Demystifying the border of depth-3

Joint works with Pranjal Dutta & Prateek Dwivedi. [CCC'21, FOCS'21, FOCS'22]

Nitin Saxena
CSE, IIT Kanpur

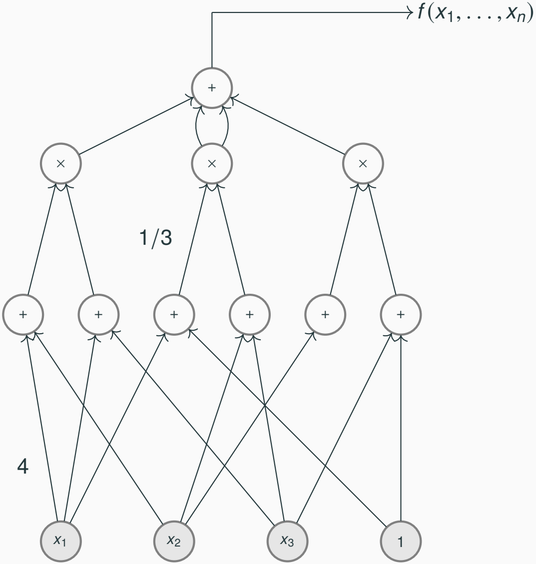
August, 2023

WAC @ Göteborg, Sweden

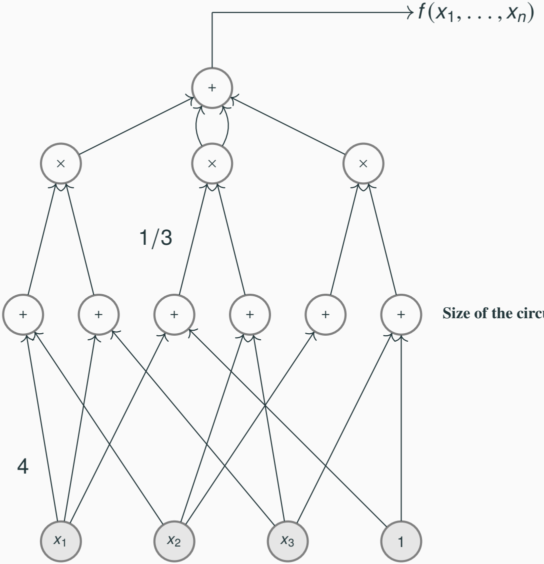
Table of Contents

Basic Definitions and Terminologies

Algebraic circuits– VP

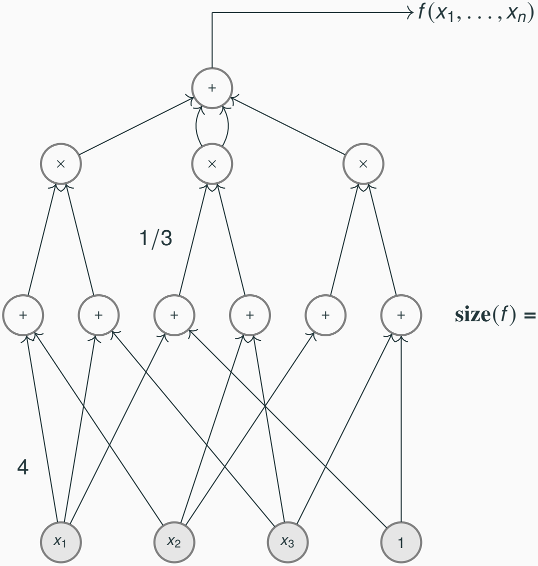


Algebraic circuits– VP



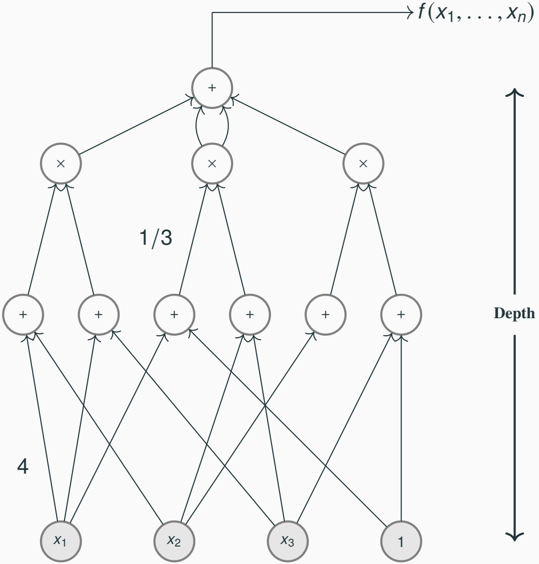
Size of the circuit = number of nodes + edges

Algebraic circuits– VP



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

Algebraic circuits– VP



The determinant polynomial– VBP

The determinant polynomial– VBP

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

$$\mathrm{det}_s \ := \ \mathrm{det}(X_s) \ = \ \sum_{\pi \in S_s} \mathrm{sgn}(\pi) \prod_{i=1}^{s} x_{i,\pi(i)} \ .$$

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

$$\mathsf{det}_s := \mathsf{det}(X_s) = \sum_{\pi \in S_s} \operatorname{sgn}(\pi) \prod_{i=1}^{s} x_{i,\pi(i)} .$$

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

The determinant polynomial– VBP

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

$$\mathrm{det}_s := \mathrm{det}(X_s) = \sum_{\pi \in S_s} \mathrm{sgn}(\pi) \prod_{i=1}^{s} x_{i,\pi(i)} .$$

- VBP: The class VBP is defined as the set of all sequences of polynomials $(f_n)_n$ with polynomially bounded determinantal-complexity $\mathrm{dc}(f_n)$.
- Relates tightly to Algebraic Branching Programs ABP, or IMM: Iterated Matrix Multiplication.

‘Hard’ polynomials?

- Hard polynomial family $(f_n)_n$ such that it cannot be computed by an n^c-size determinant, for *every* constant c? i.e. $\text{size}(f_n) = n^{\omega(1)}$?

'Hard' polynomials?

- Hard polynomial family $(f_n)_n$ such that it cannot be computed by an n^c-size determinant, for *every* constant c? i.e. $\mathrm{size}(f_n) = n^{\omega(1)}$?
- A *random* polynomial with 0-1 coefficient is **hard** [Hrubeš-Yehudayoff ToC'11].

‘Hard’ polynomials?

- Hard polynomial family $(f_n)_n$ such that it cannot be computed by an n^c-size determinant, for *every* constant c? i.e. $\text{size}(f_n) = n^{\omega(1)}$?
- A *random* polynomial with 0-1 coefficient is **hard** [Hrubeš-Yehudayoff ToC’11]. Challenge: Find an **explicit** one!

‘Hard’ polynomials?

- Hard polynomial family $(f_n)_n$ such that it cannot be computed by an n^c-size determinant, for *every* constant c? i.e. $\text{size}(f_n) = n^{\omega(1)}$?
- A *random* polynomial with 0-1 coefficient is **hard** [Hrubeš-Yehudayoff ToC'11]. Challenge: Find an **explicit** one!
- Candidate hard polynomial:

$$\text{perm}(X_s) = \sum_{\pi \in S_s} \prod_{i=1}^{s} x_{i,\pi(i)} .$$

‘Hard’ polynomials?

- Hard polynomial family $(f_n)_n$ such that it cannot be computed by an n^c-size determinant, for *every* constant c? i.e. $\text{size}(f_n) = n^{\omega(1)}$?
- A *random* polynomial with 0-1 coefficient is **hard** [Hrubeš-Yehudayoff ToC’11]. Challenge: Find an **explicit** one!
- Candidate hard polynomial:

$$\text{perm}(X_s) = \sum_{\pi \in S_s} \prod_{i=1}^{s} x_{i,\pi(i)} .$$

- The minimum dimension of the matrix X_s to compute f, is called the **permanental complexity** $\text{pc}(f)$.

Valiant's Conjecture– VNP

VNP = "explicit" (but "hard to compute"?) [Valiant 1979]

The class VNP is defined as the set of all sequences of polynomials $(f_n(x_1, \ldots, x_n))_{n \geq 1}$ such that $\mathrm{pc}(f_n)$ is bounded by n^c for some constant c.

Valiant's Conjecture– VNP

VNP = "explicit" (but "hard to compute"?) [Valiant 1979]

The class VNP is defined as the set of all sequences of polynomials $(f_n(x_1, \ldots, x_n))_{n \geq 1}$ such that $\mathsf{pc}(f_n)$ is bounded by n^c for some constant c.

- $\mathsf{VBP} \subseteq \mathsf{VP} \subseteq \mathsf{VNP}$.

Valiant's Conjecture– VNP

VNP = "explicit" (but "hard to compute"?) [Valiant 1979]

The class VNP is defined as the set of all sequences of polynomials $(f_n(x_1, \ldots, x_n))_{n \geq 1}$ such that $\mathsf{pc}(f_n)$ is bounded by n^c for some constant c.

- $\mathsf{VBP} \subseteq \mathsf{VP} \subseteq \mathsf{VNP}$.

Valiant's Conjecture [Valiant 1979]

$\mathsf{VBP} \neq \mathsf{VNP}$ & $\mathsf{VP} \neq \mathsf{VNP}$.

Equivalently, $\mathsf{dc}(\mathsf{perm}_n)$ and size(perm_n) are both $n^{\omega(1)}$.

Border Complexity and GCT

Border complexity

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

Border complexity

- Let Γ be any sensible measure. Eg. it can be size, dc and so on.
- For Γ, we can define the border complexity measure $\overline{\Gamma}$ via:

  $\overline{\Gamma}(h)$ is the *smallest* s such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\varepsilon(\boldsymbol{x})$ with $\Gamma(h_\varepsilon) \leq s$.

Border complexity

- Let Γ be any sensible measure. Eg. it can be size, dc and so on.
- For Γ, we can define the border complexity measure $\overline{\Gamma}$ via:

  $\overline{\Gamma}(h)$ is the *smallest* s such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\varepsilon(\boldsymbol{x})$ with $\Gamma(h_\varepsilon) \leq s$. In other words,

$$\lim_{\varepsilon \to 0} h_\varepsilon = h \quad (\textit{least}\text{-coefficient-wise}) .$$

Border complexity

- Let Γ be any sensible measure. Eg. it can be size, dc and so on.
- For Γ, we can define the border complexity measure $\overline{\Gamma}$ via:

  $\overline{\Gamma}(h)$ is the *smallest* s such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\varepsilon(\boldsymbol{x})$ with $\Gamma(h_\varepsilon) \leq s$. In other words,

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

  Eg. $\lim_{\varepsilon \to 0} \left(\varepsilon z + \varepsilon^{-1} z^2 x_1\right) = \lim_{\varepsilon \to 0} \left(\varepsilon^2 z + z^2 x_1\right) = z^2 x_1$.

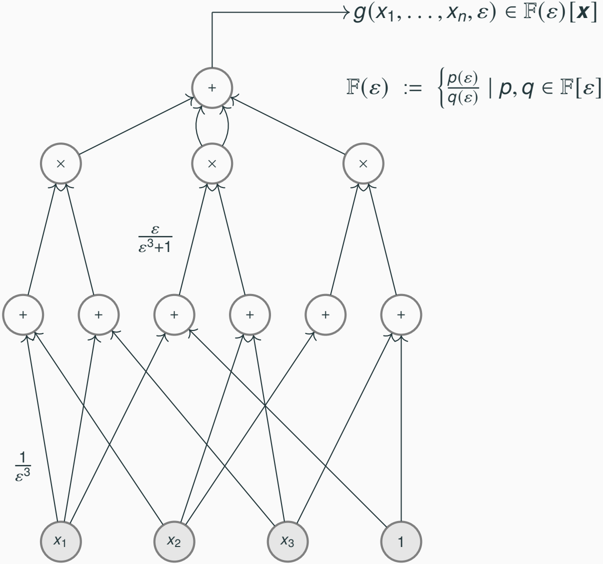
Border complexity

- Let Γ be any sensible measure. Eg. it can be size, dc and so on.
- For Γ, we can define the border complexity measure $\overline{\Gamma}$ via:

  $\overline{\Gamma}(h)$ is the *smallest* s such that $h(\boldsymbol{x})$ can be approximated arbitrarily closely by polynomials $h_\varepsilon(\boldsymbol{x})$ with $\Gamma(h_\varepsilon) \le s$. In other words,

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

  Eg. $\lim_{\varepsilon \to 0} \left(\varepsilon z + \varepsilon^{-1} z^2 x_1\right) = \lim_{\varepsilon \to 0} \left(\varepsilon^2 z + z^2 x_1\right) = z^2 x_1$.
- This motivates a new model: *'approximative circuit'*.



$\mathbb{F}(\varepsilon) := \left\{ \frac{p(\varepsilon)}{q(\varepsilon)} \mid p, q \in \mathbb{F}[\varepsilon], q(\varepsilon) \neq 0 \right\}$ is the fn.field

Algebraic approximation

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

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

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

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

  - How easy is g_0 in comparison to g?

Algebraic approximation

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

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

  - How easy is g_0 in comparison to g?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\varepsilon = 0$?!

Algebraic approximation

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

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

  - How easy is g_0 in comparison to g?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\varepsilon = 0$?! Setting $\varepsilon = 0$ *may not* be 'valid' as it could be using $1/\varepsilon$ in the wire. Though it is well-defined!

Algebraic approximation

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

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

  - How easy is g_0 in comparison to g?
- Obvious attempt:
  - Since, $g(\boldsymbol{x}, 0) = g_0$, why not just set $\varepsilon = 0$?! Setting $\varepsilon = 0$ *may not* be 'valid' as it could be using $1/\varepsilon$ in the wire. Though it is well-defined!
- Summary: g_0 is **non-trivially** 'approximated' by the circuit, since $\lim_{\varepsilon \to 0} g(\boldsymbol{x}, \varepsilon) = g_0$.

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

Algebraic approximation— $\overline{\mathrm{VP}}$

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\mathbf{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\mathbf{x}]$, of size s, and an error polynomial $S(\mathbf{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\mathbf{x}]$ such that
$g(\mathbf{x}, \varepsilon) = h(\mathbf{x}) + \varepsilon \cdot S(\mathbf{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020].

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020]. [Bezout's degree theorem.]

Algebraic approximation— $\overline{\mathrm{VP}}$

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020]. [Bezout's degree theorem.]
- Let us assume that $g(\boldsymbol{x}, \varepsilon) = \sum_{i=0}^{M} g_i \varepsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that $g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020]. [Bezout's degree theorem.]
- Let us assume that $g(\boldsymbol{x}, \varepsilon) = \sum_{i=0}^{M} g_i \varepsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - Pick $M + 1$ many values from $\mathbb{F}$ *randomly* and interpolate ε;

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that
$g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020]. [Bezout's degree theorem.]

- Let us assume that $g(\boldsymbol{x}, \varepsilon) = \sum_{i=0}^{M} g_i \varepsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - Pick $M + 1$ many values from $\mathbb{F}$ *randomly* and interpolate ε;

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

Algebraic approximation— $\overline{\mathrm{VP}}$

Algebraic Approximation [Bürgisser 2004]

A polynomial $h \in \mathbb{F}[\boldsymbol{x}]$ has approximative complexity s, if there is a $g \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$, of size s, and an error polynomial $S(\boldsymbol{x}, \varepsilon) \in \mathbb{F}[\varepsilon][\boldsymbol{x}]$ such that
$g(\boldsymbol{x}, \varepsilon) = h(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$.

Informally we write, $\lim_{\varepsilon \to 0} g = h$.

- If g has circuit of size s over $\mathbb{F}(\varepsilon)$, then the degree of ε in g is at most *exponential*, 2^{s^2} [Lehmkuhl, Lickteig 1989] [Bürgisser 2004, erratum-2020]. [Bezout's degree theorem.]
- Let us assume that $g(\boldsymbol{x}, \varepsilon) = \sum_{i=0}^{M} g_i \varepsilon^i$, where $M = 2^{s^2}$. Note: $g_0 = h$.
  - ➢ Pick $M + 1$ many values from $\mathbb{F}$ *randomly* and interpolate ε;
- $\overline{\text{size}}(h) \le \text{size}(h) \le \exp(\overline{\text{size}}(h))$.
- Curious eg.: Complexity of degree-s factor of a size-s polynomial? VP? VNP?

Border 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. $a_0 + a_1x_1 + \ldots + a_nx_n$, 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 $\Sigma^{[2]}\Pi\Sigma$ circuits? 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 $\Sigma^{[2]}\Pi\Sigma$ circuits? Are they *universal*?
- Impossibility result: The *Inner Product* polynomial $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1y_1 + x_2y_2 + x_3y_3$ **cannot** be written as a $\Sigma^{[2]}\Pi\Sigma$ circuit,

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 $\Sigma^{[2]}\Pi\Sigma$ circuits? Are they *universal*?
- Impossibility result: The *Inner Product* polynomial $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := x_1y_1 + x_2y_2 + x_3y_3$ **cannot** be written as a $\Sigma^{[2]}\Pi\Sigma$ 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 $\Sigma^{[2]}\Pi\Sigma$ circuits? Are they *universal*?
- Impossibility result: The *Inner Product* polynomial $\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + x_2 y_2 + x_3 y_3$ **cannot** be written as a $\Sigma^{[2]}\Pi\Sigma$ circuit, *regardless* of the product fan-in (even allowing $\exp(n)$ product fan-in!).
- How about $\overline{\Sigma^{[2]}\Pi\Sigma}$?

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

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

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

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

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

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

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

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

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

Let P be *any* n-variate degree d polynomial. Then, $P \in \overline{\Sigma^{[2]}\Pi\Sigma}$,

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

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

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

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

Let P be *any* n-variate degree d polynomial. Then, $P \in \overline{\Sigma^{[2]}\Pi\Sigma}$, where the first product has fanin $\exp(n, d)$ and the second is merely constant !

Proof of Kumar's result

Proof.

skip proof

1. Let $\mathsf{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

Proof of Kumar's result

Proof.

⏩ skip proof

1. Let $\mathsf{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n,d)] \ .$$

Proof of Kumar's result

Proof.

skip proof

1. Let $\mathrm{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n,d)]\ .$$

2. Consider $A(\boldsymbol{x}) := \prod_{i=1}^{m}(1+\ell_i^d) = \prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j+\ell_i)$, for $\alpha_j \in \mathbb{C}$.

Proof of Kumar's result

Proof.

skip proof

1. Let $\mathsf{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n,d)] \ .$$

2. Consider $A(\boldsymbol{x}) := \prod_{i=1}^{m}(1 + \ell_i^d) = \prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \ell_i)$, for $\alpha_j \in \mathbb{C}$. Note that

$$A(\boldsymbol{x}) = 1 + P + B \text{ where } \deg(B) \geq 2d \ .$$

Proof of Kumar's result

Proof.

⏩ skip proof

1. Let $\mathsf{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n, d)] .$$

2. Consider $A(\boldsymbol{x}) := \prod_{i=1}^{m} (1 + \ell_i^d) = \prod_{i=1}^{m} \prod_{j=1}^{d} (\alpha_j + \ell_i)$, for $\alpha_j \in \mathbb{C}$. Note that

$$A(\boldsymbol{x}) = 1 + P + B \text{ where } \deg(B) \geq 2d .$$

3. Replace x_i by $\varepsilon \cdot x_i$ to get that

Proof of Kumar's result

Proof.

skip proof

1. Let $\mathsf{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n,d)]\ .$$

2. Consider $A(\boldsymbol{x}) := \prod_{i=1}^{m}(1+\ell_i^d) = \prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \ell_i)$, for $\alpha_j \in \mathbb{C}$. Note that

$$A(\boldsymbol{x}) = 1 + P + B \text{ where } \deg(B) \geq 2d\ .$$

3. Replace x_i by $\varepsilon \cdot x_i$ to get that

$$\prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \varepsilon \cdot \ell_i) = 1 + \varepsilon^d \cdot P + \varepsilon^{2d} \cdot R(\boldsymbol{x}, \varepsilon)\ .$$

Proof of Kumar's result

Proof.

⏩ skip proof

1. Let $\mathrm{WR}(P) =: m$. Then, there are linear forms ℓ_i such that

$$P = \sum_{i=1}^{m} \ell_i^d \qquad [m \text{ can be as large as } \exp(n,d)] \ .$$

2. Consider $A(\boldsymbol{x}) := \prod_{i=1}^{m}(1 + \ell_i^d) = \prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \ell_i)$, for $\alpha_j \in \mathbb{C}$. Note that

$$A(\boldsymbol{x}) = 1 + P + B \text{ where } \deg(B) \geq 2d \ .$$

3. Replace x_i by $\varepsilon \cdot x_i$ to get that

$$\prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \varepsilon \cdot \ell_i) = 1 + \varepsilon^d \cdot P + \varepsilon^{2d} \cdot R(\boldsymbol{x}, \varepsilon) \ .$$

4. Divide by ε^d and rearrange to get

$$P + \varepsilon^d \cdot R(\boldsymbol{x}, \varepsilon) = -\varepsilon^{-d} + \varepsilon^{-d} \cdot \prod_{i=1}^{m}\prod_{j=1}^{d}(\alpha_j + \varepsilon \cdot \ell_i) \in \Sigma^{[2]}\Pi^{[md]}\Sigma \ .$$

Proving Upper Bounds

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

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

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

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

Border of poly-size depth-3 top-fanin-2 circuits are 'easy' [Dutta-Dwivedi-Saxena FOCS'21].

$\overline{\Sigma^{[2]}\Pi\Sigma} \subseteq \mathsf{VBP}$, for polynomial-sized $\overline{\Sigma^{[2]}\Pi\Sigma}$-circuits.

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

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

Border of poly-size depth-3 top-fanin-2 circuits are 'easy' [Dutta-Dwivedi-Saxena FOCS'21].

$\overline{\Sigma^{[2]}\Pi\Sigma} \subseteq \mathsf{VBP}$, for polynomial-sized $\overline{\Sigma^{[2]}\Pi\Sigma}$-circuits.

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

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- ❑ $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- ❑ $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- ❑ Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.
  - ➢ The variable z is the "degree counter",

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.
  - The variable z is the "degree counter",
  - α_i ensures "unit" : If $\ell \mid T_i$, then $\Phi(\ell)|_{z=0} = \ell(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}(\varepsilon)^*$.

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.
  - The variable z is the "degree counter",
  - α_i ensures "unit" : If $\ell \mid T_i$, then $\Phi(\ell)|_{z=0} = \ell(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}(\varepsilon)^*$.
- There's *no* loss if we study $\Phi(f) \mod z^{d+1}$.

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- ❑ $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- ❑ Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.
  - ➢ The variable z is the "degree counter",
  - ➢ α_i ensures "unit" : If $\ell \mid T_i$, then $\Phi(\ell)|_{z=0} = \ell(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}(\varepsilon)^*$.
- ❑ There's *no* loss if we study $\Phi(f) \mod z^{d+1}$. [Truncation by degree.]

Proof sketch for $k = 2$

Grand Idea: Reduce to $k = 1$!

- $T_1 + T_2 = f(\boldsymbol{x}) + \varepsilon \cdot S(\boldsymbol{x}, \varepsilon)$, where $T_i \in \Pi\Sigma \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$. Assume $\deg(f) = d$.
- Apply a map Φ, defined by $\Phi : x_i \mapsto z \cdot x_i + \alpha_i$, where $\alpha_i \in \mathbb{F}$ are *random*.
  - The variable z is the "degree counter",
  - α_i ensures "unit" : If $\ell \mid T_i$, then $\Phi(\ell)|_{z=0} = \ell(\alpha_1, \ldots, \alpha_n) \in \mathbb{F}(\varepsilon)^*$.
- There's *no* loss if we study $\Phi(f) \mod z^{d+1}$. [Truncation by degree.]
- We devise a technique called DiDIL - **Di**vide, **D**erive, **I**nduct with **L**imit.

$k = 2$ proof continued: *Di*vide and *D*erive

- $\mathsf{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathsf{val}_z(h) = 1$.

$k = 2$ proof continued: *Di*vide and *De*rive

- $\mathsf{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathsf{val}_z(h) = 1$.
- **Analysis trivia:** $\mathsf{val}_z(h) = 0$ makes $1/h$ a power-series in $\mathbb{F}(\varepsilon, \boldsymbol{x})[[z]]$.

$k = 2$ proof continued: *Di*vide and *De*rive

- $\mathsf{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathsf{val}_z(h) = 1$.

- **Analysis trivia:** $\mathsf{val}_z(h) = 0$ makes $1/h$ a power-series in $\mathbb{F}(\varepsilon, \boldsymbol{x})[[z]]$.

- Wlog $\mathsf{val}_z(\Phi(T_2)) \leq \mathsf{val}_z(\Phi(T_1))$, else we can rearrange.

$k = 2$ proof continued: *Di*vide and *De*rive

- $\mathrm{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathrm{val}_z(h) = 1$.
- **Analysis trivia:** $\mathrm{val}_z(h) = 0$ makes $1/h$ a power-series in $\mathbb{F}(\varepsilon, \boldsymbol{x})[[z]]$.
- Wlog $\mathrm{val}_z(\Phi(T_2)) \leq \mathrm{val}_z(\Phi(T_1))$, else we can rearrange.
- Divide both sides by $\Phi(T_2)$ and take partial derivative with respect to z, to get:

$k = 2$ proof continued: *Di*vide and *De*rive

- $\mathsf{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathsf{val}_z(h) = 1$.
- **Analysis trivia:** $\mathsf{val}_z(h) = 0$ makes $1/h$ a power-series in $\mathbb{F}(\varepsilon, \boldsymbol{x})[[z]]$.
- Wlog $\mathsf{val}_z(\Phi(T_2)) \leq \mathsf{val}_z(\Phi(T_1))$, else we can rearrange.
- Divide both sides by $\Phi(T_2)$ and take partial derivative with respect to z, to get:

$$\begin{aligned} \Phi(f) + \varepsilon \cdot \Phi(S) &= \Phi(T_1) + \Phi(T_2) \\ \implies \Phi(f/T_2) + \varepsilon \cdot \Phi(S/T_2) &= \Phi(T_1/T_2) + 1 \\ \implies \partial_z \Phi(f/T_2) + \varepsilon \cdot \partial_z \Phi(S/T_2) &= \partial_z \Phi(T_1/T_2) =: g_1 \, . \end{aligned} \tag{1}$$

$k = 2$ proof continued: *Di*vide and *De*rive

- $\mathrm{val}_z(\cdot)$ denotes the highest power of z dividing it (= least one across monomials). E.g., $h = \varepsilon z + \varepsilon^{-1} z^2 x_1 = (\varepsilon z) \cdot (1 + \varepsilon^{-2} z x_1)$. Then, $\mathrm{val}_z(h) = 1$.
- **Analysis trivia:** $\mathrm{val}_z(h) = 0$ makes $1/h$ a power-series in $\mathbb{F}(\varepsilon, \boldsymbol{x})[[z]]$.
- Wlog $\mathrm{val}_z(\Phi(T_2)) \leq \mathrm{val}_z(\Phi(T_1))$, else we can rearrange.
- Divide both sides by $\Phi(T_2)$ and take partial derivative with respect to z, to get:

$$\begin{aligned} \Phi(f) + \varepsilon \cdot \Phi(S) &= \Phi(T_1) + \Phi(T_2) \\ \implies \Phi(f/T_2) + \varepsilon \cdot \Phi(S/T_2) &= \Phi(T_1/T_2) + 1 \\ \implies \partial_z \Phi(f/T_2) + \varepsilon \cdot \partial_z \Phi(S/T_2) &= \partial_z \Phi(T_1/T_2) =: g_1 . \end{aligned} \tag{1}$$

- $\lim_{\varepsilon \to 0} g_1 = \lim_{\varepsilon \to 0} \partial_z \Phi(T_1/T_2) = \lim_{\varepsilon \to 0} \partial_z \Phi(f/T_2)$.

$k = 2$ proof continued: *Linearize* the product

- First target: compute $\lim_{\varepsilon \to 0} g_1 = \partial_z \Phi(f/T_2)$.

$k = 2$ proof continued: *Linearize* the product

- First target: compute $\lim_{\varepsilon \to 0} g_1 = \partial_z \Phi(f/T_2)$.
- Logarithmic derivative: $\mathrm{dlog}_z(h) := \partial_z(h)/h$.

$k = 2$ proof continued: *Linearize* the product

- First target: compute $\lim_{\varepsilon \to 0} g_1 = \partial_z \Phi(f/T_2)$.
- 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: *Linearize* the product

- First target: compute $\lim_{\varepsilon \to 0} g_1 = \partial_z \Phi(f/T_2)$.
- 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:

$$
\begin{aligned}
\partial_z \Phi(T_1/T_2) &= \Phi(T_1/T_2) \cdot \mathrm{dlog}\Phi(T_1/T_2) \\
&= \Phi(T_1/T_2) \cdot (\mathrm{dlog}\Phi(T_1) - \mathrm{dlog}\Phi(T_2)) \ .
\end{aligned}
$$

$k = 2$ proof continued: *Linearize* the product

- First target: compute $\lim_{\varepsilon \to 0} g_1 = \partial_z \Phi(f/T_2)$.
- 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:

$$\begin{aligned}\partial_z \Phi(T_1/T_2) &= \Phi(T_1/T_2) \cdot \mathrm{dlog}\Phi(T_1/T_2) \\ &= \Phi(T_1/T_2) \cdot (\mathrm{dlog}\Phi(T_1) - \mathrm{dlog}\Phi(T_2)) \ .\end{aligned}$$

- Both $\Phi(T_1)$ and $\Phi(T_2)$ have $\Pi\Sigma$ circuits (they have z and ε).

$k = 2$ proof: Why does dlog deborder?

- What is $\mathsf{dlog}(\ell)$?

$k = 2$ proof: Why does dlog deborder?

- What is $\mathrm{dlog}(\ell)$? Note, $\ell =: A - z \cdot B$, where $A \in \mathbb{F}(\varepsilon)^*$, $B \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$.

$k = 2$ proof: Why does dlog deborder?

- What is $\mathrm{dlog}(\ell)$? Note, $\ell =: A - z \cdot B$, where $A \in \mathbb{F}(\varepsilon)^*$, $B \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$.

$$\mathrm{dlog}(A - zB) = -\frac{B}{A\,(1 - z \cdot B/A)}$$

$$= -\frac{B}{A} \cdot \sum_{j=0}^{d-1} \left(\frac{z \cdot B}{A}\right)^j$$

$$\in \ \Sigma \wedge \Sigma \ .$$ [*Magic trick*]

$k = 2$ proof: Why does dlog deborder?

- What is $\mathrm{dlog}(\ell)$? Note, $\ell =: A - z \cdot B$, where $A \in \mathbb{F}(\varepsilon)^*$, $B \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$.

$$\begin{aligned}
\mathrm{dlog}(A - zB) &= -\frac{B}{A\left(1 - z \cdot B/A\right)} \\
&= -\frac{B}{A} \cdot \sum_{j=0}^{d-1} \left(\frac{z \cdot B}{A}\right)^j \\
&\in \Sigma \wedge \Sigma \,. \quad [\textit{Magic trick}]
\end{aligned}$$

- Thus,

$k = 2$ proof: Why does dlog deborder?

- What is $\mathrm{dlog}(\ell)$? Note, $\ell =: A - z \cdot B$, where $A \in \mathbb{F}(\varepsilon)^*$, $B \in \mathbb{F}(\varepsilon)[\boldsymbol{x}]$.

$$\begin{aligned} \mathrm{dlog}(A - zB) &= -\frac{B}{A\,(1 - z \cdot B/A)} \\ &= -\frac{B}{A} \cdot \sum_{j=0}^{d-1} \left(\frac{z \cdot B}{A}\right)^j \\ &\in \Sigma \wedge \Sigma\,. \quad [\textit{Magic trick}] \end{aligned}$$

- Thus,

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

Finishing the proof– Induct and Limit

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

Finishing the proof– Induct and Limit

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

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

Finishing the proof– Induct and Limit

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

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

- Integrate g_1 (i.e. interpolate $\partial_z \Phi(T_1/T_2)$ wrt z), eliminate division, to get $\Phi(f)/(\lim_{\varepsilon \to 0} \Phi(T_2)) = \mathsf{ABP} \implies \Phi(f) = \mathsf{ABP} \implies f = \mathsf{ABP}$.

Finishing the proof– Induct and Limit

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

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

- Integrate g_1 (i.e. interpolate $\partial_z \Phi(T_1/T_2)$ wrt z), eliminate division, to get $\Phi(f)/(\lim_{\varepsilon \to 0} \Phi(T_2)) = \mathsf{ABP} \implies \Phi(f) = \mathsf{ABP} \implies f = \mathsf{ABP}$.
- Note: *Definite* integration requires setting $z = 0$ in $\Phi(T_1/T_2) + 1$; that's why we need power-series in z.

Proving Lower Bounds

Looking for finer lower bounds

▸▸ skip the section

Looking for finer lower bounds

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?

Looking for finer lower bounds

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?

Looking for finer lower bounds

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This (impossibility) is already known in the classical setting!

Looking for finer lower bounds

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This (impossibility) is already known 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

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This (impossibility) is already known 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 in *border*, 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

skip the section

- Can we show an *exponential* gap between $\overline{\Sigma^{[2]}\Pi\Sigma}$ and VBP?
- Ambitious goal: Can we separate $\overline{\Sigma^{[k]}\Pi\Sigma}$ and $\overline{\Sigma^{[k+1]}\Pi\Sigma}$?
- Note: This (impossibility) is already known 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 in *border*, 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!)?

Our results

[Dutta-Saxena FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;

Our results

[Dutta-Saxena FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, f **requires** $2^{\Omega(n)}$**-size** $\overline{\Sigma^{[k]}\Pi\Sigma}$ **circuits**.

Our results

[Dutta-Saxena FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, 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 FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, 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 FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, 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 FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, 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 FOCS'22]

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 $\Sigma^{[k+1]}\Pi\Sigma$ circuit of size $O(n)$;
but, 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, border is about *optimality*.

Conclusion

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).
- Can we show $\overline{\Sigma^{[k]}\Pi\Sigma} \subseteq \Sigma\Pi\Sigma$ (resp. VF)?

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).
- Can we show $\overline{\Sigma^{[k]}\Pi\Sigma} \subseteq \Sigma\Pi\Sigma$ (resp. VF)?
- What is the border of the *sum of two* products of univariate matrices (ROABPs)?

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).
- Can we show $\overline{\Sigma^{[k]}\Pi\Sigma} \subseteq \Sigma\Pi\Sigma$ (resp. VF)?
- What is the border of the *sum of two* products of univariate matrices (ROABPs)?
- Is the border of VP explicit; i.e. in VNP?

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).
- Can we show $\overline{\Sigma^{[k]}\Pi\Sigma} \subseteq \Sigma\Pi\Sigma$ (resp. VF)?
- What is the border of the *sum of two* products of univariate matrices (ROABPs)?
- Is the border of VP explicit; i.e. in VNP?
- Are degree-s factors of a size-s polynomial explicit?

Concluding remarks

- ROABP core gives us many PIT results (see our two papers).
- Can we show $\overline{\Sigma^{[k]}\Pi\Sigma} \subseteq \Sigma\Pi\Sigma$ (resp. VF)?
- What is the border of the *sum of two* products of univariate matrices (ROABPs)?
- Is the border of VP explicit; i.e. in VNP?
- Are degree-s factors of a size-s polynomial explicit?

Thank you! Questions?