Additive combinatorics: Probabilistic perspective

Huy Tuan Pham

California Institute of Technology

NZMRI - January 2026

Recap

Definition (Cayley graph)

For a group G and symmetric subset $S \subset G$, the *Cayley graph* G_S has vertex set G and distinct x, y are adjacent if $xy^{-1} \in S$.

Given $p \in (0,1)$, the random Cayley graphs $G(p)$ is the Cayley graph G_S where each $\{g, g^{-1}\}$ is included independently in S with probability p.

Recap

Definition (Cayley graph)

For a group G and symmetric subset $S \subset G$, the *Cayley graph* G_S has vertex set G and distinct x, y are adjacent if $xy^{-1} \in S$.

Given $p \in (0,1)$, the random Cayley graphs $G(p)$ is the Cayley graph G_S where each $\{g, g^{-1}\}$ is included independently in S with probability p.

Theorem (Conlon-Fox-P.-Yepremyan '24)

With high probability, the independence and clique number of a uniform random Cayley graph on any group G of order N is $O(\log N \log\log N)$.

First moment of large independent sets in random Cayley graphs $\leftrightarrow$ Counting sets with small doubling $|AA^{-1}|/|A|$.

Recap

Definition (Cayley graph)

For a group G and symmetric subset $S \subset G$, the *Cayley graph* G_S has vertex set G and distinct x, y are adjacent if $xy^{-1} \in S$.

Given $p \in (0,1)$, the random Cayley graphs $G(p)$ is the Cayley graph G_S where each $\{g, g^{-1}\}$ is included independently in S with probability p.

Theorem (Conlon-Fox-P.-Yepremyan '24)

With high probability, the independence and clique number of a uniform random Cayley graph on any group G of order N is $O(\log N \log\log N)$.

First moment of large independent sets in random Cayley graphs $\leftrightarrow$ Counting sets with small doubling $|AA^{-1}|/|A|$.

Theorem (Conlon-Fox-P.-Yepremyan '24)

In any group G of order N, the number of subsets $A \subset G$ with $|A| = n$ and $|AA^{-1}| \leq Kn$ is at most $N^{C(K+\log n)}(CK)^n$.

Recap

Theorem (Conlon-Fox-P.-Yepremyan '24)

In any group G of order N, the number of subsets $A \subset G$ with $|A| = n$ and $|AA^{-1}| \leq Kn$ is at most $N^{C(K+\log n)}(CK)^n$.

Proof via the main combinatorial lemma.

Recap

Theorem (Conlon-Fox-P.-Yepremyan '24)

In any group G of order N, the number of subsets $A \subset G$ with $|A| = n$ and $|AA^{-1}| \leq Kn$ is at most $N^{C(K+\log n)}(CK)^n$.

Proof via the main combinatorial lemma.

Theorem (Fox-P. '25+)

If G is an abelian group with exponent r and $A \subseteq G$ is so that $|A+A| \leq K|A|$, then A is contained in subgroup H of size $|H| \leq r^{(2+o(1))K}|A|$.

In particular, Ruzsa's conjecture holds.

Recap

Theorem (Conlon-Fox-P.-Yepremyan '24)

In any group G of order N, the number of subsets $A \subset G$ with $|A| = n$ and $|AA^{-1}| \leq Kn$ is at most $N^{C(K+\log n)}(CK)^n$.

Proof via the main combinatorial lemma.

Theorem (Fox-P. '25+)

If G is an abelian group with exponent r and $A \subseteq G$ is so that $|A+A| \leq K|A|$, then A is contained in subgroup H of size $|H| \leq r^{(2+o(1))K}|A|$.

In particular, Ruzsa's conjecture holds.

Further applications:

- Dimension of sets with small doubling
- Robust Freiman-Ruzsa lemma
- Random sumset extractors

Perspective

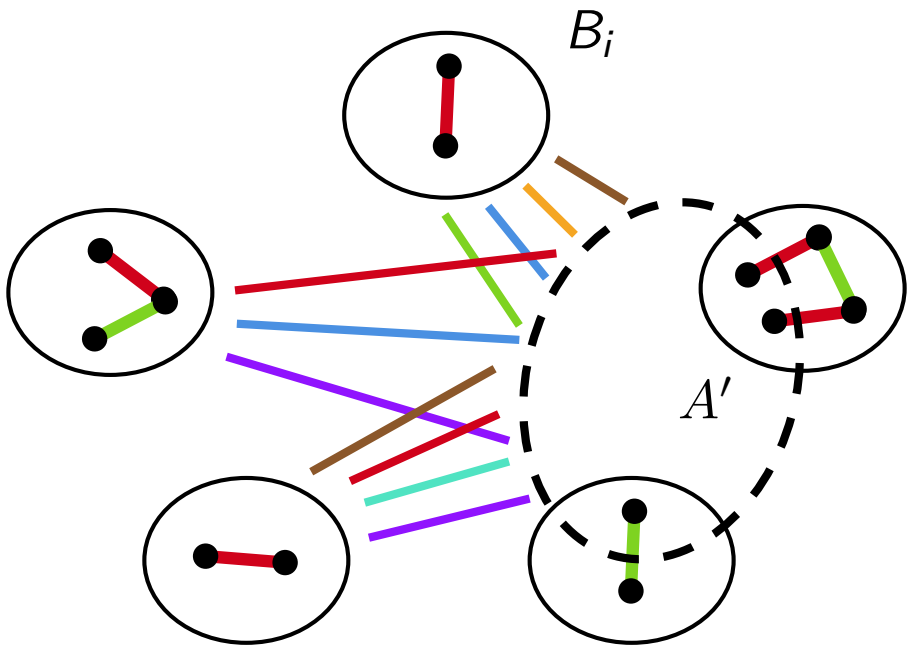
Combinatorial perspective: Main combinatorial lemma identifies novel combinatorial structures underlying sets with small doubling.

Perspective

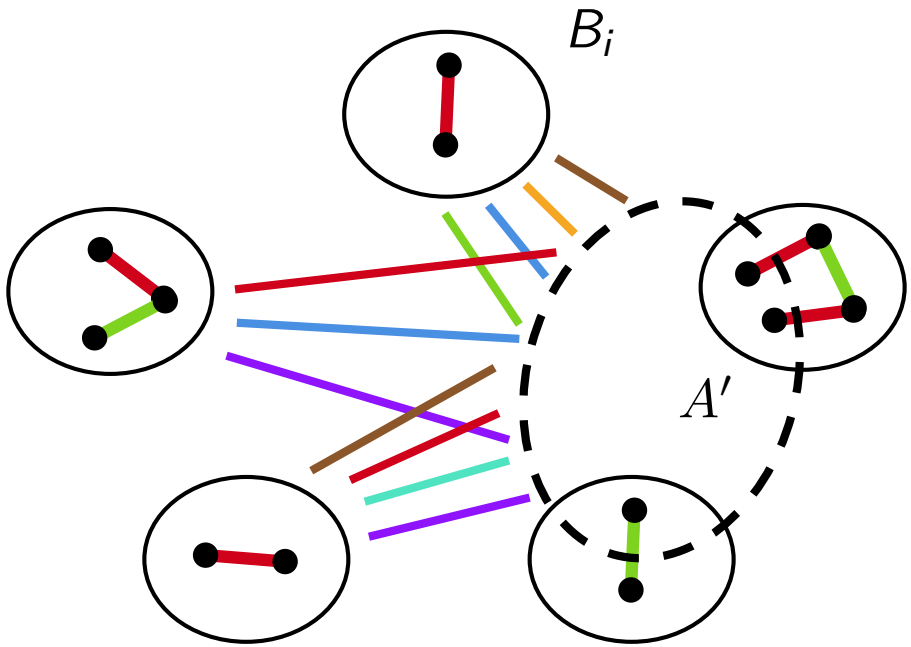
Combinatorial perspective: Main combinatorial lemma identifies novel combinatorial structures underlying sets with small doubling.

New combinatorial approach to additive combinatorics:

- Main combinatorial lemma: expanding structures in sets with small doubling.
- Refinement: Use expanding structures to probe information about the entire set.

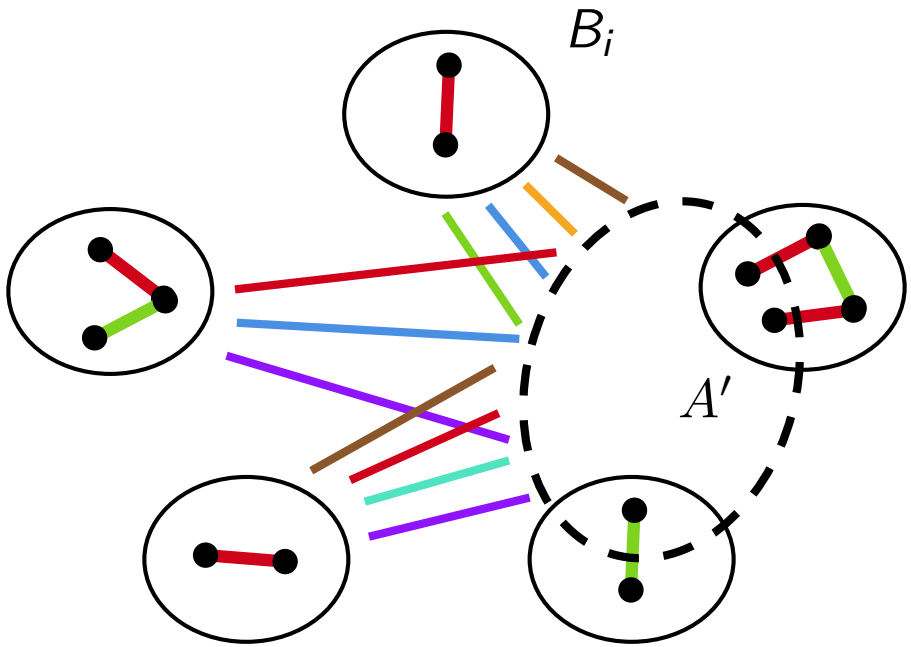


Perspective



Perspective: Instead of zooming in on explicit structures inside the set, the combinatorial lemma provides first a template that is "as random-like as possible".

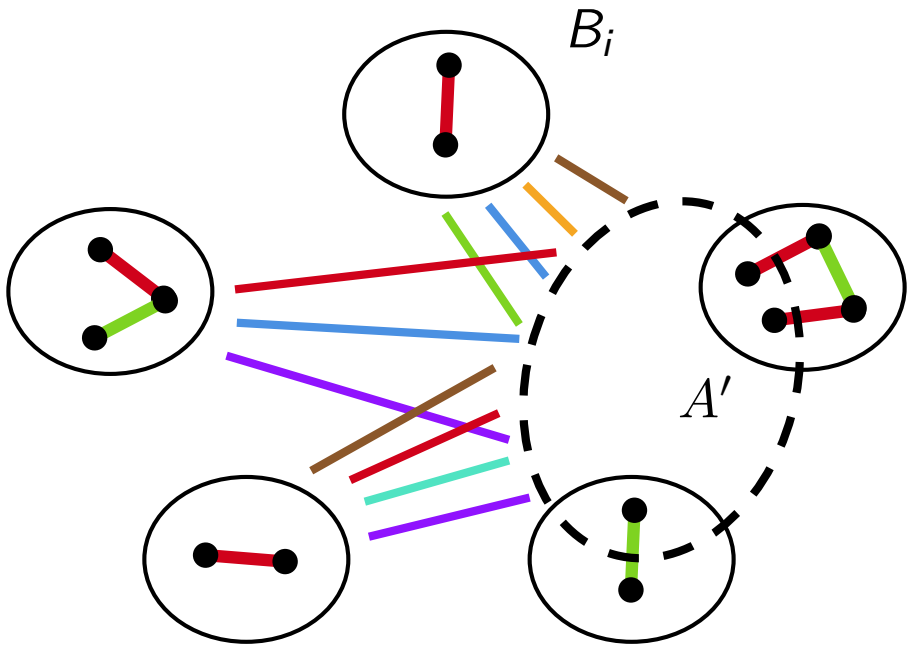
Perspective



Perspective: Instead of zooming in on explicit structures inside the set, the combinatorial lemma provides first a template that is "as random-like as possible".

- The proof follows from an intricate random exploration process.

Perspective



Perspective: Instead of zooming in on explicit structures inside the set, the combinatorial lemma provides first a template that is "as random-like as possible".

- The proof follows from an intricate random exploration process.

Robust and applicable in wide generality.

Alon's conjecture - Going beyond uniform random

Alon's conjecture

Conjecture (Alon '89)

There is a constant C such that every finite group has a Cayley graph which is C-Ramsey.

Alon's conjecture - Going beyond uniform random

Over $\mathbb{F}_p^d$, a uniformly random Cayley graph is not Ramsey w.h.p.

Alon's conjecture - Going beyond uniform random

Over $\mathbb{F}_p^d$, a uniformly random Cayley graph is not Ramsey w.h.p.

We define an alternative distribution of random Cayley graphs to "remove" the problematic cliques.

Theorem (Conlon-Fox-P.-Yepremyan '24)

For $p \geq 5$, there exists Cayley graphs over $\mathbb{F}_p^d$ with clique and independence number $(2 + o(1)) \log_2 N$ where $N = p^d$.
For $p = 1 \pmod 4$, these Cayley graphs are self-complementary.

Answer a question of Alon and Orlitsky ('95) motivated by zero-error capacity and dual-source coding.

Alon's conjecture - Going beyond uniform random

Over $\mathbb{F}_p^d$, a uniformly random Cayley graph is not Ramsey w.h.p.

We define an alternative distribution of random Cayley graphs to "remove" the problematic cliques.

Theorem (Conlon-Fox-P.-Yepremyan '24)

For $p \geq 5$, there exists Cayley graphs over $\mathbb{F}_p^d$ with clique and independence number $(2 + o(1)) \log_2 N$ where $N = p^d$.
For $p = 1 \pmod 4$, these Cayley graphs are self-complementary.

Answer a question of Alon and Orlitsky ('95) motivated by zero-error capacity and dual-source coding.

Theorem (Conlon-Fox-P.-Yepremyan '24)

For almost all N, all abelian groups G of order N have a Cayley graph which is C-Ramsey.

Alon's conjecture - Going beyond uniform random

Over $\mathbb{F}_p^d$, a uniformly random Cayley graph is not Ramsey w.h.p.

We define an alternative distribution of random Cayley graphs to "remove" the problematic cliques.

Theorem (Conlon-Fox-P.-Yepremyan '24)

For $p \geq 5$, there exists Cayley graphs over $\mathbb{F}_p^d$ with clique and independence number $(2 + o(1)) \log_2 N$ where $N = p^d$.
For $p = 1 \pmod 4$, these Cayley graphs are self-complementary.

Answer a question of Alon and Orlitsky ('95) motivated by zero-error capacity and dual-source coding.

Theorem (Conlon-Fox-P.-Yepremyan '24)

For almost all N, all abelian groups G of order N have a Cayley graph which is C-Ramsey.

Recent generalization to all groups of order coprime to 6 by Schildkraut.

Dependent random graphs

New direction: Dependent random graphs

The combinatorial approach to clique number of dense random Cayley graphs suggests intriguing universal behaviors in general random graph models with significant dependencies.

New direction: Dependent random graphs

The combinatorial approach to clique number of dense random Cayley graphs suggests intriguing universal behaviors in general random graph models with significant dependencies.

Consider an edge-coloring c of a complete graph, where each color class has degree at most Δ.

New direction: Dependent random graphs

The combinatorial approach to clique number of dense random Cayley graphs suggests intriguing universal behaviors in general random graph models with significant dependencies.

Consider an edge-coloring c of a complete graph, where each color class has degree at most Δ.

Definition (Random entangled graph)

An *entangled graph* is the edge-union of some of the color classes.

New direction: Dependent random graphs

The combinatorial approach to clique number of dense random Cayley graphs suggests intriguing universal behaviors in general random graph models with significant dependencies.

Consider an edge-coloring c of a complete graph, where each color class has degree at most Δ.

Definition (Random entangled graph)

An *entangled graph* is the edge-union of some of the color classes.

The *random entangled graph* $G_c(p)$ is formed by including each color class with probability p independently.

New direction: Dependent random graphs

The combinatorial approach to clique number of dense random Cayley graphs suggests intriguing universal behaviors in general random graph models with significant dependencies.

Consider an edge-coloring c of a complete graph, where each color class has degree at most Δ.

Definition (Random entangled graph)

An *entangled graph* is the edge-union of some of the color classes.

The *random entangled graph* $G_c(p)$ is formed by including each color class with probability p independently.

Examples:

- Erdős-Rényi random graphs.
- Random Cayley graphs.
- Random Latin square graphs: Color class $C_k = \{\{i,j\} : L_{ij} = k\}$ for a Latin square L.

New direction: Random entangled graphs

Theorem 1 (Conlon-Fox-P.-Yepremyan '24)

In a Δ-bounded edge-coloring of the complete graph on N vertices, the number of n-vertex subsets with at most Kn colors is at most

$$N^{C\Delta(K+\log n)}(C\Delta K)^n.$$

New direction: Random entangled graphs

Theorem 1 (Conlon-Fox-P.-Yepremyan '24)

In a Δ-bounded edge-coloring of the complete graph on N vertices, the number of n-vertex subsets with at most Kn colors is at most

$$N^{C\Delta(K+\log n)}(C\Delta K)^n.$$

Theorem 2 (Conlon-Fox-P.-Yepremyan '24)

If an edge-coloring c of K_N is Δ-bounded, then with high probability,

$$\alpha(G_c(p)) = O_{p,\Delta}(\log N \log\log N).$$

New direction: Random entangled graphs

Theorem 1 (Conlon-Fox-P.-Yepremyan '24)

In a Δ-bounded edge-coloring of the complete graph on N vertices, the number of n-vertex subsets with at most Kn colors is at most

$$N^{C\Delta(K+\log n)}(C\Delta K)^n.$$

Theorem 2 (Conlon-Fox-P.-Yepremyan '24)

If an edge-coloring c of K_N is Δ-bounded, then with high probability,

$$\alpha(G_c(p)) = O_{p,\Delta}(\log N \log\log N).$$

- From Theorem 1, a careful union bound yields Theorem 2.
- Theorem 2 solves a conjecture of Christofides and Markström ('11) on the independence number of random Latin square graphs.

New direction: Δ-independent random graphs

New direction: Δ-independent random graphs

How much independence is needed?

New direction: Δ-independent random graphs

How much independence is needed?

The phenomena extend to a significantly broader class of random graphs.

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

New direction: Δ-independent random graphs

How much independence is needed?

The phenomena extend to a significantly broader class of random graphs.

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Significant weakening of usual condition in the Lovász Local Lemma!

New direction: Δ-independent random graphs

How much independence is needed?

The phenomena extend to a significantly broader class of random graphs.

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Significant weakening of usual condition in the Lovász Local Lemma!

All random entangled graphs defined by a Δ-bounded edge coloring are Δ-independent random graphs.

New direction: Δ-independent random graphs

How much independence is needed?

The phenomena extend to a significantly broader class of random graphs.

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Significant weakening of usual condition in the Lovász Local Lemma!

All random entangled graphs defined by a Δ-bounded edge coloring are Δ-independent random graphs.

Theorem (Conlon-Fox-P.-Yepremyan '26+)

Consider a Δ-independent random graph G where the probability of appearance of each edge is $\Theta(p)$. Then, with high probability,

$$\alpha(G) = O_{p,\Delta}(\log N \log\log N).$$

New direction: Δ-independent random graphs

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

New direction: Δ-independent random graphs

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Theorem (Conlon-Fox-P.-Yepremyan '26+)

Consider a symmetric Δ-independent random graph G where the probability of appearance of each edge is p. Then, with high probability, all nontrivial eigenvalues of G are bounded by $O(\sqrt{pN \log N})$.

New direction: Δ-independent random graphs

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Theorem (Conlon-Fox-P.-Yepremyan '26+)

Consider a symmetric Δ-independent random graph G where the probability of appearance of each edge is p. Then, with high probability, all nontrivial eigenvalues of G are bounded by $O(\sqrt{pN \log N})$.

As a corollary, we obtain that G is Hamiltonian with high probability for $p \gg \log N$.

New direction: Δ-independent random graphs

Definition (Δ-independent graph)

An ensemble of random graphs is said to be Δ-independent if for each edge e, there is a graph G_e of maximum degree Δ such that the appearance of e is independent of all edges outside G_e.

Theorem (Conlon-Fox-P.-Yepremyan '26+)

Consider a symmetric Δ-independent random graph G where the probability of appearance of each edge is p. Then, with high probability, all nontrivial eigenvalues of G are bounded by $O(\sqrt{pN \log N})$.

As a corollary, we obtain that G is Hamiltonian with high probability for $p \gg \log N$.

Open direction

Study interesting properties of Δ-independent graphs (random entangled graphs, random Cayley graphs).

Sparse random Cayley graphs & First moment obstructions

Sparse random Cayley graphs

Definition (Cayley sum graph)

For an abelian group G and subset $S \subset G$, the *Cayley sum graph* G_S has vertex set G and distinct x, y are adjacent if $x + y \in S$.

Sparse random Cayley graphs

Definition (Cayley sum graph)

For an abelian group G and subset $S \subset G$, the *Cayley sum graph* G_S has vertex set G and distinct x, y are adjacent if $x + y \in S$.

Given $p \in (0, 1)$, the random Cayley sum graph $G(p)$ is the Cayley sum graph G_S where each x is included independently in S with probability p.

Sparse random Cayley graphs

Definition (Cayley sum graph)

For an abelian group G and subset $S \subset G$, the *Cayley sum graph* G_S has vertex set G and distinct x, y are adjacent if $x + y \in S$.

Given $p \in (0, 1)$, the random Cayley sum graph $G(p)$ is the Cayley sum graph G_S where each x is included independently in S with probability p.

Conjecture (Alon '07, '13)

The independence number of $G(p)$ is with high probability $\tilde{O}(p^{-1})$.

The random Cayley sum graph $G(p)$ behave like a random pN-regular graph.

First moment obstructions and Additive combinatorics

Naive first moment for independent sets in Cayley sum graphs:

$$\mathbb{E}[\#\text{ independent sets of size } t \text{ in } G(p)] = \sum_{|A|=t} (1-p)^{|A+A|} \approx \sum_{|A|=t} \exp(-p|A+A|).$$

First moment obstructions and Additive combinatorics

Naive first moment for independent sets in Cayley sum graphs:

$$\mathbb{E}[\#\text{ independent sets of size } t \text{ in } G(p)] = \sum_{|A|=t} (1-p)^{|A+A|} \approx \sum_{|A|=t} \exp(-p|A+A|).$$

Large for any $t = o(|G|)$ and $p = o(1)$: Exponentially many $|A| = t$ with $|A + A| = O(t)$.

First moment obstructions and Additive combinatorics

Naive first moment for independent sets in Cayley sum graphs:

$$\mathbb{E}[\#\text{ independent sets of size } t \text{ in } G(p)] = \sum_{|A|=t} (1-p)^{|A+A|} \approx \sum_{|A|=t} \exp(-p|A+A|).$$

Large for any $t = o(|G|)$ and $p = o(1)$: Exponentially many $|A| = t$ with $|A + A| = O(t)$.

Need better First moment obstructions.

First moment obstructions

General setup:

- Finite set X.
- Collection of target sets $\mathcal{H} \subseteq 2^X$.
- Random subset X_p: each element of X included independently with probability p.

First moment obstructions

General setup:

- Finite set X.
- Collection of target sets $\mathcal{H} \subseteq 2^X$.
- Random subset X_p: each element of X included independently with probability p.

Denote $\langle \mathcal{H} \rangle$ the collection of subsets $W \in 2^X$ containing at least one target set $H \in \mathcal{H}$:

Question

Does X_p contain a target set $H \in \mathcal{H}$?

What are obstructions (certificates) which imply that $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle)$ is small?

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.
  - In Erdős' case, H is taken to be a clique on $n = 2\log_2 N$ vertices.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.
  - In Erdős' case, H is taken to be a clique on $n = 2\log_2 N$ vertices.

Naive obstruction:

$$\sum_{H\in\mathcal{H}} p^{|H|} < 1/2 \qquad \Rightarrow \qquad \mathbb{P}(X_p \in \langle\mathcal{H}\rangle) < 1/2.$$

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.
  - In Erdős' case, H is taken to be a clique on $n = 2\log_2 N$ vertices.

Naive obstruction:

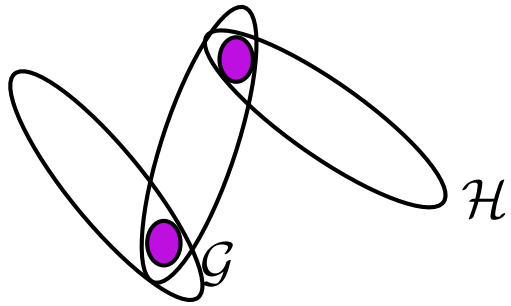
$$\sum_{H\in\mathcal{H}} p^{|H|} < 1/2 \qquad \Rightarrow \qquad \mathbb{P}(X_p \in \langle\mathcal{H}\rangle) < 1/2.$$

Example: Erdős' proof that random graphs $G(N,1/2)$ do not have cliques of size $2\log_2 N$.

First moment obstructions

Definition

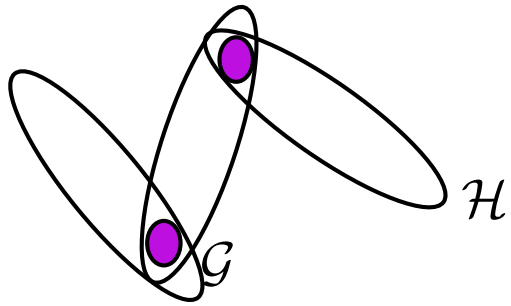
We say that $\mathcal{G}$ is a cover for $\mathcal{H}$ if every $H \in \mathcal{H}$ contains some $G \in \mathcal{G}$.



First moment obstructions

Definition

We say that $\mathcal{G}$ is a cover for $\mathcal{H}$ if every $H \in \mathcal{H}$ contains some $G \in \mathcal{G}$.



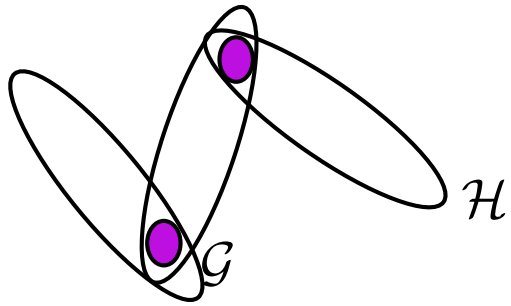
First moment obstruction:

$$\sum_{G \in \mathcal{G}} p^{|G|} < 1/2 \quad \Rightarrow \quad \mathbb{P}(X_p \in \langle \mathcal{H} \rangle) \leq \mathbb{P}(X_p \in \langle \mathcal{G} \rangle) < 1/2.$$

First moment obstructions

Definition

We say that $\mathcal{G}$ is a cover for $\mathcal{H}$ if every $H \in \mathcal{H}$ contains some $G \in \mathcal{G}$.



First moment obstruction:

$$\sum_{G \in \mathcal{G}} p^{|G|} < 1/2 \quad \Rightarrow \quad \mathbb{P}(X_p \in \langle \mathcal{H} \rangle) \leq \mathbb{P}(X_p \in \langle \mathcal{G} \rangle) < 1/2.$$

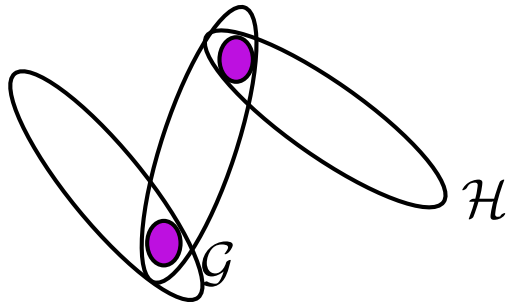
Definition

We call a cover $\mathcal{G}$ with $\sum_{G \in \mathcal{G}} p^{|G|} < 1/2$ a first moment obstruction for $\mathcal{H}$.

First moment obstructions

Definition

We say that $\mathcal{G}$ is a cover for $\mathcal{H}$ if every $H \in \mathcal{H}$ contains some $G \in \mathcal{G}$.



First moment obstruction:

$$\sum_{G \in \mathcal{G}} p^{|G|} < 1/2 \quad \Rightarrow \quad \mathbb{P}(X_p \in \langle \mathcal{H} \rangle) \leq \mathbb{P}(X_p \in \langle \mathcal{G} \rangle) < 1/2.$$

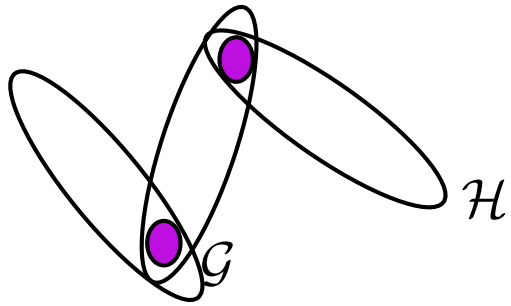
Definition

We call a cover $\mathcal{G}$ with $\sum_{G \in \mathcal{G}} p^{|G|} < 1/2$ a first moment obstruction for $\mathcal{H}$. We say that $\mathcal{H}$ is p-small if a first moment obstruction exists.

First moment obstructions

Definition

We say that $\mathcal{G}$ is a cover for $\mathcal{H}$ if every $H \in \mathcal{H}$ contains some $G \in \mathcal{G}$.



First moment obstruction:

$$\sum_{G \in \mathcal{G}} p^{|G|} < 1/2 \quad \Rightarrow \quad \mathbb{P}(X_p \in \langle \mathcal{H} \rangle) \leq \mathbb{P}(X_p \in \langle \mathcal{G} \rangle) < 1/2.$$

Definition

We call a cover $\mathcal{G}$ with $\sum_{G \in \mathcal{G}} p^{|G|} < 1/2$ a first moment obstruction for $\mathcal{H}$. We say that $\mathcal{H}$ is p-small if a first moment obstruction exists.

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

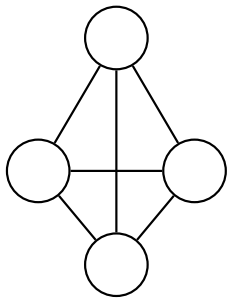


Figure 1: $H = K_4$

Example: $\mathcal{H} = \{\text{copies of } K_4\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-2/3})$-small.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

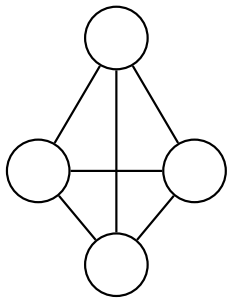


Figure 1: $H = K_4$

Example: $\mathcal{H} = \{\text{copies of } K_4\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-2/3})$-small.
- For $p < cn^{-2/3}$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } K_4) < 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

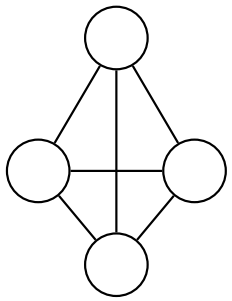


Figure 1: $H = K_4$

Example: $\mathcal{H} = \{\text{copies of } K_4\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-2/3})$-small.
- For $p < cn^{-2/3}$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } K_4) < 1/2$.
- $\mathcal{H}$ is not $Cn^{-2/3}$-small, and $\mathbb{P}(G(N, Cn^{-2/3}) \text{ contains } K_4) > 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

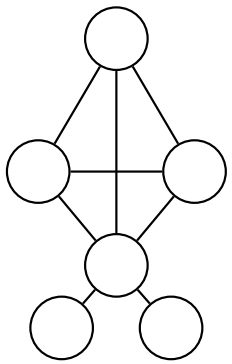


Figure 2: $H = K_4$ with 2 pendant edges

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-3/4})$-small.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

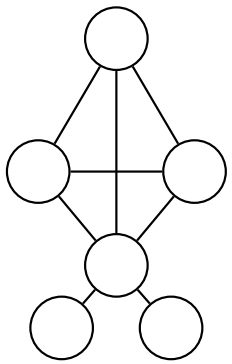


Figure 2: $H = K_4$ with 2 pendant edges

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-3/4})$-small.
- Better covers: $\mathcal{G}$ consisting of copies of K_4: $\mathcal{H}$ is $(cn^{-2/3})$-small.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

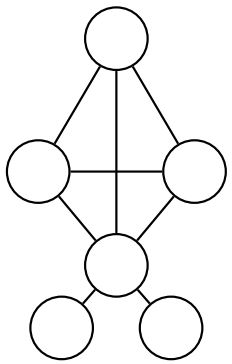


Figure 2: $H = K_4$ with 2 pendant edges

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-3/4})$-small.
- Better covers: $\mathcal{G}$ consisting of copies of K_4: $\mathcal{H}$ is $(cn^{-2/3})$-small.
- For $p < cn^{-2/3}$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } H) < 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

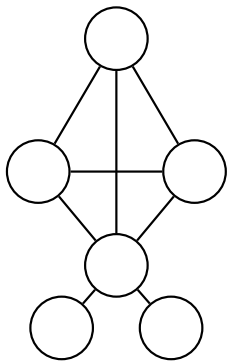


Figure 2: $H = K_4$ with 2 pendant edges

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is $(cn^{-3/4})$-small.
- Better covers: $\mathcal{G}$ consisting of copies of K_4: $\mathcal{H}$ is $(cn^{-2/3})$-small.
- For $p < cn^{-2/3}$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } H) < 1/2$.
- $\mathcal{H}$ is not $Cn^{-2/3}$-small, and $\mathbb{P}(G(N, Cn^{-2/3}) \text{ contains } H) > 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N, p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

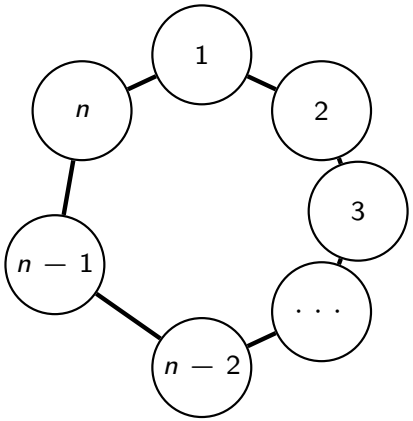


Figure 3:
$H =$ Hamiltonian cycle

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is (c/n)-small.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

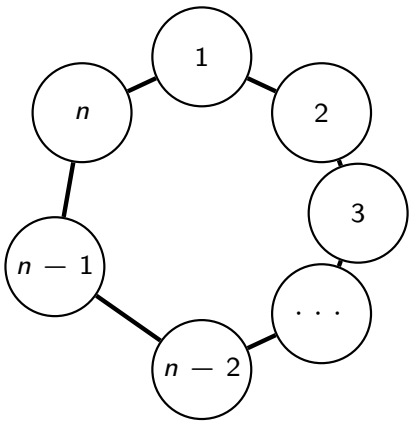


Figure 3:
$H =$ Hamiltonian cycle

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is (c/n)-small.
- For $p < c/n$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } H) < 1/2$.

First moment obstructions

Example:

- $X = \binom{[N]}{2}$, then $X_p \sim G(N,p)$.
- $\mathcal{H}$ is the collection of isomorphic copies of a graph H on N vertices.

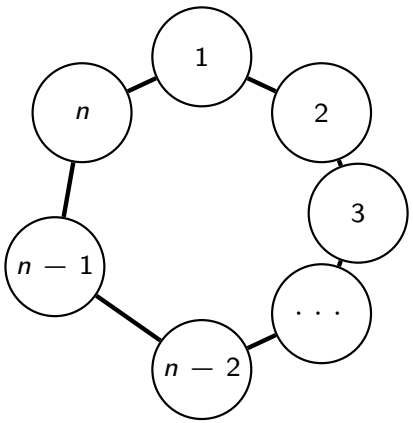


Figure 3:
$H =$ Hamiltonian cycle

Example: $\mathcal{H} = \{\text{copies of } H\}$.

- "Naive" cover: $\mathcal{H}$ is a cover for itself, so $\mathcal{H}$ is (c/n)-small.
- For $p < c/n$, there is a first moment obstruction showing that $\mathbb{P}(G(N,p) \text{ contains } H) < 1/2$.
- Other covers: all edges adjacent to a fixed vertex.
- $\mathcal{H}$ is not C/n-small, and $\mathbb{P}(G(N, C(\log n)/n) \text{ contains } H) > 1/2$.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

New philosophy:

Inexistence of first moment obstructions implies appearance of structure.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

New philosophy:

Inexistence of first moment obstructions implies appearance of structure.

Conjecture (Kahn-Kalai conjecture '06)

If $\mathcal{H}$ is not p-small, then $\mathbb{P}(X_{Cp \log |X|} \in \langle \mathcal{H} \rangle) \geq 1/2$.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

New philosophy:

Inexistence of first moment obstructions implies appearance of structure.

Conjecture (Kahn-Kalai conjecture '06)

If $\mathcal{H}$ is not p-small, then $\mathbb{P}(X_{Cp \log |X|} \in \langle \mathcal{H} \rangle) \geq 1/2$.

Theorem (Park-P. '24, Kahn-Kalai conjecture)

If $\mathcal{H}$ be not p-small with $\ell = \max_{H \in \mathcal{H}} |H|$, then $\mathbb{P}(X_{Cp \log \ell} \in \langle \mathcal{H} \rangle) \geq 1/2$.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

New philosophy:

Inexistence of first moment obstructions implies appearance of structure.

Conjecture (Kahn-Kalai conjecture '06)

If $\mathcal{H}$ is not p-small, then $\mathbb{P}(X_{Cp \log |X|} \in \langle \mathcal{H} \rangle) \geq 1/2$.

Theorem (Park-P. '24, Kahn-Kalai conjecture)

If $\mathcal{H}$ be not p-small with $\ell = \max_{H \in \mathcal{H}} |H|$, then $\mathbb{P}(X_{Cp \log \ell} \in \langle \mathcal{H} \rangle) \geq 1/2$.

Determining the threshold where target structures emerge is essentially the same as detecting existence of First moment obstructions.

From Obstructions to Structures

If $\mathcal{H}$ is p-small, then $\mathbb{P}(X_p \in \langle \mathcal{H} \rangle) < 1/2$.

What about the other direction? Does inexistence of first moment obstructions imply the appearance of structure?

New philosophy:

Inexistence of first moment obstructions implies appearance of structure.

Conjecture (Kahn-Kalai conjecture '06)

If $\mathcal{H}$ is not p-small, then $\mathbb{P}(X_{Cp \log |X|} \in \langle \mathcal{H} \rangle) \geq 1/2$.

Theorem (Park-P. '24, Kahn-Kalai conjecture)

If $\mathcal{H}$ be not p-small with $\ell = \max_{H \in \mathcal{H}} |H|$, then $\mathbb{P}(X_{Cp \log \ell} \in \langle \mathcal{H} \rangle) \geq 1/2$.

Determining the threshold where target structures emerge is essentially the same as detecting existence of First moment obstructions.

Many other interesting manifestation of the central philosophy in probability, high-dimensional geometry, etc.

First moment obstructions in Additive combinatorics

New perspective:

First moment obstructions = Existence of large, low-complexity substructures.

First moment obstructions in Additive combinatorics

New perspective:

First moment obstructions = Existence of large, low-complexity substructures.

First moment obstruction in *random Cayley graphs*:

- $X = G$.
- Target structures: $\mathcal{H}$ is the collection of sumsets $A + A$ where $|A| = t$.

First moment obstructions in Additive combinatorics

New perspective:

First moment obstructions = Existence of large, low-complexity substructures.

First moment obstruction in *random Cayley graphs*:

- $X = G$.
- Target structures: $\mathcal{H}$ is the collection of sumsets $A + A$ where $|A| = t$.

Question

Does $\mathcal{H} = \{A + A : |A| = t\}$ admit a small cover?

First moment obstructions in Additive combinatorics

New perspective:

First moment obstructions = Existence of large, low-complexity substructures.

First moment obstruction in *random Cayley graphs*:

- $X = G$.
- Target structures: $\mathcal{H}$ is the collection of sumsets $A + A$ where $|A| = t$.

Question

Does $\mathcal{H} = \{A + A : |A| = t\}$ admit a small cover?

Is there a small collection of *large* sets $\mathcal{F}$ which covers all sumsets $A + A$?

First moment obstructions in Additive combinatorics

New perspective:

First moment obstructions = Existence of large, low-complexity substructures.

First moment obstruction in *random Cayley graphs*:

- $X = G$.
- Target structures: $\mathcal{H}$ is the collection of sumsets $A + A$ where $|A| = t$.

Question

Does $\mathcal{H} = \{A + A : |A| = t\}$ admit a small cover?

Is there a small collection of *large* sets $\mathcal{F}$ which covers all sumsets $A + A$?

Perspective: New ways to quantify the structure of sets with small doubling.

First moment obstructions & Low-complexity substructures

Define the complexity of a collection of sets $\mathcal{F}$ as $\log |\mathcal{F}|$.

Question

Do sumsets $A + A$ contain large low-complexity subsets?

First moment obstructions & Low-complexity substructures

Define the complexity of a collection of sets $\mathcal{F}$ as $\log|\mathcal{F}|$.

Question

Do sumsets $A+A$ contain large low-complexity subsets?

Theorem (Alon '07)

The independence number of $G(p)$ is with high probability $\tilde{O}(p^{-2})$.

First moment obstructions & Low-complexity substructures

Define the complexity of a collection of sets $\mathcal{F}$ as $\log |\mathcal{F}|$.

Question

Do sumsets $A + A$ contain large low-complexity subsets?

Theorem (Alon '07)

The independence number of $G(p)$ is with high probability $\tilde{O}(p^{-2})$.

Key observation: There exists a collection $\mathcal{F}$ of sets of size t with complexity $t^{1/2}$ that covers $\mathcal{H}$.

- The collection of sumsets $A' + A'$ of a random subset A' of size $t^{1/2}$ of A satisfies this property.

First moment obstructions & Low-complexity substructures

Define the complexity of a collection of sets $\mathcal{F}$ as $\log |\mathcal{F}|$.

Question

Do sumsets $A + A$ contain large low-complexity subsets?

Theorem (Alon '07)

The independence number of $G(p)$ is with high probability $\tilde{O}(p^{-2})$.

Key observation: There exists a collection $\mathcal{F}$ of sets of size t with complexity $t^{1/2}$ that covers $\mathcal{H}$.

- The collection of sumsets $A' + A'$ of a random subset A' of size $t^{1/2}$ of A satisfies this property.

Fundamental barrier at $t^{1/2}$:

- For A with small doubling $K = \frac{|A+A|}{|A|}$, $|A + A| \leq |A|^{1+\delta}$, any improvement must leverage suitably additional structure.

First moment obstructions & Low-complexity substructures

Question

Among $|A| = t$ with small sumset $A + A$, does $A + A$ contain large low-complexity subsets?

First moment obstructions & Low-complexity substructures

Question

Among $|A| = t$ with small sumset $A + A$, does $A + A$ contain large low-complexity subsets?

While sets with small sumsets are structured (Freiman), the quantitative bounds are too weak for us.

First moment obstructions & Low-complexity substructures

Question

Among $|A| = t$ with small sumset $A + A$, does $A + A$ contain large low-complexity subsets?

While sets with small sumsets are structured (Freiman), the quantitative bounds are too weak for us.

Beside weak quantitative bounds, even when A is dense in G, existence of low-complexity cover is open.

Question (Lovett)

For $G = \mathbb{F}_2^d$, does there exist a collection of dense subsets of G with complexity $d^{O(1)}$ which covers the collection of sumsets $A + A$ where $|A| = \Omega(2^d)$?

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Sumsets of sets with small doubling contain large low-complexity subsets.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Sumsets of sets with small doubling contain large low-complexity subsets.

The case $K = O(1)$ resolves positively the question of Lovett.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Sumsets of sets with small doubling contain large low-complexity subsets.

The case $K = O(1)$ resolves positively the question of Lovett.

The collection $A + A$ for $|A| = t \approx p^{-3/2}$ is $(1 - p)$-small.

Theorem (Alon-P. '25+)

The independence number of $G(p)$ is with high probability $\tilde{O}(p^{-3/2})$.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Probabilistic approximation:

- Both bounds rely on approximation of large level sets of the convolution $A * A(x) = \mathbb{E}_y[A(y)A(x - y)]$.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Probabilistic approximation:

- Both bounds rely on approximation of large level sets of the convolution $A * A(x) = \mathbb{E}_y[A(y)A(x - y)]$.
- In the range K large, use a random sampling argument in physical space.

The main covering lemma

Efficient covering lemma (informal version)

There exists a collection of sets $\mathcal{C}$ such that:

- $|\mathcal{C}| \leq \exp(\tilde{O}(\min(K^2, \sqrt{Kt}))$.
- Every $C \in \mathcal{C}$ has $|C| \geq \tilde{\Omega}(Kt)$.
- For every A with $|A| = t$ and $|A + A| \leq K|A|$, there exists $C \in \mathcal{C}$ such that $C \subseteq A + A$.

Probabilistic approximation:

- The hard range (K small) performs approximation in the Fourier space:

$$A * A(x) = \sum_{\chi} \hat{A}(\chi)^2 \chi(x).$$

- Sample $\approx K^2$ random characters χ according to $|\hat{A}(\chi)|^2$, and construct a suitable Fourier-sparse pointwise approximation of $A * A$.

Low-complexity structures in sets with small doubling

First moment obstructions $\longleftrightarrow$ Existence of low-complexity substructures.

Low-complexity structures in sets with small doubling

First moment obstructions $\longleftrightarrow$ Existence of low-complexity substructures.

New perspectives on the structures of sets with small doubling:

- Construction of low-complexity subsets of sumsets.

Low-complexity structures in sets with small doubling

First moment obstructions $\longleftrightarrow$ Existence of low-complexity substructures.

New perspectives on the structures of sets with small doubling:

- Construction of low-complexity subsets of sumsets.
- Explicit low-complexity representation - Resolve questions of independent interest.

Low-complexity structures in sets with small doubling

First moment obstructions $\longleftrightarrow$ Existence of low-complexity substructures.

New perspectives on the structures of sets with small doubling:

- Construction of low-complexity subsets of sumsets.
- Explicit low-complexity representation - Resolve questions of independent interest.

While achieving the desired $O(1)$ complexity for doubling $K = O(1)$, our dependence on K is suboptimal.

Low-complexity structures in sets with small doubling

First moment obstructions $\longleftrightarrow$ Existence of low-complexity substructures.

New perspectives on the structures of sets with small doubling:

- Construction of low-complexity subsets of sumsets.
- Explicit low-complexity representation - Resolve questions of independent interest.

While achieving the desired $O(1)$ complexity for doubling $K = O(1)$, our dependence on K is suboptimal.

Question

Can we achieve a low-complexity approximation with optimal dependence on K?

Optimally counting sets with small doubling

Optimal complexity for approximating sumsets

Enumeration of sets with small doubling

Question

What do typical sets with small doubling look like?
What is the number of sets A of size t with doubling K?

Enumeration of sets with small doubling

Question

What do typical sets with small doubling look like?
What is the number of sets A of size t with doubling K?

Structural results are often (necessarily) weak quantitatively.

Enumeration of sets with small doubling

Question

What do typical sets with small doubling look like?
What is the number of sets A of size t with doubling K?

Structural results are often (necessarily) weak quantitatively.

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Enumeration of sets with small doubling

Question

What do typical sets with small doubling look like?
What is the number of sets A of size t with doubling K?

Structural results are often (necessarily) weak quantitatively.

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Motivated from Erdős-Cameron conjecture ('88) on enumeration of sum-free sets.

Enumeration of sets with small doubling

Question

What do typical sets with small doubling look like?
What is the number of sets A of size t with doubling K?

Structural results are often (necessarily) weak quantitatively.

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Motivated from Erdős-Cameron conjecture ('88) on enumeration of sum-free sets.

Except for very small K ($K < 3$), structural method is too weak for meaningful enumeration.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.
- The sum of an arithmetic progression of length t/K and K generic elements.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.
- The sum of an arithmetic progression of length t/K and K generic elements.
- In a general abelian group G: Subsets of a $\lfloor (Kt + b)/(2b) \rfloor$-progression of subgroups of size $b \leq Kt$.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.
- The sum of an arithmetic progression of length t/K and K generic elements.
- In a general abelian group G: Subsets of a $\lfloor (Kt + b)/(2b) \rfloor$-progression of subgroups of size $b \leq Kt$.
- In $\mathbb{F}_p^d$: subgroups of size $t \sim d \log d$.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.
- The sum of an arithmetic progression of length t/K and K generic elements.
- In a general abelian group G: Subsets of a $\lfloor (Kt + b)/(2b) \rfloor$-progression of subgroups of size $b \leq Kt$.
- In $\mathbb{F}_p^d$: subgroups of size $t \sim d \log d$.

The examples show that the conjecture can only hold for $K \ll t/(\log N \log\log N)$.

Optimally counting sets with small doubling

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$.

Examples:

- Subsets of an arithmetic progression of length $Kt/2$.
- The sum of an arithmetic progression of length t/K and K generic elements.
- In a general abelian group G: Subsets of a $\lfloor (Kt + b)/(2b) \rfloor$-progression of subgroups of size $b \leq Kt$.
- In $\mathbb{F}_p^d$: subgroups of size $t \sim d \log d$.

The examples show that the conjecture can only hold for $K \ll t/(\log N \log\log N)$.

Conjecture (Alon-Balogh-Morris-Samotij '14)

Consider an abelian group G of order N. Let $K \leq t/(\log N \log\log N)$. The number of $A \subseteq G$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{(Kt+s)/2}{t}$, where s is the maximum size of a subgroup of size at most Kt.

Optimally counting sets with small doubling in abelian groups

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

Optimally counting sets with small doubling in abelian groups

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

Progress:

- Green-Morris '16: $K = O(1)$.
- Campos '20, Campos-Collares-Morris-Morrison-Souza '22: $K \ll t/(\log N)^3$.
- Liu-Mattos-Szabó '25: $K \ll t/(\log N)^2$.

Optimally counting sets with small doubling in abelian groups

Conjecture (Alon-Balogh-Morris-Samotij '14)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

Progress:

- Green-Morris '16: $K = O(1)$.
- Campos '20, Campos-Collares-Morris-Morrison-Souza '22: $K \ll t/(\log N)^3$.
- Liu-Mattos-Szabó '25: $K \ll t/(\log N)^2$.

Technique:

- Hypergraph container lemma.
- Regularity-type method.

Optimally counting sets with small doubling in abelian groups

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

Optimally counting sets with small doubling in abelian groups

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

New perspective:

$$\text{First moment obstruction} \quad \to \quad \text{Low-complexity structures in } A + A.$$

Optimally counting sets with small doubling in abelian groups

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

New perspective:

$$\text{First moment obstruction} \quad \to \quad \text{Low-complexity structures in } A + A.$$

$$\text{Low-complexity } \textbf{approximation} \text{ of } A + A \quad \to \quad \text{Enumeration of } A + A.$$

Optimally counting sets with small doubling in abelian groups

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

New perspective:

First moment obstruction $\rightarrow$ Low-complexity structures in $A + A$.

Low-complexity **approximation** of $A + A$ $\rightarrow$ Enumeration of $A + A$.

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F \Delta (A + A)| \leq o(Kt)$.

Optimally counting sets with small doubling in abelian groups

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| = Kt$ is $\exp(o(t))\binom{Kt/2}{t}$ for $K \ll t/(\log N \log\log N)$.

New perspective:

$$\text{First moment obstruction} \quad \to \quad \text{Low-complexity structures in } A + A.$$

$$\text{Low-complexity } \textbf{approximation} \text{ of } A + A \quad \to \quad \text{Enumeration of } A + A.$$

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F \Delta (A + A)| \leq o(Kt)$.

More accurately, the approximation applies to the subset B of $A + A$ consisting of elements x with at least $\epsilon t/K$ representations as $a_1 + a_2$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A + A)| \leq o(Kt)$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Every sumset $A+A$ of a set A with doubling K can be approximated by a set of complexity $\tilde{O}(K)$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Every sumset $A+A$ of a set A with doubling K can be approximated by a set of complexity $\tilde{O}(K)$.

This complexity is optimal.

- Consider K generic translates of an arithmetic progression of length t/K.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Every sumset $A+A$ of a set A with doubling K can be approximated by a set of complexity $\tilde{O}(K)$.

This complexity is optimal.

- Consider K generic translates of an arithmetic progression of length t/K.

By relaxing the one-sided covering condition, we can attain optimal complexity for approximating $A+A$.

This is sufficient for enumeration of sets with small doubling.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A + A)| \leq o(Kt)$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A + A)| \leq o(Kt)$.

Recover enumeration of A from Approximation lemma:

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Recover enumeration of A from Approximation lemma:

- Find F of low-complexity which approximates $A+A$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Recover enumeration of A from Approximation lemma:

- Find F of low-complexity which approximates $A+A$.
- Based on F, construct a superset $X \supset A$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

Recover enumeration of A from Approximation lemma:

- Find F of low-complexity which approximates $A+A$.
- Based on F, construct a superset $X \supset A$.
- Via a graph container algorithm, show that we can efficiently refine until $|X| \leq (1+o(1))Kt/2$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A + A)| \leq o(Kt)$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A + A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A + A)| \leq o(Kt)$.

The approximation lemma relies on a probabilistic approximation simultaneously in the Fourier and physical space:

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

The approximation lemma relies on a probabilistic approximation simultaneously in the Fourier and physical space:

- Based on a small random sample of a subset of A, construct an approximation $\hat{f}$ of $\hat{A}$.

Approximation lemma

Theorem (Approximation lemma (Informal), P. '25+)

For every t and K, there exists $\mathcal{F}$ with $|\mathcal{F}| \leq \exp(K(\log N)(\log\log N))$ such that: For every $|A| = t$ with $|A+A| = Kt$, there exists $F \in \mathcal{F}$ such that $|F\Delta(A+A)| \leq o(Kt)$.

The approximation lemma relies on a probabilistic approximation simultaneously in the Fourier and physical space:

- Based on a small random sample of a subset of A, construct an approximation $\hat{f}$ of $\hat{A}$.
- Smoothen $\hat{f}$, and apply Fourier inversion on $\hat{f}^2$.

Refined counting

Question (Green-Morris '16)

What is the number of $A \subseteq \mathbb{Z}_N$ with $|A| = t = \Theta(\log N)$ and $|A + A| \leq Kt$?

Refined counting

Question (Green-Morris '16)

What is the number of $A \subseteq \mathbb{Z}_N$ with $|A| = t = \Theta(\log N)$ and $|A + A| \leq Kt$?

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| \leq Kt$ for $K \ll t/(\log t \log\log t)$ is $\max_r \exp(o(t)) N^{r+1} \binom{(Kt-(r-1)t)/2}{t-r+1}$.

Refined counting

Question (Green-Morris '16)

What is the number of $A \subseteq \mathbb{Z}_N$ with $|A| = t = \Theta(\log N)$ and $|A + A| \leq Kt$?

Theorem (P. '25+)

The number of $A \subseteq \mathbb{Z}_N$ with $|A| = t$ and $|A + A| \leq Kt$ for $K \ll t/(\log t \log\log t)$ is $\max_r \exp(o(t)) N^{r+1} \binom{(Kt-(r-1)t)/2}{t-r+1}$.

Rely on approximation framework together with a crucial additional ingredient: Robust version of Freiman-Ruzsa's lemma over $\mathbb{Z}_N$.

- Established in Alon-P. '25, for sharp asymptotics of the independence number of logarithmically sparse random Cayley graph.
- Proof relies on the main combinatorial lemma.

Conclusion

Existence of first moment obstructions forbids appearance of structures.

Inexistence of first moment obstructions implies appearance of structures.

Conclusion

Existence of first moment obstructions forbids appearance of structures.

Inexistence of first moment obstructions implies appearance of structures.

First moment obstructions suggest existence of low-complexity substructures.

Randomized approximations provide a pathway to Low-complexity approximations (and hence First moment obstructions).

Outlook: Low-complexity approximations

Our low-complexity covers and approximations provide explicit classes of structured functions that approximate large level sets of the Fourier transform.

- What properties can be further extracted from the low-complexity family of functions?

Outlook: Low-complexity approximations

Our low-complexity covers and approximations provide explicit classes of structured functions that approximate large level sets of the Fourier transform.

- What properties can be further extracted from the low-complexity family of functions?

Conjecture (Alon-P.)

There exists a collection of sets of size $\tilde{\Omega}(Kt)$ of complexity $\tilde{O}(K)$ which cover $A+A$ where $|A| = t$ and $|A+A| \leq Kt$.

Outlook: Low-complexity approximations

Our low-complexity covers and approximations provide explicit classes of structured functions that approximate large level sets of the Fourier transform.

- What properties can be further extracted from the low-complexity family of functions?

Conjecture (Alon-P.)

There exists a collection of sets of size $\tilde{\Omega}(Kt)$ of complexity $\tilde{O}(K)$ which cover $A+A$ where $|A| = t$ and $|A+A| \leq Kt$.

Directions

- Further applications of low-complexity approximations beyond the additive combinatorial context.
- Alon's conjecture in groups with exponent 2 and 3.
- First moment obstructions beyond random graphs: Study interesting properties of Δ-independent graphs (random entangled graphs, random Cayley graphs).

Thank you!