Structures in random graphs: New connections

Huy Tuan Pham

Stanford University

With David Conlon (Caltech), Jacob Fox (Stanford), Liana Yepremyan (Emory), Jinyoung Park (NYU) and Vishesh Jain (UIC)

Simons Institute - July 2023

The beginning

Theorem (Erdős 1947)

There exists graphs on N vertices with no clique or independent set of size $2\log_2 N$.

The beginning

Theorem (Erdős 1947)

There exists graphs on N vertices with no clique or independent set of size $2\log_2 N$.

- One of the first applications of the probabilistic method. Give a lower bound on Ramsey numbers.

The beginning

Theorem (Erdős 1947)

There exists graphs on N vertices with no clique or independent set of size $2\log_2 N$.

- One of the first applications of the probabilistic method. Give a lower bound on Ramsey numbers.
- Erdős shows that $G(N, 1/2)$ does not have a clique or independent set of size $n = 2\log_2 N$ by considering the first moment: The expected number of such cliques or independent sets is $\binom{N}{n}2^{-\binom{n}{2}}$ which is small.

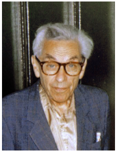
The beginning

Theorem (Erdős 1947)

There exists graphs on N vertices with no clique or independent set of size $2\log_2 N$.

- One of the first applications of the probabilistic method. Give a lower bound on Ramsey numbers.
- Erdős shows that $G(N, 1/2)$ does not have a clique or independent set of size $n = 2\log_2 N$ by considering the first moment: The expected number of such cliques or independent sets is $\binom{N}{n}2^{-\binom{n}{2}}$ which is small.
- Much more precise asymptotic understanding of the clique and independence numbers of $G(N, p)$ by Matula and Bollobás and Erdős.

Timeline



Random graphs

Ramsey

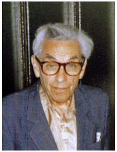
First moment prediction

- Suprema of stochastic processes
- Structure of small sets
- Random optimization
- Convex geometry

Sunflowers

Complexity / Random restriction / DNF sparsification

Timeline



Random graphs

Ramsey

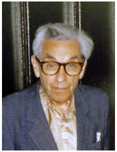
First moment
prediction / obstruction

- Suprema of stochastic processes
- Structure of small sets
- Random optimization
- Convex geometry

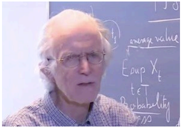
Sunflowers

Complexity / Random restriction /
DNF sparsification

Timeline



Random graphs

- Suprema of stochastic processes
- Structure of small sets
- Random optimization
- Convex geometry

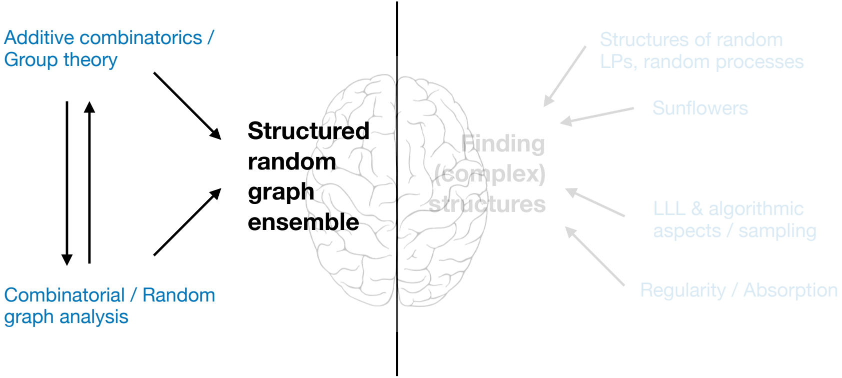
Ramsey

First moment
prediction / obstruction

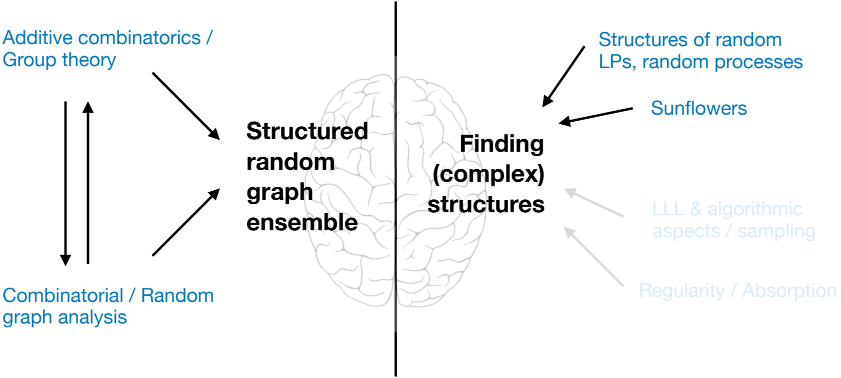
Sunflowers

Complexity / Random restriction /
DNF sparsification

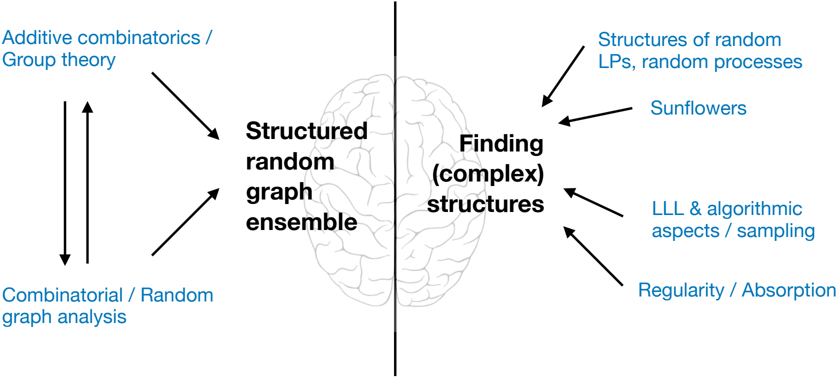
Today Roadmap



Today Roadmap



Today Roadmap



Ramsey graphs

Definition (Ramsey graphs)

A graph on N vertices is *C-Ramsey* if it has no clique or independent set of size $C \log_2 N$.

Ramsey graphs

Definition (Ramsey graphs)

A graph on N vertices is *C-Ramsey* if it has no clique or independent set of size $C \log_2 N$.

Theorem (Erdős 1947)

Almost all graphs on N vertices are 2-Ramsey.

Ramsey graphs

Definition (Ramsey graphs)

A graph on N vertices is *C-Ramsey* if it has no clique or independent set of size $C \log_2 N$.

Theorem (Erdős 1947)

Almost all graphs on N vertices are 2-Ramsey.

Theorem (Campos-Griffiths-Morris-Saharasbudhe 2023)

There is no $\frac{1}{2} + \epsilon$-Ramsey graph.

Ramsey graphs

Definition (Ramsey graphs)

A graph on N vertices is *C-Ramsey* if it has no clique or independent set of size $C \log_2 N$.

Theorem (Erdős 1947)

Almost all graphs on N vertices are 2-Ramsey.

Theorem (Campos-Griffiths-Morris-Saharasbudhe 2023)

There is no $\frac{1}{2} + \epsilon$-Ramsey graph.

Problem (Erdős)

Explicitly construct C-Ramsey graphs for some constant C.

Ramsey Cayley graphs

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 $xy^{-1} \in S$.

Ramsey Cayley graphs

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

Ramsey Cayley graphs

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

- Motivations: (Random) Cayley graphs and their applications have been extensively studied in theoretical computer science, combinatorics, group theory. Strong connections to coding theory, spectral graph theory, etc.

Ramsey Cayley graphs

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

- Motivations: (Random) Cayley graphs and their applications have been extensively studied in theoretical computer science, combinatorics, group theory. Strong connections to coding theory, spectral graph theory, etc.

Question

What is the size of the largest clique or independent set in uniform random Cayley graphs ($G(1/2)$)? Are uniform random Cayley graphs Ramsey ?

Ramsey Cayley graphs

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

- Motivations: (Random) Cayley graphs and their applications have been extensively studied in theoretical computer science, combinatorics, group theory. Strong connections to coding theory, spectral graph theory, etc.

Question

What is the size of the largest clique or independent set in uniform random Cayley graphs ($G(1/2)$)? Are uniform random Cayley graphs Ramsey ?

Conjecture (Alon 1989)

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

Clique number of random Cayley graphs

Clique number of random Cayley graphs

Theorem (Alon 1995)

The clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$ whp.

Clique number of random Cayley graphs

Theorem (Alon 1995)

The clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$ whp.

Theorem (Green 2005, Green-Morris 2016)

For N prime, the clique number of a uniform random Cayley graph on $\mathbb{Z}_N$ is $(2 + o(1)) \log_2 N$ whp.

Clique number of random Cayley graphs

Theorem (Alon 1995)

The clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$ whp.

Theorem (Green 2005, Green-Morris 2016)

For N prime, the clique number of a uniform random Cayley graph on $\mathbb{Z}_N$ is $(2 + o(1)) \log_2 N$ whp.

Theorem (Green 2005, Mrazović 2017)

The clique number of a uniform random Cayley graph on $\mathbb{F}_2^d$ with $N = 2^d$ is $\Theta(\log N \log \log N)$ whp.

Random graphs meet additive combinatorics

Theorem (Conlon-Fox-P.-Yepremyan)

The clique number of a uniform random Cayley graph on any group G of order N is almost surely $O(\log N \log \log N)$.

Random graphs meet additive combinatorics

Theorem (Conlon-Fox-P.-Yepremyan)

The clique number of a uniform random Cayley graph on any group G of order N is almost surely $O(\log N \log \log N)$.

The first moment of the number of cliques in a uniformly random Cayley graph is intimately related to the number of subsets with small product sets: The first moment must reflect the structure of the group and random graph ensemble.

Random graphs meet additive combinatorics

Theorem (Conlon-Fox-P.-Yepremyan)

The clique number of a uniform random Cayley graph on any group G of order N is almost surely $O(\log N \log\log N)$.

The first moment of the number of cliques in a uniformly random Cayley graph is intimately related to the number of subsets with small product sets: The first moment must reflect the structure of the group and random graph ensemble.

Theorem (Conlon-Fox-P.-Yepremyan)

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

- $AA^{-1} := \{ab^{-1} : a, b \in A\}$.
- Note that A is a clique in G_S if and only if $AA^{-1} \setminus \{1_G\} \subset S$.
- Previously analyzed in nice abelian groups via strong structural results/regularity method.

Alon's conjecture and expansion

Observation: Subgroups (sets with small expansion) at roughly logarithmic size are problematic and account for the $\log \log N$ factor.

Alon's conjecture and expansion

Observation: Subgroups (sets with small expansion) at roughly logarithmic size are problematic and account for the $\log\log N$ factor.

- Let $N = 5^d$. A uniform random $S \subset \mathbb{F}_5^d$ a.a.s. contains the nonzero elements of a subspace of order $\Theta(\log N \log\log N)$ and hence G_S a.a.s. contains a clique of that order.

Alon's conjecture and expansion

Observation: Subgroups (sets with small expansion) at roughly logarithmic size are problematic and account for the $\log \log N$ factor.

- Let $N = 5^d$. A uniform random $S \subset \mathbb{F}_5^d$ a.a.s. contains the nonzero elements of a subspace of order $\Theta(\log N \log \log N)$ and hence G_S a.a.s. contains a clique of that order.

Alon's conjecture and expansion

Observation: Subgroups (sets with small expansion) at roughly logarithmic size are problematic and account for the $\log \log N$ factor.

- Let $N = 5^d$. A uniform random $S \subset \mathbb{F}_5^d$ a.a.s. contains the nonzero elements of a subspace of order $\Theta(\log N \log \log N)$ and hence G_S a.a.s. contains a clique of that order.

Theorem (Conlon-Fox-P.-Yepremyan)

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

Alon's conjecture and expansion

Observation: Subgroups (sets with small expansion) at roughly logarithmic size are problematic and account for the $\log \log N$ factor.

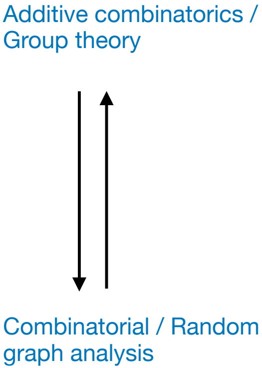
- Let $N = 5^d$. A uniform random $S \subset \mathbb{F}_5^d$ a.a.s. contains the nonzero elements of a subspace of order $\Theta(\log N \log \log N)$ and hence G_S a.a.s. contains a clique of that order.

Theorem (Conlon-Fox-P.-Yepremyan)

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

In particular, all N for which the largest factor which is a power of 2 or 3 is at most $(\log N)^{.001}$ has the above property.

Towards additive combinatorics, and back?



Our analysis combines closely purely combinatorial view and additive insights:

- Purely combinatorial view on the role of the group structure, analyzed via exploration reminiscent of classical random graph analysis.
- Relation between solutions to linear equations, expansion (large product sets) and dimension.

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

Surprise: The combinatorial constraint on the degree of color classes is sufficient!
This naturally leads to studying a more general random graph model:

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

Surprise: The combinatorial constraint on the degree of color classes is sufficient!
This naturally leads to studying a more general random graph model:

- Consider an edge-coloring c of a complete graph.
  An *entangled graph* is the edge-union of some of the color classes.

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

Surprise: The combinatorial constraint on the degree of color classes is sufficient!
This naturally leads to studying a more general random graph model:

- Consider an edge-coloring c of a complete 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.

A purely combinatorial view on Cayley graphs

What is the role of the group structure?

- Consider a group G of order N. Color the edges of the complete graph on G by assigning each edge (x, y) the color $\{xy^{-1}, yx^{-1}\}$. This edge-coloring of K_N is such that each color class is 1 or 2-regular.
  A Cayley graph on G is the edge-union of some color classes.

Surprise: The combinatorial constraint on the degree of color classes is sufficient!
This naturally leads to studying a more general random graph model:

- Consider an edge-coloring c of a complete 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.

- Constraint: Each color class has bounded degree.

Cliques in entangled graphs & Sets with small product set

Cliques in entangled graphs & Sets with small product set

Theorem 1

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

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

Cliques in entangled graphs & Sets with small product set

Theorem 1

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

If an edge-coloring c of K_N is Δ-bounded, then a.a.s.

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

Cliques in entangled graphs & Sets with small product set

Theorem 1

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

If an edge-coloring c of K_N is Δ-bounded, then a.a.s.

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

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

Sets with small product set: Greedy exploration

Theorem 1 (weaker version)

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

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

Sets with small product set: Greedy exploration

Theorem 1 (weaker version)

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

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

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree using $O(K \log n)$ colors.

Sets with small product set: Greedy exploration

Theorem 1 (weaker version)

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

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

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree using $O(K \log n)$ colors.

Greedy process to grow a large component:

- In each step, pick a color that maximally extends the size of the component.
- Guarantee that the component grows roughly by a factor $1/K$ per step. Hence, the entire set is connected in $O(K \log n)$ steps.

Sets with small product set: Greedy exploration

Theorem 1 (weaker version)

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

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

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree using $O(K \log n)$ colors.

Greedy process to grow a large component:

- In each step, pick a color that maximally extends the size of the component.
- Guarantee that the component grows roughly by a factor $1/K$ per step. Hence, the entire set is connected in $O(K \log n)$ steps.

The bound on the number of colors required in a spanning tree is tight.

Sets with small product set: Random exploration

Theorem 1

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

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

Sets with small product set: Random exploration

Theorem 1

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

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

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree with $O(K + \log n)$ colors on $(1-\epsilon)n$ vertices.

Sets with small product set: Random exploration

Theorem 1

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

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

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree with $O(K + \log n)$ colors on $(1-\epsilon)n$ vertices.

- Random exploration: Expose colors randomly and analyze the connected components formed.
- When the set of colors coming out of components is large ($\Omega(Kn)$), large components will merge in $O(\log n)$ steps to a giant component.

Sets with small product set: Random exploration

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree with $O(K + \log n)$ colors on $(1 - \epsilon)n$ vertices.

Sets with small product set: Random exploration

Lemma

For any n vertex subset with Kn colors, we can find a spanning tree with $O(K + \log n)$ colors on $(1 - \epsilon)n$ vertices.

Building large components:

- Pick random colors and grow components using all edges of these colors.
- Keep track of colors going out of each component; as the components grow, the set of colors out of the components also grows.
- Prove that the probability that a component grows increases with the size of the component. Hence, most vertices are in components of size $\Omega(K)$ after $O(K)$ samples of colors.

Alon's conjecture - Going beyond uniform random

Alon's conjecture - Going beyond uniform random

Theorem

There is a self-complementary Ramsey Cayley graph on $\mathbb{F}_5^d$.

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

Alon's conjecture - Going beyond uniform random

Theorem

There is a self-complementary Ramsey Cayley graph on $\mathbb{F}_5^d$.

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

Alon's conjecture - Going beyond uniform random

Theorem

There is a C-Ramsey self-complementary Cayley graph on $\mathbb{F}_5^d$.

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S:

Alon's conjecture - Going beyond uniform random

Theorem

There is a C-Ramsey self-complementary Cayley graph on $\mathbb{F}_5^d$.

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S:

- S is symmetric.
- G_S is self-complementary with isomorphism $\phi(x) = 2x$.
- If $x \in S$, then $2x \notin S$.

Alon's conjecture - Going beyond uniform random

Theorem

There is a C-Ramsey self-complementary Cayley graph on $\mathbb{F}_5^d$.

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S:

- S is symmetric.
- G_S is self-complementary with isomorphism $\phi(x) = 2x$.
- If $x \in S$, then $2x \notin S$.

The last condition leads to expansion of any potential clique: $|A + 2 \cdot A| = |A|^2$, so the Plünnecke-Ruzsa inequality implies

$$|A|^2 = |A + 2 \cdot A| \leq |A + A + A| \leq |A - A|^3 |A|^{-2},$$

yielding $|A - A| \geq |A|^{4/3}$.

Alon's conjecture - Going beyond uniform random

Theorem

There is a C-Ramsey self-complementary Cayley graph on $\mathbb{F}_5^d$.

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S:

- S is symmetric.
- G_S is self-complementary with isomorphism $\phi(x) = 2x$.
- If $x \in S$, then $2x \notin S$.

The last condition leads to expansion of any potential clique: $|A + 2 \cdot A| = |A|^2$, so the Plünnecke-Ruzsa inequality implies

$$|A|^2 = |A + 2 \cdot A| \leq |A + A + A| \leq |A - A|^3 |A|^{-2},$$

yielding $|A - A| \geq |A|^{4/3}$.
Expansion, together with the previous counting result, allows the union bound to work in the large K range without losing the $\log \log N$ factor.

From random graphs to additive combinatorics and back

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S.

Theorem

The random Cayley graph G_S has clique and independence number $(2 + o(1)) \log_2 N$.

From random graphs to additive combinatorics and back

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S.

Theorem

The random Cayley graph G_S has clique and independence number $(2 + o(1)) \log_2 N$.

The precise asymptotic is closely related to the constant in front of K in the $N^{O(K)}$ factor in the counting result.

- Over vector spaces, sharp dependence on K can be determined through sharp bound on the dimension of A.
- This can be obtained from results on Freiman's conjecture over $\mathbb{F}_p^n$ by Chaim Even-Zohar and Lovett: If $|A - A| \le K|A|$ and $|A| = p^{o(K)}$, then the dimension of A is at most $(1 + o(1))K$.

From random graphs to additive combinatorics and back

Model. For each nonzero $x \in \mathbb{F}_5^d$, randomly pick exactly one of $\{x, 4x\}$ or $\{2x, 3x\}$ to be a subset of the generating set S.

Theorem

The random Cayley graph G_S has clique and independence number $(2 + o(1)) \log_2 N$.

The precise asymptotic is closely related to the constant in front of K in the $N^{O(K)}$ factor in the counting result.

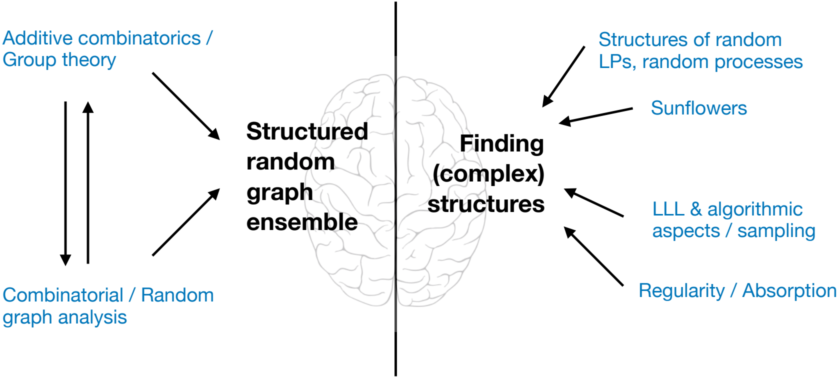
- Over vector spaces, sharp dependence on K can be determined through sharp bound on the dimension of A.
- This can be obtained from results on Freiman's conjecture over $\mathbb{F}_p^n$ by Chaim Even-Zohar and Lovett: If $|A - A| \leq K|A|$ and $|A| = p^{o(K)}$, then the dimension of A is at most $(1 + o(1))K$.

Even sharper asymptotics requires more precise understanding of the additive structure and correspondence with the combinatorial analysis: Make full leverage of the expansion condition and combinatorial insights.

Roadmap

Careful understanding of the role of structure on the first moment is crucial.

Mutual connection between additive combinatorics and random exploration/random graph view.



Interlude: More general structures

- Finite set X, random subset X_p. Collection of desired structures $\mathcal{H} \subseteq 2^X$.

Interlude: More general structures

- Finite set X, random subset X_p. Collection of desired structures $\mathcal{H} \subseteq 2^X$.
- The *threshold* $p_c(\mathcal{H})$ is the value p at which the probability that a structure in $\mathcal{H}$ appears in X_p is $1/2$.

Interlude: More general structures

- Finite set X, random subset X_p. Collection of desired structures $\mathcal{H} \subseteq 2^X$.
- The *threshold* $p_c(\mathcal{H})$ is the value p at which the probability that a structure in $\mathcal{H}$ appears in X_p is $1/2$.

Theorem (the Kahn-Kalai conjecture '06, resolved by Park-P. '22+)

The threshold $p_c(\mathcal{H})$ is closely predicted by the *expectation threshold* $p_E(\mathcal{H})$:

$$p_c(\mathcal{H}) = O(p_E(\mathcal{H}) \log |X|).$$

Interlude: More general structures

- Finite set X, random subset X_p. Collection of desired structures $\mathcal{H} \subseteq 2^X$.
- The *threshold* $p_c(\mathcal{H})$ is the value p at which the probability that a structure in $\mathcal{H}$ appears in X_p is $1/2$.

Theorem (the Kahn-Kalai conjecture '06, resolved by Park-P. '22+)

The threshold $p_c(\mathcal{H})$ is closely predicted by the *expectation threshold* $p_E(\mathcal{H})$:

$$p_c(\mathcal{H}) = O(p_E(\mathcal{H}) \log |X|).$$

The expectation threshold is defined as the largest p for which there is $\mathcal{H}'$ with

- $\mathcal{H}'$ covers $\mathcal{H}$: All $H \in \mathcal{H}$ contains some $H' \in \mathcal{H}'$.
- $\mathcal{H}'$ has a small cost: $\sum_{H' \in \mathcal{H}'} p^{|H'|} \leq 1/2$ (naive union bound/first moment).

We say $\mathcal{H}$ is p-small if there exists $\mathcal{H}'$ satisfying the above properties.

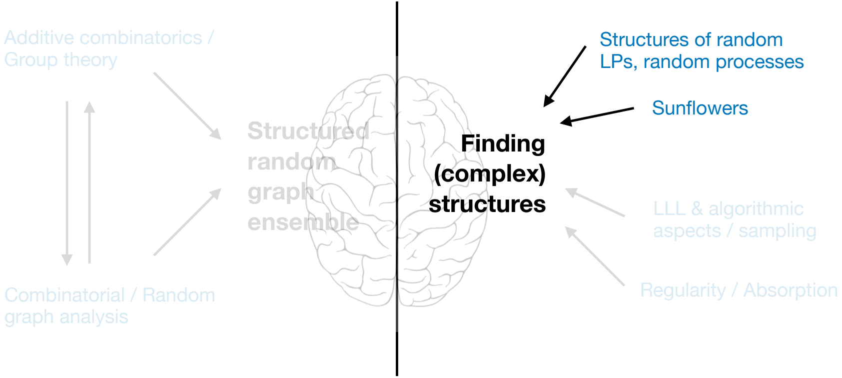
Thresholds and the Kahn-Kalai conjecture

Theorem (the Kahn-Kalai conjecture '06, resolved by Park-P. '22+)

If $\mathcal{H}$ is not p-small, then $X_{Lp \log |X|}$ contains a set from $\mathcal{H}$ with probability at least $1/2$.

Inexistence of first moment (union bound) obstruction is sufficient to guarantee emergence of structure!

Roadmap



Additive combinatorics / Group theory
Combinatorial / Random graph analysis
Structured random graph ensemble
Finding (complex) structures
Structures of random LPs, random processes
Sunflowers
LLL & algorithmic aspects / sampling
Regularity / Absorption

Thresholds and random LPs

Interesting connections to the structure of random processes and high-dimensional convex geometry (Talagrand '94, '06, '10).

Theorem (Talagrand's selector process conjecture, resolved by Park-P. '22)

Given linear functions $f_i(S) = v_i \cdot S$ on $S \subseteq X$, for which $v_i \geq 0$ and $f_i(X) \geq 1$. If the support of v_i's is not p-small, then

$$\mathbb{E} \sup_i f_i(X_p) = \Omega(1) \text{ independent of } p!$$

Thresholds and random LPs

Interesting connections to the structure of random processes and high-dimensional convex geometry (Talagrand '94, '06, '10).

Theorem (Talagrand's selector process conjecture, resolved by Park-P. '22)

Given linear functions $f_i(S) = v_i \cdot S$ on $S \subseteq X$, for which $v_i \geq 0$ and $f_i(X) \geq 1$. If the support of v_i's is not p-small, then

$$\mathbb{E} \sup_i f_i(X_p) = \Omega(1) \text{ independent of } p!$$

- $\sup_i f_i(X_p)$ - Suprema of stochastic processes: structure of tail events.
- $\sup_i f_i(X_p)$ - Fractionally subadditive/XOS functions under random domain subsampling.

Thresholds and random LPs

Interesting connections to the structure of random processes and high-dimensional convex geometry (Talagrand '94, '06, '10).

Theorem (Talagrand's selector process conjecture, resolved by Park-P. '22)

Given linear functions $f_i(S) = v_i \cdot S$ on $S \subseteq X$, for which $v_i \geq 0$ and $f_i(X) \geq 1$. If the support of v_i's is not p-small, then

$$\mathbb{E} \sup_i f_i(X_p) = \Omega(1) \text{ independent of } p!$$

- $\sup_i f_i(X_p)$ - Suprema of stochastic processes: structure of tail events.
- $\sup_i f_i(X_p)$ - Fractionally subadditive/XOS functions under random domain subsampling.
- Main idea in a simpler setup gives the proof of the Kahn-Kalai conjecture. Kahn-Kalai conjecture as "structure" of containment.

Estimating expectation threshold

- Bounding the expectation threshold by the dual certificate (Talagrand): $p_E(\mathcal{H}) \le p_f(\mathcal{H})$, the largest p for which there exists a probability measure λ supported on $\mathcal{H}$ which is *p-spread*:

$$\lambda(\{H \in \mathcal{H} : H \supseteq S\}) \le 2p^{|S|} \text{ for all } S.$$

Estimating expectation threshold

- Bounding the expectation threshold by the dual certificate (Talagrand): $p_E(\mathcal{H}) \leq p_f(\mathcal{H})$, the largest p for which there exists a probability measure λ supported on $\mathcal{H}$ which is p-spread:

$$\lambda(\{H \in \mathcal{H} : H \supseteq S\}) \leq 2p^{|S|} \text{ for all } S.$$

- The fractional Kahn-Kalai conjecture and connections to robust sunflowers (Alweiss-Lovett-Wu-Zhang '19, Frankston-Kahn-Narayanan-Park '19):

$$p_c(\mathcal{H}) = O(p_f(\mathcal{H}) \log |X|).$$

Estimating expectation threshold

- Bounding the expectation threshold by the dual certificate (Talagrand): $p_E(\mathcal{H}) \le p_f(\mathcal{H})$, the largest p for which there exists a probability measure λ supported on $\mathcal{H}$ which is p-spread:

$$\lambda(\{H \in \mathcal{H} : H \supseteq S\}) \le 2p^{|S|} \text{ for all } S.$$

- The fractional Kahn-Kalai conjecture and connections to robust sunflowers (Alweiss-Lovett-Wu-Zhang '19, Frankston-Kahn-Narayanan-Park '19):

$$p_c(\mathcal{H}) = O(p_f(\mathcal{H}) \log |X|).$$

- If one is interested in a specific family of structures $\mathcal{H}$, an imminent question is how to estimate (fractional) expectation thresholds/construct dual certificates.

Estimating expectation threshold

- Bounding the expectation threshold by the dual certificate (Talagrand): $p_E(\mathcal{H}) \leq p_f(\mathcal{H})$, the largest p for which there exists a probability measure λ supported on $\mathcal{H}$ which is p-spread:

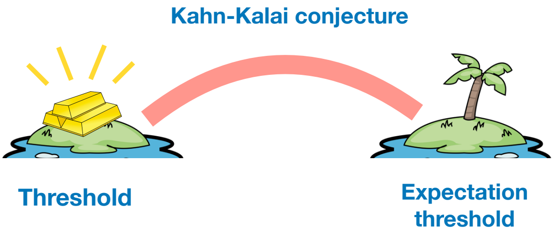
$$\lambda(\{H \in \mathcal{H} : H \supseteq S\}) \leq 2p^{|S|} \text{ for all } S.$$

- The fractional Kahn-Kalai conjecture and connections to robust sunflowers (Alweiss-Lovett-Wu-Zhang '19, Frankston-Kahn-Narayanan-Park '19):

$$p_c(\mathcal{H}) = O(p_f(\mathcal{H}) \log |X|).$$

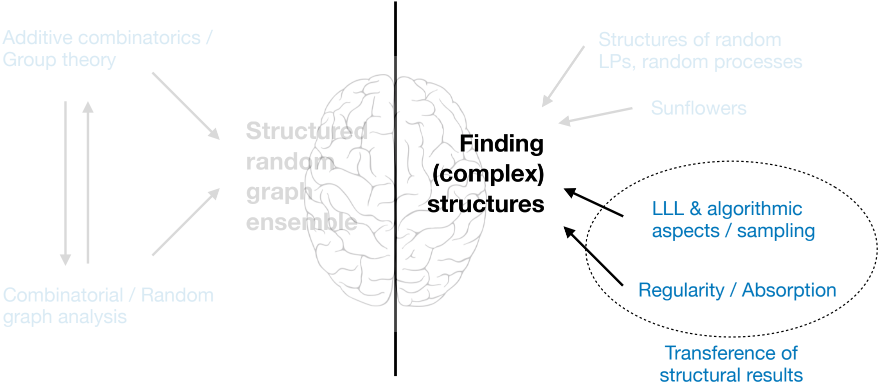
- If one is interested in a specific family of structures $\mathcal{H}$, an imminent question is how to estimate (fractional) expectation thresholds/construct dual certificates.
- Previously restricted to very simple structures in highly symmetric setting, amenable to trivial enumerations.

Where are we?


Kahn-Kalai conjecture
Threshold
Expectation threshold


YOU ARE HERE

Roadmap


Additive combinatorics / Group theory
Combinatorial / Random graph analysis
Structured random graph ensemble
Finding (complex) structures
Structures of random LPs, random processes
Sunflowers
LLL & algorithmic aspects / sampling
Regularity / Absorption
Transference of structural results

Estimating expectation threshold: Latin squares

- The complete bipartite graph $K_{n,n}$ is n-colorable. For each edge of $K_{n,n}$, pick a random list of pn colors from $[n]$.

Estimating expectation threshold: Latin squares

- The complete bipartite graph $K_{n,n}$ is n-colorable. For each edge of $K_{n,n}$, pick a random list of pn colors from $[n]$.
- When can we guarantee that with high probability there exists a proper edge-coloring of $K_{n,n}$ using the selected lists?

Estimating expectation threshold: Latin squares

- The complete bipartite graph $K_{n,n}$ is n-colorable. For each edge of $K_{n,n}$, pick a random list of pn colors from $[n]$.
- When can we guarantee that with high probability there exists a proper edge-coloring of $K_{n,n}$ using the selected lists?
- An $n \times n$ Latin square is a matrix $(x_{i,j})_{i,j \in [n]}$ with entries in $[n]$ where each entry appears exactly once in each row and column.

1	3	2
2	1	3
3	2	1

Estimating expectation threshold: Latin squares

- The complete bipartite graph $K_{n,n}$ is n-colorable. For each edge of $K_{n,n}$, pick a random list of pn colors from $[n]$.
- When can we guarantee that with high probability there exists a proper edge-coloring of $K_{n,n}$ using the selected lists?
- An $n \times n$ Latin square is a matrix $(x_{i,j})_{i,j \in [n]}$ with entries in $[n]$ where each entry appears exactly once in each row and column.

1	3	2
2	1	3
3	2	1

- For each coordinate (i, j), consider a random subset $X_{i,j}$ of $[n]$ where each element is sampled independently with probability p.
- What is the probability that there exists a Latin square with $x_{i,j} \in X_{i,j}$?

Estimating expectation threshold: Latin squares

Conjecture (Johansson '06, Keevash '14, Luria-Simkin '17)

For $p \geq C(\log n)/n$, with high probability, there exists a Latin square with $x_{i,j} \in X_{i,j}$.

- Related conjectures by Simkin, Casselgren-Häggkvist.

Theorem (Jain-P., Keevash '22+)

There exists a C/n-spread probability distribution on Latin squares. As a corollary, for $p \geq C(\log n)/n$, with high probability, there exists a Latin square with $x_{i,j} \in X_{i,j}$.

- Previous partial progress by Sah-Sawhney-Simkin ('22), Kang-Kelly-Kühn-Methuku-Osthus ('22).

Lovász Local Lemma and spread measures

- Main goal: construct a measure over Latin squares which is optimally spread. Difficulty: Latin squares are highly rigid objects.

Lovász Local Lemma and spread measures

- Main goal: construct a measure over Latin squares which is optimally spread. Difficulty: Latin squares are highly rigid objects.
- Insight: Algorithmic construction of spread measures.
- Employ ideas and tools from (iterative) absorption, Lovász Local Lemma and their algorithmic aspects.

Lovász Local Lemma and spread measures

- Main goal: construct a measure over Latin squares which is optimally spread. Difficulty: Latin squares are highly rigid objects.
- Insight: Algorithmic construction of spread measures.
- Employ ideas and tools from (iterative) absorption, Lovász Local Lemma and their algorithmic aspects.
- Summary: Existence of (rigid) object leveraging on LLL. Spread is guaranteed by local uniformity property of the distribution of solutions to the constraint satisfaction problem.

Lovász Local Lemma and spread measures

- Main goal: construct a measure over Latin squares which is optimally spread. Difficulty: Latin squares are highly rigid objects.
- Insight: Algorithmic construction of spread measures.
- Employ ideas and tools from (iterative) absorption, Lovász Local Lemma and their algorithmic aspects.
- Summary: Existence of (rigid) object leveraging on LLL. Spread is guaranteed by local uniformity property of the distribution of solutions to the constraint satisfaction problem.
- Local uniformity property: Under LLL setting, given bad events with probability at most p, and assume that the maximum degree of the dependency graph is Δ with $4p\Delta \leq 1$. For any event $\mathcal{F}$ depending on at most N bad events, the probability of $\mathcal{F}$ under a random satisfying solution is at most the probability of $\mathcal{F}$ under the product measure up to an error $\exp(pN)$.
  - Key property in previous works on algorithmic LLL and recent works on sampling algorithms for the distribution of solutions.

Constructing spread measures

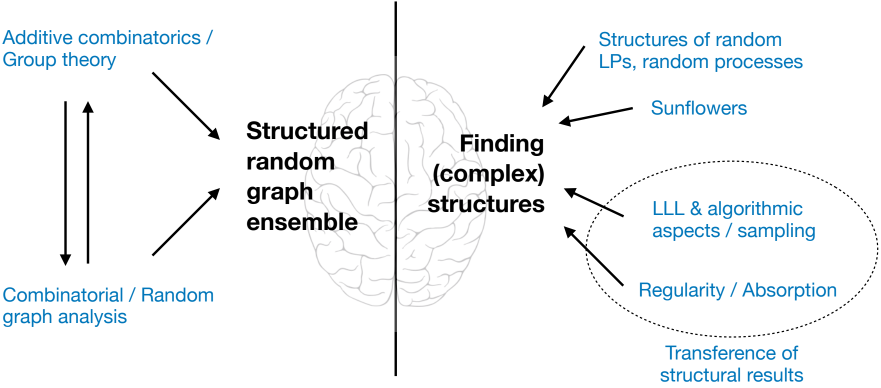
- We view Latin squares as edge decompositions of $K_{n,n}$ into perfect matchings, and construct the desired distribution in progressive steps, decomposing $K_{n,n}$ into regular subgraphs of decreasing degrees.
- Each step employs a random partition that naturally has optimal spread. However, this is not compatible with the rigid (regular) nature of the objects.
- We condition the random partition on satisfying a constraint satisfaction problem, which allows to correct the random object to a regular object.
- Show spread by using local uniformity property, and bootstrap on spread to show success of the iterations.

Constructing spread measures

- We view Latin squares as edge decompositions of $K_{n,n}$ into perfect matchings, and construct the desired distribution in progressive steps, decomposing $K_{n,n}$ into regular subgraphs of decreasing degrees.
- Each step employs a random partition that naturally has optimal spread. However, this is not compatible with the rigid (regular) nature of the objects.
- We condition the random partition on satisfying a constraint satisfaction problem, which allows to correct the random object to a regular object.
- Show spread by using local uniformity property, and bootstrap on spread to show success of the iterations.

- Interesting future direction: Obtaining robust (threshold) versions of other properties given by constraint satisfaction problems.

Roadmap



Additive combinatorics / Group theory
Combinatorial / Random graph analysis
Structured random graph ensemble
Finding (complex) structures
Structures of random LPs, random processes
Sunflowers
LLL & algorithmic aspects / sampling
Regularity / Absorption
Transference of structural results

Further connections: An open invitation

Random Linear Programs:

Further connections: An open invitation

Random Linear Programs:

- Max-Cut/Max-Bisection in $G(N, p)$: The max-cut value is $N(pN/4 + (1 + o_{pN}(1))P_* \sqrt{pN}/2)$ (Dembo-Montanari-Sen '17).

Further connections: An open invitation

Random Linear Programs:

- Max-Cut/Max-Bisection in $G(N, p)$: The max-cut value is $N(pN/4 + (1 + o_{pN}(1))P_* \sqrt{pN}/2)$ (Dembo-Montanari-Sen '17).
- Talagrand's selector process conjecture: Maxima is p-independent if support of linear functions involved is not p-small.

Further connections: An open invitation

Random Linear Programs:

- Max-Cut/Max-Bisection in $G(N, p)$: The max-cut value is $N(pN/4 + (1 + o_{pN}(1))P_* \sqrt{pN}/2)$ (Dembo-Montanari-Sen '17).
- Talagrand's selector process conjecture: Maxima is p-independent if support of linear functions involved is not p-small.
- Contention resolution/Correlation Gap in Rounding schemes: Constant matters!

Further connections: An open invitation

Random Linear Programs:

- Max-Cut/Max-Bisection in $G(N, p)$: The max-cut value is $N(pN/4 + (1 + o_{pN}(1))P_*\sqrt{pN}/2)$ (Dembo-Montanari-Sen '17).
- Talagrand's selector process conjecture: Maxima is p-independent if support of linear functions involved is not p-small.
- Contention resolution/Correlation Gap in Rounding schemes: Constant matters!

Question: How does p-dependence to p-independence transition happen?

Further connections: An open invitation

Conjecture (Alon 1989)

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

An important step in this direction is the following:

Toy Conjecture

There is a two-coloring of $\mathbb{F}_2^d \setminus \{0\}$ such that there is no subspace of size Cd whose nonzero elements are monochromatic.

The trouble in small characteristic indicates interesting relationship with Ramsey theory, additive combinatorics.

Further connections: An open invitation

Conjecture (Alon 1989)

Consider a random Cayley graph with density p. The independence number is almost surely $\tilde{O}(p^{-1})$.

- Random Cayley graphs should have similar behaviors to random regular graphs.
- Relations to spectral graph theory, random matrix theory/suprema of processes.

Thank you!