Pseudorandom Graphs and the Green-Tao Theorem

Yufei Zhao

MIT

Based on joint work with David Conlon and Jacob Fox

SIAM Conference on Discrete Mathematics
Dénes König Prize Lecture
June 5, 2018

A progression of theorems on progressions

van der Waerden's theorem (1927)

If $\mathbb{N}$ is colored with finitely many colors, then there are arbitrarily long monochromatic arithmetic progressions (AP).

A progression of theorems on progressions

van der Waerden's theorem (1927)

If $\mathbb{N}$ is colored with finitely many colors, then there are arbitrarily long monochromatic arithmetic progressions (AP).

Erdős–Turán conjecture (1936)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

A progression of theorems on progressions

van der Waerden's theorem (1927)

If $\mathbb{N}$ is colored with finitely many colors, then there are arbitrarily long monochromatic arithmetic progressions (AP).

Erdős–Turán conjecture (1936)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

Roth's theorem (1953)

Every subset of $\mathbb{N}$ with positive density contains a 3-term AP.

A progression of theorems on progressions

van der Waerden's theorem (1927)

If $\mathbb{N}$ is colored with finitely many colors, then there are arbitrarily long monochromatic arithmetic progressions (AP).

Erdős–Turán conjecture (1936)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

Roth's theorem (1953)

Every subset of $\mathbb{N}$ with positive density contains a 3-term AP.

Szemerédi's theorem (1975)

Erdős–Turán conjecture is true.

Szemerédi's theorem (1975)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

(upper) density of $A \subset \mathbb{N}$ is $\limsup_{N\to\infty} \dfrac{|A \cap [N]|}{N}$ where $[N] := \{1, 2, \ldots, N\}$

Szemerédi's theorem (1975)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

(upper) density of $A \subset \mathbb{N}$ is $\limsup_{N\to\infty} \frac{|A \cap [N]|}{N}$ where $[N] := \{1, 2, \ldots, N\}$

Conjecture (Erdős 1973)

Every $A \subset \mathbb{N}$ with $\sum_{a\in A} 1/a = \infty$ contains arbitrarily long APs.

Szemerédi's theorem (1975)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

(upper) density of $A \subset \mathbb{N}$ is $\limsup_{N\to\infty} \dfrac{|A \cap [N]|}{N}$ where $[N] := \{1, 2, \ldots, N\}$

Conjecture (Erdős 1973)

Every $A \subset \mathbb{N}$ with $\sum_{a \in A} 1/a = \infty$ contains arbitrarily long APs.

Green–Tao theorem (2008)

The primes contain arbitrarily long APs.

Szemerédi's theorem (1975)

Every subset of $\mathbb{N}$ with positive density contains arbitrarily long APs.

(upper) density of $A \subset \mathbb{N}$ is $\limsup_{N\to\infty} \dfrac{|A \cap [N]|}{N}$ where $[N] := \{1, 2, \ldots, N\}$

Conjecture (Erdős 1973)

Every $A \subset \mathbb{N}$ with $\sum_{a\in A} 1/a = \infty$ contains arbitrarily long APs.

Green–Tao theorem (2008)

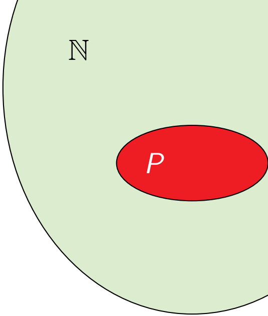
The primes contain arbitrarily long APs.

Prime number theorem: $\dfrac{\#\text{ primes up to } N}{N} \sim \dfrac{1}{\log N}$

Our main advance, then, lies not in our understanding of the primes but rather in what we can say about *arithmetic progressions*.

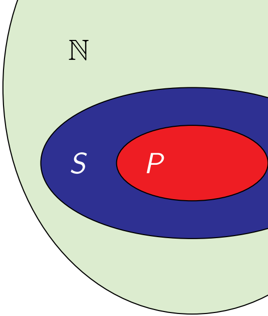
Ben Green
Clay Math Proceedings 2007

Proof strategy of Green–Tao theorem



P = prime numbers

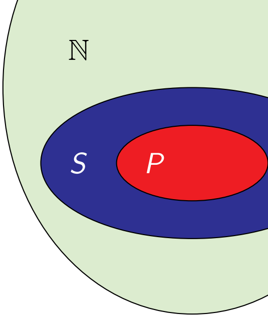
Proof strategy of Green–Tao theorem



$P =$ prime numbers, $S =$ "almost primes"

$P \subseteq S$ with positive relative density, i.e., $\dfrac{|P \cap [N]|}{|S \cap [N]|} > \delta$

Proof strategy of Green–Tao theorem



P = prime numbers, S = "almost primes"

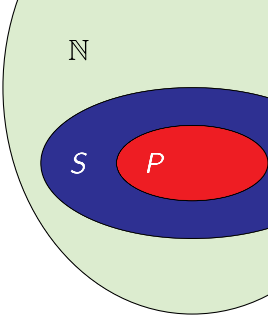
$P \subseteq S$ with positive relative density, i.e., $\dfrac{|P \cap [N]|}{|S \cap [N]|} > \delta$

Step 1:

Relative Szemerédi theorem (informally)

If $S \subset \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S with positive relative density contains long APs.

Proof strategy of Green–Tao theorem



P = prime numbers, S = "almost primes"

$P \subseteq S$ with positive relative density, i.e., $\dfrac{|P \cap [N]|}{|S \cap [N]|} > \delta$

Step 1:

Relative Szemerédi theorem (informally)

If $S \subset \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S with positive relative density contains long APs.

Step 2: Construct a superset of primes that satisfies the pseudorandomness conditions.

Relative Szemerédi theorem

Relative Szemerédi theorem (informally)

If $S \subset \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S with positive relative density contains long APs.

What pseudorandomness conditions?

Green–Tao:

1. Linear forms condition
2. Correlation condition

Relative Szemerédi theorem

Relative Szemerédi theorem (informally)

If $S \subset \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S with positive relative density contains long APs.

What pseudorandomness conditions?

Green–Tao:
1. Linear forms condition
2. Correlation condition

Question

Does relative Szemerédi theorem hold with weaker and more natural pseudorandomness hypotheses?

Relative Szemerédi theorem

Relative Szemerédi theorem (informally)

If $S \subset \mathbb{N}$ satisfies certain pseudorandomness conditions, then every subset of S with positive relative density contains long APs.

What pseudorandomness conditions?

Green–Tao:

1. Linear forms condition
2. Correlation condition $\leftarrow$ no longer needed

Question

Does relative Szemerédi theorem hold with weaker and more natural pseudorandomness hypotheses?

Theorem (Conlon–Fox–Z. '15)

Yes! A weaker linear forms condition suffices.

Relative Szemerédi theorem

k-AP-free: contains no k-term arithmetic progressions

Szemerédi's theorem (1975)

If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is k-AP-free, then $|A| = o(N)$.

Relative Szemerédi theorem (Conlon–Fox–Z.)

If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition,
and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

Earlier versions of relative Roth theorems with other pseudorandomness hypotheses:

Green, Green–Tao, Kohayakawa–Rödl–Schacht–Skokan

Relative Szemerédi theorem

k-AP-free: contains no k-term arithmetic progressions

Szemerédi's theorem (1975)

If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is k-AP-free, then $|A| = o(N)$.

Relative Szemerédi theorem (Conlon–Fox–Z.)

If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition,
and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

Earlier versions of relative Roth theorems with other pseudorandomness hypotheses:
Green, Green–Tao, Kohayakawa–Rödl–Schacht–Skokan

What does it mean for a set to be pseudorandom?
A: It resembles a random set in certain statistics

Pseudorandom graphs

In what ways can a graph look like a random graph?

Pseudorandom graphs

In what ways can a graph look like a random graph?

Fix a graph H. The H-density in a random graph with edge density p is $\approx p^{e(H)}$.

Pseudorandom graphs

In what ways can a graph look like a random graph?

Fix a graph H. The H-density in a random graph with edge density p is $\approx p^{e(H)}$.

A (sequence of) graph is pseudorandom if it satisfies some asymptotic properties, e.g., having asymptotically the same H-density as that of a typical random graph.

Pseudorandom graphs

In what ways can a graph look like a random graph?

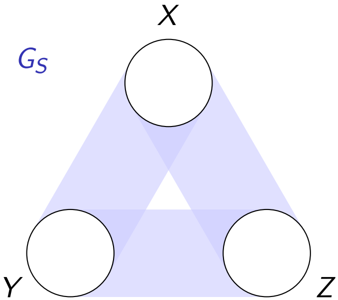
Fix a graph H. The H-density in a random graph with edge density p is $\approx p^{e(H)}$.

A (sequence of) graph is pseudorandom if it satisfies some asymptotic properties, e.g., having asymptotically the same H-density as that of a typical random graph.

Other ways that graphs can be pseudorandom: eigenvalues, edge discrepancy
Equivalent for dense graphs, but not for sparse graphs
(Thomason '87, Chung–Graham–Wilson '89)

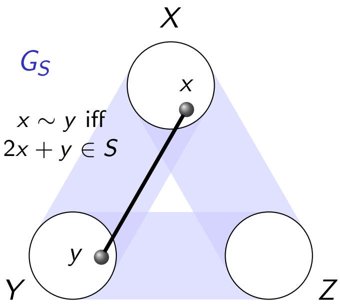
Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.



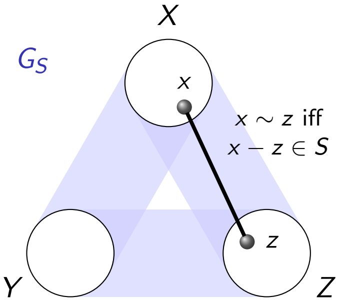
Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.



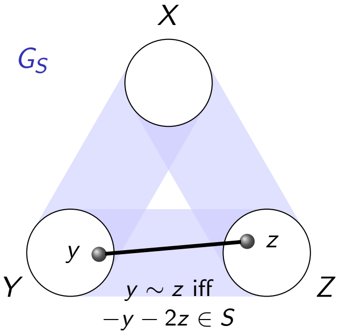
Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.



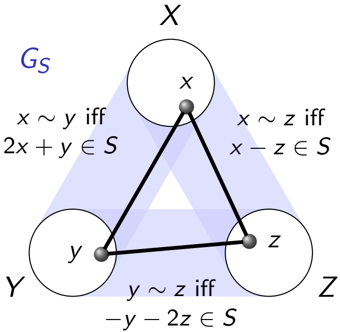
Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.



Graphs and 3-APs (3-term arithmetic progression)

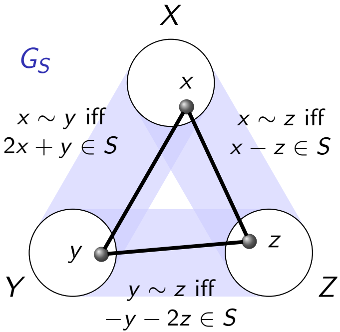
Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.



Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.

Triangle xyz in G_S $\iff$
$2x+y,\ x-z,\ -y-2z \in S$



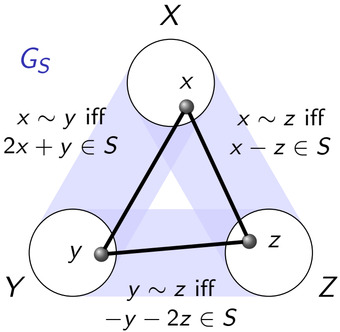
Graphs and 3-APs (3-term arithmetic progression)

Given $S \subseteq \mathbb{Z}/N\mathbb{Z}$, construct tripartite graph G_S with vertex sets $X = Y = Z = \mathbb{Z}/N\mathbb{Z}$.

Triangle xyz in G_S $\Longleftrightarrow$

$2x + y,\ x - z,\ -y - 2z \in S$

3-AP with common difference $-x - y - z$



Roth's theorem (1952)

If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon–Fox–Z.)

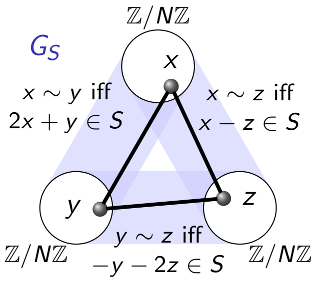
If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the 3-linear forms condition,
and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.

Roth's theorem (1952)

If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon–Fox–Z.)

If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the 3-linear forms condition,
and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.

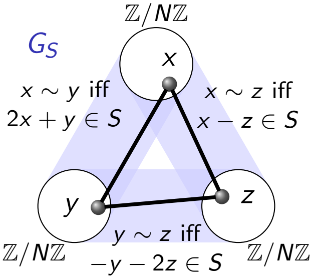


Roth's theorem (1952)

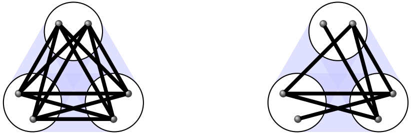
If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is 3-AP-free, then $|A| = o(N)$.

Relative Roth theorem (Conlon–Fox–Z.)

If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the 3-linear forms condition,
and $A \subseteq S$ is 3-AP-free, then $|A| = o(|S|)$.



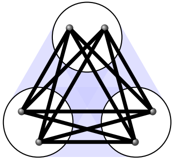
3-linear forms condition:
G_S has asymptotically the same H-density as a random graph for every $H \subseteq K_{2,2,2}$



3-linear forms condition

$S \subset \mathbb{Z}/N\mathbb{Z}$ satisfies the 3-linear forms condition if, for uniformly random $x_0, x_1, y_0, y_1, z_0, z_1 \in \mathbb{Z}/N\mathbb{Z}$, the probability that

$$\left\{ \begin{array}{lll} -y_0 - 2z_0, & x_0 - z_0, & 2x_0 + y_0, \\ -y_1 - 2z_0, & x_1 - z_0, & 2x_1 + y_0, \\ -y_0 - 2z_1, & x_0 - z_1, & 2x_0 + y_1, \\ -y_1 - 2z_1, & x_1 - z_1, & 2x_1 + y_1 \end{array} \right\} \subseteq S$$

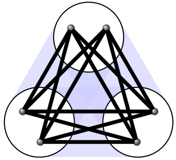


is with in $1 + o(1)$ factor of the expectation for a random S, and

3-linear forms condition

$S \subset \mathbb{Z}/N\mathbb{Z}$ satisfies the 3-linear forms condition if, for uniformly random $x_0, x_1, y_0, y_1, z_0, z_1 \in \mathbb{Z}/N\mathbb{Z}$, the probability that

$$\left\{ \begin{array}{lll} -y_0 - 2z_0, & x_0 - z_0, & 2x_0 + y_0, \\ -y_1 - 2z_0, & x_1 - z_0, & 2x_1 + y_0, \\ -y_0 - 2z_1, & x_0 - z_1, & 2x_0 + y_1, \\ -y_1 - 2z_1, & x_1 - z_1, & 2x_1 + y_1 \end{array} \right\} \subseteq S$$



is with in $1 + o(1)$ factor of the expectation for a random S, and the same is true if we erase any subset of the 12 patterns.

Relative Szemerédi theorem (Conlon–Fox–Z.)

Fix $k \geq 3$. If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

Relative Szemerédi theorem (Conlon–Fox–Z.)

Fix $k \geq 3$. If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

$k = 4$: build a 4-partite 3-uniform hypergraph

4-AP $\longleftrightarrow$ tetrahedron

Relative Szemerédi theorem (Conlon–Fox–Z.)

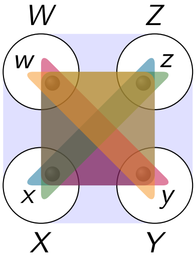
Fix $k \geq 3$. If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

$k = 4$: build a 4-partite 3-uniform hypergraph

4-AP $\longleftrightarrow$ tetrahedron

Vertex sets $W = X = Y = Z = \mathbb{Z}/N\mathbb{Z}$

$$
\begin{aligned}
wxy \in E &\iff 3w + 2x + y \in S \\
wxz \in E &\iff 2w + x - z \in S \\
wyz \in E &\iff w - y - 2z \in S \\
xyz \in E &\iff -x - 2y - 3z \in S
\end{aligned}
$$



4-AP with common diff: $-w - x - y - z$

Relative Szemerédi theorem (Conlon–Fox–Z.)

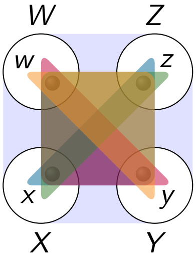
Fix $k \geq 3$. If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

$k = 4$: build a 4-partite 3-uniform hypergraph

4-AP $\longleftrightarrow$ tetrahedron

Vertex sets $W = X = Y = Z = \mathbb{Z}/N\mathbb{Z}$

$$
\begin{aligned}
wxy \in E &\iff 3w + 2x + y \in S \\
wxz \in E &\iff 2w + x - z \in S \\
wyz \in E &\iff w - y - 2z \in S \\
xyz \in E &\iff -x - 2y - 3z \in S
\end{aligned}
$$



4-AP with common diff: $-w - x - y - z$

4-linear forms condition: If H is a subgraph of the 2-blow-up of the tetrahedron, then the H-density in the above hypergraph is asymptotically same as random

Relative Szemerédi theorem (Conlon–Fox–Z.)

Fix $k \geq 3$. If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition, and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

4-linear forms condition: for uniform random $w_0, w_1, x_0, x_1, y_0, y_1, z_0, z_1 \in \mathbb{Z}/N\mathbb{Z}$, the probability that

$$\left\{ \begin{array}{llll} 3w_0 + 2x_0 + y_0, & 2w_0 + x_0 - z_0, & w_0 - y_0 - 2z_0, & -x_0 - 2y_0 - 3z_0, \\ 3w_0 + 2x_0 + y_1, & 2w_0 + x_0 - z_1, & w_0 - y_0 - 2z_1, & -x_0 - 2y_0 - 3z_1, \\ 3w_0 + 2x_1 + y_0, & 2w_0 + x_1 - z_0, & w_0 - y_1 - 2z_0, & -x_0 - 2y_1 - 3z_0, \\ 3w_0 + 2x_1 + y_1, & 2w_0 + x_1 - z_1, & w_0 - y_1 - 2z_1, & -x_0 - 2y_1 - 3z_1, \\ 3w_1 + 2x_0 + y_0, & 2w_1 + x_0 - z_0, & w_1 - y_0 - 2z_0, & -x_1 - 2y_0 - 3z_0, \\ 3w_1 + 2x_0 + y_1, & 2w_1 + x_0 - z_1, & w_1 - y_0 - 2z_1, & -x_1 - 2y_0 - 3z_1, \\ 3w_1 + 2x_1 + y_0, & 2w_1 + x_1 - z_0, & w_1 - y_1 - 2z_0, & -x_1 - 2y_1 - 3z_0, \\ 3w_1 + 2x_1 + y_1, & 2w_1 + x_1 - z_1, & w_1 - y_1 - 2z_1, & -x_1 - 2y_1 - 3z_1 \end{array} \right\} \subseteq S$$

is with in $1 + o(1)$ factor of the expectation for a random S, and the same is true if we erase any subset of the $2^3 \cdot 4 = 32$ patterns.

Roth's theorem: from one 3-AP to many 3-APs

Roth's theorem

Let $\delta > 0$, every $A \subset \mathbb{Z}/N\mathbb{Z}$ with $|A| \geq \delta N$ contains a 3-AP if N is sufficiently large.

By an averaging argument (Varnavides), we get many 3-APs:

Roth's theorem (counting version)

Every $A \subset \mathbb{Z}/N\mathbb{Z}$ with $|A| \geq \delta N$ contains $\geq c(\delta)N^2$ many 3-APs for some $c(\delta) > 0$.

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S, \qquad\qquad |A| \geq \delta\,|S|$$

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S, \qquad\qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good dense model $\widetilde{A}$ for A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S, \qquad\qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good dense model $\widetilde{A}$ for A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S, \qquad\qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good dense model $\widetilde{A}$ for A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

$$\geq cN^2 \qquad \text{[By Roth's Theorem]}$$

$\Longrightarrow$ relative Roth theorem (also works for k-AP)

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S, \qquad\qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good dense model $\widetilde{A}$ for A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

$$\geq cN^2 \qquad \text{[By Roth's Theorem]}$$

$\Longrightarrow$ relative Roth theorem (also works for k-AP)

Dense model

What does it mean for

$$\text{(dense)} \quad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}$$

to be a good approximation (dense model) of

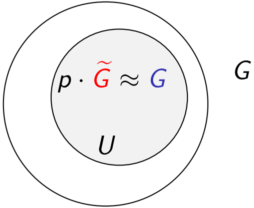
$$\text{(sparse)} \quad A \subset S \subset \mathbb{Z}/N\mathbb{Z} \ ?$$

Dense model

Let $\widetilde{G}$ (dense) and G (sparse) be two graphs on the same set of N vertices

We say that $\widetilde{G}$ is an good p-dense model of G if $p \cdot \widetilde{G} \approx G$ in terms of the number of edges when restricted to every vertex subset, i.e.,

$$\left| p \cdot e_{\widetilde{G}}(U) - e_G(U) \right| = o(pN^2) \qquad \forall U \subset V(G) = V(\widetilde{G})$$

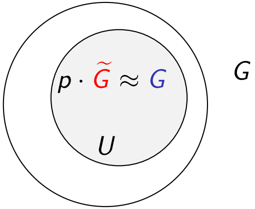


Dense model

Let $\widetilde{G}$ (dense) and G (sparse) be two graphs on the same set of N vertices

We say that $\widetilde{G}$ is an good p-dense model of G if $p \cdot \widetilde{G} \approx G$ in terms of the number of edges when restricted to every vertex subset, i.e.,

$$\left| p \cdot e_{\widetilde{G}}(U) - e_G(U) \right| = o(pN^2) \qquad \forall U \subset V(G) = V(\widetilde{G})$$



We say that $\widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}$ is a good p-dense model of $A \subset \mathbb{Z}/N\mathbb{Z}$ if $\mathrm{CayleySumGraph}(\mathbb{Z}/N\mathbb{Z}, \widetilde{A})$ is a good p-dense model of $\mathrm{CayleySumGraph}(\mathbb{Z}/N\mathbb{Z}, A)$

$\mathrm{CayleySumGraph}(G, A)$ has vertex set G, and $x \sim y$ iff $x + y \in A$

Dense model theorem

If $\mathbb{Z}/N\mathbb{Z}$ is a good p-dense model of $S \subset \mathbb{Z}/N\mathbb{Z}$ with $p = |S|/N$, then every $A \subset S$ has a good p-dense model $\widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}$.

Proof ideas: Hahn–Banach theorem/linear programming duality

Originally Green–Tao and Tao–Ziegler. Simplified by Gowers and Reingold–Trevisan–Tulsiani–Vadhan. Specialized to this form in Z.

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good p-dense model $\widetilde{A}$ of A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad\qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

$$\geq cN^2 \qquad \text{[By Roth's Theorem]}$$

$\Longrightarrow$ relative Roth theorem (also works for k-AP)

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S \subset \mathbb{Z}/N\mathbb{Z}, \qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good p-dense model $\widetilde{A}$ of A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

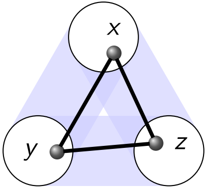
Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

$$\geq cN^2 \qquad \text{[By Roth's Theorem]}$$

$\Longrightarrow$ relative Roth theorem (also works for k-AP)

Counting lemma

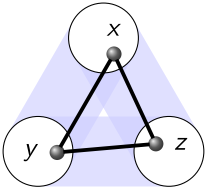


Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$

Counting lemma



Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$

Triangle counting lemma, sparse setting (Conlon–Fox–Z.)

(Sparse) $G \subset \Gamma$ and (dense) $\widetilde{G}$ are (tripartite) graphs on the same vertex set. Suppose

- "Sparse pseudorandom host graph" Γ has edge density p and satisfies the 3-linear forms condition (densities of $H \subset K_{2,2,2}$ are close to random)
- $\widetilde{G}$ is a good p-dense model of G

Then

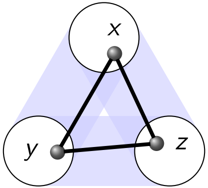
$$\text{triangle-density}(G) = p^3(\text{triangle-density}(\widetilde{G}) + o(1))$$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$

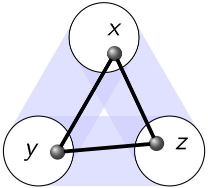


Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



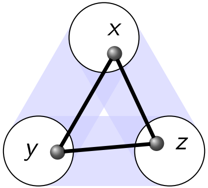
good 1-dense model: $\quad \left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



good 1-dense model: $\quad \left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

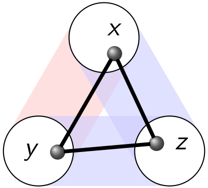
$$\text{triangle-density}(G) = \mathbb{E}[G(x,y)G(x,z)G(y,z)]$$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



good 1-dense model: $\left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

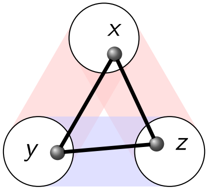
$$\begin{aligned}\text{triangle-density}(G) &= \mathbb{E}[G(x,y)G(x,z)G(y,z)] \\ &= \mathbb{E}[\widetilde{G}(x,y)G(x,z)G(y,z)] + o(1)\end{aligned}$$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



good 1-dense model: $\left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

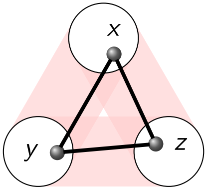
$$\begin{aligned}
\text{triangle-density(G)} &= \mathbb{E}[G(x,y)G(x,z)G(y,z)] \\
&= \mathbb{E}[\widetilde{G}(x,y)G(x,z)G(y,z)] + o(1) \\
&= \mathbb{E}[\widetilde{G}(x,y)\widetilde{G}(x,z)G(y,z)] + o(1)
\end{aligned}$$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



good 1-dense model: $\quad \left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

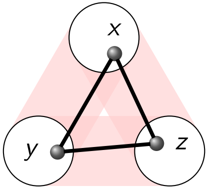
$$\begin{aligned}
\text{triangle-density}(G) &= \mathbb{E}[G(x,y)G(x,z)G(y,z)] \\
&= \mathbb{E}[\widetilde{G}(x,y)G(x,z)G(y,z)] + o(1) \\
&= \mathbb{E}[\widetilde{G}(x,y)\widetilde{G}(x,z)G(y,z)] + o(1) \\
&= \mathbb{E}[\widetilde{G}(x,y)\widetilde{G}(x,z)\widetilde{G}(y,z)] + o(1) = \text{triangle-density}(G) + o(1) \qquad \square
\end{aligned}$$

Counting lemma

Triangle counting lemma, dense setting

Let G and $\widetilde{G}$ be (tripartite) graphs on the same vertex set, such that $\widetilde{G}$ is a good 1-dense model of G. Then

$$\text{triangle-density}(G) = \text{triangle-density}(\widetilde{G}) + o(1)$$



good 1-dense model: $\quad \left|\mathbb{E}[(G(x,y) - \widetilde{G}(x,y))1_A(x)1_B(y)]\right| = o(1) \quad \forall A \subseteq X, B \subseteq Y$

$$\begin{aligned}
\text{triangle-density}(G) &= \mathbb{E}[G(x,y)G(x,z)G(y,z)] \\
&= \mathbb{E}[\widetilde{G}(x,y)G(x,z)G(y,z)] + o(1) \\
&= \mathbb{E}[\widetilde{G}(x,y)\widetilde{G}(x,z)G(y,z)] + o(1) \\
&= \mathbb{E}[\widetilde{G}(x,y)\widetilde{G}(x,z)\widetilde{G}(y,z)] + o(1) = \text{triangle-density}(G) + o(1) \qquad \square
\end{aligned}$$

Fails in the sparse setting (need $o(p^3)$ error)

Sparse counting lemma

Triangle counting lemma, sparse setting (Conlon–Fox–Z.)

(Sparse) $G \subset \Gamma$ and (dense) $\widetilde{G}$ are (tripartite) graphs on the same vertex set. Suppose

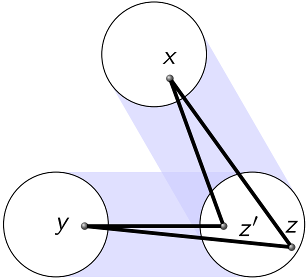
- "Sparse pseudorandom host graph" Γ has edge density p and satisfies the 3-linear forms condition (densities of $H \subset K_{2,2,2}$ are close to random)
- $\widetilde{G}$ is a good p-dense model of G

Then

$$\text{triangle-density}(G) = p^3(\text{triangle-density}(\widetilde{G}) + o(1))$$

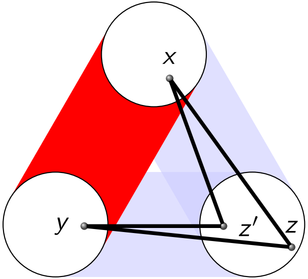
Key new proof ingredient: **densification**

Densification



$$\mathbb{E}[G(x,z)G(y,z)G(x,z')G(y,z')]$$

Densification

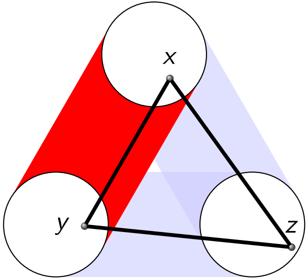


$$\mathbb{E}[G(x,z)G(y,z)G(x,z')G(y,z')] = \mathbb{E}[G'(x,y)G(x,z)G(y,z)]$$

Set $G'(x,y) := \mathrm{codeg}_G(x,y)/|Z|$

$G'(x,y) = O(p^2)$ for almost all pairs (x,y), and thus behaves like a dense weighted graph after scaling

Densification



$$\mathbb{E}[G(x,z)G(y,z)G(x,z')G(y,z')]$$
$$= \mathbb{E}[G'(x,y)G(x,z)G(y,z)]$$

Set $G'(x,y) := \mathrm{codeg}_G(x,y)/|Z|$

$G'(x,y) = O(p^2)$ for almost all pairs (x,y), and thus behaves like a dense weighted graph after scaling

Densified $G(X,Y)$. Now repeat for $G(X,Z)$ and $G(Y,Z)$.
Reduce to dense setting.

Transference

Let $S \subset \mathbb{Z}/N\mathbb{Z}$ be pseudorandom with density p, and

$$\text{(sparse)} \qquad A \subset S \subset \mathbb{Z}/N\mathbb{Z}, \qquad |A| \geq \delta\,|S|$$

Dense model theorem: One can find a good p-dense model $\widetilde{A}$ of A:

$$\text{(dense)} \qquad \widetilde{A} \subset \mathbb{Z}/N\mathbb{Z}, \qquad \frac{|\widetilde{A}|}{N} \approx \frac{|A|}{|S|} \geq \delta$$

Counting lemma:

$$\left(\frac{N}{|S|}\right)^3 |\{\text{3-APs in } A\}| \approx |\{\text{3-APs in } \widetilde{A}\}|$$

$$\geq cN^2 \qquad \text{[By Roth's Theorem]}$$

$\Longrightarrow$ relative Roth theorem (also works for k-AP)

Relative Szemerédi theorem

Szemerédi's theorem (1975)

If $A \subseteq \mathbb{Z}/N\mathbb{Z}$ is k-AP-free, then $|A| = o(N)$.

Relative Szemerédi theorem (Conlon–Fox–Z.)

If $S \subseteq \mathbb{Z}/N\mathbb{Z}$ satisfies the k-linear forms condition,
and $A \subseteq S$ is k-AP-free, then $|A| = o(|S|)$.

Green–Tao theorem

Every subset of the primes with positive relative density contains arbitrarily long APs.

Polynomial progressions in the primes

Polynomial Szemerédi theorem (Bergelson–Leibman 1996)

Every subset of $\mathbb{N}$ with positive density contains arbitrary polynomial progressions, i.e., for every $P_1, \ldots, P_k \in \mathbb{Z}[X]$ with $P_1(0) = \cdots = P_k(0) = 0$, the subset contains $x + P_1(y), \ldots, x + P_k(y)$ for some x and $y > 0$.

Polynomial Szemerédi theorem in the primes (Tao–Ziegler 2008)

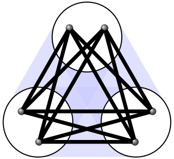
Every subset of the primes with positive relative density contains arbitrary polynomial progressions.

Using the **densification method**, Tao and Ziegler recently strengthened their result:

- (2015) existence of *narrow* progressions with polylogarithmic gaps
- (2018) asymptotics for the number of polynomial patterns in the primes

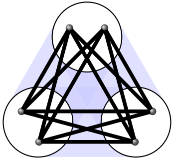
Some open problems

- Can the pseudorandomness hypotheses be further weakened?
- A multidimensional relative Szemerédi theorem?
  Linear forms conditions on $S \subset \mathbb{Z}/N\mathbb{Z}$ so that every relatively dense $A \subset S \times S$ contains a $k \times k$ square grid



Some open problems

- Can the pseudorandomness hypotheses be further weakened?
- A multidimensional relative Szemerédi theorem?
  Linear forms conditions on $S \subset \mathbb{Z}/N\mathbb{Z}$ so that every relatively dense $A \subset S \times S$ contains a $k \times k$ square grid



THANK YOU!