European Journal of Combinatorics 77 (2019) 102–113

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

A bilinear Bogolyubov theorem

Pierre-Yves Bienvenu [a], Thái Hoàng Lê [b]

[a] Institut Camille-Jordan, Université Lyon 1, 43 boulevard du 11 novembre 1918 69622 Villeurbanne cedex, France

[b] Department of Mathematics, The University of Mississippi, University, MS 38677, United States

ARTICLE INFO

Article history:
Received 8 August 2018
Accepted 15 November 2018
Available online 14 December 2018

ABSTRACT

The purpose of this note is to prove the existence of a remarkable structure in an iterated sumset derived from a set P in a Cartesian square $\mathbb{F}_p^n \times \mathbb{F}_p^n$. More precisely, we perform horizontal and vertical sums and differences on P, that is, operations on the second coordinate when the first one is fixed, or vice versa. The structure we find is the zero set of a family of bilinear forms on a Cartesian product of vector subspaces. The codimensions of the subspaces and the number of bilinear forms involved are bounded by a function $c(\delta)$ of the density $\delta = |P|/p^{2n}$ only. The proof uses various tools of additive combinatorics, such as the (linear) Bogolyubov theorem, the density increment method, as well as the Balog–Szemerédi–Gowers and Freiman–Ruzsa theorems.

1. Introduction

Let $\mathbb{F} = \mathbb{F}_p$ be a finite field of fixed prime order and V be a vector space of dimension n over $\mathbb{F}$. In this paper, the dimension n is the asymptotic parameter and all the $O(\cdot)$ and $o(\cdot)$ may depend on p but not n. By the *density* of a subset $A \subset V$ we mean $\frac{|A|}{|V|}$. The classical Bogolyubov theorem states the following:

Theorem 1 (*Bogolyubov*)**.** *If $A \subset V$ is a set of density $\alpha > 0$, then the sumset*

$$A + A - A - A := \{a_1 + a_2 - a_3 - a_4 \mid (a_1, \dots, a_4) \in A^4\}$$

contains a vector subspace of codimension $c(\alpha)$.

E-mail addresses: pbienvenu@math.univ-lyon1.fr (P.-Y. Bienvenu), leth@olemiss.edu (T.H. Lê).

https://doi.org/10.1016/j.ejc.2018.11.003
0195-6698/

The notation $A + A - A - A$ is often abbreviated as $2A - 2A$. Note that Bogolyubov's original argument [2] works in $\mathbb{Z}$ instead of vector spaces, but at least since Green's survey [4] on finite field models, it has been applied to the vector space setting, where it gives $c(\delta) = O(\delta^{-2})$. The best bound is due to Sanders [8, Theorem 11.1], who showed that $c(\delta) = O(\log^4 \delta^{-1})$. The very short proof of the polynomial bound as well as a simplified exposition of Sanders' breakthrough can be found in the excellent survey [9]. Sanders' result is usually stated and proven for $p = 2$ but holds for any p, the implied constant depending on p. We point out that the corresponding statement for $A - A$ is not true both in vector spaces (see [5, Theorem 9.4]) and in $\mathbb{Z}$ (a result of Kříž [6]).

The purpose of this note is to prove a bilinear version of Theorem 1, that is, for dense subsets of $V \times V$. Before stating our result, we first need some definitions. For a subset $P \subset V \times V$, we define vertical and horizontal additive operations on P as follows:

$$P \overset{V}{\pm} P = \{(x, y_1 \pm y_2) \mid ((x, y_1), (x, y_2)) \in P^2\}$$

and

$$P \overset{H}{\pm} P = \{(x_1 \pm x_2, y) \mid ((x_1, y), (x_2, y)) \in P^2\}$$

where V and H mean vertical and horizontal. Note that V is also the name of the ambient space, but this should not create any confusion. We denote by ϕ_V the operation

$$P \mapsto (P \overset{V}{+} P) \overset{V}{-} (P \overset{V}{+} P)$$

and define the operation ϕ_H similarly.

Theorem 2. *For any $\delta > 0$, there exists a constant $c(\delta) > 0$ such that the following holds. Let $P \subset V \times V$ have density δ. There exist subspaces $W_1 \leq V, W_2 \leq V$ of codimension r_1, r_2 and a family $\mathbf{Q} = (Q_1, \dots, Q_{r_3})$ of bilinear forms on $W_1 \times W_2$ such that*

$$\phi_H \phi_V \phi_H(P) \supset \{(x, y) \in W_1 \times W_2 \mid Q_1(x, y) = \cdots = Q_{r_3}(x, y) = 0\} \tag{1}$$

where $\max(r_1, r_2, r_3) \leq c(\delta)$. Moreover $c(\delta)$ can be taken as $O(\exp(\exp(\exp(\log^{O(1)} \delta^{-1}))))$.

Our proof actually gives $\max(r_1, r_3) = O(\log^{O(1)} \delta^{-1})$. We point out that Gowers and Milićević [3] independently proved a result very similar to Theorem 2. However, their method and bounds are different from ours. They proved $\max(r_1, r_2, r_3) = O(\exp(\exp(\log^{O(1)} \delta^{-1})))$.

In view of our bounds for r_1 and r_3 and the fact that the roles of r_1 and r_2 are symmetric, it is quite reasonable to conjecture the following.

Conjecture 3 (*Polylogarithmic Bilinear Bogolyubov*)**.** *In Theorem 2, one can take $c(\delta) = O(\log^{O(1)} \delta^{-1})$.*

If P is a Cartesian product $A \times B$ for some subsets $A, B \subset V$, then using Theorem 1 once on each coordinate, we obtain a product $A' \times B'$ of subspaces of codimension $O(\log^4 \delta^{-1})$. Also it is easy to see that $c(\delta) \gg \log \delta^{-1}$ by considering a set such as the right-hand side of Eq. (1). Conjecture 3 says that, like in the linear case, this lower bound on δ should not be too far off the truth. The conjecture remains equally interesting and useful for the application we have in mind if $O(1)$ operations ϕ_V or ϕ_H are required instead of 3.

Conjecture 3 was very recently proven by [7], at the cost of requiring a somewhat larger number of iterations of ϕ_H and ϕ_V, namely $\phi_H \phi_V \phi_H \phi_V \phi_V \phi_H$ instead of $\phi_H \phi_V \phi_H$.

A quick application

Our application concerns matrices of low rank. Suppose a two-parameter, bilinearly varying family of matrices is often of rank at most ϵ. Then it must be of rank $O(\epsilon)$ on a whole bilinear set. We now state this application precisely. Let $\mathrm{Mat}_m(\mathbb{F})$ be the space of $m \times m$ matrices with coefficients in $\mathbb{F}$.

Corollary 4. *Suppose that we have a bilinear map* $\psi : \mathbb{F}^n \times \mathbb{F}^n \to \mathrm{Mat}_m(\mathbb{F})$. *Suppose that the set*

$$P_\epsilon = \{(f, g) \in \mathbb{F}^n \times \mathbb{F}^n \mid \mathrm{rank}(\psi(f, g)) \leq \epsilon\}$$

has density $\delta > 0$. *Then the set*

$$P_{64\epsilon} = \{(f, g) \in \mathbb{F}^n \times \mathbb{F}^n \mid \mathrm{rank}(\psi(f, g)) \leq 64\epsilon\}$$

contains a set of the form (1) *where the codimensions and the cardinality of the family of bilinear forms are at most* $c(\delta)$.

The authors exploit this corollary, with the conjectured bound on $c(\delta)$ in Theorem 2, in a companion paper [1].

Proof. We apply Theorem 2 to P. Note that the set P' it produces is included in $P_{64\epsilon}$ by the bilinearity of ψ and the fact that $\mathrm{rank}(A + B) \leq \mathrm{rank}\, A + \mathrm{rank}\, B$ for any two matrices A and B. □

The organization of the paper is as follows. In Section 2 we recall some basic facts and preliminaries. The heavy lifting part of our argument is an iteration scheme, Proposition 11. In Section 3 we show how this proposition implies Theorem 2. Section 4 is devoted to proving Proposition 11.

2. Preliminaries

The symbol $\mathbb{E}$ will be at some point used in its usual probabilistic sense, but it will frequently denote an average, thus $\mathbb{E}_{x\in X} = \frac{1}{|X|}\sum_{x\in X}$. Similarly, $\mathbb{P}_{x\in X}(x \in Y) = \frac{|Y|}{|X|}$ for sets $Y \subset X$.

We now briefly recall some basic facts about the Fourier transform and convolutions. Let V be a finite $\mathbb{F}$-vector space; in fact, all spaces considered will be finite in this paper, so we may not always specify this hypothesis. Then we denote by $\widehat{V}$ its *dual*, the set of characters on V. A character $\chi \in \widehat{V}$ takes values in the pth roots of unity, that is, $1, \omega, \ldots, \omega^{p-1}$ where $\omega = \exp(2i\pi x/p)$. The trivial character is $\chi = 1$. Let $f : V \to \mathbb{C}$ be a function. Then the Fourier transform $\hat{f}$ is defined on $\widehat{V}$ by

$$\hat{f}(\chi) = \mathbb{E}_{x\in V} f(x)\chi(x).$$

In particular, if $A \subset V$ has density α and indicator function 1_A, we have $\widehat{1_A}(1) = \alpha$. Besides, we have $\widehat{1_{-A}} = \overline{\widehat{1_A}}$.

If W is an affine subspace of V of direction $\overrightarrow{W}$, thus $W = a + \overrightarrow{W}$ for some $a \in V$, and $f : W \to \mathbb{C}$ is a function, we define the function $\tilde{f}$ on the vector space $\overrightarrow{W}$ by $\tilde{f}(v) = f(a + v)$. We then define the Fourier transform of f relative to W as the Fourier transform of $\tilde{f}$ on $\overrightarrow{W}$. We will abuse notation and denote by $\widehat{W}$ the dual of $\overrightarrow{W}$. Thus the notion of Fourier transform depends on the ambient (potentially affine) space one is considering, but when no ambiguity is possible, the space considered may not be made explicit.

Besides, if $f, g : V \to \mathbb{C}$ are two functions, we define their *convolution* $f * g : V \to \mathbb{C}$ by

$$f * g(x) = \mathbb{E}_{y\in V} f(y)g(x - y).$$

We define the U^2 norm by

$$\|f\|_{U^2(V)}^4 = \mathbb{E}_{x\in V} |f * f(x)|^2 .$$

A quadruple $(x_1, x_2, x_3, x_4) \in V^4$ satisfying $x_1 + x_2 = x_3 + x_4$ is called an *additive quadruple*.

Observe that if $f = 1_A$ is the indicator function of the subset $A \subset V$, then

$$\|1_A\|_{U^2(V)}^4 = \frac{\left|\{(x_1, x_2, x_3, x_4) \in A^4 \mid x_1 + x_2 = x_3 + x_4\}\right|}{|V|^3}$$

and we refer to this quantity as the *density of additive quadruples* in A. Again if W is an affine subspace of V and $f : W \to \mathbb{C}$ is a function, we will write $\|f\|_{U^2(W)} = \|\tilde{f}\|_{U^2(\overrightarrow{W})}$. Observe that the connection with the additive quadruples of $A \subset W$ is preserved, because additive quadruples are invariant by translation. When it is obvious from the context which space one is considering, one will simply write $\|f\|_{U^2}$.

We recall without proof a few basic properties of the Fourier transform.

(1) Parseval's identity is the statement that

$$\mathbb{E}_{x\in V}\,|f(x)|^2 = \sum_{\chi\in\widehat{V}} \left|\hat{f}(\chi)\right|^2.$$

In particular, for a subset $A \subset V$ of density α, we have $\sum_{\chi\in\widehat{V}} \left|\widehat{1_A}(\chi)\right|^2 = \alpha$.

(2) The Fourier transform of a convolution is the product of the Fourier transforms, that is

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

(3) Combining the previous two points, we see that the U^2 norm of a function is the L_4 norm of its Fourier transform, that is

$$\|f\|_{U^2(V)} = \|\hat{f}\|_4.$$

In particular if $f = 1_A$ for a subset A of density α, Parseval's identity implies that

$$\alpha^4 \le \|\widehat{1_A}\|_4^4 = \alpha^4 + \sum_{\chi\in\widehat{V},\,\chi\neq 1} \left|\widehat{1_A}(\chi)\right|^4 \le \alpha^4 + \alpha \max_{\chi\in\widehat{V},\,\chi\neq 1} \left|\widehat{1_A}(\chi)\right|^2. \tag{2}$$

When a set $A \subset W$ of density α has about as few additive quadruples as it can, that is, $\alpha^4 \le \|1_A\|_{U^2(W)}^4 \le \alpha^4(1+\epsilon)$, we will call it ϵ-*pseudorandom*. In particular, A is ϵ-pseudorandom in W if $\max_{\chi\in\widehat{W},\,\chi\neq 1} \left|\widehat{1_A}(\chi)\right| \le \alpha^{3/2}\epsilon^{1/2}$.

(4) The Fourier inversion formula is the statement that

$$f = \sum_{\chi\in\widehat{V}} \hat{f}(\chi)\overline{\chi}. \tag{3}$$

Our first lemma says that if A is sufficiently pseudorandom in terms of its density then $2A - 2A$ is the whole space.

Lemma 5. *Let W be an affine subspace of V and $A \subset W$ have density α. If $\|1_A - \alpha\|_{U^2(W)} < \alpha$, or equivalently,*

$$\sum_{\chi\neq 1} \left|\widehat{1_A}(\chi)\right|^4 < \alpha^4, \tag{4}$$

then $2A - 2A = \overrightarrow{W}$. Consequently, if $\max_{\chi\in\widehat{W},\,\chi\neq 1} \left|\widehat{1_A}(\chi)\right| < \alpha^{3/2}$ then $2A - 2A = \overrightarrow{W}$.

Proof. For any $x \in \overrightarrow{W}$, by the Fourier inversion formula (3), we have

$$1_A * 1_A * 1_{-A} * 1_{-A}(x) = \sum_{\chi\in\widehat{W}} \left|\widehat{1_A}(\chi)\right|^4 \overline{\chi(x)} \ge \alpha^4 - \sum_{\chi\neq 1} \left|\widehat{1_A}(\chi)\right|^4 > 0.$$

This implies that $x \in 2A - 2A$. □

We also need the following standard fact which relates the lack of pseudorandomness to density increment.

Lemma 6 (*[4, Lemma 3.4]*)**.** *Let W be an affine subspace of V and $A \subset W$ have density α. Suppose there exists $\chi \in \widehat{W}$, $\chi \neq 1$ such that $\left|\widehat{1_A}(\chi)\right| \ge \beta$. Then there exists an affine subspace $H \le W$ of codimension 1 such that the density of $A \cap H$ on H is at least $\alpha + \beta/2$.*

Our next tool is a regularity lemma.

Lemma 7. *Let W be an affine subspace of V and $A \subset W$ have density α. Let $\epsilon > 0$. For any t, there exists an affine subspace $H \le W$ of codimension $O(t\epsilon^{-1}\log\alpha^{-1})$ such that $\left|A'\right| = \alpha'|H|$ (where $A' = A \cap H$)*

with $\alpha' \geq \alpha$ and for any affine subspace F of codimension at most t of H, $\frac{|A\cap F|}{|F|} \leq \alpha(1+\epsilon)$. Consequently, for any affine subspace F of codimension at most t of H, we also have $\frac{|A\cap F|}{|F|} \geq \alpha(1-p^t\epsilon)$.

Proof. Let us prove the first conclusion. If W does the trick already, we do nothing. If not, there exists a subspace H of codimension at most t such that $\frac{|A\cap H|}{|H|} > \alpha(1+\epsilon)$. We replace W by H, and A by $A \cap H$. And we iterate. We duplicate the density in at most ϵ^{-1} iterations. And we may duplicate up to $\log \alpha^{-1}$ times before hitting 1. At every iteration we may lose up to t dimensions. Whence the first conclusion. The second conclusion follows from summing the upper bound over all cosets of F. □

In particular, when $t = 1$, the following corollary says that we can always suppose that a set $A \subset W$ is pseudorandom, at the cost of passing to a subset in an affine subspace.

Corollary 8. *Let W be an affine subspace of V and $A \subset W$ have density α. Let $\epsilon > 0$. Then there exists an affine subspace $H \leq W$ of codimension $O((\alpha\epsilon)^{-1/2}\log\alpha^{-1})$ such that $A' = A \cap H$ has density $\geq \alpha$ and is ϵ-pseudorandom in H.*

Proof. We use Lemma 6 with $\beta = \alpha^{3/2}\epsilon^{1/2}$, and Lemma 7 with $t = 1$ and $\epsilon' = \alpha^{1/2}\epsilon^{1/2}/2$ to obtain the conclusion. □

Our next tool is a standard lemma resulting from the combination of the Balog–Szemerédi–Gowers and Freiman–Ruzsa theorems. A useful reference for this lemma is [5, Lecture 2]. We reproduce the proof as we want to incorporate the quasipolynomial bound of Sanders [8, Theorem 11.4] for the Freiman–Ruzsa theorem.

Lemma 9. *Let $W \leq V$ be $\mathbb{F}$-vector spaces and $A \subset W$ have density α. Let $c > 0$ be a constant. Suppose $\xi : A \to V$ is such that are at least $c\,|A|^3$ additive quadruples in the graph $\Gamma = \{(y, \xi(y)) \mid y \in A\}$. Then there is a subset $S \subset A$ such that $\xi_{|S}$ coincides with an affine-linear map. Moreover, the density of S in A can be taken quasipolynomial in c^{-1}, that is, $|S| \gg |A|\exp(-\log^{O(1)} c^{-1})$.*

Proof. First, the Balog–Szemerédi–Gowers theorem implies that there exists a set $A' \subset A$ satisfying $\left|A'\right| \geq C\,|A|$ that induces a subgraph $\Gamma' \subset \Gamma$ satisfying $\left|\Gamma' + \Gamma'\right| \leq C'\,|\Gamma|$, where both C and C' can be taken polynomial in c^{-1}. Using the Freiman–Ruzsa theorem (with Sanders' bounds from [8, Theorem 11.4]), we get a subgraph $\Gamma'' \subset \Gamma'$ corresponding to a subset $A'' \subset A'$ satisfying $\left|\Gamma''\right| \geq D\left|\Gamma'\right|$ and $\left|\mathrm{span}(\Gamma'')\right| \leq E\left|\Gamma''\right|$ with D polynomial and E quasipolynomial in c^{-1}. Write $H = \mathrm{span}(\Gamma'') \leq W \times V$ and $\pi : H \to W$ the canonical projection of $W \times V$ to the first coordinate restricted to H. Then $\pi(H) \supset A''$ by definition. Because $|H| \leq E\left|A''\right|$, the size of the kernel of π is at most E. Then we can partition H into at most E cosets of some subspace H' so that π is injective on each of them. By the pigeonhole principle, there exists such a coset that has a large intersection with Γ'', that is, an $x \in W$ such that

$$\left|(x + H') \cap \Gamma''\right| \geq \left|\Gamma''\right|/E.$$

Let now $\Delta = (x + H') \cap \Gamma''$ and S be the corresponding subset of A'', that is, $\Delta = \{(y, \xi(y)) \mid y \in S\}$. The map $\pi_{|x+H'}$ is a bijection onto its image, an affine space $M \leq V$. Its inverse function is an affine map $\psi : M \to W$ such that $(s, \psi(s)) \in \Gamma''$ for all $s \in S$, that is, $\psi(s) = \xi(s)$. Moreover,

$$|S| = |\Delta| \geq \left|A''\right|/E \geq K\,|A|$$

where K is quasipolynomial in c^{-1}. □

We will also need the following lemma.

Lemma 10. *Let A be a finite set and $T \subset \{(a_1, a_2, a_3, a_4) \in A^4 \mid a_i \text{are pairwise distinct}\}$. Then there is a partition $A = \cup_{i=1}^4 A_i$ such that $|T \cap A_1 \times A_2 \times A_3 \times A_4| \geq |T|/256$.*

Proof. We use the probabilistic method. For a random partition of A where each $y \in A$ is assigned a part A_i with $i \in \{1, 2, 3, 4\}$ chosen independently, uniformly with probability 1/4, we have

$$\mathbb{E}[|T \cap A_1 \times A_2 \times A_3 \times A_4|] = \mathbb{E} \sum_{(x_1,\dots,x_4)\in T} \prod_{i=1}^{4} 1_{x_i \in A_i} = \sum_{(x_1,\dots,x_4)\in T} \mathbb{E} \prod_{i=1}^{4} 1_{x_i \in A_i}. \tag{5}$$

Let $(x_1, \dots, x_4) \in T$. In particular the x_i are pairwise distinct. Then by uniform distribution and independence, for any $(m_1, \dots, m_4) \in [4]^4$, we have

$$\mathbb{P}(x_i \in A_{m_i} \text{ for each } i \in [4]) = 4^{-4}.$$

Together with Eq. (5), this implies that

$$\mathbb{E}[|T \cap A_1 \times A_2 \times A_3 \times A_4|] = |T|/256$$

so there must be a partition with $|T \cap A_1 \times A_2 \times A_3 \times A_4| \geq |T|/256$. □

3. Deducing the main theorem from an iteration scheme

Outline of the proof

We first make a reduction and give an outline of the proof of the main theorem. Let $P \subset V \times V$ have density $\delta > 0$. Write $P = \cup_{y\in V} B_y \times \{y\}$. Because P has density δ, the set A of elements $y \in V$ such that $|B_y| \geq \delta |V|/2$ has density at least $\delta/2$. Using the linear Bogolyubov theorem (Theorem 1) on each set B_y for $y \in A$, we see that $\phi_H(P)$ contains a set $P' = \cup_{y\in A} V_y \times \{y\}$ where V_y is a subspace of codimension at most $r = O(\log^4 1/\delta)$. From now on, we will assume that

$$P = \cup_{y\in A} V_y \times \{y\},$$

where $A \subset V$ has density $\alpha \geq \delta/2$. A priori the subspaces V_y have nothing to do with each other. We will start an iterative process. At each step we will find a common structure (either linear or bilinear) for the V_y. In the end the V_y are very well structured, and this will enable us to show that $\phi_H\phi_V(P)$ contains the desired bilinear structure.

The iteration scheme

Let V^* be the linear dual of V, that is, the set of linear forms on V. For $(x, \xi) \in V \times V^*$, we denote $x \cdot \xi = \xi(x)$. For a set $U \subset V$, we let $U^\perp = \{\xi \in V^* \mid \forall x \in U,\ x \cdot \xi = 0\}$. Also, for a set $T \subset V^*$, we let $T^\perp = \{x \in V \mid \forall \xi \in T,\ x \cdot \xi = 0\}$. Our iteration scheme is as follows.

Proposition 11. *Let V be an $\mathbb{F}$-vector space, and W be an affine subspace. Let $r \leq \dim V$ be an integer and $\alpha > 0$. Let $\epsilon = p^{-r}/256$. Then there exists a constant $c(r, \alpha)$ such that the following holds. Let $A \subset W$ be an ϵ-pseudorandom subset of density α. Let $P = \cup_{y\in A} V_y \times \{y\} \subset V \times W$ where each V_y is a subspace of codimension at most r. Suppose there exist $s \leq r$ and affine maps $\xi_1, \dots, \xi_s$ from W to V^* and spaces $U_y \leq V^*$ for $y \in A$ of dimension at most $r - s$ such that $V_y^\perp = \operatorname{span}(\xi_j(y))_{j\in[s]} + U_y$. Then at least one of the following statements holds.*

(1) *(Termination) The set $\phi_V(P)$ contains*

$$\{(x, y) \in X_3 \times W_2 \mid x \cdot \overrightarrow{\xi_1}(y) = \dots = x \cdot \overrightarrow{\xi_s}(y) = 0\}$$

where $\overrightarrow{\xi}$ denotes the linear part of an affine map ξ, W_2 is the direction of W and X_3 is a subset of density at least $p^{-r}/12$ in V.

(2) *(Reduction of codimension) There exist a set $S \subset A$ of density $c(r, \alpha)$, a subspace $V' \leq V$ of codimension at most $3r+10$ and subspaces $V'_y \leq V'$ of codimension at most $r-1$ (inside V') for each $y \in S$ such that $P \supset \bigcup_{y\in S} V'_y \times \{y\}$. Moreover, there exist affine maps $\xi'_1, \dots, \xi'_{s-1}$ from W to V'^* and spaces $U'_y \leq V'^*$ for $y \in S$ of dimension at most $r - s$ such that $(V'_y)^\perp = \operatorname{span}(\xi_j(y))_{j\in[s-1]} + U'_y$ (where $(V'_y)^\perp$ is defined in V').*

(3) *(Linearization) There exist a set* $S \subset A$ *of density* $c(r, \alpha)$ *and an affine map* $\xi_{s+1} : W \to V$ *and a space* $U'_y < U_y$ *such that for all* $y \in S$*, we have* $V_y^{\perp} = \mathrm{span}(\xi_1(y), \dots, \xi_{s+1}(y)) + U'_y$.

Moreover $c(r, \alpha)$ *can be taken quasipolynomial in* $\alpha^{-1}p^r$*, that is,*

$$c(r, \alpha) = \Omega(\exp(-\log^{O(1)}(\alpha^{-1}p^r))).$$

Since the statement of Proposition 11 looks complicated, an explanation is in order. We can think of the maps $\xi_1, \dots, \xi_s$ as the number of simultaneous constraints that the spaces $(V_y)_{y\in A}$ have to satisfy. (Thus at the beginning, $s = 0$ since we do not have any information on the V_y yet.) At each step, either the codimension of the V_y in V is reduced (the second alternative) or the number of constraints is increased (the third alternative). Clearly this process must stop when $r = 0$ (i.e., $V_y = V$ for all y) or when $s = r$ (i.e., all the V_y are given by r simultaneous constraints). In either case the V_y are very structured, which gives us the desired bilinear structure. We will now make this argument rigorous while keeping track of the bounds.

Proof of Theorem 2 using Proposition 11. Applying Corollary 8 with $\epsilon = p^{-r}/256$ and $r = O(\log^{O(1)} \alpha^{-1})$, we obtain an affine subspace $W^{(0)}$ of V of codimension

$$O((\epsilon\alpha)^{-1/2} \log \alpha^{-1}) = O(p^{r/2}\alpha^{-2/3})$$

such that the set $A_0 := A \cap W^{(0)}$ has density $\alpha_0 \geq \alpha$ and is ϵ-pseudorandom in W_0. We set $V^{(0)} = V, P_0 = \cup_{y\in A_0} V_y \times \{y\} \subset P$ and apply Proposition 11 with the tuple $(V^{(0)}, W^{(0)}, A_0, P_0)$ and $s_0 = 0, r_0 = r$.

If the first alternative of Proposition 11 holds, we stop.

Suppose the second alternative of Proposition 11 holds. We set $V^{(1)} \subset V^{(0)}$ to be the subspace V' given by the second alternative, of codimension $O(r)$. We are also given subspaces $V_y^{(1)} \leq V^{(1)}$ of codimension at most $r_1 = r - 1$ such that $P \supset \bigcup_{y\in A} V_y^{(1)} \times \{y\}$.

Suppose the third alternative holds. We obtain a set $S \subset A_0$ of density $c(r, \alpha_0)$ in $W^{(0)}$, an affine map $\xi_1 : W_0 \to V^*$ and subspaces $U_y^{(1)} \leq V_y^*$ of dimension at most $r_1 \leq r - 1$ such that $V_y^{\perp} = \mathrm{span}(\xi_1(y)) + U_y^{(1)}$. Then we let $V^{(1)} = V$ and $V_y^{(1)} = V_y$. We can find an affine subspace $W^{(1)} \subset W^{(0)}$ of codimension $O(p^{r/2}\alpha_0^{-2/3})$ in $W^{(0)}$ such that the set $A_1 := S \cap W^{(1)}$ is ϵ-pseudorandom and has density $\alpha_1 \geq c(r, \alpha_0)$ in $W^{(1)}$. Let $s_1 = 1$ and $r_1 = r$.

Set $P_1 = \bigcup_{y\in A_1} V_y^{(1)} \times \{y\}$. We have $P_0 \supset P_1$.

We can now apply Proposition 11 with $(V^{(1)}, W^{(1)}, A_1, P_1, r_1, s_1)$ and start an iterative process. This iterative process stops whenever one can apply the first item of Proposition 11, or when $r - s$ vanishes. When applying either of the last two alternatives, at least one of the parameters r or $r - s$ is decreased by at least one, while the other one cannot increase, so the iteration does eventually stop.

At the ith stage, we obtain a subspace $V^{(i)} \subset V$ of codimension $O(ri)$, an affine subspace $W^{(i)} \subset V$ of codimension

$$O(\exp(\log^{C^i} \alpha^{-1})),$$

where C is a constant (depending at most on p), an ϵ-pseudorandom set $A_i \subset W^{(i)}$ of density

$$\alpha_i = \Omega(\exp(-\log^{C^i} \alpha^{-1}))$$

and a set

$$P_i = \cup_{y\in A_i} V_y^{(i)} \times \{y\} \subset V^{(i)} \times W^{(i)}$$

where each $V_y^{(i)} \subset V^{(i)}$ has codimension $r_i \leq r$. The bounds for α_i and $\mathrm{codim}_V W^{(i)}$ follow from solving the recursive relations

$$\alpha_{i+1} = \Omega(\exp(-\log^{O(1)}(\alpha_i^{-1}p^{r_i}))),$$

so that

$$\log \alpha_{i+1}^{-1} \leq \log^{O(1)}(\alpha_i^{-1}p^{r_i}) + O(1) \leq \log^{C} \alpha_i^{-1}$$

(recall that $r_i \leq r = O(\log^4 \alpha^{-1})$), and

$$\operatorname{codim}_{W^{(i)}} W^{(i+1)} = O(p^{r_i/2}\alpha_i^{-2/3}).$$

Moreover, we have affine maps $\xi_1, \dots, \xi_{s_i}$ from W^i to $V^{(i)*}$ and subspaces $U_y^{(i)} \leq V^{(i)*}$ of dimension at most $r_i - s_i$ such that $(V_y^{(i)})^\perp = \operatorname{span}(\xi_1(y), \dots, \xi_{s_1}(y)) + U_y^{(i)}$. Furthermore, $P \supset P_i$. Suppose the algorithm stops after the ith iteration, where $i \leq 2r$. Note that we have $s_i \leq r$,

$$\operatorname{codim} V^{(i)} = O(r^2) = O(\log^{O(1)} \alpha^{-1}),$$

and

$$\operatorname{codim} W^{(i)} = O(\exp(\log^{C^r} \alpha^{-1})) = O(\exp(\exp(\exp(\log^{O(1)} \alpha^{-1})))).$$

There are two possibilities.

Case 1: $r_i = s_i$, and

$$P_i = \{(x, y) \in V^{(i)} \times A_i \mid x \cdot \xi_1(y) = \cdots = x \cdot \xi_{s_i}(y) = 0\}$$

where $\xi_1, \dots, \xi_{s_i}$ are affine maps from $W^{(i)}$ to $V^{(i)*}$, $A_i \subset W^{(i)}$ is a set of density $\gamma := \alpha_i = \Omega\left(\exp(-\exp(\exp(\log^{O(1)} \alpha^{-1})))\right)$.

Case 2: The first alternative of Proposition 11 holds, and

$$\phi_V(P_i) \supset \{(x, y) \in X \times W_2 \mid x \cdot \overrightarrow{\xi_1}(y) = \cdots = x \cdot \overrightarrow{\xi_{s_i}}(y) = 0\},$$

where $\xi_1, \dots, \xi_{s_i}$ are affine maps from $W^{(i)}$ to $V^{(i)*}$, $X \subset V^{(i)}$ is a set of density $\Omega\left(p^{-r_i}\right) = \Omega\left(\exp\left(-\log^{O(1)} \alpha^{-1}\right)\right)$ and W_2 is the direction of $W^{(i)}$.

Since the two cases are similar, we will work with Case 1. By translating P by $(0, a)$ for some $a \in A_i$ if necessary, we may assume that $W^{(i)}$ is a vector subspace of V. Let $\eta := \frac{1}{10}\gamma^{3/2}p^{-r-1}$. Applying Lemma 7 with $t = r + 1$, there is a subspace $H \leq W^{(i)}$ of codimension $O(r\eta^{-1}\log\gamma^{-1})$ such that $\left|A'\right| = \gamma' |H|$ (where $A' = A_i \cap H$) with $\gamma' \geq \gamma$ and for any subspace F of codimension at most $r + 1$ of H, $\frac{|A' \cap F|}{|F|} \leq \gamma(1 + \eta)$.

For each $x \in V^{(i)}$, let $B_x = \{y \in H \mid x \cdot \xi_1(y) = \cdots = x \cdot \xi_{s_i}(y) = 0\}$. Then B_x is a subspace of codimension at most r inside H. Let $A_x = A' \cap B_x$. We claim that $2A_x - 2A_x = \overrightarrow{B_x}$.

By Lemma 5, it suffices to show $\left|\widehat{1_{A_x}}(\chi)\right| < \gamma_x^{3/2}$ for any $\chi \neq 1$, where γ_x is the density of A_x in B_x.

Suppose for a contradiction that this is not true. Then Lemma 6 implies that there is a hyperplane F of B_x on which the density of A is at least $\gamma_x + \gamma_x^{3/2}/2$. From Lemma 7 we also have $\gamma_x \geq \gamma(1 - \eta p^{r+1})$. Therefore,

$$\begin{aligned}
\gamma_x + \gamma_x^{3/2}/2 &\geq \gamma(1 - \eta p^{r+1}) + \frac{1}{2}\gamma^{3/2}(1 - \eta p^{r+1})^{3/2} \\
&\geq \gamma(1 - \eta p^{r+1}) + \frac{1}{2}\gamma^{3/2}(1 - 2\eta p^{r+1}) \\
&\geq \gamma - \frac{1}{10}\gamma^{3/2} + \frac{2}{5}\gamma^{3/2} = \gamma + \frac{3}{10}\gamma^{3/2} > \gamma + \eta. \qquad (6)
\end{aligned}$$

This contradicts the assumption on H since F is a subspace of codimension at most $r+1$ of H. Therefore, $2A_x - 2A_x = \overrightarrow{B_x}$ and

$$\phi_V(P) \supset \cup_{x \in V^{(i)}} \{x\} \times \overrightarrow{B_x} = \{(x, y) \in V^{(i)} \times H \mid x \cdot \overrightarrow{\xi_1}(y) = \cdots = x \cdot \overrightarrow{\xi}_{s_i}(y) = 0\}.$$

Since the codimension of H in $W^{(i)}$ is $O(r\eta^{-1}\log\gamma^{-1})$, Theorem 2 follows in this case. In Case 2, a similar argument shows that $\phi_H\phi_V(P)$ contains the desired bilinear structure. □

4. Proof of Proposition 11

First we suppose that there exists a linear combination of the affine maps $\xi_1, \dots, \xi_s$ that has small rank as an affine map, that is, a nonzero $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_s) \in \mathbb{F}_p^s$ such that $\xi_0' = \sum_{j=1}^s \lambda_j \xi_j$ satisfies

$\mathrm{rk}\,\overrightarrow{\xi_0'} < 3r + 10$. By completing λ into a basis of $\mathbb{F}_p^s$, we obtain affine maps $\xi_0', \dots, \xi_{s-1}'$ from W to V^* such that $\mathrm{span}(\xi_0'(y), \dots, \xi_{s-1}'(y)) = \mathrm{span}(\xi_1(y), \dots, \xi_s(y))$ for any $y \in A$. Let $V' = \mathrm{span}(\xi_0'(y) \mid y \in A)^\perp$, then the codimension of V' in V is $\leq 3r + 10$. For each $y \in A$, let $V_y' < V'$ be given by $(V_y')^\perp = \mathrm{span}(\xi_1'(y), \dots, \xi_{s-1}'(y)) + U_y'$ where U_y' is the projection of U_y onto V'^*. Then $\dim U_y' \leq r - s$ and $\mathrm{codim}_{V'} V_y' \leq r - 1$. Since $V_y' < V_y$, we have $P \supset \bigcup_{y \in A} V_y' \times \{y\}$. This proves the second alternative (with $S = A$).

So let us now suppose that there exists no nonzero $\lambda \in \mathbb{F}_p^s$ such that $\xi_0' = \sum_{j \in [s]} \lambda_j \xi_j$ satisfies $\mathrm{rk}\,\overrightarrow{\xi_0'} < 3r + 10$. For $x \in V$, let $A_x = \{y \in A \mid x \in V_y\} \subset W$. Thus

$$\sum_{x \in V} |A_x| = |P| = |A||V|p^{-r} = \alpha p^{-r} |W|\,|V|. \tag{7}$$

Also, let $B_x = \{y \in W \mid x \cdot \xi_1(y) = \cdots = x \cdot \xi_s(y) = 0\}$; it is an affine subspace of codimension at most s, and $A_x \subset A \cap B_x$.

Claim 1. $\mathbb{P}_{x \in V}(\mathrm{codim} B_x < s) \leq \epsilon p^{-r}/4$. *(Recall that $\epsilon = p^{-r}/256$.)*

Proof. Note that

$$\overrightarrow{B_x} = \{y \in \overrightarrow{W} \mid y \cdot \overrightarrow{\xi_1}(x) = \cdots = y \cdot \overrightarrow{\xi_s}(x) = 0\}$$

is a subspace of codimension s unless there exists a nonzero $\lambda \in \mathbb{F}_p^s$ such that $\sum_{j=1}^s \lambda_j \overrightarrow{\xi_j}(x) = 0$. For any fixed such λ, the set of x that satisfy this relation is a linear subspace, namely the kernel K_λ of $\sum_{j=1}^s \lambda_j \overrightarrow{\xi_j}$ whose codimension equals the rank of $\sum_{j=1}^s \lambda_j \overrightarrow{\xi_j}$, hence at least $3r+10$. Hence $|K_\lambda|/|V| \leq p^{-3r-10}$. Because there are at most p^r tuples λ to consider, we conclude that $\mathbb{P}_{x \in V}(\mathrm{codim} B_x < s) \leq p^r \cdot p^{-3r-10} \leq \epsilon p^{-r}/4$, and Claim 1 is proved. □

Claim 2. *Let α_x be the density of A_x in B_x, then $\mathbb{E}_{x \in V} \alpha_x \geq \alpha p^{s-r}(1 - \epsilon/4)$.*

Proof. Let $X = \{x \in V \mid \mathrm{codim} B_x = s\}$, then we have

$$\begin{aligned}
\mathbb{E}_{x \in V} \alpha_x &= \frac{1}{|V|} \sum_{x \in V} \frac{|A_x|}{|B_x|} \geq \frac{1}{|V|} \sum_{x \in X} \frac{|A_x|}{|B_x|} \\
&= p^s \mathbb{E}_{x \in V} \frac{|A_x|}{|W|} - p^s \frac{1}{|V|} \sum_{x \in X^c} \frac{|A_x|}{|W|} \\
&\geq \alpha p^{s-r} - \alpha p^s \frac{|X^c|}{|V|} \geq \alpha p^{s-r}(1 - \epsilon/4)
\end{aligned}$$

where we have used (7) and the trivial bound $|A_x| \leq |A| = \alpha |W|$. □

Proposition 11 will follow from Lemmas 12 and 13.

Lemma 12. *At least one of the following statements holds.*

1. *For at least $p^{-r}\,|V|\,/12$ elements $x \in V$, we have $2A_x - 2A_x = \overrightarrow{B_x}$ (the direction of B_x).*
2. *Among additive quadruples $y_1 + y_2 = y_3 + y_4$ in A, a proportion at least $p^{-4r}\epsilon$ has the property that $\mathrm{codim} \bigcap_{i=1}^4 V_{y_i} < 4r - s$.*

The quantity $\mathrm{codim} \bigcap_{i=1}^4 V_{y_i}$, in a sense, measures the linear dependence between the V_{y_i}. Recall that the V_{y_i} all have codimension r and satisfy s simultaneous constraints. If $\mathrm{codim} \bigcap_{i=1}^4 V_{y_i} < 4r - s$, then we will be able to find a new linear dependence between them.

Proof. Let Q be the set of additive quadruples $\mathbf{y} = (y_1, \dots, y_4)$ of A. Let $m = \dim W$. We have

$$\begin{aligned}
\mathbb{E}_{x\in V}\|1_{A_x}\|_{U^2(B_x)}^4 &= \mathbb{E}_{x\in V}\mathbb{E}_{\substack{y_1,\dots,y_4\in B_x\\ y_1+y_2=y_3+y_4}}\prod_{i=1}^4 1_{y_i\in A_x}\\
&\le \frac{1}{p^{3(m-s)}}\sum_{\substack{(y_1,\dots,y_4)\in A^4\\ y_1+y_2=y_3+y_4}}\mathbb{E}_{x\in V}1_{\forall i\, x\in V_{y_i}} \qquad (\text{since } \dim B_x \ge m-s)\\
&= \frac{1}{p^{3(m-s)}}\sum_{\mathbf{y}\in Q} p^{-\operatorname{codim}\bigcap_i V_{y_i}}\\
&\le \alpha^4(1+\epsilon)\big(\mathbb{E}_{\mathbf{y}\in Q}(p^{3s-\operatorname{codim}\bigcap_i V_{y_i}})\big) \qquad (\text{recall that } A \text{ is } \epsilon\text{-pseudorandom})\\
&\le \alpha^4 p^{-4(r-s)}(1+\epsilon)(1+p^{4r-s}\mathbb{P}_{\mathbf{y}\in Q}(\operatorname{codim}\bigcap V_{y_i} < 4r-s)).
\end{aligned}$$

So either

$$\mathbb{P}_{\mathbf{y}\in Q}\left(\operatorname{codim}\bigcap_{i=1}^4 V_{y_i} < 4r-s\right) \ge p^{-4r+s}\epsilon \tag{8}$$

or

$$\mathbb{E}_{x\in V}\|1_{A_x}\|_{U^2(B_x)}^4 \le \alpha^4 p^{-4(r-s)}(1+\epsilon)(1+p^{4r-s}\cdot p^{-4r+s}\epsilon) = \alpha^4 p^{-4(r-s)}(1+\epsilon)^2. \tag{9}$$

Eq. (8) is exactly the second clause of Lemma 12, so assume instead that (9) holds. We infer that

$$\begin{aligned}
\mathbb{E}_{x\in V}\|1_{A_x}-\alpha_x\|_{U^2}^4 &= \mathbb{E}_{x\in V}(\|1_{A_x}\|_{U^2}^4 - \alpha_x^4)\\
&\le \mathbb{E}_{x\in V}\|1_{A_x}\|_{U^2}^4 - \alpha^4 p^{-4(r-s)}(1-\epsilon/4)^4\\
&\le \alpha^4 p^{-4(r-s)}(2\epsilon+\epsilon^2+\epsilon)\\
&\le 4\epsilon\alpha^4 p^{-4(r-s)} =: \gamma
\end{aligned}$$

where we used Jensen's inequality, the lower bound $\mathbb{E}_{x\in V}\alpha_x \ge p^{-(r-s)}(1-\epsilon p^{-r}/4)$ and the elementary inequality $(1-\epsilon/4)^4 \ge 1-\epsilon$. Thus, if $X_1 := \{x \in V \mid \|1_{A_x}-\alpha_x\|_{U^2(B_x)}^4 \le 4p^r\gamma\}$, by Markov's inequality, we have $|X_1| \ge |V|(1-p^{-r}/4)$. Also, because $\alpha_x \le \alpha p^s$ for any $x \in V$, the set $X_2 := \{x \in V \mid \alpha_x > \alpha p^{-(r-s)}/2\}$ has density at least $p^{-r}(1/2-\epsilon/4) \ge p^{-r}/3$. So $X_3 := X_1 \cap X_2$ must have density at least $p^{-r}/12$ by inclusion–exclusion.

Besides, if $\epsilon = 1/(256p^r)$, then for $x \in X_3$ we have $\|1_{A_x}-\alpha_x\|_{U^2}^4 \le 4p^r\gamma < \alpha_x^4$ and then $2A_x - 2A_x = \overrightarrow{B_x}$ by Lemma 5. □

We now prove Proposition 11. When the first outcome of Lemma 12 holds, we see that $\phi_V(P)$ contains $\{(x,y) \in X_3 \times W_2 \mid x\cdot\overrightarrow{\xi_1}(y) = \cdots = x\cdot\overrightarrow{\xi_s}(y) = 0\}$ where W_2 is the direction of W.

The real challenge lies in extracting something from the second outcome of Lemma 12. This is the purpose of the next lemma.

Lemma 13. *Suppose $r > s$ and a proportion at least κ of the additive quadruples $(y_1, \dots, y_4)$ of A have the property that $\operatorname{codim}\bigcap_{i=1}^4 V_{y_i} < 4r-s$. Then there is a subset $S \subset A$ of density $\sigma = \sigma(r,\alpha,\kappa)$ such that one of the following holds.*

(1) *There is a subspace $V' \le V$ of codimension one such that $V_y \subset V'$ for all $y \in S$. Moreover, there exist affine maps $\xi_1', \dots, \xi_{s-1}'$ from W to V'^* and spaces $U_y' \le V'^*$ of dimension at most $r-s$ such that $(V_y)^\perp = \operatorname{span}(\xi_1'(y), \dots, \xi_{s-1}'(y)) + U_y'$ (where $(V_y)^\perp$ is defined in V').*

(2) *There is an affine map* $\xi_{s+1} : W \to V^*$ *and a subspace* $U'_y \leq V^*$ *of dimension at most* $r-s-1$ *such that* $V_y^\perp = \mathrm{span}(\xi_j(y) \mid j \in [s+1]) + U'_y$.

Moreover σ *can be taken to be quasipolynomial*[1] *in* $\alpha\kappa p^{-r}$.

Applying Lemma 13 with $\kappa = p^{-4r}\epsilon = p^{-5r}/256$, the first alternative implies the second statement of Proposition 11 since $\mathrm{codim}_{V'} V_y \leq r-1$, while the second alternative yields the third one of Proposition 11. Our goal is now to prove Lemma 13.

Proof of Lemma 13. Let $\xi_{s+1}, \dots, \xi_r$ be (not necessarily linear) maps from A to V^* such that $U_y = \mathrm{span}(\xi_{s+1}(y), \dots, \xi_r(y))$ for any $y \in A$. The number of additive quadruples in A is at least $\alpha^4 |W|^3 = \alpha |A|^3$, and we assume at least $\kappa\alpha |A|^3$ of them have the property that the $4r$ vectors $\xi_j(y_i)$ satisfy at least $s+1$ linearly independent equations. For any additive quadruple in A, we already have s obvious equations

$$\xi_j(y_1) + \xi_j(y_2) = \xi_j(y_3) + \xi_j(y_4) \quad \text{for } j \in [s], \tag{10}$$

so there needs to be one more (independent) equation. Because there are only p^{4r} possible linear equations

$$\sum_{j=1}^{r} \sum_{i=1}^{4} a_{i,j} \xi_j(y_i) = 0, \tag{11}$$

the pigeonhole principle implies that we can find $(a_{i,j}) \in \mathbb{F}^{4r} \setminus \{0\}$ (linearly independent from the vectors $b_{i,j} = 1_{j=j_0}$ for $j_0 \in [s]$) such that there are at least $\kappa\alpha |A|^3 / p^{4r}$ quadruples $(y_1, \dots, y_4) \in W^4$ for which $y_1 + y_2 = y_3 + y_4$ and Eq. (11) holds. Let T be that set of quadruples. Write $\mathbf{a}_i = (a_{i,j})_{j=1,\dots,r}$. We distinguish two cases.

Case 1: One of the four families $\mathbf{a}_1, \dots, \mathbf{a}_4$, say $\mathbf{a}_4$, satisfies $a_{4,j} = 0$ for any $j > s$. Then we can use the Eqs. (10) to eliminate y_4 in Eq. (11). We obtain $\phi_1 + \phi_2 + \phi_3 = 0$ for some vectors $\phi_i \in V_{y_i}^\perp$ for $i \in [3]$, not all equal to 0. Write $r(\phi) = \left|\{y \in A \mid \phi \in V_y^\perp\}\right|$ for any $\phi \in V^*$. Then we have

$$p^{-4r}\kappa\alpha |A|^3 \leq |T| \leq \sum_{\substack{\phi_1,\phi_2 \in V^*, \\ (\phi_1,\phi_2) \neq (0,0)}} r(\phi_1) r(\phi_2) r(-\phi_1-\phi_2) \leq 2 \max_{\phi \in V^* \setminus \{0\}} r(\phi) \left(\sum_{\phi \in V} r(\phi) \right)^2.$$

A double counting argument shows that $\sum_{\phi \in V^*} r(\phi) \leq p^r |A|$. So $\max_{\phi \in V^* \setminus \{0\}} r(\phi) \geq \frac{1}{2}\kappa\alpha p^{-6r} |A|$, which implies that there exists a linear form $\phi \in V^* \setminus \{0\}$ such that for a noticeable proportion of $y \in A$, we have $\phi \in V_y^\perp$. Name $S_0 \subset A$ this set of elements $y \in A$, then $|S_0| \geq \frac{1}{2}\kappa\alpha p^{-6r} |A|$. Let $V' = \phi^\perp$, then $\mathrm{codim}_V V' = 1$ and $V_y \leq V'$ for any $y \in S_0$. Let $S_1 := \{y \in A : \phi \in \mathrm{span}(\xi_1(y), \dots, \xi_s(y))\}$ and $S_2 := S_0 \setminus S_1$. So $\max(|S_1|, |S_2|) \geq \frac{1}{4}\kappa\alpha p^{-6r} |A|$. We distinguish two subcases.

Case 1a: $|S_1| \geq \frac{1}{4}\kappa\alpha p^{-6r} |A|$. Since there are at most p^r possible ways to write ϕ as a linear combination of $\xi_1(y), \dots, \xi_s(y)$, there is a nonzero $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_s) \in \mathbb{F}_p^s$ and a subset $S_3 \subset S_1$ of size $\gg \kappa\alpha p^{-7r}$ such that $\phi = \sum_{i=1}^s \lambda_i \xi_i(y)$ for any $y \in S_3$. By completing λ into a basis of $\mathbb{F}_p^s$, we can find affine maps $\xi'_1, \dots, \xi'_{s-1} : W \to V^*$ such that $\mathrm{span}(\xi_1(y), \dots, \xi_s(y)) = \mathrm{span}(\phi, \xi'_1(y), \dots, \xi'_{s-1}(y))$ for any $y \in S_3$. Considering now V_y as a subspace of V', and thus defining its orthogonal as a subspace of V'^*, we have

$$V_y^\perp = \mathrm{span}(\xi'_1(y), \dots, \xi'_{s-1}(y)) + U'_y$$

for any $y \in S_3$, where U'_y is the projection of U_y on V'^*. Thus the first alternative of Lemma 13 follows with $S = S_3$.

Case 1b: $|S_2| \geq \frac{1}{4}\kappa\alpha p^{-6r} |A|$. For $y \in S_2$, let $U'_y \leq V_y^\perp$ be such that

$$V_y^\perp = \mathrm{span}(\xi_1(y), \dots, \xi_s(y), \phi) \oplus U'_y,$$

[1] Polynomial under the polynomial Freiman–Ruzsa conjecture.

then $\dim U'_y \leq r - s - 1$. Projecting to V', we have

$$V_y^{\perp} = \mathrm{span}(\xi_1(y), \ldots, \xi_s(y)) + U'_y$$

for any $y \in S_2$ and replacing U'_y by $U'_y + \mathrm{span}(\xi_s(y))$, the first alternative of Lemma 13 follows with $S = S_2$.

Case 2: None of the four families satisfies $a_{i,j} = 0$ for any $j > s$. We shall aim at linearity instead of constancy. Removing quadruples for which two entries are equal (there are $O(|A|^2)$ such quadruples), we still have a set T' of quadruples satisfying $\left|T'\right| \geq C\,|A|^3$ for some constant $C \gg \kappa\alpha p^{-4r}$. Applying Lemma 10, we can pick a partition $A = A_1 \cup A_2 \cup A_3 \cup A_4$ such that the set

$$T'' = T' \cap A_1 \times A_2 \times A_3 \times A_4$$

satisfies $\left|T''\right| \geq C\,|A|^3/256$. For $i \in [4]$ and $y \in A_i$, set $\xi'_{s+1}(y) = z_i \sum_{j\in[r]} a_{i,j}\xi_j(y)$ where $z_1 = z_2 = 1$ and $z_3 = z_4 = -1$. Observe that $\xi'_{s+1}(y)$ is a nonzero vector in $V_y^{\perp}$. For $i \in [4]$, let $j_i > s$ be any index such that $a_{i,j_i} \neq 0$. For $y \in A_i$, let $U'_y = \mathrm{span}(\xi_{s+1}(y), \ldots, \widehat{\xi_{j_i}(y)}, \ldots, \xi_r(y))$, where the hat denotes an omitted form; this is a space of dimension at most $r - s - 1$. We have $V_y^{\perp} = \mathrm{span}(\xi_1(y), \ldots, \xi_s(y), \xi'_{s+1}(y)) + U'_y$. Further, for a quadruple $\mathbf{y} \in T''$, we observe that $(y_i, \xi'_{s+1}(y_i))_{i\in[4]}$ is an additive quadruple. So there are at least $C\,|A|^3/256$ additive quadruples in the graph $\{(y, \xi'_{s+1}(y)) \mid y \in A\}$. We then invoke Lemma 9 to obtain a set $S \subset A$ of quasipolynomial (in C) density in A, such that ξ'_{s+1} coincides with an affine map on S. This concludes the proof. □

Acknowledgments

The authors are thankful to Ben Green and Terence Tao for suggesting the bilinear Bogolyubov theorem and sharing ideas on how it could be proven. The first author is also grateful to his supervisor Julia Wolf and Tom Sanders for useful conversations. We would like to thank the referees for a careful reading and very useful comments which help to improve the presentation of the paper. Part of this work was carried out while the first author was staying at the Simons Institute for the theory of computing and supported by a travel grant of the University of Bristol Alumni Foundation, United Kingdom. The second author was supported by National Science Foundation, United States Grant DMS-1702296 and a Ralph E. Powe Junior Faculty Enhancement Award from Oak Ridge Associated Universities, United States.

References

[1] P.-Y. Bienvenu, T.H. Lê, Linear and quadratic uniformity of the Möbius function over $\mathbb{F}_q[t]$. arXiv:1711.05349.
[2] N. Bogolioùboff, Sur quelques propriétés arithmétiques des presque-périodes, Ann. Chaire Phys. Math. Kiev 4 (1939) 185–205.
[3] W.T. Gowers, L. Milićević, A bilinear version of Bogolyubov's theorem arXiv:1712.00248.
[4] B. Green, Finite field models in additive combinatorics, in: Surveys in combinatorics 2005, in: London Math. Soc. Lecture Note Ser, vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27.
[5] B. Green, Montréal notes on quadratic Fourier analysis, in: Additive Combinatorics, in: CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc, Providence, RI, 2007, pp. 69–102.
[6] I. Kříž, Large independent sets in shift-invariant graphs: solution of bergelson's problem, Graphs Combin. 3 (2) (1987) 145–158.
[7] S. Lovett, K. Hosseini, A bilinear Bogolyubov-Ruzsa lemma with poly-logarithmic bounds, preprint, arXiv:1808:049651.
[8] T. Sanders, On the bogolyubov-Ruzsa lemma, Anal. PDE 5 (3) (2012) 627–655.
[9] J. Wolf, Finite field models in arithmetic combinatorics—ten years on, Finite Fields Appl. 32 (2015) 233–274.