Geometric Complexity Theory: No Occurrence Obstructions for Determinant vs Permanent

Peter Bürgisser

joint work with Christian Ikenmeyer and Greta Panova

Technische Universität Berlin

Simons Institute, Berkeley

December 3, 2018

Problem and Main Result

Permanent versus determinant

- How many arithmetic operations are sufficient to evaluate the permanent of an m by m matrix (x_{ij})?

$$\mathrm{per}_m := \sum_{\pi \in S_m} x_{1\pi(1)} \cdots x_{m\pi(m)}$$

- Best known algorithm: $O(m2^m)$ operations
- The determinant $\det_n$ can be evaluated with $\mathrm{poly}(n)$ operations

$$\det_n := \sum_{\pi \in S_n} sgn(\pi)\, x_{1\pi(1)} \cdots x_{n\pi(n)}$$

- Work over $\mathbb{C}$

Valiant's Conjecture

- Are there linear forms $a_{ij} = a_{ij}(x, z)$ in x_{ij} and z such that $(n \geq m)$

$$z^{n-m}\mathrm{per}_m = \det \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \qquad ? \qquad\qquad (*)$$

- Impossible for $n = m > 2$ (Polya)
- Possible for $n \leq 2^m - 1$ (Valiant, Grenet)
- $n \geq \frac{1}{2}m^2$ (Mignon & Ressayre 2004)

- $\boxed{\text{Valiant's Conjecture (1979): (*) impossible for } n = poly(m)}$

- Conjecture equivalent to the separation $\mathrm{VP}_{ws} \neq \mathrm{VNP}$ of complexity classes
- $\mathrm{P} \neq \mathrm{NP}$ implies $\mathrm{VP}_{ws} \neq \mathrm{VNP}$ under GRH (B, 2000)

Orbit closure of $\det_n$

- Approach by Mulmuley and Sohoni (2001) based on algebraic geometry and representation theory
- Idea of orbit closures already in Strassen (1987) for tensor rank
- nth symmetric power $\mathrm{Sym}^n V^*$ of dual space V^* with natural action of group $G := \mathrm{GL}(V)$
- Orbit $G \cdot f := \{gf \mid g \in G\}$ of $f \in \mathrm{Sym}^n V^*$
- Take $V := \mathbb{C}^{n\times n}$, $N = n^2$, view $\det_n$ as element of $\mathrm{Sym}^n V^*$
- Orbit closure w.r.t. Euclidean or Zariski topology

$$\Omega_n := \overline{\mathrm{GL}_{n^2} \cdot \det_n} \subseteq \mathrm{Sym}^n(\mathbb{C}^{n\times n})^*$$

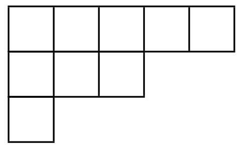
- $\Omega_2 = \mathrm{Sym}^2(\mathbb{C}^{2\times 2})^*$; Ω_3 known (Hüttenhain & Lairez '16); Ω_4 already unknown

Orbit Closure Conjecture

- Padded permanent $X_{11}^{n-m}\mathrm{per}_m \in \mathrm{Sym}^n(\mathbb{C}^{n\times n})^*$, where $n > m$
- **Orbit Closure Conjecture (M-S 2001)**

  For all $c \in \mathbb{N}_{\geq 1}$ we have $X_{11}^{m^c-m}\mathrm{per}_m \notin \Omega_{m^c}$ for infinitely many m.
- The Orbit Closure Conjecture implies Valiant's Conjecture

Splitting into irreps

- Action of group $G = \mathrm{GL}(V)$ on $\mathrm{Sym}^n V^*$ induces action on its graded coordinate ring $\mathbb{C}[\mathrm{Sym}^n V^*] = \oplus_{d \in \mathbb{N}} \mathrm{Sym}^d \mathrm{Sym}^n V$
- The plethysms $\mathrm{Sym}^d \mathrm{Sym}^n V$ splits into irreducible G-representations $\mathcal{W}_\lambda$ (Weyl modules), labeled by partitions $\lambda \vdash dn$ into at most $\dim V = n^2$ parts
- Visualize partition as Young diagram: $(5, 3, 1) \vdash 9$ write as 
- Size $|(5, 3, 1)| := 9$ is number of boxes; length $\ell(5, 3, 1) = 3$ is number of parts
- $\mathbb{C}[\Omega_n]$ denotes coordinate ring of Ω_n
- Restriction of polynomial maps to Ω_n gives surjective G-equivariant linear map:

$$\mathrm{Sym}^d \mathrm{Sym}^n V = \mathbb{C}[\mathrm{Sym}^n V^*] \twoheadrightarrow \mathbb{C}[\Omega_n]_d$$

- Say λ occurs in $\mathbb{C}[\Omega_n]_d$ if it contains a copy of $\mathcal{W}_\lambda$

Obstructions

- $Z_{n,m}$ denotes orbit closure of the padded permanent ($n > m$):

$$Z_{n,m} := \overline{\mathrm{GL}_{n^2} \cdot X_{11}^{n-m} \mathrm{per}_m} \subseteq \mathrm{Sym}^n(\mathbb{C}^{n \times n})^*. \tag{1}$$

- Suppose $X_{11}^{n-m} \mathrm{per}_m \in \Omega_n$
- Then $Z_{n,m} \subseteq \Omega_n$ and restriction gives $\mathbb{C}[\Omega_n] \twoheadrightarrow \mathbb{C}[Z_{n,m}]$
- Schur's lemma: if λ occurs in $\mathbb{C}[Z_{n,m}]$, then λ occurs in $\mathbb{C}[\Omega_n]$
- Partition λ violating this condition is called occurrence obstruction.
- Its existence would prove $Z_{n,m} \not\subseteq \Omega_n$
- Schur's lemma also gives inequality of multiplicities:

$$\mathrm{mult}_\lambda \mathbb{C}[\Omega_n] \;\geq\; \mathrm{mult}_\lambda \mathbb{C}[Z_{n,m}]$$

- Partition λ violating this inequality is called multiplicity obstruction.

Main Result

M-S suggested the following conjecture

Occurrence Obstruction Conjecture (M-S 2001)

For all $c \in \mathbb{N}_{\geq 1}$, for infinitely many m, there exists a partition λ occurring in $\mathbb{C}[Z_{m^c,m}]$ but not in $\mathbb{C}[\Omega_{m^c}]$.

Occurrence Obstruction Conjecture implies Orbit Closure Conjecture

Unfortunately, the Occurrence Obstruction Conjecture is false!

Thm. (B, Ikenmeyer, Panova, FOCS 16, J. AMS '18)

Let n, d, m be positive integers with $n \geq m^{25}$ and $\lambda \vdash nd$. If λ occurs in $\mathbb{C}[Z_{n,m}]$, then λ also occurs in $\mathbb{C}[\Omega_n]$. In particular, the Occurrence Obstruction Conjecture is false.

Before this, [IP16] (Ikenmeyer, Panova FOCS 16) had a similar result showing that the Orbit Closure Conjecture cannot be resolved via Kronecker coefficients

No occurrence obstructions for Waring rank

- Waring rank (symmetric tensor rank) of $p \in \mathrm{Sym}^n V^*$: minimum r s.t. $p = \varphi_1^n + \ldots + \varphi_r^n$ for linear forms $\varphi_i \in V^*$
- Can prove exponential lower bound on Waring rank of $\det_n$, per_n
- May think of proving lower bounds on Waring rank by studying orbit closure

$$\mathrm{PS}_n := \overline{\mathrm{GL}_{n^2} \cdot (X_1^n + \cdots + X_{n^2}^n)} \subseteq \mathrm{Sym}^n(\mathbb{C}^{n^2})^*.$$

Corollary

Let n, d, m be positive integers with $n \geq m^{25}$ and $\lambda \vdash nd$. If λ occurs in $\mathbb{C}[Z_{n,m}]$, then λ also occurs in $\mathbb{C}[\mathrm{PS}_n]$. Moreover, the permanent can be replaced by any homogeneous polynomial p of degree m in m^2 variables.

Hence strategy of occurrence obstructions cannot even be used in weak model of PS_n against padded polynomials!

Outline and Ingredients of Proof

Kadish & Landsberg's observation

- body $\bar{\lambda}$ of λ: obtained by removing the first row of λ,

> **Kadish & Landsberg '14**
>
> If $\lambda \vdash nd$ occurs in $\mathbb{C}[Z_{n,m}]_d$, then $\ell(\lambda) \leq m^2$ and $|\bar{\lambda}| \leq md$.

- $|\bar{\lambda}| \leq md$ is equivalent to $\lambda_1 \geq (n-m)d$: λ must have a very long first row if n is substantially larger than m
- This is the only information we exploit about the orbit closure $Z_{n,m}$ of the padded permanent
- Can replace the permanent by any homogeneous polynomial p of degree m in m^2 variables
- Kadish & Landsberg also crucially used in [IP16]

Semigroup property

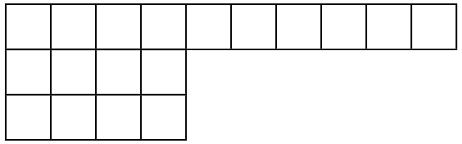
- Need to show that many partitions λ occur in $\mathbb{C}[\Omega_n]$
- For this establish the occurrence of certain basic shapes in $\mathbb{C}[\Omega_n]$
- Then get more shapes by

Semigroup Property

If λ occurs in $\mathbb{C}[\Omega_n]$ and μ occurs in $\mathbb{C}[\Omega_n]$, then $\lambda + \mu$ occurs in $\mathbb{C}[\Omega_n]$.

- Also crucially used in [IP16]

Basic building blocks

- Denote by $(k \times \ell)^{\sharp nk}$ the rectangular diagram $k \times \ell$ with k rows of length ℓ, to which a row has been appended s.t. we ge nk boxes
- $$(3 \times 4)^{\sharp 18} =$$ 

- **Prop. RER (Row Extended Rectangles)**

  Let $n \geq k\ell$ and ℓ be even. Then $(k \times \ell)^{\sharp nk}$ occurs in $\mathbb{C}[\Omega_n]_k$.

- The only property of Ω_n used in the proof is that Ω_n contains many padded power sums (follows from universality of determinant)

Prop. PPS (Padded Power Sums)

- Let $X, \varphi_1, \ldots, \varphi_k$ be linear forms on $\mathbb{C}^{n \times n}$ and assume $n \geq sk$. Then the power sum $X^{n-s}(\varphi_1^s + \cdots + \varphi_k^s)$ of k terms of degree s, padded to degree n, is contained in Ω_n.

Strategy of proof of main result

- Suppose have even $\lambda \vdash nd$ such that $n \geq m^{25}$ and λ occurs in $\mathbb{C}[Z_{n,m}]$. Want to show that λ occurs in $\mathbb{C}[\Omega_n]$.
- By [KL14] we have $\ell(\lambda) \leq m^2$ and $|\bar{\lambda}| \leq md$.
- Distinguish two cases
- CASE 1: If the degree d is large (say $d \geq 24m^6$), we proceed as in [IP16]: we decompose body $\bar{\lambda}$ into a sum of even rectangles
- Since n and d are sufficiently large in comparison with m, can write (!) λ as a sum of row extended rectangles $(k \times \ell)^{\sharp nk}$, where $n \geq k\ell$.
- By Prop. RER the row extended rectangles occur in $\mathbb{C}[\Omega_n]_k$. The semigroup property implies that λ occurs in $\mathbb{C}[\Omega_n]_d$.

Case of small degree

- CASE 2: If the degree d is small, we rely on the following crucial result. Recall $V = \mathbb{C}^{n\times n}$.

Prop. ALL

Let $\lambda \vdash nd$ be such that $|\bar{\lambda}| \leq md$ and $md^2 \leq n$ for some m.

Then every highest weight vector of weight λ in $\mathrm{Sym}^d \mathrm{Sym}^n V$, viewed as a degree d polynomial function on $\mathrm{Sym}^n V^*$, does not vanish on Ω_n.

In particular, if λ occurs in $\mathrm{Sym}^d \mathrm{Sym}^n V$, then λ occurs in $\mathbb{C}[\Omega_n]_d$.

- The proof relies on new insights on "lifting highest weight vectors" in plethysms
- This is related to known stability property of plethysms, for which we obtain new proofs
- For treating noneven partitions, need more bulding blocks (row and column extended rectangles) and more tricks

Fundamental Invariants and Lifting of Highest Weight Vectors

Highest weight vectors

- How to show that λ occurs in $\mathbb{C}[\Omega_n]$?
- $F \in \mathrm{Sym}^d \mathrm{Sym}^n \mathbb{C}^N$ called highest weight vector of weight λ if

$$\begin{pmatrix} t_1 & * & * & * \\ & t_2 & * & * \\ & & \ddots & \vdots \\ & & & t_N \end{pmatrix} \cdot F = t_1^{\lambda_1} \cdots t_N^{\lambda_N} F \qquad \text{for all } t_i \in \mathbb{C}^*$$

- F is invariant under SL_N iff λ is rectangular: $\lambda_1 = \ldots = \lambda_N$
- View F as homogeneous degree d polynomial function

$$F \colon \mathrm{Sym}^n(\mathbb{C}^N)^* \to \mathbb{C}, \quad F(p) = \langle F, p^n \rangle$$

- Essential observation:

If $F(p) \neq 0$, then λ occurs in $\mathbb{C}[\overline{\mathrm{GL}_N \cdot p}]$

Fundamental invariants

- Suppose n is even. Howe ('87) showed:
- If $d < N$, then $\mathrm{Sym}^d \mathrm{Sym}^n \mathbb{C}^N$ doesn't have a nonzero SL_N-invariant
- If $d = N$, then $\mathrm{Sym}^d \mathrm{Sym}^n \mathbb{C}^N$ has exactly one SL_N-invariant $F_{n,N}$, up to scaling, the fundamental invariant, already known to Cayley as a "hyperdeterminant"
- View $F_{n,N}$ as a homogeneous degree N polynomial map

$$F_{n,N} \colon \mathrm{Sym}^n (\mathbb{C}^N)^* \to \mathbb{C}$$

- For $p = \sum_{1 \le j_1, \dots, j_n \le N} v(j_1, \dots, j_n) X_{j_1} \cdots X_{j_n}$ with symmetric coefficients

$$F_{n,N}(p) = \sum_{\sigma_1, \dots, \sigma_n \in S_N} \mathrm{sgn}(\sigma_1) \cdots \mathrm{sgn}(\sigma_n) \prod_{i=1}^{N} v(\sigma_1(i), \dots, \sigma_n(i))$$

- For $g \in \mathrm{GL}_N$

$$F_{n,N}(g \cdot p) = \det(g)^n F_{n,N}(p)$$

- Ex. $n = 2$: $F_{2,N}(p) = N! \det(v)$ where v is symmetric matrix

Evaluating fundamental invariants

- [B, Ikenmeyer '17]: systematic investigation of fundamental invariants
- $F_{n,N}$ is a highest weight vector (weight $N \times n$)
- It is not easy to prove $F_{n,N}(p) \neq 0$
- Seemingly simple example (n even)

$$F_{n,n}(X_1 \cdots X_n) = \frac{1}{n!}\big(\#\{\text{col. even latin squares}\} - \#\{\text{col. odd latin squares}\}\big) \stackrel{?}{=} 0$$

- This is unknown: Alon-Tarsi Conjecture!
- Essential for basic building blocks: prove $F_{n,N}(X_1^n + \ldots + X_N^n) \neq 0$ by writing it as sum of squares [B, Christandl, Ikenmeyer '11]

Lifiting in plethysms

- Construct explicit injective linear lifting map for $n \geq m$

$$\kappa^d_{m,n} \colon \mathrm{Sym}^d \mathrm{Sym}^m V \to \mathrm{Sym}^d \mathrm{Sym}^n V$$

- $\kappa^d_{m,n}$ defined as d-fold symmetric power of linear map

$$M \colon \mathrm{Sym}^m V \to \mathrm{Sym}^n V,\ p \mapsto p\, e_1^{n-m}$$

  multiplication with e_1^{n-m}, 1st standard basis vector $e_1 \in V = \mathbb{C}^N$
- Use duality to show for $f \in \mathrm{Sym}^d \mathrm{Sym}^m V$, $q \in \mathrm{Sym}^n V^*$,

$$\langle \kappa^d_{m,n}(f), q^d \rangle = \langle f, M^*(q)^d \rangle$$

  Here $M^* \colon \mathrm{Sym}^n V^* \to \mathrm{Sym}^m V^*$ denotes dual map of M.
- $M^*(q)$ is $(n-m)$-fold partial derivative of q in direction e_1 (times $m!/n!$)

Highest weight vectors in plethysms

- Proved that lifting

$$\kappa^d_{m,n} \colon \mathrm{Sym}^d \mathrm{Sym}^m V \to \mathrm{Sym}^d \mathrm{Sym}^n V \,,$$

  maps highest weight vectors of weight $\mu \vdash md$ to highest weight vectors of weight $\mu^{\sharp dn}$ (μ with extended 1st row)
- Constructed system of generators v_T of space of highest weight vectors of weight μ, labelled by tableaux T of shape $\mu \vdash dm$ with d letters, each occuring m times (no letter appears more than once in a column)
- Proved: $\kappa^d_{m,n}$ maps generator v_T to generator $v_{T'}$ where T' arises from T by adding in the first row $n - m$ copies of each of the d letters
- Side result: new proof of known stability property of plethysms

Corollary on lifting

Cor. Lift

Suppose $\lambda \vdash nd$ satisfies $\lambda_2 \leq m$ and $\lambda_2 + |\bar{\lambda}| \leq md$. Then every highest weight vector of weight λ is obtained as a lifting.

Proof.

- $\lambda_2 + |\bar{\lambda}| \leq md$ is number of boxes of λ that appear in non-singleton columns
- Hence λ is obtained by extending the 1st row of some $\mu \vdash md$
- Let T' be a tableau of shape λ with d letters, each occuring m times. Since no letter appears more than once in a column, each of the d letters appears at least $n - \lambda_2 \geq n - m$ times in singleton columns. Hence T' is obtained from a tableau T of shape μ as before
- From before: $\kappa^d_{m,n}(v_T) = v_{T'}$
- Moreover, the $v_{T'}$ generate space of hwv of weight λ $\square$

Proof of Prop. ALL

Prop. ALL

$\lambda \vdash nd$ s.t. $|\bar{\lambda}| \leq md$ and $md^2 \leq n$. Then every highest weight vector of weight λ in $\mathrm{Sym}^d \mathrm{Sym}^n V$ does not vanish on Ω_n.

Proof.

- Let $h \in \mathrm{Sym}^d \mathrm{Sym}^n V$ be hwv of weight λ
- $\lambda_2 \leq |\bar{\lambda}| \leq md$ and $\lambda_2 + |\bar{\lambda}| \leq 2|\bar{\lambda}| \leq 2md \leq md \cdot d$
- Cor. Lift applied to $\mathrm{Sym}^d \mathrm{Sym}^{md} V \to \mathrm{Sym}^d \mathrm{Sym}^n V$ shows $h = \kappa^d_{md,n}(f)$ for some hwv $f \in \mathrm{Sym}^d \mathrm{Sym}^{md} V$ of weight λ
- Can show that for almost all power sums $p = \varphi_1^{md} + \cdots + \varphi_d^{md}$ we have $\langle f, p^d \rangle \neq 0$ and with $q := X_1^{n-md} p$,
$$\langle f, M^*(q)^d \rangle \neq 0$$
- Using duality
$$\langle h, q^d \rangle = \langle \kappa^d_{m,n}(f), q^d \rangle = \langle f, M^*(q)^d \rangle \neq 0.$$
By Prop. PPS, we have $q \in \Omega_n$ since $n \geq md \cdot d$. $\square$

Thank you for your attention!