choice of a base.

Following Shpilka and Wigderson [10], for polynomial $f \in F[x_1, \ldots, x_n]$, the *restriction of f to A* is defined to be the polynomial obtained by substitution of l_b for variable x_b for each $b \in B$, and is denoted by $f_{|_A}$. For a set of polynomials W, define $W_{|_A} = \{f_{|_A} \mid f \in W\}$. For a linear form $l = c_1x_1 + \ldots + c_nx_n + c_0$, we denote $l^h = c_1x_1 + \ldots + c_nx_n$. For a set S of linear forms, $S^h = \{l^h : l \in S\}$.

3 Resistance of polynomials

We state our new definition in the weakest and simplest form that suffices for the lower bounds, although the functions in our applications all meet the stronger condition of Lemma 1 below.

Definition 1. *A polynomial f in variables $x_1, x_2, \ldots, x_n$ is* (d, r, k)-resistant *if for any polynomial $g(x_1, x_2, \ldots, x_n)$ of degree at most $r-1$, for any affine linear subspace A of co-dimension k, there exists a dth order partial derivative of $f - g$ that is non-constant on A.*

For a multiset X of size d with elements taken from $\{x_1, x_2, \ldots, x_n\}$, we will use the notation $\frac{\partial^d f}{\partial X}$ to indicate the dth-order derivative with respect to the variables in X. As our applications all have $r = \deg(f)$, we call f simply (d, k)-*resistant* in this case. Then the case $d = 0$ says that f itself has full degree on any affine A of co-dimension k, and in most cases corresponds to the non-vanishing condition in [10]. We separate our notion from [10] in applications and notably in the important case of the elementary symmetric polynomials in Section 4.4 below.

The conclusion of Definition 1 is not equivalent to saying that some $(d+1)$st-order partial of $f-g$ is non-vanishing on A, because the restriction of this partial on A need not be the same as a first-partial of the restriction of the dth-order partial to A. Moreover, (d, k)-resistance need not imply $(d-1, k)$-resistance, even for $d, k = 1$: consider $f(x, y) = xy$ and A defined by $x = 0$.

Theorem 1. *Suppose $f(x_1, x_2, \ldots, x_n)$ is (d, r, k)-resistant, then*

$$\ell_3^*(f) \geq r\frac{k+1}{d+1}.$$

Proof. Consider a $\Sigma\Pi\Sigma$-formula that computes f. Remove all multiplication gates that have degree at most $r-1$. Doing so we obtain a $\Sigma\Pi\Sigma$ formula $\mathcal{F}$ computing $f - g$, where g is some polynomial of degree at most $r-1$. Say $\mathcal{F}$ has s multiplication gates. Write: $f - g = \sum_{i=1}^{s} M_i$, where $M_i = \Pi_{j=1}^{d_i} l_{i,j}$ and $l_{i,j} = c_{i,j,1}x_1 + c_{i,j,2}x_2 + \ldots + c_{i,j,n}x_n + c_{i,j,0}$. The degree of each multiplication gate in $\mathcal{F}$ is at least r, i.e. $d_i \geq r$, for each $1 \leq i \leq s$. Now select a set S of input linear forms using the following algorithm:

$S = \emptyset$
for $i = 1$ to s **do**
  repeat $d+1$ times:
    if $(\exists j \in \{1, 2, \ldots, d_i\})$: $S^h \cup \{l_{i,j}^h\}$ is a set of independent vectors
      then $S = S \cup \{l_{i,j}\}$

Let A be the set of common zeroes of the linear forms in S. Since S^h is an independent set, A is affine linear of co-dimension $|S| \leq (d+1)s$.

We claim that if at a multiplication gate M_i we picked strictly fewer than $d+1$ linear forms, then any linear form that was not picked is constant on A. Namely, each linear form l that was not picked had l^h already in the span of S^h, for the set S built up so far. Hence we can write $l = c + l^h = c + \sum_{g \in S} c_g g^h$, for certain scalars c_g. Since each g^h is constant on A, we conclude l is constant on A. This settles the claim, and yields that for each multiplication gate either

1. $(d+1)$ input linear forms vanish on A, or
2. fewer than $(d+1)$ linear forms vanish on A, with all others constant on A.

For each multiset X of size d with elements from $\{ x_1, x_2, \ldots, x_n \}$, the dth order partial derivative $\partial^d (f-g)/\partial X$ is in the linear span of the set

$$\{ \prod_{\substack{j=1 \\ j \notin J}}^{d_i} l_{ij} : \ 1 \leq i \leq s, \ J \subseteq \{1, 2, \ldots, d_i\}, \ |J| = d \ \}$$

This follows from the sum and product rules for derivatives and the fact that a first order derivative of an individual linear form l_{ij} is a constant. Consider $1 \leq i \leq s$ and $J \subseteq \{1, 2, \ldots, d_i\}$ with $|J| = d$. If item 1. holds for the multiplication gate M_i, then $\prod_{\substack{j=1 \\ j \notin J}}^{d_i} l_{ij}$ vanishes on A, since there must be one l_{ij} that vanishes on A that was not selected, given that $|J| = d$. If item 2 holds for M_i, then this product is constant on A.

Hence, we conclude that $\partial^d (f-g)/\partial X$ is constant on A. Since f is (d, r, k)-resistant, we must have that the co-dimension of A is at least $k+1$. Hence $(d+1)s \geq k+1$. Since each gate in $\mathcal{F}$ is of degree at least r, we obtain $\ell_3^*(\mathcal{F}) \geq r\frac{k+1}{d+1}$. Since $\mathcal{F}$ was obtained by removing zero or more multiplication gates from a $\Sigma\Pi\Sigma$-formula computing f, we have proven the statement of the theorem. $\square$

To prove lower bounds on resistance, we supply the following lemma:

Lemma 1. *Over fields of characteristic zero, for any $d \leq r$, $k > 0$, and any polynomial $f(x_1, x_2, \ldots, x_n)$, if for every affine linear subspace A of co-dimension k, there exists some dth order partial derivative of f such that*

$$\deg(\left(\frac{\partial^d f}{\partial X}\right)_{|_A}) \geq r - d + 1, \qquad \textit{then } f \textit{ is } (d, r+1, k)\textit{-resistant.}$$

Proof. Assume for every affine linear subspace A of co-dimension k, there exists some dth order partial derivative derivative of f such that

$$\deg(\left(\frac{\partial^d f}{\partial X}\right)_{|A}) \geq r - d + 1.$$

Let g be an arbitrary polynomial of degree r. Then

$$\left(\frac{\partial^d f - g}{\partial X}\right)_{|A} = \left(\frac{\partial^d f}{\partial X} - \frac{\partial^d g}{\partial X}\right)_{|A} = \left(\frac{\partial^d f}{\partial X}\right)_{|A} - \left(\frac{\partial^d g}{\partial X}\right)_{|A}.$$

The term $\left(\frac{\partial^d f}{\partial X}\right)_{|A}$ has degree at least $r - d + 1$, whereas the term $\left(\frac{\partial^d g}{\partial X}\right)_{|A}$ can have degree at most $r - d$. Hence $\deg(\left(\frac{\partial^d f - g}{\partial X}\right)_{|A}) \geq r - d + 1 \geq 1$. Since over fields of characteristic zero, syntactically different polynomials define different mappings, we conclude $\frac{\partial^d f - g}{\partial X}$ must be non-constant on A. □

The main difference between Lemma 1 and the original Definition 1 appears to be the order of quantifying the polynomial "g" of degree $r - 1$ out front in the former, whereas analogous considerations in the lemma universally quantify it later (making a stronger condition). We have not found a neat way to exploit this difference in any prominent application, however.

4 Applications

4.1 Sum of nth powers polynomial

Consider $f = \sum_{i=1}^{n} x_i^n$. By repeated squaring for each x_i^n, one obtains $\Sigma\Pi$ circuits (not formulas) of size $O(n \log n)$. All arithmetical circuits require size $\Omega(n \log n)$ for f [1]. The expression for f yields a *$\Sigma\Pi\Sigma$ formula* ϕ with n multiplication gates of degree n, with n^2 wires in the top linear layer fanning in to them. This works over any field, but makes $\ell(\phi) = \ell^*(\phi) = n^2$. We prove that this is close to optimal.

Theorem 2. *Over fields of characteristic zero, any $\Sigma\Pi\Sigma$-formula for $f = \sum_{i=1}^{n} x_i^n$ has multiplicative size at least $n^2/2$.*

Proof. By Theorem 1 it suffices to show f is $(1, n-1)$-resistant. Let g be an arbitrary polynomial of degree $n - 1$. Letting $g_1, \ldots, g_n$ denote the first order partial derivatives of g, we get that the *ith* partial derivative of $f - g$ equals $nx_i^{n-1} - g_i(x_1, \ldots, x_n)$. Note that the g_i's are of total degree at most $n - 2$.

We claim there is no affine linear subspace of dimension greater than zero on which all $\partial f/\partial x_i$ are constant. Consider an arbitrary affine line $x_i = c_i + d_i t$ parameterized by a variable t, where c_i and d_i are constants for all $i \in [n]$, and with at least one d_i nonzero. Then $\frac{\partial(f-g)}{\partial x_i}$ restricted to the line is given by $n(c_i + d_i t)^{n-1} - h_i(t)$, for some univariate polynomials $h_i(t)$ of degree $\leq n - 2$.

Since there must exist *some* i such that d_i is nonzero, we know some partial derivative restricted to the affine line is parameterized by a univariate polynomial of degree $n-1$, and thus, given that the field is of characteristic zero, is not constant for all t. □

In case the underlying field is the real numbers $\mathbf{R}$ and n is even, we can improve the above result to prove an absolutely tight n^2 lower bound.

Theorem 3. *Over the real numbers, for even n, any $\Sigma\Pi\Sigma$-formula for $f = \sum_{i=1}^{n} x_i^n$ has multiplicative size at least n^2.*

Proof. Since f is symmetric we can assume without loss of generality that A has the base representation $x_{k+1} = l_1(x_1, \ldots, x_k)$, $\ldots$, $x_n = l_{n-k}(x_1, \ldots, x_k)$. Then

$$f_{|_A} = x_1^n + \ldots x_k^n + l_1^n + \ldots + l_{n-k}^n.$$

Hence $f_{|_A}$ must include the term x_1^n, since each l_j^n has a non-negative coefficient for the term x_1^n and n is even. Thus via Lemma 1 we conclude that over the real numbers f is $(0, n-1)$-resistant. Hence, by Theorem 1 we get that $\ell_3^*(f) \geq \deg(f)\frac{n}{1} = n^2$. □

Let us note that $f = \sum_{i=1}^{n} x_i^n$ is an example of a polynomial that has few d-th partial derivatives, namely only n of them regardless of d. This renders the partial derivatives technique of Shpilka and Wigderson [10]—which we will describe and extend in the next section—not directly applicable.

4.2 Blocks of powers polynomials

Let the underlying field have characteristic zero, and suppose $n = m^2$ for some m. Consider the "m blocks of m powers" polynomial $f = \sum_{i=1}^{m} \prod_{j=(i-1)m+1}^{im} x_j^m$. The straightforward $\Sigma\Pi\Sigma$-formula for f, that computes each term/block using a multiplication gate of degree n, is of multiplicative size $n^{3/2}$. We will show this is tight.

Proposition 1. *The blocks of powers polynomial f is $(0, m-1)$-resistant.*

Proof. Consider an affine linear space of co-dimension $m-1$. For any base B of A, restriction to A consists of substitution of the $m-1$ variables in B by linear forms in the remaining variables X/B. This means there is at least one term/block $B_i := \prod_{j=(i-1)m+1}^{im} x_j^m$ of f whose variables are disjoint from B. This block B_i remains the same under restriction to A. Also, for every other term/block there is at least one variable that is not assigned to. As a consequence, B_i cannot be canceled against terms resulting from restriction to A of other blocks. Hence $\deg(f_{|_A}) = \deg(f)$. Hence by Lemma 1 we have that f is $(0, m-1)$-resistant. □

Corollary 1. *For the blocks of powers polynomial f, $\ell_3^*(f) \geq nm = n^{3/2}$.*

Alternatively, one can observe that by substitution of a variable y_i for each variable appearing in the ith block one obtains from a $\Sigma\Pi\Sigma$-formula $\mathcal{F}$ for f a formula for $f' = \sum_{i=1}^{m} y_i^n$ of the same size as $\mathcal{F}$. Theorem 2 generalizes to show that $\ell_3^*(f') \geq \frac{1}{2}n^{3/2}$, which implies $\ell_3^*(f) \geq \frac{1}{2}n^{3/2}$.

4.3 Polynomials depending on distance to the origin

Over the real numbers, $d_2(x) = x_1^2 + x_2^2 + \cdots + x_n^2$ is the square of the Euclidean distance of the point x to the origin. Polynomials f of the form $q(d_2(x))$ where q is a single-variable polynomial can be readily seen to have high resistance. Only the leading term of q matters. For example, consider $f = (x_1^2 + x_2^2 + \cdots + x_n^2)^m$. On any affine line L in $\mathbf{R}^n$, $\deg(f_{|_L}) = 2m$. Therefore, by Lemma 1, over the reals, f is $(0, n-1)$-resistant. Hence by Theorem 1 we get that

Proposition 2. *Over the real numbers,* $\ell_3^*((x_1^2 + x_2^2 + \cdots + x_n^2)^m) \geq 2mn$.

Observe that by reduction this means that the "mth-power of an inner product polynomial", defined by $g = (x_1y_1 + x_2y_2 + \cdots + x_ny_n)^m$, must also have $\Sigma\Pi\Sigma$-size at least $2mn$ over the reals numbers. Results for l_p norms, $p \neq 2$, are similar.

4.4 The case of symmetric polynomials

The special case of $(0, k)$-resistance is implicitly given by Shpilka [9], at least insofar as the sufficient condition of Lemma 1 is used for the special case $d = 0$ in which no derivatives are taken. For the elementary symmetric polynomial S_n^r of degree $r \geq 2$ in n variables, Theorem 4.3 of [9] implies (via Lemma 1) that S_n^r is $(0, n - \frac{n+r}{2})$-resistant. Shpilka proves for $r \geq 2$, $\ell_3^*(S_n^r) = \Omega(r(n-r))$, which can be verified using Theorem 1: $\ell_3^*(S_n^r) \geq (r+1)(n - \frac{n+r}{2}) = \Omega(r(n-r))$.

The symmetric polynomials S_n^k collectively have the "telescoping" property that every dth-order partial is (zero or) the symmetric polynomial S_{n-d}^{k-d} on an $(n-d)$-subset of the variables. Shpilka [9] devolves the analysis into the question, "What is the maximum dimension of a linear subspace of $\mathbf{C}^n$ on which S_n^r vanishes?" In Shpilka's answer, divisibility properties of r come into play as is witnessed by Theorem 5.9 of [9]. To give an example case of this theorem, one can check that S_9^2 vanishes on the 3-dimensional linear space given by

$$\{(x_1, \omega x_1, \omega^2 x_1, x_2, \omega x_2, \omega^2 x_2, x_3, \omega x_3, \omega^2 x_3) : x_1, x_2, x_3 \in \mathbf{C}\},$$

where ω can be selected to be either primitive 3rd root of unity. Let

$$\rho_0(f) = \max\{k : \text{ for any linear } A \text{ of codim. } k,\ f_{|_A} \neq 0\}.$$

Shpilka proved for $r > n/2$, that $\rho_0(S_n^r) = n - r$, and for $r \geq 2$, that $\frac{n-r}{2} < \rho_0(S_n^r) \leq n - r$. For S_9^2 we see via divisibility properties of d that the value for ρ_0 can get less than the optimum value, although the $\frac{n-r}{2}$ lower bound suffices for obtaining the above mentioned $\ell_3^*(S_n^r) = \Omega(r(n-r))$ lower bound. We have some indication from computer runs using the polynomial algebra package *Singular* [4] that the "unruly" behavior seen for ρ_0 because of divisibility properties for $r \leq n/2$ can be made to go away by considering the following notion:

$$\rho_1(f) = \max\{k : \text{ for any linear } A \text{ of codim. } k, \text{ there exists } i,\ \left(\frac{\partial f}{\partial x_i}\right)_{|_A} \neq 0\}$$

One can still see from the fact that S_n^r is homogeneous and using Lemma 1 and Theorem 1 that $\ell_3^*(S_n^r) \geq \frac{r\cdot(\rho_1(S_n^r)+1)}{2}$. Establishing the exact value of $\rho_1(S_n^r)$, which we conjecture to be $n+1-r$ at least over the rationals, seems at least to simplify obtaining the $\ell_3(S_n^r) = \Omega(d(n-d))$ lower bound. In the full version we prove that for $r \geq 2$, $\rho_1(S_{n+1}^{r+1}) \geq \rho_0(S_n^{r-1})$.

For another example, S_6^3 is made to vanish at dimension 3 not by any subspace that zeroes out 3 co-ordinates but rather by $A = \{\, (u, -u, w, -w, y, -y) : u, w, y \in \mathbf{C} \,\}$. Now add a new variable t in defining $f = S_7^4$. The notable fact is that f 1-resists the dimension-3 subspace A' obtained by adjoining $t = 0$ to the equations for A, upon existentially choosing to derive by a variable other than t, such as u. All terms of $\partial f/\partial u$ that include t vanish, leaving 10 terms in the variables v, w, x, y, z. Of these, 4 pairs cancel under the equations $x = -w$, $z = -y$, but the leftover $vwx + vyz$ part equates to $uw^2 + uy^2$, which not only doesn't cancel but also dominates any contribution from the lower-degree g. Gröbner basis runs using *Singular* imply that S_7^4 is $(1,4)$-resistant over $\mathbf{C}$ as well as the rationals and reals, though we have not yet made this a consequence of a general resistance theorem for all S_n^r.

Hence our $(1,k)$-resistance analysis for S_7^4 is not impacted by the achieved upper bound of 3 represented by A. Admittedly the symmetric polynomials f have $O(n^2)$ upper bounds on $\ell_3(f)$, so our distinction in this case does not directly help surmount the quadratic barrier. But it does show promise of making progress in our algebraic understanding of polynomials in general.

5 Bounds for +,*-Complexity

The partial derivatives technique used by Shpilka and Wigderson [10] ignores the wires of the formula present in the first layer. In the following we show how to account for them. As a result we get a sharpening of several lower bounds, though not on ℓ_3^* but on total formula size. We employ the concepts and lemmas from [10]. For $f \in F[x_1, \ldots, x_n]$, let $\partial_d(f)$ be the set of all dth order *formal* partial derivatives of f w.r.t. variables from $\{x_1, \ldots, x_n\}$. For a set of polynomials $A = \{f_1, \ldots, f_t\}$ $\mathrm{span}(A) = \{\sum_{i=1}^t c_i f_i \mid c_i \in F\}$. Write $\dim[A]$ as shorthand for $\dim[\mathrm{span}(A)]$. Note $\mathrm{span}(f_1, \ldots, f_t)_{|A} = \mathrm{span}(f_{1|A}, \ldots, f_{t|A})$, and that $\dim[W_{|A}] \leq \dim[W]$. The basic inequality from [10] then becomes:

Proposition 3. $\dim[\partial_d(c_1 f_1 + c_2 f_2)_{|A}] \leq \dim[\partial_d(f_1)_{|A}] + \dim[\partial_d(f_2)_{|A}]$.

We refine two main results in [10] $*$-complexity into results with tighter bounds but for $+,*$-complexity. In each case we compare old and new versions.

Theorem 4 ([10]). *Let $f \in F[x_1, \ldots, x_n]$. Suppose for integers d, D, κ it holds that for every affine subspace A of co-dimension κ, $\dim(\partial_d(f)_{|A}) > D$. Then $\ell_3^*(f) \geq \min(\frac{\kappa^2}{d}, \frac{D}{\binom{\kappa+d}{d}})$;*

Theorem 5 (new). *Let $f \in F[x_1, \ldots, x_n]$. Suppose for integers d, D, κ it holds that for every affine subspace A of co-dimension κ, $\sum_{i=1}^n \dim[\partial_d(\frac{\partial f}{\partial x_i})_{|A}] > D$. Then $\ell_3(f) \geq \min(\frac{\kappa^2}{d+2}, \frac{D}{\binom{\kappa+d}{d}})$.*

Proof. Consider a minimum-size $\Sigma\Pi\Sigma$-formula for f with multiplication gates $M_1, \ldots, M_s$. We have that $f = \sum_{i=1}^{s} M_i$, where for $1 \leq i \leq s$, $M_i = \Pi_{j=1}^{d_i} l_{i,j}$ and $l_{i,j} = c_{i,j,1}x_1 + c_{i,j,2}x_2 + \ldots + c_{i,j,n}x_n + c_{i,j,0}$, for certain constants $c_{i,j,k} \in F$. Computing the partial derivative of f w.r.t. variable x_k we get

$$\frac{\partial f}{\partial x_k} = \sum_{i=1}^{s} \sum_{j=1}^{d_i} c_{i,j,k} \frac{M_i}{l_{i,j}}. \tag{1}$$

Let $S = \{i : \dim[M_i^h] \geq \kappa\}$. If $|S| \geq \frac{\kappa}{d+2}$, then $\ell_3(f) \geq \frac{\kappa^2}{d+2}$. Suppose $|S| < \frac{\kappa}{d+2}$. If $S = \emptyset$, then let A be an arbitrary affine subspace of co-dimension κ. Otherwise, construct an affine space A as follows. Since $|S|(d+2) < \kappa$, and since for each $j \in S$, $\dim[M_i^h] \geq \kappa$, it is possible to pick $d+2$ input linear forms $l_{j,1}, \ldots, l_{j,d+2}$ of each multiplication gate M_j with $j \in S$, such that $\{l_{j,1}^h, \ldots, l_{j,d+2}^h | j \in S\}$ is a set of $|S|(d+2) < \kappa$ independent homogeneous linear forms. Define

$$A = \{x : \ l_{i,j}(x) = 0, \text{ for any } i \in S, \ j \in [d+2]\}.$$

We have that the co-dimension of A is at most κ. W.l.o.g. assume the co-dimension of A equals κ. For each $i \in S$, d+2 linear forms of M_i vanish on A. This implies that $\dim[\partial_d(\frac{M_i}{l_{i,j}})_{|A}] = 0$. for any $i \in S$. For any $i \notin S$, by Proposition 2.3 in [10], $\dim[\partial_d(\frac{M_i}{l_{i,j}})_{|A}] < \binom{\kappa+d}{d}$. Let $D_k = \dim[\partial_d(\frac{\partial f}{\partial x_k})_{|A}]$. By Proposition 3 and equation (1),

$$D_k \leq \sum_{i \notin S} \sum_{\substack{j \\ c_{i,j,k} \neq 0}} \dim[\partial_d(\frac{M_i}{l_{i,j}})_{|A}].$$

Hence there must be at least $\frac{D_k}{\binom{\kappa+d}{d}}$ terms on the r.h.s., i.e. there are at least that many wires from x_k to gates in the first layer. Hence in total the number of wires to the first layer is at least $\sum_{i=1}^{n} \frac{D_i}{\binom{\kappa+d}{d}} > \frac{D}{\binom{\kappa+d}{d}}$. □

Theorem 6 ([10]). *Let $f \in F[x_1, \ldots, x_n]$. Suppose for integers d, D, κ it holds that for every affine subspace A of co-dimension κ, $\dim(\partial_d(f_{|A})) > D$. Then for every $m \geq 2$, $\ell_3^*(f) \geq \min(\kappa m, \frac{D}{\binom{m}{d}})$.*

Theorem 7 (new). *Let $f \in F[x_1, \ldots, x_n]$. Suppose for integers d, D, κ with $d \geq 1$, it holds that for every affine subspace A of co-dimension κ, $\sum_{i=1}^{n} \dim[\partial_d(\frac{\partial f}{\partial x_i}_{|A})] > D$. Then for every $m \geq 2$, $\ell_3(f) \geq \min(\frac{1}{2}\kappa m, \frac{D}{\binom{m-1}{d}})$.*

The proof of Theorem 7 is analogous to above and appears in the full version.

In [10] it was proved that for $d \leq \log n$, $\ell_3^*(S_n^{2d}) = \Omega(\frac{n^{\frac{2d}{d+2}}}{d})$. Note that for $d = 2$, this lower bound is only $\Omega(n)$. We can apply Theorem 5 to prove the following stronger lower bound on the total formula size of S_n^{2d}. In particular for $d = 2$, we get an $\Omega(n^{\frac{4}{3}})$ bound.

Theorem 8. *For* $1 \leq d \leq \log n$, $\ell_3(S_n^{2d}) = \Omega(\frac{n^{\frac{2d}{d+1}}}{d})$.

Proof. For any affine subspace A of co-dimension κ and $d \geq 2$ we have that

$$\sum_{i=1}^{n} \dim[\partial_{d-1}(\frac{\partial S_n^{2d}}{\partial x_i})_{|_A}] \geq \dim[\partial_d(S_n^{2d})_{|_A}] \geq \binom{n-\kappa}{d}.$$

The latter inequality follows from Lemma 4.4 in [10]. Applying Theorem 5 we get that

$$\ell_3(S_n^{2d}) \geq \min(\frac{\kappa^2}{d+1}, \frac{\binom{n-\kappa}{d}}{\binom{\kappa+d-1}{d-1}}) = \min(\frac{\kappa^2}{d+1}, \frac{\binom{n-\kappa}{d}}{\binom{\kappa+d}{d}} \frac{\kappa+d}{d}). \tag{2}$$

Set $\kappa = \frac{1}{9} n^{\frac{d}{d+1}}$. Then we have that

$$\frac{\binom{n-\kappa}{d}}{\binom{\kappa+d}{d}} \frac{\kappa+d}{d} \geq (\frac{n-\kappa}{\kappa+d})^d \frac{\kappa+d}{d} \geq (\frac{8/9n}{2/9n^{\frac{d}{d+1}}})^d \frac{\kappa+d}{d} = 4^d n^{\frac{d}{d+1}} \frac{\kappa+d}{d} \geq \frac{4^d}{9d} n^{\frac{2d}{d+1}} \geq n^{\frac{2d}{d+1}}.$$

Hence (2) is at least $\min(\frac{n^{\frac{2d}{d+1}}}{81(d+1)}, n^{\frac{2d}{d+1}}) = \Omega(\frac{n^{\frac{2d}{d+1}}}{d})$. □

Corollary 2. $\ell_3(S_n^4) = \Omega(n^{4/3})$.

Shpilka and Wigderson defined the "product-of-inner-products" polynomial over $2d$ variable sets of size n (superscript indicate different variables, each variable has degree one) by $PIP_n^d = \prod_{i=1}^{d} \sum_{j=1}^{n} x_j^i y_j^i$.

Theorem 9. *For any constant* $d > 0$, $\ell_3(PIP_n^d) = \Omega(n^{\frac{2d}{d+1}})$.

Proof. Let $f = PIP_n^d$. Essentially we have that $\frac{\partial f}{\partial x_j^i} = y_j^i PIP_n^{d-1}$, where the PIP_n^{d-1} must be chosen on the appropriate variable set. Let A be an arbitrary affine linear subspace of co-dimension κ. Then

$$\begin{aligned}\sum_{i=1}^{d}\sum_{j=1}^{n} \dim[\partial_{d-1}(\frac{\partial f}{\partial x_j^i}|_A)] &= \sum_{i=1}^{d}\sum_{j=1}^{n} \dim[\partial_{d-1}(y_j^i PIP_n^{d-1}|_A)] \\ &\geq (dn-\kappa)\dim[\partial_{d-1}(PIP_n^{d-1}|_A)]\end{aligned}$$

The last inequality follows because at least $dn - \kappa$ of the y-variables are not assigned to with the restriction to A. From Lemma 4.9 in [10] one gets

$$\dim[\partial_{d-1}(PIP_n^{d-1}|_A) \geq n^{d-1} - 2^{2d-1}\kappa n^{d-2}.$$

Using Theorem 7 we get

$$\ell_3(f) \geq \min(\frac{\kappa^2}{2}, \frac{(dn-\kappa)(n^{d-1} - 2^{2d-1}\kappa n^{d-2})}{\binom{\kappa-1}{d-1}}).$$

Taking $\kappa = n^{\frac{d}{d+1}}$, one gets for constant d that $\ell_3(PIP_n^d) = \Omega(n^{\frac{2d}{d+1}})$. □

For comparison, in [10] one gets $\ell_3^*(PIP_n^d) = \Omega(n^{\frac{2d}{d+2}})$.

6 Conclusion

We have taken some further steps after Shpilka and Wigderson [10, 9], obtaining absolutely tight (rather than asymptotically so) multiplicative size lower bounds for some natural functions, and obtaining somewhat improved bounds on $+, *$-size for low-degree symmetric and product-of-inner-product polynomials. However, these may if anything enhance the feeling from [10, 9] that the concepts being employed may go no further than quadratic for lower bounds. One cannot after all say that a function $f(x_1, \ldots, x_n)$ is non-vanishing on an affine-linear space of co-dimension more than n. The quest then is for a mathematical invariant that scales beyond linear with the number of degree-d-or-higher multiplication gates in the formula.

Acknowledgments We thank Avi Wigderson for comments on a very early version of this work, and referees of later versions for very helpful criticism.

References

1. W. Baur and V. Strassen. The complexity of partial derivatives. *Theor. Comp. Sci.*, 22:317–330, 1982.
2. P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic Complexity Theory.* Springer Verlag, 1997.
3. M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. *Math. Sys. Thy.*, 17:13–27, 1984.
4. G.-M. Greuel, G. Pfister, and H. Schönemann. SINGULAR 3.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2005. `http://www.singular.uni-kl.de`.
5. D. Grigoriev and M. Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *Proc. 30th Annual ACM Symposium on the Theory of Computing*, pages 577–582, 1998.
6. D. Grigoriev and A.A. Razborov. Exponential lower bounds for depth 3 algebraic circuits in algebras of functions over finite fields. *Applicable Algebra in Engineering, Communication, and Computing*, 10:465–487, 2000. (preliminary version FOCS 1998).
7. J. Håstad. Almost optimal lower bounds for small-depth circuits. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 143–170. JAI Press, Greenwich, CT, USA, 1989.
8. N. Nisan and A. Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6:217–234, 1996.
9. A. Shpilka. Affine projections of symmetric polynomials. *J. Comp. Sys. Sci.*, 65:639–659, 2002.
10. A. Shpilka and A. Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. *Computational Complexity*, 10:1–27, 2001.
11. V. Strassen. Berechnung und Programm II. *Acta Informatica*, 2:64–79, 1973.
12. L. Valiant. The complexity of computing the permanent. *Theor. Comp. Sci.*, 8:189–201, 1979.
13. A. Yao. Separating the polynomial-time hierarchy by oracles. In *Proc. 26th Annual IEEE Symposium on Foundations of Computer Science*, pages 1–10, 1985.