Zero subsums in vector spaces over finite fields

Cosmin Pohoata*  Dmitriy Zakharov†

Abstract

The Olson constant $\mathcal{O}_L(\mathbb{F}_p^d)$ represents the minimum positive integer $t$ with the property that every subset $A \subset \mathbb{F}_p^d$ of cardinality $t$ contains a nonempty subset with vanishing sum. The problem of estimating $\mathcal{O}_L(\mathbb{F}_p^d)$ is one of the oldest questions in additive combinatorics, with a long and interesting history even for the case $d = 1$.

In this paper, we prove that for any fixed $d \geq 2$ and $\epsilon > 0$, the Olson constant of $\mathbb{F}_p^d$ satisfies

$$\mathcal{O}_L(\mathbb{F}_p^d) \leq (d - 1 + \epsilon)p$$

for all sufficiently large primes $p$. This settles a conjecture of Hoi Nguyen and Van Vu.

1 Introduction

For a subset $A$ of an additive group $G$, consider the set of all nonempty subsums

$$\Sigma^*(A) := \left\{ \sum_{x \in B} x \mid B \subset A, B \neq \emptyset \right\}.$$  

The Olson constant $\mathcal{O}_L(G)$ represents the minimum $t$ such that every subset $A \subset G$ of cardinality $t$ satisfies $0 \in \Sigma^*(A)$. This is a well-known quantity in additive combinatorics, which is notoriously difficult to estimate even for the most basic groups. Its nice history begins in 1964 with Erdős and Heilbronn, who in [7] proved that there exists an absolute constant $c$ such that $\mathcal{O}_L(\mathbb{F}_p) \leq c\sqrt{p}$, where $p$ is an odd prime. In the same paper, they conjectured that their result should generalize to arbitrary additive groups $G$ and that the optimal constant $c$ in the inequality above is probably $c = \sqrt{2}$. A few years later Szemerédi [21] settled the former conjecture in the affirmative. The result of Erdős and Heilbronn for $\mathbb{F}_p$ and Szemerédi’s theorem for general groups were both later refined by Olson in [16], [17] and [18], who proved that $\mathcal{O}_L(G) \leq 2\sqrt{|G|}$ and also introduced a remarkable group ring approach (which has also recently resurfaced in the context of the polynomial method developments around the cap set problem; see [19] and [20]). Olson’s result was subsequently pushed further by Hamidoune and Zemor [10], who proved that $\mathcal{O}_L(G) \leq \sqrt{2|G|} + O(|G|^{1/3} \log |G|)$, and among other things established the correct order of growth for $\mathcal{O}_L(\mathbb{F}_p)$, up to lower order terms. In 2008, Nguyen, Szemerédi and Vu [14] finally removed the lower terms in the primordial case $G = \mathbb{F}_p$, therefore proving the optimal inequality $\mathcal{O}_L(\mathbb{F}_p) \leq \sqrt{2p}$ for all sufficiently large primes $p$. This work was also further refined in two separate rounds by Balandraud in [3] and [4], who finally gave a short alternative argument which works for all odd primes $p$ based on the quantitative Combinatorial Nullstellensatz introduced by Karasev and Petrov in [12].

*Department of Mathematics, Yale University, USA. Email: andrei.pohoata@yale.edu.
†Laboratory of Combinatorial and Geometric Structures, MIPT, Russia. Email: zakharov2k@gmail.com.
In this paper, we address the problem for $G = \mathbb{F}_p^d$, where $p$ is an odd prime number, $d \geq 2$, and $\mathbb{F}_p^d$ denotes as usual the vector space of $d$-dimensional vectors with coordinates from $\mathbb{F}_p$. The situation in higher dimensions has been traditionally known to be much more complicated. For $d = 2$, the first important result only appeared in 2004. In [9], Gao, Ruzsa and Thangadurai proved that $\mathcal{O}L(\mathbb{F}_p^2) = p + O(L(\mathbb{F}_p)) - 1$ holds for all primes $p > 4.67 \times 10^{34}$, thus establishing a beautiful connection between $\mathcal{O}L(\mathbb{F}_p^2)$ and $\mathcal{O}L(\mathbb{F}_p)$. In particular, given the successful story for $\mathbb{F}_p$, this result also determines the Olson constant constant of $\mathbb{F}_p^2$ for large primes. For higher dimensions, however, not much more is known. In the same paper [9], Gao, Ruzsa and Thangadurai conjectured that

$$\mathcal{O}L(\mathbb{F}_p^d) = p + O(L(\mathbb{F}_p^{d-1})) - 1$$

should hold in general for all $d \geq 2$ and for all sufficiently large primes $p$, but this is still a (difficult) open problem. It is also perhaps worth mentioning the curiosity that the assumption that $p$ is sufficiently large is necessary this time around, see for instance the discussion from [8]. In 2011, Nguyen and Vu [15] also studied this higher dimensional problem and proposed the following asymptotic version of the conjecture: for any fixed $d \geq 2$ and $\epsilon > 0$, the Olson constant of $\mathbb{F}_p^d$ satisfies the inequality

$$\mathcal{O}L(\mathbb{F}_p^d) \leq (d - 1 + \epsilon)p$$

for all sufficiently large primes $p$. Since $\mathcal{O}L(\mathbb{F}_p) = O(\sqrt{p}) = o(p)$, it is clear that (1) implies (2), but in some sense (2) should still capture all of the difficulties around (1) when $d \geq 3$. Extending an elegant alternative approach they found for the case $d = 2$ of the Gao-Ruzsa-Thangadurai conjecture, Nguyen and Vu then established this asymptotic conjecture when $d = 3$; however, their argument has various serious limitations already starting with $d \geq 4$, and so no further progress has been made since.

Our main result is a resolution of this conjecture of Nguyen and Vu in all dimensions $d \geq 2$ by using a new approach inspired by the second author’s recent work on the Erdős-Ginzburg-Ziv problem [22].

**Theorem 1** For any fixed $d \geq 2$ and $\epsilon > 0$, the Olson constant of $\mathbb{F}_p^d$ satisfies the inequality

$$\mathcal{O}L(\mathbb{F}_p^d) \leq (d - 1 + \epsilon)p$$

for all sufficiently large primes $p$.

We include the proof of Theorem 1 in Section 3, after discussing terminology and the required preliminary results in Section 2.

Before we move on however to the technical details, we end this section with a high-level overview of the argument. Starting with a set $X \subset \mathbb{F}_p^d$ of size $(d - 1 + \epsilon)p$, where $p$ is a sufficiently large prime number, the first important idea is to prove that one can reduce the problem of finding a vanishing subsum in $\Sigma^*(X)$ to the case when $X$ lies in a translate of the form $v + [-K, K]^l \times \mathbb{F}_p^d$, for some $l \in \{1, \ldots, d - 1\}$, $v \in \mathbb{F}_p^d$, and where $[-K, K]$ stands for the interval $\{ -K, -(K - 1), \ldots, (K - 1), K \}$—regarded as a subset of $\mathbb{F}_p$ (whose size does not depend on $p$). The second idea is that if $0 \notin \Sigma^*(X)$, then one can also force $X$ to satisfy some further refined structural properties. Roughly, we’ll be able to assume among other things, for example, that $X$ must always have some positive proportion of its elements outside the set $\{ x \in \mathbb{F}_p^d \mid \xi(x) \in [-K, K] \}$, for every linear function $\xi : \mathbb{F}_p^d \to \mathbb{F}_p$ (except for some trivial cases). The absence of “linear concentration” is crucial because the third main idea is to consider the projection of this structured set $X$ onto the first $l$ coordinates. The image $Y$ of this projection is a large multiset in $\mathbb{F}_p^l$, so we can make use of tools such as the Combinatorial
Nullstellensatz to find a suitable subsequence whose sum of elements vanishes and whose elements have various prescribed multiplicities. Finally, in order to close the argument, we then need to use the rich structure of $X$ to lift this auxiliary zero sum subsequence in $Y$ from the previous step up to an actual proper subset of $X$ whose sum of elements vanishes.

While sharing a rather similar philosophy with the method of Nguyen and Vu from [15] (where projection is also important), finding the right framework to project and lift in order to capture higher dimensional (additive) information and establishing the precise structural results which allow our procedure to go through for all dimensions $d \geq 2$ requires several new ideas, with both algebraic and probabilistic ingredients.

## 2 Preliminaries

A function $\xi$ on the space $\mathbb{F}_p^d$ is called **linear** if it has the form

$$\xi(x_1, \ldots, x_d) = a_0 + a_1 x_1 + \ldots + a_d x_d,$$

for some $a_i \in \mathbb{F}_p$. Linear functions $\xi_1, \ldots, \xi_l$ are called linearly independent if their “linear parts”, i.e. the vectors $(a_1, \ldots, a_d)$ from $[\mathbb{R}]$, are linearly independent in $\mathbb{F}_p^d$.

If $\xi$ is a linear function and $K \in \mathbb{N}$ we let

$$H(\xi, K) := \left\{ x \in \mathbb{F}_p^d \mid \xi(x) \in [-K, K] \right\},$$

where $[-K, K]$ stands for the interval $\{-K, -(K-1), \ldots, (K-1), K\}$, regarded as a subset of $\mathbb{F}_p$.

Given a linear function $\xi : \mathbb{F}_p^d \to \mathbb{F}_p$ and a multiset $X$ in $\mathbb{F}_p^d$, for $K \in \mathbb{N}$ and $\delta \geq 0$ we say that $X$ is $(K, \delta)$-thick along $\xi$ if $|X \setminus H(\xi, K)| \geq \delta|\mathbb{F}_p^d|$, where the cardinality $|\mathbb{F}_p^d|$ is calculated with multiplicities.

Sometimes it will be convenient to use functions $f : 2^d \to \mathbb{R}_+$ instead of multisets $X \subset \mathbb{F}_p^d$. For a subset $S \subset \mathbb{F}_p^d$ we denote by $f(S)$ the sum of values of $f$ on elements of $S$. We say that a function $f$ is $(K, \delta)$-thick along a linear function $\xi$ if $f(H(\xi, K)) \leq (1 - \delta)f(\mathbb{F}_p^d)$.

If $A$ is a multiset then a submultiset (or subset for shortcut) of $A$ is a multiset $B$ such that the multiplicity of every element in $B$ is at most the corresponding multiplicity in $A$. For technical reasons, we will also need a slightly different notion. A labeled subset of a multiset $A$ is a set $B \subset A \times \mathbb{N}$ such that every element $(a, n) \in B$ has the property that the multiplicity of $a$ in $A$ is at least $n$. So every finite multiset $A$ has exactly $2^{|A|}$ labeled subsets. Abusing notation, we sometimes also denote by $B$ the corresponding submultiset of $A$. A pair $(a, n) \in A \times \mathbb{N}$ where $n$ is at most the multiplicity of $a$ in $A$ is called a labeled element of $A$.

In what follows, $\delta \in (0, 1)$, $K'$ and $K$ are natural numbers, and for $n \in \mathbb{N}$ we denote by $[n]$ the set $\{1, \ldots, n\}$.

### 2.1 Tube decomposition

**Definition:** A set $X \subset \mathbb{F}_p^d$ is called $(K, K', \delta)$-tubular if there exists $l \in [0, d]$ and an affine isomorphism $\psi : \mathbb{F}_p^d \to \mathbb{F}_p^d$ such that $\psi X$ is contained in the set $[-K, K]^l \times \mathbb{F}_p^{d-l}$ and satisfies the following property: $\psi X$ is $(K', \delta)$-thick along any linear function $\xi$ which is not constant on $\{0\} \times \mathbb{F}_p^{d-l}$.
So, for instance, if \( l = 0 \) then \( X \) is \((K',\delta)\)-thick along any non-constant linear function. If \( l = d \) then, after an appropriate change of coordinates, \( X \) is contained in the box \([-K,K]^d \subset \mathbb{F}_p^d\). In general, we allow some combination of the above situations.

Note that a set \( X \) is \((0,K,\delta)\)-tubular if it is \((K,\delta)\)-thick along any linear map \( \xi \) which is not constant on the affine hull of \( X \). Indeed, the space \([0] \times \mathbb{F}_p^{d-l}\) from the definition above must coincide with the affine hull of \( \psi X \). We shall say that in this case \( X \) is \((K,\delta)\)-thick in its affine hull.

**Lemma 1** Let \( g : \mathbb{N} \to \mathbb{N} \) be any increasing function and let \( K_0 \in \mathbb{N} \), \( \delta \in (0,1) \), and \( d \in \mathbb{N} \) be such that \( \delta < 2^{-d-1} \). For any set \( X \subset \mathbb{F}_p^d \), there exists \( Y \subset X \) of size at least \((1 - 2^{d+1}\delta)|X|\) which is \((K,g(K),\delta)\)-tubular where \( K = g^l(K_0) \) for some \( l \in [0,d] \).

**Proof of Lemma 1** Let \( \xi_1,\ldots,\xi_l \) be a maximal sequence of linearly independent linear functions such that \( X \) is not \((g^i(K_0),2^i\delta)\)-thick along \( \xi_i \) for any \( i = 1,\ldots,l \). Note that we can have \( l = 0 \), in which case we put \( g^0(K_0) = K_0 \).

Consider \( K = g^l(K_0) \) and let

\[
Y = X \cap \bigcap_{i=1}^l H(\xi_i,K).
\]

By the definition of the \( \xi_i \)'s we have

\[
|Y| \geq |X| - |X| \sum_{i=1}^l 2^i\delta \geq |X|(1 - (2^{l+1} - 1)\delta).
\]

Moreover, for any linear function \( \eta \) which is linearly independent from \( \xi_1,\ldots,\xi_l \), note that the set \( X \) is \((g(K),2^{l+1}\delta)\)-thick along \( \eta \). Consequently, by \([\square]\), the set \( Y \subset X \) is \((g(K),\delta)\)-thick along \( \eta \). Moreover, after an appropriate change of coordinates, we have \( Y \subset [-K,K]^l \times \mathbb{F}_p^{d-l} \), where the copy of \( \mathbb{F}_p^{d-l} \)

arises as the \((d-l)\)-dimensional intersection of the kernels of the maps \( \xi_1,\ldots,\xi_l \).

\[ \square \]

**Lemma 2** For any increasing function \( g : \mathbb{N} \to \mathbb{N} \), \( K_0 \geq 0 \), \( \varepsilon > 0 \), and \( d \geq 1 \) there is some \( N = N(K_0,d,\varepsilon,g) \in \mathbb{N} \) and functions \( \mu(\cdot,\varepsilon,d),\delta(\cdot,\varepsilon,d) : \mathbb{N} \to \mathbb{R}_{>0} \) such that the following holds: For any (multi-)set \( X \subset \mathbb{F}_p^d \) there is \( l \in [0,N] \) and a decomposition

\[
X = X_0 \cup X_1 \cup \ldots \cup X_m,
\]

such that \( |X_0| \leq \varepsilon|X| \) and for any \( i \in [m] \) we have \( |X_i| \geq \mu|X| \) and \( X_i \) is \((g(K),\delta)\)-thick in its affine hull. Here \( K = g^l(K_0) \), \( \mu = \mu(K,\varepsilon,d) \) and \( \delta = \delta(K,\varepsilon,d) \).

**Proof of Lemma 2** The proof is by induction on \( d \). Take an arbitrary (multi-)set \( X \subset \mathbb{F}_p^d \). If \( X \) is \((g(K_0),\varepsilon/2)\)-thick along any non-constant linear map, then there is nothing to prove since we can take the decomposition \( X_0 = \emptyset \), \( X_1 = X \) and \( l = 0 \). So we may assume that \( X \) is not \((g(K_0),\varepsilon/2)\)-thick. Then after a change of coordinates and removing at most \( \varepsilon/2|X| \) elements from \( X \) we may assume that \( X \subset [-g(K_0),g(K_0)] \times \mathbb{F}_p^{d-1} \). For each \( y \in [-g(K_0),g(K_0)] \) let \( X_y = X \cap \{y\} \times \mathbb{F}_p^{d-1} \). Remove from \( X \) all sets \( X_y \) such that \( |X_y| < \varepsilon|X|/8g(K_0) \), so that the size of \( X \) will decrease at most by \( \varepsilon|X|/4 \). Denote \( \varepsilon' = \varepsilon/8g(K_0) \).
Now we are going to apply the induction hypothesis to each of the remaining sets \(X_{y_i}\). Let \(X_{y_1}, \ldots, X_{y_r}\) be the list of all these sets, where \(r \leq 2g(K_0)\). We apply induction to \(X_{y_1}\) with \(\varepsilon = \varepsilon', K'_0 = g(K_0)\) and \(g = g^{N_1}\) where \(N_1\) will be determined later. In order to the number \(K = g^{N_1}(K'_0)\) from the induction hypothesis to lie in the interval \([K_0, g^N(K_0)]\) we need the following inequality:

\[ N > N(K'_0, d - 1, \varepsilon', g^{N_1}) \cdot N_1. \tag{6} \]

So we get a decomposition of the form \(X_{y_1} = \bigcup_{i=0}^{m_1} X_{1,i}\) and there is some \(l_1 \leq N(K'_0, d - 1, \varepsilon', g^{N_1})\) so that if we let \(K_1 = g^{l_1, N_1}(K'_0)\) then \(X_{1,i}\) is \((g^{N_1}(K_1), \delta_1)\)-thick in its affine hull for every \(i \in [m_1]\).

Now we apply the induction hypothesis to the set \(X_{y_2}\) with \(\varepsilon = \varepsilon', K_0 = K_1\) and \(g = g^{N_2}\) where \(N_2\) will be determined later. To apply induction we need the following inequality:

\[ N_1 > N(K_1, d - 1, \varepsilon', g^{N_2}) \cdot N_2. \]

We thus will obtain a decomposition \(X_{y_2} = \bigcup_{i=0}^{m_2} X_{2,i}\) where \(X_{2,i}\) is \((g^{N_2}(K_2), \delta_2)\)-thick in its affine hull where \(K_2 = g^{N_2 l_2}(K_1)\) for some \(l_2 \leq N(K_1, d - 1, \varepsilon', g^{N_2})\). Moreover, we have \(\delta_2 \gg K_2, \varepsilon/8 K_0, d\).

Now observe that we have the following chain of inequalities:

\[ K_1 \leq K_2 \leq g^{N_2}(K_2) \leq g^{N_1}(K_1). \]

Thus, for every \(i \in [m_1]\), the sets \(X_{1,i}\) is \((g^{N_2}(K_2), \delta_1)\)-thick in its affine hull. Also note that since \(K_2 \geq K_1\) we have \(\delta_1 \gg K_{2, \varepsilon', d}\) as well.

Applying induction in a similar manner \(r - 2\) more times to sets \(X_{y_i}\) for \(j = 3, \ldots, r\) we will eventually get some \(K_r = g^{l_r}(K_0)\) where \(l \leq N\) such that all sets \(X_{i,j}\) are \((g(K_r), \delta_r)\)-thick in their affine hulls for some \(\delta_r \gg K_r, \varepsilon, d\). We will get a chain of inequalities of the form \([5]\) which will give an upper bound on the function \(N(K_0, d, \varepsilon, g)\). This concludes the proof.

\[ \square \]

**Lemma 3** For any increasing function \(g : \mathbb{N} \rightarrow \mathbb{N}, K_0 \geq 0, \varepsilon > 0,\) and \(d \geq 1\) there is some \(N = N(K_0, d, \varepsilon, g) \in \mathbb{N}\) and functions \(\mu(\cdot, \varepsilon, d), \delta(\cdot, \varepsilon, d) : \mathbb{N} \rightarrow \mathbb{R}_{>0}\) such that the following holds: For any set \(X \subset \mathbb{F}_p^d\) there is \(l \in [0, N]\) and a decomposition

\[ X = X_0 \cup X_1 \cup \ldots \cup X_m, \]

such that \(|X_0| \leq \varepsilon|X|\) and for any \(i \in [m]\) we have \(|X_i| \geq \mu|X|\) and \(X_i\) is \((g^{d+1}(K), \delta)\)-thick in its affine hull. Here \(K = g^l(K_0), \mu = \mu(K, \varepsilon, d)\) and \(\delta = \delta(K, \varepsilon, d)\).

Moreover, for any \(S \subset [m]\) the set \(X_S = \bigcup_{i \in S} X_i\) is \((K_S, g(K_S), \mu)\)-tubular where \(K_S = g^s(K)\) for some \(s \in [0, d]\).

**Proof of Lemma 3**: Let \(g' = g^{d+1}\) and apply Lemma 2 to \(X\) with \(g'\) instead of \(g\), \(\varepsilon/2\) instead of \(\varepsilon\) and \(K_0 = K_0\). We get a decomposition of the form \([5]\) where sets \(X_i\) are \((g^{d+1}(K), \delta_0)\)-thick in their affine hulls for some \(\delta_0 \gg K, \varepsilon, d\). Also we have \(|X_i| \geq \mu_0|X|\) for some \(\mu_0 \gg K, \varepsilon, d\).

Let \(S_1, S_2, \ldots, S_{2^m - 1}\) be the list of all non-empty subsets of \([m]\) in any order. For \(j = 1, \ldots, 2^m - 1\), apply Lemma 1 consecutively to sets \(\bigcup_{i \in S} X_i\) with \(K_0 = K\), \(g = g\) and

\[ \delta_j = \varepsilon \mu_0 \delta_0 2^{-d-2} m 2^{-(d+m+4)j}. \]

Such choice of \(\delta_j\) will guarantee us the following properties:
1. The total number $R$ of removed elements will be at most

$$R \leq 2^{d+1}|X| \sum_{j=1}^{2^{m-1}} \delta_j < \varepsilon|X|/2.$$ 

2. All sets $X_i$ will be $(y^{d+1}(K), \delta_0/2)$-thick since for every $i \in [m]$:

$$R \leq 2^{d+1}|X| \sum_{j=1}^{2^{m-1}} \delta_j \leq \varepsilon \mu_0 |X| \delta_0/2 \leq |X| \delta_0/2.$$ 

3. For any $j$ the set $X_S = \bigcup_{i \in S_j} X_i$ will be $(K_S, g(K_S), \delta_j/2)$-tubular because the number of elements removed from $X_S$ at steps $j' > j$ is at most

$$2^{d+1}|X| \sum_{j' > j} \delta_{j'} \leq \varepsilon \mu_0 \delta_0 2^{-(d+m+4) j} |X| \sum_{k=1}^{\infty} 2^{-(d+m+4) k} \leq \varepsilon \mu_0 \delta_0 2^{-(d+m+3) j} |X| 2^{-d-2-m}/2 = \delta_j/2.$$ 

We clearly have $\delta_{2^{m-1}} \gg \varepsilon, \mu_0, \delta_0, m$. But $m \leq 1/\mu_0$ and so $\delta \gg K_\varepsilon \mu_0$. The lemma is proved.

2.2 From tubes to subset sums

The main auxiliary result in this section is the following Proposition inspired by the ideas from [22, Section 7.2].

**Proposition 1** Let $d \geq 1$, $K \geq 1$, $\delta > 0$ and $\mu > 0$ and let $K_2 > K$ be sufficiently large with respect to parameters $K, d, \delta, \mu$. Let $p > p_0(d, K, \delta, \mu)$ be a sufficiently large prime.

Fix some $l \in [0, d]$ and let $Y \subset [-K, K]^l$ be a non-empty set. For $y \in Y$ let $X_y = \{y\} \times \mathbb{F}_p^{d-l}$ be an arbitrary (multi-)set of size at least $\mu p$. Denote $X = \bigcup_{y \in Y} X_y$. Suppose that $X$ is $(K_2, \delta)$-thick along any linear function which is not constant on $\{0\} \times \mathbb{F}_p^{d-l}$. Then there is some $u_0 \in \mathbb{F}_p^l$ and $k \in \mathbb{N}$ such that for every $u \in \mathbb{F}_p^{d-l}$ there are subsets $S_y \subset X_y$ such that:

$$\sum_{y \in Y} \sum_{x \in S_y} x = (u_0, u),$$

and $\sum_{y \in Y} |S_y| = k$.

We present the proof of Proposition 1 below. The argument is based on the following lemma which was essentially proved by Alon and Dubiner [22, Corollary 2.3 and Proposition 2.4]. See also [22, Lemmas 3.1 and 3.2].

**Lemma 4** Let $A \subset \mathbb{F}_p^d$ be a multiset which is $(K, \delta)$-thick along any linear function $\xi$ without constant term, for some $K \geq 0$ and $\delta > 0$. Then for any set $Y \subset \mathbb{F}_p^d$ of size at most $p^d/2$ there is $a \in A$ such that

$$|(Y + a) \setminus Y| \geq \max \left\{ \frac{|Y|^{d+1}}{2^d}, \frac{K \delta}{c_0 p} |Y| \right\},$$

where $c_0 \leq 10^{10}$ is an absolute constant.
Proof: For the first inequality, we proceed like in the proof of [22, Lemma 3.3] and make use of the so-called Loomis–Whitney inequality [13].

Lemma 5 Let $X \subset \mathbb{R}^d$ be a finite set. Let $X_i$ be the projection of $X$ on the $i$-th coordinate hyperplane $\{(x_1, \ldots, x_d) \mid x_i = 0\}$. Then one has an inequality $|X|^{d-1} \leq \prod_{i=1}^d |X_i|$. 

Let $Y \subset \mathbb{F}_p^d$ and $|Y| = x_0 \leq p^d/2$. Since $A$ is $(K, \delta)$-thick with $\delta > 0$, there are vectors $e_1, \ldots, e_d \in A$ which form a basis of the space $\mathbb{F}_p^d$. Consider the standard embedding of $\mathbb{F}_p^d$ in $\mathbb{Z}^d$ with respect to the basis $e_1, \ldots, e_d$. Lemma 3 applied to the image of $Y$ yields that there is $i \in \{1, \ldots, d\}$ such that $|Y_i| \geq x_0^{d-1}$. This means that at least $x_0^{d-1}$ lines of the form $l_v = \{v + te_i \mid t \in \mathbb{F}_p\}$ intersect $Y$. For any line $l_v$ intersecting $Y$ we have either $|(Y \cup (Y + e_i)) \cap l_v| > |Y \cap l_v|$ or $l_v \not\subset Y$. But the number of the latter lines is at most $|Y|/p$ since these lines are disjoint and contain $p$ points each. Thus, since $x_0 \leq 2^{-1/d}p$,

$$|(Y + e_i) \setminus Y| \geq x_0^{d-1} - x_0^d/p \geq x_0^{d-1}/2^d.$$ 

The second inequality follows from [22, Lemma 3.1].

Proof of Proposition 1: Let us first consider the case $l = 0$. So, $X$ is $(K_2, \delta)$-thick along any non-constant linear function and $|X| \geq \mu p$. Then, the multiset $A = X - X$ is $(K_2, \delta)$-thick along any linear function without constant term. Indeed, suppose that for some linear function $\xi$ more than $(1 - \delta)|X|^2$ differences $(x_1 - x_2)$ belong to $H(\xi, K_2)$. Then, by the pigeonhole principle, there is $x_2 \in X$ such that more than $(1 - \delta)|X|$ vectors $x_1 \in X$ belong to $x_2 + H(\xi, K_2)$. But this contradicts the assumption that $X$ is $(K_2, \delta)$-thick.

Now Lemma 4 can be applied to the multiset $A$. By choosing $K_2$ sufficiently large and applying Lemma 4 iteratively one can construct a sequence of pairwise disjoint sets $\{a_1, b_1\}, \ldots, \{a_t, b_t\} \subset X$ such that

$$\{a_1, b_1\} + \ldots + \{a_t, b_t\} = \mathbb{F}_p^d.$$ 

To see this, note that at each step we can apply Lemma 4 with $Y_i = \{a_1, b_1\} + \ldots + \{a_i, b_i\}$. By the thickness of $X$, if $i \leq 0.1\mu d$ then the set $X \setminus \{a_1, b_1, \ldots, a_i, b_i\}$ is $(K_i^d, \delta/2)$-thick along any linear function without constant term. In this case, by Lemma 4 we have some $x_1 - x_2 \in X - X$ so that

$$|(Y_1 + x_1 - x_2) \setminus Y_1| \geq \max \left\{\frac{|Y_1|^{d-1}}{2^d}, \frac{K_2\delta}{2c_0p}|Y_1|\right\}.$$ 

Since $(Y_1 + x_1) \cup (Y_1 + x_2) = Y_1 \cup (Y_1 + x_1 - x_2) + x_2$, we can thus choose $\{a_{i+1}', b_{i+1}'\} = \{x_1, x_2\}$ such that $Y_{i+1} := Y_i + \{a_{i+1}', b_{i+1}'\}$ is significantly larger than $Y_i$. It can be easily checked that if we take $K_2$ sufficiently large then this process stops after at most $r := 0.01\mu dp$ steps, when we eventually reach a set $Y_r = \{a_1', b_1'\} + \ldots + \{a_r', b_r'\}$ of size $> p^{d/2}$ (unless we cover the whole space $\mathbb{F}_p^d$). On the other hand, if we let $X'$ denote the set $X \setminus \{a_1', b_1', \ldots, a_r', b_r'\}$, then $X'$ is also $(K_2, 0.9\delta)$-thick along any linear function without constant term since $|X \setminus X'| \leq 0.02\delta \mu p \leq 0.1\delta |X|$. Moreover, by a similar argument as above, the difference set $X' - X'$ is also $(K_2, 0.9\delta)$-thick along any linear function without constant term, so one can repeat the procedure to produce a sequence of disjoint pairs $\{a_{1}^{n}, b_{1}^{n}\}, \ldots, \{a_{s}^{n}, b_{s}^{n}\} \subset X'$ such that $Y_s := \{a_1', b_1'\} + \ldots + \{a_s^{n}, b_s^{n}\}$ has size $> p^{d/2}$. By applying the easy case of the Cauchy-Davenport theorem in $\mathbb{F}_p^d$ (for a reference, see for example [3]), we thus have that $Y_r + Y_s = \mathbb{F}_p^d$, so one can just take $t = r + s$ and consider $\{a_1, b_1\}, \ldots, \{a_t, b_t\}$ to be the
concatenation of the two disjoint lists \( \{a'_1, b'_1\}, \ldots, \{a'_r, b'_r\} \) and \( \{a''_1, b''_1\}, \ldots, \{a''_s, b''_s\} \) so that (2.2) is satisfied. Using this construction, the conclusion of Proposition 1 immediately follows with \( k = t \).

Now we consider the general case, that is, \( l \in [d] \) is arbitrary. In this case the multiset \( X - X \subset \mathbb{F}_p^d \) is not necessarily \((K_2, \delta)\)-thick and so we cannot apply Lemma 4. Instead, we are going to construct a certain function \( f : \{0\} \times \mathbb{F}_p^{d-l} \to \mathbb{Q}_+ \) which will be \((K', \delta')\)-thick along any linear function without constant term. Then we will apply Lemma 4 to \( f \) in a similar manner as in the case \( l = 0 \) to conclude the proof.

To define \( f \) let us consider the set \( \Lambda \subset \mathbb{Z}^Y \) consisting of all integer vectors \((\lambda_y)_{y \in Y}\) such that \( \sum \lambda_y y = 0 \) and \( \sum \lambda_y = 0 \). For a pair of vectors \( \lambda^1, \lambda^2 \in \mathbb{N}^Y \) let \( \mathcal{J}^{\lambda^1, \lambda^2} \) be the set of all pairs \((J_1, J_2)\), where \( J_1 \) and \( J_2 \) are disjoint labeled subsets of \( X \) and for every \( y \in Y \) we have

\[
|J_1 \cap X_y| = \lambda^1_y, \\
|J_2 \cap X_y| = \lambda^2_y.
\]

For a pair of (multi-)sets \((J_1, J_2)\) we denote by \( \sigma(J_1, J_2) \) the sum of elements of \( J_1 \) minus the sum of elements of \( J_2 \). Note that if \( \lambda^1 - \lambda^2 \in \Lambda \) then for any \((J_1, J_2) \in \mathcal{J}^{\lambda^1, \lambda^2}\) we have \( \sigma(J_1, J_2) \in \{0\} \times \mathbb{F}_p^{d-l} \) and \( |J_1| = |J_2| \).

For \( \lambda^1, \lambda^2 \in \mathbb{N}^Y \) such that \( \lambda^1 - \lambda^2 \in \Lambda \) let \( f_{\lambda^1, \lambda^2} : \{0\} \times \mathbb{F}_p^{d-l} \to \mathbb{N} \) be the characteristic function of the multiset of all sums \( \sigma(J_1, J_2) \) over \((J_1, J_2) \in \mathcal{J}^{\lambda^1, \lambda^2}\). We define the function \( f : \{0\} \times \mathbb{F}_p^{d-l} \to \mathbb{Q}_+ \) as follows:

\[
f = \sum_{\lambda^1, \lambda^2} f_{\lambda^1, \lambda^2} |\mathcal{J}^{\lambda^1, \lambda^2}|. \tag{7}
\]

where the sum is taken over all vectors \( \lambda^1, \lambda^2 \in \mathbb{N}^Y \) with \( l_1 \)-norm bounded by a sufficiently large constant \( T = T(d, K, \delta, \mu) > 0 \) and such that \( \lambda^1 - \lambda^2 \in \Lambda \).

Note that we have the following expression for the size of \( \mathcal{J}^{\lambda^1, \lambda^2} \), as a product of multinomial coefficients:

\[
|\mathcal{J}^{\lambda^1, \lambda^2}| = \prod_{y \in Y} \binom{|X_y|}{\lambda^1_y, \lambda^2_y}. \tag{8}
\]

It requires some work to show that \( f \) is indeed \((K', \delta')\)-thick along any linear function on \( \{0\} \times \mathbb{F}_p^{d-l} \), so we isolate this fact as a separate lemma.

**Lemma 6** The function \( f \) is \((K', \delta')\)-thick along any linear function without constant term on \( \{0\} \times \mathbb{F}_p^{d-l} \). Here \( K' \) and \( \delta' \) depend on parameters \( K, K_2, d, \delta, \mu \) in such a way that \( K' \) can be arbitrarily large compared to \( K, d, \delta, \mu \) and \( \delta' \) if one takes \( K_2 \) large enough.

A somewhat similar result was proved in [22, Lemma 7.3] but the details are different, so we include below the complete proof for our setup.

**Proof of Lemma 6**: Suppose that \( f \) is not \((K', \delta')\)-thick along some linear function \( \xi : \{0\} \times \mathbb{F}_p^{d-l} \to \mathbb{F}_p \) without constant term. Here we take \( \delta' = \min\{0.01/T, 0.1\delta\} \). The number \( K' \) will be determined at the end of the argument.

We may extend \( \xi \) to a linear function on \( \mathbb{F}_p^d \) which does not depend on the first \( l \) coordinates. Denote by \( S \subset \mathbb{Z}^Y \times \mathbb{Z}^Y \) the set of all pairs \( \lambda^1, \lambda^2 \in \mathbb{N}^Y \) satisfying \( \|\lambda^i\|_1 \leq T \) and \( \lambda^1 - \lambda^2 \in \Lambda \). Then by
we have $f(\{0\} \times \mathbb{F}_p^{d-1}) = |S|$. So we have $f(H(\xi, K')) \geq (1 - \delta')|S|$. Let $S' \subset S$ be the set of all pairs $(\lambda_1, \lambda_2) \in S$ such that $f_{\lambda_1, \lambda_2}(H(\xi, K')) \geq (1 - 2\delta')|\mathcal{J}_{\lambda_1, \lambda_2}|$, that is, the function $f_{\lambda_1, \lambda_2}$ is not $(K', 2\delta')$-thick along $\xi$. Note that

$$\delta'|S| \geq f(\mathbb{F}_p^d \setminus H(\xi, K)) \geq \sum_{(\lambda_1, \lambda_2) \in S \setminus S'} \frac{f_{\lambda_1, \lambda_2}(\mathbb{F}_p^d \setminus H(\xi, K))}{|\mathcal{J}_{\lambda_1, \lambda_2}|} \geq 2\delta'|S \setminus S'|,$$

and so $|S'| \geq |S|/2$.

**Claim 1** For any $y \in Y$ there is $r_y \in \mathbb{F}_p$ such that $|X_y \cap H(\xi - r_y, 2K')| \geq (1 - 10\delta')|X_y|$.

**Proof:** Recall that $\Lambda$ is a lattice defined by equations with coefficients bounded by $K$. Thus, if $T$ is large enough compared to $K$ then any subspace $V \subset \mathbb{R}^Y \times \mathbb{R}^Y$ either contains $S$ or intersects it in at most $|S|/10$ elements. In particular, for any subspace $V$ we have $S' \not\subset V$ or $S \subset V$.

For $y \in Y$ consider the subspace $V = \{(\lambda_1, \lambda_2) \in \mathbb{R}^Y \times \mathbb{R}^Y \mid \lambda_1 = \lambda_2 = 0\}$. Then we clearly have $S \not\subset V$ because the vector $(e_y, e_y)$, where $e_y$ has only one non-zero coordinate in $y$-th place, belongs to $S$. So there is a pair $(\lambda_1, \lambda_2) \in S'$ such that $\lambda_1^i \neq 0$ for some $i \in \{1, 2\}$. Without loss of generality we may assume that $i = 1$.

For labeled elements $x_1, x_2$ of $X_y$ denote by $\mathcal{J}_{x_1, x_2}^{\lambda_1, \lambda_2}$ the subfamily in $\mathcal{J}_{\lambda_1, \lambda_2}$ consisting of all pairs $(J_1, J_2)$ such that $x_1 \in J_1, x_2 \notin J_1$ and $x_1, x_2 \notin J_2$. Define a graph $G$ on the set of labeled elements of $X_y$ where a pair $x_1, x_2$ of labeled elements of $X_y$ forms an edge if $\xi(x_1 - x_2) \notin [-2K', 2K']$. Let $x_1, x_2$ be an edge and take $(J_1, J_2) \in \mathcal{J}_{x_1, x_2}^{\lambda_1, \lambda_2}$. Denote $J'_1 = J_1 \setminus \{x_1\} \cup \{x_2\}$, then we have $(J'_1, J_2) \in \mathcal{J}_{x_2, x_1}^{\lambda_1, \lambda_2}$ and

$$\sigma(J_1, J_2) - \sigma(J'_1, J_2) = x_1 - x_2.$$

Thus, one of the elements $\xi(\sigma(J_1, J_2))$ or $\xi(\sigma(J'_1, J_2))$ does not belong to the interval $[-K', K']$. So at least half of the sums $\sigma(J_1, J_2), \sigma(J'_1, J_2)$ over $(J_1, J_2) \in \mathcal{J}_{x_1, x_2}^{\lambda_1, \lambda_2}$ does not belong to the strip $H(\xi, K')$. We are going to combine this observation with the fact that the multiset of sums $\sigma(J_1, J_2)$ over $(J_1, J_2) \in \mathcal{J}_{x_1, x_2}^{\lambda_1, \lambda_2}$ is not $(K', \delta')$-thick along $\xi$ to conclude that the graph $G$ has a very large independence number.

Indeed, suppose that the independence number of $G$ is at most $|X_y| - r$ for some $r \geq 1$. Then one can find $r$ disjoint edges $\{x_1, y_1\}, \ldots, \{x_r, y_r\}$ in the graph $G$. By an elementary double counting argument, for any $i \neq j$, $i, j \in [r]$, we have

$$|\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2} \cap \mathcal{J}_{x_j, y_j}^{\lambda_1, \lambda_2}| = \frac{(\lambda_1^1 - 1)(|X_y| - \lambda_1^1 - \lambda_2^1 - 1)}{|X_y| - 2} |\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}| \leq \frac{\lambda_1^1}{|X_y|} |\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}|.$$

Applying Bonferroni’s inequality to the sets $\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2} \cup \mathcal{J}_{x_i, x_i}^{\lambda_1, \lambda_2}, i = 1, \ldots, r'$, for some $r' \leq r$, we can estimate the number $N$ of pairs $(J_1, J_2) \in \mathcal{J}_{\lambda_1, \lambda_2}$ such that $\xi(\sigma(J_1, J_2)) \notin [-K', K']$:

$$N \geq |\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}| \left(r' - 4 \binom{r'}{2} \frac{\lambda_1^1}{|X_y|} \right).$$

A double counting argument shows that

$$|\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}| = \frac{\lambda_1^1(|X_y| - \lambda_1^1 - \lambda_2^1)}{|X_y||X_y| - 1} |\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}| \geq \frac{\lambda_1^1}{2|X_y|} |\mathcal{J}_{x_1, y_1}^{\lambda_1, \lambda_2}|.$$
since $|X_y| \geq \mu p$, $\lambda_y^1 \leq \|\lambda^1\|_1 \leq T$ and $p$ is large enough compared to $\mu$ and $T$. We conclude that for any $r' \leq r$:

$$N \geq \frac{\lambda_y^1}{2|X_y|} \left( r' - 2r'^2 \frac{\lambda_y^1}{|X_y|} \right) |J^{\lambda^1,\lambda^2}|.$$  

On the other hand, since $(\lambda^1,\lambda^2) \in S'$ we have $N \leq 2\delta'|J^{\lambda^1,\lambda^2}|$. It is easy to see that the last two inequalities are incompatible if we put $r' = 10\delta|\lambda_y^1/|X_y||$ and note that $\delta' \leq 0.01$.

Thus, the independence number of the graph $G$ is at least $(1 - 10\delta')|X_y|$. Let $Z'_y$ be an independent set of labeled elements of $X_y$ of maximal size. Pick an arbitrary $z \in Z'_y$ and put $r_y = \xi(z) \in \mathbb{F}_p$. Then by the definition of $G$ we have

$$Z'_y \subset H(\xi - r_y, 2K'),$$

and $|Z'_y| \geq (1 - 10\delta')|X_y|$ and the claim follows. \hfill $\square$

Denote $Z_y = X_y \cap H(\xi - r_y, 2K')$ so that we have $|Z_y| \geq (1 - 10\delta')|X_y|$. The next step is to show that the value $r_y$ can be approximated by a linear function of $y \in Y$. This will then contradict the thickness condition on the multiset $X$.

For a given $(\lambda^1,\lambda^2) \in S$ denote by $J_0^{\lambda^1,\lambda^2}$ a subfamily in $J^{\lambda^1,\lambda^2}$ consisting of those pairs $(J_1,J_2)$ such that for every $(x,n) \in J_1 \cup J_2$ we have $x \in \bigcup_{y \in Y} Z_y$. From $[S]$ and a similar expression for the size of $J_0^{\lambda^1,\lambda^2}$ we see that

$$\frac{|J_0^{\lambda^1,\lambda^2}|}{|J^{\lambda^1,\lambda^2}|} \geq 1 - 20\delta'(|\lambda^1|_1 + |\lambda^2|_1) \geq 1 - 40\delta'T.$$  

Here we used the fact that $|X_y| \geq \mu p$ and $p$ is large enough which implies that $\binom{|X_y|}{a,b} = (1+o_p(1))\binom{|X|}{a,b}$ for $a,b \leq T$. Since $\delta' \leq 0.01/T$, if $(\lambda^1,\lambda^2) \in S'$ then there is a pair $(J_1,J_2) \in J_0^{\lambda^1,\lambda^2}$ such that $\sigma(J_1,J_2) \subset H(\xi, K')$.

But $|r_y - x| \leq 2K'$ for every $x \in Z_y$. And so, slightly abusing notation, we can estimate:

$$\left| \sum_{y \in Y} r_y (\lambda_y^1 - \lambda_y^2) \right| \leq 4K'T + \sum_{(x,n) \in J_1} \xi(x) - \sum_{(x,n) \in J_2} \xi(x) \leq 4K'T + K' \leq 5K'T.$$

In other words, for any $(\lambda^1,\lambda^2) \in S'$ we have $\langle r, \lambda^1 - \lambda^2 \rangle \in [-5K'T, 5K'T]$ where $\langle \cdot, \cdot \rangle$ denotes the standard pairing on $\mathbb{F}_p^{Y}$. Using the fact that $|S'| \geq |S|$ and some elementary linear algebra one can find a constant $C \ll_{K',|Y|} 1$ such that for any vector $\lambda \in \Lambda$ with $l_1$-norm at most $T$ we have $\langle r, \lambda \rangle \in [-C,C]$.

Let $Y' \subset Y$ be a minimal set whose affine hull contains $Y$. Then for every $y \in Y \setminus Y'$ there is a unique up to scaling vector $\lambda(y) \in \Lambda$ such that $\langle \lambda(y),y \rangle \neq 0$ and $\lambda(y)\bar{y} = 0$ for any $\bar{y} \not\in Y' \cup \{y\}$. We fix $\lambda(y)$ in such a way that $\lambda(y)_y$ has minimal positive value. Note that if $p$ is large enough then the number $\lambda(y)_y$ is not divisible by $p$. Now we define a new vector $\tilde{r} = (\tilde{r}_y) \in \mathbb{F}_p^{Y}$ as follows. For $y \in Y'$ we let $\tilde{r}_y = r_y$ and for $y \in Y \setminus Y'$ we define

$$\tilde{r}_y = -\frac{1}{\lambda(y)_y} \sum_{y' \in Y'} \lambda(y)_y r_{y'}.$$  

It is now clear that $\lambda(y)_y \tilde{r}_y - \lambda(y)_y r_y \in [-C,C]$ (considered as an element of $\mathbb{F}_p$). Let $\eta : \mathbb{F}_p^l \to \mathbb{F}_p$ be a linear function such that $\eta(y) = r_y$ for $y \in Y'$. Then from linearity we have $\eta(y) = \tilde{r}_y$ for every

$$\eta(y) = \tilde{r}_y$$
Lemma 8. Let \( y \in Y \). Denote \( L = \prod_{y \in Y \setminus y'} \lambda(y)y \). We may view \( \eta \) as a function on \( \mathbb{F}^d_p \) which does not depend on the last \( d-l \) coordinates. Then for any \( x \in Z_y \) we have
\[
|L\xi(x) - L\eta(x)| \leq 2K'L + |Lr_y - L\eta| \leq 2K'L + LC =: B,
\]
i.e. if we let \( \xi' = L\xi - L\eta \) then for any \( x \in Z_y \) we have \( \xi'(x) \in [-B, B] \). Therefore, the multiset \( X \) is not \((B, 10\delta')\)-thick along \( \xi' \). It is easy to see that \( B \) is bounded in terms of parameters \( K', K, |Y| \) and \( d \). Define \( K' \) to be the largest number such that \( B \leq K_2 \). Then it is clear that \( K' \) can be made arbitrarily large compared to \( K, \delta \) and \( \delta' \) if one takes \( K_2 \) sufficiently large. Indeed, recall that \( \delta' = \min\{0.01/T, 0.1\delta\} \) where \( T \) depends on \( K \) and \( d \) only. With these parameters, the multiset \( X \) is not \((K_2, \delta')\)-thick along the linear function \( \xi' \). But \( \xi' \) coincides with \( L\xi \) on \( \{0\} \times \mathbb{F}^d_p \) and therefore is not constant on this subspace. This contradicts to the initial assumption on \( X \).

\[ \square \]

Returning to the proof of Proposition 7 we can now similarly apply Lemma 4 repeatedly to the multiset \( A \) like in the case \( l = 0 \), and construct recursively a sequence of pairs \((J_1^1, J_2^1), \ldots, (J_1^r, J_2^r)\) where all sets \( J_i^t \) are pairwise disjoint as labeled subsets of \( X \) and such that
\[
\{\sigma(J_1^1), \sigma(J_2^1)\} + \ldots + \{\sigma(J_1^t), \sigma(J_2^t)\} = \{u_0\} \times \mathbb{F}^d_p,
\]
for some fixed vector \( u_0 \) (by first producing two pairwise disjoint lists \((J_{1,1}^1, J_{1,2}^1), \ldots, (J_{1,1}^r, J_{1,2}^r)\) and \((J_{2,1}^1, J_{2,2}^1), \ldots, (J_{2,1}^r, J_{2,2}^r)\) such that \(|\{\sigma(J_{1,1}^1), \sigma(J_{1,2}^1)\} + \ldots + \{\sigma(J_{1,1}^r), \sigma(J_{1,2}^r)\}| > p^{d-l}/2\) and \(|\{\sigma(J_{2,1}^1), \sigma(J_{2,2}^1)\} + \ldots + \{\sigma(J_{2,1}^r), \sigma(J_{2,2}^r)\}| > p^{d-l}/2\), and then using again the Cauchy-Davenport theorem in \( \{u_0\} \times \mathbb{F}^d_p \) to see that the concatenation of the two lists (of length \( t = r + s \)) satisfies (2.2). Using this construction, Proposition 7 immediately follows with \( k \) equal to \(|J_1^1| + \ldots + |J_1^r|\) (note that by definition \(|J_1^1| = |J_2^1|\)).

\[ \square \]

2.3 High multiplicity case

The final lemma is a result about zero sums in sequences which can be regarded as a generalization of Olson’s main result from [10].

Lemma 7. Let \( Y \subseteq \mathbb{F}^d_p \) be an arbitrary set and let \( w : Y \rightarrow \mathbb{N} \) be a function such that \( \sum_{y \in Y} w(y) \geq d(p-1) + 2r|Y| + 1 \) for some \( r \geq 0 \). Then, there exist coefficients \( a_y \in \mathbb{N} \), one for each \( y \in Y \), such that \( a_y \in \{0\} \cup \{r, w(y) - r\} \) and \( \sum_{y \in Y} a_y y = 0 \), while not all \( a_y \)’s are simultaneously zero.

When \( r = 0 \), notice that this indeed immediately implies that if \( n > d(p-1) \), then among any \( n \) elements \( v_1, \ldots, v_n \) of \( \mathbb{F}^d_p \) there exists a nonempty subsequence with a zero subsum. To prove Lemma 4 we will make use of Alon’s Combinatorial Nullstellensatz [1] Theorem 1.2], which we recall for the reader’s convenience.

Lemma 8. Let \( \mathbb{F} \) be an arbitrary field, and let \( f = f(x_1, \ldots, x_n) \) be a polynomial in \( \mathbb{F}[x_1, \ldots, x_n] \). Suppose the degree \( \deg(f) \) of \( f \) is \( \sum_{i=1}^n t_i \), where each \( t_i \) is a nonnegative integer, and suppose that the coefficient of \( \prod_{i=1}^n x_i^{t_i} \) in \( f \) is nonzero. Then, if \( S_1, \ldots, S_n \) are the subsets of \( F \) with \(|S_i| > t_i \), there exist \( s_1 \in S_1, \ldots, s_n \in S_n \) so that
\[
f(s_1, \ldots, s_n) \neq 0.
\]
Proof of Lemma 7: For $y \in Y$ denote $A_y = \{0\} \cup [r, w(y) - r]$ and consider the following $|Y|$-variate polynomial in $(\alpha_y)_{y \in Y}$:

$$P(\alpha_y \mid y \in Y) = \prod_{i=1}^{d} \left(1 - \left(\sum_{y \in Y} \alpha_y y_i\right)^{p-1}\right),$$

where $y_i$ denotes the $i$-th coordinate of $y$ as an element in $\mathbb{F}_p^d$. Note that $P(\alpha_y \mid y \in Y)$ is non-zero if and only if $\sum_{y \in Y} \alpha_y y_i = 0$, so the zero vector $\vec{0}$ in $\mathbb{F}_p^{|Y|}$ is certainly not a zero of the polynomial $P$. On the other hand, observe that

$$\sum_{y \in Y} (|A_y| - 1) \geq \sum_{y \in Y} w(y) - 2r|Y| > d(p - 1),$$

so by Lemma applied in a slightly smaller cartesian product which is strictly contained in $\prod_{y \in Y} A_y$ and which does not contain $\vec{0}$, it follows that $P$ must take some other non-zero value at a vector in $\prod_{y \in Y} A_y$ that does not have all coordinates equal to 0. This completes the proof of Lemma 7.

\[\square\]

3 Proof of Theorem 1

Let $X \subset \mathbb{F}_p^d$ be an arbitrary set of size $(d - 1 + \varepsilon)p$ where $p$ is a sufficiently large prime number. Let $g : \mathbb{N} \rightarrow \mathbb{N}$ be a sufficiently fast growing function.

Apply Lemma 3 to $X$ with $\varepsilon' = \varepsilon/2d$ and $g = g$, $K = K_0$. After removing $X_0$ from $X$ we will obtain a set $X$ of size at least $(d - 1 + \varepsilon/2)p$ and a decomposition $X = X_1 \cup \ldots \cup X_m$ with several important properties. In particular, if $U_i$ denotes the affine hull of the set $X_i$, recall that $X_i$ is $(g^{d+1}(K), \delta)$-thick in $U_i$ for some $\delta \gg K, \varepsilon, d$ 1, for each $i \in \{1, \ldots, m\}$. Moreover, for every $i$ we also have that $|X_i| \geq \mu|X|$, where $\mu \gg K, \varepsilon, d$ 1. Note that we may assume that $\mu$ is small enough, namely, $\mu m < \varepsilon/100$. Furthermore, note that since $X$ is a set, all spaces $U_i$ are non-zero dimensional.

Let $H \subset \mathbb{F}_p^d$ be a generic hyperplane passing through the origin which intersects all affine spaces $U_i$. Such $H$ exists since $m \ll K, d, \varepsilon$ 1 and $p$ is large enough (indeed, a random hyperplane $H$ intersects an affine subspace of dimension $\geq 1$ with probability $\geq 1 - \frac{1}{p}$).

Fix an arbitrary vector $x_i \in H \cap U_i$ for each $i \in [m]$. Assign the weight $w_i = |X_i|$ and apply Lemma 7 to the set $Y = \{x_1, \ldots, x_m\} \subset H$ with the weight $w$ and $r = \mu|X|/3$. We have $\mu < \varepsilon/10m$ and so

$$\sum_{i=1}^{m} w_i = |X| \geq (d - 1 + \varepsilon/2)p > (d - 1)p + 2\mu|X|m/3.$$ 

Thus, there are non-negative, not all zero, coefficients $a_i \in \{0\} \cup [\mu|X|/3, w_i - \mu|X|/3]$ such that $\sum_{i=1}^{m} a_i x_i = 0$. Let $S \subset [m]$ be the set of $i \in [m]$ for which $a_i > 0$. By Lemma 3 the set $X_S = \bigcup_{i \in S} X_i$ is $(K_S, g(K_S), \mu)$-tubular for $K_S = g'(K)$ and $l \in [0, d]$. So after a linear change of coordinates, there is a vector $v \in \mathbb{F}_p^d \times \{0\}$ such that

$$X_S \subset v + [-K_S, K_S]^d \times \mathbb{F}_p^{d-l}.$$
Denote by \( \pi \) the projection onto first \( l \) coordinates. The condition that \( X_i \) is \((g^{d+1}(K), \delta)\)-thick in \( U_i \) and the fact that \( K_S < g^{d+1}(K) \) implies that \( \pi(U_i) \) is a single point for any \( i \in S \). Indeed, for \( j = 1, \ldots, l \) consider the linear function \( \xi_j(x) = x_j - v_j \). Then for any \( i \in S \) the set \( X_i \) is not \((K_S, 0)\)-thick along \( \xi_j \). This implies that \( \xi_j \) is constant on \( U_i \) and so \( \pi(U_i) \subset \mathbb{F}_p^l \) is a single point. Denote this point by \( y_i \in v + [-K_S, K_S]^l \) and observe that
\[
\sum_{i=1}^{m} a_i y_i = 0, \tag{9}
\]
since \( \pi \) is a linear operator.

Denote by \( Y \subset [-K_S, K_S]^l \) the set obtained from the projection \( \pi(X) \) and by shifting by \( v \). Note that for any \( y \in Y \) the set \( X_y = X_S \cap \{y\} \times \mathbb{F}_p^{d-l} \) has size at least \( \mu|X| \geq \mu|X_S| \) by Lemma 3. Moreover, the set \( X_S \) is \((g(K_S), \delta)-\text{thick along any linear function which is non-constant on } \{0\} \times \mathbb{F}_p^{d-l} \). For \( y \in Y \) denote by \( a_y \) the sum of all numbers \( a_i \) over all \( i \) is such that \( y = y_i - v \).

**Proposition 2** If \( p \) is large enough then there are sets \( Z_y \subset X_y \), such that for any \( y \in Y \) we have \( |Z_y| \in [\mu|X_y|/20, \mu|X_y|/10] \) and the set \( Z = \bigcup_{y \in Y} Z_y \) is \((g(K_S), \delta/4)-\text{thick along any linear function which is not constant on } \{0\} \times \mathbb{F}_p^{d-l} \).

We prove Proposition 2 by using a probabilistic argument, where we make use of the following Chernoff bound (see, for example, [11, Corollary 21.7]).

**Lemma 9** Let \( X \) be a random variable with the binomial distribution \( \text{Bin}(N, p) \) and \( \eta \in (0, 1) \). Then
\[
\Pr(X \leq (1 - \eta)\mathbb{E}[X]) \leq \exp \left(-\frac{\eta^2}{2} \mathbb{E}[X]\right)
\]
\[
\Pr(X \geq (1 + \eta)\mathbb{E}[X]) \leq \exp \left(-\frac{\eta^2}{3} \mathbb{E}[X]\right)
\]

**Proof of Proposition 2** Choose sets \( Z_y \subset X_y \) at random according to the binomial distribution \( \text{Bin}(|X_y|, \mu/15) \). It follows from Lemma 9 and the fact that \( |X_y| \gg p \) that with high probability we have \( |Z_y| \in [\mu|X_y|/20, \mu|X_y|/10] \) for all \( y \in Y \). We will show that for any fixed linear function the event that \( Z = \bigcup_{y \in Y} Z_y \) is not \((g(K_S), \delta/4)\)-thick has probability exponentially small in \( p \). Since there are only \( O(p^d) \) linear functions on \( \mathbb{F}_p^d \) this will be enough to prove Proposition 2. Fix a linear function \( \xi \) which is not constant on \( \{0\} \times \mathbb{F}_p^{d-l} \) and denote \( X'_y = X_y \setminus H(\xi, g(K_S)) \). Since the set \( X_S \) is \((g(K_S), \delta)-\text{thick along } \xi \), the set \( X' = \bigcup_{y \in Y} X'_y \) has size at least \( \delta|X_S| \). Note that the expected size of the intersection \( Z_y \cap X'_y \) is asymptotically equal to \( \frac{|Z_y||X'_y|}{|X_S|} \) and so, provided that \( |X'_y| \gg p \), by Lemma 9, the probability of the event that \( |Z_y \cap X'_y| < \frac{|Z_y||X'_y|}{1.5|X_S|} \) is at most \( e^{-cp} \) for some \( c \gg 1 \). Therefore, the probability that
\[
\sum_{y \in Y} |Z_y \cap X'_y| < \frac{1}{1.5} \sum_{y \in Y} \frac{|Z_y||X'_y|}{|X_S|}
\]
is at most \( |Y| e^{-cp} \). But since the set \( Z_y \) has size in the interval \( [\mu|X_y|/20, \mu|X_y|/10] \) the right hand side is at least
\[
(\mu/20) \sum_{y \in Y} \frac{|X'_y|}{1.5} \geq (\mu/20) \delta|X|/1.5 > \delta|Z|/4.
\]
which means that $Z$ is $(g(K_S), \delta/4)$-thick along $\xi$ with probability at least $1 - |Y|e^{-cp}$. This completes the proof.

Fix sets $Z_y$ as in Proposition 2. Now let the function $g$ grow so fast that we have $g(K) > K_2$ where $K_2 = K_2(K, d, \delta/4, \mu^2/10)$ is the function from Proposition 1. We can then apply Proposition 1 to sets $Z_y - v$ with $u = 0$ to get some $k_y \in \mathbb{N}$ such that

$$\sum_{y \in Y} k_y(y - v) = u_0, \quad \sum_{y \in Y} k_y = k,$$

where $u_0$ and $k$ are from the statement of Proposition 1. Note that $k_y \leq |Z_y| \leq \mu |X_y|/10$.

For each $y \in Y$ fix a subset $A_y \subset X_y \setminus Z_y$ of size $a_y - k_y$ (which is possible thanks to the estimates on $a_i$). Let $u = (u_1, u_2) \in \mathbb{F}_p^d$ denote the following vector:

$$u = \sum_{y \in Y} \sum_{x \in A_y} x.$$

From (9) and from the definition of the $k_y$’s we see that, in fact, $u_1 = -u_0 - kv$.

By the conclusion of Proposition 1 applied to the vector $-u_2 \in \mathbb{F}_p^{d-1}$, we obtain some sets $S_y \subset Z_y$ such that $\sum_{y \in Y} |S_y| = k$ and

$$\sum_{y \in Y} \sum_{x \in S_y} (x - v) = (u_0, -u_2),$$

After rearranging, this rewrites as

$$\sum_{y \in Y} \sum_{x \in S_y} x = (u_0 + kv, -u_2) = -\sum_{y \in Y} \sum_{x \in A_y} x.$$

So we see that the set $B = \bigcup_{y \in Y} A_y \cup S_y$ has zero sum. Theorem 1 is proved.

References

[1] N. Alon, Combinatorial Nullstellensatz, *Combin. Probab. Comput.* 8 (1999), 7–29.

[2] N. Alon, M. Dubiner, A lattice point problem and additive number theory, *Combinatorica* 15.3 (1995): 301-309.

[3] É. Balandraud, An addition theorem and maximal zero-sum free sets in $\mathbb{Z}/p\mathbb{Z}$, *Israel Journal of Mathematics*, 188 (2012), 405–429.

[4] É. Balandraud, Addition Theorems in $\mathbb{F}_p$ via the Polynomial Method, arXiv:1702.06419

[5] G. Bhowmik, J.-C. Schlage-Puchta, An improvement on Olson’s constant for $\mathbb{Z}_n \oplus \mathbb{Z}_n$, *Acta Arithmetica*. 141(4) (2010), 311–319.

[6] S. Eliahou and M. Kervaire, Sumsets in vector spaces over finite fields, *J. Number Theory* 71 (1998), 12–39.

[7] P. Erdős, H. Heilbronn, On the addition of residue classes modulo p, *Acta Arithmetica*. 9 (1964), 149–159.
[8] W. D. Gao, A. Geroldinger, Zero-sum problems in finite abelian groups: a survey, *Expo. Math.* 24 (2006), 337–369.

[9] W. D. Gao, I. Z. Ruzsa and R. Thangadurai, Olson’s constant for the group $\mathbb{F}_p \oplus \mathbb{F}_p$, *J. of Combinatorial Theory*, Series A, 107 (2004), 49–67.

[10] Y. O. Hamidoune and G. Zémor, On zero-free subset sums, *Acta Arithmetica*, 78 2 (1996), 143–152.

[11] S. Janson, T. Łuczak, A. Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[12] R. N. Karasev, F. Petrov, Partitions of nonzero elements of a finite field into pairs, *Israel Journal of Mathematics* 192 (1), 143–156.

[13] L. H. Loomis, H. Whitney, An inequality related to the isoperimetric inequality, *Bulletin of the American Mathematical Society*, 55.10 (1949): 961–962.

[14] H. H. Nguyen, E. Szemerédi, V. H. Vu, Subset sums modulo a prime, *Acta Arithmetica*. 131 (2008), 303–316.

[15] H. H. Nguyen, V. Vu, A characterization of incomplete sequences in $\mathbb{F}_p^d$, *Journal of Combinatorial Theory*, Series A, 119 (2012).

[16] J. E. Olson, A combinatorial problem on finite abelian groups I, *J. Number theory*, 1(1) (1969) 8–11.

[17] J. E. Olson, A combinatorial problem on finite abelian groups II, *J. Number theory*, 1(2) (1969) 195–199.

[18] J. E. Olson, Sum of sets of group elements, *Acta Arithmetica*, 28 (1975), 147–156.

[19] F. Petrov, Combinatorial results implied by many zero divisors in a group ring, [arXiv:1606.03256](https://arxiv.org/abs/1606.03256)

[20] F. Petrov, C. Pohoata, Improved bounds for progression-free sets in $C_n^m$, *Israel Journal of Mathematics*, 236 (2020).345–363.

[21] E. Szemerédi, On a conjecture of Erdős and Heilbronn, *Acta Arithmetica*, 17 (1970), 227–229.

[22] D. Zakharov, Convex geometry and the Erdős-Ginzburg-Ziv problem, [arXiv:2002.09892](https://arxiv.org/abs/2002.09892)