Multilinear-$NC_1 \neq$ Multilinear-$NC_2$

Ran Raz*
Weizmann Institute
ran.raz@weizmann.ac.il

Abstract

An arithmetic circuit or formula is multilinear if the polynomial computed at each of its wires is multilinear. We give an explicit example for a polynomial $f(x_1, ..., x_n)$, with coefficients in $\{0, 1\}$, such that over any field:

1. $f$ can be computed by a polynomial-size multilinear circuit of depth $O(\log^2 n)$.
2. Any multilinear formula for $f$ is of size $n^{O(\log n)}$.

This gives a super-polynomial gap between multilinear circuit and formula size, and separates multilinear $NC_1$ circuits from multilinear $NC_2$ circuits.

1 Introduction

An outstanding open problem in arithmetic circuit complexity is to understand the relative power of circuits and formulas. Surprisingly, any arithmetic circuit of size $s$ for a polynomial of degree $d$ can be translated into an arithmetic formula of size quasi-polynomial in $s$ and $d$ [H, VSBR]. Can such a circuit be translated into a formula of size polynomial in $s$ and $d$?

In this paper, we answer this question for multilinear circuits and formulas. An arithmetic circuit (or formula) is multilinear if the polynomial computed at each of its wires is multilinear (as a formal polynomial), that is, in each of its monomials the power of every input variable is at most one.

1.1 Multilinear Circuits

Let $F$ be a field, and let $\{x_1, ..., x_n\}$ be a set of input variables. An arithmetic circuit is a directed acyclic graph with nodes of in-degree 0 or 2. Every leaf of the graph (i.e., a node of in-degree 0) is labelled with either an input variable or a field element. Every other node

---

*Research supported by Israel Science Foundation (ISF) grant.

1 Moreover, if $s, d$ are both polynomial in the number of input variables $n$, then the circuit can be translated into a polynomial-size circuit of depth $O(\log^2 n)$, that is, an $NC_2$ circuit [VSBR].
of the graph is labelled with either + or × (in the first case the node is a plus gate and in
the second case a product gate). We assume that there is only one node of out-degree zero,
called the root. The circuit is a formula if its underlying graph is a (binary) tree (with edges
directed from the leaves to the root).

An arithmetic circuit computes a polynomial in the ring \( F[x_1, \ldots, x_n] \) in the following way.
A leaf just computes the input variable or field element that labels it. A plus gate computes
the sum of the two polynomials computed by its sons. A product gate computes the product
of the two polynomials computed by its sons. The output of the circuit is the polynomial
computed by the root. For a circuit \( \Phi \), we denote by \( \hat{\Phi} \) the output of the circuit, that is, the
polynomial computed by the circuit. The size of a circuit \( \Phi \) is defined to be the number of
nodes in the graph, and is denoted by \( |\Phi| \). The depth of a circuit is defined to be the maximal
distance between the root and a leaf in the graph.

A polynomial in the ring \( F[x_1, \ldots, x_n] \) is multilinear if in each of its monomials the power
of every input variable is at most one. An arithmetic circuit (or formula) is multilinear if the
polynomial computed by each gate of the circuit is multilinear.

### 1.2 Background

Multilinear circuits (and formulas) were formally defined by Nisan and Wigderson in [NW].
Obviously, multilinear circuits can only compute multilinear functions. Moreover, multilinear
circuits are restricted, as they do not allow the intermediate use of higher powers of variables
in order to finally compute a certain multilinear function. Note, however, that for many
multilinear functions, circuits that are not multilinear are very counter-intuitive, as they
require a "magical" cancellation of all high powers of variables. For many multilinear functions,
it seems "obvious" that the smallest circuits and formulas should be multilinear. Moreover,
for many multilinear functions, all (or almost all) known circuits are multilinear.

Super-polynomial lower bounds for the size of multilinear formulas were recently proved [R].
In particular, it was proved that over any field, any multilinear formula for the permanent
or the determinant of an \( n \times n \) matrix is of size \( n^{O(\log n)} \). Note, however, that all known
multilinear circuits for the permanent or the determinant are of exponential size, and hence
these bounds don’t give any separation between multilinear circuit and formula size.

For more background and motivation for the study of multilinear circuits and formulas
see [NW, R, A]. For general background on algebraic complexity theory see [G, BCS].

### 1.3 Our Results

We give an explicit example for a (multilinear) polynomial \( f(x_1, \ldots, x_n) \), with coefficients in
\( \{0, 1\} \), such that over any field:

1. \( f \) can be computed by a polynomial-size multilinear circuit of depth \( O(\log^2 n) \), that is,
a multilinear \( NC_2 \) circuit.
2. Any multilinear formula for $f$ is of size $n^{\Omega(\log n)}$. In particular, $f$ cannot be computed by a polynomial-size multilinear circuit of depth $O(\log n)$, that is, a multilinear $NC_1$ circuit.\footnote{Note that any (multilinear) circuit of depth $O(\log n)$ can trivially be translated into a polynomial size (multilinear) formula (of depth $O(\log n)$).}

This gives a super-polynomial gap between multilinear circuit and formula size, and separates multilinear $NC_1$ circuits from multilinear $NC_2$ circuits.

For the proof of our lower bound on the multilinear formula size of $f$, we use methods from [R]. The main contribution of this paper is the construction of a polynomial $f$ that can be computed by small multilinear circuits, and for which these methods can be applied.

## 2 Syntactic Multilinear Circuits

Let $\Phi$ be an arithmetic circuit over the set of variables $\{x_1, \ldots, x_n\}$. For every node $v$ in the circuit, denote by $\Phi_v$ the sub-circuit with root $v$, and denote by $X_v$ the set of variables that appear in the circuit $\Phi_v$. We say that an arithmetic circuit $\Phi$ is syntactic multilinear if for every product gate $v$ of $\Phi$, with sons $v_1, v_2$, the sets of variables $X_{v_1}$ and $X_{v_2}$ are disjoint.

Note that any syntactic multilinear circuit is clearly multilinear. At the other hand, a multilinear circuit is not necessarily syntactic multilinear. Nevertheless, the following proposition shows that without loss of generality we can assume that a multilinear formula is syntactic multilinear.

**Proposition 2.1** [R] For any multilinear formula, there exists a syntactic multilinear formula of the same size that computes the same polynomial.

**Proof:**

Let $\Phi$ be a multilinear formula. Let $v$ be a product gate in $\Phi$, with sons $v_1, v_2$, and assume that $X_{v_1}$ and $X_{v_2}$ both contain the same variable $x_i$. Since $\Phi$ is multilinear, $\Phi_v$ is a multilinear polynomial and hence in at least one of the polynomials $\Phi_{v_1}, \Phi_{v_2}$ the variable $x_i$ doesn’t appear. W.l.o.g. assume that in the polynomial $\Phi_{v_1}$ the variable $x_i$ doesn’t appear. Then every appearance of $x_i$ in $\Phi_{v_1}$ can be replaced by the constant 0. By repeating this for every product gate in the formula, as many times as needed, we obtain a syntactic multilinear formula that computes the same polynomial. \hfill $\square$

## 3 Lower Bounds for Multilinear Formulas

In this section, we prove general lower bounds for the size of multilinear formulas. To prove these bounds we use techniques from [R]. As in [R], our starting point is the partial derivatives method of [N, NW]. As in [R], to handle sets of partial derivatives, we make use of the partial derivatives matrix (first used in [N]).
3.1 The Partial-Derivatives Matrix

Let $f$ be a multilinear polynomial over the set of variables $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$. For a multilinear monomial $p$ in the set of variables $\{y_1, \ldots, y_m\}$ and a multilinear monomial $q$ in the set of variables $\{z_1, \ldots, z_m\}$, denote by $M_f(p, q)$ the coefficient of the monomial $pq$ in the polynomial $f$. Since the number of multilinear monomials in a set of $m$ variables is $2^m$, we can think of $M_f$ as a $2^m \times 2^m$ matrix, with entries in the field $F$. We will be interested in the rank of the matrix $M_f$ over the field $F$.

The following two propositions give some basic facts about the partial derivatives matrix.

**Proposition 3.1** Let $f, f_1, f_2$ be three multilinear polynomials over the set of variables $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$, such that $f = f_1 + f_2$. Then $M_f = M_{f_1} + M_{f_2}$.

**Proof:** Immediate from the definition of the partial derivatives matrix. \qed

**Proposition 3.2** Let $f, f_1, f_2$ be three multilinear polynomials over the set of variables $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$, such that $f = f_1 \cdot f_2$, and such that the set of variables that $f_1$ depends on and the set of variables that $f_2$ depends on are disjoint. Then, $\text{Rank}(M_f) = \text{Rank}(M_{f_1}) \cdot \text{Rank}(M_{f_2})$.

**Proof:** Note that the matrix $M_f$ is the tensor product of $M_{f_1}$ and $M_{f_2}$ (where all matrices are restricted to rows and columns that are non-zero). Hence, the rank of $M_f$ is the product of the rank of $M_{f_1}$ and the rank of $M_{f_2}$. \qed

Let $\Phi$ be a multilinear formula over the set of variables $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$. Recall that the output $\hat{\Phi}$ of the formula $\Phi$ is a multilinear polynomial over $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$. For simplicity, we denote the matrix $M_{\hat{\Phi}}$ also by $M_\Phi$. We will be interested in bounding the rank of the matrix $M_\Phi$ over the field $F$. (Note, however, that the rank of $M_\Phi$ may be as large as $2^m$ (i.e., full rank), even if the formula $\Phi$ is of linear size).

3.2 Unbalanced Nodes

Let $\Phi$ be a syntactic multilinear formula over the set of variables $\{y_1, \ldots, y_m\} \cup \{z_1, \ldots, z_m\}$. For every node $v$ in the formula, denote by $Y_v$ the set of variables in $\{y_1, \ldots, y_m\}$ that appear in the formula $\Phi_v$, and denote by $Z_v$ the set of variables in $\{z_1, \ldots, z_m\}$ that appear in the formula $\Phi_v$.

Denote by $b(v)$ the average of $|Y_v|$ and $|Z_v|$ and denote by $a(v)$ their minimum. Denote, $d(v) = b(v) - a(v)$. We say that a node $v$ is $k$-unbalanced if $d(v) \geq k$.

Let $\gamma$ be a simple path from a leaf $w$ to a node $v$ of the formula $\Phi$. We say that $\gamma$ is $k$-unbalanced if it contains at least one $k$-unbalanced node. We say that $\gamma$ is central if for

---

3We only consider monomials with coefficient 1.
every $u, u_1$ on the path $\gamma$, such that $u_1$ is a direct son of $u$ (i.e., there is an edge from $u_1$ to $u$), we have $b(u) \leq 2b(u_1)$. Note that for every node $u$ in the formula, with sons $u_1, u_2$, we have $b(u) \leq b(u_1) + b(u_2)$. Hence, by induction, for every node $u$ in the formula, there exists at least one central path that reaches $u$. In particular, at least one central path reaches the root.

We say that the formula $\Phi$ is $k$-weak if every central path that reaches the root of the formula contains at least one $k$-unbalanced node. The following lemma from [R] shows that if the formula $\Phi$ is $k$-weak then the rank of the matrix $M_\Phi$ can be bounded.

**Lemma 3.3** [R] Let $\Phi$ be a syntactic multilinear formula over the set of variables $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$, and assume that $\Phi$ is $k$-weak. Then,

$$\text{Rank}(M_\Phi) \leq |\Phi| \cdot 2^{m-k/2}.$$ 

### 3.3 Random Partition

Let $n = 2m$. Let $\Phi$ be a syntactic multilinear formula over the set of variables $X = \{x_1, ..., x_n\}$. Let $A$ be a random partition of the variables in $X$ into $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$. Formally, $A$ is a (randomly chosen) one to one function from the set of variables $X$ to the set of variables $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$.

Denote by $\Phi_A$ the formula $\Phi$ after replacing every variable of $X$ by the variable assigned to it by $A$. Obviously, $\Phi_A$ is a syntactic multilinear formula over the set of variables $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$.

The following lemma shows that if $|\Phi|$ is small then with high probability $\Phi_A$ is $k$-weak for $k = n^{1/8}$. We will give the proof of the lemma in the next section.

**Lemma 3.4** Let $n = 2m$. Let $\Phi$ be a syntactic multilinear formula over the set of variables $X = \{x_1, ..., x_n\}$, such that every variable in $X$ appears in $\Phi$, and such that $|\Phi| \leq n^{\epsilon \log n}$, where $\epsilon$ is a small enough universal constant (e.g., $\epsilon = 10^{-6}$). Let $A$ be a random partition of the variables in $X$ into $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$. Then, with probability of at least $1 - n^{-\Omega(\log n)}$ the formula $\Phi_A$ is $k$-weak, for $k = n^{1/8}$.

### 3.4 The Lower Bounds

Lower bounds for the size of multilinear formulas can be proved as a corollary of Lemma 3.3 and Lemma 3.4. We will prove lower bounds for functions that satisfy the following high rank property.\(^4\)

**Definition 3.5 (High Rank)** Let $n = 2m$. Let $f$ be a multilinear polynomial (over a field $F$) over the set of variables $X = \{x_1, ..., x_n\}$. We say that $f$ is of high rank over $F$ if the following is satisfied: Let $A$ be a random partition of the variables in $X$ into $\{y_1, ..., y_m\} \cup \{z_1, ..., z_m\}$. Then, with probability of at least $n^{-\omega(\log n)}$,

\(^4\)Note that the functions $f$ used in this paper will actually satisfy a much stronger property. Namely, for any partition $A$, we will have $\text{Rank}(M_f_A) = 2^m$ (where all notations are as in Definition 3.5).
\[ \text{Rank}(M_{f_A}) \geq 2^{m - m^{1/8}/2}, \]

where the rank is over \( F \), and \( f_A \) denotes the polynomial \( f \) after replacing every variable in \( X \) by the variable assigned to it by \( A \).

The following corollary is our basic lower bound.

**Corollary 3.6** Let \( n = 2m \). Let \( f \) be a multilinear polynomial (over a field \( F \)) over the set of variables \( X = \{x_1, ..., x_n\} \). If \( f \) is of high rank over \( F \) (see Definition 3.5) then for any multilinear formula \( \Phi \) for \( f \),

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

**Proof:**

By Proposition 2.1, we can assume w.l.o.g. that \( \Phi \) is syntactic multilinear. Note also that we can assume w.l.o.g. that all the variables in \( X \) appear in \( \Phi \), as we can always add variables multiplied by 0. Assume for a contradiction that \( |\Phi| \leq n^{\epsilon \log n} \), where \( \epsilon \) is the universal constant from Lemma 3.4. Let \( A \) be a random partition of the variables in \( X \) into \( \{y_1, ..., y_m\} \cup \{z_1, ..., z_m\} \). Then, by Lemma 3.4, with probability of at least \( 1 - n^{-\Omega(\log n)} \) the formula \( \Phi_A \) is \( k \)-weak, for \( k = n^{1/8} \).

Hence, by Lemma 3.3, with probability of at least \( 1 - n^{-\Omega(\log n)} \),

\[ \text{Rank}(M_{\Phi_A}) < 2^{m - m^{1/8}/2}. \]

Thus \( \Phi \) cannot be a formula for the high rank function \( f \).

We will now consider multilinear polynomials \( f \) (over a field \( F \)) over two sets of variables: \( X = \{x_1, ..., x_n\} \) and \( X' = \{x'_1, ..., x'_l\} \). We think of the variables in \( X' \) as auxiliary variables. Let \( A' : X' \rightarrow F \) be an assignment of values in \( F \) to all the variables in \( X' \). We denote by \( f_{A'} \) the polynomial \( f \), after substituting in every variable in \( X' \) the value assigned to it by \( A' \). Note that \( f_{A'} \) is a multilinear polynomial over the set of variables \( X \).

**Corollary 3.7** Let \( n = 2m \). Let \( f \) be a multilinear polynomial (over a field \( F \)) over the sets of variables \( X = \{x_1, ..., x_n\} \) and \( X' = \{x'_1, ..., x'_l\} \). If for some assignment \( A' : X' \rightarrow F \) the polynomial \( f_{A'} \) is of high rank over \( F \) (see Definition 3.5) then for any multilinear formula \( \Phi \) for \( f \),

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

**Proof:**

Denote by \( \Phi_{A'} \) the formula \( \Phi \) after replacing every variable of \( X' \) by the value assigned to it by \( A' \). Then, \( \Phi_{A'} \) is a formula for \( f_{A'} \), and \( |\Phi_{A'}| = |\Phi| \). Hence, by Corollary 3.6, \( |\Phi| = |\Phi_{A'}| \geq n^{\Omega(\log n)}. \)

In some cases, in order to find an assignment \( A' \) such that the polynomial \( f_{A'} \) is of high rank, we will need to consider extensions \( G \) of the field \( F \). Note that any polynomial \( f \) over \( F \) is also a polynomial over any field extending \( F \).
Corollary 3.8 Let \( n = 2m \). Let \( f \) be a multilinear polynomial (over a field \( F \)) over the sets of variables \( X = \{ x_1, \ldots, x_n \} \) and \( X' = \{ x'_1, \ldots, x'_1 \} \). If for some field \( G \supset F \) there exists an assignment \( A' : X' \rightarrow G \), such that the polynomial \( f_{A'} \) is of high rank over \( G \) (see Definition 3.5) then for any multilinear formula \( \Phi \) for \( f \) (over the field \( F \)),
\[
|\Phi| \geq n^{\Omega(\log n)}.
\]

Proof:
Any multilinear formula for \( f \) over the field \( F \) is also a multilinear formula for \( f \) over the field \( G \). The proof hence follows by Corollary 3.7. \( \square \)

4 Proof of Lemma 3.4

Let us first give a brief sketch of the proof. Note that the intuition and the basic structure of the proof are the same as in [R], but the details here are much simpler.

Intuitively, since \( A \) is random, every node \( v \) with large enough \( X_v \) will be \( k \)-unbalanced with high probability. The probability that such \( v \) is not \( k \)-unbalanced is smaller than \( O(n^{-\delta}) \), for some constant \( \delta \). This may not be enough since the number of central paths is possibly as large as \( n^{\log n} \). Nevertheless, each central path contains \( \Omega(\log n) \) nodes so we can hope to prove that the probability that none of them is \( k \)-unbalanced is as small as \( n^{-\Omega(\log n)} \).

This, however, is not trivial since there are dependencies between the different nodes. We will identify \( \Omega(\log n) \) nodes, \( v_1, \ldots, v_l \), on the path (that will be "far enough" from each other). We will show that for every \( v_i \), the probability that \( v_i \) is not \( k \)-unbalanced is smaller than \( O(n^{-\delta}) \), even when conditioning on the event that \( v_1, \ldots, v_{i-1} \) are not \( k \)-unbalanced.

4.1 Notations

For any integer \( n \), denote \( [n] = \{ 1, \ldots, n \} \).

To simplify notations, we denote in this section the formula \( \Phi_A \) by \( \Psi \). There is a one to one correspondence between the nodes of \( \Phi \) and the nodes of \( \Psi \). For every node \( v \) in \( \Phi \), there is a corresponding node in \( \Psi \) and vice versa. For simplicity, we denote both these nodes by \( v \), and we think of them as the same node. Hence, \( X_v \) denotes the set of variables in \( X \) that appear in the formula \( \Phi_v \), while \( Y_v \) denotes the set of variables in \( \{ y_1, \ldots, y_m \} \) that appear in the formula \( \Psi_v \), and \( Z_v \) denotes the set of variables in \( \{ z_1, \ldots, z_m \} \) that appear in \( \Psi_v \). Denote,
\[
\alpha(v) = |X_v|/n.
\]

For three integers \( M_1, M_2 \leq N \), denote by \( H(N, M_1, M_2) \) the hypergeometric distribution with parameters \( N, M_1, M_2 \), that is, the distribution of the size of the intersection of a random set of size \( M_2 \) and a set of size \( M_1 \) in a universe of size \( N \).
Proposition 4.1 Let \( \chi \) be a random variable that has the hypergeometric distribution \( \mathcal{H}(N, M_1, M_2) \), such that \( N/4 \leq M_2 \leq 3N/4 \), and \( N^{1/2} \leq M_1 \leq N/2 \). Then, \( \chi \) gets any specific value with probability of at most \( O(N^{-1/4}) \).

Proof:
Follows by the definition of the hypergeometric distribution and standard bounds on binomial coefficients. \( \square \)

4.2 Central Paths are Unbalanced

Let \( \gamma \) be a simple path from a leaf to a node in \( \Phi \). Note that \( \gamma \) is central in \( \Psi \) iff for every \( u, u_1 \) on the path \( \gamma \), such that \( u_1 \) is a direct son of \( u \), we have \( \alpha(u) \leq 2\alpha(u_1) \). Since this property doesn’t depend on the partition \( A \), we say in this case that \( \gamma \) is central in \( \Phi \). We will show that if \( \gamma \) is central then with high probability \( \gamma \) is unbalanced in the formula \( \Psi \).

Claim 4.2 Let \( \gamma \) be a central path from a leaf to the root of \( \Phi \). Then,
\[
\Pr[\gamma \text{ is not } k\text{-unbalanced in } \Psi] \leq n^{-\Omega(\log n)}.
\]

Proof:
Recall that the first node of \( \gamma \) is a leaf and hence \( \alpha(v) \) for that node is at most \( 1/n \), and the last node of \( \gamma \) is the root and hence \( \alpha(v) \) for that node is 1. Note that \( \alpha(v) \) is monotonically increasing along \( \gamma \). Let \( v_1, \ldots, v_l \) be nodes on \( \gamma \), chosen by the following process: Let \( v_1 \) be the first node on \( \gamma \), such that \( \alpha(v_1) \geq n^{-1/2} \). For every \( i \), let \( v_{i+1} \) be the first node on \( \gamma \), such that \( \alpha(v_{i+1}) \geq 2 \cdot \alpha(v_i) \). Stop when \( \alpha(v_{i+1}) > 1/4 \). Denote by \( l \) the index \( i \) of the last \( v_i \) in this process.

Since \( \gamma \) is central, for every \( u, u' \) on \( \gamma \), such that \( u' \) is a direct son of \( u \), we have \( \alpha(u) \leq 2\alpha(u') \). Hence, for every \( i \in [l-1] \), we have \( \alpha(v_{i+1}) < 4 \cdot \alpha(v_i) \). Hence, the process above continues for \( \Omega(\log n) \) steps. To summarize, we have \( l = \Omega(\log n) \) and nodes \( v_1, \ldots, v_l \) on \( \gamma \), such that for every \( i \in \{2, \ldots, l\} \),
\[
1/4 \geq \alpha(v_i) \geq 2 \cdot \alpha(v_{i-1}) \geq n^{-1/2}.
\]

Denote by \( \mathcal{E} \) the event that \( \gamma \) is not \( k \)-unbalanced in the formula \( \Psi \). For every \( i \in [l] \), denote by \( \mathcal{E}_i \) the event that the node \( v_i \) is not \( k \)-unbalanced in the formula \( \Psi \). Since \( \mathcal{E} \subset \cap_{i \in [l]} \mathcal{E}_i \),
\[
\Pr[\mathcal{E}] \leq \Pr \left[ \bigcap_{i \in [l]} \mathcal{E}_i \right] = \prod_{i \in [l]} \Pr \left[ \mathcal{E}_i \mid \cap_{i' \in [i-1]} \mathcal{E}_{i'} \right]
\]
We will bound for every \( i > 1 \) the conditional probability \( \Pr[\mathcal{E}_i \mid \cap_{i' \in [i-1]} \mathcal{E}_{i'}] \).

Fix \( i \in \{2, \ldots, l\} \). Note that \( X_{v_{i-1}} \subset X_{v_i} \). Given the set \( Y_{v_{i-1}} \), we can write,
\[
|Y_{v_i}| = |Y_{v_{i-1}}| + \chi,
\]

8
Multilinear-

In this section, we present our construction for a multilinear polynomial which is smaller than any specific value with probability larger than $O(n^{-1/4})$, even when conditioning on (the content of) the set $Y_{v_{i-1}}$.

Note that the event $\cap_{i' \in [i-1]} E_{i'}$ depends only on the content of the set $Y_{v_{i-1}}$. Therefore, $|Y_{v_i}|$, and hence also $d(v_i)$, do not get any specific value with probability larger than $O(n^{-1/4})$, even when conditioning on the event $\cap_{i' \in [k-1]} E_{i'}$. Recall that $v_i$ is not $k$-unbalanced iff $d(v_i) < k$. Since $d(v_i)$ is integer, the probability for that is at most $O(k \cdot n^{-1/4}) = O(n^{-1/8})$, even when conditioning on the event $\cap_{i' \in [k-1]} E_{i'}$. That is,

$$\Pr \left[ \mathcal{E}_i \mid \cap_{i' \in [i-1]} E_{i'} \right] \leq O(n^{-1/8})$$

We can now bound

$$\Pr[\mathcal{E}] \leq \prod_{i \in [l]} \Pr \left[ \mathcal{E}_i \mid \cap_{i' \in [i-1]} E_{i'} \right] = n^{-\Omega(\log n)}$$

We can now complete the proof of Lemma 3.4. By Claim 4.2, if $\gamma$ is a central path from a leaf to the root of $\Phi$, then $\gamma$ is not $k$-unbalanced (in $\Psi$) with probability of at most $n^{-\Omega(\log n)}$. The number of paths from a leaf to the root of $\Phi$ is the same as the number of leaves in $\Phi$, which is smaller than $n^{\epsilon \log n}$ (and we assumed that $\epsilon$ is small enough). Hence, by the union bound, with probability of at least $1 - n^{-\Omega(\log n)}$ all central paths from a leaf to the root of $\Psi$ are $k$-unbalanced, that is, the formula $\Psi$ is $k$-weak.

\[\square\]

## 5 Multilinear-$NC_1 \neq$ Multilinear-$NC_2$

In this section, we present our construction for a multilinear polynomial $f$ that has polynomial-size multilinear circuits and doesn’t have polynomial-size multilinear formulas. Let us start with some notations.

Denote, $[n] = \{1, ..., n\}$. For every $i, j \in [n]$ such that $i \leq j$, denote by $[i, j]$ the interval of $[n]$ starting at $i$ and ending at $j$, that is, $[i, j] = \{i, i+1, ..., j\}$. Denote by $\mathcal{S}$ the set of all such intervals, including the empty interval (which is denoted by $\emptyset$). For $s_1, s_2 \in \mathcal{S}$, such that $s_1, s_2$ are disjoint and $s_2$ is consecutive\(^\text{5}\) to $s_1$, denote by $s_1 \circ s_2$ their concatenation, that is, if $s_1 = [i, j]$, and $s_2 = [j+1, j']$ then $s_1 \circ s_2 = [i, j']$.

Denote by $\mathcal{T}$ the set of (ordered) pairs of disjoint intervals in $\mathcal{S}$, that is,

$$\mathcal{T} = \{(s_1, s_2) \in \mathcal{S} \times \mathcal{S} : s_1 \cap s_2 = \emptyset\}.$$ 

\(^5\)We think of the empty interval as consecutive to every interval, and every interval is consecutive to it.

\[9\]
For $t_1, t_2 \in \mathcal{T}$, such that, $t_1 = (s_{1,1}, s_{1,2}), t_2 = (s_{2,1}, s_{2,2})$, and such that $s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}$ are all disjoint and $s_{2,1}$ is consecutive to $s_{1,1}$ and $s_{2,2}$ is consecutive to $s_{1,2}$, denote by $t_1 \circ t_2$ their pairwise concatenation, that is, $t_1 \circ t_2 = (s_{1,1} \circ s_{2,1}, s_{1,2} \circ s_{2,2}) \in \mathcal{T}$.

For every $s \in \mathcal{S}$, denote by $l(s)$ its length (i.e., the number of elements in it). For $t = (s_1, s_2) \in \mathcal{T}$, denote $l(t) = l(s_1) + l(s_2)$. For $t = (s_1, s_2) \in \mathcal{T}$, define $L(t) = l(t)$ if both $s_1, s_2$ are non-empty, and $L(t) = 0.75 \cdot l(t)$ if either $s_1$ or $s_2$ is empty.

For $t_1, t_2, t_3 \in \mathcal{T}$, such that, $t_1 = (s_{1,1}, s_{1,2}), t_2 = (s_{2,1}, s_{2,2}), t_3 = (s_{3,1}, s_{3,2}), t = (s_1, s_2)$, we say that $\{t_1, t_2\}$ is a partition of $t$ if $\{s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}\}$ is a partition of $s_1 \cup s_2$ as sets. We say that the partition is proper if $t = t_1 \circ t_2$, and $l(t_1), l(t_2) > 0$. In the same way, $\{t_1, t_2, t_3\}$ is a partition of $t$ if $\{s_{1,1}, s_{1,2}, s_{2,1}, s_{2,2}, s_{3,1}, s_{3,2}\}$ is a partition of $s_1 \cup s_2$ as sets. The partition is proper if $t = t_1 \circ t_2 \circ t_3$, and $l(t_1), l(t_2), l(t_3) > 0$.

For a function $A : [n] \rightarrow \{1, -1\}$ and for $s \in \mathcal{S}$, denote by $A(s)$ the sum of $A$ on the elements in $s$. In the same way, for $t \in \mathcal{T}$, denote by $A(t)$ the sum of $A$ on the elements in the union of the two intervals in $t$. We say that $A$ is balanced on $s \in \mathcal{S}$ if $A(s) = 0$, and in the same way, $A$ is balanced on $t \in \mathcal{T}$ if $A(t) = 0$. Denote by $\mathcal{B}_A$ the set of all $t \in \mathcal{T}$ on which $A$ is balanced, that is,

$$\mathcal{B}_A = \{t \in \mathcal{T} : A(t) = 0\}.$$  

Obviously, the length $l(t)$ of every $t \in \mathcal{B}_A$ is even.

**Lemma 5.1** Let $A$ be a function $A : [n] \rightarrow \{1, -1\}$. Let $t \in \mathcal{B}_A$ be such that $l(t) > 2$. Then, there exist $t_1, t_2, t_3 \in \mathcal{B}_A$, such that $\{t_1, t_2, t_3\}$ is a partition of $t$, and $L(t_1), L(t_2), L(t_3) \leq 0.75 \cdot L(t)$.

For any $t \in \mathcal{T}$, denote by $\mathcal{P}(t)$ the set of all $\{t_1, t_2, t_3\}$, such that, $t_1, t_2, t_3 \in \mathcal{T}$, and $\{t_1, t_2, t_3\}$ is a partition of $t$, and $L(t_1), L(t_2), L(t_3) \leq 0.75 \cdot L(t)$.

### 5.1 Proof of Lemma 5.1

Before giving the proof of Lemma 5.1, we will need to prove two other lemmas.

**Lemma 5.2** Let $A$ be a function $A : [n] \rightarrow \{1, -1\}$. Let $t = (s_1, s_2) \in \mathcal{B}_A$ be such that $l(t) > 2$ and $l(s_1), l(s_2) > 0$. Then, there exist $t_1, t_2 \in \mathcal{B}_A$, such that $\{t_1, t_2\}$ is a proper partition of $t$.

**Proof:**

Denote $s_1 = [i_1, j_1], s_2 = [i_2, j_2]$.

Since $t \in \mathcal{B}_A$, we have $A(s_1) + A(s_2) = 0$. If $A(s_1) = A(s_2) = 0$ then we can define $t_1 = (s_1, \emptyset), t_2 = (\emptyset, s_2)$. Otherwise, we can assume w.l.o.g. that $A(s_1)$ is negative and $A(s_2)$ is positive.

If $A(i_1) \neq A(i_2)$, we can define $t_1 = ([i_1, i_1], [i_2, i_2]), t_2 = ([i_1 + 1, j_1], [i_2 + 1, j_2])$. Otherwise, we can assume w.l.o.g. that $A(i_1) = A(i_2) = 1$.

Since $A(s_1)$ is negative and $A(i_1) = 1$, there must exist $j' \in s_1$, such that $A([i_1, j']) = 0$. We can then define $t_1 = ([i_1, j'], \emptyset), t_2 = ([j' + 1, j_1], s_2)$.
Since we required \( l(t) > 2 \) and \( l(s_1), l(s_2) > 0 \), we have in all cases \( l(t_1), l(t_2) > 0 \), and hence \( \{t_1, t_2\} \) is a proper partition of \( t \).

\[ \square \]

Lemma 5.3 Let \( A \) be a function \( A : [n] \to \{1, -1\} \). Let \( t = (s_1, s_2) \in B_A \) be such that \( l(t) > 2 \) and \( l(s_1), l(s_2) > 0 \). Then, there exist \( t_1, t_2, t_3 \in B_A \), such that \( \{t_1, t_2, t_3\} \) is a partition of \( t \), and

1. \( L(t_1), L(t_3) \leq 0.5 \cdot L(t) \).
2. \( L(t_2) \leq 0.75 \cdot L(t) \).
3. \( l(t_2) \leq \max(l(s_1), l(s_2)) \).

\[ \text{Proof:} \]

First note that since \( l(s_1), l(s_2) > 0 \), we have \( L(t) = l(t) \), and since \( l(t) > 2 \) and is even, \( L(t) = l(t) \geq 4 \). We will describe a procedure for finding \( t_1, t_2, t_3 \) with the required properties.

We start with \( \hat{t}_1 = (\emptyset, \emptyset) \), \( \hat{t}_2 = (s_1, s_2) \) and \( \hat{t}_3 = (\emptyset, \emptyset) \). Note that \( t = \hat{t}_1 \circ \hat{t}_2 \circ \hat{t}_3 \).

Claim 5.4 Let \( t_1', t_2', t_3' \in B_A \) be such that \( t = t_1' \circ t_2' \circ t_3' \). Assume that \( l(t_1'), l(t_3') \leq 0.5 \cdot l(t) \) and that both intervals in \( t_2' \) are non-empty, and \( l(t_2') > 2 \). Then, there exist \( t_1'', t_2'', t_3'' \in B_A \), such that \( t = t_1'' \circ t_2'' \circ t_3'' \), and \( l(t_1''), l(t_3'') \leq 0.5 \cdot l(t) \), and \( l(t_2'') < l(t_2') \).

\[ \text{Proof:} \]

By Lemma 5.2 (applied to \( t_2' \) ), there exist \( \tilde{t}_1, \tilde{t}_3 \in B_A \), such that \( \{\tilde{t}_1, \tilde{t}_3\} \) is a proper partition of \( t_2' \). Since \( t = t_1' \circ t_2' \circ t_3' \) and since \( t_2' = \tilde{t}_1 \circ \tilde{t}_3 \), we have \( t = t_1' \circ \tilde{t}_1 \circ \tilde{t}_3 \circ t_3' \).

If \( l(t_1') + l(\tilde{t}_1) \leq 0.5 \cdot l(t) \) then we can define \( t_1'' = t_1' \circ \tilde{t}_1 \), \( t_2'' = \tilde{t}_2 \), \( t_3'' = t_3' \). Otherwise, \( l(\tilde{t}_3) + l(t_3') \leq 0.5 \cdot l(t) \), and we can define \( t_1'' = t_1', \ t_2'' = \tilde{t}_2 \), \( t_3'' = \tilde{t}_3 \circ t_3' \).

Since \( \{\tilde{t}_1, \tilde{t}_3\} \) is a proper partition of \( t_2' \), in both cases \( l(t_2'') < l(t_2') \).

We now continue with the proof of Lemma 5.3. We apply Claim 5.4 on \( t_1' = \hat{t}_1, \ t_2' = \hat{t}_2, \ t_3' = \hat{t}_3 \), and we substitute (i.e., redefine) \( \hat{t}_1 = t_1'', \hat{t}_2 = t_2'', \hat{t}_3 = t_3'' \). We keep applying Claim 5.4 and substituting in \( \hat{t}_1, \hat{t}_2, \hat{t}_3 \), until the conditions of Claim 5.4 are not satisfied by \( \hat{t}_1, \hat{t}_2, \hat{t}_3 \), namely, either \( l(\hat{t}_2) \leq 2 \) or one of the intervals in \( \hat{t}_2 \) is empty. (Note that the process must stop because \( l(\hat{t}_2) \) keeps decreasing). At this point we can define \( t_1 = \hat{t}_1, \ t_2 = \hat{t}_2, \ t_3 = \hat{t}_3 \).

Since \( t_1, t_2, t_3 \) are the output of Claim 5.4, \( t_1, t_2, t_3 \in B_A \), and \( \{t_1, t_2, t_3\} \) is a partition of \( t \), and \( L(t_1), L(t_3) \leq 0.5 \cdot l(t) = 0.5 \cdot L(t) \). It remains to prove that \( L(t_2) \leq 0.75 \cdot L(t) \), and \( l(t_2) \leq \max(l(s_1), l(s_2)) \). Recall that there were two possibilities: either \( l(t_2) \leq 2 \) or one of the intervals in \( t_2 \) is empty.

In the first case, \( L(t_2) \leq l(t_2) \leq 2 \). Since \( L(t) = l(t) \geq 4 \), we have in the first case, \( L(t_2) \leq 0.5 \cdot L(t) \), and \( l(t_2) \leq 0.5 \cdot l(t) \leq \max(l(s_1), l(s_2)) \).

In the second case, \( L(t_2) = 0.75 \cdot l(t_2) \leq 0.75 \cdot l(t) = 0.75 \cdot L(t) \). Since the non-empty interval of \( t_2 \) is a sub-interval of either \( s_1 \) or \( s_2 \), we have \( l(t_2) \leq \max(l(s_1), l(s_2)) \).
Proof of Lemma 5.1:
Denote \( t = (s_1, s_2) \). If \( l(s_1), l(s_2) > 0 \), then the proof follows by Lemma 5.3. Otherwise, one of the intervals \( s_1, s_2 \) is empty. W.l.o.g. assume that \( s_1 \) is empty. Then, since \( t \in B_A \), we know that \( l(s_2) = l(t) \) is even. Partition \( s_2 \) into two intervals \( \{s'_1, s'_2\} \) with \( l(s'_1) = l(s'_2) = 0.5 \cdot l(s_2) \). The proof now follows by applying Lemma 5.3 on \( t' = (s'_1, s'_2) \) as follows.

Note that \( L(t) = 0.75 \cdot l(t) = 0.75 \cdot l(t') = 0.75 \cdot L(t') \). By Lemma 5.3 there exist \( t_1, t_2, t_3 \in B_A \), such that \( \{t_1, t_2, t_3\} \) is a partition of \( t' \) (and hence also of \( t \)), and \( L(t_1), L(t_3) \leq 0.5 \cdot L(t') < 0.75 \cdot L(t) \), and \( L(t_2) \leq l(t_2) \leq \max(l(s'_1), l(s'_2)) = 0.5 \cdot l(t) < 0.75 \cdot L(t) \). \( \square \)

### 5.2 The Construction

We will now define our multilinear polynomial \( f \) (with coefficients in \( \{0, 1\} \)), such that over any field, \( f \) can be computed by a polynomial-size multilinear circuit and cannot be computed by a polynomial-size multilinear formula. \( f \) will be defined over the set of variables \( X = \{x_1, \ldots, x_n\} \) (where \( n \) is even) and a set of (auxiliary) variables

\[ X' = \left\{ x'_{t, t_1, t_2, t_3} \right\}_{t, t_1, t_2, t_3 \in T} \]

That is, for every \( t, t_1, t_2, t_3 \in T \) we have an (auxiliary) variable \( x'_{t, t_1, t_2, t_3} \). Note that the total number of auxiliary variables is polynomial in \( n \).

\( f \) will be defined in the following way. For every \( t \in T \), such that \( l(t) \) is even, we will define a multilinear polynomial \( f_t \). We then define

\[ f = f_{\{0, \emptyset\}}. \]

We define the polynomials \( f_t \) by induction on \( L(t) \):

**Case 1:** \( L(t) = l(t) = 0 \). We define in this case, \( f_t = 1 \).

**Case 2:** \( 0 < L(t) \leq 2 \). Since \( l(t) \) is even, \( l(t) = 2 \). Hence, the union of the two intervals in \( t \) contains two indices. Denote these indices by \( i_t, j_t \). We define in this case,

\[ f_t = x_{i_t} \cdot x_{j_t} + 1. \]

Note that for the two possible partitions of \( \{x_{i_t}, x_{j_t}\} \) into \( \{y_1\} \cup \{z_1\} \), the partial derivatives matrix of \( f_t \) is the identity matrix of size \( 2 \times 2 \) and is hence of rank 2 (i.e., full rank).

**Case 3:** \( L(t) > 2 \). Since \( l(t) \) is even, \( l(t) \) is at least 4. We define in this case,

\[ f_t = \sum_{\{t_1, t_2, t_3\} \in P(t)} x_{t, t_1, t_2, t_3} \cdot f_{t_1} \cdot f_{t_2} \cdot f_{t_3}. \]

Observe that (by the inductive definition) for any \( \{t_1, t_2, t_3\} \) that give a partition of \( t \), the polynomials \( f_{t_1}, f_{t_2}, f_{t_3} \) depend on disjoint sets of variables. Hence, since we only sum over \( \{t_1, t_2, t_3\} \) that give partitions of \( t \), it follows by induction that the polynomial \( f_t \) is multilinear.
5.3 Upper Bound

The inductive definition of \( f \) gives a syntactic multilinear circuit for \( f \). Note that since we defined an arithmetic circuit to be of fan-in (i.e., in-degree) 2 (see Subsection 1.1), we need to replace the sum in the definition of each \( f_t \) by a tree of depth \( O(\log n) \) of addition gates (of in-degree 2).

The final circuit is of size polynomial in \( n \), since the size of \( T \) (and hence also the size of \( X' \) and the size of \( P(t) \) for every \( t \in T \)) is polynomial in \( n \).

The circuit is of depth \( O(\log^2 n) \), since in the definition of \( f_t \) we only sum over \( \{t_1, t_2, t_3\} \) with \( L(t_1), L(t_2), L(t_3) \leq 0.75 \cdot L(t) \) and since \( L([n], \emptyset) < n \). (Note that this gives a depth of \( O(\log n) \), but since we replace every sum by a tree of depth \( O(\log n) \) of addition gates we get another factor of \( O(\log n) \)).

**Corollary 5.5** Over any field \( F \), the polynomial \( f \) (as defined above) can be computed by a polynomial-size syntactic multilinear circuit of depth \( O(\log^2 n) \).

5.4 Lower Bound

We will now show that any multilinear formula for \( f \), over any field \( F \), is of size \( n^\Omega(\log n) \). For the proof, we use Corollary 3.8.

Denote \( n = 2m \). Let \( G \) be a field extending \( F \), such that the transcendental dimension of \( G \) over \( F \) is infinite, that is, \( G \) contains an infinite number of elements that are algebraically independent over \( F \). Define \( A' : X' \rightarrow G \) to be such that the variables in \( X' \) are mapped to elements that are algebraically independent over \( F \).

Let \( A \) be any partition of the variables in \( X \) into \( \{y_1, ..., y_m\} \cup \{z_1, ..., z_m\} \). Denote by \( f_{A',A} \) the polynomial \( f \) after substituting in every variable in \( X' \) the value assigned to it by \( A' \) and after replacing every variable in \( X \) by the variable assigned to it by \( A \).

**Claim 5.6** Over the field \( G \),

\[
\text{Rank}(M_{f_{A',A}}) = 2^m.
\]

**Proof:**

In this proof, the \textbf{Rank} function is always taken over the field \( G \). For simplicity, we denote in this proof by \( g \) the polynomial \( f_{A',A} \), and for every \( t \) we denote by \( g_t \) the polynomial \( f_{t,A',A} \) (i.e., the polynomial \( f_t \) after substituting in every variable in \( X' \) the value assigned to it by \( A' \) and after replacing every variable in \( X \) by the variable assigned to it by \( A \)).

Define the function \( \tilde{A} : [n] \rightarrow \{1, -1\} \) by \( \tilde{A}(i) = 1 \) if \( A(x_i) \in \{y_1, ..., y_m\} \) and \( \tilde{A}(i) = -1 \) if \( A(x_i) \in \{z_1, ..., z_m\} \). For simplicity, we denote the set \( B_A \) also by \( B_{\tilde{A}} \). We will prove by induction on \( L(t) \) that for every \( t \in B_{\tilde{A}} \),

\[
\text{Rank}(M_{g_t}) \geq 2^{l(t)/2}.
\]

For \( L(t) = l(t) = 0 \), we defined \( g_t = 1 \). Hence, \( M_{g_t} \) is the \( 1 \times 1 \) identity matrix and its
rank is 1. For $0 < L(t) \leq 2$, we know that $l(t) = 2$, and we defined $f_t = x_{it} \cdot x_{jt} + 1$. Since $t \in B_A$, the matrix $M_{g_t}$ is the $2 \times 2$ identity matrix and its rank is 2.

For $L(t) > 2$, 
$$f_t = \sum_{\{t_1, t_2, t_3\} \in \mathcal{P}(t)} x'_{t_1, t_2, t_3} \cdot f_{t_1} \cdot f_{t_2} \cdot f_{t_3}. $$

Hence, 
$$g_t = \sum_{\{t_1, t_2, t_3\} \in \mathcal{P}(t)} A'(x'_{t_1, t_2, t_3}) \cdot g_{t_1} \cdot g_{t_2} \cdot g_{t_3}, $$

and by Proposition 3.1,
$$M_{g_t} = \sum_{\{t_1, t_2, t_3\} \in \mathcal{P}(t)} A'(x'_{t_1, t_2, t_3}) \cdot M_{g_{t_1}} \cdot M_{g_{t_2}} \cdot M_{g_{t_3}}. $$

Therefore, since every $A'(x'_{t_1, t_2, t_3})$ is algebraically independent (over $F$) of all the other elements in the domain of $A'$ and all the coefficients that appear in any of the matrices in the sum\(^6\),
$$\text{Rank}(M_{g_t}) \geq \max_{\{t_1, t_2, t_3\} \in \mathcal{P}(t)} \text{Rank}(M_{g_{t_1}} \cdot M_{g_{t_2}} \cdot M_{g_{t_3}}).$$

By Lemma 5.1, there exist $\hat{t}_1, \hat{t}_2, \hat{t}_3 \in B_A$, such that $\{\hat{t}_1, \hat{t}_2, \hat{t}_3\} \in \mathcal{P}(t)$. Thus, by Proposition 3.1 and by the inductive hypothesis for $t_1, t_2, t_3$,
$$\text{Rank}(M_{g_t}) \geq \text{Rank}(M_{g_{\hat{t}_1}} \cdot M_{g_{\hat{t}_2}} \cdot M_{g_{\hat{t}_3}}) = \text{Rank}(M_{g_{\hat{t}_1}}) \cdot \text{Rank}(M_{g_{\hat{t}_2}}) \cdot \text{Rank}(M_{g_{\hat{t}_3}}) \geq 2^{l(\hat{t}_1)/2} \cdot 2^{l(\hat{t}_2)/2} \cdot 2^{l(\hat{t}_3)/2} = 2^{l(t)/2}.$$ 

Since this is true for every $t \in B_A$, we can apply it for $t = ([n], \emptyset) \in B_A$ and get
$$\text{Rank}(M_g) \geq 2^m.$$ 

Since $M_g$ is a matrix of size $2^m \times 2^m$, we actually have an equality in the last formula. \(\square\)

Corollary 5.7 Over any field $F$, any multilinear formula for the polynomial $f$ (as defined above) is of size $n^{\Omega(\log n)}$.

Proof: 
Follows immediately from Corollary 3.8 and Claim 5.6. \(\square\)

Acknowledgment

I would like to thank Amir Shpilka for very helpful conversations.

\(^6\)More precisely, $A'(x'_{t_1, t_2, t_3})$ is transcendental over the field $F$ extended by every other element in the domain of $A'$. That field obviously contains any coefficient that appears in any of the matrices in the sum.
References


[R] R. Raz. Multi-Linear Formulas for Permanent and Determinant are of Super-Polynomial Size. STOC 2004