A Lower Bound for the Size of Syntactically Multilinear Arithmetic Circuits

Ran Raz [*] Amir Shpilka [†] Amir Yehudayoff [‡]

Abstract

We construct an explicit polynomial $f(x_1, \ldots, x_n)$, with coefficients in $\{0, 1\}$, such that the size of any syntactically multilinear arithmetic circuit computing f is at least $\Omega(n^{4/3}/\log^2 n)$. The lower bound holds over any field.

1 Introduction

Arithmetic circuits are the standard model for computing polynomials (see Section 1.1 for definition). Roughly speaking, given a set of variables $X = \{x_1, \ldots, x_n\}$, an arithmetic circuit uses additions and multiplications to compute a polynomial f in the set of variables X. Given a polynomial f, we are interested in the number of operations needed to compute f.

The best lower bound known for the size of arithmetic circuits is the classical $\Omega(n \log n)$ of Strassen [S73], and of Baur and Strassen [BS83]. Proving better lower bounds is an outstanding open problem. Better lower bounds are not known even for arithmetic circuits of depth 4 (over fields of characteristic different than 2). In this paper, we focus on a restricted class of arithmetic circuits, the class of syntactically

[*]Faculty of Mathematics and Computer Science, Weizmann Institute, Rehovot, Israel, and Microsoft Research. 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 Computer Science, Technion, Israel. Email: shpilka@cs.technion.ac.il.

[‡]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), and the Minerva Foundation.

multilinear arithmetic circuits. We prove an $\Omega(n^{4/3}/\log^2 n)$ lower bound for the size of syntactically multilinear arithmetic circuits computing an explicit polynomial.

1.1 Syntactically Multilinear Arithmetic Circuits

An *arithmetic circuit* Φ over the field $\mathbb{F}$ and the set of variables $X = \{x_1, \ldots, x_n\}$ 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 in Φ of in-degree 0 is labelled by either a variable in X or a field element in $\mathbb{F}$. Every vertex v in Φ 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).

Let Φ be an arithmetic circuit over the field $\mathbb{F}$ and the set of variables X. The vertices of 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 an *addition gate*. Every gate of out-degree 0 is called an *output gate*. For two gates u and v in Φ, if (u, v) is an edge in Φ, then u is called a *son* of v, and v is called a *father* of u. The *size* of Φ, denoted $|\Phi|$, is the number of edges in Φ. Since the in-degree of Φ is at most 2, the size of Φ is at most twice the number of gates in Φ.

For a gate v in Φ, define Φ_v to be the sub-circuit of Φ rooted at v as follows: 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. The edges and labels of Φ_v are the same edges and labels of Φ (restricted to the set of gates of 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 an addition 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 $f = \widehat{\Phi}_v$. For a polynomial $f \in \mathbb{F}[X]$, we say that Φ *computes* f if there exists a gate u in Φ that computes f.

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 *(semantically) multilinear* if every gate in Φ computes a multilinear polynomial. An arithmetic circuit Φ is called *syntactically multilinear* if for every product gate v in Φ with sons v_1 and v_2, the set of variables that occur in Φ_{v_1} and the set of variables that occur in Φ_{v_2} are disjoint.

1.2 Background and Motivation

There are two ways to define multilinear arithmetic circuits: a syntactic definition, and a semantic definition, as described above. The semantic definition is a natural one, but the syntactic definition is

more convenient to work with. Note, for example, that given an arithmetic circuit Φ, deciding whether Φ is syntactically multilinear is straightforward, whereas it is not clear if one can decide whether Φ is semantically multilinear in deterministic polynomial time. We note also that similar distinction between semantic and syntactic definitions occur in other places in computer science (e.g., read k-times branching programs).

Multilinear arithmetic circuits were defined by Nisan and Wigderson in [NW96]. The model of syntactically multilinear arithmetic formulas was defined in [R04A]. In [R04A], it is shown that any multilinear arithmetic *formula* computing the determinant (or the permanent) of an $n \times n$ matrix must be of size $n^{\Omega(\log n)}$. Prior to our work, no lower bounds (better than the $\Omega(n \log n)$ lower bound of Strassen, and of Baur and Strassen) for the size of syntactically multilinear arithmetic *circuits* were known.

The techniques of [R04A] for proving super-polynomial lower bounds for the size of multilinear arithmetic formulas fail for circuits. In fact, [R04B] used these techniques to prove that syntactically multilinear arithmetic circuits are super-polynomially more powerful than multilinear arithmetic formulas. More specifically, there exists a polynomial f such that every multilinear arithmetic formula computing f is of size $n^{\Omega(\log n)}$, and on the other hand, there exists a polynomial size syntactically multilinear arithmetic circuit computing f.

Every multilinear polynomial f can be computed by a syntactically multilinear arithmetic circuit Φ, but Φ might not be the smallest arithmetic circuit computing f. However, computing a multilinear polynomial by an arithmetic circuit that is *not* syntactically multilinear is usually less intuitive, as cancellations of monomials are needed.

A syntactically multilinear arithmetic circuit is semantically multilinear as well. However, it is still not known whether there is an efficient way to transform a semantically multilinear arithmetic circuit to a syntactically multilinear circuit. We note that a semantically multilinear arithmetic *formula* can be transformed *without changing its size* to a syntactically multilinear arithmetic formula that computes the same polynomial (see Section 2 in [R04A]). We do not know of any significant example of a semantically multilinear arithmetic circuit that is *not* syntactically multilinear.

Finally, we note that two known classes of arithmetic circuits are contained in the class of syntactically multilinear arithmetic circuits: Pure arithmetic circuits (as defined by Nisan and Wigderson in [NW96], see also [RS05]) are a restricted type of syntactically multilinear arithmetic circuits. Monotone arithmetic circuits computing a multilinear polynomial are also syntactically multilinear (see [NW96]).

1.3 Results and Methods

Our main result is a construction of an explicit polynomial f such that any syntactically multilinear arithmetic circuit computing f is of size $\Omega(n^{4/3}/\log^2 n)$. Formally,

Theorem 1.1. *Let $f \in \mathbb{F}[X, \Omega]$ be the polynomial defined in Section 6, where $\mathbb{F}$ is any field, and X and Ω are two sets of variables of size n each. Let Φ be a syntactically multilinear arithmetic circuit over the field $\mathbb{F}$ and the sets of variables X and Ω computing f . Then,*

$$|\Phi| = \Omega\left(\frac{n^{4/3}}{\log^2 n}\right).$$

The paper has three main parts: Section 3 investigates the method of Baur and Strassen for computing all partial derivatives of a polynomial. Section 5, which is the heart of our proof, gives a simple characterization of a polynomial for which our lower bound applies. Section 6 constructs a polynomial for which the lower bound applies.

In [BS83], Baur and Strassen showed that given an arithmetic circuit Ψ computing a polynomial $f \in \mathbb{F}[X]$, there exists an arithmetic circuit Ψ' computing all the partial derivatives of f, such that $|\Psi'| = O(|\Psi|)$. In Section 3, we apply the method of Baur and Strassen for syntactically multilinear arithmetic circuits. We show that if Ψ is syntactically multilinear, then Ψ' is syntactically multilinear as well. Furthermore, every variable $x \in X$ does not occur in the computation of $\frac{\partial f}{\partial x}$ in Ψ'.

In Section 5, we use the results of Section 3, to show that the rank of the partial derivative matrix of a polynomial computed by a "small" syntactically multilinear arithmetic circuit is not full (see Theorem 5.1). We use techniques that were previously used in [R04A] and [R04B] together with some new ideas. In particular, we use the partial derivative method of Nisan and Wigderson, and the partial derivative matrix of Nisan. We mainly study the rank of the partial derivative matrix. We also use the notion of unbalanced gates, and the notion of partitions of the variables.

In Section 6, we construct a multilinear polynomial f such that the rank of the partial derivative matrix of f is full. As in [R04B], to show that the partial derivative matrix of f has full rank, we think of f as a polynomial over some extension field. We also show that f is explicit in the sense that f is in the class VNP, which is Valiant's algebraic analogue of NP (see Section 6.3 for more details).

Our lower bound follows from Sections 5 and 6: since the rank of the partial derivative matrix of a polynomial computed by a "small" syntactically multilinear arithmetic circuit is not full, and since the rank of the partial derivative matrix of f, the polynomial defined in Section 6, is full, it follows that any syntactically multilinear arithmetic circuit computing f is "large".

2 Preliminaries

2.1 Notation

For an integer $n \in \mathbb{N}$, denote $[n] = \{1, \ldots, n\}$. For a set $B \subseteq [n]$, denote $\overline{B} = [n] \setminus B$, the complement set of B. Similarly, for a set of variables X, and a set $X' \subseteq X$, denote $\overline{X'} = X \setminus X'$, the complement set of X'. For a polynomial f in the set of variables X, and a variable $x \in X$, denote by $d_x(f)$ the degree of x in f. We say that x *occurs* in f if $d_x(f) > 0$.

2.2 Arithmetic Circuits - Some More Definitions

Let Φ be an arithmetic circuit over a field $\mathbb{F}$ and a set of variables X. For a variable $x \in X$, we say that x *occurs* in Φ if x labels one of the input gates of Φ. Recall that, for a gate v in Φ, we defined Φ_v to be the sub-circuit of Φ rooted at v. For a gate v in Φ, define X_v to be the set of variables that occur in Φ_v. That is,

$$X_v = \begin{cases} \emptyset & v \text{ is an input gate labelled by a field element} \\ \{x\} & v \text{ is an input gate labelled by a variable } x \in X \\ X_{v_1} \cup X_{v_2} & v \text{ has sons } v_1 \text{ and } v_2 \end{cases}$$

For a variable $x \in X$ and a gate v in Φ, define $d_x(v)$, the *algebraic degree* of x in v, to be the degree of x in the polynomial $\widehat{\Phi}_v$. For a variable $x \in X$ and a gate v in Φ, define $sd_x(v)$, the *syntactic degree* of x in v, to be the degree of x in v when one ignores cancellations of monomials (in other words, if all the constants in Φ are replaced by 1's, and the field $\mathbb{F}$ is replaced by $\mathbb{R}$, then the syntactic degree of x in v is the algebraic degree of x in v). More precisely, define $sd_x(v)$ inductively as follows: If v is an input gate labelled by $\alpha \in \mathbb{F} \cup X$, then

$$sd_x(v) = \begin{cases} 1 & \alpha = x \\ 0 & \alpha \neq x \end{cases}$$

If v is a product gate with sons v_1 and v_2, then $sd_x(v) = sd_x(v_1) + sd_x(v_2)$. If v is an addition gate with sons v_1 and v_2, then $sd_x(v) = \max(sd_x(v_1), sd_x(v_2))$. The syntactic degree is a non-decreasing function, while the algebraic degree can decrease (in addition gates). That is, for every variable $x \in X$ and gates u and v in Φ, if there exists a directed path from u to v in Φ, then $sd_x(u) \leq sd_x(v)$. On the other hand, if v is an addition gate, and u is a son of v, it might be the case that $d_x(u) > d_x(v)$.

An arithmetic circuit Φ is called a *multilinear* arithmetic circuit if the polynomial computed at each gate in Φ is multilinear; that is, if for all $x \in X$ and v in Φ, it holds that $d_x(v) \leq 1$. An arithmetic circuit Φ is

called a *syntactically multilinear* arithmetic circuit, if for all $x \in X$ and v in Φ, it holds that $sd_x(v) \leq 1$. By induction, for all $x \in X$ and v in Φ, it holds that $d_x(v) \leq sd_x(v)$. Hence, indeed, every syntactically multilinear arithmetic circuit is a multilinear arithmetic circuit as well.

2.3 Partial Derivatives

Let f be a polynomial over the field $\mathbb{F}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. For $i \in [n]$, define $\frac{\partial f}{\partial x_i}$, the *partial derivative of f with respect to x_i*, as follows: If f is a monomial in the variables $X \setminus \{x_i\}$, then $\frac{\partial f}{\partial x_i} = 0$. If f is a monomial of the form $f = x_i^d g$, where d is a positive integer, and g is a monomial in the variables $X \setminus \{x_i\}$, then

$$\frac{\partial f}{\partial x_i} = \frac{\partial (x_i^d g)}{\partial x_i} = \underbrace{(1 + 1 + \cdots + 1)}_{d \text{ times}} x_i^{d-1} g.$$

If f is a sum of monomials $f = \sum_j m_j$, where m_j is a monomial in $\mathbb{F}[X]$, then $\frac{\partial f}{\partial x_i} = \sum_j \frac{\partial m_j}{\partial x_i}$.

The following lemma is known as the *chain rule of partial derivatives* (we state the lemma without giving a proof).

Lemma 2.1. *Let g be a polynomial over the field $\mathbb{F}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. Let h be a polynomial over the field $\mathbb{F}$ and the set of variables $X \cup \{x_0\}$. Let $f \in \mathbb{F}[X]$ be the polynomial h after substituting x_0 by g; that is, $f = h\Big|_{x_0=g}$. Then, for all $i \in [n]$,*

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\bigg|_{x_0=g} + \frac{\partial h}{\partial x_0}\bigg|_{x_0=g} \frac{\partial g}{\partial x_i}.$$

2.4 Multiplication of Variables in an Arithmetic Circuit

Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. For a variable $x_i \in X$, and a product gate v in Ψ with sons v_1 and v_2, define $\mathcal{M}_v(x_i)$, the *set of variables multiplying x_i in v*, by

$$\mathcal{M}_v(x_i) = \begin{cases} \emptyset & x_i \notin X_{v_1},\ x_i \notin X_{v_2} \\ X_{v_2} & x_i \in X_{v_1},\ x_i \notin X_{v_2} \\ X_{v_1} & x_i \notin X_{v_1},\ x_i \in X_{v_2} \\ X_{v_1} \cup X_{v_2} & x_i \in X_{v_1},\ x_i \in X_{v_2} \end{cases}$$

For a variable $x_i \in X$, define $\mathcal{M}_\Psi(x_i)$, the *set of variables multiplying x_i in Ψ*, by

$$\mathcal{M}_\Psi(x_i) = \bigcup_v \mathcal{M}_v(x_i),$$

where the union is over all product gates v in Ψ. For two variables $x_i, x_j \in X$, if x_i multiplies x_j in Ψ, then x_j multiplies x_i in Ψ, and vice versa; that is

$$x_i \in \mathcal{M}_\Psi(x_j) \Leftrightarrow x_j \in \mathcal{M}_\Psi(x_i).$$

Thus, for two variables $x_i, x_j \in X$, we say that x_i and x_j are multiplied in Ψ if $x_i \in \mathcal{M}_\Psi(x_j)$. Note that the following are equivalent:

- Ψ is a syntactically multilinear arithmetic circuit over the set of variables X.
- Ψ is an arithmetic circuit over the set of variables X such that every $x_i \in X$ admits $x_i \notin \mathcal{M}_\Psi(x_i)$.

2.5 The Partial Derivative Matrix

Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be two sets of variables. Let $f \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the variables Y and Z. Define L_f to be the $2^m \times 2^m$ *partial derivative matrix* of f as follows: for $p \in \mathbb{G}(Y)$ a monic[1] multilinear monomial in Y, and $q \in \mathbb{G}(Z)$ a monic multilinear monomial in Z, define $L_f(p, q)$ to be the coefficient of the monomial $p \cdot q$ in f. Thus, the rows of L_f correspond to monic multilinear monomials in Y, and the columns of L_f correspond to monic multilinear monomials in Z. We are mainly interested in the rank of the partial derivative matrix.

The following propositions bound the rank of the partial derivative matrix in different cases.

Proposition 2.2. *Let $f \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the sets of variables $Y' \subseteq Y$ and $Z' \subseteq Z$. Let $a = \min(|Y'|, |Z'|)$. Then,*

$$\mathit{Rank}(L_f) \leq 2^a.$$

Proof. There are two cases:

[1] A monic monomial is a monomial whose coefficient is 1.

1. If $|Y'| \leq |Z'|$, then $a = |Y'|$. Thus, L_f has at most 2^a non-zero rows.

2. If $|Y'| > |Z'|$, then $a = |Z'|$. Thus, L_f has at most 2^a non-zero columns. □[Proposition 2.2]

Proposition 2.3. *Let $f_1, f_2 \in \mathbb{G}[Y, Z]$ be two multilinear polynomials over the field $\mathbb{G}$ and the sets of variables Y and Z. Then,*

$$\mathit{Rank}(L_{f_1+f_2}) \leq \mathit{Rank}(L_{f_1}) + \mathit{Rank}(L_{f_2}).$$

Proof. Note that $L_{f_1+f_2} = L_{f_1} + L_{f_2}$. For any two matrices A and B, it holds that $\mathrm{Rank}(A + B) \leq \mathrm{Rank}(A) + \mathrm{Rank}(B)$. Thus,

$$\mathrm{Rank}(L_{f_1+f_2}) \leq \mathrm{Rank}(L_{f_1}) + \mathrm{Rank}(L_{f_2}).$$

□[Proposition 2.3]

Proposition 2.4. *Let $f_1 \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the sets of variables $Y_1 \subseteq Y$ and $Z_1 \subseteq Z$. Let $f_2 \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the sets of variables $Y_2 \subseteq Y$ and $Z_2 \subseteq Z$. Assume $Y_1 \cap Y_2 = \emptyset$ and $Z_1 \cap Z_2 = \emptyset$. Then,*

$$\mathit{Rank}(L_{f_1 \cdot f_2}) = \mathit{Rank}(L_{f_1}) \cdot \mathit{Rank}(L_{f_2}).$$

Proof. We think of L_{f_1} as a matrix of size $2^{|Y_1|} \times 2^{|Z_1|}$ and not of size $2^{|Y|} \times 2^{|Z|}$ (an entry in L_{f_1} that corresponds to a monomial that is not in the variables Y_1 and Z_1 is zero). Similarly, we think of L_{f_2} as a matrix of size $2^{|Y_2|} \times 2^{|Z_2|}$, and we think of $L_{f_1 \cdot f_2}$ as a matrix of size $2^{|Y_1 \cup Y_2|} \times 2^{|Z_1 \cup Z_2|}$. Since $Y_1 \cap Y_2 = \emptyset$ and $Z_1 \cap Z_2 = \emptyset$,

$$L_{f_1 \cdot f_2} = L_{f_1} \otimes L_{f_2},$$

where $\otimes$ denotes tensor product of matrices. For any two matrices A and B, it holds that $\mathrm{Rank}(A \otimes B) = \mathrm{Rank}(A) \cdot \mathrm{Rank}(B)$. Thus,

$$\mathrm{Rank}(L_{f_1 \cdot f_2}) = \mathrm{Rank}(L_{f_1}) \cdot \mathrm{Rank}(L_{f_2}).$$

□[Proposition 2.4]

Proposition 2.5. *Let $f \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the sets of variables Y and Z, where $|Y| = |Z| = m$. Let $t \in Y \cup Z$ be a variable, and let $g = \frac{\partial f}{\partial t}$. Assume that $Rank(L_f) = 2^m$. Then,*

$$Rank(L_g) = 2^{m-1}.$$

Proof. Assume without loss of generality that $t \in Z$. Assume without loss of generality that the columns of L_f are ordered such that $L_f = (A\ B)$, where A is a $2^m \times 2^{m-1}$ matrix whose columns correspond to all monomials q in which t does not occur, and B is a $2^m \times 2^{m-1}$ matrix whose columns correspond to all monomials q in which t occurs. Since $g = \frac{\partial f}{\partial t}$, we have $L_g = (B\ 0)$ (where 0 is a $2^m \times 2^{m-1}$ matrix of zeros). Since L_f is of full rank, the rank of B is 2^{m-1}. Hence, $\text{Rank}(L_g) = 2^{m-1}$. □[Proposition 2.5]

Proposition 2.6. *Let $f \in \mathbb{G}[Y, Z]$ be a multilinear polynomial over the field $\mathbb{G}$ and the sets of variables Y and Z, where $|Y| = |Z| = m$. Assume that the total degree of f is at most $T \in \mathbb{N}$. Then,*

$$Rank(L_f) \leq 2^{(T+1)\log m}.$$

Proof. Since the total degree of f is at most T, there are at most $\sum_{i=0}^{T} \binom{m}{i} \leq 2^{(T+1)\log m}$ non-zero rows in L_f.

□[Proposition 2.6]

2.6 Unbalanced Gates

Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the variables $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$. Let v be a gate in Ψ. Define Y_v to be the set of Y variables that occur in Ψ_v, and Z_v to be the set of Z variables that occur in Ψ_v. Define $b(v)$ to be the *average* of $|Y_v|$ and $|Z_v|$; that is, $b(v) = (|Y_v| + |Z_v|)/2$. Define $a(v)$ to be the *minimum* of $|Y_v|$ and $|Z_v|$; that is, $a(v) = \min(|Y_v|, |Z_v|)$. Define $d(v)$, the *balance gauge* of v, by $d(v) = b(v) - a(v)$. For an integer $k \in \mathbb{N}$, the gate v is called *k-unbalanced* if $d(v) \geq k$. Note that if Ψ is a syntactically multilinear arithmetic circuit, and u is a product gate in Ψ with sons u_1 and u_2, then $b(u) = b(u_1) + b(u_2)$.

3 Computing Partial Derivatives of Syntactically Multilinear Arithmetic Circuits

Let f be a polynomial over the field $\mathbb{G}$ and the variables $X = \{x_1, \ldots, x_n\}$. In [BS83], Baur and Strassen showed that the complexity of computing all n partial derivatives of f is (up to a constant factor) not more than computing f. More precisely, given an arithmetic circuit Ψ computing f, one can construct an arithmetic circuit Ψ' computing $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ such that $|\Psi'| = O(|\Psi|)$ (moreover, the depth[2] of Ψ' is up to a constant factor the same as the depth of Ψ). Later, Morgenstern ([M85]) simplified the construction of such a Ψ'.

Let Ψ be a *syntactically multilinear* arithmetic circuit over the field $\mathbb{G}$ and the set of variables X computing f. Let Ψ' be the arithmetic circuit computing $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ (as constructed by Baur and Strassen, and Morgenstern). For every $i \in [n]$, denote by v_i the gate computing $\frac{\partial f}{\partial x_i}$ in Ψ'. Since Ψ is a syntactically multilinear arithmetic circuit, f is a multilinear polynomial. Hence, for all $i \in [n]$ the degree of x_i in $\frac{\partial f}{\partial x_i}$ is 0; that is, $d_{x_i}(v_i) = 0$. The next theorem shows that (in addition to what Baur and Strassen showed)

- Ψ' is a *syntactically* multilinear arithmetic circuit.
- For all $i \in [n]$, the *syntactic* degree of x_i in v_i is 0; that is, $sd_{x_i}(v_i) = 0$. In other words, for all $i \in [n]$ the variable x_i does not occur in Ψ'_{v_i} (recall that, Ψ'_{v_i} is the sub-circuit of Ψ' rooted at v_i); that is, $x_i \notin X_{v_i}$.

Theorem 3.1. *Let Ψ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$ computing f. Then, there exists an arithmetic circuit Ψ' over the field $\mathbb{G}$ and the set of variables X such that*

1. *Ψ' computes all n partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.*
2. *$|\Psi'| \leq 5\,|\Psi|$.*
3. *Ψ' is a syntactically multilinear arithmetic circuit.*
4. *For every $i \in [n]$, it holds that $x_i \notin X_{v_i}$, where v_i is the gate computing $\frac{\partial f}{\partial x_i}$ in Ψ'.*

[2]The depth of an arithmetic circuit Ψ is the length of the longest directed path in Ψ.

We defer the proof of Theorem 3.1 to Section 3.1. We do not know if Theorem 3.1 holds for multilinear arithmetic circuits (that are not *syntactically* multilinear). In particular, there exists a *multilinear* arithmetic circuit Ψ (which is *not* syntactically multilinear) such that Ψ' (as constructed by Baur and Strassen, and Morgenstern) does not satisfy Theorem 3.1.

For the proof of our lower bound (Theorem 1.1) we need the following corollary of Theorem 3.1. The corollary shows that a gate v in Ψ' such that X_v is large (that is, many variables occur in Ψ'_v) is connected by directed paths to few output gates in Ψ' (the output gates of Ψ' are the gates computing $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$).

Corollary 3.2. *Let Ψ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$ computing f. Let Ψ' be the arithmetic circuit computing $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$, as described in Theorem 3.1. For $i \in [n]$, denote by v_i the gate computing $\frac{\partial f}{\partial x_i}$ in Ψ'. Then, for every gate v in Ψ',*

$$|X_v| \leq n - \left|\left\{i \in [n] \ : \ v \text{ is a gate in } \Psi'_{v_i}\right\}\right|.$$

Note that $\left|\left\{i \in [n] \ : \ v \text{ is a gate in } \Psi'_{v_i}\right\}\right|$ is the number of output gates that v is connected to by directed paths in Ψ'.

Proof. Let v be a gate in Ψ'. Let $i \in [n]$ be such that there exists a directed path from v to v_i in Ψ'; that is, v is a gate in Ψ'_{v_i}. Thus, $X_v \subseteq X_{v_i}$. By property 4 of Theorem 3.1, $x_i \notin X_{v_i}$. Hence, $x_i \notin X_v$. Therefore, $X_v \cap \left\{x_i \in X \ : \ v \text{ is a gate in } \Psi'_{v_i}\right\} = \emptyset$. Thus,

$$|X_v| \leq |X| - \left|\left\{x_i \in X \ : \ v \text{ is a gate in } \Psi'_{v_i}\right\}\right| = n - \left|\left\{i \in [n] \ : \ v \text{ is a gate in } \Psi'_{v_i}\right\}\right|.$$

□[Corollary 3.2]

3.1 Proof of Theorem 3.1

The following lemma is a generalization of Theorem 3.1.

Lemma 3.3. *Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X computing f. Then, there exists an arithmetic circuit Ψ' over the field $\mathbb{G}$ and the set of variables X such that*

1. *Ψ' computes all n partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.*
2. *$|\Psi'| \leq 5\,|\Psi|$.*

3. *For every $i \in [n]$, it holds that $\mathcal{M}_{\Psi'}(x_i) \subseteq \mathcal{M}_{\Psi}(x_i)$.*

4. *For every $i \in [n]$, it holds that $X_{v_i} \subseteq \mathcal{M}_{\Psi}(x_i)$, where v_i is the gate computing $\frac{\partial f}{\partial x_i}$ in Ψ'.*

We defer the proof of Lemma 3.3 to Section 3.2. First we give some intuition for Lemma 3.3. Let Ψ be an arithmetic circuit computing f. Let Ψ' be the arithmetic circuit computing all n partial derivatives of f (as constructed by Baur and Strassen, and Morgenstern). Then,

- Properties 1 and 2 of Lemma 3.3 were shown by Baur and Strassen.
- Property 3 of Lemma 3.3 states that if two variables x_i and x_j in X are not multiplied in Ψ, then x_i and x_j are not multiplied in Ψ' either.
- Let $i \in [n]$. Denote by v_i the gate computing $\frac{\partial f}{\partial x_i}$ in Ψ'. Property 4 of Lemma 3.3 states that Ψ'_{v_i} depends only on variables that multiply x_i in Ψ.

Theorem 3.1 follows from Lemma 3.3:

Proof of Theorem 3.1. Let Ψ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables X computing f. Since Ψ is a syntactically multilinear arithmetic circuit, $x_i \notin \mathcal{M}_{\Psi}(x_i)$. Let Ψ' be the arithmetic circuit given by Lemma 3.3. Then, Ψ' is an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X such that

1. Ψ' computes all n partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.
2. $|\Psi'| \leq 5\,|\Psi|$.
3. Let $i \in [n]$. Then, $\mathcal{M}_{\Psi'}(x_i) \subseteq \mathcal{M}_{\Psi}(x_i)$. Thus, as $x_i \notin \mathcal{M}_{\Psi}(x_i)$, it holds that $x_i \notin \mathcal{M}_{\Psi'}(x_i)$. Hence, Ψ' is a syntactically multilinear arithmetic circuit.
4. Let $i \in [n]$. Then, $X_{v_i} \subseteq \mathcal{M}_{\Psi}(x_i)$. Thus, as $x_i \notin \mathcal{M}_{\Psi}(x_i)$, it holds that $x_i \notin X_{v_i}$.

□[Theorem 3.1]

3.2 Proof of Lemma 3.3

Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X computing f. The proof of the lemma is by induction on the size of Ψ. The proof has four parts:

1. Induction Base.
2. Using Ψ to define a *smaller* arithmetic circuit Φ. Using induction to conclude the existence of an arithmetic circuit Φ' that has properties 1,2,3 and 4.
3. Using Φ' to construct Ψ'.
4. Proving that Ψ' has all the needed properties.

Induction Base

Assume that Ψ has no edges; that is, $|\Psi| = 0$. Thus, $f = \alpha$, for some $\alpha \in \mathbb{G} \cup X$. Hence, for all $i \in [n]$, it holds that $\frac{\partial f}{\partial x_i} \in \{0, 1\}$. Define Ψ' to be an arithmetic circuit with two input gates: an input gate labelled 1, and an input gate labelled 0. Then, Ψ' has properties 1,2,3, and 4.

Defining a smaller arithmetic circuit Φ

Let r be the gate computing f in Ψ. Assume that r is the unique output gate in Ψ (otherwise, consider Ψ_r). Assume that Ψ has at least two edges. Let v^* in Ψ be a gate with sons v_1^* and v_2^* such that v_1^* and v_2^* are input gates. Let $\alpha_1 \in \mathbb{G} \cup X$ be the label of v_1^* in Ψ. Let $\alpha_2 \in \mathbb{G} \cup X$ be the label of v_2^* in Ψ. Let x_0 be a new variable associated with the gate v^*, and denote $X^0 = X \cup \{x_0\}$. Let X^* be the set of X variables labelling v_1^* and v_2^* in Ψ; that is, $X^* = \{\alpha_1, \alpha_2\} \cap X = X_{v^*}$. Denote by g the polynomial computed by v^* in Ψ; that is, if v^* is an addition gate in Ψ, then $g = \alpha_1 + \alpha_2$, and if v^* is a product gate in Ψ, then $g = \alpha_1 \cdot \alpha_2$.

If $X^* = \emptyset$, then define a *smaller* arithmetic circuit Φ *computing* f as follows: Φ is obtained from Ψ by deleting the edges (v_1^*, v^*) and (v_2^*, v^*), and labelling v^* (which is an input gate) by $g(\alpha_1, \alpha_2) \in \mathbb{G}$. Thus, Φ computes f, and $|\Phi| < |\Psi|$. By induction, there exists an arithmetic circuit Φ' with properties $1, 2, 3$ and 4. Set Ψ' to be Φ'. Then, Ψ' has properties $1, 2, 3$ and 4, and the proof is complete.

Hence, we can assume that $X^* \neq \emptyset$.

Let Φ be the arithmetic circuit over the field $\mathbb{G}$ and the set of variables X^0 obtained from Ψ by deleting the edges (v_1^*, v^*) and (v_2^*, v^*), and labelling v^* (which is an input gate) by x_0. For every gate v in Ψ

there is a corresponding gate $u = u(v)$ in Φ, and vice versa. For a gate v in Ψ, we think of the gate $u = u(v)$ in Φ as the same gate as v.

Let $u(r)$ be the gate corresponding to r in Φ. Set h to be the polynomial computed by $u(r)$ in Φ. Thus, h is a polynomial in the variables X^0. By the construction of Φ, it follows that upon substituting x_0 by g in h we obtain f; that is, $f(X) = h(X^0)\Big|_{x_0=g}$.

Since $|\Phi| = |\Psi| - 2$, it follows by induction that there exists an arithmetic circuit Φ' over the field $\mathbb{G}$ and the set of variables X^0 such that

1. Φ' computes all $n+1$ partial derivatives $\frac{\partial h}{\partial x_0}, \frac{\partial h}{\partial x_1}, \ldots, \frac{\partial h}{\partial x_n}$.
2. $|\Phi'| \leq 5|\Phi|$.
3. For every $j \in \{0, \ldots, n\}$, it holds that $\mathcal{M}_{\Phi'}(x_j) \subseteq \mathcal{M}_{\Phi}(x_j)$.
4. For every $j \in \{0, \ldots, n\}$, it holds that $X^0_{u_j} \subseteq \mathcal{M}_{\Phi}(x_j)$, where u_j is the gate computing $\frac{\partial h}{\partial x_j}$ in Φ'.

Using Φ' to construct Ψ'

We construct Ψ' by adding a *few* gates and edges to Φ'.

Step one: gates and edges added to Φ' to substitute x_0 by g - constructing Ψ_1

Assume without loss of generality that in Φ' there is a unique input gate v' labelled x_0 (otherwise, join all input gates labelled x_0 to a single input gate labelled x_0). Denote by Ψ_1 the arithmetic circuit obtained by the following changes to Φ'

- Add two input gates to Φ': an input gate v'_1 labelled α_1, and an input gate v'_2 labelled α_2.
- Add two edges to Φ': the edge (v'_1, v'), and the edge (v'_2, v').
- Label v' by the same label of v^*.

Thus, Ψ_1 is an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X. Moreover, Ψ_1 is Φ' after substituting x_0 by g. We note that every gate in Φ' can also be thought of as a gate in Ψ_1.

Step two: gates and edges added to Ψ_1 to compute $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$

Fix $i \in [n]$. We describe what gates and edges are added to Ψ_1 in order for Ψ' to compute $\frac{\partial f}{\partial x_i}$. By Lemma 2.1 (the chain rule of partial derivatives), since $f = h\Big|_{x_0=g}$,

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\Big|_{x_0=g} + \frac{\partial h}{\partial x_0}\Big|_{x_0=g} \frac{\partial g}{\partial x_i}. \tag{3.1}$$

As Ψ_1 is Φ' after substituting x_0 by g, it follows that $\frac{\partial h}{\partial x_i}\Big|_{x_0=g}$ is computed by u_i (as a gate in Ψ_1), and $\frac{\partial h}{\partial x_0}\Big|_{x_0=g}$ is computed by u_0 (as a gate in Ψ_1). Consider the following four cases:

Case one: $\alpha_1 = x_i$ *and* $\alpha_2 \neq x_i$. Consider the following two cases:

1. If v^* is an addition gate, then $\frac{\partial g}{\partial x_i} = 1$. Hence, by (3.1),

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\Big|_{x_0=g} + \frac{\partial h}{\partial x_0}\Big|_{x_0=g}.$$

   To construct Ψ' add one gate and two edges as follows: add an addition gate v_i, and add the edges (u_0, v_i) and (u_i, v_i). Thus, v_i computes $\frac{\partial f}{\partial x_i}$.

2. If v^* is a product gate, then $\frac{\partial g}{\partial x_i} = \alpha_2$. Hence, by (3.1),

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\Big|_{x_0=g} + \frac{\partial h}{\partial x_0}\Big|_{x_0=g} \alpha_2.$$

   Recall that, v'_2 is an input gate labelled α_2 in Ψ_1. To construct Ψ' add two gates and four edges as follows:

   - A product gate w_1, and the edges (u_0, w_1) and (v'_2, w_1). Thus, w_1 computes $\frac{\partial h}{\partial x_0}\Big|_{x_0=g} \alpha_2$.
   - An addition gate v_i, and the edges (u_i, v_i) and (w_1, v_i).

   Thus, v_i computes $\frac{\partial f}{\partial x_i}$.

Case two: $\alpha_1 \neq x_i$ *and* $\alpha_2 = x_i$. We do the same as in case one (replacing 1 and 2). Note that in this case, when v^* is a product gate, we add a product gate w_2 to Ψ_1.

Case three: $\alpha_1 = \alpha_2 = x_i$. Consider the following two cases:

1. If v^* is an addition gate, then $\frac{\partial g}{\partial x_i} = 1 + 1$. Hence, by (3.1),

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\bigg|_{x_0=g} + \frac{\partial h}{\partial x_0}\bigg|_{x_0=g} (1+1).$$

   To construct Ψ' add three gates and four edges as follows:

   - An input gate w_3 labelled by $1 + 1$.
   - A product gate w_4, and the edges (w_3, w_4) and (u_0, w_4). Thus, w_4 computes $\frac{\partial h}{\partial x_0}\Big|_{x_0=g} (1+1)$.
   - An addition gate v_i, and the edges (u_i, v_i) and (w_4, v_i).

   Thus, v_i computes $\frac{\partial f}{\partial x_i}$.

2. If v^* is a product gate, then $\frac{\partial g}{\partial x_i} = (1+1)x_i$. Hence, by (3.1),

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\bigg|_{x_0=g} + \frac{\partial h}{\partial x_0}\bigg|_{x_0=g} (1+1)x_i.$$

   Recall that, v'_2 is an input gate labelled $x_i = \alpha_2$ in Ψ_1. To construct Ψ' add four gates and six edges as follows:

   - An input gate w_3 labelled by $1 + 1$.
   - A product gate w_4, and the edges (w_3, w_4) and (v'_2, w_4). Thus, w_4 computes $(1+1)x_i$.
   - A product gate w_5, and the edges (u_0, w_5) and (w_4, w_5). Thus, w_5 computes $\frac{\partial h}{\partial x_0}\Big|_{x_0=g} (1+1)x_i$.
   - An addition gate v_i, and the edges (u_i, v_i) and (w_5, v_i).

   Thus, v_i computes $\frac{\partial f}{\partial x_i}$.

Case four: $\alpha_1 \neq x_i$ *and* $\alpha_2 \neq x_i$. In this case no gates or edges are added. As g is a function of α_1 and α_2, it follows that $\frac{\partial g}{\partial x_i} = 0$. Hence, by (3.1), $\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial x_i}\Big|_{x_0=g}$. Denote by v_i the gate u_i in Ψ'. Thus, v_i computes $\frac{\partial f}{\partial x_i}$.

Ψ' has the needed properties

Let Ψ' be the arithmetic circuit over the field $\mathbb{G}$ and the set of variables X as constructed above. The following claims show that Ψ' has the needed properties. To prove the claims we make use of the following

- The construction of Φ from Ψ.
- The properties of Φ' (which we know by induction).
- The construction of Ψ' from Φ'.

In some of the proofs there are several cases to consider. Some of the cases are similar, but we consider all the cases for completeness.

The following claim shows that Ψ' satisfies property 1.

Claim 3.4. *Ψ' computes all n partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.*

Proof. Let $i \in [n]$. By the construction of Ψ' there exists a gate v_i in Ψ' computing $\frac{\partial f}{\partial x_i}$.

□[Claim 3.4]

The following claim shows that Ψ' satisfies property 2.

Claim 3.5. $|\Psi'| \leq 5|\Psi|$.

Proof. By the construction of Φ from Ψ, it follows that $|\Phi| = |\Psi| - 2$. By induction, $|\Phi'| \leq 5|\Phi|$. By the construction of Ψ' from Φ', there are at most ten more edges in Ψ' than in Φ'; that is, $|\Psi'| \leq |\Phi'| + 10$. Hence,

$$|\Psi'| \leq 5(|\Psi| - 2) + 10 = 5|\Psi|.$$

□[Claim 3.5]

Let $j \in \{0, \ldots, n\}$. Think of u_j both as the gate computing $\frac{\partial h}{\partial x_j}\Big|_{x_0=g}$ in Ψ', and as the gate computing $\frac{\partial h}{\partial x_j}$ in Φ'. Thus, X_{u_j} is the set of X variables that occur in Ψ'_{u_j}, and $X^0_{u_j}$ is the set of X^0 variables that occur in Φ'_{u_j}. We use the following claim.

Claim 3.6. *1. For every $i \in [n]$,*

$$X_{u_i} \subseteq \mathcal{M}_\Psi(x_i).$$

2. For every variable α in X^*,

$$X_{u_0} \subseteq \mathcal{M}_\Psi(\alpha).$$

Proof. For every $j \in \{0, \ldots, n\}$, by property 4 of Φ', $X^0_{u_j} \subseteq \mathcal{M}_\Phi(x_j)$, and by the construction of Ψ',

$$X_{u_j} = \begin{cases} X^0_{u_j} & x_0 \notin X^0_{u_j} \\ (X^0_{u_j} \setminus \{x_0\}) \cup X^* & x_0 \in X^0_{u_j} \end{cases}$$

Proof of 1: Fix $i \in [n]$. Since every variable that multiplies x_i in Φ except (possibly) x_0 also multiplies x_i in Ψ, it follows that $\mathcal{M}_\Phi(x_i) \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(x_i)$. Thus, since $X^0_{u_i} \subseteq \mathcal{M}_\Phi(x_i)$, it follows that $X^0_{u_i} \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(x_i)$. Consider the following two cases:

1. $x_0 \notin X^0_{u_i}$. Then, $X_{u_i} = X^0_{u_i} = X^0_{u_i} \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(x_i)$.
2. $x_0 \in X^0_{u_i}$. Since $X^0_{u_i} \subseteq \mathcal{M}_\Phi(x_i)$, it follows that x_0 multiplies x_i in Φ; that is, $x_0 \in \mathcal{M}_\Phi(x_i)$. Thus, $X^* \subseteq \mathcal{M}_\Psi(x_i)$. Hence, $X_{u_i} = (X^0_{u_i} \setminus \{x_0\}) \cup X^* \subseteq \mathcal{M}_\Psi(x_i)$.

Proof of 2: Fix $\alpha \in X^*$. Since every variable that multiplies x_0 in Φ except (possibly) x_0 also multiplies α in Ψ, it follows that $\mathcal{M}_\Phi(x_0) \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(\alpha)$. Thus, as $X^0_{u_0} \subseteq \mathcal{M}_\Phi(x_0)$, we have $X^0_{u_0} \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(\alpha)$. Consider the following two cases:

1. $x_0 \notin X^0_{u_0}$. Then, $X_{u_0} = X^0_{u_0} = X^0_{u_0} \setminus \{x_0\} \subseteq \mathcal{M}_\Psi(\alpha)$.
2. $x_0 \in X^0_{u_0}$. Since $X^0_{u_0} \subseteq \mathcal{M}_\Phi(x_0)$, it follows that x_0 multiplies x_0 in Φ; that is, $x_0 \in \mathcal{M}_\Phi(x_0)$. Thus, as $\alpha \in X^*$, we have $X^* \subseteq \mathcal{M}_\Psi(\alpha)$. Hence, $X_{u_0} = (X^0_{u_0} \setminus \{x_0\}) \cup X^* \subseteq \mathcal{M}_\Psi(\alpha)$.

□[Claim 3.6]

The following claim shows that Ψ' satisfies property 3.

Claim 3.7. *For all $i \in [n]$, it holds that $\mathcal{M}_{\Psi'}(x_i) \subseteq \mathcal{M}_\Psi(x_i)$.*

Proof. Let $i, j \in [n]$ be such that x_i and x_j are multiplied in Ψ'; that is, $x_i \in \mathcal{M}_{\Psi'}(x_j)$. To prove the claim it is enough to show that $x_i \in \mathcal{M}_\Psi(x_j)$. By property 3 of Φ', for all $\ell \in \{0, \ldots, n\}$,

$$\mathcal{M}_{\Phi'}(x_\ell) \subseteq \mathcal{M}_\Phi(x_\ell).$$

Recall that, X^* is the set of X variables that occur in Ψ_{v^*}. Consider two cases:

x_i and x_j are not in X^*. Assume that $x_i \notin X^*$ and $x_j \notin X^*$. Thus, as $x_i \in \mathcal{M}_{\Psi'}(x_j)$ and by the construction of Ψ', it follows that x_i and x_j are multiplied in Φ'; that is, $x_i \in \mathcal{M}_{\Phi'}(x_j)$. Thus, since

$\mathcal{M}_{\Phi'}(x_j) \subseteq \mathcal{M}_{\Phi}(x_j)$, it follows that x_i and x_j are multiplied in Φ; that is, $x_i \in \mathcal{M}_{\Phi}(x_j)$. Hence, by the construction of Φ (as $x_i \neq x_0$ and $x_j \neq x_0$), it follows that x_i and x_j are multiplied in Ψ; that is, $x_i \in \mathcal{M}_{\Psi}(x_j)$.

At least one of x_i and x_j is in X^*. Assume without loss of generality that $x_j = \alpha_1$ (recall that, $x_i \in \mathcal{M}_{\Psi}(x_j) \Leftrightarrow x_j \in \mathcal{M}_{\Psi}(x_i)$. Similar arguments hold for $x_j = \alpha_2$). Let v be a gate in Ψ' in which x_i and $x_j = \alpha_1$ are multiplied; that is, $x_i \in \mathcal{M}_v(\alpha_1)$. In the following we think of u_0 as the gate computing $\frac{\partial h}{\partial x_0}\Big|_{x_0=g}$ in Ψ'. Consider the following cases:

Case one: v is v'. Since $v = v'$ is a product gate in Ψ', we have that v^* is a product gate in Ψ. Since v' and v^* are the "same" gate, $\mathcal{M}_{v'}(\alpha_1) = \mathcal{M}_{v^*}(\alpha_1)$. Thus, $x_i \in \mathcal{M}_{v'}(\alpha_1) = \mathcal{M}_{v^*}(\alpha_1) \subseteq \mathcal{M}_{\Psi}(\alpha_1)$.

Case two: v is w_1. Recall that, w_1 is the product gate added to Ψ_1 in order for Ψ' to compute $\frac{\partial f}{\partial \alpha_1}$. The two sons of v in Ψ' are u_0 and v'_2. Since w_1 is added only if $\alpha_1 \neq \alpha_2$, it follows that $\mathcal{M}_v(\alpha_1) \subseteq X_{v'_2} \subseteq \{\alpha_2\}$. Thus, $x_i = \alpha_2$. Hence, since w_1 is added only if v^* is a product gate (that multiplies α_1 and α_2) in Ψ, $x_i = \alpha_2 \in \mathcal{M}_{v^*}(\alpha_1) \subseteq \mathcal{M}_{\Psi}(\alpha_1)$.

Case three: v is w_2. Recall that, w_2 is the product gate added to Ψ_1 in order for Ψ' to compute $\frac{\partial f}{\partial \alpha_2}$. The two sons of v in Ψ' are u_0 and v'_1, and v computes $\frac{\partial h}{\partial x_0}\Big|_{x_0=g} \alpha_1$. By definition of $\mathcal{M}_v(\alpha_1)$,

$$\mathcal{M}_v(\alpha_1) = \begin{cases} X_{u_0} & \alpha_1 \notin X_{u_0} \\ \{\alpha_1\} \cup X_{u_0} & \alpha_1 \in X_{u_0} \end{cases}$$

Thus, $\mathcal{M}_v(\alpha_1) = X_{u_0}$. Hence, by Claim 3.6,

$$x_i \in X_{u_0} \subseteq \mathcal{M}_{\Psi}(\alpha_1).$$

Case four: v is w_4. Recall that, w_4 is added to Ψ_1 in the case that $\alpha_1 = \alpha_2$. Since one of v's sons is an input gate labelled by a field element, $\mathcal{M}_v(\alpha_1) = \emptyset$. Hence, this case can not happen.

Case five: v is w_5. Recall that, w_5 is the product gate added to Ψ_1 in order for Ψ' to compute $\frac{\partial f}{\partial \alpha_1}$ (when $\alpha_1 = \alpha_2$). Thus, v computes $\frac{\partial h}{\partial x_0}\Big|_{x_0=g} (1+1)\alpha_1$. By definition of $\mathcal{M}_v(\alpha_1)$,

$$\mathcal{M}_v(\alpha_1) = \begin{cases} X_{u_0} & \alpha_1 \notin X_{u_0} \\ \{\alpha_1\} \cup X_{u_0} & \alpha_1 \in X_{u_0} \end{cases}$$

Thus, $\mathcal{M}_v(\alpha_1) = X_{u_0}$. Hence, by Claim 3.6,

$$x_i \in X_{u_0} \subseteq \mathcal{M}_{\Psi}(\alpha_1).$$

Case six: v is also a product gate in Φ'. There are two cases to consider:

1. If $x_i \notin \mathcal{M}_{\Phi'}(\alpha_1)$. Then, as $x_i \in \mathcal{M}_{\Psi'}(\alpha_1)$, it holds that $x_i \in \mathcal{M}_{\Phi'}(x_0)$. Hence, as $\mathcal{M}_{\Phi'}(x_0) \subseteq \mathcal{M}_{\Phi}(x_0)$, it follows that $x_i \in \mathcal{M}_{\Phi}(x_0)$. Since every variable that multiplies x_0 except (possibly) x_0 in Φ also multiplies α_1 in Ψ, we have $\mathcal{M}_{\Phi}(x_0) \setminus \{x_0\} \subseteq \mathcal{M}_{\Psi}(\alpha_1)$. Hence, as $x_i \neq x_0$,
$$x_i \in \mathcal{M}_{\Psi}(\alpha_1).$$
2. If $x_i \in \mathcal{M}_{\Phi'}(\alpha_1)$. Then, $x_i \in \mathcal{M}_{\Phi'}(\alpha_1) \subseteq \mathcal{M}_{\Phi}(\alpha_1)$. Since every variable that multiplies α_1 in Φ except (possibly) x_0 also multiplies α_1 in Ψ, it follows that $\mathcal{M}_{\Phi}(\alpha_1) \setminus \{x_0\} \subseteq \mathcal{M}_{\Psi}(\alpha_1)$. Thus, as $x_i \neq x_0$,
$$x_i \in \mathcal{M}_{\Psi}(\alpha_1).$$

□[Claim 3.7]

The following claim shows that Ψ' satisfies property 4.

Claim 3.8. *For all $i \in [n]$, it holds that $X_{v_i} \subseteq \mathcal{M}_{\Psi}(x_i)$.*

Proof. Fix $i \in [n]$. For simplicity of notation, denote $v = v_i, u = u_i$, and $u' = u_0$. We think of u and u' both as gates in Φ' and as gates in Ψ'. By the construction of Ψ', we have to consider the following cases:

Case one: $\alpha_1 = x_i$ or $\alpha_2 = x_i$. Note that this case applies both for $\alpha_1 = \alpha_2$ and $\alpha_1 \neq \alpha_2$. Assume without loss of generality that $\alpha_1 = x_i$. Consider the following two cases:

v^* is an addition gate in Ψ. Since u and u' are the sons of v in Ψ', it follows that $X_v = X_u \cup X_{u'}$. By Claim 3.6, it follows that $X_u \subseteq \mathcal{M}_{\Psi}(x_i)$ and $X_{u'} \subseteq \mathcal{M}_{\Psi}(\alpha_1) = \mathcal{M}_{\Psi}(x_i)$. Hence,
$$X_v \subseteq \mathcal{M}_{\Psi}(x_i).$$

v^* is a product gate. By the construction of Ψ' (loosely speaking, the sons of v in Ψ' are: u, u', and v'_2), it follows that $X_v = X_u \cup X_{u'} \cup X_{v'_2}$. As v^* is a product gate in Ψ (v^* multiplies $x_i = \alpha_1$ and α_2), it follows that $X_{v'_2} \subseteq \mathcal{M}_{\Psi}(x_i)$. By Claim 3.6, it follows that $X_u \subseteq \mathcal{M}_{\Psi}(x_i)$ and $X_{u'} \subseteq \mathcal{M}_{\Psi}(\alpha_1) = \mathcal{M}_{\Psi}(x_i)$. Hence,
$$X_v \subseteq \mathcal{M}_{\Psi}(x_i).$$

Case two: $\alpha_1 \neq x_i$ and $\alpha_2 \neq x_i$. By the construction of Ψ', it follows that $X_v = X_u$. By Claim 3.6, it follows that $X_u \subseteq \mathcal{M}_{\Psi}(x_i)$. Hence,
$$X_v \subseteq \mathcal{M}_{\Psi}(x_i).$$

□[Claim 3.8]

Thus, Ψ' is an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X such that

1. By Claim 3.4, Ψ' computes all n partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.
2. By Claim 3.5, $|\Psi'| \leq 5 \cdot |\Psi|$.
3. By Claim 3.7, for every $i \in [n]$, it holds that $\mathcal{M}_{\Psi'}(x_i) \subseteq \mathcal{M}_{\Psi}(x_i)$.
4. By Claim 3.8, for every $i \in [n]$, it holds that $X_{v_i} \subseteq \mathcal{M}_{\Psi}(x_i)$.

□[Lemma 3.3]

4 Partitions of the Variables of an Arithmetic Circuit

In this section we define a distribution $\mathcal{D}$ on partitions of the variables of an arithmetic circuit. We show that by the distribution $\mathcal{D}$, a specific gate is unbalanced with high probability.

4.1 Definitions

Let $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_m\}$, and $Z = \{z_1, \ldots, z_m\}$ be three sets of variables (where $n = 2m$). A one-to-one function $A : X \to Y \cup Z$ is called a *partition* of X to Y and Z. For a partition A of X to Y and Z and $X' \subseteq X$ a subset of X, denote $A(X') = \{A(x) \ : \ x \in X'\}$.

Let $X_1 \subset X$ be a subset of X of size $n/4$. The distribution $\mathcal{D}(X_1)$ on partitions A of X to Y and Z is the uniform distribution on all partitions A such that $A(X_1) \subset Y$. We write $A \sim \mathcal{D}(X_1)$ if A is a partition of X to Y and Z chosen by the distribution $\mathcal{D}(X_1)$.

Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the set of variables X computing a polynomial f. Let A be a partition of X to Y and Z. Denote by Ψ_A the arithmetic circuit Ψ after substituting every $x \in X$ by $A(x) \in Y \cup Z$. Denote by f^A the polynomial f after substituting every $x \in X$ by $A(x) \in Y \cup Z$. Note that Ψ_A is an arithmetic circuit over the field $\mathbb{G}$ and the sets of variables Y and Z computing f^A. For every gate in Ψ there is a corresponding gate in Ψ_A, and vice versa. We think of a gate v in Ψ and v's corresponding gate in Ψ_A as the same gate, and we denote both of them by v.

4.2 The Probability that a Gate is Unbalanced

The following proposition bounds the probability that a gate (with a certain number of variables) is not unbalanced.

Proposition 4.1. *Let Ψ be an arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be two sets of variables (where $n = 2m$ and m is even). Let $X_1 \subset X$ be a subset of X of size $n/4$. Let $A \sim \mathcal{D}(X_1)$ be a random partition of X to Y and Z such that $A(X_1) \subset Y$. Let β be such that $0 < \beta < 1$, and let v be a gate in Ψ such that $n^\beta < |X_v| < n - n^\beta$. Then, for any integer $k \in \mathbb{N}$,*

$$\Pr_{A \sim \mathcal{D}(X_1)}[v \text{ is not } k\text{-unbalanced in } \Psi_A] = O\left(kn^{-\beta/2}\right).$$

To prove Proposition 4.1 we need some property of the hypergeometric distribution. We defer the proof of Proposition 4.1 to Section 4.4.

4.3 The Hypergeometric Distribution

Let $N, M_1, M_2 \in \mathbb{N}$ be three integers such that $M_1 \leq N$ and $M_2 \leq N$. Denote by $\mathcal{H}(N, M_1, M_2)$ the hypergeometric distribution with parameters M_1, M_2 and N; that is, $\mathcal{H}(N, M_1, M_2)$ is the distribution of $|S_1 \cap S_2|$, where S_1 is a random subset of $[N]$ of size M_1 (chosen uniformly at random from all subsets of $[N]$ of size M_1), and S_2 is a fixed subset of $[N]$ of size M_2.

The following proposition shows that a hypergeometric random variable does not take any specific value with high probability (for a certain range of the parameters).

Proposition 4.2. *Let $n \in \mathbb{N}$ be an integer such that $n/4$ is an integer as well. Let β be such that $0 < \beta < 1$, and let χ be a random variable that has the hypergeometric distribution $\mathcal{H}(3n/4, n/4, M)$, where*

$$n^\beta/4 < M < 3n/4 - n^\beta/4. \tag{4.1}$$

Then, every $j \in \mathbb{N}$ admits $Pr[\chi = j] = O(n^{-\beta/2})$.

Proof. Denote $j_{\max} = \min(M, n/4)$, the maximal value that χ takes. For every $j \in \mathbb{N}$, denote $P(j) = \Pr[\chi = j]$. Thus, for every $j \in \{0, \ldots, j_{\max}\}$, we have

$$P(j) = \frac{\binom{M}{j}\binom{3n/4-M}{n/4-j}}{\binom{3n/4}{n/4}}, \tag{4.2}$$

and for every $j \notin \{0, \ldots, j_{\max}\}$, we have $P(j) = 0$. Set $j^* \in \{0, \ldots, j_{\max}\}$ to be the integer that maximizes $P(j)$; that is, every $j \in \mathbb{N}$ admits $P(j) \leq P(j^*)$. To find j^* consider $P(j+1)/P(j)$. By (4.2), for all $j \in \{0, \ldots, j_{\max} - 1\}$,

$$\frac{P(j+1)}{P(j)} = \frac{(M-j)(n/4-j)}{(j+1)(n/2-M+j+1)},$$

which implies that

$$P(j) \leq P(j+1) \Leftrightarrow j \leq \frac{Mn/4 + M - n/2 - 1}{3n/4+2} = \frac{M}{3} + \frac{M - 3n/2 - 3}{3(3n/4+2)}.$$

Since $0 < M < 3n/4$, we have $-1 < (M - 3n/2 - 3)/(3(3n/4+2)) < 0$. Thus, $j^* \in \{\lfloor M/3 \rfloor, \lceil M/3 \rceil\}$. Using Stirling's formula, we have

$$\binom{N}{\lfloor N/3 \rfloor} = \Theta\left(\frac{1}{\sqrt{N}} 2^{N \cdot H(1/3)}\right) \text{ and } \binom{N}{\lceil N/3 \rceil} = \Theta\left(\frac{1}{\sqrt{N}} 2^{N \cdot H(1/3)}\right), \tag{4.3}$$

where $H(1/3) = -(1/3)\log_2(1/3) - (2/3)\log_2(2/3)$. Hence, by (4.2), using (4.3) for N equals $M, 3n/4 - M$, and $3n/4$,

$$P(j^*) = \frac{\Theta\left(\frac{1}{\sqrt{M}} 2^{M \cdot H(1/3)}\right) \Theta\left(\frac{1}{\sqrt{3n/4-M}} 2^{(3n/4-M) \cdot H(1/3)}\right)}{\Theta\left(\frac{1}{\sqrt{3n/4}} 2^{(3n/4) \cdot H(1/3)}\right)} = \Theta\left(\frac{\sqrt{3n/4}}{\sqrt{M}\sqrt{3n/4-M}}\right) = \Theta\left(n^{-\beta/2}\right),$$

where the last equality follows from (4.1). Hence, every $j \in \mathbb{N}$ admits $P(j) \leq P(j^*) = O(n^{-\beta/2})$.

□[Proposition 4.2]

4.4 Proof of Proposition 4.1

If $k > n^\beta/4$, then the proposition holds (as $kn^{-\beta/2} > 1$). Thus, assume that $k \leq n^\beta/4$. Let $A \sim \mathcal{D}(X_1)$ be a random partition of X to Y and Z such that $A(X_1) \subset Y$. By the definition of $\mathcal{D}(X_1)$, we think of A as obtained by the following randomized process: let X_2 be a random subset of $\overline{X_1} = X \setminus X_1$ of size $n/4$, then let A be a random partition such that $A(X_1 \cup X_2) = Y$. Recall that, Y_v is the set of Y variables that occur in the sub-circuit of Ψ_A rooted at v. Thus, $|Y_v| = |X_v \cap X_1| + |X_v \cap X_2|$. Hence,

$$|X_v \cap X_1| \leq |Y_v| \leq |X_v \cap X_1| + n/4.$$

There are three cases to consider. In the first two cases X_v either has small intersection with $\overline{X_1}$ or has large intersection with $\overline{X_1}$, and then v is always unbalanced. In the third case we use Proposition 4.2 to show that v is unbalanced with high probability.

Case one: Assume $\left|X_v \cap \overline{X_1}\right| \leq n^\beta/4$. Then, $|X_v| = \left|X_v \cap \overline{X_1}\right| + |X_v \cap X_1| \leq n^\beta/4 + |X_v \cap X_1|$. Thus, since $|X_v| \geq n^\beta$, it follows that $|Y_v| \geq |X_v \cap X_1| \geq |X_v| - n^\beta/4 \geq |X_v|/2 + n^\beta/4$. Hence, $a(v) \leq |Z_v| = |X_v| - |Y_v| \leq |X_v|/2 - n^\beta/4$. Thus, $d(v) = |X_v|/2 - a(v) \geq n^\beta/4$. Therefore, since $k \leq n^\beta/4$, it follows that v is not $k-$unbalanced in Ψ_A with probability 0.

Case two: Assume $\left|X_v \cap \overline{X_1}\right| \geq 3n/4 - n^\beta/4$. Then, $|X_v| = \left|X_v \cap \overline{X_1}\right| + |X_v \cap X_1| \geq 3n/4 - n^\beta/4 + |X_v \cap X_1|$. Hence, $a(v) \leq |Y_v| \leq |X_v \cap X_1| + n/4 \leq |X_v| - n/2 + n^\beta/4$. Thus, since $|X_v| \leq n - n^\beta$, it follows that $d(v) = |X_v|/2 - a(v) \geq -|X_v|/2 + n/2 - n^\beta/4 \geq n^\beta/4$. Thus, since $k \leq n^\beta/4$, it follows that v is not $k-$unbalanced in Ψ_A with probability 0.

Case three: Assume

$$n^\beta/4 < \left|X_v \cap \overline{X_1}\right| < 3n/4 - n^\beta/4. \tag{4.4}$$

Denote $Y_1 = X_v \cap X_1$ and $Y_2 = X_v \cap X_2$. Thus, $|Y_v| = |Y_1| + |Y_2|$. Note that Y_1 is a fixed subset of X, and $|Y_2|$ has the geometric distribution $\mathcal{H}(|\overline{X_1}|, n/4, |X_v \cap \overline{X_1}|)$. Thus, by (4.4) and Proposition 4.2, $|Y_2|$ takes any specific value with probability $O(n^{-\beta/2})$. Denote $\mu = |X_v|/2$. Hence,

$$\begin{aligned}
\Pr_{A\sim\mathcal{D}(X_1)}[v \text{ is not } k\text{-unbalanced in } \Psi_A] &\leq \Pr_{A\sim\mathcal{D}(X_1)}[\mu - k \leq |Y_v| \leq \mu + k] \\
&\leq \sum_{j=\lfloor\mu-k\rfloor}^{\lceil\mu+k\rceil} \Pr_{A\sim\mathcal{D}(X_1)}[|Y_v| = j] = O\left(kn^{-\beta/2}\right).
\end{aligned}$$

□[Proposition 4.1]

5 Small Syntactically Multilinear Arithmetic Circuits Compute Polynomials of "Low Rank"

In this section we prove that a small syntactically multilinear arithmetic circuit computes a polynomial whose partial derivative matrix is not of full rank (for some partition of the variables). Formally,

Theorem 5.1. *Let Ψ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$ computing f. Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be two sets of variables (where $n = 2m$ and m is even). If for all partitions A of X to Y and Z*

$$Rank\left(L_{f^A}\right) = 2^m,$$

then

$$|\Psi| = \Omega\left(\frac{n^{4/3}}{\log^2 n}\right).$$

The rest of this section is devoted for the proof of Theorem 5.1. In Section 5.1 we introduce the notion of levelled gates, and state a lemma. In Section 5.2 we use the lemma to prove Theorem 5.1.

5.1 Few Levelled Gates Means Low Rank

Let Φ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the variables $X = \{x_1, \ldots, x_n\}$. Fix $\tau = 3 \log n$. Define $\mathcal{L}(\Phi, \tau)$, the set of *lower levelled* gates in Φ, by

$$\mathcal{L}(\Phi, \tau) = \{u \text{ is a gate in } \Phi \ : \ 2\tau < |X_u| < n - 2\tau \text{ and } u \text{ has a father } u' \text{ such that } |X_{u'}| \geq n - 2\tau\}.$$

The following lemma shows that, if the set of lower levelled gates in a circuit is small, then the partial derivative matrix of a polynomial computed by the circuit is not of full rank (for some partition of the variables).

Lemma 5.2. *Let Φ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$ computing f. Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be two sets of variables (where $n = 2m$ and m is even). Let $\tau = 3 \log n$, and let $\mathcal{L} = \mathcal{L}(\Phi, \tau)$ be the set of lower levelled gates in Φ (as defined above). Let $c > 0$ be a small enough constant ($c = 1/1000$ suffices). Assume $|\mathcal{L}| < \frac{c}{\tau} n^{1/3}$. Then, there exists a partition A of X to Y and Z such that*

$$Rank(L_{f^A}) < 2^{m-1}.$$

We defer the proof of Lemma 5.2 to Section 5.4. We use Lemma 5.2 to prove Theorem 5.1.

5.2 Proof of Theorem 5.1

Let Ψ' be the arithmetic circuit computing all n partial derivatives of f given by Theorem 3.1. Let $\tau = 3\log n$, and let $\mathcal{L} = \mathcal{L}(\Psi', \tau)$ be the set of lower levelled gates in Ψ' (as defined in Section 5.1). Define $\mathcal{U} = \mathcal{U}(\Psi', \tau)$, the set of *upper levelled* gates in Ψ', by

$$\mathcal{U} = \{u' \text{ is a gate in } \Psi' \ : \ n - 2\tau \leq |X_{u'}| \text{ and } u' \text{ has a son } u \text{ such that } u \in \mathcal{L}\}.$$

To prove the theorem, we will bound from bellow the size of $\mathcal{U}$.

Let $i \in [n]$. Set $g_i = \frac{\partial f}{\partial x_i}$. Let v_i be the gate computing g_i in Ψ'. Denote by Ψ'_i the arithmetic circuit Ψ'_{v_i}. Define $\mathcal{L}_i = \mathcal{L}(\Psi'_i, \tau)$ to be the set of lower levelled gates in Ψ'_i. The following claim gives two properties of $\mathcal{L}_i$.

Claim 5.3. *For every $i \in [n]$,*

1. $\mathcal{L}_i \subseteq \mathcal{L}$.
2. $|\mathcal{L}_i| \geq \frac{c}{\tau} n^{1/3}$, *where c is the constant from Lemma 5.2.*

Proof. Proof of 1: Note that for every gate u in Ψ'_i, the set X_u in Ψ' and in Ψ'_i is the same set. Let $u \in \mathcal{L}_i$. Thus, $2\tau < |X_u| < n - 2\tau$ and u has a father u' in Ψ'_i such that $|X_{u'}| \geq n - 2\tau$. Since u' is a father of u in Ψ', we have $u \in \mathcal{L}$. Hence, $\mathcal{L}_i \subseteq \mathcal{L}$.

Proof of 2: For every partition A of X to Y and Z, we have $g_i^A = \left(\frac{\partial f}{\partial x_i}\right)^A = \frac{\partial f^A}{\partial A(x_i)}$, which implies using Proposition 2.5 (since L_{f^A} is of full rank) that $\mathrm{Rank}(L_{g_i^A}) = 2^{m-1}$. Hence, by Lemma 5.2, since Ψ'_i computes g_i, $|\mathcal{L}_i| \geq \frac{c}{\tau} n^{1/3}$, where c is the constant from Lemma 5.2. □[Claim 5.3]

For a gate v in Ψ', define

$$C_v = |\{i \in [n] \ : \ v \text{ is a gate in } \Psi'_i\}|.$$

For $i \in [n]$, define

$$\mathcal{U}_i = \{u' \in \mathcal{U} \ : \ u' \text{ is a gate in } \Psi'_i\}.$$

Thus, for all $i \in [n]$, we have $\mathcal{U}_i \subseteq \mathcal{U}$. Hence,

$$\sum_{i \in [n]} |\mathcal{U}_i| = |\{(u', i) \ : \ u' \in \mathcal{U} \text{ and } i \in [n] \text{ are such that } u' \text{ is a gate in } \Psi'_i\}| = \sum_{u' \in \mathcal{U}} C_{u'}. \tag{5.1}$$

Let $i \in [n]$. By property 1 of Claim 5.3, $\mathcal{L}_i \subseteq \mathcal{L}$. Hence, for every gate $u \in \mathcal{L}_i$, there is a corresponding gate $u' \in \mathcal{U}_i$, which is a father of u. Thus, since the in-degree of the gates in $\mathcal{U}_i$ is two, we have

$$|\mathcal{L}_i| \leq 2|\mathcal{U}_i|. \tag{5.2}$$

Recall that, for all $u' \in \mathcal{U}$, it holds that $|X_{u'}| \geq n - 2\tau$. Thus, by Corollary 3.2, every $u' \in \mathcal{U}$ admits $C_{u'} \leq n - |X_{u'}| \leq n - (n - 2\tau) = 2\tau$. Thus, by (5.2), (5.1), and by property 2 of Claim 5.3,

$$\frac{c}{\tau} n^{4/3} \leq \sum_{i \in [n]} |\mathcal{L}_i| \leq 2 \sum_{i \in [n]} |\mathcal{U}_i| = 2 \sum_{u' \in \mathcal{U}} C_{u'} \leq 2|\mathcal{U}| \cdot 2\tau.$$

Hence, by property 2 of Theorem 3.1, since $\tau = 3 \log n$,

$$|\Psi| = \Omega\left(|\Psi'|\right) = \Omega\left(|\mathcal{U}|\right) = \Omega\left(\frac{n^{4/3}}{\log^2 n}\right).$$

□[Theorem 5.1]

5.3 Unbalancing the Lower Levelled Gates of a Small Arithmetic Circuit

In the rest of this section we prove Lemma 5.2. First we prove that the set of lower levelled gates in a small arithmetic circuit can be made simultaneously unbalanced.

Proposition 5.4. *Let Φ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables $X = \{x_1, \ldots, x_n\}$. Let $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be two sets of variables (where $n = 2m$ and m is even). Let $\tau = 3 \log n$, and let $\mathcal{L} = \mathcal{L}(\Phi, \tau)$ be the set of lower levelled gates in Φ (as defined in Section 5.1). Let $c > 0$ be a small enough constant ($c = 1/1000$ suffices). Assume $|\mathcal{L}| < \frac{c}{\tau} n^{1/3}$. Then, there exists a partition A of X to Y and Z such that every $u \in \mathcal{L}$ is $\tau-$unbalanced in Φ_A.*

Proof. For a gate v in Φ define $\tilde{X}_v$, by

$$\tilde{X}_v = \begin{cases} X_v & |X_v| \leq n/2 \\ X \setminus X_v & |X_v| > n/2 \end{cases}.$$

Every partition A of X to Y and Z defines a partition of $\tilde{X}_v$ to $\tilde{Y}_v$ and $\tilde{Z}_v$:

$$\tilde{Y}_v = \left\{ y \in Y \ : \ A^{-1}(y) \in \tilde{X}_v \right\} \text{ and } \tilde{Z}_v = \left\{ z \in Z \ : \ A^{-1}(z) \in \tilde{X}_v \right\}.$$

For every partition A of X to Y and Z and for every gate v in Φ_A,

$$d(v) = \frac{|Y_v| + |Z_v|}{2} - \min(|Y_v|, |Z_v|) \quad = \quad \frac{|\tilde{Y}_v| + |\tilde{Z}_v|}{2} - \min(|\tilde{Y}_v|, |\tilde{Z}_v|),$$

which implies that

$$v \text{ is } \tau-\text{unbalanced in } \Phi_A \Leftrightarrow \frac{|\tilde{Y}_v| + |\tilde{Z}_v|}{2} - \min(|\tilde{Y}_v|, |\tilde{Z}_v|) \geq \tau. \tag{5.3}$$

Partition $\mathcal{L}$ to two sets:

$$\mathcal{L}_{small} = \left\{ v \in \mathcal{L} \ : \ |\tilde{X}_v| \leq n^{2/3} \right\} \quad , \quad \mathcal{L}_{big} = \mathcal{L} \setminus \mathcal{L}_{small} = \left\{ v \in \mathcal{L} \ : \ |\tilde{X}_v| > n^{2/3} \right\}.$$

For all $u \in \mathcal{L}_{small}$, it holds that $|\tilde{X}_u| \leq n^{2/3}$. Define $X'_1 = \bigcup_{u \in \mathcal{L}_{small}} \tilde{X}_u$. Since $|\mathcal{L}| < \frac{c}{\tau} n^{1/3}$, it follows that $|X'_1| \leq \frac{c}{\tau} n^{1/3} n^{2/3} < n/4$. Hence, there exists a set $X_1 \subseteq X$ of size $n/4$ such that for all $u \in \mathcal{L}_{small}$, we have $\tilde{X}_u \subseteq X_1$ (X_1 is some superset of X'_1 of size $n/4$). Let $A \sim \mathcal{D}(X_1)$ be a random partition of X to Y and Z such that $A(X_1) \subset Y$ (see Section 4 for definition of $\mathcal{D}(X_1)$).

For all $u \in \mathcal{L}_{small}$, it holds that $|\tilde{Y}_u| + |\tilde{Z}_u| = |\tilde{X}_u| \geq 2\tau$ and $\tilde{Z}_u = \emptyset$ (as $A(\tilde{X}_u) \subseteq A(X_1) \subset Y$). Thus, by (5.3), every $u \in \mathcal{L}_{small}$ is $\tau-$unbalanced in Φ_A (with probability 1). Note that every $u \in \mathcal{L}_{big}$ admits $n^{2/3} < |X_u| < n - n^{2/3}$. Thus, by Proposition 4.1 for $\beta = 2/3$, and since $|\mathcal{L}| < \frac{c}{\tau} n^{1/3}$,

$$\mathop{\mathbb{E}}_{A \sim \mathcal{D}(X_1)} [|\{u \in \mathcal{L} \ : \ u \text{ is not } \tau\text{-unbalanced in } \Phi_A\}|] \leq O\left(|\mathcal{L}_{big}| \tau n^{-1/3}\right) < 1$$

(for c small enough). Hence, there exists a partition A such that every $u \in \mathcal{L}$ is $\tau-$unbalanced in Φ_A. □[Proposition 5.4]

5.4 Proof of Lemma 5.2

Let r be a gate computing f in Φ. Assume $|X_r| < n - 2$. Let A be a partition of X to Y and Z. Thus, in Φ_A we have $a(r) \leq b(r) < m - 1$. Hence, by Proposition 2.2, $\text{Rank}(L_{f^A}) < 2^{m-1}$, which proves the lemma. Thus, assume $|X_r| \geq n - 2$.

Since $|\mathcal{L}| < \frac{c}{\tau} n^{1/3}$, by Proposition 5.4 there exists a partition A of X to Y and Z such that every $u \in \mathcal{L}$ is $\tau-$unbalanced in Φ_A. Denote by Ψ the arithmetic circuit Φ_A. In the rest of the proof we focus on Ψ. Recall that, for a gate u in Ψ, Y_u is the set of Y variables that occur in Ψ_u, and Z_u is the set of Z variables that occur in Ψ_u.

Define an order on $\mathcal{L}$ that respects the order of Ψ; that is, $\mathcal{L} = \{u_1, \ldots, u_\ell\}$, where $\ell = |\mathcal{L}|$, and for every $i, j \in [\ell]$ such that $i < j$, there is no directed path from u_i to u_j in Ψ. For $i \in [\ell]$, denote $h_i = \widehat{\Psi}_{u_i}$, the polynomial computed by u_i in Ψ, denote $Y_i = Y_{u_i}$, and denote $Z_i = Z_{u_i}$. For a gate v in Ψ, we say that v is *substituted* by α (which is a field element or a variable) in Ψ, if the edges going into v are deleted, and v is relabelled by α.

The following proposition shows how to write the polynomial $f^A \in \mathbb{G}[Y, Z]$ (i.e., the polynomial computed by r in Ψ) as a function of the polynomials $h_1, \ldots, h_\ell \in \mathbb{G}[Y, Z]$ (i.e., the polynomials computed by $u_1, \ldots, u_\ell$ in Ψ).

Proposition 5.5.

$$f^A = \sum_{i \in [\ell]} g_i h_i + g,$$

where $g, g_1, \ldots, g_\ell \in \mathbb{G}[Y, Z]$ are multilinear polynomials such that

1. *For all $i \in [\ell]$, the set of variables that occur in g_i and the set $Y_i \cup Z_i$ are disjoint.*
2. *g is the polynomial computed by r in Ψ, after substituting (in Ψ) each $u \in \mathcal{L}$ by 0.*

Proof. To prove the proposition we follow an inductive process that is based on the following claim.

Claim 5.6. *Let Υ be a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the sets of variables Y and Z. Let $\tilde{r}$ be a gate in Υ computing a polynomial $\tilde{f}$. Let v be a gate in Υ such that $Y_v \cup Z_v \neq \emptyset$. Denote by $\tilde{h} = \widehat{\Upsilon}_v$, the polynomial computed by v in Υ. Then, there exist two multilinear polynomials $\tilde{g}_1, \tilde{g}_2 \in \mathbb{G}[Y, Z]$ such that $\tilde{f} = \tilde{g}_1 \tilde{h} + \tilde{g}_2$, where*

1. *The set of variables that occur in $\tilde{g}_1$ and the set $Y_v \cup Z_v$ are disjoint.*
2. *$\tilde{g}_2$ is the polynomial computed by $\tilde{r}$ in Υ, after substituting (in Υ) v by 0.*

Proof. Let t be a new variable. Let Υ_1 be the arithmetic circuit Υ, after substituting (in Υ) v by t. Let $G \in \mathbb{G}[Y, Z, t]$ be the polynomial computed by $\tilde{r}$ in Υ_1. Since Υ is syntactically multilinear and since $Y_v \cup Z_v \neq \emptyset$, it follows that G is linear in t; that is, G is of the form $G = \tilde{g}_1 t + \tilde{g}_2$, where $\tilde{g}_1$ and $\tilde{g}_2$ are two multilinear polynomials in $\mathbb{G}[Y, Z]$. Since $\tilde{f}$ is G, after substituting (in G) t by $\tilde{h}$, we have $\tilde{f} = \tilde{g}_1 \tilde{h} + \tilde{g}_2$. Recall that, $\mathcal{M}_{\Upsilon_1}(t)$ is the set of variables that multiply t in Υ_1 (see Section 2.4). The set of variables that occur in $\tilde{g}_1$ is a subset of $\mathcal{M}_{\Upsilon_1}(t)$. Since Υ is syntactically multilinear, the sets $\mathcal{M}_{\Upsilon_1}(t)$ and $Y_v \cup Z_v$ are disjoint. Hence, the set of variables that occur in $\tilde{g}_1$ and the set $Y_v \cup Z_v$ are disjoint. Since $\tilde{g}_2 = G\big|_{t=0}$, we have that $\tilde{g}_2$ is the polynomial computed by $\tilde{r}$ in Υ, after substituting (in Υ) v by 0. □[Claim 5.6]

For $i \in [\ell]$, define Ψ^i to be the arithmetic circuit Ψ, after substituting (in Ψ) $u_1, \ldots, u_i$ by 0. We now describe the inductive process:

First step: Recall that, r computes f^A in Ψ, and u_1 computes h_1 in Ψ. Hence, by Claim 5.6 (for $\Upsilon = \Psi$, $\tilde{r} = r$ and $v = u_1$), since $Y_1 \cup Z_1 \neq \emptyset$, there exist two multilinear polynomials $g_1, g'_1 \in \mathbb{G}[Y, Z]$ such that

$$f^A = g_1 h_1 + g'_1,$$

where the set of variables that occur in g_1 and the set $Y_1 \cup Z_1$ are disjoint, and g'_1 is the polynomial computed by r in Ψ^1. We continue in a similar manner.

Inductive step: Assume $g_1, \ldots, g_i$ are already defined (where $i \in [\ell - 1]$), and r computes g'_i in Ψ^i. Since there are no directed pathes from the gates $u_1, \ldots, u_i$ to u_{i+1} in Ψ, the polynomial computed by u_{i+1} in Ψ^i is h_{i+1}, and the set of variables that occur in $\Psi^i_{u_{i+1}}$ is the same as the set of variables that occur in $\Psi_{u_{i+1}}$. Hence, by Claim 5.6 (for $\Upsilon = \Psi^i$, $\tilde{r} = r$ and $v = u_{i+1}$), since $Y_{i+1} \cup Z_{i+1} \neq \emptyset$, there exist two multilinear polynomials $g_{i+1}, g'_{i+1} \in \mathbb{G}[Y, Z]$ such that

$$g'_i = g_{i+1} h_{i+1} + g'_{i+1},$$

where the set of variables that occur in g_{i+1} and the set $Y_{i+1} \cup Z_{i+1}$ are disjoint, and g'_{i+1} is the polynomial computed by r in Ψ^{i+1}.

Thus,

$$f^A = g_1 h_1 + g'_1 = g_1 h_1 + g_2 h_2 + g'_2 = \cdots = \sum_{i \in [\ell]} g_i h_i + g,$$

where for all $i \in [\ell]$, the set of variables that occur in g_i and the set $Y_i \cup Z_i$ are disjoint, and $g = g'_\ell$ is the polynomial computed by r in Ψ, after substituting (in Ψ) $u_1, \ldots, u_\ell$ by 0. □[Proposition 5.5]

The following claim shows that the partial derivative matrices of $g_1 h_1, \ldots, g_\ell h_\ell$ are of low rank.

Claim 5.7. *For every $i \in [\ell]$, Rank$(L_{g_i h_i}) \leq 2^{m-\tau}$.*

Proof. Fix $i \in [\ell]$. Denote by Y' the set of Y variables that occur in g_i, by Z' the set of Z variables that occur in g_i, and denote $a' = \min(|Y'|, |Z'|)$. By property 1 of Proposition 5.5, $(Y' \cup Z') \cap (Y_i \cup Z_i) = \emptyset$. Thus, $|Y'| + |Z'| \leq n - 2b(u_i)$, which implies $a' \leq m - b(u_i)$. Hence, by Proposition 2.2,

$$\text{Rank}(L_{g_i}) \leq 2^{a'} \leq 2^{m - b(u_i)}.$$

Since $u_i \in \mathcal{L}$, u_i is τ-unbalanced. Thus, $d(u_i) = b(u_i) - a(u_i) \geq \tau$. Hence, by Proposition 2.2,

$$\text{Rank}(L_{h_i}) \leq 2^{a(u_i)} \leq 2^{b(u_i) - \tau}.$$

Thus, since $(Y' \cup Z') \cap (Y_i \cup Z_i) = \emptyset$, by Proposition 2.4,

$$\mathrm{Rank}(L_{g_i h_i}) \leq 2^{m-b(u_i)+b(u_i)-\tau} = 2^{m-\tau}.$$

□[Claim 5.7]

The following proposition shows that the total degree of g is small.

Proposition 5.8. *The total degree of g is at most 4τ.*

Proof. Denote by Ψ^ℓ the arithmetic circuit Ψ, after substituting (in Ψ) each gate $u \in \mathcal{L}$ by 0. For a gate v in Ψ^ℓ, denote by $td(v)$ the total degree of the polynomial computed by v in Ψ^ℓ. Every gate in Ψ^ℓ is also a gate in Φ. For a gate v in Ψ^ℓ, define X_v to be the set of X variables that occur in Φ_v. Note that, every gate v in Ψ^ℓ admits $td(v) \leq |X_v|$.

The following claim shows that, if X_v is large , then $td(v)$ is small (where v is a gate in Ψ^ℓ).

Claim 5.9. *Let v be a gate in Ψ^ℓ, and let $k = n - |X_v|$. Assume that $k \leq 2\tau$. Then, $td(v) \leq 4\tau - k$.*

Proof. The proof is by induction on the structure of Ψ^ℓ (that is, we consider a gate v only after considering v's two sons). Since $|X_v| = n - k \geq n - 2\tau$, it follows that v is not an input gate in Ψ^ℓ. Let v_1 and v_2 be the two sons of v in Ψ^ℓ. Let $k_1 = n - |X_{v_1}|$ and $k_2 = n - |X_{v_2}|$. Since $X_v = X_{v_1} \cup X_{v_2}$, it follows that $k \leq k_1$ and $k \leq k_2$. Consider the following two cases:

Case one: v is an addition gate. First, we claim that $td(v_1) \leq 4\tau - k$ and $td(v_2) \leq 4\tau - k$. Consider v_1 without loss of generality. There are three cases:

a. Assume $n - 2\tau \leq |X_{v_1}|$. Thus, $k_1 \leq 2\tau$. Hence, by induction, $td(v_1) \leq 4\tau - k_1 \leq 4\tau - k$.

b. Assume $2\tau < |X_{v_1}| < n - 2\tau$. Since $|X_v| \geq n - 2\tau$, it follows that $v_1 \in \mathcal{L}$. Hence, v_1 is an input gate labelled by 0 in Ψ^ℓ, which implies $td(v_1) = 0 \leq 4\tau - k$.

c. Assume $|X_{v_1}| \leq 2\tau$. Since $k \leq 2\tau$, we have $td(v_1) \leq |X_{v_1}| \leq 2\tau \leq 4\tau - k$.

Hence, since v is an addition gate, $td(v) \leq \max(td(v_1), td(v_2)) \leq 4\tau - k$.

Case two: v is a product gate. Assume without loss of generality that $|X_{v_1}| \geq |X_{v_2}|$. Since $X_v = X_{v_1} \cup X_{v_2}$, it follows that $|X_{v_1}| \geq |X_v|/2 > 2\tau$ (for large enough n). Hence, there are two cases:

a. Assume $n - 2\tau \leq |X_{v_1}|$. Thus, $k_1 \leq 2\tau$. Hence, by induction, $td(v_1) \leq 4\tau - k_1$. Since Φ is syntactically multilinear, $|X_v| = |X_{v_1}| + |X_{v_2}|$, which implies $|X_{v_2}| = k_1 - k$. Thus, $td(v_2) \leq |X_{v_2}| \leq k_1 - k$. Hence, $td(v) = td(v_1) + td(v_2) \leq 4\tau - k_1 + k_1 - k = 4\tau - k$.

b. Assume $2\tau < |X_{v_1}| < n - 2\tau$. Since $|X_v| \geq n - 2\tau$, it follows that $v_1 \in \mathcal{L}$. Hence, v_1 is an input gate labelled by 0 in Ψ^ℓ, which implies $td(v) = 0 \leq 4\tau - k$. □[Claim 5.9]

Since $|X_r| \geq n - 2$, by Claim 5.9, it follows that $td(r) \leq 4\tau$. Since r computes g in Ψ^ℓ, the proposition follows. □[Proposition 5.8]

By Proposition 5.8, and Proposition 2.6, since $\tau = 3 \log n$, we have

$$\text{Rank}(L_g) \leq 2^{(4\tau+1)\log m} \leq 2^{\tau^3}.$$

By Proposition 5.5, $f^A = \sum_{i\in[\ell]} g_i h_i + g$. Thus, by Claim 5.7, and by Proposition 2.3,

$$\text{Rank}(L_{f^A}) \leq \sum_{i\in[\ell]} 2^{m-\tau} + 2^{\tau^3} < 2^{m-1},$$

where the last inequality holds for large enough n, as $\ell = |\mathcal{L}| < \frac{c}{\tau} n^{1/3}$ and $\tau = 3 \log n$.

□[Lemma 5.2]

6 The Construction

For a field $\mathbb{F}$ and a set of variables T, we denote by $\mathbb{F}[T]$ the *ring of polynomials* over the field $\mathbb{F}$ and the set of variables T, and we denote by $\mathbb{F}(T)$ the *field of rational functions* over $\mathbb{F}$ in the set of variables T. Let $X = \{x_1, \ldots, x_n\}$, $\Omega = \{\omega_1, \ldots, \omega_n\}$, $Y = \{y_1, \ldots, y_m\}$ and $Z = \{z_1, \ldots, z_m\}$ be four sets of variables (where $n = 2m$). Let $\mathbb{F}$ be a field, and let $\mathbb{G} = \mathbb{F}(\Omega)$ be the field of rational functions over $\mathbb{F}$ in the set of variables Ω. Note that, a polynomial in $\mathbb{F}[X, \Omega]$ can also be thought of as a polynomial in $\mathbb{G}[X]$.

In this section, we construct a polynomial $f \in \mathbb{F}[X, \Omega]$ such that

- Thinking of f as a polynomial in $\mathbb{G}[X]$, for every partition A of X to Y and Z, the partial derivative matrix of f^A has full rank over $\mathbb{G}$ (recall that, $f^A \in \mathbb{G}[Y, Z]$ is the polynomial f, after substituting every $x \in X$ by $A(x) \in Y \cup Z$).
- f is explicit in the sense that f is in the class VNP, which is Valiant's algebraic analogue of NP. Moreover, the coefficient of every monomial in f, as a polynomial in $\mathbb{F}[X, \Omega]$, is either 0 or 1.

6.1 Definition of f

For a set $B \subset [n]$ of size m, denote by $i_1, \ldots, i_m$ the elements of B in an increasing order (that is, $B = \{i_1, \ldots, i_m\}$ and $i_1 < \cdots < i_m$), and denote by $j_1, \ldots, j_m$ the elements of $[n] \setminus B$ in an increasing order (that is, $[n] \setminus B = \{j_1, \ldots, j_m\}$ and $j_1 < \cdots < j_m$). Define r_B, a multilinear monomial in $\mathbb{F}[\Omega]$, by $r_B = \prod_{\ell \in B} \omega_\ell$, and define g_B, a multilinear polynomial in $\mathbb{F}[X]$, by $g_B = \prod_{\ell \in [m]} (x_{i_\ell} + x_{j_\ell})$. Define

$$f = \sum_B r_B g_B, \tag{6.1}$$

where the sum is over all sets $B \subset [n]$ of size m. Thus, $f \in \mathbb{F}[X, \Omega]$ is a multilinear polynomial over the field $\mathbb{F}$ and the sets of variables X and Ω. We think of f also as a polynomial in $\mathbb{G}[X]$.

6.2 The Partial Derivative Matrix of f^A Has Full Rank

The following theorem states that, thinking of f as a polynomial in $\mathbb{G}[X]$, for any partition A of X to Y and Z, the partial derivative matrix of f^A has full rank. Formally,

Theorem 6.1. *Let $f \in \mathbb{G}[X]$ be the polynomial defined in (6.1). Then, for any partition A of X to Y and Z, the partial derivative matrix of f^A has full rank (over $\mathbb{G}$).*

We defer the proof of Theorem 6.1 to Section 6.4. We remark that, the larger the set Ω is, the simpler it is to construct a polynomial f that satisfies Theorem 6.1. For the purpose of our lower bound, we need Ω to be as small as possible (and Ω such that $|\Omega| = |X|$ suffices).

6.3 f is Explicit

In [V79], Valiant defined an algebraic theory, analogues to the theory of NP-completeness. The analogue of NP, according to Valiant's theory, is called VNP. In this section, we show that f is in the class VNP.

For simplicity, instead of using the formal definition of VNP, we use a criterion (given by Valiant) for a polynomial to be in VNP. Valiant's criterion states that a polynomial f is in VNP if the coefficient of a monomial in f can be computed efficiently; that is, there exists a polynomial time Turing machine M such that given as input the degrees of the variables in a monomial p, M outputs the coefficient of p in f (in fact, Valiant's criterion is stronger. For more details see Proposition 2.20 in [B]).

We use Valiant's criterion to show that f is in VNP. Let $r \in \mathbb{F}[\Omega]$ and $g \in \mathbb{F}[X]$ be two monic multilinear monomials. To prove that f is in VNP, we describe an efficient algorithm that outputs the coefficient of rg in f. The algorithm is as follows:

If the total degree of r is not m, then the coefficient of rg in f is 0. Therefore, assume that the total degree of r is m. Thus, r is of the form $r = \prod_{\ell \in B} \omega_\ell$, for a set $B \subset [n]$ of size m. Let $i_1, \ldots, i_m$ be the elements of B in an increasing order, and let $j_1, \ldots, j_m$ be the elements of $[n] \setminus B$ in an increasing order. Since

$$g_B = \prod_{\ell \in [m]} (x_{i_\ell} + x_{j_\ell}) = \sum_{D \subseteq [m]} \prod_{\ell \in D} x_{i_\ell} \prod_{\ell \in [m] \setminus D} x_{j_\ell},$$

it follows that the coefficient of rg in f is 1 iff g is of the form $g = \prod_{\ell \in D} x_{i_\ell} \prod_{\ell \in [m] \setminus D} x_{j_\ell}$, for some $D \subseteq [m]$ (otherwise, the coefficient is 0). Checking whether g is of the form $g = \prod_{\ell \in D} x_{i_\ell} \prod_{\ell \in [m] \setminus D} x_{j_\ell}$ is straightforward. Hence, f is in VNP.

6.4 Proof of Theorem 6.1

We first prove the following lemma, which is a generalization of Theorem 6.1.

Lemma 6.2. *Let T be a set of variables, and let $\mathbb{F}(T)$ be the field of rational functions over $\mathbb{F}$ in the set of variables T. Let $k \in \mathbb{N}$. Let $s_1, \ldots, s_k \in \mathbb{F}[T]$ be different monic multilinear monomials. Let $h_1, \ldots, h_k \in \mathbb{F}[Y, Z]$ be multilinear polynomials. Denote $h = \sum_{i \in [k]} s_i h_i$, a polynomial in $\mathbb{F}[Y, Z, T]$. Thus, h can be viewed also as a polynomial in $\mathbb{F}(T)[Y, Z]$. Assume that there exists $\ell \in [k]$ such that the partial derivative matrix of h_ℓ has full rank over $\mathbb{F}$. Then, the partial derivative matrix of h has full rank over $\mathbb{F}(T)$.*

Proof. The proof is by induction on the size of T.

Induction base: Assume $|T| = 0$. Since $s_1, \ldots, s_k$ are *different monic* monomials in $\mathbb{F}[T] = \mathbb{F}$, it follows that $k = 1$, and $s_1 = 1$. Hence, $h = h_1$, and the partial derivative matrix of h_1 has full rank over $\mathbb{F} = \mathbb{F}(T)$, which proves the lemma.

The induction step is based on the following claim.

Claim 6.3. *Let P and Q be two $M \times M$ matrices with entries in a field $\mathbb{H}$. Let t be a variable. Then,*

$$\exists\, a_1, \ldots, a_{M-1} \in \mathbb{H} \;:\; \det(tQ + P) = t^M \det(Q) + \det(P) + \sum_{d=1}^{M-1} a_d t^d,$$

where $\det(\cdot)$ is the determinant.

Proof. The claim follows by induction on M. For $M = 1$, we have $\det(tQ + P) = t\det(Q) + \det(P)$. Assume $M \geq 2$. For $i \in [M]$, denote by Q^i the matrix formed by eliminating the first row and the i'th column from Q, and denote by P^i the matrix formed by eliminating the first row and the i'th column from P. Thus, by induction, for every $i \in [M]$, there exist $a^i_1, \ldots, a^i_{M-2} \in \mathbb{H}$ such that

$$\begin{aligned}
\det(tQ + P) &= \sum_{i\in[M]} (-1)^{i+1}(tQ_{1,i} + P_{1,i}) \det(tQ^i + P^i) \\
&= \sum_{i\in[M]} (-1)^{i+1}(tQ_{1,i} + P_{1,i})\Big(t^{M-1}\det(Q^i) + \det(P^i) + \sum_{d=1}^{M-2} a^i_d t^d\Big) \\
&= t^M \det(Q) + \det(P) + \sum_{d=1}^{M-1} a_d t^d,
\end{aligned}$$

where $a_1, \ldots, a_{M-1} \in \mathbb{H}$. □[Claim 6.3]

Induction step: Assume $|T| > 0$. Let $t \in T$ be a variable. Denote $T' = T \setminus \{t\}$, and denote by $\mathbb{F}(T')$ the field of rational functions over $\mathbb{F}$ in the set of variables T'. Since $s_1, \ldots, s_k$ are multilinear monomials, assume without loss of generality that $h = t\sum_{i=1}^{k'} s'_i h_i + \sum_{i=k'+1}^{k} s_i h_i$, where $k' \in [k]$, $s'_1, \ldots, s'_{k'} \in \mathbb{F}[T']$ are different monic multilinear monomials, and $s_{k'+1}, \ldots, s_k \in \mathbb{F}[T']$ are different monic multilinear monomials. Recall that, the partial derivative matrix of h_ℓ has full rank over $\mathbb{F}$. Assume without loss of generality that $\ell \leq k'$ (similar arguments hold for $\ell > k'$). Denote $h' = \sum_{i=1}^{k'} s'_i h_i$ and $h'' = \sum_{i=k'+1}^{k} s_i h_i$. By induction, it follows that $L_{h'}$ has full rank over $\mathbb{F}(T')$, which implies that $\det(L_{h'}) \neq 0$. Hence, since $L_h = tL_{h'} + L_{h''}$, using Claim 6.3 (for $M = 2^m$, $Q = L_{h'}$, $P = L_{h''}$, and $\mathbb{H} = \mathbb{F}(T')$), we have that $\det(L_h) \neq 0$, which implies that L_h has full rank over $\mathbb{F}(T)$. □[Lemma 6.2]

Fix a partition A of X to Y and Z, and let $B_0 = \{i \in [n] : A(x_i) \in Y\}$. Recall that g_{B_0} is the polynomial defined in Section 6.1. The following claim shows that the partial derivative matrix of $g^A_{B_0}$ is a permutation matrix (a permutation matrix is a matrix obtained by permuting the rows of the identity matrix).

Claim 6.4. *The partial derivative matrix of $g^A_{B_0}$ is a permutation matrix.*

Proof. Denote by $i_1, \ldots, i_m$ the elements of B_0 in an increasing order, and denote by $j_1, \ldots, j_m$ the elements of $[n] \setminus B_0$ in an increasing order. By definition of g_{B_0},

$$g^A_{B_0} = \Big(\prod_{\ell\in[m]} \big(x_{i_\ell} + x_{j_\ell}\big)\Big)^A = \prod_{\ell\in[m]} \big(A(x_{i_\ell}) + A(x_{j_\ell})\big).$$

Note that, for every $\ell \in [m]$, we have $A(x_{i_\ell}) \in Y$ and $A(x_{j_\ell}) \in Z$. Thus, there exists a permutation $\pi : [m] \to [m]$ such that $g^A_{B_0} = \prod_{\ell \in [m]} \left(y_\ell + z_{\pi(\ell)}\right)$. Hence, the partial derivative matrix of $g^A_{B_0}$ is a permutation matrix. □[Claim 6.4]

By Claim 6.4, the set B_0 is such that the partial derivative matrix of $g^A_{B_0}$ has full rank over $\mathbb{F}$. Hence, using Lemma 6.2 (for $T = \Omega$), since $\{r_B\}_{B \subset [n]:|B|=m}$ is a set of different monic multilinear monomials in $\mathbb{F}[\Omega]$, it follows that the partial derivative matrix of f^A has full rank over the field $\mathbb{G}$. Since A is an arbitrary partition of X to Y and Z, the theorem follows. □[Theorem 6.1]

7 The Lower Bound - Proof of Theorem 1.1

Denote $\mathbb{G} = \mathbb{F}(\Omega)$, the field of rational functions over $\mathbb{F}$ in the set of variables Ω. We can think of Φ as a syntactically multilinear arithmetic circuit over the field $\mathbb{G}$ and the set of variables X: every input gate in Φ labelled by $\omega \in \Omega$ is thought of as labelled by a field element in $\mathbb{G}$, and every other input gate in Φ is labelled by either a field element in $\mathbb{F} \subseteq \mathbb{G}$ or by a variable in X. The number of variables in Φ is n (instead of $2n$). The polynomial computed by Φ is f, but we think of f as a polynomial in $\mathbb{G}[X]$. By Theorem 6.1, for all partitions A of X to Y and Z,

$$\mathrm{Rank}(L_{f^A}) = 2^m,$$

where the rank is over $\mathbb{G}$. Hence, by Theorem 5.1,

$$|\Phi| = \Omega\left(\frac{n^{4/3}}{\log^2 n}\right).$$

□[Theorem 1.1]

References

[BS83] W. Baur and V. Strassen. The Complexity of Partial Derivatives. *Theoretical Computer Science*, 22: 317-330, 1983.

[B] P. Burgisser. Completeness and Reduction in Algebraic Comlexity Theory. *Algorithms and Computation in Mathematics*, Volume 7, Springer - Verlag Berlin Heidelberg, 2000.

[M85] J. Morgenstern. How to Compute Fast a Function and all Its Derivatives, a Variation on the Theorem of Baur-Strassen. *SIGACT News*, 16: 60–62, 1985.

[N91] N. Nisan. Lower Bounds for Non-Commutative Computation. *Proceeding of the 23th STOC*: 410-418, 1991.

[NW96] N. Nisan and A. Wigderson. Lower Bound 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. *Proceeding of the 45th FOCS*: 344-351, 2004 (title: "Multilinear-$NC_1 \neq$ Multilinear-NC_2").

[RS05] R. Raz and A. Shpilka. Deterministic Polynomial Identity Testing in Non Commutative Models. *Journal of Computational Complexity*, 14:1-19, 2005.

[S73] V. Strassen. Die Berechnungskomplexitat von Elementarsymmetrischen Funktionen und von Interpolationskoeffizienten. *Numerische Mathematik*, 20: 238-251, 1973.

[V79] L. G. Valiant. Completeness Classes in Algebra. *Proceeding of the 11th STOC*: 249-261, 1979.