Arithmetic Circuits Lower Bounds via (Polynomial) Partial Derivatives Matrices

Mrinal Kumar (Rutgers Univ.)
Gaurav Maheshwari (Goldman Sachs)
Jayalal Sarma (IIT Madras)

June 28, 2013

IMSc, Chennai

Arithmetic Circuits

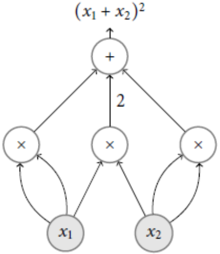
Basic Objects : $\{f_n : f(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots x_n], n \in \mathbb{N}\}$

Arithmetic Circuits

Basic Objects : $\{f_n : f(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots x_n], n \in \mathbb{N}\}$
Adversaries : Circuits with $+, \times$ as gates computes a polynomial in $\mathbb{F}[X]$

Parameters:

- Size: # of gates in the circuit.
- Depth: Longest path from any leaf to root.

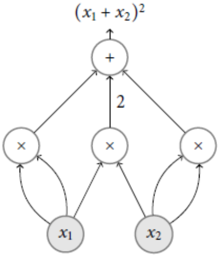


Arithmetic Circuits

Basic Objects : $\{f_n : f(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, x_2, \ldots x_n], n \in \mathbb{N}\}$
Adversaries : Circuits with $+, \times$ as gates computes a polynomial in $\mathbb{F}[X]$

Parameters:

- Size: # of gates in the circuit.
- Depth: Longest path from any leaf to root.



Two natural Questions :

- Are there polynomials that are "hard" in terms of size?
- Are there polynomials that are "hard" in terms of depth?

A Central Question and Two Fundamental Polynomials

VP : Set of polynomials of poly degree computed by polysized arithmetic circuits.

$$Det(X) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i \in [n]} x_{ij}$$

Determinant polynomial of the generic matrix is complete for VP.

VNP : Set of polynomials of expressible as an exponential sum of a polynomial in VP.

$$Perm(X) = \sum_{\sigma \in S_n} \prod_{i \in [n]} x_{ij}$$

Permanent polynomial of the generic matrix is complete for VNP.

VP vs VNP Problem. $\equiv$ **Permanent vs Determinant Problem.**

Are there polynomials in VNP that requires super polynomial size for any arithmetic circuit computing them?

What is known? - Structurally Limited Circuits

Restriction	Bound	Reference
Depth-2 circuits	$2^{\Omega(n \log n)}$	Trivial
Depth-3 circuits (over finite fields)	$2^{\Omega(n)}$	Grigoriev-Karpinski(1998)
Depth-3 circuits	$\Omega(n^2)$	Shpilka-Wigderson (2001)
General circuits	$\Omega(n \log n)$	Baur-Strassen(1983)
General formulas	$\Omega(n^3)$	Kalorkoti(1985)

What is known? - Structurally Limited Circuits

Restriction	**Bound**	**Reference**
Depth-2 circuits	$2^{\Omega(n \log n)}$	Trivial
Depth-3 circuits (over finite fields)	$2^{\Omega(n)}$	GRIGORIEV-KARPINSKI(1998)
Depth-3 circuits	$\Omega(n^2)$	SHPILKA-WIGDERSON (2001)
General circuits	$\Omega(n \log n)$	BAUR-STRASSEN(1983)
General formulas	$\Omega(n^3)$	KALORKOTI(1985)

- We are stuck for the case of constant depth circuits (even for depth three !).
- What can we assume in general about the depth of the circuit?

Depth reductions till 2010

Valiant-Skyum-Berkowitz-Rackoff(1983) If f of polynomial degree can be computed with a circuit of polynomial size, then f can be computed in polynomial size and depth $O(\log^2 n)$. Thus,

$$\mathrm{VP} = \mathrm{VNC}^2$$

Depth reductions till 2010

VALIANT-SKYUM-BERKOWITZ-RACKOFF(1983) If f of polynomial degree can be computed with a circuit of polynomial size, then f can be computed in polynomial size and depth $O(\log^2 n)$. Thus,

$$\mathrm{VP} = \mathrm{VNC}^2$$

AGRAWAL-VINAY(2008), KOIRAN(2010): If f can be computed by polynomial size circuits, then f can be computed in size $2^{O(\sqrt{n}\log^2 n)}$ by a depth 4 circuit.

Depth reductions till 2010

Valiant-Skyum-Berkowitz-Rackoff(1983) If f of polynomial degree can be computed with a circuit of polynomial size, then f can be computed in polynomial size and depth $O(\log^2 n)$. Thus,

$$\mathrm{VP} = \mathrm{VNC}^2$$

Agrawal-Vinay(2008), Koiran(2010): If f can be computed by polynomial size circuits, then f can be computed in size $2^{O(\sqrt{n}\log^2 n)}$ by a depth 4 circuit.

Conclusion : For separating VNP from VP, it suffices to show that there is a polynomial with n variables in VNP which requires size $2^{\omega(\sqrt{n}\log^2 n)}$ for any depth 4 circuit computing it.

Observations on the Candidate Hard Polynomial - I

They are *homogeneous* : Each monomial is of the same degree.

Homogeneous circuit : It computes a homogeneous polynomial at each gate.

Observations on the Candidate Hard Polynomial - I

They are *homogeneous* : Each monomial is of the same degree.

Homogeneous circuit : It computes a homogeneous polynomial at each gate.

QUESTION 1 : Can we prove lower bounds against homogeneous circuits?

QUESTION 2 : Does non-homogeneity help in super polynomial size reduction?

Observations on the Candidate Hard Polynomial - I

They are *homogeneous* : Each monomial is of the same degree.

Homogeneous circuit : It computes a homogeneous polynomial at each gate.
QUESTION 1 : Can we prove lower bounds against homogeneous circuits?
QUESTION 2 : Does non-homogeneity help in super polynomial size reduction? No, in general, but not known for constant depth.

Observations on the Candidate Hard Polynomial - I

They are *homogeneous* : Each monomial is of the same degree.

Homogeneous circuit : It computes a homogeneous polynomial at each gate.
QUESTION 1 : Can we prove lower bounds against homogeneous circuits?
QUESTION 2 : Does non-homogeneity help in super polynomial size reduction? No, in general, but not known for constant depth.

AGRAWAL-VINAY(2008), KOIRAN(2010): If f can be computed by polynomial size circuits, then f can be computed in size $2^{O(\sqrt{n}\log n)}$ by a depth 4 **homogeneous** circuit.

Conclusion : Suffices to prove lower bounds of the form $2^{\omega(\sqrt{n}\log n)}$ against depth 4 homegenous circuits.

In the Homogeneous World ...

Iterated Matrix Multiplication (IMM) Given d, $n \times n$ generic matrices, compute the product matrix. Polynomial is the one computed at $(1,1)$-entry of the resulting matrix.

$$\begin{pmatrix} x_{11}^{(1)} & \cdots & x_{1n}^{(1)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(1)} & \cdots & x_{nn}^{(1)} \end{pmatrix} \begin{pmatrix} x_{11}^{(2)} & \cdots & x_{1n}^{(2)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(2)} & \cdots & x_{nn}^{(2)} \end{pmatrix} \cdots \begin{pmatrix} x_{11}^{(d)} & \cdots & x_{1n}^{(d)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(d)} & \cdots & x_{nn}^{(d)} \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

Nisan-Wigderson (1995): Any **depth three homogeneous** circuit computing the IMM polynomial must have size $\Omega\left(\frac{n^{d-1}}{d!}\right)$.

In the Homogeneous World . . .

ITERATED MATRIX MULTIPLICATION (IMM) Given d, $n \times n$ generic matrices, compute the product matrix. Polynomial is the one computed at $(1,1)$-entry of the resulting matrix.

$$\begin{pmatrix} x_{11}^{(1)} & \cdots & x_{1n}^{(1)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(1)} & \cdots & x_{nn}^{(1)} \end{pmatrix} \begin{pmatrix} x_{11}^{(2)} & \cdots & x_{1n}^{(2)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(2)} & \cdots & x_{nn}^{(2)} \end{pmatrix} \cdots \begin{pmatrix} x_{11}^{(d)} & \cdots & x_{1n}^{(d)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(d)} & \cdots & x_{nn}^{(d)} \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

NISAN-WIGDERSON (1995): Any **depth three homogeneous** circuit computing the IMM polynomial must have size $\Omega\left(\frac{n^{d-1}}{d!}\right)$.

Technique : Partial Derivatives Method.

There is a depth two homogeneous circuit computing the IMM polynomial of size $O(n^{d-1})$.

Meanwhile in Bangalore ... Strong Lower Bounds against Homogeneous depth-4 circuits

Gupta-Kamath-Kayal-Saptarishi (2012): Any homogeneous depth four arithmetic circuit with bottom fanin bounded by $\sqrt{n}$ computing permanent must of size $2^{\Omega(\sqrt{n}\log n)}$.

Meanwhile in Bangalore ... Strong Lower Bounds against Homogeneous depth-4 circuits

GUPTA-KAMATH-KAYAL-SAPTARISHI (2012): Any homogeneous depth four arithmetic circuit with bottom fanin bounded by $\sqrt{n}$ computing permanent must of size $2^{\Omega(\sqrt{n}\log n)}$.

Technique : Shifted Partial Derivatives Method.

Good news : If this size lower bound is quantitatively improved to $2^{\omega(\sqrt{n}\log^2 n)}$, then we separate VP from VNP .

Meanwhile in Bangalore ... Strong Lower Bounds against Homogeneous depth-4 circuits

GUPTA-KAMATH-KAYAL-SAPTARISHI (2012): Any homogeneous depth four arithmetic circuit with bottom fanin bounded by $\sqrt{n}$ computing permanent must of size $2^{\Omega(\sqrt{n}\log n)}$.

Technique : Shifted Partial Derivatives Method.

Good news : If this size lower bound is quantitatively improved to $2^{\omega(\sqrt{n}\log^2 n)}$, then we separate VP from VNP .

Bad news : Whatever is known, works even for the determinant.

Observations on the Candidate Hard Polynomial - II

They are *multilinear* : Each monomial has variables occuring in individual degree at most 1.
Circuit is *multilinear* if each gate computes a multilinear polynomial.

Observations on the Candidate Hard Polynomial - II

They are *multilinear* : Each monomial has variables occuring in individual degree at most 1.
Circuit is *multilinear* if each gate computes a multilinear polynomial.

QUESTION 1 : Can we prove lower bounds against multilinear circuits?
QUESTION 2 : Does non-multilinearity help in super polynomial size reduction?

Observations on the Candidate Hard Polynomial - II

They are *multilinear* : Each monomial has variables occuring in individual degree at most 1.
Circuit is *multilinear* if each gate computes a multilinear polynomial.

QUESTION 1 : Can we prove lower bounds against multilinear circuits?
QUESTION 2 : Does non-multilinearity help in super polynomial size reduction?

RAZ 2005 : Multilinear formulas computing determinant and permanent of $n \times n$ matrices require $n^{\Omega(\log n)}$ size.

Can we extend the above lower bound technique to the case of non-multilinear circuits?

Product Dimension

Consider a depth three circuit ($\Sigma\Pi\Sigma$). Let the top-fanin be k. $\forall 1 \leq i \leq k$, d_i denote the fanin of the i^{th} product gate Q_i.

Product Dimension(Q_i) $= dim\{span\{L_{ij} : j \in [d_i]\}\}$.
Product Dimension(C) $= \max_i\{$Product Dimension(Q_i)$\}$.

Product Dimension

Consider a depth three circuit $(\Sigma\Pi\Sigma)$. Let the top-fanin be k. $\forall 1 \leq i \leq k$, d_i denote the fanin of the i^{th} product gate Q_i.

Product Dimension$(Q_i) = dim\{span\{L_{ij} : j \in [d_i]\}\}$.
Product Dimension(C) $= \max_i\{$Product Dimension$(Q_i)\}$.

- Product Dimension 1 : Diagonal Circuits - we know lower bounds against them (Saxena(2008)).
- Product Dimension n : General depth three circuits. Our main adversary.
- Product Dimension vs Rank of a circuit. The latter is a strong restriction.

Our Main Results

We generalize Raz's method to non-multilinear setting. And apply it to prove the following results:

Theorem

Any homogeneous depth three circuit computing an entry in the product of d, $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

Theorem

There is an explicit polynomial $p(x_1, \ldots, x_n)$ of degree at most $\frac{n}{2}$ in VNP *such that any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ computing it has size $2^{\Omega(n)}$.*

Extending to product dimension n settles the depth three lowerbounds question over infinite fields.

Surprise, surprise ... the chasm is at depth three.

Gupta-Kamath-Kayal-Saptarishi (2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d\log(d)}\log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n}\log^{\frac{3}{2}} n)}$.

Surprise, surprise ... the chasm is at depth three.

Gupta-Kamath-Kayal-Saptarishi (2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d \log(d)} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log^{\frac{3}{2}} n)}$.

Tavenas(2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log n)}$.

Surprise, surprise ... the chasm is at depth three.

Gupta-Kamath-Kayal-Saptarishi (2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d \log(d)} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log^{\frac{3}{2}} n)}$.

Tavenas(2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log n)}$.

Revised Goal : Show lower bounds of the kind $2^{\omega(\sqrt{n} \log n)}$ against depth three circuits.   .

Surprise, surprise ... the chasm is at depth three.

Gupta-Kamath-Kayal-Saptarishi (2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d \log(d)} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log^{\frac{3}{2}} n)}$.

Tavenas(2013): If an n-variate polynomial of degree d ($d = n^{O(1)}$) is computable by an arithmetic circuit of polynomial size then it can also be computed by a depth **three** circuit of size $2^{O(\sqrt{d} \log(n))}$. If $d \leq n$, this is $2^{O(\sqrt{n} \log n)}$.

Revised Goal : Show lower bounds of the kind $2^{\omega(\sqrt{n} \log n)}$ against depth three circuits of product dimension n.

We will show this for product dimension $\frac{n}{10}$.

Our Results contd.

(s, d)-product-sparse formulas. Each product gate is having one of the inputs as 2^s-sparse, number of non-syntactic-multilinear violations in any path is at most d.
Syntactic multilinear formulas, and skew formulas are special cases.

Theorem (Generalizing Multilinear Formulas)

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

Theorem (Generalizing Ordered Branching Programs)

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

Partial Derivative Matrix : from Multilinear World

$X = \{x_1, x_2, \ldots x_n\}$ be the set of variables.
$X = Y \cup Z$.

M_f: for any $f \in \mathbb{F}[Y, Z]$; rows and cols indexed by subsets of Y and Z resp.

$M_f(p, q) = c$, where c is the coefficient of the multilinear monomial pq in f.

Partial Derivative Matrix : from Multilinear World

$X = \{x_1, x_2, \ldots x_n\}$ be the set of variables.
$X = Y \cup Z$.

M_f: for any $f \in \mathbb{F}[Y, Z]$; rows and cols indexed by subsets of Y and Z resp.

$M_f(p, q) = c$, where c is the coefficient of the multilinear monomial pq in f.
RAZ (2005): RANK(M_f) can be used as a complexity measure for **multilinear** circuits polynomials.

Partial Derivative Matrix : from Multilinear World

$X = \{x_1, x_2, \ldots x_n\}$ be the set of variables.
$X = Y \cup Z$.

M_f: for any $f \in \mathbb{F}[Y, Z]$; rows and cols indexed by subsets of Y and Z resp.

$M_f(p, q) = c$, where c is the coefficient of the multilinear monomial pq in f.
Raz (2005): Rank(M_f) can be used as a complexity measure for **multilinear** circuits polynomials.

- For any multilinear formula of polynomials size, there is a partition such that the polynomial at the output has "low" rank for M_f.
- For any partition, the M_f of permanent and determinant has "large" rank.

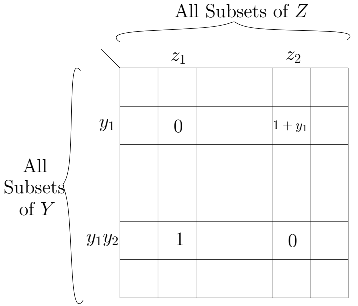
Our Main Tool: Polynomial Coefficient Matrix

$X = Y \cup Z, |Y| = |Z|$
$\mathrm{Var}(h)$: Variables appearing in h.

M_f: for any $f \in \mathbb{F}[Y, Z]$

$M_f(p, q) = h$, where

- $f = h.pq + r$
- $h, r \in \mathbb{F}[Y \cup Z]$
- $\mathrm{Var}(h) \subseteq \mathrm{Var}(pq)$.
- pq does not divide any monomial in r with $Var(r) \subseteq Var(pq)$.



$$f = y_1y_2z_1 + y_1^2z_2 + y_1z_2$$
$$= y_1y_2z_1 + y_1z_2(y_1 + 1)$$

Our Main Tool: Polynomial Coefficient Matrix

$X = Y \cup Z, |Y| = |Z|$

$\mathrm{Var}(h)$: Variables appearing in h.

M_f: for any $f \in \mathbb{F}[Y, Z]$

$M_f(p, q) = h$, where

- $f = h.pq + r$
- $h, r \in \mathbb{F}[Y \cup Z]$
- $\mathrm{Var}(h) \subseteq \mathrm{Var}(pq)$.
- pq does not divide any monomial in r with $Var(r) \subseteq Var(pq)$.

All Subsets of Z

All Subsets of Y

		z_1		z_2	
y_1		0		$1 + y_1$	
$y_1 y_2$		1		0	

$$f = y_1 y_2 z_1 + y_1^2 z_2 + y_1 z_2 = y_1 y_2 z_1 + y_1 z_2 (y_1 + 1)$$

$$f = \sum_{p,q} M_f(p, q) pq$$

. If f is multilinear then M_f is same as PDM.

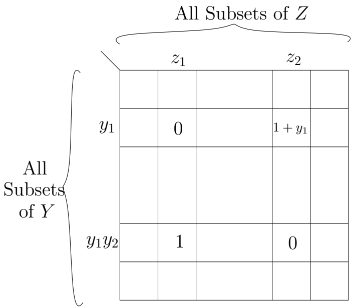
Our Main Tool: Polynomial Coefficient Matrix

$X = Y \cup Z, |Y| = |Z|$
$\mathsf{Var}(h)$: Variables appearing in h.

M_f: for any $f \in \mathbb{F}[Y, Z]$

$M_f(p, q) = h$, where

- $f = h.pq + r$
- $h, r \in \mathbb{F}[Y \cup Z]$
- $\mathsf{Var}(h) \subseteq \mathsf{Var}(pq)$.
- pq does not divide any monomial in r with $Var(r) \subseteq Var(pq)$.



$$f = y_1y_2z_1 + y_1^2z_2 + y_1z_2$$
$$= y_1y_2z_1 + y_1z_2(y_1 + 1)$$

Complexity Measure:

$$\textsc{Max-Rank}(f) = \max_{S: Y \cup Z \to \mathbb{F}} \{\textsc{Rank}(M_f|_S)\}$$

Properties of $\textsc{Max-Rank}(f)$

- $\textsc{Max-Rank}(f_v) \leq 2^{\min\{|Y_v|,|Z_v|\}}$.

Properties of $\textsc{Max-Rank}(f)$

- $\textsc{Max-Rank}(f_v) \leq 2^{\min\{|Y_v|,|Z_v|\}}$.
- With Addition: $h = f + g \implies M_h = M_f + M_g$
  $\textsc{Max-Rank}(h) \leq \textsc{Max-Rank}(f) + \textsc{Max-Rank}(g)$.

Properties of $\textsc{Max-Rank}(f)$

- $\textsc{Max-Rank}(f_v) \leq 2^{\min\{|Y_v|,|Z_v|\}}$.
- With Addition: $h = f + g \implies M_h = M_f + M_g$
  $\textsc{Max-Rank}(h) \leq \textsc{Max-Rank}(f) + \textsc{Max-Rank}(g)$.
- With Multiplication: $h = f \times g$
  - $X_f \cap X_g = \phi \implies M_h = M_f \otimes M_g$.
    $\textsc{Max-Rank}(h) \leq \textsc{Max-Rank}(f) \times \textsc{Max-Rank}(g)$
  - $g \in \mathbb{F}[Y]$, then $\textsc{Max-Rank}(h) \leq \textsc{Max-Rank}(f)$
  - $\text{Support}(g) \leq r \implies \textsc{Max-Rank}(h) \leq r \cdot \textsc{Max-Rank}(f)$
  - If g is an affine form, $\textsc{Max-Rank}(h) \leq 2 \cdot \textsc{Max-Rank}(f)$.

Lemma

Lemma

If $f \in \mathbb{F}[Y, Z]$ *and* $g \in \mathbb{F}[Y]$, *then* $\textsc{Max-Rank}(M_{fg}) \leq \textsc{Max-Rank}(M_f)$.

Lemma

Lemma

If $f \in \mathbb{F}[Y, Z]$ and $g \in \mathbb{F}[Y]$, then
$\textsc{Max-Rank}(M_{fg}) \leq \textsc{Max-Rank}(M_f)$.

Proof.

- Consider a simple case. $g = y$.
- Conider the row of $M_{f \cdot y}$ indexed by a monomial p (denote it by $M_f(p)$) will either be zero (if $y \notin Var(p)$) or will be expressible as $M_{fg}(p) = y \cdot M_f(p) + M_f(p/y)$.
- Under any substitution, rowspace of $M_{fy}|_S$ is contained in rowspace of $M_f|_S$.
- Hence $\textsc{Max-Rank}(M_{fy}) \leq \textsc{Max-Rank}(M_f)$.

□

Lemma

Lemma

If $f \in \mathbb{F}[Y, Z]$ and $g \in \mathbb{F}[Y]$, then
$\text{MAX-RANK}(M_{fg}) \leq \text{MAX-RANK}(M_f)$.

Proof.

- Let g be a monomial. Let $T \subseteq Y$. Let y^T denote the corresponding monomial.
- The row of $M_{y^T.f}$ indexed by a monomial p will either be zero (if $T \not\subseteq Var(p)$) or will be expressible as

$$M_{y^T.f}(p) = \sum_{T' \subseteq T} y^{T \setminus T'} M_f(p/y^{T'})$$

- Under any substitution, rowspace of $M_{fy^T}|_S$ is contained in rowspace of $M_f|_S$.
- Hence $\text{MAX-RANK}(M_{fg}) \leq \text{MAX-RANK}(M_f)$.

Lemma

Lemma

If $f \in \mathbb{F}[Y, Z]$ and $g \in \mathbb{F}[Y]$, then $\textsc{Max-Rank}(M_{fg}) \leq \textsc{Max-Rank}(M_f)$.

Proof.

- Consider $g = \sum_{i \in [r]} m_i$ where r is the number of monomials in g. $M_{fg} = \sum_{i \in [r]} M_{fm_i}$.
- Under any substitution, rowspace of $M_{fg}|_S$ is contained in rowspace of $M_f|_S$.
- Hence $\textsc{Max-Rank}(M_{fg}) \leq \textsc{Max-Rank}(M_f)$.

□

Lemma

Lemma

If $f \in \mathbb{F}[Y, Z]$ *and* $g \in \mathbb{F}[Y]$*, then* $\textsc{Max-Rank}(M_{fg}) \leq \textsc{Max-Rank}(M_f)$.

Corollary

Let $f, g \in \mathbb{F}[Y, Z]$*:*

- *If* g *is a linear form then* $\textsc{Max-Rank}(M_{fg}) \leq 2. \textsc{Max-Rank}(M_f)$.
- *If* $g = \sum_{i \in [r]} g_i h_i$ *where* $g_i \in \mathbb{F}[Y]$ *and* $h_i \in \mathbb{F}[Z]$*, then* $\textsc{Max-Rank}(M_{fg}) \leq r. \textsc{Max-Rank}(M_f)$.
- *If* g *has* r *monomials, then* $\textsc{Max-Rank}(M_{fg}) \leq r \cdot \textsc{Max-Rank}(M_f)$.

First Application : Homogeneous frontier

Theorem

Any homogeneous depth three circuit computing an entry in the product of d $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

First Application : Homogeneous frontier

Theorem

Any homogeneous depth three circuit computing an entry in the product of d $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

Proof.

1. Let C be the depth three circuit with formal degree d and top fan-in k. Fix and arbitrary partition,
2. C can be written as $\sum_i P_i$ where $P_i = \prod_{j=1}^{deg(P_i)} \ell_{ij}$ where ℓ_{ij} is a homogeneous linear form.

First Application : Homogeneous frontier

Theorem

Any homogeneous depth three circuit computing an entry in the product of d $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

Proof.

1. Let C be the depth three circuit with formal degree d and top fan-in k. Fix and arbitrary partition,
2. C can be written as $\sum_i P_i$ where $P_i = \prod_{j=1}^{deg(P_i)} \ell_{ij}$ where ℓ_{ij} is a homogeneous linear form. $\textsc{Max-Rank}(P_i) \leq 2^d$.

First Application : Homogeneous frontier

Theorem

Any homogeneous depth three circuit computing an entry in the product of d $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

Proof.

1. Let C be the depth three circuit with formal degree d and top fan-in k. Fix and arbitrary partition,
2. C can be written as $\sum_i P_i$ where $P_i = \prod_{j=1}^{deg(P_i)} \ell_{ij}$ where ℓ_{ij} is a homogeneous linear form. $\text{Max-Rank}(P_i) \leq 2^d$.

$$\text{Max-Rank}(C) \leq k.2^d$$

First Application : Homogeneous frontier

Theorem

Any homogeneous depth three circuit computing an entry in the product of d $n \times n$ matrices has size $\Omega(\frac{n^{d-1}}{2^d})$.

Proof.

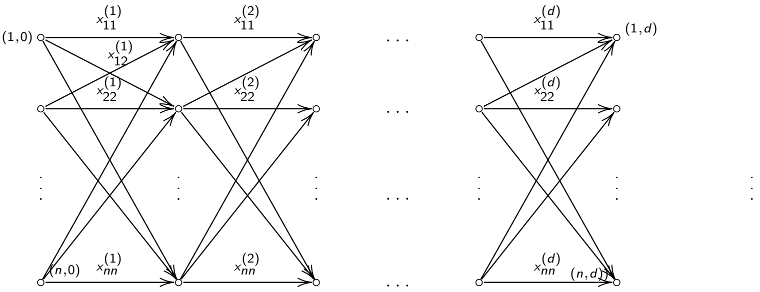
1. Let C be the depth three circuit with formal degree d and top fan-in k. Fix and arbitrary partition,
2. C can be written as $\sum_i P_i$ where $P_i = \prod_{j=1}^{deg(P_i)} \ell_{ij}$ where ℓ_{ij} is a homogeneous linear form. $\textsc{Max-Rank}(P_i) \leq 2^d$.

$$\textsc{Max-Rank}(C) \leq k.2^d$$

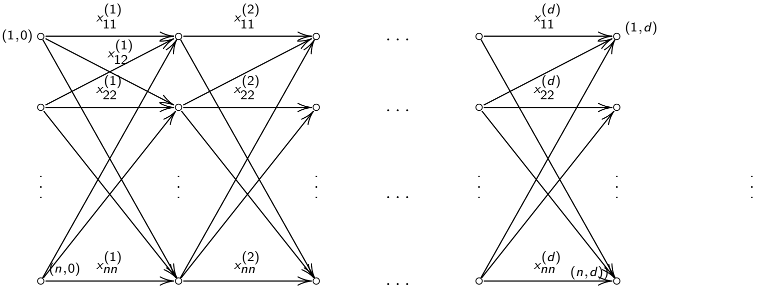
3. $\textsc{Max-Rank}(IMM(d,n)) = n^{d-1}$.

Max-Rank$(IMM(d,n)) = n^{d-1}$

$$\begin{pmatrix} x_{11}^{(1)} & \cdots & x_{1n}^{(1)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(1)} & \cdots & x_{nn}^{(1)} \end{pmatrix} \begin{pmatrix} x_{11}^{(2)} & \cdots & x_{1n}^{(2)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(2)} & \cdots & x_{nn}^{(2)} \end{pmatrix} \cdots \begin{pmatrix} x_{11}^{(d)} & \cdots & x_{1n}^{(d)} \\ \vdots & \vdots & \vdots \\ x_{n1}^{(d)} & \cdots & x_{nn}^{(d)} \end{pmatrix} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$



$\textsc{Max-Rank}(IMM(d,n)) = n^{d-1}$



Partition: $Y(Z)$ as variables in the odd(even) indexed matrices.
Observe : while constructing a path, if we fix the edges from the odd layers, the edges from the even layers are unique.

$\textsc{Max-Rank}(IMM(d, n)) = n^{d-1}$

- The matrix M_f will have only one non-zero entry in the row chosen, at the column(T) indexed by the corresponding even indexed variables.
- The same set of edges (the column T) from even indexed layers will not form a path with any other set of edges from the odd indexed layers.
- Thus the matrix M_f simply has the identity matrix of size n^{d-1} up to permutation.
- Hence rank of M_f is exactly n^{d-1}.

Application 2 : Depth Three circuits with Product Dimension $\frac{n}{10}$

Theorem

There is an explicit polynomial P in n variables and degree at most $\frac{n}{2}$ such that any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ computing it has size $2^{\Omega(n)}$.

Proof.

1. For a $\Sigma\Pi\Sigma$ circuit C with top-fanin k and product dimension r, computing a degree d polynomial, for any equipartition, $\textsc{Max-Rank}(C) \leq k\binom{d+r}{r}(d+1)$.

Application 2 : Depth Three circuits with Product Dimension $\frac{n}{10}$

Theorem

There is an explicit polynomial P in n variables and degree at most $\frac{n}{2}$ such that any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ computing it has size $2^{\Omega(n)}$.

Proof.

1. For a $\Sigma\Pi\Sigma$ circuit C with top-fanin k and product dimension r, computing a degree d polynomial, for any equipartition, $\textsc{Max-Rank}(C) \leq k\binom{d+r}{r}(d+1)$.
2. There is a polynomial P of degree $\frac{n}{2}$ such that there is a partition for which $\textsc{Max-Rank}(P) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.

Application 2 : Depth Three circuits with Product Dimension $\frac{n}{10}$

Theorem

There is an explicit polynomial P in n variables and degree at most $\frac{n}{2}$ such that any $\Sigma\Pi\Sigma$ circuit C of product dimension at most $\frac{n}{10}$ computing it has size $2^{\Omega(n)}$.

Proof.

1. For a $\Sigma\Pi\Sigma$ circuit C with top-fanin k and product dimension r, computing a degree d polynomial, for any equipartition, $\textsc{Max-Rank}(C) \leq k\binom{d+r}{r}(d+1)$.
2. There is a polynomial P of degree $\frac{n}{2}$ such that there is a partition for which $\textsc{Max-Rank}(P) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.
3. Hence, $k \geq 2^{\Omega(n)}$, if $r \leq \frac{n}{10}$.

Step 1 : Upper bound from the model.

Lemma

For a $\Sigma\Pi\Sigma$ circuit C with product dimension r, computing a degree d polynomial, for any equipartition, $\textsc{Max-Rank}(C) \leq (\textit{top fanin})\binom{d+r}{r}(d+1)$.

Proof.

- Consider a product gate $Q = \prod_{i=1}^{t} \ell_i$.
- Let ℓ_i's (for this Q) be spanned by the affine forms $m_1, \ldots m_r$.

Step 1 : Upper bound from the model.

Lemma

For a ΣΠΣ circuit C with product dimension r, computing a degree d polynomial, for any equipartition, $\textsc{Max-Rank}(C) \leq (\textit{top fanin})\binom{d+r}{r}(d+1)$.

Proof.

- Consider a product gate $Q = \prod_{i=1}^{t} \ell_i$.
- Let ℓ_i's (for this Q) be spanned by the affine forms $m_1, \ldots m_r$.

$$Q = \prod_{i=1}^{t} (\ell_i' + \beta_i)$$

where $\ell_i' = \ell_i - \beta_i$ is the homog. part of the affine form ℓ_i.

- Difficulty 1 : s could be as large as $2^t r^d$.
- Difficulty 2 : $\textsc{Max-Rank}(\prod_i^d m_{ij}')$ can be as large as 2^d.

Step 1 : Upper bound from the model.

Lemma

For a $\Sigma\Pi\Sigma$ circuit C with product dimension r, computing a degree d polynomial, for any equipartition,
$\textsc{Max-Rank}(C) \leq (\textit{top fanin})\binom{d+r}{r}(d+1)$.

Proof.

- Consider a product gate $Q = \prod_{i=1}^{t} \ell_i$.
- Let ℓ_i's (for this Q) be spanned by the affine forms $m_1, \ldots m_r$.

$$Q = \sum_{j=1}^{2^t} c_j \left(\prod_{i=1}^{\leq t} \ell'_i \right)$$

where ℓ'_i is the homog. part of the affine form ℓ_i.

- Difficulty 1 : s could be as large as $2^t r^d$.
- Difficulty 2 : $\textsc{Max-Rank}(\prod_i^d m'_{ij})$ can be as large as 2^d.

Step 1 : Upper bound from the model.

Lemma

For a $\Sigma\Pi\Sigma$ circuit C with product dimension r, computing a degree d polynomial, for any equipartition,
$\textsc{Max-Rank}(C) \leq (\textit{top fanin})\binom{d+r}{r}(d+1)$.

Proof.

- Consider a product gate $Q = \prod_{i=1}^{t} \ell_i$.
- Let ℓ_i's (for this Q) be spanned by the affine forms $m_1, \ldots m_r$.

$$Q = \sum_{j=1}^{2^t} c_j \left(\prod_{i=1}^{\leq t} (\alpha_{i1} m'_1 + \alpha_{i2} m'_2 + \ldots + \alpha_{ir} m'_r) \right)$$

where m'_i is the homogenous part of the linear form m_i.

- Difficulty 1 : s could be as large as $2^t r^d$.
- Difficulty 2 : $\textsc{Max-Rank}(\prod_i^d m'_{ij})$ can be as large as 2^d.

Step 1 : Upper bound from the model.

Lemma

For a $\Sigma\Pi\Sigma$ circuit C with product dimension r, computing a degree d polynomial, for any equipartition,
$\textsc{Max-Rank}(C) \leq (\textit{top fanin})\binom{d+r}{r}(d+1)$.

Proof.

- Consider a product gate $Q = \prod_{i=1}^{t} \ell_i$.
- Let ℓ_i's (for this Q) be spanned by the affine forms $m_1, \ldots m_r$.

$$Q = \sum_{j=1}^{s} c'_j \prod_{i=1}^{d} m'_{ij}$$

  where s could be as large as $2^t r^d$.

- Difficulty 1 : s could be as large as $2^t r^d$.
- Difficulty 2 : $\textsc{Max-Rank}(\prod_i^d m'_{ij})$ can be as large as 2^d.

Step 1 contd

- Observe : $\textsc{Max-Rank}(\ell^d) \leq d + 1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.

Step 1 contd

- Observe : $\textsc{Max-Rank}(\ell^d) \leq d + 1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.
- $\textsc{Shpilka}$ (2001): Any monomial of degree d can be written as the sum of d^{th} powers of 2^d linear forms - the linear forms are $\sum_{x \in S} x$ for $S \subseteq [d]$.

Step 1 contd

- Observe : $\textsc{Max-Rank}(\ell^d) \leq d+1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.
- $\textsc{Shpilka}$ (2001): Any monomial of degree d can be written as the sum of d^{th} powers of 2^d linear forms - the linear forms are $\sum_{x \in S} x$ for $S \subseteq [d]$.
- $S = \prod_{i=1}^{d} \ell_i$ to $S = \sum_{t=1}^{2^d} (L_t)^d$.

Step 1 contd

- Observe : $\textsc{Max-Rank}(\ell^d) \leq d+1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.
- $\textsc{Shpilka}$ (2001): Any monomial of degree d can be written as the sum of d^{th} powers of 2^d linear forms - the linear forms are $\sum_{x \in S} x$ for $S \subseteq [d]$.
- $S = \prod_{i=1}^{d} \ell_i$ to $S = \sum_{t=1}^{2^d} (L_t)^d$.

  Each L_t is $\sum_{i \in [r]} \alpha_i \ell_i$ such that $\sum \alpha_i \leq d$.

Step 1 contd

- Observe : $\text{MAX-RANK}(\ell^d) \leq d+1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.
- SHPILKA (2001): Any monomial of degree d can be written as the sum of d^{th} powers of 2^d linear forms - the linear forms are $\sum_{x \in S} x$ for $S \subseteq [d]$.
- $S = \prod_{i=1}^{d} \ell_i$ to $S = \sum_{t=1}^{2^d} (L_t)^d$.

  Each L_t is $\sum_{i \in [r]} \alpha_i \ell_i$ such that $\sum \alpha_i \leq d$.
- Thus, $Q = \sum_{q=1}^{m} c_q \cdot (L_q)^d$ where $m = \binom{d+r}{r}$.

$$\text{MAX-RANK}(Q) \leq (d+1)\binom{d+r}{r}$$

Step 1 contd

- Observe : $\textsc{Max-Rank}(\ell^d) \leq d+1$. The idea is to express express a product of linear forms as a sum of product of powers of linear forms.
- $\textsc{Shpilka}$ (2001): Any monomial of degree d can be written as the sum of d^{th} powers of 2^d linear forms - the linear forms are $\sum_{x \in S} x$ for $S \subseteq [d]$.
- $S = \prod_{i=1}^{d} \ell_i$ to $S = \sum_{t=1}^{2^d} (L_t)^d$.

  Each L_t is $\sum_{i \in [r]} \alpha_i \ell_i$ such that $\sum \alpha_i \leq d$.
- Thus, $Q = \sum_{q=1}^{m} c_q \cdot (L_q)^d$ where $m = \binom{d+r}{r}$.

$$\textsc{Max-Rank}(C) \leq k(d+1)\binom{d+r}{r}$$

Step 2: Constructing the hard polynomial

Lemma

There is a polynomial P of degree $\frac{n}{2}$ and a partition such that there is a partition for which $\textsc{Max-Rank}(P) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.

Proof.

- Fix $Y = \{x_1, x_2, \ldots, x_{\frac{n}{2}}\}$ and $Z = \{x_{\frac{n}{2}+1}, \ldots, x_n\}$.
- Let $S_1 \ldots S_\ell$ and $T_1 \ldots T_\ell$ be canonically ordered subsets of Y and Z of size exactly $\frac{n}{4}$ where $\ell = \binom{n/2}{n/4}$.

Step 2: Constructing the hard polynomial

Lemma

There is a polynomial P of degree $\frac{n}{2}$ and a partition such that there is a partition for which $\textsc{Max-Rank}(P) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.

Proof.

- Fix $Y = \{x_1, x_2, \ldots, x_{\frac{n}{2}}\}$ and $Z = \{x_{\frac{n}{2}+1}, \ldots, x_n\}$.
- Let $S_1 \ldots S_\ell$ and $T_1 \ldots T_\ell$ be canonically ordered subsets of Y and Z of size exactly $\frac{n}{4}$ where $\ell = \binom{n/2}{n/4}$.

$$P = \sum_{i=1}^{\ell} \prod_{y \in S_i} \prod_{z \in T_i} (yz)$$

Step 2: Constructing the hard polynomial

Lemma

There is a polynomial P of degree $\frac{n}{2}$ and a partition such that there is a partition for which $\textsc{Max-Rank}(P) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.

Proof.

- Fix $Y = \{x_1, x_2, \ldots, x_{\frac{n}{2}}\}$ and $Z = \{x_{\frac{n}{2}+1}, \ldots, x_n\}$.
- Let $S_1 \ldots S_\ell$ and $T_1 \ldots T_\ell$ be canonically ordered subsets of Y and Z of size exactly $\frac{n}{4}$ where $\ell = \binom{n/2}{n/4}$.

$$P = \sum_{i=1}^{\ell} \prod_{y \in S_i} \prod_{z \in T_i} (yz)$$

- In the matrix, only the diagonal entries of these corresponding subsets will be non-zero. Thus, $\textsc{Max-Rank}(P) \geq \binom{\frac{n}{2}}{\frac{n}{4}} \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$.

□

Choosing the parameters

$$k \times \binom{d+r}{r}(d+1) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$$

$d = \frac{n}{2}$, $r = \frac{n}{10}$ gives, $k \geq 2^{cn}$ for some constant $c > 0$.

Lemma

The polynomial P can be computed by a diagonal circuit (hence product dimension 1) of size 2^n.

Choosing the parameters

$$k \times \binom{d+r}{r}(d+1) \geq \frac{2^{\frac{n}{2}}}{\sqrt{n}}$$

$d = \frac{n}{2}$, $r = \frac{n}{10}$ gives, $k \geq 2^{cn}$ for some constant $c > 0$.

Lemma

The polynomial P can be computed by a diagonal circuit (hence product dimension 1) of size 2^n.

Proof.

- Express the polynomial as a sum of monomials.
- Express each monomial as a sum of powers of linear forms.
- Each product gate has product dimension 1.
- The resulting circuit is of depth d and of size $2^{O(n)}$.

□

Concluding remarks

- We showed lower bounds against depth three homogeneous circuits, depth three circuits of product dimension $\frac{n}{10}$.

Concluding remarks

- We showed lower bounds against depth three homogeneous circuits, depth three circuits of product dimension $\frac{n}{10}$. (Follow up : can be improved to $\frac{n}{4}$).

Concluding remarks

- We showed lower bounds against depth three homogeneous circuits, depth three circuits of product dimension $\frac{n}{10}$. (Follow up : can be improved to $\frac{n}{4}$).
- So close, yet so far : the techniques so far do not distinguish between determinant and permanent. What makes them distinct? Properties?

Concluding remarks

- We showed lower bounds against depth three homogeneous circuits, depth three circuits of product dimension $\frac{n}{10}$. (Follow up : can be improved to $\frac{n}{4}$).
- So close, yet so far : the techniques so far do not distinguish between determinant and permanent. What makes them distinct? Properties?
- Open Problem : Is there a chasm at depth three for finite fields?
- Open Problem : Is there a depth reduction to depth three homogeneous circuits?
- Open Problem : Unify our method with the shifted partial derivatives method of $GKKS12$.

so close . . . yet so far . . .

Thanks !

Questions?