# Algebraic Circuit Complexity

## Methods to Prove Superpolynomial Lower Bounds

Sankalp Mittal

UGP-1

April 13, 2024

# Outline

# Introduction

This presentation will deal with the research papers on lower bounds

- Paper by Ran Raz, which deals with *Multi-Linear Formulas for Permanent and Determinant* [Raz09]
- Paper by LST, which deals with *Low-Depth Algebraic Circuits* [LST21]

This explores new methods for transforming circuits and proving super-polynomial lower bounds for ABP's and Low Depth Circuits

# Multilinear Formulas

## Multilinear Polynomial

A polynomial is *multilinear* if, in each of its monomials, the power of every input variable is at most **one**.

Next, define a multilinear formula using the above definition

## Multilinear Formula

An arithmetic formula is *multilinear* if the polynomial computed by each gate of the formula is *multilinear*.

# Multilinear Formulas

## Syntactic Multilinear Formula

If in a formula, at each *multiplication node*($v$), the subformula($\Phi_v$) is such that the set of variables at the subnodes are disjoint, the formula is called *syntactic multilinear*.

# Syntactic Multilinear

**Convert multilinear to syntactic multilinear**

For any multilinear formula, there exists a syntactic multilinear formula of the same size that computes the same polynomial.

# The Result

**Superpolynomial Lower Bound**

The proof will be that any multilinear arithmetic formula for the permanent or the determinant of an $n \times n$ matrix is of size $n^{\Omega(\log n)}$ [Raz09]

# Partial Derivative Matrix

## Partial Derivative Matrix

Let $f$ be a multilinear polynomial over the set of variables $\{y_1, \cdots, y_m\} \cup \{z_1, \cdots, z_m\}$. For a multilinear monomial $p$ in the set of variables $\{y_1, \cdots, y_m\}$ and a multilinear monomial $q$ in the set of variables $\{z_1, \cdots, z_m\}$, 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 $F$.

# Partial Derivative Matrix

## Some Definitions

For each node $v$ denote by $Y_v$ the set of variables in $\{y_1, \cdots, y_m\}$ that appear in $\Phi_v$ and similarly for $Z_v$.

$$b(v) = \frac{|Y_v| + |Z_v|}{2}$$

$$a(v) = min(|Y_v| + |Z_v|)$$

$$d(v) = b(v) - a(v)$$

# Rank of Matrix

We are interested in the rank of the matrix $M_v$. The following are a set of results regarding the same

## Some Results

- $\textbf{Rank}(M_v) \leq 2^{a(v)}$
- $\textbf{Rank}(M_v) \leq \textbf{Rank}(M_{v_1}) + \textbf{Rank}(M_{v_2})$
- $\textbf{Rank}(M_v) = \textbf{Rank}(M_{v_1}) \cdot \textbf{Rank}(M_{v_2})$

# Unbalanced Nodes

## k-unbalanced node

We say that a node $v$ is *k-unbalanced* if $d(v) \geq k$

## k-unbalanced path

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

## Central Path

We say that $\gamma$ is *central* if for 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 $b(u) \leq 2b(u_1)$.

# k-weak paths

## k-weak

We say that a node $v$ of the formula is *k-weak* if every central path that reaches $v$ is k-unbalanced.

The following lemma shows that if a node $v$ is *k-weak* then the rank of the matrix $M_v$ can be bounded.

## Lemma

Let $\Phi$ be a syntactic multilinear arithmetic formula over the set of variables $\{y_1, \cdots, y_m\} \cup \{z_1, \cdots, z_m\}$ and let $v$ be a node in $\Phi$. If $v$ is *k-weak* then,

$$\textbf{Rank}(M_v) \leq |\Phi_v| \cdot 2^{b(v)-k/2}$$

# Random Partition

We think of X as a matrix of variables with $n$ rows and $n$ columns. Let $m = \lceil n^{1/3} \rceil$

## Assignment

For each variable in $X$, we assign either a value in $\{0, 1\}$ or a variable in $\{y_1, \cdots, y_m\} \cup \{z_1, \cdots, z_m\}$. The assignment will have the property that for each variable in $\{y_1, \cdots, y_m\} \cup \{z_1, \cdots, z_m\}$ there is exactly one variable in $X$ that is assigned that value

# Random Partition

First, choose, uniformly at random, for every $i \in [m]$ two values $q_i, r_i \in [n]$, such that all the $2m$ chosen values are different. (We think of $q_i, r_i$ as the indices of rows). Then choose, uniformly at random, for every $i \in [m]$ two additional values $s_i, t_i \in [n]$, such that all the $2m$ chosen values are different. (We think of $s_i, t_i$ as the indices of columns). For every $i$, we consider the four variables $x_{q_i,s_i}, x_{q_i,t_i}, x_{r_i,s_i}, x_{r_i,t_i}$ .
With probability half, we assign

$$A(x_{q_i,s_i}) = y_i, A(x_{q_i,t_i}) = z_i, A(x_{r_i,s_i}) = 1, A(x_{r_i,t_i}) = 1$$

and with probability half, we assign

$$A(x_{q_i,s_i}) = y_i, A(x_{q_i,t_i}) = 1, A(x_{r_i,s_i}) = z_i, A(x_{r_i,t_i}) = 1$$

All other variables in $X$ are assigned values in $\{0, 1\}$

# Probability of k-weak

**Lemma**

Let $\Phi$ be a syntactic multilinear arithmetic formula over the set of variables $X = \{x_{i,j}\}_{i,j\in[n]}$, 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 assignment to the variables in $X$, as above. Then, with probability of $1 - o(1)$ the formula $A$ is *k-weak*, for $k = n^{1/32}$.

# Probability of k-weak

## Proof.

Intuitively, since the random assignment $A$ has a lot of randomness, every node $v$ with large enough $X_v$ will be k-unbalanced with high probability. It can be proved that 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 $n^{\epsilon \log n}$. 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 $n^{-\Omega(\log n)}$

# Probability of k-weak

**Proof.**

This, however, is not trivial since there are dependencies between the different nodes. We will identify $\Omega(\log n)$ nodes, $v_1, \cdots, v_l$, 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 $v_1, \cdots, v_{i-1}$ are not k-unbalanced. □

# Chernoff Bound

## Chernoff Bound

Let $\chi_i, \cdots, \chi_l$ be mutually independent random variables, such that $\mathbf{Pr}[\chi_i = 1] = p$ and $\mathbf{Pr}[\chi_i = 0] = 1 - p$. Then for any $c > 0$,

$$\mathbf{Pr}\left[|\sum_{i=1}^{l} \chi_i - pl| > cpl\right] < 2e^{-2(cp)^2 l}$$

This is used to prove bounds on the new matrix after mapping the variables, and then the mapped formula $\Psi$ is also proved to be *k-weak*.

# Final Proof

Permanent and determinant follow similar proofs, we use proof by contradiction, assume that $|\Phi| \leq n^{\epsilon \log n}$.

## Proof.

Let $A$ be a random assignment to the variables in X, as defined earlier. Then, $\Phi_A$ is a syntactic multilinear arithmetic formula over the set of variables $\{y_1, \cdots, y_m\} \cup \{z_1, \cdots, z_m\}$. With probability of $1 - o(1)$ the formula $\Phi_A$ is k-weak, for $k = n^{1/32}$. Hence with probability of $1 - o(1)$

$$\textbf{Rank}(M_{\Phi_A}) \leq n^{\epsilon \log n} \cdot 2^{m-k/2} < 2^m$$

# Final Proof

## Proof.

At the other hand, since the output of $\Phi$ is the permanent of $X$, the output of $\Phi_A$ must be the permanent of the matrix $\{A(x_{i,j})\}_{i,j\in[n]}$. By the definition of $A$ the *permanent* of that matrix is always

$$f(y_1, \cdots, y_m, z_1, \cdots, z_m) \equiv \prod_{i=1}^{m}(y_i + z_i)$$

And the *determinant* of that matrix is,

$$g(y_1, \cdots, y_m, z_1, \cdots, z_m) \equiv \prod_{i=1}^{m}(y_i - z_i)$$

# Final Proof

**Proof.**

Note that $M_f$ is a permutation matrix, because for every multilinear monomial $p$ in the set of variables $\{y_1, \cdots, y_m\}$ there is exactly one multilinear monomial $q$ in the set of variables $\{z_1, \cdots, z_m\}$ such that $M_f(p, q) = 1$ (and otherwise $M_f(p, q) = 0$), and vice-versa. Hence,

$$\textbf{Rank}(M_{\Phi_A}) = \textbf{Rank}(M_f) = 2^m$$

□

This is a direct contradiction of one of our results, and hence our initial assumption was wrong.

# Background on LST Paper

There were no known lower bounds known for general Algebraic Circuits of *product-depth 1* over fields of large size, and no superpolynomial lower bounds against general algebraic circuits of *product-depth more than 1*. This paper proves the first superpolynomial lower bounds for algebraic circuits of *constant product-depth*.

# Hardness Escalation

## Hardness Escalation

This is a technique used to prove lower bounds on more general circuits by proving lower bounds on restricted forms on arithmetic circuits.
This is very widely used in proving results in many areas of computational complexity.

# PIT

## Polynomial Identity Testing (PIT)

PIT is the problem of efficiently determining whether two multivariate polynomials are identical. More formally, a PIT algorithm is given a circuit that computes $p$ in a field, and decides whether $p$ is the *zero polynomial*. Finding *deterministic algorithms* for PIT, is one of the most important open problems in algebraic computing complexity.

# Results

**Theorem 1 (Main Result)**

Let $N, d, \Delta$ be growing parameters with $d = o(\log N)$. Assume $\mathbb{F}$ has characteristic 0 or greater than $d$. There is an explicit polynomial $P_{N,d}(x_1, \cdots, x_N)$ that has no algebraic circuits of product-depth $\Delta$ and size at most $N^{d^{exp(-O(\Delta))}}$. [LST21]

# Results

**Theorem 2 (Lower bound for set-multilinear circuits)**

Assume $d \leq (\log n)/100$. For any product-depth $\Delta \geq 1$, any *set-multilinear* circuit $C$ computing $IMM_{n,d}$ of product-depth at most $\Delta$ must have size at least $n^{d^{exp(-O(\Delta))}}$. In the particular case that $\Delta = 2$, the size of C must be at least $n^{\Omega(\sqrt{d})}$. [LST21]

# Results

Any homogeneous circuit computing a set-multilinear polynomial can be converted a set-multilinear circuit of the same depth and size $s \cdot d^{O(d)}$, combining with *Theorem 2* we get,

**Corollary 3 (Lower bound for homogeneous circuits)**

Assume $d \leq (\log n)/100$. For any product-depth $\Delta \geq 1$, any *homogeneous* circuit $C$ computing $IMM_{n,d}$ of product-depth at most $\Delta$ must have size at least $n^{d^{exp(-O(\Delta))}}$. In the particular case that $\Delta = 2$, the size of C must be at least $n^{\Omega(\sqrt{d})}$. [LST21]

# Results

## Convert general circuits to homogeneous

Any (possibly non-homogeneous) algebraic circuit of product depth $\Delta$ and size $s$ computing a homogeneous polynomial $P$ of degree $d$ can be converted to a homogeneous circuit for $P$ of product-depth $2\Delta$ and size $\text{poly}(s) \cdot d^{O(d)}$. This conversion assumes that the underlying field has characteristic 0 or greater than $d$.

# Results

**Corollary 4 (Lower bound for general circuits)**

Assume $d \leq (\log n)/100$ and $\text{char}(\mathbb{F}) = 0$ or *greater than* $d$. For any product-depth $\Delta \geq 1$, any *homogeneous* circuit $C$ computing $IMM_{n,d}$ of product-depth at most $\Delta$ must have size at least $n^{d^{exp(-O(\Delta))}}$. In the particular case that $\Delta = 2$, the size of C must be at least $n^{\Omega(\sqrt{d})}$. [LST21]

For the case of $\Delta = 1$, the bound is actually tight

# Results

*Theorem 2* also allows us to prove the following theorem for constant-depth Algebraic circuits. Which means that circuits of depth $\Gamma$ are *superpolynomially* more powerful than circuits of depth $\Gamma - 1$

**Theorem 5 (Depth Hierarchy theorem)**

Assume that the underlying field $\mathbb{F}$ has characteristic 0. For any constant $\Gamma \geq 2$ and $s$ a growing parameter, there exists a set-multilinear polynomial $Q_\Gamma$ of depth $\Gamma$ and size $s$ such that any depth $(\Gamma - 1)$ circuit computing $Q_\Gamma$ must have size $s^{\omega(1)}$

The main idea behind proving such a result is to design an explicit set-multilinear polynomial for which the lower bound implied by the techniques of Theorem 2 is *tight*.

# Results

## PIT and Lower Bounds

People have showed that superpolynomial lower bounds for general algebraic circuits imply deterministic sub-exponential time algorithms for general PIT.
It has also been shown that the hardness of constant depth circuits implies deterministic PIT for constant depth circuits

# Results

## Corollary 6

Let $\mu > 0$ be a real number and $\mathbb{F}$ a field of characteristic 0. Let $C$ be an algebraic circuit of size $s \leq \text{poly}(n)$, depth $\Delta = o(\log \log \log n)$ computing a polynomial on $n$ variables, then there is a deterministic algorithm that can check whether the polynomial computed by $C$ is identically zero or not in time $(s^{\Delta+1} \cdot n)^{O(n^{\mu})}$.

# References



Nutan Limaye, Srikanth Srinivasan, and Sebastien Tavenas.
Superpolynomial lower bounds against low-depth algebraic circuits.
2021.


Ran Raz.
Multi-linear formulas for permanent and determinant are of super-polynomial size.
2009.