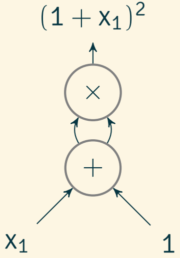
Multilinear Formula Lower Bounds for Sparse Determinants

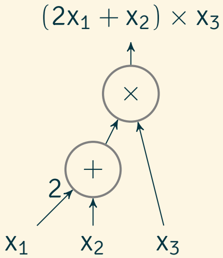
Pruthvi Boyapati Suryajith Chillara Pratyush Vempati

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD

CCC 2026, Lisbon
06.08.2026

Algebraic/Arithmetic circuits





An **Arithmetic Circuit** is a directed acyclic graph where

- leaf nodes: labelled by constants or variables,
- internal nodes: labelled by either $\times$ or $+$,
- edges: labelled by constants.

Circuit size: number of nodes present in it. [Measure of complexity]

Circuit depth: length of the longest leaf to root path. [Measure of parallelizability]

Formulas: circuits where computations are not reused, i.e., directed tree.

Multilinear polynomials & formulas

Multilinear polynomial

$$f \in \mathbb{F}[X] : \quad \deg_{x_i}(f) \leqslant 1 \text{ for every } x_i.$$

Examples: Determinant, Permanent, elementary symmetric polynomials.

Multilinear polynomials & formulas

Multilinear polynomial

$$f \in \mathbb{F}[X] : \quad \deg_{x_i}(f) \leqslant 1 \text{ for every } x_i.$$

Examples: Determinant, Permanent, elementary symmetric polynomials.

Syntactically multilinear (SML) formula

- Every $\times$ gate: variable-disjoint factors
- So no x_i^2 is ever formed

$$(x_1 + x_2)(x_1 + x_3) - (x_1 + x_4)(x_1 + x_2) = x_1x_3 + x_2x_3 - x_1x_4 - x_2x_4$$

Output multilinear; computes x_1^2 inside: not syntactically so.

Multilinear polynomials & formulas

Multilinear polynomial

$$f \in \mathbb{F}[X] : \quad \deg_{x_i}(f) \leqslant 1 \text{ for every } x_i.$$

Examples: Determinant, Permanent, elementary symmetric polynomials.

Syntactically multilinear (SML) formula

- Every $\times$ gate: variable-disjoint factors
- So no x_i^2 is ever formed

$$(x_1 + x_2)(x_1 + x_3) - (x_1 + x_4)(x_1 + x_2) = x_1x_3 + x_2x_3 - x_1x_4 - x_2x_4$$

Output multilinear; computes x_1^2 inside: not syntactically so.

For formulas: ML $=$ SML, w.l.o.g. [Raz, JACM 2009]

Power of cancellation

Determinant polynomial

$$\mathrm{Det}_n(X) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)}\,.$$

All known efficient algorithms exploit higher-degree cancellations:

- Samuelson-Berkowitz [Ann. Math. Statist. 1942; IPL 1984]
- Mahajan-Vinay [CJTCS 1997]
- ...

Question

Can multilinear formulas compute Det_n efficiently?

Raz's lower bound

Theorem [Raz, JACM 2009]

Every multilinear formula computing Det_n (or Perm_n) has size $n^{\Omega(\log n)}$.

Raz's lower bound

Theorem [Raz, JACM 2009]

Every multilinear formula computing Det_n (or Perm_n) has size $n^{\Omega(\log n)}$.

Key ideas

- Structural decomposition of ML formulas
- Measure: low on small formulas, high on Det_n

Raz's lower bound

Theorem [Raz, JACM 2009]

Every multilinear formula computing Det_n (or Perm_n) has size $n^{\Omega(\log n)}$.

Key ideas

- Structural decomposition of ML formulas
- Measure: low on small formulas, high on Det_n

This work

Is there a weaker structure than the full determinant that provides full hardness against ML formulas?

Known ML formula lower bounds (unrestricted depth)

All known bounds: $n^{\Omega(\log n)}$, all via Raz's technique.

- Det_n, Perm_n [Raz, JACM 2009]
- Full-rank f in ML-VP [Raz, ToC 2006]
- Arc-full-rank polynomial in ML-VBP
  [Dvir-Malod-Perifel-Yehudayoff, STOC 2012]

Known ML formula lower bounds (unrestricted depth)

All known bounds: $n^{\Omega(\log n)}$, all via Raz's technique.

- Det_n, Perm_n [Raz, JACM 2009]
- Full-rank f in ML-VP [Raz, ToC 2006]
- Arc-full-rank polynomial in ML-VBP
  [Dvir-Malod-Perifel-Yehudayoff, STOC 2012]

Weaker structures before us?

- DMPY: weaker class (ML-ABPs), tight separation
- But: polynomial engineered around arc-partitions
- No natural weakening of Det_n was known hard

Our result

Main Theorem

There exist $n \times n$ sparse symbolic matrices X_G:

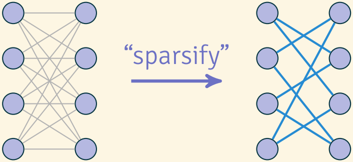
- $\Theta(\log^6 n)$ nonzero entries per row and column
- $\mathrm{Det}(X_G) \neq 0$
- ML formulas for $\mathrm{Det}(X_G)$ need size $n^{\Omega(\log n)}$

Our result

Main Theorem

There exist $n \times n$ sparse symbolic matrices X_G:

- $\Theta(\log^6 n)$ nonzero entries per row and column
- $\mathrm{Det}(X_G) \neq 0$
- ML formulas for $\mathrm{Det}(X_G)$ need size $n^{\Omega(\log n)}$



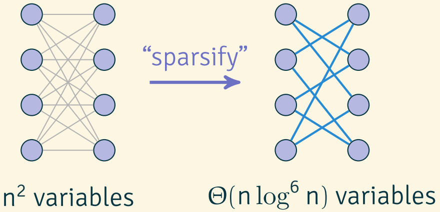
n^2 variables $\qquad$ $\Theta(n \log^6 n)$ variables

Our result

Main Theorem

There exist $n \times n$ sparse symbolic matrices X_G:

- $\Theta(\log^6 n)$ nonzero entries per row and column
- $\mathrm{Det}(X_G) \neq 0$
- ML formulas for $\mathrm{Det}(X_G)$ need size $n^{\Omega(\log n)}$



lower bound: $n^{\Omega(\log n)}$ — the best possible

Why doesn't this follow from [DMPY 2012]?

Tempting route:

- Arc-full-rank poly $\in$ VBP $\equiv$ projections of sparse determinants
- Arc-full-rank poly $\equiv \mathrm{Det}(X_{\mathrm{sparse}}) \circ \sigma$, some substitution σ
- Hardness of arc poly $\Rightarrow$ hardness of $\mathrm{Det}(X_{\mathrm{sparse}})$?

Why doesn't this follow from [DMPY 2012]?

Tempting route:

- Arc-full-rank poly $\in$ VBP $\equiv$ projections of sparse determinants
- Arc-full-rank poly $\equiv \mathrm{Det}(X_{\text{sparse}}) \circ \sigma$, some substitution σ
- Hardness of arc poly $\Rightarrow$ hardness of $\mathrm{Det}(X_{\text{sparse}})$?

No: reductions must preserve multilinearity

- σ repeats a variable across many entries
- ML formula $\circ\, \sigma$: repeated leaves — not ML anymore
- ML lower bounds don't transfer through projections in general

Why sparse determinants?

- **Want:** weaker structures, with superpolynomial ML hardness
- **Hope:** small general formulas for them
- Both together: ML $\subsetneq$ general formulas [Open Question 14, Shpilka-Yehudayoff, FnT-TCS 2010]
- Such separations are known at depth $o(\log n)$ [**C.**-Limaye-Srinivasan, SICOMP 2019]

Why sparse determinants?

- **Want:** weaker structures, with superpolynomial ML hardness
- **Hope:** small general formulas for them
- Both together: ML $\subsetneq$ general formulas [Open Question 14, Shpilka-Yehudayoff, FnT-TCS 2010]
- Such separations are known at depth $o(\log n)$ [**C.**-Limaye-Srinivasan, SICOMP 2019]

But sparse determinants are universal

- Every Det reduces to a sparse Det
- Via ABP staggering [Mahajan-Vinay, CJTCS 1997]; polynomial blow-up
- So worst-case sparse Det is VBP-hard

Why sparse determinants?

- **Want:** weaker structures, with superpolynomial ML hardness
- **Hope:** small general formulas for them
- Both together: ML $\subsetneq$ general formulas [Open Question 14, Shpilka-Yehudayoff, FnT-TCS 2010]
- Such separations are known at depth $o(\log n)$ [**C.**-Limaye-Srinivasan, SICOMP 2019]

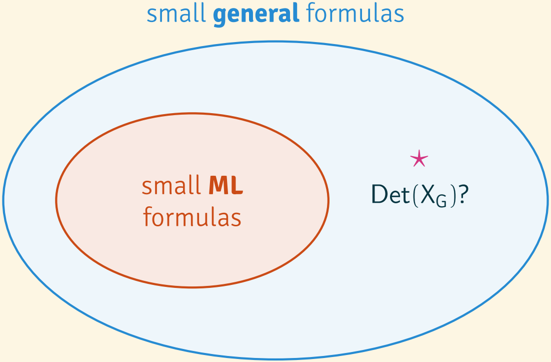
But sparse determinants are universal

- Every Det reduces to a sparse Det
- Via ABP staggering [Mahajan-Vinay, CJTCS 1997]; polynomial blow-up
- So worst-case sparse Det is VBP-hard

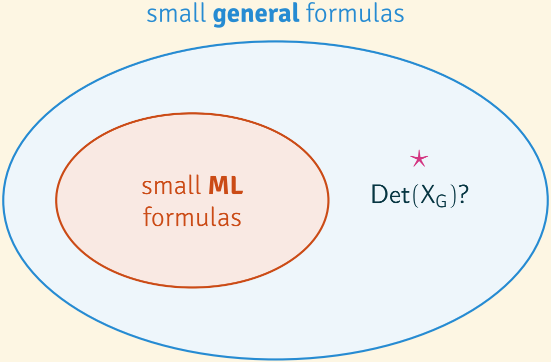
Question

Worst case is hard. Is our biased distribution formula-easy?

The target region



The target region



If $\mathrm{Det}(X_G)$ lands on $\star \Rightarrow$ separation.

Do ML-ABPs have small general formulas?

Probably not, in the worst case:

- $\mathsf{IMM}_{n,n}$: fresh variables per layer $\Rightarrow$ ML-ABP
- Yet $\mathsf{IMM}_{n,n}$ is VBP-complete
- Small formulas $\Rightarrow$ VF $=$ VBP
- Arc polynomial itself: formula size unclear

Do ML-ABPs have small general formulas?

Probably not, in the worst case:

- $\mathrm{IMM}_{n,n}$: fresh variables per layer $\Rightarrow$ ML-ABP
- Yet $\mathrm{IMM}_{n,n}$ is VBP-complete
- Small formulas $\Rightarrow \mathrm{VF} = \mathrm{VBP}$
- Arc polynomial itself: formula size unclear

The separation hope rides on non-complete instances.

Why this route, not the arc-full-rank polynomial route?

Arc-full-rank polynomial

- One rigid target with not many handles to exploit
- Full-rank by design, making it plausibly formula-hard

Why this route, not the arc-full-rank polynomial route?

Arc-full-rank polynomial

- One rigid target with not many handles to exploit
- Full-rank by design, making it plausibly formula-hard

Sparse determinants

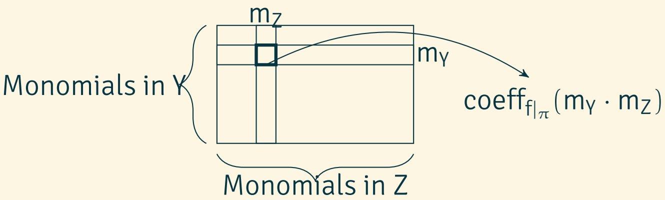
- A varied family & Raz's proof is robust
- Rich algebraic toolbox for Det_n algorithm design
- Reason for hope is due to results like: bounded treewidth $\Rightarrow$ poly formulas for Perm_n [Flarup-Koiran-Lyaudet, ISAAC 2007]

Raz's proof

Tool: Partial Derivative Matrix

- Partition $\pi : X \mapsto Y \sqcup Z$, with $|Y| = |Z|$

$M_{(Y,Z)}(f|_\pi)$:



- Measure: $\mu_\pi(f) = \mathrm{rank}(M_{(Y,Z)}(f|_\pi))$
- Trivially, $\mu_\pi(f) \leqslant 2^{|Y|}$

Rank depends on partitions (I)

$$f = (x_1 + x_2)(x_3 + x_4) = x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4$$

Split **each factor evenly**: $x_1, x_3 \mapsto y_1, y_2; \quad x_2, x_4 \mapsto z_1, z_2.$

$$f = y_1y_2 + y_1z_2 + y_2z_1 + z_1z_2$$

$$M_{(Y,Z)}(f) = \begin{array}{c|cccc} & 1 & z_1 & z_2 & z_1z_2 \\ \hline 1 & 0 & 0 & 0 & 1 \\ y_1 & 0 & 0 & 1 & 0 \\ y_2 & 0 & 1 & 0 & 0 \\ y_1y_2 & 1 & 0 & 0 & 0 \end{array}$$

Rank depends on partitions (I)

$$f = (x_1 + x_2)(x_3 + x_4) = x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4$$

Split **each factor evenly**: $x_1, x_3 \mapsto y_1, y_2; \quad x_2, x_4 \mapsto z_1, z_2.$

$$f = y_1y_2 + y_1z_2 + y_2z_1 + z_1z_2$$

$$M_{(Y,Z)}(f) = \begin{array}{c|cccc} & 1 & z_1 & z_2 & z_1z_2 \\ \hline 1 & 0 & 0 & 0 & 1 \\ y_1 & 0 & 0 & 1 & 0 \\ y_2 & 0 & 1 & 0 & 0 \\ y_1y_2 & 1 & 0 & 0 & 0 \end{array}$$

Permutation matrix: rank $= 4 = 2^{|Y|}$, full.

Rank depends on partitions (II)

Same f. Now factor-aligned: $x_1, x_2 \mapsto y_1, y_2; \quad x_3, x_4 \mapsto z_1, z_2$.

$$f = (y_1 + y_2)(z_1 + z_2)$$

$$M_{(Y,Z)}(f) = \begin{array}{c|cccc} & 1 & z_1 & z_2 & z_1z_2 \\ \hline 1 & 0 & 0 & 0 & 0 \\ y_1 & 0 & 1 & 1 & 0 \\ y_2 & 0 & 1 & 1 & 0 \\ y_1y_2 & 0 & 0 & 0 & 0 \end{array}$$

Identical rows: rank $= 1$, far from full.

Workplan

- Disjoint factors: $\mu_\pi(f) = \prod_j \mu_\pi(f_j)$
- Rank losses multiply across factors
- Small loss, many factors $\Rightarrow$ small total rank

Workplan

- Disjoint factors: $\mu_\pi(f) = \prod_j \mu_\pi(f_j)$
- Rank losses multiply across factors
- Small loss, many factors $\Rightarrow$ small total rank

Don't wreck every factor. Nick many of them.

Bounding a formula's rank

Log-product decomposition: $F = \sum_{i=1}^{s+1} f_i, \quad f_i = \prod_{j=1}^{t} f_{i,j}.$

- Subadditivity: $\mu_\pi(F) \leqslant \sum_i \mu_\pi(f_i)$
- Multiplicativity: $\mu_\pi(f_i) = \prod_j \mu_\pi(f_{i,j})$
- Suppose ℓ factors each lose $2^{-\delta}$:

$$\mu_\pi(f_i) \leqslant 2^{|X_i|/2} \cdot 2^{-\ell\delta} \quad \Longrightarrow \quad \mu_\pi(F) \leqslant (s+1) \cdot 2^{|X|/2} \cdot 2^{-\ell\delta}$$

Bounding a formula's rank

Log-product decomposition: $F = \sum_{i=1}^{s+1} f_i, \quad f_i = \prod_{j=1}^{t} f_{i,j}$.

- Subadditivity: $\mu_\pi(F) \leqslant \sum_i \mu_\pi(f_i)$
- Multiplicativity: $\mu_\pi(f_i) = \prod_j \mu_\pi(f_{i,j})$
- Suppose ℓ factors each lose $2^{-\delta}$:

$$\mu_\pi(f_i) \leqslant 2^{|X_i|/2} \cdot 2^{-\ell\delta} \quad \Longrightarrow \quad \mu_\pi(F) \leqslant (s+1) \cdot 2^{|X|/2} \cdot 2^{-\ell\delta}$$

$s+1$ summands cost only a union bound.

Complexity measure for Det_n

- $|\mathrm{vars}(\mathsf{Det}_n)| = n^2$; sparsity $\approx 2^{n \log n}$
- Equipartition: matrix has $2^{n^2/2}$ rows
- Rank can never approach $2^{n^2/2}$

Complexity measure for Det_n

- $|\mathrm{vars}(\mathsf{Det}_n)| = n^2$; sparsity $\approx 2^{n \log n}$
- Equipartition: matrix has $2^{n^2/2}$ rows
- Rank can never approach $2^{n^2/2}$

First reduce Det_n to something that can be full.

Reduce, then partition

- Random restriction ρ: set most variables to $\{0, 1\}$
- Target: $\mathrm{Det}_n\big|_\rho = \prod_{i=1}^{m}(x_{i_1} - x_{i_2})$
- Equipartition π: $x_{i_1} \mapsto y_i \in Y, \quad x_{i_2} \mapsto z_i \in Z$
- Under π: target becomes $\prod_i (y_i - z_i)$
- Only 2m variables left: full rank now possible

Reduce, then partition

- Random restriction ρ: set most variables to $\{0,1\}$
- Target: $\mathrm{Det}_n\big|_\rho = \prod_{i=1}^{m}(x_{i_1} - x_{i_2})$
- Equipartition π: $x_{i_1} \mapsto y_i \in Y, \quad x_{i_2} \mapsto z_i \in Z$
- Under π: target becomes $\prod_i (y_i - z_i)$
- Only 2m variables left: full rank now possible

A **good** pair (ρ, π)

- $\mu_\pi(\mathrm{Det}_n|_\rho) = 2^m$ full
- $\mu_\pi(F|_\rho) \ll 2^m$ deficient

Reduce, then partition

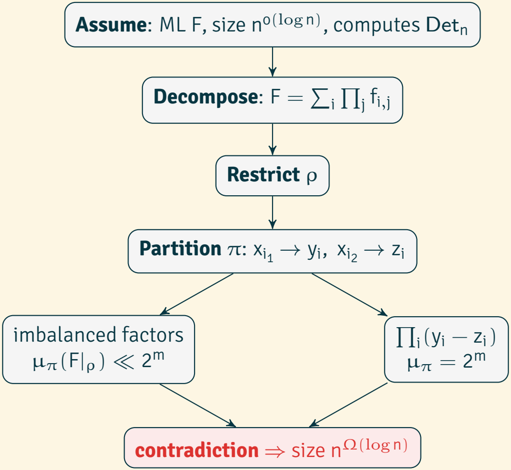
- Random restriction ρ: set most variables to $\{0, 1\}$
- Target: $\mathrm{Det}_n \big|_\rho = \prod_{i=1}^{m} (x_{i_1} - x_{i_2})$
- Equipartition π: $x_{i_1} \mapsto y_i \in Y, \quad x_{i_2} \mapsto z_i \in Z$
- Under π: target becomes $\prod_i (y_i - z_i)$
- Only $2m$ variables left: full rank now possible

A **good** pair (ρ, π)

- $\mu_\pi(\mathrm{Det}_n |_\rho) = 2^m$ full
- $\mu_\pi(F|_\rho) \ll 2^m$ deficient

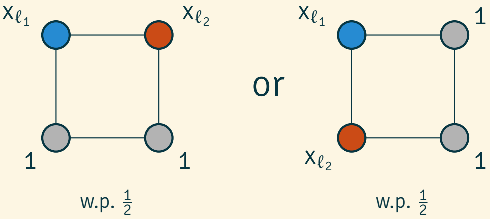
Goal: a good pair exists for every small F.

Raz's proof: the pipeline



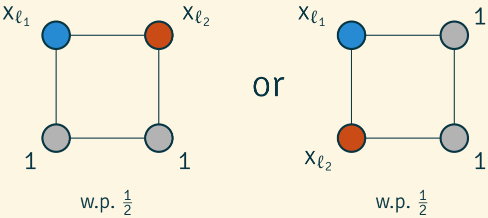
Raz's reduction: one block

$m = o(n)$ disjoint blocks on a **generalized diagonal**; name the active pair (x_{ℓ_1}, x_{ℓ_2}):



Raz's reduction: one block

$m = o(n)$ disjoint blocks on a **generalized diagonal**; name the active pair (x_{ℓ_1}, x_{ℓ_2}):



Either way: $\det \begin{pmatrix} x_{\ell_1} & \cdot \\ \cdot & 1 \end{pmatrix} = x_{\ell_1} - x_{\ell_2}$.

Why randomize the x_{ℓ_2} corner?

- F's decomposition is fixed before sampling
- Fixed corners: factors align, swallow both actives
- Aligned blocks contribute no imbalance
- Coin flip: alignment cannot be pre-planned

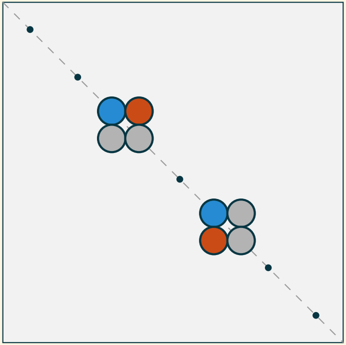
Why randomize the x_{ℓ_2} corner?

- F's decomposition is fixed before sampling
- Fixed corners: factors align, swallow both actives
- Aligned blocks contribute no imbalance
- Coin flip: alignment cannot be pre-planned

Randomness defeats every formula, not just known ones.

Assembling the reduction

Blocks on a **generalized diagonal**: no shared rows or columns.



diagonal blocks active; diagonal rest $= 1$; elsewhere $= 0$

$$\mathrm{Det}_n\big|_\rho = \prod_{i=1}^{m}(x_{i_1} - x_{i_2}) \xrightarrow{\pi} \prod_{i=1}^{m}(y_i - z_i)$$

$\mu_\pi = 2^m$ (always full)

Putting it together

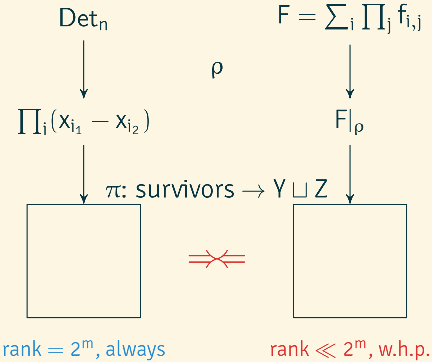
- Fix F of size $s \leqslant n^{o(\log n)}$
- Each summand: $\Theta(\log n)$ large factors
- Each large factor: imbalanced with high probability
- Concentration: $\Theta(\log n)$ imbalanced factors per summand
- Union bound over $s + 1$ summands
- Meanwhile: $\mu(\mathrm{Det}_n|_\rho) = 2^m$, always

Putting it together

- Fix F of size $s \leqslant n^{o(\log n)}$
- Each summand: $\Theta(\log n)$ large factors
- Each large factor: imbalanced with high probability
- Concentration: $\Theta(\log n)$ imbalanced factors per summand
- Union bound over $s + 1$ summands
- Meanwhile: $\mu(\mathrm{Det}_n|_\rho) = 2^m$, always

Good (ρ, π) exists — probabilistic method.

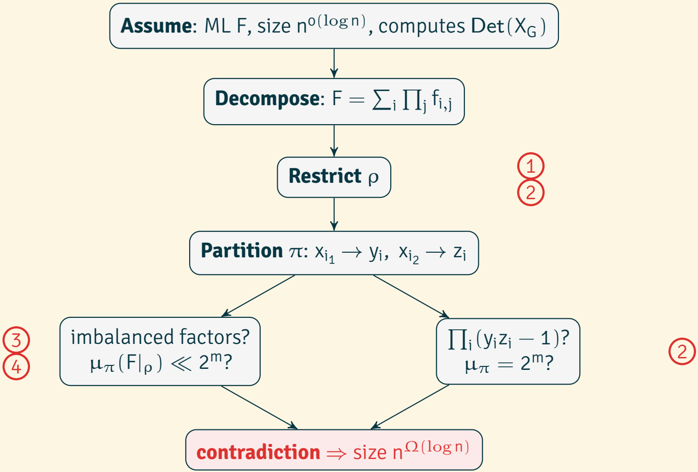
Raz's proof, in one picture



No ML formula of size $n^{o(\log n)}$.

What breaks for sparse?

Sparse case: where the pipeline breaks



(1) blocks in G? (2) matching, $\mathrm{Det} \neq 0$? (3) alignment? (4) dependence?

Where does the proof fail for $\mathrm{Det}(X_G)$?

(1) Blocks must exist inside G

(2) Residual graph must keep a perfect matching

(3) Few blocks per edge: alignment attacks possible

(4) Shared edges $\Rightarrow$ dependent randomness

Where does the proof fail for $\mathrm{Det}(X_G)$?

(1) Blocks must exist inside G

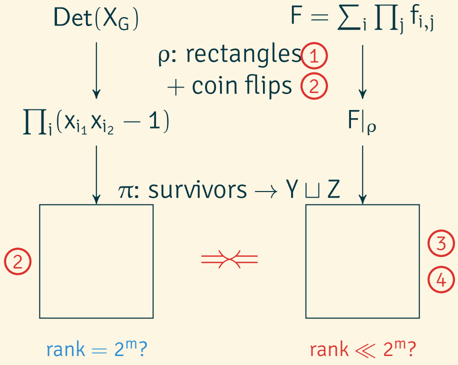
(2) Residual graph must keep a perfect matching

(3) Few blocks per edge: alignment attacks possible

(4) Shared edges $\Rightarrow$ dependent randomness

In the dense case, all four are free.

The same picture, annotated



(1) blocks exist in G? — fix: sample G from rectangles

(2) residual matching, Det $\neq 0$? — fix: Hall + deletion invariance

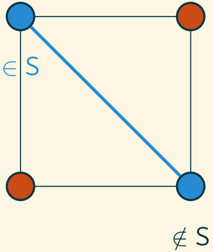
(3) alignment attacks? — fix: diagonal imbalance via random σ

(4) dependence? — fix: McDiarmid, read-2 Chernoff [Gavinsky et al. 2015]

Fixes 1 & 2: rectangle sampling

- Sample each $K_{2,2}$ of $K_{n,n}$ w.p. $\frac{\log^6 n}{n^3}$
- Keep every edge in a sampled rectangle
- G: degree $\Theta(\log^6 n)$, built from rectangles
- Round 2: sample m rectangles from $\mathcal{R}$ – they exist
- Deletion invariance: residual $\sim$ fresh sampling
- Hall's condition $\Rightarrow$ matching survives, Det $\neq 0$

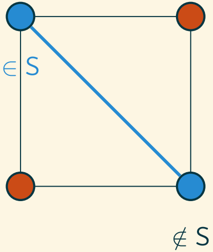
Fix 3: why diagonal imbalance?



exactly one diagonal corner in S

- Dense: each edge in $(n-1)^2$ blocks
- Sparse: as few as one
- Adversarial S can align with corners
- Fix: random permutation σ of rows/columns
- Every large S: $\approx |S|/2$ diagonal rectangles

Fix 3: why diagonal imbalance?



exactly one diagonal corner in S

- Dense: each edge in $(n-1)^2$ blocks
- Sparse: as few as one
- Adversarial S can align with corners
- Fix: random permutation σ of rows/columns
- Every large S: $\approx |S|/2$ diagonal rectangles

Diagonal rectangles are what create imbalance.

Fix 4: taming dependence

McDiarmid

- Diagonal count: function of dependent permutation σ
- Bounded differences: one swap moves $O(\log^6 n)$ rectangles

Fix 4: taming dependence

McDiarmid

- Diagonal count: function of dependent permutation σ
- Bounded differences: one swap moves $O(\log^6 n)$ rectangles

Read-2 Chernoff

- Coin flip C_ℓ touches two active variables
- So C_ℓ affects at most **2** factors
- Balance events: a read-2 family
- Tail bound: $e^{-\Omega(\log^2 n)}$ per summand

[Gavinsky-Lovett-Saks-Srinivasan, RS&A 2015]

Fix 4: taming dependence

McDiarmid

- Diagonal count: function of dependent permutation σ
- Bounded differences: one swap moves $O(\log^6 n)$ rectangles

Read-2 Chernoff

- Coin flip C_ℓ touches two active variables
- So C_ℓ affects at most **2** factors
- Balance events: a read-2 family
- Tail bound: $e^{-\Omega(\log^2 n)}$ per summand

[Gavinsky-Lovett-Saks-Srinivasan, RS&A 2015]

Union bound over $s + 1$ summands: done.

The final count

- Restricted determinant: $\prod_i (y_i z_i - 1)$, rank 2^m
- Each summand: rank $\leqslant 2^m \cdot 2^{-\Omega(n^{1/32} \log n)}$
- Total: $(s+1) \cdot 2^m \cdot 2^{-\Omega(n^{1/32} \log n)} \ll 2^m$

The final count

- Restricted determinant: $\prod_i (y_i z_i - 1)$, rank 2^m
- Each summand: rank $\leqslant 2^m \cdot 2^{-\Omega(n^{1/32} \log n)}$
- Total: $(s+1) \cdot 2^m \cdot 2^{-\Omega(n^{1/32} \log n)} \ll 2^m$

ML formulas for $\mathrm{Det}(X_G)$ need size $n^{\Omega(\log n)}$.

Thank you!