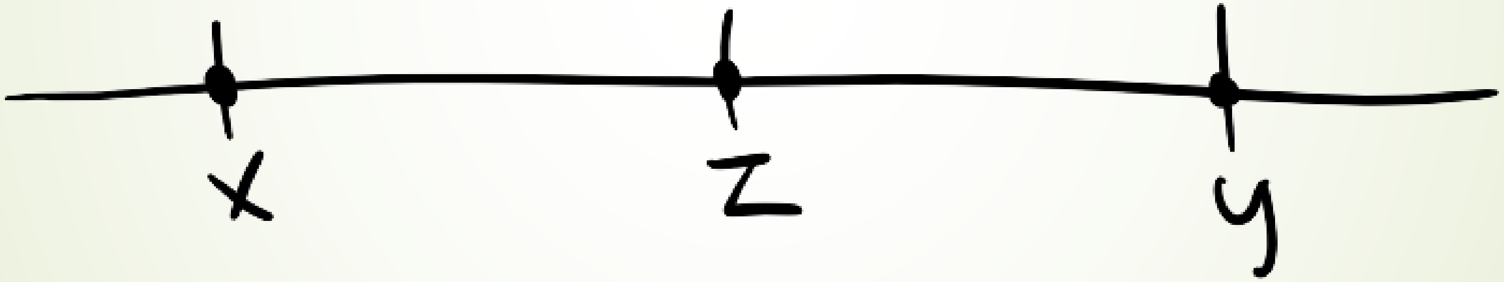
Strong Bounds for 3-Progressions

Zander Kelley (UIUC), Raghu Meka (UCLA)

3-Term Arithmetic Progressions

- Triple (x, z, y) with $x + y = 2z$



- "trivial" when $x = y = z$

3-Term Arithmetic Progressions

Theorem (Roth '53)

If $A \subseteq \{1, 2, \dots, N\}$ is dense enough*,
where density $\delta := \frac{|A|}{N}$,
then A must have a (nontrivial) 3-progression.

* (density threshold $\delta \approx {}^{1}/_{\log\log N}$)

History ($A \subseteq [N], A \geq \delta N \Rightarrow$ 3-progression)

$\delta \approx 1/\log\log N$	(Roth '53)
$\delta \approx 1/\log(N)^c, c > 0$	(Heath-Brown '87) (Szemerédi '90)
$\delta \approx 1/\log(N)^{2/3}$	(Bourgain '08)
$\delta \approx (\log\log N)^{O(1)}/\log(N)$	(Sanders '12)
$\delta \approx 1/\log(N)^{1+c}, c > 0$	(Bloom-Sisask '20)

Our Result

Theorem (K-Meka '23)

If $A \subseteq [N]$ is dense enough*, then A must have a (nontrivial) 3-progression.

* (density threshold $\delta \approx 2^{-\log(N)^{1/12}}$)

- Compare to lower bound, $\delta \approx 2^{-\log(N)^{1/2}}$

Dense sets have many 3-progressions

Theorem (K-Meka '23)

If $A \subseteq [N]$, $|A| \geq 2^{-d}N$ then A has $\sim 2^{-d^{12}}N^2$ solutions to $x + y = 2z$

- (At most $|A| \leq N$ trivial solutions)

3-Progression over finite abelian G

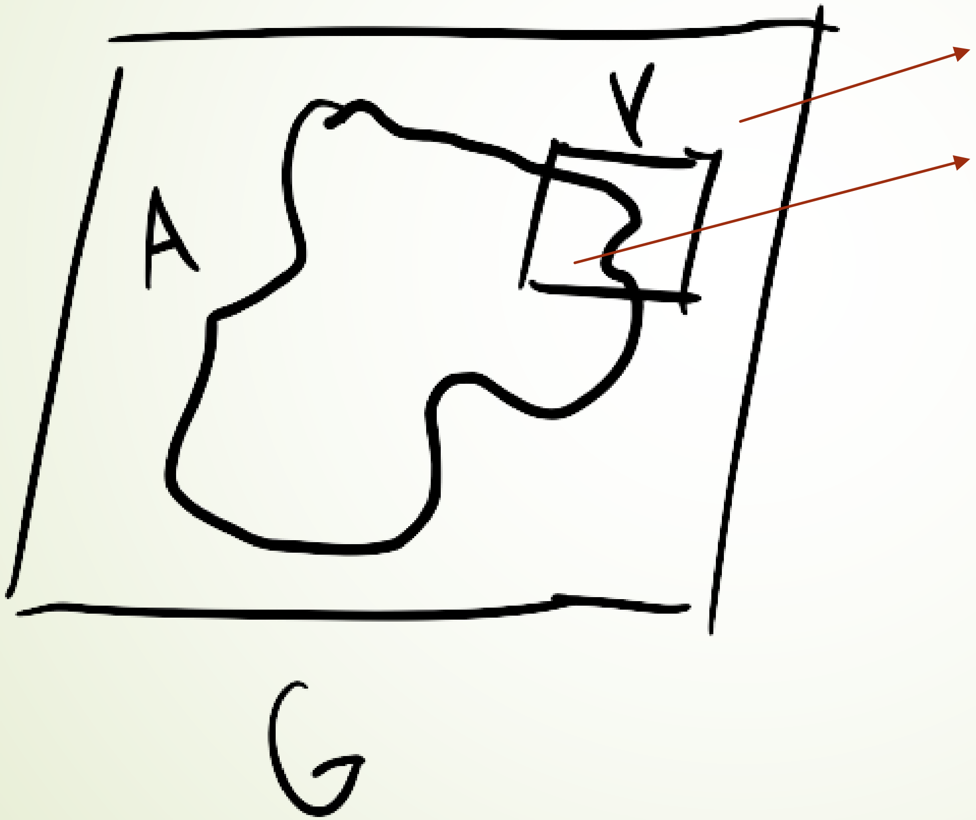
- If $A \subseteq G$, we can ask if A must have many solutions to $x + y = 2z$ (in G).
  - ($A \subseteq [N]$, $|A| \geq 2^{-d}N \;\Rightarrow\; 2^{-d^{12}}N^2$ solutions.)
  - $A \subseteq [N]$, $|A| \geq 2^{-d}|\mathbb{F}_q^n| \;\Rightarrow\; 2^{-d^9}|\mathbb{F}_q^n|^2$ solutions.
  - $A \subseteq G$, $|A| \geq 2^{-d}|G| \;\Rightarrow\; 2^{-d^{12}}|G|^2$ solutions. [BS ‘23]

↳ ($G = \mathbb{Z}_n$ is roughly equivalent to $[N]$)

The "Analytic" Approach ($A \subseteq G$)

- Find $\mathrm{A}' \subseteq A$, with $\approx \frac{|A'|^3}{|G|}$ solutions to $x + y = 2z$.
- (Want A' large)
- E.g. try $A' = A \cap V$, V structured:
  - V = translate of some approximate subgroup:
    - Subgroup
    - Bohr set
    - Generalized Arithmetic Progression

The "Analytic" approach



- V = structured set.
- $A' = A \cap V$ has the "right" number of solutions to $x + y = 2z$

$$(= (1 \pm \epsilon)\frac{|A'|^3}{|V|}.)$$

(ϵ is some small constant, like 1/10)

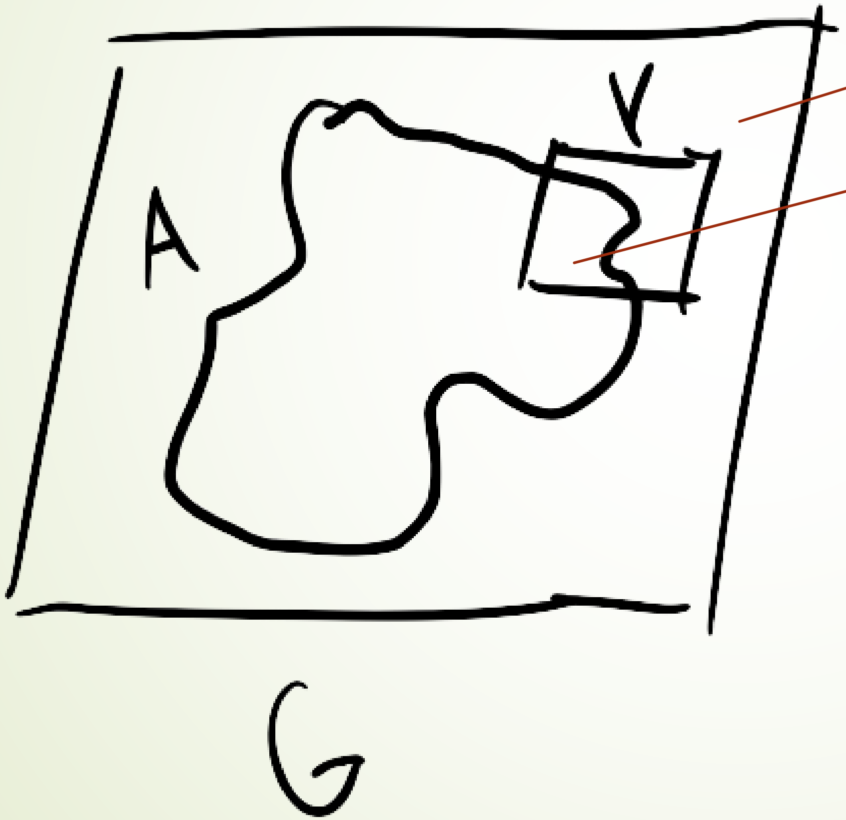
Approximate Subgroups

- Example: $I = [-m, m] \subseteq \mathbb{Z}$.
- For generic sets $S \subseteq \mathbb{Z}$, we expect
$$|S + S| \approx |S|^2$$
- In contrast, $|I + I| = 2|I|$:

  "approximately closed under addition"

$A \subseteq \mathbb{F}_q^n$, V = subgroup		$A \subseteq [N]$, $V \approx$ subgroup
$\delta = 1/\log(N)$ (Roth)	→	$\delta = 1/\log\log(N)$ $\delta = 1/\log(N)^c$, c > 0 $\delta = 1/\log(N)^{2/3}$ $\delta = \log\log(N)^{O(1)}/\log(N)$
$\delta = 1/\log(N)^{1+c}$ (BK '12)	→	$\delta = 1/\log(N)^{1+c}$ (BS '20)
$\delta = 2^{-\log(N)^{1/9}}$ (KM '23)	→	$\delta = 2^{-\log(N)^{1/12}}$ (KM '23)

The "Analytic" approach



- V = structured set.
- $A' = A \cap V$ has the "right" number of solutions to $x + y = 2z$

A' is "pseudorandom".

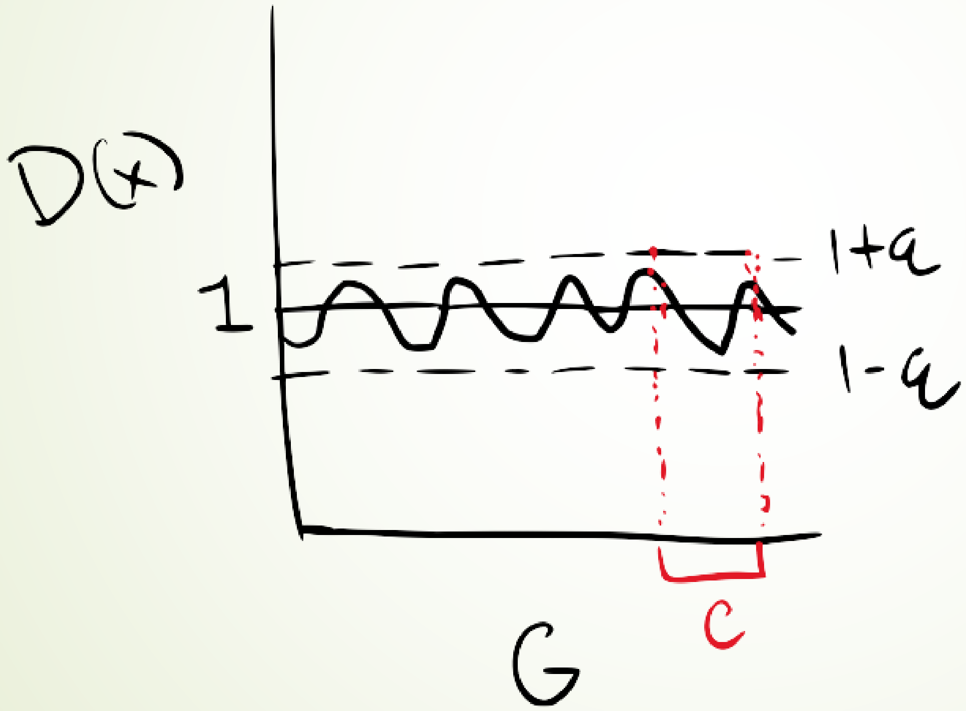
Notion of Pseudorandomness ($A \subseteq G$)

- Draw $a, a' \sim A$ (uniformly) at random
- ★ Say that A is pseudorandom if:

$a + a'$ is near-uniform over G.

Notion of Pseudorandomness $(\mathrm{a}, \mathrm{a}' \sim A)$

- Let $D(x) = \mathrm{PDF}(a + a')$



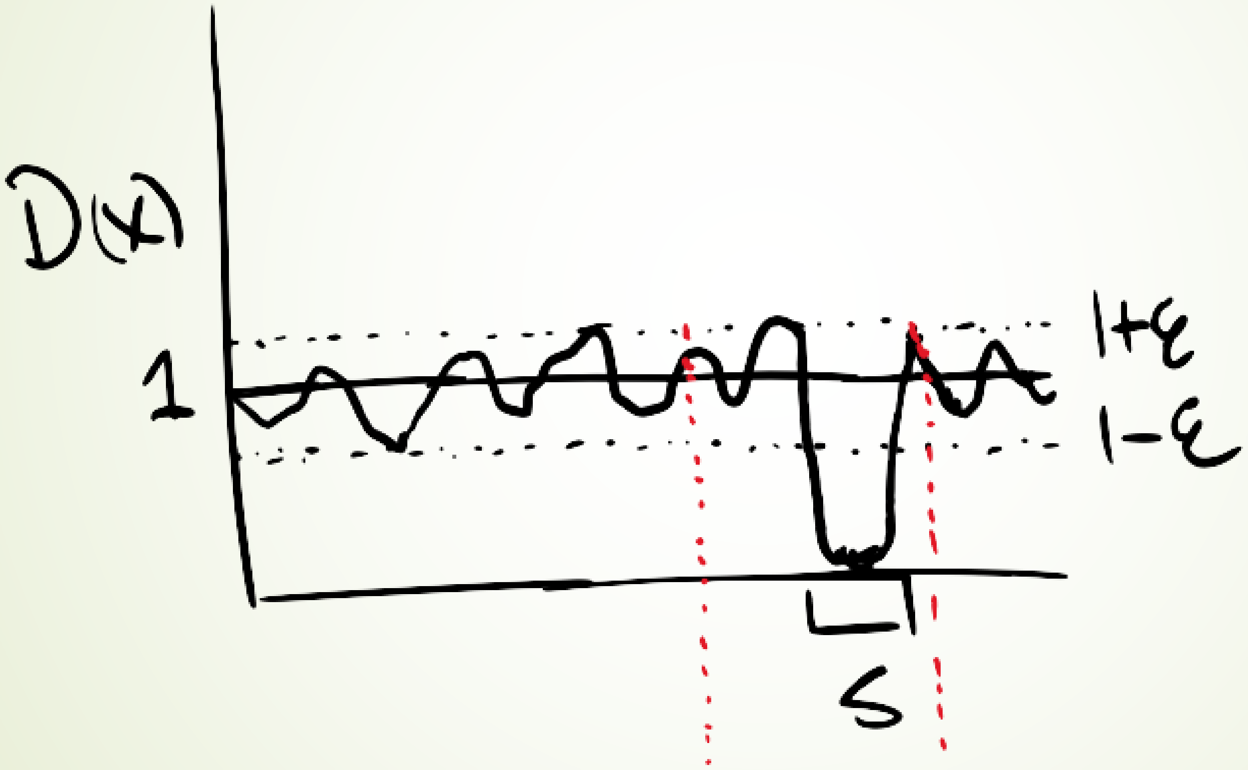
$\Rightarrow$ for any $C \subseteq G$,

$$\#\mathrm{sol}(a + a' = c) = (1 \pm \epsilon)\frac{|A|^2|C|}{|G|}$$

$\Rightarrow$(e.g. $C = \{2z \mid z \in A\}$)

Definition of "near-uniform"

$|A| = 2^{-d}|G|,$
$D(x) = PDF(a + a')$
$C = \{2z \mid z \in A\}$



If $p := d + 1$, then $\#\mathrm{sol}(a + a' = c) \geq \frac{1}{4}\frac{|A|^3}{|G|}$

$(\|D - 1\|_p \leq \epsilon \;\cong) \quad \frac{|S|}{|G|} \leq 2^{-p}$

Notation For (Min)-Entropy Deficit

- Write

$$\Delta(A) = d$$

$$iff$$

$$|A| = 2^{-d}|G|$$

Main Lemma (for general G)

- Let $A \subseteq G$, $\Delta(A) \leq d$.
- Either

  I. $PDF(a + a')$ is near-uniform, or

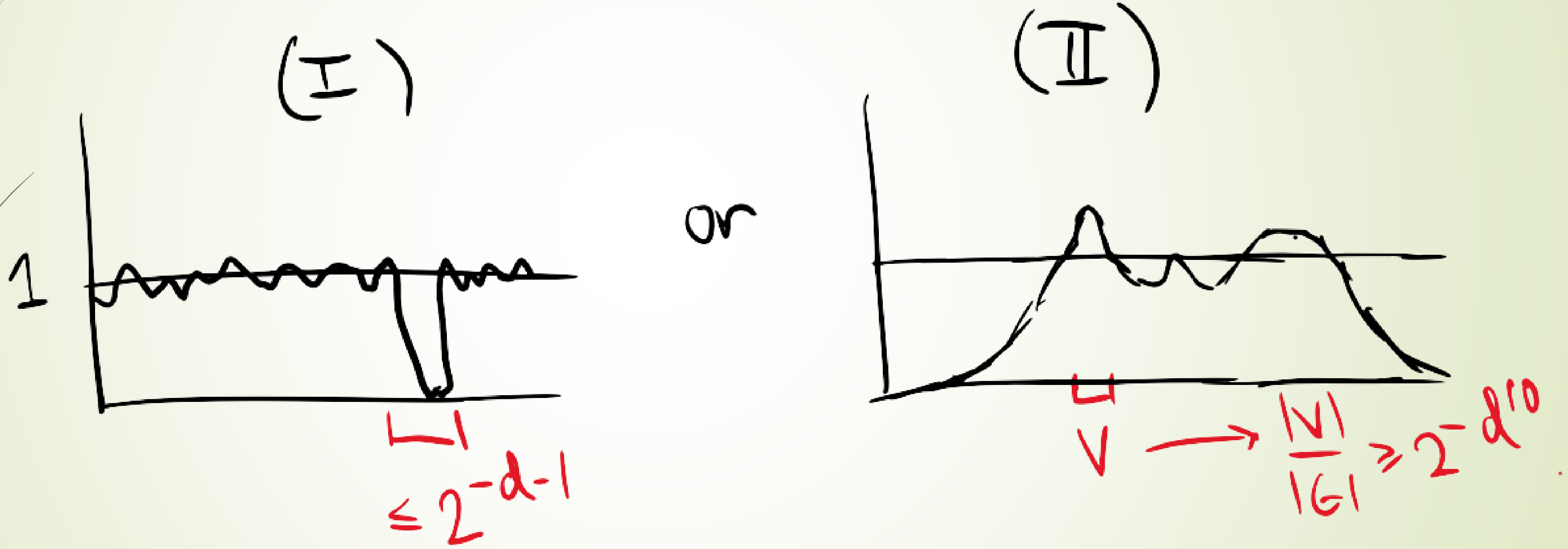
  II. $\frac{|A \cap V|}{|V|} \geq (1+\epsilon)\frac{|A|}{|G|}$,

  for some approximate subgroup V,

  $\Delta(V) \leq poly(d,p)$.

Main Lemma (visualized)

$D(x) = \mathrm{PDF}(a + a')$



Plan for (II): Zoom in on $A' = A \cap V$ until it looks like (I)

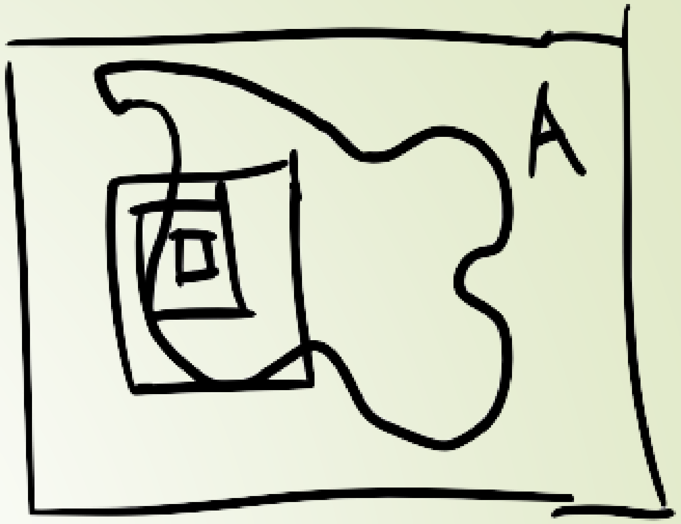
Main Lemma (for $G = \mathbb{F}_q^n$)

- Let $A \subseteq \mathbb{F}_q^n$, $\Delta(A) \leq d$.
- Either
  I. $PDF(a + a')$ is near-uniform, or
  II. $\frac{|A \cap V|}{|V|} \geq (1+\epsilon)\frac{|A|}{|\mathbb{F}_q^n|}$,

  for some affine subspace V,

  $\text{Codim}(V) \leq O(d^4 p^4)$.

Density Increments



- Initialize $A_0 = A$, $V_0 = \mathbb{F}_q^n$.
- If A_i is not pseudorandom, pass to some

$$A_{i+1} \coloneqq A_i \cap V_{i+1}, \qquad \frac{|A_{i+1}|}{|V_{i+1}|} \geq (1+\epsilon)\frac{|A_i|}{|V_i|}.$$

- If $\frac{|A_t|}{|V_t|} \geq (1+\epsilon)^t \frac{|A|}{|\mathbb{F}_q^n|} \geq 2^{\epsilon t - d}$, then $t \leq d/\epsilon$, and $\Delta(A_t) \leq O(td^8) = O(d^9)$.

Proof of Main Lemma: Setup

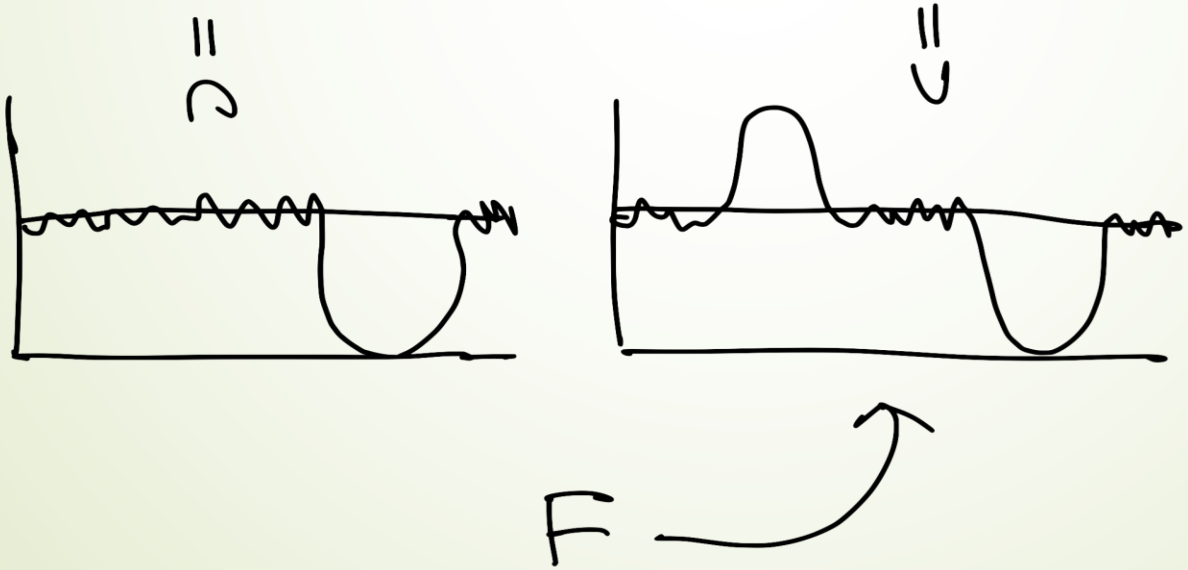
- Let $D(x) = \mathrm{PDF}(a + a')$.
- Assume D is not near-uniform: $\|D - 1\|_p \geq \epsilon$.
- We want to find a large V, $\mathbb{E}_V[\mathbf{1}_A] \geq (1+\epsilon)\mathbb{E}_{\mathbb{F}_q^n}[\mathbf{1}_A]$.
- Actually, we will find a “density increment”

$$\mathbb{E}_V[D] \geq 1 + \epsilon$$

Main Idea #1: Spectral Positivity

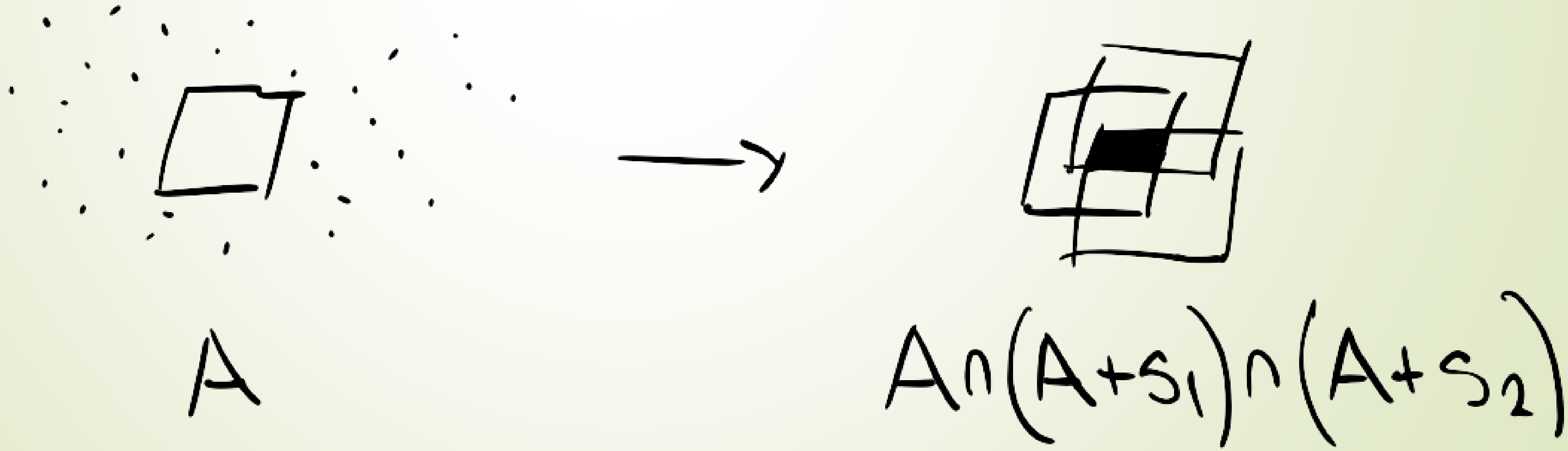
- Let $D = \mathrm{PDF}(a + a')$, $F = \mathrm{PDF}(a - a')$.
- ★ $\|D - 1\|_p \leq \|F - 1\|_p$.
- ★ $\|(F - 1)_-\|_p \leq \|(F - 1)_+\|_{\mathrm{p}}$.

because $F(x - y) \succcurlyeq 0$.



Main Idea #2: Sifting

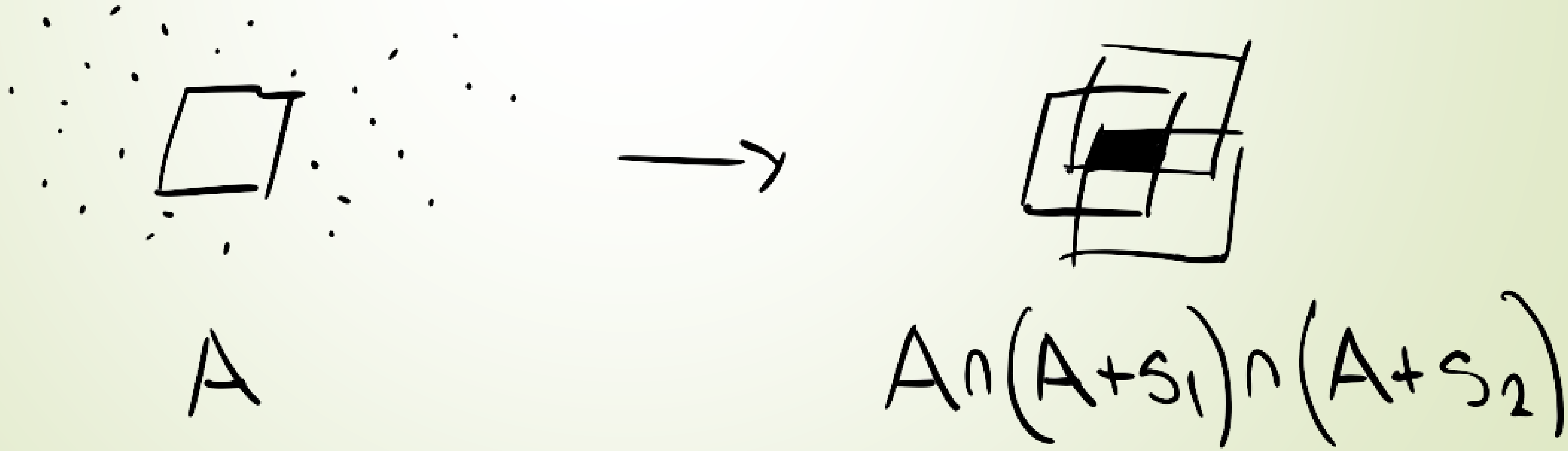
- Hard case: A is mostly pseudorandom, but with a "planted" (strong but rare) structured part.
- Suppose $A = V \cup R$, for some subspace V and a random set R. How to find V?



Main Idea #2: Sifting

- Let $F(x) = \mathrm{PDF}(a - a')$ and assume $\|F\|_p \geq 1 + \epsilon$.
- We use sifting to find a set $B = \cap_{\{i=1\}}^{p} (A + s_i)$,
  - of size roughly $|B| \geq 2^{-dp}|A|$,
  - witnessing

$$\mathbb{E}_{\mathrm{b,b'} \in B}[\mathrm{F(b - b')}] \geq 1 + \epsilon/2$$



Rough Proof Outline

$F(x) = \mathrm{PDF}(a - a')$

Spectral positivity (from line 1 to line 2)

Sifting (from line 2 to line 3)

$$\mathbb{E}_{\mathrm{C}}[|\mathrm{F} - 1|] \geq \epsilon, \qquad \Delta(C) \leq p$$

$$\Rightarrow \mathbb{E}_{\mathrm{S}}[\mathrm{F}] \geq 1 + \epsilon/2, \qquad \Delta(S) \leq O(p)$$

$$\Rightarrow \mathbb{E}_{\mathrm{b,b'} \in B}[\mathrm{F}(\mathrm{b} - \mathrm{b'})] \geq 1 + \epsilon/4, \qquad \Delta(B) \leq O(pd)$$

$$\Rightarrow \mathbb{E}_{\mathrm{V}}[\mathrm{F}] \geq 1 + \epsilon/8, \qquad \Delta(V) \leq O(p^4 d^4).$$

Thanks for listening !