Additive Combinatorics Without Addition

Jacob Fox

Based on joint work with
David Conlon, Huy Tuan Pham, and Liana Yepremyan

ICMS workshop on Additive Combinatorics

July 26, 2024

Ramsey theory

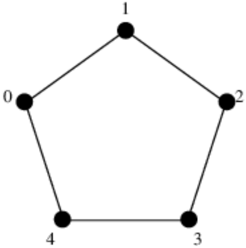
Ramsey theory contains many deep results which show that every very large structure contains a large well-organized substructure.

Ramsey's theorem guarantees that every very large graph contains a large clique or independent set.

Ramsey numbers

Definition (Ramsey number)

The *Ramsey number* $r(n)$ is the minimum N such that every graph on N vertices contains a clique or independent set of size n.



The 5-cycle has no clique or independent set of size 3.

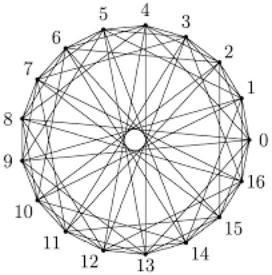
Every 6-vertex graph has a clique or independent set of size 3.

Hence, $r(3) = 6$.

Ramsey number

Definition (Ramsey number)

The *Ramsey number* $r(n)$ is the minimum N such that every graph on N vertices contains a clique or independent set of size n.



The Paley graph P_{17} has no clique or independent set of size 4. Furthermore, $r(4) = 18$.

Ramsey numbers

Definition (Ramsey number)

The *Ramsey number* $r(n)$ is the minimum N such that every graph on N vertices contains a clique or independent set of size n.

What about $r(5)$? $r(6)$?

Ramsey numbers

Definition (Ramsey number)

The *Ramsey number* $r(n)$ is the minimum N such that every graph on N vertices contains a clique or independent set of size n.

What about $r(5)$? $r(6)$?

Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack.

-Paul Erdős

Ramsey numbers

Definition (Ramsey number)

The *Ramsey number* $r(n)$ is the minimum N such that every graph on N vertices contains a clique or independent set of size n.

What about $r(5)$? $r(6)$?

Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack.

-Paul Erdős

$$43 \leq r(5) \leq 48.$$
$$102 \leq r(6) \leq 147.$$

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-Szekeres 1935)

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

Theorem (Campos-Griffiths-Morris-Sahasrabudhe 2023+)

There is $\varepsilon > 0$ such that there is no $(\frac{1}{2} + \varepsilon)$-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.

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.

Proof: Let $n = 2 \log_2 N$. Consider a random N-vertex graph G.

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.

Proof: Let $n = 2 \log_2 N$. Consider a random N-vertex graph G. The probability that a given set of n vertices is a clique is $2^{-\binom{n}{2}}$.

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.

Proof: Let $n = 2\log_2 N$. Consider a random N-vertex graph G. The probability that a given set of n vertices is a clique is $2^{-\binom{n}{2}}$.

$$\mathbb{P}[G \text{ is not 2-Ramsey}] \le \mathbb{E}[\# \text{ cliques or ind. sets of order } n]$$

$$= 2^{1-\binom{n}{2}}\binom{N}{n} = o(1).$$

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 (Campos-Griffiths-Morris-Sahasrabudhe 2023+)

There is $\varepsilon > 0$ such that there is no $(\frac{1}{2} + \varepsilon)$-Ramsey graph.

Theorem (Erdős 1947)

Almost all graphs on N vertices are 2-Ramsey.

Problem (Erdős)

Explicitly construct C-Ramsey graphs for some constant C.

Searching for hay in a haystack

Sven Sachsalber hunts for a needle in a haystack in a 2014 performance art piece. Photo: Palais de Tokyo, Paris.

Paley graphs

Definition (Paley graph)

For a prime $N \equiv 1 \pmod{4}$, the *Paley graph* P_N has vertex set $\mathbb{Z}_N$ and vertices $x \neq y$ are adjacent if $x - y$ is a quadratic residue.

Paley graphs

Definition (Paley graph)

For a prime $N \equiv 1 \pmod 4$, the *Paley graph* P_N has vertex set $\mathbb{Z}_N$ and vertices $x \neq y$ are adjacent if $x - y$ is a quadratic residue.

P_5, P_{17} give the tight lower bound on $r(3)$ and $r(4)$, respectively.

Paley graphs

Definition (Paley graph)

For a prime $N \equiv 1 \pmod 4$, the *Paley graph* P_N has vertex set $\mathbb{Z}_N$ and vertices $x \neq y$ are adjacent if $x - y$ is a quadratic residue.

P_5, P_{17} give the tight lower bound on $r(3)$ and $r(4)$, respectively.

Theorem (Montgomery 1972)

Assuming GRH, $\omega(P_N) \geq c \log N \log\log N$ for infinitely many N.

Theorem (Graham-Ringrose 1990)

$\omega(P_N) \geq c \log N \log\log\log N$ for infinitely many N.

Theorem (Hanson-Petrides, Di Benedetto-Solymosi-White 2021)

$\omega(P_N) \leq (\sqrt{2N-1}+1)/2$ for all N.

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

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

Conjecture (Alon 1989)

There is a constant C such that every finite group has a Cayley graph which is 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 $xy^{-1} \in S$.

Conjecture (Alon 1989)

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

Question

Are uniform random Cayley graphs Ramsey?

Clique number of random Cayley graphs

Clique number of random Cayley graphs

Theorem (Alon)

Asymptotically almost surely, the clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$.

Clique number of random Cayley graphs

Theorem (Alon)

Asymptotically almost surely, the clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$.

Theorem (Green 2005, Green-Morris 2016)

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

Clique number of random Cayley graphs

Theorem (Alon)

Asymptotically almost surely, the clique number of a uniform random Cayley graph on any group G of order N is $O(\log^2 N)$.

Theorem (Green 2005, Green-Morris 2016)

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

Theorem (Green 2005, Mrazović 2017)

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

Clique number of random Cayley graphs

Theorem

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

On Alon's conjecture

Theorem

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

Counting sets with small product set

Counting sets with small product set

$$AA^{-1} := \{ab^{-1} : a, b \in A\}.$$

Counting sets with small product set

$AA^{-1} := \{ab^{-1} : a, b \in A\}$.

$A \subset G$ is a clique in the Cayley graph G_S iff $AA^{-1} \setminus \{1\} \subset S$.

Counting sets with small product set

$AA^{-1} := \{ab^{-1} : a, b \in A\}$.

$A \subset G$ is a clique in the Cayley graph G_S iff $AA^{-1} \setminus \{1\} \subset S$.

Theorem

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

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

This naturally leads to studying a more general graph model:

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

This naturally leads to studying a more general graph model:

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

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

This naturally leads to studying a more general 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.

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

This naturally leads to studying a more general 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.

c is Δ-*bounded* if each color class has maximum degree $\leq \Delta$.

From additive combinatorics to edge-colored graphs

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 is 1 or 2-regular. A Cayley graph on G is the edge-union of some color classes.

This naturally leads to studying a more general 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.

c is Δ-*bounded* if each color class has maximum degree $\leq \Delta$.

What can we say about $\omega(G_c(p))$ if c is Δ-bounded?

Counting sets with few colors

Theorem

In a Δ-bounded edge-coloring of K_N, the number of n-vertex subsets with at most Kn colors is at most

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

Counting sets with few colors

Theorem

In a Δ-bounded edge-coloring of K_N, the number of n-vertex subsets with at most Kn colors is at most

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

By applying Vizing's theorem, we may assume $\Delta = 1$ (that is, the edge-coloring is proper).

Theorem

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

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

Improved communication through repetition

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

Improved communication through repetition

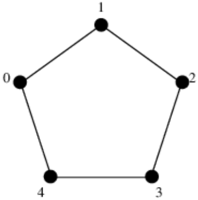
The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

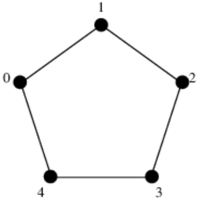


$\alpha(C_5) = 2$ realized by the independent set $\{0, 2\}$.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.



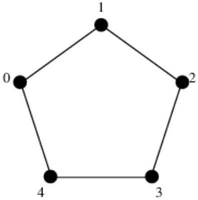
$\alpha(C_5) = 2$ realized by the independent set $\{0, 2\}$.

$\alpha(C_5^2) \geq \alpha(C_5)^2 = 4$ realized by the independent set $\{0, 2\}^2$.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.



$\alpha(C_5) = 2$ realized by the independent set $\{0, 2\}$.

$\alpha(C_5^2) \geq \alpha(C_5)^2 = 4$ realized by the independent set $\{0, 2\}^2$.

$\alpha(C_5^2) = 5$ given by the ind. set $\{(0,0), (1,2), (2,4), (3,1), (4,3)\}$.
More generally, if G is self-complementary, then $\alpha(G^2) \geq |G|$.
Indeed, $\{(x, \pi(x)) : x \in V(G)\}$ for π an isomorphism from G to its complement is an independent set in G^2.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

So $c_n(G) := \alpha(G^n)^{1/n}$ is the maximum number of messages per use of the channel in n uses.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

So $c_n(G) := \alpha(G^n)^{1/n}$ is the maximum number of messages per use of the channel in n uses.

$c(G) := \lim_{n\to\infty} c_n(G)$ is the Shannon capacity of the channel.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

So $c_n(G) := \alpha(G^n)^{1/n}$ is the maximum number of messages per use of the channel in n uses.

$c(G) := \lim_{n\to\infty} c_n(G)$ is the Shannon capacity of the channel.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

So $c_n(G) := \alpha(G^n)^{1/n}$ is the maximum number of messages per use of the channel in n uses.

$c(G) := \lim_{n\to\infty} c_n(G)$ is the Shannon capacity of the channel.

Alon and Orlitsky proved that there are self-complementary Ramsey graphs, so $c_1(G) = \Theta(\log|G|)$ and $c_2(G) \geq \sqrt{|G|}$.

Improved communication through repetition

The n^{th} power G^n of a graph $G = (V, E)$ has vertex set V^n and $(u, v) \in E(G^n)$ if $u \neq v$ and for each i, $u_i = v_i$ or $(u_i, v_i) \in E(G)$.

$\alpha(G^n)$ is the maximum number of messages a channel with confusion graph G can communicate without error in n uses.

So $c_n(G) := \alpha(G^n)^{1/n}$ is the maximum number of messages per use of the channel in n uses.

$c(G) := \lim_{n\to\infty} c_n(G)$ is the Shannon capacity of the channel.

Alon and Orlitsky proved that there are self-complementary Ramsey graphs, so $c_1(G) = \Theta(\log |G|)$ and $c_2(G) \geq \sqrt{|G|}$.

They made the following stronger conjecture, as it would give an analogous result for the Witsenhausen rate for dual source coding.

Conjecture (Alon and Orlitsky '95)

There exists self-complementary Ramsey Cayley graphs.

Going beyond uniform random

Going beyond uniform random

Let $N = 5^d$. A uniform random symmetric $S \subset \mathbb{F}_5^d$ 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.

Going beyond uniform random

Let $N = 5^d$. A uniform random symmetric $S \subset \mathbb{F}_5^d$ 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

There are self-complementary Ramsey Cayley graphs on $\mathbb{F}_5^d$.

Going beyond uniform random

Let $N = 5^d$. A uniform random symmetric $S \subset \mathbb{F}_5^d$ 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

There are self-complementary Ramsey Cayley graphs on $\mathbb{F}_5^d$.

Theorem

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

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

Going beyond uniform random

Let $N = 5^d$. A uniform random symmetric $S \subset \mathbb{F}_5^d$ 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

There are self-complementary Ramsey Cayley graphs on $\mathbb{F}_5^d$.

Theorem

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

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

It suffices to pick a random self-complementary Cayley graph on $\mathbb{F}_5^d$ in which each possible clique A has $|A-A| \geq |A|\log|A|$ and A has probability of being a clique at most $2^{-\Omega(|A-A|)}$.

Going beyond uniform random

Going beyond uniform random

Theorem

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

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. This guarantees:

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

(3) implies if A is a clique, then $|A + 2 \cdot A| = |A|^2$, so the Plünnecke-Ruzsa inequality implies

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

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

Theorem

Every finite vector space of characteristic at least five has a $(2 + o(1))$-Ramsey Cayley graph.

Theorem

Every finite vector space of characteristic $\equiv 1 \pmod 4$ has a self-complementary $(2 + o(1))$-Ramsey Cayley graph.

General random graph models

General random graph models

Definition: Δ-independent random graphs

Suppose a random graph G is such that each pair e of vertices appears as an edge of G with probability p_e, and appears independently of all edges apart from those in a graph G_e.
We say G is Δ-*independent* if $\Delta(G_e) \leq \Delta$ for each pair e.

General random graph models

Definition: Δ-independent random graphs

Suppose a random graph G is such that each pair e of vertices appears as an edge of G with probability p_e, and appears independently of all edges apart from those in a graph G_e.
We say G is Δ-*independent* if $\Delta(G_e) \leq \Delta$ for each pair e.

Examples: Erdős-Renyi random graphs, random Cayley graphs, random Latin square graphs, random entangled graphs, ...

General random graph models

Definition: Δ-independent random graphs

Suppose a random graph G is such that each pair e of vertices appears as an edge of G with probability p_e, and appears independently of all edges apart from those in a graph G_e.
We say G is Δ-*independent* if $\Delta(G_e) \leq \Delta$ for each pair e.

Examples: Erdős-Renyi random graphs, random Cayley graphs, random Latin square graphs, random entangled graphs, ...

Theorem

Suppose $0 < p < 1$ is fixed. Let G be a Δ-independent random graph on N vertices with $p_e = p$ for all pairs e.

1. If $\Delta = N^{o(1)}$, then $\omega(G) \geq (2 - o(1)) \log_{1/p} N$ a.a.s.
2. If $\Delta = O(1)$, then $\omega(G) \leq O(\log N \log\log N)$ a.a.s.

Conjectures

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.

Thank you