Balancing Syntactically Multilinear Arithmetic Circuits

Ran Raz [*] Amir Yehudayoff [†]

Abstract

In their seminal paper, Valiant, Skyum, Berkowitz and Rackoff proved that arithmetic circuits can be balanced [VSBR]. That is, [VSBR] showed that for every arithmetic circuit Φ of size s and degree r, there exists an arithmetic circuit Ψ of size $\mathrm{poly}(r, s)$ and depth $O(\log(r)\log(s))$ computing the same polynomial. In the first part of this paper, we follow the proof of [VSBR] and show that syntactically multilinear arithmetic circuits can be balanced. That is, we show that if Φ is syntactically multilinear, then so is Ψ.

Recently, [R04b] proved a super-polynomial separation between multilinear arithmetic formula and circuit size. In the second part of this paper, we use the result of the first part to simplify the proof of this separation. That is, we construct a (simpler) polynomial $f(x_1, \ldots, x_n)$ such that

- Every multilinear arithmetic formula computing f is of size $n^{\Omega(\log(n))}$.
- There exists a syntactically multilinear arithmetic circuit of size $\mathrm{poly}(n)$ and depth $O(\log^2(n))$ computing f.

1 Introduction

Arithmetic circuits use sums and products to compute multivariate polynomials. The size of an arithmetic circuit is the number of operations it uses. The depth of an arithmetic circuit can be thought of as the

[*] Faculty of Mathematics and Computer Science, Weizmann Institute, Rehovot, Israel. Email: ran.raz@weizmann.ac.il. Research supported by grants from the Binational Science Foundation (BSF), the Israel Science Foundation (ISF), and the Minerva Foundation.

[†] Faculty of Mathematics and Computer Science, Weizmann Institute, Rehovot, Israel. Email: amir.yehudayoff@weizmann.ac.il. Research supported by grants from the Binational Science Foundation (BSF), the Israel Science Foundation (ISF), the Minerva Foundation, and the Israel Ministry of Science (IMOS) - Eshkol Fellowship.

parallel time it takes for the circuit to complete its computation. The degree of an arithmetic circuit is the maximal degree of a polynomial computed by it. Surprisingly, Valiant, Skyum, Berkowitz and Rackoff showed that the computation of polynomial-size polynomial-degree arithmetic circuits can be done in short parallel time using a polynomial number of operations [VSBR]. More precisely, they showed that for every arithmetic circuit of size s and degree r, there exists an arithmetic circuit of size $\text{poly}(r, s)$ and depth $O(\log(r)\log(s))$ computing the same polynomial.

Previously to [VSBR], Hyafil [H] showed that for every polynomial-size polynomial-degree arithmetic circuit, there exists an arithmetic formula of quasi-polynomial size computing the same polynomial. It remains an outstanding open problem to decide whether for every polynomial-size polynomial-degree arithmetic circuit, there exists a polynomial-size arithmetic formula computing the same polynomial. Recently, [R04B] solved this problem for the multilinear case. [R04B] showed a super-polynomial separation between the size of multilinear arithmetic formulas and multilinear arithmetic circuits.

In this paper, we follow the proof of [VSBR] and show that syntactically multilinear computations can be done in short parallel time using a small number of operations (without violating the syntactical multilinearity condition). That is, we show that for every syntactically multilinear arithmetic circuit of size s and degree r, there exists a syntactically multilinear arithmetic circuit of size $\text{poly}(s)$ and depth $O(\log(r)\log(s))$ computing the same polynomial. We use this to simplify the proof of the super-polynomial separation between the size of multilinear formula and circuit size.

1.1 Arithmetic Circuits and Formulas

An *arithmetic circuit* Φ over the field $\mathbb{F}$ and the set of variables X is a directed acyclic graph as follows: Every vertex v in Φ is either of in-degree 0 or of in-degree 2. Every vertex v of in-degree 0 is labelled by either a variable in X or a field element in $\mathbb{F}$. Every vertex v of in-degree 2 is labelled by either $\times$ or $+$. An arithmetic circuit Φ is called an *arithmetic formula*, if Φ is a directed binary tree (the edges of an arithmetic formula are directed from the leaves to the root).

The vertices of an arithmetic circuit Φ are also called *gates*. Every gate of in-degree 0 is called an *input gate*. Every gate of in-degree 2 labelled by $\times$ is called a *product gate*. Every gate of in-degree 2 labelled by $+$ is called a *sum gate*. Every gate of out-degree 0 is called an *output gate*. For two gates u and v, if (u, v) is an edge in Φ, then u is called a *son* of v. The *size* of Φ, denoted $|\Phi|$, is the number of gates in Φ. The *depth* of the gate v is the maximal length of a directed path reaching v. The *depth* of Φ is the maximal depth of a gate in Φ.

For a gate v, define Φ_v to be the sub-circuit of Φ rooted at v. That is, the gates of Φ_v are all the gates u in Φ such that there exists a directed path from u to v in Φ, and the edges and labels of Φ_v are the same edges and labels of Φ (restricted to the set of gates of Φ_v). We say that a variable $x \in X$ occurs in Φ_v, if x labels one of the input gates of Φ_v. Define X_v to be the set of variables that occur in Φ_v.

An arithmetic circuit computes a polynomial in a natural way. For a gate v in Φ, define $\widehat{\Phi}_v \in \mathbb{F}[X]$ to be the polynomial computed by Φ_v as follows: If v is an input gate labelled by $\alpha \in \mathbb{F} \cup X$, then $\widehat{\Phi}_v = \alpha$. If v is a product gate with sons v_1 and v_2, then $\widehat{\Phi}_v = \widehat{\Phi}_{v_1} \cdot \widehat{\Phi}_{v_2}$. If v is a sum gate with sons v_1 and v_2, then $\widehat{\Phi}_v = \widehat{\Phi}_{v_1} + \widehat{\Phi}_{v_2}$. For a polynomial $f \in \mathbb{F}[X]$ and a gate v in Φ, we say that v *computes* f, if $f = \widehat{\Phi}_v$. We say that Φ *computes* f, if one of the output gates of Φ computes f.

For a gate v, define the *degree* of v to be the total degree of the polynomial $\widehat{\Phi}_v$. We denote the degree of v by $\deg(v)$. The *degree* of Φ is the maximal degree of a gate in Φ.

A polynomial $f \in \mathbb{F}[X]$ is called *multilinear*, if the degree of each variable in f is at most one. An arithmetic circuit Φ is called *multilinear*, if every gate in Φ computes a multilinear polynomial. Φ is called *syntactically multilinear*, if every product gate v in Φ with sons v_1 and v_2 admits $X_{v_1} \cap X_{v_2} = \emptyset$.

1.2 Background

The model of multilinear arithmetic circuits was first considered in [NW]. It is a restricted model, as it does not allow the use of high powers of variables during a computation. However, it is a natural model for computing multilinear polynomials, as the use of high powers of variables in order to compute multilinear polynomials is often not intuitive. Furthermore, the best known circuits for many multilinear polynomials are indeed multilinear.

In [R04a] it is shown that every multilinear formula computing the determinant or the permanent of an n by n matrix is of size $n^{\Omega(\log(n))}$. Later [R04b] used the techniques of [R04a] to prove a super-polynomial separation between multilinear formula and circuit size. Recently, [RSY] showed an $\Omega(n^{1+\varepsilon})$ lower bound (for some $\varepsilon > 0$) for the size of a syntactically multilinear arithmetic circuit (where n is the number of variables).

1.3 Results

Our first result is a syntactically multilinear version of [VSBR]. We show that for every syntactically multilinear arithmetic circuit of size s and degree r over the field $\mathbb{F}$ and over the set of variables X

computing the polynomial f, there exists a syntactically multilinear arithmetic circuit of size poly(s) and depth $O(\log(r)\log(s))$ over the field $\mathbb{F}$ and over the set of variables X computing f as well (see Theorem 3.1). Basically, our proof is the same as the proof of [VSBR]. We note that the proof does not work for the multilinear model (the proof uses *syntactical* multilinearity).

Our second result is a simpler proof of the following result of [R04B] (see Theorem 4.4): There exists a multilinear polynomial $f(x_1, \ldots, x_n)$ such that (over every field)

1. Every multilinear arithmetic formula computing f is of size $n^{\Omega(\log(n))}$.
2. There exists a syntactically multilinear arithmetic circuit of size poly(n) and depth $O(\log^2(n))$ computing f.

To prove this super-polynomial separation between multilinear formula and circuit size, we use the techniques of [R04A, R04B] and our first result, together with a new construction of f.

1.4 Discussion

Our first result implies that for every syntactically multilinear arithmetic circuit of size poly(n), there exists a syntactically multilinear arithmetic formula of size $n^{O(\log(n))}$ computing the same polynomial. Thus, a proof of an $n^{\omega(\log(n))}$ lower bound for the size of multilinear arithmetic formulas automatically gives a super-polynomial lower bound for the size of syntactically multilinear arithmetic *circuits*. So, no better lower bounds than that of [R04A, R04B] are available without a proof of a super-polynomial lower bound for the size of syntactically multilinear arithmetic circuits. We also recall a *conjecture* from [NW]: Multilinear arithmetic circuits require $\Omega(n)$ depth to compute the determinant of an n by n matrix. A proof of the conjecture automatically gives an exponential lower bound for the *size* of syntactically multilinear arithmetic circuits for the determinant.

As mentioned above, the proof of our first result is, basically, the same as the proof of [VSBR]. We will now discuss the proof of our second result (note that the following statements are not accurate). We find a polynomial f with properties 1. and 2. above. The proof has two parts:

Proof of property 1. A polynomial is of *full rank*, if the partial derivative matrix of the polynomial is of full rank (see Section 4.2.2 for definitions). [R04B] shows that small multilinear arithmetic formulas can't compute polynomials of full rank. So, to prove property 1., we show that f is of full rank.

Proof of property 2. We find a multilinear arithmetic circuit of polynomial size and linear depth computing f. Using our first result, property 2. follows. We note that [R04B] constructed a polynomial of full rank that is computed explicitly by a multilinear arithmetic circuit of size poly(n) and depth $O(\log^2(n))$. Much of the complexity of [R04B] comes from the explicitness of the small depth.

2 Preliminaries

In the following $\mathbb{G}$ denotes a field, and $X = \{x_1, \ldots, x_n\}$ denotes a set of variables. $\mathbb{G}[X]$ denotes the ring of polynomials over the field $\mathbb{G}$ and over the set of variables X. Arithmetic circuits are denoted by either Φ or Ψ. For an integer $n \in \mathbb{N}$, denote $[n] = \{1, \ldots, n\}$. For two integers $i \in \mathbb{N}$ and $j \in \mathbb{N}$, denote

$$[i,j] = \{k \in \mathbb{N} \ : \ i \leq k \text{ and } k \leq j\}.$$

We say that a variable *occurs* in a polynomial f, if the degree of the variable in f is greater than 0. We denote by $\mathbf{V}(f)$ the set of variables that occur in f. We say that a monomial *occurs* in f, if the coefficient of the monomial in f is non-zero.

2.1 Homogeneous Arithmetic Circuits

A polynomial f is called *homogeneous*, if all the monomials that occur in f have the same total degree. For an integer $i \in \mathbb{N}$, we define the *homogeneous part of degree i* of f to be the restriction of f to the set of monomials of total degree i. We say that an arithmetic circuit Φ is *homogeneous*, if for every gate v in Φ, the polynomial $\widehat{\Phi}_v$ is homogeneous.

The following theorem shows that syntactically multilinear arithmetic circuits can be made homogeneous. The theorem is well known for general arithmetic circuits (e.g., [NW]). Here we show that the theorem applies for the syntactically multilinear case as well.

Theorem 2.1. *Let Φ be a syntactically multilinear arithmetic circuit of size s and degree r over the field $\mathbb{G}$ and over the set of variables X computing the polynomial f. Then there exists a syntactically multilinear homogeneous arithmetic circuit Ψ of size $O(r^2 s)$ and degree r over the field $\mathbb{G}$ and over the set of variables X computing f as well.*

Proof. For every gate v in Φ and for every $i \in [0, r]$, define the pair (v, i) as a gate in Ψ. The gate (v, i) will compute the homogeneous part of degree i of the polynomial $\widehat{\Phi}_v$. We will construct Ψ by induction on the structure of Φ.

Let v be a gate in Φ.

If v is an input gate, then $\widehat{\Phi}_v$ is a homogeneous polynomial. So for all $i \in [0, r]$, define

$$\Psi_{(v,i)} = \begin{cases} \Phi_v & \deg(v) = i \\ 0 & \deg(v) \neq i. \end{cases}$$

Otherwise, let v_1 and v_2 be the two sons of v in Φ.

If v is a product gate, then for every $i \in [0, r]$, define

$$\Psi_{(v,i)} = \sum_{j \in [0,i]} \Psi_{(v_1,j)} \times \Psi_{(v_2,i-j)}.$$

If v is a sum gate, then for every $i \in [0, r]$, define

$$\Psi_{(v,i)} = \Psi_{(v_1,i)} + \Psi_{(v_2,i)}.$$

It follows by induction that for every gate v in Φ and for every $i \in [0, r]$, the gate (v, i) in Ψ computes the homogeneous part of degree i of $\widehat{\Phi}_v$. So, Ψ is homogeneous. Furthermore, it follows that

$$X_{(v,i)} \subseteq X_v,$$

where $X_{(v,i)}$ is the set of variables that occur in $\Psi_{(v,i)}$, and X_v is the set of variables that occur in Φ_v.

Since each gate in Φ adds at most $O(r^2)$ vertices to Ψ, the size of Ψ is $O(r^2 s)$. Also the degree of Ψ is at most r.

Finally, we claim that Ψ is syntactically multilinear. Indeed, let v' be a product gate in Ψ. So, $\Psi_{v'}$ is of the form

$$\Psi_{v'} = \Psi_{(v_1,j)} \times \Psi_{(v_2,i-j)},$$

where v_1 and v_2 are the two sons of a product gate v in Φ, and $i, j \in [0, r]$. Since v is a product gate in Φ and since Φ is syntactically multilinear, $X_{v_1} \cap X_{v_2} = \emptyset$. So, since $X_{(v_1,j)} \subseteq X_{v_1}$ and $X_{(v_2,i-j)} \subseteq X_{v_2}$, we have

$$X_{(v_1,j)} \cap X_{(v_2,i-j)} = \emptyset.$$

□

3 Balancing Arithmetic Circuits

In this section we follow the proof of [VSBR], and show how to balance syntactically multilinear arithmetic circuits. In particular, we show that poly(n)-size syntactically multilinear arithmetic circuits are without loss of generality of depth $O(\log^2(n))$ (where n is the number of variables).

Theorem 3.1. *Let Φ be a syntactically multilinear arithmetic circuit of size s and degree r over the field $\mathbb{G}$ and over the set of variables $X = \{x_1, \ldots, x_n\}$ computing the polynomial f. Then there exists a syntactically multilinear arithmetic circuit Ψ of size $O(r^6 s^3)$, of depth $O(\log(r)\log(s))$, and of degree r over the field $\mathbb{G}$ and over the set of variables X computing f as well.*

We defer the proof of Theorem 3.1 to Section 3.3.

3.1 Preliminaries

In the following Φ denotes a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and over the set of variables $X = \{x_1, \ldots, x_n\}$. For a gate v in Φ, we denote

$$f_v = \widehat{\Phi}_v \in \mathbb{G}[X],$$

the polynomial computed by v in Φ.

3.1.1 Non-redundant Circuits

We say that Φ is *non-redundant*, if every gate v in Φ admits $f_v \neq 0$.

Claim 3.2. *Let Φ be a homogeneous non-redundant arithmetic circuit over the field $\mathbb{G}$ and over the set of variables X. Let v be a sum gate in Φ with sons v_1 and v_2. Then,*

$$\deg(v_1) = \deg(v_2) = \deg(v).$$

Proof. Since Φ is non-redundant, all the polynomials f_v, f_{v_1} and f_{v_2} are non-zero. Since Φ is homogeneous, all the polynomials f_v, f_{v_1} and f_{v_2} are homogeneous. So, since $f_v = f_{v_1} + f_{v_2}$, the claim follows. □

3.1.2 The Partial Derivative of f_v by w

For two gates v and w in Φ such that $X_w \neq \emptyset$, we define $\partial_w f_v$, the partial derivative of f_v by w, as follows: Substitute the gate w in Φ by a new variables y, and denote by f'_v the 'new' polynomial computed by v. So, f'_v is a polynomial in $\mathbb{G}[y, X]$. Since Φ is syntactically multilinear, and since $X_w \neq \emptyset$, f'_v is linear in y; that is,

$$f'_v = h_{w,v} \cdot y + g_{w,v},$$

where $h_{w,v}$ and $g_{w,v}$ are polynomials in $\mathbb{G}[X]$. So, $f_v = h_{w,v} \cdot f_w + g_{w,v}$, and we define

$$\partial_w f_v = h_{w,v} \left(= \frac{\partial f'_v}{\partial y} \right).$$

We shall use the following properties of $\partial_w f_v$.

Claim 3.3. *Let Φ be a syntactically multilinear homogeneous non-redundant arithmetic circuit over the field $\mathbb{G}$ and over the set of variables X. Let v and w be two gates in Φ such that $X_w \neq \emptyset$ and $\partial_w f_v \neq 0$. Then, $\partial_w f_v$ is a homogeneous polynomial over the set of variables $X_v \setminus X_w$ of total degree* $\deg(v) - \deg(w)$.

Proof. We will use the following claim, which follows by induction. Let Ψ be a homogeneous arithmetic circuit, and let Ψ' be the arithmetic circuit obtained from Ψ by substituting a gate in Ψ by 0. Then, for every gate u in Ψ' (u is also a gate in Ψ), the polynomial $\widehat{\Psi}'_u$ is either 0 or homogeneous of the same degree as $\widehat{\Psi}_u$.

Since Φ is homogeneous, the polynomials f_v and f_w are homogenous. Furthermore, since $g_{w,v}$ is the polynomial computed by v after substituting w by 0, it follows that $g_{w,v}$ is either 0 or homogenous of degree $\deg(v)$. Since Φ is non-redundant, $f_w \neq 0$. So, since $\partial_w f_v \cdot f_w = f_v - g_{w,v}$, and since $\partial_w f_v \neq 0$, we have that $\partial_w f_v$ is a homogenous polynomial of total degree $\deg(v) - \deg(w)$.

Let $x \in X$ be a variable that occurs in $\partial_w f_v$. By the definition of $\partial_w f_v$, there exists a product gate u in Φ with sons u_1 and u_2 such that (without loss of generality) $x \in X_{u_1}$ and w is in Φ_{u_2}. Note that $X_w \subseteq X_{u_2}$. Since Φ is syntactically multilinear, $X_{u_1} \cap X_{u_2} = \emptyset$. So, the set of variables that occur in $\partial_w f_v$ and X_w are disjoint. □

Claim 3.4. *Let Φ be a syntactically multilinear homogeneous non-redundant arithmetic circuit over the field $\mathbb{G}$ and over the set of variables X. Let v and w be two gates in Φ such that $X_w \neq \emptyset$, and*

$$\deg(v) < 2\deg(w).$$

Assume that v is a product gate with sons v_1 and v_2 such that $\deg(v_1) \geq \deg(v_2)$. Then,

$$\partial_w f_v = f_{v_2} \cdot \partial_w f_{v_1}.$$

Proof. Since v is a product gate, $\deg(v) = \deg(v_1) + \deg(v_2)$. So, since $\deg(v) < 2\deg(w)$, we have $\deg(v_2) < \deg(w)$, which implies (by Claim 3.3) that $\partial_w f_{v_2} = 0$. So, by the rules of partial derivatives,

$$\partial_w f_v = f_{v_1} \cdot \partial_w f_{v_2} + f_{v_2} \cdot \partial_w f_{v_1} = f_{v_2} \cdot \partial_w f_{v_1}.$$

□

3.2 Representing f_v and $\partial_w f_v$ Differently

In this section we prove two claims that are the main tool in the proof of Theorem 3.1.

For an integer $m \in \mathbb{N}$, denote by $\mathbf{G}_m$ the set of product gates t in Φ with sons t_1 and t_2 such that

$$m < \deg(t) \text{ and } \deg(t_1) \leq m \text{ and } \deg(t_2) \leq m.$$

For a gate v, the following claim shows how to represent f_v as a function of the gates in $\mathbf{G}_m$ (for the appropriate m).

Claim 3.5. *Let Φ be a syntactically multilinear homogeneous non-redundant arithmetic circuit over the field $\mathbb{G}$ and over the set of variables X. Let $m > 0$ be an integer, and let v be a gate in Φ such that $m < \deg(v) \leq 2m$. Then,*

$$f_v = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_v$$

(since $m > 0$, every $t \in \mathbf{G}_m$ admits $X_t \neq \emptyset$).

Proof. Since $m < \deg(v)$, there is a directed path from $\mathbf{G}_m$ to v in Φ. We prove the claim by induction on the length of the longest directed path from $\mathbf{G}_m$ to v. Let v_1 and v_2 be the two sons of v in Φ.

Induction Base: Assume that $v \in \mathbf{G}_m$. Thus, $\deg(v_1) \leq m$ and $\deg(v_2) \leq m$. So, every $t \in \mathbf{G}_m$ different than v is not in Φ_v, which implies $\partial_t f_v = 0$. Hence, since $\partial_v f_v = 1$,

$$f_v = f_v \cdot 1 + \sum_{t \in \mathbf{G}_m : t \neq v} f_t \cdot 0 = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_v.$$

Induction Step: Consider the following two cases:

Case one: Assume that v is a sum gate. By Claim 3.2,

$$\deg(v_1) = \deg(v_2) = \deg(v).$$

So, by induction,

$$f_{v_1} = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_{v_1} \text{ and } f_{v_2} = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_{v_2}.$$

Hence, by the rules of partial derivatives,

$$f_v = f_{v_1} + f_{v_2} = \sum_{t \in \mathbf{G}_m} f_t \cdot (\partial_t f_{v_1} + \partial_t f_{v_2}) = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_v.$$

Case two: Assume that v is a product gate, and assume without loss of generality that $\deg(v_1) \geq \deg(v_2)$. Since $v \notin \mathbf{G}_m$, we have $m < \deg(v_1) \leq 2m$. So, by induction,

$$f_{v_1} = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_{v_1}.$$

For all gates $t \in \mathbf{G}_m$, we have $\deg(v) \leq 2m < 2\deg(t)$. So, using Claim 3.4,

$$f_v = f_{v_1} \cdot f_{v_2} = \sum_{t \in \mathbf{G}_m} f_t \cdot (f_{v_2} \cdot \partial_t f_{v_1}) = \sum_{t \in \mathbf{G}_m} f_t \cdot \partial_t f_v.$$

□

For a gate v in Φ, the following claim shows how to represent $\partial_w f_v$ as a function of the gates in $\mathbf{G}_m$ (for the appropriate choice of w and m).

Claim 3.6. *Let Φ be a syntactically multilinear homogeneous non-redundant arithmetic circuit over the field $\mathbb{G}$ and over the set of variables X. Let $m > 0$ be an integer, and let w be a gate in Φ such that $X_w \neq \emptyset$ and $\deg(w) \leq m < 2\deg(w)$. Let v be a gate in Φ such that $m < \deg(v) < 2\deg(w)$. Then,*

$$\partial_w f_v = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_v$$

(since $m > 0$, every $t \in \mathbf{G}_m$ admits $X_t \neq \emptyset$).

Proof. Since $m < \deg(v)$, there is a directed path from $\mathbf{G}_m$ to v in Φ. We prove the claim by induction on the length of the longest directed path from $\mathbf{G}_m$ to v. Let v_1 and v_2 be the two sons of v in Φ.

Induction Base: Assume that $v \in \mathbf{G}_m$. Thus, $\deg(v_1) \leq m$ and $\deg(v_2) \leq m$. So, every gate t in $\mathbf{G}_m$ different than v is not in Φ_v, which implies $\partial_t f_v = 0$. Hence, since $\partial_v f_v = 1$,

$$\partial_w f_v = \partial_w f_v \cdot 1 + \sum_{t \in \mathbf{G}_m : t \neq v} \partial_w f_t \cdot 0 = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_v.$$

Induction Step: Consider the following two cases:

Case one: Assume that v is a sum gate. By Claim 3.2,

$$\deg(v_1) = \deg(v_2) = \deg(v).$$

So, by induction,

$$\partial_w f_{v_1} = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_{v_1} \text{ and } \partial_w f_{v_2} = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_{v_2}.$$

Hence, using the rules of partial derivatives,

$$\partial_w f_v = \partial_w f_{v_1} + \partial_w f_{v_2} = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot (\partial_t f_{v_1} + \partial_t f_{v_2}) = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_v.$$

Case two: Assume that v is a product gate, and assume without loss of generality that $\deg(v_1) \geq \deg(v_2)$. Since $v \notin \mathbf{G}_m$, we have $m < \deg(v_1) < 2\deg(w)$. So, by induction,

$$\partial_w f_{v_1} = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_{v_1}.$$

Note that $\deg(v) < 2\deg(w)$, and for all gates $t \in \mathbf{G}_m$, we have $\deg(v) < 2m < 2\deg(t)$. So, using Claim 3.4,

$$\partial_w f_v = f_{v_2} \cdot \partial_w f_{v_1} = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot (f_{v_2} \cdot \partial_t f_{v_1}) = \sum_{t \in \mathbf{G}_m} \partial_w f_t \cdot \partial_t f_v.$$

□

3.3 Proof of Theorem 3.1

First, we assume without loss of generality that $s \geq n$. Second, using Theorem 2.1, we assume without loss of generality that Φ is a homogeneous arithmetic circuit of size $s' = O(r^2 s)$. Third, we assume without loss of generality that Φ is non-redundant (otherwise, we can make Φ smaller). Note that Φ 'remains' syntactically multilinear.

To prove the theorem we construct Ψ. The construction is done in steps. For an integer $i \geq 0$, at the i'th step we have the following

1. We compute all polynomials f_v, for gates v in Φ such that $2^{i-1} < \deg(v) \leq 2^i$. To compute f_v we add $O(s')$ gates, and we increase the depth by $O(\log(s'))$.
2. We compute all polynomials $\partial_w f_v$, for gates v and w in Φ such that $X_w \neq 0$,
$$2^{i-1} < \deg(v) - \deg(w) \leq 2^i \quad \text{and} \quad \deg(v) < 2\deg(w).$$
To compute $\partial_w f_v$ we add $O(s')$ gates, and we increase the depth by $O(\log(s'))$.
3. All the product gates added 'multiply' disjoint sets of variables; that is, every product gate u in Ψ with sons u_1 and u_2 admits $X_{u_1} \cap X_{u_2} = \emptyset$.

Before describing the construction we show why it completes the proof. Since the degree of Φ is r, the total degree of f is at most r. So, we can end the process in $O(\log(r))$ steps. Since in each step we increase the depth by at most $O(\log(s'))$, the depth of Ψ is $O(\log(r)\log(s))$ (note that $r \leq n \leq s$). Since for each gate v in Φ, we add $O(s')$ gates to compute f_v, and since for each pair of gates v and w (with the appropriate properties), we add $O(s')$ gates to compute $\partial_w f_v$, it follows that the size of Ψ is $O(s'^3) = O(r^6 s^3)$. Finally, by 3., Ψ is syntactically multilinear.

The Construction of Ψ

We use the following conventions: Gates in Φ are denoted by v, t and w. For a gate v, we denote by v' the gate in Ψ computing f_v. For two gates v and w in Φ such that $X_w \neq \emptyset$ and $\deg(v) < 2\deg(w)$, we denote by (w, v) the gate in Ψ computing $\partial_w f_v$.

Throughout the process we maintain the following two properties:

A For every gate v in Φ,
$$X_{v'} \subseteq X_v,$$
where $X_{v'}$ is the set of variables that occur in $\Psi_{v'}$, and X_v is the set of variables that occur in Φ_v.

B For every two gates v and w in Φ such that $X_w \neq \emptyset$ and $\deg(v) < 2\deg(w)$,

$$X_{(w,v)} \subseteq X_v \setminus X_w,$$

where $X_{(w,v)}$ is the set of variables that occur in $\Psi_{(w,v)}$, X_v is the set of variables that occur in Φ_v, and X_w is the set of variables that occur in Φ_w.

Step 0: For every gate v in Φ such that $\deg(v) \leq 1$, the polynomial f_v is linear. So, since $s' \geq n$, we can compute f_v with an arithmetic circuit of size $O(s')$ and depth $O(\log(s'))$. Furthermore, we can have $X_{v'} \subseteq X_v$ (property **A**).

For every two gates v and w in Φ such that $X_w \neq \emptyset$ and $\deg(v) - \deg(w) \leq 1$, using Claim 3.3, we have that $\partial_w f_v$ is linear. So, we can compute $\partial_w f_v$ with an arithmetic circuit of size $O(s')$ and depth $O(\log(s'))$. Furthermore, by Claim 3.3, the set of variables that occur in $\partial_w f_v$ is a subset of $X_v \setminus X_w$, so we can have $X_{(w,v)} \subseteq X_v \setminus X_w$ (property **B**).

Step i+1: Assume that we already computed all polynomials f_v, for gates v such that $\deg(v) \leq 2^i$, and all polynomials $\partial_w f_v$, for gates v and w such that $X_w \neq \emptyset$, $\deg(v) - \deg(w) \leq 2^i$ and $\deg(v) \leq 2\deg(w)$. Furthermore, assume that properties **A** and **B** hold so far.

Recall that for an integer $m \in \mathbb{N}$, we defined $\mathbf{G}_m$ to be the set of product gates t in Φ with sons t_1 and t_2 such that

$$m < \deg(t) \ \text{ and } \ \deg(t_1) \leq m \text{ and } \ \deg(t_2) \leq m.$$

The $i+1$ step is done in two parts:

First part: computing f_v.

Let v be a gate of degree $2^i < \deg(v) \leq 2^{i+1}$, and denote

$$m = 2^i.$$

Recall that if a gate t is not in Φ_v, then $\partial_t f_v = 0$. Thus, by Claim 3.5,

$$f_v = \sum_{t \in T} f_t \cdot \partial_t f_v = \sum_{t \in T} f_{t_1} \cdot f_{t_2} \cdot \partial_t f_v,$$

where T is the set of gates $t \in \mathbf{G}_m$ with sons t_1 and t_2 such that t is in Φ_v.

Let $t \in T$ be a gate with sons t_1 and t_2. Thus, $m < \deg(t) \leq 2m$, $\deg(t_1) \leq m$ and $\deg(t_2) \leq m$. So, $\deg(v) - \deg(t) \leq 2^{i+1} - 2^i = 2^i$ and $\deg(v) \leq 2^{i+1} < 2\deg(t)$. Therefore, f_{t_1}, f_{t_2} and $\partial_t f_v$ are already computed. Thus, to compute f_v (using the polynomials computed so far) we add $O(s')$ gates, and we increase the depth by at most $O(\log(s'))$.

We will now show that property **A** still holds. Indeed, recall that

$$X_{t'_1} \subseteq X_{t_1} \text{ and } X_{t'_2} \subseteq X_{t_2} \text{ and } X_{(t,v)} \subseteq X_v \setminus X_t.$$

Since t is in Φ_v,

$$X_{t_1} \cup X_{t_2} = X_t \subseteq X_v.$$

Thus,

$$X_{v'} = \bigcup_{t \in T} X_{t'_1} \cup X_{t'_2} \cup X_{(t,v)} \subseteq X_v \text{ (property } \mathbf{A}\text{)}.$$

Finally, we claim that every product gate added 'multiplies' disjoint sets of variables. Indeed, a product gate added is of the form $f_{t_1} \cdot f_{t_2} \cdot \partial_t f_v$, where $t \in T$ is product a gate with sons t_1 and t_2. Since Φ is syntactically multilinear, $X_{t_1} \cap X_{t_2} = \emptyset$. So,

$$X_{t'_1} \cap X_{t'_2} = \emptyset \quad \text{and} \quad X_{(t,v)} \cap (X_{t'_1} \cup X_{t'_2}) = \emptyset.$$

Second part: computing $\partial_w f_v$.

Let v and w be two gates in Φ such that $X_w \neq \emptyset$,

$$2^i < \deg(v) - \deg(w) \leq 2^{i+1} \text{ and } \deg(v) < 2\deg(w).$$

Now, denote

$$m = 2^i + \deg(w).$$

Thus, $\deg(w) \leq m < \deg(v) < 2\deg(w)$. Recall that if a gate t is not in Φ_v, then $\partial_t f_v = 0$. Also, by Claim 3.3, if a gate t admits $\deg(t) > \deg(v)$, then $\partial_t f_v = 0$. Hence, by Claim 3.6,

$$\partial_w f_v = \sum_{t \in T} \partial_w f_t \cdot \partial_t f_v,$$

where T is the set of gates $t \in \mathbf{G}_m$ such that t is in Φ_v and $\deg(t) \leq \deg(v)$. For a gate $t \in T$, we have $\deg(t) \leq \deg(v) < 2\deg(w)$. Hence, using Claim 3.4, we have

$$\partial_w f_v = \sum_{t \in T} f_{t_2} \cdot \partial_w f_{t_1} \cdot \partial_t f_v,$$

where for all $t \in T$, we denote its sons by t_1 and t_2, where (without loss of generality) w is in Φ_{t_1}, $\deg(w) \leq \deg(t_1)$ and $\deg(t_1) \geq \deg(t_2)$.

Let $t \in T$ be a gate with sons t_1 and t_2. We will now show that all the polynomials f_{t_2}, $\partial_w f_{t_1}$ and $\partial_t f_v$ are already computed (including the first part of the $i+1$ step).

Since

$$\deg(v) \leq 2^{i+1} + \deg(w) \leq 2^{i+1} + \deg(t_1) = 2^{i+1} + \deg(t) - \deg(t_2),$$

we have

$$\deg(t_2) \leq 2^{i+1} + \deg(t) - \deg(v) \leq 2^{i+1}.$$

So, f_{t_2} is already computed (including in the first part of the $i+1$ step).

Since $\deg(t_1) \leq m = 2^i + \deg(w)$, we have $\deg(t_1) - \deg(w) \leq 2^i$. So, since $\deg(t_1) \leq \deg(t) \leq \deg(v) < 2\deg(w)$, the polynomial $\partial_w f_{t_1}$ is already computed.

Since $\deg(t) > m = 2^i + \deg(w)$, we have

$$\deg(v) - \deg(t) < \deg(v) - 2^i - \deg(w) \leq 2^{i+1} - 2^i = 2^i.$$

So, since

$$\deg(v) \leq 2^{i+1} + \deg(w) \leq 2(2^i + \deg(w)) < 2\deg(t),$$

the polynomial $\partial_t f_v$ is already computed.

Thus, to compute $\partial_w f_v$ (using the polynomials computed so far) we add $O(s')$ gates, and we increase the depth by $O(\log(s'))$.

We will now show that property **B** still holds. Let $t \in T$ be a gate with sons t_1 and t_2. Recall that

$$X_{t'_2} \subseteq X_{t_2} \text{ and } X_{(w,t_1)} \subseteq X_{t_1} \setminus X_w \text{ and } X_{(t,v)} \subseteq X_v \setminus X_t.$$

Since Φ is syntactically multilinear,

$$X_{t_1} \cap X_{t_2} = \emptyset.$$

Since t is in Φ_v,

$$X_{t_1} \cup X_{t_2} = X_t \subseteq X_v.$$

Since w is in Φ_{t_1},

$$X_w \subseteq X_{t_1}.$$

So,

$$X_{(w,v)} = \bigcup_{t \in T} X_{t'_2} \cup X_{(w,t_1)} \cup X_{(t,v)} \subseteq X_v \setminus X_w \text{ (property } \mathbf{B}\text{)}.$$

Finally, we claim that every product gate added 'multiplies' disjoint sets of variables. Indeed, a product gate added in the second part is of the form $f_{t_2} \cdot \partial_w f_{t_1} \cdot \partial_t f_v$, where $t \in T$ is a gate with sons t_1 and t_2. Thus,

$$X_{t'_2} \cap X_{(w,t_1)} = \emptyset \quad \text{and} \quad (X_{t'_2} \cup X_{(w,t_1)}) \cap X_{(t,v)} = \emptyset.$$

□

4 Separation Between Multilinear Formula and Circuit Size

In this section we show a super-polynomial separation between multilinear arithmetic formula and circuit size. More specifically, we give an explicit polynomial $f(x_1, \ldots, x_n)$ such that

1. Every multilinear arithmetic formula computing f is of size $n^{\Omega(\log(n))}$.
2. There exists a syntactically multilinear arithmetic circuit of $O(\log^2(n))$-depth and poly(n)-size computing f.

Such a separation was already proved in [R04b]. We use the fact that syntactically multilinear arithmetic circuits can be balanced in order to simplify the construction of f, and the proof.

4.1 The Definition of f

Let $n \in \mathbb{N}$ be an integer. Let $X = \{x_1, \ldots, x_{2n}\}$ and $\mathcal{W} = \{\omega_{i,\ell,j}\}_{i,\ell,j \in [2n]}$ be two sets of variables. Recall that for two integers $i \in \mathbb{N}$ and $j \in \mathbb{N}$, we denote $[i,j] = \{k \in \mathbb{N} \ : \ i \leq k \text{ and } k \leq j\}$. We call $[i,j]$ an *interval*. We call $|[i,j]|$ the *length* of the interval $[i,j]$. Note that for $j < i$, the interval $[i,j] = \emptyset$ is of even length. Denote by $X_{i,j}$ the set of variables x_m, where $m \in [i,j]$. Denote by $\mathcal{W}_{i,j}$ the set of variables $w_{i',\ell,j'}$, where $i', \ell, j' \in [i,j]$.

For every interval $[i,j] \subseteq [2n]$ of even length, we define a polynomial $f_{i,j} \in \mathbb{F}[X, \mathcal{W}]$ inductively as follows:

If the length of $[i,j]$ is 0, then define

$$f_{i,j} = 1.$$

If the length of $[i, j]$ is greater than 0, then define

$$f_{i,j} = (1 + x_i x_j) f_{i+1,j-1} + \sum_{\ell} \omega_{i,\ell,j} f_{i,\ell} f_{\ell+1,j},$$

where the sum is over $\ell \in [i+1, j-2]$ such that the interval $[i, \ell]$ is of even length (so, the length of $[\ell+1, j]$ is even as well). Since the lengths of the intervals $[i, \ell]$ and $[\ell+1, j]$ are even and smaller than the length of $[i, j]$, both $f_{i,\ell}$ and $f_{\ell+1,j}$ are already defined. Similarly, $f_{i+1,j-1}$ is already defined.

Finally, we define

$$f = f_{1,2n}.$$

Let $[i, j] \subseteq [2n]$ be an interval of even length. Recall that $\mathbf{V}(f_{i,j})$ is the set of X and $\mathcal{W}$ variables that occur in $f_{i,j}$. By induction, we have

$$\mathbf{V}(f_{i,j}) \subseteq X_{i,j} \cup \mathcal{W}_{i,j}. \tag{4.1}$$

So, by the definition of f, there exists a syntactically multilinear arithmetic circuit of size $\text{poly}(n)$ computing f. In particular, f is multilinear. We note that it will be useful to think of f also as a polynomial in $\mathbb{G}[X]$, where $\mathbb{G} = \mathbb{F}(\mathcal{W})$ is the field of rational functions over the field $\mathbb{F}$ and over the set of variables $\mathcal{W}$.

4.2 Preliminaries

4.2.1 Partitions of the Set of Variables

Let $X = \{x_1, \ldots, x_{2n}\}$, $Y = \{y_1, \ldots, y_n\}$ and $Z = \{z_1, \ldots, z_n\}$ be three sets of variables. We call a one-to-one mapping $A : X \to Y \cup Z$ a *partition* of X to Y and Z. When X, Y and Z are clear, we call A a partition. For a polynomial $g \in \mathbb{G}[X]$, we denote by g^A the polynomial g, after substituting each $x \in X$ by $A(x) \in Y \cup Z$. So, $g^A \in \mathbb{G}[Y, Z]$.

4.2.2 The Partial Derivative Matrix

Let $Y = \{y_1, \ldots, y_n\}$ and $Z = \{z_1, \ldots, z_n\}$ be two sets of variables. Let $g \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the two sets of variables Y and Z. A monomial whose coefficient is 1 is called a *monic* monomial. Define M_g, the *partial derivative matrix* of g, as follows: for p a monic multilinear monomial in Y and q a monic multilinear monomial in Z, define $M_g(p, q)$ to be the coefficient of the monomial $p \cdot q$ in g. Thus, the rows of M_g correspond to monic multilinear monomials in Y, and the columns of M_g correspond to monic multilinear monomials in Z. The size of M_g is $2^n \times 2^n$.

We say that the polynomial g is of *full rank*, if for every partition A of $X = \{x_1, \dots, x_{2n}\}$ to Y and Z, the rank of M_{g^A} is full.

[R04a, R04b] proved the following theorem, which shows that a polynomial-size multilinear formula can't compute a polynomial of full rank.

Theorem 4.1. *Let Φ be a multilinear arithmetic formula over the field $\mathbb{G}$ and over the set of variables $X = \{x_1, \dots, x_{2n}\}$ computing a polynomial g. If g is of full rank, then*

$$|\Phi| \geq n^{\Omega(\log(n))}.$$

Having Theorem 4.1 in mind, to prove a super-polynomial separation between multilinear formula and circuit size, it is enough to find a full rank polynomial that is computed by a polynomial-size multilinear arithmetic circuit.

4.3 f is of Full Rank

In this section we prove that f is of full rank.

Theorem 4.2. *Let $n \in \mathbb{N}$ be an integer. Let $X = \{x_1, \dots, x_{2n}\}$ and $\mathcal{W} = \{\omega_{i,\ell,j}\}_{i,\ell,j \in [2n]}$ be two sets of variables. Let $\mathbb{G} = \mathbb{F}(\mathcal{W})$ be the field of rational functions over the field $\mathbb{F}$ and over the set of variables $\mathcal{W}$. Let $f \in \mathbb{G}[X]$ be the polynomial defined in Section 4.1. Then f is of full rank (over the field $\mathbb{G}$).*

To prove Theorem 4.2 we shall need some definitions, and Lemma 4.3. Recall that a partition A of $X = \{x_1, \dots, x_{2n}\}$ to $Y = \{y_1, \dots, y_n\}$ and $Z = \{z_1, \dots, z_n\}$ is a one-to-one mapping from X to $Y \cup Z$. For a partition A and an interval $[i, j] \subseteq [2n]$, define

$$\mathcal{D}_{i,j}(A) = |A(X_{i,j}) \cap Y| - |A(X_{i,j}) \cap Z|\,.$$

We say that the interval $[i, j]$ is *balanced* on A, if $\mathcal{D}_{i,j}(A) = 0$.

The following lemma shows that for every interval $[i, j] \subseteq [2n]$ that is balanced on A, the polynomial $f^A_{i,j}$ is of full rank.

Lemma 4.3. *Let $n \in \mathbb{N}$ be an integer. Let $X = \{x_1, \ldots, x_{2n}\}$, $\mathcal{W} = \{\omega_{i,\ell,j}\}_{i,\ell,j\in[2n]}$, $Y = \{y_1, \ldots, y_n\}$ and $Z = \{z_1, \ldots, z_n\}$ be four sets of variables. Let A be a partition of X to Y and Z. Let $\mathbb{G} = \mathbb{F}(\mathcal{W})$ be the field of rational functions over the field $\mathbb{F}$ and over the set of variables $\mathcal{W}$. Let $m \in [0, n]$ and let $[i, j] \subseteq [2n]$ be an interval of length $2m$ that is balanced on A. Let $f_{i,j} \in \mathbb{G}[X]$ be the polynomial defined in Section 4.1. Then*

$$\mathit{Rank}(M_{f^A_{i,j}}) = 2^m,$$

where the rank is over the field $\mathbb{G}$.

Proof of Theorem 4.2. Let A be a partition of X to $Y = \{y_1, \ldots, y_n\}$ and $Z = \{z_1, \ldots, z_n\}$. Thus, the interval $[1, 2n]$ is balanced on A. So, by Lemma 4.3, the partial derivative matrix of f^A is of full rank. □

4.3.1 Proof of Lemma 4.3

The proof is by induction on m.

Induction Base: If $m = 0$, then $f_{i,j} = 1$, which implies

$$\text{Rank}(M_{f^A_{i,j}}) = 1 = 2^m.$$

Induction Step: Assume that $m > 0$ and that the lemma holds for every interval of length smaller than $2m$. Consider the following two cases:

Case one: For every $\ell \in [i+1, j-2]$ such that the interval $[i, \ell]$ is of even length, we have $\mathcal{D}_{i,\ell}(A) \neq 0$.

Note that if an interval is balanced on A, then the interval is of even length. Thus, for every $\ell \in [i+1, j-1]$, we have $\mathcal{D}_{i,\ell}(A) \neq 0$.

Assume without loss of generality that $\mathcal{D}_{i,i}(A) = 1$ (the case $\mathcal{D}_{i,i}(A) = -1$ is similar). For every $\ell \in [i+1, j-1]$, we have

$$\mathcal{D}_{i,\ell}(A) - \mathcal{D}_{i,\ell+1}(A) \in \{1, -1\}.$$

So for all $\ell \in [i+1, j-1]$, we have $\mathcal{D}_{i,\ell}(A) > 0$. Hence, since $\mathcal{D}_{i,j}(A) = 0$, we have $\mathcal{D}_{j,j}(A) = -1$. Therefore, the interval $[i+1, j-1]$ is balanced on A. Thus, by induction,

$$\text{Rank}(M_{f^A_{i+1,j-1}}) = 2^{m-1}.$$

Since $\mathcal{D}_{i,i} = 1$ and $\mathcal{D}_{j,j} = -1$,

$$\text{Rank}(M_{1+A(x_i)A(x_j)}) = 2.$$

Since, by (4.1), both x_i and x_j do not occur in $f_{i+1,j-1}$,

$$M_{(1+A(x_i)A(x_j))f^A_{i+1,j-1}} = M_{1+A(x_i)A(x_j)} \otimes M_{f^A_{i+1,j-1}}, \tag{4.2}$$

where we think of $M_{(1+A(x_i)A(x_j))f^A_{i+1,j-1}}$ as a $2^m \times 2^m$ matrix, we think of $M_{1+A(x_i)A(x_j)}$ as a 2×2 matrix, we think of $M_{f^A_{i+1,j-1}}$ as a $2^{m-1} \times 2^{m-1}$ matrix, and $\otimes$ denotes tensor product of matrices. Therefore,

$$\text{Rank}(M_{(1+A(x_i)A(x_j))f^A_{i+1,j-1}}) = \text{Rank}(M_{1+A(x_i)A(x_j)}) \cdot \text{Rank}(M_{f^A_{i+1,j-1}}) = 2^m.$$

Recall that by (4.1), for all $\ell \in [i+1, j-1]$, the variable $\omega_{i,\ell,j}$ does not occur in $f_{i+1,j-1}$. Thus, substituting $\omega_{i,\ell,j} = 0$, for all $\ell \in [i+1, j-1]$, in $f^A_{i,j}$, we have

$$f^A_{i,j}\Big|_{\omega_{i,\ell,j}=0 \ \forall \ell \in [i+1,j-1]} = (1 + A(x_i)A(x_j))f^A_{i+1,j-1},$$

which implies

$$\text{Rank}(M_{f^A_{i,j}}) \geq \text{Rank}(M_{(1+A(x_i)A(x_j))f^A_{i+1,j-1}}).$$

Therefore, since $\text{Rank}(M_{f^A_{i,j}}) \leq 2^m$,

$$\text{Rank}(M_{f^A_{i,j}}) = 2^m.$$

Case two: There exists $\ell' \in [i+1, j-2]$ such that the interval $[i, \ell']$ is of even length and $\mathcal{D}_{i,\ell'}(A) = 0$. The interval $[\ell'+1, j]$ is of even length as well. Furthermore,

$$D_{\ell'+1,j}(A) = D_{i,j}(A) - D_{i,\ell'}(A) = 0.$$

So, by induction,

$$\text{Rank}(M_{f^A_{i,\ell'}}) = 2^{|[i,\ell']|/2} \text{ and } \text{Rank}(M_{f^A_{\ell'+1,j}}) = 2^{|[\ell'+1,j]|/2}.$$

By (4.1), similarly to (4.2), we have

$$M_{f^A_{i,\ell'} f^A_{\ell'+1,j}} = M_{f^A_{i,\ell'}} \otimes M_{f^A_{\ell'+1,j}}.$$

So, since $|[i, \ell']| + |[\ell'+1, j]| = 2m$,

$$\text{Rank}(M_{f^A_{i,\ell'} f^A_{\ell'+1,j}}) = \text{Rank}(M_{f^A_{i,\ell'}}) \cdot \text{Rank}(M_{f^A_{\ell'+1,j}}) = 2^m.$$

Write $f^A_{i,j}$ as

$$f^A_{i,j} = \omega_{i,\ell',j} f^A_{i,\ell'} f^A_{\ell'+1,j} + \underbrace{\left((1 + A(x_i)A(x_j))f^A_{i+1,j-1} + \sum_{\ell \neq \ell'} \omega_{i,\ell,j} f^A_{i,\ell} f^A_{\ell+1,j}\right)}_{f'},$$

where, by (4.1), the variable $\omega_{i,\ell',j}$ does not occur in f'. So,

$$\mathrm{Rank}(M_{f^A_{i,j}}) \geq \mathrm{Rank}(M_{f^A_{i,\ell'} f^A_{\ell'+1,j}}).$$

Therefore, since $\mathrm{Rank}(M_{f^A_{i,j}}) \leq 2^m$,

$$\mathrm{Rank}(M_{f^A_{i,j}}) = 2^m.$$

□

4.4 Proof of Separation

In this section we prove the following theorem.

Theorem 4.4. *Let $n \in \mathbb{N}$ be an integer. Let $X = \{x_1, \ldots, x_{2n}\}$ and $\mathcal{W} = \{\omega_{i,\ell,j}\}_{i,\ell,j \in [2n]}$ be two sets of variables. Let f be the polynomial over the field $\mathbb{F}$ and the two sets of variables X and $\mathcal{W}$ defined in Section 4.1. Then*

1. *Let Φ be a multilinear arithmetic formula over the field $\mathbb{F}$ and the two sets of variables X and $\mathcal{W}$ computing f. Then*
$$|\Phi| \geq n^{\Omega(\log(n))}.$$

2. *There exists a syntactically multilinear arithmetic circuit over the field $\mathbb{F}$ and the two sets of variables X and $\mathcal{W}$ of depth $O(\log^2(n))$ and of size poly(n) computing f.*

Proof. We prove the two properties separately.

1. Denote by $\mathbb{G} = \mathbb{F}(\mathcal{W})$ the field of rational functions over the field $\mathbb{F}$ and the set of variables $\mathcal{W}$. We think of Φ as an arithmetic formula over the field $\mathbb{G}$ and over the set of variables X. By Theorem 4.2, it follows that f is of full rank. So, by Theorem 4.1,
$$|\Phi| \geq n^{\Omega(\log(n))}.$$

2. The definition of f gives a syntactically multilinear arithmetic circuit of size $\mathrm{poly}(n)$ computing f. So, by Theorem 3.1, there exists a syntactically multilinear arithmetic circuit of depth $O(\log^2(n))$ and size $\mathrm{poly}(n)$ computing f.

□

References

[H] L. Hyafil. On the Parallel Evaluation of Multivariate Polynomials. *SIAM J. Comput.* 8(2): 120-123, 1979.

[N] N. Nisan. Lower Bounds for Non-Commutative Computation. *Proceeding of the 23th STOC*: 410-418, 1991.

[NW] N. Nisan and A. Wigderson. Lower Bounds on Arithmetic Circuits via Partial Derivatives. *Computational Complexity*, 6: 217-234, 1996 (preliminary version in Proceeding of the 36th FOCS 1995).

[R04a] R. Raz. Multi-Linear Formulas for Permanent and Determinant are of Super-Polynomial Size. *Proceeding of the 36th STOC*: 633-641, 2004.

[R04b] R. Raz. Separation of Multilinear Circuit and Formula Size. *Theory Of Computing*, 2: article 6, 2006. (preliminary version in Proceeding of the 45th FOCS: 344-351, 2004 (title: "Multilinear-NC_1 $\neq$ Multilinear-NC_2")).

[RSY] R. Raz, A. Shpilka and A. Yehudayoff. A Lower Bound for the Size of Syntactically Multilinear Arithmetic Circuits. *ECCC Report* TR06-060.

[VSBR] L. G. Valiant, S. Skyum, S. Berkowitz, C. Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM J. Comput. 12(4)*: 641-644 (1983)