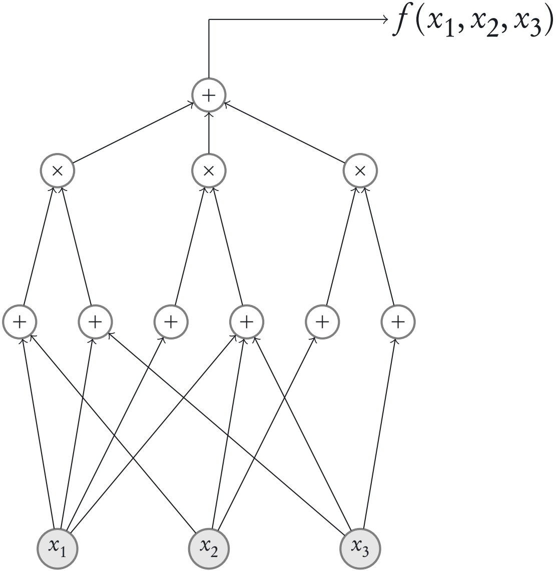
Exponential lower bounds for hom. depth-5 circuits over finite fields

Mrinal Kumar
Rutgers → Harvard

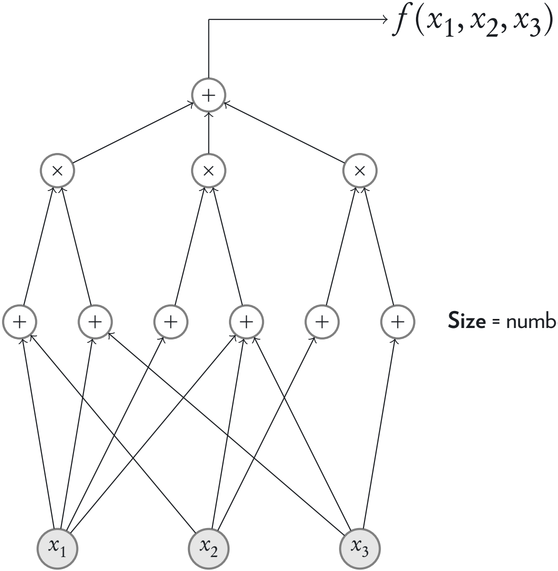
Ramprasad Saptharishi
TIFR, Mumbai

CCC 2017
Riga

Algebraic Circuits

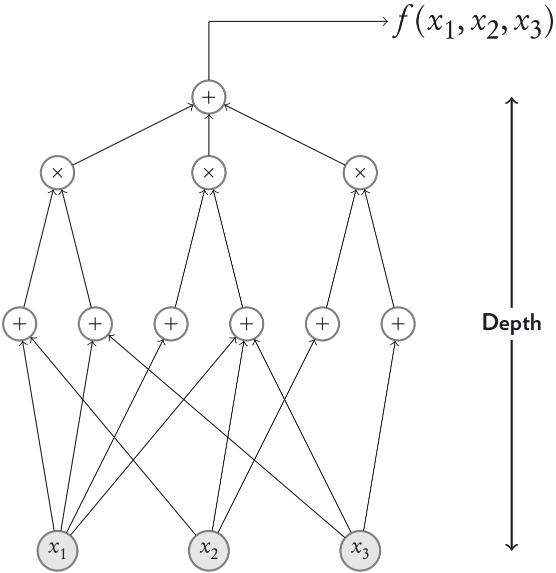


Algebraic Circuits

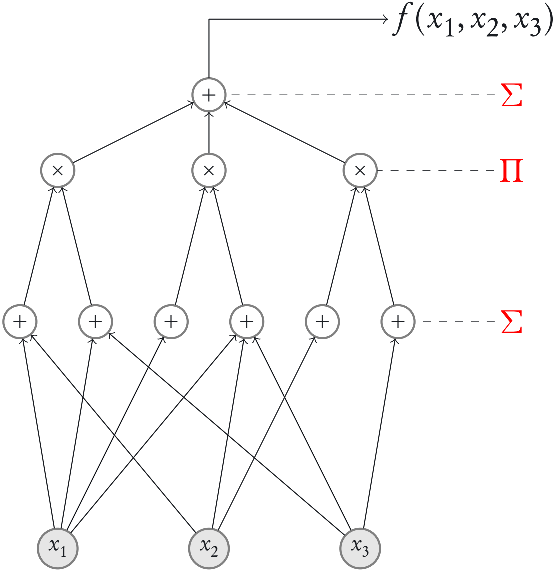


Size = number of gates

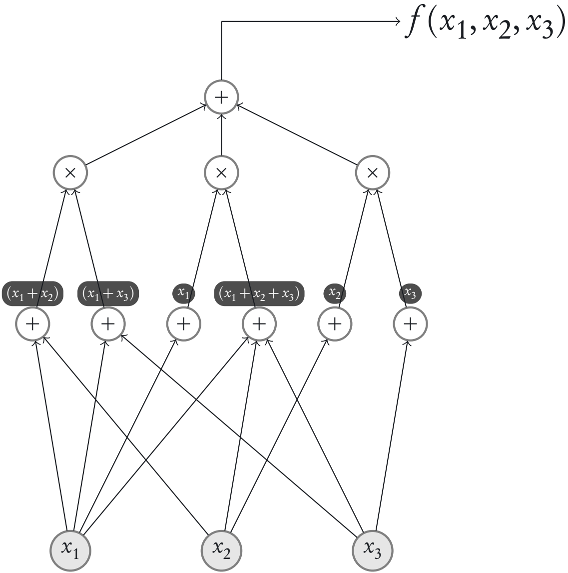
Algebraic Circuits



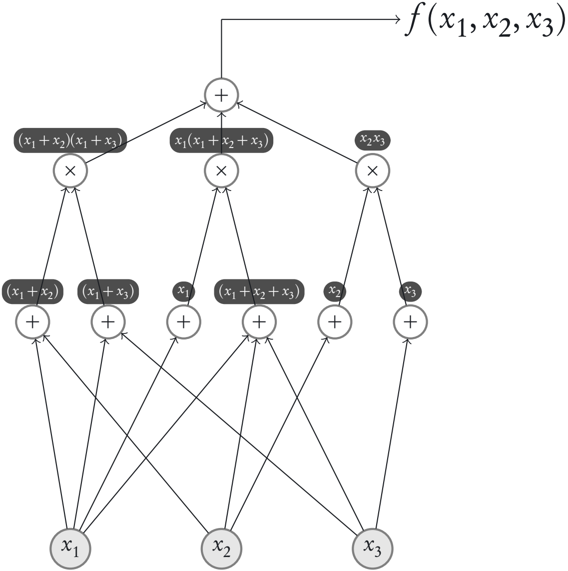
Algebraic Circuits



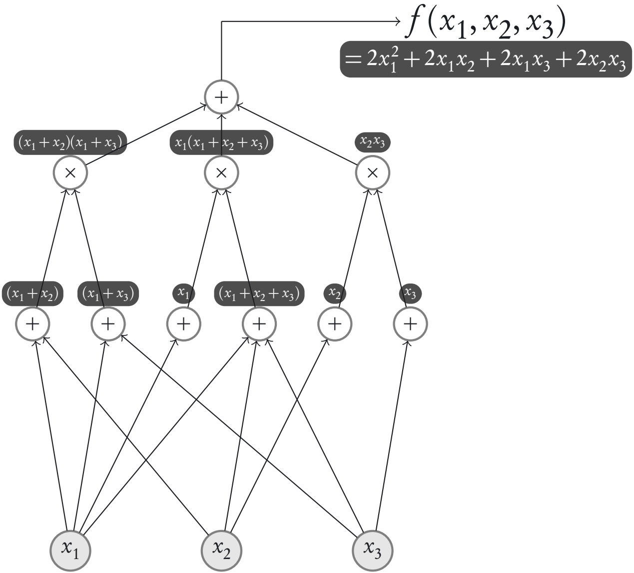
Algebraic Circuits



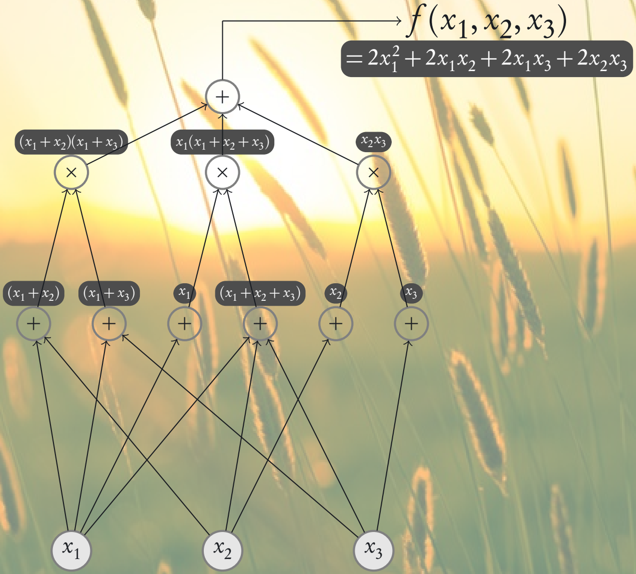
Algebraic Circuits



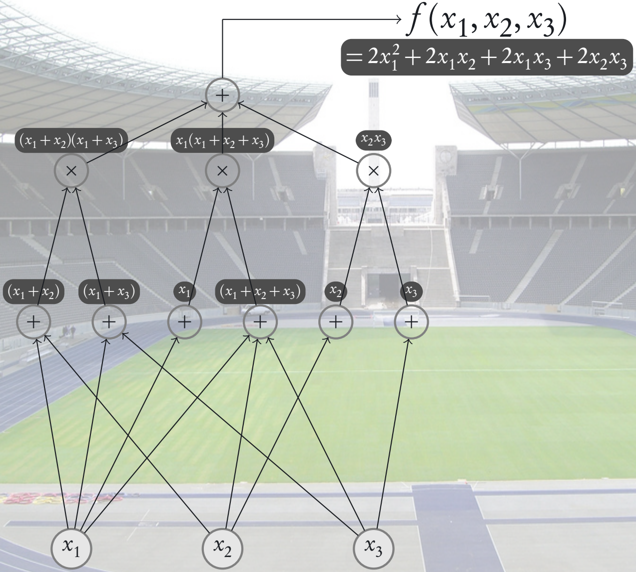
Algebraic Circuits



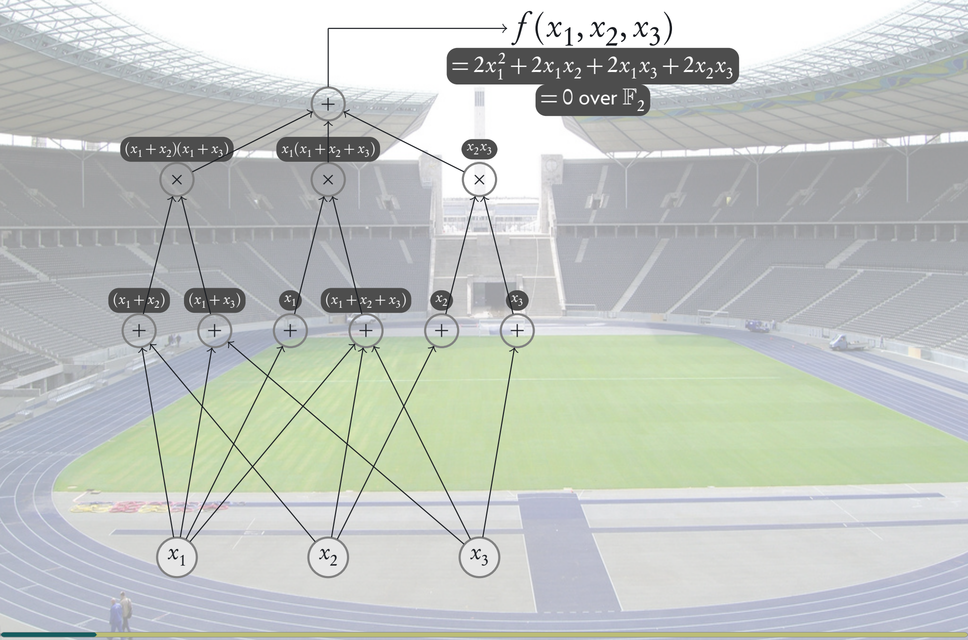
Algebraic Circuits



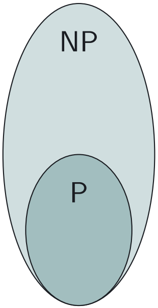
Algebraic Circuits



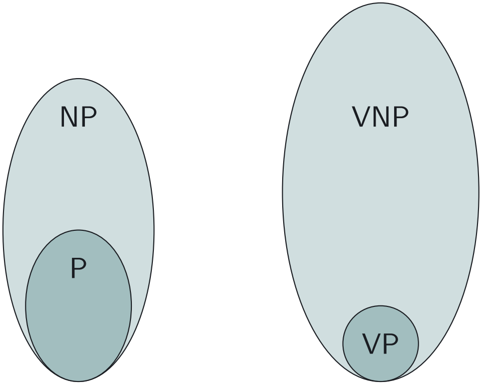
Algebraic Circuits



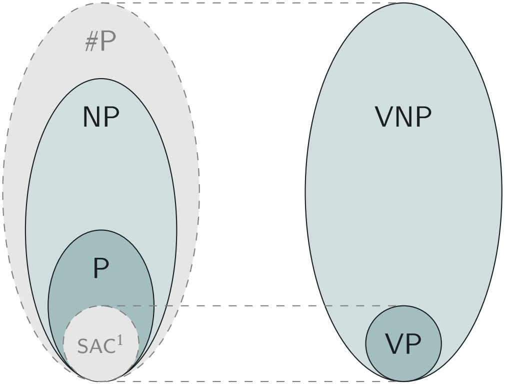
The Open Problem(s)



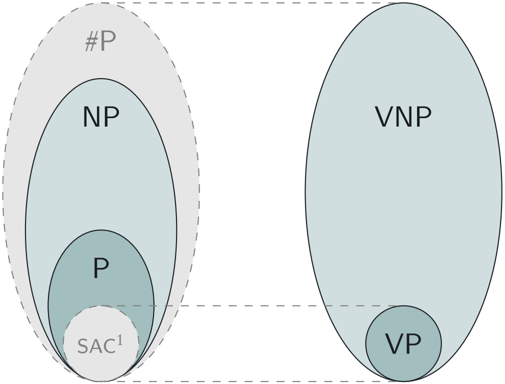
The Open Problem(s)



The Open Problem(s)

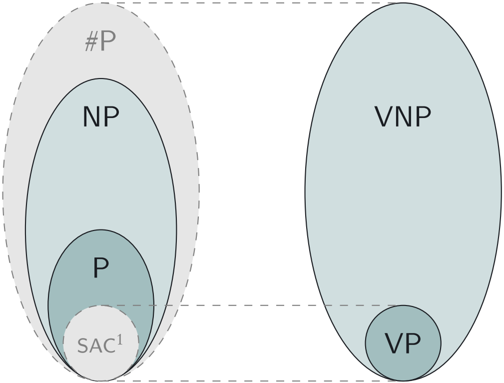


The Open Problem(s)



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.

Depth Reduction

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

Can be computed by		*Can be computed by*
algebraic 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
algebraic circuits	$\longrightarrow$	$\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ *circuits*
of $\mathrm{poly}(n,d)$ *size*		*of* $n^{O(\sqrt{d})}$ *size*

Depth Reduction

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

Can be computed by		*Can be computed by*
algebraic circuits	$\longrightarrow$	$\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ *circuits*
of $\mathrm{poly}(n,d)$ *size*		*of* $n^{O(\sqrt{d})}$ *size*

(Or)

Cannot be computed by		*Cannot be computed by*
algebraic circuits	$\longleftarrow$	$\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ *circuits*
of $\mathrm{poly}(n,d)$ *size*		*of* $n^{O(\sqrt{d})}$ *size*

A brief history of related results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

A brief history of related results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

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

A brief history of related results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

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

Theorem ([Grigoriev-Karpinski, Grigoriev-Razborov])
A $2^{\Omega_q(d)}$ lower bound $\Sigma\Pi\Sigma$ circuits over any fixed finite field $\mathbb{F}_q$

A brief history of related results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

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

Theorem ([Grigoriev-Karpinski, Grigoriev-Razborov])
A $2^{\Omega_q(d)}$ *lower bound* $\Sigma\Pi\Sigma$ *circuits over any fixed finite field* $\mathbb{F}_q$

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

A brief history of related results

Goal: To prove an $n^{\omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

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

Theorem ([Grigoriev-Karpinski, Grigoriev-Razborov])
A $2^{\Omega_q(d)}$ lower bound $\Sigma\Pi\Sigma$ circuits over any fixed finite field $\mathbb{F}_q$

Theorem ([Gupta-Kamath-Kayal-S])
A $2^{\Omega(\sqrt{d})}$ lower bound for $\Sigma\Pi^{[\sqrt{d}]}\Sigma\Pi^{[\sqrt{d}]}$ circuits.

Theorem ([Kayal-Limaye-Saha-Srinivasan])
A $n^{\Omega(\sqrt{d})}$ lower bound for homogeneous depth-4 circuits.

Our results

Theorem

An explicit polynomial $f(x_1, \ldots, x_n)$ of degree d with $0/1$ coefficients such that, for any fixed finite field $\mathbb{F}_q$, any homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f must have size $2^{\Omega_q(\sqrt{d})}$.

Our results

Theorem

An explicit polynomial $f(x_1, \ldots, x_n)$ of degree d with $0/1$ coefficients such that, for any fixed finite field $\mathbb{F}_q$, any homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f must have size $2^{\Omega_q(\sqrt{d})}$.

Ingredients for the proof:

[Kayal-Limaye-Saha-Srinivasan] + [Grigoriev-Karpinski]
+ a good amount of sweat

Our results

Theorem

An explicit polynomial $f(x_1, \ldots, x_n)$ of degree d with $0/1$ coefficients such that, for any fixed finite field $\mathbb{F}_q$, any homogeneous $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f must have size $2^{\Omega_q(\sqrt{d})}$.

Ingredients for the proof:

[Kayal-Limaye-Saha-Srinivasan] + [Grigoriev-Karpinski]
+ a good amount of sweat

... ought to have been easier than this

How are such bounds proved?

Natural proof strategies

Construct a map $\Gamma : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{N}$, that assigns a number to every polynomial such that:

1. If f is computable by "small" circuits, then $\Gamma(f)$ is "small".
2. For the desired polynomial f we wish to show a lower bound, then $\Gamma(f)$ is "large".

How are such bounds proved?

Natural proof strategies

Construct a map $\Gamma : \mathbb{F}[x_1, \ldots, x_n] \to \mathbb{N}$, that assigns a number to every polynomial such that:
Typically $\Gamma(f)$ is the rank of some associated linear space.

1. If f is computable by "small" circuits, then $\Gamma(f)$ is "small".

2. For the desired polynomial f we wish to show a lower bound, then $\Gamma(f)$ is "large".

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are just $\binom{d}{k}$ linearly independent k-th order partial derivatives of $\ell_1 \cdots \ell_d$.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are just $\binom{d}{k}$ linearly independent k-th order partial derivatives of $\ell_1 \cdots \ell_d$.
  For a generic polynomial, you would all partial derivatives to be linearly independent.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are just $\binom{d}{k}$ linearly independent k-th order partial derivatives of $\ell_1 \cdots \ell_d$.
  For a generic polynomial, you would all partial derivatives to be linearly independent.

$$\partial^{=k}(\ell_1 \cdots \ell_d) \subseteq \operatorname{span}\left\{ \prod_{i \in S} \ell_i \ : \ S \subseteq [d], |S| = d - k \right\}$$

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are just $\binom{d}{k}$ linearly independent k-th order partial derivatives of $\ell_1 \cdots \ell_d$.
  For a generic polynomial, you would all partial derivatives to be linearly independent.

$$\partial^{=k}(\ell_1 \cdots \ell_d) \subseteq \operatorname{span}\left\{ \prod_{i \in S} \ell_i \ : \ S \subseteq [d], |S| = d - k \right\}$$

  For $f = \mathrm{Det}_d$, the symbolic determinant of a $d \times d$ matrix, we have $\binom{d}{k}^2$ linearly independent $(d-k) \times (d-k)$ minors.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sums of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are just $\binom{d}{k}$ linearly independent k-th order partial derivatives of $\ell_1 \cdots \ell_d$.
  For a generic polynomial, you would all partial derivatives to be linearly independent.

$$\partial^{=k}(\ell_1 \cdots \ell_d) \subseteq \text{span}\left\{\prod_{i \in S} \ell_i \ : \ S \subseteq [d], |S| = d-k\right\}$$

  For $f = \text{Det}_d$, the symbolic determinant of a $d \times d$ matrix, we have $\binom{d}{k}^2$ linearly independent $(d-k) \times (d-k)$ minors.
  Therefore, if $\text{Det}_d = \sum_{i=1}^{s} \ell_{i1} \cdots \ell_{id}$, then $s \geq \binom{d}{d/2}$. $\square$

Examples

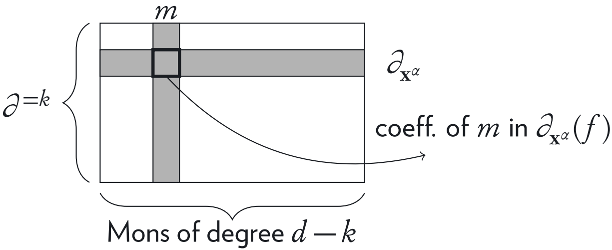
- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.



Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

$$\partial_x(Q_1 \cdots Q_r) \quad = \quad \partial_x(Q_1) \cdot Q_2 \cdots Q_r + \cdots + Q_1 \cdots Q_{r-1} \cdot \partial_x(Q_r)$$

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

$$\partial_x(Q_1 \cdots Q_r) \quad = \quad \partial_x(Q_1) \cdot Q_2 \cdots Q_r + \cdots + Q_1 \cdots Q_{r-1} \cdot \partial_x(Q_r)$$

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

$$\partial_{x}(Q_1 \cdots Q_r) \quad = \quad \mathrm{span}\left\{ \mathbf{x}^{=\sqrt{d}} \cdot \prod_{i \in S} Q_i \; : \; S \subset [r], \, |S| = r-1 \right\}$$

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

$$\partial^{=k}(Q_1 \cdots Q_r) \quad = \quad \operatorname{span}\left\{ \mathbf{x}^{=k\sqrt{d}} \cdot \prod_{i \in S} Q_i \; : \; S \subset [r],\, |S| = r - k \right\}$$

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^d\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

$$\partial^{=k}(Q_1 \cdots Q_r) \quad = \quad \operatorname{span}\left\{ \mathbf{x}^{=k\sqrt{d}} \cdot \prod_{i \in S} Q_i \ : \ S \subset [r], |S| = r - k \right\}$$

Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

$$\Gamma(f) = \dim \left\{ \mathbf{x}^{=\ell} \partial^{=k}(f) \right\}$$

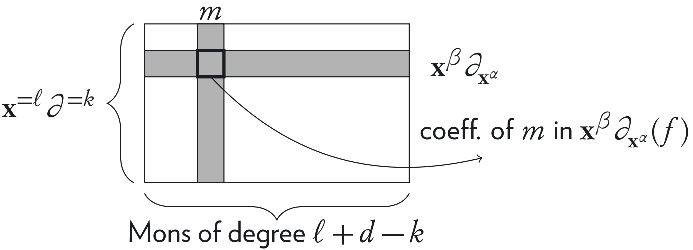
Dimension of shifted partial derivatives

Examples

- [Nisan-Wigderson-95]: $\Sigma\Pi^{d}\Sigma$ circuits, sum of terms of the form $\ell_1 \cdots \ell_d$.

  Key observation: There are "few" linearly independent partial derivatives of $\ell_1 \cdots \ell_d$.

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.



Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.		High degree mons.

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	High degree mons.
✓ [GKKS-12]	

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	High degree large support mons.	High degree small support mons.
✓ [GKKS-12]	Eg. $x_1 \cdots x_d$	Eg. $x_1^{d/2} x_2^{d/2}$

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	High degree large support mons.	High degree small support mons.
✓ [GKKS-12]	Eg. $x_1 \cdots x_d$	Eg. $x_1^{d/2} x_2^{d/2}$

  - Idea 1 - Random restrictions: Randomly set a small number of variables to zero

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	~~High degree large support mons.~~	High degree small support mons.
✓ [GKKS-12]	~~Eg. $x_1 \cdots x_d$~~	Eg. $x_1^{d/2} x_2^{d/2}$

  - Idea 1 - Random restrictions: Randomly set a small number of variables to zero

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	~~High degree large support mons.~~	High degree small support mons.
✓ [GKKS-12]	~~Eg. $x_1 \cdots x_d$~~	Eg. $x_1^{d/2} x_2^{d/2}$

  - Idea 1 - Random restrictions: Randomly set a small number of variables to zero
  - Idea 2 - Multilinear projection: Discard all non-multilinear monomials

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.	~~High degree large support mons.~~	~~High degree small support mons.~~
✓ [GKKS-12]	~~Eg. $x_1 \cdots x_d$~~	~~Eg. $x_1^{d/2} x_2^{d/2}$~~

  - Idea 1 - Random restrictions: Randomly set a small number of variables to zero
  - Idea 2 - Multilinear projection: Discard all non-multilinear monomials

Examples...

- [Gupta-Kamath-Kayal-S-13]: $\Sigma\Pi^{\sqrt{d}}\Sigma\Pi^{\sqrt{d}}$ circuits, terms of the form $Q_1 \cdots Q_{\sqrt{d}}$.

  Key observation: Many *low-degree* combinations of partial derivatives are zero if all Q_is have low degree.

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]: hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

Low degree mons.
✓
[GKKS-12]

~~High degree large support mons.~~
~~Eg. $x_1 \cdots x_d$~~

~~High degree small support mons.~~
~~Eg. $x_1^{d/2} x_2^{d/2}$~~

- Idea 1 - Random restrictions: Randomly set a small number of variables to zero
- Idea 2 - Multilinear projection: Discard all non-multilinear monomials

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]:
hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

  - Idea 1 - Random restrictions: Randomly set a small number of variables to zero
  - Idea 2 - Multilinear projection: Discard all non-multilinear monomials

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]:
  hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.
  - IDEA 1 - RANDOM RESTRICTIONS: Randomly set a small number of variables to zero
  - IDEA 2 - MULTILINEAR PROJECTION: Discard all non-multilinear monomials

$$\Gamma(f) \quad = \quad \dim(\mathbf{x}^{=\ell}\partial^{=k}(f))$$

Dimension of shifted partials of f.

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]:
  hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

  - IDEA 1 - RANDOM RESTRICTIONS: Randomly set a small number of variables to zero
  - IDEA 2 - MULTILINEAR PROJECTION: Discard all non-multilinear monomials

$$\Gamma(f) \quad = \quad \dim(\mathbf{x}^{=\ell}\partial^{=k}(\rho(f)))$$

Dimension of shifted partials of a random restriction of f.

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]: hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.
  - IDEA 1 - RANDOM RESTRICTIONS: Randomly set a small number of variables to zero
  - IDEA 2 - MULTILINEAR PROJECTION: Discard all non-multilinear monomials

$$\Gamma(f) \quad = \quad \dim(\mathrm{mult} \circ \mathbf{x}^{=\ell} \partial^{=k}(\rho(f)))$$

Dimension of projected shifted partials of a random restriction of f.

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]:
  hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

$$\Gamma(f) \quad = \quad \dim(\mathrm{mult} \circ \mathbf{x}^{=\ell} \partial^{=k}(\rho(f)))$$

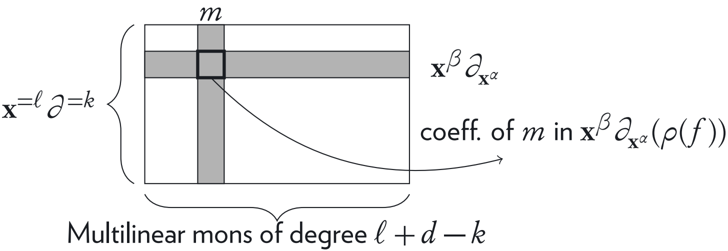
Dimension of projected shifted partials of a random restriction of f.

Examples...

- [Kayal-Limaye-Saha-Srinivasan-13], [Kumar-Saraf-13]:
  hom. $\Sigma\Pi\Sigma\Pi$ circuits: terms like $Q_1 \cdots Q_b$ with total degree d.

$$\Gamma(f) \quad = \quad \dim(\mathrm{mult} \circ \mathbf{x}^{=\ell} \partial^{=k}(\rho(f)))$$

Dimension of projected shifted partials of a random restriction of f.



Handling depth-5 circuits

Handling depth-5 circuits

We already have a complexity measure Γ_{PSPD} for hom. depth-4 circuits.

Handling depth-5 circuits

We already have a complexity measure Γ_{PSPD} for hom. depth-4 circuits.

How large is Γ_{PSPD} for a *generic depth*-5 *circuit*?

Handling depth-5 circuits

We already have a complexity measure Γ_{PSPD} for hom. depth-4 circuits.

How large is Γ_{PSPD} for a *generic depth*-5 *circuit*?

Small $\Rightarrow$ a lower bound against depth-5 circuits.

Handling depth-5 circuits

We already have a complexity measure Γ_{PSPD} for hom. depth-4 circuits.

How large is Γ_{PSPD} for a *generic depth*-5 *circuit*?

Small $\Rightarrow$ a lower bound against depth-5 circuits.

Large $\Rightarrow$ separation between depth-5 and depth-4 circuits.

Handling depth-5 circuits

We already have a complexity measure Γ_{PSPD} for hom. depth-4 circuits.

How large is Γ_{PSPD} for a *generic depth-5 circuit*?

Small $\Rightarrow$ a lower bound against depth-5 circuits.

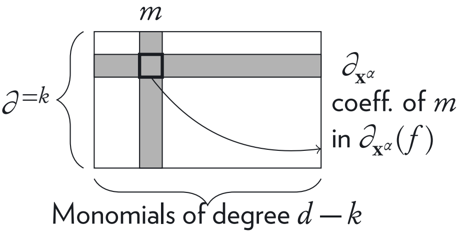
Large $\Rightarrow$ separation between depth-5 and depth-4 circuits.

... still don't know

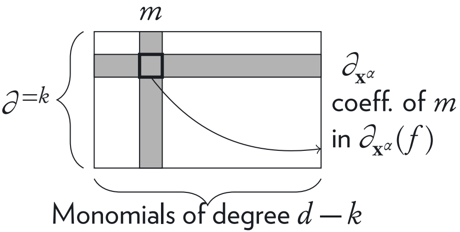
Evaluating the complexity measure

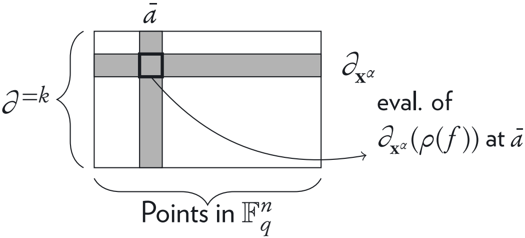
$$\Gamma_k(f) = \dim \left\{ \partial^{=k}(f) \right\}$$

Evaluating the complexity measure

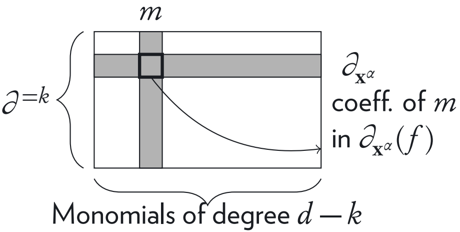


Evaluating the complexity measure



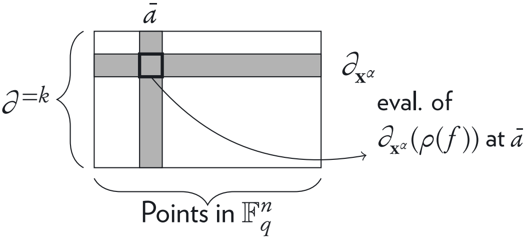


Evaluating the complexity measure



Small rank

$\Downarrow$



Small rank

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.

High degree terms.

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.	High degree high rank terms	High degree low rank terms

Eg. $\ell_1^{d/3} \ell_2^{d/3} (\ell_1 + 3\ell_2)^{d/3}$

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.	High degree high rank terms	High degree low rank terms

✓

[NW-95]

Eg. $\ell_1^{d/3} \ell_2^{d/3} (\ell_1 + 3\ell_2)^{d/3}$

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.	High degree high rank terms	High degree low rank terms

✓ [NW-95]

Eg. $\ell_1^{d/3} \ell_2^{d/3} (\ell_1 + 3\ell_2)^{d/3}$

✓ [NW-95]

[Grigoriev-Karpinski]

$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.	High degree high rank terms	High degree low rank terms
✓ [NW-95]		Eg. $\ell_1^{d/3} \ell_2^{d/3} (\ell_1 + 3\ell_2)^{d/3}$ ✓ [NW-95]

Observation

If $\dim \{\ell_1, \cdots, \ell_r\}$ is large, then almost all evaluations of it on $\mathbb{F}_q^n$ are zero.

[Grigoriev-Karpinski]

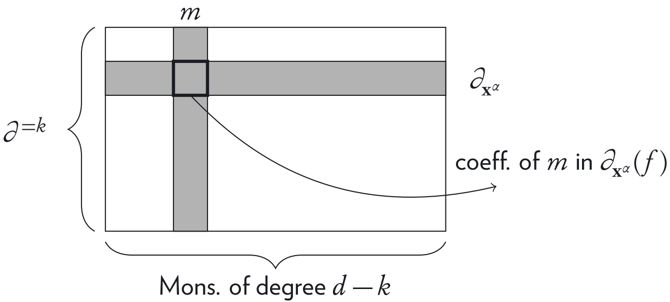
$$f = \ell_{11} \cdots \ell_{1d_1} + \cdots + \ell_{s1} \cdots \ell_{sd_s}$$

Low degree terms.	~~High degree high rank terms~~	High degree low rank terms
✓ [NW-95]		Eg. $\ell_1^{d/3} \ell_2^{d/3} (\ell_1 + 3\ell_2)^{d/3}$ ✓ [NW-95]

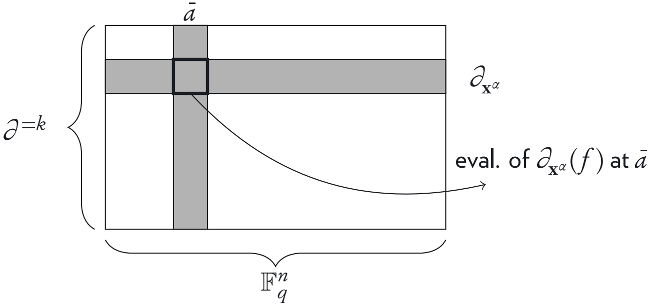
Observation

If $\dim\{\ell_1, \cdots, \ell_r\}$ is large, then almost all evaluations of it on $\mathbb{F}_q^n$ are zero.

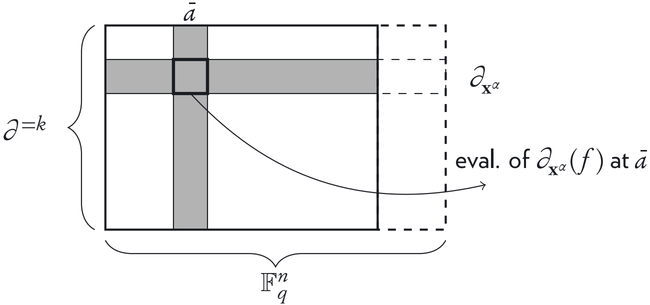
[Grigoriev-Karpinski]



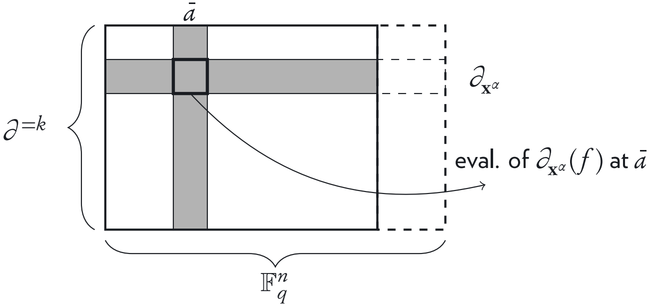
[Grigoriev-Karpinski]



[Grigoriev-Karpinski]



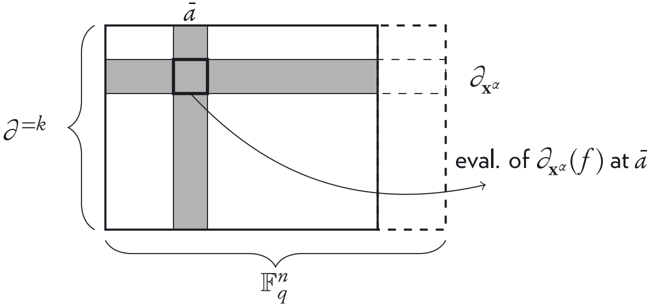
[Grigoriev-Karpinski]



Lemma

If f is computable by a small $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_q$, then there the above matrix has small rank when a certain small set of columns are removed.

[Grigoriev-Karpinski]



Lemma

If f is computable by a small $\Sigma\Pi\Sigma$ circuit over $\mathbb{F}_q$, then there the above matrix has small rank when a certain small set of columns are removed.

Lemma

For Det_n or Perm_n the above matrix remains full rank, as long as we removed only few columns.

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

Low degree products.	High degree products.

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

Low degree products.	High degree products.
✓ [GKKS]	

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

Low degree products.	High degree, large rank products.	High degree, small rank products.
✓ [GKKS]	Eg. $\ell_1 \cdots \ell_d$	Eg. $\ell_1^{d/2} \ell_2^{d/2}$

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

Low degree products.	High degree, large rank products.	High degree, small rank products.
$\checkmark$ [GKKS]	Eg. $\ell_1 \cdots \ell_d$	Eg. $\ell_1^{d/2} \ell_2^{d/2}$ $\checkmark$

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

Low degree products.	High degree, large rank products.	High degree, small rank products.
$\checkmark$ [GKKS]	Eg. $\ell_1 \cdots \ell_d$	Eg. $\ell_1^{d/2} \ell_2^{d/2}$ $\checkmark$

Observation

If $\dim\{\ell_1, \cdots, \ell_r\}$ is large, then almost all evaluations of it on $\mathbb{F}_q^n$ are zero.

Lifting to depth five

$\Sigma\Pi\Sigma\Pi\Sigma$

Types of products of linear polynomials:

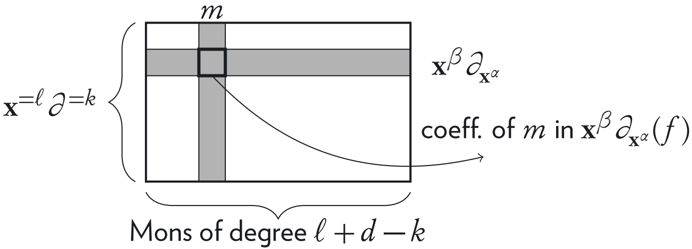
Low degree products.	~~High degree, large rank products.~~	High degree, small rank products.
$\checkmark$ [GKKS]	~~Eg. $\ell_1 \cdots \ell_d$~~	Eg. $\ell_1^{d/2}\ell_2^{d/2}$ $\checkmark$

Observation

If $\dim\{\ell_1, \cdots, \ell_r\}$ is large, then almost all evaluations of it on $\mathbb{F}_q^n$ are zero.

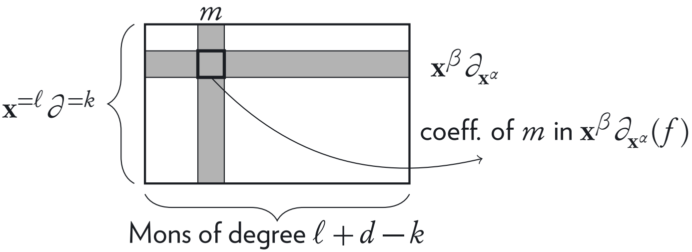
Rank of the eval. version of PSPD

WE KNOW THIS RANK IS LARGE:

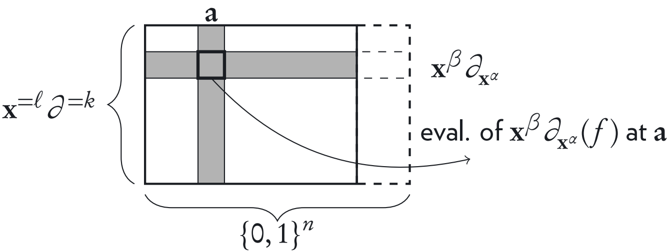


Rank of the eval. version of PSPD

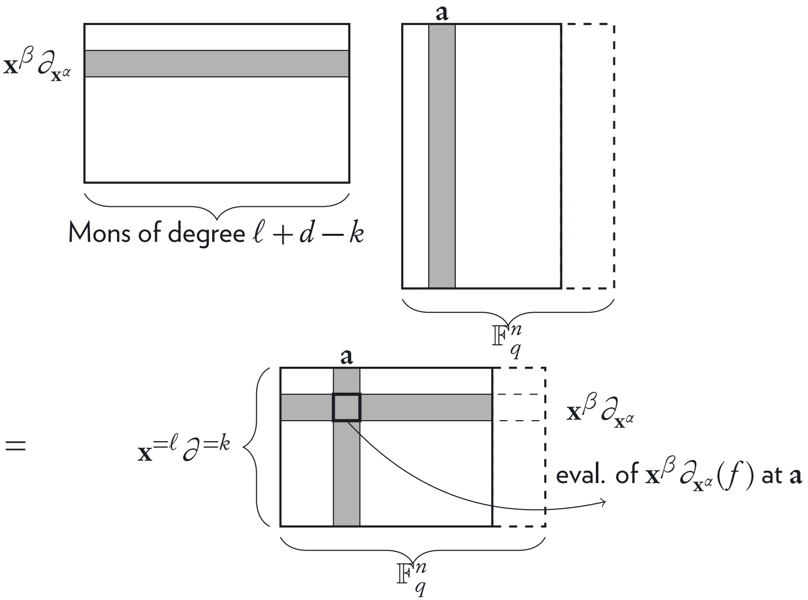
WE KNOW THIS RANK IS LARGE:



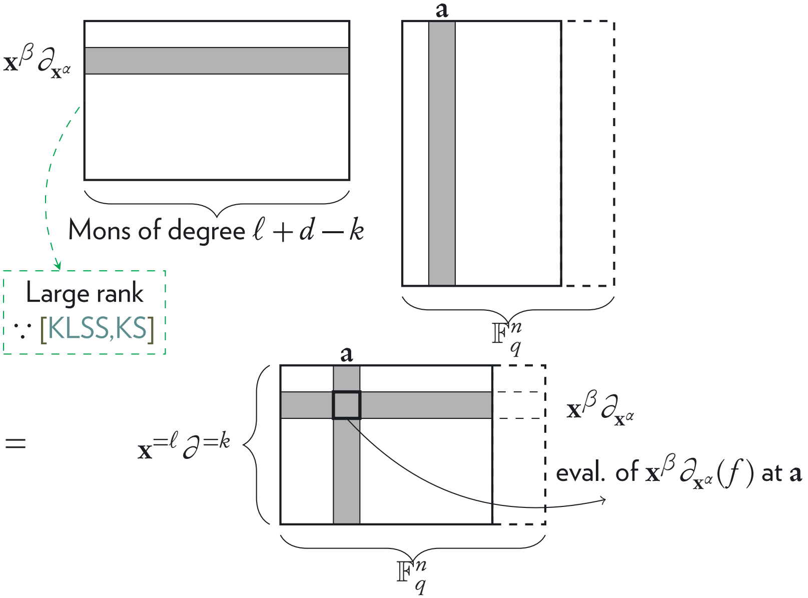
NEED TO SHOW *this* RANK IS LARGE:



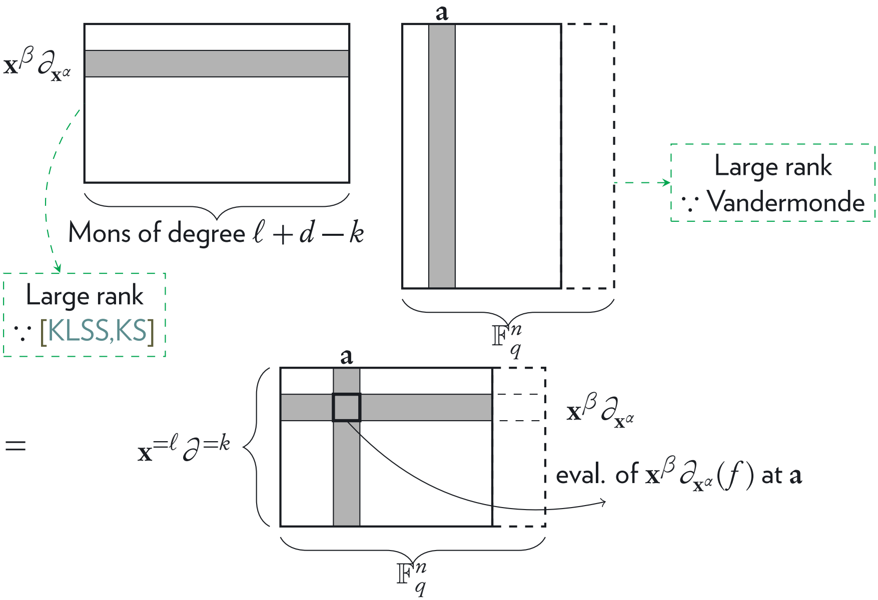
Switching to the evaluation perspective



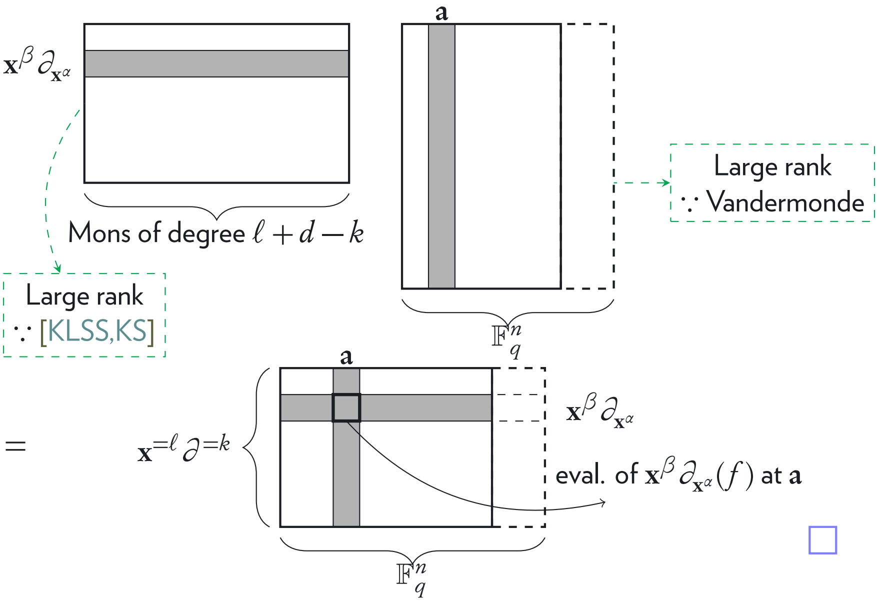
Switching to the evaluation perspective



Switching to the evaluation perspective



Switching to the evaluation perspective



Issues to be resolved

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0,1\}^n$.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0, 1\}^n$.

Issue 2: But then $(x_1 + 1) \cdots (x_n + 1)$, over $\mathbb{F}_3$, is *never* zero over $\{0, 1\}^n$.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0,1\}^n$.

Issue 2: But then $(x_1+1)\cdots(x_n+1)$, over $\mathbb{F}_3$, is *never* zero over $\{0,1\}^n$.

Fix: Ok fine. Work with $\bar{c}+\{0,1\}^n$ for some random $\bar{c}\in\mathbb{F}_q^n$.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0,1\}^n$.

Issue 2: But then $(x_1+1)\cdots(x_n+1)$, over $\mathbb{F}_3$, is *never* zero over $\{0,1\}^n$.

Fix: Ok fine. Work with $\bar{c}+\{0,1\}^n$ for some random $\bar{c}\in\mathbb{F}_q^n$.

Issue 3: Even with $\bar{c}+\{0,1\}^n$ the matrix is still slightly fat and the Vandermonde is slightly tall.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0,1\}^n$.

Issue 2: But then $(x_1+1)\cdots(x_n+1)$, over $\mathbb{F}_3$, is *never* zero over $\{0,1\}^n$.

Fix: Ok fine. Work with $\bar{c}+\{0,1\}^n$ for some random $\bar{c}\in\mathbb{F}_q^n$.

Issue 3: Even with $\bar{c}+\{0,1\}^n$ the matrix is still slightly fat and the Vandermonde is slightly tall.

Fix: Prove a *really* good rank lower bound on the left matrix.

Issues to be resolved

Issue 1: [Fat matrix] $\times$ [Tall matrix] could be zero, even if both are full rank.

Fix: Make the matrix slimmer by only considering evaluations on $\{0,1\}^n$.

Issue 2: But then $(x_1+1)\cdots(x_n+1)$, over $\mathbb{F}_3$, is *never* zero over $\{0,1\}^n$.

Fix: Ok fine. Work with $\bar{c}+\{0,1\}^n$ for some random $\bar{c}\in\mathbb{F}_q^n$.

Issue 3: Even with $\bar{c}+\{0,1\}^n$ the matrix is still slightly fat and the Vandermonde is slightly tall.

Fix: Prove a *really* good rank lower bound on the left matrix.
(Barely manages to work for a specific explicit polynomial. Phew!) $\square$

Summary

Theorem

There is a polynomial $f \in$ *VNP such that, for every finite field* $\mathbb{F}_q$*, any hom.* $\Sigma\Pi\Sigma\Pi\Sigma$ *circuit computing* f *over* $\mathbb{F}_q$ *must have size* $\exp(\Omega_q(\sqrt{d}))$.

Summary

Theorem

There is a polynomial $f \in$ VNP such that, for every finite field $\mathbb{F}_q$, any hom. $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f over $\mathbb{F}_q$ must have size $\exp(\Omega_q(\sqrt{d}))$.

REMARKS AND OPEN PROBLEMS:

- [Grigoriev-Karpinski] meets [Kayal-Limaye-Saha-Srinivasan]. Delicate analysis.

Summary

Theorem

There is a polynomial $f \in$ VNP such that, for every finite field $\mathbb{F}_q$, any hom. $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f over $\mathbb{F}_q$ must have size $\exp(\Omega_q(\sqrt{d}))$.

REMARKS AND OPEN PROBLEMS:

- [Grigoriev-Karpinski] meets [Kayal-Limaye-Saha-Srinivasan]. Delicate analysis.
- The proof ought to work for IMM also but we don't have a tight enough analysis (yet).

Summary

Theorem

There is a polynomial $f \in VNP$ such that, for every finite field $\mathbb{F}_q$, any hom. $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f over $\mathbb{F}_q$ must have size $\exp(\Omega_q(\sqrt{d}))$.

REMARKS AND OPEN PROBLEMS:

- [Grigoriev-Karpinski] meets [Kayal-Limaye-Saha-Srinivasan]. Delicate analysis.
- The proof ought to work for IMM also but we don't have a tight enough analysis (yet).
- After this, [Kumar-S] *did* manage to separate depth-4 and depth-5 in the low-degree regime, but via a different complexity measure.

Summary

Theorem

There is a polynomial $f \in$ VNP such that, for every finite field $\mathbb{F}_q$, any hom. $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f over $\mathbb{F}_q$ must have size $\exp(\Omega_q(\sqrt{d}))$.

REMARKS AND OPEN PROBLEMS:

- [Grigoriev-Karpinski] meets [Kayal-Limaye-Saha-Srinivasan]. Delicate analysis.
- The proof ought to work for IMM also but we don't have a tight enough analysis (yet).
- After this, [Kumar-S] *did* manage to separate depth-4 and depth-5 in the low-degree regime, but via a different complexity measure.
- Other fields?

Summary

Theorem

There is a polynomial $f \in$ VNP such that, for every finite field $\mathbb{F}_q$, any hom. $\Sigma\Pi\Sigma\Pi\Sigma$ circuit computing f over $\mathbb{F}_q$ must have size $\exp(\Omega_q(\sqrt{d}))$.

REMARKS AND OPEN PROBLEMS:

- [Grigoriev-Karpinski] meets [Kayal-Limaye-Saha-Srinivasan]. Delicate analysis.
- The proof ought to work for IMM also but we don't have a tight enough analysis (yet).
- After this, [Kumar-S] *did* manage to separate depth-4 and depth-5 in the low-degree regime, but via a different complexity measure.
- Other fields?

\end{document}

References

- [Agrawal-Vinay]:
  "Arithmetic Circuits: A Chasm at Depth Four"
  Foundations of Computer Science, 2008
- [Koiran]:
  "Arithmetic circuits: The chasm at depth four gets wider"
  Theoretical Computer Science, 2012
- [Tavenas]:
  "Improved bounds for reduction to depth 4 and depth 3"
  Information and Computation, 2015
- [Nisan-Wigderson]:
  "Lower Bounds on Arithmetic Circuits Via Partial Derivatives"
  Computational Complexity, 1997

References

- [Grigoriev-Karpinski]:
  "An Exponential Lower Bound for Depth 3 Arithmetic Circuits"
  Symposium on Theory of Computing, 1998

- [Gupta-Kamath-Kayal-S]:
  "Approaching the Chasm at Depth Four"
  Journal of the ACM, 2014

- [Kayal-Limaye-Saha-Srinivasan]:
  "An Exponential Lower Bound for Homogeneous Depth Four Arithmetic Formulas"
  SIAM Journal of Computing, 2017

- [Kumar-Saraf]:
  "On the Power of Homogeneous Depth 4 Arithmetic Circuits"
  SIAM Journal of Computing, 2017

References

- [Grigoriev-Razborov]:
  "Exponential Lower Bounds for Depth 3 Arithmetic Circuits in Algebras of Functions over Finite Fields"
  Appl. Algebra Eng. Commun. Comput. , 2000

- [Kumar-S]:
  "Finer Separations Between Shallow Arithmetic Circuits"
  Foundations of Software Technology and Theoretical Computer Science, 2016