Exponential-gap fanin-hierarchy for border depth-3 circuits

Joint work with Nitin Saxena (IIT Kanpur).

Pranjal Dutta
Google Ph.D. Fellow, CMI & Visiting Research Scholar, IIT Kanpur

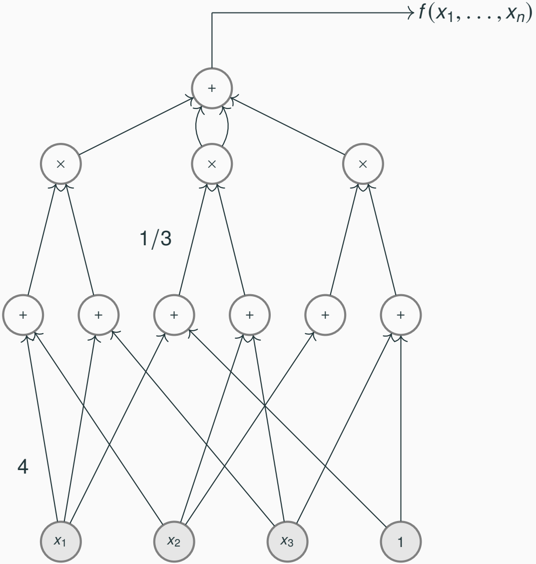
27th January, 2022

GCT 2022 (Online)

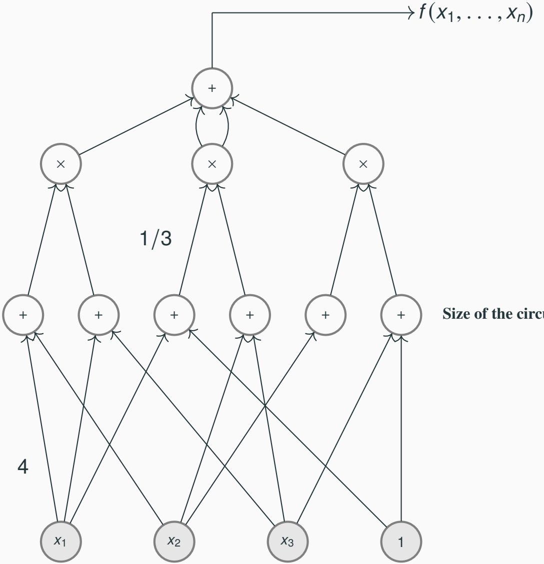
Table of Contents

Basic definitions and terminologies

Algebraic circuits

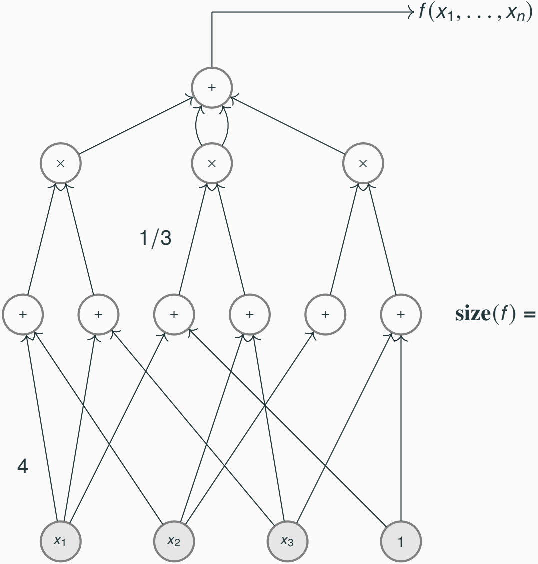


Algebraic circuits



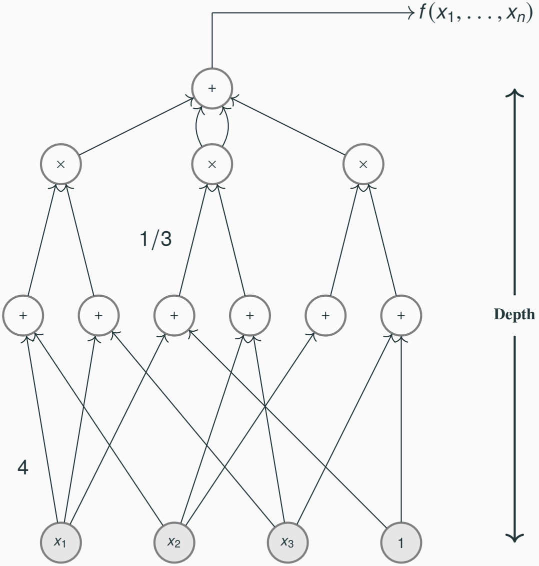
Size of the circuit = number of nodes + edges

Algebraic circuits



size(f) = min size of the circuit computing f

Algebraic circuits



The determinant polynomial

The determinant polynomial

- Let $X_n = [x_{i,j}]_{1 \leq i,j \leq n}$ be a $n \times n$ matrix of distinct variables $x_{i,j}$. Let $S_n := \{\pi \mid \pi : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ such that π is bijective $\}$. Define

$$f_n := \det(X_n) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \prod_{i=1}^{n} x_{i,\pi(i)} .$$

The determinant polynomial

- Let $X_n = [x_{i,j}]_{1 \le i,j \le n}$ be a $n \times n$ matrix of distinct variables $x_{i,j}$. Let $S_n := \{\pi \mid \pi : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ such that π is bijective $\}$. Define

$$f_n := \det(X_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} x_{i,\pi(i)} .$$

- $\det$ is *universal*, i.e. any polynomial $f(\boldsymbol{x})$ can be computed as a determinant of a square matrix whose entries are polynomials of degree ≤ 1.

The determinant polynomial

- Let $X_n = [x_{i,j}]_{1 \leq i,j \leq n}$ be a $n \times n$ matrix of distinct variables $x_{i,j}$. Let $S_n := \{\pi \mid \pi : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ such that π is bijective $\}$. Define

$$f_n := \det(X_n) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \prod_{i=1}^{n} x_{i,\pi(i)} .$$

- $\det$ is *universal*, i.e. any polynomial $f(\boldsymbol{x})$ can be computed as a determinant of a square matrix whose entries are polynomials of degree ≤ 1.
- The minimum dimension of the matrix to compute f, is called the **determinantal complexity** $\mathrm{dc}(f)$.

The determinant polynomial

- Let $X_n = [x_{i,j}]_{1 \le i,j \le n}$ be a $n \times n$ matrix of distinct variables $x_{i,j}$. Let $S_n := \{\pi \mid \pi : \{1,\dots,n\} \longrightarrow \{1,\dots,n\}$ such that π is bijective $\}$. Define

$$f_n := \det(X_n) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \prod_{i=1}^{n} x_{i,\pi(i)} .$$

- $\det$ is *universal*, i.e. any polynomial $f(\boldsymbol{x})$ can be computed as a determinant of a square matrix whose entries are polynomials of degree ≤ 1.
- The minimum dimension of the matrix to compute f, is called the **determinantal complexity** $\mathrm{dc}(f)$.
- E.g. $\mathrm{dc}(x_1 \cdots x_n) = n$, since

$$x_1 \cdots x_n = \det \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & x_n \end{pmatrix} .$$

The determinant polynomial

- Let $X_n = [x_{i,j}]_{1 \le i,j \le n}$ be a $n \times n$ matrix of distinct variables $x_{i,j}$. Let $S_n := \{\pi \mid \pi : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ such that π is bijective $\}$. Define

$$f_n := \det(X_n) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) \prod_{i=1}^{n} x_{i,\pi(i)} .$$

- $\det$ is *universal*, i.e. any polynomial $f(\boldsymbol{x})$ can be computed as a determinant of a square matrix whose entries are polynomials of degree ≤ 1.
- The minimum dimension of the matrix to compute f, is called the **determinantal complexity** $\mathrm{dc}(f)$.
- E.g. $\mathrm{dc}(x_1 \cdots x_n) = n$, since

$$x_1 \cdots x_n = \det \begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & x_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & x_n \end{pmatrix} .$$

- VBP: The class VBP is defined as the set of all sequences of polynomials $(f_n)_n$ with polynomially bounded $\mathrm{dc}(f_n)$.

Border complexity

Border complexity

- Let Γ be any sensible measure. It can be size, dc and so on.

Border complexity

- Let Γ be any sensible measure. It can be size, dc and so on.
- For any Γ, we can define the border complexity measure $\overline{\Gamma}$ via:
  $\overline{\Gamma}(h)$ is the *smallest* n such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\epsilon(\boldsymbol{x})$ with $\Gamma(h_\epsilon) \leq n$.

Border complexity

- Let Γ be any sensible measure. It can be size, dc and so on.
- For any Γ, we can define the border complexity measure $\overline{\Gamma}$ via: $\overline{\Gamma}(h)$ is the *smallest* n such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\epsilon(\boldsymbol{x})$ with $\Gamma(h_\epsilon) \leq n$. In other words,

$$\lim_{\epsilon \to 0} h_\epsilon = h \text{ (coefficient-wise)} .$$

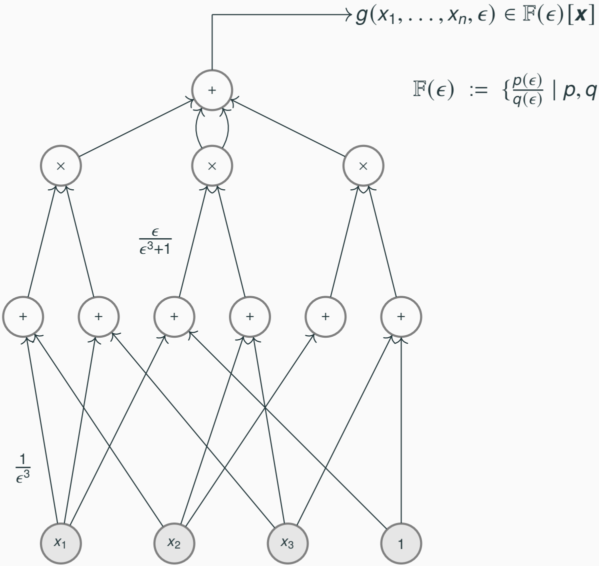
Border complexity

- Let Γ be any sensible measure. It can be size, dc and so on.
- For any Γ, we can define the border complexity measure $\overline{\Gamma}$ via:
  $\overline{\Gamma}(h)$ is the *smallest* n such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\epsilon(\boldsymbol{x})$ with $\Gamma(h_\epsilon) \leq n$. In other words,

$$\lim_{\epsilon \to 0} h_\epsilon = h \text{ (coefficient-wise)} .$$

- We will work with *'approximative circuits'*.

Approximative circuits



$$\mathbb{F}(\epsilon) := \{\tfrac{p(\epsilon)}{q(\epsilon)} \mid p, q \in \mathbb{F}[\epsilon], q(\epsilon) \neq 0\}$$

Algebraic approximation

Algebraic approximation

- Suppose, we assume the following:
  - $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}[x_1, \ldots, x_n, \epsilon]$, i.e. it is a polynomial of the form

$$g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i(x_1, \ldots, x_n) \cdot \epsilon^i ,$$

Algebraic approximation

- Suppose, we assume the following:
  - $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}[x_1, \ldots, x_n, \epsilon]$, i.e. it is a polynomial of the form

$$g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i(x_1, \ldots, x_n) \cdot \epsilon^i ,$$

  - Can we say anything about the complexity of g_0?

Algebraic approximation

- Suppose, we assume the following:
  - $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}[x_1, \ldots, x_n, \epsilon]$, i.e. it is a polynomial of the form

$$g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i(x_1, \ldots, x_n) \cdot \epsilon^i ,$$

  - Can we say anything about the complexity of g_0?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\epsilon = 0$?!

Algebraic approximation

- Suppose, we assume the following:
  - $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}[x_1, \ldots, x_n, \epsilon]$, i.e. it is a polynomial of the form

$$g(\boldsymbol{x}, \epsilon) \;=\; \sum_{i=0}^{M} g_i(x_1, \ldots, x_n) \cdot \epsilon^i \;,$$

  - Can we say anything about the complexity of g_0?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\epsilon = 0$?! Setting $\epsilon = 0$ *may not* be 'legal' as it could be using $1/\epsilon$ in the wire. Though it is well-defined!

Algebraic approximation

- Suppose, we assume the following:
  - $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}[x_1, \ldots, x_n, \epsilon]$, i.e. it is a polynomial of the form

$$g(\boldsymbol{x}, \epsilon) \ = \ \sum_{i=0}^{M} g_i(x_1, \ldots, x_n) \cdot \epsilon^i \ ,$$

  - Can we say anything about the complexity of g_0?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\epsilon = 0$?! Setting $\epsilon = 0$ *may not* be 'legal' as it could be using $1/\epsilon$ in the wire. Though it is well-defined!
- Summary: g_0 is really something **non-trivial** and being 'approximated' by the circuit since $\lim_{\epsilon \to 0} g(\boldsymbol{x}, \epsilon) = g_0$.

Algebraic approximation

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$.

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\mathbf{x}]$ approximative complexity s, if there is a $g(\mathbf{x}, \epsilon) \in \mathbb{F}(\epsilon)[\mathbf{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\mathbf{x}, \epsilon) \in \mathbb{F}[\epsilon][\mathbf{x}]$ such that $g(\mathbf{x}, \epsilon) = h(\mathbf{x}) + \epsilon \cdot S(\mathbf{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$
- If g has circuit of size s over $\mathbb{F}(\epsilon)$, then one can assume that the highest degree of ϵ in g can be *exponentially large* 2^{s^2} [Bürgisser 2004, 2020].

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$
- If g has circuit of size s over $\mathbb{F}(\epsilon)$, then one can assume that the highest degree of ϵ in g can be *exponentially large* 2^{s^2} [Bürgisser 2004, 2020].
- Let us assume that $g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i \epsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$
- If g has circuit of size s over $\mathbb{F}(\epsilon)$, then one can assume that the highest degree of ϵ in g can be *exponentially large* 2^{s^2} [Bürgisser 2004, 2020].
- Let us assume that $g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i \epsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - Pick $M + 1$ many distinct values from $\mathbb{F}$ *randomly* and interpolate;

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$
- If g has circuit of size s over $\mathbb{F}(\epsilon)$, then one can assume that the highest degree of ϵ in g can be *exponentially large* 2^{s^2} [Bürgisser 2004, 2020].
- Let us assume that $g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i \epsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - Pick $M + 1$ many distinct values from $\mathbb{F}$ *randomly* and interpolate;
  - $\text{size}(h) \leq \exp(\overline{\text{size}}(h))$.

Algebraic approximation

Algebraic Approximation [Bürgisser 2004]

A polynomial $h(x) \in \mathbb{F}[\boldsymbol{x}]$ approximative complexity s, if there is a $g(\boldsymbol{x}, \epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, of size s, over $\mathbb{F}(\epsilon)$, and a polynomial $S(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\epsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$. In other words, $\lim_{\epsilon \to 0} g = h$.

- $\overline{\text{size}}(h) \leq \text{size}(h)$. $[h = h + \epsilon \cdot 0.]$
- If g has circuit of size s over $\mathbb{F}(\epsilon)$, then one can assume that the highest degree of ϵ in g can be *exponentially large* 2^{s^2} [Bürgisser 2004, 2020].
- Let us assume that $g(\boldsymbol{x}, \epsilon) = \sum_{i=0}^{M} g_i \epsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - Pick $M + 1$ many distinct values from $\mathbb{F}$ *randomly* and interpolate;
  - $\text{size}(h) \leq \exp(\overline{\text{size}}(h))$.
- $\overline{\text{size}}(h) \leq \text{size}(h) \leq \exp(\overline{\text{size}}(h))$.

Border Waring Rank

Border Waring Rank

Border Waring rank

The border Waring rank $\overline{\mathsf{WR}}(h)$, of a d-form h is defined as the smallest s such that $h = \lim_{\epsilon\to 0} \sum_{i\in[s]} \ell_i^d$, where $\ell_i \in \mathbb{F}(\epsilon)[\mathbf{x}]$, are homogeneous linear forms.

Border Waring Rank

Border Waring rank

The border Waring rank $\overline{\mathsf{WR}}(h)$, of a d-form h is defined as the smallest s such that $h = \lim_{\epsilon \to 0} \sum_{i \in [s]} \ell_i^d$, where $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, are homogeneous linear forms.

- When ℓ_i are non-homogeneous, we will write this as $h \in \overline{\Sigma^{[s]} \wedge \Sigma}$.

Border Waring rank

The border Waring rank $\overline{\mathsf{WR}}(h)$, of a d-form h is defined as the smallest s such that $h = \lim_{\epsilon \to 0} \sum_{i \in [s]} \ell_i^d$, where $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, are homogeneous linear forms.

- When ℓ_i are non-homogeneous, we will write this as $h \in \overline{\Sigma^{[s]} \wedge \Sigma}$.
- $\overline{\mathsf{WR}}(x^{d-1}y) = 2$, since, $d \cdot x^{d-1}y + O(\epsilon) = (x + \epsilon y)^d - x^d$.

Lower bounds for border depth-2 circuits

- A few known upper bound/lower bound results on depth-2:

Lower bounds for border depth-2 circuits

- A few known upper bound/lower bound results on depth-2:
  - $\overline{\Sigma^{[s]}\Pi} = \Sigma^{[s]}\Pi$ and $\overline{\Pi\Sigma} = \Pi\Sigma$.

Lower bounds for border depth-2 circuits

- A few known upper bound/lower bound results on depth-2:
  - $\overline{\Sigma^{[s]}\Pi} = \Sigma^{[s]}\Pi$ and $\overline{\Pi\Sigma} = \Pi\Sigma$.
  - det_n is irreducible and $\mathsf{exp}(n)$-sparse. So, det_n *requires* exponential-size border depth-2 circuits!

Lower bounds for border depth-2 circuits

- A few known upper bound/lower bound results on depth-2:
  - $\overline{\Sigma^{[s]}\Pi} = \Sigma^{[s]}\Pi$ and $\overline{\Pi\Sigma} = \Pi\Sigma$.
  - det_n is irreducible and $\mathsf{exp}(n)$-sparse. So, det_n *requires* exponential-size border depth-2 circuits!
  - What about border depth-3 circuits?

Border depth-3 circuits

Depth-3 circuits

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.
- They compute polynomials (*not necessarily* homogeneous) of the form $\sum_{i=1}^{k} \prod_{j=1}^{d_i} \ell_{ij}$, where ℓ_{ij} are linear polynomials (i.e. $a_0 + a_1 x_1 + \ldots + a_n x_n$, for $a_i \in \mathbb{F}$).

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.
- They compute polynomials (*not necessarily* homogeneous) of the form $\sum_{i=1}^{k}\prod_{j=1}^{d_i}\ell_{ij}$, where ℓ_{ij} are linear polynomials (i.e. $a_0 + a_1x_1 + \ldots + a_nx_n$, for $a_i \in \mathbb{F}$).
- Product fan-in $= \max d_i$.

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.
- They compute polynomials (*not necessarily* homogeneous) of the form $\sum_{i=1}^{k}\prod_{j=1}^{d_i}\ell_{ij}$, where ℓ_{ij} are linear polynomials (i.e. $a_0 + a_1x_1 + \ldots + a_nx_n$, for $a_i \in \mathbb{F}$).
- Product fan-in $= \max d_i$.
- How powerful are $\Sigma^{[k]}\Pi\Sigma$ circuits, for constant k? Are they *universal*?

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.
- They compute polynomials (*not necessarily* homogeneous) of the form $\sum_{i=1}^{k}\prod_{j=1}^{d_i}\ell_{ij}$, where ℓ_{ij} are linear polynomials (i.e. $a_0 + a_1x_1 + \ldots + a_nx_n$, for $a_i \in \mathbb{F}$).
- Product fan-in $= \max d_i$.
- How powerful are $\Sigma^{[k]}\Pi\Sigma$ circuits, for constant k? Are they *universal*?
- Impossibility result: The *Inner Product* polynomial $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = x_1y_1 + \ldots + x_{k+1}y_{k+1}$ **cannot** be written as a $\Sigma^{[k]}\Pi\Sigma$ circuit, *regardless* of the product fan-in (even allowing $\exp(n)$ product fan-in!).

Depth-3 circuits

- Depth-3 circuits with top fan-in k, are denoted as $\Sigma^{[k]}\Pi\Sigma$.
- They compute polynomials (*not necessarily* homogeneous) of the form $\sum_{i=1}^{k} \prod_{j=1}^{d_i} \ell_{ij}$, where ℓ_{ij} are linear polynomials (i.e. $a_0 + a_1 x_1 + \ldots + a_n x_n$, for $a_i \in \mathbb{F}$).
- Product fan-in $= \max d_i$.
- How powerful are $\Sigma^{[k]}\Pi\Sigma$ circuits, for constant k? Are they *universal*?
- Impossibility result: The *Inner Product* polynomial $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + \ldots + x_{k+1} y_{k+1}$ **cannot** be written as a $\Sigma^{[k]}\Pi\Sigma$ circuit, *regardless* of the product fan-in (even allowing $\exp(n)$ product fan-in!).
- The same holds if we replace by $\det_n$.

Lower bound for Constant fanin depth-3 circuits

- det_n *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

Lower bound for Constant fanin depth-3 circuits

- det_n *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

Lower bound for Constant fanin depth-3 circuits

- $\mathsf{det_n}$ *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.

Lower bound for Constant fanin depth-3 circuits

- det_n *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.
  - LHS $\bmod \mathcal{I}_1$, becomes $\sum_{i\in[k-1]} T_i'$ while $\mathsf{det}_n \bmod \mathcal{I}_1 \neq 0$.

Lower bound for Constant fanin depth-3 circuits

- $\mathsf{det_n}$ *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.
  - LHS $\mod \mathcal{I}_1$, becomes $\sum_{i\in[k-1]} T_i'$ while $\mathsf{det}_n \mod \mathcal{I}_1 \neq 0$.
  - Keep repeating and each time include an independent ℓ_i in the ideal $\mathcal{I}_i$.

Lower bound for Constant fanin depth-3 circuits

- $\mathsf{det_n}$ *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.
  - LHS $\mod \mathcal{I}_1$, becomes $\sum_{i\in[k-1]} T_i'$ while $\mathsf{det}_n \mod \mathcal{I}_1 \neq 0$.
  - Keep repeating and each time include an independent ℓ_i in the ideal $\mathcal{I}_i$.
  - At the end, there is some $\mathcal{I}_k$ such that RHS has become 0 while LHS is *non-zero* (because $k << n$).

Lower bound for Constant fanin depth-3 circuits

- det_n *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.
  - LHS $\mod \mathcal{I}_1$, becomes $\sum_{i\in[k-1]} T_i'$ while $\mathsf{det}_n \mod \mathcal{I}_1 \neq 0$.
  - Keep repeating and each time include an independent ℓ_i in the ideal $\mathcal{I}_i$.
  - At the end, there is some $\mathcal{I}_k$ such that RHS has become 0 while LHS is *non-zero* (because $k << n$).

- So, $\Sigma^{[k]}\Pi\Sigma \subsetneq \mathsf{VBP}$!

Lower bound for Constant fanin depth-3 circuits

- det_n *cannot* be computed by a $\Sigma^{[k]}\Pi\Sigma$-circuit, *regardless* of the product fan-in.

  Proof sketch: Let $\sum_{i\in[k]} T_i = \mathsf{det}_n$, where $T_i \in \Pi\Sigma$, product of linear polynomials.

  - Pick ℓ_1 from T_1 and consider $\mathcal{I}_1 := \langle \ell_1 \rangle$.
  - LHS $\mod \mathcal{I}_1$, becomes $\sum_{i\in[k-1]} T_i'$ while $\mathsf{det}_n \mod \mathcal{I}_1 \neq 0$.
  - Keep repeating and each time include an independent ℓ_i in the ideal $\mathcal{I}_i$.
  - At the end, there is some $\mathcal{I}_k$ such that RHS has become 0 while LHS is *non-zero* (because $k << n$).

- So, $\Sigma^{[k]}\Pi\Sigma \subsetneq \mathsf{VBP}$!

- How about $\overline{\Sigma^{[k]}\Pi\Sigma}$?

Power of border depth-3 circuits

Power of border depth-3 circuits

- Recall: $h \in \overline{\Sigma^{[k]}\Pi\Sigma}$ of size s if there exists a polynomial g such that

Power of border depth-3 circuits

- Recall: $h \in \overline{\Sigma^{[k]}\Pi\Sigma}$ of size s if there exists a polynomial g such that

$$g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon) \ ,$$

Power of border depth-3 circuits

- Recall: $h \in \overline{\Sigma^{[k]}\Pi\Sigma}$ of size s if there exists a polynomial g such that

$$g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon) \ ,$$

  where g can be computed by a $\Sigma^{[k]}\Pi\Sigma$ circuit, over $\mathbb{F}(\epsilon)$, of size s.

Power of border depth-3 circuits

- Recall: $h \in \overline{\Sigma^{[k]}\Pi\Sigma}$ of size s if there exists a polynomial g such that

$$g(\boldsymbol{x}, \epsilon) = h(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon) \ ,$$

where g can be computed by a $\Sigma^{[k]}\Pi\Sigma$ circuit, over $\mathbb{F}(\epsilon)$, of size s.

Border depth-3 fan-in 2 circuits are 'universal' [Kumar 2020]

Let P be *any* n-variate degree d polynomial. Then, $P \in \overline{\Sigma^{[2]}\Pi^{[D]}\Sigma}$, where $D := \exp(n, d)$.

De-bordering $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits

De-bordering $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits

- If h is approximated by a $\Sigma^{[2]}\Pi\Sigma$ circuit with product fanin, bounded by $\mathrm{poly}(n)$, what's the *exact* complexity of h?

- If h is approximated by a $\Sigma^{[2]}\Pi\Sigma$ circuit with product fanin, bounded by $\mathsf{poly}(n)$, what's the *exact* complexity of h?

Border of polynomial-sized depth-3 top-fanin-2 circuits are 'easy' [Dutta-Dwivedi-Saxena FOCS 2021].

$\overline{\Sigma^{[2]}\Pi\Sigma} \subseteq \mathsf{VBP}$, for polynomial-sized $\overline{\Sigma^{[2]}\Pi\Sigma}$-circuits. In particular, any polynomial in the border of top-fanin-2 size-s depth-3 circuits, can also be exactly computed by a linear projection of a $\mathsf{poly}(s) \times \mathsf{poly}(s)$ determinant.

De-bordering $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits

- If h is approximated by a $\Sigma^{[2]}\Pi\Sigma$ circuit with product fanin, bounded by $\mathsf{poly}(n)$, what's the *exact* complexity of h?

Border of polynomial-sized depth-3 top-fanin-2 circuits are 'easy' [Dutta-Dwivedi-Saxena FOCS 2021].

$\overline{\Sigma^{[2]}\Pi\Sigma} \subseteq \mathsf{VBP}$, for polynomial-sized $\overline{\Sigma^{[2]}\Pi\Sigma}$-circuits. In particular, any polynomial in the border of top-fanin-2 size-s depth-3 circuits, can also be exactly computed by a linear projection of a $\mathsf{poly}(s) \times \mathsf{poly}(s)$ determinant.

Remark. The result holds if one replaces the top-fanin-2 by arbitrary constant k.

Lifting classical lower bound in the border

Lifting classical lower bound in the border

- Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?

Lifting classical lower bound in the border

- Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
  - [Limaye-Srinivasan-Tavenas, FOCS 2021] showed that $\mathsf{IMM}_{n,d}$ with $d = o(\log n)$ requires $n^{\omega(1)}$-size depth-3 circuits.

Lifting classical lower bound in the border

- Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
  - [Limaye-Srinivasan-Tavenas, FOCS 2021] showed that $\mathsf{IMM}_{n,d}$ with $d = o(\log n)$ requires $n^{\omega(1)}$-size depth-3 circuits.
  - Rank-based lower bounds can be lifted in the border!

Lifting classical lower bound in the border

- Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
  - [Limaye-Srinivasan-Tavenas, FOCS 2021] showed that $\mathsf{IMM}_{n,d}$ with $d = o(\log n)$ requires $n^{\omega(1)}$-size depth-3 circuits.
  - Rank-based lower bounds can be lifted in the border!
  - Since, $\mathsf{IMM}_{n,d} \in \mathsf{VBP}$, $\overline{\Sigma^{[k]}\Pi\Sigma} \neq \mathsf{VBP}$.

Looking for finer lower bounds

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This is already known (impossibility) in the classical setting!

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This is already known (impossibility) in the classical setting!
- $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1}$ cannot be computed by $\Sigma^{[k]}\Pi\Sigma$ circuits!

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This is already known (impossibility) in the classical setting!
- $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1}$ cannot be computed by $\Sigma^{[k]}\Pi\Sigma$ circuits!
- Catch: $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1}$ does not work anymore since, $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1} \in \overline{\Sigma^{[2]}\Pi^{O(k)}\Sigma}$!

Looking for finer lower bounds

- Can we show an *exponential* gap between $\overline{\Sigma^{[k]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This is already known (impossibility) in the classical setting!
- $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1}$ cannot be computed by $\Sigma^{[k]}\Pi\Sigma$ circuits!
- Catch: $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1}$ does not work anymore since, $x_1 \cdot y_1 + \ldots + x_{k+1} \cdot y_{k+1} \in \overline{\Sigma^{[2]}\Pi^{O(k)}\Sigma}$!
- What does work (if at all!)?

Hierarchy theorem

Our results

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

- Fix $k = 2$. Define the polynomial $P_d := x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d$, a degree-$d$ polynomial on $n = 3d$-variables.

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

- Fix $k = 2$. Define the polynomial $P_d := x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d$, a degree-$d$ polynomial on $n = 3d$-variables.
- P_d has trivial fanin-3 depth-3 circuit (and hence in border too!).

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

- Fix $k = 2$. Define the polynomial $P_d := x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d$, a degree-$d$ polynomial on $n = 3d$-variables.
- P_d has trivial fanin-3 depth-3 circuit (and hence in border too!).
- We will show that P_d requires $2^{\Omega(d)}$-size $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits.

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

- Fix $k = 2$. Define the polynomial $P_d := x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d$, a degree-$d$ polynomial on $n = 3d$-variables.
- P_d has trivial fanin-3 depth-3 circuit (and hence in border too!).
- We will show that P_d requires $2^{\Omega(d)}$-size $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits.
- Kumar's proof establishes that P_d has a $2^{O(d)}$-size $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits, showing *optimality*!

Our results

[Dutta-Saxena 2022, Preprint]

Fix any constant $k \geq 1$. There is an explicit n-variate and $< n$ degree polynomial f such that f can be computed by a $\overline{\Sigma^{[k+1]}\Pi\Sigma}$ circuit of size $O(n)$ such that f requires $2^{\Omega(n)}$-size $\overline{\Sigma^{[k]}\Pi\Sigma}$ circuits.

- Fix $k = 2$. Define the polynomial $P_d := x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d$, a degree-$d$ polynomial on $n = 3d$-variables.
- P_d has trivial fanin-3 depth-3 circuit (and hence in border too!).
- We will show that P_d requires $2^{\Omega(d)}$-size $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits.
- Kumar's proof establishes that P_d has a $2^{O(d)}$-size $\overline{\Sigma^{[2]}\Pi\Sigma}$ circuits, showing *optimality*!
- Classical is about *impossibility* while in border, it is about *optimality*.

Why $k = 2$ is hard to analyze?

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.
- E.g., $T_1 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \ldots)$, $T_2 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_4 + \ldots)$.

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.

- E.g., $T_1 := \epsilon^{-3}(1+\epsilon x_1+\epsilon^2 x_2+\epsilon^3 x_3+\ldots)$, $T_2 := \epsilon^{-3}(1+\epsilon x_1+\epsilon^2 x_2+\epsilon^3 x_4+\ldots)$. Note, $\lim_{\epsilon\to 0}\ (T_1 - T_2) = (x_3 - x_4)$.

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.
- E.g., $T_1 := \epsilon^{-3}(1+\epsilon x_1+\epsilon^2 x_2+\epsilon^3 x_3+\ldots)$, $T_2 := \epsilon^{-3}(1+\epsilon x_1+\epsilon^2 x_2+\epsilon^3 x_4+\ldots)$. Note, $\lim_{\epsilon\to 0}\ (T_1 - T_2)\ =\ (x_3 - x_4)$.
- Note $x^2 \equiv (x - \epsilon^{M/2}\cdot a)(x+\epsilon^{M/2}\cdot a) \mod \epsilon^M$, for *any* $a \in \mathbb{F}$.

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.
- E.g., $T_1 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \ldots)$, $T_2 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_4 + \ldots)$.
  Note, $\lim_{\epsilon \to 0} \ (T_1 - T_2) \ = \ (x_3 - x_4)$.
- Note $x^2 \equiv (x - \epsilon^{M/2} \cdot a)(x + \epsilon^{M/2} \cdot a) \mod \epsilon^M$, for *any* $a \in \mathbb{F}$.
- Moreover,

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon^M} \cdot \left(x^2 \ - \ (x - \epsilon^{M/2} \cdot a)(x + \epsilon^{M/2} \cdot a)\right) \ = \ a^2 \ .$$

Why $k = 2$ is hard to analyze?

- Non-trivial cancellations for $k = 2$ make things harder.
- E.g., $T_1 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \ldots)$, $T_2 := \epsilon^{-3}(1 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_4 + \ldots)$. Note, $\lim_{\epsilon \to 0} (T_1 - T_2) = (x_3 - x_4)$.
- Note $x^2 \equiv (x - \epsilon^{M/2} \cdot a)(x + \epsilon^{M/2} \cdot a) \mod \epsilon^M$, for *any* $a \in \mathbb{F}$.
- Moreover,

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon^M} \cdot \left(x^2 - (x - \epsilon^{M/2} \cdot a)(x + \epsilon^{M/2} \cdot a)\right) = a^2 .$$

- *Infinitely* many factorizations may give *infinitely* many limits.

Non-homogeneity is 'bad'

Non-homogeneity is ‘bad’

- Let $\ell_1 := 1 + \epsilon x_1$. What does taking $\mathsf{mod}\ \ell_1$ in the ‘border’ ($\epsilon \to 0$) mean? Essentially we are eventually setting $x_1 = -1/\epsilon$ (and then $\epsilon \to 0$)!

Non-homogeneity is ‘bad’

- Let $\ell_1 := 1 + \epsilon x_1$. What does taking $\mathsf{mod}\ \ell_1$ in the ‘border’ ($\epsilon \longrightarrow 0$) mean? Essentially we are eventually setting $x_1 = -1/\epsilon$ (and then $\epsilon \longrightarrow 0$)!
- In other words, work with $\mathcal{I} := \langle \ell_1, \epsilon \rangle = \langle 1 \rangle$!

Non-homogeneity is 'bad'

- Let $\ell_1 := 1 + \epsilon x_1$. What does taking $\mathsf{mod}\ \ell_1$ in the 'border' ($\epsilon \longrightarrow 0$) mean? Essentially we are eventually setting $x_1 = -1/\epsilon$ (and then $\epsilon \longrightarrow 0$)!
- In other words, work with $\mathcal{I} := \langle \ell_1, \epsilon \rangle = \langle 1 \rangle$!
- Lesson: Taking mod blindly fails *miserably*!

Non-homogeneity is ‘bad’

- Let $\ell_1 := 1 + \epsilon x_1$. What does taking $\mathrm{mod}\ \ell_1$ in the ‘border’ ($\epsilon \longrightarrow 0$) mean? Essentially we are eventually setting $x_1 = -1/\epsilon$ (and then $\epsilon \longrightarrow 0$)!
- In other words, work with $\mathcal{I} := \langle \ell_1, \epsilon \rangle = \langle 1 \rangle$!
- Lesson: Taking mod blindly fails *miserably*!
- The worst case:

$$f + \epsilon S \;=\; T_1 + T_2\,,$$

where T_i has each linear factor of the form $1 + \epsilon\ell$!

Non-homogeneity is all we need to care

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),
  - Case III (all-non-homogeneous): Each T_i has all the factors non-homogeneous.

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),
  - Case III (all-non-homogeneous): Each T_i has all the factors non-homogeneous.
- For the first case, take $\mathcal{I} := \langle \ell_1, \ell_2, \epsilon \rangle$ (and $1 \notin \mathcal{I}$) and show that $x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d = P_d \bmod \mathcal{I} \neq 0$, while RHS $\equiv 0 \bmod \mathcal{I}$.

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),
  - Case III (all-non-homogeneous): Each T_i has all the factors non-homogeneous.
- For the first case, take $\mathcal{I} := \langle \ell_1, \ell_2, \epsilon \rangle$ (and $1 \notin \mathcal{I}$) and show that $x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d = P_d \bmod \mathcal{I} \neq 0$, while RHS $\equiv 0 \bmod \mathcal{I}$.
- For the second case, take $\mathcal{I} := \langle \ell_1, \epsilon \rangle$. Then, RHS $\bmod \mathcal{I} = \overline{\Pi\Sigma} = \Pi\Sigma$, while $P_d \bmod \mathcal{I} \notin \Pi\Sigma$.

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),
  - Case III (all-non-homogeneous): Each T_i has all the factors non-homogeneous.
- For the first case, take $\mathcal{I} := \langle \ell_1, \ell_2, \epsilon \rangle$ (and $1 \notin \mathcal{I}$) and show that $x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d = P_d \bmod \mathcal{I} \neq 0$, while RHS $\equiv 0 \bmod \mathcal{I}$.
- For the second case, take $\mathcal{I} := \langle \ell_1, \epsilon \rangle$. Then, RHS $\bmod \mathcal{I} = \overline{\Pi\Sigma} = \Pi\Sigma$, while $P_d \bmod \mathcal{I} \notin \Pi\Sigma$.
- So, all-non-homogeneous is all we have to care!

Non-homogeneity is all we need to care

- Three cases to consider:
  - Case I: Each T_1 and T_2 has one linear form $\ell_i \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ as a factor,
  - Case II (intermediate): T_1 has one linear form has factor, (ℓ_1) and all factors in T_2 is non-homogeneous (in $\boldsymbol{x}$),
  - Case III (all-non-homogeneous): Each T_i has all the factors non-homogeneous.
- For the first case, take $\mathcal{I} := \langle \ell_1, \ell_2, \epsilon \rangle$ (and $1 \notin \mathcal{I}$) and show that $x_1 \cdots x_d + y_1 \cdots y_d + z_1 \cdots z_d = P_d \bmod \mathcal{I} \neq 0$, while RHS $\equiv 0 \bmod \mathcal{I}$.
- For the second case, take $\mathcal{I} := \langle \ell_1, \epsilon \rangle$. Then, RHS $\bmod \mathcal{I} = \overline{\Pi\Sigma} = \Pi\Sigma$, while $P_d \bmod \mathcal{I} \notin \Pi\Sigma$.
- So, all-non-homogeneous is all we have to care!
- For $k > 2$, we need a technical lemma, but still can be reduced!

Lower bound for all-non-homogeneous $k = 2$

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.
- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.
- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.
- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.
- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.
- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.
  - The variable z is the "degree counter".

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.

- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.

- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.

  - The variable z is the "degree counter".

- We use DiDIL - **Di**vide, **D**erive, **I**nterpolate with **L**imit, introduced in [Dutta-Dwivedi-Saxena, FOCS 2021].

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.
- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.
- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.
  - The variable z is the "degree counter".
- We use DiDIL - **Di**vide, **D**erive, **I**nterpolate with **L**imit, introduced in [Dutta-Dwivedi-Saxena, FOCS 2021].
- DiDIL shows: If P_d has s-size (comes from product fanin) border depth-3 fanin-2 all-non-homogeneous circuit, then P_d has $\mathsf{poly}(s)$ border-waring rank.

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.
- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.
- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.
  - The variable z is the "degree counter".
- We use DiDIL - **Di**vide, **D**erive, **I**nterpolate with **L**imit, introduced in [Dutta-Dwivedi-Saxena, FOCS 2021].
- DiDIL shows: If P_d has s-size (comes from product fanin) border depth-3 fanin-2 all-non-homogeneous circuit, then P_d has $\mathsf{poly}(s)$ border-waring rank.
- Partial-derivative measure shows that $\overline{\mathsf{WR}}(P_d) = 2^{\Omega(d)}$.

Lower bound for all-non-homogeneous $k = 2$

- $T_1 + T_2 = P_d(\boldsymbol{x}) + \epsilon \cdot S(\boldsymbol{x}, \epsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ have all-non-homogeneous factors.

- Very broad idea: reduce fanin 2 to 1 with a 'nice' form.

- Apply a scaling map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i$.

  - The variable z is the "degree counter".

- We use DiDIL - **Di**vide, **D**erive, **I**nterpolate with **L**imit, introduced in [Dutta-Dwivedi-Saxena, FOCS 2021].

- DiDIL shows: If P_d has s-size (comes from product fanin) border depth-3 fanin-2 all-non-homogeneous circuit, then P_d has $\mathsf{poly}(s)$ border-waring rank.

- Partial-derivative measure shows that $\overline{\mathsf{WR}}(P_d) = 2^{\Omega(d)}$.

- Thus, $s \geq 2^{\Omega(d)}$!

Why direct DiDIL fails

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.
- Shifting is required for non-homogenity for DiDIL to succeed, since $1/(1-z) \mod z^d$ exists but $1/z \mod z^d$ does not!

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.
- Shifting is required for non-homogenity for DiDIL to succeed, since $1/(1-z) \mod z^d$ exists but $1/z \mod z^d$ does not!
- If one blindly does that, we get $\partial_z(P_d/\Pi\Sigma) \equiv (\Pi\Sigma) \cdot (\overline{\Sigma \wedge \Sigma}) \mod z^d$. How to show lower bound?

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.
- Shifting is required for non-homogenity for DiDIL to succeed, since $1/(1-z) \mod z^d$ exists but $1/z \mod z^d$ does not!
- If one blindly does that, we get $\partial_z(P_d/\Pi\Sigma) \equiv (\Pi\Sigma) \cdot (\overline{\Sigma \wedge \Sigma}) \mod z^d$. How to show lower bound?
- For $k > 2$, lifting the lower bound via interpolation makes it even harder!

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.
- Shifting is required for non-homogenity for DiDIL to succeed, since $1/(1-z) \mod z^d$ exists but $1/z \mod z^d$ does not!
- If one blindly does that, we get $\partial_z(P_d/\Pi\Sigma) \equiv (\Pi\Sigma) \cdot (\overline{\Sigma \wedge \Sigma}) \mod z^d$. How to show lower bound?
- For $k > 2$, lifting the lower bound via interpolation makes it even harder!
- Shifting is problematic for showing lower bound!

Why direct DiDIL fails

- DiDIL works with $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, for *random* α_i.
- Shifting is required for non-homogenity for DiDIL to succeed, since $1/(1-z) \mod z^d$ exists but $1/z \mod z^d$ does not!
- If one blindly does that, we get $\partial_z(P_d/\Pi\Sigma) \equiv (\Pi\Sigma) \cdot (\overline{\Sigma \wedge \Sigma}) \mod z^d$. How to show lower bound?
- For $k > 2$, lifting the lower bound via interpolation makes it even harder!
- Shifting is problematic for showing lower bound!
- So, the current proof is about pre-processing (technical lemma, reducing to all-non-homogeneous) & DiDIL.

Proof sketch for $k = 2$

$k = 2$ lower bound proof: Divide and Derive

$k = 2$ lower bound proof: Divide and Derive

- Let $\Phi(T_i) =: \epsilon^{a_i} \cdot \tilde{T}_i$, for $i \in [2]$, where $a_i := \mathsf{val}_\epsilon\,(\Phi(T_i))$.
- $\mathsf{val}_\epsilon\,(\cdot)$ denotes the highest power of ϵ dividing it.

$k = 2$ lower bound proof: Divide and Derive

- Let $\Phi(T_i) =: \epsilon^{a_i} \cdot \tilde{T}_i$, for $i \in [2]$, where $a_i := \mathsf{val}_\epsilon\,(\Phi(T_i))$.
- $\mathsf{val}_\epsilon\,(\cdot)$ denotes the highest power of ϵ dividing it.
- **Di**vide and **D**erive:

$k = 2$ lower bound proof: Divide and Derive

- Let $\Phi(T_i) =: \epsilon^{a_i} \cdot \tilde{T}_i$, for $i \in [2]$, where $a_i := \mathsf{val}_\epsilon\, (\Phi(T_i))$.
- $\mathsf{val}_\epsilon\, (\cdot)$ denotes the highest power of ϵ dividing it.
- **Di**vide and **D**erive:

$$
\begin{aligned}
P_d + \epsilon \cdot S =\; & T_1 + T_2 \\
\implies \Phi(P_d) + \epsilon \cdot \Phi(S) =\; & \Phi(T_1) + \Phi(T_2) \\
\implies \Phi(P_d)/\tilde{T}_2 + \epsilon \cdot \Phi(S)/\tilde{T}_2 =\; & \epsilon^{a_2} + \Phi(T_1)/\tilde{T}_2 \\
\implies \partial_z\left(\Phi(P_d)/\tilde{T}_2\right) + \epsilon \cdot \partial_z\left(\Phi(S)/\tilde{T}_2\right) =\; & \partial_z\left(\Phi(T_1)/\tilde{T}_2\right) =: g_1 \;. \qquad (1)
\end{aligned}
$$

$k = 2$ lower bound proof: Divide and Derive

- Let $\Phi(T_i) =: \epsilon^{a_i} \cdot \tilde{T}_i$, for $i \in [2]$, where $a_i := \mathrm{val}_\epsilon\,(\Phi(T_i))$.
- $\mathrm{val}_\epsilon\,(\cdot)$ denotes the highest power of ϵ dividing it.
- **Di**vide and **D**erive:

$$\begin{aligned} P_d + \epsilon \cdot S &= T_1 + T_2 \\ \implies \Phi(P_d) + \epsilon \cdot \Phi(S) &= \Phi(T_1) + \Phi(T_2) \\ \implies \Phi(P_d)/\tilde{T}_2 + \epsilon \cdot \Phi(S)/\tilde{T}_2 &= \epsilon^{a_2} + \Phi(T_1)/\tilde{T}_2 \\ \implies \partial_z \left(\Phi(P_d)/\tilde{T}_2\right) + \epsilon \cdot \partial_z \left(\Phi(S)/\tilde{T}_2\right) &= \partial_z \left(\Phi(T_1)/\tilde{T}_2\right) =: g_1 \,. \end{aligned} \tag{1}$$

- $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$, where $\mathbb{F} \ni t_2 := \lim_{\epsilon \to 0} \tilde{T}_2$, because $\lim_{\epsilon \to 0} \prod(1 + \epsilon \ell_i) = 1$.

$k = 2$ proof continued

$k = 2$ proof continued

- Target: Compute $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$.

$k = 2$ proof continued

- Target: Compute $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$.
- Logarithmic derivative: $\mathrm{dlog}_z(h) := \partial_z(h)/h$.

$k = 2$ proof continued

- Target: Compute $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$.
- Logarithmic derivative: $\mathrm{dlog}_z(h) := \partial_z(h)/h$.
- dlog *linearizes* product: $\mathrm{dlog}(h_1 h_2) = \mathrm{dlog}(h_1) + \mathrm{dlog}(h_2)$.

$k = 2$ proof continued

- Target: Compute $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$.
- Logarithmic derivative: $\mathrm{dlog}_z(h) := \partial_z(h)/h$.
- dlog *linearizes* product: $\mathrm{dlog}(h_1 h_2) = \mathrm{dlog}(h_1) + \mathrm{dlog}(h_2)$. Note:

$$\begin{aligned}\partial_z\left(\Phi(T_1)/\tilde{T}_2\right) &= \Phi(T_1)/\tilde{T}_2 \cdot \mathrm{dlog}\left(\Phi(T_1)/\tilde{T}_2\right)\\ &= (\Pi\Sigma/\Pi\Sigma) \cdot \mathrm{dlog}\,(\Pi\Sigma/\Pi\Sigma)\\ &= \Pi\Sigma/\Pi\Sigma \cdot \left(\pm \sum \mathrm{dlog}(\Sigma)\right).\end{aligned}$$

$k = 2$ proof continued

- Target: Compute $\lim_{\epsilon \to 0} g_1 = 1/t_2 \cdot \partial_z(\Phi(P_d))$.
- Logarithmic derivative: $\mathrm{dlog}_z(h) := \partial_z(h)/h$.
- dlog *linearizes* product: $\mathrm{dlog}(h_1 h_2) = \mathrm{dlog}(h_1) + \mathrm{dlog}(h_2)$. Note:

$$\begin{aligned}\partial_z\left(\Phi(T_1)/\tilde{T}_2\right) &= \Phi(T_1)/\tilde{T}_2 \cdot \mathrm{dlog}\left(\Phi(T_1)/\tilde{T}_2\right) \\ &= (\Pi\Sigma/\Pi\Sigma) \cdot \mathrm{dlog}\,(\Pi\Sigma/\Pi\Sigma) \\ &= \Pi\Sigma/\Pi\Sigma \cdot \left(\pm \sum \mathrm{dlog}(\Sigma)\right).\end{aligned}$$

- Here Σ means just a linear polynomial ℓ.

$k = 2$ proof continued: Quick recap

$k = 2$ proof continued: Quick recap

- Recap: $1/t_2 \cdot \partial_z(\Phi(P_d)) = \lim_{\epsilon \to 0} g_1 = \lim_{\epsilon \to 0} (\Pi\Sigma/\Pi\Sigma) \cdot (\pm \sum \mathrm{dlog}(\Sigma))$.

$k = 2$ proof continued: Quick recap

- Recap: $1/t_2 \cdot \partial_z(\Phi(P_d)) = \lim_{\epsilon \to 0} g_1 = \lim_{\epsilon \to 0} (\Pi\Sigma/\Pi\Sigma) \cdot (\pm \sum \mathrm{dlog}(\Sigma))$.
- $\deg(P_d) = d \implies \deg_z(\Phi(f)) = d \implies \deg_z(\partial_z(\Phi(P_d))) = d - 1$.

$k = 2$ proof continued: Quick recap

- Recap: $1/t_2 \cdot \partial_z(\Phi(P_d)) = \lim_{\epsilon \to 0} g_1 = \lim_{\epsilon \to 0} (\Pi\Sigma/\Pi\Sigma) \cdot (\pm \sum \mathrm{dlog}(\Sigma))$.
- $\deg(P_d) = d \implies \deg_z(\Phi(f)) = d \implies \deg_z(\partial_z(\Phi(P_d))) = d - 1$.
- Suffices to compute $\lim_{\epsilon \to 0} g_1 \mod z^d$.

$k = 2$ **proof:** dlog **strikes!**

$k = 2$ **proof:** dlog **strikes!**

- What is $\mathrm{dlog}(\ell)$ for a linear polynomial $\ell = 1 - z \cdot \ell$?

$k = 2$ proof: dlog strikes!

- What is $\mathrm{dlog}(\ell)$ for a linear polynomial $\ell = 1 - z \cdot \ell$?

$$\begin{aligned} \mathrm{dlog}(1 - z\ell) &= -\frac{\ell}{(1 - z \cdot \ell)} \\ &= \sum_{j=0}^{d-1} \ell \cdot (z \cdot \ell)^j \mod z^d \\ &\in \Sigma \wedge \Sigma \,. \end{aligned}$$

$k = 2$ proof: dlog strikes!

- What is $\mathsf{dlog}(\ell)$ for a linear polynomial $\ell = 1 - z \cdot \ell$?

$$\begin{aligned}\mathsf{dlog}(1 - z\ell) &= -\frac{\ell}{(1 - z \cdot \ell)} \\ &= \sum_{j=0}^{d-1} \ell \cdot (z \cdot \ell)^j \mod z^d \\ &\in \Sigma \wedge \Sigma \,.\end{aligned}$$

Thus,

$$\begin{aligned}\lim_{\epsilon \to 0} g_1 \mod z^d &\equiv \lim_{\epsilon \to 0} (\Pi\Sigma/\Pi\Sigma) \cdot \left(\sum \mathsf{dlog}(\Sigma)\right) \mod z^d \\ &\equiv \lim_{\epsilon \to 0} (\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma) \mod z^d \\ &\in \overline{(\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma)} \mod z^d \,.\end{aligned}$$

Finishing the proof

Finishing the proof

- $\overline{\mathcal{C} \cdot \mathcal{D}} = \overline{\mathcal{C}} \cdot \overline{\mathcal{D}}$. Therefore,

Finishing the proof

- $\overline{\mathcal{C} \cdot \mathcal{D}} = \overline{\mathcal{C}} \cdot \overline{\mathcal{D}}$. Therefore,

$$\begin{aligned}\overline{(\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma)} &= \overline{(\Pi\Sigma/\Pi\Sigma)} \cdot (\overline{\Sigma \wedge \Sigma}) \\ &\subseteq \overline{\Sigma \wedge \Sigma}\,.\end{aligned}$$

Finishing the proof

- $\overline{\mathcal{C} \cdot \mathcal{D}} = \overline{\mathcal{C}} \cdot \overline{\mathcal{D}}$. Therefore,

$$\begin{aligned}\overline{(\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma)} &= \overline{(\Pi\Sigma/\Pi\Sigma)} \cdot (\overline{\Sigma \wedge \Sigma}) \\ &\subseteq \overline{\Sigma \wedge \Sigma}\,.\end{aligned}$$

- The above is because $\lim_{\epsilon \to 0} \prod (1 + \epsilon \ell) = 1$.

Finishing the proof

- $\overline{\mathcal{C} \cdot \mathcal{D}} = \overline{\mathcal{C}} \cdot \overline{\mathcal{D}}$. Therefore,

$$\overline{(\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma)} = \overline{(\Pi\Sigma/\Pi\Sigma)} \cdot (\overline{\Sigma \wedge \Sigma}) \subseteq \overline{\Sigma \wedge \Sigma}\,.$$

- The above is because $\lim_{\epsilon \to 0} \prod (1 + \epsilon\ell) = 1$.

- Thus, $1/t_2 \cdot \partial_z(\Phi(P_d)) = \lim_{\epsilon \to 0} g_1 \in \overline{\Sigma \wedge \Sigma}$.

Finishing the proof

- $\overline{\mathcal{C} \cdot \mathcal{D}} = \overline{\mathcal{C}} \cdot \overline{\mathcal{D}}$. Therefore,

$$\begin{aligned}\overline{(\Pi\Sigma/\Pi\Sigma) \cdot (\Sigma \wedge \Sigma)} &= \overline{(\Pi\Sigma/\Pi\Sigma)} \cdot (\overline{\Sigma \wedge \Sigma}) \\ &\subseteq \overline{\Sigma \wedge \Sigma}\,.\end{aligned}$$

- The above is because $\lim_{\epsilon \to 0} \prod (1 + \epsilon \ell) = 1$.
- Thus, $1/t_2 \cdot \partial_z(\Phi(P_d)) = \lim_{\epsilon \to 0} g_1 \in \overline{\Sigma \wedge \Sigma}$.
- Thus, $\Phi(P_d) \in \overline{\Sigma \wedge \Sigma} \implies P_d \in \overline{\Sigma \wedge \Sigma}$.

Conclusion

Concluding remarks

Concluding remarks

- Can we show exponential lower bound for $\overline{\Sigma^{[o(n)]}\Pi\Sigma}$-circuits? The current method gives subexponential lower bound only as long as $k = o(\log n)$.

Concluding remarks

- Can we show exponential lower bound for $\overline{\Sigma^{[o(n)]}\Pi\Sigma}$-circuits? The current method gives subexponential lower bound only as long as $k = o(\log n)$.
- Can we show exponential lower bound for $\overline{\Sigma^{[k]}\Pi\Sigma\wedge}$-circuits (i.e. rather special depth-4)?

Concluding remarks

- Can we show exponential lower bound for $\overline{\Sigma^{[o(n)]}\Pi\Sigma}$-circuits? The current method gives subexponential lower bound only as long as $k = o(\log n)$.
- Can we show exponential lower bound for $\overline{\Sigma^{[k]}\Pi\Sigma\wedge}$-circuits (i.e. rather special depth-4)?
- Can we extend the hierarchy theorem to bounded (top & bottom fanin) depth-4 circuits? i.e., for a *fixed* constant δ, is $\overline{\Sigma^{[1]}\Pi\Sigma\Pi^{[\delta]}} \subsetneq \overline{\Sigma^{[2]}\Pi\Sigma\Pi^{[\delta]}} \subsetneq \overline{\Sigma^{[3]}\Pi\Sigma\Pi^{[\delta]}} \ldots$, where the respective gaps are exponential? Clearly, $\delta = 1$ holds, from this work.

Concluding remarks

- Can we show exponential lower bound for $\overline{\Sigma^{[o(n)]}\Pi\Sigma}$-circuits? The current method gives subexponential lower bound only as long as $k = o(\log n)$.
- Can we show exponential lower bound for $\overline{\Sigma^{[k]}\Pi\Sigma\wedge}$-circuits (i.e. rather special depth-4)?
- Can we extend the hierarchy theorem to bounded (top & bottom fanin) depth-4 circuits? i.e., for a *fixed* constant δ, is $\overline{\Sigma^{[1]}\Pi\Sigma\Pi^{[\delta]}} \subsetneq \overline{\Sigma^{[2]}\Pi\Sigma\Pi^{[\delta]}} \subsetneq \overline{\Sigma^{[3]}\Pi\Sigma\Pi^{[\delta]}} \ldots$, where the respective gaps are exponential? Clearly, $\delta = 1$ holds, from this work.

Thank you! Questions?