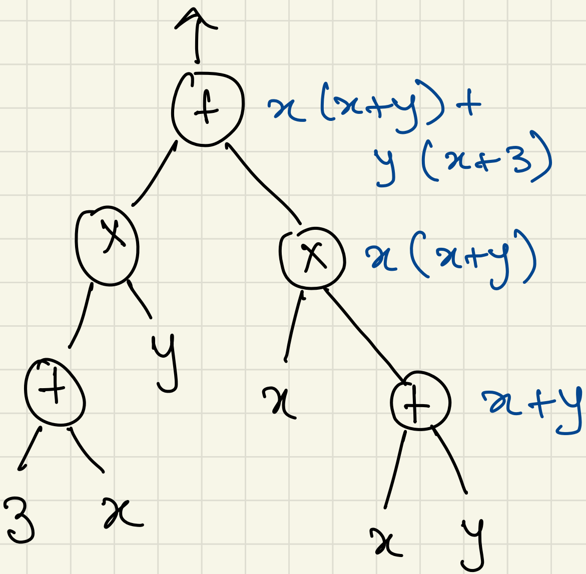
An Almost Quadratic Lower Bound Against Formulas for a Constant-Degree Polynomial.

MM Fest, IIT Gandhinagar.

December 15, 2024.

Algebraic Formula.



- Underlying graph is a tree.
- Fan-in: 2.
- Size: No. of leaves.

[Kalorkoti]: Any formula computing the n^2-variate determinant polynomial requires size $\Omega(n^3)$.

What we will see a full proof of

[Shpilka-Yehudayoff] : Any formula computing

$$\sum_{i=1}^{n} \sum_{j=1}^{n} y_j \, x_i^{\,j}$$ requires size $\Omega(n^2)$.

Can be modified to prove

For any $k \in \mathbb{N}$ with $k \geq 3$, there exists a degree k polynomial, such that any formula computing it requires size $\Omega(n^{2-1/k})$.

Algebraic Independence.

$f_1, \dots, f_k \in \mathbb{F}[x_1, \dots, x_n]$: Algebraically dependent if there exists a polynomial $A \in \mathbb{F}[y_1, \dots, y_k]$ such that $A(f_1, \dots, f_k) = 0$.

If no such A exists, then $\{f_1, \dots, f_k\}$ is said to be algebraically independent.

Algebraic Rank.

The algebraic rank of a set of polynomials $\{f_1, \dots, f_k\}$ is the size of the largest subset of $\{f_1, \dots, f_k\}$ that is algebraically independent.

The Measure.

Given a polynomial $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$, let $X_1, X_2, \ldots, X_k$ be a partition of the set of underlying variables. That is, $X_1 \sqcup \cdots \sqcup X_k = \{x_1, \ldots, x_n\}$.

For any $i \in [k]$, let $f = \sum_{\bar{x} \in X_i} f_{\bar{e}} \cdot \bar{x}^{\bar{e}}$.

Then we define $\Gamma_{X_i}(f) = \text{alg rank}(\{f_{\bar{e}}\})$ and

$$\Gamma(f) = \sum_{i=1}^{k} \Gamma_{X_i}(f).$$

Upper Bounding the Measure.

If a polynomial f is computable by a formula of size s, then $\Gamma(f) \leq s$.

Completing the Proof (assuming the upper bound)

Recall the polynomial: $\sum_{i=1}^{n} \sum_{j=1}^{n} y_j \, x_i^{\,j}$

Consider the partition: $X_1 \sqcup X_2 \sqcup \cdots \sqcup X_n \sqcup Y$ where

$\forall i \in [n]$, $X_i = \{x_i\}$ and $Y = \{y_1, \ldots, y_n\}$.

For any $i \in [n]$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} y_j \, x_i^{\,j} =$$

$$\left(\sum_{k \in [n] \setminus \{i\}} y_j \, x_k^{\,j} \right) + y_1 x_i + y_2 x_i^2 + \cdots + y_n x_i^n$$

and so $\Gamma_{x_i} \geqslant \Omega(n)$.

$$\Rightarrow \Gamma\left(\sum_{i=1}^{n} \sum_{j=1}^{n} y_j \, x_i^{\,j} \right) \geqslant \Omega(n^2)$$

An $\Omega(n^{3/2})$ lower bound for a degree - 3 polynomial.

Consider the polynomial $\sum_{k=1}^{\sqrt{n}} \sum_{i=0}^{\sqrt{n}-1} \sum_{j=0}^{\sqrt{n}-1} z_{i\sqrt{n}+j} \, x_{k,i} \, y_{k,j}$.

Also consider the partition

$$X_k = \{x_{k,0}, \dots, x_{k,\sqrt{n}-1}, y_{k,0}, \dots, y_{k,\sqrt{n}-1}\} \text{ for every } k \in [\sqrt{n}]$$

$$\text{and } Z = \{z_0, \dots, z_{n-1}\}.$$

No. of Variables : $\sqrt{n}\,(2\sqrt{n}) + n = 3n$

Degree $= 3$.

For any $k_0 \in [\sqrt{n}]$,

$$\sum_{k=1}^{\sqrt{n}} \sum_{i=0}^{\sqrt{n}-1} \sum_{j=0}^{\sqrt{n}-1} z_{i\sqrt{n}+j} \; x_{k,i} \, y_{k,j} =$$

$$\left(\sum_{l \in [\sqrt{n}] \setminus \{k_0\}} \sum_{i=0}^{\sqrt{n}-1} \sum_{j=0}^{\sqrt{n}-1} z_{i\sqrt{n}+j} \; x_{k,i} \, y_{k,j} \right) + z_0 \, x_{k_0,0} \, y_{k_0,0}$$

$$+ z_1 \, x_{k_0,0} \, y_{k_0,1} + \cdots + z_{n-2} \, x_{k_0,\sqrt{n}-1} \, y_{k_0,\sqrt{n}-2} + z_{n-1} \, x_{k_0,\sqrt{n}-1} \, y_{k_0,\sqrt{n}-1}.$$

and so $\Gamma_{x_i} \geq \Omega(n)$

$$\Rightarrow \Gamma \left(\sum_{k=1}^{\sqrt{n}} \sum_{i=0}^{\sqrt{n}-1} \sum_{j=0}^{\sqrt{n}-1} z_{i\sqrt{n}+j} \; x_{k,i} \, y_{k,j} \right) \geq \Omega\left(n^{3/2}\right)$$

Extending the Result to an Arbitrary Constant.

For any constant d, consider the degree-d polynomial:

$$\sum_{k=1}^{n^{1-1/d}} \sum_{\substack{i_0,\ldots,i_{d-1} \\ \{0,\ldots,n^{1/d}-1\}}} y_{\sum_{j=0}^{d-1} i_j n^{j/d}} \prod_{j=0}^{d-1} x^{(j)}_{k,i_j}$$

Partition: $X_0 \sqcup X_1 \sqcup \cdots \sqcup X_{d-1} \sqcup Y$ where

$Y = \{y_0, y_1, \ldots, y_{n-1}\}$ and for any $k \in [n^{1-1/d}]$

$$X_k = \{x^{(0)}_{k,0}, x^{(0)}_{k,1}, \ldots, x^{(0)}_{k,n^{1/d}-1}, x^{(1)}_{k,0}, \ldots, x^{(1)}_{k,n^{1/d}-1}, \ldots, x^{(d-1)}_{k,0}, \ldots, x^{(d-1)}_{k,n^{1/d}-1}\}.$$

No. of Variables $= n + n^{1-1/d}\left(d \cdot n^{1/d}\right)$.

$$= n + d.n = (d+1)\,n.$$

Further, for any $k \in [n^{1-1/d}]$, $\Gamma_{x_i} \geq \Omega(n)$ and so

$$\Gamma\left(\sum_{k=1}^{n^{1-1/d}} \sum_{\substack{i_0,\ldots,i_{d-1} \\ \{0,\ldots,n^{1/d}-1\}}} y_{\sum_{j=0}^{d-1} i_j n^{j/d}} \prod_{j=0}^{d-1} x^{(j)}_{k,i_j}\right) \geq \Omega\left(n^{2-1/d}\right)$$

That is,

for any constant $\varepsilon > 0$, there is a polynomial of degree $1/\varepsilon$ such that any formula computing it has size $\Omega(n^{2-\varepsilon})$.

Upper Bounding the Measure.

If a polynomial f is computable by a formula of size s, then $\Gamma(f) \leq s$.

Proof: Υ : Formula of size s computing f.

For any $v \in \Upsilon$,

$L_{X_i}(v)$: Leaves of the sub-tree of Υ rooted at v that are in the set X_i.

Claim: $\Gamma_{X_i}(f) \leq O(|L_{X_i}(\text{root})|)$.

Note that this is enough to prove the lemma.

Claim: $\Gamma_{x_i}(f) \leq O(|L_{x_i}(\text{root})|)$.

Proof: Let $Y = x_i$ for some i.

Main Idea:

Modify the formula so that any polynomial on variables outside of Y can be computed "freely".

- The new formula, Υ_Y, will be computing the same poly.
- $|L_Y(\text{root}(\Upsilon))| = |L_Y(\text{root}(\Upsilon_Y))|$.
- Total no. of leaves in $\Upsilon_Y \leq O(|L_Y(\text{root}(\Upsilon_Y))|)$.

Let $f = \sum f_i \, m_i$ where $f_i \in \mathbb{F}[X \setminus Y]$ and m_i s are monomials over Y.

Since Υ_Y computes f, each f_i is some polynomial combination of the leaves in Υ_Y.

Thus, $\text{algrank}_Y(f) \leq \#$ leaves in Υ_Y

$$\leq O\left(|L_Y(\text{root}(\Upsilon_Y))|\right)$$

$$\leq O\left(|L_Y(\text{root}(\Upsilon))|\right)$$

So if we can construct such a Υ_Y, then we would be done.

Let Υ be a formula of size s that computes f.

Define

$$V_0 = \{ v \in \Upsilon : |L_Y(v)| = 0 \text{ and } |L_Y(\text{Parent}(v))| \geq 1 \}$$

$$V_1 = \{ v \in \Upsilon : |L_Y(v)| = 1 \text{ and } |L_Y(\text{Parent}(v))| \geq 2 \}$$

$$V_2 = \{ v \in \Upsilon : |L_Y(v)| \geq 2 \}.$$

- For any $v \in V_0$, replace the subtree at v by a leaf labelled by the poly. f_v.

Note that $f_v \in \mathbb{F}[X \setminus Y]$.

- For any $v \in V_1$, the polynomial computed at the node has the form $f_1 y_v + f_0$ where $f_0, f_1 \in \mathbb{F}[X \setminus Y]$ and $y_v \in Y$.

So we replace the subtree with the gadget

$$(l_1 \times y_v) + l_0$$

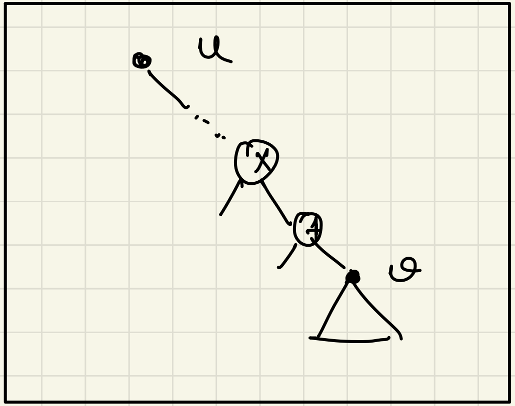
where l_1 is a leaf labelled by the poly. f_1

l_0 is a leaf labelled by the poly. f_0.

<u>Observation 1</u>: $|V_1| \leq |L_Y(\mathrm{root}(\gamma))|$

- Let u be an arbitrary node, and v be another node in the subtree rooted at u with $L_Y(u) = L_Y(v)$. Then $g_u = f_1 g_v + f_0$ where $f_1, f_0 \in \mathbb{F}[X \setminus Y]$.

  So we replace the entire chain from u to v by this gadget.



What we get:

For distinct $u, v \in V_2$, $L_Y(u) \neq L_Y(v)$.

<u>Observation 2</u>: $|V_2| \leq |L_Y(\mathrm{root}(\Upsilon))|$.

So we now have a new formula Υ_y computing f, where

- $|L_y(\mathrm{root}(\Upsilon))| = |L_y(\mathrm{root}(\Upsilon_y))|$.
- $|V_1| \leq |L_y(\mathrm{root}(\Upsilon))|$.
- $|V_2| \leq |L_y(\mathrm{root}(\Upsilon))|$.
- No. of leaves in Υ_y $= 3(|V_1| + |V_2|)$

$$= O(|L_y(\mathrm{root}(\Upsilon))|)$$

This completes the proof. ∎

Conclusion.

- Any formula computing $\sum_{i=1}^{n} \sum_{j=1}^{n} y_j \, x_i^{\,j}$ requires size $\Omega(n^2)$.

- For any $k \in \mathbb{N}$ with $k \geq 3$, there exists a degree k polynomial, such that any formula computing it requires size $\Omega(n^{2-1/k})$.

- If we get $\Omega(n^{2+\epsilon})$ lower bound against formulas in the border sense, for a constant degree polynomial, then we would have sub-exponential PIT for linear sized circuits.

Thank You !!