On arithmetic progressions in symmetric sets in finite field model

Jan Hązła[*]

School of Computer and Communication Sciences
EPFL
Lausanne, Switzerland

jan.hazla@epfl.ch

Submitted: Dec 24, 2019; Accepted: Aug 4, 2020; Published: Sep 18, 2020

Abstract

We consider two problems regarding arithmetic progressions in symmetric sets in the finite field (product space) model.

First, we show that a symmetric set $S \subseteq \mathbb{Z}_q^n$ containing $|S| = \mu \cdot q^n$ elements must contain at least $\delta(q, \mu) \cdot q^n \cdot 2^n$ arithmetic progressions $x, x+d, \ldots, x+(q-1) \cdot d$ such that the difference d is restricted to lie in $\{0,1\}^n$.

Second, we show that for prime p a symmetric set $S \subseteq \mathbb{F}_p^n$ with $|S| = \mu \cdot p^n$ elements contains at least $\mu^{C(p)} \cdot p^{2n}$ arithmetic progressions of length p. This establishes that the qualitative behavior of longer arithmetic progressions in symmetric sets is the same as for progressions of length three.

Mathematics Subject Classifications: 11B25

1 Introduction

In this paper we consider problems in the finite field model in additive combinatorics. This model has been a fruitful area of research, originally considered as a "playground" for classical problems over integers, but subsequently becoming a source of many results that are interesting on their own. The reader can consult two surveys [Gre05a, Wol15] that are removed in time by ten years.

The most well-known problem in this setting concerns arithmetic progressions: Given a subset $S \subseteq \mathbb{Z}_q^n$ with density $\mu(S) := |S|/q^n$, what are the bounds on the number of arithmetic progressions of length k contained in S? The case $q = k = 3$ is called the

[*] Supported by awards ONR N00014-16-1-2227, NSF CCF-1665252 and ARO MURI W911NF-12-1-0509.

https://doi.org/10.37236/9242

capset problem. There, it has long been known [Rot53, Mes95] that any subset of $\mathbb{F}_3^n$ of constant density must contain an arithmetic progression of length three for large enough n. Subsequent improvements culminating in recent breakthrough applying the polynomial method [CLP17, EG17] establish that (contrary to the integer case as evidenced by the Behrend's construction) the largest progression-free set in $\mathbb{F}_3^n$ has density that is exponentially small in n. It is also well known that this last statement is equivalent to the following: There exists a constant $C > 0$ such that every set $S \subseteq \mathbb{F}_3^n$ with density μ contains at least $\mu^C \cdot 9^n$ arithmetic progressions of length three (including among 9^n progressions degenerate ones with difference zero).

As for longer progressions, while it is known (for example using the density Hales-Jewett theorem [FK91]) that dense subsets of $\mathbb{F}_p^n$ contain a dense proportion of progressions of any length k, the quantitative bounds are quite weak with the exception of progressions of length four (see [GT12]), where it has been established by Green and Tao that a set of density μ contains at least an $\exp(-\operatorname{poly}(1/\mu))$ proportion of all progressions.

We present a result that achieves a μ^C type of bound for arbitrarily long progressions, at the expense of restricting ourselves to *symmetric* sets: Subsets $S \subseteq \mathbb{F}_p^n$ where membership $x \in S$ is invariant under permutations of coordinates. More formally, for $x \in \mathbb{Z}_q^n$ and $a \in \mathbb{Z}_q$ we define the *weight* $w_a(x) := |\{i \in [n] : x_i = a\}|$. We say that $S \subseteq \mathbb{Z}_q^n$ is symmetric if membership $x \in S$ depends only on the weight tuple $(w_0(x), \ldots, w_{q-1}(x))$.

In fact, we prove a more general removal lemma. In the following we find it useful to frame our statements in terms of probabilities. For that purpose, let $X^{(1)}, \ldots, X^{(p)} \in \mathbb{F}_p^n$ be random variables representing a uniformly random arithmetic progression of length p, i.e., $X^{(j)} = (X_1^{(j)}, \ldots, X_n^{(j)})$ for $j = 1, \ldots, p$, and tuples $(X_i^{(1)}, \ldots, X_i^{(p)})$ for $i = 1, \ldots, n$ are independent and distributed uniformly among p^2 possible progressions $x, x+d, \ldots, x+(p-1)d$ for $x, d \in \mathbb{F}_p$.

Theorem 1. *Let $p \geqslant 3$ be prime and $0 < \mu < 1$. There exist n_0 and $C > 0$ such that for all $n \geqslant n_0$, if $S^{(1)}, \ldots, S^{(p)}$ are symmetric subsets of $\mathbb{F}_p^n$ satisfying*

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \ldots \wedge X^{(p)} \in S^{(p)}\right] < \mu^C , \tag{1}$$

then there exist symmetric sets $S'^{(1)}, \ldots, S'^{(p)} \subseteq \mathbb{F}_p^n$, each of density at most μ, such that letting $T^{(j)} := S^{(j)} \setminus S'^{(j)}$ we have

$$\Pr\left[X^{(1)} \in T^{(1)} \wedge \ldots \wedge X^{(p)} \in T^{(p)}\right] = 0 .$$

Taking $S^{(1)} = \cdots = S^{(p)} = S$ and noting that due to trivial progressions with difference $d = 0^n$ the probability $\Pr[X^{(1)}, \ldots, X^{(p)} \in S] = 0$ if and only if S is empty, it follows that a symmetric set of density μ contains at least $(\mu/p)^C \cdot p^{2n}$ progressions.[1] We remark that Theorem 1 has a weakness in that its conclusion holds only for large enough n after fixing the density μ. The technical reason is that we apply a version of the local limit theorem without an explicit error bound. We do not attempt to fix this deficiency in this work.

[1] We note that the question of existence of an arithmetic progression in a symmetric set has a positive answer for a less interesting reason: If a symmetric set S contains an element x with all weights $w_a(x)$ non-zero, then it is easy to find a progression in S consisting only of permutations of coordinates of x.

The second problem we consider concerns arithmetic progressions in $\mathbb{Z}_q^n$ with the difference restricted to lie in $\{0,1\}^n$. Accordingly, we call them *restricted progressions*. Again, an application of the density Hales-Jewett theorem establishes that a dense set $S \subseteq \mathbb{Z}_q^n$ contains a non-trivial restricted progression of length q for large enough n. However, the author is not aware of a proof that a dense set contains a dense fraction of all such progressions.

Our second result is a removal lemma for symmetric sets with respect to restricted progressions:

Theorem 2. *Let $q \geqslant 3$ and $\mu > 0$. There exists $\delta := \delta(q,\mu) > 0$ such that for every tuple of symmetric sets $S^{(1)}, \dots, S^{(q)} \subseteq \mathbb{Z}_q^n$ the following holds: Letting $X^{(1)}, \dots, X^{(q)}$ be a random arithmetic progression in $\mathbb{Z}_q^n$ with a difference restricted to $\{0,1\}^n$, if*

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \cdots \wedge X^{(q)} \in S^{(q)}\right] < \delta \;,$$

then there exists a symmetric set S' with density at most μ such that for $T^{(j)} := S^{(j)} \setminus S'$ we have

$$\Pr\left[X^{(1)} \in T^{(1)} \wedge \cdots \wedge X^{(q)} \in T^{(q)}\right] = 0 \;.$$

Similar as in the case of Theorem 1, it follows that a symmetric set S of density μ contains a dense fraction of all restricted progressions.

1.1 Proof idea

The proofs of Theorems 1 and 2 are applications of the same technique, proceeding in two stages. First, we use a local central limit theorem to show that those theorems are implied (in fact, equivalent to) certain additive combinatorial statements over the integers (more precisely, over $\mathbb{Z}^{q-1}$). Since the membership $x \in S$ depends only on the weight tuple $(w_0(x), \dots, w_{q-2}(x))$ (we omit $w_{q-1}(x)$, since, knowing n, it can be inferred from the other components), we can think in terms of weight tuples in $\mathbb{Z}^{q-1}$ rather than vectors in $\mathbb{Z}_q^n$. Furthermore, the CLT argument shows that sampling a random arithmetic progression of length q can be approximated by sampling q random weight tuples uniformly, under some additional constraints.

In case of Theorem 2 these constraints have the form of linear equations with integer coefficients. For illustration, we show below a statement for $q = 3$, which is equivalent to the same-set case of Theorem 2 (a more general statement that we need for full Theorem 2 is given as Theorem 17). For $N > 0$, let $[-N, N]$ denote the set $\{n \in \mathbb{Z} : |n| \leqslant N\}$.

Theorem 3. *Let $\mu > 0$ and let A_1, B_1, A_2, B_2 be i.i.d. uniform in $[-N, N]$. There exists $\delta := \delta(\mu) > 0$ such that for every subset $R \subseteq [-N, N]^2$ with density $\mu(R) := |R|/(2N+1)^2 \geqslant \mu$ we have*

$$\Pr\left[(A_1, B_1) \in R \wedge (A_2, B_2) \in R \wedge (A_1 + B_1 - B_2, A_2 + B_2 - A_1) \in R\right] \geqslant \delta \;.$$

To prove Theorems 17 and 3 we use a (hyper)graph removal lemma argument, similar as in the classical proof of Szemerédi's theorem or in works on removal lemmas for sets of linear equations [KSV09, Sha10, KSV12]. This application of graph removal makes the constant δ to be very small compared to the density μ and we do not attempt to make it explicit.

Moving to Theorem 1, it turns out that, since there are more progressions to choose from, there is a larger collection of possible arrangements of p weight tuples. As a result, the constraints in the relevant problem over $\mathbb{Z}^{p-1}$ turn out to be linear equations modulo p, which can be directly handled in an easier and more abstract fashion, at least for prime p.

While the restriction to symmetric sets is a strong one and the application of central limit theorem might be considered quite natural, the author finds it interesting that this technique results in an easier proof and a better bound in Theorem 1 as compared to Theorem 2.

1.2 Correlated spaces

One can view the finite field problems we consider as instances in a more general framework of *correlated product spaces*. Namely, let Ω be a finite set, $\ell \geqslant 2$ and $\mathcal{P}$ a probability distribution over Ω^ℓ such that all of its ℓ marginals are uniform over Ω. We call such a distribution $\mathcal{P}$ an *ℓ-step correlated space*. We consider the product probability space with n i.i.d. coordinates, where coordinate $i \in [n]$ gives rise to a random tuple $X_i^{(1)}, \ldots, X_i^{(\ell)}$ distributed according to $\mathcal{P}$.

The random variables $X_i^{(j)}$ form ℓ random vectors $X^{(j)} := \left(X_1^{(j)}, \ldots, X_n^{(j)}\right)$. Each of those vectors is individually uniform in Ω^n, but their joint distribution exhibits correlation across the steps. We consider a setting with fixed correlated space and n going to infinity.

Most generally, given sets $S^{(1)}, \ldots, S^{(\ell)} \subseteq \Omega^n$ with densities $\mu^{(1)}, \ldots, \mu^{(\ell)}$ we want to study the probability

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \ldots \wedge X^{(\ell)} \in S^{(\ell)}\right] .$$

For example, one can ask about the *same-set* case $S^{(1)} = \ldots = S^{(\ell)} = S$ with $\mu := \mu(S) > 0$. That is, for a given correlated space we can ask if there exists a bound

$$\Pr\left[X^{(1)} \in S \wedge \ldots \wedge X^{(\ell)} \in S\right] \geqslant c\,(\mathcal{P}, \mu) > 0\ ? \tag{2}$$

This problem was introduced in [HHM18] and we call a space satisfying (2) *same-set hitting*. Note that indeed the capset problem is captured by the same-set hitting on the three-step correlated space where $\Omega = \mathbb{F}_3$ and $\mathcal{P}$ is uniform in the set of progressions of length three, i.e., $\{000, 111, 222, 012, 120, 201, 021, 102, 210\}$.

Considering "dictator" sets, for which the membership depends on a single coordinate, it is easy to see that a necessary condition for same-set hitting is that the diagonal

$$\operatorname{diag}(\Omega) := \{(\omega, \ldots, \omega) : \omega \in \Omega\}$$

is contained in the support of $\mathcal{P}$. In [HHM18] we proved that this condition is sufficient for $\ell = 2$. As a matter of fact, we state the following conjecture:

Conjecture 4. Every correlated space with $\operatorname{diag}(\Omega) \subseteq \operatorname{supp}(\mathcal{P})$ is same-set hitting.

Generalizing Theorem 2 to arbitrary sets would confirm Conjecture 4 in case of restricted arithmetic progressions. A related, more general question is if general removal lemma holds for correlated product spaces:

Question 5. Is it the case that for every correlated space $\mathcal{P}$ and every $\mu > 0$ there exists $\delta(\mathcal{P}, \mu) > 0$ such that if

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \ldots \wedge X^{(\ell)} \in S^{(\ell)}\right] < \delta \ ,$$

then it is possible to remove a set S' of density at most μ from $S^{(1)}, \ldots, S^{(\ell)}$ and obtain $T^{(j)} := S^{(j)} \setminus S'$ with

$$\Pr\left[X^{(1)} \in T^{(1)} \wedge \ldots \wedge X^{(\ell)} \in T^{(\ell)}\right] = 0 \ ?$$

For all the author knows, we cannot even exclude a positive answer to Question 5 with $\delta > \mu^{C(\mathcal{P})}$ for every correlated space $\mathcal{P}$. On the other hand, while it is plausible that with some effort Theorems 1 and 2 can be generalized to hold for arbitrary correlated spaces, in this work we leave even this problem unresolved. We also leave open the problem of characterizing the correlated spaces which allow a stronger polynomial bound from Theorem 1. The class of spaces for which we can confirm Conjecture 4 is limited and we discuss known results in the following section.

1.3 Related works

We mention here some works that we find most relevant to our results and proofs.

As we said before, one well-studied example of a correlated space corresponds to the problem of arithmetic progressions in the finite field model. Extensive recent line of work based on the polynomial method [Gre05b, BX15, FK14, BCC+17, KSS18, Nor19, Peb18, FL17, FLS18, LS19] culminated in establishing that for random *k-cycles*, i.e., solutions to the equation $x_1 + \cdots + x_k = 0$ over finite field $\mathbb{F}_p$ indeed the removal lemma holds with $\delta > \mu^C$.

More generally, another interesting instance of a correlated space arises when we take $\Omega = G$ for a group G and $\mathcal{P}$ is uniform over solutions to some (full-rank) fixed linear equation system over G. For example, a random arithmetic progression $a_1, \ldots, a_q$ over $\mathbb{Z}_q$ is a random solution of the equation system $\{a_j + a_{j+2} = 2a_{j+1}\}_{j \in \{1,\ldots,q-2\}}$. Green [Gre05b] established such removal lemma for a single equation and any abelian group G (not necessarily in the product setting) and further work by Shapira [Sha10] and Král', Serra and Vena [KSV12] extended it to systems of equations over finite fields, and it can be seen that their results carry over to the product model $\mathbb{F}_p^n$.

Our proof of Theorem 17, which is the second part of the proof of Theorem 2, is related to this previous work on removal lemmas in systems of linear equations in the following way: On the one hand, the statement of Theorem 17 is a removal lemma for a particular type of a system of linear equations. Since it is a special system with some additional structure, more involved constructions from [Sha10] and [KSV12] are not required and we make a simpler argument, similar as in the proof of Szemerédi's theorem or in [KSV09]. On the other hand, since we consider subsets $W \subseteq \mathbb{Z}^{q-1}$, our result is not directly covered by [Sha10] or [KSV12], which concern only $W \subseteq \mathbb{Z}$.

Regarding Theorem 2, the problem of restricted progressions in $\mathbb{Z}_q^n$ seems to be particularly challenging, with most relevant questions being wide open. For example, it follows from known results that a *linear subspace* of $\mathbb{F}_3^n$ free of restricted progressions has dimension at most $n/2$ and that a subset of $\mathbb{F}_3^n$ that does not contain a restricted progression *of length two* must have size at most 2^n [Lev18]. It is also known [HHM18] that every subset of $\mathbb{F}_3^n$ of density μ contains at least $\delta(\mu) \cdot 6^n$ restricted progressions of length two, where δ is a triply exponentially small in μ.

More generally, a paper by Cook and Magyar [CM12] shows that a set of constant density $S \subseteq \mathbb{F}_p^n$ in the finite field model contains a constant proportion of arithmetic progressions with differences restricted to lie in a sufficiently well-behaved algebraic set. However, the author does not see how to apply their result in a very restricted setting of differences from $\{0,1\}^n$.

One reason we find the framework of correlated spaces interesting is that it encompasses some important problems from analysis of discrete functions, with applications in computer science. A canonical example of this setting are two steps $\ell = 2$ over binary alphabet $\Omega = \{0,1\}$ with $\mathcal{P}(00) = \mathcal{P}(11) = (1-p)/2$, $\mathcal{P}(01) = \mathcal{P}(10) = p/2$ for some $p \in [0,1]$. More generally, one can take any correlated space $\mathcal{P}$ and add to it a small amount of uniform noise, e.g., taking $\mathcal{P}' := (1-\varepsilon) \cdot \mathcal{P} + \varepsilon \cdot \mathcal{U}$, where $\mathcal{U}$ is the uniform distribution over Ω^ℓ.

It turns out that the theory of reverse hypercontractivity [MOS13] can be used to show that in such setting (and, more generally, whenever $\mathrm{supp}(\mathcal{P}) = \Omega^\ell$), one gets a general *set hitting*: For any $S^{(1)}, \ldots, S^{(\ell)}$ with $\mu(S^{(j)}) \geqslant \mu$ it holds that

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \ldots \wedge X^{(\ell)} \in S^{(\ell)}\right] \geqslant \mu^C .$$

More generally, [HHM18] established Conjecture 4 for $\ell = 2$, as well as whenever a certain *correlation* value is bounded with $\rho(\mathcal{P}) < 1$. The latter condition intuitively corresponds to the following: For all possible assignments of values to $\ell - 1$ of the steps in $\mathcal{P}$, the value of the remaining step is not determined. Note that this is quite a different regime that what is usually encountered in additive combinatorics. For example, the condition does not hold for full-rank systems of r linear equations over m variables, where fixing $m - r$ variables determines the values of the remaining r variables. [HHM18] is based on the invariance principle by Mossel [Mos10], which together with a follow-up work[2] [Mos20] establishes set-hitting and, more precisely, Gaussian bounds, in spaces with $\rho(\mathcal{P}) < 1$ for sets with small low-degree Fourier coefficients.

A work by Friedgut and Regev [FR18] applied the invariance principle and previous work with Dinur [DFR08] to establish a removal lemma in the two-step case $\ell = 2$. This

[2] The technique from [Mos20] implies another proof of same-set hitting for spaces with $\rho(\mathcal{P}) < 1$.

removal lemma has tower-type dependence between μ and δ, which is worth contrasting with [HHM18] which established an easier property of same-set hitting but with "only" triply exponential dependence between μ and δ. [DFR08] and [FR18] also studied the structure of sets with hitting probability zero, establishing that any such set must be almost contained in a junta.

The invariance principle can be compared with the Fourier-analytic approach to Szemerédi's theorem due to Gowers [Gow98, Gow01], which takes as its starting point the fact that the space of arithmetic progressions of length k is set hitting for all sets with low *Gowers uniformity norm* U_k.

Finally, we note a work by Austrin and Mossel [AM13] that established set hitting for low-Fourier degree sets with small Fourier coefficients in all correlated spaces where the distribution $\mathcal{P}$ is pairwise independent.

Organization of the paper In the following we prove Theorems 1 and 2. In Section 2 we introduce some notation, as well as the local limit theorem we use in the remaining proofs.

For convenience of a less committed reader, in Section 3 we prove the same-set case of Theorem 1 for $q = 3$. This proof utilizes main ideas of our technique, while being somewhat less technical and lighter in notation.

We proceed to prove Theorem 1 in Section 4 and Theorem 2 in Section 5. Each of the latter three sections is intended to be self-contained.

2 Preliminaries

We use both $O(\cdot)$ and $\Omega(\cdot)$ asymptotic notation, as well as constants $C > 0$ that will vary from time to time. All such implicit constants are allowed to depend on the alphabet size denoted by p or q.

Given $x \in \mathbb{Z}_q^n$ and $a \in \mathbb{Z}_q$, we define the respective *weight* to be $w_a(x) := |\{i \in [n] : x_i = a\}|$. In the context of arithmetic progressions $x^{(1)}, \ldots, x^{(q)}$ of length q, we will often speak of *weight tuples* $w^{(j)} = (w_0^{(j)}, \ldots, w_{q-2}^{(j)})$, where coordinates of $w^{(j)}$ will be weights $w_a(x^{(j)})$ shifted by a normalizing term approximately equal to n/q. A collection of q weight tuples $w = (w^{(1)}, \ldots, w^{(q)})$ will be referred to as a *weight arrangement.*

In some of the estimates we employ standard notation $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$ and $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$ for $x \in \mathbb{R}^n$.

We will apply several times the following corollary of the local multidimensional central limit theorem (see, e.g., Chapter 5 in [BR10] or Section 7 in [Spi76]):

Theorem 6. *Let $W_1, \ldots, W_n$ be i.i.d. random tuples such that each $W_i = (W_i^{(1)}, \ldots, W_i^{(\ell)})$ is distributed uniformly in*

$$\{0^\ell\} \cup \left\{0^j 1 0^{\ell-j-1} : 0 \leqslant j \leqslant \ell - 1\right\} .$$

For any tuple $w \in \mathbb{Z}^\ell$ with $d := w - \frac{n}{\ell+1} \cdot (1, \ldots, 1)$ we have

$$\begin{aligned} \Pr\left[\sum_{i=1}^n W_i = w\right] &= \frac{(\ell+1)^{(\ell+1)/2}}{(2\pi)^{\ell/2}} \cdot \frac{1}{n^{\ell/2}} \cdot \exp\left(-\frac{\ell+1}{2n}\left(\|d\|_2^2 + \left(\sum_{j=1}^{\ell} d_j\right)^2\right)\right) \\ &\quad + o\left(\frac{1}{n^{\ell/2}} \cdot \min\left(1, \frac{n}{\|d\|_2^2}\right)\right) , \end{aligned} \tag{3}$$

where the error term converges uniformly in w. In particular, we have

$$\frac{1}{Cn^{\ell/2}} \cdot \left(\exp\left(-\frac{C\|d\|_2^2}{n}\right) + o(1)\right) \leqslant \Pr\left[\sum_{i=1}^n W_i = w\right] \leqslant \frac{C}{n^{\ell/2}} \tag{4}$$

for some $C > 0$ that depends only on ℓ.

3 Restricted Progressions of Length Three

In this section we prove the same-set case of Theorem 2 for $q = 3$. For simplicity of exposition we additionally assume that n is divisible by six. We first show that our result is implied by Theorem 3 and then prove Theorem 3 via the triangle removal lemma. Let us start with the statement of the theorem.

Theorem 7. *Let $X, Y, Z \in \mathbb{Z}_3^n$ be a random arithmetic progression with difference restricted to lie in $\{0, 1\}^n$, where n is a multiple of six. For every symmetric set $S \subseteq \mathbb{Z}_3^n$ with density $\mu(S) \geqslant \mu > 0$ we have*

$$\Pr[X \in S \wedge Y \in S \wedge Z \in S] \geqslant \delta(\mu) > 0 . \tag{5}$$

The crucial part of the proof is a lemma that characterizes which weight arrangements are likely to be sampled in a random restricted progression. For this purpose, it is useful to introduce two random tuples. The first one is

$$M := (M_{000}, M_{111}, M_{012}, M_{120}, M_{201}) ,$$

where $M_{abc} := |\{i \in [n] : (X_i, Y_i, Z_i) = (a, b, c)\}| - n/6$ for $(a, b, c) = (x, x+d, x+2d), x \in \mathbb{Z}_3, d \in \{0, 1\}$. That is, the tuple M expresses normalized counts of six restricted progressions across n coordinates. Note that M_{222} is omitted, since it can be inferred from the remaining components of M.

The second random tuple represents weight arrangements of elements of the restricted progression:

$$W := (W_X, W_Y, W_Z) := (w_0(X), w_1(X), w_0(Y), w_1(Y), w_0(Z), w_1(Z)) - (n/3, \ldots, n/3) .$$

Again, we omit weights $w_2(\cdot)$, since they can be deduced from the rest. Note that W is determined by M, but, as it turns out, not the other way around.

Lemma 8. *Let (X, Y, Z) be a random restricted progression with n divisible by six. Let $w := (x_0, x_1, y_0, y_1, z_0, z_1) \in [-N, N]^6$ with $N = C_1\sqrt{n}$ for some $C_1 > 0$. Then,*

$$\Pr[W = w] \text{ is } \begin{cases} \text{at least } C_2/N^4 & \text{if } (z_0, z_1) = (x_0 + x_1 - y_1, y_0 + y_1 - x_0) \ , \\ 0 & \text{otherwise,} \end{cases}$$

for some $C_2 := C_2(C_1) > 0$ and N large enough (also depending on C_1).

Proof. Let us call a tuple w that satisfies $(z_0, z_1) = (x_0 + x_1 - y_1, y_0 + y_1 - x_0)$ *feasible.* If w is not feasible, then clearly $\Pr[W = w] = 0$, since restricting the progression difference to $\{0, 1\}^n$ implies that

$$\begin{aligned} w_0(Z) &= M_{000} + M_{120} + n/3 = w_0(X) + w_1(X) - w_1(Y) \ , \\ w_1(Z) &= M_{111} + M_{201} + n/3 = w_0(Y) + w_1(Y) - w_0(X) \ . \end{aligned}$$

For a feasible w, one can see that a tuple m gives rise to the weight arrangement w if and only if it is an integer solution of the equation system

$$\begin{cases} x_0 = m_{000} + m_{012} \\ x_1 = m_{111} + m_{120} \\ y_0 = m_{000} + m_{201} \\ y_1 = m_{111} + m_{012} \end{cases} ,$$

and these solutions are given as

$$m = (m_{000}, m_{111}, m_{012}, m_{120}, m_{201}) = (k, y_1 - x_0 + k, x_0 - k, x_1 - y_1 + x_0 - k, y_0 - k)$$

for $k \in \mathbb{Z}$. Now we can calculate

$$\begin{aligned} \Pr[W = w] &= \sum_{k \in \mathbb{Z}} \Pr\left[M = (k, y_1 - x_0 + k, x_0 - k, x_1 - y_1 + x_0 - k, y_0 - k)\right] \\ &\geqslant \sum_{k=0}^{\lfloor C_1\sqrt{n} \rfloor} \Pr\left[M = (k, y_1 - x_0 + k, x_0 - k, x_1 - y_1 + x_0 - k, y_0 - k)\right] \ . \end{aligned} \tag{6}$$

Each tuple $m = m(k)$ in the summation (6) is contained in $[-4N, 4N]^5$ and therefore satisfies $\|m\|_2^2 = O(C_1^2 \cdot n)$. Applying lower bound in (4) to M and m, each term in the summation (6) must be at least $C/n^{5/2}$, where C depends on C_1 and n is large enough. Finally, summing up over k we get

$$\Pr[W = w] \geqslant \frac{C}{n^2} = \frac{C_2}{N^4} \ ,$$

as claimed. □

Theorem 3 implies Theorem 7. Let $\mu > 0$ and $S \subseteq \mathbb{Z}_3^n$ be a symmetric set with $\mu(S) \geqslant \mu > 0$. Note that we can assume that n is big enough.

Recall the random tuple $W = (W_X, W_Y, W_Z)$ and for $x \in \mathbb{Z}_3^n$ let $W_x := (w_0(x) - n/3, w_1(x) - n/3)$. Since S is symmetric, there exists a set $R := R(S) \subseteq \mathbb{Z}^2$ such that $x \in S$ if and only if $W_x \in R$. Since, by a standard concentration bound, we can find $C(\mu)$ such that

$$\Pr\left[\, \|W_X\|_\infty > C\sqrt{n}\right] < \mu/2 \, ,$$

from now on we will assume w.l.o.g. that $R \subseteq [-N, N]^2$ for $N := C\sqrt{n}$.

Observe that, due to upper bound in (4) applied to random variable W_X, for each $w \in R$ we have

$$\Pr\left[W_X = w\right] \leqslant O(1/n) \, ,$$

and therefore $|R| = \Omega(\mu/n)$, implying $\frac{|R|}{(2N+1)^2} \geqslant c(\mu) > 0$. Now, Theorem 3 gives

$$\frac{\#\left\{(x_0, x_1, y_0, y_1) : (x_0, x_1) \in R \wedge (y_0, y_1) \in R \wedge (x_0 + x_1 - y_1, y_0 + y_1 - x_0) \in R\right\}}{(2N+1)^4} \geqslant c(\mu) > 0 \, ,$$

but that, due to Lemma 8, yields

$$\begin{aligned}\Pr[X \in S \wedge Y \in S \wedge Z \in S] &= \Pr[W_X \in R \wedge W_Y \in R \wedge W_Z \in R] \\ &= \sum_{\substack{(x_0,x_1)\in R \\ (y_0,y_1)\in R}} \mathbb{1}_R(x_0 + x_1 - y_1, y_0 + y_1 - x_0) \cdot \\ &\quad \cdot \Pr\left[W = (x_0, x_1, y_0, y_1, x_0 + x_1 - y_1, y_0 + y_1 - x_0)\right] \\ &\geqslant c(\mu) \cdot N^4 \cdot \frac{C_2(\mu)}{N^4} \geqslant \delta(\mu) > 0. \qquad \square\end{aligned}$$

We proceed to the proof of Theorem 3, which we restate here for convenience.

Theorem 3. *Let $\mu > 0$ and let A_1, B_1, A_2, B_2 be i.i.d. uniform in $[-N, N]$. There exists $\delta := \delta(\mu) > 0$ such that for every subset $R \subseteq [-N, N]^2$ with density $\mu(R) := |R|/(2N+1)^2 \geqslant \mu$ we have*

$$\Pr\left[(A_1, B_1) \in R \wedge (A_2, B_2) \in R \wedge (A_1 + B_1 - B_2, A_2 + B_2 - A_1) \in R\right] \geqslant \delta \, .$$

The proof is a variation on the triangle removal proof of Roth's theorem. Let us start by stating the removal lemma:

Theorem 9 (Triangle removal lemma). *For every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that if a simple graph $G = (V, E)$ contains at most $\delta|V|^3$ triangles, then it is possible to make G triangle-free by removing from it at most $\varepsilon|V|^2$ edges.*

Proof of Theorem 3. Let $N \in \mathbb{N}$ and $R \subseteq [-N,N]^2$ with density $\mu(R) \geqslant \mu > 0$. As before, we will call a triple of points $(a_1,b_1),(a_2,b_2),(a_3,b_3) \in \mathbb{Z}^2$ *feasible* if $(a_3,b_3) = (a_1+b_1-b_2, a_2+b_2-a_1)$. We define a tripartite graph G as follows:

- There are three groups of vertices V_1, V_2, V_3. In each group the vertices are labeled with elements of $[-M,M]^2$ for $M := 3N$. Note that the total number of vertices of G is $|V| = 3(2M+1)^2$.
- Edge adjacency is defined by:

$$V_1 \ni (i_a,i_b) \sim (j_a,j_b) \in V_2 \text{ iff } (a_1,b_1) := (i_b-j_b, j_a-i_a+j_b-i_b) \in R\ , \tag{7}$$

$$V_2 \ni (j_a,j_b) \sim (k_a,k_b) \in V_3 \text{ iff } (a_2,b_2) := (k_a-j_a+k_b-j_b, j_a-k_a) \in R\ , \tag{8}$$

$$V_1 \ni (i_a,i_b) \sim (k_a,k_b) \in V_3 \text{ iff } (a_3,b_3) := (k_a-i_a, k_b-i_b) \in R\ . \tag{9}$$

Given a triple of vertices $(i_a,i_b),(j_a,j_b),(k_a,k_b)$, we associate with it a triple of points $(a_1,b_1),(a_2,b_2),(a_3,b_3) \in \mathbb{Z}^2$ given by the right-hand sides of equations in (7) to (9). One checks that this triple of points is feasible. Furthermore, by definition, whenever $(i_a,i_b),(j_a,j_b),(k_a,k_b)$ form a triangle, the points $(a_1,b_1),(a_2,b_2),(a_3,b_3)$ all belong to R.

Conversely, given a point $(a,b) \in R$, we can see that each triple of vertices $(i_a,i_b),(i_a+a+b, i_b-a),(i_a+a,i_b+b)$ for $(i_a,i_b) \in [-N,N]^2$ forms a triangle. Therefore, the graph G contains at least $\mu\cdot(2N+1)^4 \geqslant \mu\cdot\left(\frac{2M+1}{3}\right)^4 = \frac{\mu}{3^6}|V|^2$ triangles. Furthermore, it is clear that all those triangles are edge-disjoint. Hence, G requires at least $\frac{\mu}{3^6}|V|^2$ edge deletions to become triangle-free and, by triangle removal lemma, contains at least $\delta(\mu)|V|^3$ triangles.

Finally, we note that each feasible triple of points $(a_1,b_1),(a_2,b_2),(a_3,b_3) \in R$ gives rise to at most $(2M+1)^2$ triangles. This is because each vertex $(i_a,i_b) \in V_1$ determines at most one triangle associated with this triple. Since G contains at least $\delta|V|^3$ triangles, the number of feasible triples contained in R must be at least

$$\frac{\delta|V|^3}{(2M+1)^2} = 27\delta(2M+1)^4 \geqslant \delta(2N+1)^4\ ,$$

but this means

$$\Pr\left[(A_1,B_1) \in R \wedge (A_2,B_2) \in R \wedge (A_1+B_1-B_2, A_2+B_2-A_1) \in R\right] \geqslant \delta > 0\ ,$$

as we wanted. □

4 Proof of Theorem 1

In this section we prove Theorem 1.

As a very preliminary point, note that it suffices to consider only densities $\mu \leqslant 1-1/p$ that are bounded away from one. If μ is greater than $1-1/p$, one can, for example, take $T^{(j)} := S^{(j)} \cap \{x : \sum_{i=1}^n x_i = 0 \pmod p\}$ for $j \neq p$ and $T^{(p)} := S^{(p)} \cap \{x : \sum_{i=1}^n x_i = 1 \pmod p\}$ in order to obtain progression-free sets $T^{(1)},\ldots,T^{(p)}$.

As in the proof of Theorem 7, we start with observing that there exist sets $R^{(1)}, \dots, R^{(p)} \subseteq \mathbb{F}^{p-1}$ such that $x \in S^{(j)}$ if and only if $W(x) := (w_0(x) - p\lfloor n/p^2 \rfloor, \dots, w_{p-2}(x) - p\lfloor n/p^2 \rfloor) \in R^{(j)}$ (the choice of the $p\lfloor n/p^2 \rfloor$ shift will become apparent in the next paragraph). Furthermore, using standard concentration bound

$$\Pr\left[|w_a(X) - n/p| > tn\right] \leqslant 2\exp(-2nt^2)$$

for any fixed $a \in \mathbb{F}_p$ together with the union bound, we also establish that there exists some $C_1 > 0$ such that

$$\Pr\left[\left\|W(X^{(j)})\right\|_\infty > C_1\sqrt{n \cdot \ln 1/\mu}\right] \leqslant \frac{\mu}{2} ,$$

and therefore we can remove from each of $S^{(1)}, \dots, S^{(p)}$ a symmetric set of density at most $\mu/2$ and assume from now on that the weight sets $R^{(1)}, \dots, R^{(p)}$ have limited range: If $w \in R^{(j)}$, then $\|w\|_\infty \leqslant C_1\sqrt{n \cdot \ln 1/\mu}$.

For $a, d \in \mathbb{F}_p$, define random variables

$$M(a,d) := \left|\left\{i \in [n] : x_i^{(1)} = a \wedge x_i^{(2)} - x_i^{(1)} = d\right\}\right| - \lfloor n/p^2 \rfloor .$$

Consider the random tuple M consisting of $p^2 - 1$ coordinates $M(a,d)$ except for $M(p-1, 0)$. We note for future reference that M can be written as a sum of i.i.d. random tuples $M = \sum_{i=1}^n M_i$ such that Theorem 6 is applicable. We also note that there exists a matrix $A \in \{0,1\}^{p(p-1)\times(p^2-1)}$ such that letting $W := (W(X^{(1)}), \dots, W(X^{(p)}))$ we can write a linear system of equations $W = AM$.

At this point we need to understand how solutions to the system $W = AM$ look like. This is done in the following lemma:

Lemma 10. *A general solution to the equation system $w = Am$ for $w \in \mathbb{R}^{p(p-1)}$ is given by*

$$m = \frac{1}{p}Bw + K \cdot \mathbb{R}^{p-1} \tag{10}$$

for some integer-valued *matrices $B \in \mathbb{Z}^{(p^2-1)\times p(p-1)}$ and $K \in \mathbb{Z}^{(p^2-1)\times(p-1)}$. In particular, matrix A has full rank over reals.*

Furthermore, if the vector $\frac{1}{p}Bw$ is not integer for some integer $w \in \mathbb{Z}^{p(p-1)}$, then the system $w = Am$ does not have an integer solution.

Proof. We start by showing that matrix A has full rank with a solution given as $m = \frac{1}{p}B'w$ for some $B' \in \mathbb{Z}^{(p^2-1)\times p(p-1)}$. To do this, we explicitly construct the columns of B'. That is, for every $j \in [p]$ and $a \in \mathbb{F}_p, a \neq p-1$, we find a solution $m_{j,a} \in \mathbb{R}^{p^2-1}$ to equation $w_{j,a} = Am_{j,a}$ where $w_{j,a}$ has value one at the coordinate corresponding to $W_a(x^{(j)})$ and zero everywhere else. Furthermore, this solution satisfies $p \cdot m_{j,a} \in \mathbb{Z}^{p^2-1}$. We give $m_{j,a}$

as:

$$m_{j,a}(b,d) := \begin{cases} -(p-2)/p & \text{if } b = a, d = 0\ , \\ -(p-1)/p & \text{if } b \neq a, d = 0\ , \\ 2/p & \text{if } d \neq 0, b + (j-1)d = a\ , \\ 1/p & \text{if } d \neq 0, b + (j-1)d \notin \{a, p-1\}\ , \\ 0 & \text{if } d \neq 0, b + (j-1)d = p-1\ . \end{cases} \tag{11}$$

As a sanity check we can convince ourselves that $m_{j,a}$ features one coordinate with value $-(p-2)/p$, $p-2$ coordinates with value $-(p-1)/p$, $p-1$ coordinates with value $2/p$, $(p-1)(p-2)$ coordinates with value $1/p$ and $p-1$ coordinates with value zero. Indeed, we have $w_{j,a} = Am_{j,a}$ as can be seen by indexing coordinates of $w = Am_{j,a}$ as $w(j',b)$ with $j' \in [p], b \in \mathbb{F}_p \setminus \{p-1\}$ and computing

$$\begin{aligned} w(j,a) &= -\frac{p-2}{p} + (p-1) \cdot \frac{2}{p} = 1 \\ w(j,b) &= -\frac{p-1}{p} + (p-1) \cdot \frac{1}{p} = 0 && \text{for } b \notin \{a, p-1\}\ , \\ w(j',a) &= -\frac{p-2}{p} + (p-2) \cdot \frac{1}{p} = 0 && \text{for } j' \neq j\ , \\ w(j',b) &= -\frac{p-1}{p} + \frac{2}{p} + (p-3) \cdot \frac{1}{p} = 0 && \text{for } b \notin \{a, p-1\}, j' \neq j\ . \end{aligned}$$

By a similar check we can characterize the $(p-1)$-dimensional kernel of the linear operator A concluding that $Am = 0$ holds if

$$m(b,d) := \begin{cases} -\sum_{d'=1}^{p-1} \alpha_{d'} & \text{if } d = 0\ , \\ \alpha_d & \text{if } d \neq 0\ , \end{cases} \tag{12}$$

for $\alpha_1, \ldots, \alpha_{p-1} \in \mathbb{R}$. Since matrix A is full rank, its kernel has dimension $p-1$ and equation (12) represents all elements in the kernel.

Combining (11) and (12) allows us to write a general solution to $w = Am$ as

$$m = \frac{1}{p} B' w + K \cdot \mathbb{R}^{p-1}$$

for $B' \in \mathbb{Z}^{(p^2-1) \times p(p-1)}$ and $K \in \mathbb{Z}^{(p^2-1) \times (p-1)}$.

As for the "furthermore" claim, notice that another general solution to $w = Am$ can be obtained by adding an arbitrary kernel vector Kv to one of the columns of B'. Applying this observation repeatedly together with (12), we obtain another integer matrix B such that we still have the equation

$$m = \frac{1}{p} B w + K \cdot \mathbb{R}^{p-1}$$

and, additionally, for every column $b_{j,a}$ of B we have $b_{j,a}(0,d) = 0$ for $d = 1, \ldots, p-1$. But this ensures that a solution $m = \frac{1}{p} Bw + Kv$ has $m(0,d) = v_d$, so v must be integer in order for m to be integer, which implies that $\frac{1}{p} Bw$ is integer as well. □

Consider an integer tuple $w \in \mathbb{Z}^{p(p-1)}$ and a tuple $w \pmod{p} \in \mathbb{F}_p^{p(p-1)}$ which consists of entries of w reduced modulo p. Since by Lemma 10 $w = Am$ has an integer solution $m \in \mathbb{Z}^{p^2-1}$ if and only if $\frac{1}{p}Bw$ is integer, in particular this property depends only on $w \pmod{p}$.

Furthermore, if w with $\|w\|_\infty \leqslant D$ has at least one integer solution, then taking $m = \frac{1}{p}Bw + K \cdot [D]^{p-1}$ we get D^{p-1} integer solutions m with bounded norm $\|m\|_\infty \leqslant O(D)$ and, consequently, $\|m\|_2^2 \leqslant O(D^2)$. Taking $D = C_1\sqrt{n \ln 1/\mu}$ and applying lower bound in (4) to random tuple M, we see that, for n big enough, we will have for each such m

$$\Pr[M = m] \geqslant \frac{\mu^C}{n^{(p^2-1)/2}} \ ,$$

and, as a result,

$$\Pr[W = w] \geqslant \frac{\mu^C}{n^{(p^2-1)/2}} \cdot \left(C_1\sqrt{n \ln 1/\mu}\right)^{p-1} \geqslant \frac{\mu^C}{n^{p(p-1)/2}} \ . \tag{13}$$

To finish the proof, divide each set $S^{(j)}$ into p^{p-1} "congruence classes" $S_v^{(j)}$ indexed by $v = (v_0, \ldots, v_{p-2}) \in \mathbb{F}_p^{p-1}$ and defined as

$$S_v^{(j)} := S^{(j)} \cap \left\{x : w_a(x) - p\lfloor n/p^2 \rfloor = v_a \pmod{p}, \quad a = 0, \ldots, p-2\right\} \ .$$

Now, if $|S_v^{(j)}| \leqslant \frac{\mu}{2p^{p-1}} \cdot p^n$, we remove $S_v^{(j)}$ from $S^{(j)}$. Clearly, we removed from each $S^{(j)}$ a symmetric set of density at most μ. The final claim is that there can be no tuple arrangement $v^{(1)}, \ldots, v^{(p)}$ such that:

1. None of $S_{v^{(j)}}^{(j)}$ has been removed.
2. There exists an integer solution to $v = Am$, where $v = (v^{(1)}, \ldots, v^{(p)})$.

Otherwise, each of the sets $S_{v^{(j)}}^{(j)}$ has density at least $\mu/2p^{p-1}$. A set $S_{v^{(j)}}^{(j)}$ is symmetric with the corresponding set of tuples $R_{v^{(j)}}^{(j)} \subseteq \mathbb{Z}^{p-1}$ bounded by $\|w^{(j)}\|_\infty \leqslant C_1\sqrt{n \ln 1/\mu}$ for $w^{(j)} \in R_{v^{(j)}}^{(j)}$. Applying Theorem 6 to the random tuple $W^{(j)} = (W_0^{(j)}, \ldots, W_{p-2}^{(j)})$, we see that each such tuple $w^{(j)} \in R_{v^{(j)}}^{(j)}$ has

$$\Pr[W^{(j)} = w^{(j)}] \leqslant O\left(\frac{1}{n^{(p-1)/2}}\right) \ ,$$

and therefore, we can bound the size of $R_{v^{(j)}}^{(j)}$ by

$$\left|R_{v^{(j)}}^{(j)}\right| \geqslant \Omega\left(\mu \cdot n^{(p-1)/2}\right) \ .$$

As a result, we get a set of $\Omega\left(\mu^p \cdot n^{p(p-1)/2}\right)$ weight arrangements $w = (w^{(1)}, \ldots, w^{(p)})$ with $\|w\|_\infty \leqslant C_1\sqrt{n \ln 1/\mu}$ and for each of them the system $w = Am$ has an integer-valued solution and, by (13), $\Pr[W = w] \geqslant \frac{\mu^C}{n^{p(p-1)/2}}$, which finally gives us

$$\Pr\left[X^{(1)} \in S^{(1)} \wedge \ldots \wedge X^{(p)} \in S^{(p)}\right] \geqslant \frac{1}{C} \cdot \mu^p \cdot n^{p(p-1)/2} \cdot \frac{\mu^C}{n^{p(p-1)/2}} \geqslant \mu^C \ ,$$

which contradicts assumption (1) if the constant C is chosen large enough.

We established that there are no "mod p" weight arrangements $(v^{(1)},\dots,v^{(p)})$ that satisfy the two conditions above. But it follows that there are no weight arrangements $w=(w^{(1)},\dots,w^{(p)})$ left in $R^{(1)}\times\cdots\times R^{(p)}$ for which there is an integer solution to $w=Am$, and therefore no arithmetic progressions left in the product set $S^{(1)}\times\cdots\times S^{(p)}$, and we are done. □

Remark 11. We make no attempt to precisely estimate the constant C in the exponent, but following the argument above one can see that it is bounded by a polynomial function of p.

5 Proof of Theorem 2

In the following we prove Theorem 2. We start with some definitions:

Definition 12. For $q\geqslant 3$ and $n\geqslant 1$, we let

$$\mathcal{P}:=\mathcal{P}(q,n):=\left\{(x,x+d,\dots,x+(q-1)d):x\in\mathbb{Z}_q^n,d\in\{0,1\}^n\right\}$$

Note that $|\mathcal{P}(q,n)|=(2q)^n$. We will call an element of $\mathcal{P}(q,n)$ a *restricted progression.* We will restate our removal lemma now:

Theorem 13. *For all $\mu>0$ there exists $\delta:=\delta(q,\mu)>0$ such that for all symmetric sets $S^{(1)},\dots,S^{(q)}\subseteq\mathbb{Z}_q^n$:*

If $S^{(1)}\times\ldots\times S^{(q)}$ contains at most $\delta\cdot(2q)^n$ restricted progressions, then it is possible to remove a total number of at most μq^n elements from $S^{(1)},\dots,S^{(q)}$ and obtain symmetric sets $T^{(1)},\dots,T^{(q)}$ such that $T^{(1)}\times\ldots\times T^{(q)}$ contains no *restricted progressions.*

As before, the proof consists of two parts: First, we make a CLT argument reducing Theorem 13 to a variation on removal lemma for certain linear equations over $\mathbb{Z}_N^{q-1}$. Then, we apply the hypergraph removal lemma to establish the linear equation removal property. To state the removal property over $\mathbb{Z}_N^{q-1}$ we need another definition specifying the allowed weight arrangements of restricted progressions.

Definition 14. For $x\in\mathbb{Z}_q^n$, we define the *weight tuple* of x as $W(x):=(W_1(x),\dots,W_{q-1}(x))\in\mathbb{Z}^{q-1}$, where

$$W_a(x):=w_a(x)-2\lfloor n/2q\rfloor=|\{i\in[n]:x_i=a\}|-2\lfloor n/2q\rfloor\ .$$

Definition 15. An arrangement of tuples $(w_1^{(1)},\dots,w_{q-1}^{(1)}),\dots,(w_1^{(q)},\dots,w_{q-1}^{(q)})$ (understood, depending on the context, as element of $\mathbb{Z}^{q(q-1)}$ or $\mathbb{Z}_N^{q(q-1)}$) is *feasible* if

$$\forall j=3,\dots,q:$$

$$w_1^{(j)}=\sum_{a=1}^{q-1}w_a^{(j-2)}-\sum_{a=2}^{q-1}w_a^{(j-1)} \tag{14}$$

$$\forall a=2,\dots,q-1:w_a^{(j)}=-w_{a-1}^{(j-2)}+w_{a-1}^{(j-1)}+w_a^{(j-1)} \tag{15}$$

For $q \geqslant 3$ and $N \geqslant 1$ we let $\mathcal{E} := \mathcal{E}(q, N) \subseteq \mathbb{Z}_N^{q(q-1)}$ to be the set of all feasible arrangements of tuples.

Note that $|\mathcal{E}(q, N)| = N^{2(q-1)}$. The definition of a feasible tuple is motivated by the following claim, which can be seen to be true by inspection:

Claim 16. *Let $x^{(1)}, \dots, x^{(q)} \in \mathbb{Z}_q^n$ be a restricted progression. Then, the weight arrangement $w(x^{(1)}), \dots, w(x^{(q)}) \in \mathbb{Z}^{q(q-1)}$ is feasible.*

Finally, we are ready to state the removal property for feasible arrangements. In this case it seems slightly more convenient (but not much different) to work in the cyclic group $\mathbb{Z}_N$ rather than in $\mathbb{Z}$.

Theorem 17. *For all $\mu > 0$ there exists $\delta := \delta(q, \mu) > 0$ such that for all sets $R^{(1)}, \dots, R^{(q)} \subseteq \mathbb{Z}_N^{q-1}$:*

If the product $R^{(1)} \times \ldots \times R^{(q)}$ contains at most $\delta N^{2(q-1)}$ feasible arrangements, then it is possible to remove a total number of at most μN^{q-1} tuples from $R^{(1)}, \dots, R^{(q)}$ and obtain sets $R'^{(1)}, \dots, R'^{(q)}$ such that the product $R'^{(1)} \times \ldots \times R'^{(q)}$ contains no feasible tuple arrangements.

5.1 Theorem 17 implies Theorem 13

The CLT argument that we use to prove that Theorem 13 is implied by Theorem 17 can be encapsulated in the following lemma that will be proved last. Before stating the lemma we need one more definition:

Definition 18. Let $w \in \mathbb{Z}^{q-1}$ be a weight tuple and $w^{(1)}, \dots, w^{(q)} \in \mathbb{Z}^{q(q-1)}$ a weight arrangement. We let

$$\#w := \left|\left\{x \in \mathbb{Z}_q^n : w(x) = w\right\}\right|$$
$$\#\left(w^{(1)}, \dots, w^{(q)}\right) := \left|\left\{\left(x^{(1)}, \dots x^{(q)}\right) \in \mathcal{P} : \ \forall j = 1, \dots, q : w(x^{(j)}) = w^{(j)}\right\}\right|$$

Lemma 19. *Let $q \geqslant 3$, $C_1 > 0$ and let $N := C_1\sqrt{n}$. For any weight tuple $w \in [-N, N]^{q-1}$ and any weight arrangement $w^{(1)}, \dots, w^{(q)} \in [-N, N]^{q(q-1)}$ we have the following:*

1. *If $w^{(1)}, \dots, w^{(q)}$ is not feasible, then $\#\left(w^{(1)}, \dots, w^{(q)}\right) = 0$.*
2. *If $w^{(1)}, \dots, w^{(q)}$ is feasible, then*

$$\frac{1}{C} \leqslant \#\left(w^{(1)}, \dots, w^{(q)}\right) \cdot \frac{N^{2(q-1)}}{(2q)^n} \leqslant C\ , \tag{16}$$

   for large enough n and some $C > 0$ that may depend on C_1.

3. *Similarly, $\frac{1}{C} \leqslant \#w \cdot \frac{N^{q-1}}{q^n} \leqslant C$ for large enough n.*

Proof of Theorem 13 assuming Theorem 17 and Lemma 19. Let $S^{(1)}, \dots, S^{(q)} \subseteq \mathbb{Z}_q^n$ be the sets from the statement. Since they are symmetric, there are sets $R^{(1)}, \dots, R^{(q)} \subseteq \mathbb{Z}^{q-1}$ such hat

$$x \in S^{(j)} \iff W(x) \in R^{(j)} .$$

The first observation is that we can assume without loss of generality that n is large and that the weights are restricted such that $R^{(j)} \in [-N, N]^{q-1}$ for $N := C_1\sqrt{n}$ for some $C_1 := C_1(q, \mu) > 0$. This is because by a standard concentration bound

$$\left|\left\{x : W(x) \notin [-N, N]^{q-1}\right\}\right| \leqslant \frac{\mu}{2q} q^n$$

for a big enough C_1 and therefore it takes at most $\mu/2 \cdot q^n$ removals to get rid of all the elements giving rise to weight tuples outside $[-N, N]^{q-1}$.

By Lemma 19.2, there exists some $C := C(q, C_1) > 0$ such that each feasible arrangement in $R^{(1)} \times \ldots \times R^{(q)}$ induces at least $\frac{(2q)^n}{C \cdot N^{2(q-1)}}$ restricted progressions in $S^{(1)} \times \ldots \times S^{(q)}$. Let $\mu' := \frac{\mu}{2C(2q)^{q-1}}$ and let $\delta'(\mu') > 0$ be given by Theorem 17. We set $\delta(\mu) := (2q)^{2(q-1)}\delta'/C$.

Since, by assumption, $S^{(1)} \times \ldots \times S^{(q)}$ contains at most $\delta(2q)^n$ restricted progressions, $R^{(1)} \times \ldots \times R^{(q)}$ contains at most $\delta C N^{2(q-1)}$ feasible arrangements (understood as elements of $\mathbb{Z}^{q(q-1)}$). Furthermore, taking $N' := 2qN$ and inspecting Definition 15, we conclude that $R^{(1)} \times \ldots \times R^{(q)}$ contains at most $\frac{\delta C}{(2q)^{2(q-1)}} N'^{2(q-1)} = \delta' N'^{2(q-1)}$ feasible arrangements understood as elements of $\mathbb{Z}_{N'}^{q(q-1)}$.

Applying Theorem 17 for N' and μ', we get that one can remove at most $\mu' N'^{q-1} = \frac{\mu}{2C} N^{q-1}$ elements from $R^{(1)}, \dots, R^{(q)}$ and obtain $R'^{(1)}, \dots, R'^{(q)} \subseteq [-N, N]^{q-1}$ such that $R'^{(1)} \times \ldots \times R'^{(q)}$ contains no feasible arrangements (understood either as elements of $\mathbb{Z}_{N'}^{q(q-1)}$ or $\mathbb{Z}^{q(q-1)}$). Finally, due to Lemma 19.3, we can remove at most $\frac{\mu}{2} q^n$ elements from the sets $S^{(1)}, \dots, S^{(q)}$ to obtain symmetric sets $T^{(1)}, \dots, T^{(q)}$ such that, by Lemma 19.1, the product $T^{(1)} \times \ldots \times T^{(q)}$ contains no restricted progressions. □

It remains to prove Lemma 19. We achieve this by utilizing Theorem 6.

Proof of Lemma 19. Point 1 is just a restatement of Claim 16.

We turn to Point 3 next. Consider $X \in \mathbb{Z}_q^n$ sampled uniformly at random. Recall our notation $W(x) = (W_1(x), \dots, W_{q-1}(x)) = (w_1(x) - 2\lfloor n/2q \rfloor, \dots, w_{q-1}(x) - 2\lfloor n/2q \rfloor)$ for $x \in \mathbb{Z}_q^n$ and the random variable $W = W(X)$. Clearly, we can apply (4) to W and obtain

$$\frac{1}{Cn^{(q-1)/2}} \leqslant \Pr[W = w] = \frac{\#w}{q^n} \leqslant \frac{C}{n^{(q-1)/2}} ,$$

which yields the conclusion after rearranging the terms.

As for Point 2, consider a choice of uniform random restricted progression $X^{(1)}, \ldots, X^{(q)}$. We will apply Theorem 6 to random variables

$$\begin{aligned}
M_{\mathrm{same}}(a) &:= \left|\left\{i \in [n] : X_i^{(1)}, \ldots, X_i^{(q)} = a, a, \ldots, a\right\}\right| - \left\lfloor \frac{n}{2q} \right\rfloor \ , \\
M_{\mathrm{cycle}}(a) &:= \left|\left\{i \in [n] : X_i^{(1)}, \ldots, X_i^{(q)} = a, a+1, \ldots, a-1\right\}\right| - \left\lfloor \frac{n}{2q} \right\rfloor \ .
\end{aligned}$$

We let $M := (M_{\mathrm{same}}(1), \ldots, M_{\mathrm{same}}(q-1), M_{\mathrm{cycle}}(0), \ldots, M_{\mathrm{cycle}}(q-1))$ (note that $M \in \mathbb{Z}^{2q-1}$). Now we need to specify a relation between possible values of M and feasible weight arrangements $w^{(1)}, \ldots, w^{(q)}$. Observe that each possible value m of M uniquely determines a feasible weight arrangement $w^{(1)}, \ldots, w^{(q)}$. It turns out that there is a reasonably simple characterization of the set of tuples m that give rise to a given arrangement $w^{(1)}, \ldots, w^{(q)}$. Namely, we check that these values form a linear one-dimensional solution space with triangular structure given by

$$m_{\mathrm{cycle}}(0) = k \ , \tag{17}$$

$$m_{\mathrm{same}}(a) = w_a^{(2)} - m_{\mathrm{cycle}}(a-1) \ , \quad a = 1, \ldots, q-1 \ , \tag{18}$$

$$m_{\mathrm{cycle}}(a) = w_a^{(1)} - m_{\mathrm{same}}(a) \ , \qquad a = 1, \ldots, q-1 \ . \tag{19}$$

for every $k \in \mathbb{Z}$. Let us denote each solution given by (17)-(19) by $m(w^{(1)}, \ldots, w^{(q)}; k)$. Therefore, we have

$$\#\left(w^{(1)}, \ldots, w^{(q)}\right) = (2q)^n \sum_{k \in \mathbb{Z}} \Pr\left[M = m\left(w^{(1)}, \ldots, w^{(q)}; k\right)\right] \ . \tag{20}$$

To establish (16) we will separately bound this sum from below and from above. For the lower bound, first observe that as long as $|k| \leqslant N$, then also $|w_a^{(j)}| \leqslant N$, we can bound absolute values of all elements of the tuple $m\left(w^{(1)}, \ldots, w^{(q)}; k\right)$ with

$$|m_{\mathrm{cycle}}(a)|, |m_{\mathrm{same}}(a)| \leqslant 2qN$$

and consequently obtain bounds on the 2-norm and use (4) to bound the probability in (20):

$$\begin{aligned}
\left\|m\left(w^{(1)}, \ldots, w^{(q)}; k\right)\right\|_2^2 &\leqslant 8q^3N^2 \ , \\
\Pr\left[M = m\left(w^{(1)}, \ldots, w^{(q)}; k\right)\right] &\geqslant \frac{1}{Cn^{(2q-1)/2}} \ .
\end{aligned}$$

Consequently,

$$\begin{aligned}
\#\left(w^{(1)}, \ldots, w^{(q)}\right) &\geqslant (2q)^n \sum_{k \in [-N, N]} \Pr\left[M = m\left(w^{(1)}, \ldots, w^{(q)}; k\right)\right] \\
&\geqslant \frac{(2q)^n N}{Cn^{(2q-1)/2}} \geqslant \frac{(2q)^n}{CN^{2(q-1)}} \ ,
\end{aligned}$$

and rearranging gives the lower bound in (16). For the upper bound, first note that clearly

$$\left\|m\left(w^{(1)},\ldots,w^{(q)};k\right)\right\|_2^2 \geqslant m_{\text{cycle}}(0)^2 = k^2 \ .$$

This time we need to use the more precise error bound from (3). Continuing the computation,

$$\begin{aligned}
\frac{\#\left(w^{(1)},\ldots,w^{(q)}\right)}{(2q)^n} &= \sum_{k\in\mathbb{Z}} \Pr\left[M = m\left(w^{(1)},\ldots,w^{(q)};k\right)\right] \\
&\leqslant \frac{1}{n^{(2q-1)/2}} \cdot \left(\sum_{k\in\mathbb{Z}} O\left(\exp\left(-\frac{k^2}{Cn}\right)\right) + o\left(\min\left(1,\frac{n}{k^2}\right)\right)\right) \\
&\leqslant \frac{1}{n^{(2q-1)/2}} \cdot \left(\sum_{D=0}^{\infty} \sum_{D\sqrt{n}\leqslant k<(D+1)\sqrt{n}} O\left(\exp\left(-D^2/C\right)\right) + o\left(1/D^2\right)\right) \\
&\leqslant \frac{1}{n^{q-1}} \cdot \left(\sum_{D=0}^{\infty} O\left(\exp\left(-D^2/C\right)\right) + o\left(1/D^2\right)\right) \leqslant \frac{C}{N^{2(q-1)}} \ ,
\end{aligned}$$

and another rearrangement of terms finishes the proof. □

Remark 20. Strictly speaking, we did not need slightly more complicated upper bound in (16) to establish that Theorem 17 implies Theorem 13. However, this upper bound allows us to reverse the reasoning and obtain also that Theorem 13 implies Theorem 17. We omit the details, but the proof is a straightforward reversal of the "forward" argument.

5.2 Proof of Theorem 17

To prove Theorem 17 we will need the hypergraph removal lemma originally used in the proof of Szemerédi's theorem. To state the removal lemma we first define hypergraphs and simplices.

Definition 21. A *k-uniform hypergraph* is a pair $H = (V, E)$, where E is a set of subsets of size k (*edges*) of a finite set of *vertices* V. A *k-simplex* is the unique k-uniform hypergraph with $k+1$ vertices and $k+1$ edges.

Theorem 22 ([RS04, NRS06, Gow07])**.** *For every $k \geqslant 2$ and every $\varepsilon > 0$ there exists $\delta := \delta(k,\varepsilon) > 0$ such that for all k-uniform hypergraphs H with N vertices: If H contains at most δN^{k+1} simplices, then it is possible to remove at most εN^k edges from H and obtain a hypergraph that does not contain any simplices.*

Note that a 2-uniform hypergraph is a simple graph, a 2-simplex is a triangle and Theorem 22 restricted to $k = 2$ is the triangle removal lemma. With Theorem 22 we are ready to prove the removal property for feasible arrangements.

Let $R^{(1)},\ldots,R^{(q)} \subseteq \mathbb{Z}_N^{q-1}$. We define a $(q-1)$-uniform hypergraph $H = (X, E)$. The set of vertices of H consists of q disjoint parts $X = X^{(1)} \cup \ldots \cup X^{(q)}$ with each of the parts indexed by $\mathbb{Z}_N^{q-1}$. Therefore, H has qN^{q-1} vertices.

The edge set also consists of q disjoint parts $E = E^{(1)} \cup \ldots \cup E^{(q)}$ such that

$$E^{(j)} \subseteq X^{(1)} \times \ldots \times X^{(j-1)} \times X^{(j+1)} \times \ldots \times X^{(q)} .$$

Therefore, every simplex in H must contain one vertex from each $X^{(j)}$ and one edge from each $E^{(j)}$.

Recall that a vertex $x^{(j)} \in X^{(j)}$ is of the form $x^{(j)} = \left(x^{(j)}_1, \ldots, x^{(j)}_{q-1}\right) \in \mathbb{Z}_N^{q-1}$. To define the edges of H it will be useful to let $x^{(j)}_0 := -\sum_{a=1}^{q-1} x^{(j)}_a$. With that in mind, we say that

$$\left(x^{(1)}, \ldots, x^{(j-1)}, x^{(j+1)}, \ldots, x^{(q)}\right) \in E^{(j)} \iff \left(w^{(j)}_1, \ldots, w^{(j)}_{q-1}\right) \in R^{(j)} ,$$

where

$$w^{(j)}_a := \sum_{t=1}^{q-1} \sum_{b=a-t+1}^{a} x^{(j-t)}_b . \tag{21}$$

In the expression above the indices b and $j-t$ are understood to "wrap around" modulo q. To clarify by example (which might be useful to keep in mind throughout the proof), for $q = 4$, $j = 2$ and $a = 1$ we get

$$w^{(2)}_1 = x^{(1)}_1 + x^{(4)}_0 + x^{(4)}_1 + x^{(3)}_3 + x^{(3)}_0 + x^{(3)}_1 .$$

We use the tuple arrangement $\left(w^{(j)}_1, \ldots, w^{(j)}_{q-1}\right) \in R^{(j)}$ defined in (21) as a label of the edge $\left(x^{(1)}, \ldots, x^{(j-1)}, x^{(j+1)}, \ldots, x^{(q)}\right) \in E^{(j)}$.

We now proceed to checking that simplices in H and feasible arrangements in $R^{(1)} \times \ldots \times R^{(q)}$ correspond to each other. To that end we start with some preparation. First, observing that by definition $\sum_{a=0}^{q-1} x^{(j)}_a = 0$, we can rewrite (21) as

$$w^{(j)}_a = \sum_{t=0}^{q-1} \sum_{b=a-t+1}^{a} x^{(j-t)}_b , \qquad a = 1, \ldots, q-1, \ j = 1, \ldots, q . \tag{22}$$

Let $w^{(j)}_0 := -\sum_{a=1}^{q-1} w^{(j)}_a$ and check that (22) naturally extends to

$$w^{(j)}_0 = \sum_{t=0}^{q-1} \sum_{b=-t+1}^{0} x^{(j-t)}_b .$$

We now make two claims going from simplices to feasible arrangements and vice versa.

Claim 23. *If some q vertices $x^{(1)}, \ldots, x^{(q)}$ of the hypergraph H form a simplex, then the corresponding arrangement formed by edge labels $\left(w^{(1)}, \ldots, w^{(q)}\right) \in R^{(1)} \times \ldots \times R^{(q)}$ is feasible.*

Proof. The feasibility requirement from (14) and (15) can be rewritten using $w_0^{(j)}$ as

$$\forall j = 1, \ldots, q-2 : \forall a = 1, \ldots, q-1 : w_a^{(j+2)} = -w_{a-1}^{(j)} + w_{a-1}^{(j+1)} + w_a^{(j+1)} \ . \tag{23}$$

To verify (23) we compute (taking special care for $t \in \{0, q-1\}$ and still keeping in mind $\sum_{a=0}^{q-1} x_a^{(j)} = 0$)

$$\begin{aligned} -w_{a-1}^{(j)} + w_{a-1}^{(j+1)} + w_a^{(j+1)} &= \sum_{t=0}^{q-1} \left(-\sum_{b=a-t}^{a-1} x_b^{(j-t)} + \sum_{b=a-t-1}^{a-1} x_b^{(j-t)} + \sum_{b=a-t}^{a} x_b^{(j-t)} \right) \\ &= \sum_{t=0}^{q-1} \sum_{b=a-t-1}^{a} x_b^{(j-t)} = \sum_{t=0}^{q-1} \sum_{b=a-t+1}^{a} x_b^{(j+2-t)} = w_a^{(j+2)}. \qquad \square \end{aligned} \tag{24}$$

Claim 24. *For every feasible arrangement $w^{(1)}, \ldots, w^{(q)} \in R^{(1)} \times \cdots \times R^{(q)}$ there exist exactly $N^{(q-1)(q-2)}$ simplices in H labeled with $w^{(1)}, \ldots, w^{(q)}$. Furthermore, these simplices are edge disjoint.*

Proof. We will show that every $x^{(2)}, \ldots, x^{(q-1)} \in \mathbb{Z}_N^{(q-1)(q-2)}$ can be extended to a simplex $x^{(1)}, \ldots, x^{(q)}$ labeled with the feasible arrangement $w^{(1)}, \ldots, w^{(q)} \in R^{(1)} \times \ldots \times R^{(q)}$. First, by inspection we see that for fixed $x^{(2)}, \ldots, x^{(q-1)}, w^{(1)}, \ldots, w^{(q)}$ the value of $x^{(1)}$ can be determined from the formula for $w^{(2)}$ given by (21). Similarly, the value of $x^{(2)}$ can be determined from (21) for $w^{(1)}$.

We still need to check that the vertices $x^{(1)}, \ldots, x^{(q)}$ obtained in this way satisfy (21) for $j = 3, \ldots, q$. But this follows by induction, using (23) (recall that arrangement $w^{(1)}, \ldots, w^{(q)}$ is feasible) and a rearrangement of the computation in (24).

To argue that the simplices are edge disjoint, we first observe that two different simplices $x^{(1)}, \ldots, x^{(q)}$ and $y^{(1)}, \ldots, y^{(q)}$ with the same label $w^{(1)}, \ldots, w^{(q)}$ have to differ on at least two vertices. This is because for two simplices $x^{(1)}, \ldots, x^{(j-1)}, x, x^{(j+1)}, \ldots, x^{(q)}$ and $x^{(1)}, \ldots, x^{(j-1)}, y, x^{(j+1)}, \ldots, x^{(q)}$ with $x_a \neq y_a$, the formula (21) implies that

$$w_a^{(j+1)} \left(x^{(1)}, \ldots, x, \ldots, x^{(q)} \right) \neq w_a^{(j+1)} \left(x^{(1)}, \ldots, y, \ldots, x^{(q)} \right) \ .$$

Therefore, any two $(q-1)$-hyperedges of simplices $x^{(1)}, \ldots, x^{(q)}$ and $y^{(1)}, \ldots, y^{(q)}$ with the same label must differ on at least one vertex. $\square$

With the two claims the proof is almost finished: For $\mu > 0$ we let $\mu' := \mu / q^{q+1}$ and take $\delta := \delta(q-1, \mu')$ from the hypergraph removal lemma. Let $R^{(1)}, \ldots, R^{(q)} \subseteq \mathbb{Z}_N^{q-1}$ be such that the product $W^{(1)} \times \ldots \times W^{(q)}$ contains at most $\delta N^{2(q-1)}$ feasible arrangements. By Claims 23 and 24, the hypergraph H contains at most $\delta N^{q(q-1)} = \frac{\delta}{q^q} \cdot |X|^q \leqslant \delta |X|^q$ simplices. By the hypergraph removal lemma, it is possible to remove at most $\mu' |X|^{q-1} = \mu' q^{q-1} N^{(q-1)^2}$ edges from H to make it simplex-free. Let $\tilde{E}$ be the set of removed edges.

We define

$$Z^{(j)} := \left\{ w^{(j)} \in R^{(j)} : \text{at least } N^{(q-1)(q-2)}/q \text{ edges in } \tilde{E} \text{ are labeled with } w^{(j)} \right\}$$

and let $V^{(j)} := R^{(j)} \setminus Z^{(j)}$. Observe that $\left|Z^{(j)}\right| \leqslant \mu' q^q N^{q-1} = \mu/q \cdot N^{q-1}$, and therefore indeed the set of removed arrangements has total density at most μ.

We argue that the product $V^{(1)} \times \cdots \times V^{(q)}$ does not contain a feasible arrangement. Indeed, let $w^{(1)}, \ldots, w^{(q)}$ be a feasible arrangement in $R^{(1)} \times \cdots \times R^{(q)}$. By Claim 24, the hypergraph H contains $N^{(q-1)(q-2)}$ edge disjoint simplices labeled with $w^{(1)}, \ldots, w^{(q)}$. Since those simplices disappear from H after removing $\tilde{E}$, each of them must intersect $\tilde{E}$ on at least one edge. By averaging, there must exist j such that $\tilde{E}$ contains at least $N^{(q-1)(q-2)}/q$ edges labeled with $w^{(j)}$. But that implies that $w^{(j)}$ was removed from $R^{(j)}$ and the arrangement $w^{(1)}, \ldots, w^{(q)}$ does not occur in $V^{(1)} \times \cdots \times V^{(q)}$. □

Acknowledgements

I am grateful to Thomas Holenstein for suggesting the problem and to Elchanan Mossel for helpful discussions and encouragement to write this paper. I also thank the referee for helpful recommendations.

References

[AM13] Per Austrin and Elchanan Mossel. Noise correlation bounds for uniform low degree functions. *Arkiv för Matematik*, 51(1):29–52, 2013.

[BCC+17] Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. On cap sets and the group-theoretic approach to matrix multiplication. *Discrete Analysis*, 3, 2017.

[BR10] Rabi N. Bhattacharya and R. Ranga Rao. *Normal Approximation and Asymptotic Expansions*. Society for Industrial and Applied Mathematics, 2010.

[BX15] Arnab Bhattacharyya and Ning Xie. Lower bounds for testing triangle-freeness in Boolean functions. *Computational Complexity*, 24(1):65–101, 2015.

[CLP17] Ernie Croot, Vsevolod F. Lev, and Péter Pál Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, 185:331–337, 2017.

[CM12] Brian Cook and Ákos Magyar. On restricted arithmetic progressions over finite fields. *Online Journal of Analytic Combinatorics*, 7(1):1–10, 2012.

[DFR08] Irit Dinur, Ehud Friedgut, and Oded Regev. Independent sets in graph powers are almost contained in juntas. *Geometric & Functional Analysis GAFA*, 18(1):77–97, 2008.

[EG17] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics*, 185:339–343, 2017.

[FK91] Harry Furstenberg and Yitzhak Katznelson. A density version of the Hales-Jewett theorem. *Journal d'Analyse Mathématique*, 57(1):64–119, 1991.

[FK14] Hu Fu and Robert D. Kleinberg. Improved lower bounds for testing triangle-freeness in Boolean functions via fast matrix multiplication. In *APPROX-RANDOM*, volume 28 of *LIPIcs*, pages 669–676, 2014.

[FL17] Jacob Fox and László Miklós Lovász. A tight bound for Green's arithmetic triangle removal lemma in vector spaces. *Advances in Mathematics*, 321:287–297, 2017.

[FLS18] Jacob Fox, László Miklós Lovász, and Lisa Sauermann. A polynomial bound for the arithmetic k-cycle removal lemma in vector spaces. *Journal of Combinatorial Theory, Series A*, 160:186–201, 2018.

[FR18] Ehud Friedgut and Oded Regev. Kneser graphs are like Swiss cheese. *Discrete Analysis*, 2, 2018.

[Gow98] W. Timothy Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geometric & Functional Analysis GAFA*, 8(3):529–551, 1998.

[Gow01] W. Timothy Gowers. A new proof of Szemerédi's theorem. *Geometric & Functional Analysis GAFA*, 11(3):465–588, 2001.

[Gow07] W. Timothy Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Annals of Mathematics*, 166(3):897–946, 2007.

[Gre05a] Ben Green. *Finite field models in additive combinatorics*, pages 1–28. Surveys in Combinatorics 2005. Cambridge University Press, 2005.

[Gre05b] Ben Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geometric & Functional Analysis GAFA*, 15(2):340–376, 2005.

[GT12] Ben Green and Terence Tao. New bounds for Szemerédi's theorem, Ia: Progressions of length 4 in finite field geometries revisited. arXiv:1205.1330, 2012.

[HHM18] Jan Hązła, Thomas Holenstein, and Elchanan Mossel. Product space models of correlation: Between noise stability and additive combinatorics. *Discrete Analysis*, 20, 2018.

[KSS18] Robert Kleinberg, Will Sawin, and David Speyer. The growth rate of tricolored sum-free sets. *Discrete Analysis*, 12, 2018.

[KSV09] Daniel Král̆, Oriol Serra, and Lluı́s Vena. A combinatorial proof of the removal lemma for groups. *Journal of Combinatorial Theory, Series A*, 116(4):971–978, 2009.

[KSV12] Daniel Král̆, Oriol Serra, and Lluı́s Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics*, 187(1):193–207, 2012.

[Lev18] Vsevolod Lev. Personal communication, 2018.

[LS19] László Miklós Lovász and Lisa Sauermann. A lower bound for the k-multicolored sum-free problem in $\mathbb{Z}_m^n$. *Proceedings of the London Mathematical Society*, 119(1):55–103, 2019.

[Mes95] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *Journal of Combinatorial Theory, Series A*, 71(1):168–172, 1995.

[Mos10] Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric & Functional Analysis GAFA*, 19(6):1713–1756, 2010.

[MOS13] Elchanan Mossel, Krzysztof Oleszkiewicz, and Arnab Sen. On reverse hypercontractivity. *Geometric & Functional Analysis GAFA*, 23(3):1062–1097, 2013.

[Mos20] Elchanan Mossel. Gaussian bounds for noise correlation of resilient functions. *Israel Journal of Mathematics*, 235(1):111–137, 2020.

[Nor19] Sergey Norin. A distribution on triples with maximum entropy marginal. *Forum of Mathematics, Sigma*, 7:e46, 2019.

[NRS06] Brendan Nagle, Vojtěch Rödl, and Mathias Schacht. The counting lemma for regular k-uniform hypergraphs. *Random Structures & Algorithms*, 28(2):113–179, 2006.

[Peb18] Luke Pebody. Proof of a conjecture of Kleinberg-Sawin-Speyer. *Discrete Analysis*, 13, 2018.

[Rot53] Klaus F. Roth. On certain sets of integers. *Journal of the London Mathematical Society*, s1-28(1):104–109, 1953.

[RS04] Vojtěch Rödl and Jozef Skokan. Regularity lemma for k-uniform hypergraphs. *Random Structures & Algorithms*, 25(1):1–42, 2004.

[Sha10] Asaf Shapira. A proof of Green's conjecture regarding the removal properties of sets of linear equations. *Journal of the London Mathematical Society*, 81(2):355–373, 2010.

[Spi76] Frank Spitzer. *Principles of Random Walk*. Springer-Verlag New York, 2nd edition, 1976.

[Wol15] Julia Wolf. Finite field models in arithmetic combinatorics – ten years on. *Finite Fields and Their Applications*, 32(C):233–274, 2015.