Arithmetic Circuit Lower Bounds via MaxRank*

Mrinal Kumar[1], Gaurav Maheshwari[2], and Jayalal Sarma M.N.[2]

[1] Department of Computer Science, Rutgers University, Piscataway, NJ 08855, USA
[2] Department of Computer Science & Engineering, IIT Madras, Chennai 36, India

Abstract. We introduce the polynomial coefficient matrix and identify maximum rank of this matrix under variable substitution as a complexity measure for multivariate polynomials. We use our techniques to prove super-polynomial lower bounds against several classes of non-multilinear arithmetic circuits. In particular, we obtain the following results :

- As our first main result, we prove that any homogeneous depth-3 circuit for computing the product of d matrices of dimension $n \times n$ requires $\Omega(n^{d-1}/2^d)$ size. This improves the lower bounds in [9] for $d = \omega(1)$.
- As our second main result, we show that there is an explicit polynomial on n variables and degree at most $\frac{n}{2}$ for which any depth-3 circuit C of product dimension at most $\frac{n}{10}$ (dimension of the space of affine forms feeding into each product gate) requires size $2^{\Omega(n)}$. This generalizes the lower bounds against diagonal circuits proved in [14]. Diagonal circuits are of product dimension 1.
- We prove a $n^{\Omega(\log n)}$ lower bound on the size of product-sparse formulas. By definition, any multilinear formula is a product-sparse formula. Thus, this result extends the known super-polynomial lower bounds on the size of multilinear formulas [11].
- We prove a $2^{\Omega(n)}$ lower bound on the size of partitioned arithmetic branching programs. This result extends the known exponential lower bound on the size of ordered arithmetic branching programs [7].

1 Introduction

Arithmetic circuits are a fundamental model of computation for polynomials. Establishing the limitations of polynomial sized arithmetic circuits is a central open question in the area of algebraic complexity(see [17] for a detailed survey). One of the recent surprises in the area was the result due to Agrawal and Vinay [2] where they show that if a polynomial in n variables of degree d (linear in n) can be computed by arithmetic circuits of size $2^{o(n)}$, then it can also be computed by depth-4 circuits of size $2^{o(n)}$. The parameters of this result was further tightened by Koiran [8]. These results explained the elusiveness of proving lower bounds against even depth-4 circuits. For depth-3 circuits, the best known general result (over finite fields) is an exponential lower bound due to Grigoriev

* The full version of the paper is available as a technical report at the ECCC. Technical report No. ECCC-TR-13-028. See `http://eccc.hpi-web.de/report/2013/028/`

F.V. Fomin et al. (Eds.): ICALP 2013, Part I, LNCS 7965, pp. 661–672, 2013.

and Karpinski [5] and Grigoriev and Razborov [4]. Over infinite fields, obtaining such strong lower bounds is a long-standing open problem. Lower bounds for restricted classes of depth-3 and depth-4 circuits are studied in [1,9,16] .

One class of models which has been extensively studied is when the gates are restricted to compute multilinear polynomials. Super-polynomial lower bounds are known for the size of multilinear formulas computing the permanent or determinant polynomial [12]. However, even under this restriction proving super-polynomial lower bounds against arbitrary multilinear arithmetic circuits is an open problem (see [17] and references there in). The parameter identified by [11], which showed the limitations of multilinear formulas, was the rank of a matrix associated with the circuit - namely the partial derivatives matrix[1]. The method showed that there exists a partition of variables into two sets such that the rank of the partial derivatives matrix of any polynomial computed by the model is upper bounded by a function of the size of the circuit. But there are explicit polynomials for which the rank of the partial derivatives matrix is high. This program has been carried out for several classes of multilinear polynomials and several variants of multilinear circuits [3,7,10,11,12,13]. However, the partial derivatives matrix, in the form that was studied, was known to yield lower bounds only for multilinear circuits.

In this work, we generalize this framework to prove lower bounds against several classes of non-multilinear arithmetic circuits. This generalization also shows that the multilinearity restriction in the above proof strategy can possibly be eliminated from the circuit model side. Hence it can also be seen as an approach towards proving lower bounds against the general arithmetic circuits.

We introduce a variant of the partial derivatives matrix where the entries will be polynomials instead of constants - which we call the *polynomial coefficient matrix*. Instead of rank of the partial derivatives matrix, we analyze the max-rank - the maximum rank of the polynomial coefficient matrix[2] under any substitution for the variables from the underlying field. We first prove how the max-rank changes under arithmetic operations. These tools are combined to prove upper bounds on max-rank of various restrictions of arithmetic circuits.

In [9], it was proved that any homogeneous depth-3 circuit for multiplying d $n \times n$ matrices (Iterated Matrix Multiplication, IMM_d^n) requires $\Omega\left(n^{d-1}/d!\right)$ size. We use our techniques to improve this result in terms of the lower bound. Our methods are completely different from [9] and this demonstrates the power of this method beyond the reach of the original partial derivatives matrix method due to Raz [11]. As our first main result, we prove the following.

Theorem 1. *Any homogeneous depth-*3 *circuit for computing the product of d matrices of dimension $n \times n$ requires $\Omega(n^{d-1}/2^d)$ size.*

[1] An exponential sized matrix associated with the multilinear polynomial with respect to a partition of the variables into two sets. See Section 2 for the formal definition.

[2] When it is clear from the context, we drop the matrix as well as the partition. By the term, max-rank of a polynomial, we denote the maximum rank of the polynomial coefficient matrix corresponding to the polynomial with respect to the partition in the context.

Notice that compared to the bounds in [9], our bounds are stronger when $d = \omega(1)$. Very recently, Gupta et al. [6] studied the model of homogeneous circuits and proved a strong lower bound parameterized by the bottom fan-in. They studied depth-4 circuits ($\Sigma\Pi\Sigma\Pi$) and showed that if the fan-in of the bottom level product gate of the circuits is t, then any homogeneous depth-4 circuit computing the permanent (and the determinant) of $n \times n$ matrices must have size $2^{\Omega(\frac{n}{t})}$. In particular, this implies a $2^{\Omega(n)}$ lower bound for any depth-3 homogeneous circuit computing the permanent (and the determinant) of $n \times n$ matrices (n^2 variables). However, we remark that Theorem 1 is addressing the iterated matrix multiplication polynomial and hence is not directly subsumed by the above result. Moreover, the techniques used in [6] are substantially different from ours.

We apply our method to depth-3 circuits where space of the affine forms feeding into each product gate in the circuit is of limited dimension. Formally, a depth-3 $\Sigma\Pi\Sigma$ circuit C is said to be of product dimension r if for each product gate P in C, where $P = \Pi_{i=1}^{d} L_i$, where L_i is an affine form for each i, the dimension of the span of the set $\{L_i\}_{i\in[d]}$ is at most r. As our second main result, we prove exponential lower bounds on the size (in fact, the top fan in) of depth-3 circuits of bounded product dimension for computing an explicit polynomial.

Theorem 2. *There is an explicit polynomial on n variables and degree $\leq \frac{n}{2}$ for which any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ requires size $2^{\Omega(n)}$.*

In [14], the author studies diagonal circuits, which are depth-3 circuits where each product gate is an exponentiation gate. Clearly, such a product gate can be visualized as a product gate with the same affine form being fed into it multiple times. Thus, these circuits are of product dimension 1, and our lower bound result generalizes size lower bounds against diagonal circuits.

Note that the product dimension of a depth-3 circuit is different from the dimension of the span of all affine forms computed at the bottom sum gates of a $\Sigma\Pi\Sigma$ circuit. It can be easily seen that, when this parameter, which we refer to as the total dimension of the circuit, when bounded, the model non-universal.

For our next result, we generalize the model of syntactic multilinear formulas to product-sparse formulas. We formally define product-sparse formulas and full max-rank polynomials in Section 2. These formulas can compute non-multilinear polynomials as well. We show the following theorem regarding this model using our methods.

Theorem 3. *Let X be a set of $2n$ variables and let $f \in \mathbb{F}[X]$ be a full max-rank polynomial. Let Φ be any (s, d)-product-sparse formula of size $n^{\epsilon \log n}$, for a constant ϵ. If $sd = o(n^{1/8})$, then f cannot be computed by Φ.*

As our fourth result, we define partitioned arithmetic branching programs which are generalizations of ordered ABPs. We prove an exponential lower bound for partitioned ABPs extending results in [7].

Theorem 4. *Let X be a set of $2n$ variables and $\mathbb{F}$ be a field. For any full max-rank homogeneous polynomial f of degree n over X and $\mathbb{F}$, the size of any partitioned ABP computing f must be $2^{\Omega(n)}$.*

2 Preliminaries

In this section, we define some of the models we study. For more detailed account of models and the results we refer the reader to the survey [17].

An arithmetic circuit Φ over the field $\mathbb{F}$ and the set of variables $X = \{x_1, x_2, \ldots, x_n\}$ is a directed acyclic graph $G = (V, E)$. The vertices of G with in-degree 0 are called *input* gates and are labelled by variables in X or constants from the field $\mathbb{F}$. The vertices of G with out-degree 0 are called *output* gates. Rest all vertices are referred to as internal vertices. Every internal vertex is either a plus gate or a product gate. We will study arithmetic circuits with a single output gate. Thus, the polynomial computed by the arithmetic circuit is the polynomial associated with the output gate. The size of Φ is defined to be the number of gates in Φ. For a vertex $v \in V$, we denote the set of variables that occur in the subgraph rooted at v by X_v.

We consider depth restricted circuits. A $\Sigma\Pi\Sigma$ circuit is a levelled depth-3 circuit with a plus gate at the top, multiplication gates at the middle level and plus gates at the bottom level. The fan-in of the top plus gate is referred to as top fan-in. A $\Sigma\Pi\Sigma$ circuit is said to be *homogeneous* if the plus gate at the bottom level compute homogeneous linear forms only.

An important restricted model of arithmetic circuits is multilinear circuits. A polynomial $f \in \mathbb{F}[X]$ is called *multilinear* if the degree of every variable in f is at most one. An arithmetic circuit is called *multilinear* if the polynomial computed at every gate is multilinear. An arithmetic circuit is called *syntactic multilinear* if for every product gate v with children v_1 and v_2, $X_{v_1} \cap X_{v_2} = \phi$. An arithmetic circuit is called an *arithmetic formula* if the underlying undirected graph is acyclic i.e. fan-out of every vertex is at most one.

Let Φ be a formula defined over the set of variables X and a field $\mathbb{F}$. For a product gate v in Φ with children v_1 and v_2, let us define the following properties:

Disjoint v is said to be *disjoint* if $X_{v_1} \cap X_{v_2} = \phi$.
Sparse v is said to be *s-sparse* if the number of monomials in the polynomial computed by at least one of its input gates is at most 2^s.

For a node v, let us define the product-sparse depth of v to be equal to the maximum number of non-disjoint product gates in any path from a leaf to v.

Definition 1. *A formula is said to be a (s, d)-product-sparse if every product gate v is either disjoint or s-sparse, where d is the product-sparse depth of the root node.*

Clearly, any syntactic multilinear formula is a $(s, 0)$-product-sparse formula for any s. Thus, proving lower bounds for product-sparse formulas will be a strengthening of known results.

An Arithmetic Branching Program (ABP) B is a levelled graph $G(V, E)$ in which V can be partitioned into levels $L_0, L_1, \ldots, L_d$ such that $L_0 = \{s\}$ and $L_d = \{t\}$ and edges can only go between consecutive levels. s and t are called the *source* and *sink* respectively. The weight function w assigns affine forms to E. For a path p, extend the weight function by $w(p) = \prod_{e \in p} w(e)$. B computes the polynomial $\sum_p w(p)$ where p runs over all source-sink paths. B is said to be *homogeneous* if all edge labels are homogeneous linear forms and naturally computes a homogeneous polynomial. For any $i, j \in V$, $P_{i,j}$ denotes all paths from i to j in G, $X_{i,j}$ denotes the variables occuring in those paths and $f_{i,j}$ denotes the polynomial $\sum_{p \in P_{i,j}} w(p)$.

Definition 2. *Let B be a homogeneous ABP over a field $\mathbb{F}$ and set of variables $X = \{x_1, x_2, \ldots, x_{2n}\}$. B is said to be π-partitioned for a permutation $\pi : [2n] \to [2n]$ if there exists an $i = 2\alpha n$ for some constant α such that the following condition is satisfied, $\forall v \in L_i$:*

- *Either, $X_{s,v} \subseteq \{x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}\}$ and $|X_{v,t}| \leq 2n(1-\alpha)$.*
- *Or, $X_{v,t} \subseteq \{x_{\pi(n+1)}, x_{\pi(n+2)}, \ldots, x_{\pi(2n)}\}$ and $|X_{s,v}| \leq 2n(1-\alpha)$*

We say that B is partitioned with respect to the level L_i. B is said to be a partitioned ABP if it is π-partitioned for some $\pi : [2n] \to [2n]$.

We now introduce the main tool used in the paper and prove its properties. Let $Y = \{y_1, y_2, \ldots, y_m\}$ and $Z = \{z_1, z_2, \ldots, z_m\}$ be two sets of variables. Let $f \in \mathbb{F}[Y, Z]$ be a multilinear polynomial. Define L_f to be the $2^m \times 2^m$ *partial derivatives matrix* as follows: for monic multilinear monomials $p \in \mathbb{F}[Y], q \in \mathbb{F}[Z]$, define $L_f(p, q)$ to be the coefficient of the monomial pq in f. Let us denote the rank of L_f by $\mathrm{rank}(L_f)$. We extend the partial derivatives matrix to non-multilinear polynomials.

Definition 3 (Polynomial Coefficient Matrix). *For $f \in \mathbb{F}[Y, Z]$, define M_f to be the $2^m \times 2^m$* polynomial coefficient matrix *with each entry from $\mathbb{F}[Y, Z]$ defined as follows. For monic multilinear monomials p and q in Y and Z respectively, $M_f(p, q) = G$ if and only if f can be uniquely written as $f = pq(G) + Q$, where $G, Q \in \mathbb{F}[Y, Z]$ such that G does not contain any variable other than those present in p and q, Q does not have any monomial m which is divisible by pq and which contains only variables that are present in p and q.*

Observe that we can write, $f = \sum\limits_{p,q} M_f(p, q)pq$ and for a multilinear polynomial f, M_f is same as L_f. For any function $S : Y \cup Z \to \mathbb{F}$, let us denote by $M_f|_S$ the matrix obtained by substituting each variable x by $S(x)$ at each entry in M_f. Let us define $\text{max-rank}(M_f) = \max\limits_{S:Y \cup Z \to \mathbb{F}} \{\mathrm{rank}(M_f|_S)\}$. The following proposition bounds the max-rank of the matrix (similar bounds on the rank of partial derivatives matrix for some cases have been proved in [13]). We defer the proof to the full version of the paper.

Proposition 1. *Let $f, g \in \mathbb{F}[Y, Z]$, $h \in \mathbb{F}[Y]$ and $w \in F[Z]$.*

1.1 If f contains variables $Y' \subseteq Y$ and $Z' \subseteq Z$ only, then $\text{max-rank}(M_f) \leq 2^a$ *where $a = \min\{|Y'|, |Z'|\}$.*

1.2 $\text{max-rank}(M_{f+g}) \leq \text{max-rank}(M_f) + \text{max-rank}(M_g)$.

1.3 Let $Y_1, Y_2 \subseteq Y$ and $Z_1, Z_2 \subseteq Z$ such that $Y_1 \cap Y_2 = \phi$ and $Z_1 \cap Z_2 = \phi$. If $f \in \mathbb{F}[Y_1, Z_1]$ and $g \in \mathbb{F}[Y_2, Z_2]$, then $\text{max-rank}(M_{fg}) = \text{max-rank}(M_f) \cdot \text{max-rank}(M_g)$.

1.4 max-rank$(M_{fh}) \leq$ max-rank(M_f) *and* max-rank$(M_{fw}) \leq$ max-rank(M_f)*.*
1.5 If g is a linear form, then max-rank$(M_{fg}) \leq 2 \cdot$ max-rank(M_f)*.*
1.6 If g can be expressed as $\sum_{i\in[r]} h_i w_i$ where $h_i \in \mathbb{F}[Y]$ and $w_i \in \mathbb{F}[Z]$, then max-rank$(M_{fg}) \leq r \cdot$ max-rank(M_f)*.*
1.7 If g has r monomials, then max-rank$(M_{fg}) \leq r \cdot$ max-rank(M_f)*.*

Full Rank Polynomials: Let $X = \{x_1, \cdots, x_{2n}\}, Y = \{y_1, \cdots, y_n\}$ and $Z = \{z_1, \cdots, z_n\}$ be sets of variables and $f \in \mathbb{F}[X]$. f is said to be a *full rank* polynomial if for any partition $A : X \to Y \cup Z$, rank$(L_{f^A}) = 2^n$, where f^A is the polynomial obtained from f after substituting every variable x by $A(x)$. We say that f is a *full max-rank* polynomial if max-rank$(M_{f^A}) = 2^n$ for any partition A. Any full rank polynomial is also a full max-rank polynomial. Many full rank polynomials have been studied in the literature [7,11,12].

3 Lower Bounds against Homogeneous Depth-3 Circuits

Let Φ be a homogeneous $\Sigma\Pi\Sigma$ circuit with top fan-in k defined over the set of variables X and field $\mathbb{F}$ computing a homogeneous polynomial $f = \sum_{i=1}^{k} P_i$, where $P_i = \prod_{j=1}^{deg(P_i)} l_{i,j}$, each $l_{i,j}$ is a linear form and $deg(P_i)$ is the fan-in of the i^{th} multiplication gate. For a partition $A : X \to Y \cup Z$, denote by Φ^A the circuit obtained after replacing every variable x by $A(x)$ and the corresponding polynomial by f^A. We prove the following upper bound on the max-rank(M_{f^A}).

Lemma 1. *Let Φ be a homogeneous $\Sigma\Pi\Sigma$ circuit as defined above and the degree of f be d. Then, for any partition $A : X \to Y \cup Z$,* max-rank$(M_{f^A}) \leq k \cdot 2^d$*.*

Proof. Let us denote by $l^A_{i,j}$ and P^A_i the polynomials obtained after substitution of x by $A(x)$ in the polynomials $l_{i,j}$ and P_i respectively.
Since each $l_{i,j}$ is a homogeneous linear form, a multiplication gate P_i computes a homogeneous polynomial of degree $deg(P_i)$. Thus if $deg(P_i) \neq d$ then the multiplication gate P_i does not contribute any monomial in the output polynomial f. Hence, it can be assumed without loss of generality that $deg(P_i) = d$ for all $i \in [k]$.

Since $l_{i,j}$ is a homogeneous linear form, max-rank$(M_{l^A_{i,j}}) \leq 2$. Thus, using Proposition 1.5,$\forall i \in [k]$: max-rank$(M_{P^A_i}) \leq 2^d$. Hence, using Proposition 1.2, max-rank$(M_{f^A}) \leq \sum_{i\in[k]}$ max-rank$(M_{P^A_i}) \leq k \cdot 2^d$.

In [9], it was proved that any homogeneous $\Sigma\Pi\Sigma$ circuit for multiplying d $n \times n$ matrices requires $\Omega(n^{d-1}/d!)$ size. We prove a better lower bound using our techniques. Formally, let $X^1, X^2, \ldots, X^d$ be disjoint sets of variables of size n^2 each, with $X = \cup_{i\in[d]} X^i$. The variables in X^i will be denoted by x^i_{jk} for $j, k \in [n]$. We will be looking at the problem of multiplying d $n \times n$ matrices $A^1, A^2, \ldots, A^d$

where $(j,k)^{th}$ entry of matrix A^i, denoted by A^i_{jk}, is defined to be equal x^i_{jk} for all $i \in [d]$ and $j,k \in [n]$. The output polynomial, that we are interested in, is the $(1,1)^{th}$ entry of $\prod_{i\in[d]} A^i$ denoted by f. We also refer to f by IMM^n_d. f is clearly a homogeneous multilinear polynomial of degree d. Moreover, any monomial in f contains one variable each from the sets $X^1, X^2, \ldots, X^d$.

We first prove an important lemma below.

Lemma 2. *For the polynomial f as defined above, there exists a bijective partition $B : X \to Y \cup Z$ such that* $\text{max-rank}(M_{f^B}) = n^{d-1}$.

Proof. We fix some notations first. For $i < j$, let us denote the set $\{i, i+1, \ldots, j\}$ by $[i,j]$. Let us also denote the pair $((k,i),(k+1,j))$ by e_{ijk} for any i,j,k. Construct a directed graph $G(V,E)$ on the set of vertices $V = [0,d] \times [1,n]$ and consisting of edges $E = \{e_{ijk} \mid k \in [0,d-1], i,j \in [1,n]\}$. Note that the edges e_{ijk} and e_{jik} are two distinct edges for fixed values of i,j,k when $i \neq j$. Let us also define a weight function $w : E \to X$ such that $w(e_{ijk}) = x^{k+1}_{ij}$.

It is easy to observe that the above graph encodes the matrices $A_1, A_2, \ldots, A_d$. The weights on the edges are the variables in the matrices. For example, a variable x^{k+1}_{ij} in the matrix A_{k+1} is the weight of the edge e_{ijk}. Let us denote the set of paths in G from the vertex $(0,1)$ to the vertex $(d,1)$ by $\mathcal{P}$. Let us extend the weight function and define $w(p) = \prod_{e\in p} w(e)$ for any $p \in \mathcal{P}$. Since, all paths in $\mathcal{P}$ are of length equal to d, the weights corresponding to each of these paths are monomials of degree d.

Let us define the partition $B : X \to Y \cup Z$ as follows: all the variables in odd numbered matrices are assigned variables in Y and all the variables in even numbered matrices are assigned variables in Z. Let us denote the variable assigned by B to x^{2k-1}_{ij} by y^{2k-1}_{ij} and the variable assigned to x^{2k}_{ij} by z^{2k}_{ij}.

It follows from the matrix multiplication properties that for any path $p \in \mathcal{P}$, the monomial $w(p)$ is a monomial in the output polynomial. Each such path is uniquely specified once we specify the odd steps in the path. Now, specifying odd steps in the path corresponds to specifying a variable from each of the odd numbered matrices. To count number of such ways, let us first consider the case when d is even. There are $d/2$ odd numbered matrices and we have n^2 ways to choose a variable from each of these $d/2$ matrices except for the first matrix for which we can only choose a variable from the 1^{st} row since our output polynomial is the $(1,1)^{th}$ entry. Thus, there are n^{d-1} number of ways to specify one variable each from the odd numbered matrices, the number of such paths is also n^{d-1}. We get the same count for the case when d is odd using a similar argument. Since once the odd steps are chosen, there is only one way to choose the even steps, all these n^{d-1} monomials give rise to non-zero entries in different rows and columns in the matrix M_{f^B}. Hence, the matrix is an identity block of dimension n^{d-1} upto a permutation of rows and columns and thus it has rank n^{d-1}.

Theorem 5. *Any homogeneous $\Sigma\Pi\Sigma$ circuit for computing the product of d $n \times n$ matrices requires $\Omega(n^{d-1}/2^d)$ size.*

Proof. Let $\varPhi$ be a homogeneous $\varSigma\varPi\varSigma$ circuit computing f. Then, using Lemma 1, for any partition A, max-rank$(M_{f^A}) \leq k \cdot 2^d$. From Lemma 2, we know that there exists a partition B such that max-rank$(M_{f^B}) = n^{d-1}$. Hence, $k \geq n^{d-1}/2^d$.

It is worth noting that there exists a depth-2 circuit of size n^{d-1} computing IMM_d^n polynomial. As observed in Lemma 2, there are n^{d-1} monomials in the IMM_d^n polynomial. Hence, the sum of monomials representation for IMM_d^n will have top fan-in equal to n^{d-1}. We remark that when the number of matrices is a constant, the upper and lower bounds for IMM_d^n polynomial asymptotically match.

4 Lower Bounds against Depth-3 Circuits of Bounded Product Dimension

If a depth-3 circuit is not homogeneous, the fan-in of a product gate can be arbitrarily larger than the degree of the polynomial being computed. Hence the techniques in the previous section fails to give non-trivial size lower bounds. In this section, we study depth-3 circuits with bounded product dimension - where the affine forms feeding into every product gate are from a linear vector space of small dimension and prove exponential size lower bounds for such circuits.

We will first prove an upper bound on the max-rank of the polynomial coefficient matrix for the polynomial computed by a depth-3 circuit of product dimension r, parameterized by r. Let C be a $\varSigma\varPi\varSigma$ circuit of product dimension r and top fan in k. Let P^j be the product gates in C for $j \in [k]$, given by $P^j = \varPi_{i=1}^{s} L_i^j$. Without loss of generality, let us assume that the vectors $L_1^j, L_2^j, \ldots, L_r^j$ form a basis for the span of $\{L_1^j, L_2^j, \ldots, L_s^j\}$. Let l_i^j be the homogeneous part of L_i^j for each i. So, clearly the set $\{l_i^j\}_{i\in[r']}$ spans the set $\{l_i^j\}_{i\in[s]}$, where $r' \leq r$. To simplify the notation, we will refer to r' as r. In the following presentation, we will always use d to refer to the degree of the homogeneous polynomial computed by the circuit under consideration. Now, let us express each l_i^j as a linear combination of $\{l_i^j\}_{i\in[r]}$. Let us expand the product P^j into a sum of product of homogeneous linear forms coming from $\{l_i^j\}_{i\in r}$. Let P_d^j be the slice of P^j of degree exactly d, for each $j \in [k]$.

Observation 6. *Let $C_d = \varSigma_{i\in[k]} P_d^i$. If C computes a homogeneous polynomial of degree d, then C_d computes the same polynomial.*

Proof. We know that, $C = \varSigma_{i\in[k]} P^i$. Now, writing each product gate as a sum of product of homogeneous linear forms as described in the paragraph above, we get $C = \varSigma_{i\in[k]} \varSigma_{j\in[d]} P_j^i$. Now, equating the degree d parts of the polynomial in both sides of the equality, we obtain $C_d = \varSigma_{i\in[k]} P_d^i$. If C computes a homogeneous polynomial of degree d, $C = C_d$ and the lemma follows.

We know that for each $P_d^j = \varSigma_i \varPi_{u=1}^{d} l_{\alpha_{iu}}$ where $\alpha_{iu} \in [r]$ and $j \in [k]$. We now use the following lemma to simplify the inner product terms $\varPi_{u=1}^{d} l_{\alpha_{iu}}$ in the expression for P_d^j.

Lemma 3. *([15]) Any monomial of degree d can be written as a linear combination of d^{th} power of some 2^d linear forms. Further, each of the 2^d linear forms in the expression corresponds to $\Sigma_{x \in S} x$ for a subset S of $[d]$.*

By applying to each product term $\Pi_{u=1}^{d} l_{\alpha_{iu}}$ in P_d^j, we obtain the following:

Lemma 4. *If $P_d^j = \Sigma_i \Pi_{u=1}^{d} l_{\alpha_{iu}}$ where $\alpha_{iu} \in [r]$, then $P_d^j = \Sigma_{q=1}^{v} c_q {L_q}^d$ for some homogeneous linear forms L_q, constants c_q and $v \leq \binom{d+r}{r}$.*

Proof. Consider any product term in the sum of products expansion P_d^j as described, say $S = \Pi_{u=1}^{d} l_{\alpha_{iu}}$. From Lemma 3, we know that S can be written as $S = \Sigma_{t=1}^{2^d} {L_t}^d$, where for every subset U of $[d]$, there is a $\beta \in [2^d]$ such that $L_\beta = \Sigma_{u \in U} l_{\alpha_{iu}}$. In general, each L_t can be written as $L_t = \Sigma_{i \in [r]} \gamma_i l_i$ for non-negative integers γ_i satisfying $\Sigma_{i \in [r]} \gamma_i \leq d$. Now, each of the product terms in P_d^j can be expanded in a similar fashion into d^{th} powers of linear forms, each from the set $\{\Sigma_{i \in [r]} \gamma_i l_i : \gamma_i \in \mathbb{Z}^{\geq 0} \wedge \Sigma_{i \in [r]} \gamma_i \leq d\}$. The number of distinct such linear forms is at most $\binom{d+r}{r}$. Hence, the lemma follows.

We now bound the max-rank of the power of a homogeneous linear form which in turn will give us a bound for P_d^j due to the subadditivity of max-rank.

Lemma 5. *Given a linear form l and any positive integer t, the* max-rank *of l^t is at most $t+1$ for any partition of the set X of variables into Y and Z.*

Proof. Partition the linear form l into two parts, $l = l_y + l_z$, where l_y consists of all variables in l from the set Y and l_z consists of the variables which come from the set Z. By the binomial theorem, $l^t = \Sigma_{i=0}^{t} \binom{t}{i} l_y^i l_z^{t-i}$. Now, l_y^i is a polynomial just in Y variables and hence its max-rank can be bounded above by 1, and multiplication by l_z^{t-i} does not increase the max-rank any further, by Proposition 1.4. Hence, the max-rank of each term in the sum is at most 1 and there are at most $t+1$ terms, so, by using the subadditivity of max-rank, we get an upper bound of $t+1$ on the max-rank of the sum.

The following lemma gives an upper bound on the max-rank of P_d^j and follows from Lemma 4, Lemma 5 and the subadditivity of max-rank.

Lemma 6. *The* max-rank *of P_d^j is at most $(d+1)\binom{d+r}{r}$ for any partition of the set X of variables into Y and Z.*

Now we are ready to prove the theorem.

Theorem 7. *There is an explicit polynomial in n variables and degree $\frac{n}{2}$ for which any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ requires size $2^{\Omega(n)}$.*

Proof. We describe the explicit polynomial $Q(X)$ first. Fix an equal sized partition A of X into Y and Z. Order all subsets of Y and Z of size exactly $\frac{n}{4}$ in any order, say $S_1, S_2, \ldots, S_w$ and $T_1, T_2, \ldots, T_w$, where $w = \binom{\frac{n}{2}}{\frac{n}{4}}$. Let us define the polynomial $Q^A(Y,Z)$ for the partition A as follows: $Q^A(Y,Z) = \Sigma_{i=1}^{w} \Pi_{y \in S_i} \Pi_{z \in T_i} yz$.

We obtain the polynomial $Q(X)$ by replacing variables in Y and Z in $Q^A(Y,Z)$ by $A^{-1}(Y)$ and $A^{-1}(Z)$ respectively.

Now we prove the size lower bound. The polynomial coefficient matrix of Q with respect to the partition Y and Z is simply the diagonal submatrix, and the rank is at least $\binom{\frac{n}{2}}{\frac{n}{4}} \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$. Since C computes the polynomial, the top fan in k should be at least $\frac{\frac{2^{\frac{n}{2}}}{\sqrt{n}}}{\binom{d+r}{r}(d+1)}$. For $d = \frac{n}{2}$, and $r = \frac{n}{10}$, we have a lower bound of 2^{cn}, for a constant $c > 0$.

5 Lower Bounds against Product-Sparse Formulas

Let $Y = \{y_1, y_2, \ldots, y_m\}$ and $Z = \{z_1, z_2, \ldots, z_m\}$. Let Φ be a (s,d)-product-sparse formula defined over the field $\mathbb{F}$ and the variables $Y \cup Z$. For a node v, let us denote by Φ_v the sub-circuit rooted at v, and denote by Y_v and Z_v, the set of variables in Y and Z that appear in Φ_v respectively. Let us define, $a(v) = \min\{|Y_v|, |Z_v|\}$ and $b(v) = (|Y_v| + |Z_v|)/2$. We say that a node v is k-unbalanced if $b(v) - a(v) \geq k$. Let γ be a simple path from a leaf to the node v. We say that γ is k-unbalanced if it contains at least one k-unbalanced node. We say that γ is central if for every u, u_1 on the path γ such that there is an edge from u_1 to u in Φ, $b(u) \leq 2b(u_1)$. v is said to be k-weak if every central path that reaches v is k-unbalanced.

We prove that if v is k-weak then the max-rank of the matrix M_v can be bounded. The proof goes via induction on $|\Phi_v|$ and follows the same outline as that of [12]. It only differs in the case of non-disjoint product gates which we include in full detail below. The proofs of the rest of cases is easy to see.

Lemma 7. *Let Φ be a (s,d)-product-sparse formula over the set of variables $Y \cup Z$, and let v be a node in Φ. Denote the product-sparse depth of v by $d(v)$. If v is k-weak,* $\text{max-rank}(M_v) \leq 2^{s \cdot d(v)} \cdot |\Phi_v| \cdot 2^{b(v)-k/2}$.

Proof. Consider the case when v is a s-sparse product gate with children v_1 and v_2. Without loss of generality it can be assumed that v is not disjoint.

Let us suppose that the product-sparse depth of v is d. Without loss of generality, assume that v_2 computes a sparse polynomial having at most 2^s number of monomials. Thus using Proposition 1.7, $\text{max-rank}(M_v) \leq 2^s \cdot \text{max-rank}(M_{v_1})$ Clearly, product-sparse depth of v_1 is at most $d - 1$. Consider the following cases: Case 1 : If $b(v) \leq 2b(v_1)$, then v_1 is also k-weak. Therefore, by induction hypothesis, $\text{max-rank}(M_{v_1}) \leq 2^{s(d-1)} \cdot |\Phi_{v_1}| \cdot 2^{b(v_1)-k/2} \leq 2^{s(d-1)} \cdot |\Phi_v| \cdot 2^{b(v)-k/2}$. Thus, $\text{max-rank}(M_v) \leq 2^{sd} \cdot |\Phi_v| \cdot 2^{b(v)-k/2}$. Case 2 : If $b(v) > 2b(v_1)$, then $b(v_1) < b(v)/2 < b(v) - k/2$ since $b(v) \geq k$. Therefore using Proposition 1.1, $\text{max-rank}(M_{v_1}) \leq 2^{a(v_1)} \leq 2^{b(v_1)} < 2^{b(v)-k/2}$. Therefore, $\text{max-rank}(M_v) \leq 2^s \cdot 2^{b(v)-k/2} \leq 2^{sd} \cdot |\Phi_v| \cdot 2^{b(v)-k/2}$.

Now, to prove a lower bound for (s,d)-product-sparse formulas computing a full max-rank polynomial, we only need to show that there exists a partition that makes the formula k-weak with suitable values of s, d and k.

In [11], Raz proved that for syntactic multilinear formulas of size at most $n^{\epsilon \log n}$, where ϵ is a small enough universal constant, there exists such a partition that makes the formula k-weak for $k = n^{1/8}$. We observe that this result also holds for product-sparse formulas, the proof given in [11] is not specific to just syntactic multilinear formulas and holds for any arithmetic formula. With above lemma, the following theorem is easy to derive.

Theorem 8. *Let X be a set of $2n$ variables and let $f \in \mathbb{F}[X]$ be a full max-rank polynomial. Let Φ be any (s,d)-product-sparse formula of size $n^{\epsilon \log n}$ for a constant ϵ (same as in [11]). If $sd = o(n^{1/8})$, then f cannot be computed by Φ.*

6 Lower Bounds against Partitioned Arithmetic Branching Programs

In the preliminaries section, we defined partitioned arithmetic branching programs which are a generalization of ordered ABPs. By definition, any polynomial computed by a partitioned ABP is homogenous. In [7], a full rank homogenous polynomial was constructed. Thus, to prove lower bounds for partitioned ABP, we only need to upper bound the max-rank of the polynomial coefficient matrix for any polynomial being computed by a partitioned ABP. Now we prove such an upper bound and use it to prove exponential lower bound on the size of partitioned ABPs, thus extending result in [7].

Theorem 9. *Let X be a set of $2n$ variables and $\mathbb{F}$ be a field. For any full max-rank homogenous polynomial f of degree n over X and $\mathbb{F}$, the size of any partitioned ABP computing f must be $2^{\Omega(n)}$.*

Proof. Let B be a π-partitioned ABP computing f for a permutation $\pi : [2n] \to [2n]$. Let $L_0, L_1, \ldots, L_n$ be the levels of B. Consider any partition A that assigns all n y-variables to $\{x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}\}$ and all n z-variables to $\{x_{\pi(n+1)}, x_{\pi(n+2)}, \ldots, x_{\pi(2n)}\}$. Let us denote by f^A the polynomial obtained from f after substituting each variable x by $A(x)$. Let B is partitioned with respect to the level L_i for $i = 2\alpha n$. We can write, $f = f_{st} = \sum_{v \in L_i} f_{s,v} f_{v,t}$. Consider a node $v \in L_i$. By definition, there are following two cases:
Case 1: $X_{s,v} \subseteq \{x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}\}$ and $|X_{v,t}| \leq 2n(1-\alpha)$. Thus, $f^A_{s,v} \in \mathbb{F}[Y]$. Hence, using Proposition 1.4 and 1.1,
max-rank$(M_{f^A_{s,v} f^A_{v,t}}) \leq$ max-rank$(M_{f^A_{v,t}}) \leq 2^{|X_{v,t}|/2} \leq 2^{n(1-\alpha)}$
Case 2: $X_{v,t} \subseteq \{x_{\pi(n+1)}, x_{\pi(n+2)}, \ldots, x_{\pi(2n)}\}$ and $|X_{s,v}| \leq 2n(1-\alpha)$. Thus, $f^A_{v,t} \in \mathbb{F}[Z]$. Hence, again using Proposition 1.4 and 1.1,
max-rank$(M_{f^A_{s,v} f^A_{v,t}}) \leq$ max-rank$(M_{f^A_{s,v}}) \leq 2^{|X_{s,v}|/2} \leq 2^{n(1-\alpha)}$ Thus, in any case, max-rank$(M_{f^A_{s,v} f^A_{v,t}}) \leq 2^{n(1-\alpha)}$ for all $v \in L_i$. Using Proposition 1.2, max-rank$(M_{f^A}) \leq |L_i| \cdot 2^{n(1-\alpha)}$. Since f is a full max-rank polynomial, we get $|L_i| \geq 2^{\alpha n}$.

References

1. Agrawal, M., Saha, C., Saptharishi, R., Saxena, N.: Jacobian Hits Circuits: Hitting-sets, Lower Bounds for Depth-d Occur-k Formulas & Depth-3 Transcendence Degree-k Circuits. In: Proceedings of ACM Symposium on Theory of Computing (STOC), pp. 599–614 (2012)
2. Agrawal, M., Vinay, V.: Arithmetic Circuits: A Chasm at Depth Four. In: Proceedings of Symposium on Foundations of Computer Science (FOCS), pp. 67–75 (2008)
3. Dvir, Z., Malod, G., Perifel, S., Yehudayoff, A.: Separating Multilinear Branching Programs and Formulas. In: Proceedings of Symposium on Theory of Computing (STOC), pp. 615–624 (2012)
4. Grigoriev, D., Razborov, A.: Exponential Complexity Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields. In: Proceedings of Symposium on Foundations of Computer Science (FOCS), pp. 269–278 (1998)
5. Grigoriev, D., Karpinski, M.: An Exponential Lower Bound for Depth 3 Arithmetic Circuits. In: Proceedings of Symposium on Theory of Computing (STOC), pp. 577–582 (1998)
6. Gupta, A., Kamath, P., Kayal, N., Saptharishi, R.: Approaching the chasm at depth four. To Appear in Conference on Computational Complexity (2013)
7. Jansen, M.J.: Lower Bounds for Syntactically Multilinear Algebraic Branching Programs. In: Ochmański, E., Tyszkiewicz, J. (eds.) MFCS 2008. LNCS, vol. 5162, pp. 407–418. Springer, Heidelberg (2008)
8. Koiran, P.: Arithmetic Circuits: The Chasm at Depth Four Gets Wider. Theor. Comput. Sci. 448, 56–65 (2012)
9. Nisan, N., Wigderson, A.: Lower Bounds on Arithmetic Circuits via Partial Derivatives. In: Proceedings of Symposium on Foundations of Computer Science (FOCS), pp. 16–25 (1995)
10. Raz, R., Yehudayoff, A.: Lower Bounds and Separations for Constant Depth Multilinear Circuits. In: Proceedings of Conference on Computational Complexity, pp. 128–139 (June 2008)
11. Raz, R.: Separation of Multilinear Circuit and Formula Size. Theory of Computing 2(1), 121–135 (2006)
12. Raz, R.: Multi-linear Formulas for Permanent and Determinant are of Super-polynomial Size. Journal of ACM 56, 8:1–8:17 (2009)
13. Raz, R., Shpilka, A., Yehudayoff, A.: A Lower Bound for the Size of Syntactically Multilinear Arithmetic Circuits. SIAM Journal of Computing 38(4), 1624–1647 (2008)
14. Saxena, N.: Diagonal Circuit Identity Testing and Lower Bounds. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 60–71. Springer, Heidelberg (2008)
15. Shpilka, A.: Affine Projections of Symmetric Polynomials. In: Proceedings of Conference on Computational Complexity, pp. 160–171 (2001)
16. Shpilka, A., Wigderson, A.: Depth-3 Arithmetic Circuits over Fields of Characteristic Zero. Computational Complexity 10(1), 1–27 (2001)
17. Shpilka, A., Yehudayoff, A.: Arithmetic Circuits: A Survey of Recent Results and Open Questions. Foundations and Trends in Theoretical Computer Science 5(3-4), 207–388 (2010)