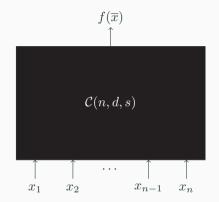
Finding the Order of an ROABP

Can we proper-learn ROABPs?

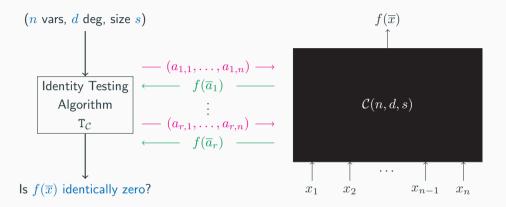
Anamay Tengse TIFR. June 11, 2025

joint work with Vishwas Bhargava, Pranjal Dutta and Sumanta Ghosh

Polynomial Identity Testing (PIT)

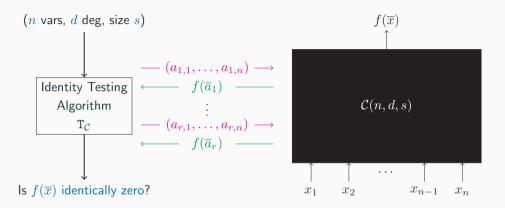


Polynomial Identity Testing (PIT)



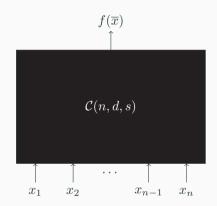
1

Polynomial Identity Testing (PIT)

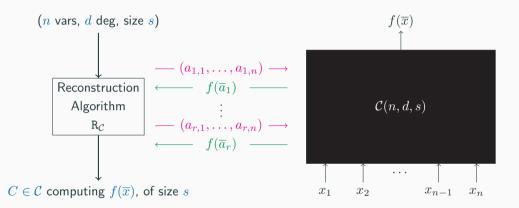


Known. Randomized, poly(n, d, s)-time algorithm for circuits. Open. Deterministic algorithm for circuits, making fewer than d^n queries.

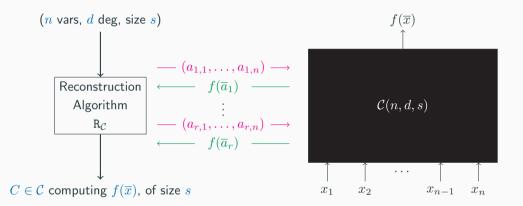
Reconstruction ("Proper-Learning") of Polynomials



Reconstruction ("Proper-Learning") of Polynomials



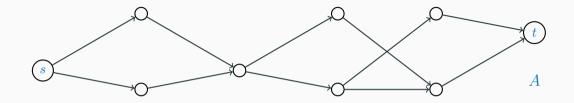
Reconstruction ("Proper-Learning") of Polynomials

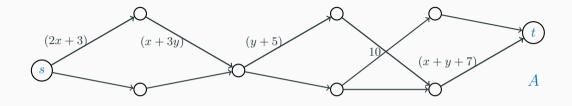


 $R_{\mathcal{C}}$ exists: "The class \mathcal{C} can be proper-learnt in <complexity-of-R>"

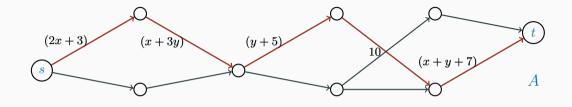
The Question

Read-once Oblivious ABPs

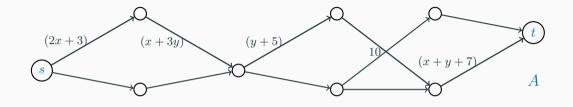




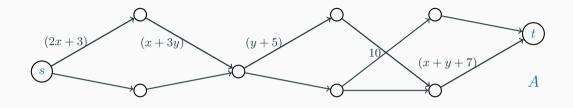
• Label on each edge: An linear polynomial in $\{x_1, x_2, \ldots, x_n\}$



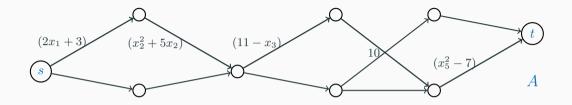
- Label on each edge: An linear polynomial in $\{x_1, x_2, \ldots, x_n\}$
- Polynomial computed by the path p = wt(p): Product of the edge labels on p

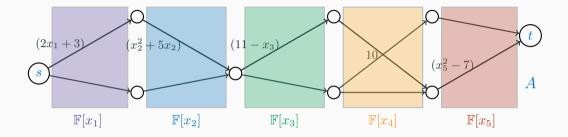


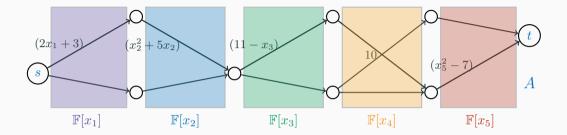
- Label on each edge: An linear polynomial in $\{x_1, x_2, \dots, x_n\}$
- Polynomial computed by the path p = wt(p): Product of the edge labels on p
- Polynomial computed by the ABP: $f_A(\overline{x}) = \sum_p \operatorname{wt}(p)$ (sum over s to t paths)



- Label on each edge: An linear polynomial in $\{x_1, x_2, \dots, x_n\}$
- Polynomial computed by the path p = wt(p): Product of the edge labels on p
- Polynomial computed by the ABP: $f_A(\overline{x}) = \sum_p wt(p)$ (sum over s to t paths)
- Size of the ABP: total number of vertices (9 in the example)

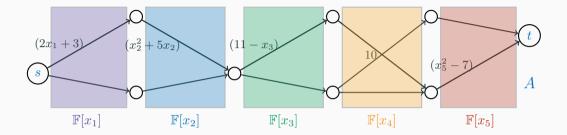






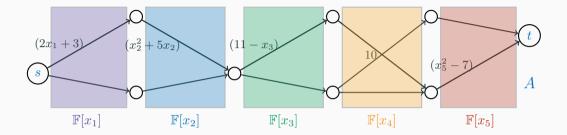
ROABP (*n*-variate, degree-*d*, width-*w*)

• On each s to t path, every variable is read-once, oblivious of the others.



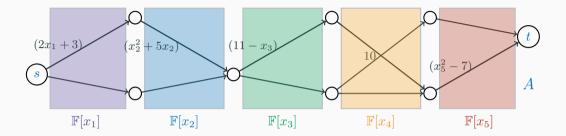
ROABP (*n*-variate, degree-d, width-w)

- On each s to t path, every variable is read-once, oblivious of the others.
- The i^{th} layer of edges only has degree-d univariates in x_i as labels.



ROABP (*n*-variate, degree-d, width-w)

- On each s to t path, every variable is read-once, oblivious of the others.
- The i^{th} layer of edges only has degree-d univariates in x_i as labels.
- Width of the ROABP: Maximum number of vertices in any layer.



ROABP (*n*-variate, degree-*d*, width-*w*, order σ)

- On each s to t path, every variable is read-once, oblivious of the others.
- The i^{th} layer of edges only has degree-d univariates in $x_{\sigma(i)}$ as labels.
- Width of the ROABP: Maximum number of vertices in any layer.
- Order of the ROABP: permutation $\sigma \in s_n$ in which the variables are read.

Example. Width depends on Order

 $F(\overline{x}, \overline{y})$ has a width 2 ROABP in the order $(x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$, but requires width 2^n in the order $(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n)$.

$$F(\overline{x},\overline{y}) := (x_1 + y_1)(x_2 + y_2)\cdots(x_n + y_n)$$

• Lower Bounds. Easy, due to an explicit characterization (Nisan 1991) to compute the optimal width in each layer exactly.

- Lower Bounds. Easy, due to an explicit characterization (Nisan 1991) to compute the optimal width in each layer exactly.
- Identity Testing.
 - Algorithm can access circuit ("Whitebox"): Deterministic time poly(n, d, w).
 - Algorithm can only query ("Blackbox"): Deterministic time $(ndw)^{O(\log n)}$.

- Lower Bounds. Easy, due to an explicit characterization (Nisan 1991) to compute the optimal width in each layer exactly.
- Identity Testing.
 - Algorithm can access circuit ("Whitebox"): Deterministic time poly(n, d, w).
 - Algorithm can only query ("Blackbox"): Deterministic time $(ndw)^{O(\log n)}$.
- **Reconstruction.** Randomized polytime, and deterministic time $(ndw)^{O(\log n)}$, when the algorithm is given the order.

- Lower Bounds. Easy, due to an explicit characterization (Nisan 1991) to compute the optimal width in each layer exactly.
- Identity Testing.
 - Algorithm can access circuit ("Whitebox"): Deterministic time poly(n, d, w).
 - Algorithm can only query ("Blackbox"): Deterministic time $(ndw)^{O(\log n)}$.
- **Reconstruction.** Randomized polytime, and deterministic time $(ndw)^{O(\log n)}$, when the algorithm is given the order.

Q. What is the complexity of finding the order?

Order Finding Problem

Given parameters $n, d, w \in \mathbb{N}$ and a polynomial $f(\overline{x})$, find some order σ in which f has an ROABP of width at most w.

Order Finding Problem

Given parameters $n, d, w \in \mathbb{N}$ and a polynomial $f(\overline{x})$, find some order σ in which f has an ROABP of width at most w.

Note. Using reconstruction, we can check if the given w is correct. So we can assume WLOG that f has an ROABP of width w in some order, hence the following is a simpler problem.

Order Finding Problem

Given parameters $n, d, w \in \mathbb{N}$ and a polynomial $f(\overline{x})$, find some order σ in which f has an ROABP of width at most w.

Note. Using reconstruction, we can check if the given w is correct. So we can assume WLOG that f has an ROABP of width w in some order, hence the following is a simpler problem.

Order Finding Problem (Decision)

Given an *n*-variate, degree-*d* polynomial $f(\overline{x})$, and a parameter $w \in \mathbb{N}$, determine if f has an ROABP of width at most w in some order σ .

NP hardness (Algebraic Circuit Minimization)

Order finding problem is NP-hard, even when f is given as an algebraic circuit.

NP hardness (Algebraic Circuit Minimization)

Order finding problem is NP-hard, even when f is given as an algebraic circuit.

NP hardness for constant degree (Algebraic MCSP)

For any constant $\Delta \ge 6$, order finding for *n*-variate, degree- Δ polynomials is NP-hard, even when *f* is given in the dense representation (algebraic analogue of a truth table).

NP hardness (Algebraic Circuit Minimization)

Order finding problem is NP-hard, even when f is given as an algebraic circuit.

NP hardness for constant degree (Algebraic MCSP)

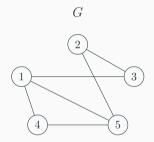
For any constant $\Delta \ge 6$, order finding for *n*-variate, degree- Δ polynomials is NP-hard, even when f is given in the dense representation (algebraic analogue of a truth table).

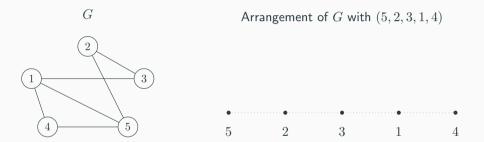
Average-case algorithm

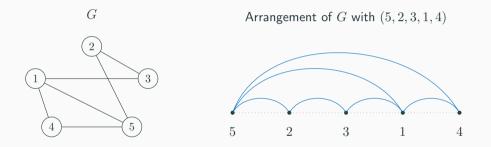
Randomized order-finding algorithm that runs in polytime for a random/generic ROABP.

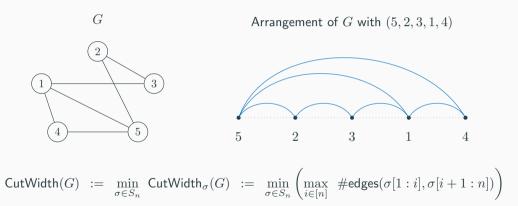
Proof Ideas

NP-hardness



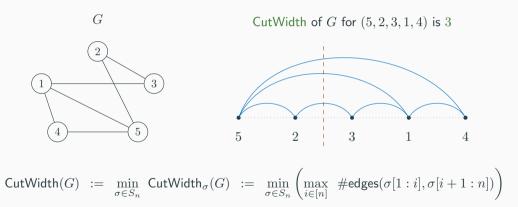






Linear Arrangement of Graphs

Goal. Show that finding an optimal order is hard.



$$\mathsf{CutWidth}(G) := \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# \mathsf{edges}(\sigma[1:i], \sigma[i+1:n]) \right)$$

$$\begin{aligned} \mathsf{CutWidth}(G) &:= \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# \mathsf{edges}(\sigma[1:i], \sigma[i+1:n]) \right) \\ \\ \mathsf{ROABPwidth}(f) &:= \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# \{ \mathsf{vertices in layer } i \} \right) \end{aligned}$$

$$CutWidth(G) := \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# edges(\sigma[1:i], \sigma[i+1:n]) \right)$$
$$ROABPwidth(f) := \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# \{ \text{vertices in layer } i \} \right)$$

Theorem (implied by [Nisan 1991])

For any $f(\overline{x})$, the optimal ROABP for f in the order σ has exactly $w_i = \operatorname{rk}\left(M_f^{(\sigma,i)}\right)$ vertices in layer i,

$$\begin{aligned} \mathsf{CutWidth}(G) &:= \min_{\sigma \in S_n} \left(\max_{i \in [n]} \# \mathsf{edges}(\sigma[1:i], \sigma[i+1:n]) \right) \end{aligned}$$
$$\begin{aligned} \mathsf{ROABPwidth}(f) &:= \min_{\sigma \in S_n} \quad \left(\max_{i \in [n]} \# \{\mathsf{vertices in layer } i\} \right) \end{aligned}$$

Theorem (implied by [Nisan 1991])

For any $f(\overline{x})$, the optimal ROABP for f in the order σ has exactly $w_i = \operatorname{rk}\left(M_f^{(\sigma,i)}\right)$ vertices in layer i, where $M_f^{(\sigma,i)}$ is as follows, for $\overline{x}_L = \{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\}, \overline{x}_R = \{x_{\sigma(i+1)}, \ldots, x_{\sigma(n)}\}.$

$$\forall m \in \operatorname{mons}(\overline{x}_L), m' \in \operatorname{mons}(\overline{x}_R), \qquad M_f^{(\sigma, i)}[m, m'] = \operatorname{coeff}_f(m \cdot m')$$

Lemma (Bhargava-Dutta-Ghosh-T. 2024)

Given any graph G = (V, E), there is a polynomial $f_G(x_1, \ldots, x_n)$ such that:

• n = |V|,

•
$$\forall \sigma \in S_n, i \in [n], \operatorname{rk}\left(M_{f_G}^{(\sigma,i)}\right) = \#\operatorname{edges}(\sigma[1:i], \sigma[i+1:n]) + 2,$$

Lemma (Bhargava-Dutta-Ghosh-T. 2024)

Given any graph G = (V, E), there is a polynomial $f_G(x_1, \ldots, x_n)$ such that:

- n = |V|,
- $\forall \sigma \in S_n, i \in [n], \operatorname{rk}\left(M_{f_G}^{(\sigma,i)}\right) = \#\operatorname{edges}(\sigma[1:i], \sigma[i+1:n]) + 2,$
- $\deg(f_G) = 2 \cdot \deg(G)$ and $\operatorname{ideg}(f_G) = \deg(G) + 1$,
- f_G has |E| + |V| + 1 monomials.

Lemma (Bhargava-Dutta-Ghosh-T. 2024)

Given any graph G = (V, E), there is a polynomial $f_G(x_1, \ldots, x_n)$ such that:

- n = |V|,
- $\forall \sigma \in S_n, i \in [n], \operatorname{rk}\left(M_{f_G}^{(\sigma,i)}\right) = \#\operatorname{edges}(\sigma[1:i], \sigma[i+1:n]) + 2,$
- $\deg(f_G) = 2 \cdot \deg(G)$ and $\operatorname{ideg}(f_G) = \deg(G) + 1$,
- f_G has |E| + |V| + 1 monomials.

Fact (Monien-Sudborough 1988)

CutWidth is NP-complete, even for planar graphs of degree 3.

Theorem (Algebraic MCSP)

For any constant $\Delta \ge 6$, order finding for *n*-variate, degree- Δ polynomials is NP-hard, even when *f* is given in the dense representation (algebraic analogue of a truth table). Proof. Truth table has length $\binom{n+\Delta}{\Delta} = poly(n)$ for constant Δ .

Theorem (Algebraic MCSP)

For any constant $\Delta \ge 6$, order finding for *n*-variate, degree- Δ polynomials is NP-hard, even when *f* is given in the dense representation (algebraic analogue of a truth table). Proof. Truth table has length $\binom{n+\Delta}{\Delta} = poly(n)$ for constant Δ .

Theorem (Algebraic Circuit Minimization)

Order finding problem is NP-hard, even when f is given as an algebraic circuit.

Proof. CircuitSize $(f_G) = O(n^3)$. (Truth table length $\sim 2^n$ for degree $\Omega(n)$.)

Proof Ideas

E-time worst-case algorithm

Theorem (Nisan's characterization)

$$\begin{split} & \mathsf{ROABPwidth}_{\sigma}(f(x_1,\ldots,x_n)) \leq w, \text{ iff} \\ & \mathrm{rk}\left(M_f^{(\sigma,i)}\right) \leq w \text{ for all } 1 < i < n. \end{split}$$

E.g. For
$$n = 5$$
, $\sigma = (5, 2, 3, 1, 4)$,
 $\operatorname{rk}(M_f^{\{5\}})$, $\operatorname{rk}(M_f^{\{2,5\}})$, $\operatorname{rk}(M_f^{\{2,3,5\}})$ and
 $\operatorname{rk}(M_f^{\{1,2,3,5\}})$ are all at most w .

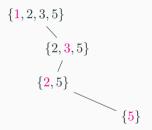
 M_f^S has mons in x_S and $x_{\overline{S}}$ as rows and columns.

Theorem (Nisan's characterization)

 $\begin{aligned} \mathsf{ROABPwidth}_{\sigma}(f(x_1, \dots, x_n)) &\leq w, \text{ iff} \\ \mathrm{rk}\left(M_f^{(\sigma, i)}\right) &\leq w \text{ for all } 1 < i < n. \end{aligned}$

E.g. For $n=5,\,\sigma=(5,2,3,1,4),\, \mathrm{rk}(M_f^{\{5\}}),\,\mathrm{rk}(M_f^{\{2,5\}}),\,\mathrm{rk}(M_f^{\{2,3,5\}})$ and $\mathrm{rk}(M_f^{\{1,2,3,5\}})$ are all at most w.

 M_f^S has mons in x_S and $x_{\overline{S}}$ as rows and columns.

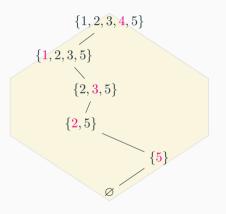


Theorem (Nisan's characterization)

 $\begin{aligned} \mathsf{ROABPwidth}_{\sigma}(f(x_1, \dots, x_n)) &\leq w, \text{ iff} \\ \mathrm{rk}\left(M_f^{(\sigma, i)}\right) &\leq w \text{ for all } 1 < i < n. \end{aligned}$

E.g. For $n=5,\,\sigma=(5,2,3,1,4),\,\operatorname{rk}(M_f^{\{5\}}),\,\operatorname{rk}(M_f^{\{2,5\}}),\,\operatorname{rk}(M_f^{\{2,3,5\}})$ and $\operatorname{rk}(M_f^{\{1,2,3,5\}})$ are all at most w.

 M_f^S has mons in x_S and $x_{\overline{S}}$ as rows and columns.



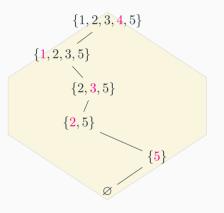
Theorem (Nisan's characterization)

 $\begin{aligned} \mathsf{ROABPwidth}_{\sigma}(f(x_1, \dots, x_n)) &\leq w, \text{ iff} \\ \mathrm{rk}\left(M_f^{(\sigma, i)}\right) &\leq w \text{ for all } 1 < i < n. \end{aligned}$

E.g. For $n=5,\,\sigma=(5,2,3,1,4),\, \mathrm{rk}(M_f^{\{5\}}),\,\mathrm{rk}(M_f^{\{2,5\}}),\,\mathrm{rk}(M_f^{\{2,3,5\}})$ and $\mathrm{rk}(M_f^{\{1,2,3,5\}})$ are all at most w.

 M_f^S has mons in x_S and $x_{\overline{S}}$ as rows and columns.

Observation. ROABPwidth_{σ} $(f(x_1, ..., x_n)) \le w$ iff ' σ traces an \varnothing to [n] path in the graph $H_w(f)$ '. $H_w(f)$: induced subgraph of hypercube, where $S \in H_w(f)$ if and only if $\operatorname{rk}(M_f^S) \le w$.



Observation

 $\mathsf{ROABPwidth}_{\sigma}(f(x_1,\ldots,x_n)) \leq w$ iff σ traces an \varnothing to [n] path in the graph $H_w(f)$.

Observation

 $\mathsf{ROABPwidth}_{\sigma}(f(x_1,\ldots,x_n)) \leq w$ iff σ traces an \varnothing to [n] path in the graph $H_w(f)$.

Fact

For any $f(x_1, \ldots, x_n)$ of deg d, and $S \subseteq [n]$, checking if $\operatorname{rk}(M_f^S) \leq w$ reduces to PIT.

Observation

 $\mathsf{ROABPwidth}_{\sigma}(f(x_1,\ldots,x_n)) \leq w$ iff σ traces an \varnothing to [n] path in the graph $H_w(f)$.

Fact

For any $f(x_1, \ldots, x_n)$ of deg d, and $S \subseteq [n]$, checking if $\operatorname{rk}(M_f^S) \leq w$ reduces to PIT.

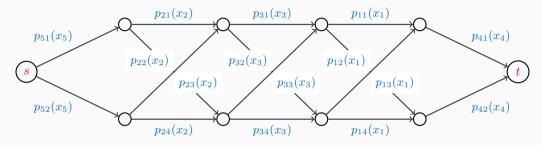
Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \varnothing (and above fact).
- 2. Output any σ that traces an \emptyset to [n] path in $H_w(f)$.

Proof Ideas

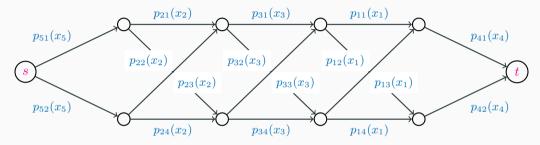
Algorithm for the generic case

Random/generic ROABPs



Generic ROABP for n = 5, w = 2 and $\sigma = (5, 2, 3, 1, 4)$: random coeffs for p_{ij} s.

Random/generic ROABPs



Generic ROABP for n = 5, w = 2 and $\sigma = (5, 2, 3, 1, 4)$: random coeffs for p_{ij} s.

Definition ((n, d, w, σ, D) -Generic ROABP)

ROABP in order σ with all coefficients of edge labels ($\sim ndw^2$) iid according to \mathcal{D} .

Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \varnothing (and above fact).
- 2. Output any σ that traces an \varnothing to [n] path in the populated graph.

Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \varnothing (and above fact).
- 2. Output any σ that traces an \varnothing to [n] path in the populated graph.

Key idea. Bad inputs are special (i.e. not generic).

Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \emptyset (and above fact).
- 2. Output any σ that traces an \varnothing to [n] path in the populated graph.

Key idea. Bad inputs are special (i.e. not generic).

• Bad input $f: H_w(f)$ has many vertices, but very few \emptyset to [n] paths. DFS has to backtrack from several blocked paths.

Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \emptyset (and above fact).
- 2. Output any σ that traces an \varnothing to [n] path in the populated graph.

Key idea. Bad inputs are special (i.e. not generic).

- Bad input f: H_w(f) has many vertices, but very few Ø to [n] paths.
 DFS has to backtrack from several blocked paths.
- We show: for generic $f \in \mathsf{ROABP}(n, d, w, \sigma)$, $H_w(f)$ only has the obvious vertices. $\mathrm{rk}(M_f^S) \leq w$ when S is a prefix of σ , or |S| is too small (M_f^S is skewed).

Algorithm. FindOrder(f,w)

- 1. PopulateGraph(f,w): Find $H_w(f)$ using a DFS starting at \varnothing (and above fact).
- 2. Output any σ that traces an \varnothing to [n] path in the populated graph.

Key idea. Bad inputs are special (i.e. not generic).

- Bad input f: H_w(f) has many vertices, but very few Ø to [n] paths.
 DFS has to backtrack from several blocked paths.
- We show: for generic $f \in \mathsf{ROABP}(n, d, w, \sigma)$, $H_w(f)$ only has the obvious vertices. $\mathrm{rk}(M_f^S) \leq w$ when S is a prefix of σ , or |S| is too small (M_f^S is skewed).
- For "inconsistent" S, fs with $\operatorname{rk}(M_f^S) \leq w$ form a strict subvariety of $\operatorname{ROABP}(n, d, w, \sigma)$. [SZ lemma]: $H_w(f)$ has $n^{O(\log_d(w))}$ vertices w.h.p., for any large-enough domain.

Theorem (Average-case algorithm)

Over all sets D of size 2^{10n} , and for any n, d, w, σ , PopulateGraph runs in randomized time $n^{O(\log_d(w))} \cdot \operatorname{poly}(d, w)$ on a random/generic input from ROABP (n, d, w, σ) w.h.p., where the coeffs are drawn from D.

- Polynomial time when $w = d^{O(1)}$.
- Quasi-polynomial time even when d = O(1).

Theorem (Average-case algorithm)

Over all sets D of size 2^{10n} , and for any n, d, w, σ , PopulateGraph runs in randomized time $n^{O(\log_d(w))} \cdot \operatorname{poly}(d, w)$ on a random/generic input from ROABP (n, d, w, σ) w.h.p., where the coeffs are drawn from D.

- Polynomial time when $w = d^{O(1)}$.
- Quasi-polynomial time even when d = O(1).

Remark. We need $|D| \sim 2^n$ due to a union bound over all inconsistent S.

Summary

• ROABPs can be proper-learnt efficiently when order is known.

- ROABPs can be proper-learnt efficiently when order is known.
- Order-finding problem is NP-hard even in simple, "white-box" settings. (Only other algebraic MCSP results are for tensor and Waring ranks [Håstad'90].)

- ROABPs can be proper-learnt efficiently when order is known.
- Order-finding problem is NP-hard even in simple, "white-box" settings. (Only other algebraic MCSP results are for tensor and Waring ranks [Håstad'90].)
- Order-finding can be solved in average case in (quasi-)polynomial time.

- ROABPs can be proper-learnt efficiently when order is known.
- Order-finding problem is NP-hard even in simple, "white-box" settings. (Only other algebraic MCSP results are for tensor and Waring ranks [Håstad'90].)
- Order-finding can be solved in average case in (quasi-)polynomial time.
- Approximation algorithms.
 - o ROABPwidth is hard to approximate up to any constant factor under SSE conjecture.
 - o Unconditionally, any constant approximation for ROABPwidth leads to a PTAS.

• Average-case algorithm.

o Polynomial time for constant individual degree? Will require a different approach.

o Better dependence on domain-size.

Different argument that bypasses the union bound.

• Average-case algorithm.

o Polynomial time for constant individual degree?

Will require a different approach.

o Better dependence on domain-size.

Different argument that bypasses the union bound.

• Hardness of approximation.

- Is CutWidth hard to approximate up to a constant factor (without SSE)?
- Is ROABPwidth hard to approximate (for some other reason)?

Thank you!

ABP-figure credits: Prerona Chatterjee