Algebraic Geometry with Applications to TEnsors and Secants
Algebraic geometry and complexity theory workshop
Warsaw, 2022-Nov-14

Homogeneous algebraic computation

Christian Ikenmeyer
joint work with Pranjal Dutta, Fulvio Gesmundo, Gorav Jindal, and Vladimir Lysikov

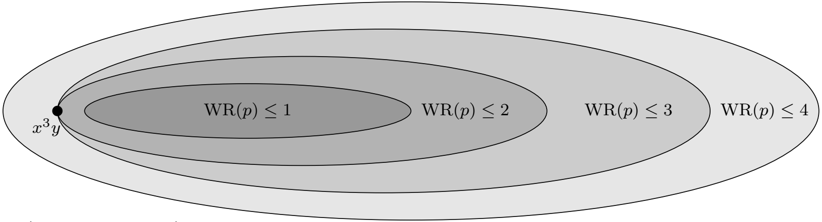
Name of the paper: Border complexity via elementary symmetric polynomials

WARWICK
THE UNIVERSITY OF WARWICK

For a homogeneous degree d polynomial p define the **Waring rank** $\mathrm{WR}(p)$ as the smallest r such that there exist homogeneous linear polynomials with

$$p = \sum_{i=1}^{r} (\ell_i)^d.$$

$12x^3y = (x+y)^4 + i^3(x+iy)^4 + i^2(x+i^2y)^4 + i(x+i^3y)^4$, hence $\mathrm{WR}(x^3y) \le 4$. In fact, $\mathrm{WR}(x^3y) = 4$.



$\frac{1}{\varepsilon}\Big((x+\varepsilon y)^4 - x^4\Big) = 4x^3y + \varepsilon(6x^2y^2 + 4\varepsilon xy^3 + \varepsilon^2 y^4) \xrightarrow{\varepsilon \to 0} 4x^3y$

The **border Waring rank** $\underline{\mathrm{WR}}(p)$ is defined as the smallest r such that p can be approximated arbitrarily closely by polynomials of Waring rank $\le r$. For example, $\underline{\mathrm{WR}}(x^3y) \le 2$.

Theorem (works in high generality)

Let $V = \mathbb{C}[\vec{x}]_d$. Zariski closure and Euclidean closure coincide:

$$\{p \in V \mid \underline{\mathrm{WR}}(p) \le k\} = \overline{\{p \in V \mid \mathrm{WR}(p) \le k\}}^{\mathbb{C}} = \overline{\{p \in V \mid \mathrm{WR}(p) \le k\}}^{\mathsf{Zar}}.$$

(secant variety of the Veronese variety)

Analogously: The Chow rank

For a homogeneous degree d polynomial p define the **Chow rank** $\mathrm{CR}(p)$ as the smallest r such that there exist homogeneous linear polynomials $\ell_{i,j}$ with

$$p = \sum_{i=1}^{r} \prod_{j=1}^{d} \ell_{i,j}.$$

The **border Chow rank** $\underline{\mathrm{CR}}(p)$ is defined as the smallest r such that p can be approximated arbitrarily closely by polynomials of Chow rank $\leq r$.

(analogous theorem with secant variety of the Chow variety (i.e., variety of products of homogeneous linear forms))

For a homogeneous degree d polynomial p the **Waring rank** $\mathrm{WR}(p)$ is defined as the smallest r such that $\exists$ linear forms with

$$p = \begin{pmatrix}\ell_1 & \ell_2 & \cdots & \ell_r\end{pmatrix} \begin{pmatrix}\ell_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_r\end{pmatrix} \begin{pmatrix}\ell_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_r\end{pmatrix} \cdots \begin{pmatrix}\ell_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_r\end{pmatrix} \begin{pmatrix}\ell_1 \\ \vdots \\ \ell_r\end{pmatrix}$$

For a homogeneous degree d polynomial p the **Chow rank** $\mathrm{CR}(p)$ is defined as the smallest r such that $\exists$ linear forms with

$$p = \begin{pmatrix}\ell_{1,1} & \ell_{2,1} & \cdots & \ell_{r,1}\end{pmatrix} \begin{pmatrix}\ell_{1,2} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_{r,2}\end{pmatrix} \begin{pmatrix}\ell_{1,3} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_{r,3}\end{pmatrix} \cdots \begin{pmatrix}\ell_{1,d-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \ell_{r,d-1}\end{pmatrix} \begin{pmatrix}\ell_{1,d} \\ \vdots \\ \ell_{r,d}\end{pmatrix}$$

For a homogeneous degree d polynomial p the **width** $\mathrm{w}(p)$ is defined as the smallest r such that $\exists$ linear forms with

$$p = \begin{pmatrix}\ell_{1,1,1} & \ell_{1,2,1} & \cdots & \ell_{1,r,1}\end{pmatrix} \begin{pmatrix}\ell_{1,1,2} & \cdots & \ell_{1,r,2} \\ \vdots & \ddots & \vdots \\ \ell_{r,1,2} & \cdots & \ell_{r,r,2}\end{pmatrix} \begin{pmatrix}\ell_{1,1,3} & \cdots & \ell_{1,r,3} \\ \vdots & \ddots & \vdots \\ \ell_{r,1,3} & \cdots & \ell_{r,r,3}\end{pmatrix} \cdots \begin{pmatrix}\ell_{1,1,d-1} & \cdots & \ell_{1,r,d-1} \\ \vdots & \ddots & \vdots \\ \ell_{r,1,d-1} & \cdots & \ell_{r,r,d-1}\end{pmatrix} \begin{pmatrix}\ell_{1,1,d} \\ \vdots \\ \ell_{1,r,d}\end{pmatrix}$$

This is also called the **iterated matrix multiplication complexity** or the **algebraic branching program width**.

$\underline{\mathrm{w}}$ is defined analogously to $\underline{\mathrm{WR}}$ and $\underline{\mathrm{CR}}$.

Two parameters: d and r. This makes this a **general linear group orbit closure containment** question.

$$\underline{\mathrm{WR}}(p) \le r \quad \text{iff} \quad p \in \overline{\mathsf{GL}_N(x_1^d + x_2^d + \cdots + x_r^d)}.$$

$$\underline{\mathrm{CR}}(p) \le r \quad \text{iff} \quad p \in \overline{\mathsf{GL}_N(\prod_{i=1}^d x_{1,i} + \prod_{i=1}^d x_{2,i} + \cdots + \prod_{i=1}^d x_{r,i})}.$$

$$\mathsf{IMM}_r^{(d)} := \begin{pmatrix} x_{1,1,1} & x_{1,2,1} & \cdots & x_{1,r,1} \end{pmatrix} \begin{pmatrix} x_{1,1,2} & \cdots & x_{1,r,2} \\ \vdots & \ddots & \vdots \\ x_{r,1,2} & \cdots & x_{r,r,2} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1,d-1} & \cdots & x_{1,r,d-1} \\ \vdots & \ddots & \vdots \\ x_{r,1,d-1} & \cdots & x_{r,r,d-1} \end{pmatrix} \begin{pmatrix} x_{1,1,d} \\ \vdots \\ x_{1,r,d} \end{pmatrix}$$

homogeneous of degree d in $N := (d-2)r^2 + 2r$ variables.

$$\underline{\mathrm{w}}(p) \le r \quad \text{iff} \quad p \in \overline{\mathsf{GL}_N \mathsf{IMM}_r^{(d)}}.$$

Non-homogeneous computation via the determinant (Valiant 1979):

$$x_{11}x_{22} + x_{12}x_{21} + 3x_{11}x_{21} \;=\; \det \begin{pmatrix} x_{11} & x_{12} & 0 \\ 0 & x_{22} & x_{21} \\ 1 & 3 & 1 \end{pmatrix}$$

In fact, every polynomial can be written as the determinant of a matrix with **affine** linear entries. The smallest size is called the determinantal complexity $\mathsf{dc}(p)$.

For determinantal complexity we need padding or the general affine group: $\underline{\mathsf{dc}}(p) \le r$ iff $x_0^{r-\deg(p)} p \in \overline{\mathsf{GL}_{r^2} \det_r}$.

Issues with non-homogeneity pointed out in: [Kadish-Landsberg 2012], [I-Panova 2015], [Bürgisser-I-Panova 2016]

- A sequence $(c_n)_{n\in\mathbb{N}}$ of integers is **polynomially bounded** if $\exists$ a polynomial t such that $\forall n \in \mathbb{N}:\ c_n \le t(n)$.
- A sequence $(c_n)_{n\in\mathbb{N}}$ of integers is **quasipolynomially bounded** if $\exists$ a polynomial t such that $\forall n \ge 2:\ c_n \le n^{t(\log n)}$.

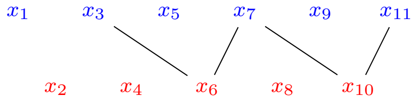
The permanent polynomial $\text{per}_n := \sum_{\pi\in\mathfrak{S}_n}\prod_{i=1}^n x_{i,\pi(i)}$. Grenet 2012: $\mathrm{w}(\text{per}_n) \le \binom{n}{n/2} \sim 2^n/\sqrt{\pi n/2}$.

Conjectures, all very similar:

Valiant's conjecture:	w(per) is not polynomially bounded.
Extended Valiant's conjecture (Bürgissser):	w(per) is not quasipolynomially bounded.
Mulmuley-Sohoni conjecture:	$\underline{\mathrm{w}}$(per) is not polynomially bounded.
Extended Mulmuley-Sohoni conjecture:	$\underline{\mathrm{w}}$(per) is not quasipolynomially bounded.

The parity-alternating elementary symm. polyn.:

$$C_r^{(d)} = \sum_{\substack{0\le i_1<\ldots<i_d\le r \\ \text{parity-alternating} \\ \text{starts with odd}}} x_{i_1}\cdot x_{i_2}\cdots x_{i_d}$$



$x_3 \cdot x_6 \cdot x_7 \cdot x_{10} \cdot x_{11}$

c(p) is the smallest r such that $\exists$ **homogeneous linear** ℓ_i with $p = C_r^{(d)}(\ell_1,\ldots,\ell_r)$ (not always finite).

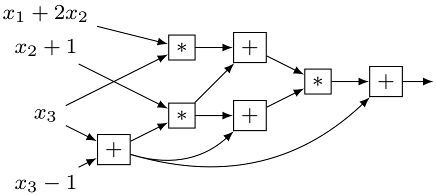
Main theorem of this talk

For every homogeneous p we have $\underline{\text{c}}(p)$ finite.
The extended Mulmuley-Sohoni conjecture is equivalent to

"$\underline{\text{c}}$(per) is not quasipolynomially bounded".

This looks much simpler than IMM: $\text{per}_d \overset{?}{\in} \overline{\mathsf{GL}_r C_r^{(d)}}$

Research direction: Are there even simpler polynomials than $C_r^{(d)}$ that have this property?



An arithmetic circuit with affine linear inputs, size 7, depth 5.

An arithmetic **formula** is a circuit whose graph is a tree.

- For $p \in \mathbb{C}[\vec{x}]$ let $\mathrm{acc}(p)$ be the size of the smallest circuit computing p.
- For $p \in \mathbb{C}[\vec{x}]$ let $\mathrm{afc}(p)$ be the size of the smallest formula computing p.

Recall Valiant's conjecture and extension:

Valiant's conjecture:	$\mathrm{w}(\mathrm{per})$ is not polynomially bounded.
Extended Valiant's conjecture:	$\mathrm{w}(\mathrm{per})$ is not quasipolynomially bounded.

consider two more conjectures:

Valiant's conjecture for formulas:	$\mathrm{afc}(\mathrm{per})$ is not polynomially bounded.
Valiant's conjecture for circuits:	$\mathrm{acc}(\mathrm{per})$ is not polynomially bounded.

Theorem (Equivalent formulations of Valiant's extended conjecture)

The following are equivalent:

- $\mathrm{w}(\mathrm{per})$ is quasipolynomially bounded,
- $\mathrm{afc}(\mathrm{per})$ is quasipolynomially bounded,
- $\mathrm{acc}(\mathrm{per})$ is quasipolynomially bounded.

Completely analogously for the border measures, i.e., Mulmuley-Sohoni conjectures.

Definition Kc (Kumar's complexity)

Let $p \in \mathbb{C}[\vec{x}]$ with $p(0) = 0$. Define $\mathrm{Kc}(p)$ to be the smallest r such that $\exists \alpha \in \mathbb{C} \setminus \{0\}$ and $\exists \ell_i \in \mathbb{C}[\vec{x}]_1$, $1 \leq i \leq r$, such that

$$\alpha^{-1} \cdot p = (1 + \ell_1) \cdot (1 + \ell_2) \cdots (1 + \ell_r) - 1.$$

If this is impossible, then $\mathrm{Kc}(p) = \infty$.

Lemma: If p is homogeneous and $\mathrm{Kc}(p) < \infty$, then p is a power of a linear form.

Kumar's theorem (2020)

$\forall p \in \mathbb{C}[\vec{x}]_d$ we have $\underline{\mathrm{Kc}}(p) \leq d \cdot \underline{\mathrm{WR}}(p)$.

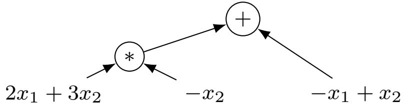
The proof is an application of Shpilka's 2002 paper "Affine projections of symmetric polynomials".

Our "Kumar-reverse" Theorem

If $p \in \mathbb{C}[\vec{x}]_d$ is not a product of homogeneous linear forms, then $\underline{\mathrm{WR}}(p) \leq \underline{\mathrm{Kc}}(p)$.

Proof: case distinction, depending on the most significant exponent of ε in $\alpha \in \mathbb{C}[\varepsilon^{-1}, \varepsilon]$.

The interesting case is when $|\alpha| \xrightarrow{\varepsilon \to 0} \infty$. All degrees $< d$ must converge to 0.
Hence all elem. symm. functions of degree $< d$ in the ℓ_i converge to 0.
Hence all symm. functions of degree $< d$ in the ℓ_i converge to 0.
Hence $e_d(\ell_1, \ldots, \ell_r)$ and $p_d(\ell_1, \ldots, \ell_r)$ coincide in the limit (up to scale). □



An algebraic formula is **input homogeneous linear** (IHL) if all leaf labels are homogeneous linear (no constants allowed).

$f(0) = 0$, but that is the only requirement.

Proposition (Rescaling)

If p is computed by a size s IHL formula, then αp is also computed by a size s IHL formula.

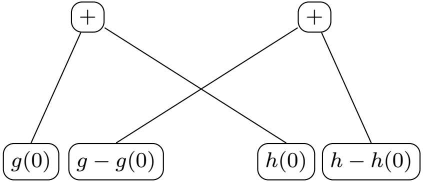
Proof: Rescale the root recursively, both childs for addition gates, one child for product gates.

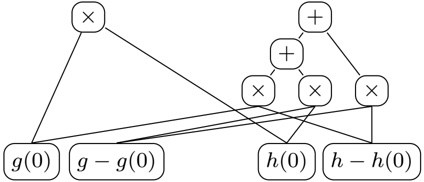
Proposition (Conversion)

Given a size s formula for p, we find a $\text{poly}(s)$ size IHL formula for p, same depth up to factor of 3.

Proof:

1. Brent's depth reduction to depth $O(\log s)$.
2. Split each gate into two: constant part and deg ≥ 1 part.





3. Unroll the circuit into a formula; use the rescaling proposition on gates that are rescaled by constants; delete the constant root subformula.

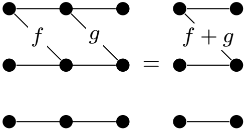
□

Homogeneous Ben-Or & Cleve, similar structure to Kumar's complexity

Let p have a depth δ IHL formula. Then there exist $r \leq 4^\delta$ many 3×3 matrices $A_1, \ldots, A_r$ with **homogeneous linear entries** such that

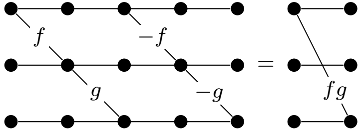
$$\begin{pmatrix} 0 & p & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = (\mathrm{id}_3 + A_1)(\mathrm{id}_3 + A_2)\cdots(\mathrm{id}_3 + A_r) - \mathrm{id}_3.$$

Proof: In the original Ben-Or & Cleve the addition and multiplication gadgets can be realized via $\mathrm{id}_3 + A$:



$$\begin{pmatrix} 1 & f & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & g & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & f+g & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & f & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & g \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -f & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -g \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & fg \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



□

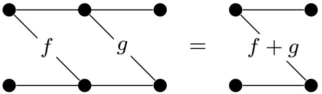
Let $A_i = \begin{pmatrix} 0 & x_{1,2,i} & x_{1,3,i} \\ x_{2,1,i} & 0 & x_{2,3,i} \\ x_{3,1,i} & x_{3,2,i} & 0 \end{pmatrix}$ and $D_r^{(d)} := \left(\sum_{1 \leq i_1 < i_2 < \cdots < i_d \leq r} A_{i_1} \cdots A_{i_d}\right)_{1,2}$.

Let $D(p)$ be the smallest r such that $\exists$ homogeneous linear ℓ_i with $p = D_r^{(d)}(\ell_1, \ldots, \ell_{6r})$.

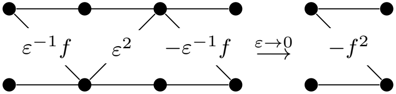
Corollary

- Valiant's conjecture for formula size is equivalent to "D(per) is not polynomially bounded".
- Valiant's extended conjecture is equivalent to "D(per) is not quasipolynomially bounded".

We try the same approach for 2×2 matrices (similar to [Bringmann-I-Zuiddam 2018]):

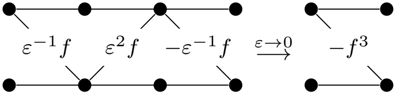


In [BIZ18] the product is simulated by squares: $fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right)$.



But this is affine!

Homogeneous:



We only have arity 3 products available: $fgh = \frac{1}{24}\left((f+g+h)^3 - (f+g-h)^3 - (f-g+h)^3 + (f-g-h)^3\right)$

In the construction we use these matrices (for $f \in \mathbb{C}[\vec{x}]_d$):

$$\begin{aligned} \begin{pmatrix} 0 & f \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & \ell_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \ell_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & \ell_3 \\ 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & \ell_r \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \sum_{\substack{0 \le i_1 < \ldots < i_d \le r \\ \text{parity-alternating} \\ \text{starts with odd}}} \ell_{i_1} \cdot \ell_{i_2} \cdots \ell_{i_d} \end{aligned}$$

Arity 3 products are surprisingly subtle

- In a HIL formula with arity 3 products, arity 2 products **cannot be simulated**!
- Formulas over the arity 3 basis can efficiently be simulated by formulas over the standard basis (trivial).
- Formulas can be simulated efficiently by **circuits** over the arity 3 basis.

Pictorially:

$$\begin{array}{ccccccc} \mathbf{V3F} & \subseteq & \mathbf{VF} & \subseteq & \mathbf{VBP} & \subseteq & \mathbf{VP}, \\ & & \cap\!\!\!| & & & & \\ & & \mathbf{V3P} & & & & \end{array}$$

- We go to the quasipolynomial versions and show $\mathbf{VQ3F} = \mathbf{VQ3P}$, which implies what we wanted:

$$\mathbf{VQ3F} = \mathbf{VQF} = \mathbf{VQBP} = \mathbf{VQP} = \mathbf{VQ3P}.$$

Theorem

$\mathbf{VQ3F} = \mathbf{VQ3P}$

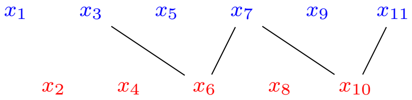
Proof:

1. **V3P** has polylog depth circuits (via a modified Valiant-Skyum-Berkowitz-Rackoff depth reduction).
2. Expansion of the circuit as a (quasipolynomially large) formula: $\mathbf{V3P} \subseteq \mathbf{VQ3F}$.
3. Now close **V3P** and **VQ3F** under quasipolynomial reductions.

□

The parity-alternating elementary symm. polyn.:

$$C_r^{(d)} = \sum_{\substack{0 \le i_1 < \ldots < i_d \le r \\ \text{parity-alternating} \\ \text{starts with odd}}} x_{i_1} \cdot x_{i_2} \cdots x_{i_d}$$



$x_3 \cdot x_6 \cdot x_7 \cdot x_{10} \cdot x_{11}$

c(p) is the smallest r such that $\exists$ **homogeneous linear** ℓ_i with $p = C_r^{(d)}(\ell_1, \ldots, \ell_r)$.

Main theorem of this talk

For every homogeneous p we have $\underline{\text{c}}(p)$ finite.
The extended Mulmuley-Sohoni conjecture is equivalent to

"$\underline{\text{c}}$(per) is not quasipolynomially bounded".

Open questions:

- Are there even nicer polynomials that achieve this?
- What do we get when we take the elementary symmetric polynomial?

Thank you for your attention!

7th Workshop on Algebraic Complexity Theory (WACT) in Warwick: 2023, March 27–31

https://www.dcs.warwick.ac.uk/~u2270030/wact