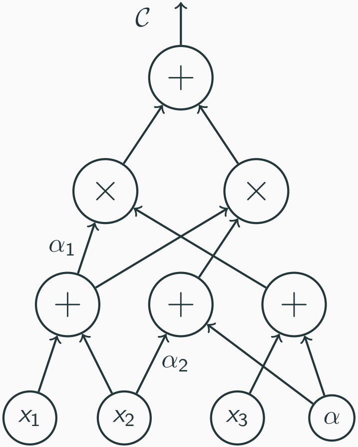
A Quadratic Lower Bound against Homogeneous Non-Commutative Circuits

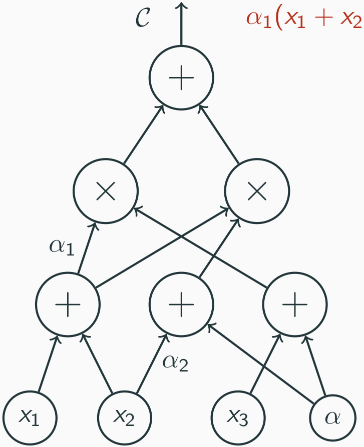
Prerona Chatterjee [joint work with Pavel Hrubeš (Institute of Mathematics, CAS)]
Tel Aviv University

March, 31, 2023

Algebraic Circuits

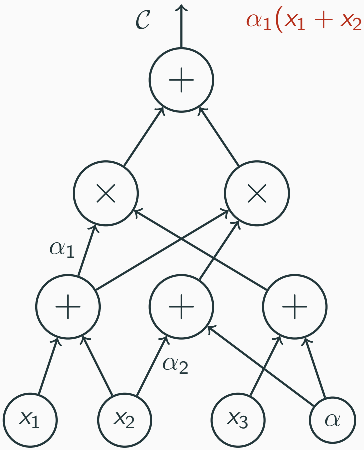


Algebraic Circuits



$\alpha_1(x_1 + x_2)(x_3 + \alpha) + (x_1 + x_2)(\alpha_2 x_2 + \alpha)$

Algebraic Circuits

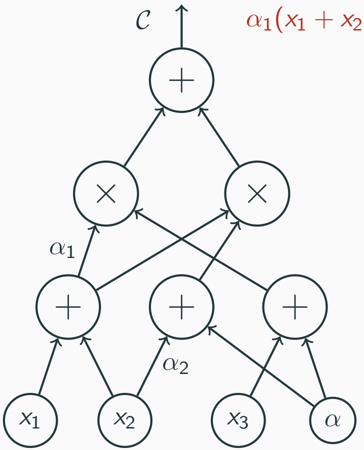


$\alpha_1(x_1 + x_2)(x_3 + \alpha) + (x_1 + x_2)(\alpha_2 x_2 + \alpha)$

Objects of Study

Polynomials over n variables of degree d.

Algebraic Circuits



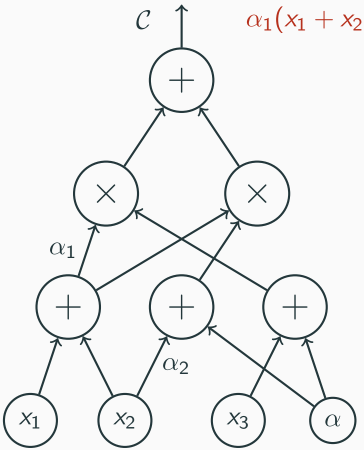
$\alpha_1(x_1 + x_2)(x_3 + \alpha) + (x_1 + x_2)(\alpha_2 x_2 + \alpha)$

Objects of Study

Polynomials over n variables of degree d.

Central Question: Find explicit polynomials that cannot be computed by efficient circuits.

Algebraic Circuits



$\alpha_1(x_1 + x_2)(x_3 + \alpha) + (x_1 + x_2)(\alpha_2 x_2 + \alpha)$

Objects of Study

Polynomials over n variables of degree d.

Central Question: Find explicit polynomials that cannot be computed by efficient circuits.

[Baur-Strassen]: Any algebraic circuit computing $\sum_{i=1}^{n} x_i^d$ requires $\Omega(n \log d)$ wires.

The Non-Commutative Setting

$$f(x, y) = (x + y) \times (x + y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

The Non-Commutative Setting

$$f(x,y) = (x+y) \times (x+y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

Non-Commutative Circuits: The multiplication gates, additionally, respect the order.

The Non-Commutative Setting

$$f(x,y) = (x+y)\times(x+y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

Non-Commutative Circuits: The multiplication gates, additionally, respect the order.

Can we do something better in this setting?

The Non-Commutative Setting

$$f(x,y) = (x+y)\times(x+y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

Non-Commutative Circuits: The multiplication gates, additionally, respect the order.

Can we do something better in this setting?

[Nisan]

$\mathrm{VBP_{nc}} \subsetneq \mathrm{VP_{nc}}$

The Non-Commutative Setting

$$f(x, y) = (x + y) \times (x + y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

Non-Commutative Circuits: The multiplication gates, additionally, respect the order.

Can we do something better in this setting?

[Nisan]	**[Tavenas-Limaye-Srinivasan]**
$\mathrm{VBP}_{\mathrm{nc}} \subsetneq \mathrm{VP}_{\mathrm{nc}}$	$\mathrm{VF}_{\text{nc, hom}} \subsetneq \mathrm{VBP}_{\text{nc, hom}}$.

The Non-Commutative Setting

$$f(x, y) = (x + y) \times (x + y) = x^2 + xy + yx + y^2 \neq x^2 + 2xy + y^2$$

Non-Commutative Circuits: The multiplication gates, additionally, respect the order.

Can we do something better in this setting?

[Nisan]

$\mathrm{VBP}_{\mathrm{nc}} \subsetneq \mathrm{VP}_{\mathrm{nc}}$

[Tavenas-Limaye-Srinivasan]

$\mathrm{VF}_{\mathrm{nc,\ hom}} \subsetneq \mathrm{VBP}_{\mathrm{nc,\ hom}}$.

[Carmossino-Impagliazzo-Lovett-Mihajlin]

$\Omega(n^{\omega+\varepsilon})$ for $f_{n,c} \implies \Omega(2^n)$ for $f'_{n,n}$.

Our Result

The best lower bound against NC circuits continues to be $\Omega(n \log d)$.

Our Result

The best lower bound against NC circuits continues to be $\Omega(n \log d)$.

Can we do better at least in the homogeneous case?

Our Result

The best lower bound against NC circuits continues to be $\Omega(n \log d)$.

Can we do better at least in the homogeneous case?

Theorem: Any homogeneous non-commutative circuit computing

$$\mathrm{OSym}_{n,d}(\mathbf{x}) = \sum_{1 \le i_1 < \cdots < i_d \le n} x_{i_1} \cdots x_{i_d}$$

has size $\Omega(nd)$ for $d \le \frac{n}{2}$.

Our Result

The best lower bound against NC circuits continues to be $\Omega(n \log d)$.

Can we do better at least in the homogeneous case?

Theorem: Any homogeneous non-commutative circuit computing

$$\mathrm{OSym}_{n,d}(\mathbf{x}) = \sum_{1 \leq i_1 < \cdots < i_d \leq n} x_{i_1} \cdots x_{i_d}$$

has size $\Omega(nd)$ for $d \leq \frac{n}{2}$.The lower bound is tight for homogeneous non-commutative circuits.

Our Result

The best lower bound against NC circuits continues to be $\Omega(n \log d)$.

Can we do better at least in the homogeneous case?

Theorem: Any homogeneous non-commutative circuit computing

$$\mathrm{OSym}_{n,d}(\mathbf{x}) = \sum_{1 \leq i_1 < \cdots < i_d \leq n} x_{i_1} \cdots x_{i_d}$$

has size $\Omega(nd)$ for $d \leq \frac{n}{2}$.The lower bound is tight for homogeneous non-commutative circuits.

Further, there is a non-commutative circuit of size $O(n \log^2 n)$ that computes $\mathrm{OSym}_{n,n/2}(\mathbf{x})$.

Our Measure

f: Hom. non-commutative polynomial of degree d.

Our Measure

f: Hom. non-commutative polynomial of degree d.

$f^{(i)}$: Polynomial got from f by setting variables in positions other than i, $i+1$ to 1.

Our Measure

f: Hom. non-commutative polynomial of degree d.

$f^{(i)}$: Polynomial got from f by setting variables in positions other than i, $i+1$ to 1.

Example: $f = x_1 \cdots x_d + x_d \cdots x_1$

Our Measure

f: Hom. non-commutative polynomial of degree d.

$f^{(i)}$: Polynomial got from f by setting variables in positions other than i, $i+1$ to 1.

Example: $f = x_1 \cdots x_d + x_d \cdots x_1 \implies f^{(1)} = x_1 x_2 + x_d x_{d-1}$.

Our Measure

f: Hom. non-commutative polynomial of degree d.

$f^{(i)}$: Polynomial got from f by setting variables in positions other than i, $i+1$ to 1.

Example: $f = x_1 \cdots x_d + x_d \cdots x_1 \implies f^{(1)} = x_1 x_2 + x_d x_{d-1}$.

$$\mu(f) = \text{rank}\left(\text{span}_{\mathbb{F}}\left(\left\{f^{(0)}, f^{(1)}, \ldots, f^{(d)}\right\}\right)\right).$$

Our Measure

f: Hom. non-commutative polynomial of degree d.

$f^{(i)}$: Polynomial got from f by setting variables in positions other than i, $i+1$ to 1.

Example: $f = x_1 \cdots x_d + x_d \cdots x_1 \implies f^{(1)} = x_1 x_2 + x_d x_{d-1}$.

$$\mu(f) = \text{rank}\left(\text{span}_{\mathbb{F}}\left(\left\{f^{(0)}, f^{(1)}, \ldots, f^{(d)}\right\}\right)\right).$$

Main Observation: For any f that is computable by a homogeneous non-commutative circuit of size s,

$$\mu(f) \leq s + 1.$$

A simple proof of an obvious fact

$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \mathsf{rank}\left(\mathsf{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

A simple proof of an obvious fact

$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \mathsf{rank}\left(\mathsf{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

$$\boxed{\mu(\mathcal{C}) \leq \mathsf{size}(\mathcal{C}) + 1}$$

A simple proof of an obvious fact

$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \operatorname{rank}\left(\operatorname{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

$$\boxed{\mu(\mathcal{C}) \leq \operatorname{size}(\mathcal{C}) + 1}$$

Idea: Use induction

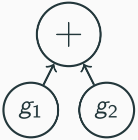
A simple proof of an obvious fact

$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \text{rank}\left(\text{span}_{\mathbb{F}}\left(\bigcup_{g \in \mathcal{C}} \left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

$$\boxed{\mu(\mathcal{C}) \leq \text{size}(\mathcal{C}) + 1}$$

Idea: Use induction



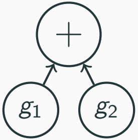
A simple proof of an obvious fact

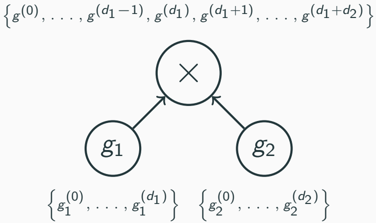
$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \mathsf{rank}\left(\mathsf{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

$$\boxed{\mu(\mathcal{C}) \leq \mathsf{size}(\mathcal{C}) + 1}$$

Idea: Use induction





A simple proof of an obvious fact

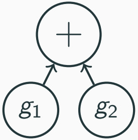
$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \operatorname{rank}\left(\operatorname{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

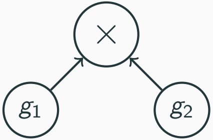
$$\boxed{\mu(\mathcal{C}) \leq \operatorname{size}(\mathcal{C}) + 1}$$

$$\boxed{\mu(f_{\mathcal{C}}) \leq \mu(\mathcal{C})}$$

Idea: Use induction



$$\left\{g^{(0)}, \ldots, g^{(d_1-1)}, g^{(d_1)}, g^{(d_1+1)}, \ldots, g^{(d_1+d_2)}\right\}$$



$$\left\{g_1^{(0)}, \ldots, g_1^{(d_1)}\right\} \quad \left\{g_2^{(0)}, \ldots, g_2^{(d_2)}\right\}$$

A simple proof of an obvious fact

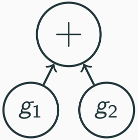
$\mathcal{C}$: Homogeneous non-commutative circuit.

$$\mu(\mathcal{C}) = \operatorname{rank}\left(\operatorname{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

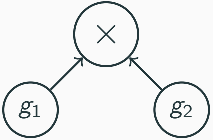
$$\boxed{\mu(\mathcal{C}) \leq \operatorname{size}(\mathcal{C}) + 1}$$

$$\boxed{\mu(f_{\mathcal{C}}) \leq \mu(\mathcal{C})}$$

Idea: Use induction



$$\left\{g^{(0)}, \ldots, g^{(d_1-1)}, g^{(d_1)}, g^{(d_1+1)}, \ldots, g^{(d_1+d_2)}\right\}$$



$$\left\{g_1^{(0)}, \ldots, g_1^{(d_1)}\right\} \quad \left\{g_2^{(0)}, \ldots, g_2^{(d_2)}\right\}$$

$$f = x_1 \cdots x_n$$

$$\Downarrow$$

$$\mu(f) = n + 1.$$

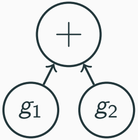
A simple proof of an obvious fact

$\mathcal{C}$: Homogeneous non-commutative circuit.

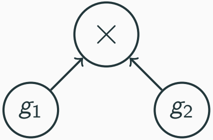
$$\mu(\mathcal{C}) = \text{rank}\left(\text{span}_{\mathbb{F}}\left(\bigcup_{g\in\mathcal{C}}\left\{g^{(0)}, g^{(1)}, \ldots, g^{(d)}\right\}\right)\right).$$

$$\boxed{\mu(\mathcal{C}) \le \text{size}(\mathcal{C}) + 1}$$

Idea: Use induction



$\left\{g^{(0)}, \ldots, g^{(d_1-1)}, g^{(d_1)}, g^{(d_1+1)}, \ldots, g^{(d_1+d_2)}\right\}$



$\left\{g_1^{(0)}, \ldots, g_1^{(d_1)}\right\}$ $\left\{g_2^{(0)}, \ldots, g_2^{(d_2)}\right\}$

$$\boxed{\mu(f_{\mathcal{C}}) \le \mu(\mathcal{C})}$$

$$f = x_1 \cdots x_n$$

$$\Downarrow$$

$$\mu(f) = n + 1.$$

Therefore, $\mu(\mathcal{C}_f) \ge n$.

Using it to prove a “not so obvious” fact

Theorem: There exists an explicit monomial over $\{x_0, x_1\}$ of degree d such that any homogeneous non-commutative circuit computing it must have size $\Omega\left(\frac{d}{\log d}\right)$.

Using it to prove a "not so obvious" fact

Theorem: There exists an explicit monomial over $\{x_0, x_1\}$ of degree d such that any homogeneous non-commutative circuit computing it must have size $\Omega\left(\frac{d}{\log d}\right)$.

The tweak: For a homogeneous non-commutative polynomial f of degree d, define

$f^{(i)}$ by setting, in f, variables in positions other than $\{i, i+1, \ldots i + \log d\}$ to 1.

Using it to prove a "not so obvious" fact

Theorem: There exists an explicit monomial over $\{x_0, x_1\}$ of degree d such that any homogeneous non-commutative circuit computing it must have size $\Omega\left(\frac{d}{\log d}\right)$.

The tweak: For a homogeneous non-commutative polynomial f of degree d, define

$f^{(i)}$ by setting, in f, variables in positions other than $\{i, i+1, \ldots i + \log d\}$ to 1.

In this case, if $\mathcal{C}$ is a homogeneous non-commutative circuit of size s, then $\mu_\ell(\mathcal{C}) \leq O(s \log d)$.

Using it to prove a "not so obvious" fact

Theorem: There exists an explicit monomial over $\{x_0, x_1\}$ of degree d such that any homogeneous non-commutative circuit computing it must have size $\Omega\left(\frac{d}{\log d}\right)$.

The tweak: For a homogeneous non-commutative polynomial f of degree d, define

$$f^{(i)} \text{ by setting, in } f\text{, variables in positions other than } \{i, i+1, \ldots i+\log d\} \text{ to } 1.$$

In this case, if $\mathcal{C}$ is a homogeneous non-commutative circuit of size s, then $\mu_\ell(\mathcal{C}) \leq O(s \log d)$.

Therefore all we need is a monomial, f, over $\{x_0, x_1\}$ of degree d such that $\mu_\ell(f) \geq \Omega(d)$.

A monomial with high measure

de Bruijn Sequence (of order $\log d$**)**: It is a cyclic sequence in the alphabet $\{0, 1\}$ in which every string of length $\log d$, occurs exactly once as a substring.

A monomial with high measure

de Bruijn Sequence (of order $\log d$): It is a cyclic sequence in the alphabet $\{0, 1\}$ in which every string of length $\log d$, occurs exactly once as a substring.

Fact: There is a length-d de Bruijn sequence of order $\log d$.

A monomial with high measure

de Bruijn Sequence (of order $\log d$): It is a cyclic sequence in the alphabet $\{0, 1\}$ in which every string of length $\log d$, occurs exactly once as a substring.

Fact: There is a length-d de Bruijn sequence of order $\log d$.

Therefore, if B_d is the monomial corresponding to this de Bruijn sequence, then $\mu(B_d) \geq \Omega(d)$.

A monomial with high measure

de Bruijn Sequence (of order $\log d$**)**: It is a cyclic sequence in the alphabet $\{0, 1\}$ in which every string of length $\log d$, occurs exactly once as a substring.

Fact: There is a length-d de Bruijn sequence of order $\log d$.

Therefore, if B_d is the monomial corresponding to this de Bruijn sequence, then $\mu(B_d) \geq \Omega(d)$.

How can non-homogeneity possibly help in computing a monomial?

A monomial with high measure

de Bruijn Sequence (of order $\log d$**)**: It is a cyclic sequence in the alphabet $\{0, 1\}$ in which every string of length $\log d$, occurs exactly once as a substring.

Fact: There is a length-d de Bruijn sequence of order $\log d$.

Therefore, if B_d is the monomial corresponding to this de Bruijn sequence, then $\mu(B_d) \geq \Omega(d)$.

How can non-homogeneity possibly help in computing a monomial?

Question: Can we prove the same lower bound against general non-commutative circuits?

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- Suppose a similar result was true in the homogeneous non-commutative setting.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- Suppose a similar result was true in the homogeneous non-commutative setting.
- Suppose there is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- Suppose a similar result was true in the homogeneous non-commutative setting.
- Suppose there is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Then we would have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- A similar result is true in the homogeneous non-commutative setting.
- Suppose there is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Then we would have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- A similar result is true in the homogeneous non-commutative setting.
- Suppose there is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Then we would have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Note: $f = x_1 B_d(x_0^{(1)}, x_1^{(1)}) + \cdots + x_n B_d(x_0^{(n)}, x_1^{(n)})$ already (almost) has the required property.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- A similar result is true in the homogeneous non-commutative setting.
- There is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Then we would have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- A similar result is true in the homogeneous non-commutative setting.
- There is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Therefore we have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Getting back to the main result

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

- A similar result is true in the homogeneous non-commutative setting.
- There is an n-variate, degree-d polynomial f such that

$$\mu(\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}) \geq \Omega(nd).$$

Therefore we have an $\Omega(nd)$ lower bound against homogeneous non-commutative circuits.

Note: f has a non-homogeneous non-commutative circuit of size $O(n \log^2 d)$.

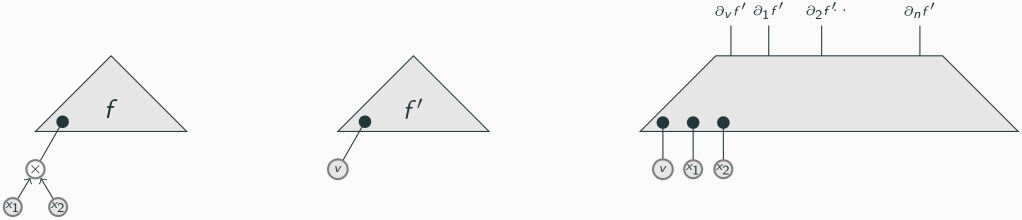
Proof of [Baur-Strassen]

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Proof of [Baur-Strassen]

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

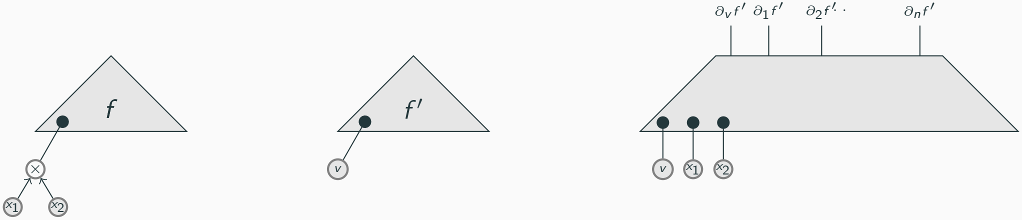
Step 1:



Proof of [Baur-Strassen]

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Step 1:

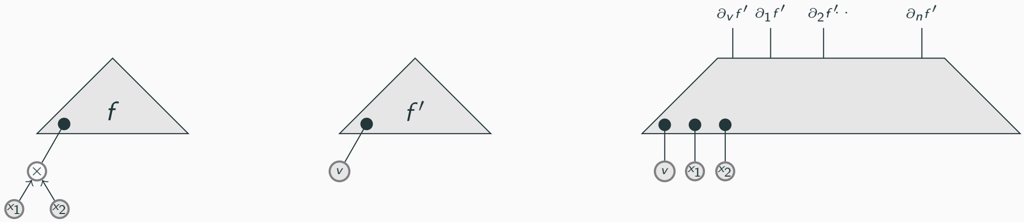


Step 2: Write each of $\{\partial_i f\}_i$ using $\partial_v f'$ and $\{\partial_i f'\}_i$.

Proof of [Baur-Strassen]

[Baur-Strassen]: If there is a circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Step 1:



Step 2: Write each of $\{\partial_i f\}_i$ using $\partial_v f'$ and $\{\partial_i f'\}_i$. Add (the ≤ 10 extra) edges accordingly.

Making [Baur-Strassen] work in the homogeneous setting

Target: If there is a homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Making [Baur-Strassen] work in the homogeneous setting

Target: If there is a homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Weights: $w_i = \mathrm{wt}(x_i)$.

Making [Baur-Strassen] work in the homogeneous setting

Target: If there is a homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Weights: $w_i = \mathrm{wt}(x_i)$.

Given $\mathbf{w} = (w_1, \ldots, w_n)$, define $\mathbf{w}$-homogeneous.

Making [Baur-Strassen] work in the homogeneous setting

Target: If there is a homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Weights: $w_i = \mathrm{wt}(x_i)$. Given $\mathbf{w} = (w_1, \ldots, w_n)$, define $\mathbf{w}$-homogeneous.

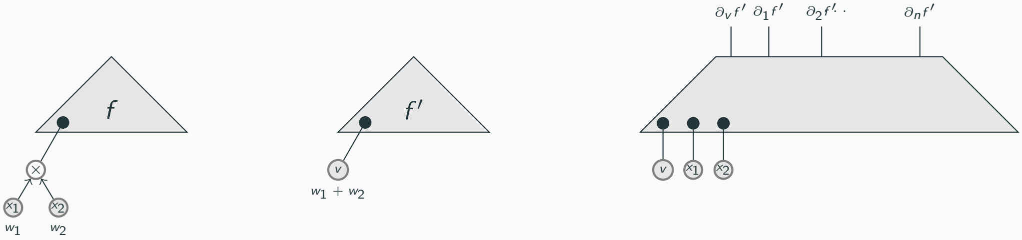
Lemma: If there is a $\mathbf{w}$-homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a $\mathbf{w}$-homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Making [Baur-Strassen] work in the homogeneous setting

Target: If there is a homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.

Weights: $w_i = \mathrm{wt}(x_i)$. Given $\mathbf{w} = (w_1, \ldots, w_n)$, define $\mathbf{w}$-homogeneous.

Lemma: If there is a $\mathbf{w}$-homogeneous circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a $\mathbf{w}$-homogeneous circuit of size at most $5s$ that simultaneously compute $\{\partial_{x_1} f, \partial_{x_2} f, \ldots, \partial_{x_n} f\}$.



Making [Baur-Strassen] work in the non-commutative setting

Formal derivatives (with respect to the first position)

Given a polynomial f and a variable x, f can be uniquely written as

$$f = x \cdot f_0 + f_1$$

where no monomial in f_1 contains x in the first position.

Making [Baur-Strassen] work in the non-commutative setting

Formal derivatives (with respect to the first position)

Given a polynomial f and a variable x, f can be uniquely written as

$$f = x \cdot f_0 + f_1$$

where no monomial in f_1 contains x in the first position.

We can then define the formal derivative to be $\partial_{1,x} f := f_0$.

Making [Baur-Strassen] work in the non-commutative setting

Formal derivatives (with respect to the first position)

Given a polynomial f and a variable x, f can be uniquely written as

$$f = x \cdot f_0 + f_1$$

where no monomial in f_1 contains x in the first position.

We can then define the formal derivative to be $\partial_{1,x} f := f_0$.

Chain rules can be defined formally as well.

Making [Baur-Strassen] work in the non-commutative setting

Formal derivatives (with respect to the first position)

Given a polynomial f and a variable x, f can be uniquely written as

$$f = x \cdot f_0 + f_1$$

where no monomial in f_1 contains x in the first position.

We can then define the formal derivative to be $\partial_{1,x} f := f_0$.

Chain rules can be defined formally as well.

Lemma: If there is a homogeneous NC circuit of size s computing $f \in \mathbb{F}[\mathbf{x}]$, then there is a homogeneous NC circuit of size at most $5s$ that simultaneously compute $\{\partial_{1,x_1} f, \ldots, \partial_{1,x_n} f\}$.

Polynomial with a large measure

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1 \le i_1 < \cdots < i_{\frac{n}{2}+1} \le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j} \right)}$$

Polynomial with a large measure

$x_{\frac{n}{2}+1}x_{\frac{n}{2}+2} \quad \cdots \quad x_2x_{\frac{n}{2}+2} \quad \cdots \quad \cdots \quad x_{n-2}x_{n-1} \quad \cdots \quad x_{\frac{n}{2}-1}x_{n-1} \quad x_{n-1}x_n \quad \cdots \quad x_{\frac{n}{2}}x_n$

$(1, \frac{n}{2})$

$\vdots$

$(1, 1)$

$\vdots$

$\vdots$

$(\frac{n}{2}-2, \frac{n}{2})$

$\vdots$

$(\frac{n}{2}-2, 1)$

$(\frac{n}{2}-1, \frac{n}{2})$

$\vdots$

$(\frac{n}{2}-1, 1)$

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1 \le i_1 < \cdots < i_{\frac{n}{2}+1} \le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j} \right)}$$

Polynomial with a large measure

$x_{\frac{n}{2}+1}x_{\frac{n}{2}+2} \quad \cdots \quad x_2x_{\frac{n}{2}+2} \quad \cdots \quad \cdots \quad x_{n-2}x_{n-1} \quad \cdots \quad x_{\frac{n}{2}-1}x_{n-1} \quad x_{n-1}x_n \quad \cdots \quad x_{\frac{n}{2}}x_n$

$(1, \frac{n}{2})$

$\vdots$

$(1, 1)$

$\vdots$

$\vdots$

$(\frac{n}{2} - 2, \frac{n}{2})$

$\vdots$

$(\frac{n}{2} - 2, 1)$

$(\frac{n}{2} - 1, \frac{n}{2})$

$\vdots$

$(\frac{n}{2} - 1, 1)$

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1\le i_1<\cdots<i_{\frac{n}{2}+1}\le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j}\right)}$$

Defining the matrix $\mathcal{M}(f)$

$x_k x_l$

(j, i) $\quad \boxed{\mathrm{coeff}_{x_k x_l}(\partial_i f^{(j)})}$

Polynomial with a large measure

$x_{\frac{n}{2}+1}x_{\frac{n}{2}+2} \quad \cdots \quad x_2x_{\frac{n}{2}+2} \quad \cdots \quad \cdots \quad x_{n-2}x_{n-1} \quad \cdots \quad x_{\frac{n}{2}-1}x_{n-1} \quad x_{n-1}x_n \quad \cdots \quad x_{\frac{n}{2}}x_n$

$(1, \frac{n}{2})$

$\vdots$

$(1, 1)$

$\vdots$

$\vdots$

$(\frac{n}{2}-2, \frac{n}{2})$

$\vdots$

$(\frac{n}{2}-2, 1)$

$(\frac{n}{2}-1, \frac{n}{2})$

$\vdots$

$(\frac{n}{2}-1, 1)$

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1\le i_1<\cdots<i_{\frac{n}{2}+1}\le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j}\right)}$$

Defining the matrix $\mathcal{M}(f)$

x_kx_l

$(j, i) \quad \boxed{\mathsf{coeff}_{x_kx_l}(\partial_i f^{(j)})}$

$\mu(\mathbb{D}(f)) \geq \mathsf{rank}(\mathcal{M}(f))$

Polynomial with a large measure

$x_{\frac{n}{2}+1}x_{\frac{n}{2}+2} \quad \cdots \quad x_2x_{\frac{n}{2}+2} \quad \cdots \quad \cdots \quad x_{n-2}x_{n-1} \quad \cdots \quad x_{\frac{n}{2}-1}x_{n-1} \quad x_{n-1}x_n \quad \cdots \quad x_{\frac{n}{2}}x_n$

$(1, \frac{n}{2})$
$\vdots$
$(1, 1)$
$\vdots$
$\vdots$
$(\frac{n}{2}-2, \frac{n}{2})$
$\vdots$
$(\frac{n}{2}-2, 1)$
$(\frac{n}{2}-1, \frac{n}{2})$
$\vdots$
$(\frac{n}{2}-1, 1)$

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1 \le i_1 < \cdots < i_{\frac{n}{2}+1} \le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j} \right)}$$

Defining the matrix $\mathcal{M}(f)$

$x_k x_l$

$(j, i) \qquad \boxed{\mathsf{coeff}_{x_k x_l}(\partial_i f^{(j)})}$

This matrix is lower triangular with 1s on the diagonal.

$$\mu(\mathbb{D}(f)) \ge \mathsf{rank}(\mathcal{M}(f))$$

Polynomial with a large measure

$x_{\frac{n}{2}+1}x_{\frac{n}{2}+2} \quad \cdots \quad x_2x_{\frac{n}{2}+2} \quad \cdots \quad \cdots \quad x_{n-2}x_{n-1} \quad \cdots \quad x_{\frac{n}{2}-1}x_{n-1} \quad x_{n-1}x_n \quad \cdots \quad x_{\frac{n}{2}}x_n$

$(1, \frac{n}{2})$
$\vdots$
$(1, 1)$
$\vdots$
$\vdots$
$(\frac{n}{2}-2, \frac{n}{2})$
$\vdots$
$(\frac{n}{2}-2, 1)$
$(\frac{n}{2}-1, \frac{n}{2})$
$\vdots$
$(\frac{n}{2}-1, 1)$

$$\boxed{f = \mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x}) = \sum_{1 \le i_1 < \cdots < i_{\frac{n}{2}+1} \le n} \left(\prod_{j=1}^{\frac{n}{2}+1} x_{i_j} \right)}$$

Defining the matrix $\mathcal{M}(f)$

$x_k x_l$

$(j, i) \qquad \boxed{\mathsf{coeff}_{x_k x_l}(\partial_i f^{(j)})}$

This matrix is lower triangular with 1s on the diagonal.

$$\mu(\mathbb{D}(f)) \ge \mathsf{rank}(\mathcal{M}(f)) = \Omega(n^2).$$

The lower bound is tight

There is a homogeneous non-commutative circuit of size $O(n^2)$ that computes $\mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x})$.

The lower bound is tight

There is a homogeneous non-commutative circuit of size $O(n^2)$ that computes $\mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x})$.

How?

The lower bound is tight

There is a homogeneous non-commutative circuit of size $O(n^2)$ that computes $\mathrm{OSym}_{n,\frac{n}{2}+1}(\mathbf{x})$.

How?

Use the following fact recursively.

$$\mathrm{OSym}_{n,d}(x_1,\ldots,x_n) = \mathrm{OSym}_{n-1,d-1}(x_1,\ldots,x_{n-1}) \cdot x_n + \mathrm{OSym}_{n-1,d}(x_1,\ldots,x_{n-1}).$$

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

$\mathrm{OSym}_{n,d}(x_1, \ldots, x_n) = \mathsf{coeff}_{t^d} \left(\prod_{i=1}^{n} (1 + t x_i) \right)$

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n\log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

$$\mathrm{OSym}_{n,d}(x_1,\ldots,x_n) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{n}(1+tx_i)\right) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{\frac{n}{2}}(1+tx_i)\cdot\prod_{i=\frac{n}{2}+1}^{n}(1+tx_i)\right).$$

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

$$\mathrm{OSym}_{n,d}(x_1, \ldots, x_n) = \mathrm{coeff}_{t^d}\left(\prod_{i=1}^{n}(1 + tx_i)\right) = \mathrm{coeff}_{t^d}\left(\prod_{i=1}^{\frac{n}{2}}(1 + tx_i) \cdot \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i)\right).$$

Think of $\quad f = \prod_{i=1}^{\frac{n}{2}}(1 + tx_i), g = \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i) \in \mathbb{F}\langle \mathbf{x} \rangle [t].$

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

$$\mathrm{OSym}_{n,d}(x_1, \ldots, x_n) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{n}(1 + tx_i)\right) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{\frac{n}{2}}(1 + tx_i) \cdot \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i)\right).$$

Think of $\qquad f = \prod_{i=1}^{\frac{n}{2}}(1 + tx_i), g = \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i) \in \mathbb{F}\langle \mathbf{x} \rangle [t].$

Do polynomial multiplication recursively $\log n$ times.

Better Upper bound in the non-homogeneous setting

There is a non-commutative circuit of size $O(n \log^2 n)$ that computes all the elementary symmetric polynomials simultaneously.

How?

$$\mathrm{OSym}_{n,d}(x_1, \ldots, x_n) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{n}(1 + tx_i)\right) = \mathsf{coeff}_{t^d}\left(\prod_{i=1}^{\frac{n}{2}}(1 + tx_i) \cdot \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i)\right).$$

Think of $\quad f = \prod_{i=1}^{\frac{n}{2}}(1 + tx_i), g = \prod_{i=\frac{n}{2}+1}^{n}(1 + tx_i) \in \mathbb{F}\langle \mathbf{x} \rangle [t].$

Do polynomial multiplication recursively $\log n$ times. Note that polynomial multiplication can be done in time $O(n \log n)$ using FFT.

Open Questions

- Can we show a $\tilde{\Omega}(d)$ lower bound against general non-commutative circuits?

Open Questions

- Can we show a $\tilde{\Omega}(d)$ lower bound against general non-commutative circuits?
- Can we show a quadratic lower bound for a constant variate polynomial?

Open Questions

- Can we show a $\tilde{\Omega}(d)$ lower bound against general non-commutative circuits?
- Can we show a quadratic lower bound for a constant variate polynomial?

Conjecture: If

$$f = x_1 x_0^{d-1} f_1 + x_0 x_1 x_0^{d-2} f_2 + \cdots + x_0^{d-1} x_1 f_d$$

can be computed by a non-commutative circuit of size s, then $\{f_1, \ldots, f_d\}$ can be simultaneously computed by a non-commutative circuit of size $O(s + d)$.

Open Questions

- Can we show a $\tilde{\Omega}(d)$ lower bound against general non-commutative circuits?
- Can we show a quadratic lower bound for a constant variate polynomial?

Conjecture: If

$$f = x_1 x_0^{d-1} f_1 + x_0 x_1 x_0^{d-2} f_2 + \cdots + x_0^{d-1} x_1 f_d$$

can be computed by a non-commutative circuit of size s, then $\{f_1, \ldots, f_d\}$ can be simultaneously computed by a non-commutative circuit of size $O(s + d)$.

If true, then the answer to the second question is "yes".

Thank you!