Sums of Powers of Linear Forms, Partial Derivatives, and Commuting Matrices

Anamay Tengse

(University of Haifa)

Based on a joint work with C. Ramya (IMSc, Chennai)

WACT 2023

ALGEBRAIC BRANCHING PROGRAMS

$$f(x_1, \ldots, x_n) = \left(\begin{bmatrix} & & \\ & M_1 & \\ & & \end{bmatrix}_{r \times r} \cdot \begin{bmatrix} & & \\ & M_2 & \\ & & \end{bmatrix}_{r \times r} \cdots \begin{bmatrix} & & \\ & M_n & \\ & & \end{bmatrix} \right)_{(1, r)}$$

ALGEBRAIC BRANCHING PROGRAMS

$$f(x_1,\dots,x_n) = \left(\underbrace{\begin{bmatrix} M_1(x_1) \end{bmatrix}_{r\times r}}_{\text{deg-d univariates in } x_1} \cdot \underbrace{\begin{bmatrix} M_2(x_2) \end{bmatrix}_{r\times r}}_{\text{deg-d univariates in } x_2} \cdots \underbrace{\begin{bmatrix} M_n(x_n) \end{bmatrix}}_{\text{deg-d univariates in } x_n} \right)_{(1,r)}$$

READ-ONCE, OBLIVIOUS, ALGEBRAIC BRANCHING PROGRAMS

$$f(x_1,\dots,x_n) = \left(\underbrace{\begin{bmatrix} M_1(x_1) \end{bmatrix}_{r\times r}}_{\text{deg-}d\text{ univariates in }x_1} \cdot \underbrace{\begin{bmatrix} M_2(x_2) \end{bmatrix}_{r\times r}}_{\text{deg-}d\text{ univariates in }x_2} \cdots \underbrace{\begin{bmatrix} M_n(x_n) \end{bmatrix}}_{\text{deg-}d\text{ univariates in }x_n} \right)_{(1,r)}$$

ROABP for f

"Each variable is **read once**, **oblivious** to the others."

(Introduced in [Forbes-Shpilka '12])

READ-ONCE, OBLIVIOUS, ALGEBRAIC BRANCHING PROGRAMS

ROABP for f

$$f(x_1,\dots,x_n) = \Bigg(\underbrace{\begin{bmatrix} \\ M_1(x_1) \\ \\ \end{bmatrix}_{r\times r}}_{\text{deg-}d\text{ univariates in }x_1} \cdot \underbrace{\begin{bmatrix} \\ M_2(x_2) \\ \\ \end{bmatrix}_{r\times r}}_{\text{deg-}d\text{ univariates in }x_2} \cdots \underbrace{\begin{bmatrix} \\ M_n(x_n) \\ \\ \end{bmatrix}}_{\text{deg-}d\text{ univariates in }x_n} \Bigg)_{(1,r)}$$

▷ "Each variable is read once, oblivious to the others."

(Introduced in [Forbes-Shpilka '12])

▷ '$f(x_1,\dots,x_n)$ has width-r ROABPs in order $(x_1,x_2,\dots,x_n)$.'

ROABPs : ORDER MATTERS

$$FR(x_1,\dots,x_n,y_1,\dots,y_n) = \left(\underbrace{\begin{bmatrix} 1 & x_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_1 & 0 \end{bmatrix}}_{\begin{bmatrix} 1+x_1y_1 & 0 \\ 0 & 0 \end{bmatrix}} \underbrace{\begin{bmatrix} 1 & x_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_2 & 0 \end{bmatrix}}_{\begin{bmatrix} 1+x_2y_2 & 0 \\ 0 & 0 \end{bmatrix}} \cdots \begin{bmatrix} 1 & 0 \\ y_n & 0 \end{bmatrix} \right)_{(1,1)}$$

$\dots$

$$FR(\bar{x},\bar{y}) = \sum_{\bar{b} \in \{0,1\}^n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \cdot y_1^{b_1} y_2^{b_2} \cdots y_n^{b_n}$$

ROABPs : ORDER MATTERS

$$FR(x_1,\dots,x_n,y_1,\dots,y_n) = \left(\begin{bmatrix} 1 & x_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_1 & 0 \end{bmatrix} \begin{bmatrix} 1 & x_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_2 & 0 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ y_n & 0 \end{bmatrix} \right)_{(1,1)}$$

$$FR(\bar{x},\bar{y}) = \sum_{\bar{b}\in\{0,1\}^n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \cdot y_1^{b_1} y_2^{b_2} \cdots y_n^{b_n}$$

▷ Any ROABP for $FR(\bar{x},\bar{y})$ in $(x_1,\dots,x_n,y_1,\dots,y_n)$ has width 2^n.

ROABPs : ORDER MATTERS

$$FR(x_1,\ldots,x_n,y_1,\ldots,y_n) = \left(\begin{bmatrix} 1 & x_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_1 & 0 \end{bmatrix} \begin{bmatrix} 1 & x_2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y_2 & 0 \end{bmatrix} \cdots \begin{bmatrix} 1 & 0 \\ y_n & 0 \end{bmatrix} \right)_{(1,1)}$$

$$FR(\bar{x},\bar{y}) = \sum_{\bar{b}\in\{0,1\}^n} x_1^{b_1} x_2^{b_2} \cdots x_n^{b_n} \cdot y_1^{b_1} y_2^{b_2} \cdots y_n^{b_n}$$

▷ Any ROABP for $FR(\bar{x},\bar{y})$ in $(x_1,\ldots,x_n,y_1,\ldots,y_n)$ has width 2^n.

Q. When does an $f(\bar{x})$ have ROABPs in every order?

SOME NOTATION

$$f(x_1, \dots, x_n) = \left(\Big[\quad M_1(x_1) \Big]_{r \times r} \cdot \Big[\quad M_2(x_2) \Big]_{r \times r} \cdots \Big[\quad M_n(x_n) \Big] \right)_{(1,r)}$$

SOME NOTATION

$$f(x_1, \dots, x_n) = \left(\begin{bmatrix} & & \\ & M_1(x_1) & \\ & & \end{bmatrix}_{n \times n} \cdot \begin{bmatrix} & & \\ & M_2(x_2) & \\ & & \end{bmatrix}_{n \times n} \cdots \begin{bmatrix} & & \\ & M_n(x_n) & \\ & & \end{bmatrix} \right)_{(1,n)}$$

$$\begin{bmatrix} x^2-1 & x \\ 1 & x+3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix} \cdot 1 + \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \cdot x + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot x^2$$

$$M_i(x_i) = M_{i,0} \cdot 1 + M_{i,1} \cdot x_i + \cdots + M_{i,d} \cdot x_i^d \ ;$$

SOME NOTATION

$$f(x_1,\ldots,x_n) = \left(\begin{bmatrix} \\ M_1(x_1) \\ \\ \end{bmatrix}_{r\times r} \cdot \begin{bmatrix} \\ M_2(x_2) \\ \\ \end{bmatrix}_{r\times r} \cdots \begin{bmatrix} \\ M_n(x_n) \\ \\ \end{bmatrix} \right)_{(1,r)}$$

$$\begin{bmatrix} x^2-1 & x \\ 1 & x+3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 3 \end{bmatrix} \cdot 1 + \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \cdot x + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot x^2$$

$$M_i(x_i) = M_{i,0} \cdot 1 + M_{i,1} \cdot x_i + \cdots + M_{i,d} \cdot x_i^d \ ;$$

$\{M_{i,j}\}_{[n]\times[d]}$ — "coefficient matrices" of ROABP

ROABPs in every order

ROABPs in every order

Diagonal ROABPs : all coeff matrices are diagonal

ROABPs in every order

Commutative ROABPs : all coeff matrices commute with each other

Diagonal ROABPs : all coeff matrices are diagonal

ROABPs in every order

Commutative ROABPs : all coeff matrices commute with each other

$\leq_{poly}$ $\uparrow$

Diagonal ROABPs : all coeff matrices are diagonal

ROABPs in every order

Commutative ROABPs : all coeff matrices commute with each other

$\leq_{poly}$ $\uparrow$ $\downarrow$?

Diagonal ROABPs : all coeff matrices are diagonal

Q. Suppose $f(x_1, \ldots, x_n)$ has a width-r commutative ROABP. How large should a diagonal ROABP for f be?

(I) $\mathrm{poly}(n, d, w)$

(II) "superpoly" (n, d, w)

...., YOU WON'T BELIEVE WHAT HAPPENS NEXT !?!

..., YOU WON'T BELIEVE WHAT HAPPENS NEXT !?!

Theorem [Ramya, T.]

Super-polynomial separation between commutative-ROABPs and diagonal-ROABPs

$$\Downarrow$$

Super-polynomial separation between Waring rank and dimension-of-partial-derivatives.

WARING RANK AND PARTIAL DERIVATIVES

$WR(f)$: smallest s s.t. $f(\bar{x}) = \sum_{i=1}^{s} \beta_i \cdot l_i(\bar{x})^d$

WARING RANK AND PARTIAL DERIVATIVES

$WR(f)$: smallest s s.t. $f(\bar{x}) = \sum_{i=1}^{s} \beta_i \cdot l_i(\bar{x})^d$

$- \dfrac{\partial\, l(\bar{x})^d}{\partial x_j} = \alpha_j \cdot l(\bar{x})^{d-1}$; $\quad \text{span}\left\{ \dfrac{\partial l_i^d}{\partial x_1}, \dots, \dfrac{\partial l_i^d}{\partial x_n} \right\} = \text{span}\left\{ l_i^{d-1} \right\}$

$\therefore\ \text{span}\left\{ \dfrac{\partial^{|m|} l_i^d}{\partial m} : \text{monomial } m \right\} = \text{span}\left\{ l_i^d, l_i^{d-1}, \dots, l_i, 1 \right\}$

WARING RANK AND PARTIAL DERIVATIVES

$WR(f)$: smallest s s.t. $f(\bar{x}) = \sum_{i=1}^{s} \beta_i \cdot l_i(\bar{x})^d$

– $\frac{\partial \, l(\bar{x})^d}{\partial x_j} = \alpha_j \cdot l(\bar{x})^{d-1}$; $\text{span}\left\{ \frac{\partial l_i^d}{\partial x_1}, \ldots, \frac{\partial l_i^d}{\partial x_n} \right\} = \text{span}\{ l_i^{d-1} \}$

$\therefore \text{span}\left\{ \frac{\partial^{|m|} l_i^d}{\partial m} : \text{monomial } m \right\} = \text{span}\{ l_i^d, l_i^{d-1}, \ldots, l_i, 1 \}$

Theorem [Nisan-Wigderson '95]

$$WR(f) \leq s \implies \dim\left(\text{span}\left\{ \frac{\partial^{|m|} f}{\partial m} \right\} \right) \leq s \cdot (d+1)$$

Q. Is the converse true (up to poly. factors)?

Theorem [Ramya, T.]

$$\forall f_{n,d} : \dim\left(\operatorname{span}\left\{\frac{\partial^{|m|} f}{\partial m}\right\}\right) \leq T \quad \Rightarrow \quad \mathrm{WR}(f) \leq (nd\,T)^{\alpha}$$

$$\Downarrow$$

$\forall g_n$: commutative-ROABP of width-r for g,

imply diagonal-ROABP for g of width $O(n \cdot r^{5\alpha})$.

PROOF IDEAS - I

[Ben-Or]: $$\mathrm{coeff}_{t^d}\left((1+tx_1)(1+tx_2)\cdots(1+tx_n)\right) = \sum_{\substack{S\subseteq[n]\\ |S|=d}} \left(\prod_{j\in S} x_j\right) = \mathrm{ESym}_{n,d}$$

PROOF IDEAS - I

[Ben-Or]: $\operatorname{coeff}_{t^d}\left((1+tx_1)(1+tx_2)\cdots(1+tx_n)\right) = \sum_{\substack{S\subseteq[n]\\|S|=d}} \left(\prod_{j\in S} x_j\right) = \mathrm{ESym}_{n,d}$

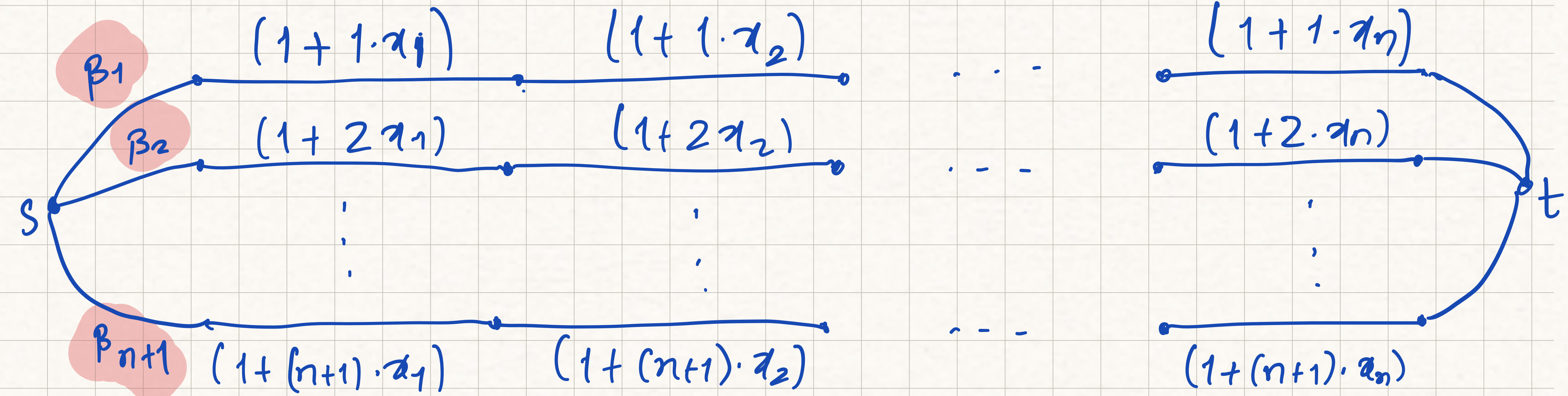
$$t \leftarrow \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}_{(d+1)\times(d+1)} \quad : \quad \mathrm{ESym}_{n,d} = \left(\begin{bmatrix} 1 & x_1 & & & \\ & 1 & x_1 & & \\ & & 1 & \ddots & \\ & & & \ddots & x_1 \\ & & & & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & x_2 & & & \\ & 1 & x_2 & & \\ & & 1 & \ddots & \\ & & & \ddots & x_2 \\ & & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & x_n & & & \\ & 1 & \ddots & & \\ & & \ddots & & x_n \\ & & & & 1 \end{bmatrix} \right)_{(1,d+1)}$$

PROOF IDEAS - I

[Ben-Or]: $$\mathrm{coeff}_{t^d}\Big((1+tx_1)(1+tx_2)\cdots(1+tx_n)\Big) = \sum_{\substack{S\subseteq[n]\\ |S|=d}} \Big(\prod_{j\in S} x_j\Big) = \mathrm{ESym}_{n,d}$$

$$t \leftarrow \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}_{(d+1)\times(d+1)} \quad : \quad \mathrm{ESym}_{n,d} = \left(\begin{bmatrix} 1 & x_1 & & \\ & 1 & x_1 & \\ & & 1 & \ddots \\ & & & \ddots & x_1 \\ & & & & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & x_2 & & \\ & 1 & x_2 & \\ & & 1 & \ddots \\ & & & \ddots & x_2 \\ & & & & 1 \end{bmatrix} \cdots \begin{bmatrix} 1 & x_n & & \\ & 1 & \ddots & \\ & & \ddots & x_n \\ & & & 1 \end{bmatrix} \right)_{(1,d+1)}$$

Interpolation for coeff_{t^d} :



PROOF IDEAS - II

$$\tilde{f}(\bar{t}, \bar{x}) = \prod_{i=1}^{n} f_i(\bar{t}, x_i), \qquad |\bar{t}| \le r^2, \quad \bar{t}\text{-deg}(f_i) \le r$$

commutative ROABP
of width r

PROOF IDEAS - II

$$\tilde{f}(\bar{t}, \bar{x}) = \prod_{i=1}^{n} f_i(\bar{t}, x_i),$$

$|\bar{t}| \leq r^2$

$\bar{t}\text{-deg}(f_i) \leq r$

$$f(\bar{x}) = \sum_{m} \alpha_m \cdot \text{coeff}_m(\tilde{f}),$$

$\partial_{<\infty}\left(\sum_m \alpha_m \cdot m\right) \leq r$

commutative ROABP

of width r

PROOF IDEAS - II

$$\tilde{f}(\bar{t}, \bar{x}) = \prod_{i=1}^{n} f_i(\bar{t}, x_i), \qquad |\bar{t}| \leq r^2, \quad \bar{t}\text{-}\deg(f_i) \leq r$$

$$f(\bar{x}) = \sum_{m} \alpha_m \cdot \mathrm{coeff}_m(\tilde{f}), \qquad \partial_{<\infty}\Big(\underbrace{\sum_m \alpha_m \cdot m}_{h}\Big) \leq r$$

commutative ROABP of width r

diagonal ROABP of width $\mathrm{poly}(n, WR(h))$

PROOF IDEAS - II

$$\tilde{f}(\bar{t}, \bar{x}) = \prod_{i=1}^{n} f_i(\bar{t}, x_i), \qquad |\bar{t}| \leq r^2, \quad \bar{t}\text{-}\deg(f_i) \leq r$$

$$f(\bar{x}) = \sum_{m} \alpha_m \cdot \mathrm{coeff}_m(\tilde{f}), \qquad \underbrace{\partial_{<\infty}\Big(\sum_m \alpha_m \cdot m\Big)}_{h} \leq r$$

[MS'95]
[MMM'93]

e.g. [Pratt'19]

commutative ROABP of width r

diagonal ROABP of width $\mathrm{poly}(n, WR(h))$

LET US STUDY ROABPs !!!

* Lower bounds for commutative, diagonal ROABPs (that do not extend to 'every-order-ROABPs').

* PIT for diagonal or commutative ROABPs (different from [AGKS'15], [GKS'16]).

  ↳ for $n \sim O(\log(dr))$, implications to $\Sigma\wedge\Sigma$.

THANK YOU.

Thanks to Premona Chatterjee, Ramprasad Saptharishi.