Rectangular Kronecker coefficients and plethysms in GCT

Christian Ikenmeyer

TEXAS A&M UNIVERSITY®

Flagship example: Writing the permanent as a determinant

$$\mathrm{per}_m := \sum_{\pi \in \mathfrak{S}_m} \prod_{i=1}^{m} x_{i,\pi(i)}.$$

- VNP-complete as a polynomial; #P-complete as a function
- Grenet 2011: We can write per_m as a determinant of a matrix of size $2^m - 1$.

Example: $$\mathrm{per}_3 = \det \begin{pmatrix} x_{11} & x_{12} & x_{13} & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & x_{32} & x_{33} & 0 & 0 \\ 0 & 1 & 0 & x_{31} & 0 & x_{33} & 0 \\ 0 & 0 & 1 & 0 & x_{31} & x_{32} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & x_{23} \\ 0 & 0 & 0 & 0 & 1 & 0 & x_{22} \\ 0 & 0 & 0 & 0 & 0 & 1 & x_{21} \end{pmatrix}$$

  Proof: Explicit construction of the algebraic branching program.
- With Hüttenhain 2015 (constant free model); also Alper, Bogart, Velasco: For $m = 3$ there is no smaller such matrix.
- Valiant: Every polynomial h can be written as a determinant. Let $\mathrm{dc}(h)$ denote the smallest size possible. So $\mathrm{dc}(\mathrm{per}_3) = 7$.
- Best known lower bound: $\mathrm{dc}(\mathrm{per}_m) \geq m^2/2$ (Mignon and Ressayre 2004)
- det vs per first studied by Pólya (1913) in a toy case

From combinatorics to geometry: Approximations

We are now working over the field $\mathbb{C}$ of complex numbers.

$$\text{Example: } h := X_{3,1} + X_{1,1}X_{2,3}X_{3,1} + X_{1,1}X_{2,2}X_{3,3}.$$

The matrix

$$A_\varepsilon := \begin{pmatrix} 1 & 1 & \varepsilon X_{1,1} & 0 \\ \varepsilon^{-1} & \varepsilon^{-1} & 0 & 1 \\ 0 & X_{2,2} & 1 & X_{2,3} \\ X_{3,1} & 0 & 0 & X_{3,3} \end{pmatrix}$$

has determinant

$$\det(A_\varepsilon) = h + \varepsilon X_{1,1}X_{2,2}X_{3,1}.$$

So

$$\lim_{\varepsilon \to 0}(\det(A_\varepsilon)) = h.$$

- In other words, h can be approximated arbitrarily closely by determinants of size 4.
- Let $\underline{\mathrm{dc}}(h)$ denote the size of the smallest matrix sequence whose determinant approximates h.
- In this example $\underline{\mathrm{dc}}(h) \leq 4$. It might be $\mathrm{dc}(h) > 4$.
- Landsberg, Manivel, Ressayre 2010: $\underline{\mathrm{dc}}(\mathrm{per}_m) \geq m^2/2$.
- Open question: $5 \leq \underline{\mathrm{dc}}(\mathrm{per}_3) \leq 7$.

Approximations?

- Clearly $\underline{\mathrm{dc}}(\mathrm{per}_m) \le \mathrm{dc}(\mathrm{per}_m)$. But how large is the gap?
- **As far as we know** $\mathrm{dc}(\mathrm{per}_m)$ could grow **superpolynomially** (Valiant's conjecture) while at the same time $\underline{\mathrm{dc}}(\mathrm{per}_m)$ could grow **just polynomially**.
- Mulmuley and Sohoni's conjecture: $\underline{\mathrm{dc}}(\mathrm{per}_m)$ grows superpolynomially.
- Could we prove at least the following implication:

Conjecture (Valiant's conjecture = Mulmuley and Sohoni's conjecture)

If $\underline{\mathrm{dc}}(\mathrm{per}_m)$ is polynomially bounded, then $\mathrm{dc}(\mathrm{per}_m)$ is polynomially bounded.

Remarks:

- In the setting of bilinear complexity one can show that the transition to approximations is harmless: Rank and border rank of matrix multiplication grow with the same order of magnitude ω.
- Most lower bound techniques cannot distsinguish between dc and $\underline{\mathrm{dc}}$.

How lower bounds on <u>dc</u> must look like

- Let V^m denote the vector space of polynomials in m^2 variables of degree m.
- $\mathrm{per}_m \in V^m$.
- $\dim(V^m) = \binom{m^2+m-1}{m}$.
- A basis of V^m is given by the monomials.
- Since V^m is a finite dimensional vector space (with a chosen basis) we have the usual metric on V^m. In particular we can talk about continuous functions $f : V^m \to \mathbb{C}$.
- Elementary point-set topology gives:

Proposition

If $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$, then there exists a **continuous** function $f : V^m \to \mathbb{C}$ such that

- $f(h) = 0$ for all $h \in V^m$ with $\underline{\mathrm{dc}}(h) \leq n$
- and $f(\mathrm{per}_m) \neq 0$.

- Algebraic geometry gives something even stronger:

Proposition

If $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$, then there exists a **polynomial** function $f : V^m \to \mathbb{C}$ such that

- $f(h) = 0$ for all $h \in V^m$ with $\underline{\mathrm{dc}}(h) \leq n$
- and $f(\mathrm{per}_m) \neq 0$.

- And representation theory will give an even stronger proposition on later slides.

Polynomials on spaces of polynomials: A toy example

- A quadratic homogeneous polynomial

$$h := ax^2 + bxy + cy^2, \quad a, b, c \in \mathbb{C}$$

  is the square of a linear form

$$h = (\alpha x + \beta y)^2, \quad \alpha, \beta \in \mathbb{C}$$

  iff its **discriminant** vanishes:

$$f(a, b, c) := b^2 - 4ac = 0.$$

- The case $y = 1$ from high school: $ax^2 + bx + c$ has a double root iff $b^2 - 4ac = 0$.
- The discriminant f is a polynomial whose variables are the coefficients of other polynomials: The polynomial h is interpreted as its coefficient vector $(a, b, c) \in \mathbb{C}^3$.
- **Complexity lower bound (toy version, symmetric rank):**
  If $f(h) \neq 0$, then we need at least 2 summands to write h:

$$h = (\alpha_1 x + \beta_1 y)^2 + (\alpha_2 x + \beta_2 y)^2.$$

Next steps

Recall:

Proposition

If $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$, then there exists a **polynomial** function $f : V^m \to \mathbb{C}$ such that

- $f(h) = 0$ for all $h \in V^m$ with $\underline{\mathrm{dc}}(h) \leq n$
- and $f(\mathrm{per}_m) \neq 0$.

- Better: We can restrict ourselves to homogeneous polynomials f (like the discriminant).
- Representation theory can make an even stronger statement!

Representation Theory

- Recall $V^m = \mathbb{C}[x_{1,1}, x_{1,2}, \dots, x_{m,m}]_m$.
- Let
$$\mathbb{C}[V^m]_d$$
denote the space of homogeneous degree d polynomials whose variables are the degree m monomials in m^2 variables.
- The dimension is very high: $\dim(\mathbb{C}[V^m]_d) = \binom{d+\binom{m^2+m-1}{m}-1}{d}$.
- But: These spaces can be studied with representation theory!
- Example:
For $V = \mathbb{C}[x, y]_2$ we have $\dim(V) = 3$ with basis $a := x^2$, $b := xy$, $c := y^2$.
$\dim(\mathbb{C}[V]_2) = 6$ with basis a^2, ab, ac, b^2, bc, c^2.

Isotypic components

- $\mathbb{C}[V^m]_d$ decomposes uniquely into the sum of **isotypic components** $\mathscr{W}_\lambda$ and we only have to search for f inside isotypic components:

$$\mathbb{C}[V^m]_d = \bigoplus_\lambda \mathscr{W}_\lambda$$

  The sum is over all **partitions** λ of dm into at most m^2 parts.
- For example, if $d = 5$, $m = 2$, then $(5, 3, 1, 1)$ is a partition of 10 into 4 parts.
- In each isotypic component we only have to look at so-called **highest weight vectors** if we want to prove lower bounds.
- Example: For $V = \mathbb{C}[x, y]$ the vector space $\mathbb{C}[V]_2$ decomposes into two isotypic components.
  - The discriminant $b^2 - 4ac$ is a highest weight vector living in a 1-dim isotypic component. Here $\lambda = (2, 2)$.
  - The polynomial a^2 is another one, living in a 5-dim isotypic component. Here $\lambda = (4, 0)$.

Group actions

- Recall the example $f = b^2 - 4ac, \quad a = x^2,\ b = xy,\ c = y^2$.
- Let us permute x and y in f and write

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} f = b^2 - 4ca = f,$$

  so f does not change if we permute x and y.
- Let us scale x by $\gamma \in \mathbb{C}$ and y by $\delta \in \mathbb{C}$:

$$\begin{pmatrix} \gamma & 0 \\ 0 & \delta \end{pmatrix} f = (\gamma\delta b)^2 - 4(\gamma^2 a \delta^2 c) = \gamma^2\delta^2 f,$$

  so under this operation f gets scaled by $\gamma^2\delta^2$.
  The vector of scaling exponents is $(2, 2)$.
- The scaling exponent of a^2 is $(4, 0)$.
- Upper triangular matrices fix a^2:

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} a^2 = a^2$$

  because this matrix sends x to x and y to $\alpha x + y$.
- For any matrix $g \in \mathrm{GL}(\mathbb{C}^{m^2})$ and a polynomial $f \in V^m$ we can define gf in a natural way.

Isotypic components and highest weight vectors

Definition

$f \in \mathbb{C}[V^m]_d$ is called a **highest weight vector** if:

- f does not change under the action of any upper triangular matrices,
- and f gets scaled under the action of diagonal matrices.

The vector of scaling exponents is called the **type** λ of f.

- In the example, $b^2 - 4ac$ is a highest weight vector of type $(2, 2)$ and a^2 is a highest weight vector of type $(4, 0)$.
- Remark: Highest weight vectors of the same type form a vector space.
- Highest weight vectors of type λ lie in the isotypic component $\mathscr{W}_\lambda$.

Proposition (Lower bounds are always given by highest weight vectors)

If $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$, then there exists a a highest weight vector f of some type λ that vanishes on all $h \in V^m$ with $\underline{\mathrm{dc}}(h) \leq n$ and $(gf)(\mathrm{per}_m) \neq 0$ for some matrix $g \in \mathrm{GL}(\mathbb{C}^{m^2})$.

There are concrete algorithms for constructing highest weight vectors via multilinear algebra.

How could we find obstructions?

Let $V^m(n)$ denote the set of points $h \in V^m$ with $\underline{\mathrm{dc}}(h) \leq n$.

- To simplify the study of highest weight vectors Mulmuley and Sohoni introduced the following approach:

Proposition (Occurrence Obstruction Approach)

For a partition λ, if **all** highest weight vectors f of type λ vanish on $V^m(n)$, and if one of them satisfies $(gf)(\mathrm{per}_m) \neq 0$ for some matrix $g \in \mathrm{GL}(\mathbb{C}^{m^2})$, then $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$.

- The vanishing of all highest weight vectors could be easier to study than analyzing specific highest weight vectors.
- A sufficient criterion for the vanishing of all highest weight vectors is also given:

Theorem (Mulmuley and Sohoni)

If the **rectangular Kronecker coefficient** $g(\lambda, d, m)$ is zero, then all highest weight vectors of type λ vanish on $V^m(n)$.

Def. (via representation theory): $g(\lambda, d, m)$ is the multiplicity of the irreducible Specht module $[\lambda]$ in the tensor product $[d^m] \otimes [d^m]$.

Kronecker coefficients

Theorem (Mulmuley and Sohoni)

If the **rectangular Kronecker coefficient** $g(\lambda, d, m)$ is zero, then all highest weight vectors of type λ vanish on $V^m(n)$.

- Studied since the 1950s, many papers that treat special cases, but mostly not understood.

Theorem (with Mulmuley and Walter, August 2015)

Deciding positivity of the Kronecker coefficient is NP-hard.

- Proof: In a certain subcase we can interpret the Kronecker coeff. combinatorially and show NP-hardness.
- Open question: Is the function $g(\lambda, d, n)$ in #P?
- For the general Kronecker coefficient, containment in #P is problem 10 in Stanley's (2000) list of "outstanding open problems in algebraic combinatorics related to positivity"

No superquartic lower bounds using the vanishing of Kronecker coefficients

Theorem (Mulmuley and Sohoni)

If the **rectangular Kronecker coefficient** $g(\lambda, d, m)$ is zero and if at least one highest weight vectors f of type λ satisfies $(gf)(\mathrm{per}_m) \neq 0$ for some $g \in \mathrm{GL}(\mathbb{C}^{m^2})$, then $\underline{\mathrm{dc}}(\mathrm{per}_m) > n$.

- The recent paper with Greta Panova shows that this **does not give lower bounds** better than $\Omega(m^4)$:

Theorem (with Panova, December 2015)

The vanishing of rectangular Kronecker coefficients $g(\lambda, d, m)$ cannot be used to prove superquartic lower bounds on $\underline{\mathrm{dc}}(\mathrm{per}_m)$.

Proof idea: Show that in all relevant cases either $g(\lambda, d, m) > 0$ or no highest weight vector of type λ exists.

No superquartic lower bounds using the vanishing of Kronecker coefficients

Theorem (with Panova, December 2015)

The vanishing of rectangular Kronecker coefficients $g(\lambda, d, m)$ cannot be used to prove superquartic lower bounds on $\underline{\mathrm{dc}}(\mathrm{per}_m)$.

Proof idea: Show that in all relevant cases either $g(\lambda, d, m) > 0$ or no highest weight vector of type λ exists.

The main ingredients of the proof:

- A result by Kadish and Landsberg (2012) about "padded" polynomials. This gives significant restrictions on the possible λ that could be used for proving lower bounds.
- Stability properties of the Kronecker coefficient (Manivel 2011) enabled us to prove a lower bound on the possible degree d: If there are obstructions in a low degree, then $\underline{\mathrm{dc}}(\mathrm{per}_m)$ is infinite.
- A representation theoretic result by Bessenrodt and Behns (2004), proved using character theory.
- The Kronecker semigroup property.

Open Questions

- Indeed, the approach via vanishing Kronecker coefficients does not work.
- Even worse: Many partial results from [Ik., Panova 2015] also work for other multiplicities. So it is unlikely that any vanishing of multiplicities can prove superpolynomial lower bounds on $\underline{\mathrm{dc}}(\mathrm{per}_m)$.
- But we still know: If $\underline{\mathrm{dc}}(\mathrm{per}_m)$ grows superpolynomially, then there are highest weight vectors proving this.
- The best lower bound by Landsberg, Manivel, Ressayre specifies the highest weight vector.

Open Question

Given m, d, λ, what is the dimension of the vector space of highest weight vectors of type λ in $\mathbb{C}[V^m]_d$?
This dimension is called the **plethysm coefficient**.
Containment in #P is problem 9 in Stanley's list from 2000.

Separating using multiplicities

- Since the determinant polynomial and the permanent polynomial are characterized by their symmetries, one could conjecture that the **dimensions** of the highest weight vector spaces should be sufficient to prove superpolynomial growth of $\underline{\mathrm{dc}}(\mathrm{per}_m)$.
- Results that study this possibility are rare: Larsen and Pink (1990) study group orbits.
- Sometimes this works, as seen in the lower bound for the border rank of matrix multiplication (with Bürgisser 2011, 2013).

Task

Are these dimensions enough to prove lower bounds? Can we settle this question in a similar but simpler setting than det vs per, for example for tensor rank or symmetric rank.
This will not be easy: One can construct artificial settings where this does **not** work, but those settings are very different in nature from det vs per. In particular the points there are not defined by their symmetries.

Separating using multiplicities

- Since the determinant polynomial and the permanent polynomial are characterized by their symmetries, one could conjecture that the **dimensions** of the highest weight vector spaces should be sufficient to prove superpolynomial growth of $\underline{\mathrm{dc}}(\mathrm{per}_m)$.
- Results that study this possibility are rare: Larsen and Pink (1990) study group orbits.
- Sometimes this works, as seen in the lower bound for the border rank of matrix multiplication (with Bürgisser 2011, 2013).

Task

Are these dimensions enough to prove lower bounds? Can we settle this question in a similar but simpler setting than det vs per, for example for tensor rank or symmetric rank.
This will not be easy: One can construct artificial settings where this does **not** work, but those settings are very different in nature from det vs per. In particular the points there are not defined by their symmetries.

Thank you.