Geometric Complexity Theory: No Occurrence Obstructions for Determinant vs Permanent

Peter Bürgisser
joint work with Christian Ikenmeyer and Greta Panova

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 \((x_{ij})\)?

\[
\text{per}_m := \sum_{\pi \in S_m} x_{1\pi(1)} \cdots x_{m\pi(m)}
\]

- Best known algorithm: \( O(m2^m) \) operations

- The determinant \( \text{det}_n \) can be evaluated with \( \text{poly}(n) \) operations

\[
\text{det}_n := \sum_{\pi \in S_n} \text{sgn}(\pi) x_{1\pi(1)} \cdots x_{n\pi(n)}
\]

- Work over \( \mathbb{C} \)
Valiant’s Conjecture

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

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

- 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)

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

- Conjecture equivalent to the separation $\text{VP}_{ws} \neq \text{VNP}$ of complexity classes
- $\text{P} \neq \text{NP}$ implies $\text{VP}_{ws} \neq \text{VNP}$ 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
- $n$th symmetric power $\text{Sym}^n V^*$ of dual space $V^*$ with natural action of group $G := \text{GL}(V)$
- Orbit $G \cdot f := \{gf \mid g \in G\}$ of $f \in \text{Sym}^n V^*$
- Take $V := \mathbb{C}^{n \times n}$, $N = n^2$, view $\det_n$ as element of $\text{Sym}^n V^*$
- Orbit closure w.r.t. Euclidean or Zariski topology

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

- $\Omega_2 = \text{Sym}^2(\mathbb{C}^{2 \times 2})^*$; $\Omega_3$ known (Hüttenhain & Lairez ’16); $\Omega_4$ already unknown
Orbit Closure Conjecture

- **Padded permanent** $X_{11}^{n-m} \text{per}_m \in \text{Sym}^n(\mathbb{C}^{n \times n})^*$, where $n > m$

- **Orbit Closure Conjecture (M-S 2001)**

  For all $c \in \mathbb{N}_{\geq 1}$ we have $X_{11}^{m^c-m} \text{per}_m \notin \Omega_{m^c}$ for infinitely many $m$.

- The Orbit Closure Conjecture implies Valiant’s Conjecture
Splitting into irreps

- Action of group $G = \text{GL}(V)$ on $\text{Sym}^n V^*$ induces action on its graded coordinate ring $\mathbb{C}[\text{Sym}^n V^*] = \bigoplus_{d \in \mathbb{N}} \text{Sym}^d \text{Sym}^n V$

- The plethysms $\text{Sym}^d \text{Sym}^n V$ splits into irreducible $G$-representations $W_\lambda$ (Weyl modules), labeled by partitions $\lambda \vdash dn$ into at most $\text{dim } V = n^2$ 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 $\Omega_n$

- Restriction of polynomial maps to $\Omega_n$ gives surjective $G$-equivariant linear map:

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

- Say $\lambda$ occurs in $\mathbb{C}[\Omega_n]_d$ if it contains a copy of $W_\lambda$
No Occurrence Obstructions in GCT

Problem and Main Result

Obstructions

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

\[
Z_{n,m} := \overline{\GL_{n^2} \cdot X_{11}^{n-m} \per_m} \subseteq \Sym^n(\mathbb{C}^{n \times n})^*. \tag{1}
\]

- Suppose \( X_{11}^{n-m} \per_m \in \Omega_n \)
- Then \( Z_{n,m} \subseteq \Omega_n \) and restriction gives \( \mathbb{C}[\Omega_n] \rightarrow \mathbb{C}[Z_{n,m}] \)
- Schur’s lemma: if \( \lambda \) occurs in \( \mathbb{C}[Z_{n,m}] \), then \( \lambda \) occurs in \( \mathbb{C}[\Omega_n] \)
- Partition \( \lambda \) 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:

\[
\mult_{\lambda} \mathbb{C}[\Omega_n] \geq \mult_{\lambda} \mathbb{C}[Z_{n,m}]
\]

- Partition \( \lambda \) 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 $\lambda$ 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 $\lambda$ occurs in $\mathbb{C}[Z_{n,m}]$, then $\lambda$ 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 \( p \in \text{Sym}^n V^* \): 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

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

**Corollary**

Let \( n, d, m \) be positive integers with \( n \geq m^{25} \) and \( \lambda \vdash nd \). If \( \lambda \) occurs in \( \mathbb{C}[\mathbb{Z}_n,m] \), then \( \lambda \) also occurs in \( \mathbb{C}[\text{PS}_n] \). 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 \( \text{PS}_n \) against padded polynomials!
Outline and Ingredients of Proof
Kadish & Landsberg’s observation

- body $\lambda$ of $\lambda$: obtained by removing the first row of $\lambda$,

**Kadish & Landsberg ’14**

If $\lambda \vdash nd$ occurs in $\mathbb{C}[Z_{n,m}]_d$, then $\ell(\lambda) \leq m^2$ and $|\lambda| \leq md$.

- $|\lambda| \leq md$ is equivalent to $\lambda_1 \geq (n - m)d$: $\lambda$ 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 $\lambda$ 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 $\lambda$ occurs in $\mathbb{C}[\Omega_n]$ and $\mu$ 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 \((k \times \ell)^{nk}\) the rectangular diagram \(k \times \ell\) with \(k\) rows of length \(\ell\), to which a row has been appended s.t. we get \(nk\) boxes

\[
(3 \times 4)^{18} = \begin{array}{cccc}
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\end{array}
\]

Prop. RER (Row Extended Rectangles)

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

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 \(\Omega_n\).
No Occurrence Obstructions in GCT

Outline and Ingredients

Strategy of proof of main result

- Suppose have even $\lambda \vdash nd$ such that $n \geq m^{25}$ and $\lambda$ occurs in $\mathbb{C}[Z_{n,m}]$. Want to show that $\lambda$ 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 (!) $\lambda$ as a sum of row extended rectangles $(k \times \ell)^{nk}$, where $n \geq k\ell$.
- By Prop. RER the row extended rectangles occur in $\mathbb{C}[\Omega_n]_k$. The semigroup property implies that $\lambda$ 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 $|\lambda| \leq md$ and $md^2 \leq n$ for some $m$. Then every highest weight vector of weight $\lambda$ in $\text{Sym}^d \text{Sym}^n V$, viewed as a degree $d$ polynomial function on $\text{Sym}^n V^*$, does not vanish on $\Omega_n$.

In particular, if $\lambda$ occurs in $\text{Sym}^d \text{Sym}^n V$, then $\lambda$ 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 building blocks (row and column extended rectangles) and more tricks
Fundamental Invariants and Lifting of Highest Weight Vectors
Highest weight vectors

- How to show that $\lambda$ occurs in $\mathbb{C}[\Omega_n]$?
- $F \in \text{Sym}^d \text{Sym}^n \mathbb{C}^N$ called highest weight vector of weight $\lambda$ if
  \[
  \begin{pmatrix}
  t_1 & * & * & * \\
  * & t_2 & * & * \\
  * & * & \ddots & * \\
  * & * & * & t_N
  \end{pmatrix} \cdot F = t_1^{\lambda_1} \cdots t_N^{\lambda_N} F \quad \text{for all } t_i \in \mathbb{C}^*
  \]

- $F$ is invariant under $\text{SL}_N$ iff $\lambda$ is rectangular: $\lambda_1 = \ldots = \lambda_N$
- View $F$ as homogeneous degree $d$ polynomial function
  \[
  F : \text{Sym}^n(\mathbb{C}^N)^* \to \mathbb{C}, \quad F(p) = \langle F, p^n \rangle
  \]

- Essential observation:
  If $F(p) \neq 0$, then $\lambda$ occurs in $\mathbb{C}[\text{GL}_N \cdot p]$
Fundamental invariants

- Suppose $n$ is even. Howe (’87) showed:
- If $d < N$, then $\text{Sym}^d \text{Sym}^n \mathbb{C}^N$ doesn’t have a nonzero $\text{SL}_N$-invariant
- If $d = N$, then $\text{Sym}^d \text{Sym}^n \mathbb{C}^N$ has exactly one $\text{SL}_N$-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

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

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

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

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

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

- 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)

\[
F_{n,n}(X_1 \cdots X_n) = \frac{1}{n!} \left( \#\{\text{col. even latin squares}\} - \#\{\text{col. odd latin squares}\} \right) \equiv 0
\]

- 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]
Lifting in plethysms

- **Construct** explicit injective linear **lifting map** for \( n \geq m \)

\[
\kappa_{m,n}^d : \text{Sym}^d \text{Sym}^m V \rightarrow \text{Sym}^d \text{Sym}^n V
\]

- \( \kappa_{m,n}^d \) defined as \( d \)-fold symmetric power of linear map

\[
M : \text{Sym}^m V \rightarrow \text{Sym}^n V, \ p \mapsto p e_1^{n-m}
\]

multiplication with \( e_1^{n-m} \), 1st standard basis vector \( e_1 \in V = \mathbb{C}^N \)

- Use duality to show for \( f \in \text{Sym}^d \text{Sym}^m V, \ q \in \text{Sym}^n V^* \),

\[
\langle \kappa_{m,n}^d(f), q^d \rangle = \langle f, M^*(q)^d \rangle
\]

Here \( M^* : \text{Sym}^n V^* \rightarrow \text{Sym}^m V^* \) denotes dual map of \( M \).

- \( M^*(q) \) is \((n - m)\)-fold partial derivative of \( q \) in direction \( e_1 \) (times \( m!/n! \))
No Occurrence Obstructions in GCT

Fund. Invariants and Lifting

Highest weight vectors in plethysms

- Proved that lifting

\( \kappa_{m,n}^d : \text{Sym}^d \text{Sym}^m V \rightarrow \text{Sym}^d \text{Sym}^n V, \)

maps highest weight vectors of weight \( \mu \vdash md \) to highest weight vectors of weight \( \mu \vdash^d n \) (\( \mu \) with extended 1st row)

- Constructed system of generators \( v_T \) of space of highest weight vectors of weight \( \mu \), 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: \( \kappa_{m,n}^d \) 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 \( \lambda_2 + |\lambda| \leq md \). Then every highest weight vector of weight \( \lambda \) is obtained as a lifting.

Proof.

1. \( \lambda_2 + |\lambda| \leq md \) is number of boxes of \( \lambda \) that appear in non-singleton columns
2. Hence \( \lambda \) is obtained by extending the 1st row of some \( \mu \vdash md \)
3. Let \( T' \) be a tableau of shape \( \lambda \) with \( d \) letters, each occurring \( m \) 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 \( \mu \) as before
4. From before: \( \kappa_{m,n}^d(v_T) = v_{T'} \)
5. Moreover, the \( v_{T'} \) generate space of hwv of weight \( \lambda \) \( \square \)
Proof of Prop. ALL

**Prop. ALL**

\[ \lambda \vdash nd \text{ s.t. } |\bar{\lambda}| \leq md \text{ and } md^2 \leq n. \] Then every highest weight vector of weight \( \lambda \) in \( \text{Sym}^d \text{Sym}^n V \) does not vanish on \( \Omega_n \).

**Proof.**

- Let \( h \in \text{Sym}^d \text{Sym}^n V \) be hwv of weight \( \lambda \)
- \( \lambda_2 \leq |\bar{\lambda}| \leq md \) and \( \lambda_2 + |\bar{\lambda}| \leq 2|\bar{\lambda}| \leq 2md \leq md \cdot d \)
- Cor. Lift applied to \( \text{Sym}^d \text{Sym}^{md} V \to \text{Sym}^d \text{Sym}^n V \) shows \( h = \kappa_{md,n}^d(f) \) for some hwv \( f \in \text{Sym}^d \text{Sym}^{md} V \) of weight \( \lambda \)
- 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
  \[ \langle h, q^d \rangle = \langle \kappa_{md,n}^d(f), q^d \rangle = \langle f, M^*(q)^d \rangle \neq 0. \]

By Prop. PPS, we have \( q \in \Omega_n \) since \( n \geq md \cdot d \). \( \square \)
Thank you for your attention!