

Kronecker products and iterated matrix multiplication

Christian Ikenmeyer



Foundations of Computational Mathematics, Vienna, 2026-Jul-09

- 1 Overview
- 2 Tensor products and Kronecker products
- 3 Iterated matrix multiplication
- 4 Kronecker products of matrices of polynomials

- 1 Overview
- 2 Tensor products and Kronecker products
- 3 Iterated matrix multiplication
- 4 Kronecker products of matrices of polynomials

Overview

- Valiant (1979): determinant vs permanent question: Given m , what is the smallest n such that $\exists y_{i,j} \in \mathbb{F} \cup \{x_{i,j}\}$ with

$$\det \begin{pmatrix} y_{1,1} & \cdots & y_{1,n} \\ \vdots & \ddots & \vdots \\ y_{n,1} & \cdots & y_{n,n} \end{pmatrix} = \sum_{\pi \in \mathfrak{S}_m} \underbrace{\prod_{i=1}^m x_{i,\pi(i)}}_{=: \text{per}_m}$$

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

- Conj.: n grows superpolynomially in m .
- Known to be false in characteristic 2 because $\text{per}_m = \det_m$. Hence, cannot be proved via field-independent methods
- Replacements per_m that work over every field come from combinatorics, not from algebra
- Mulmuley–Sohoni (2001): Geometric complexity theory (GCT) to resolve det vs per
- Kadish–Landsberg (2012), Bürgisser–I–Panova (2016): Some GCT conjectures are false, since $\deg(\det_n) \neq \deg(\text{per}_m)$
- Several different ideas to use iterated matrix multiplication to handle the degrees, each with issues

We fix this. Works over all commutative semirings.
Works for polynomials, tensors, and exterior algebra.

Nice corollaries of our work (unintended):

- Generalization of the Cayley–Hamilton theorem, new record ABP construction for the determinant (Mahajan–Vinay 1997), answer a question from Mahajan–Vinay (1999) about combinatorial algorithms for the determinant
- Gurvits' (2004, 2005) VNP-hardness results for the mixed discriminant and Cayley's first hyperdeterminant hold over every commutative ring (Gurvits had fields of char $\neq 2$).
- We can translate Christandl–Fawzi–Ta–Zuiddam (2022) about asymptotic spectrum to arbitrary characteristic

$$\left(\begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,m}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{m,1}^{(1)} & \cdots & x_{m,m}^{(1)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d)} & \cdots & x_{1,m}^{(d)} \\ \vdots & \ddots & \vdots \\ x_{m,1}^{(d)} & \cdots & x_{m,m}^{(d)} \end{pmatrix} \right)^{\boxtimes 2} \leq \begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,n}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(1)} & \cdots & x_{n,n}^{(1)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d)} & \cdots & x_{1,n}^{(d)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(d)} & \cdots & x_{n,n}^{(d)} \end{pmatrix}$$

- 1 Overview
- 2 Tensor products and Kronecker products
- 3 Iterated matrix multiplication
- 4 Kronecker products of matrices of polynomials

- Tensor product $\otimes$ increases the tensor order:

▶ $\mathbb{F}^n \otimes \mathbb{F}^m = \mathbb{F}^{n \times m}$ matrices.

▶ $\mathbb{F}^n \otimes \mathbb{F}^m \otimes \mathbb{F}^\ell = \mathbb{F}^{n \times m \times \ell}$ order 3 tensors.

- Kronecker product $\boxtimes$ preserves the tensor order:

▶ $\mathbb{F}^n \boxtimes \mathbb{F}^m = \mathbb{F}^{nm}$.

▶ $\mathbb{F}^{n_1 \times n_2} \boxtimes \mathbb{F}^{m_1 \times m_2} = \mathbb{F}^{(n_1 m_1) \times (n_2 m_2)}$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

$$A \boxtimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix} = \begin{pmatrix} ae & af & be & bf \\ ag & ah & bg & bh \\ ce & cf & de & df \\ cg & ch & dg & dh \end{pmatrix}$$

▶ $\boxtimes$ works for every fixed degree d .

Kronecker product and the matrix multiplication exponent

- The matrix multiplication tensor

$$M_n = \sum_{i,j,k \in \{1, \dots, n\}} (e_i \boxtimes e_j) \otimes (e_j \boxtimes e_k) \otimes (e_k \boxtimes e_i) \in \mathbb{F}^{n^2} \otimes \mathbb{F}^{n^2} \otimes \mathbb{F}^{n^2}$$

- $M_n \boxtimes M_m = M_{nm}$
- $M_n \boxtimes M_m = M_{nm}$ was first used by Strassen (1969) for subcubic matrix multiplication algorithms.
- Used today for the Coppersmith-Winograd tensor to obtain the fastest matrix multiplication algorithms.
- Many tasks in computational linear algebra run in roughly the same time:
 - ▶ matrix multiplication
 - ▶ solving systems of linear equations
 - ▶ determinant
 - ▶ matrix inverse
 - ▶ LU decomposition
 - ▶ etc

- 1 Overview
- 2 Tensor products and Kronecker products
- 3 Iterated matrix multiplication**
- 4 Kronecker products of matrices of polynomials

Iterated matrix multiplication as a machine model

For $f \in \mathbb{R}[\mathbf{x}]_d$, the width $w(f)$ is defined as the smallest n such that $\exists \ell_{i,j}^{(k)} \in \mathbb{R}[\mathbf{x}]_1$ such that

$$f = (\ell_{1,1}^{(1)} \cdots \ell_{1,n}^{(1)}) \begin{pmatrix} \ell_{1,1}^{(2)} & \cdots & \ell_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(2)} & \cdots & \ell_{n,n}^{(2)} \end{pmatrix} \begin{pmatrix} \ell_{1,1}^{(3)} & \cdots & \ell_{1,n}^{(3)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(3)} & \cdots & \ell_{n,n}^{(3)} \end{pmatrix} \cdots \begin{pmatrix} \ell_{1,1}^{(d-1)} & \cdots & \ell_{1,n}^{(d-1)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(d-1)} & \cdots & \ell_{n,n}^{(d-1)} \end{pmatrix} \begin{pmatrix} \ell_{1,1}^{(d)} \\ \ell_{2,1}^{(d)} \\ \vdots \\ \ell_{n,1}^{(d)} \end{pmatrix}$$

Nisan's 1990 result

- Nisan (1991) proved:

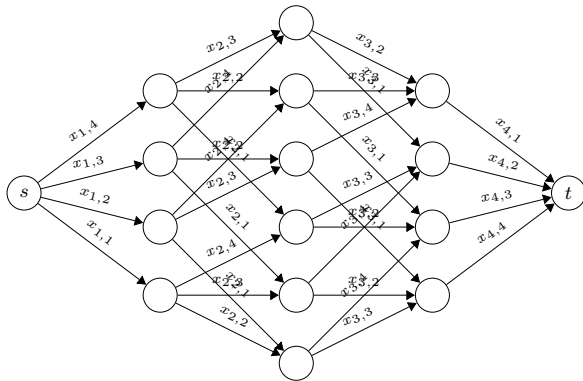
$$w(f) = \max\{\text{rk}(\mathcal{F}_j(f)) \mid 1 \leq j \leq d-1\}$$

where $\mathcal{F}_j(f)$ is the rank of f interpreted as a matrix in $(\bigotimes^j M) \otimes (\bigotimes^{d-j} M)$.

This machine model works for arbitrary commutative semirings, and for polynomials or tensors or the exterior algebra.

Bürgisser–Cucker–Lairez (2020): width to describe structured systems of polynomial eqns, and solve Smale's 17th problem.

$$\begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} & 0 & 0 \end{pmatrix}
 \begin{pmatrix} x_{2,2} & x_{2,3} & x_{2,4} & 0 & 0 & 0 \\ x_{2,1} & 0 & 0 & x_{2,3} & x_{2,4} & 0 \\ 0 & x_{2,1} & 0 & x_{2,2} & 0 & x_{2,4} \\ 0 & 0 & x_{2,1} & 0 & x_{2,2} & x_{2,3} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & x_{3,4} & x_{3,3} & 0 & 0 \\ 0 & x_{3,4} & 0 & x_{3,2} & 0 & 0 \\ 0 & x_{3,3} & x_{3,2} & 0 & 0 & 0 \\ x_{3,4} & 0 & 0 & x_{3,1} & 0 & 0 \\ x_{3,3} & 0 & x_{3,1} & 0 & 0 & 0 \\ x_{3,2} & x_{3,1} & 0 & 0 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} x_{4,1} \\ x_{4,2} \\ x_{4,3} \\ x_{4,4} \\ 0 \\ 0 \end{pmatrix}$$



$$\sum_{s-t\text{-path } p} \prod_{e \in p} \text{label}(e)$$

Algebraic Branching Program (ABP)

Computing Determinants

$$X_n := \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix}$$

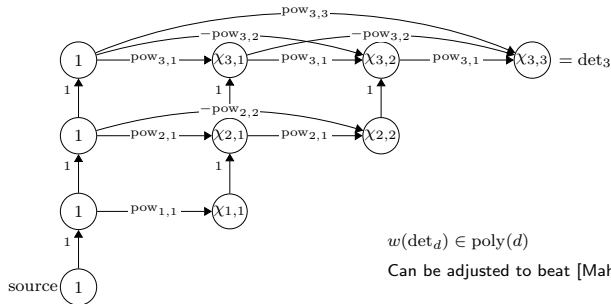
$$\det(X_n + \text{diag}(t, \dots, t)) = \sum_{d \geq 0} \chi_{n,d} \cdot t^{n-d}$$

Strengthening of the Cayley–Hamilton theorem

For all $n, d \in \mathbb{N}$ we have $(\nabla \chi_{n,d+1})^T = \sum_{i=0}^d (-1)^i \chi_{n,d-i} (X_n)^i$.

Take the bottom-right entry: $\text{pow}_{n,i} := [(X_n)^i]_{n,n}$ $\chi_{n-1,d} = \sum_{i=0}^d (-1)^i \chi_{n,d-i} \text{pow}_{n,i}$. Rearrange:

$$\chi_{n,d} = \chi_{n-1,d} + \sum_{i=1}^d (-1)^{i+1} \chi_{n,d-i} \text{pow}_{n,i}.$$



$w(\det_d) \in \text{poly}(d)$

Can be adjusted to beat [Mahajan–Vinay 1997]

Restrictions

For $f \in \mathbb{R}[\mathbf{x}]_d$, the width $w(f)$ is defined as the smallest n such that $\exists \ell_{i,j}^{(k)} \in \mathbb{R}[\mathbf{x}]_1$ such that

$$f = (\ell_{1,1}^{(1)} \cdots \ell_{1,n}^{(1)}) \begin{pmatrix} \ell_{1,1}^{(2)} & \cdots & \ell_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(2)} & \cdots & \ell_{n,n}^{(2)} \end{pmatrix} \begin{pmatrix} \ell_{1,1}^{(3)} & \cdots & \ell_{1,n}^{(3)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(3)} & \cdots & \ell_{n,n}^{(3)} \end{pmatrix} \cdots \begin{pmatrix} \ell_{1,1}^{(d-1)} & \cdots & \ell_{1,n}^{(d-1)} \\ \vdots & \ddots & \vdots \\ \ell_{n,1}^{(d-1)} & \cdots & \ell_{n,n}^{(d-1)} \end{pmatrix} \begin{pmatrix} \ell_{1,1}^{(d)} \\ \ell_{2,1}^{(d)} \\ \vdots \\ \ell_{n,1}^{(d)} \end{pmatrix}$$

- We write $f \leq h$ if $f = h(\ell_1, \dots, \ell_N)$ for homogeneous linear ℓ_i .

$$\Gamma_{n,d} := (x_{1,1}^{(1)} \cdots x_{1,n}^{(1)}) \begin{pmatrix} x_{1,1}^{(2)} & \cdots & x_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(2)} & \cdots & x_{n,n}^{(2)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(3)} & \cdots & x_{1,n}^{(3)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(3)} & \cdots & x_{n,n}^{(3)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d-1)} & \cdots & x_{1,n}^{(d-1)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(d-1)} & \cdots & x_{n,n}^{(d-1)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(d)} \\ x_{2,1}^{(d)} \\ \vdots \\ x_{n,1}^{(d)} \end{pmatrix}$$

$$w(f) \leq n \text{ iff } f \leq \Gamma_{n,d}$$

For $f \in \text{Sym}^d \mathbb{C}^N$ let

$$\text{stab}(f) = \{g \in \text{GL}_N \mid gf = f\}.$$

These all have reductive stabilizers: $x_1 x_2 \cdots x_d$, $x_1^d + x_2^d + \cdots + x_d^d$, $\det_n$, per_m

BUT $\text{stab}(\Gamma_{n,d})$ is not reductive!

$$\Gamma_{n,d} = (x_{1,1}^{(1)} \cdots x_{1,r}^{(1)}) \begin{pmatrix} x_{1,1}^{(2)} & \cdots & x_{1,r}^{(2)} \\ \vdots & \ddots & \vdots \\ x_{r,1}^{(2)} & \cdots & x_{r,r}^{(2)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(3)} & \cdots & x_{1,r}^{(3)} \\ \vdots & \ddots & \vdots \\ x_{r,1}^{(3)} & \cdots & x_{r,r}^{(3)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d-1)} & \cdots & x_{1,r}^{(d-1)} \\ \vdots & \ddots & \vdots \\ x_{r,1}^{(d-1)} & \cdots & x_{r,r}^{(d-1)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(d)} \\ x_{2,1}^{(d)} \\ \vdots \\ x_{r,1}^{(d)} \end{pmatrix}$$

$$\Xi_{n,d} := \begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,n}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(1)} & \cdots & x_{n,n}^{(1)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(2)} & \cdots & x_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(2)} & \cdots & x_{n,n}^{(2)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d)} & \cdots & x_{1,n}^{(d)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(d)} & \cdots & x_{n,n}^{(d)} \end{pmatrix}$$

	Nisan's result	reductive stabilizer
$\Gamma_{n,d}$	✓	☹
$\text{tr}(\Xi_{n,d})$	☹	✓
$\Xi_{n,d}$	✓	✓

Note: Matrices of polynomials in n and d are the main players in the generalization of the Cayley–Hamilton theorem.

- 1 Overview
- 2 Tensor products and Kronecker products
- 3 Iterated matrix multiplication
- 4 Kronecker products of matrices of polynomials**

$$\Xi_{n,d} := \begin{pmatrix} x_{1,1}^{(1)} & \cdots & x_{1,n}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(1)} & \cdots & x_{n,n}^{(1)} \end{pmatrix} \begin{pmatrix} x_{1,1}^{(2)} & \cdots & x_{1,n}^{(2)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(2)} & \cdots & x_{n,n}^{(2)} \end{pmatrix} \cdots \begin{pmatrix} x_{1,1}^{(d)} & \cdots & x_{1,n}^{(d)} \\ \vdots & \ddots & \vdots \\ x_{n,1}^{(d)} & \cdots & x_{n,n}^{(d)} \end{pmatrix}$$

This section: Define $\Xi_{n,d} \boxtimes \Xi_{n,d}$

Kronecker product of polynomials in characteristic 0

- $\text{polar} : \mathbb{R}[x_1, \dots, x_N]_d \rightarrow (\bigotimes^d \mathbb{R}^N)^{\mathfrak{S}_d}$
- $\text{resti} : \bigotimes^d \mathbb{R}^N \rightarrow \mathbb{R}[x_1, \dots, x_N]_d$

Definition in char 0

Let $h, f \in \mathbb{R}[x_1, \dots, x_N]_d$.

$$h \boxtimes f := \frac{1}{d!} \text{resti}(\text{polar}(h) \boxtimes \text{polar}(f))$$

$\text{mon}_d = x_1 x_2 \cdots x_d$.

- The permanent polynomial: $\text{mon}_d \boxtimes \text{mon}_d$.
- The hyperpermanent: $\text{mon}_d^{\boxtimes m}$.
- The rectangular permanent: $\text{mon}_d \boxtimes e_{n,d}$.
- The mixed discriminant: $\det_d \boxtimes \text{mon}_d$.
- Cayley's first hyperdeterminant: $\det_d^{\boxtimes m}$.

⊠ over commutative semirings R

- Let $\text{Sym}_d(V) := (\bigotimes^d V)^{\mathfrak{S}_d}$
- R -bilinear map $\boxtimes : \bigotimes^d V \times \bigotimes^d W \rightarrow \bigotimes^d (V \boxtimes W)$.
- Restrict to $\psi : \text{Sym}_d(V) \times \bigotimes^d W \rightarrow \bigotimes^d (V \boxtimes W)$.
- Fix $s \in \text{Sym}_d(V)$.
- $\mathfrak{S}_d$ -equivariant R -linear map $\psi'_s : \bigotimes^d W \rightarrow \bigotimes^d (V \boxtimes W)$, $f \mapsto \psi(s, f)$
- Apply coinvariant functor: $\psi_s : \text{Sym}^d(W) \rightarrow \text{Sym}^d(V \boxtimes W)$, $[f] \mapsto [\psi(s, f)]$.
- R -linear map $\kappa : \text{Sym}_d(V) \rightarrow \text{Hom}(A_d(W), \text{Sym}^d(V \boxtimes W))$, $s \mapsto \psi_s$
- Altogether: R -bilinear map

$$\boxtimes : \text{Sym}_d(V) \times \text{Sym}^d(W) \rightarrow \text{Sym}^d(V \boxtimes W), \quad (s, [f]) \mapsto \psi_s([f]) = [\psi(s, f)]$$

Definition (Kronecker product of polynomials)

Let $h, f \in \mathbb{R}[x_1, \dots, x_N]_d$.

$h \boxtimes f := \text{resti}(\text{polar}(h) \boxtimes f)$

$(h \boxtimes f = f \boxtimes h \text{ up to renaming variables})$

On monomials: $(x_1 \cdots x_d) \boxtimes (y_1 \cdots y_d) = \sum_{\pi \in \mathfrak{S}_d} (x_1 \boxtimes y_{\pi(1)}) \cdots (x_d \boxtimes y_{\pi(d)})$.

This definition fixes the issues with the characteristic in Christandl–Fawzi–Ta–Zuiddam (2022).

The hypercomputant $\mathbf{h}\Xi_{n,d}$

$$\mathbf{h}\Xi_{n,d} := \Xi_{n,d} \boxtimes \Xi_{n,d}$$

an $n^2 \times n^2$ matrix of homog. degree d polynomials. Its bottom-right entry: $\mathbf{h}\Gamma_{n,d} = \Gamma_{n,d} \boxtimes \Gamma_{n,d}$
Hamiltonian cycle polynomial:

$$\mathbf{HC}_d := \sum_{\phi \in \mathfrak{S}_{d-1}} x_{d,\phi(1)}^{(1)} \cdot x_{\phi(1),\phi(2)}^{(2)} \cdots x_{\phi(d-1),d}^{(d)}$$

Theorem

$$\mathbf{HC}_d \leq \mathbf{h}\Gamma_{d,d}.$$

Proof:

$$\Gamma_{d,d} = \sum_{\phi: [d-1] \rightarrow [d]} x_{d,\phi(1)}^{(1)} \cdot x_{\phi(1),\phi(2)}^{(2)} \cdot x_{\phi(2),\phi(3)}^{(3)} \cdots x_{\phi(d-1),d}^{(d)}.$$

$$\Gamma_{d,d} \boxtimes \Gamma_{d,d} \geq \text{mon}_d \boxtimes \Gamma_{d,d} = \sum_{\substack{\phi: [d-1] \rightarrow [d] \\ \pi \in \mathfrak{S}_d}} x_{\pi(1),d,\phi(1)}^{(1)} \cdot x_{\pi(2),\phi(1),\phi(2)}^{(2)} \cdots x_{\pi(d),\phi(d-1),d}^{(d)}.$$

Let $T \in \text{End}$ as follows: $T(x_{a,b,c}^{(k)}) = x_{b,c}^{(k)}$ if $a = c$, and $T(x_{a,b,c}^{(k)}) = 0$ otherwise. We calculate

$$T(\text{mon}_d \boxtimes \Gamma_{d,d}) = \sum_{\substack{\phi \in \mathfrak{S}_d \\ \phi(d)=d}} x_{d,\phi(1)}^{(1)} \cdot x_{\phi(1),\phi(2)}^{(2)} \cdots x_{\phi(d-1),d}^{(d)} = \mathbf{HC}_d. \quad \square$$

This lifts two hardness results of Gurvits (2004, 2005) to all commutative rings:

$$\mathbf{HC}_d \leq \text{mon}_d \boxtimes \Gamma_{d,d} \leq_p \text{mon}_d \boxtimes \det_d \leq \det_d \boxtimes \det_d$$

$$\xi_d := \inf\{e \mid w(\mathbf{h}\Xi_{n,d}) \in O(n^e)\},$$

- Robust: Same exponent for ABP width, ABP size, or for determinantal complexity

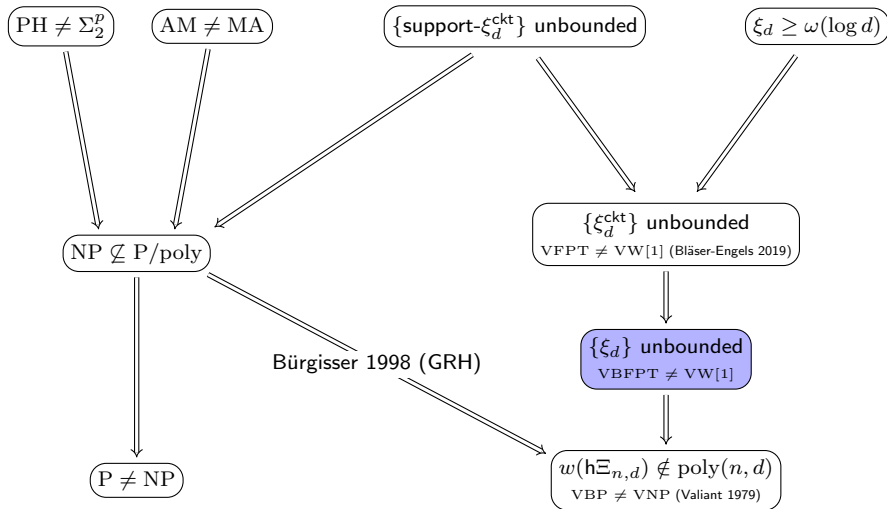
d	2	3	4	5	6	7		
ξ_d for tensors over $\mathbb{R}$	4	4	6	6	8	8	$\dots$	$\Theta(d)$ (Nisan)
ξ_d for polynomials over $\mathbb{R}_{\geq 0}$	4	4	6	6	6	8	$\dots$	$\Theta(d)$ (pathwidth)

Main open question

Is the set of exponents $\{\xi_d \mid d \in \mathbb{N}\}$ unbounded for polynomials over $\mathbb{R}$?

$\xi_d \geq 2$. Best algorithm in the general case is **monotone**.

For polynomials over $\mathbb{C}$ or $\mathbb{R}$:



Summary

- Guided by fixing some rough edges in the definitions of geometric complexity theory (GCT).
- Instead of polynomials, study **matrices** of **double-index** polynomials (n and d).
- This gives a short existence proof for polysized ABPs for $\det_d$, gives a new record ABP for $\det_d$, and answers a question by Mahajan–Vinay.
- The stabilizer of $\Xi_{n,d}$ is reductive and the tensor variant is Zariski-closed (only had these separately).
- The definition of $\boxtimes$ for polynomials solves issues with the characteristic (Christandl–Fawzi–Ta–Zuiddam).
- The equivariance of $\boxtimes$ gives a short proof for lifting two hardness results of Gurvits to all commutative rings.
- And possibly more . . .

Thank you for your attention!