Electronic Colloquium on Computational Complexity, Report No. 23 (1999)

Depth-3 Arithmetic Formulae over Fields of Characteristic Zero

Amir Shpilka Avi Wigderson [*]

Institute of Computer Science, Hebrew University, Jerusalem, Israel
E-mail:

Abstract

In this paper we prove near quadratic lower bounds for depth-3 arithmetic formulae over fields of characteristic zero. Such bounds are obtained for the elementary symmetric functions, the (trace of) iterated matrix multiplication, and the determinant. As corollaries we get the first nontrivial lower bounds for computing polynomials of constant degree, and a gap between the power depth-3 arithmetic formulas and depth-4 arithmetic formulas.

The main technical contribution relates the complexity of computing a polynomial in this model to the wealth of partial derivatives it has on every affine subspace of small co-dimension. Lower bounds for related models utilize an algebraic analog of Nečhiporuk lower bound on Boolean formulae.

1. Introduction

Arithmetic circuits are a very natural model for computing polynomials. Like most computational models, almost no lower bounds are known for this one. The best size lower bound known is the classical $\Omega(n \log d)$ (for some natural degree d polynomials over n variables) of [2]. No nontrivial lower bounds are known for depth. For a survey of known results see [16, 21] and the introduction to [12].

Our intuition suggests that arithmetic circuits (being more “structured”) are weaker than Boolean circuits, and thus lower bounds for the former should be easier to prove. Our experience with monotone analogs of both models certainly justifies this intuition. However, it is shattered by the simple problem of computing majority in depth-3 circuits. We know that in the Boolean model this requires exponential size. However, Ben-Or proved that the majority polynomial has simple, quadratic-size depth-3 arithmetic formulae!

In this paper we deal mainly with depth-3 arithmetic formulae. Such a formula can be viewed as a sum of products of linear functions of the variables. This is a very restricted model, but it can clearly compute any multivariate polynomial, some of which surprisingly cheaply. Being the “simplest” nontrivial model, it has received significant attention as detailed below.

Despite it’s innocence, no superlinear lower bounds are known for this model when the field is large, except of the degree lower bound of [2]. This state of affairs is in contrast with what is known for Boolean circuits with $mod\ q$ gates, and arithmetic circuits over finite fields. In the first model [13] and [15] proved exponential lower bounds e.g. for the majority function for any constant depth. In the second model (for depth-3 only) [5] and [6] recently proved exponential lower bounds for some symmetric functions.

These techniques cannot be extended to give analogous results for large fields, as by the above mentioned result of Ben-Or [3], they are simply false. One path which was taken to handle large fields was to further restrict the model [11, 12]. They consider homogeneous circuits, which for depth-3 circuits amounts to requiring a homogeneous linear function at every gate in the bottom level. In that model [12] were able to prove exponential lower bounds, even for the majority polynomial. Thus, the general model is sometimes exponentially more powerful than its homogeneous variant.

In this paper we study the general model. We prove near quadratic lower bounds for a number of natural functions, such as the elementary symmetric functions, the (trace of) iterated matrix product, and the determinant. We show in particular the following lower bounds (where n denotes the total number of variables):

[*] This research was supported by grant number 69/96 of the Israel Science Foundation, founded by the Israel Academy for Sciences and Humanities.

ISSN 1433-8092

- $\Omega(n^2)$ lower bound for the elementary symmetric function of degree $\Omega(n)$ (Thus proving that the construction in [3] is essentially optimal).
- $n^{2-\epsilon}$ lower bound for a polynomials of constant degree (No superlinear bound was known before).
- $\tilde{\Omega}(n^2)$ lower bound for the determinant (But we believe that the real bound is exponential).

The proof combines the idea of partial derivatives from [12] and the idea of approximating "high rank" multiplication gates from [5]. Some of these lower bounds, (together with upper bounds) provide near quadratic separation between depth-3 and depth-4 formulae. Of special interest is a new (depth-6) formula of size $O(nd^3 \log d)$ for the elementary symmetric polynomials of degree d (which beats the Ben-Or construction for small degrees).

We prove two general theorems that express a lower bound for the depth-3 complexity of a polynomial f, in terms of parameters k and D, where D measures the wealth of partial derivatives f has on every affine subspace of co-dimension k. These are stated and proved in Section 3. Specific lower bounds follow from proving lower bounds on D for specific polynomials and well chosen values of k. These are stated and proved in Section 4.

In Section 5 we look at different (but related) models of computation. The first is a depth-3 arithmetic formula where the bottom plus gates can take only linear functions of one variable. The reason to consider this model is that it is strong enough for the construction of [3] for the elementary symmetric polynomials. Here we can prove an exponential lower bound for the Determinant.

The second model is formulae without any depth restriction, whose inputs are allowed to be not only the variables, but also arbitrary polynomials in single variables. For this model we prove a quadratic lower bound for the Discriminant function (which is the determinant of a Vandermunde matrix). The proof uses an algebraic analog of a Nečhiporuk-like argument (see [10]).

Finally, we study the computation of polynomials of the form $t^m f$, where f is our target (homogeneous) polynomial, and t is a new indeterminate. We show general computation of such products of f is no stronger than *homogeneous* computation of f. This result has the same flavor of results on the monotone vs general computation of "slice functions". It yields an exponential gap between the depth-3 homogeneous complexity of a polynomial and the homogeneous complexity of its derivative with respect to a single variable.

2. Definitions and Tools

2.1. Arithmetic Circuits

Definition 2.1 *An* arithmetic circuit *is a labeled directed acyclic graph. The inputs (nodes of in-degree zero) are labeled from the set of variables X. A constant from F (the base field) can label an edge, which means the polynomial computed at its tail is multiplied by this constant. The internal nodes are labeled by addition or multiplication gates, computing the sum and product, resp, of the polynomials on the tails of incoming edges. (Subtraction is obtained using the constant -1.) The output is the polynomial computed at the output node. A* formula *is a circuit which all it's nodes have out-degree one (namely, a tree). We consider unbounded fan-in formulas. The* size *of a circuit is the number of gates in it, and the size of a formula is the number of it's leafs. The* depth *of the circuit is the length of the longest path between the output node and an input node.*

The main model we shall deal with in the paper is the following.

Definition 2.2 *A $\Sigma\Pi\Sigma$ formula is a leveled depth-3 formula with a plus gate at the top, multiplication gates at the middle level, and plus gates at the bottom. A homogeneous $\Sigma\Pi\Sigma$ formula is a $\Sigma\Pi\Sigma$ formula that is allowed to compute only homogeneous linear functions in the bottom level.*
For a polynomial f, we denote by $L_3(f)$ the size of the smallest $\Sigma\Pi\Sigma$ formula computing f, and by $L_3^H(f)$ the size of the smallest homogeneous $\Sigma\Pi\Sigma$ formula computing f.

From the definition it is clear that if f is computed by such formula then f has a representation of the form: $f = \sum_{j=1}^{s} M_j$ where $M_j = \prod_{i=1}^{deg(M_j)} \ell_{i,j}$, each $\ell_{i,j}$ is a linear function in the variables, and $deg(M_j)$ is the fan-in of the jth multiplication gate. We stress that this linear function may involve a constant term (and indeed without this ability the model is homogeneous). It will be useful for us to separate out the homogeneous part of such functions,

Definition 2.3 *For a linear function ℓ we let ℓ^h be the homogeneous part of ℓ, and $\ell^0 = \ell - \ell^h$ is the constant term. Let $M = \prod_{i=1}^{deg(M)} \ell_i$. We denote $M^h = \{\ell_1^h, ..., \ell_{deg(M)}^h\}$ and let $\dim(M^h)$ be the dimension of the linear span of the set M^h.*

All our lower bounds for this model will trade off the number of (plus) gates at the bottom level with the number of (multiplication) gates in the middle level. We will use the obvious

Proposition 2.1 *Every $\Sigma\Pi\Sigma$ formula with multiplication gates $\{M_1, ..., M_s\}$ has size $\Omega(\sum_{j=1}^s deg(M_j))$.*

2.2. Partial Derivatives

We now recall basic definitions and results from [12] on partial derivatives, which will be key for our lower bounds.

Let F be a field of characteristic zero. We will consider polynomials over a set of variables X.

Definition 2.4 *For any set of polynomials $V \subseteq F[X]$ we use $dim(V)$ for the dimension of the linear span of V (in other words the maximum number of linearly independent polynomials in V over F).*

Remark 1 *Observe that $dim(V)$ is invariant under any full rank linear transformation on the set of variables X.*

Definition 2.5 *[12] Let f be a polynomial and d an integer. We let $\partial_d(f)$ denote the set of partial derivatives of order d of f.*

Example 1

$$\partial_2(x^2y) = \{0, 2x, 2y\}$$

A fundamental observation of [12] is that this dimension commutes with the arithmetic operations! Here we'll use it only for addition gates.

Proposition 2.2 *[12] For every $f_1, f_2, \cdots, f_r \in F[X]$ and $\alpha \in F$, $\alpha \neq 0$ we have:*

- $dim(\partial_d(\alpha f_1)) = dim(\partial_d(f_1))$.
- $dim(\partial_d(\sum_i f_i)) \leq \sum_i dim(\partial_d(f_i))$.

For multiplication gates, we'll use a stronger bounds than those in [12], which applies only to product of linear functions.

Proposition 2.3 *For a multiplication gate M with $\dim(M^h) = m$, and for every d, $dim(\partial_d(M)) \leq \binom{m+d}{d}$.*

Proof: From the definition it is clear that M is a function of m independent linear functions (wlog - the first m), i.e $M = M(\ell_1, ..., \ell_m)$. Therefore $\frac{\partial M}{\partial x} = \sum_{i=1}^m \frac{\partial M}{\partial \ell_i} \frac{\partial \ell_i}{\partial x}$. Since ℓ_i is a linear function, $\frac{\partial \ell_i}{\partial x}$ is a scalar, so $\frac{\partial M}{\partial x}$ is a linear combination of the $\frac{\partial M}{\partial \ell_i}$-s. A similar thing happens when we look at derivatives of order d, each derivatives lies in the linear span of order d derivatives of M with respect to the ℓ_i-s. Therefore to bound $dim(\partial_d(M))$ it is sufficient to bound the number of such derivatives. Since M is a polynomial the order in which we take derivatives doesn't matter, i.e $\frac{\partial^2 M}{\partial \ell_1 \partial \ell_2} = \frac{\partial^2 M}{\partial \ell_2 \partial \ell_1}$. Therefore the number of order d partial derivatives is at most the number of ways to write d as a sum of m integers (the value of the i'th integer corresponds to the order we take derivatives of M w.r.t ℓ_i). □

Proposition 2.4 *For a multiplication gate M with $deg(M) = m$, and for every d, $dim(\partial_d(M)) \leq \binom{m}{d}$.*

Proof: Write $M = \prod_{i=1}^m \ell_i$. Since each ℓ_i is a linear function we get,

$$\partial_d(M) \subset \; span \; \{\prod_{i \in T} \ell_i \mid T \subset [m], |T| = m - d\} \; .$$

The result follows since

$$\dim(\; span \; \{\prod_{i \in T} \ell_i \mid T \subset [m], |T| = m - d\} \;) \leq \binom{m}{d}.$$

□

2.3. Restrictions to affine subspaces

A key ingredient of our lower bound technique will be to study the dimension of a set of partial derivatives not over the whole vector space F^n, but over affine subspaces of it.

Definition 2.6 *Let A be an affine subspace of F^n with the set of coordinates $(x_1, ..., x_n)$. Call $B \subset X$ a base for A if A can be represented by the set of equations:*

$$\{x_b = \ell_b \mid b \in B\}$$

where ℓ_b is a linear function on the set of variables $X \setminus B$. Note that every A has such a base B, so lets fix one such base. Define $\phi_B : F[X] \mapsto F[X \setminus B]$ to be the homomorphism which assigns to every variable x_b with $b \in B$ the linear function ℓ_b (and leaves the other variables untouched). This map is extended by mutiplicativity to monomials and then by additivity to polynomials.

Definition 2.7 *Let A be an affine subspace of F^n, for a polynomial f we denote by $f|_A$ the restriction $\phi_B(f)$ of f to A. For a set of polynomials V, $V|_A = \{f|_A : f \in V\}$.*

Definition 2.8 *Let C be an arithmetic formula, and ϕ_B as above. Then $C|_A$ is the formula obtained by applying ϕ_B to the inputs, and removing subformulae whose output becomes identically zero.*

Clearly, restrictions commute with arithmetic operations. Thus

Proposition 2.5 *Assume C computes f. Then for every affine subspace A (chosen together with a base B), $C|_A$ computes $f|_A$.*

The complexity of computing a polynomial f will be related in the next section to the dimension of one of two different sets of polynomials derived from f.

1. $(\partial_d(f))|_A$, is the set of restrictions of all order d partial derivatives on f to A.
2. $\partial_d(f|_A)$ is the set defined by first restricting f itself to A, and then taking all order d partial derivatives. To formally define it, we need an "inverse" to ϕ_B below.

Definition 2.9 *Assume that $X \setminus B = \{x_{i_1}, ..., x_{i_{n-k}}\}$. Define $\ell_B : F^{n-k} \mapsto F^n$ by*

$$\ell_{B\,(i)} = \begin{cases} l_b & i = b \in B \\ x_j & i = i_j \end{cases}$$

Definition 2.10 *Let $\partial_d(f|_A) = \partial_d(f \circ \ell_B)$.*

We conclude this Subsection with a remark on the arbitrariness of the choice of the base B of the affine subspace A. Clearly, this choice affect the sets defined above, as well as the restricted formulae, at least in the obvious sense that the restricted polynomials will be over different sets of indeterminates.

Still, the choice of B has no affect on the proofs, so from this point of view it can certainly be arbitrary. Moreover, it turns out that while the sets of polynomials defined by (1) and (2) above change with different choices of B, the dimension of each is invariant under this choice.

3. Results

3.1. General Results

The following theorems will be our main tool in deriving the lower bounds. Observe that in both theorems, the lower bound is the minimum of two functions. It will be clear from their proofs that they both actually give a trade-off: either the first function lower bounds the number of (addition) gates in the bottom level, or the second lower bounds the number of (multiplication) gates in the middle level.

Theorem 3.1 *Let f be a polynomial. Assume that for some integers d, k, D, for every affine subspace A of co-dimension k, $dim((\partial_d(f))|_A) > D$. Then*

$$L_3(f) \geq \min(\frac{k^2}{d}, \frac{D}{\binom{k+d}{d}})$$

.

Theorem 3.2 *Let f be a polynomial. Assume that for some integers d, k, D, for every affine subspace A of co-dimension k, $dim(\partial_d(f|_A)) > D$. Then for every m*

$$L_3(f) \geq \min(km, \frac{D}{\binom{m}{d}}).$$

Using these theorems will follow proofs that in specific polynomials D is large for appropriate values of d, k. We note that both theorems are incomparable – each can give a better lower bound than the other for some polynomials.

3.2. Proofs

To prove these theorems we will first need a technical lemma. It basically shows how to define an affine subspace that nullifies high rank multiplication gates.

Lemma 3.3 *Fix an integer z. Let $\mathcal{F}$ be a $\Sigma\Pi\Sigma$ formula, and let $S = \{j \mid \dim(M_j \geq k\}$. If $|S| < \frac{k}{z}$ then there is an affine subspace A of co-dimension k such that in $\mathcal{F}|_A$ all the multiplication gates are of dimension less then k. Moreover, in A all the multiplication gates in S have a zero of order $\geq z$.*

Proof: Let's assume w.l.o.g that $S = \{1, 2, ..., r\}$ for some $0 \leq r < \frac{k}{z}$. Now lets take for each $1 \leq j \leq r$, z linear functions, $\ell_{j,1}, ..., \ell_{j,z}$, from M_j, such that all the $\ell^h_{j,i}$-s are linearly independent. It's possible to make such a choice since $r < \frac{k}{z}$ and $\dim(M_j) \geq k$. Because of the independence there's an affine subspace A of co-dimension at most rz such that

$$\forall \vec{x} \in A, \;\; \ell^h_{j,i}(\vec{x}) = \ell^0_{j,i} \;\; \text{(i.e } \ell_{j,i}(\vec{x}) = 0 \text{ for the relevant } (j,i)\text{-s).}$$

So now in $\mathcal{F}|_A$ all the multiplication gates are of dimension less then k, and $M_1, ..., M_r$ each have a zero of order $\geq z$ on A. □

We can now turn to the proofs of the main theorems.

Proof of Theorem 3.1: Fix any k and d, and set $z = d + 1$. Assume that $\mathcal{F}$ is a depth-3 formula computing f. If the assumption of Lemma 3.3 does not hold, then by Proposition 2.1 we have a $\frac{k^2}{d}$ lower bound on the number of addition gates. Otherwise the assumption holds, and let A be the subspace guaranteed by the lemma.

Now we know that all the gates in $\mathcal{F}$ with dimension $\geq k$ have a zero of degree $d + 1$ when restricted to A. Thus for each such multiplication gate M, $dim(\partial_d(M)|_A) = 0$. By Proposition 2.3 all the other multiplication gates have $dim(\partial_d(M)|_A) \leq \binom{k+d}{d}$ so by Proposition 2.2 there must be at least $\frac{D}{\binom{k+d}{d}}$ multiplication gates in $\mathcal{F}$. □

Proof of Theorem 3.2: Assume that we are given d, k, m. We now look at the best $\Sigma\Pi\Sigma$ formula for f. If there are more then k multiplication gates of degree greater then m then we are done (again by Proposition 2.1) . Otherwise we can find an affine subspace A of co-dimension k so if we restrict the formula to A all the remaining gates will be of degree $\leq m$. Using Propositions 2.2, 2.4 we see that there are at least $\frac{D}{\binom{m}{d}}$ multiplication gates in the restricted formula. □

4. Lower Bounds

We prove lower bounds for natural functions such as the elementary symmetric functions, matrix multiplication and determinant. All the lower bounds will be stated in terms of parameters of the function we deal with.

4.1. Elementary symmetric functions

Here we obtain tight lower and upper bounds.

Definition 4.1

$$S_n^d(X) = \sum_{\substack{T \subset [n] \\ |T| = d}} \prod_{i \in T} x_i$$

X is the set of variables $\{x_1, ..., x_n\}$. This is the d'th elementary symmetric function.

Theorem 4.1

$$L_3(S_n^{2d}) \geq \max(\Omega(\frac{n^{\frac{2d}{d+2}}}{d}), \Omega(nd)),\ \forall d \leq 4n/9.$$

This lower bound is tight (for large values of d) in view of the following theorem:

Theorem 4.2 (Ben-Or) *For every d, $L_3(S_n^d) \leq O(n^2)$.*

In contrast, allowing depth-6 or even depth-4, the Ben-Or construction can be greatly improved for small d. With our lower bound above for $d = \log n$ or $d = constant$, it provides a near quadratic separation between depth-3 and depth-4 formulae.

Theorem 4.3 *For every d, $L_4(S_n^d) \leq O(nd2^{\sqrt{d}})$.*

Theorem 4.4 *For every d, $L_6(S_n^d) \leq O(nd^3 \log d)$.*

4.2. Product of inner products

Definition 4.2

$$PIP_n^d(X,Y) = \prod_{i=1}^{d} \sum_{j=1}^{n} x_j^i y_j^i$$

This is the product of d inner products. $X = \cup_{i=1}^d X^i$ and $Y = \cup_{i=1}^d Y^i$ with each X^i and Y^i containing n variables.

We have $2nd$ variables, and the lower bound that we show is strongest when $d \leq \log n$.

Theorem 4.5 *$L_3(PIP_n^d) \geq \Omega(n^{2\frac{d}{d+2}})$ for $d \leq \frac{1}{2}\sqrt{\log n}$.*
$L_3(PIP_n^d) \geq \Omega(n^{2-\frac{4}{\sqrt{\log n}}})$ for $d > \frac{1}{2}\sqrt{\log n}$.

Remark 2 *We showed a lower bound for $\Sigma\Pi\Sigma$ formulas calculating $PIP_n^d(X,Y)$, but there is a trivial $\Pi\Sigma\Pi$ formula for it of linear size. This is an unusual example where $\Pi\Sigma\Pi$ formula is more efficient then $\Sigma\Pi\Sigma$ formula, and provides a (near) quadratic gap between the two.*

Another separation between depth-4 formulas and depth-3 formulas is:

Theorem 4.6

$$\forall d,\ L_4(PIP_n^d(X,Y) + PIP_n^d(W,Z)) = O(nd)$$

$$for\ \ d \leq \frac{1}{2}\sqrt{\log n},$$

$$L_3(PIP_n^d(X,Y) + PIP_n^d(W,Z)) \geq \Omega(n^{2\frac{d}{d+2}})$$

4.3. (Trace of) Iterated matrix multiplication

Another function which has a similar lower bound, with a similar proof technique is the trace of matrix multiplication.

Definition 4.3

$$TR_n^{2d}(X^1, ..., X^n) = \sum_{i_1,\ldots,i_{2d}} x^1_{i_1,i_2} \cdot x^2_{i_2,i_3} \cdots x^{2d}_{i_{2d},i_1}$$

This is the trace of the multiplication of $2d$ $n \times n$ matrices (each X^i is an $n \times n$ matrix).

Here we have $2dn^2$ variables and the lower bound depends on d and n.

Theorem 4.7 $L_3(TR_n^{2d}) \geq \Omega(n^{4\frac{d}{d+2}})$ *for* $d \leq \frac{1}{2}\sqrt{\log n}$.
$L_3(TR_n^{2d}) \geq \Omega(n^{4-\frac{8}{\sqrt{\log n}}})$ *for* $d > \frac{1}{2}\sqrt{\log n}$.

4.4. Determinant

The last lower bound that we show is for the determinant function. We show an almost n^4 lower bound, where the number of variables is n^2. Note that it is a very weak lower bound compared to what is known over finite fields. Getting a superpolynomial lower bound for this function seems to be the main next challenge.

Definition 4.4

$$DET_n(X) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)}$$

Where S_n is the symmetric group on n variables.

Here X is $n \times n$ matrix.

Theorem 4.8 $L_3(DET_n(X)) \geq \Omega(n^4/\log n)$.

Denote by $L_3^m(f)$ the size of the smallest depth-3 formula computing f in which the degree of every multiplication gate is at most m.

Theorem 4.9 $L_3^m(DET_n(X)) \geq \Omega(e^{n^2/m})$.

4.5. Proofs

The lower bounds for a polynomial f are expressed in terms of the dimensions of the sets $\partial_d(f|_A)$ and $(\partial_d f)|_A$, for some arbitrary affine subspace A. These sets can be quite complicated even for very simple polynomials. To bound the dimensions, we establish two simple lemmas relating these sets to each other and to the more easily understood set, $\partial_d(f)$.

Lemma 4.10 *For every polynomial f and affine subspace A with a basis B we have*

$$dim((\partial_d f)|_A) \geq dim(\partial_d(f) \cap F[X \setminus B])$$

Proof: The polynomials in $\partial_d(f)$ which do not depend on variable from B are not effected by the restriction to A. □

Definition 4.5 *For a set of variables B, let $\partial_d^B f$ denote all order d derivatives of f, where at least one of the d derivations is with respect to a variable in B.*

Lemma 4.11 *For every polynomial f and affine subspace A with a basis B we have*

$$dim(\partial_d(f|_A)) \geq dim((\partial_d f)|_A) - dim((\partial_d^B f)|_A)$$

Proof: We will demonstrate the proof for $|B| = 1$ and $d = 1$. W.l.o.g assume that $B = \{x_1\}$ and A is given by $x_1 = \ell(x_2, ..., x_n)$. According to the chain rule we get:

$$\partial_1(f|_A) = \{\frac{\partial f(\ell, x_2, ..., x_n)}{\partial x_i} \mid 1 < i\} =$$

$$\{\frac{\partial f}{\partial x_1}(\ell, x_2, ..., x_n)\frac{\partial \ell}{\partial x_i} + \frac{\partial f}{\partial x_i}(\ell, x_2, ..., x_n) \mid 1 < i\}$$

Since $\partial_1^B f = \frac{\partial f}{\partial x_1}$, the set $(\partial_1 f)|_A$ (which is actually the set $\{\frac{\partial f}{\partial x_i}(\ell, x_2, ..., x_n) \mid 1 \leq i\}$) is spanned by the set

$$\partial_1(f|_A) \cup \{\frac{\partial f}{\partial x_1}(\ell, x_2, ..., x_n)\}.$$

The proof for larger values of $|B|, d$ is achieved using similar arguments. View the restriction to A as a composition of a polynomial with a set of linear functions. Then from the chain rule of partial derivatives, and from the fact that all derivatives of linear functions are constants, it follows that the set $(\partial_d f)|_A$ is spanned by the two sets $\partial_d(f|_A)$ and $(\partial_d^B f)|_A$. □

4.5.1 Elementary Symmetric Function

Upper Bounds:

Proof of Theorem 4.4: Let $T_n^d(X) = \sum_{i=1}^n x_i^d$. The well known Newton Identities provide a polynomial relation expressing each S_n^d in terms of $\{T_n^k : k \leq d\}$. More precisely, let the degree d truncation of the series for e^y be the polynomial $E^d(y) = \sum_{j=0}^d y^j/j!$. Then $S_n^d(X)$ is the coefficient of z^d in the polynomial

$$\Pi_{k=1}^d E^{\lfloor d/k \rfloor}((-1)^{k-1}T_n^k z^k) \qquad (1)$$

See e.g [7] Section 1.2.9 on Generating Fnctions. Clearly, this is a $\Pi\Sigma$ formula of size $d \log d$ in the variables T_n^k, each of which is a $\Pi\Sigma\Pi$ formula of size nd in the original X variables. Finally, interpolation over the variable z (as in Ben-Or's construction) requires $d + 1$ values for this variable and gives a $\Sigma\Pi\Sigma\Pi$ formula of size $nd^3 \log d$ for S_n^d. □

Proof of Theorem 4.3: The coefficient of z^d in Equation 1 is:

$$\sum_{m_1+2m_2+...+dm_d=d} \prod_{k=1}^d \frac{(-1)^{(k-1)m_k}}{m_k!} T_n^{k\,m_k}$$

There are $2^{O(\sqrt{d})}$ (see [20] Theorem 15.7) summands (these are all the partitions of d), each is a $\Pi\Sigma\Pi$ formula of size $\leq nd$, so the result follows. □

Lower Bounds:

The proof of Theorem 4.1 is split into two parts:

Theorem 4.12

$$L_3(S_n^{2d}) \geq \Omega(\frac{n^{\frac{2d}{d+2}}}{d}), \ \forall d \leq n/10.$$

Theorem 4.13 $L_3(S_n^{2d}) \geq \Omega(dn)$ *For* $8 \log n \leq d \leq 4n/9$.

It is clear that these theorems imply Theorem 4.1.

Lemma 4.14 *For every* n, k, d *and for every affine subspace* A *of co-dimension* k,

$$dim((\partial_d(S_n^{2d}))|_A) \geq \binom{n-k}{d}$$

Proof of Lemma 4.14: The set $\partial_d(S_n^{2d})$ is actually the set $\{S_{n-d}^d(X-T)\}$ (recall Definition 4.1) where T ranges over the $\binom{n}{d}$ subsets of $[n]$ of size d. By [4] (as used in [12]), this set is spanned by all degree d multilinear monomials. There are $\binom{n-k}{d}$ linear independent degree d multilinear monomials in the variables of the set $X \setminus B$, using Lemma 4.10 the result follows. □

Proof of Theorems 4.12: Plug Lemma 4.14 into Theorem 3.1, to get a lower bound of

$$L_3(S_n^{2d}) \geq \min(\frac{k^2}{d}, \frac{\binom{n-k}{d}}{\binom{k+d}{d}}) \ \forall k, d.$$

Estimate $\frac{\binom{n-k}{d}}{\binom{k+d}{d}}$ by $(\frac{n-k}{k+d})^d$ and take $k = n^{\frac{d}{d+2}}/9$. We get

$$(\frac{n-k}{k+d})^d \geq \left\{ \begin{array}{ll} (\frac{8n/9}{2n^{\frac{d}{d+2}}/9})^d & d \leq \log n \\ (\frac{8n/9}{2n/9})^{\log n} & \log n \leq d \leq n/10 \end{array} \right.$$

and the lower bound follows. □

Lemma 4.15 *For every n, k, d, and for every subspace A of co-dimension k with $k \leq \min(d, n-2d)$*

$$dim((\partial_{d-k}(S_n^{2d}))|_A) \geq \binom{n-2k}{d-k}$$

We will derive this lemma from the following one:

Lemma 4.16 *For every affine subspace A of co-dimension $k \leq \min(d, n-2d)$ we have $(S_n^{2d})|_A \neq 0$.*

To prove this we need some new notations:

Definition 4.6 *Let M be a monomial, say $M = \prod_i {x_i}^{\alpha_i}$, define:*

- $\alpha(M) = (\alpha_1, \alpha_2, ..., \alpha_n)$
- $|\alpha(M)| = \sum_i \alpha_i$

For two monomials M_1, M_2 denote $M_1 > M_2$ if one of the following happens:

- $|\alpha(M_1)| > |\alpha(M_2)|$.
- $|\alpha(M_1)| = |\alpha(M_2)|$ *and* $\alpha(M_1) > \alpha(M_2)$ *in the lexicographical order.*

For a polynomial $f = \sum_j c_j M_j$ denote $LM(f) = \max_{>}\{M_j\}$.

Proof of Lemma 4.16: Let A be an affine subspace of co-dimension k and let B be a basis for A (as in Definition 2.6). Denote by $Z \subset B$ the variables in B not assigned zero, and assume that $|Z| = z$. So B is actually:

$$x_i = \left\{ \begin{array}{ll} 0 & i \in B \setminus Z \\ \ell_i(X \setminus B) & i \in Z \end{array} \right.$$

The first $k - z$ equations reduce $S_n^{2d}(X)$ to the function $S_{n-(k-z)}^{2d}(X \setminus (B \setminus Z))$. So assume w.l.o.g that $Z = B$. Denote $Y = \{j \mid \exists i \in Z,\ x_j = LM(\ell_i)\}$, $|Y| = t$ and $\tilde{X} = X \setminus (Y \cup B)$, $|\tilde{X}| = \tilde{n}$ (notice that $\tilde{n} = n - (k+t)$). Also define $M = (\prod_{i \in B} x_i)(\prod_{j \in Y} x_j)$. It is clear that we can write

$$S_n^{2d} = M \cdot S_{\tilde{n}}^{2d-(k+t)}(\tilde{X}) + f$$

with f not divisible by M. After the restriction to A we get

$$M = (\prod_{i \in B} \ell_i)(\prod_{j \in Y} x_j) \ ,$$

$$LM(M) = (\prod_{i \in B} LM(\ell_i))(\prod_{j \in Y} x_j) \ and$$

$$(S_n^{2d})|_A = LM(M) \cdot S_{\tilde{n}}^{2d-(k+t)}(\tilde{X}) + \tilde{f}.$$

It is also easy to see that every monomial $\mathcal{M}$ appearing in $\tilde{f}$ is not divisible by $LM(M)$. Therefore $(S_n^{2d})|_A \neq 0$. □

Proof of Lemma 4.15: In the notations of the proof above take all derivatives of order $r = d - \frac{k+t}{2}$ with respect to the variables from $\tilde{X}$. These derivatives include the following set of polynomials:

$$\left\{ LM(M) \cdot S_{\tilde{n}-r}^{r}(\tilde{X} \setminus R) \ \mid \ R \subset \tilde{X}, \ \mid R \mid = r \right\}.$$

These polynomials are all independent, so the dimension of their span is $\binom{\tilde{n}}{r}$ and the worst case is when $t = k$. □

Proof of Theorem 4.13: All we have to do is plug Lemma 4.15 into Theorem 3.2 with $k = d/4, \quad m = n/2$ to get: $L_3(S_n^{2d}) \geq \min(dn/2, \frac{\binom{n-d/2}{\frac{3d}{4}}}{\binom{n/2}{\frac{3d}{4}}}) \geq \min(dn/2, \frac{\binom{7n/9}{\frac{3d}{4}}}{\binom{n/2}{\frac{3d}{4}}}) \geq \min(dn/2, 2^{d/4}) \geq \Omega(nd)$ □

4.5.2 Determinant

Proof of Theorem 4.9: This is an immediate corollary of Proposition 2.4 and the fact that $\partial_d(DET_n) = \binom{n}{d}^2$. These facts give:

$$L_3^m(Det_n) \geq \frac{\binom{n}{d}^2}{\binom{m}{d}} \approx (\frac{n^2}{md})^d.$$

Maximizing d we get $d = \frac{n^2}{me}$ and the lower bound follows. □

Proposition 4.17 *If $\mathcal{F}$ is a $\Sigma\Pi\Sigma$ formula that computes $f(X)$, Where $f(X)$ is a linear function in $x \in X$, i.e $f(X) = xg(X \setminus \{x\}) + h(X \setminus \{x\})$, and x appears in N multiplication gates, each of degree at most m, then there is a $\Sigma\Pi\Sigma$ formula that computes $g(X \setminus \{x\})$ with $2N$ multiplication gates each of degree at most m.*

Proof: Since $g(X \setminus \{x\}) = f(x+1, X \setminus \{x\}) - f(x, X \setminus \{x\})$, we can construct a new formula from $\mathcal{F}$ in the obvious manner, and all the gates not including x will be canceled. □

Proof of Theorem 4.8: The proof is by induction on n. We begin with $x_{1,n}$, if all the multiplication gates including it are of degree $< \frac{n^2}{4e \log n}$ then we make the following assignment:

$$x_{i,n} = 0 \ \ \forall i > 1.$$

Using Proposition 4.17 we get a new formula of degree $< \frac{n^2}{4e \log n}$ for DET_{n-1} and according to Theorem 4.9 there are at least $\frac{1}{2}n^4$ multiplication gates in our formula. Otherwise we can find a multiplication gate of degree $\geq \frac{n^2}{4e \log n}$ and a linear function $x_{1,n} = \ell_n(X \setminus \{x_{1,n}\})$ that nullifies it. We do this for $x_{1,n}...x_{1,2}$. If at any stage there is no multiplication gate of degree $\geq \frac{n^2}{4e \log n}$ involving the variable we look at, then we make the appropriate assignment (at the $k'th$ stage we put $x_{i,n-k+1} = 0 \ \ \forall i > 1$). After doing so for $x_{1,2}$ we put $x_{i,1} = 0 \ \ \forall i > 1, \ \ x_{1,1} = 1$. Thus our restricted formula computes DET_{n-1}. Therefor we get the recursion:

$$L_3(DET_n) \geq min(DET_{n-1} + (n-1)\frac{n^2}{4e \log n}, n^4)$$

Solving we get $DET_n \geq \Omega(\frac{n^4}{\log n})$.

□

5. Strange models

5.1. Restricting the fan-in of the gates on level 1

Since it is hard to prove lower bounds even for such a restricted model, we will restrict the model even further, we will allow the linear functions that are computed at the first level to consist of only one variable, i.e every function will have the form:

$$\ell = \alpha_i x_i + \ell^0$$

For this model we can prove an exponential lower bound for DET_n.

Theorem 5.1 *Every restricted depth-3 formula that computes DET_n must have at least $\Omega(\frac{2^n}{n})$.*

Proof: Write each multiplication gate in the form:

$$M = \prod_{i=1}^{n} \prod_{j=1}^{n} (\alpha_{i,j} x_{i,j} + \alpha_{i,j}^0 \ + \ P_{i,j}(x_{i,j})).$$

(Where every monomial in $P_{i,j}(x_{i,j})$ is of degree ≥ 2). But, in DET_n no two variables from the same column appear in the same monomial, therefore from each of the gates we only need to collect the multilinear monomials in which there are no two variables from the same column. Therefore from each of the terms: $\prod_{j=1}^{n} (\alpha_{i,j} x_{i,j} + \alpha_{i,j}^0 \ + \ P_{i,j}(x_{i,j}))$, only the part of degree 1 can be used to create a monomial that will appear in the result. So we replace each $\prod_{j=1}^{n} (\alpha_{i,j} x_{i,j} + \alpha_{i,j}^0 + P_{i,j}(x_{i,j}))$ with it's degree one monomial: $\sum_{j=1}^{n} \beta_{i,j} x_{i,j}$. Now each multiplication gate looks like

$$M = \prod_{i=1}^{n} \sum_{j=1}^{n} \beta_{i,j} x_{i,j}.$$

The rest of the proof is similar to the proof of Theorem 2.5 from [12]. By Lemma 2.2 for each multiplication gate M, $dim(\partial_{\frac{n}{2}}(M)) \leq 2^n$, and we know that $dim(\partial_{\frac{n}{2}}(DET_n)) \geq \binom{n}{\frac{n}{2}}^2$ therefore Lemma 2.2 gives us a lower bound on the number of multiplication gate of $\frac{\binom{n}{\frac{n}{2}}^2}{2^n} \approx \frac{2^n}{n}$. □

Remark 3 *This proof goes without any change for $\Sigma\Pi\Sigma$ circuits as well.*

5.2. A Nečhiporuk-like lower bound

We will now define a different kind of formula, for which we will prove quadratic lower bounds.

Definition 5.1 *A PolyFormula in the set of variables X over the field F is a tree who's internal nodes are labeled by addition or multiplication, it's inputs are labeled by polynomials in one of the variables of X or by constants from the field. The size of this formula is the number of input nodes that are labeled with polynomials in the variables. We denote the smallest size of a PolyFormula that computes a polynomial f with $\hat{L}(f)$.*

Notice the difference between the definition of PolyFormula and the definition of a Formula where we allow edges to be labeled with constants.

Claim 5.2 *If we change the definition of a Formula to a definition where we do not allow constants on edges then this variant changes only a factor of 2 in the size.*

Proof: We can push constants down towards the inputs until we are left with constants only on edges leaving inputs. Any constant on such edge can be replaced by an input gate labeled with it which fans out to a multiplication gate that will multiply this constant with the appropriate input. □

Since we count only the number of polys in x in the definition of PolyFormula, one may assume that there might be many constants too, but the following lemma bound the number of constants in terms of $\hat{L}$.

Lemma 5.3 *The number of constants in a PolyFormula is at most 3 times the number of inputs labeled with polys in X.*

Proof: Each constant can enter a plus or a times gate with an X-input, or to a plus gate that adds it to a multiplication gate of two X-polys, for otherwise we could delete this constant and change a bit the other constants (we are just throwing away constant subformulas). □

Definition 5.2

$$Disc(X) = \prod_{i<j}(x_j - x_i)$$

X is the set of variables $x_1, ..., x_n$. This is the discriminant polynomial and it is the determinant of the Vandermonde matrix ($a_{i,j} = {x_i}^{j-1}$).

Theorem 5.4 $\hat{L}(Disc_n) \geq (1/3)n^2$.

Definition 5.3 *A $d-scheme$ is a PolyFormula in one variable, x, with a set of constants $C = \{c_1, ..., c_k\}$, so if we run over all the values for C in F^k we get all deg d primitive[1] polynomials in x.*

We call its size $s(d) =$ the min number of x inputs required for such a scheme.

A trivial upper bound for $s(d)$ is d.

Theorem 5.5 *If F is algebraically closed then $s(3d) \leq 2d$.*

Proof: Let's look at the following PolyFormula:

$$(x^2 + b)(x + a) + (c - ab) = x^3 + ax^2 + bx + c.$$

Thus using only two polys in x we can generate every primitive deg 3 polynomial. Since F is algebraically closed every primitive polynomial of degree $3d$ is the multiplication of d deg 3 primitive polynomials who's coefficients are in F. Therefore we can take d distinct copies of this formula, multiply all of them, to get a general deg $3d$ primitive polynomial. □

From now on we will assume that F is either $\mathcal{R}$ or $\mathcal{C}$.

Definition 5.4 *We say that a PolyFormula in one variable x, with a set of constants $C = \{c_1, ..., c_k\}$ is of dimension d if running over all possible assignments to C yields a manifold of polynomials in x of dimension d.*

An easy observation is:

Claim 5.6 *A PolyFormula of dimension d is a polynomial mapping from F^k to a manifold of dimension d, where k is the number of constants in the PolyFormula.*

Since every polynomial mapping is a C^∞ mapping and F^k is a manifold of dimension k, we get by Theorem 5.7 (see [9]) that the dimension of a PolyFormula with k constants is at most k.

Theorem 5.7 *If M,N are manifolds of dimensions m, n respectively, where $m < n$ and $\Phi : M \rightarrow N$ is a C^1 mapping, then $\Phi(M)$ has measure 0 (Provided M has only countably many components).*

Together with Lemma 5.3 we get the following corollary.

corollary 5.1 $\hat{L} \geq d/3$ *for every PolyFormula of dimension d.*

We can now deduce the following lower bound.

Theorem 5.8 $s(d) \geq \frac{d}{3}$.

Proof: Follows from Definition 5.1, Corollary 5.1 and the fact that every d-scheme is a PolyFormula of dimension d. □

We now prove our main theorem.

Proof of Theorem 5.4: Let's view the PolyFormula for $DISC(x_1, ..., x_n)$ as a PolyFormula in x_n over the field $F(x_1, ..., x_{n-1})$ and remove all unnecessary gates. Clearly it is a PolyFormula of dimension $n-1$. According to Corollary 5.1 the number of input gates labeled with polys in x_n is at least $\frac{n-1}{3}$. Since we could do it for every x_i we get $\hat{L}(Disc_n) \geq (1/3)n(n-1)$. □

[1] A primitive polynomial is a polynomial who's leading coefficient equals 1

5.3. The derived homogeneous formula

Theorem 5.9 (observation) *If $f(X)$ is an homogeneous polynomial computed by a $\Sigma\Pi\Sigma$ formula of size s then there's a m such that $g(X,t) = t^m f(X)$ can be computed by an* **homogeneous** *$\Sigma\Pi\Sigma$ formula of size $\leq s^2$.*

Proof: Let d be the highest degree of all multiplication gates in the $\Sigma\Pi\Sigma$ formula computing f. We will now change the formula in the following way:

- Replace each linear function $\ell = \sum_{i=1}^{n} \alpha_i x_i + \ell^0$ with $\hat{\ell} = \sum_{i=1}^{n} \alpha_i x_i + \ell^0 t$.
- Multiply each multiplication gate of degree $d - k$ with t^k.

Each monomial of degree k in the former formula is now multiplied by t^{d-k}, since f is homogeneous the result will be $t^{d-deg(f)} f$. □

Definition 5.5 *Let f be an homogeneous polynomial, the derived homogeneous formula of f is the homogeneous formula of smallest size that computes $t^m f$ for some m. We call it $\mathcal{F}_D(f)$ and we denote by $L_3^D(f)$ the size of the smallest $\mathcal{F}_D(f)$ for f.*

corollary 5.2 *For S_n^{2d}*
1. $L_3^D(S_n^{2d}) \leq n^2$
2. $L_3^H(S_n^{2d}) \geq \binom{n}{d}/2^{2d}$

Proof: According to Theorem 4.2 there is a $\Sigma\Pi\Sigma$ formula of size n^2 that computes S_n^{2d}, therefore according to Theorem 5.9 there is an homogeneous $\Sigma\Pi\Sigma$ formula of size n^2 (the size remains the same under this transformation) that computes $t^{n-2d} S_n^{2d}$.

In [12] it was proved that each homogeneous $\Sigma\Pi\Sigma$ formula that computes S_n^{2d} requires size $\geq \binom{n}{d}/2^{2d}$. □

corollary 5.3 *There is an exponential gap between computing homogeneously an homogeneous polynomial f, and computing homogeneously $t^m f$ for some m.*

corollary 5.4 *Taking derivatives of an homogeneous polynomial f with respect to a single variable or assigning a fixed value (say 1) to a variable may exponentiate the size of the depth-3 homogeneous formula computing it.*

The corollary above is an unconditional result (for a weaker model) of a phenomenon which was observed before. In [16] it was shown that assuming PERM is hard, taking multiple derivatives wrt many variables may exponentiate size of general circuits. One way of seeing it is by looking at the following restriction of $PIP(X,Y)$:

Definition 5.6 $PIP(X, \vec{y}) = \prod_{i=1}^{n} \sum_{j=1}^{n} x_{i,j} y_j$.

Now we get $PERM(X) = \frac{\partial^n}{\partial y_1 \cdots \partial y_n} PIP(X, \vec{y})$.

6 Conclusion

To conclude we will represent some related open problems.

6.1 Depth-3 reductions

While quite basic and powerful, depth-3 formulae seem too weak to perform reductions between interesting functions. Most known reductions when depth is not restricted do not seem to carry through when depth is restricted.

The first problem concerns self reductions of elementary symmetric functions. We have proved a lower bound for the $\log n$ elementary symmetric function. The proof doesn't work for S_n^d where $d \geq \log n$, and so we have no lower bound for higher d. In the Boolean setting there is a trivial reduction from smaller to higher values of d (setting variables to 1). Is there an algebraic analog?

Problem 6.1 *Does $L_3(S_n^d) \leq L_3(S_n^r)$ for $d < r$?*

For polynomialy related functions we define below a natural notion of reduction, and list the most frustrating problems of this type.

Definition 6.1 *For two polynomials f, g we say that $f \leq_3 g$ if $L_3(f) \leq L_3(g)^{O(1)}$.*

A very interesting and nontrivial reduction for formulas without depth restriction is between the Determinant and Iterated Matrix Multiplication (see [21]). Again, we do not know if it extends to the bounded depth case.

Problem 6.2 *Does $DET \leq_3 IMM$?*

Another curiosity is with the Permanent function.

Definition 6.2

$$PERM_n(X) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i,\sigma(i)}$$

Where X is an $n \times n$ matrix and S_n is the symmetric group on n variables.

PERM is known to be $\#P$ complete, and so we expect it to be harder than DET in any model. Surprisingly we have a more efficient depth-3 formula for PERM than for DET.
The following theorem is due to it Ryser, see e.g [20].

Theorem 6.1 (Ryser) *There is an homogeneous $\Sigma\Pi\Sigma$ formula for $PERM(X)$ of size $O(n2^n)$.*

Proof: $PERM(X) = \sum_{T \subset [n]} (-1)^{|T|} \prod_{i=1}^{n} \sum_{j \in T} x_{i,j}$. □

Problem 6.3 *Does $DET \leq_3 PERM$, or maybe $PERM \leq_3 DET$?*

Problem 6.4 *Find a simply exponential ($2^{O(n)}$ size) $\Sigma\Pi\Sigma$ formula for DET.*

6.2 Strange models

In Theorem 5.1 we showed a lower bound for depth-3 formulas which use linear functions of only one variable. What happens if we allow linear functions of two variables ?

Problem 6.5 *Show an exponential lower bound for $\Sigma\Pi\Sigma$ formulas of the form:*

$$\sum_{j=1}^{m} \prod_{i=1}^{deg(M_j)} (\alpha_{i,j,1} x_{k_1} + \alpha_{i,j,2} x_{k_2} + c_{i,j}).$$

7. Acknowledgments

We would like to thank Michael Ben-Or, Dima Grigoriev and Alexander A. Razborov for helpful discussions during various stages of this work. A.S would also like to thank Carmit Benbenishty for her love and support.

References

[1] W. Baur, M. O. Rabin. *Linear disjointness and algebraic complexity,* Logic and Algorithmic, L'Enseignement Mathématique, pp.35-46, 1982.

[2] W. Baur, V. Strassen. *The complexity of partial derivatives,* TCS 22, pp.317-330, 1982.

[3] M. Ben-Or. *Private communication.*

[4] G. H. Gottlieb, *A certain class of incidence matrices,* Proc. AMS 17, pp.1233-1237, 1966.

[5] D. Grigoriev, M. Karpinski. *An exponential lower bound for depth-3 arithmetic circuits*, STOC, pp.577-582, 1998.

[6] D. Grigoriev, A. A. Razborov. *Exponential lower bounds for depth-3 arithmetic circuits in algebras of functions over finite fields*, to appear FOCS, 1998.

[7] R. E. Knuth. *The Art of Computer Programming : Fundamental Algorithms (Vol 1, 3rd Ed)*, Addison-Wesley, 1997.

[8] S. Lang. *Algebra (3rd edition)*, Addison-Wesley, 1993.

[9] J. W. Milnor. *Topology from the differentiable viewpoint*, Princeton Univ Pr., 1997.

[10] E. I. Nečhiporuk. *A boolean function*, Sov. Math. Dokl. 7:4, pp.999-1000, 1966.

[11] N. Nisan. *Lower bounds for non-commutative computation*, 23rd STOC, pp.410-418, 1991.

[12] N. Nisan, A. Wigderson. *Lower bound on arithmetic circuits via partial derivatives*, Comput. comp. Vol 6, pp.217-234, 1996.

[13] A. A. Razborov. *Lower bounds on the size of bounded depth circuits over a complete basis with logical addition*, Math. Notes 41, pp.333-338, 1987.

[14] E. Shamir, M. Snir. *On the depth complexity of formulas*, Math. Systems theory 13, pp.301-322, 1980.

[15] R. Smolensky. *Algebraic methods in the theory of lower bounds for boolean circuit complexity*, 19th STOC, pp.77-82, 1987.

[16] V. Strassen. *Algebraic complexity theory* Handbook of Th. Comp. Science vol. A, pp.634-672, 1990.

[17] P. Tiwari, M. Tompa. *A direct version of Shamir and Snir's lower bounds on monotone circuit depth*, Information Processing Letters 49, 1994.

[18] L. Valiant. *Negation can be exponentially powerful*, TCS 12, pp.303-314, 1980.

[19] L. Valiant, S. Skyum, S. Berkowitz, and C. Rackoff. *Fast parallel computation of polynomials using few processors*, Siam J. Comp. 12, pp.641-644, 1983.

[20] J. H. Van Lint, R. M. Wilson. *A course in combinatorics* Cambridge Univ Pr, 1992.

[21] J. Von Zur Gathen. *Algebraic complexity theory*, Ann. Rev. Comp. Sci. 3, pp.317-347, 1988.