Lower Bounds Against Sums of Ordered Set-Multilinear ABPs

Prerona Chatterjee [with Deepanshu Kush (UoT), Shubhangi Saraf (UoT), Amir Shpilka (TAU)] Tel Aviv University

March 13, 2024

Answer: Using Horner's rule, O(d) in general.

Answer: Using Horner's rule, O(d) in general. But for $f(x) = x^d$, $O(\log d)$.

Answer: Using Horner's rule, O(d) in general. But for $f(x) = x^d$, $O(\log d)$.

Fact: There exist polynomials $f(x) \in \mathbb{F}[x]$, for which the answer is $\Omega(\sqrt{d})$.

Answer: Using Horner's rule, O(d) in general. But for $f(x) = x^d$, $O(\log d)$.

Fact: There exist polynomials $f(x) \in \mathbb{F}[x]$, for which the answer is $\Omega(\sqrt{d})$. In general the answer must be $\Omega(\log d)$.

Answer: Using Horner's rule, O(d) in general. But for $f(x) = x^d$, $O(\log d)$.

Fact: There exist polynomials $f(x) \in \mathbb{F}[x]$, for which the answer is $\Omega(\sqrt{d})$. In general the answer must be $\Omega(\log d)$.

Open Question: Describe $f(x) \in \mathbb{F}[x]$ of degree *d* for which the answer is $\omega(\log d)$.

Answer: Using Horner's rule, O(d) in general. But for $f(x) = x^d$, $O(\log d)$.

Fact: There exist polynomials $f(x) \in \mathbb{F}[x]$, for which the answer is $\Omega(\sqrt{d})$. In general the answer must be $\Omega(\log d)$.

Open Question: Describe $f(x) \in \mathbb{F}[x]$ of degree *d* for which the answer is $\omega(\log d)$.

Theorem [Shamir 79, Lipton 94]: If $h(x) = \prod_{i=1}^{d} (x - i)$ can be computed using poly(log d) additions and multiplications, then integer factoring is in P/ poly.







Polynomials over n variables of degree d.

Objects of Study



$$\alpha_{1}(x_{1}+x_{2})(x_{3}+\alpha)+(x_{1}+x_{2})(\alpha_{2}x_{2}+\alpha)$$

Objects of Study Polynomials over *n* variables of degree *d*.

Easy: Most polynomials require exp(n, d) sized 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*.

Easy: Most polynomials require exp(n, d) sized circuits.

Central Question

 $VP \stackrel{?}{=} VNP$: Find explicit polynomials that

cannot be computed by circuits of size poly(n,d).





• Label on each edge: An affine linear form in $\{x_1, x_2, \dots, x_n\}$



- Label on each edge: An affine linear form in $\{x_1, x_2, \dots, x_n\}$
- Weight of path p = wt(p): Product of the edge labels on p



- Label on each edge: An affine linear form in $\{x_1, x_2, \dots, x_n\}$
- Weight of path p = wt(p): Product of the edge labels on p
- Polynomial computed by the ABP: $\sum_{p} wt(p)$



- Label on each edge: An affine linear form in $\{x_1, x_2, \dots, x_n\}$
- Weight of path p = wt(p): Product of the edge labels on p
- Polynomial computed by the ABP: $\sum_{p} wt(p)$

In this talk: Is there an explicit *n*-variate, degree *d* polynomial that can not be represented by an ABP of size poly(n, d)? For general ABPs, the best lower bound is just quadratic.

[C-Kumar-She-Volk]: Any ABP computing $\sum_{i=1}^{n} x_i^d$ requires $\Omega(nd)$ vertices.

For general ABPs, the best lower bound is just quadratic.

[C-Kumar-She-Volk]: Any ABP computing $\sum_{i=1}^{n} x_i^d$ requires $\Omega(nd)$ vertices.

Recently Bhargav, Dwivedi and Saxena showed that there is a different line of attack.

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $O\left(\frac{\log n}{\log \log n}\right)$ implies super-polynomial lower bound against ABPs.

For general ABPs, the best lower bound is just quadratic.

[C-Kumar-She-Volk]: Any ABP computing $\sum_{i=1}^{n} x_i^d$ requires $\Omega(nd)$ vertices.

Recently Bhargav, Dwivedi and Saxena showed that there is a different line of attack.

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $O\left(\frac{\log n}{\log \log n}\right)$ implies super-polynomial lower bound against ABPs.

The Question: Can we prove lower bounds against a general $\sum \text{osmABP}$?

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $d = O\left(\frac{\log n}{\log \log n}\right) \implies$ super-polynomial lower bound against ABPs.

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $d = O\left(\frac{\log n}{\log \log n}\right) \implies$ super-polynomial lower bound against ABPs.

Our Main Result: For $\omega(\log n) = d \le n$, there is a polynomial $G_{n,d}(\mathbf{x})$ which is set-multilinear w.r.t $\mathbf{x} = {\mathbf{x}_1, \ldots, \mathbf{x}_d}$, where $|\mathbf{x}_i| \le n$ for every $i \in [d]$, such that:

- $G_{n,d}$ is computable by a set-multilinear ABP of size poly(n),
- any $\sum \text{osmABP}$ computing $G_{n,d}$ must have super-polynomial total-width.

Set-Multilinearity and Ordered Set-Multilinearity

The variable set is divided into buckets.

$$\mathbf{x} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_d$$
 where $\mathbf{x}_i = \{x_{i,1}, \dots, x_{i,n_i}\}$.

The variable set is divided into buckets.

$$\mathbf{x} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_d$$
 where $\mathbf{x}_i = \{x_{i,1}, \dots, x_{i,n_i}\}$.

f is set-multilinear with respect to $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$ if

every monomial in f has exactly one variable from \mathbf{x}_i for each $i \in [d]$.

The variable set is divided into buckets.

$$\mathbf{x} = \mathbf{x}_1 \cup \cdots \cup \mathbf{x}_d$$
 where $\mathbf{x}_i = \{x_{i,1}, \dots, x_{i,n_i}\}$.

f is set-multilinear with respect to $\{x_1, \ldots, x_d\}$ if

every monomial in f has exactly one variable from \mathbf{x}_i for each $i \in [d]$.

An ABP is set-multilinear with respect to $\{\mathbf{x}_1, \dots, \mathbf{x}_d\}$ if every path in it

computes a set-multilinear monomial with respect to $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$.

Ordered Set-Multilinear ABPs (osmABPs)

An ABP is ordered set-multilinear with respect to $\{x_1, \ldots, x_d\}$ if

• there are *d* layers in the ABP

An ABP is ordered set-multilinear with respect to $\{x_1, \ldots, x_d\}$ if

- there are *d* layers in the ABP
- there is a permutation $\sigma \in S_d$ such that

every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

An ABP is ordered set-multilinear with respect to $\{x_1,\ldots,x_d\}$ if

- there are *d* layers in the ABP
- there is a permutation $\sigma \in S_d$ such that

every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

Example:

$$\text{IMM}_{n,d} = \sum_{1 \le i_1, \dots, i_{d-1} \le n} x_{1,i_1}^{(1)} \cdot \left(\prod_{j=2}^{d-1} x_{i_{j-1},i_j}^{(j)}\right) \cdot x_{i_{d-1},i_d}^{(d)}$$

- there are *d* layers in the ABP
- there is a permutation $\sigma \in S_d$ such that

every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

Example: For $\mathbf{x}_k = \left\{ x_{i,j}^{(k)} : i \in [n], j \in [n] \right\}$, $\mathrm{IMM}_{n,d} = \sum_{1 \le i_1, \dots, i_{d-1} \le n} x_{1,i_1}^{(1)} \cdot \left(\prod_{j=2}^{d-1} x_{i_{j-1},i_j}^{(j)} \right) \cdot x_{i_{d-1},i_d}^{(d)}.$ An ABP is ordered set-multilinear with respect to $\{\textbf{x}_1,\ldots,\textbf{x}_d\}$ if

- there are *d* layers in the ABP
- there is a permutation $\sigma \in S_d$ such that

every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

Example: For
$$\mathbf{x}_{k} = \left\{ x_{i,j}^{(k)} : i \in [n], j \in [n] \right\}$$
,

$$\mathrm{IMM}_{n,d} = \sum_{1 \le i_{1}, \dots, i_{d-1} \le n} x_{1,i_{1}}^{(1)} \cdot \left(\prod_{j=2}^{d-1} x_{i_{j-1},i_{j}}^{(j)} \right) \cdot x_{i_{d-1},i_{d}}^{(d)}.$$

has an osmABP of size O(nd) for $\sigma \in S_d$ being the identity permutation.

- there are *d* layers in the ABP
- every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

- there are *d* layers in the ABP
- every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

 \sum osmABP: Sum of ordered set-multilinear ABPs, each with a possibly different ordering.

- there are *d* layers in the ABP
- every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

 \sum osmABP: Sum of ordered set-multilinear ABPs, each with a possibly different ordering.

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $d = O\left(\frac{\log n}{\log \log n}\right) \implies$ super-polynomial lower bound against ABPs.

- there are *d* layers in the ABP
- every edge in layer *i* is labelled by a homogeneous linear form in $\mathbf{x}_{\sigma(i)}$

 \sum osmABP: Sum of ordered set-multilinear ABPs, each with a possibly different ordering.

[Bhargav-Dwivedi-Saxena]: Super polynomial lower bound against total-width of \sum osmABP for a polynomial of degree $d = O\left(\frac{\log n}{\log \log n}\right) \implies$ super-polynomial lower bound against ABPs.

[C-Kush-Saraf-Shpilka]: Super polynomial lower bound against total-width of $\sum \text{osmABP}$ for a polynomial of degree $d = \omega(\log n)$ that is computable by polynomial-sized ABPs.

Our Other Results

Sum of Ordered Set-Multilinear ABPs:

There is a polynomial $G_{n,n}(\mathbf{x})$ which is set-multilinear with respect to $\mathbf{x} = {\mathbf{x}_1, \ldots, \mathbf{x}_n}$, where $|\mathbf{x}_i| \le n$ for each $i \in [n]$, such that:

- it has a set-multilinear branching program of size poly(n),
- but any $\sum \text{osmABP}$ computing $G_{n,n}(\mathbf{x})$ requires total-width $\exp(\Omega(n^{1/1000}))$.

Sum of Ordered Set-Multilinear ABPs:

There is a polynomial $G_{n,n}(\mathbf{x})$ which is set-multilinear with respect to $\mathbf{x} = {\mathbf{x}_1, \ldots, \mathbf{x}_n}$, where $|\mathbf{x}_i| \le n$ for each $i \in [n]$, such that:

- it has a set-multilinear branching program of size poly(n),
- but any $\sum \text{osmABP}$ computing $G_{n,n}(\mathbf{x})$ requires total-width $\exp(\Omega(n^{1/1000}))$.

A single Ordered Set-Multilinear ABP:

There is a polynomial $G_{n,d}(\mathbf{x})$ which is set-multilinear with respect to $\mathbf{x} = {\mathbf{x}_1, \dots, \mathbf{x}_d}$, where $|\mathbf{x}_i| \le n$ for each $i \in [d]$, such that:

- it has a set-multilinear branching program of size poly(n, d),
- but any ordered set-multilinear branching program computing $G_{n,d}$ requires width $n^{\Omega(d)}$.

any $\sum \text{osmABP}$ computing it requires total-width $\exp(\Omega(n^{1/3}))$.

any $\sum \text{osmABP}$ computing it requires total-width $\exp(\Omega(n^{1/3}))$.

The same result also holds for the Nisan-Wigderson polynomial family, which is in VNP.

any $\sum \text{osmABP}$ computing it requires total-width $\exp(\Omega(n^{1/3}))$.

The same result also holds for the Nisan-Wigderson polynomial family, which is in VNP.

Low Degree Regime: For $\omega(\log n) = d \le n$, there is polynomial family $\{F_{n,d}(\mathbf{x})\}$, in VP, which is set-multilinear with respect to a set of $\Theta(d)$ buckets, each of size $\Theta(n)$, such that

 $F_{n,d}$ cannot be computed by a $\sum \text{osmABP}$ of total-width poly(n).

any $\sum \text{osmABP}$ computing it requires total-width $\exp(\Omega(n^{1/3}))$.

The same result also holds for the Nisan-Wigderson polynomial family, which is in VNP.

Low Degree Regime: For $\omega(\log n) = d \le n$, there is polynomial family $\{F_{n,d}(\mathbf{x})\}$, in VP, which is set-multilinear with respect to a set of $\Theta(d)$ buckets, each of size $\Theta(n)$, such that

 $F_{n,d}$ cannot be computed by a $\sum \text{osmABP}$ of total-width poly(n).

The same result also holds for the Nisan-Wigderson polynomial family, which is in VNP.

[Bhargav-Dwivedi-Saxena]

Any $\sum \text{osmABP}$ computing $\text{IMM}_{n,n}$ which has max-width $n^{o(1)}$ must have $2^{\Omega(n)}$ summands.

[Bhargav-Dwivedi-Saxena]

Any $\sum \text{osmABP}$ computing $\text{IMM}_{n,n}$ which has max-width $n^{o(1)}$ must have $2^{\Omega(n)}$ summands.

[Arvind-Raja]

Any $\sum_{i=1}^{t}$ osmABP computing the $n \times n$ permanent polynomial has max-width $2^{\Omega(n/t)}$.

[Ramya-Rao]

There exists an explicit polynomial family $\{g_n\}_n \in VP$, g_n being defined on the variable set $\{x_{1,0}, x_{1,1}\} \cup \cdots \cup \{x_{n,0}, x_{n,1}\}$, such that any $\sum \text{osmABP}$ computing it has total width $2^{\Omega\left(\frac{n^{1/6}}{\log n}\right)}$.

[Ramya-Rao]

There exists an explicit polynomial family $\{g_n\}_n \in VP$, g_n being defined on the variable set $\{x_{1,0}, x_{1,1}\} \cup \cdots \cup \{x_{n,0}, x_{n,1}\}$, such that any $\sum \text{osmABP}$ computing it has total width $2^{\Omega\left(\frac{n^{1/6}}{\log n}\right)}$.

[Ghoshal-Rao]

There exists an explicit polynomial family $\{g_n\}_n \in VBP$, g_n being defined on the variable set $\{x_{1,0}, x_{1,1}\} \cup \cdots \cup \{x_{n,0}, x_{n,1}\}$, such that any $\sum \text{osmABP computing } g_n$ that has max-width poly(n) must have total width $2^{\Omega(n^{1/500})}$.

Proof Overviews











Every path corresponds to a sequence of d/2 pairs.



Every path corresponds to a sequence of d/2 pairs. $\mathcal{P}_{d/2}$: Set of all such sequences of pairs.



f is a set-multilinear poly. w.r.t $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$.



f is a set-multilinear poly. w.r.t $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$.

[Nisan]: For every $1 \le k \le d$, the number of vertices in the *k*-th layer of the smallest osmABP(σ) computing *f* is equal to the rank of $M_{f,\sigma}(k)$.



f is a set-multilinear poly. w.r.t $\{\mathbf{x}_1, \ldots, \mathbf{x}_d\}$.

[Nisan]: For every $1 \le k \le d$, the number of vertices in the *k*-th layer of the smallest osmABP(σ) computing *f* is equal to the rank of $M_{f,\sigma}(k)$.

If \mathcal{A} is the smallest osmABP computing f, then

$$\operatorname{size}(\mathcal{A}) = \sum_{i=1}^{d} \operatorname{rank}(M_{f,\sigma}(k)).$$

Lower Bound for a single osmABP (contd.)

$$G_{n,d} = \sum_{\mathcal{P} \in \mathcal{P}_{d/2}} \prod_{(i,j) \in \mathcal{P}} y_{i,j} y_{j,i} \cdot \left(\sum_{k=1}^n x_{i,k} x_{j,k} \right).$$

Lower Bound for a single osmABP (contd.)

$$G_{n,d} = \sum_{\mathcal{P} \in \mathcal{P}_{d/2}} \prod_{(i,j) \in \mathcal{P}} y_{i,j} y_{j,i} \cdot \left(\sum_{k=1}^n x_{i,k} x_{j,k} \right).$$

Properties:

• *G_{n,d}* is computable by a set-multilinear ABP of size poly(*n*, *d*).



$$G_{n,d} = \sum_{\mathcal{P} \in \mathcal{P}_{d/2}} \prod_{(i,j) \in \mathcal{P}} y_{i,j} y_{j,i} \cdot \left(\sum_{k=1}^n x_{i,k} x_{j,k} \right).$$

Properties:

- *G_{n,d}* is computable by a set-multilinear ABP of size poly(*n*, *d*).
- For every $\sigma \in S_d$, there is some \mathcal{P} such that for at least d/8 of the $P = (i, j) \in \mathcal{P}$, $i \in$ $\{\sigma(1), \ldots \sigma(\frac{d}{2})\} \& j \in \{\sigma(1 + \frac{d}{2})), \ldots \sigma(d)\}.$



$$G_{n,d} = \sum_{\mathcal{P} \in \mathcal{P}_{d/2}} \prod_{(i,j) \in \mathcal{P}} y_{i,j} y_{j,i} \cdot \left(\sum_{k=1}^n x_{i,k} x_{j,k} \right).$$

Properties:

- *G_{n,d}* is computable by a set-multilinear ABP of size poly(*n*, *d*).
- For every $\sigma \in S_d$, there is some \mathcal{P} such that for at least d/8 of the $P = (i,j) \in \mathcal{P}$, $i \in \{\sigma(1), \ldots \sigma(\frac{d}{2})\}$ & $j \in \{\sigma(1 + \frac{d}{2})), \ldots \sigma(d)\}$.

Therefore,

$$\operatorname{rank}(M_{G_{n,d},\sigma}(d/2)) = \Omega(n^{d/8}).$$

• $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.

- $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.
- If $G_{n,d}$ is computed by a sum of t osmABPs, then

$$G_{n,d} = \sum_{i=1}^{t} g_i$$
 where $g_i = \sum_{u_1,...,u_{q-1}} \prod_{j=1}^{q} g_{u_{j-1},u_j}^{(i)}$

- $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.
- If $G_{n,d}$ is computed by a sum of t osmABPs, then

$$G_{n,d} = \sum_{i=1}^t g_i \quad ext{where} \quad g_i = \sum_{u_1, \dots, u_{q-1}} \prod_{j=1}^q g_{u_{j-1}, u_j}^{(i)}.$$

- $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.
- If $G_{n,d}$ is computed by a sum of t osmABPs, then

$$G_{n,d} = \sum_{i=1}^t g_i \quad ext{where} \quad g_i = \sum_{u_1, \dots, u_{q-1}} \prod_{j=1}^q g_{u_{j-1}, u_j}^{(i)}.$$

for every *i*, w.h.p. there are many *j*s, for which $M_w(g_{u_{i-1},u_i}^{(i)})$ is far from full rank

- $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.
- If $G_{n,d}$ is computed by a sum of t osmABPs, then

$$G_{n,d} = \sum_{i=1}^t g_i \quad ext{where} \quad g_i = \sum_{u_1, \dots, u_{q-1}} \prod_{j=1}^q g_{u_{j-1}, u_j}^{(i)}.$$

for every *i*, w.h.p. there are many *j*s, for which $M_w(g_{u_{j-1},u_j}^{(i)})$ is far from full rank \implies for every *i*, w.h.p. $M_w(g_i)$ is far from full rank

- $\{M_w(f) : w \in S\}$ is a set of matrices such that $M_w(G_{n,d})$ has full rank for every $w \in S$.
- If $G_{n,d}$ is computed by a sum of t osmABPs, then

$$G_{n,d} = \sum_{i=1}^t g_i \quad ext{where} \quad g_i = \sum_{u_1, \dots, u_{q-1}} \prod_{j=1}^q g_{u_{j-1}, u_j}^{(i)}.$$

for every *i*, w.h.p. there are many *j*s, for which $M_w(g_{u_{j-1},u_j}^{(i)})$ is far from full rank

 \implies for every *i*, w.h.p. $M_w(g_i)$ is far from full rank

 $\implies M_w(G_{n,d})$ is far from full rank unless *t* is large.

Open Questions

1. PIT for $\sum osmABP$?

- 1. PIT for $\sum \text{osmABP}$?
- 2. Super-quadratic lower bounds against smABPs?

- 1. PIT for $\sum \text{osmABP}$?
- 2. Super-quadratic lower bounds against smABPs?

Thank you!