Finer separations between shallow arithmetic circuits

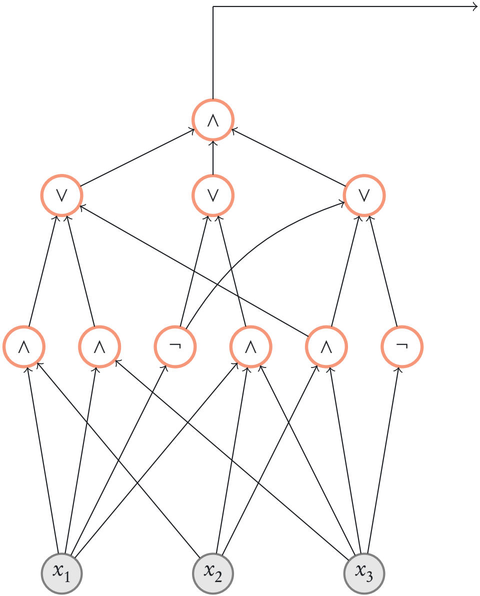
Mrinal Kumar
Rutgers University

Ramprasad Saptharishi
TIFR

FSTTCS 2016
Chennai

(Work done while in Tel Aviv University)

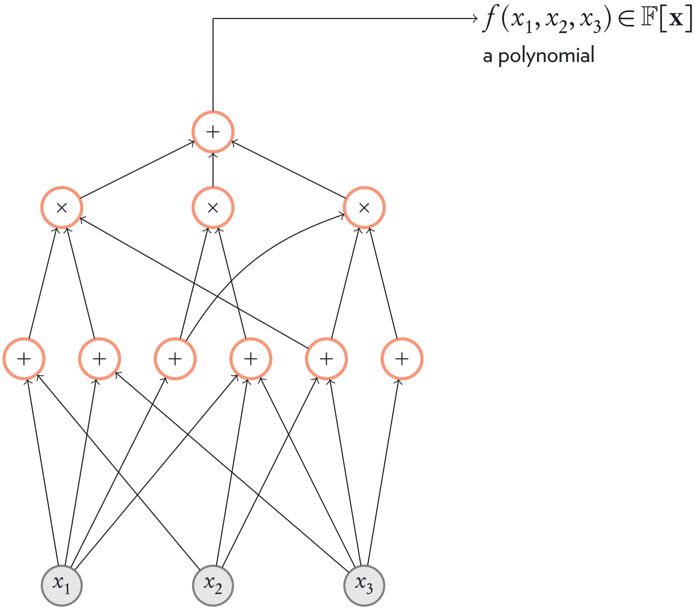
Boolean circuits



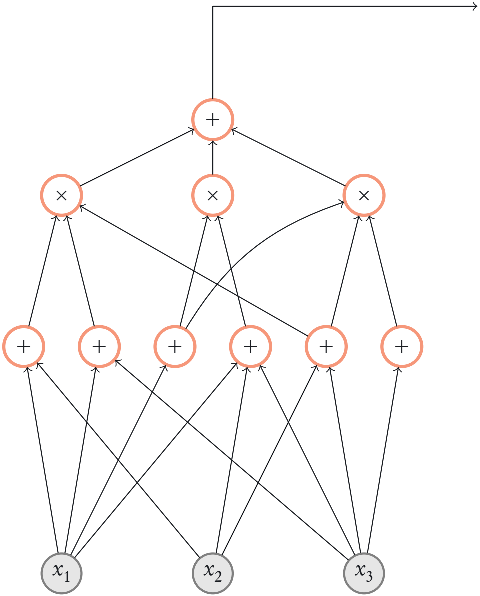
$f(x_1, x_2, x_3) : \{0,1\}^n \to \{0,1\}$

a boolean function

Arithmetic circuits



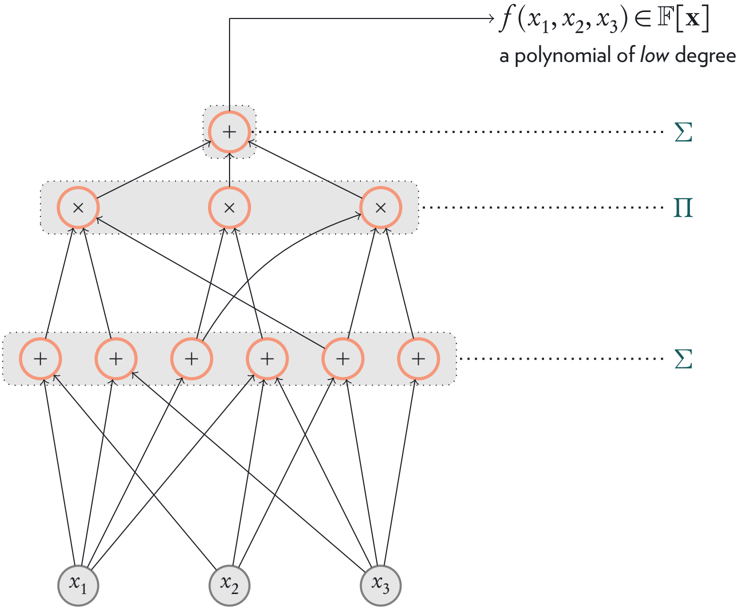
Arithmetic circuits



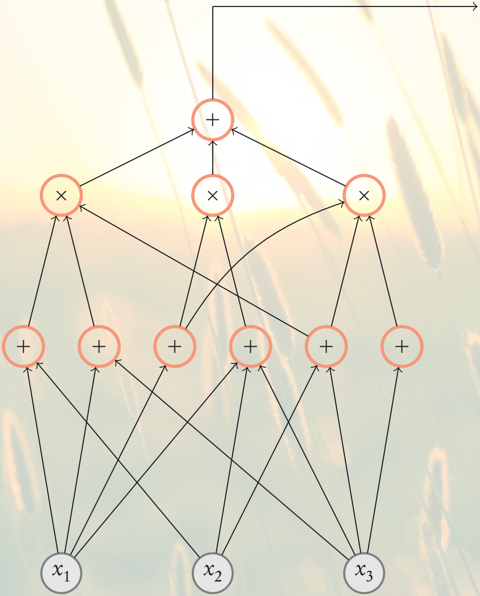
$f(x_1, x_2, x_3) \in \mathbb{F}[\mathbf{x}]$

a polynomial of *low* degree

Arithmetic circuits



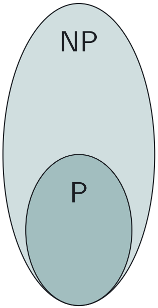
Arithmetic circuits



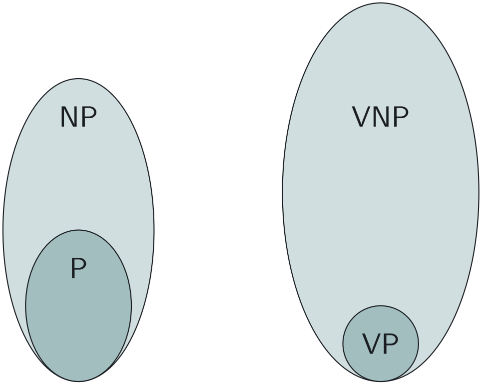
$f(x_1, x_2, x_3) \in \mathbb{F}[\mathbf{x}]$

a polynomial of *low* degree

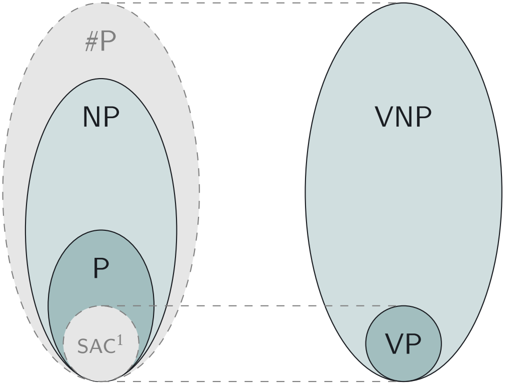
The Open Problem(s)



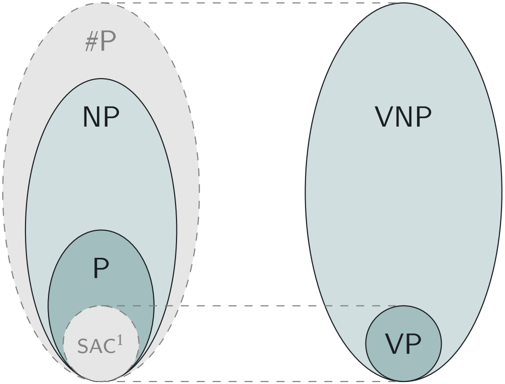
The Open Problem(s)



The Open Problem(s)

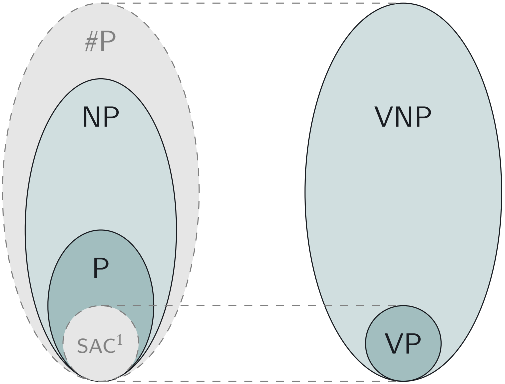


The Open Problem(s)



VP $\neq$ VNP is simpler to prove than P $\neq$ NP.

The Open Problem(s)



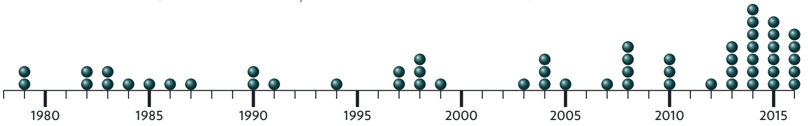
VP $\neq$ VNP is simpler to prove than P $\neq$ NP.

Ultimate goal: Find an explicit n-variate degree d polynomial that requires large arithmetic circuits to compute it.

Recent activity in algebraic complexity

A recent surge in optimism in the field.

Someone even conjectured that VP $\neq$ VNP would be resolved by 2018...

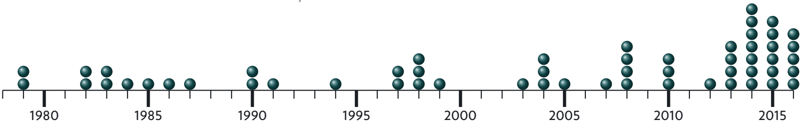


Each ● is one result in algebraic complexity lower bounds.

Recent activity in algebraic complexity

A recent surge in optimism in the field.

Someone even conjectured that VP $\neq$ VNP would be resolved by 2018...



Each ● is one result in algebraic complexity lower bounds.

Point of attack: Lower bounds for very shallow circuits

Completely different situation from the boolean world!

Depth Reduction

Theorem ([Agrawal-Vinay + Koiran, Tavenas])

Can be computed by		*Can be computed by*
arithmetic circuits	$\longrightarrow$	*depth-4 circuits**
of "small" size		*of "not-too-large" size*

Depth Reduction

Theorem ([Agrawal-Vinay + Koiran, Tavenas])

Can be computed by		*Can be computed by*
arithmetic circuits	$\longrightarrow$	*depth-4 circuits**
of $\mathrm{poly}(n, d)$ *size*		*of* $n^{O(\sqrt{d})}$ *size*

Depth Reduction

Theorem ([Agrawal-Vinay + Koiran, Tavenas])

Can be computed by *arithmetic circuits* *of* $\mathrm{poly}(n,d)$ *size*	$\longrightarrow$	*Can be computed by* *depth-4 circuits** *of* $n^{O(\sqrt{d})}$ *size*

(Or)

Cannot be computed by *arithmetic circuits* *of* $\mathrm{poly}(n,d)$ *size*	$\longleftarrow$	*Cannot be computed by* *depth-4 circuits** *of* $n^{O(\sqrt{d})}$ *size*

A brief history of recent results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for depth-4 circuits*.

A brief history of recent results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for depth-4 circuits*.

Theorem ([Nisan-Wigderson])

A $2^{\Omega(d)}$ lower bound for depth-3 circuits.*

A brief history of recent results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for depth-4 circuits*.

Theorem ([Nisan-Wigderson])

A $2^{\Omega(d)}$ lower bound for depth-3 circuits*.

Theorem ([Kayal])

A $2^{\Omega(d)}$ lower bound for a sub-class of depth-4 circuits*.

A brief history of recent results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for depth-4 circuits*.

Theorem ([Nisan-Wigderson])

A $2^{\Omega(d)}$ lower bound for depth-3 circuits.*

Theorem ([Kayal])

A $2^{\Omega(d)}$ lower bound for a sub-class of depth-4 circuits.*

Theorem ([Gupta-Kamath-Kayal-Saptharishi])

A $2^{\Omega(\sqrt{d})}$ lower bound for depth-4 circuits.*

A brief history of recent results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for depth-4 circuits*.

Theorem ([Nisan-Wigderson])
*A $2^{\Omega(d)}$ lower bound for depth-3 circuits**.*

Theorem ([Kayal])
*A $2^{\Omega(d)}$ lower bound for a sub-class of depth-4 circuits**.*

Theorem ([Gupta-Kamath-Kayal-Saptharishi])
*A $2^{\Omega(\sqrt{d})}$ lower bound for depth-4 circuits**.*

... and the deluge began –
[Fournier-Limaye-Malod-Srinivasan], [Kayal-Saha-Saptharishi], [Chillara-Mukhopadhyay], [Kumar-Saraf], [Kayal-Limaye-Saha-Srinivasan], [Kumar-Saraf], [Kayal-Saha], [Kumar-Saraf], [Kayal-Saha], [Bera-Chakrabarti], [Kumar-Saraf], [Kumar-Saptharishi], [Forbes-Kumar-Saptharishi], [Kumar-Saraf],

Three types of improvements

Theorem

An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

Three types of improvements

Theorem

An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

- **Better size lower bound:**

Three types of improvements

Theorem

An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

- **Better size lower bound:**

  Theorem ([Kayal-Saha-Saptharishi])

  An explicit polynomial f such that any depth-4 circuit* computing it must have size $n^{\Omega(\sqrt{d})}$.

  "Can we cross the $n^{\Omega(\sqrt{d})}$ threshold and get a $n^{\omega(\sqrt{d})}$ lower bound (and hence VP $\neq$ VNP)?"

Three types of improvements

Theorem

An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

- **Lower bound for more general classes:**

Three types of improvements

Theorem
An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

- **Lower bound for more general classes:**

  Theorem ([Kayal-Limaye-Saha-Srinivasan])
  An explicit polynomial f such that any hom. depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

  "Can we expect a more efficient depth reduction to this general class?"

Three types of improvements

Theorem

*An explicit polynomial f such that any depth-4 circuit** computing it must have size* $2^{\Omega(\sqrt{d})}$.

- **Lower bounds for *easier* polynomials:**

Three types of improvements

Theorem
An explicit polynomial f such that any depth-4 circuit* computing it must have size $2^{\Omega(\sqrt{d})}$.

- **Lower bounds for *easier* polynomials:**

  Theorem ([Kumar-Saraf])
  An explicit polynomial f, which is in VP, such that any depth-4 circuit* computing it must have size $n^{\Omega(\sqrt{d})}$.

  "Can we possibly hope to use this measure *as is* separate VP and VNP?"
  "Can we get a more efficient depth reduction? What about formulas, or depth-10 circuits?"

Three types of improvements

Theorem

An explicit polynomial f such that any depth-4 circuit * *computing it must have size* $2^{\Omega(\sqrt{d})}$.

- **Lower bounds for *easier* polynomials:**

  Theorem ([Kumar-Saraf])

  An explicit polynomial f, which is in VP, such that any depth-4 circuit * *computing it must have size* $n^{\Omega(\sqrt{d})}$.

  "Can we possibly hope to use this measure *as is* separate VP and VNP?"
  "Can we get a more efficient depth reduction? What about formulas, or depth-10 circuits?"

... this talk is in this category

Our results

Theorem (Separation between depth-4 and 5)

There is an explicit n-variate degree $d = O(\log^2 n)$ polynomial f, that is computable by depth-5 circuits, such that any homogeneous depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

Our results

Theorem (Separation between depth-4 and 5)
There is an explicit n-variate degree $d = O(\log^2 n)$ polynomial f, that is computable by depth-5 circuits, such that any homogeneous depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

So, no; you can't hope for a more efficient depth reduction to hom. depth-4 circuit from even depth-5 circuits.

Our results

Theorem (Separation between depth-4 and 5)

There is an explicit n-variate degree $d = O(\log^2 n)$ polynomial f, that is computable by depth-5 circuits, such that any homogeneous depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

Theorem (Separation between non hom. depth-3 and hom. depth-4)

The above polynomial actual has a $\mathrm{poly}(n)$-sized non-hom. depth-3 circuit computing it.

Our results

Theorem (Separation between depth-4 and 5)
There is an explicit n-variate degree $d = O(\log^2 n)$ polynomial f, that is computable by depth-5 circuits, such that any homogeneous depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

Theorem (Separation between non hom. depth-3 and hom. depth-4)
The above polynomial actual has a $\text{poly}(n)$-sized non-hom. depth-3 circuit computing it.

The polynomial happens to have a poly-sized non-hom. depth-3 circuit.

Our results

Theorem (Separation between depth-4 and 5)
There is an explicit n-variate degree $d = O(\log^2 n)$ polynomial f, that is computable by depth-5 circuits, such that any homogeneous depth-4 circuit computing it must have size $n^{\Omega(\sqrt{d})}$.

Theorem (Separation between non hom. depth-3 and hom. depth-4)
The above polynomial actual has a $\text{poly}(n)$-sized non-hom. depth-3 circuit computing it.

Open: Get rid of the $d = O(\log^2 n)$ caveat.

How does one prove lower bounds?

Goal: *To show that a certain target polynomial cannot be computed efficiently in* $\mathscr{C}$.

How does one prove lower bounds?

Goal: *To show that a certain target polynomial cannot be computed efficiently in* $\mathscr{C}$.

- Complexity measure: Assigns a number to every polynomial.
  Bigger the number, *harder* the polynomial.

How does one prove lower bounds?

Goal: *To show that a certain target polynomial cannot be computed efficiently in $\mathscr{C}$.*

- Complexity measure: Assigns a number to every polynomial.
  Bigger the number, *harder* the polynomial.
- Typically, complexity measures are ranks of some linear operators applied on these polynomials.
  Evaluations, derivatives, *shifts*, projections, etc.

How does one prove lower bounds?

Goal: *To show that a certain target polynomial cannot be computed efficiently in $\mathscr{C}$.*

- Complexity measure: Assigns a number to every polynomial.
  Bigger the number, *harder* the polynomial.
- Typically, complexity measures are ranks of some linear operators applied on these polynomials.
  Evaluations, derivatives, *shifts*, projections, etc.
- Somehow show that the measure is *small* for polynomials in $\mathscr{C}$, and *large* on the target polynomial.

Examples of complexity measures

- Hom. depth-3 circuits (or $\Sigma\Pi^{[d]}\Sigma$ circuits):

$$\Gamma_k(f) \quad := \quad \dim\left\{\partial^{=k}(f)\right\}$$

Examples of complexity measures

- Hom. depth-3 circuits (or $\Sigma\Pi^{[d]}\Sigma$ circuits):

$$\Gamma_k(f) \quad := \quad \dim\left\{\partial^{=k}(f)\right\}$$

- Depth-4 circuits* (or $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits):

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\mathbf{x}^{=\ell}\,\partial^{=k}(f)\right\}$$

Examples of complexity measures

- Hom. depth-3 circuits (or $\Sigma\Pi^{[d]}\Sigma$ circuits):

$$\Gamma_k(f) \quad := \quad \dim\left\{\partial^{=k}(f)\right\}$$

- Depth-4 circuits* (or $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits):

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\mathbf{x}^{=\ell}\,\partial^{=k}(f)\right\}$$

- Hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\mathrm{mult}\left(\mathbf{x}^{=\ell}\,\partial^{=k}(\rho(f))\right)\right\}$$

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\text{mult}\left(\mathbf{x}^{=\ell}\,\partial^{=k}(\rho(f))\right)\right\}$$

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\text{mult}\left(\mathbf{x}^{=\ell}\,\partial^{=k}(\rho(f))\right)\right\}$$

Hence, if f is computable by a depth-4 circuit the $\Gamma_{k,\ell}(f)$ is *small*.

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\mathrm{mult}\left(\mathbf{x}^{=\ell}\partial^{=k}(\rho(f))\right)\right\}$$

Hence, if f is computable by a depth-4 circuit the $\Gamma_{k,\ell}(f)$ is *small*.

Question: What is $\Gamma_{k,\ell}(f)$ for a *generic depth-5 circuit*?

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\mathrm{mult}\left(\mathbf{x}^{=\ell}\partial^{=k}(\rho(f))\right)\right\}$$

Hence, if f is computable by a depth-4 circuit the $\Gamma_{k,\ell}(f)$ is *small*.

Question: What is $\Gamma_{k,\ell}(f)$ for a *generic depth-5 circuit*?

The answer *ought to be* large, but we do not know how to prove it.

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\text{mult}\left(\mathbf{x}^{=\ell}\,\partial^{=k}(\rho(f))\right)\right\}$$

Hence, if f is computable by a depth-4 circuit the $\Gamma_{k,\ell}(f)$ is *small*.

Question: What is $\Gamma_{k,\ell}(f)$ for a *generic depth-5 circuit*?

The answer *ought to be* large, but we do not know how to prove it.

[K-S]: a closely related measure is *small* when the field is small!

A strategy for separating d-4 and d-5

Complexity measure for hom. depth-4 circuits:

$$\Gamma_{k,\ell}(f) \quad := \quad \dim\left\{\text{mult}\left(\mathbf{x}^{=\ell}\,\partial^{=k}(\rho(f))\right)\right\}$$

Hence, if f is computable by a depth-4 circuit the $\Gamma_{k,\ell}(f)$ is *small*.

Question: What is $\Gamma_{k,\ell}(f)$ for a *generic depth-5 circuit*?

The answer *ought to be* large, but we do not know how to prove it.

[K-S]: a closely related measure is *small* when the field is small!

Baby steps: What's the effect of a random restriction ρ on a depth-4 and a depth-5 circuit?

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Effects of random restrictions

ρ : set each x_i to zero with probability $1 - \frac{1}{n^{0.1}}$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Observation

Let $r = \sqrt{d}$. Then,

$$\Pr[\rho(x_1 \cdots x_r) \neq 0] \leq \frac{1}{n^{0.1r}}$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Observation

Let $r = \sqrt{d}$. Then,

$$\Pr[\rho(x_1 \cdots x_r) \neq 0] \quad \leq \quad \frac{1}{n^{0.1r}}$$

Lemma

If a depth-4 circuit C has size s, then

$$\Pr[\rho(C) \text{ is not a bottom-support-} r \text{ circuit}] \quad \leq \quad \frac{s}{n^{0.1r}}$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Observation

Let $r = \sqrt{d}$. Then,

$$\Pr[\rho(x_1 \cdots x_r) \neq 0] \quad \leq \quad \frac{1}{n^{0.1r}}$$

Lemma

If a depth-4 circuit C has size $s = n^{0.01r}$, then

$$\Pr[\rho(C) \text{ is not a bottom-support-}r \text{ circuit}] \quad \leq \quad \frac{s}{n^{0.1r}}$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Observation

Let $r = \sqrt{d}$. Then,

$$\Pr[\rho(x_1 \cdots x_r) \neq 0] \quad \leq \quad \frac{1}{n^{0.1r}}$$

Lemma

If a depth-4 circuit C has size $s = n^{0.01r}$, then

$$\Pr[\rho(C) \text{ is not a bottom-support-}r\text{ circuit}] \quad \leq \quad \frac{n^{0.01r}}{n^{0.1r}}$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Observation

Let $r = \sqrt{d}$. Then,

$$\Pr[\rho(x_1 \cdots x_r) \neq 0] \quad \leq \quad \frac{1}{n^{0.1r}}$$

Lemma

If a depth-4 circuit C has size $s = n^{0.01r}$, then

$$\Pr[\rho(C) \text{ is not a bottom-support-} r \text{ circuit}] \quad \leq \quad o(1)$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-4 circuits: $\Sigma\Pi\Sigma\Pi$

Bottom level consists of monomials.

Take away

If C is a small-ish $\Sigma\Pi\Sigma\Pi$ circuit, then $\rho(C)$ is a low-bottom-support depth-4 circuit whp.

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-5 circuits: $\Sigma\Pi\Sigma\Pi\Sigma$

Bottom level consists of linear polynomials.

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-5 circuits: $\Sigma\Pi\Sigma\Pi\Sigma$

Bottom level consists of linear polynomials.

Observation

If $\ell = x_1 + \cdots + x_m$, then

$$\Pr[\, \rho(\ell) = 0 \,] \quad \leq \quad \left(1 - \frac{1}{n^{0.1}}\right)^m$$

Effects of random restrictions

ρ : set each x_i to zero with probability $1 - \frac{1}{n^{0.1}}$

Depth-5 circuits: $\Sigma\Pi\Sigma\Pi\Sigma$

Bottom level consists of linear polynomials.

Observation

If $\ell = x_1 + \cdots + x_m$ with $m = n^{0.2}$, then

$$\Pr[\,\rho(\ell) = 0\,] \leq \left(1 - \frac{1}{n^{0.1}}\right)^m$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-5 circuits: $\Sigma\Pi\Sigma\Pi\Sigma$

Bottom level consists of linear polynomials.

Observation

If $\ell = x_1 + \cdots + x_m$ with $m = n^{0.2}$, then

$$\Pr[\,\rho(\ell) = 0\,] \quad \leq \quad \left(1 - \frac{1}{n^{0.1}}\right)^m \quad = \quad \frac{1}{\exp(n^{0.1})}$$

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Depth-5 circuits: $\Sigma\Pi\Sigma\Pi\Sigma$

Bottom level consists of linear polynomials.

Observation

If $\ell = x_1 + \cdots + x_m$ with $m = n^{0.2}$, then

$$\Pr[\,\rho(\ell) = 0\,] \quad \leq \quad \left(1 - \frac{1}{n^{0.1}}\right)^m \quad = \quad \frac{1}{\exp(n^{0.1})}$$

Hence, if C is a wide $\Sigma\Pi\Sigma\Pi\Sigma$ circuit of $\mathrm{poly}(n)$ size, then w.h.p. none of the bottom linear polynomials are killed.

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Take away

If C is a small-ish $\Sigma\Pi\Sigma\Pi$ circuit, then $\rho(C)$ is a low-bottom-support depth-4 circuit whp.

If C is a small and wide $\Sigma\Pi\Sigma\Pi\Sigma$ circuit, then $\rho(C)$ still has within it a high-bottom-support depth-4 circuit whp.

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Take away

If C is a small-ish $\Sigma\Pi\Sigma\Pi$ circuit, then $\rho(C)$ is a low-bottom-support depth-4 circuit whp.

If C is a small and wide $\Sigma\Pi\Sigma\Pi\Sigma$ circuit, then $\rho(C)$ still has within it a high-bottom-support depth-4 circuit whp.

Task reduced to separating high-bottom-support and low-bottom-support depth-4 circuits.

Effects of random restrictions

$$\rho \quad : \quad \text{set each } x_i \text{ to zero with probability } 1 - \frac{1}{n^{0.1}}$$

Take away

If C is a small-ish $\Sigma\Pi\Sigma\Pi$ circuit, then $\rho(C)$ is a low-bottom-support depth-4 circuit whp.

If C is a small and wide $\Sigma\Pi\Sigma\Pi\Sigma$ circuit, then $\rho(C)$ still has within it a high-bottom-support depth-4 circuit whp.

Task reduced to separating high-bottom-support and low-bottom-support depth-4 circuits.

[Kumar-Saraf] separated low-bottom-**degree** and high-bottom-**degree** depth-4 circuits.

Exploiting more structure

A generic $\Pi\Sigma\Pi$ circuit of high-bottom-support ($\approx 100\sqrt{d}$)

$$C \quad = \quad \prod_{i=1}^{\sqrt{d}/100} \sum_{j=1}^{m} \prod_{k=1}^{100\sqrt{d}} x_{ijk}$$

Observation

This polynomial is set-multilinear. That is, if

$$X_{i*k} \quad = \quad \left\{ x_{ijk} \,:\, j \in [m] \right\}$$

then every monomial in C contains exactly one variable in each X_{i*k}.

The complexity measure

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{ x_{ij} \,:\, j \in [m] \right\}$.

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{ x_{ij} \,:\, j \in [m] \right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{ x_{ij} \ : \ j \in [m] \right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad =$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{x_{ij} \, : \, j \in [m]\right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \qquad\qquad\qquad P'$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{x_{ij} \,:\, j \in [m]\right\}$.

Given a $P(X_1,\ldots,X_d)$, define the polynomial

$$P'(X_1,\ldots,X_d,y_1,\cdots,y_d) \quad = \quad P(y_1X_1,\ldots,y_dX_d).$$

$$\Gamma_{k,\ell}(P) \quad = \qquad\qquad \partial_{\mathbf{x}}^{=k}\left(P'\right)$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{ x_{ij} \,:\, j \in [m] \right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \quad \mathrm{mult}_{\mathbf{y}}\left(\partial_{\mathbf{x}}^{=k}\left(P'\right)\right)$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{x_{ij} \,:\, j \in [m]\right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \quad \mathbf{x}^{=\ell}\, \mathrm{mult}_{\mathbf{y}}\left(\partial_{\mathbf{x}}^{=k}\left(P'\right)\right)$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{x_{ij} \;:\; j \in [m]\right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \quad \dim\left\{\mathbf{x}^{=\ell}\,\mathrm{mult}_{\mathbf{y}}\left(\partial_{\mathbf{x}}^{=k}\left(P'\right)\right)\right\}$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{ x_{ij} \,:\, j \in [m] \right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \quad \dim \left\{ \mathbf{x}^{=\ell} \, \mathrm{mult}_{\mathbf{y}} \left(\partial_{\mathbf{x}}^{=k} \left(P' \right) \right) \right\}$$

The complexity measure

Say $X = X_1 \sqcup \cdots \sqcup X_d$ where $X_i = \left\{x_{ij} \,:\, j \in [m]\right\}$.

Given a $P(X_1, \ldots, X_d)$, define the polynomial

$$P'(X_1, \ldots, X_d, y_1, \cdots, y_d) \quad = \quad P(y_1 X_1, \ldots, y_d X_d).$$

$$\Gamma_{k,\ell}(P) \quad = \quad \dim\left\{\mathbf{x}^{=\ell}\,\mathrm{mult}_{\mathbf{y}}\left(\partial_{\mathbf{x}}^{=k}\left(P'\right)\right)\right\}$$

This measure is not new — a very similar used by [Kayal-Limaye-Saha-Srinivasan] to prove an $n^{\Omega(\sqrt{d})}$ lower bound for hom. depth-4 circuits computing IMM when $d = O(\log^2 n)$.

The proof

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \leq s \cdot 2^{2d} \cdot \binom{d}{k} \cdot \binom{n+\ell+k\sqrt{d}}{n}$$

The proof

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \text{ is small.}$$

The proof

Lemma (Upper bound for circuit class)

If C *is a size* s *hom.* $\Sigma\Pi\Sigma\Pi$ *circuit of low-bottom-support* ($\leq \sqrt{d}$)*, then*

$$\Gamma_{k,\ell}(C) \text{ is small.}$$

Proof.

Follow the proof of [Kayal-Limaye-Saha-Srinivasan].

(The 2^{2d} term comes from a *set-multilinearization* step on C.) □

The proof

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \text{ is small.}$$

Proof.

Follow the proof of [Kayal-Limaye-Saha-Srinivasan].

(The 2^{2d} term comes from a *set-multilinearization* step on C.) □

Lemma (Lower bound for target polynomial)

If C is a generic wide (bottom degree $\approx 100\sqrt{d}$) $\Pi\Sigma\Pi$ circuit, then

$$\Gamma_{k,\ell}(C) \geq \frac{1}{4} \cdot \binom{n+\ell}{n}^{\frac{d-\sqrt{d}}{2}} \cdot \binom{n+\ell-1}{n}$$

The proof

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \text{ is small.}$$

Proof.

Follow the proof of [Kayal-Limaye-Saha-Srinivasan].

(The 2^{2d} term comes from a *set-multilinearization* step on C.) □

Lemma (Lower bound for target polynomial)

If C is a generic wide (bottom degree $\approx 100\sqrt{d}$) $\Pi\Sigma\Pi$ circuit, then

$$\Gamma_{k,\ell}(C) \text{ is large.}$$

The proof

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \text{ is small.}$$

Proof.

Follow the proof of [Kayal-Limaye-Saha-Srinivasan].

(The 2^{2d} term comes from a *set-multilinearization* step on C.) □

Lemma (Lower bound for target polynomial)

If C is a generic wide (bottom degree $\approx 100\sqrt{d}$) $\Pi\Sigma\Pi$ circuit, then

$$\Gamma_{k,\ell}(C) \text{ is large.}$$

Proof.

Follow the proof of [Kumar-Saraf]. □

Here be dragons!

Get me out of here!

A few proof details

Lemma (Upper bound for circuit class)

If C is a size s hom. $\Sigma\Pi\Sigma\Pi$ circuit of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(C) \quad \leq \quad s \cdot 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$Q_1 \cdots Q_r$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$x_{ij} \mapsto y_j x_{ij}$

$$Q_1 \cdots Q_r$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ *are polynomials of low-bottom-support* $(\leq \sqrt{d})$*, then*

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n+\ell+k\sqrt{d}}{n}$$

Proof.

$x_{ij} \mapsto y_j x_{ij}$

$$Q_1' \cdots Q_r'$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$\partial^{=k}(Q_1' \cdots Q_r')$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$\partial^{=k}(Q_1' \cdots Q_r') \in \left\{ \partial_\alpha(Q_A') \cdot Q_{\overline{A}}' \; : \; \begin{matrix} |\alpha| = k\,,\, |A| \leq k \\ A \subseteq [r] \end{matrix} \right\}$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ are polynomials of low-bottom-support ($\leq \sqrt{d}$), then

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$\text{mult}_{\mathbf{y}}\left(\partial^{=k}(Q_1' \cdots Q_r')\right) \in \left\{ \text{mult}_{\mathbf{y}}\left(\partial_\alpha(Q_A') \cdot Q_{\overline{A}}'\right) \; : \; \begin{array}{c} |\alpha| = k \, , \, |A| \leq k \\ A \subseteq [r] \end{array} \right\}$$

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ *are polynomials of low-bottom-support* ($\leq \sqrt{d}$)*, then*

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$\mathrm{mult}_{\mathbf{y}}\left(\partial^{=k}(Q_1' \cdots Q_r')\right) \in \left\{ \mathrm{mult}_{\mathbf{y}}\left(\partial_\alpha(Q_A') \cdot Q_{\overline{A}}'\right) \;:\; \begin{array}{c} |\alpha| = k\,,\, |A| \leq k \\ A \subseteq [r] \end{array} \right\}$$

... a factor of 2^d shows up in *set-multilinearizing* Q_A' and $Q_{\overline{A}}'$.

A few proof details

Lemma (Upper bound for circuit class)

If $Q_1 \cdots Q_r$ *are polynomials of low-bottom-support* ($\leq \sqrt{d}$)*, then*

$$\Gamma_{k,\ell}(Q_1 \cdots Q_r) \quad \leq \quad 2^{2d} \cdot \binom{d}{k} \cdot \binom{n + \ell + k\sqrt{d}}{n}$$

Proof.

$$\mathrm{mult}_{\mathbf{y}}\left(\partial^{=k}(Q_1' \cdots Q_r')\right) \in \left\{ \mathrm{mult}_{\mathbf{y}}\left(\partial_\alpha(Q_A') \cdot Q_{\overline{A}}'\right) \ : \ \begin{array}{c} |\alpha| = k \,,\, |A| \leq k \\ A \subseteq [r] \end{array} \right\}$$

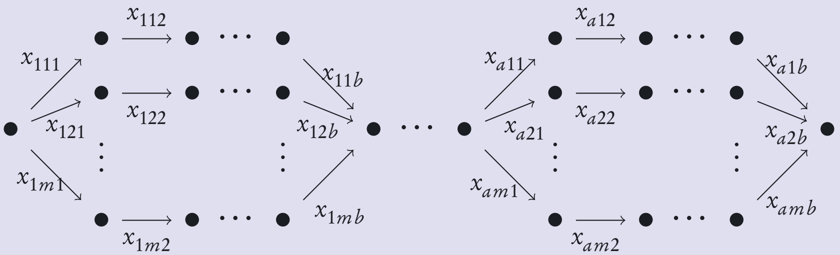
... a factor of 2^d shows up in *set-multilinearizing* Q_A' and $Q_{\overline{A}}'$. □

Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a} \sum_{j=1}^{m} \prod_{k=1}^{b} x_{ijk}$$

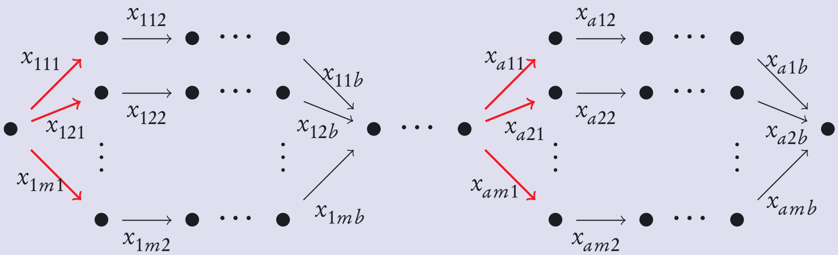
Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



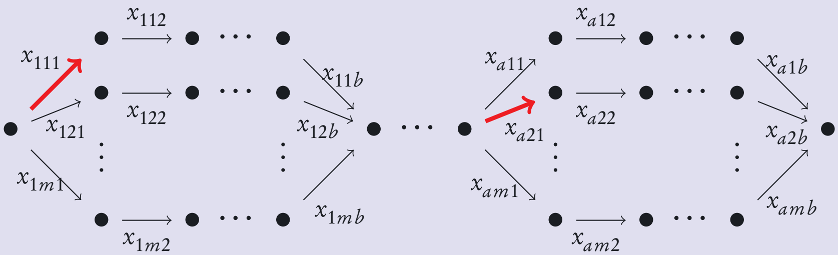
Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



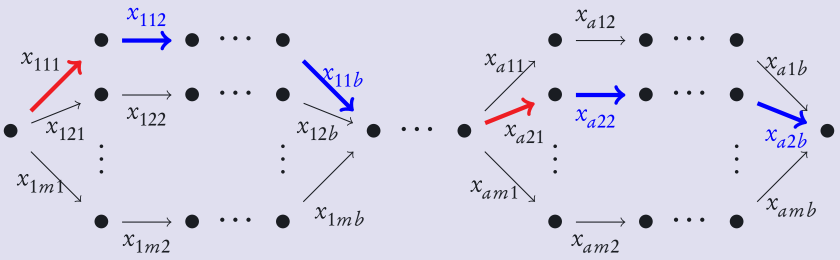
Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



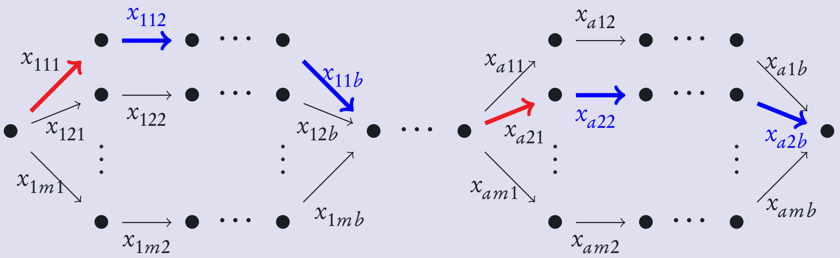
Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



Analysis for a generic $\Pi\Sigma\Pi$ circuit

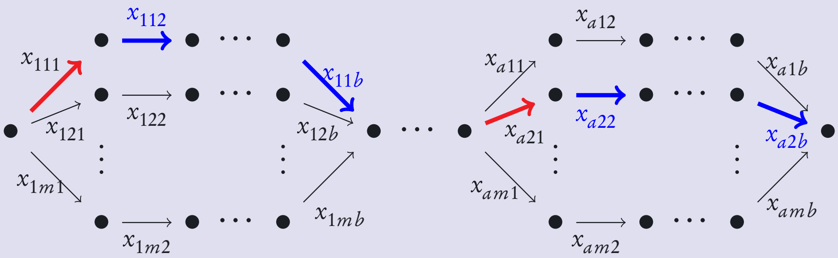
$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



[Chillara-Mukhopadhyay]: If there are many *far* monomials, then there are many distinct shifts and hence $\Gamma_{k,\ell}(C)$ will be large.

Analysis for a generic $\Pi\Sigma\Pi$ circuit

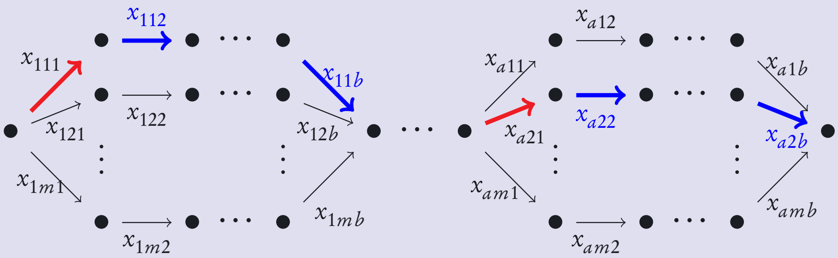
$$\prod_{i=1}^{a} \sum_{j=1}^{m} \prod_{k=1}^{b} x_{ijk}$$



Observation. If derivatives are *far*, then so are the resulting monomials.

Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$

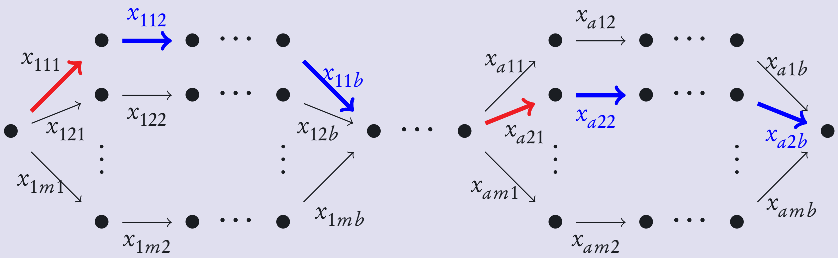


Observation. If derivatives are *far*, then so are the resulting monomials.

Use a *code*!

Analysis for a generic $\Pi\Sigma\Pi$ circuit

$$\prod_{i=1}^{a}\sum_{j=1}^{m}\prod_{k=1}^{b} x_{ijk}$$



Observation. If derivatives are *far*, then so are the resulting monomials.

Use a *code*! □

Summary

Optimal separation between homogeneous depth-4 and homogeneous depth-5, but alas only when $d = O(\log^2 n)$.

Summary

Optimal separation between homogeneous depth-4 and homogeneous depth-5, but alas only when $d = O(\log^2 n)$.

Same for non-hom. depth-3.

Summary

Optimal separation between homogeneous depth-4 and homogeneous depth-5, but alas only when $d = O(\log^2 n)$.

Same for non-hom. depth-3.

Question

Hom. depth-4 lower bounds use the *projected shifted partial derivatives*.

How large *are* these for a generic depth-5 circuit?

Summary

Optimal separation between homogeneous depth-4 and homogeneous depth-5, but alas only when $d = O(\log^2 n)$.

Same for non-hom. depth-3.

Question

Hom. depth-4 lower bounds use the *projected shifted partial derivatives.*

How large *are* these for a generic depth-5 circuit?
You would get a result either way!

Summary

Optimal separation between homogeneous depth-4 and homogeneous depth-5, but alas only when $d = O(\log^2 n)$.

Same for non-hom. depth-3.

Question

Hom. depth-4 lower bounds use the *projected shifted partial derivatives.*

How large *are* these for a generic depth-5 circuit?
You would get a result either way!

\end{document}