Electronic Colloquium on Computational Complexity, Report No. 142 (2018)

A bilinear Bogolyubov-Ruzsa lemma with poly-logarithmic bounds

Kaave Hosseini*
University of California, San Diego
skhossei@ucsd.edu

Shachar Lovett†
University of California, San Diego
slovett@ucsd.edu

August 14, 2018

Abstract

The Bogolyubov-Ruzsa lemma, in particular the quantitative bounds obtained by Sanders, plays a central role in obtaining effective bounds for the inverse U^3 theorem for the Gowers norms. Recently, Gowers and Milićević applied a bilinear Bogolyubov-Ruzsa lemma as part of a proof of the inverse U^4 theorem with effective bounds. The goal of this note is to obtain quantitative bounds for the bilinear Bogolyubov-Ruzsa lemma which are similar to those obtained by Sanders for the Bogolyubov-Ruzsa lemma.

We show that if a set $A \subset \mathbb{F}^n \times \mathbb{F}^n$ has density α, then after a constant number of horizontal and vertical sums, the set A would contain a bilinear structure of co-dimension $r = \log^{O(1)} \alpha^{-1}$. This improves the results of Gowers and Milićević which obtained similar results with a weaker bound of $r = \exp(\exp(\log^{O(1)} \alpha^{-1}))$ and by Bienvenu and Lê which obtained $r = \exp(\exp(\exp(\log^{O(1)} \alpha^{-1})))$.

1 Introduction

One of the key ingredients in the proof of quantitative inverse theorem for Gowers U^3 norm over finite fields, due to Green and Tao [GT08] and Samorodnitsky [Sam07], is an inverse theorem on the structure of sumsets. More concretely, the tool that gives the best bounds is the improved Bogolyubov-Ruzsa lemma due to Sanders [San12]. Before introducing it, we set some common notation. We assume that $\mathbb{F} = \mathbb{F}_p$ is a prime field where p is a fixed constant, and suppress the exact dependence on p in the bounds. Given a subset $A \subset \mathbb{F}^n$ its density is $\alpha = |A|/|\mathbb{F}|^n$. The sumset of A is $2A = A + A = \{a + a' : a, a' \in A\}$ and its difference set is $A - A = \{a - a' : a, a' \in A\}$.

*Supported by NSF grant CCF-1614023.
†Supported by NSF grant CCF-1614023.

ISSN 1433-8092

Theorem 1.1. *([San12]) Let $A \subset \mathbb{F}^n$ be a subset of density α. Then there exists a subspace $V \subset 2A - 2A$ of co-dimension $O(\log^4 \alpha^{-1})$.*

In fact the link between the inverse U^3 theorem and inverse sumset theorems is deeper. It was shown in [GT10, Lov12] that an inverse U^3 conjecture with polynomial bounds is equivalent to the polynomial Freiman-Ruzsa conjecture, one of the central open problems in additive combinatorics. Given this, one can not help but wonder whether there is a more general inverse sumset phenomena that would naturally correspond to quantitative inverse theorems for U^k norms. In a recent breakthrough, Gowers and Milićević [GM17b] showed that this is indeed the case, at least for the U^4 norm. They used a *bilinear* generalization of Theorem 1.1 to obtain a quantitative inverse U^4 theorem.

To be able to explain this result we need to introduce some notation. Let $A \subset \mathbb{F}^n \times \mathbb{F}^n$. Define two operators, capturing subtraction on horizontal and vertical fibers as follows:

$$\phi_{\mathrm{h}}(A) := \{(x_1 - x_2, y) : (x_1, y), (x_2, y) \in A\},$$
$$\phi_{\mathrm{v}}(A) := \{(x, y_1 - y_2) : (x, y_1), (x, y_2) \in A\}.$$

Given a word $w \in \{\mathrm{h}, \mathrm{v}\}^k$ define $\phi_w = \phi_{w_1} \circ \ldots \circ \phi_{w_k}$ to be their composition. A *bilinear variety* $B \subset \mathbb{F}^n \times \mathbb{F}^n$ of co-dimension $r = r_1 + r_2 + r_3$ is a set defined as follows:

$$B = \{(x, y) \in V \times W : b_1(x, y) = \ldots = b_{r_3}(x, y) = 0\},$$

where $V, W \subset \mathbb{F}^n$ are subspaces of co-dimension r_1, r_2, respectively, and $b_1, \ldots, b_{r_3} : \mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$ are bilinear forms.

Gowers and Milićević [GM17a] and independently Bienvenu and Lê [BL17] proved the following, although [BL17] obtained a weaker bound of $r = \exp(\exp(\exp(\log^{O(1)} \alpha^{-1})))$.

Theorem 1.2 ([GM17a, BL17])**.** *Let $A \subset \mathbb{F}^n \times \mathbb{F}^n$ be of density α and let $w = \mathrm{hhvvhh}$. Then there exists a bilinear variety $B \subset \phi_w(A)$ of co-dimension $r = \exp(\exp(\log^{O(1)} \alpha^{-1}))$.*

To be fair, it was not Theorem 1.2 directly but a more analytic variant of it that was used (combined with many other ideas) to prove the inverse U^4 theorem in [GM17b]. However, we will not discuss that analytical variant here.

The purpose of this note is to improve the bound in Theorem 1.2 to $r = \log^{O(1)} \alpha^{-1}$. Our proof is arguably simpler and is obtained only by invoking Theorem 1.1 a few times, without doing any extra Fourier analysis. The motivation behind this work — other than obtaining the right form of bound — is to employ this result in a more algebraic framework to obtain a modular and simpler proof of an inverse U^4 theorem.

One more remark before explaining the result is that Theorem 1.2 generalizes Theorem 1.1 because given a set $A \subset \mathbb{F}^n$, one can apply Theorem 1.2 to the set $A' = \mathbb{F}^n \times A$ and find $\{x\} \times V \subset \phi_w(A')$ where x is arbitrary, and V a subspace of co-dimension $3r$. This implies $V \subset 2A - 2A$.

Theorem 1.3 (**Main theorem**)**.** *Let $A \subset \mathbb{F}^n \times \mathbb{F}^n$ be of density α and let $w = \mathrm{hvvhvvvhh}$. Then there exists a bilinear variety $B \subset \phi_w(A)$ of co-dimension $r = O(\log^{80} \alpha^{-1})$.*

Note that the choice of the word w in Theorem 1.3 is $w = \mathrm{hvvhvvvhh}$ which is slightly longer than in Theorem 1.2 being hhvvhh. However, for applications this usually does not matter and any constant length w would do the job. In fact allowing w to be longer is what enables us to obtain a result with a stronger bound.

1.1 A robust analog of Theorem 1.3

Going back to the theorem of Sanders, there is a more powerful variant of Theorem 1.1 which guarantees that V enjoys a stronger property rather than just being a subset of $2A - 2A$. The stronger property is that every element $y \in V$ can be written in many ways as $y = a_1 + a_2 - a_3 - a_4$, with $a_1, a_2, a_3, a_4 \in A$. This stronger property of V has a number of applications such as obtaining upper bounds for Roth theorem in four variables. We refer the reader to [SS16] where Theorem 3.2 is similarly obtained from Theorem 1.1 and also for the noted application.

Theorem 1.4 ([San12, SS16])**.** *Let $A \subset \mathbb{F}^n$ be a subset of density α. Then there exists a subspace $V \subset 2A - 2A$ of co-dimension $O(\log^4 \alpha^{-1})$ such that the following holds. Every $y \in V$ can be expressed as $y = a_1 + a_2 - a_3 - a_4$ with $a_1, a_2, a_3, a_4 \in A$ in at least $\alpha^{O(1)}|\mathbb{F}|^{3n}$ many ways.*

In Section 3 we also prove a statistical analog of Theorem 1.4 by slightly modifying the proof of Theorem 1.3. To explain it, we need just a bit more notation.

Fix an arbitrary $(x, y) \in \mathbb{F}^n \times \mathbb{F}^n$, and note that (x, y) can be written as $(x, y) = \phi_{\mathrm{h}}((x + x_1, y), (x_1, y))$ for any $x_1 \in \mathbb{F}^n$. Moreover, for any fixed x_1, each of the points $(x + x_1, y), (x_1, y)$ can be written as $(x + x_1, y) = \phi_{\mathrm{v}}((x + x_1, y + y_1), (x + x_1, y_1))$ and $(x_1, y) = \phi_{\mathrm{v}}((x_1, y + y_2), (x_1, y_2))$ for arbitrary $y_1, y_2 \in \mathbb{F}^n$. So over all, the point (x, y) can be written using the operation ϕ_{vh} in exactly $|\mathbb{F}^n|^3$ many ways, namely, the total number of two-dimensional parallelograms $(x+x_1, y+y_1), (x+x_1, y_1), (x_1, y+y_2), (x_1, y_2)$ where (x, y) is fixed. We can continue this and consider an arbitrary word $w \in \{\mathrm{h}, \mathrm{v}\}^k$. Then (x, y) can be written using the operation ϕ_w in exactly $|\mathbb{F}^n|^{2^k-1}$ many ways.

Now, we have a set $A \subset \mathbb{F}^n \times \mathbb{F}^n$ and fix a word $w \in \{\mathrm{h}, \mathrm{v}\}^k$. Define $\phi_w^\varepsilon(A)$ to be the set of all elements $(x, y) \in \mathbb{F}^n \times \mathbb{F}^n$ that can be obtained in at least $\varepsilon|\mathbb{F}^n|^{2^k-1}$ many ways by applying the operation $\phi_w(A)$.

The following is an extension of Theorem 1.3 similar in spirit to Theorem 1.4.

Theorem 1.5. *Let $A \subset \mathbb{F}^n \times \mathbb{F}^n$ be of density α and let $w = \mathrm{hvvhvvvhh}$ and $\varepsilon = \exp(-O(\log^{20} \alpha^{-1}))$. Then there exists a bilinear variety $B \subset \phi_w^\varepsilon(A)$ of co-dimension $r = O(\log^{80} \alpha^{-1})$.*

As a final comment, we remark that if one keeps track of dependence on the field size in the proofs, then the bound in Theorem 1.3 and Theorem 1.5 is $r = O(\log^{80} \alpha^{-1} \cdot \log^{O(1)} |\mathbb{F}|)$.

Paper organization. We prove Theorem 1.3 in Section 2 and Theorem 1.5 in Section 3.

2 Proof of Theorem 1.3

We prove Theorem 1.3 in six steps. It corresponds to applying chain of operators $\phi_{\mathrm{h}} \circ \phi_{\mathrm{vv}} \circ \phi_{\mathrm{h}} \circ \phi_{\mathrm{v}} \circ \phi_{\mathrm{vv}} \circ \phi_{\mathrm{hh}}$ to A. In the proof, we invoke Theorem 1.1 (or Theorem 1.4, or the Freiman-Ruzsa theorem which is a corollary of Theorem 1.1), four times in total, in steps 1,2,4, and 5.

We will assume that $A \subset \mathbb{F}^m \times \mathbb{F}^n$, where initially $m = n$ but where throughout the proof we update m, n independently when we restrict x or y to large subspaces. It also helps readability, as we will always have that x and related sets or subspaces are in $\mathbb{F}^m$, while y and related sets or subspace are in $\mathbb{F}^n$.

We use three variables r_1, r_2, r_3 that hold the total number of linear forms on x, linear forms on y, and bilinear forms on (x, y) that are being fixed throughout the proof, respectively. Initially, $r_1 = r_2 = r_3 = 0$, but their values will be updated as we go along and at the end, $r = \max(r_1, r_2, r_3)$ will be the codimension of the final bilinear variety.

Step 1. Decompose $A = \bigcup_{y \in \mathbb{F}^n} A_y \times \{y\}$ with $A_y \subset \mathbb{F}^m$. Define $A^1 := \phi_{\mathrm{hh}}(A)$, so that

$$A^1 = \bigcup_{y \in \mathbb{F}^n} (2A_y - 2A_y) \times \{y\}.$$

Let α_y denote the density of A_y. By Theorem 1.1, there exists a linear subspace $V'_y \subset 2A_y - 2A_y$ of co-dimension $O(\log^4 \alpha_y^{-1})$. Let $S := \{y : \alpha_y \geq \alpha/2\}$, where by averaging S has density $\geq \alpha/2$. Note that for every $y \in S$ the co-dimension of each V'_y is $O(\log^4 \alpha^{-1})$. We have

$$B^1 := \bigcup_{y \in S} V'_y \times \{y\} \subset A^1.$$

Step 2. Consider $A^2 := \phi_{\mathrm{vv}}(B^1)$. It satisfies

$$A^2 = \bigcup_{y_1, y_2, y_3, y_4 \in S} \left(V'_{y_1} \cap V'_{y_2} \cap V'_{y_3} \cap V'_{y_4}\right) \times \{y_1 + y_2 - y_3 - y_4\}.$$

By Theorem 1.1, there is a subspace $W' \subset 2S - 2S$ of co-dimension $O(\log^4 \alpha^{-1})$. Note that the co-dimension of W', as well as the co-dimension of each $V'_{y_1} \cap V'_{y_2} \cap V'_{y_3} \cap V'_{y_4}$, is at most $O(\log^4 \alpha^{-1})$. We thus have

$$B^2 := \bigcup_{y \in W'} V_y \times \{y\} \subset A^2,$$

where $V_y = V'_{y_1} \cap V'_{y_2} \cap V'_{y_3} \cap V'_{y_4}$ for some $y_1, y_2, y_3, y_4 \in S$ which satisfy $y = y_1 + y_2 - y_3 - y_4$.

Update $r_2 := \text{co-dim}(W')$, where we restrict $y \in W'$. To simplify notations, identify $W' \cong \mathbb{F}^{n - \text{co-dim}(W')}$ and update $n := n - \text{co-dim}(W')$. Thus we assume from now that

$$B^2 := \bigcup_{y \in \mathbb{F}^n} V_y \times \{y\},$$

where each V_y has co-dimension $d = O(\log^4 \alpha^{-1})$.

Step 3. Consider $A^3 := \phi_{\mathrm{v}}(B^2)$. It satisfies

$$A^3 = \bigcup_{y,z \in \mathbb{F}^n} (V_z \cap V_{y+z}) \times \{y\}.$$

Step 4. Consider $A^4 := \phi_{\mathrm{h}}(A^3)$. It satisfies

$$A^4 = \bigcup_{y,z,w \in \mathbb{F}^n} ((V_z \cap V_{y+z}) + (V_w \cap V_{y+w})) \times \{y\}.$$

Define $U_y := V_y^{\perp}$, so that $\dim(U_y) = d$ and

$$A^4 = \bigcup_{y,z,w \in \mathbb{F}^n} ((U_z + U_{y+z}) \cap (U_w + U_{y+w}))^{\perp} \times \{y\}.$$

Next, observe that if $(U_z + U_{y+z}) \cap (U_w + U_{y+w}) = \{0\}$ for some z, w, then $\mathbb{F}^m \times \{y\} \subset A^4$. If this is true for a typical y, then A^4 has constant density in $\mathbb{F}^m \times \mathbb{F}^n$. Our goal is to get to that situation by fixing a few linear forms on x and bi-linear forms on (x, y).

The following lemma identifies common structure in the subspaces U_y in the case that for a typical y, z, w, $(U_z + U_{y+z}) \cap (U_w + U_{y+w}) \neq \{0\}$. We recall that an affine map $L : \mathbb{F}^n \to \mathbb{F}^m$ is $L(y) = My + b$ where $M \in \mathbb{F}^{m \times n}, b \in \mathbb{F}^m$.

Lemma 2.1. *For each $y \in \mathbb{F}^n$ let $U_y \subset \mathbb{F}^m$ be a subspace of dimension d. Assume that*

$$\Pr_{y,z,w \in \mathbb{F}^n} [(U_z + U_{y+z}) \cap (U_w + U_{y+w}) \neq \{0\}] \geq \frac{1}{2}.$$

Then there exists an affine function $L : \mathbb{F}^n \to \mathbb{F}^m$ such that

$$\Pr_{y \in \mathbb{F}^n} [L(y) \in U_y \setminus \{0\}] \geq \exp(-O(d^4)).$$

To prove Lemma 2.1, we use the Freiman-Ruzsa theorem, being a consequence of Theorem 1.1, which we quote below. We refer the reader to [Gre05] for details on how it is derived from Theorem 1.1.

Theorem 2.2. *Let $f : \mathbb{F}^n \to \mathbb{F}^m$ be a function such that*

$$\Pr_{y,z,w \in \mathbb{F}^n} [f(z) + f(y+z) = f(w) + f(y+w)] \geq \alpha.$$

Then there exists an affine map $L : \mathbb{F}^n \to \mathbb{F}^m$ so that

$$\Pr_{y \in \mathbb{F}^n} [f(y) = L(y)] \geq \exp(-O(\log^4 \alpha^{-1})).$$

Proof of Lemma 2.1. First assume that

$$\Pr_{y,z,w\in\mathbb{F}^n}[(U_z\setminus\{0\}+U_{y+z}\setminus\{0\})\cap(U_w\setminus\{0\}+U_{y+w}\setminus\{0\})\neq\{0\}]\geq\frac{1}{4}. \tag{1}$$

Choose $f:\mathbb{F}^n\to\mathbb{F}^m$ by picking $f(y)\in U_y\setminus\{0\}$ uniformly and independently for each $y\in\mathbb{F}^n$. Then

$$\Pr_{y,z,w\in\mathbb{F}^n,f}[f(z)+f(y+z)=f(w)+f(y+w)]\geq\frac{1}{4}\cdot|\mathbb{F}|^{-4d}.$$

Fix f where the above bound holds. By Theorem 2.2, there exists an affine function $L:\mathbb{F}^n\to\mathbb{F}^m$ such that

$$\Pr_{y\in\mathbb{F}^n}[f(y)=L(y)]\geq\exp(-O(d^4)).$$

This concludes the proof, assuming Equation (1) holds. Otherwise, if Equation (1) does not hold, then we have

$$\Pr_{y,z,w\in\mathbb{F}^n}[U_z\cap(U_w+U_y)\neq\{0\}]\geq\frac{1}{4}.$$

This implies that either

$$\Pr_{y,z,w\in\mathbb{F}^n}[(U_z\setminus\{0\})\cap(U_w\setminus\{0\}+U_y\setminus\{0\})\neq\{0\}]\geq\frac{1}{8}$$

or that

$$\Pr_{y,w\in\mathbb{F}^n}[(U_z\setminus\{0\})\cap(U_w\setminus\{0\})]\geq\frac{1}{8}.$$

In the first case, choose the most popular w,y and then elements of $U_w\setminus\{0\},U_y\setminus\{0\}$ to obtain a constant map $L\equiv b$ that satisfies the lemma. The second case is similar. $\square$

Next, we proceed as follows. As long as

$$\Pr_{y,z,w\in\mathbb{F}^n}[(U_z+U_{y+z})\cap(U_w+U_{y+w})\neq\{0\}]\geq\frac{1}{2},$$

apply Lemma 2.1 to find an affine function $L:\mathbb{F}^n\to\mathbb{F}^m$. For each y that satisfies $L(y)\in U_y$ replace U_y with $U'_y=U_y/\langle L(y)\rangle$, which is a subspace of co-dimension 1 in U_y. By Lemma 2.1, this process needs to stop after $t=\exp(O(d^4))$ many steps. Let $L_1,\ldots,L_t:\mathbb{F}^n\to\mathbb{F}^m$ be the affine maps obtained in this process.

We pause for a moment to introduce one useful notation. Given a set of maps $\mathcal{F}=\{f_i:\mathbb{F}^n\to\mathbb{F}^m,i\in[k]\}$ and $y\in\mathbb{F}^n$, let $\mathcal{F}(y)=\{f_1(y),\ldots,f_k(y)\}\subset\mathbb{F}^m$, and also let $\overline{\mathcal{F}}(y)$ denote the linear span of $\mathcal{F}(y)$.

Using this notation, set $\mathcal{F}=\{L_1,\ldots,L_t\}$ and note that $\overline{\mathcal{F}}(y)$ is a subspace of dimension at most t for each $y\in\mathbb{F}^n$. For every subspace U_y there is a set $\mathcal{F}_y\subset\mathcal{F}$ with $|\mathcal{F}_y|\leq d$ such that the final subspace obtained in the process is $U_y/\overline{\mathcal{F}_y}(y)$. This implies that

$$\Pr_{y,z,w\in\mathbb{F}^n}\left[\left(U_z/\overline{\mathcal{F}_z}(z)+U_{y+z}/\overline{\mathcal{F}_{y+z}}(y+z)\right)\cap\left(U_w/\overline{\mathcal{F}_w}(w)+U_{y+w}/\overline{\mathcal{F}_{y+w}}(y+w)\right)=\{0\}\right]\geq\frac{1}{2}.$$

Consider the most popular quadruple $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4 \subset \mathcal{F}$ so that

$$\Pr_{y,z,w\in\mathbb{F}^n}\left[\left(U_z/\overline{\mathcal{F}_1}(z) + U_{y+z}/\overline{\mathcal{F}_2}(y+z)\right)\cap\left(U_w/\overline{\mathcal{F}_3}(w) + U_{y+w}/\overline{\mathcal{F}_4}(y+w)\right) = \{0\}\right] \geq \frac{1}{2}\times\binom{t}{d}^{-4}.$$

Let $\mathcal{L} := \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3 \cup \mathcal{F}_4$. Recall that $t = \exp(O(d^4))$ so that $\binom{t}{d} = \exp(O(d^5))$. We have

$$\Pr_{y,z,w\in\mathbb{F}^n}\left[(U_z + U_{y+z}) \cap (U_w + U_{y+w}) \subset \overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(y+z) + \overline{\mathcal{L}}(w) + \overline{\mathcal{L}}(y+w)\right] \geq \exp(-O(d^5)).$$

By averaging, there is some choice of z, w such that

$$\Pr_{y\in\mathbb{F}^n}\left[(U_z + U_{y+z}) \cap (U_w + U_{y+w}) \subset \overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(y+z) + \overline{\mathcal{L}}(w) + \overline{\mathcal{L}}(y+w)\right] \geq \exp(-O(d^5)).$$

Recall that each L_i is an affine map and that $|\mathcal{L}| \leq 4d$. Thus, $\overline{\mathcal{L}}(z), \overline{\mathcal{L}}(y+z), \overline{\mathcal{L}}(w), \overline{\mathcal{L}}(y+w) \subset \overline{\mathcal{L}}(y) + Q$ where $Q \subset \mathbb{F}^m$ is a linear subspace of dimension $O(d)$. We thus have

$$B^4 := \bigcup_{y\in T} (\overline{\mathcal{L}}(y) + Q)^{\perp} \times \{y\} \subset A^4,$$

where $T \subset \mathbb{F}^n$ has density $\exp(-O(d^5))$.

To simplify the presentation, we would like to assume that the maps in $\mathcal{L}$ are linear maps instead of affine maps, that is, that they do not have a constant term. This can be obtained by restricting x to the subspace orthogonal to Q and to the constant term in the affine maps in $\mathcal{L}$. Correspondingly, we update $r_1 := r_1 + \dim(Q) + |\mathcal{L}| = O(d)$.

So, from now we assume that $\mathcal{L}$ is defined by $4d$ linear maps, and that

$$B^4 := \bigcup_{y\in T} \overline{\mathcal{L}}(y)^{\perp} \times \{y\} \subset A^4,$$

where $T \subset \mathbb{F}^n$ has density $\exp(-O(d^5))$.

Step 5. Consider $A^5 := \phi_{\text{vv}}(B^4)$ so that

$$A^5 = \bigcup_{y_1,y_2,y_3,y_4\in T} \left(\overline{\mathcal{L}}(y_1)^{\perp} \cap \overline{\mathcal{L}}(y_2)^{\perp} \cap \overline{\mathcal{L}}(y_3)^{\perp} \cap \overline{\mathcal{L}}(y_4)^{\perp}\right) \times \{y_1 + y_2 - y_3 - y_4\}.$$

By Theorem 1.1 there exists a subspace $W \subset 2T - 2T$ with co-dimension $O(d^{20})$. However, this time, this is not enough for us. We need to use Theorem 1.4 instead. The following equivalent formulation of Theorem 1.4 will be more convenient for us: there is a subspace $W \subset \mathbb{F}^n$ of co-dimension $O(\log^4 \alpha^{-1})$ such that, for each $y \in W$ there is a set $S_y \subset (\mathbb{F}^n)^3$ of density $\alpha^{O(1)}$, for which

$$\forall (a_1, a_2, a_3) \in S_y : \quad a_1, a_2, a_3, a_1 + a_2 - a_3 - y \in A.$$

Apply Theorem 1.4 to the set T to obtain the subspace W and the sets S_y. We have

$$B^5 := \bigcup_{y \in W} \left(\bigcup_{(y_1,y_2,y_3) \in S_y} \left(\overline{\mathcal{L}}(y_1) + \overline{\mathcal{L}}(y_2) + \overline{\mathcal{L}}(y_3) + \overline{\mathcal{L}}(y_1 + y_2 - y_3 - y) \right)^{\perp} \right) \times \{y\} \subset A^5.$$

To simplify the presentation we introduce the notation $\overline{\mathcal{L}}(y_1, y_2, y_3) := \overline{\mathcal{L}}(y_1) + \overline{\mathcal{L}}(y_2) + \overline{\mathcal{L}}(y_3)$. Next, observe that for any $y, y' \in \mathbb{F}^n$, $\overline{\mathcal{L}}(y') + \overline{\mathcal{L}}(y + y') \subset \overline{\mathcal{L}}(y') + \overline{\mathcal{L}}(y)$. Thus we can simplify the expression of B^5 to

$$B^5 = \bigcup_{y \in W} \left(\bigcup_{(y_1,y_2,y_3) \in S_y} \left(\overline{\mathcal{L}}(y_1, y_2, y_3) + \overline{\mathcal{L}}(y) \right)^{\perp} \right) \times \{y\},$$

which can be re-written as

$$B^5 = \bigcup_{y \in W} \left(\bigcup_{(y_1,y_2,y_3) \in S_y} \overline{\mathcal{L}}(y_1, y_2, y_3)^{\perp} \cap \overline{\mathcal{L}}(y)^{\perp} \right) \times \{y\}.$$

Step 6. Consider $A^6 := \phi_{\mathrm{h}}(B^5)$. It satisfies

$$A^6 = \bigcup_{y \in W} \left(\left(\bigcup_{\substack{(y_1,y_2,y_3) \in S_y \\ (y'_1,y'_2,y'_3) \in S_y}} \overline{\mathcal{L}}(y_1, y_2, y_3)^{\perp} + \overline{\mathcal{L}}(y'_1, y'_2, y'_3)^{\perp} \right) \cap \overline{\mathcal{L}}(y)^{\perp} \right) \times \{y\}$$

In order to complete the proof, we will find a large subspace V such that for every $y \in W$,

$$V \subset \bigcup_{\substack{(y_1,y_2,y_3) \in S_y \\ (y'_1,y'_2,y'_3) \in S_y}} \overline{\mathcal{L}}(y_1, y_2, y_3)^{\perp} + \overline{\mathcal{L}}(y'_1, y'_2, y'_3)^{\perp}.$$

In fact, we will prove something stronger: there is a large subspace V such that for each $y \in W$, there is a choice of $(y_1, y_2, y_3), (y'_1, y'_2, y'_3) \in S_y$ for which

$$V \subset \overline{\mathcal{L}}(y_1, y_2, y_3)^{\perp} + \overline{\mathcal{L}}(y'_1, y'_2, y'_3)^{\perp}.$$

The following lemma is key. Given a set $\mathcal{L}$ of linear maps from $\mathbb{F}^n$ to $\mathbb{F}^m$, let $\dim(\overline{\mathcal{L}})$ denote the dimension of linear span of $\mathcal{L}$ as a vector space over $\mathbb{F}$.

Lemma 2.3. *Fix $\delta > 0$. Let $\mathcal{L}$ be a set of linear maps from $\mathbb{F}^n$ to $\mathbb{F}^m$ with $\dim(\overline{\mathcal{L}}) = k$. Then there is a subspace $V \subset \mathbb{F}^m$ of co-dimension at most $(k+1)^2 \log \delta^{-1}$ such that the following holds. For every subset $S \subset \mathbb{F}^n$ of density at least δ, at least half the pairs $y, y' \in S$ satisfy that*

$$V \subset \overline{\mathcal{L}}(y)^{\perp} + \overline{\mathcal{L}}(y')^{\perp}.$$

Proof. The proof is by induction on $\dim(\overline{\mathcal{L}})$. Consider first the base case of $\dim(\overline{\mathcal{L}}) = 1$. Take some $M \in \overline{\mathcal{L}} \setminus \{0\}$. If $\operatorname{rank}(M) \le \log \delta^{-1} + 3$, then set $V = \operatorname{Im}(M)^\perp$ and notice that $\operatorname{Im}(M)^\perp \subset \overline{\mathcal{L}}(y)^\perp$ for any $y \in \mathbb{F}^n$. Otherwise do as follows. Fix arbitrary $L, L' \in \overline{\mathcal{L}} \setminus \{0\}$ and observe that

$$\Pr_{y,y' \in S}[L(y) = L'(y')] \le |\mathbb{F}|^{-(\log \delta^{-1}+3)}\delta^{-1}.$$

By applying the union bound over all pairs of $L, L' \in \overline{\mathcal{L}} \setminus \{0\}$, we obtain that

$$\Pr_{y,y' \in S}\left[\overline{\mathcal{L}}(y) \cap \overline{\mathcal{L}}(y') \neq \{0\}\right] \le |\mathbb{F}|^2 |\mathbb{F}|^{-(\log \delta^{-1}+3)}\delta^{-1} \le \frac{1}{2}.$$

The claim then holds for $V = \mathbb{F}^m$.

Now suppose $\dim(\overline{\mathcal{L}}) = k$. Again, if there is some $M \in \overline{\mathcal{L}} \setminus \{0\}$ with rank at most $2k + \log \delta^{-1} + 1$, then project every map down to $\operatorname{Im}(M)^\perp$. That is, consider the new family of maps

$$\mathcal{L}' = \{\operatorname{Proj}_{\operatorname{Im}(M)^\perp} L : L \in \mathcal{L}\}.$$

Note that $\overline{\mathcal{L}'}$ has dimension $k-1$ and so by induction hypothesis, there exists a subspace V' of co-dimension at most $k^2 \log \delta^{-1}$ such that, for at least half the pairs $y, y' \in S$ it holds that

$$V' \subset \overline{\mathcal{L}'}(y)^\perp + \overline{\mathcal{L}'}(y')^\perp.$$

The claim then holds for $V = V' \cap \operatorname{Im}(M)^\perp$.

Otherwise, similar to the base case, observe that

$$\Pr_{y,y' \in S}\left[\overline{\mathcal{L}}(y) \cap \overline{\mathcal{L}}(y') \neq \{0\}\right] \le |\overline{\mathcal{L}}|^2 |\mathbb{F}|^{-(2k+\log \delta^{-1}+1)}\delta^{-1} \le |\mathbb{F}|^{2k} |\mathbb{F}|^{-(2k+\log \delta^{-1}+1)}\delta^{-1} \le \frac{1}{2}.$$

In this case the claim holds for $V = \mathbb{F}^m$.

$\square$

We note that for Theorem 1.3 we only need a weaker form of Lemma 2.3, which states that at least one pair $y, y' \in S$ exists; however, we would need the stronger version for Theorem 1.5.

We apply Lemma 2.3 as follows. Define a new family of linear maps $\mathcal{L}^*$ from $\mathbb{F}^{3n}$ to $\mathbb{F}^m$ as follows. For each $L \in \mathcal{L}$ define three linear maps L_i, $i \in \{1, 2, 3\}$ by:

$$L_i : (\mathbb{F}^n)^3 \to \mathbb{F}^m, L_i(y_1, y_2, y_3) = L(y_i)$$

and let

$$\mathcal{L}^* := \{L_i : L \in \mathcal{L}, i \in [3]\}.$$

Apply Lemma 2.3 to the family $\mathcal{L}^*$ with $\delta = \exp(-O(d^5))$ and obtain a subspace $V \subset \mathbb{F}^m$ of codimension $O(d^2 \log(\exp(-O(d^5))) = O(d^7)$ so that, for every $S_y \subset (\mathbb{F}^n)^3$ with $y \in W$, there exist $(y_1, y_2, y_3), (y_1', y_2', y_3') \in S_y$ for which

$$V \subset \overline{\mathcal{L}^*}((y_1, y_2, y_3))^\perp + \overline{\mathcal{L}^*}((y_1', y_2', y_3'))^\perp.$$

This directly implies that

$$V \subset \overline{\mathcal{L}}(y_1, y_2, y_3)^\perp + \overline{\mathcal{L}}(y_1', y_2', y_3')^\perp.$$

Define

$$B^6 := \bigcup_{y \in W} \left(V \cap \overline{\mathcal{L}}(y)^\perp\right) \times \{y\} \subset A^6.$$

Observe that B^6 is a bilinear variety defined by co-dim(V) linear equations on x, co-dim(W) linear equations on y and $|\mathcal{L}|$ bilinear equations on (x, y).

To complete the proof we calculate the quantitative bounds obtained. We have $d = O(\log^4 \alpha^{-1})$ where α was the density of the original set A, and

$$\begin{aligned} r_1 &= O(d) + \text{co-dim}(V) = O(d^7), \\ r_2 &= O(d) + \text{co-dim}(W) = O(d^{20}), \\ r_3 &= |\mathcal{L}| = O(d). \end{aligned}$$

Together these give the final bound of $r = \max(r_1, r_2, r_3) = O(\log^{80} \alpha^{-1})$.

3 Proof of Theorem 1.5

In this section we prove Theorem 1.5 by slightly modifying the proof of Theorem 1.3. We point out the necessary modifications to proof of Theorem 1.3.

Step 1. In this step, we use Theorem 1.4 instead of Theorem 1.1 and directly obtain

$$B^1 \subset \phi_{\mathrm{hh}}^{\varepsilon_1}(A) \tag{2}$$

for $\varepsilon_1 = \alpha^{O(1)}$.

Step 2. Similarly in this step as well, using Theorem 1.4 instead of Theorem 1.1 gives

$$B^2 \subset \phi_{\mathrm{vv}}^{\varepsilon_2}(B^1) \tag{3}$$

with $\varepsilon_2 = \alpha^{O(1)}$. To recall, we assume for simplicity of exposition from now on that $B^2 = \bigcup_{y \in \mathbb{F}^n} V_y \times \{y\}$.

Steps 3 and 4. This step is slightly different than steps 1 and 2. Here, we are not able to directly produce some set B^4 that would satisfy $B^4 \subset \phi_{\mathrm{hv}}^{\varepsilon_4}(B^2)$. But what we can do is to apply the remaining operation ϕ_{hvvhv} altogether to B^2 and obtain the final bilinear structure B^6 that satisfies what we want, which is

$$B^6 \subset \phi_{\mathrm{hvvhv}}^{\varepsilon_6}(B^2) \tag{4}$$

for $\varepsilon_6 = \exp(-\text{poly}\log \alpha^{-1})$. Combining Equations (2) to (4) gives

$$B^6 \subset \phi^{\varepsilon}_{\text{hvvhvvvhh}}(A)$$

for $\varepsilon = \exp(-\text{poly}\log \alpha^{-1})$.

We establish Equation (4) in the rest of the proof. Recall that previously we showed that the following holds: there is a set of affine maps $\mathcal{L}$, with $|\mathcal{L}| = O(d)$, such that

$$\Pr_{y,w,z \in \mathbb{F}^n}\left[\left(\overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(y+z) + \overline{\mathcal{L}}(w) + \overline{\mathcal{L}}(y+w)\right)^{\perp} \subset \left(V_z^{\perp} \cap V_{y+z}^{\perp}\right) + \left(V_w^{\perp} \cap V_{y+w}^{\perp}\right) \right] \geq \exp(-O(d^5))$$

and consequently

$$\Pr_{y,w,z \in \mathbb{F}^n}\left[\left(\overline{\mathcal{L}}(y) + \overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(w)\right)^{\perp} \subset \left(V_z^{\perp} \cap V_{y+z}^{\perp}\right) + \left(V_w^{\perp} \cap V_{y+w}^{\perp}\right) \right] \geq \exp(-O(d^5)).$$

Remember that $d = O(\log^4 \alpha^{-1})$. Furthermore, we may assume the maps in $\mathcal{L}$ are linear (instead of affine) after we update $r_1 := r_1 + |\mathcal{L}| = O(d)$.

Then what we did in the proof of Theorem 1.3 was to fix one popular choice of w, z. However, here we can't do that, as we need many pairs of w, z. Let T be the set of y's that satisfy

$$\Pr_{w,z \in \mathbb{F}^n}\left[\left(\overline{\mathcal{L}}(y) + \overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(w)\right)^{\perp} \subset \left(V_z^{\perp} \cap V_{y+z}^{\perp}\right) + \left(V_w^{\perp} \cap V_{y+w}^{\perp}\right) \right] \geq \exp(-O(d^5)), \tag{5}$$

and so T has density $\exp(-O(d^5))$. We deduce something stronger from Equation (5) but we need to introduce some notation first.

For $A, B \subset \mathbb{F}^n$ let $A -_{\eta} B$ denote the set of all elements $c \in A - B$ that can be written in at least $\eta|\mathbb{F}^n|$ many ways as $c = a - b$ for $a \in A, b \in B$. To use this notation, note that if A, B are two subspaces of co-dimension k, then $A - B = A -_{\eta} B$ for $\eta = \exp(-O(k))$. This is because every element $c \in A - B$ can be written as $c = (a + v) - (b + v)$ where v is an arbitrary element in the subspace $A \cap B$ of codimension at most $2k$. So we can improve the Equation (5) to

$$\Pr_{w,z \in \mathbb{F}^n}\left[\left(\overline{\mathcal{L}}(y) + \overline{\mathcal{L}}(z) + \overline{\mathcal{L}}(w)\right)^{\perp} \subset \left(V_z^{\perp} \cap V_{y+z}^{\perp}\right) -_{\eta} \left(V_w^{\perp} \cap V_{y+w}^{\perp}\right) \right] \geq \exp(-O(d^5)), \tag{6}$$

for $\eta = \exp\left(-O(d)\right)$

Step 5. Similar to before, consider the subspace $W \subset 2T - 2T$ of co-dimension $O(d^{20})$ that is given by Theorem 1.4. This subspace W has the following property: fix arbitrary $y \in W$. Sample $y_1, y_2, y_3 \in \mathbb{F}^n$ uniformly and independently, and set $y_4 = -y + y_1 + y_2 - y_3$. Then with probability at least $\exp(-O(d^5))$ we have $y_1, y_2, y_3, y_4 \in T$. This means that if we furthermore sample $w_1, w_2, w_3, w_4, z_1, z_2, z_3, z_4 \in \mathbb{F}^n$ uniformly and independently, then, with probability at least $\exp(-O(d^5))$, the following four equations simultaneously hold:

$$\left(\overline{\mathcal{L}}(y_i) + \overline{\mathcal{L}}(z_i) + \overline{\mathcal{L}}(w_i)\right)^{\perp} \subset \left(V_{z_i}^{\perp} \cap V_{y_i+z_i}^{\perp}\right) -_{\eta} \left(V_{w_i}^{\perp} \cap V_{y_i+w_i}^{\perp}\right) \qquad \forall i = 1, \ldots, 4.$$

By computing the intersection of the left hand sides and the right hand sides we obtain that with probability at least $\exp(-O(d^5))$, it holds that

$$\left(\overline{\mathcal{L}}(y)+\sum_{i=1}^{3}\overline{\mathcal{L}}(y_i)+\sum_{i=1}^{4}\overline{\mathcal{L}}(z_i)+\sum_{i=1}^{4}\overline{\mathcal{L}}(w_i)\right)^{\perp}\subset\bigcap_{i=1}^{4}\left(\left(V_{z_i}^{\perp}\cap V_{y_i+z_i}^{\perp}\right)-_{\eta}\left(V_{w_i}^{\perp}\cap V_{y_i+w_i}^{\perp}\right)\right). \quad (7)$$

For a given $y\in\mathbb{F}^n, \mathbf{s}=(y_1,y_2,y_3,w_1,w_2,w_3,w_4,z_1,z_2,z_3,z_4)\in(\mathbb{F}^n)^{11}$, let

$$\mathcal{V}_{y,\mathbf{s}}=\bigcap_{i=1}^{4}\left(\left(V_{z_i}^{\perp}\cap V_{y_i+z_i}^{\perp}\right)-_{\eta}\left(V_{w_i}^{\perp}\cap V_{y_i+w_i}^{\perp}\right)\right),$$

where to recall $y_4=-y+y_1+y_2-y_3$. Observe that for any $\mathbf{s}$,

$$\bigcup_{y\in W}\mathcal{V}_{y,\mathbf{s}}\times\{y\}\subset\phi_{\mathrm{vvhv}}(B^2).$$

We rewrite Equation (7) more compactly as

$$\Pr_{\mathbf{s}}\left[\left(\overline{\mathcal{L}}(y)+\overline{\mathcal{L}}(\mathbf{s})\right)^{\perp}\subset\mathcal{V}_{y,\mathbf{s}}\right]\geq\exp(-O(d^5)), \quad (8)$$

where we use the notation $\overline{\mathcal{L}}(\mathbf{s})=\sum_{i=1}^{3}\overline{\mathcal{L}}(y_i)+\sum_{i=1}^{4}\overline{\mathcal{L}}(z_i)+\sum_{i=1}^{4}\overline{\mathcal{L}}(w_i)$.

Step 6. Now we consider applying the operation hvvhv altogether to B^2. Only the last operation h remains to be applied, which after doing so, we will find a subspace $V\subset\mathbb{F}^m$ of co-dimension $O(d^7)$ that satisfies the following: for any $y\in W$, choose $\mathbf{s_1},\mathbf{s_2}\in(\mathbb{F}^n)^{11}$ uniformly and randomly. Then with probability $\exp(-O(d^5))$,

$$V\cap\overline{\mathcal{L}}(y)^{\perp}\subset\mathcal{V}_{y,\mathbf{s_1}}-_{\eta}\mathcal{V}_{y,\mathbf{s_2}}.$$

where to recall $\eta=\exp(-O(d))$.

To do so, fix $y\in W$ and let S_y be the set of all tuples $\mathbf{s}=(y_1,y_2,y_3,w_1,w_2,w_3,w_4,z_1,z_2,z_3,z_4)\in(\mathbb{F}^n)^{11}$ that satisfy Equation (8). Note that the density of each S_y is at least $\exp(-O(d^5))$. To simplify notation denote $\mathbf{s}=(s_1,\dots,s_{11})$. We call up Lemma 2.3 in a similar way as we did before. Define a family $\mathcal{L}^*$ of linear maps, containing linear maps L_i for each $L\in\mathcal{L}$ and $i=1,\dots,11$, where

$$L_i:(\mathbb{F}^n)^{11}\to\mathbb{F}^m, L_i(\mathbf{s})=L(s_i).$$

Apply Lemma 2.3 to $\mathcal{L}^*$ and density parameter $\exp(-O(d^5))$. So, we obtain a subspace $V\subset\mathbb{F}^m$ of co-dimension $O(d^7)$ such that for each $y\in W$,

$$\Pr_{\mathbf{s_1},\mathbf{s_2}\in S_y}\left[V\subset\overline{\mathcal{L}}(\mathbf{s_1})^{\perp}+\overline{\mathcal{L}}(\mathbf{s_2})^{\perp}\right]\geq\frac{1}{2}, \quad (9)$$

which implies

$$\Pr_{\mathbf{s_1},\mathbf{s_2}\in(\mathbb{F}^n)^{11}}\left[V\cap\overline{\mathcal{L}}(y)^\perp\subset\mathcal{V}_{y,\mathbf{s_1}}-_\eta\mathcal{V}_{y,\mathbf{s_2}}\right]\geq\exp(-O(d^5)). \tag{10}$$

Define the final bilinear structure as

$$B^6:=\bigcup_{y\in W}\left(V\cap\overline{\mathcal{L}}(y)^\perp\right)\times\{y\}.$$

It satisfies

$$B^6\subset\phi^{\varepsilon_6}_{\mathrm{hvvhv}}(B^2)$$

for $\varepsilon_6=\exp(-O(d^5))$ and so over all

$$B^6\subset\phi^{\varepsilon}_{\mathrm{hvvhvvvhh}}(A)$$

for $\varepsilon=\exp(-O(d^5))$.

References

[BL17] Pierre-Yves Bienvenu and Thái Hoàng Lê. A bilinear Bogolyubov theorem. *arXiv preprint arXiv:1711.05349*, 2017.

[GM17a] WT Gowers and Luka Milićević. A bilinear version of Bogolyubov's theorem. *arXiv preprint arXiv:1712.00248*, 2017.

[GM17b] WT Gowers and Luka Milićević. A quantitative inverse theorem for the U^4 norm over finite fields. *arXiv preprint arXiv:1712.00241*, 2017.

[Gre05] B Green. Notes on the polynomial Freiman-Ruzsa conjecture. *preprint*, 2005. `http://people.maths.ox.ac.uk/greenbj/papers/PFR.pdf`.

[GT08] Ben Green and Terence Tao. An inverse theorem for the Gowers U^3 norm. *Proceedings of the Edinburgh Mathematical Society*, 51(1):73–153, 2008.

[GT10] Ben Green and Terence Tao. An equivalence between inverse sumset theorems and inverse conjectures for the U^3 norm. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 149, pages 1–19. Cambridge University Press, 2010.

[Lov12] Shachar Lovett. Equivalence of polynomial conjectures in additive combinatorics. *Combinatorica*, 32(5):607–618, 2012.

[Sam07] Alex Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 506–515. ACM, 2007.

[San12] Tom Sanders. On the Bogolyubov-Ruzsa lemma. *Analysis & PDE*, 5(3):627–655, 2012.

[SS16] Tomasz Schoen and Olof Sisask. Roth's theorem for four variables and additive structures in sums of sparse sets. In *Forum of Mathematics, Sigma*, volume 4. Cambridge University Press, 2016.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il