Corner-free sets in groups - A Part III essay

May 19, 2024

Contents

1 Introduction

This essay is concerned with sets which do not contain triples of the form $\{(x, y), (x + d, y), (x, y + d)\}$ for $x, y, d \in G$ for some abelian group G. We call such a triple a **corner**. The corner is **non-trivial** if $d \neq 0$. The main question is as follows.

Question 1. What is the maximum size $r_\llcorner(G)$ of a set $A \subseteq G \times G$ containing no non-trivial corner?

Erdős and Graham [14] first asked the similar question with *corners* replaced by *squares*, conjecturing that a square-free set in $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ must have vanishingly small density as $N \to \infty$. Ajtai and Szemerédi [1], unable to answer this question, proved the corresponding conjecture for corner-free sets. The motivation of Erdős and Graham was to find a higher dimensional analog to arithmetic progressions of length three after the following question had seen much progress by Roth [28].

Question 2. What is the maximum size $r_k(G)$ of a set $A \subseteq G$ containing no k-term arithmetic progression?

Roth proved that $r_3(\mathbb{Z}/N\mathbb{Z}) \ll N/\log\log N$ using a Fourier analytic argument. There was hope that the argument could be extended to longer arithmetic progressions and to higher dimensions, but all attempts were thwarted. Eventually, Szemerédi [32] proved his much celebrated theorem that $r_k(\mathbb{Z}/N\mathbb{Z}) = o(N)$, which was then extended to higher dimensions using ergodic theory by Furstenberg and Katznelson [15]. However all these proofs were basically ineffective, yielding no sensible bound on the rate at which the density goes to 0. The bounds became reasonable and the relationship to Fourier analyis clarified only after the groundbreaking work of Gowers introducing higher order Fourier analysis [16]. In this regard, k-term arithmetic progressions belong to the realm of $k - 2$ order Fourier analysis, squares to second order Fourier analysis, and corners to usual first order Fourier analysis. This justifies seeing corners as "3.5-term arithmetic progressions" and motivates Fourier analytic approaches to upper bounding $r_\llcorner(G)$. This is exactly what Shkredov [29] did, by heavily adapting Roth's original proof for 3-term arithmetic progressions to corners.

Questions 1 and 2 were first asked for subsets of $\{1, \dots, N\}$. Up to a constant factor, this is the same as considering $G = \mathbb{Z}/N\mathbb{Z}$. We call this the **integer case**. Perhaps surprisingly, those questions are much more easily answered when $G = \mathbb{F}_p^n$, a finite dimensional vector space over the finite field with p elements. We call this the **finite field case**. The main reason for this simplification is that $\mathbb{F}_p^n$ has many subspaces which look the same as itself, meaning that we can run iterative arguments that "zoom in" on subspaces. From this point of view, $\mathbb{Z}/N\mathbb{Z}$ is a rather generic group, and indeed our arguments in the integer case will work just as well for any finite abelian group G.

Conversely, the finite field case informs the integer case through an idea of Bourgain [7]: If we can't find many subspaces, maybe we can instead find many "approximate subspaces". These approximate subspaces are not closed under addition anymore, but only closed under addition of "small elements". With this in mind, one can attempt to replace every subspace appearing in a finite field case proof by an "approximate subspace" and more often than not this will yield an integer case proof. The cost is mostly paid in clarity as replacing algebraic objects by analytic objects induces a technical overhead.

More recently, links were found between sets avoiding some additive structure, like arithmetic progressions or corners, and communication complexity. In the Number On the Forehead (NOF) model, k participants want to compute $F(x_1, \dots, x_k)$ where $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0, 1\}$ with the least amount of communication, where participant i knows x_j for all $j \neq i$. Large corner-free sets in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ translate to efficient communication protocols for the NOF model when $F(x_1, x_2, x_3) = 1_{x_1+x_2+x_3=0}$. This motivated a series of improvements on large corner-free sets constructions, with the state of the art being Christandl, Fawzi, Ta, Zuiddam [9].

The general paradigm for finding large corner-free sets in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ is to first construct large corner-free sets in $\mathbb{F}_p^m \times \mathbb{F}_p^m$ for some small fixed m and then use the supermultiplicativity of corner-free sets ($r_\llcorner(G_1 \times G_2) \geq r_\llcorner(G_1) \times r_\llcorner(G_2)$). Christandl, Fawzi, Ta, Zuiddam [9] used a tool from algebraic complexity theory to reduce the construction of those large corner-free sets in $\mathbb{F}_p^m \times \mathbb{F}_p^m$ to an integer linear programming problem, making it tractable to computer optimisation.

The essay is structured as follows: Section 2 sets up basic notation. Section 3 quickly proves a non-trivial upper bound on $r_3(\mathbb{F}_p^n)$ due to Meshulam [26] and explains what breaks when naïvely applying the same method to corners. Section 4 patches the breakage using Shkredov's ideas to upper-bound corner-free sets in the finite field case. We will then describe in Section 5 how to extend the argument to the integer case. Finally, Section 6 explains the integer linear the recent lower bound on $r_{\llcorner}(\mathbb{F}_p^n)$ due to Christandl, Fawzi, Ta, Zuiddam [9].

2 Notation and basic definitions

Let's agree on some notation first.

Throughout this essay, p will be a fixed prime. $\mathbb{F}_p$ will be the finite field with p elements. We will use $\mathbb{F}_p^n$ as a placeholder for a vector field of dimension n over $\mathbb{F}_p$. We write $[n] = \{1, \dots, n\}$. G will be an ambient finite abelian group fixed by the context and $N = |G|$ will be its cardinality. Asymptotic statements will be taken as $N \to \infty$.

We denote $\widehat{G}$ the **Pontryagin dual** of G, namely the set of group homomorphisms from G to $\mathbb{C}^\times$, the unit complex numbers. We use the compact normalisation on physical space G and the discrete normalisation on Fourier space $\widehat{G}$. As such, the **convolution** and **difference convolution** of two functions $f, g : G \to \mathbb{C}$ are

$$\begin{aligned} f * g : G &\to \mathbb{C} & \qquad f \circ g : G &\to \mathbb{C} \\ x &\mapsto \mathop{\mathbb{E}}_{y+z=x} f(y)g(z) & \qquad x &\mapsto \mathop{\mathbb{E}}_{y-z=x} f(y)\overline{g(z)} \end{aligned}$$

and the **Fourier transform** of a function $f : G \to \mathbb{C}$ is

$$\begin{aligned} \hat{f} : \widehat{G} &\to \mathbb{C} \\ \gamma &\mapsto \mathop{\mathbb{E}}_{x} f(x)\gamma(x) \end{aligned}$$

For ease of notation, we will identify sets and their characteristic functions. For a set A,

$$A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

As we use the compact normalisation on physical space, we will talk much more about the density of sets rather than their cardinality. For two sets A, B, we define

$$\delta_B(A) = \delta(A) = |A \cap B| \,/\, |B|$$

Define the **balanced function** of A as

$$f_A = A - \delta(A)$$

Here is a list of facts we will use repeatedly.

$$\widehat{f * g} = \widehat{f}\widehat{g}$$

$$\widehat{f \circ g} = \widehat{f}\widehat{\overline{g}}$$

$$\sum_{\psi} \widehat{f}(\psi)\overline{\widehat{g}(\psi)} = \mathop{\mathbb{E}}_{x} f(x)\overline{g(x)} \tag{Plancherel}$$

$$\sum_{\psi} \left|\widehat{f}(\psi)\right|^2 = \mathop{\mathbb{E}}_{x} |f(x)|^2 \tag{Parseval}$$

$$\widehat{A}(0) = \delta(A)$$

$$\widehat{f_A}(0) = 0$$

A set A is ε-**uniform** if $\left|\widehat{A}(\psi)\right| \leq \varepsilon$ for all $\psi \neq 0$. This is equivalent to $\left\|\widehat{f_A}\right\| \leq \varepsilon$.

For a fixed natural number k, we call a set of the form $\{a, \dots, a+(k-1)d\}$ an **arithmetic progression of length** k, and denote these by kAP. Similarly to corners, we call a kAP **non-trivial** if $d \neq 0$.

3 Arithmetic progressions of length 3

This section sketches the proof of Meshulam's theorem [26], which is a finite field version of Roth's celebrated theorem.

Theorem 3.1 (Meshulam, 1995)**.** Let $A \subseteq G = \mathbb{F}_p^n$ be a set containing no non-trivial arithmetic progression of length 3. Then

$$|A| \ll N/\log N$$

Two examples of a set A are instructive to consider:

- $A = H$ is a subspace of $\mathbb{F}_p^n$. Then *any* $a, b \in H$ can be completed to a 3AP a, b, c in H since $c = 2b - a \in H$.
- $A = R$ is a set of density α picked uniformly at random. Then any element a lies in R with probability α, so any 3AP a, b, c lies in R with probability roughly α^3. Summing over all $|G|^2$ 3APs, the expectation of 3APs in R is $\alpha^3 |G|^2$. Of these, $\alpha |G|$ 3APs are trivial. Hence we have a non-trivial 3AP unless $\alpha^3 |G|^2 \ll \alpha |G|$, namely $|G| \ll \alpha^{-2}$.

These two cases are very different in nature, so the hope is that we can push *any* set A towards one of the two extremes. The key idea is that of a **density increment**. Either:

1. A is "random", ie it contains the same number of 3APs as a random set of the same density. Then either:
    - A contains a non-trivial 3AP.
    - G is very small ($|G| \ll \alpha^{-2}$).
2. There is a coset C of some subspace V of codimension 1 such that A is significantly denser on C than it was on G.

Simply put, every time our set displays some structure, we "zoom in" on that structure and see the same picture as before but with a denser set. The density increment acts as a ticking clock: Since the density is bounded above by 1 and the density increeases by some fixed positive quantity ε, we can't zoom in more than ε^{-1} times.

Let us prove one of the key steps: that a set is either random or has an uneven distribution along some subspace V. We already said we measured randomness of A as the number of 3APs in A compared to the expected number of 3APs of a random set of the same size. We will measure even distribution along cosets using the Fourier transform: If $\psi \neq 0$ and

$$\left|\widehat{A}(\psi)\right| = \left|\mathop{\mathbb{E}}_{x} A(x)\psi(x)\right|$$

is big, then the contributions of the different cosets of $\langle t\rangle^{\perp}$ to the expectation must be unequal, from which one can derive the density increment.

Lemma 3.2. [Randomness-structure dichotomy for 3APs] If $A \subseteq \mathbb{F}_p^n$ is of density α and $\left|\widehat{A}(t)\right| \leq \varepsilon$ for all $\psi \neq 0$, then the number of 3APs in A is $(\alpha^3 + o(1))\,|G|^2$.

Proof. Write

$$T_3 = \mathop{\mathbb{E}}_{x,d} A(x-d)A(x)A(x+d) = \langle A * A, 1_{2\cdot A}\rangle = \left\langle \widehat{A}^2, 1_{2\cdot A}\right\rangle$$

the normalised number of 3APs in A. We calculate

$$\left|T_3 - \alpha^3\right| = \left|\sum_{\psi \neq 0} \widehat{A}(\psi)^2 1_{2\cdot A}(\psi)\right| \leq \sup_{\psi \neq 0} 1_{2\cdot A}(\psi) \sum_{t} \left|\widehat{A}(\psi)\right|^2 \leq \varepsilon \left\|\widehat{A}\right\|_2^2 = \varepsilon\alpha^2$$

□

4 Shkredov's upper bound on corners

In this section, we prove an analog of Shkredov's upper bound in the case $G = \mathbb{F}_p^n$ following Ben Green's exposition [18].

Theorem 4.1 (Shkredov)**.** If $A \subseteq \mathbb{F}_p^n \times \mathbb{F}_p^n$ contains no non-trivial corners, then $|A| \ll N^2/(\log\log N)^{1/20}$.

One might hope that the same density increment strategy as for 3APs could work for corners. Unfortunately, we encounted a major hurdle: corners break our existing randomness-structure dichotomy.

Example 4.2. Consider a set $E \subseteq \mathbb{F}_p^n$ of density $\sqrt{\alpha} > 0$ picked uniformly at random. Then $A := E \times E$ has density α in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ and its expected number of corners is

$$\sum_{x,y,d} \mathbb{P}((x+d,y),(x,y),(x,y+d) \in E \times E) = \sum_{x,y,d} \mathbb{P}(x,y,x+d,y+d \in E) \approx \alpha^2 N^3$$

In comparison, the expected number of corners in a random set $A \in \mathbb{F}_p^n \times \mathbb{F}_p^n$ of density α is

$$\sum_{x,y,d} \mathbb{P}((x+d,y),(x,y),(x,y+d) \in A) \approx \alpha^3 N^3$$

Namely, there are sets A which are neither structured (in the sense that there is some subspace H of $\mathbb{F}_p^n$ such that A is denser on $H \times H$) nor random (in the sense that A has the expected number of corners).

However we remark two features of the troublesome set A:

- A contains around $\alpha^2 N^4$ rectangles, which is *more* than the expected $\alpha^3 N^4$. This suggests we should use some "rectangle-aware" norm.
- $A = E \times E$ does have increased density on some product set, namely $E \times E$ itself. This suggests we will have to be a bit looser in our definition of "structured".

To address the first point, we make the following definition.

Definition 4.3. For a set $S = E_1 \times E_2$, the **rectangle norm** of a function $f : S \to \mathbb{R}$ is given by

$$\|f\|_{\square}^4 = \mathop{\mathbb{E}}_{x,x' \in E_1, y,y' \in E_2} f(x,y)f(x,y')f(x',y)f(x',y')$$

$N^4 \|A\|_{\square}^4$ is the number of rectangles contained in A, hence $\|f_A\|_{\square}^4$ is a measure of how closely the number of rectangles in A is close to its expectation. The rectangle norm is a norm, but we won't need that fact.

To address the second point, one idea would be to call any product set $E_1 \times E_2$ structured. However this is too loose. $\|f_A\|_{\square}^4$ being small on an arbitrary product set doesn't imply that A has many corners, namely we can't establish what Green [18] and Tao call a *generalised von Neumann theorem* for such a wide class of structured sets.

We need a middle ground between the too sparse class of products of subspaces and too large class of products of arbitrary sets. Shkredov's insight is that this middle ground can be taken as the class of products of adequately uniform sets. The uniformity we impose depends on the density α of A. This is in notable contrast to the 3AP story where structured sets are the same regardless of α.

Definition 4.4. For $\alpha > 0$, a set S is **α-structured of codimension d and density β** if there exists a subspace $H \leq \mathbb{F}_p^n$ such that $\operatorname{codim} H = d$, $\delta_{H \times H}(S) = \beta$ and $S = E_1 \times E_2$ for some $2^{-14}\beta^6\alpha^{15}$-uniform subsets E_1, E_2 of H.

In what follows, we will not abbreviate "α-structured" to "structured" in order to draw the distinction between the informal idea of structure and its formal incarnation relevant specifically to corners.

Let's now state the randomness-structure dichotomy and apply it to get Theorem 4.1. We relegate its proof to the three subsections.

Proposition 4.5 (Randomness-structure dichotomy for corners)**.** Let S be an α-structured set of codimension d and density β. Suppose that $A \subseteq S$ has density α. Then either

- $\|f_A\|_\square^4 \leq 2^{-8}\alpha^{10}$ and A contains at least $\alpha^3\beta^2N^3/2$ corners.
- $\|f_A\|_\square^4 \geq 2^{-8}\alpha^{10}$ and there exists an α-structured set $S' \subseteq S$ of codimension at most $d + 2^{91}\alpha^{-70}\beta^{-13}$ and density at least $2^{-60}\alpha^{40}\beta$ such that A has density at least $\alpha + 2^{-31}\alpha^{20}$ on S'.

From there, deducing the end result is formally analogous to the 3AP case.

Proof of Theorem 4.1. Let k be maximal such that there exist sequences of sets $A_0, \dots, A_k$, α-structured sets $S_0, \dots, S_k$ with corresponding vector subspaces $H_0, \dots, H_k$ such that

1. A_i is a subset of density α_i of S_i
2. S_i is a subset of density β_i of H_i
3. A_i contains no non-trivial corner
4. $A_0 = A, S_0 = H_0 = \mathbb{F}_p^n$
5. $\alpha_{i+1} \geq \alpha_i + 2^{-31}\alpha_i^{20}$
6. $\beta_{i+1} \geq 2^{-60}\alpha_i^{40}\beta_i$
7. $\operatorname{codim} H_{i+1} \leq \operatorname{codim} H_i + 2^{91}\alpha_i^{-70}\beta_i^{-13}$

How large can k be? By induction,

$$\alpha_i \geq \alpha + 2^{-31}\alpha^{20}i$$

In particular, we get $\alpha_i \geq 2\alpha$ after $\lceil 2^{31}\alpha^{-19} \rceil$ steps. From there, another $\lceil 2^{31}(2\alpha)^{-19} \rceil$ steps bring us to $\alpha_i \geq 4\alpha$, etc. In general, the density is at least $2^r\alpha$ after

$$\sum_{i=0}^{r} \lceil 2^{31-19i}\alpha^{-19} \rceil$$

steps. Since the density is bounded above by 1, this means that $r \leq \log_2 \alpha^{-1}$ and

$$k \leq \sum_{i=0}^{\log_2 \alpha^{-1}} \lceil 2^{31-19i}\alpha^{-19} \rceil \leq \sum_{i=0}^{\infty} 2^{31-19i}\alpha^{-19} + \log_2 \alpha^{-1} \leq 2^{32}\alpha^{-19}$$

We can lower bound β_i using a simple induction:

$$\beta_i \geq 2^{-60i}\alpha^{40i}$$

By maximality, Proposition 4.5 applied to $A_k \subseteq S_k \subseteq H_k$ tells us that $\|f_{A_k}\|_\square^4 \leq 2^{-8}\alpha_k^{10}$ and A_k contains at least $\alpha_k^3\beta_k^2 |H_k|^3/2$ corners. By assumption, all those corners are trivial. So

$$|H_k|^2 \geq \alpha_k^3\beta_k^2 |H_k|^3/2 \geq 2^{-120i}\alpha^{3+80i} |H_k|^3$$

We can bound the size of H_k by computing

$$\operatorname{codim} H_{i+1} - \operatorname{codim} H_i \leq 2^{91}\alpha_i^{-70}\beta_i^{-13} \leq 2^{720i+91}\alpha^{-70-520i}$$

which gives

$$\operatorname{codim} H_k \leq \sum_{i=0}^{k} 2^{720i+91}\alpha^{-70-520i} \leq 2^{720k+92}\alpha^{-70-520k}$$

Hence

$$2^{121}\alpha^{-83} \geq |H_k| = p^{n-\operatorname{codim} H_k} \geq p^{n-2^{720k+92}\alpha^{-70-520k}}$$

We are now in the home-stretch. To finish, notice that one of the following holds.

- $\alpha^{-83} \gg p^{n/2} = N^{1/2}$. Then $\alpha \ll N^{-1/166} \ll (\log\log N)^{-1/20}$.

- $k \geq (\log\log N)^{19/20}$. Then $\alpha \ll (\log\log N)^{-1/20}$ since $k \ll \alpha^{-19}$.
- $\alpha^{-83} \ll p^{n/2}, k \leq (\log\log N)^{19/20}$. Then

  $$2^{720k+92}\alpha^{-70-520k} \geq n/2 = \log N/2$$

  So

  $$\alpha \ll (\log N)^{-1/520k} \leq e^{-(\log\log N)^{1/20}/520} \ll (\log\log N)^{-1/20}$$

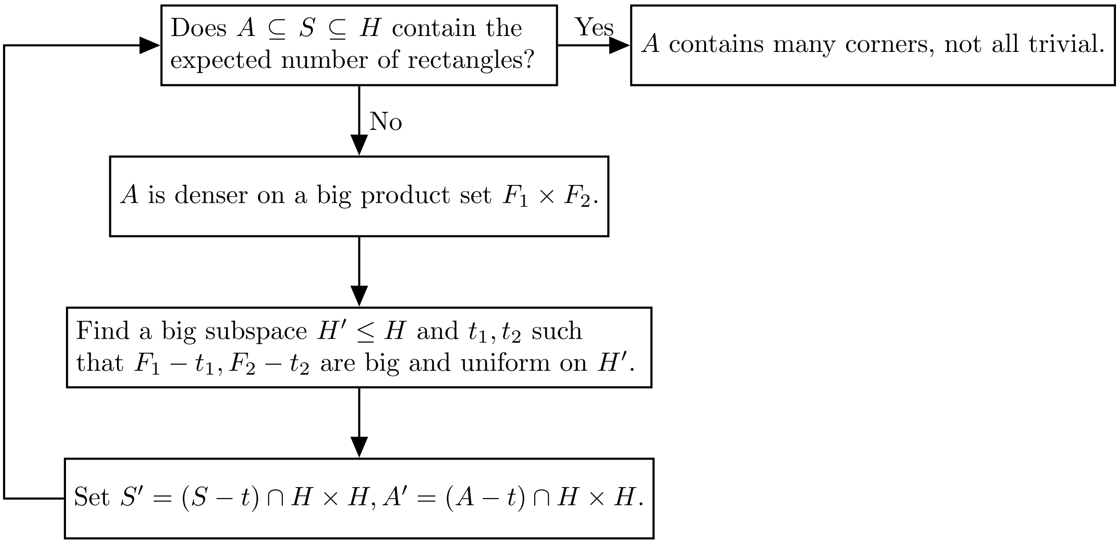
□

The following subsections prove the various parts of Proposition 4.5. Before doing so, let us remark one last feature of the proof.

We enlarged structured sets with some product sets in order to find a density increment, but restricted to sufficiently uniform sets in order to find a generalised von Neumann theorem. As such, the uniformity requirement will not be useful for us to establish the density increment. In fact, it will work against us: A moderately easy graph theoretic argument (Proposition 4.8) will show that a big rectangle norm implies a density increment on a **product set**. However we want a density increment on an **α-structured set**. We need to *uniformise* our product set to make it α-structured. This is what Proposition 4.12 does.

This uniformisation argument is interesting in that it itself requires an application of the iterative method. This makes the overall proof a doubly iterative method.

We now know the full sketch of the proof:



4.1 Generalised von Neumann theorem

In this subsection, we prove that uniform sets with large rectangle norm contain many corners.

Proposition 4.6 (Generalised von Neumann theorem)**.** Let S be an α-structured set of density β. Suppose that $A \subseteq S$ has density α and $\|f_A\|_\square^4 \leq 2^{-8}\alpha^{10}$. Then A contains at least $\alpha^3\beta^2N^3/2$ corners.

We first need a lemma telling us that most translates of a collection of uniform sets intersect in the expected number of elements.

Lemma 4.7. Let G be a finite group. Let $\tau, \kappa \in]0,1[$. Write $\varepsilon_k = \kappa 2^k \tau^{-k/2}$. If $S_1, \dots, S_k \subseteq G$ are ε_{k-1}-uniform sets of density $\sigma_1, \dots, \sigma_k$, then for all but a proportion of $k\tau$ tuples $(x_1, \dots, x_k) \in G^k$ we have

$$\left|\delta_{(S_1 - x_1)\cap\dots\cap(S_k - x_k)} - \sigma_1 \dots \sigma_k\right| \leq \varepsilon_k$$

Proof. Induct on k. The result is clear when $k = 0$. For $k+1$, write

$$\delta_{(S_1 - x_1)\cap\dots\cap(S_{k+1} - x_{k+1})} = \mathop{\mathbb{E}}_{d} \underbrace{S_1(x_1 + d)\dots S_k(x_k + d)}_{=:F_x(d)} S_{k+1}(x_{k+1} + d) = (S_{k+1} \circ F_x)(x_{k+1})$$

Parseval gives

$$\mathop{\mathbb{E}}_{x_{k+1}}(S_{k+1}\circ F_x)(x_{k+1})^2=\sum_{\psi}\left|\widehat{S_{k+1}\circ F_x}(\psi)\right|^2=\sigma_{k+1}^2\left(\mathop{\mathbb{E}}_{d}F_x(d)\right)^2+\sum_{\psi\neq 0}\left|\widehat{S_{k+1}}(\psi)\right|^2\left|\widehat{F_x}(\psi)\right|^2$$

Hence

$$\mathop{\mathbb{E}}_{x_{k+1}}\left((S_{k+1}\circ F_x)(x_{k+1})-\sigma_{k+1}\mathop{\mathbb{E}}_{d}F_x(d)\right)^2=\sum_{\psi\neq 0}\left|\widehat{S_{k+1}}(\psi)\right|^2\left|\widehat{F_x}(\psi)\right|^2\leq\varepsilon_k^2$$

By induction hypothesis, the proportion of tuples $(x_1,\dots,x_k)$ such that

$$\left|\mathop{\mathbb{E}}_{d}F_x(d)-\sigma_1\dots\sigma_k\right|\leq\varepsilon_k\tau$$

is at least $1-k\tau$. For these tuples,

$$\mathop{\mathbb{E}}_{x_{k+1}}((S_{k+1}\circ F_x)(x_{k+1})-\sigma_1\dots\sigma_{k+1})^2\leq(2+2\sigma_{k+1}^2\tau^2)\varepsilon_k^2\leq 4\varepsilon_k^2$$

and the proportion of "bad" x_{k+1} for which $|(S_{k+1}\circ F_x)(x_{k+1})-\sigma_1\dots\sigma_{k+1}|>\varepsilon_{k+1}$ is no more than

$$\frac{4\varepsilon_k^2}{\varepsilon_{k+1}^2}=\tau$$

Averaging over $x_1,\dots,x_k$, the total proportion of "bad" tuples $(x_1,\dots,x_{k+1})$ is at most $(k+1)\tau$, as wanted. □

Proof of Proposition 4.6. Write β the density of S and write S as $E_1\times E_2$ where E_1,E_2 are $2^{-14}\beta^6\alpha^{15}$-uniform subsets of some subgroup $H\leq\mathbb{F}_p^n$. All expectations in this proof will range over H. Write $\eta=2^{-8}\alpha^{10}$ and $f=f_A=A-\alpha$.

For functions $f_1,f_2,f_3:H\to\mathbb{R}$, write $T_{\llcorner}(f_1,f_2,f_3)=\mathbb{E}_{x,y,d}f_1(x+d,y)f_2(x,y)f_3(x,y+d)$. This operator is trilinear and $|H|^3\,T_{\llcorner}(A,A,A)$ is the number of corners in A. Write[1]

$$T_{\llcorner}(A,A,A)=T_{\llcorner}(A,f,A)+\alpha T_{\llcorner}(A,1,A)\geq\alpha T_{\llcorner}(A,1,A)-|T_{\llcorner}(A,f,A)|$$

We can lower-bound the first term as follows:

$$T_{\llcorner}(A,1,A)=\mathop{\mathbb{E}}_{x,y,d}A(x+d,y)A(x,y+d)=\mathop{\mathbb{E}}_{z}\left(\mathop{\mathbb{E}}_{s+t=z}A(s,t)\right)^2\geq\left(\mathop{\mathbb{E}}_{z,s+t=z}A(s,t)\right)^2=\alpha^2\beta^2$$

And now we want to show the second term is small in absolute value. This is where we use that $\|f\|_\square$ is small. To make the $\|f\|_\square^4$ term appear, we will use Cauchy-Schwarz twice to each time duplicate the number of appearances of f in the expression. For this to work, we need one variable to not appear as an argument of each function fed to T. A neat way to do this is to write corners in the form $(x,z-x),(x,y),(z-y,y)$. We get

$$T_{\llcorner}(A,f,A)=\mathop{\mathbb{E}}_{x,y,z}A(x,z-x)f(x,y)A(z-y,y)=\mathop{\mathbb{E}}_{y,z}A(z-y,y)E_1(z-y)\mathop{\mathbb{E}}_{x}f(x,y)A(x,z-x)$$

Hence Cauchy-Schwarz once gives

$$\begin{aligned}T_{\llcorner}(A,f,A)^2&\leq\left(\mathop{\mathbb{E}}_{y,z}A(z-y,y)\right)\mathop{\mathbb{E}}_{y,z}\left(E_1(z-y)\mathop{\mathbb{E}}_{x}f(x,y)A(x,z-x)\right)^2\\&=\alpha\beta\underbrace{\mathop{\mathbb{E}}_{x,x',y,z}E_1(z-y)f(x,y)A(x,z-x)f(x',y)A(x',z-x')}_{(*)}\end{aligned}$$

---

[1] It is tempting to try expanding $T_{\llcorner}(f,f,f)=T_{\llcorner}(A,A,A)+\dots$ and use $|T_{\llcorner}(f,f,f)|\leq\|f\|_\square^4$ just as we did in the 3AP case. However we don't have enough control over the cross terms $T_{\llcorner}(1,A,A),T_{\llcorner}(A,1,A),T_{\llcorner}(A,A,1)$, so we are forced to do this more delicate argument instead.

Cauchy-Schwarz once more gives

$$\begin{aligned}(*)^2 &= \left(\mathop{\mathbb{E}}_{x,x',z} A(x,z-x)A(x',z-x')E_2(z-x)E_2(z-x')\mathop{\mathbb{E}}_{y} f(x,y)f(x',y)\right)^2 \\ &\le \underbrace{\left(\mathop{\mathbb{E}}_{x,x',z} A(x,z-x)A(x',z-x')\right)}_{(\dagger)}\underbrace{\mathop{\mathbb{E}}_{x,x',y,y'} \omega(x,x',y,y')f(x,y)f(x,y')f(x',y)f(x',y')}_{(\ddagger)}\end{aligned}$$

where

$$\omega(x,x',y,y') = \mathop{\mathbb{E}}_{z} E_1(z-y)E_1(z-y')E_2(z-x)E_2(z-x')$$

$\omega(x,x',y,y')$ is the density of the intersection of four $2^{-2}\beta^6\eta^{3/2}$-uniform sets. Hence Lemma 4.7 tells us it is basically constant. Precisely, with $k=4, \tau = 2^{-2}\beta^4\eta, \kappa = 2^{-8}\beta^{12}\eta^3$, we get that $\left|\omega(x,x',y,y') - \beta^2\right| \le \beta^4\eta$ with probability at least $1-\beta^4\eta$. We can therefore approximate $(\ddagger)$ by $\beta^4 \left\|f\right\|_{\square}^4$ as such:[2]

$$2\beta^4\eta \ge \left|\mathop{\mathbb{E}}_{x,x',y,y'} (\omega(x,x',y,y')-\beta^2)f(x,y)f(x,y')f(x',y)f(x',y')\right| = \left|(\ddagger) - \beta^4\left\|f\right\|_{\square}^4\right| \ge |(\ddagger)| - \beta^4\eta$$

Further, Parseval and uniformity of E_1, E_2 tell us

$$(\dagger) \le \mathop{\mathbb{E}}_{x,x',z} E_1(x)E_2(z-x)E_1(x')E_2(z-x') = \mathop{\mathbb{E}}_{z}(E_1 * E_2)(z)^2 \le \beta^2 + 2^{-2}\beta^6\eta^{3/2} \le 2\beta^2$$

Together, these inequalities say

$$T_{\llcorner}(A,f,A)^4 \le 2\alpha^2\beta^4 \mathop{\mathbb{E}}_{x,x',y,y'} \omega(x,x',y,y')f(x,y)f(x,y')f(x',y)f(x',y') \le 6\alpha^2\beta^8\eta$$

Therefore,

$$T_{\llcorner}(A,A,A) \ge \alpha T_{\llcorner}(A,1,A) - |T_{\llcorner}(A,f,A)| \ge \alpha^3\beta^2 - 2\alpha^{1/2}\beta^2\eta^{1/4} \ge \alpha^3\beta^2/2$$

as wanted. □

4.2 Unstructured density increment

In this subsection, we prove that sets with many rectangles are denser on some product set.

Proposition 4.8 (Unstructured density increment)**.** *Let $S = E_1 \times E_2$ be a product set. Suppose that $A \subseteq S$ has density α in S and $\|f_A\|_{\square}^4 \ge 2^{-8}\alpha^{10}$. Then there exist $F_1 \subseteq E_1, F_2 \subseteq E_2$ of density at least $2^{-15}\alpha^{10}$ such that A has density at least $\alpha + 2^{-29}\alpha^{20}$ on $F_1 \times F_2$.*

Remark. This α^{20} term is directly related to the exponent of $\log\log N$ in Theorem 4.1. A bound of the form $\delta_{F_1\times F_2}(A) \ge \alpha + C\alpha^m$ gives a final bound in Theorem 4.1 of $\alpha \ll (\log\log N)^{-1/(m-1+c)}$ for all $c > 0$.

Shkredov's argument in [30] is Fourier-analytic. We instead present the simple double-counting argument from [17], stated graph theoretically.

Our graph of interest is the bipartite graph G on parts X, Y with edges A. The density of A within S becomes the edge density of G. More generally, the density of A within $X' \times Y' \subseteq S$ is $\delta(X', Y')$, the edge density between X' and Y'. Rectangles in A become $K_{2,2}$ in G. Our goal is to find large subsets $X' \subseteq X, Y' \subseteq Y$ such that $\delta(X', Y')$ is big. For brevity, we will write $\eta = 2^{-8}\alpha^{10}$.

The proof idea is to pick X', Y' to be the neighborhoods of some y, x whose degree is close to the expected value and maximising $\delta(X', Y')$. The expectation of $\delta(X', Y')$ over x and y is controlled by $\|A\|_{\square}^4$. What we have is a condition on $\|f_A\|_{\square}$, so we must relate it to $\|A\|_{\square}$. Both steps will work unless many $x \in X$ or many $y \in Y$ have degree far from the average. But in this case we have many vertices of unexpectedly high degree and can make X', Y' out of them.

---

[2] $\mathbb{E}_{x,x',y,y'}$ is an expectation over $H \times H$ while $\|f\|_{\square}^4$ is an expectation over $S \times S$. This is why there is an extra β^2 term appearing when expanding out.

Lemma 4.9 (Imbalanced case)**.** Let $\varepsilon_1, \varepsilon_2 > 0$. Suppose that a proportion ε_1 of vertices $x \in X$ are such that $|\delta(N(x)) - \alpha| > \varepsilon_2$. Then there exists a subset $X' \subseteq X$ of density at least $\min(\varepsilon_1, \varepsilon_2)/2$ such that $\delta(X', Y) \geq \alpha + \varepsilon_1\varepsilon_2/2$.

Proof. By the assumption, there is either a proportion of $\varepsilon_1/2$ vertices $x \in X$ such that $\delta(N(x)) > \alpha + \varepsilon_2$ or a proportion of $\varepsilon_1/2$ vertices $x \in X$ such that $\delta(N(x)) < \alpha - \varepsilon_2$.

In the first case, define X' to be the set of such vertices. Then $\delta(X', Y) \geq \alpha + \varepsilon_2 \geq \alpha + \varepsilon_1\varepsilon_2/2$.

In the second case, define X' to be the complement of the set of such vertices. Counting edges, we find

$$(\alpha - \varepsilon_2)(1 - \delta(X')) + \delta(X') \geq \alpha$$

Hence

$$\delta(X') \geq \frac{\varepsilon_2}{1 + \alpha - \varepsilon_2} \geq \frac{\varepsilon_2}{2}$$

By assumption, $\delta(X') \leq 1 - \varepsilon_1/2$, so

$$\delta(X', Y) \geq \frac{\alpha\,|X| - (\alpha - \varepsilon_2)(1 - |X'|)}{|X'|} = \alpha + (\delta(X')^{-1} - 1)\varepsilon_2 \geq \alpha + \frac{\varepsilon_1\varepsilon_2}{2}$$

□

Lemma 4.10 (Relating $\|f_A\|_\square$ and $\|A\|_\square$)**.** Suppose that $\|f_A\|_\square^4 \geq \eta$ and that the proportion of $x \in X$ such that $|\delta(N(x)) - \alpha| > 2^{-6}\eta$ is at most $2^{-6}\eta$. Then $\|A\|_\square^4 \geq \alpha^4 + \eta/2$.

Proof. In this proof, all expectations will be over X or Y. Expand

$$\|A\|_\square^4 = \|\alpha + f_A\|_\square^4 = \alpha^4 + \underbrace{\cdots}_{14\text{ terms}} + \|f_A\|_\square^4$$

Changing variables, we may write each of the 14 middle terms in the form

$$\alpha \mathop{\mathbb{E}}_{x,x',y,y'} f_A(x,y)g(x',y)h(x,y')$$

where $\|g\|_\infty, \|h\|_\infty \leq 1$. We bound each term using

$$\left|\mathop{\mathbb{E}}_{x,x',y,y'} f_A(x,y)g(x',y)h(x,y')\right| \leq \mathop{\mathbb{E}}_{x}\left|\mathop{\mathbb{E}}_{y} f_A(x,y)\right| = \mathop{\mathbb{E}}_{x}|\delta(N(x)) - \alpha| \leq 2^{-5}\eta$$

where the last inequality holds by the second assumption. Therefore

$$\|A\|_\square^4 \geq \alpha^4 + \|f_A\|_\square^4 - 14 \cdot 2^{-5}\alpha\eta \geq \alpha^4 + \eta/2$$

□

Lemma 4.11. Suppose that $\|A\|_\square^4 \geq \alpha^4 + \eta/2$ and that there's at most a proportion of $2^{-6}\eta$ of $x \in X$ such that $|\delta(N(x)) - \alpha| > 2^{-6}\eta$. Then there exist subsets $X' \subseteq X, Y' \subseteq Y$ of density at least $2^{-5}\eta$ such that $\delta(X', Y') \geq \alpha + \eta/8$.

Proof. For $x \in X, y \in Y$, write $e(x,y)$ the number of edges between $N(x)$ and $N(y)$. Observe that

$$|X|^{-1}\,|Y|^{-1} \mathop{\mathbb{E}}_{x,y} 1_{x\sim y}e(x,y) = \|A\|_\square^4 \geq \alpha^4 + \eta/2$$

Define X_0 to be the set of $x \in X$ and $|\delta(N(x)) - \alpha| \leq 2^{-8}\eta$ and Y_0 similarly. We want to restrict G to X_0 and Y_0. We compute

$$\begin{aligned}\mathop{\mathbb{E}}_{x,y} 1_{x\sim y}e(x,y) - \mathop{\mathbb{E}}_{x,y} X_0(x)Y_0(y)1_{x\sim y}e(x,y) &= \mathop{\mathbb{E}}_{x,y}(1 - X_0(x)Y_0(y))1_{x\sim y}e(x,y)\\ &\leq \mathop{\mathbb{E}}_{x,y}(X_0^c(x) + Y_0^c(y))\,|X|\,|Y| \leq 2^{-5}\eta\,|X|\,|Y|\end{aligned}$$

where the last inequality holds since $\delta(X_0^c), \delta(Y_0^c) \leq 2^{-6}\eta$ by assumption. Therefore

$$|X|^{-1}|Y|^{-1} \mathop{\mathbb{E}}_{x \in X_0, y \in Y_0, x \sim y} e(x,y) \geq \alpha^{-1}|X|^{-1}|Y|^{-1} \mathop{\mathbb{E}}_{x,y} X_0(x)Y_0(y)1_{x \sim y} e(x,y) \geq \alpha^4 + \eta/4$$

and we can find $x \in X_0, y \in Y_0$ such that $x \sim y$ and

$$|X|^{-1}|Y|^{-1} e(x,y) \geq \alpha^3 + \alpha^{-1}\eta/4$$

Since $x \in X_0$, we have

$$2^{-7}\eta \leq \alpha - 2^{-5}\eta \leq \delta(X') \leq \alpha + 2^{-5}\eta$$

and similarly for Y'. Hence X' and Y' are decently big and

$$\delta(X', Y') \geq \frac{\alpha^3 + \alpha^{-1}\eta/4}{(\alpha + 2^{-5}\eta)^2} \geq \alpha + \eta/8$$

□

Proof of Proposition 4.8. Suppose that the proportion of $x \in X$ such that $|\delta(N(x)) - \alpha| > 2^{-6}\eta$ is at least $2^{-6}\eta$. Then Lemma 4.9 gives us $F_1 \subseteq E_1$ of density at least $2^{-7}\eta$ such that the density of A on $F_1 \times E_2$ is at least $\alpha + 2^{-13}\eta^2$.

If instead the proportion of $y \in Y$ such that $|\delta(N(y)) - \alpha| > 2^{-6}\eta$ is at least $2^{-6}\eta$, swap X and Y and apply Lemma 4.9 again.

If both proportions are smaller than $2^{-6}\eta$, apply Lemmas 4.10 and 4.11 in succession to get $F_1 \subseteq E_1, F_2 \subseteq E_2$ of density at least $2^{-7}\eta$ such that the density of A on $F_1 \times F_2$ is at least $\alpha + \eta/8 \geq \alpha + 2^{-13}\eta^2$. □

4.3 Uniformisation

This subsection proves that any product set may be intersected with a large subspace of $G = \mathbb{F}_p^n$ on which it is uniform.

Proposition 4.12 (Uniformisation)**.** Let $\tau, \sigma, \alpha, \delta \in]0,1[$. Let $H \leq \mathbb{F}_p^n$ be a subspace.[3] Let $T = F_1 \times F_2$ be a product subset of $H \times H$ of density δ. Let $A \subseteq T$ be a subset of density $\alpha + \tau$ inside T. Then there exist $H' \leq H$ a subspace of codimension at most $16\sigma^{-2}\tau^{-1}\delta^{-1}$ and elements $t_1, t_2 \in H$ such that

1. $E_1' := (F_1 - t_1) \cap H', E_2' := (F_2 - t_2) \cap H'$ are σ-uniform subsets of H'
2. $S' := E_1' \times E_2'$ has density at least $\tau\delta/2$ inside $H' \times H'$
3. $A' := A - (t_1, t_2)$ has density at least $\alpha + \tau/4$ inside $S' \times S'$.

Following the 3AP case, we could try turning the non-uniformity of T on $H \times H$ into a density increment on a coset of $W \times W$ where W is some one-codimensional subspace of H. If we iterated this argument naïvely here, we would find a subspace $H' \leq H$ of small codimension on which T is uniform and dense. Namely we would achieve Properties 1 and 2, but not Property 3 as we might restrict T to a coset of H' containing little to none of A.

Since we don't know which coset of $W \times W$ will contain a density increment for A, we might as well look at all cosets simultaneously, iterate in each of them, and only at the end choose the coset with the biggest density increment for A.[4]

At step k, we will have a partition $\mathcal{C}_k$ of $H \times H$ into cosets of subspaces. In order to measure the density increment of T and ensure termination, we need a way to combine the densities $\delta_C(T)$ for $C \in \mathcal{C}_k$. In the 3AP case, the strategy amounted to maximising $\sup_{C \in \mathcal{C}_k} \delta_C(T)$, but now we need some way to measure the density increment even if T is already dense and uniform on some $C \in \mathcal{C}_k$. Therefore we will instead look at the L^2 average $\sum_{C \in \mathcal{C}_k} \delta_H(C)\delta_C(T)^2$.

Let's first show that we can increase this L^2 average on one coset, here assimilated with $\mathbb{F}_p^n$ itself.

---

[3] In [17], Green assumes a lower bound on $|H|$, but all that lower bound does is ensuring that H' is not the zero subspace. This doesn't actually matter for the rest of the proof.

[4] `https://www.koalastothemax.com` provides an evocative illustration of the iteration when $p = 2$ and $A =$ koala.

Lemma 4.13. Let $F \subseteq \mathbb{F}_p^n$ be a set of density δ. Suppose $\psi \neq 0$ is such that $\left|\widehat{F}(\psi)\right| \geq \sigma$. Let $W = \langle \psi \rangle^\perp$. Then

$$p^2 \mathop{\mathbb{E}}_x (F * W)(x)^2 \geq \delta^2 + \sigma^2$$

Remark. $p(F * W)(x)$ is the density of F on the coset $x + W$.

Proof. By Parseval,

$$\mathop{\mathbb{E}}_x (F * W)(x)^2 = \sum_\xi \left|\widehat{F}(\xi)\right|^2 \left|\widehat{W}(\xi)\right|^2 \geq \left|\widehat{F}(0)\right|^2 \left|\widehat{W}(0)\right|^2 + \left|\widehat{F}(\psi)\right|^2 \left|\widehat{W}(\psi)\right|^2 \geq p^{-2}(\delta^2 + \sigma^2)$$

□

We now want to run our recursive uniformisation procedure. We can't expect it to partition $H \times H$ into cosets of big subspaces such that S is uniform on *every* part. Indeed, this would require so many steps that some parts would be really small. Instead we might want to ensure uniformity on *most* parts, but that's impossible either since the way we get rid of a non-uniform part is by subdividing it into p^2 parts, only one of which is guaranteed to be a bit more uniform. One way out would be to weigh the cosets according to their size, which can neatly be described as partitioning $H \times H$ into cosets of big subspaces such that *most* of $H \times H$ lies in cosets on which T is uniform.

Definition 4.14. For a partition $C_i, i \in \mathcal{I}$ of $H \times H$ indexed by a set $\mathcal{I}$, we define the **measure** of a subset $\mathcal{J} \subseteq \mathcal{I}$ by

$$\operatorname{meas}(\mathcal{J}) = \sum_{i \in \mathcal{J}} \delta_{H \times H}(C_i)$$

Lemma 4.15. Let $\varepsilon, \sigma \in]0, 1[$. Let $H \leq \mathbb{F}_p^n$ be a subspace and let $F_1, F_2 \subseteq H$. Then there is a partition of $H \times H$ indexed by some set $\mathcal{I}$ into cosets $C_i = t_i + W_i \times W_i$ of subspaces $W_i \leq H$ of codimension at most $16\sigma^{-2}\varepsilon^{-1}$ such that, if we denote by $\mathcal{N}$ the set of $i \in \mathcal{I}$ for which $\delta_{C_i}(T) \geq \varepsilon/2$ and either $F_1 - t_{i,1}$ or $F_2 - t_{i,2}$ is not σ-uniform on W_i, then

$$\operatorname{meas}(\mathcal{N}) < \frac{\varepsilon}{4}$$

Remark. As expected from the proof sketch, A is completely irrelevant in this statement. It will come into play when we decide on a part of the partition.

Proof. We describe an algorithm. At step k of the algorithm, we have an index set $\mathcal{I}_k$ indexing a partition of $H \times H$ into cosets $C_i = t_i + W_i \times W_i$ of subspaces $W_i \leq H$ of codimension at most k. For each $i \in \mathcal{I}_k$, write

$$\gamma_i = \delta_{H \times H}(C_i), \delta_{i,1} = \delta_{t_{i,1} + W_i}(F_1), \delta_{i,2} = \delta_{t_{i,2} + W_i}(F_2), \delta_i = \delta_{C_i}(T) = \delta_{i,1}\delta_{i,2}$$

We classify the parts into three classes:

- If $\delta_i < \varepsilon/2$, then we say i is **expired**.
- If $\delta_i \geq \varepsilon/2$ and $F_1 - t_{i,1}, F_2 - t_{i,2}$ are both σ-uniform on W_i, then we say i is **uniform**.
- If $\delta_i \geq \varepsilon/2$ and one of $F_1 - t_{i,1}, F_2 - t_{i,2}$ is not σ-uniform on W_i, then we say i is **non-uniform**.

Denote by $\mathcal{E}_k, \mathcal{U}_k, \mathcal{N}_k$ the corresponding sets of indices. At $k = 0$, we start with the trivial partition of H in one part.

If the overall size of the non-uniform parts becomes small, say

$$\operatorname{meas}(\mathcal{N}_k) < \frac{\varepsilon}{4}$$

then we stop the algorithm at step k.

Else we go onto step $k + 1$, during which we will subdivide the cosets in $\mathcal{N}_k$ and leave the cosets in $\mathcal{E}_k$ and $\mathcal{U}_k$ unchanged. Since the algorithm hasn't stopped at step k, we have

$$\operatorname{meas}\{i \in \mathcal{I} \mid F_1 - t_{i,1} \ \sigma\text{-uniform on } W_i\} + \operatorname{meas}\{i \in \mathcal{I} \mid F_2 - t_{i,2} \ \sigma\text{-uniform on } W_i\} \geq \operatorname{meas}(\mathcal{N}_k) \geq \frac{\varepsilon}{4}$$

WLOG the first set has measure[5] at least $\varepsilon/8$. Call that set $\mathcal{N}_k'$. For each $i \in \mathcal{N}_k'$, use Lemma 4.13 to find a codimension one subspace $W_i' \leq W_i$ such that

$$p^2 \mathop{\mathbb{E}}_{x \in W_i} (F_1 * W_i')(x)^2 \geq \delta_{i,1}^2 + \sigma^2$$

Write

$$\begin{aligned} t_{i,1} + W_i &= (u_1 + W_i') \cup \cdots \cup (u_p + W_i') \\ t_{i,2} + W_i &= (v_1 + W_i') \cup \cdots \cup (v_p + W_i') \end{aligned}$$

Replace $i \in \mathcal{I}_k$, by (j_1, j_2) for $j_1, j_2 \in [p]$, with associated subspace $W_{(j_1,j_2)} := W_i'$ and coset

$$C_{(j_1,j_2)} := \underbrace{(u_{j_1}, v_{j_2})}_{=:t_{(j_1,j_2)}} + W_i' \times W_i'$$

Let's show the algorithm terminates after at most $16\sigma^{-2}\varepsilon^{-1}$ steps. We do so by looking at the sum of the L^2 averages of the densities of F_1 and F_2 over parts of the partition. For a partition of $H \times H$ indexed by $\mathcal{I}$, its **index** is

$$\operatorname{ind}(\mathcal{I}) = \sum_{i \in \mathcal{I}} \gamma_i (\delta_{i,1}^2 + \delta_{i,2}^2)$$

At step $k+1$, reusing our notation from the above paragraph, we write

$$\delta_{j_1,1}' = \delta_{u_{j_1} + W_i'}(F_1), \delta_{j_2,2}' = \delta_{v_{j_2} + W_i'}(F_2)$$

and compute

$$\begin{aligned} \mathop{\mathbb{E}}_{j_1} \delta_{j_1,1}'^2 &= p^2 \mathop{\mathbb{E}}_{x \in W_i} (F_1 * W_i')(x)^2 \geq \delta_{i,1}^2 + \sigma^2 \\ \mathop{\mathbb{E}}_{j_2} \delta_{j_2,2}'^2 &\geq \left(\mathop{\mathbb{E}}_{j_2} \delta_{j_2,2}' \right)^2 = \delta_{i,2}^2 \end{aligned}$$

Noticing that $\gamma_{(j_1,j_2)} = p^{-2}\gamma_i$ ($W_i' \times W_i' \leq W_i \times W_i$ with codimension 2), we get

$$\begin{aligned} \operatorname{ind}(\mathcal{I}_{k+1}) - \operatorname{ind}(\mathcal{I}_k) &= \sum_{i \in \mathcal{N}_k', j_1, j_2} \gamma_{(j_1,j_2)} (\delta_{j_1,1}'^2 + \delta_{j_2,2}'^2) - \sum_{i \in \mathcal{N}_k'} \gamma_i (\delta_{i,1}^2 + \delta_{i,2}^2) \\ &= \sum_{i \in \mathcal{N}_k'} \gamma_i \left(\mathop{\mathbb{E}}_{j_1} \delta_{j_1,1}'^2 + \mathop{\mathbb{E}}_{j_2} \delta_{j_2,2}'^2 - \delta_{i,1}^2 - \delta_{i,2}^2 \right) \\ &\geq \sigma^2 \operatorname{meas}(\mathcal{N}_k') \geq \sigma^2 \varepsilon/8 \end{aligned}$$

Since $0 \leq \operatorname{ind}(\mathcal{I}) \leq 2$ for all $\mathcal{I}$, the algorithm indeed terminates after at most $16\sigma^{-2}\varepsilon^{-1}$ steps, at which point all cosets in the partition have codimension at most $16\sigma^{-2}\varepsilon^{-1}$. ☐

We are now ready to use the information about the density of A across the parts of the partition provided by Lemma 4.15 to decide which part will become H'.

Proof of Proposition 4.12. Using Lemma 4.15 with $\varepsilon = \tau\delta$, find $\mathcal{I}$ a set indexing a partition into cosets C_i of subspaces $W_i \leq H$ of codimension at most $16\sigma^{-2}\tau^{-1}\delta^{-1}$ and such that

$$\operatorname{meas}(\mathcal{N}) < \frac{\tau\delta}{4}$$

---

[5] Since each part in $\mathcal{N}_k$ must have been subdivided once during each of the first k steps, we know that C_i for $i \in \mathcal{N}_k$ has codimension exactly $n - k$, which means that $|\mathcal{N}_k| \geq p^{2k}\varepsilon/4$. However, we won't need that information and will only use the bound on $\operatorname{meas}(\mathcal{N}_k)$.

where $\mathcal{N}$ is the set of $i \in \mathcal{I}$ for which $\delta_{C_i}(T) \geq \tau\delta/2$ and either $F_1 - t_{i,1}$ or $F_2 - t_{i,2}$ is not σ-uniform on W_i. Similarly to the proof of Lemma 4.15, write $\mathcal{E}$ the set of $i \in \mathcal{I}$ for which $\delta_{C_i}(T) < \tau\delta/2$ and $\mathcal{U} = (\mathcal{E} \cup \mathcal{N})^c$ the set of $i \in \mathcal{I}$ for which $\delta_{C_i}(T) \geq \tau\delta/2$ and $F_1 - t_{i,1}, F_2 - t_{i,2}$ are both σ-uniform on W_i. Define

$$\delta_i = \delta_{C_i}(T), \alpha_i = \delta_{C_i \cap T}(A)$$

Note that $\delta_{C_i}(A) = \delta_i\alpha_i$ and $\delta_{H\times H}(C_i \cap T) = \gamma_i\delta_i$.

First observe that

$$\sum_{i\in\mathcal{E}} \gamma_i\delta_i\alpha_i \leq \frac{\tau\delta}{2}\sum_{i\in\mathcal{E}}\gamma_i \leq \frac{\tau\delta}{2} \text{ and } \sum_{i\in\mathcal{N}}\gamma_i\delta_i\alpha_i \leq \operatorname{meas}(\mathcal{N}) < \frac{\tau\delta}{4}$$

Combined with

$$\sum_{i\in\mathcal{E}}\gamma_i\delta_i\alpha_i + \sum_{i\in\mathcal{N}}\gamma_i\delta_i\alpha_i + \sum_{i\in\mathcal{U}}\gamma_i\delta_i\alpha_i = \delta_{H\times H}(A) = (\alpha+\tau)\delta$$

we get

$$\sum_{i\in\mathcal{U}}\gamma_i\delta_i\alpha_i \geq (\alpha+\tau)\delta - \frac{\tau\delta}{2} - \frac{\tau\delta}{4} = \left(\alpha + \frac{\tau}{4}\right)\delta$$

Hence

$$\sum_{i\in\mathcal{I}}\gamma_i\delta_i 1_{i\in\mathcal{U}}\alpha_i = \sum_{i\in\mathcal{U}}\gamma_i\delta_i\alpha_i \geq \left(\alpha+\frac{\tau}{4}\right)\delta = \sum_{i\in\mathcal{I}}\gamma_i\delta_i\left(\alpha+\frac{\tau}{4}\right)$$

and there must be some $i \in \mathcal{I}$ for which $1_{i\in\mathcal{U}}\alpha_i \geq \alpha + \tau/4$. In particular, $i \in \mathcal{U}$. Take $H' = W_i, t = t_i$. We check that

1. $E_1' := (F_1 - t_1) \cap H', E_2' := (F_2 - t_2) \cap H'$ are σ-uniform subsets of H' since $i \in \mathcal{U}$
2. $S' := E_1' \times E_2'$ has density $\delta_i \geq \tau\delta/2$ on $H' \times H'$ since $i \in \mathcal{U}$
3. $A' := A - (t_1, t_2)$ has density at least $\alpha_i \geq \alpha + \tau/4$ on H'.

□

Note that we never made use of the fact that we were in two dimensions. The same uniformisation argument works for any dimension, up to constants.

5 Adapting Shkredov's upper bound to the integer setting

In this section, we sketch out how to translate the upper bound on corners in $\mathbb{F}_p^n$ described in Section 4 to the integer setting.

Theorem 5.1 (Shkredov)**.** If $A \subseteq \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ contains no non-trivial corners, then

$$|A| \ll N^2/(\log\log N)^{1/73}$$

We first explain what new definitions we need, before contrasting the main steps of the proof to their finite field counterparts.

5.1 New framework

In this subsection, we set up the new definitions needed for the integer case.

Following our program, we will replace every occurrence of a subspace in the finite field case proof by an "approximate subspace", which we will take to be a Bohr set.

Definition 5.2. Let $\Gamma \subseteq \widehat{G}$. The **Bohr set** of **frequencies** Γ and **width**[6] ρ is

$$B(\Gamma, \rho) = \{x \in G \mid \forall\gamma \in \Gamma, |\gamma(x) - 1| \leq \rho\}$$

$|\Gamma|$ is the **rank** of the Bohr set.

---

[6] Sometimes authors allow a different width for each frequency, but we will only ever need constant width Bohr sets here.

Remark. Shkredov further requires $|x| \leq N$ where N is an extra parameter. This is because he works with $G = \mathbb{Z}$. We spare ourselves this parameter by working in $\mathbb{Z}/N\mathbb{Z}$ instead.

The notion of a Bohr set being "closed under addition of small elements" is formalised as such.

Lemma 5.3. If $x \in B(\Gamma, \rho_1)$ and $y \in B(\Gamma, \rho_2)$, then

$$x + y \in B(\Gamma, \rho_1 + \rho_2)$$

We will use this simple fact repeatedly in the case where ρ_2 is much smaller than ρ_1.

Definition 5.4. We say $B(\Gamma, \rho_2)$ is an η**-attendant** of $B(\Gamma, \rho_1)$ if $\rho_2 \leq \eta\rho_1$.

Now every time we add two elements of G together in the proof, we need to make sure one of them comes from a Bohr set B and the other one from an η-attendant B' of B. For this notion to be useful, we also need $B + B'$ to be comparable in size to B.

Definition 5.5. Let $0 < \kappa < 1$ be an ambient parameter. A Bohr set B of rank d is κ**-regular** if, for all ρ' such that

$$|\rho' - \rho| < \frac{\kappa}{100d}\rho$$

we have

$$1 - \kappa \leq \frac{|B(\Gamma, \rho')|}{|B(\Gamma, \rho)|} \leq 1 + \kappa$$

Not all Bohr sets are regular, but they can all be scaled by a small amount to become regular.

Lemma 5.6 (Bohr set regularity)**.** For all Γ and ρ, there exists ρ' such that $\rho/2 < \rho' < \rho$ and $B(\Gamma, \rho')$ is regular.

For the rest of this section, **all Bohr sets will be assumed regular**. We will always precise what κ is in context.

One major difference between Bohr sets and subspaces is that we can't partition a Bohr set into cosets of smaller Bohr sets. Luckily, if B is a Bohr set and B' an η-attendant, we can cover B by copies of B' such that each element $b \in B$ lies within roughly the same number of copies. In this sense, η-attendants play the role of codimension 1 subspaces. The following lemma, which we will use repeatedly, incarnates that principle by saying that $B * B'$ is close to B in L^1.

Lemma 5.7. Let $0 < \kappa < 1$ be an ambient parameter. Let B be a Bohr set of rank d and B' a $\kappa/100d$-attendant of B. Then

$$\|B * B' - B\|_1 < 2\kappa$$

where the L^1 norm is taken over B.

Let's now try to find the analog of uniformity. One blatant issue with Bohr sets is that we can't directly talk about the Fourier transform of $f : B \to \mathbb{C}$ where B is a Bohr set. We could simply talk about the Fourier transform of f extended to a function $G \to \mathbb{C}$, but this would fail to capture the fact that within B we are only allowed to see additive structure on a small scale. We must therefore probe f with an η-attendant of B. Else we will never manage to produce uniform sets in the uniformisation procedure.

Definition 5.8. For an ambient Bohr set B and an η-attendant B' of B, a set A is ε**-Bohr-uniform** if

$$\delta_B\{b \mid \|\widehat{f_{A,b+B'}}\|_\infty \geq \varepsilon\} \leq \varepsilon, \quad \mathop{\mathbb{E}}_{b \in B} |\delta_{b+B'}(A) - \delta_B(A)|^2 \leq \alpha^2$$

where $f_{A,b+B'} = A \cap (b + B') - \delta_B(A)(b + B')$ is the **balanced function of** A **on** $b + B'$.

Structured sets are now defined analogously to the finite field case.

Definition 5.9. For $\alpha > 0$, a set S is α**-Bohr-structured of rank** d **and density** β if there exist Bohr sets B_1, B_2 such that $\operatorname{rank} B_1 = \operatorname{rank} B_2 = d$, $\delta_{B_1 \times B_2}(S) = \beta$ and $S = E_1 \times E_2$ for some $2^{-2000}\beta^{48}\alpha^{96}$-Bohr-uniform subsets E_1, E_2 of B_1, B_2 with $\eta = 2^{-10000}\beta^{192}\alpha^{384}d^{-1}$.

Similarly to uniformity, the rectangle norm captures more than just the small scale behavior, meaning that the failure of a set to be rectangularly uniform won't provide us with a good enough unstructured density argument. Hence we season our norm with some η-attendants.

Definition 5.10. For two ambient Bohr sets B_1, B_2 and η-attendants B_1' of B_1 and B_2' of B_2, the **Bohr rectangle norm** of a function $f : G \times G \to \mathbb{R}$ is defined by

$$\|f\|_{B_1 \times B_2}^4 = \mathop{\mathbb{E}}_{\substack{i \in B_1 \\ j \in B_2}} \mathop{\mathbb{E}}_{d} \mathop{\mathbb{E}}_{\substack{x,x' \in i+d+B_1' \\ y,y' \in j-d+B_2'}} f(x,y)f(x,y')f(x',y)f(x',y')$$

For a further ambient η-attendant B_1'' of B_1' and an ambient product subset $S = E_1 \times E_2 \subseteq B_1 \times B_2$ of density β, a set $A \subseteq S$ of density α is **ε-Bohr-rectangularly-uniform** if

$$\delta_{B_1}\{b \mid \|f_{A,S}\|_{b+B_1',B_2}^4 \geq \varepsilon\beta^2\} \leq \varepsilon$$

Warning. We now have mentioned four kinds of uniformity:

- ε-uniformity of subsets of H (one-dimensional)
- ε-Bohr-uniformity of subsets of H (one-dimensional)
- ε-rectangular-uniformity of subsets of $H \times H$ (two-dimensional)
- ε-Bohr-rectangular-uniformity of subsets of $H \times H$ (two-dimensional)

This might get a little confusing, but we will need the last three notions in the proof.

5.2 New proofs

In the previous subsection, we covered all the definitions that need changing to tackle the integer case. In this subsection, we will investigate the proofs. Let's scan through Section 4 to find uses of subspaces and physical space addition.

The generalised von Neumann theorem consumes a subspace. We need to replace that subspace with a Bohr set and track a bunch of η-attendants, but not much else changes. We can still prove that the size of the intersection of uniform sets concentrates around its mean (this was Lemma 4.7 for the finite field case and is Theorem 3.1 in [29]), which we again use to estimate the weight function that appears when expanding (a suitably modified version of) $T_\llcorner$ along one direction.

Proposition 5.11 (Generalised von Neumann theorem)**.** Let S be an α-Bohr-structured set of density β. Suppose that $A \subseteq S$ has density α and is $2^{-100}\alpha^{12}$-Bohr-rectangularly-uniform. Then A contains at least $2^{-8}\alpha^3\beta^2 |B_1'|^2 |B_2|$ corners.

On its own, our previous unstructured density increment (Proposition 4.8) works out just fine. It is a purely combinatorial statement that doesn't consume nor produce a subspace, and the rectangle norm does not mention addition. However, recall that we have sprinkled η-attendants in our definition of uniformity in order for the upcoming uniformisation argument to go through. We must therefore work harder to eliminate those η-attendants. In fact, we will work strictly harder since we will reduce the new density increment to Proposition 4.8.

Proposition 5.12 (Unstructured density increment)**.** Let S be an α-Bohr-structured set of density β. Suppose that $A \subseteq S$ has density α and is not $2^{-100}\alpha^{12}$-Bohr-rectangularly-uniform. Then there exist a Bohr set $\tilde{B}$, an element $t \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ and subsets $F_1 \subseteq E_1 \cap (t_1 + \tilde{B}), F_2 \subseteq E_2 \cap (t_2 + \tilde{B})$ such that, writing $T = F_1 \times F_2$,

$$\delta_{\tilde{B} \times \tilde{B}}(T) \geq 2^{-250}\alpha^{24}\beta$$

and

$$\delta_T(A) \geq \alpha + 2^{-500}\alpha^{37}$$

Remark. As opposed to Proposition 4.8, we do need that S is structured and not merely a product set.

Proof sketch. The idea is to pass to large subsets $\tilde{E}_1 \subseteq E_1, \tilde{E}_2 \subseteq E_2$ on which A is not just non-ε-Bohr-rectangularly-uniform but has actually big rectangle norm. Proposition 4.8 then applies.

To do so, we must exclude (many) pathological cases where the density of some set on a product set with some η-attendant is too small or too large. But in those cases, we (painfully) notice that our life is actually easier (not as hard) because we can directly (in a page or so) turn the imbalance into a density increment. ☐

Last but not least, we must translate uniformisation. Recall that, in the finite field case, we proved uniformisation (Proposition 4.12) by creating a partition of the ambient subgroup H into cosets of big subspaces, recursively splitting cosets on which S was not uniform, and at the end picking a coset depending on the density of A. As we said earlier, we can't partition a Bohr set into smaller Bohr sets. Instead, we will create a tree of products of Bohr cosets by replacing at each step a bunch of products of Bohr cosets on which A is not Bohr-rectangularly-uniform by a collection of η-attendants of those Bohr cosets. With this idea in mind, everything works out the same as in the finite field case, except that our L^2 average density loses $O(\kappa)$ at every step (due to applications of Lemma 5.7). This is where we get to set κ to something adequately small.

Proposition 5.13 (Uniformisation)**.** Let $\tau, \sigma, \alpha, \beta, \eta, \kappa \in]0,1[$ be such that $\kappa = 2^{-100}\sigma^3\tau^5\delta^5, \sigma \leq 2^{-100}\tau\delta, \eta \leq \kappa/100d$. Let $B = B(\Gamma, \rho)$ be a Bohr set of rank d and width ρ. Let $T = F_1 \times F_2$ be a product subset of $B \times B$ of density δ. Let $A \subseteq T$ be a subset of density $\alpha + \tau$ inside T. Then there exist a Bohr set B' of rank $d' \leq d + 2^{30}\sigma^{-3}\tau^{-5}\beta^{-5}$ and width $\rho' \geq (2^{-10}\eta)^{d'}\rho$ along with elements $t_1, t_2 \in B$ such that

1. $E'_1 := (F_1 - t_1) \cap B', E'_2 := (F_2 - t_2) \cap B'$ are σ-Bohr-uniform subsets of B'
2. $S' := E'_1 \times E'_2$ has density at least $\tau\delta/16$ inside $B' \times B'$
3. $A' := A - (t_1, t_2)$ has density at least $\alpha + \tau/16$ inside $S' \times S'$.

6 Lower bound on corners

In this section, we describe the recent method developed by Christandl, Fawzi, Ta, Zuiddam [9] to find lower bounds on corner-free sets in $\mathbb{F}_p^n$ and the connections to the NOF model.

We will first explain how large corner-free sets give rise to efficient communication protocols in a specific instance of the NOF model. Then we will translate the corners problem into an hypergraph Ramsey problem. Finally, we will use our newly gained perspective to describe Christandl, Fawzi, Ta and Zuiddam's main advances and give explicit lower bounds on $r_{\llcorner}(\mathbb{F}_2^n)$ and $r_{\llcorner}(\mathbb{F}_3^n)$.

6.1 The communication complexity perspective

Let's describe the communication problem relevant to corner-free sets.

Recall that in the NOF model we have k participants, k data $x_i \in \mathcal{X}_i$ and an objective function $F : \mathcal{X}_1 \times \cdots \times \mathcal{X}_k \to \{0, 1\}$. Participant i knows x_j for all $j \neq i$ (think of the datum x_i as being written on participant i's forehead). The participants agree on a strategy beforehead. Then, at time t, participant $t \bmod k$ broadcasts one bit of information. The goal is for all k participants to know the value of $F(x_1, \ldots, x_k)$ as quickly as possible.

The NOF model was introduced by Chandra, Furst and Lipton [8] to study branching programs. When $k = 2$, the model is equivalent to the standard two-party communication model. When $k \geq 3$ however, the model becomes much more powerful thanks to the shared information between players. A problem of great interest is to find some explicit objective function F for $k \geq 3$ such that randomised protocols are much more efficient than deterministic ones. The Eval function was put forward as a potential candidate. It is defined by

$$\mathrm{Eval}_G(x_1, x_2, x_3) = \begin{cases} 1 & \text{if } x_1 + x_2 + x_3 = 0 \\ 0 & \text{if } x_1 + x_2 + x_3 \neq 0 \end{cases}$$

for $x_1, x_2, x_3 \in G$. The communication problem associated to Eval_G in the NOF model can be solved by a *randomised* algorithm in $O(1)$. However we know very little about its *deterministic* communication complexity $D(\mathrm{Eval}_G)$. For example, the state of our knowledge for $G = \mathbb{F}_2^n$ is

$$\log \log n \ll D(\mathrm{Eval}_{\mathbb{F}_2^n}) \leq 0.24n + O(\log n)$$

The lower bound is due to Lacey and McClain [22] and the upper bound is given by Theorem 6.10.

Chandra, Furst and Lipton [8] were the first to notice that many deterministic communication complexity problems in the NOF model have a Ramsey theoretic interpretation. This is where it gets interesting to us. Large corner-free sets correspond to efficient deterministic communication protocols for the Eval problem.

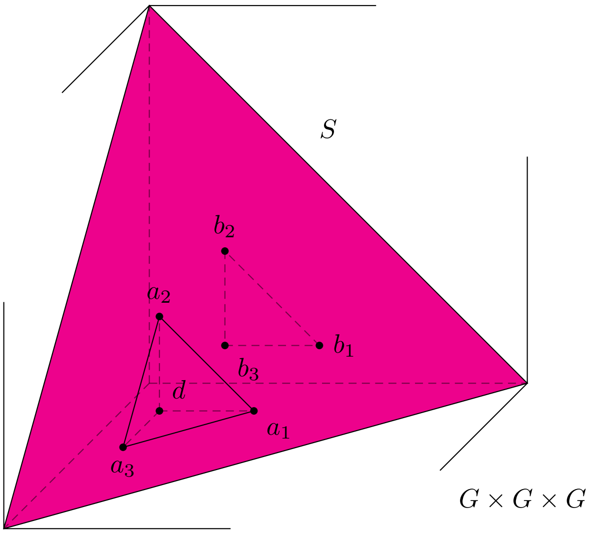


Figure 1: Part of the solution set S inside the cube $G \times G \times G$. (a_1, a_2, a_3) is a non-trivial forbidden pattern for d, with corresponding non-trivial corner (b_1, b_2, b_3).

Theorem 6.1 (Chandra, Furst, Lipton)**.** Let $\chi_\llcorner(G)$ be the number of colors required to color $G \times G$ such that no corner is monochromatic. Then the deterministic communication complexity of Eval_G is

$$D(\mathrm{Eval}_G) = \log_2 \chi_\llcorner(G) + O(1)$$

Remark. The same is true for more participants than just 3. In general, the communication complexity of the Eval_G problem for k participants is related to the number of colors required to color G^{k-1} in such a way that no *multidimensional corner* $(a_1 + d, \dots, a_{k-1} + d)$ is monochromatic.

Proof. We need a more symmetric onlook on the situation. Consider $G \times G \times G$ the space of all possible data (x_1, x_2, x_3). Within that space, there is the **solution hyperplane**

$$S = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 0\}$$

By definition, $\mathrm{Eval}(x) = 1 \iff x \in S$. We identify S with $G \times G$ under projection onto the first two coordinates. Under that identification, corners become what we will call **forbidden patterns**, namely tuples (a_1, a_2, a_3) such that there is $d \in G \times G \times G$ which, for each i, has the same coordinates as a_i except for the i-th one. Equivalently, $a_{i,k} = a_{j,k}$ for all i, j and all $k \neq i, j$. We say that (a_1, a_2, a_3) **is a forbidden pattern for** d, and that is **non-trivial** if the a_i are different.

First, let's prove that $D(\mathrm{Eval}_G) \geq \log_2 \chi_\llcorner(G)$. For $x \in G \times G \times G$, define $B_t(x)$ the sequence of bits the participants broadcast until time t when x_i is written on the forehead of participant i, $T(x)$ the time at which all participants know the value of $\mathrm{Eval}(x_1, x_2, x_3)$, and write $B_\infty(x) = B_{T(x)}(x)$.

Notice that if (a_1, a_2, a_3) is a forbidden pattern for d and $B_\infty(a_1) = B_\infty(a_2) = B_\infty(a_3)$, then $B_\infty(d)$ is also equal to them. Indeed, induction on t shows that $B_t(d) = B_t(a_t)$ (where the index of a is taken mod 3): If $B_t(d) = B_t(a_t)$, then participant t mod 3 can't distinguish between the broadcast history starting at d from the broadcast history starting at a_t and therefore broadcasts the same bit in either case, whence $B_{t+1}(d) = B_{t+1}(a_t) = B_{t+1}(a_{t+1})$. In particular, if we further have $a_1, a_2, a_3 \in S$, then $d \in S$ as well (assuming participants follow a communication protocol that correctly computes Eval) and $d = a_1 = a_2 = a_3$.

This means that B_∞ colors S in such a way that no non-trivial forbidden pattern is monochromatic. Therefore B_∞ takes at least $C_\llcorner(G)$ different values and there must be some x such that $T(x) \geq \log_2 \chi_\llcorner(G)$, as wanted.

Now, let's prove that $D(\mathrm{Eval}_G) \leq \log_2 \chi_\llcorner(G) + 4$. We describe an algorithm. Partition S into $\chi_\llcorner(G)$ colors such that no non-trivial forbidden pattern is monochromatic. Assume x_i is written on the forehead of participant i. Participants 1, 2, 3 respectively compute

$$a_1 = (-x_2 - x_3, x_2, x_3), \quad a_2 = (x_1, -x_3 - x_1, x_3), \quad a_3 = (x_1, x_2, -x_1 - x_2) \in S$$

From time $t = 1$ to time $t = \lceil \log_2 \chi_{\llcorner}(G) \rceil$, participant $t \bmod 3$ broadcasts the t-th bit of the color of a_t. The $\lceil \log_2 \chi_{\llcorner}(G) \rceil$ bits make up a *broadcasted color*. Then, until time $t = \lceil \log_2 \chi_{\llcorner}(G) \rceil + 3$, participant $t \bmod 3$ broadcasts 1 iff a_t is of the broadcasted color. The answer is then "Yes" iff the final 3 broadcasts were all 1. Since a_1, a_2, a_3 form a forbidden pattern, we have

$$\text{answer is "Yes"} \iff a_1, a_2, a_3 \text{ all have the same color} \iff a_1 = a_2 = a_3 \iff x \in S$$

Therefore the algorithm correctly computes $\mathrm{Eval}_G(x_1, x_2, x_3)$. ☐

Theorem 6.1 is stated in terms of $\chi_{\llcorner}(G)$ but we are interested in $r_{\llcorner}(G)$ instead. Luckily, those quantities are easily seen to be approximately the same.

Proposition 6.2 ([8])**.** Let G be a finite abelian group. Then

$$\frac{|G|^2}{r_{\llcorner}(G)} \leq c_{\llcorner}(G) \leq 2 \log |G| \, \frac{|G|^2}{r_{\llcorner}(G)}$$

Proof. On one hand, we clearly have $|G|^2 \leq r_{\llcorner}(G) c_{\llcorner}(G)$ since each element of $G \times G$ belongs to some corner-free set of size at most $r_{\llcorner}(G)$.

On the other hand, if A is a corner-free set of density $\alpha = r_{\llcorner}(G) / |G|^2$, then we can pick $m = 2 \log |G| / \alpha$ translates of A randomly and the expected number of elements not covered by any translate will be

$$|G|^2 (1 - \alpha)^m < 1$$

Namely, there is some collection of m translates of A that covers all of $G \times G$. So

$$c_{\llcorner}(G) \leq m = 2 \log |G| \, \frac{|G|^2}{r_{\llcorner}(G)}$$

☐

6.2 The Shannon capacity of hypergraphs perspective

How is avoiding corners an hypergraph Ramsey problem? And what does Shannon capacity have to do with it?

Let's first recall a few definitions which are perhaps not so standard in the context of hypergraphs.

Definition 6.3. $H = (V, E)$ is a **(directed) 3-uniform hypergraph** if E is a set of (ordered) triples of V. Elements of E are called **hyperedges** of H (or just **edges** for short).

An **independent set** in H is a set $A \subseteq V$ such that no hyperedge $e \in E$ is a subset of A (as an unordered triple). An **induced matching** is a triple of sets $A_1, A_2, A_3 \subseteq V$ such that (a_1, a_2, a_3) is not an hyperedge for any $a_1 \in A_1, a_2 \in A_2, a_3 \in A_3$. Note that A is an independent set iff A, A, A is an induced matching. We will therefore think of independent sets as symmetric induced matchings.

The **independence number** $\alpha(H)$ of H is the size of its largest independent set.

The **strong product** $H_1 \boxtimes H_2$ of two directed 3-uniform hypergraph $H_1 = (V_1, E_1), H_2 = (V_2, E_2)$ is the directed 3-uniform hypergraph on $V_1 \times V_2$ whose edges are those $((a_1, b_1), (a_2, b_2), (a_3, b_3))$ for which one of the following holds:

1. $a_1 = a_2 = a_3$ and $(b_1, b_2, b_3) \in V_2$.
2. $(a_1, a_2, a_3) \in V_1$ and $b_1 = b_2 = b_3$.
3. $(a_1, a_2, a_3) \in V_1$ and $(b_1, b_2, b_3) \in V_2$.

Denote by $H^{\boxtimes n}$ the strong product of H with itself n times.

We note that the independence number is *supermultiplicative*, in the sense that

$$\alpha(H_1)\alpha(H_2) \leq \alpha(H_1 \boxtimes H_2), \quad \alpha(H)^n \leq \alpha(H^{\boxtimes n})$$

since an independent set in H_1 can be combined with an independent set in H_2 to give an independent set in $H_1 \boxtimes H_2$. Fekete's lemma therefore tells us that $\lim_{n \to \infty} \alpha(H^{\boxtimes n})^{1/n}$ exists, is equal to

$\sup_{n\to\infty} \alpha(H^{\boxtimes n})^{1/n}$ and is also finite if V is since $\alpha(H^{\boxtimes n}) \leq |V|^n$. This quantity has a name in the literature.

Definition 6.4. The **Shannon capacity** of a directed hypergraph H is

$$\Theta(H) := \lim_{n\to\infty} \alpha(H^{\boxtimes n})^{1/n} = \sup_{n\to\infty} \alpha(H^{\boxtimes n})^{1/n}$$

We now can define the hypergraph relevant to corner-free sets.

Definition 6.5. For a finite abelian group G, define $H_{\llcorner}(G)$ to be the directed 3-uniform hypergraph on vertices $G \times G$ with hyperedges of the form

$$((x+d, y), (x, y), (x, y+d))$$

with $d \neq 0$.

Remark. Similarly, there is an hypergraph $H_3(G)$ relevant to 3APs.

By construction, $r_{\llcorner}(G) = \alpha(H_{\llcorner}(G))$ and $H_{\llcorner}(G_1 \times G_2) = H_{\llcorner}(G_1) \boxtimes H_{\llcorner}(G_2)$. Therefore we see that the asymptotic behavior of $r_{\llcorner}(G^n)$ is controlled by the Shannon capacity of $H_{\llcorner}(G)$.

Lemma 6.6. For G a finite abelian group, we have

$$r_{\llcorner}(G^n) = \Theta(H_{\llcorner}(G))^{n-o(n)}$$

Note that nothing really happened here. The translation from and to the world of hypergraphs is done without any significant loss[7]. We insist on staying in the world of hypergraphs in order to emphasize that the method applies to other problems (eg 3AP-free sets) as long as they can be interpreted as an hypergraph Ramsey problem.

6.3 Combinatorial degeneration

Now that we have reduced the search for large corner-free sets to lower-bounding the Shannon capacity of some hypergraph, let's describe Christandl, Fawzi, Ta and Zuiddam's machinery to produce such lower bounds.

We need a concept from algebraic complexity theory, referred to as *combinatorial degeneration*. It was introduced by Strassen in [31, Section 6] on his way to constructing fast matrix multiplication algorithms.

Definition 6.7. Let $I_1, \ldots, I_k$ be finite sets. Let $\phi \subseteq \psi I_1 \times \cdots \times I_k$ be two sets. We say ϕ is a **combinatorial degeneration** of ψ if there exist maps $u_i : I_i \to \mathbb{Z}$ such that

1. For all $x \in \phi$, we have $\sum_i u_i(x_i) = 0$
2. For all $x \in \psi \setminus \phi$, we have $\sum_i u_i(x_i) > 0$

Remark. ψ should be thought of as a universe of "possible outcomes" and ϕ as the "goal set" that we would like to land in.

Combinatorial degenerations have already been used in combinatorial contexts to build large *induced matchings* [3, 2, 10]. The insight of Christandl, Fawzi, Ta and Zuiddam is that they can actually be used to construct large *independent sets.*

Theorem 6.8 (Combinatorial degeneration method)**.** Let $H = (V, E)$ be a directed k-uniform hypergraph. Let $S \subseteq V$ be a subset of vertices such that

$$\phi = \{(v, \ldots, v) \mid v \in S\}$$

is a combinatorial degeneration of

$$\psi = E \cup \{(v, \ldots, v) \mid v \in V\}$$

Then $\theta(H) \geq |S|$.

---

[7] This is a bit of a lie, since we have actually lost a $o(n)$ term in the exponent in Lemma 6.6. Compare this to Sections 4 and 5 whose entire purpose was to bound that $o(n)$ term. This loss is justified by the fact that we are now in a very different regime: we don't even know what the base of the exponent is.

Remark. S should be regarded as being *almost* an independent set in H, from which we build *actual* independent sets in $H^{\boxtimes n}$. Precisely, the combinatorial degeneration from ψ to ϕ lets us carve an independent set out of S^n of size approaching $|S|^n$.

Proof. For a multiple n of $|S|$, call a tuple $x = (x_1, \dots, x_n) \in S^n$ **uniform** if every $s \in S$ appears exactly $n / |S|$ times in x.

Claim. Uniform tuples in S^n form an independent set of vertices of $H^{\boxtimes n}$.

Proof of the claim. By the combinatorial degeneration assumption, find maps $u_i : I_i \to \mathbb{Z}$ such that

$$\forall x \in \phi, \sum_i u_i(x_i) = 0, \quad \forall x \in \psi \setminus \phi, \sum_i u_i(x_i) > 0$$

Assume $x_1, \dots, x_k \in S^n$ are adjacent in $H^{\boxtimes n}$. Then $(x_{1,j}, \dots, x_{k,j}) \in \psi$ for all j. In particular,

$$\sum_i u_i(x_{i,j}) \geq 0$$

Since $x_1, \dots, x_k$ are not all equal, there also exists some $j_0 \in [n]$ such that $(x_{1,j_0}, \dots, x_{k,j_0}) \notin \{(v, \dots, v) \mid v \in V\} \supseteq \phi$, whence

$$0 < \sum_i u_i(x_{i,j_0}) \leq \sum_{i,j} u_i(x_{i,j})$$

But if $x_1, \dots, x_k$ were all uniform we would instead have

$$\sum_{i,j} u_i(x_{i,j}) = \sum_i \frac{n}{|S|} \sum_{s \in S} u_i(s) = \frac{n}{|S|} \sum_{s \in S} \sum_i u_i(s) = 0$$

since $(s, \dots, s) \in \phi$ for all $s \in S$. Contradiction. □

Now, the number of uniform tuples in S^n is

$$\binom{n}{\frac{n}{|S|}, \dots, \frac{n}{|S|}} \geq \frac{|S|^n}{(n+1)^{|S|}}$$

where the last inequality holds by

$$|S|^n = \sum_{a_1 + \dots + a_{|S|} = n} \binom{n}{a_1, \dots, a_{|S|}} \leq (n+1)^{|S|} \binom{n}{\frac{n}{|S|}, \dots, \frac{n}{|S|}}$$

by concavity of multinomial coefficients and the fact that the number of partitions of n into $|S|$ parts is at most $(n+1)^{|S|}$. Taking $n \to \infty$,

$$\theta(H) \geq \lim_{n \to \infty} \left(\frac{|S|^n}{(n+1)^{|S|}} \right)^{1/n} = |S|$$

□

Remark. Overall in this essay, the only point of using Fekete's lemma (through Lemma 6.6) is to be able to assume that $|S|$ divides n in the above proof. The proof otherwise directly translates to a construction of corner-free sets in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ for n any multiple of $|S|$.

This is all well and good, but combinatorial degenerations are difficult to construct by hand and we don't have a good understanding of them. Gladly, they can be encoded as a linear integer program.

Proposition 6.9 (Christandl, Fawzi, Ta, Zuiddam [9])**.** Let $H = (V, E)$ be a directed k-uniform hypergraph. The size of the largest set $S \subseteq V$ of vertices such that $\{(v, \dots, v) \mid v \in S\}$ is a combinatorial degeneration of $E \cup \{(v, \dots, v) \mid v \in S\}$ is the solution to the following linear integer program where $s : V \to \{0, 1\}$, $u_1, \dots, u_k : V \to \mathbb{Z}$ and $M \in \mathbb{N}$.

$$\boxed{\begin{array}{lrll} \text{max} & \sum_{v \in V} s(v) & & \\ \text{subject to} & u_1(v_1) + \dots + u_k(v_k) & \geq 1 & \forall (v_1, \dots, v_k) \in E, \\ & 1 - s(v) \leq u_1(v) + \dots + u_k(v) & \leq M(1 - s(v)) & \forall v \in V \end{array}} \tag{1}$$

Proof. On one hand, if $(s, u_1, \dots, u_k, M)$ is a feasible solution to the program, then $S := \{v \in V \mid s(v) = 1\}$ is so that $u_1, \dots, u_k$ form a combinatorial degeneration from $\{(v, \dots, v) \mid v \in S\}$ to $E \cup \{(v, \dots, v) \mid v \in V\}$. Indeed, if $v \in S$ then $s(v) = 1$ and

$$0 \le u_1(v) + \dots + u_k(v) \le 0$$

and if instead $(v_1, \dots, v_k) \in E \cup \{(v, \dots, v) \mid v \in V \setminus S\}$ then either the second or third condition gives

$$1 \le u_1(v_1) + \dots + u_k(v_k)$$

as wanted.

On the other hand, if $u_1, \dots u_k$ form a combinatorial degeneration from $\{(v, \dots, v) \mid v \in S\}$ to $E \cup \{(v, \dots, v) \mid v \in V\}$, then $(S, u_1, \dots, u_k, \max_v u_1(v) + \dots + u_k(v))$ is a feasible solution to the program by a similar argument. ☐

We can now run this program and find big explicit combinatorial degenerations.

Theorem 6.10 (Christandl, Fawzi, Ta, Zuiddam [9])**.** For the corner and Eval problems over $\mathbb{F}_2^n$, we have

- $\theta(H_\llcorner(\mathbb{F}_2)) \ge \sqrt[3]{39}$
- $r_\llcorner(\mathbb{F}_2^n) \ge \sqrt[3]{39}^{\,n-o(n)} \ge 3.39^{n-o(1)}$
- $D(\mathrm{Eval}_{\mathbb{F}_2^n}) \le 0.24n + o(n)$

For the corner and Eval problems over $\mathbb{F}_3^n$, we have

- $\theta(H_\llcorner(\mathbb{F}_3)) \ge 7$
- $r_\llcorner(\mathbb{F}_3^n) \ge 7^{n-o(n)}$
- $D(\mathrm{Eval}_{\mathbb{F}_3^n}) \le 0.37n + o(n)$

Proof. Apply Lemma 6.6 and Theorem 6.8 to the explicit combinatorial degenerations for $H_\llcorner(\mathbb{F}_2^3)$ and $H_\llcorner(\mathbb{F}_3)$ given in [9]. ☐

7 Conclusion

This essay covered a reasonable subset of the current state of the art in determining the size of the largest corner-free set in finite abelian groups. Let's take a few lessons home and mention avenues of research.

Very general results, like multidimensional Szemerédi or density Hales-Jewett, apply but yield completely intractable upper bounds. For corners, all reasonable upper bounds come from Fourier analysis and the iterative method.

Shkredov's main contribution here is an appropriate description of the structured sets showing up in the iteration. Since those structured sets are required to be uniform, a new proof step is needed in order to turn the usual density increment over an *arbitrary* set into a density increment over a *structured* set. This is what we have called *uniformisation*. The $\log \log N$ term in the final bound (Theorems 4.1 and 5.1) essentially comes from the fact that the density of our structured set inside its ambient subgroup decreases during the iteration. This doesn't happen with 3APs: A structured set is a subspace, hence has density 1 in ambient subspace. And indeed the bound in Meshulam's theorem (Theorem 3.1) has a $\log N$ term instead[8].

Shkredov's strategy can be tweaked in various ways. For example, Lacey and McClain [22] obtained a bound of the form $r_\llcorner(\mathbb{F}_2^n) \ll N^2 \log \log \log N / \log \log N$ (note the exponent of $\log \log N$) by taking into account the third natural "coordinate" of the problem, namely the diagonal $(x, y) \mapsto x + y$. Precisely, if we write $X \times^{\mathrm{diag}} Y = \{(x, x + y) \mid x \in X, y \in Y\}$, they replaced the structured sets of the form $E_1 \times E_2$ with E_1, E_2 uniform by structured sets of the form $E_1 \times E_2 \cap E_1 \times^{\mathrm{diag}} E_3$ with E_1, E_2, E_3 uniform.

---

[8] Roth's theorem states $r_3(\mathbb{Z}/N\mathbb{Z}) \ll N/\log \log N$, but this poorer bound compared to the finite field case does not have the same origin as in Shkredov's argument. It is instead due to the use of arithmetic progressions as structured sets. Replacing arithmetic progressions with Bohr sets yields a superior bound.

It however seems very difficult to translate the dramatic improvement due to Kelley and Meka [21, 6] from 3APs to corners: The Kelley-Meka improvement stems from replacing very elaborate Fourier space arguments (see eg Bloom [5]) by plainer physical space arguments, including almost periodicity and a dependent random choice. Fourier analysis is relegated to the initial bootstrapping step. In Shkredov's argument however, Fourier analysis is critical throughout since structured sets depend on the Fourier transform through their uniformity.

As underlined by Green [18], the finite field case is a fruitful ground for understanding Fourier-analytic techniques on our way to attacking the integer setting. We have seen this with Shkredov's argument which, when translated from its original integer setting to the finite field setting, goes from quite off-putting to somewhat palatable. A second reason that we have not seen so far is that it allows for algebraic techniques. These include the polynomial method of Croot-Lev-Pach [11] used by Ellenberg-Gijswijzt [13] to prove $r_3(\mathbb{F}_3^n) = o(2.756^n)$, but also slice-rank and subrank methods.

A second reason for the finite field case to be providing so many new ideas is that one rather sizeable community is interested in one specific finite field, namely computer scientists and $\mathbb{F}_2^n$. The communication complexity of evaluating a function with an additive interpretation in the NOF model can be rephrased as an additive combinatorics problem, eg efficient communication protocols for the Eval function correspond to large corner-free sets, and improvements have historically gone both ways.

Historically, all corner-free set constructions in $\mathbb{F}_p^n$ come from explicit corner-free sets in small dimension, and the $r_\llcorner(\mathbb{F}_2^n) \geq 3.39^{n-o(1)}$ bound from Christandl, Fawzi, Ta, Zuiddam [9] is no exception[9]. There is a natural barrier to this technique: since $\mathbb{F}_p^n \times \mathbb{F}_p^n$ itself is not corner-free, no explicit construction in small dimension will achieve $r_\llcorner(\mathbb{F}_p^n) \geq p^{2n-o(n)}$. Interestingly, Naslund [27] recently announced a bound of $r_\llcorner(\mathbb{F}_2^n) \geq 3.40^{n-o(1)}$ which he obtained via a recursive procedure. In particular it does *not* come from a small dimension construction.

Strong upper bounds of the form $r_\llcorner(\mathbb{F}_p^n) \leq c^{2n}$ for $c < p^2$ seem unattainable at the moment. Upper bounds for the Shannon capacity $\Theta(H)$ are very hard to come by. In fact there is essentially only one non-trivial graph whose Shannon capacity we know exactly, namely $\Theta(C_5) = \sqrt{5}$ thanks to a sublime argument of Lovász [24, 25]. One reason for this poor state of matters is that all techniques that are known to yield strong bounds on *independent sets* in hypergraph powers (polynomial method, slice-rank, subrank) work just as well to yield strong bounds on *induced matchings*. Their generality is their demise: $H_\llcorner(G^n)$ contains induced matchings of size $|G|^{2n-o(n)}$ [9]. We are in need of an algebraic technique that can tell independent sets and induced matchings apart.

One final topic that we omitted from this essay is lower bounds in the integer case. Behrend's construction [4] for 3APs, as reworked by Elkin [12], Green and Wolf [20], gives

$$r_\llcorner(\mathbb{Z}/N\mathbb{Z}) \geq 2^{-(c+o(1))\sqrt{\log_2 N}} N^2$$

where $c = 2\sqrt{2} \approx 2.828$. By constructing a better communication protocol for the Eval problem in the NOF model, Linial and Shraibman [23] achieved $c = 2\sqrt{\log_2 e} \approx 2.402$. Green [19] simplified their argument by eliminating the back-and-forth to communication complexity and achieved $c = 2\sqrt{2\log_2 \frac{4}{3}} \approx 1.822$.

References

[1] M. Ajtai and E. Szemerédi. Sets of lattice points that form no squares. *Studia Sci. Math. Hungar.*, 9:9–11, 1974.

[2] Josh Alman and Virginia Vassilevska Williams. Further limitations of the known approaches for matrix multiplication. In *9th Innovations in Theoretical Computer Science*, volume 94 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 25, 15. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2018.

[3] Noga Alon, Amir Shpilka, and Christopher Umans. On sunflowers and matrix multiplication. *Comput. Complexity*, 22(2):219–243, 2013.

---

[9] Technically, Theorem 6.10 does construct bigger and bigger corner-free sets in greater and greater dimension, but those corner-free sets are limited in size by another combinatorial object of small dimension instead, namely the set S involved in the combinatorial degeneration.

[4] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U.S.A.*, 32:331–332, 1946.

[5] Thomas F. Bloom and Olof Sisask. Breaking the logarithmic barrier in roth's theorem on arithmetic progressions, 2021.

[6] Thomas F. Bloom and Olof Sisask. The Kelley-Meka bounds for sets free of three-term arithmetic progressions. *Essent. Number Theory*, 2(1):15–44, 2023.

[7] J. Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.

[8] Ashok K. Chandra, Merrick L. Furst, and Richard J. Lipton. Multi-party protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, STOC '83, page 94–99, New York, NY, USA, 1983. Association for Computing Machinery.

[9] Matthias Christandl, Omar Fawzi, Hoang Ta, and Jeroen Zuiddam. Larger corner-free sets from combinatorial degenerations. In *13th Innovations in Theoretical Computer Science Conference*, volume 215 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages Art. No. 48, 20. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2022.

[10] Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal points in the asymptotic spectrum of tensors. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 289–296. ACM, New York, 2018.

[11] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Ann. of Math. (2)*, 185(1):331–337, 2017.

[12] Michael Elkin. An improved construction of progression-free sets. *Israel J. Math.*, 184:93–128, 2011.

[13] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017.

[14] P. Erdős. Problems and results on combinatorial number theory. In *A survey of combinatorial theory (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colo., 1971)*, pages 117–138. North-Holland, Amsterdam-London, 1973.

[15] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.*, 34:275–291, 1978.

[16] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[17] Ben Green. An argument of Shkredov in the finite field setting.

[18] Ben Green. Finite field models in additive combinatorics. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 1–27. Cambridge Univ. Press, Cambridge, 2005.

[19] Ben Green. Lower bounds for corner-free sets. *New Zealand J. Math.*, 51:1–2, 2021.

[20] Ben Green and Julia Wolf. A note on Elkin's improvement of Behrend's construction. In *Additive number theory*, pages 141–144. Springer, New York, 2010.

[21] Zander Kelley and Raghu Meka. Strong bounds for 3-progressions. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 933–973. IEEE Computer Soc., Los Alamitos, CA, [2023] ©2023.

[22] Michael T. Lacey and William McClain. On an argument of Shkredov on two-dimensional corners. *Online J. Anal. Comb.*, pages Art. 2, 21, 2007.

[23] Nati Linial and Adi Shraibman. Larger corner-free sets from better NOF exactly-N protocols. *Discrete Anal.*, pages Paper No. 19, 9, 2021.

[24] László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.

[25] Jiří Matoušek. *Thirty-three miniatures*, volume 53 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2010. Mathematical and algorithmic applications of linear algebra.

[26] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.

[27] Eric Naslund. Lower bounds for the capacity of hypergraphs. *In preparation*, 2024+.

[28] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[29] I. D. Shkredov. On a generalization of Szemerédi's theorem. *Proc. London Math. Soc. (3)*, 93(3):723–760, 2006.

[30] I. D. Shkredov. On a problem of Gowers. *Izv. Ross. Akad. Nauk Ser. Mat.*, 70(2):179–221, 2006.

[31] V. Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Math.*, 375/376:406–443, 1987.

[32] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.