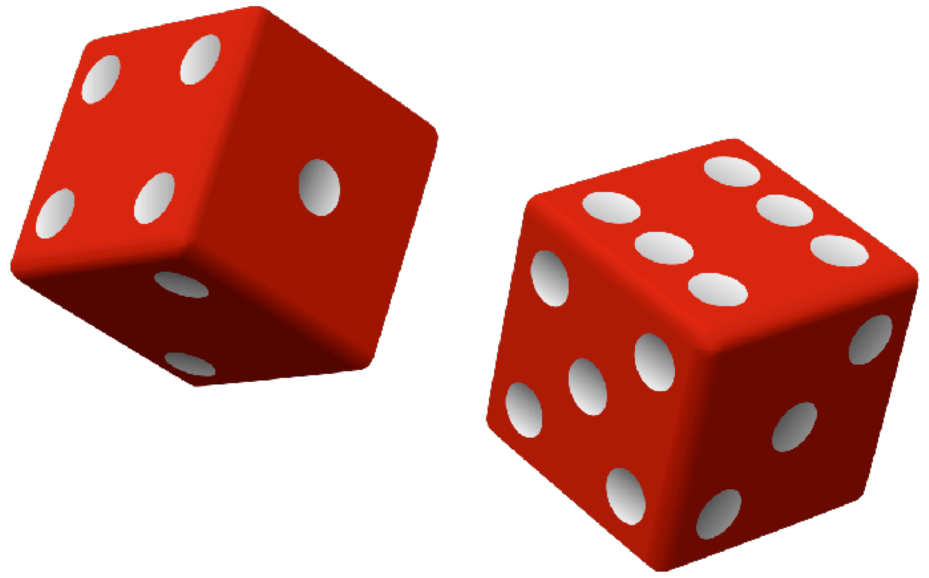
Low-degree polynomials are good extractors

Omar Alrabiah
UC Berkeley

Jesse Goodman
UT Austin

Jonathan Mosheiff
Ben-Gurion U



João Ribeiro
U Lisboa

How biased is a random function?

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a uniformly random function

$$\text{bias}(f) = \Pr_{x \sim \mathbb{F}_2^n}[f(x) = 0] - \Pr_{x \sim \mathbb{F}_2^n}[f(x) = 1]$$

Most functions are nearly unbiased:

$$\Pr_f[\,|\text{bias}(f)| > \varepsilon] \le 2^{-\Omega(\varepsilon^2 2^n)}$$

How biased is a random low-degree polynomial?

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a random degree d polynomial

$$f(x) = \sum_{S \subseteq [n], |S| \leq d} \alpha_S x^S, \quad \text{with i.i.d. } \alpha_S \sim \mathbb{F}_2$$

f is very far from a uniformly random function!

Bias of random low-degree polynomials

[Ben-Eliezer, Hod, Lovett 2008]

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a random degree d polynomial

$$\Pr_f[\,|\text{bias}(f)| > 2^{-cn/d}] \leq 2^{-c\binom{n}{\leq d}}$$

Moment argument. Very roughly,

- t-th moment of $|\text{bias}(f)|$ is probability that $p(x_1) + \cdots + p(x_t) = 0$ for all degree-d polynomials p, with $x_1, \ldots, x_t \sim \mathbb{F}_2^n$.
- This probability is controlled by dimension of puncturing of Reed-Muller code to t random coordinates.

Some applications

[Ben-Eliezer, Hod, Lovett 2008]

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a random degree d polynomial

$$\Pr_f[\,|\text{bias}(f)| > 2^{-cn/d}] \le 2^{-c\binom{n}{\le d}}$$

- Concentration bounds for weight distribution of Reed-Muller codes.
- Most degree d polynomials are hard to approximate by degree $d-1$ polynomials.
- Time-space tradeoffs for learning low-degree polynomials from random evaluations.

Generalizing "bias"

There are many notions of "bias" beyond "behavior on uniform input"!
In particular, can consider behavior on input $x \sim \mathbf{X}$.

$$\text{bias}_{\mathbf{X}}(f) = \Pr_{x \sim \mathbf{X}}[f(x) = 0] - \Pr_{x \sim \mathbf{X}}[f(x) = 1]$$

$\Pr[\mathbf{X} = x] \leq 2^{-k}$ for all $x \in \mathbb{F}_2^n$

Most functions are nearly unbiased on a k-**source** $\mathbf{X}$:

$$\Pr_f[\,|\text{bias}_{\mathbf{X}}(f)| > \varepsilon] \leq 2^{-\Omega(\varepsilon^2 2^k)}$$

How biased is a random low-degree polynomial on a k-source?

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a random degree d polynomial

Simple example: Take $\mathbf{X}$ uniform over k-dimensional subspace $V \subseteq \mathbb{F}_2^n$.

Restriction of f to V is random k-variate polynomial of degree d.

$$\implies \qquad \Pr_f[\,|\,\mathrm{bias}_{\mathbf{X}}(f)\,| > 2^{-ck/d}] \le 2^{-c\binom{k}{\le d}}$$

How biased is a random low-degree polynomial on a k-source?

$f : \mathbb{F}_2^n \to \mathbb{F}_2$ a random degree d polynomial

Bias on uniform input generalizes easily to all "affine sources".

How about arbitrary k-sources?

For any k-source $\mathbf{X}$:

$$\Pr_f[\,|\,\mathrm{bias}_{\mathbf{X}}(f)\,| > 2^{-ck/d}] \leq 2^{-c\binom{k}{\leq d}}$$

Let f be a random degree-d polynomial. Then, for any k-source $\mathbf{X}$:

$$\Pr_f[\,|\mathsf{bias}_{\mathbf{X}}(f)| > 2^{-ck/d}] \le 2^{-c\binom{k}{\le d}}$$

Proof idea: We generically reduce to the "uniform input" case.

1. For any linear map $L : \mathbb{F}_2^n \to \mathbb{F}_2^m$ and $g : \mathbb{F}_2^m \to \mathbb{F}_2$ a random degree-d polynomial,

$$\text{moments of } |\mathsf{bias}_{\mathbf{X}}(f)| \le \text{moments of } |\mathsf{bias}_{\mathbf{L(X)}}(g)|$$

2. By leftover hash lemma, there is L with $m \approx k$ such that $L(\mathbf{X}) \approx U_m$.

3. Apply rest of the Ben-Eliezer, Hod, Lovett argument for uniform input.

Low-degree polynomials as extractors

With high prob, random degree-d polynomial is nearly unbiased on any small enough class of sources $\mathscr{C}$. **In other words, f is a low-error extractor for $\mathscr{C}$.**

Direct via union bound!

Examples:

- Affine sources
- Locally-samplable sources
- Polynomial sources
- Variety sources

Concurrent work:

Golovnev, Guo, Hatami, Nagargoje, Yan (RANDOM 2024) obtained similar results with polynomially-small error.

Can we take this even further?

We saw that random degree-d polynomials are good extractors for all small classes of sources.

What about large *but structured* classes of sources?

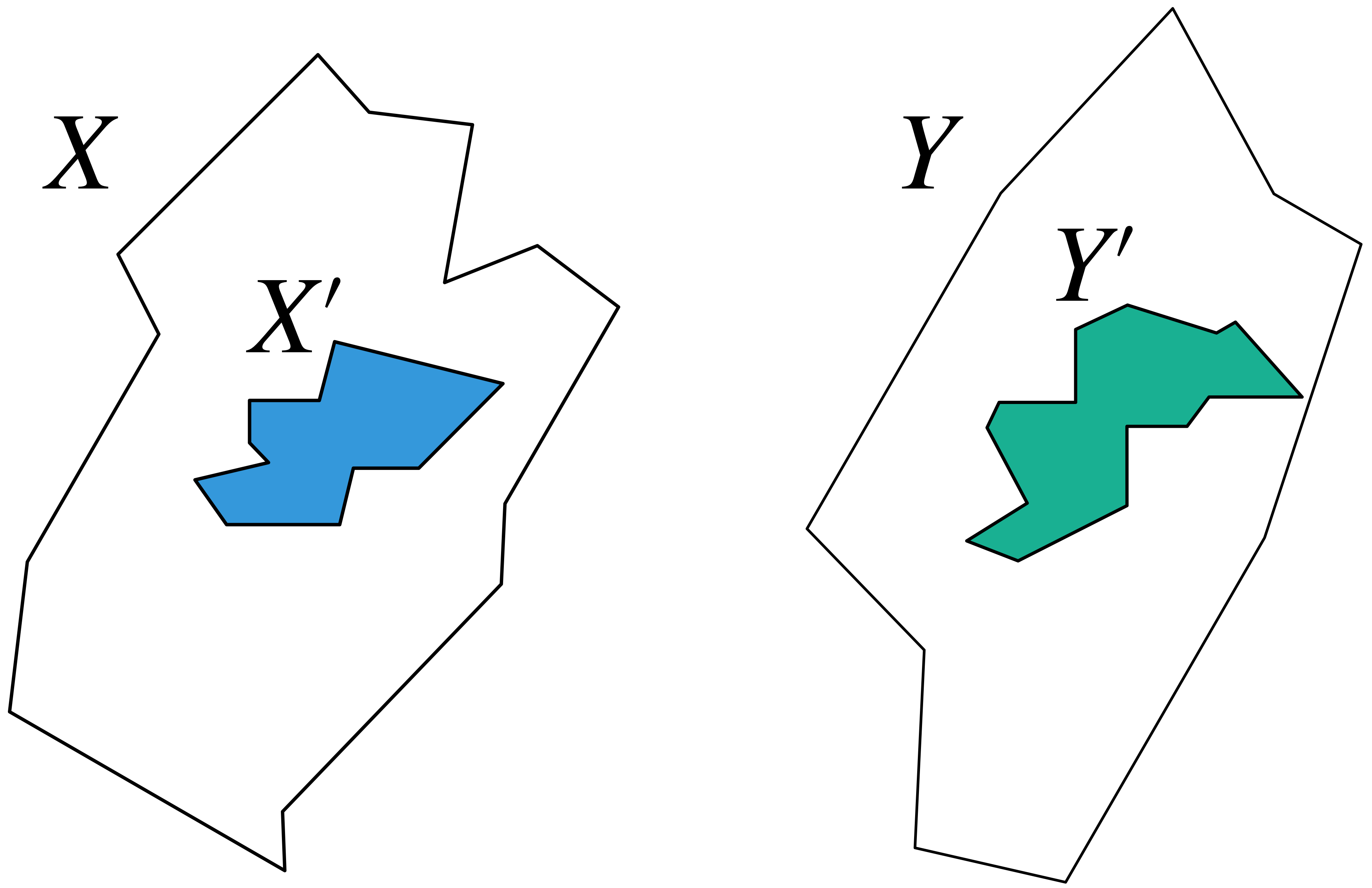
- Two independent sources: $(\mathbf{X}, \mathbf{Y})$
- Sumset sources: $\mathbf{W} = \mathbf{X} + \mathbf{Y}$ the most general so far

Some of the best explicit **low-error** extractors we know for these classes are low-degree polynomials over small fields.

How biased is a random function vs sumset sources?

Not so easy anymore…

Naive application of probabilistic method fails. There are $\approx 2^{n2^k}$ pairs of sets (X, Y) each of size 2^k, but $X + Y$ can also have size 2^k.



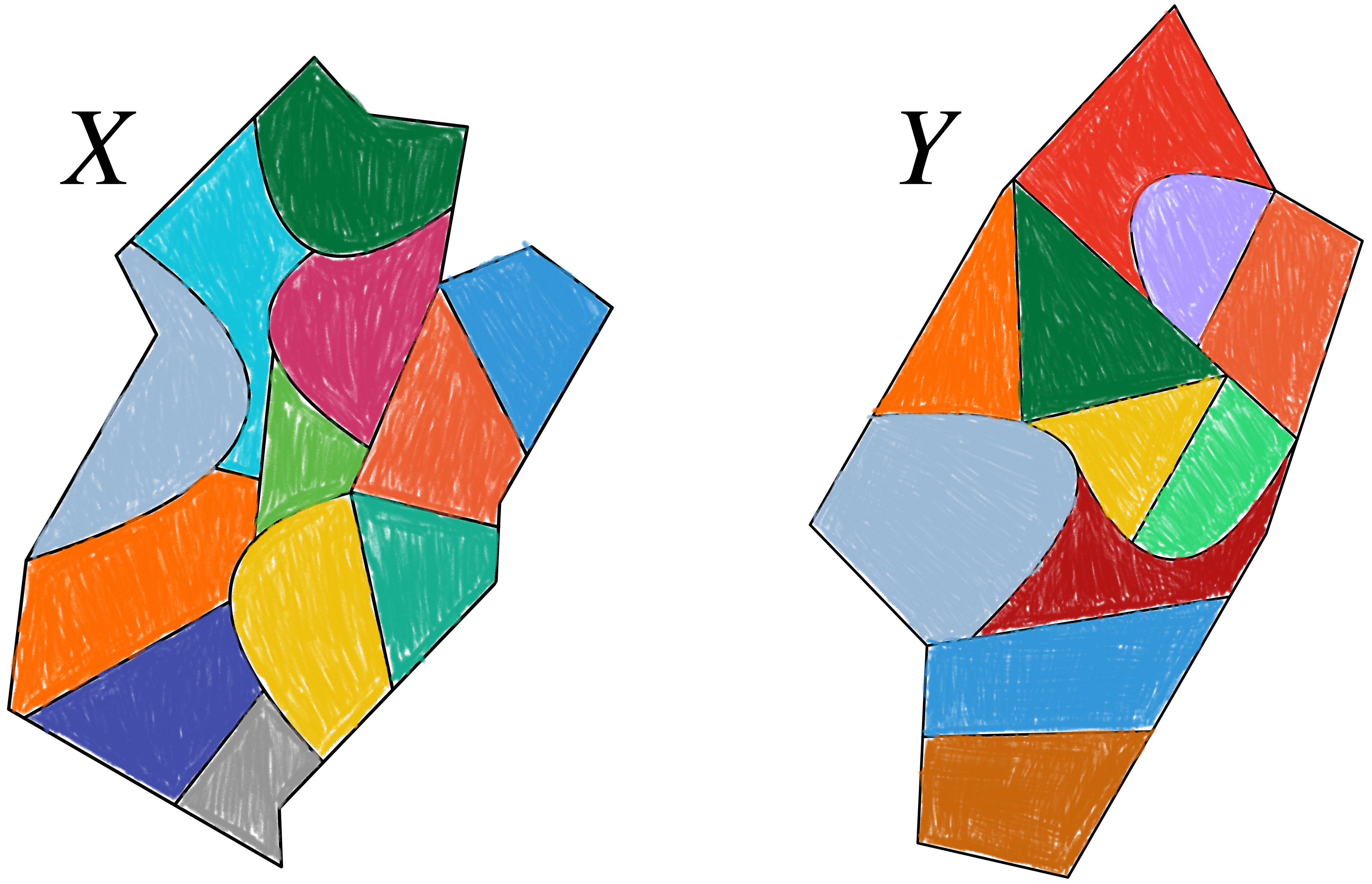
Idea: Find not-too-small $X' \subseteq X$ and $Y' \subseteq Y$ such that $|X' + Y'| \approx |X'| \cdot |Y'|$.

[Mrazović 2016]
Take random subsets of X and Y!
Can achieve $|X'| \approx \sqrt{|X|}$, $|Y'| \approx \sqrt{|Y|}$.

How biased is a random function vs sumset sources?

Not so easy anymore…

Naive application of probabilistic method fails. There are $\approx 2^{n2^k}$ pairs of sets (X, Y) each of size 2^k, but $X + Y$ can have size 2^k.



In fact, can always partition X and Y into not-too-small (X_i) and (Y_j) such that $|X_i + Y_j| \approx |X_i| \cdot |Y_j|$, for all i, j.

Take independent random partitions of X and Y into equal-size subsets!

Low-degree polynomials vs sumset sources

For even d, with high prob a random degree-d polynomial $f : \mathbb{F}_2^n \to \mathbb{F}_2$ has bias ε on the class of k-sumset sources with entropy $k \approx d(n/\varepsilon^2)^{2/d}$.

Some interesting regimes:

- For fixed degree d, get bias $\varepsilon = o(1)$ and min-entropy $k \approx dn^{2/d}$.
- $k = \Omega(dn^{\frac{1}{d-1}})$ is necessary even for constant bias ε. [Cohen-Tal 2015]
- Get min-entropy $k = O(\log(n/\varepsilon))$ with degree $d = O(\log(n/\varepsilon))$, for any ε.

An easier special case

For even d, with high prob a random degree-d polynomial $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is **non-constant** on every k-sumset $X + Y$ with $k \approx dn^{2/d}$.

How to control $\Pr_f[f(W) \equiv 0]$ for a set W ?

$$M_d^W = \begin{pmatrix} w_1^{S_1} & w_1^{S_2} & \cdots \\ w_2^{S_1} & w_2^{S_2} & \cdots \\ \vdots & \vdots & \vdots \end{pmatrix} \in \mathbb{F}_2^{|W| \times \binom{n}{\le d}}$$

$f(W) = M_d^W \times v_f$ → unif. random coeff. vector

$$\mathrm{rank}_d(W) = \mathrm{rank}(M_d^W)$$

$$\Pr_f[f(W) \equiv 0] \le 2^{-\mathrm{rank}_d(W)}$$

Naive union bound is hopeless…

There are $\approx 2^{2n2^k}$ choices for (X, Y), but $\mathrm{rank}_d(X+Y) \le \binom{n}{\le d} \le dn^d$.

An easier special case

For even d, with high prob a random degree-d polynomial $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is **non-constant** on every k-sumset $X + Y$ with $k \approx dn^{2/d}$.

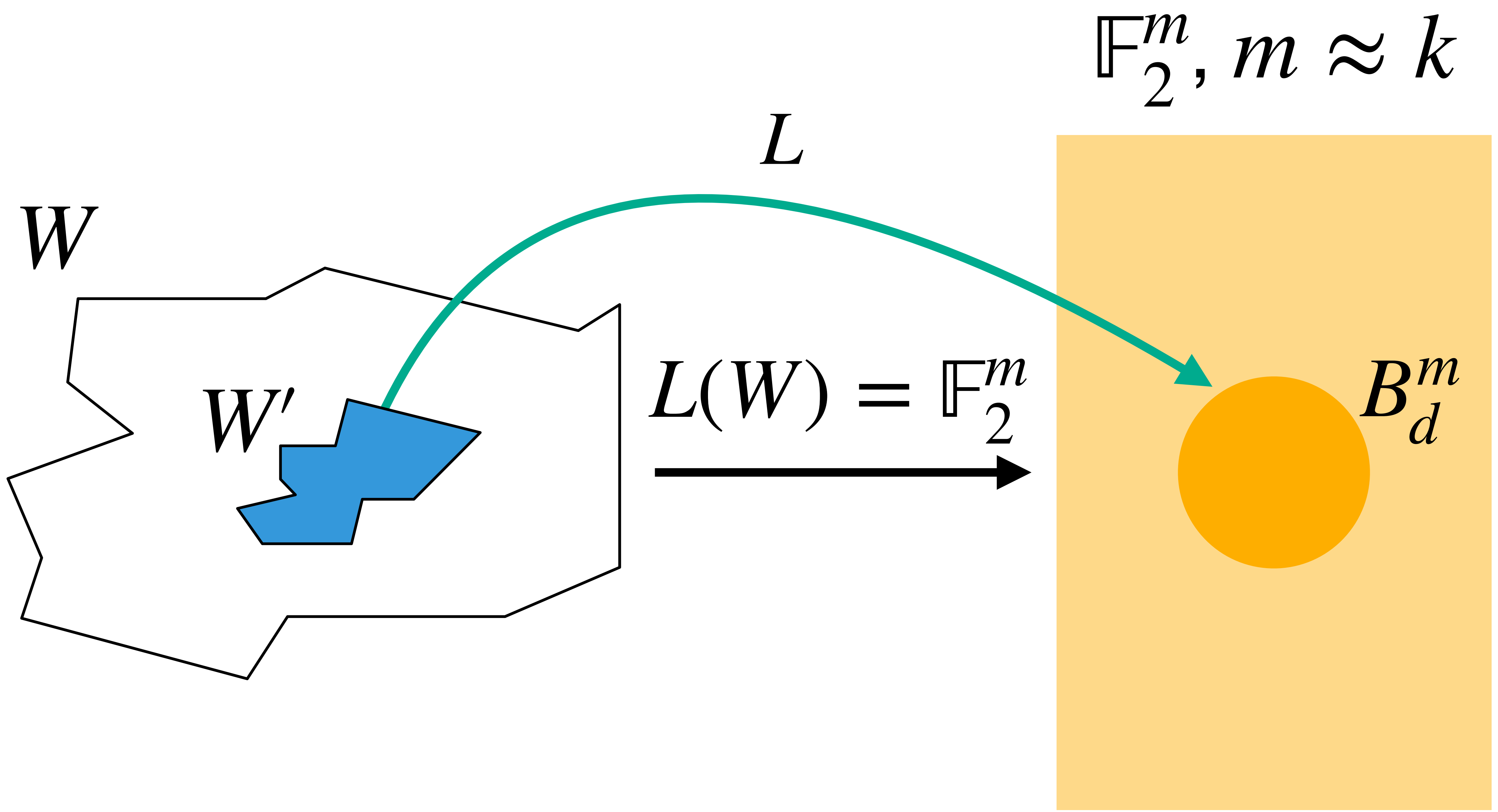
Proof idea: Find large $X' \subseteq X$ and $Y' \subseteq Y$ such that $\mathrm{rank}_d(X' + Y')$ is large.

[Keevash-Sudakov 2005] For every $W \subseteq \mathbb{F}_2^n$ of size 2^k there is $W' \subseteq W$ of size $\binom{k}{\leq d}$ such that $\mathrm{rank}_d(W') = |W'|$.

But we need W' to be a sumset!

A simple proof of ≈ Keevash-Sudakov

Goal: For W of size 2^k, find $W' \subseteq W$ of size $\approx \binom{k}{\leq d}$ such that $\text{rank}_d(W') = |W'|$.



$$\begin{aligned} &\text{rank}_d(W') \\ &\geq \text{rank}_d(L(W')) \\ &= \text{rank}_d(B_d^m) \\ &= \binom{m}{\leq d} \approx \binom{k}{\leq d} \end{aligned}$$

Upgrading to sumsets with large rank_d

Goal: Find large $X' \subseteq X$ and $Y' \subseteq Y$ such that $\mathrm{rank}_d(X' + Y') \approx |X'| \cdot |Y'|$.

Warmup: $X = Y = \mathbb{F}_2^k$

$B_d^k =$ radius-d Hamming ball in $\mathbb{F}_2^k$

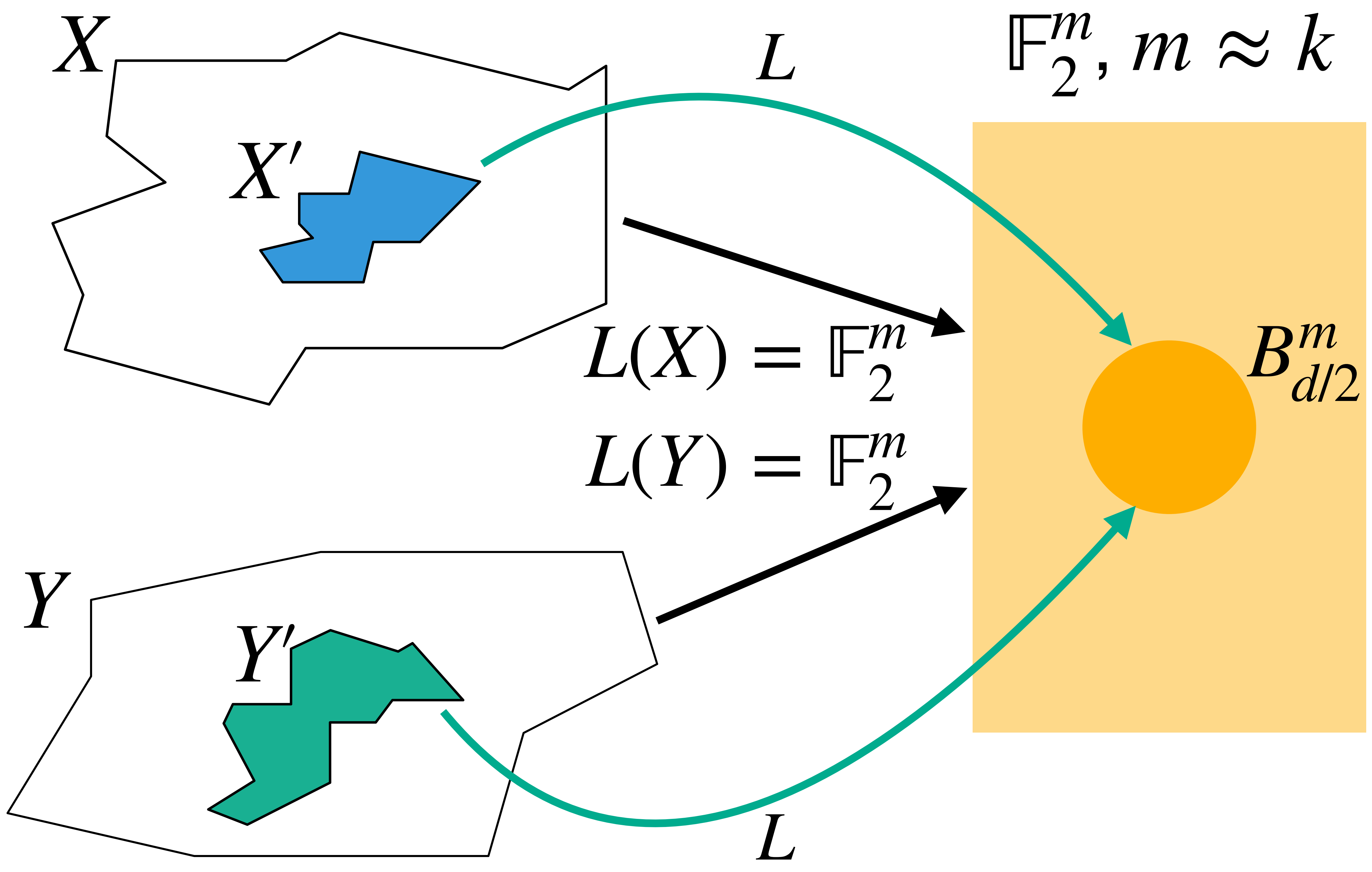
$B_{d/2}^k + B_{d/2}^k = B_d^k$

$$\mathrm{rank}_d(B_{d/2}^k + B_{d/2}^k) = \mathrm{rank}_d(B_d^k) = |B_d^k| = \binom{k}{\le d} \approx |B_{d/2}^k|^2$$

Can we generalize this?

Upgrading to sumsets with large rank_d

Goal: Find large $X' \subseteq X$ and $Y' \subseteq Y$ such that $\mathrm{rank}_d(X' + Y') \approx |X'| \cdot |Y'|$.



$$\begin{aligned} &\mathrm{rank}_d(X' + Y') \\ &\geq \mathrm{rank}_d(L(X' + Y')) \\ &= \mathrm{rank}_d(B^m_{d/2} + B^m_{d/2}) \\ &= \mathrm{rank}_d(B^m_d) = |B^m_d| \approx |X'| \cdot |Y'| \end{aligned}$$

Now the union bound works if $k \geq dn^{2/d}$

There exist subsets X', Y' of size $\approx \sqrt{dk^d}$ such that $\mathrm{rank}_d(X' + Y')) \approx |X'| \cdot |Y'|$.

Number of choices for X' and Y' is

$$\binom{2^n}{|X'|} \cdot \binom{2^n}{|Y'|} \leq 2^{n\sqrt{dk^d}}$$

Random degree-d polynomial f is constant on $X' + Y'$ with probability

$$\leq 2^{-\mathrm{rank}_d(X'+Y')} \approx 2^{-|X'|\cdot|Y'|} \approx 2^{-dk^d}$$

Achieving smaller bias vs sumsets

Previous strategy shows that most degree-d polynomials are high-error sumset extractors. We can extend this to lower bias.

Idea: Show that $\mathbf{X}$ and $\mathbf{Y}$ are close to convex combinations $(\mathbf{X}_i)$ and $(\mathbf{Y}_j)$ with $\mathrm{rank}_d(X_i + Y_j) = |X_i| \cdot |Y_j|$ for all i, j.

But… Independently and randomly selecting $X' \subseteq X$ and $Y' \subseteq Y$ doesn't work anymore! Need to choose $X' \subseteq X$ and $Y' \subseteq Y$ in a correlated manner.

Bonus

- Most degree-4 polynomials are 2-source extractors with exponentially-small error for min-entropy $k \approx n/\log n$.
  Polynomial Freiman-Ruzsa + Approximate Duality [Ron-Zewi—Ben-Sasson 2011] + subspace-evasive sets from degree-2 polynomials.

- Improved impossibility results for sumset dispersers vs. polynomial sources.

Wrapping up

- Random low-degree polynomials are unbiased in a very general sense.
- **Small classes of sources:** Most low-degree polynomials are low-error extractors.
- **Sumset sources:** Most low-degree polynomials are high-error extractors

Open problems:

- Constant-degree polynomials compute low-error sumset extractors?
- Constant-degree polynomials compute low-error 2-source extractors for min-entropy $\ll n/\log n$?

Thanks!