COMPOSITIO MATHEMATICA

# Subsets of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ without L-shaped configurations

Sarah Peluse

Compositio Math. **160** (2024), 176–236.

doi:10.1112/S0010437X2300756X



FOUNDATION
COMPOSITIO
MATHEMATICA

LONDON
MATHEMATICAL
SOCIETY
EST. 1865

Compositio Math. 160 (2024) 176–236
doi:10.1112/S0010437X2300756X

# Subsets of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ without L-shaped configurations

Sarah Peluse

Abstract

Fix a prime $p \geq 11$. We show that there exists a positive integer $m$ such that any subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ containing no nontrivial configurations of the form $(x, y)$, $(x, y + z)$, $(x, y + 2z)$, $(x + z, y)$ must have density $\ll 1/\log_m n$, where $\log_m$ denotes the $m$-fold iterated logarithm. This gives the first reasonable bound in the multidimensional Szemerédi theorem for a two-dimensional four-point configuration in any setting.

## 1. Introduction

Szemerédi's famous theorem on arithmetic progressions, which states that any subset of the integers with positive upper density contains arbitrarily long arithmetic progressions, has the following multidimensional generalization due to Furstenberg and Katznelson [FK78].

Theorem 1.1. *Let $X$ be a finite, nonempty subset of $\mathbb{Z}^d$. If $S \subset [N]^d$ contains no nontrivial homothetic copy $a + bX$ of $X$, then $|S| = o(N^d)$.*

Here we use the standard notation $[N] := \{1, \dots, N\}$. There has been great interest over the past few decades in proving a quantitative version of this theorem with reasonable bounds, i.e. with an upper bound for $|S|$ whose savings over the trivial bound of $N^d$ grows at least as quickly as a finite number of iterated logarithms. Indeed, Gowers has posed the problem of proving such a result on several occasions [Gow98a, Gow01a, Gow01b], and others, such as Graham [Gra97], have asked for bounds for sets lacking particular multidimensional configurations. While reasonable bounds are known in Szemerédi's theorem due to work of Gowers [Gow98b, Gow01b], none are known in the Furstenberg–Katznelson theorem in general. Furstenberg and Katznelson's original proof, which was via ergodic theory, produces no explicit bounds, while the hypergraph regularity proofs of Nagle, Rödl, Schacht, and Skokan [NRS06, RS04], Gowers [Gow07], and Tao [Tao06] each give a saving over the trivial bound of inverse Ackermann type.

Reasonable bounds in Theorem 1.1 are currently known for only one genuinely multidimensional configuration: two-dimensional corners

$$(x, y), \quad (x, y + z), \quad (x + z, y), \tag{1}$$

(and, thus, their linear images) due to the work of Shkredov [Shk06a, Shk06b], who proved that any subset of $[N] \times [N]$ containing no nontrivial corners has size at most $\ll N^2/(\log\log N)^c$ for some absolute constant $c > 0$. No reasonable bounds are known for any two-or-more-dimensional

Received 4 July 2022, accepted in final form 26 July 2023, published online 4 December 2023.
*2020 Mathematics Subject Classification* 11B30 (primary), 05D99 (secondary).
*Keywords:* multidimensional Szemerédi theorem, corners theorem, finite field model setting.

four-point configuration, such as three-dimensional corners,

$$(x, y, z), \quad (x, y, z + w), \quad (x, y + w, z), \quad (x + w, y, z), \tag{2}$$

or axis-aligned squares,

$$(x, y), \quad (x, y + z), \quad (x + z, y), \quad (x + z, y + z). \tag{3}$$

The latter of these two configurations is the topic of a conjecture of Graham [Gra97], which states that any subset $S \subset \mathbb{N} \times \mathbb{N}$ for which $\sum_{(x,y)\in S} 1/(x^2 + y^2)$ diverges must contain an axis-aligned square. Graham also conjectured, more generally, that if $\sum_{(x,y)\in S} 1/(x^2 + y^2)$ diverges, then $S$ must contain a homothetic copy of $[m] \times [m]$ for every positive integer $m$. This is a two-dimensional generalization of the famous and still unresolved conjecture of Erdős that every subset $T \subset \mathbb{N}$ for which $\sum_{n\in T} 1/n$ diverges must contain arbitrarily long arithmetic progressions.

Proving reasonable bounds for sets lacking the four-point configurations (2) and (3) seems to be out of reach. This is because no one has managed yet to prove anything useful about a certain two-dimensional directional uniformity norm that naturally appears in the study of these configurations. Details on this difficulty can be found in the work of Austin [Aus13a, Aus13b], where he demonstrates how enormously complicated and difficult even 100% and 99% inverse theorems can be for directional uniformity norms.

The purpose of this paper is to identify the first two-dimensional four-point configuration for which reasonable bounds in the multidimensional Szemerédi theorem can be proven, and to prove such bounds in the finite field model setting. We will study the configuration

$$(x, y), \quad (x, y + z), \quad (x, y + 2z), \quad (x + z, y), \tag{4}$$

which, when plotted on a two-dimensional integer grid, takes the shape of the capital letter 'L'. Because of this, we refer to (4) as an *L-shaped configuration*, and an L-shaped configuration with $z \neq 0$ as a *nontrivial* L-shaped configuration.

Theorem 1.2. *There exists a natural number $m$ and a constant $C > 0$ such that the following holds. Fix a prime $p \geq 11$, and set $N := p^n$. If $n \geq C$, then all $S \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ containing no nontrivial L-shaped configurations satisfy*

$$|S| \ll \frac{N^2}{\log_m N}.$$

The $m$ obtained in the theorem is huge, so we do not attempt to compute it. The bulk of the size of $m$ comes from our use of a recent quantitative inverse theorem for the $U^{10}$-norm on $\mathbb{F}_p^n$ due to Gowers and Milićević, who in [GM20] give a rough upper bound for the number of iterated exponentials appearing in their result. Based on this, $m$ is likely at least 24 trillion. The use of this inverse theorem is necessary in our proof, and no amount of care to argue efficiently in the rest of the argument can reduce $m$ by much. Thus, we have not tried to optimize the proof of Theorem 1.2, choosing instead to present the simplest argument that gives a reasonable upper bound.

It is likely that the proof of Theorem 1.2 can be adapted to the integer setting to prove a reasonable bound for subsets of $[N] \times [N]$ lacking L-shaped configurations, with the bound obtained being far more reasonable than the bound in Theorem 1.2. This is because the quantitative aspects of Manners's [Man18] inverse theorem for the $U^s$-norm on cyclic groups are better than those of Gowers and Milićević's inverse theorem when $s > 4$. It is also likely that



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Theorem 1.2 can be extended to more general L-shaped configurations with a longer vertical 'leg',

$$(x, y), \quad (x, y + z), \dots, (x, y + mz), \quad (x + z, y),$$

in both the finite field model and integer settings. We expect, however, that understanding L-shaped configurations with two longer 'legs',

$$(x, y), \quad (x, y + z), \dots, (x, y + mz), \quad (x + z, y), \dots, (x + \ell z, y),$$

is significantly more difficult, for some of the same reasons that proving reasonable bounds for sets lacking three-dimensional corners or axis-aligned squares seems out of reach.

While progress in proving a quantitative version of the multidimensional Szemerédi theorem has so far been extremely limited, there has been a bit more success in proving reasonable bounds for sets lacking multidimensional configurations with more degrees of freedom than those in Theorem 1.1. Prendiville [Pre15] has proven reasonable bounds for subsets of $[N]^d$ lacking any sufficiently nondegenerate three- or four-term matrix progression, and one consequence of his work is that any subset of $[N] \times [N]$ containing no four vertices of any square (not necessarily axis-aligned) has size at most $\ll N^2/(\log\log N)^{c'}$ for some absolute constant $c' > 0$.

The remainder of this paper is organized as follows. In § 2, we give a detailed outline of our proof of Theorem 1.2, including statements of the three main components of the density-increment argument: control of the count of L-shaped configurations by directional uniformity norms, obtaining a density increment on a structured set, and pseudorandomizing the structured set obtained previously. After introducing additional technical preliminaries in §§ 3 and 4, we prove these three main components in §§ 5, 6, and 7, respectively. We then carry out the density increment argument in § 8, proving Theorem 1.2.

## 2. Outline of the proof of Theorem 1.2

We begin this section by introducing the minimum amount of notation and preliminaries needed to understand our proof outline. We will use the standard asymptotic notation $O, \Omega$, and $o$, along with Vinogradov's notation $\ll, \gg$, and $\asymp$. For any two quantities $A$ and $B$, the relations $A = O(B)$, $B = \Omega(A)$, $A \ll B$, and $B \gg A$ all mean that $|A| \leq C|B|$ for some absolute constant $C > 0$. We will write $O(B)$ to represent a quantity that is $\ll B$ and $\Omega(A)$ to represent a quantity that is $\gg A$. When any of these asymptotic symbols appears with a subscript, the implied constant is allowed to depend on the parameters in the subscript. Since we fix a prime $p$ in Theorem 1.2, the implied constants appearing throughout the paper will sometimes depend on $p$ even though we will not alert the reader to this with a subscript. We will use $\log_m$ to denote the $m$-fold iterated logarithm, so that $\log_1 := \log$ and $\log_i := \log \circ \log_{i-1}$ for all $i > 1$, as well as $\exp^m$ to denote the $m$-fold iterated exponential, so that $\exp^1 = \exp$ and $\exp^i := \exp \circ \exp^{i-1}$ for all $i > 1$.

We will frequently denote the indicator function of a set $A$ by the letter $A$ itself, so that

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

For any pair of finite sets $X \subset Y$ with $Y \neq \emptyset$, we denote the density of $X$ in $Y$ by

$$\mu_Y(X) := \frac{|X|}{|Y|}.$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

For any function $f : X \to \mathbb{C}$, we denote the average of $f$ over $X$ by

$$\mathbb{E}_{x\in X} f(x) := \frac{1}{|X|}\sum_{x\in X} f(x).$$

When $X = \mathbb{F}_p^n$, we will usually drop '$\in X$' and just write $\mathbb{E}_x$ for $\mathbb{E}_{x\in\mathbb{F}_p^n}$. Whenever $f$ satisfies $|f(x)| \le 1$ for all $x$ in its domain, we say that it is 1-*bounded*. Note that the indicator function of any set is 1-bounded.

For any $f : \mathbb{F}_p^n \to \mathbb{C}$ and $\xi \in \mathbb{F}_p^n$, we define the Fourier coefficient of $f$ at $\xi$ using the normalization

$$\widehat{f}(\xi) := \mathbb{E}_x f(x) e_p(-\xi \cdot x),$$

where $e_p(z) := e^{2\pi i z/p}$ and $\cdot$ denotes the usual dot product in $\mathbb{F}_p^n$. With this choice of normalization, the Fourier inversion formula and Parseval's identity read

$$f(x) = \sum_{\xi\in\mathbb{F}_p^n} \widehat{f}(\xi) e_p(\xi \cdot x)$$

and

$$\mathbb{E}_x |f(x)|^2 = \sum_{\xi\in\mathbb{F}_p^n} |\widehat{f}(\xi)|^2,$$

respectively.

Let $H$ be any abelian group and $g : H \to \mathbb{C}$. For any $h \in H$, we define the function $\Delta_h g : H \to \mathbb{C}$ by

$$\Delta_h g(x) := g(x)\overline{g(x+h)},$$

and, for any $h_1, \dots, h_s \in H$, define the $s$-fold iterated differencing operator $\Delta_{h_1,\dots,h_s}$ by

$$\Delta_{h_1,\dots,h_s} g := \Delta_{h_1} \cdots \Delta_{h_s} g.$$

Note that $\Delta_{h_1,\dots,h_s} g = \Delta_{h_{\sigma(1)},\dots,h_{\sigma(s)}} g$ for any permutation $\sigma$ of $\{1,\dots,s\}$.

Now we can recall the definition of the Gowers uniformity norms.

DEFINITION 2.1. Let $s \in \mathbb{N}$, $H$ be an abelian group, and $f : H \to \mathbb{C}$. The $U^s$-norm of $f$ is defined by

$$\|f\|_{U^s(H)}^{2^s} := \mathbb{E}_{x,h_1,\dots,h_s\in H} \Delta_{h_1,\dots,h_s} f(x).$$

The basic properties of these norms can be found in [Tao12]. One such fact needed in the upcoming outline is the inverse theorem for the $U^2$-norm, which is a simple consequence of Fourier inversion and Parseval's identity.

LEMMA 2.2. *Let $H$ be an abelian group and $f : H \to \mathbb{C}$ be 1-bounded. If $\|f\|_{U^2(H)} \ge \delta$, then there exists a $\psi \in \widehat{H}$ such that*

$$|\mathbb{E}_{x\in H} f(x)\psi(x)| \ge \delta^2.$$

We will also sometimes need the notion of the $U^2$-norm on an affine subspace $w + V$ of $\mathbb{F}_p^n$, which is defined by $\|f\|_{U^2(w+V)} := \|f(\cdot - w)\|_{U^2(V)}$. The corresponding inverse theorem for these norms follows from Lemma 2.2.

### 2.1 A review of Shkredov's argument in the finite field model setting

Before we outline the proof of Theorem 1.2, it will be instructive to review Shkredov's argument for corners (1) in the finite field model setting. A detailed account of the argument can be found in the expositions of Green [Gre05a, Gre05b].



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Shkredov's proof proceeds via a density-increment argument. As in all analytic approaches to Szemerédi's theorem and its generalizations, we begin by defining a multilinear average over the configuration of interest. For $g_0, g_1, g_2 : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{C}$, set

$$\Lambda_{\llcorner}(g_0, g_1, g_2) := \mathbb{E}_{x,y,z} g_0(x,y) g_1(x, y+z) g_2(x+z, y).$$

Then, for any $S \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$, the quantity $\Lambda_{\llcorner}(S,S,S)$ equals the normalized count,

$$\frac{\#\{x,y,z \in \mathbb{F}_p^n : (x,y), (x,y+z), (x+z,y) \in S\}}{p^3},$$

of the number of corners in $S$. Setting $N := p^n = |\mathbb{F}_p^n|$, we let $\sigma := |S|/N^2$ denote the density of $S$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ and $g_S := S - \sigma$ denote the balanced function of $S$. It follows from the trilinearity of $\Lambda_{\llcorner}$ that

$$\Lambda_{\llcorner}(S,S,S) = \sigma \Lambda_{\llcorner}(1,S,S) + \Lambda_{\llcorner}(g_S, S, S).$$

Since $\Lambda_{\llcorner}(1,S,S) \geq \sigma^2$ by the Cauchy–Schwarz inequality, if the normalized count of corners in $S$ is far below the $\sim \sigma^3$ expected for a random set of density $\sigma$, which is the case when $S$ has no nontrivial corners and $N$ is sufficiently large in terms of $\sigma$, then $|\Lambda_{\llcorner}(g_S, S, S)|$ must be large.

It can then be shown, by an appropriate sequence of applications of the Cauchy–Schwarz inequality, that $g_S$ must have large *box norm*

$$\|g_S\|_{\square} := \big(\mathbb{E}_{x,y,x',y'} g_S(x,y) \overline{g_S(x,y') g_S(x',y)} g_S(x',y')\big)^{1/4}.$$

If $\|g_S\|_{\square}$ is large, it follows by an averaging argument that $S$ has density at least $\sigma + \Omega(\sigma^{O(1)})$ on a product set $A \times B$ for some large $A, B \subset \mathbb{F}_p^n$.

One may then hope to continue the density-increment argument by proving the following generalization of the result just sketched: if $S$ is a subset of density $\sigma$ of a product set $T = A \times B$ and contains no nontrivial corners, then $S$ has density at least $\sigma + \Omega(\sigma^{O(1)})$ on a product set $T'$ contained in $T$.

It turns out, however, that the Cauchy–Schwarz argument mentioned previously yields a lower bound on the box norm of large enough size only when $A$ and $B$ are sufficiently *Fourier pseudorandom*, meaning that their balanced functions $A - |A|/N$ and $B - |B|/N$ both have small $U^2$-norm. The components of the product set just obtained are essentially arbitrary aside from being large. They are, in particular, not guaranteed to be Fourier pseudorandom.

To overcome this difficulty, Shkredov introduced a *pseudorandomizing* step into his proof. He used an energy increment argument incorporating the $U^2$-inverse theorem to partition $\mathbb{F}_p^n \times \mathbb{F}_p^n$ into products of large affine subspaces of the form

$$(u + V) \times (w + V), \tag{5}$$

for most of which the sets $(A - u) \cap V$ and $(B - w) \cap V$ are Fourier pseudorandom in $V$. By an averaging argument, there must exist such a product of affine subspaces (5) on which the restrictions of $A$ and $B$ are both sufficiently dense and Fourier pseudorandom, and such that $S$ still has increased density $\sigma + \Omega(\sigma^{O(1)})$ on the intersection of $T$ with $(u+V) \times (w+V)$.

By passing to this product of cosets and using that corners are preserved by translation and invertible linear transformations of the form $(x,y) \mapsto (Ex, Ey)$, one can then continue the density-increment argument with $\mathbb{F}_p^n \times \mathbb{F}_p^n$ replaced by $\mathbb{F}_p^{n'} \times \mathbb{F}_p^{n'}$, where $n' = \dim V$. If $S \subset T$ contains no nontrivial corners and $A$ and $B$ are sufficiently Fourier pseudorandom, then $g_S$ must have large box norm localized to $T$. One must then prove that $S$ has a further density increment on a product set contained in $T$, which is, fortunately, of exactly the same difficulty whether $T = \mathbb{F}_p^n \times \mathbb{F}_p^n$ or some other large product set. By applying the pseudorandomizing procedure



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

to the factors of the product set just produced, one can then deduce that if $S$ is a subset of density $\sigma$ of a product set $T = A \times B$, where $A$ and $B$ are large and sufficiently Fourier pseudorandom, and $S$ contains no nontrivial corners, then $S$ has density at least $\sigma + \Omega(\sigma^{O(1)})$ on a product set $T' = A' \times B'$ contained in $T$, where $A'$ and $B'$ are also large and sufficiently Fourier pseudorandom. The density increment iteration can be carried out repeatedly to produce a good bound for subsets of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ lacking corners.

### 2.2 An outline of our argument

The obstructions to uniformity for L-shaped configurations are not just (skew) product sets, as was the case for corners, but also very general sets of the form

$$\{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : y \in u_x + V_x\},$$

where each $u_x + V_x$ is an affine subspace of $\mathbb{F}_p^n$. For example, assume that $n \geq 3$, and consider the set

$$\{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : x \cdot y = 0\}.$$

This set has density

$$\frac{(N-1)N/p + N}{N^2} \sim \frac{1}{p}$$

in $\mathbb{F}_p^n \times \mathbb{F}_p^n$, but

$$[(N-1) - (p-1)]\left(\frac{N}{p} - 1\right)\frac{N}{p^2} + \left(\frac{N}{p} - 1\right)(p-1)\frac{N}{p} + N + 2(N-1)\frac{N}{p} \sim \frac{N^3}{p^3}$$

L-shaped configurations, in contrast to the $\sim N^3/p^4$ expected in a random subset of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ of density $1/p$. Similarly, the number of L-shaped configurations in the sets

$$\{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : \phi(x) \cdot y = 0\}$$

and

$$\{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : y_1 = u(x)\},$$

where $\phi(x) \in \mathbb{F}_p^n$ and $u(x) \in \mathbb{F}_p$ are now chosen uniformly at random, is also $\sim N^3/p^3$ with high probability, while the sets have density $\sim 1/p$ with high probability. These new sorts of obstructions to uniformity are the main reason why the study of L-shaped configurations is significantly more difficult than that of corners, and must be taken into account to prove Theorem 1.2.

For any functions $g_0, g_1, g_2, g_3 : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{C}$, we define

$$\Lambda(g_0, g_1, g_2, g_3) := \mathbb{E}_{x,y,z} g_0(x, y) g_1(x, y + z) g_2(x, y + 2z) g_3(x + z, y), \tag{6}$$

so that $\Lambda(S, S, S, S)$ equals the normalized count of L-shaped configurations in any subset $S$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$. The multilinearity of $\Lambda$ implies that

$$|\Lambda(S, S, S, S) - \sigma^4| \leq \sigma^2 |\Lambda(1, 1, g_S, S)| + \sigma |\Lambda(1, g_S, S, S)| + |\Lambda(g_S, S, S, S)|, \tag{7}$$

where, as before, $g_S = S - \sigma$ is the balanced function of $S$. Thus, if the normalized count of L-shaped configurations in $S$ is far from the random normalized count $\sigma^4$, one of $|\Lambda(1, 1, g_S, S)|$, $|\Lambda(1, g_S, S, S)|$, or $|\Lambda(g_S, S, S, S)|$ must be large. In particular, when $S$ contains no nontrivial L-shaped configurations and $N$ is sufficiently large in terms of $\sigma$, one of these quantities will be larger than $\sigma^4/2$. It then follows from several applications of the Cauchy–Schwarz inequality



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

that one of the following directional uniformity norms of $g_S$ must be larger than $\sigma^4/2$:

$$\|g\|_{\star_1} := (\mathbb{E}_{x,y,h_1,h_2,h_3}\Delta_{(0,h_1),(0,h_2),(h_3,0)}g(x,y))^{1/8}, \tag{8}$$

$$\|g\|_{\star_2} := (\mathbb{E}_{x,y,h_1,h_2}\Delta_{(0,h_1),(-h_2,h_2)}g(x,y))^{1/4}, \tag{9}$$

or

$$\|g\|_{\star_3} := (\mathbb{E}_{x,y,h_1}\Delta_{(-h_1,2h_1)}g(x,y))^{1/2}. \tag{10}$$

Here $\|\cdot\|_{\star_3}$ is only a semi-norm, while $\|\cdot\|_{\star_1}$ and $\|\cdot\|_{\star_2}$ are genuine norms. Since these are all Gowers box norms, one can find a proof that they are (semi-)norms in Appendix B of [GT10]. The norm $\|\cdot\|_{\star_1}$ had been studied previously, in the setting of cyclic groups, in work of Shkredov [Shk09].

Directional uniformity norms with two differencing parameters,

$$[\mathbb{E}_{x,y,h,k\in H}\Delta_{hv_1,kv_2}g(x,y)]^{1/4},$$

for fixed nonzero $v_1, v_2 \in H \times H$, are well-understood. Either $v_1$ and $v_2$ are scalar multiples of each other, in which case the norm is just the $U^2$-norm on $\langle v_1\rangle$ averaged over cosets of $\langle v_1\rangle$, or they are linearly independent, as in the definition of $\|\cdot\|_{\star_2}$, in which case the norm is, after a change of variables, equivalent to the two-dimensional box norm. Directional uniformity norms with three differencing parameters,

$$[\mathbb{E}_{x,y,h_1,h_2,h_3\in H}\Delta_{h_1v_1,h_2v_2,h_3v_3}g(x,y)]^{1/8},$$

for fixed nonzero $v_1, v_2, v_3 \in H \times H$ analogously fall into one of three cases: either $v_1, v_2$, and $v_3$ are collinear, lie on exactly two lines, or are in general position. In the first case, the norm is just the $U^3$-norm on $\langle v_1\rangle$ averaged over cosets of $\langle v_1\rangle$. In the third case, the norm is linearly equivalent to the intractable norm that arises in the study of three-dimensional corners and axis-aligned squares. The norm $\|\cdot\|_{\star_1}$ we encounter falls into the second case, and the study and fruitful use of this norm turns out to be possible (though still complicated) due to its structure as a '$U^1 \times U^2$-norm'.

The upshot is that if $S$ contains no nontrivial L-shaped configurations, then it must have density at least $\sigma + \Omega(\sigma^{O(1)})$ on a set of the form

$$T := \{(x,y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : B(y)C(x+y)D(2x+y)\Phi(x,y) = 1\}, \tag{11}$$

where $\Phi \subset A \times \mathbb{F}_p^n$ is of the form

$$\Phi := \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\}, \tag{12}$$

for some element $u \in \mathbb{F}_p^n$ and collection of subspaces $\{V_x : x \in A\}$ of $\mathbb{F}_p^n$, where $A, B, C, D \subset \mathbb{F}_p^n$ are large and $\operatorname{codim} V_x$ is small for each $x \in A$. Note that this set $\Phi$ is not quite as general as the one appearing at the very beginning of this subsection, as the element $u$ of $\mathbb{F}_p^n$ does not vary with $x$. It takes some extra work to show that we can guarantee $\Phi$ to be of this special form, which turns out to be necessary for our density-increment iteration. We say more about this point in § 6.

We would like to continue the density-increment iteration and show that $S' := T \cap S$, which also lacks L-shaped configurations, has a further density increment of at least the same size as the first on a subset $T'$ of $T$ of the same general form (11). Analogously to Shkredov's argument for corners, we can only hope to do this if $A, B, C, D$, and $\Phi$ are sufficiently pseudorandom, for some appropriate notions of pseudorandomness. We will need to control the count of L-shaped configurations by the norms $\|\cdot\|_{\star_1}$, $\|\cdot\|_{\star_2}$, and $\|\cdot\|_{\star_3}$ defined in (8), (9), and (10) with no loss



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

of density factors, i.e. show that

$$\frac{|\Lambda(f_0, f_1, f_2, f_3)|}{\Lambda(T, T, T, T)} \geq \delta \implies \frac{\|f_0\|_{\star_1}}{\|T\|_{\star_1}} \gg_\delta 1, \tag{13}$$

$$\frac{|\Lambda(T, f_1, f_2, f_3)|}{\Lambda(T, T, T, T)} \geq \delta \implies \frac{\|f_1\|_{\star_2}}{\|T\|_{\star_2}} \gg_\delta 1, \tag{14}$$

and

$$\frac{|\Lambda(T, T, f_2, f_3)|}{\Lambda(T, T, T, T)} \geq \delta \implies \frac{\|f_2\|_{\star_3}}{\|T\|_{\star_3}} \gg_\delta 1, \tag{15}$$

and also obtain a density increment with no loss of density factors when some localized norm $\|\cdot\|_{\star_i}$ of the balanced function of a set is large, i.e. show that if

$$\frac{\|g_S\|_{\star_1}}{\|T\|_{\star_1}}, \quad \frac{\|g_S\|_{\star_2}}{\|T\|_{\star_2}}, \quad \text{or} \quad \frac{\|g_S\|_{\star_3}}{\|T\|_{\star_3}} \geq \delta,$$

where now $g_S := S - \sigma T$, then there exists a subset $T' \subset T$ of the same general form,

$$T' := \{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : B'(y)C'(x + y)D'(2x + y)\Phi'(x, y) = 1\},$$

as $T$ on which $S$ has a density increment

$$\mathbb{E}_{(x,y)\in T'} S(x, y) \geq \sigma + \Omega_\delta(1)$$

depending only on $\delta$. Such results are needed so that the density increment obtained at each step of the iteration is independent of the step. If one is not sufficiently careful, it is easy to end up with a density increment that gets smaller as the subset $T$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ gets sparser, which is not enough to close the density increment iteration.

To carry out these arguments, we will need $A, B, C$, and $D$ to be pseudorandom with respect to the $U^{10}(\mathbb{F}_p^n)$-norm. The situation for $\Phi$ is more complicated, and deciding on a good measure of pseudorandomness for $\Phi$ that is amenable to a Shkredov-like pseudorandomization procedure and can also be used to analyze the various averages appearing throughout our argument is one of the challenges of the proof of Theorem 1.2. A suitable condition on $\Phi$ turns out to be that it is pseudorandom with respect to the $U^8(\mathbb{F}_p^n \times \mathbb{F}_p^n)$-norm. This condition is not, on its own, immediately useful in the arguments of §§ 5 and 6, since the various averages that appear are not controllable by the $U^8(\mathbb{F}_p^n \times \mathbb{F}_p^n)$-norm of $\Phi$. It takes a bit of work to show that it implies a roughly equivalent statement about the typical codimensions of certain affine subspaces obtained from $\Phi$. We prove this in § 4, deriving some new results on the combinatorics of approximate polynomials along the way.

The proof of the implications (13), (14), and (15) when $A, B, C, D$, and $\Phi$ are sufficiently pseudorandom consists of many careful applications of the Cauchy–Schwarz inequality, along with appeals to standard facts about the number of linear configurations of bounded Cauchy–Schwarz complexity in products of pseudorandom sets intersected with subspaces of bounded codimension. We carry out this argument in § 5.

Obtaining a large enough density increment when $\|g_S\|_{\star_1}$ is large for $S \subset T$ requires some new ideas and a significant amount of extra work beyond the proof of the non-localized case, in contrast to the situation for the box norm localized to product sets, where the argument is the same as the non-localized case. In order to get such a density increment that only depends on $\delta$ and not on the densities of $A, B, C, D$, or $\Phi$, one of the key ingredients is a density-preserving inverse theorem for the $U^2(\Phi(x, \cdot))$-norms on pseudorandom sets derived from $A, B, C$, and $D$, which we prove using a version of the transference principle. We carry out this argument in § 6.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

As was the case for corners, the sets $A', B', C', D'$, and $\Phi'$ obtained in the previous paragraph are not guaranteed to be pseudorandom. We must also carry out a pseudorandomizing procedure to locate a product of large affine subspaces of the form $(u+V)\times(w+V)$ on which $A', B', C', D'$, and $\Phi'$ are sufficiently pseudorandom and $S$ still has a large density increment on $T'\cap[(u+V)\times(w+V)]$. Our pseudorandomization procedure is similar to Shkredov's, but with some new complications coming from our desire for $A', B', C'$, and $D'$ and $\Phi'$ to be pseudorandom with respect to the $U^{10}(\mathbb{F}_p^n)$- and $U^8(\mathbb{F}_p^n\times\mathbb{F}_p^n)$-norms, respectively, and from $\Phi'$'s particular structure as a union of affine subspaces in the second factor of $\mathbb{F}_p^n\times\mathbb{F}_p^n$. To handle the first complication, we use a recent quantitative inverse theorem of Gowers and Milićević [GM20] for the $U^s$-norms on vector spaces over finite fields, combined with a result of Cohen and Tal [CT15] that allows us to partition $\mathbb{F}_p^n$ into large affine subspaces on which any finite collection of bounded degree polynomials are all constant. The structure of $\Phi'$ has the potential to cause issues in a Shkredov-like pseudorandomization argument, since the intersection of $\Phi'$ with a cell may no longer be the union of affine subspaces all having the same dimension. We will explain how this complication is dealt with in § 7, since it requires a bit of set up.

### 2.3 Key intermediate results

We finish this section by stating the key intermediate results needed to prove Theorem 1.2 that we just described in the outline. Recall that $g_S = S-\sigma$ denotes the balanced function of $S$.

Lemma 2.3 (Estimation of $\Lambda(T,T,T,S)$). *There exist absolute constants $0<c_1<1<c_2$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A,B,C,D\subset\mathbb{F}_p^n$ have densities $\alpha,\beta,\gamma,\delta$, respectively, and that $\Phi\subset\mathbb{F}_p^n\times\mathbb{F}_p^n$ takes the form*

$$\Phi=\{(x,y)\in A\times\mathbb{F}_p^n : y\in u+V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Define $T\subset\mathbb{F}_p^n\times\mathbb{F}_p^n$ by (11) and suppose that $S\subset T$ has density $\sigma$ in $T$. Let $\varepsilon\le c_1(\sigma\alpha\beta\gamma\delta\rho)^{c_2}$ and assume that*

$$\|A-\alpha\|_{U^5(\mathbb{F}_p^n)},\quad \|B-\beta\|_{U^5(\mathbb{F}_p^n)},\quad \|C-\gamma\|_{U^5(\mathbb{F}_p^n)},\quad \|D-\delta\|_{U^5(\mathbb{F}_p^n)},\quad \|\Phi-\alpha\rho\|_{U^2(\mathbb{F}_p^n\times\mathbb{F}_p^n)}<\varepsilon.$$

*Then*

$$\Lambda(T,T,T,S)\gg\sigma\alpha^2\beta^3\gamma^3\delta^3\rho^3.$$

As a consequence, we get that if $\varepsilon$ is small enough, $n$ is large enough, and $S\subset T$ has no nontrivial L-shaped configurations, then

$$\max(|\Lambda(g_S,S,S,S)|,|\Lambda(T,g_S,S,S)|,|\Lambda(T,T,g_S,S)|)\gg\sigma^4\alpha^2\beta^3\gamma^3\delta^3\rho^3.$$

Lemma 2.4 (Control by $\|\cdot\|_{\star_i}$ norms). *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A,B,C,D\subset\mathbb{F}_p^n$ have densities $\alpha,\beta,\gamma,\delta$, respectively, and that $\Phi\subset\mathbb{F}_p^n\times\mathbb{F}_p^n$ takes the form*

$$\Phi=\{(x,y)\in A\times\mathbb{F}_p^n : y\in u+V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Assume that*

$$\|A-\alpha\|_{U^8(\mathbb{F}_p^n)},\quad \|B-\beta\|_{U^8(\mathbb{F}_p^n)},\quad \|C-\gamma\|_{U^8(\mathbb{F}_p^n)},\quad \|D-\delta\|_{U^8(\mathbb{F}_p^n)},\quad \|\Phi-\alpha\rho\|_{U^6(\mathbb{F}_p^n\times\mathbb{F}_p^n)}<\varepsilon.$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*Define $T \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ by (11) and suppose that $f_0, f_1, f_2, f_3 : \mathbb{F}_p^n \times \mathbb{F}_p^n \to \mathbb{C}$ are 1-bounded functions supported on $T$. Then*

$$|\Lambda(f_0, f_1, f_2, f_3)|^8 \le \alpha^{14}\beta^{20}\gamma^{16}\delta^{16}\rho^{18}\|f_0\|_{\star_1}^8 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right), \tag{16}$$

$$|\Lambda(T, f_1, f_2, f_3)|^4 \le \alpha^{8}\beta^{8}\gamma^{10}\delta^{8}\rho^{8}\|f_1\|_{\star_2}^4 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right), \tag{17}$$

*and*

$$|\Lambda(T, T, f_2, f_3)|^2 \le \alpha\beta^{3}\gamma^{3}\delta^{4}\rho^{3}\|f_2\|_{\star_3}^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right). \tag{18}$$

Thus, if $\varepsilon$ is small enough, $n$ is large enough, and $S \subset T$ has no nontrivial L-shaped configurations, then one of

$$\frac{\|g_S\|_{\star_1}}{\|T\|_{\star_1}}, \quad \frac{\|g_S\|_{\star_2}}{\|T\|_{\star_2}}, \quad \text{or} \quad \frac{\|g_S\|_{\star_3}}{\|T\|_{\star_3}}$$

is $\gg \sigma^{O(1)}$.

THEOREM 2.5 ($\|g_S\|_{\star_1}$, $\|g_S\|_{\star_2}$, or $\|g_S\|_{\star_3}$ large implies a density increment). *There exist absolute constants $0 < c_1 < 1 < c_2, c_3$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\sigma, \tau > 0$ and*

$$\varepsilon \le c_1(\sigma\tau\alpha\beta\gamma\delta\rho)^{c_2}\exp(-(64/\tau^8)^{c_3}),$$

*and assume that*

$$\|A - \alpha\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|B - \beta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|\Phi - \alpha\rho\|_{U^{8}(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < \varepsilon.$$

*Define $T \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ by (11) and assume that $S \subset T$ has density $\sigma$ in $T$. Suppose that*

$$\|g_S\|_{\star_1} \ge \tau\alpha^{1/4}\beta^{1/2}\gamma\delta\rho^{3/4},$$
$$\|g_S\|_{\star_2} \ge \tau\alpha^{1/2}\beta\gamma^{1/2}\delta\rho,$$

*or*

$$\|g_S\|_{\star_3} \ge \tau\alpha\beta\gamma\delta^{1/2}\rho.$$

*Then, $S$ has density at least $\sigma + \Omega(\tau^{O(1)})$ on a subset $T' \subset T$ of the form*

$$T' := \{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : B'(y)C'(x + y)D'(2x + y)\Phi'(x, y) = 1\},$$

*where the densities of $A', B', C', D' \subset \mathbb{F}_p^n$ are all $\gg (\sigma\tau\alpha\beta\gamma\delta\rho)^{O(1)}$, and the set $\Phi' \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi' = \{(x, y) \in A' \times \mathbb{F}_p^n : y \in u' + V_x'\},$$

*where each $V_x'$ is a subspace of $\mathbb{F}_p^n$ of codimension $d + 1$.*

The first three lemmas combined tell us that if $S$ has density $\sigma$ and contains no nontrivial L-shaped configurations, then one can find a subset $T$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ of the form (11) on which $S$ has density at least $\sigma + \Omega(\sigma^{O(1)})$. The next lemma tells us that, after restricting to a product of



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

large affine subspaces, we can get this same conclusion with $A, B, C, D$, and $\Phi$ as pseudorandom as we need, which will allow us to continue the density-increment iteration.

Lemma 2.6. *There exist absolute constants $0 < c_1 < 1 < c_2, c, c'$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Let $\varepsilon' > 0$, and suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ and takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Define $T \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ by (11), and assume that $S \subset T$ has density $\sigma + \tau$ in $T$, as well as that*

$$n \geq \exp^2\left(c_2 \frac{\exp^c(c'/\varepsilon')}{d\tau\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)}\right).$$

*Then there exists a subspace $V \leq \mathbb{F}_p^n$ of dimension*

$$\dim V \gg n^{c_1^{O(\exp^c(c'/\varepsilon')/d\tau\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T))}},$$

*$u, w \in \mathbb{F}_p^n$, and $0 \leq i \leq d$ such that, on setting $\mathcal{C} = (u + V) \times (w + V)$:*

- $B' := B \cap (w + V)$;
- $C' := C \cap (u + w + V)$;
- $D' := D \cap (2u + w + V)$;
- $\Psi' := \Phi \cap \mathcal{C}$ *and* $\Phi' := \{(x, y) \in \Psi' : \mathbb{E}_{z \in w+V} \Psi'(x, z) = p^{-i}\}$;
- $A' := \{x \in u + V : \mathbb{E}_{z \in w+V} \Phi'(x, z) \neq 0\}$;
- $\alpha' = \mu_{u+V}(A')$;
- $\beta' := \mu_{w+V}(B')$;
- $\gamma' := \mu_{u+w+V}(C')$;
- $\delta' := \mu_{2u+w+V}(D')$;
- $\rho' := p^{-i}$; *and*
- $T' := \{(x, y) \in \mathcal{C} : B'(y)C'(x + y)D'(2x + y)\Phi'(x, y) = 1\}$;

*we have*

$$\|A' - \rho'\|_{U^{10}(u+V)}, \quad \|B' - \beta'\|_{U^{10}(w+V)}, \quad \|C' - \gamma'\|_{U^{10}(u+w+V)}, \quad \|D' - \delta'\|_{U^{10}(2u+w+V)},$$
$$\|\Phi' - \alpha'\rho'\|_{U^8(\mathcal{C})} < 2\varepsilon,$$

*$\alpha', \beta', \gamma', \delta' \gg \tau\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/4$, and*

$$\mu_{\mathcal{C}}(S \cap T') \geq \left(\sigma + \frac{\tau}{4}\right)\mu_{\mathcal{C}}(T').$$

By combining the previous four lemmas and using that L-shaped configurations are preserved by translation and invertible linear transformations of the form $(x, y) \mapsto (Ex, Ey)$, we thus deduce the following density-increment lemma, which we will iterate in § 8 to prove Theorem 1.2.

Lemma 2.7. *There exist absolute constants $0 < c_1 < 1 < c_2, c_3, c_4, c, c'$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Define $T$ by (11) and let $S \subset T$ have density $\sigma$ in $T$. Let $\varepsilon \leq (\sigma\alpha\beta\gamma\delta\rho)^{c_2} \exp(-(64/\sigma)^{c_3})$, and assume that*

$$\|A-\alpha\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|B-\beta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|C-\gamma\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|D-\delta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|\Phi-\alpha\rho\|_{U^8(\mathbb{F}_p^n\times\mathbb{F}_p^n)} < \varepsilon.$$

*Let $\varepsilon' > 0$, and suppose that $S$ has no nontrivial L-shaped configurations. Then either:*

(i) $n < \exp^2(c_4 \exp^c(c'/\varepsilon')/(\sigma\alpha\beta\gamma\delta\rho)^{c_2})$; *or*

(ii) *there exists natural numbers $n'$ and $d'$ satisfying*

$$n' \gg n^{c_1^{\exp^c(c'/\varepsilon')/(\sigma\alpha\beta\gamma\delta\rho)^{c_2}}}$$

*and $0 \leq d' \leq d+1$, subsets $A', B', C', D' \subset \mathbb{F}_p^{n'}$ of densities $\alpha', \beta', \gamma', \delta'$, respectively, a subset $\Phi' \subset \mathbb{F}_p^{n'} \times \mathbb{F}_p^{n'}$ of the form*

$$\Phi' = \{(x,y) \in A' \times \mathbb{F}_p^{n'} : y \in u' + V_x'\},$$

*where each $V_x'$ is a subspace of $\mathbb{F}_p^{n'}$ of codimension $d'$ (so that $\Phi'$ has density $\alpha'\rho'$, where $\rho' := p^{-d'}$), and a subset $S' \subset T'$, where*

$$T' := \{(x,y) \in \mathbb{F}_p^{n'} \times \mathbb{F}_p^{n'} : B'(y)C'(x+y)D'(2x+y)\Phi'(x,y) = 1\},$$

*of density at least $\sigma + \Omega(\sigma^{O(1)})$ in $T'$, such that*

$$\|A'-\alpha'\|_{U^{10}(\mathbb{F}_p^{n'})}, \quad \|B'-\beta'\|_{U^{10}(\mathbb{F}_p^{n'})}, \quad \|C'-\gamma'\|_{U^{10}(\mathbb{F}_p^{n'})},$$
$$\|D'-\delta'\|_{U^{10}(\mathbb{F}_p^{n'})}, \quad \|\Phi'-\alpha'\rho'\|_{U^8(\mathbb{F}_p^{n'}\times\mathbb{F}_p^{n'})} < \varepsilon',$$

*$\alpha', \beta', \gamma', \delta' \geq (\sigma\alpha\beta\gamma\delta\rho)^{c_2}$, and $S'$ contains no nontrivial L-shaped configurations.*

## 3. Additional preliminaries

In this section, we present some more preliminaries that were not needed for the outline of the proof of Theorem 1.2, but will be convenient to have for the proof itself. We begin with the notion of *Cauchy–Schwarz complexity*, first defined by Green and Tao in [GT10].

DEFINITION 3.1 (Cauchy–Schwarz complexity). Let $\psi_1, \dots, \psi_d : (\mathbb{F}_p^n)^r \to \mathbb{F}_p^n$ be a collection of linear forms in $r$ variables. We say that $\psi_1, \dots, \psi_d$ has Cauchy–Schwarz complexity at most $s$ if, for every $j \in [d]$, there exists a partition of $\{\psi_1, \dots, \psi_d\} \setminus \{\psi_j\}$ into at most $s+1$ subsets such that $\psi_j$ is not contained in the linear span of any of the subsets.

The smallest $s$ such that $\{\psi_1, \dots, \psi_d\}$ has Cauchy–Schwarz complexity at most $s$ is called the Cauchy–Schwarz complexity of $\{\psi_1, \dots, \psi_d\}$.

For example, four term arithmetic progressions,

$$x, x+y, x+2y, x+3y,$$

have Cauchy–Schwarz complexity 2.

Any system of linear forms of complexity at most $s$ can be shown to be controlled by the $U^{s+1}$-norm using repeated applications of the Cauchy–Schwarz inequality. In particular, carrying out the proof of the generalized von Neumann theorem of Green and Tao in [GT10] in the finite field model setting (where the technical details are much simpler) produces the following useful result.

THEOREM 3.2. *Let $\psi_1, \dots, \psi_d : (\mathbb{F}_p^n)^r \to \mathbb{F}_p^n$ be a collection of linear forms in $r$ variables with Cauchy–Schwarz complexity at most $s$. For any 1-bounded functions $f_1, \dots, f_d : \mathbb{F}_p^n \to \mathbb{C}$,*



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*we have*

$$\left|\mathbb{E}_{x_1,\dots,x_r}\prod_{j=1}^{d} f_j(\psi_j(x_1,\dots,x_r))\right| \le \min_{1\le j\le d}\|f_j\|_{U^{s+1}(\mathbb{F}_p^n)}.$$

*Further, if all of $f_1,\dots,f_d$ are supported on a set $A\subset\mathbb{F}_p^n$ of density $\alpha$ that satisfies $\|A-\alpha\|_{U^{s+1}(\mathbb{F}_p^n)}<\alpha\varepsilon$, then*

$$\left|\mathbb{E}_{x_1,\dots,x_r}\prod_{j=1}^{d} f_j(\psi_j(x_1,\dots,x_r))\right| \le \alpha^{d-1}\min_{1\le j\le d}\|f_j\|_{U^{s+1}(\mathbb{F}_p^n)} + O_{d,s}(\varepsilon^{\Omega_{d,s}(1)}).$$

We will use the following immediate corollary of Theorem 3.2 numerous times throughout the proof of Theorem 1.2.

Corollary 3.3. *Let $\psi_1,\dots,\psi_d:(\mathbb{F}_p^n)^r\to\mathbb{F}_p^n$ be a collection of linear forms in $r$ variables with Cauchy–Schwarz complexity at most $s$. Suppose that $f_1,\dots,f_d:\mathbb{F}_p^n\to\mathbb{C}$ are 1-bounded functions having average values $\alpha_1,\dots,\alpha_d$, respectively, and that*

$$\|f_j-\alpha_j\|_{U^{s+1}(\mathbb{F}_p^n)}\le\varepsilon_j$$

*for all $1\le j\le d$. Then*

$$\left|\mathbb{E}_{x_1,\dots,x_r}\prod_{j=1}^{d} f_j(\psi_j(x_1,\dots,x_r)) - \prod_{j=1}^{d}\alpha_j\right| \le d\max_{1\le j\le d}\varepsilon_j.$$

We also need the Gowers–Cauchy–Schwarz inequality.

Lemma 3.4 (Gowers–Cauchy–Schwarz inequality). *Let $s$ be a natural number, $H$ be an abelian group, and $f_\omega:H\to\mathbb{C}$ for every $\omega\in\{0,1\}^s$. We have*

$$\left|\mathbb{E}_{x,h_1,\dots,h_s\in H}\prod_{\omega\in\{0,1\}^s} f_\omega(x+\omega\cdot(h_1,\dots,h_s))\right| \le \prod_{\omega\in\{0,1\}^s}\|f_\omega\|_{U^s(H)}.$$

Next, we record the basic fact that a function on $\mathbb{F}_p^n$ with small $U^2$-norm has small average on affine subspaces of small codimension.

Lemma 3.5. *Let $f:\mathbb{F}_p^n\to\mathbb{C}$ be a 1-bounded function satisfying $\|f\|_{U^2(\mathbb{F}_p^n)}<\varepsilon$ and $w+V\subset\mathbb{F}_p^n$ be an affine subspace of codimension $d$. Then*

$$|\mathbb{E}_{x\in w+V}f(x)|<p^d\varepsilon.$$

*Proof.* The indicator function of $V$ can be written as

$$\frac{1}{p^d}\sum_{\xi\in V^\perp} e_p(\xi\cdot x),$$

so we have that

$$|\mathbb{E}_x f(x-w)V(x)| \le \frac{1}{p^d}\sum_{\xi\in V^\perp}|\mathbb{E}_x f(x)e_p(\xi\cdot x)| \le \|f\|_{U^2(\mathbb{F}_p^n)},$$

by the Gowers–Cauchy–Schwarz inequality. Since $|V|=|\mathbb{F}_p^n|/p^d$, the desired result follows. □

The last lemma of this section will be used to analyze the various averages appearing in the proofs of Lemmas 2.3, 2.4, and 2.5.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

LEMMA 3.6. *Let $\psi_1, \dots, \psi_d \in \mathbb{F}_p[x_1, \dots, x_t, y]$ be a collection of linear forms such that the coefficient of $y$ in each of $\psi_1, \dots, \psi_d$ is nonzero, and $F : (\mathbb{F}_p^n)^t \times \mathbb{F}_p^n \to [0,1]$ be a function of the form*

$$F(\mathbf{x}, y) = \prod_{j=1}^{d} f_j(\psi_j(\mathbf{x}, y))$$

*for some 1-bounded functions $f_j : \mathbb{F}_p^n \to \mathbb{C}$ with average value $\beta_j$, each satisfying $\|f_j - \beta_j\|_{U^s(\mathbb{F}_p^n)} < \varepsilon$. If the Cauchy–Schwarz complexity of the set of linear forms*

$$\bigcup_{j=1}^{d} \{\psi_j(\mathbf{x}, y), \psi_j(\mathbf{x}, y+h), \psi_j(\mathbf{x}, y+k), \psi_j(\mathbf{x}, y+h+k)\} \tag{19}$$

*in the variables $x_1, \dots, x_t, y, h, k$, is at most $s-1$, then*

$$\mathbb{P}\left(\mathbf{x} \in (\mathbb{F}_p^n)^t : \left\| F(\mathbf{x}, \cdot) - \prod_{j=1}^{d} \beta_j \right\|_{U^2(\mathbb{F}_p^n)} \geq \varepsilon^{1/8}\right) \ll_d \sqrt{\varepsilon}.$$

*Proof.* Set $\beta := \prod_{j=1}^{d} \beta_j$. Note that $\mathbb{E}_{\mathbf{x} \in (\mathbb{F}_p^n)^t} \|F(\mathbf{x}, \cdot) - \beta\|_{U^2(\mathbb{F}_p^n)}^4$ equals

$$\sum_{\omega \in \{0,1\}^4} (-\beta)^{|\omega|} \mathbb{E}_{\mathbf{x} \in (\mathbb{F}_p^n)^t} \mathbb{E}_{y,h,k} f_1^\omega(\mathbf{x}, y) f_2^\omega(\mathbf{x}, y+h) f_3^\omega(\mathbf{x}, y+k) f_4^\omega(\mathbf{x}, y+h+k),$$

where

$$f_i^\omega(\mathbf{x}, y) = \begin{cases} 1 & \omega_i = 1 \\ F(\mathbf{x}, y) & \omega_i = 0 \end{cases}$$

for each $\omega \in \{0,1\}^4$ and $1 \leq i \leq 4$. Since (19) has Cauchy–Schwarz complexity at most $s-1$ by hypothesis, Corollary 3.3 implies that

$$\mathbb{E}_{\mathbf{x} \in (\mathbb{F}_p^n)^t} \mathbb{E}_{y,h,k} f_1^\omega(\mathbf{x}, y) f_2^\omega(\mathbf{x}, y+h) f_3^\omega(\mathbf{x}, y+k) f_4^\omega(\mathbf{x}, y+h+k) = \beta^{4-|\omega|} + O_d(\varepsilon)$$

for every $\omega \in \{0,1\}^4$. Thus,

$$\mathbb{E}_{\mathbf{x} \in (\mathbb{F}_p^n)^t} \|F(\mathbf{x}, \cdot) - \beta\|_{U^2(\mathbb{F}_p^n)}^4 \ll_d \varepsilon$$

Markov's inequality then gives

$$\mathbb{P}(\mathbf{x} \in (\mathbb{F}_p^n)^t : \|F(\mathbf{x}, \cdot) - \beta\|_{U^2(\mathbb{F}_p^n)} \geq r) \ll_d \frac{\varepsilon}{r^4}$$

for every $r > 0$. Taking $r = \varepsilon^{1/8}$ gives the conclusion of the lemma. □

## 4. Pseudorandomness of $\Phi$

The main purpose of this section is to show that if $\|\Phi - \alpha\rho\|_{U^{2s+2}(\mathbb{F}_p^n \times \mathbb{F}_p^n)}$ is small and $\Phi$ is a subset of the form

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

where each $V_x \leq \mathbf{F}_p^n$ is a subspace of density $\rho$ in $\mathbf{F}_p^n$ and $A \subset \mathbf{F}_p^n$ has density $\alpha$ in $\mathbf{F}_p^n$, then, whenever $\psi_1, \dots, \psi_r \in \mathbb{F}_p[x_1, \dots, x_m]$ is a collection of linear forms of Cauchy–Schwarz complexity at



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

most $s$ and $w_1, \dots, w_r \in \mathbb{F}_p^n$, the affine subspaces

$$\left\{ y \in \mathbb{F}_p^n : \prod_{i=1}^{r} \Phi(\psi_i(\mathbf{x}), y + w_i) = 1 \right\}$$

typically have maximum possible codimension. This allows us to transform the condition that $\Phi$ is $U^8(\mathbb{F}_p^n \times \mathbb{F}_p^n)$-pseudorandom into a more useful property for evaluating the various averages that arise in the proof of Theorem 1.2.

Lemma 4.1. *For each nonnegative integer $s$ and positive integer $r$, there exist constants $C_{s,r}, c_{s,r} > 0$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Let $\delta > 0$, $A \subset \mathbb{F}_p^n$ have density $\alpha$, and $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ be a set of the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x \leq \mathbb{F}_p^n$ is a subspace of codimension $d$. Assume that*

$$\|\Phi - \alpha\rho\|_{U^{2s+2}(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < C_{s,r}(\alpha\delta\rho)^{c_{s,r}}.$$

*Let $\psi_1, \dots, \psi_r \in \mathbb{F}_p[x_1, \dots, x_m]$ be a collection of linear forms of Cauchy–Schwarz complexity at most $s$. Then, for all but at most a $\delta$-proportion of $m$-tuples $\mathbf{x} \in (\mathbb{F}_p^n)^m$ for which $\psi_1(\mathbf{x}), \dots, \psi_r(\mathbf{x}) \in A$, we must have that*

$$\operatorname{codim} \left\{ y \in \mathbb{F}_p^n : \prod_{i=1}^{r} \Phi(\psi_i(\mathbf{x}), y + w_i) = 1 \right\} = rd$$

*for all $w_1, \dots, w_r \in \mathbb{F}_p^n$.*

We begin by showing that if $\Phi$ is pseudorandom with respect to the $U^s(\mathbb{F}_p^n \times \mathbb{F}_p^n)$-norm, then $A$ is pseudorandom with respect to the $U^s(\mathbb{F}_p^n)$-norm. This result will also be useful at a few other points in the proof of Theorem 1.2.

Lemma 4.2. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Let $A \subset \mathbb{F}_p^n$ have density $\alpha$ and $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ be of the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Then, for every natural number $s$, we have*

$$\|A - \alpha\|_{U^s(\mathbb{F}_p^n)} \leq \frac{1}{\rho} \|\Phi - \alpha\rho\|_{U^s(\mathbb{F}_p^n \times \mathbb{F}_p^n)}.$$

*Proof.* We write

$$\|A - \alpha\|_{U^s(\mathbb{F}_p^n)}^{2^s} = \mathbb{E}_{x, h_1, \dots, h_s} \prod_{\omega \in \{0,1\}^s} (A - \alpha)(x + \omega \cdot (h_1, \dots, h_s))$$

and then, for each $\omega \in \{0,1\}^s$, insert the identity

$$(A - \alpha)(x + \omega \cdot (h_1, \dots, h_s)) = \frac{1}{\rho} \mathbb{E}_{k_\omega} (\Phi - \alpha\rho)(x + \omega \cdot (h_1, \dots, h_s), k_\omega)$$

to get that $\|A - \alpha\|_{U^s(\mathbb{F}_p^n)}^{2^s}$ equals

$$\frac{1}{\rho^{2^s}} \mathbb{E}_{x, h_1, \dots, h_s} \mathbb{E}_{\substack{k_\omega \in \mathbb{F}_p^n \\ \omega \in \{0,1\}^s}} \prod_{\omega \in \{0,1\}^s} (\Phi - \alpha\rho)(x + \omega \cdot (h_1, \dots, h_s), k_\omega).$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

We can average the above quantity over $y, \ell_1, \dots, \ell_s$ and make the change of variables $k_\omega \mapsto k_\omega + y + \omega \cdot (\ell_1, \dots, \ell_s)$ to get that it equals

$$\frac{1}{\rho^{2^s}} \mathbb{E}_{\substack{k_\omega \in \mathbb{F}_p^n \\ \omega \in \{0,1\}^s}} \mathbb{E}_{x,y,h_1,\dots,h_s,\ell_1,\dots,\ell_s} \prod_{\omega \in \{0,1\}^s} (\Phi - \alpha\rho)((x, k_\omega + y) + \omega \cdot ((h_1, \ell_1), \dots, (h_s, \ell_s))). \quad (20)$$

Applying the Gowers–Cauchy–Schwarz inequality bounds (20) above by

$$\frac{1}{\rho^{2^s}} \mathbb{E}_{\substack{k_\omega \in \mathbb{F}_p^n \\ \omega \in \{0,1\}^s}} \prod_{\omega \in \{0,1\}^s} \|(\Phi - \alpha\rho)(\cdot, k_\omega + \cdot)\|_{U^s(\mathbb{F}_p^n \times \mathbb{F}_p^n)} = \frac{1}{\rho^{2^s}} \|\Phi - \alpha\rho\|_{U^s(\mathbb{F}_p^n \times \mathbb{F}_p^n)}^{2^s},$$

which gives us the conclusion of the lemma. □

To prove Lemma 4.1, we will need the notion of an *approximate polynomial* of bounded degree, which we define using the additive discrete difference operator $\partial$. For $\phi : H \to H$ and $h \in H$, define $\partial_h \phi : H \to H$ by

$$\partial_h \phi(x) := \phi(x) - \phi(x + h),$$

and, for $h_1, \dots, h_s \in H$, the $s$-fold additive difference operator $\partial_{h_1,\dots,h_s}$ by

$$\partial_{h_1,\dots,h_s} f := \partial_{h_1} \cdots \partial_{h_s} f.$$

DEFINITION 4.3. Let $H$ be an abelian group, $A \subset H$, and $\phi : A \to H$. We say that $\phi$ is an *$\varepsilon$-approximate polynomial of degree at most $s-1$* on $A$ if

$$\partial_{h_1,\dots,h_s} \phi(x) = 0$$

for at least an $\varepsilon$-proportion of $(s+1)$-tuples $(x, h_1, \dots, h_s) \in H^{s+1}$ for which $x + \omega \cdot (h_1, \dots, h_s) \in A$ for all $\omega \in \{0,1\}^s$.

We will also need the following result, which is the key combinatorial input into the proof of Lemma 4.1.

LEMMA 4.4. *For each nonnegative integer $s$, there exist constants $C_s, c_s > 0$ such that the following holds. Let $A \subset \mathbb{F}_p^n$ have density $\alpha$, with*

$$\|A - \alpha\|_{U^{2s+2}(\mathbb{F}_p^n)} < C_s(\alpha\delta)^{c_s}$$

*and $\phi : A \to \mathbb{F}_p^n$ be a $\delta$-approximate polynomial of degree at most $s$ on $A$. Then, for at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $(2s+2)$-dimensional parallelopipeds*

$$(x + \omega \cdot (h_1, \dots, h_{2s+2}))_{\omega \in \{0,1\}^{2s+2}}$$

*in $A^{2^{2s+2}}$, the derivative of $\phi$ on the $(2s+1)$-dimensional face $(x + \omega \cdot (h_1, \dots, h_{2s+2}))_{\substack{\omega \in \{0,1\}^{2s+2} \\ \omega_i = \epsilon}}$,*

$$\sum_{\substack{\omega \in \{0,1\}^{2s+2} \\ \omega_i = \epsilon}} (-1)^{|\omega|} \phi(x + \omega \cdot (h_1, \dots, h_{2s+2})),$$

*vanishes for all $1 \le i \le 2s+2$ and $\epsilon = 0, 1$.*

*Proof.* We proceed by induction on $s$, beginning with the case $s = 0$. If $\phi$ is a $\delta$-approximate polynomial of degree at most 0 on $A$, then $\mathbb{E}_{x,y \in A} 1_{\phi(x) = \phi(y)} \ge \delta$, so that, by the pigeonhole principle, there exists some $z \in \mathbb{F}_p^n$ for which $\mu_A(\{x \in A : \phi(x) = z\}) \ge \delta$.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Set $X := \{x \in A : \phi(x) = z\}$, and consider the set of quadruples

$$X' := \{(x, x+h, x+k, x+h+k) \in A^4 : x, x+h, x+k, x+h+k \in X\}.$$

Note that if $(x, x+h, x+k, x+h+k) \in X'$, then

$$\phi(x) = \phi(x+h) = \phi(x+k) = \phi(x+h+k) = z,$$

so certainly the derivatives

$$\phi(x) - \phi(x+h), \phi(x) - \phi(x+k), \phi(x+h) - \phi(x+h+k), \quad \text{and} \quad \phi(x+k) - \phi(x+h+k)$$

of $\phi$ on each of the 1-dimensional faces of the parallelopiped $(x, x+h, x+k, x+h+k)$ vanish. Since

$$\alpha\delta \leq \|X\|_{U^1(\mathbb{F}_p^n)} \leq \|X\|_{U^2(\mathbb{F}_p^n)} = \left(\frac{|X'|}{p^{3n}}\right)^{1/4},$$

we must have $|X'| \geq (\alpha\delta)^4 p^{3n}$. Taking $c_0 = 8$, the total number of quadruples $(x, x+h, x+k, x+h+k)$ in $A^4$ is $(\alpha^4 + O(C_0\alpha^8))p^{3n}$ by Corollary 3.3, which means that $X'$ consists of at least a $\delta^4/2$-proportion of parallelograms $(x, x+h, x+k, x+h+k)$ in $A^4$ if $C_0$ is chosen small enough. Thus, for at least a $\delta^4/2$-proportion of parallelograms $(x, x+h, x+k, x+h+k)$ in $A^4$, the derivative of $\phi$ on each 1-dimensional face vanishes, as desired.

Now suppose that the result holds for a general degree $s-1 \geq 0$, and let $\phi$ be a $\delta$-approximate polynomial of degree at most $s$ on $A$. By Corollary 3.3,

$$\mathbb{E}_{h \in \mathbb{F}_p^n} \|\Delta_h A - \alpha^2\|_{U^{2s}(\mathbb{F}_p^n)}^{2^{2s}} \ll_s C_s(\alpha\delta)^{c_s},$$

so that, as long as $C_s$ is sufficiently small and $c_s$ is sufficiently large, it follows from Markov's inequality that

$$\|\Delta_h A - \alpha^2\|_{U^{2s}(\mathbb{F}_p^n)} < \frac{C_{s-1}(\alpha\delta/2)^{2c_{s-1}}}{2},$$

and, thus,

$$|\mu_{\mathbb{F}_p^n}(A \cap (A-h)) - \alpha^2| < \frac{C_{s-1}(\alpha\delta/2)^{2c_{s-1}}}{2}$$

as well, for all but a $O(\delta^2)$-proportion of $h \in \mathbb{F}_p^n$. Thus, for at least a $\Omega(\delta)$-proportion of $h_{s+1} \in \mathbb{F}_p^n$, we have $\|\Delta_{h_{s+1}} A - \alpha^2\|_{U^{2s}(\mathbb{F}_p^n)} < C_{s-1}(\alpha\delta/2)^{2c_{s-1}}/2$, $|\mu_{\mathbb{F}_p^n}(A \cap (A - h_{s+1})) - \alpha^2| < C_{s-1}(\alpha\delta/2)^{2c_{s-1}}/2$, and that the function $\partial_{h_{s+1}}\phi$ is a $O(\delta)$-approximate polynomial of degree at most $s-1$ on $A \cap (A - h_{s+1})$. Denoting the set of such $h_{s+1}$ by $H$, so that $\mu_{\mathbb{F}_p^n}(H) \gg \delta$, the induction hypothesis then says that, for each $h \in H$, there are at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $2s$-dimensional parallelopipeds $(x + \omega \cdot (h_1, \dots, h_{2s}))_{\omega \in \{0,1\}^{2s}}$ in $(A \cap (A-h))^{2^{2s}}$ for which the derivative of $\partial_h\phi$ on each $(2s-1)$-dimensional face vanishes.

Summing over all $h \in H$, it follows that, for at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $(2s+2)$-tuples $(x, y, h_1, \dots, h_{2s})$ for which $(x + \omega \cdot (h_1, \dots, h_{2s}))_{\omega \in \{0,1\}^{2s}}$ and $(y + \omega \cdot (h_1, \dots, h_{2s}))_{\omega \in \{0,1\}^{2s}}$ are both in $A^{2^{2s}}$, one has

$$\partial_{h_1, \dots, \widehat{h_i}, \dots, h_{2s}} \phi(x + \epsilon h_i) = \partial_{h_1, \dots, \widehat{h_i}, \dots, h_{2s}} \phi(y + \epsilon h_i)$$

for all $i = 1, \dots, 2s$ and $\epsilon = 0, 1$. By Corollary 3.3,

$$\mathbb{E}_{h_1, \dots, h_{2s}} \|\Delta_{h_1, \dots, h_{2s}} A - \alpha^{2^{2s}}\|_{U^1(\mathbb{F}_p^n)}^2 \ll_s C_s(\alpha\delta)^{c_s},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

and so, by Markov's inequality,

$$\mu_{\mathbb{F}_p^n}\Big(\bigcap_{\omega\in\{0,1\}^{2s}}[A-\omega\cdot(h_1,\dots,h_{2s})]\Big)=\alpha^{2^{2s}}+O([C_s(\alpha\delta)^{c_s}]^{\Omega(1)}) \tag{21}$$

for all but a $O([C_s(\alpha\delta)^{c_s}]^{\Omega(1)})$-proportion of $(h_1,\dots,h_{2s})$ in $(\mathbb{F}_p^n)^{2s}$. By taking $C_s$ small enough and $c_s$ large enough, there are therefore at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $2s$-tuples $(h_1,\dots,h_{2s})$ in $(\mathbb{F}_p^n)^{2s}$ for which (21) holds and, for at least a $\Omega_s(\delta^{O_s(1)})$-proportion of pairs $(x,y)\in A^2$ such that $(x+\omega\cdot(h_1,\dots,h_{2s}))_{\omega\in\{0,1\}^{2s}}$ and $(y+\omega\cdot(h_1,\dots,h_{2s}))_{\omega\in\{0,1\}^{2s}}$ are both in $A^{2^{2s}}$, one also has

$$\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(x+\epsilon h_i)=\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(y+\epsilon h_i)$$

for all $i=1,\dots,2s$ and $\epsilon=0,1$. For each such $2s$-tuple $\mathbf{h}$, it follows from the pigeonhole principle that there exists a $y_{\mathbf{h}}$ in the set

$$A_{\mathbf{h}}:=\bigcap_{\omega\in\{0,1\}^{2s}}(A-\omega\cdot(h_1,\dots,h_{2s}))$$

such that, for at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $x\in A_{\mathbf{h}}$, one has

$$\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(x+\epsilon h_i)=\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(y_{\mathbf{h}}+\epsilon h_i)$$

for all $i=1,\dots,2s$ and $\epsilon=0,1$.

Now set $v_{i,\epsilon,\mathbf{h}}:=\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(y_{\mathbf{h}}+\epsilon h_i)$,

$$X_{\mathbf{h}}:=\{x\in A_{\mathbf{h}}:\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(x+\epsilon h_i)=v_{i,\epsilon,\mathbf{h}}\text{ for all }i=1,\dots,2s\text{ and }\epsilon=0,1\},$$

so that $\mu_{A_{\mathbf{h}}}(X_{\mathbf{h}})\gg_s\delta^{O_s(1)}$, and

$$X'_{\mathbf{h}}:=\{(x,x+k,x+k',x+k+k')\in A_{\mathbf{h}}^4:x,x+k,x+k',x+k+k'\in X_{\mathbf{h}}\}.$$

Note that $(x,x+k,x+k',x+k+k')\in X'_{\mathbf{h}}$ if and only if

$$\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s}}\phi(x+\epsilon h_i+\omega'\cdot(k,k'))=v_{i,\epsilon,\mathbf{h}}$$

for all $i=1,\dots,2s$, $\epsilon=0,1$, and $\omega'\in\{0,1\}^2$. Thus, whenever $(x,x+k,x+k',x+k+k')\in X'_{\mathbf{h}}$, we have

$$\begin{aligned}\partial_{h_1,\dots,h_{2s},k}\phi(x)=\partial_{h_1,\dots,h_{2s},k'}\phi(x)=\partial_{h_1,\dots,h_{2s},k}\phi(x+k')&=\partial_{h_1,\dots,h_{2s},k'}\phi(x+k)\\&=v_{1,0,\mathbf{h}}-v_{1,1,\mathbf{h}}-(v_{1,0,\mathbf{h}}-v_{1,1,\mathbf{h}})\\&=0\end{aligned}$$

and

$$\partial_{h_1,\dots,\widehat{h_i},\dots,h_{2s},k,k'}\phi(x+\epsilon h_i)=v_{i,\varepsilon,\mathbf{h}}-v_{i,\varepsilon,\mathbf{h}}-v_{i,\varepsilon,\mathbf{h}}+v_{i,\varepsilon,\mathbf{h}}=0$$

for all $i=1,\dots,2s$ and $\epsilon=0,1$. That is, the derivative of $\phi$ vanishes on all $(2s+1)$-dimensional faces of the $(2s+2)$-dimensional parallelopiped $(x+\omega\cdot(h_1,\dots,h_{2s},k,k'))_{\omega\in\{0,1\}^{2s+2}}$.

As in the $s=0$ case,

$$\delta^{O_s(1)}\mu_{\mathbb{F}_p^n}(A_{\mathbf{h}})\ll_s\|X_{\mathbf{h}}\|_{U^1(\mathbb{F}_p^n)}\le\|X_{\mathbf{h}}\|_{U^2(\mathbb{F}_p^n)}=\left(\frac{\#X'_{\mathbf{h}}}{p^{3n}}\right)^{1/4},$$

so that

$$\#X'_{\mathbf{h}}\gg_s\delta^{O_s(1)}\mu_{\mathbb{F}_p^n}(A_{\mathbf{h}})^4p^{3n}$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

for at least a $\Omega_s(\delta^{O_s(1)})$-proportion of $2s$-tuples $\mathbf{h}$. Each ordered quadruple $(x, \mathbf{h}, k, k')$ for which $(x, x+k, x+k', x+k+k') \in X'_{\mathbf{h}}$ corresponds to a unique $(2s+2)$-dimensional parallelopiped

$$P(x, \mathbf{h}, k, k') = (x + \omega \cdot (h_1, \dots, h_{2s}, k, k'))_{\omega \in \{0,1\}^{2s+2}}$$

in $A^{2s+2}$. Thus,

$$\begin{aligned}&\#\{P(x, \mathbf{h}, k, k') : \mathbf{h} \in (\mathbb{F}_p^n)^{2s} \text{ and } (x, x+k, x+k', x+k+k') \in X'_{\mathbf{h}}\}\\ &\gg_s \delta^{O_s(1)}(\alpha^{2^{2s+2}} + O([C_s(\alpha\delta)^{c_s}]^{\Omega(1)}))p^{(2s+3)n}.\end{aligned}$$

In comparison, the number of $(2s+2)$-dimensional parallelopipeds in $A$ is

$$(\alpha^{2^{2s+2}} + O(C_s(\alpha\delta)^{c_s}))p^{(2s+3)n}$$

by Corollary 3.3. The conclusion of the lemma now follows as long as $C_s$ is sufficiently small and $c_s$ is sufficiently large. □

With a bit more work, it is possible to prove a version of Lemma 4.4 with $(2s+2)$-dimensional parallelopipeds replaced by $(s+2)$-dimensional parallelopipeds (which is optimal), and thus a version of Lemma 4.1 with the $U^{2s+2}$-norm replaced by the $U^{s+2}$-norm, but this would make a negligible difference in Theorem 1.2.

Now we can prove Lemma 4.1.

*Proof of Lemma 4.1.* We proceed by induction on $r$ and $s$, beginning with the $r = 1$, $s = 0$ case[1]. Since $\operatorname{codim}\{y \in \mathbb{F}_p^n : \Phi(x, y) = 1\} = d$ for all $x \in A$, certainly

$$\operatorname{codim}\{y \in \mathbb{F}_p^n : \Phi(\psi_1(\mathbf{x}), y + w_1) = 1\} = \operatorname{codim}(\{y \in \mathbb{F}_p^n : \Phi(\psi_1(\mathbf{x}), y) = 1\} - w_1) = d$$

for all $\mathbf{x} \in (\mathbb{F}_p^n)^m$ for which $\psi_1(\mathbf{x}) \in A$, and this case follows trivially without even needing the assumption that $\|\Phi - \alpha\rho\|_{U^2(\mathbb{F}_p^n \times \mathbb{F}_p^n)}$ is small.

Now let $r \geq 2$ or $s \geq 1$, and assume that the result holds for all pairs of integers $(r', s')$ satisfying:

(i) $0 \leq r' < r$ and $1 \leq s' \leq s$; or
(ii) $1 \leq s' < s$;

and let $C'$ be at most the minimum of $C_{r',s'}$ and $c'$ be at least the maximum of $c_{r',s'}$ over all such pairs with $r' < \max(r, 2^{s+1})$. As long as $\|\Phi - \alpha\rho\|_{U^{2s+2}(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < C'(\alpha\delta\rho^{2r}/2)^{c'2^r}$, it follows from the induction hypothesis that for all but a $O(\delta)$-proportion of $\mathbf{x} \in (\mathbb{F}_p^n)^m$ for which $\psi_1(\mathbf{x}), \dots, \psi_r(\mathbf{x}) \in A$, we must have

$$\mathbb{E}_y \prod_{i=1}^r (\Phi - \rho)(\psi_i(\mathbf{x}), y + w_i) = \mathbb{E}_y \prod_{i=1}^r \Phi(\psi_i(\mathbf{x}), y + w_i) - \rho^r$$

for all $w_1, \dots, w_r \in \mathbb{F}_p^n$. If the codimension of some $\{y \in \mathbb{F}_p^n : \prod_{i=1}^r \Phi(\psi_i(\mathbf{x}), y + w_i) = 1\}$ is not $rd$ for one of these typical $\mathbf{x}$, then it is either $n$ or at most $rd - 1$, which means that

$$\left|\mathbb{E}_y \prod_{i=1}^r (\Phi - \rho)(\psi_i(\mathbf{x}), y + w_i)\right| \geq \frac{\rho^r}{2} \tag{22}$$

in either case, since $p \geq 2$. Squaring both sides of (22), multiplying by $\prod_{i=1}^r A(\psi_i(\mathbf{x}))$, averaging over all $\mathbf{x} \in (\mathbb{F}_p^n)^m$, swapping the order of summation, and applying Lemma 4.2 (to deduce the

[1] Note that if a system of $r$ linear forms has finite Cauchy–Schwarz complexity, then it has Cauchy–Schwarz complexity at most $r - 1$.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

uniformity of $A$) and Theorem 3.2 yields

$$\mathbb{E}_{y,z}\|(\Phi-\rho A)(\cdot,y)(\Phi-\rho A)(\cdot,z)\|_{U^{s+1}(\mathbb{F}_p^n)} \gg \delta\rho^{2r}(\alpha + O_{s,r}((C')^{\Omega_s(1)}\alpha^{c'-1})).$$

By Hölder's inequality, we then have

$$\mathbb{E}_{x,h_1,\dots,h_{s+1}}|\mathbb{E}_y\Delta_{(h_1,0),\dots,(h_{s+1},0)}(\Phi-\rho A)(x,y)|^2 \gg_s (\delta\rho^{2r})^{2^{s+1}}(\alpha^{2^{s+1}} + O_{s,r}((C')^{\Omega_s(1)}\alpha^{c'-1})).$$

It follows from this, the induction hypothesis, our assumption that $\|\Phi-\alpha\rho\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)} < C'(\alpha\delta\rho^{2r}/2)^{c'2^r}$, and Lemma 4.2 that

$$\mathbb{E}_{x,h_1,\dots,h_{s+1}}|\mathbb{E}_y\Delta_{(h_1,0),\dots,(h_{s+1},0)}\Phi(x,y) - \rho^{2^{s+1}}A(x)|^2 \gg_s (\delta\rho^{2r})^{2^{s+1}}(\alpha^{2^{s+1}} + O_{s,r}((C')^{\Omega_s(1)}\alpha^{c'-1})),$$

since the Cauchy–Schwarz complexity of any proper subset of

$$\{x+\omega\cdot(h_1,\dots,h_{s+1}) : \omega\in\{0,1\}^{s+1}\}$$

is at most $s-1$.

Thus, by taking $C'$ sufficiently small and $c'$ sufficiently large, we get that

$$\operatorname{codim}\{y\in\mathbb{F}_p^n : \Delta_{(h_1,0),\dots,(h_{s+1},0)}\Phi(x,y)=1\} \neq 2^{s+1}d \tag{23}$$

for at least an $\Omega_s((\delta\rho)^{O_s(1)})$-proportion of $(s+2)$-tuples $(x,h_1,\dots,h_{s+1})\in(\mathbb{F}_p^n)^{s+2}$ for which $(x+\omega\cdot(h_1,\dots,h_{s+1}))_{\omega\in\{0,1\}^{s+1}}\in A^{2^{s+1}}$. The condition (23) implies that, for each such $(x,h_1,\dots,h_{s+1})$, there exist vectors $v_{x,\mathbf{h},\omega}\in V^{\perp}_{x+\omega\cdot\mathbf{h}}$, $\omega\in\{0,1\}^{s+1}$, not all of which are zero, such that

$$\sum_{\omega\in\{0,1\}^{s+1}} v_{x,\mathbf{h},\omega} = 0.$$

Fix a basis $\{\gamma_{z,1},\dots,\gamma_{z,d}\}$ of $V_z^{\perp}$ for each $z\in\mathbb{F}_p^n$. We can write every vector $v_{x,\mathbf{h},\omega}$ in terms of this basis, giving us that

$$\sum_{\omega\in\{0,1\}^{s+1}}\sum_{j=1}^{d} b_{x,\mathbf{h},\omega,j}\gamma_{x+\omega\cdot\mathbf{h},j} = 0$$

for some $2^{s+1}d$-tuple of constants $(b_{x,\mathbf{h},\omega,j})_{\omega\in\{0,1\}^{s+1},\ 1\le j\le d}$, not all of which are zero.

We apply the pigeonhole principle to deduce that there is some $2^{s+1}d$-tuple of constants $(b_{\omega,j})_{\omega\in\{0,1\}^{s+1},\ 1\le j\le d}$, not all of which are zero, such that

$$\sum_{\omega\in\{0,1\}^{s+1}}\sum_{j=1}^{d} b_{\omega,j}\gamma_{x+\omega\cdot\mathbf{h},j} = 0$$

for at least a $\Omega_s((\delta\rho)^{O_s(1)})$-proportion of $(s+2)$-tuples $(x,h_1,\dots,h_{s+1})\in(\mathbb{F}_p^n)^{s+2}$ for which $(x+\omega\cdot(h_1,\dots,h_{s+1}))_{\omega\in\{0,1\}^{s+1}}\in A^{2^{s+1}}$. Defining

$$\phi_\omega(z) := \sum_{j=1}^{d} b_{\omega,j}\gamma_{z,j}$$

for each $\omega\in\{0,1\}^{s+1}$, the above says that, for at least a $\Omega_s((\delta\rho)^{O_s(1)})$-proportion of $(s+2)$-tuples $(x,h_1,\dots,h_{s+1})\in(\mathbb{F}_p^n)^{s+2}$ for which $(x+\omega\cdot(h_1,\dots,h_{s+1}))_{\omega\in\{0,1\}^{s+1}}\in A^{2^{s+1}}$,



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

we must have

$$\sum_{\omega\in\{0,1\}^{s+1}}\phi_\omega(x+\omega\cdot\mathbf{h})=0,$$

where at least one of the functions $\phi_\omega:\mathbb{F}_p^n\to\mathbb{F}_p^n$ does not have the zero vector in its image (because $\phi_\omega(z)$ is always a nontrivial linear combination of linearly independent vectors). Let $\phi$ be any such $\phi_\omega$. It then follows by applying Corollary 3.3 to the inside average of

$$\mathbb{E}_y\mathbb{E}_{x,h_1,\dots,h_{s+1}}e_p\bigg(y\cdot\sum_{\omega\in\{0,1\}^{s+1}}\phi_\omega(x+\omega\cdot\mathbf{h})\bigg)A(\phi_\omega(x+\omega\cdot\mathbf{h}))$$

that $\partial_{h_1,\dots,h_{s+1}}\phi(x)=0$ for at least a $\Omega_s((\delta\rho)^{O_s(1)})$-proportion of $(s+2)$-tuples $(x,h_1,\dots,h_{s+1})\in(\mathbb{F}_p^n)^{s+2}$ for which $(x+\omega\cdot(h_1,\dots,h_{s+1}))_{\omega\in\{0,1\}^{s+1}}\in A^{2^{s+1}}$, again provided that $C'$ is sufficiently small and $c'$ is sufficiently large. That is, $\phi$ is a $\Omega_s((\delta\rho)^{O_s(1)})$-approximate polynomial of degree at most $s$ on $A$.

If $C'$ is small enough and $c'$ is large enough, Lemmas 4.2 and 4.4, then imply that for at least a $\Omega_s((\alpha\delta\rho)^{O_s(1)})$-proportion of $(2s+2)$-dimensional parallelopipeds $(x+\omega\cdot(h_1,\dots,h_{2s+2}))_{\omega\in\{0,1\}^{2s+2}}$ in $A^{2^{2s+2}}$, the derivative of $\phi$ on each $(2s+1)$-dimensional face vanishes, i.e.

$$\sum_{\substack{\omega\in\{0,1\}^{2s+2}\\ \omega_i=\epsilon}}(-1)^{|\omega|}\phi(x+\omega\cdot(h_1,\dots,h_{2s+2}))=0$$

for all $i=1,\dots,2s+2$ and $\epsilon=0,1$. Call the set of $(2s+3)$-tuples $(x,h_1,\dots,h_{2s+2})$ corresponding to such $(2s+2)$-dimensional parallelopipeds $X$. Consider $\|\Phi-\rho A\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)}^{2^{2s+2}}$, which, plugging in the expression

$$\rho A(x)\sum_{0\neq v\in V_x^\perp}e_p(v\cdot(y-u))$$

for $\Phi-\rho A$, equals

$$\rho^{2^{2s+2}}\mathbb{E}_{x,h_1,\dots,h_{2s+2}}\Delta_{h_1,\dots,h_{2s+2}}A(x)\sum_{\substack{0\neq v_\omega\in V_{x+\omega\cdot\mathbf{h}}^\perp\\ \omega\in\{0,1\}^{2s+2}}}\prod_{\substack{1\le i\le 2s+2\\ \epsilon=0,1}}1_{\mathbf{0}}\bigg(\sum_{\substack{\omega\in\{0,1\}^{2s+2}\\ \omega_i=\epsilon}}(-1)^{|\omega|}v_\omega\bigg).$$

The above has size $\gg_s(p-1)(\alpha\rho)^{2^{2s+2}}(\alpha\delta\rho)^{O_s(1)}\gg_s(\alpha\delta\rho)^{O_s(1)}$, coming from the contribution of $v_\omega=\lambda\phi(x+\omega\cdot(h_1,\dots,h_{2s+2}))$ for each $\lambda\in\mathbb{F}_p^\times$ and $(x,h_1,\dots,h_{2s+2})\in X$. On the other hand, we have

$$\begin{aligned}\|\Phi-\rho A\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)}&=\|\Phi-\rho\alpha+\rho\alpha-\rho A\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)}\\&\le\|\Phi-\rho\alpha\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)}+\rho\|A-\alpha\|_{U^{2s+2}(\mathbb{F}_p^n)},\end{aligned}$$

so that

$$(\alpha\delta\rho)^{O_s(1)}\ll_s\|\Phi-\rho A\|_{U^{2s+2}(\mathbb{F}_p^n\times\mathbb{F}_p^n)}<2C_{s,r}(\alpha\delta\rho)^{c_{s,r}}.$$

Taking $C_{s,r}$ sufficiently small and $c_{s,r}$ sufficiently large will thus yield a contradiction if

$$\operatorname{codim}\bigg\{y\in\mathbb{F}_p^n:\prod_{i=1}^r\Phi(\psi_i(\mathbf{x}),y+w_i)=1\bigg\}\neq rd$$

for some $w_1,\dots,w_r\in\mathbb{F}_p^n$ for a $\delta$-proportion of $\mathbf{x}\in(\mathbb{F}_p^n)^m$ for which $\psi_1(\mathbf{x}),\dots,\psi_r(\mathbf{x})\in A$. □



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

## 5. Control by directional uniformity norms

This section is devoted to proving Lemmas 2.3 and 2.4. Our arguments mostly consist of careful, repeated applications of the Cauchy–Schwarz inequality to ensure that there is no loss of density factors and using the results of §§ 3 and 4 to analyze the resulting averages. As a simple warm-up, we begin by showing that if $T \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ has the form (11) and $B, C$, and $D$ are sufficiently pseudorandom, then $T$ has density close to the product density $\alpha\beta\gamma\delta\rho$.

LEMMA 5.1. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Let $\varepsilon > 0$ and assume that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma$, and $\delta$, respectively, and satisfy*

$$\|B - \beta\|_{U^4(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^4(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^4(\mathbb{F}_p^n)} < \varepsilon$$

*and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Then the set $T \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ defined by (11) has density*

$$\alpha\beta\gamma\delta\rho + O(\varepsilon^{1/8}).$$

*Proof.* The density of $T$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ can be written as

$$\mathbb{E}_{x,y} F(x, y)\Phi(x, y),$$

where $F(x, y) := B(y)C(x + y)D(2x + y)$. Set $L(x, y) := \{y, x + y, 2x + y\}$. Lemma 3.6 says that

$$\mathbb{P}(x \in \mathbb{F}_p^n : \|F(x, \cdot) - \beta\gamma\delta\|_{U^2(\mathbb{F}_p^n)} > \varepsilon^{1/8}) \ll \sqrt{\varepsilon},$$

since

$$L(x, y) \cup L(x, y + h) \cup L(x, y + k) \cup L(x, y + h + k)$$

has Cauchy–Schwarz complexity at most 3. Thus, Lemma 3.5 yields

$$\mathbb{E}_y F(x, y)\Phi(x, y) = (\beta\gamma\delta\rho + O(\varepsilon^{1/8}))A(x)$$

for all but a $O(\sqrt{\varepsilon})$-proportion of $x \in \mathbb{F}_p^n$, so that

$$\mathbb{E}_{x,y} F(x, y)\Phi(x, y) = \alpha\beta\gamma\delta\rho + O(\varepsilon^{1/8}). \qquad \square$$

Now we can prove Lemma 2.3.

*Proof of Lemma 2.3.* The quantity of interest $\Lambda(T, T, T, S)$ is

$$\mathbb{E}_{x,y,z}(B(y + z)B(y + 2z)C(x + y)C(x + y + 2z)D(2x + y)D(2x + y + z)$$
$$\Phi(x, y)\Phi(x, y + z)S(x + z, y)),$$

which, after a change of variables, can be written as

$$\mathbb{E}_{x,y} S(x, y)\mu(x, y),$$

where $\mu(x, y)$ equals

$$\mathbb{E}_z(B(y + z)B(y + 2z)C(x + y - z)C(x + y + z)D(2x + y - 2z)D(2x + y - z)$$
$$\Phi(x - z, y)\Phi(x - z, y + z)).$$

We will show that $\mu(x, y)$ is very close to the constant value $\alpha\beta^2\gamma^2\delta^2\rho^2$ for almost every pair $(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n$, from which it will then follow that $\Lambda(T, T, T, S)$ is close to $\sigma\alpha^2\beta^3\gamma^3\delta^3\rho^3$.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

The first moment $\mathbb{E}_{x,y}\mu(x,y)$ equals

$$\mathbb{E}_{x,y,z}B(y+z)B(y+2z)C(x+y)C(x+y+2z)D(2x+y)D(2x+y+z)\Phi(x,y)\Phi(x,y+z).$$

Applying Lemma 3.6 yields

$$\mathbb{P}((x,y)\in\mathbb{F}_p^n\times\mathbb{F}_p^n : \|F(x,y,\cdot)-\beta^2\gamma\delta\|_{U^2(\mathbb{F}_p^n)}\geq\varepsilon^{1/8})\ll\sqrt{\varepsilon},$$

where

$$F(x,y,z):=B(y+z)B(y+2z)C(x+y+2z)D(2x+y+z).$$

It therefore follows from Lemma 3.5 that

$$\mathbb{E}_z F(x,y,z)\Phi(x,y+z)=(\beta^2\gamma\delta\rho+O(\varepsilon^{1/8}))A(x)$$

for all but a $O(\sqrt{\varepsilon})$-proportion of $(x,y)\in\mathbb{F}_p^n\times\mathbb{F}_p^n$. Thus,

$$\mathbb{E}_{x,y}\mu(x,y)=\beta^2\gamma\delta\rho\mathbb{E}_{x,y}C(x+y)D(2x+y)\Phi(x,y)+O(\varepsilon^{1/8}).$$

By arguing as in the proof of Lemma 5.1, we have

$$\mathbb{E}_{x,y}C(x+y)D(2x+y)\Phi(x,y)=\alpha\gamma\delta\rho+O(\varepsilon^{1/8}),$$

and, thus, conclude that

$$\mathbb{E}_{x,y}\mu(x,y)=\alpha\beta^2\gamma^2\delta^2\rho^2+O(\varepsilon^{1/8}).$$

The second moment $\mathbb{E}_{x,y}\mu(x,y)^2$ equals

$$\begin{aligned}\mathbb{E}_{x,y,z,h}(&\Delta_h B(y+z)\Delta_{2h}B(y+2z)\Delta_{-h}C(x+y)\Delta_h C(x+y+2z)\\&\Delta_{-2h}D(2x+y)\Delta_{-h}D(2x+y+z)\Delta_{(-h,0)}\Phi(x,y)\Delta_{(-h,h)}\Phi(x,y+z)).\end{aligned}$$

Applying Lemma 3.6 again yields

$$\mathbb{P}((x,y,h)\in\mathbb{F}_p^n\times\mathbb{F}_p^n\times\mathbb{F}_p^n : \|G(x,y,h,\cdot)-\beta^4\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon},$$

where

$$G(x,y,h,z):=\Delta_h B(y+z)\Delta_{2h}B(y+2z)\Delta_h C(x+y+2z)\Delta_{-h}D(2x+y+z),$$

and applying Lemma 4.1 yields

$$\mathbb{P}((x,x-h)\in A\times A : \operatorname{codim}\{z\in\mathbb{F}_p^n : \Delta_{(-h,h)}\Phi(x,y+z)=1\}\neq 2d)\ll\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}$$

for all $y\in\mathbb{F}_p^n$. Thus, by Lemma 3.5,

$$\mathbb{E}_y G(x,y,h,z)\Delta_{(-h,h)}\Phi(x,y+z)=\beta^4\gamma^2\delta^2\rho^2+O(\varepsilon^{1/8})$$

for all but a $O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$-proportion of $(x,x+h,y)\in A\times A\times\mathbb{F}_p^n$, so that

$$\mathbb{E}_{x,y}\mu(x,y)^2=\beta^4\gamma^2\delta^2\rho^2\mathbb{E}_{x,y,h}\Delta_{-h}C(x+y)\Delta_{-2h}D(2x+y)\Delta_{(-h,0)}\Phi(x,y)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

By Lemmas 3.6 and 4.1 again, we have

$$\mathbb{P}((x,h)\in\mathbb{F}_p^n\times\mathbb{F}_p^n : \|H(x,h,\cdot)-\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon}$$

and

$$\mathbb{P}((x,x-h)\in A\times A : \operatorname{codim}\{y\in\mathbb{F}_p^n : \Delta_{(-h,0)}\Phi(x,y)=1\}\neq 2d)\ll\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

where

$$H(x,h,y) := \Delta_{-h}C(x+y)\Delta_{-2h}D(2x+y),$$

so that

$$\mathbb{E}_{x,y,h}\Delta_{-h}C(x+y)\Delta_{-2h}D(2x+y)\Delta_{(-h,0)}\Phi(x,y) = \gamma^2\delta^2\rho^2\mathbb{E}_{x,h}\Delta_{-h}A(x) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)$$

by Lemma 3.5. Using Corollary 3.3 to estimate $\mathbb{E}_{x,h}\Delta_{-h}A(x)$, we thus conclude that

$$\mathbb{E}_{x,y}\mu(x,y)^2 = \alpha^2\beta^4\gamma^4\delta^4\rho^4 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

Our estimates for the first and second moments of $\mu$ imply that $\mu$ has variance $\mathbb{E}_{x,y}|\mu(x,y) - \alpha\beta^2\gamma^2\delta^2\rho|^2 \ll \varepsilon^{\Omega(1)}/\rho^{O(1)}$. It follows that

$$\begin{aligned}\mathbb{E}_{x,y}S(x,y)\mu(x,y) &= \sigma\alpha^2\beta^3\gamma^3\delta^3\rho^3 + O(\mathbb{E}_{x,y}|\mu(x,y) - \alpha\beta^2\gamma^2\delta^2\rho^2|)\\ &= \sigma\alpha^2\beta^3\gamma^3\delta^3\rho^3 + O([\mathbb{E}_{x,y}|\mu(x,y) - \alpha\beta^2\gamma^2\delta^2\rho^2|^2]^{1/2})\\ &= \sigma\alpha^2\beta^3\gamma^3\delta^3\rho^3 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).\end{aligned}$$

When $c_1$ is sufficiently small and $c_2$ is sufficiently large, this gives the desired lower bound for $\Lambda(T,T,T,S)$. □

To finish this section, we prove Lemma 2.4.

*Proof of Lemma 2.4.* We prove (16), (17), and then (18), proceeding in decreasing order of the number of applications of Cauchy–Schwarz required. For (16), we make the change of variables $z \mapsto z - x - y$ to write $\Lambda(f_0,f_1,f_2,f_3)$ as

$$\mathbb{E}_{x,y,z}f_0(x,y)f_1(x,z-x)f_2(x,2z-2x-y)f_3(z-y,y),$$

which, by applying the Cauchy–Schwarz inequality in the $x$ and $z$ variables, has modulus squared bounded by

$$\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T)\cdot\mathbb{E}_{x,z}B(z-x)D(x+z)|\mathbb{E}_y f_0(x,y)f_2(x,2z-2x-y)f_3(z-y,y)|^2.$$

The first factor equals $\alpha\beta\gamma\delta\rho + O(\varepsilon^{1/8})$ by Lemma 5.1. Expanding the square and making a change of variables, the second factor equals

$$\mathbb{E}_{x,y,z,h_1}B(z-x)D(x+z)\Delta_{(0,h_1)}f_0(x,y)\Delta_{(0,-h_1)}f_2(x,2z-2x-y)\Delta_{(-h_1,h_1)}f_3(z-y,y).$$

By another application of the Cauchy–Schwarz inequality in the $y,z$, and $h_1$ variables, the modulus squared of this is at most

$$\mathbb{E}_{y,z,h_1}\Delta_{h_1}B(y)C(z+y)\Delta_{-h_1}D(2z+y)\Delta_{(-h_1,h_1)}\Phi(z,y) \tag{24}$$

times

$$\begin{aligned}\mathbb{E}_{y,z,h_1}(C(z)\Delta_{(-h_1,h_1)}\Phi(z-y,y)|\mathbb{E}_x B(z-x)D(x+z)\\ \Delta_{(0,h_1)}f_0(x,y)\Delta_{(0,-h_1)}f_2(x,2z-2x-y)|^2).\end{aligned}$$

The first factor (24) can be estimated in the same manner as the averages appearing in the proof of Lemma 2.3, and equals $\alpha^2\beta^2\gamma\delta^2\rho^2 + O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$. Expanding the square and making



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

a change of variables, we get that the second factor equals

$$\mathbb{E}_{x,y,z,h_1,h_2}(\Delta_{-h_2}B(y-z)C(x+y-z)\Delta_{h_2}D(2x+y-z)\Delta_{(-h_1,h_1)}\Phi(x+z,y-2z)$$
$$\Delta_{(0,h_1),(h_2,0)}f_0(x,y-2z)\Delta_{(0,-h_1),(h_2,-2h_2)}f_2(x,y)).$$

A final application of the Cauchy–Schwarz inequality in the $x, y, h_1$, and $h_2$ variables bounds the modulus squared of this by

$$\mathbb{E}_{x,y,h_1,h_2}\Delta_{-h_1,-2h_2}B(y)\Delta_{-h_1,-h_2}C(x+y)\Delta_{-h_1}D(2x+y)\Delta_{(0,-h_1),(h_2,-2h_2)}\Phi(x,y) \tag{25}$$

times

$$\mathbb{E}_{x,y,h_1,h_2}\Delta_{-h_1,-2h_2}B(y)\Delta_{-h_1,-h_2}C(x+y)\Delta_{-h_1}D(2x+y)\Delta_{(h_2,-2h_2)}\Phi(x,y)$$
$$\cdot|\mathbb{E}_z\Delta_{-h_2}B(y-z)C(x+y-z)\Delta_{h_2}D(2x+y-z)\Delta_{(-h_1,h_1)}\Phi(x+z,y-2z)$$
$$\Delta_{(0,h_1),(h_2,0)}f_0(x,y-2z)|^2$$

By Lemmas 3.6 and 4.1, we have

$$\mathbb{P}((x,y,h_2)\in\mathbb{F}_p^n\times\mathbb{F}_p^n\times\mathbb{F}_p^n:\|F(x,y,h_2,\cdot)-\beta^2\gamma^2\delta\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon},$$
$$\mathbb{P}((x,h_2)\in\mathbb{F}_p^n\times\mathbb{F}_p^n:\|G(x,h_2,\cdot)-\beta^2\gamma^2\delta\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon},$$
$$\mathbb{P}((x,x+h_2)\in A\times A:\operatorname{codim}\{h_1\in\mathbb{F}_p^n:\Delta_{(h_2,-2h_2)}\Phi(x,y-h_1)=1\}\neq 2d)\ll\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}$$

for all $y\in\mathbb{F}_p^n$, and

$$\mathbb{P}((x,x+h_2)\in A\times A:\operatorname{codim}\{y\in\mathbb{F}_p^n:\Delta_{(h_2,-2h_2)}\Phi(x,y)=1\}\neq 2d)\ll\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

where

$$F(x,y,h_2,h_1):=\Delta_{-2h_2}B(y-h_1)\Delta_{-h_2}C(x+y-h_1)D(2x+y-h_1)$$

and

$$G(x,h_2,y):=\Delta_{-2h_2}B(y)\Delta_{-h_2}C(x+y)D(2x+y).$$

It then follows from Lemma 3.5 that (25) equals

$$\beta^2\gamma^2\delta\rho^2\mathbb{E}_{x,y,h_1,h_2}\Delta_{-2h_2}B(y)\Delta_{-h_2}C(x+y)D(2x+y)\Delta_{(h_2,-2h_2)}\Phi(x,y)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

which equals

$$\beta^4\gamma^4\delta^2\rho^4\mathbb{E}_{x,h_2}\Delta_{h_2}A(x)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)=\alpha^2\beta^4\gamma^4\delta^2\rho^4+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

It remains to relate

$$\mathbb{E}_{x,y,h_1,h_2}\Delta_{-h_1,-2h_2}B(y)\Delta_{-h_1,-h_2}C(x+y)\Delta_{-h_1}D(2x+y)\Delta_{(h_2,-2h_2)}\Phi(x,y)$$
$$\cdot|\mathbb{E}_z\Delta_{-h_2}B(y-z)C(x+y-z)\Delta_{h_2}D(2x+y-z)\Delta_{(-h_1,h_1)}\Phi(x+z,y-2z)$$
$$\Delta_{(0,h_1),(h_2,0)}f_0(x,y-2z)|^2$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

to $\|f_0\|_{\star_1}$. Expanding the square and making a change of variables yields

$$\begin{aligned}\mathbb{E}_{x,y,z,h_1,h_2,h_3}(&\Delta_{-h_1,-2h_2}B(y+2z)\Delta_{-h_2,-h_3}B(y+z)\\&\Delta_{-h_1,-h_2}C(x+y+2z)\Delta_{-h_3}C(x+y+z)\\&\Delta_{-h_1}D(2x+y+2z)\Delta_{h_2,-h_3}D(2x+y+z)\\&\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3)}\Phi(x+z,y)\\&\Delta_{(0,h_1),(h_2,0),(0,-2h_3)}f_0(x,y)),\end{aligned}$$

which can be written as

$$\mathbb{E}_{x,y,h_1,h_2,h_3}\Delta_{(0,h_1),(h_2,0),(0,-2h_3)}f_0(x,y)\mu(x,y,h_1,h_2,h_3),$$

where

$$\begin{aligned}\mu(x,y,h_1,h_2,h_3) := \mathbb{E}_z(&\Delta_{-h_1,-2h_2}B(y+2z)\Delta_{-h_2,-h_3}B(y+z)\\&\Delta_{-h_1,-h_2}C(x+y+2z)\Delta_{-h_3}C(x+y+z)\\&\Delta_{-h_1}D(2x+y+2z)\Delta_{h_2,-h_3}D(2x+y+z)\\&\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3)}\Phi(x+z,y)).\end{aligned}$$

We will show that, for almost every 5-tuple $(x,y,h_1,h_2,h_3) \in (\mathbb{F}_p^n)^5$ for which $x, x+h_2 \in A$, $\mu(x,y,h_1,h_2,h_3)$ is very close to the constant value $\alpha^4\beta^8\gamma^6\delta^6\rho^6$. Indeed, the first moment $\mathbb{E}_{\substack{x,y,h_1,h_2,h_3\\ x,x+h_2\in A}}\mu(x,y,h_1,h_2,h_3)$ is

$$\begin{aligned}\frac{1}{\alpha^2+O(\varepsilon)}\mathbb{E}_{x,y,z,h_1,h_2,h_3}(&\Delta_{-h_1,-2h_2}B(y+2z)\Delta_{-h_2,-h_3}B(y+z)\\&\Delta_{-h_1,-h_2}C(x+y+2z)\Delta_{-h_3}C(x+y+z)\\&\Delta_{-h_1}D(2x+y+2z)\Delta_{h_2,-h_3}D(2x+y+z)\\&\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3)}\Phi(x+z,y)).\end{aligned}$$

Lemmas 3.6 and 4.1 tell us that

$$\mathbb{P}((x,z,h_1,h_2,h_3) \in (\mathbb{F}_p^n)^5 : \|H(x,z,h_1,h_2,h_3,\cdot) - \beta^8\gamma^6\delta^6\|_{U^2(\mathbb{F}_p^n)} > \varepsilon^{1/8}) \ll \sqrt{\varepsilon}$$

and

$$\begin{aligned}&\mathbb{P}((x,x+h_2,x+z,x+z-h_1,x+z+h_3,x+z-h_1+h_3) \in A^6\\&\quad : \operatorname{codim}\{y \in \mathbb{F}_p^n : \Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3)}\Phi(x+z,y) = 1\} \neq 6d) \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\end{aligned}$$

where

$$\begin{aligned}H(x,y,h_1,h_2,h_3,z) := (&\Delta_{-h_1,-2h_2}B(y+2z)\Delta_{-h_2,-h_3}B(y+z)\Delta_{-h_1,-h_2}C(x+y+2z)\\&\Delta_{-h_3}C(x+y+z)\Delta_{-h_1}D(2x+y+2z)\Delta_{h_2,-h_3}D(2x+y+z)),\end{aligned}$$

so it follows from Lemma 3.5 that the first moment equals

$$\frac{1}{\alpha^2+O(\varepsilon)}\beta^8\gamma^6\delta^6\rho^6\mathbb{E}_{x,z,h_1,h_2,h_3}\Delta_{h_2}A(x)\Delta_{-h_1,h_3}A(x+z) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

which equals

$$\alpha^4\beta^8\gamma^6\delta^6\rho^6 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

by Corollary 3.3. The second moment $\mathbb{E}_{\substack{x,y,h_1,h_2,h_3 \\ x,x+h_2\in A}}\mu(x,y,h_1,h_2,h_3)^2$ is

$$\frac{1}{\alpha^2+O(\varepsilon)}\mathbb{E}_{x,y,z,h_1,h_2,h_3,k}(\Delta_{-h_1,-2h_2,2k}B(y+2z)\Delta_{-h_2,-h_3,k}B(y+z)$$
$$\Delta_{-h_1,-h_2,2k}C(x+y+2z)\Delta_{-h_3,k}C(x+y+z)$$
$$\Delta_{-h_1,2k}D(2x+y+2z)\Delta_{h_2,-h_3,k}D(2x+y+z)$$
$$\Delta_{(h_2,-2h_2),(0,2k)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3),(k,0)}\Phi(x+z,y)),$$

which, noting that

$$\Delta_{(h_2,-2h_2),(0,2k)}\Phi(x,y+2z)=\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(h_2,-2h_2)}\Phi(x,2k-u),$$

we can write as

$$\mathbb{E}_{x,y,z,h_1,h_2,h_3,k}(\Delta_{-h_1,-2h_2,2k}B(y+2z)\Delta_{-h_2,-h_3,k}B(y+z)$$
$$\Delta_{-h_1,-h_2,2k}C(x+y+2z)\Delta_{-h_3,k}C(x+y+z)$$
$$\Delta_{-h_1,2k}D(2x+y+2z)\Delta_{h_2,-h_3,k}D(2x+y+z)$$
$$\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3),(k,0)}\Phi(x+z,y)$$
$$\Delta_{(h_2,-2h_2)}\Phi(x,2k-u)),$$

Lemmas 3.6 and 4.1, analogously to the case of the first moment, tell us that

$$\mathbb{P}((x,z,h_1,h_2,h_3,k)\in(\mathbb{F}_p^n)^6:\|I(x,z,h_1,h_2,h_3,k,\cdot)-\beta^{16}\gamma^{12}\delta^{12}\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon}$$

and

$$\mathbb{P}\Bigg(x\in A\cap(A-h_2)\text{ and }x+z\in\bigcap_{\omega\in\{0,1\}^8}(A-\omega\cdot(-h_1,h_3,k))$$
$$:\operatorname{codim}\{y\in\mathbb{F}_p^n:\Delta_{(h_2,-2h_2)}\Phi(x,y+2z)\Delta_{(-h_1,h_1),(h_3,-2h_3),(k,0)}\Phi(x+z,y)=1\}\neq 10d\Bigg)$$
$$\ll\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

where

$$I(x,z,h_1,h_2,h_3,k,y):=(\Delta_{-h_1,-2h_2,2k}B(y+2z)\Delta_{-h_2,-h_3,k}B(y+z)$$
$$\Delta_{-h_1,-h_2,2k}C(x+y+2z)\Delta_{-h_3,k}C(x+y+z)$$
$$\Delta_{-h_1,2k}D(2x+y+2z)\Delta_{h_2,-h_3,k}D(2x+y+z)),$$

so it follows from Lemma 3.5 that the second moment equals $(\alpha^2+O(\varepsilon))^{-1}$ times

$$\beta^{16}\gamma^{12}\delta^{12}\rho^{10}\mathbb{E}_{x,z,h_1,h_2,h_3,k}\Delta_{-h_1,h_3,k}A(x+z)\Delta_{(h_2,-2h_2)}\Phi(x,2k-u)+O\Bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\Bigg).\tag{26}$$

To estimate the main term of (26), we note that

$$\|\Phi(\cdot,2\cdot-u)-\alpha\rho\|_{U^6(\mathbb{F}_p^n\times\mathbb{F}_p^n)}=\|\Phi-\alpha\rho\|_{U^6(\mathbb{F}_p^n\times\mathbb{F}_p^n)}$$

and apply Lemmas 3.6 and 4.1 again to get that

$$\mathbb{P}((x,z,h_1,h_3)\in(\mathbb{F}_p^n)^4:\|J(x,z,h_1,h_3,\cdot)-\alpha^4\|_{U^2(\mathbb{F}_p^n)}>\varepsilon^{1/8})\ll\sqrt{\varepsilon}$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

and

$$\mathbb{P}((x, x+h_2) \in A \times A : \operatorname{codim}\{k \in \mathbb{F}_p^n : \Delta_{(h_2,-2h_2)}\Phi(x, 2k-u) = 1\} \neq 2d) \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

where

$$J(x, z, h_1, h_3, k) := \Delta_{-h_1,h_3}A(x+z+k).$$

Thus, by Lemma 3.5, we have that

$$\mathbb{E}_{x,z,h_1,h_2,h_3,k}\Delta_{-h_1,h_3,k}A(x+z)\Delta_{(h_2,-2h_2)}\Phi(x, 2k-u)$$

equals

$$\alpha^4\rho^2\mathbb{E}_{x,z,h_1,h_2,h_3}\Delta_{h_2}A(x)\Delta_{-h_1,h_3}A(x+z) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

which equals

$$\alpha^{10}\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)$$

by Corollary 3.3. It therefore follows that the second moment is

$$\alpha^8\beta^{16}\gamma^{12}\delta^{12}\rho^{12} + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

Thus,

$$\mathbb{E}_{\substack{x,y,h_1,h_2,h_3 \\ x,x+h_2\in A}}|\mu(x, y, h_1, h_2, h_3) - \alpha^4\beta^8\gamma^6\delta^6\rho^6|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

and we conclude that

$$\mathbb{E}_{x,y,h_1,h_2,h_3}\Delta_{(0,h_1),(h_2,0),(0,-2h_3)}f_0(x, y)\mu(x, y, h_1, h_2, h_3) = \alpha^4\beta^8\gamma^6\delta^6\rho^6\|f_0\|_{\star_1}^8 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

Putting everything together gives

$$|\Lambda(f_0, f_1, f_2, f_3)|^8 \leq \alpha^{14}\beta^{20}\gamma^{16}\delta^{16}\rho^{18}\|f_0\|_{\star_1}^8 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

as desired.

For (17), we make a change of variables and apply the Cauchy–Schwarz inequality to bound $|\Lambda(T, f_1, f_2, f_3)|^2$ by

$$\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T) \cdot \mathbb{E}_{x,y}B(y)C(x+y)\Phi(x, y) \cdot |\mathbb{E}_z C(x+y-2z)D(2x+y-2z)f_1(x, y-z)f_3(x+z, y-2z)|^2.$$

Expanding the square in the second quantity and making a change of variables yields

$$\mathbb{E}_{x,y,z,h_1}B(y+2z)C(x+y+z)\Delta_{-2h_1}C(x+y-z)\Delta_{-2h_1}D(2x+y-2z)\Phi(x-z, y+2z)$$
$$\Delta_{(0,-h_1)}f_1(x-z, y+z)\Delta_{(h_1,-2h_1)}f_3(x, y).$$

By applying the Cauchy–Schwarz inequality again in the $x, y$, and $h_1$ variables, the modulus squared of this is bounded above by

$$\mathbb{E}_{x,y,h_1}\Delta_{-2h_1}B(y)\Delta_{-h_1}C(x+y)D(2x+y)\Delta_{(h_1,-2h_1)}\Phi(x, y) \tag{27}$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

times

$$\mathbb{E}_{x,y,h_1}(\Delta_{-2h_1}B(y)D(2x+y)\Delta_{(h_1,-2h_1)}\Phi(x,y)$$
$$|\mathbb{E}_z[B(y+2z)C(x+y+z)\Delta_{-2h_1}C(x+y-z)\Delta_{-2h_1}D(2x+y-2z)\Phi(x-z,y+2z)$$
$$\Delta_{(0,-h_1)}f_1(x-z,y+z)]|^2).$$

Expanding the square and making a change of variables, the second factor equals

$$\mathbb{E}_{x,y,z,h_1,h_2}\Delta_{2h_2}B(y+z)\Delta_{-2h_1}B(y-z)\Delta_{-2h_1,-h_2}C(x+y-z)\Delta_{h_2}C(x+y+z)$$
$$\Delta_{-2h_1,-2h_2}D(2x+y-z)D(2x+y+z)\Delta_{(-h_2,h_2)}\Phi(x,y+z)\Delta_{(h_1,-2h_1)}\Phi(x+z,y-z)$$
$$\Delta_{(0,-h_1),(-h_2,h_2)}f_1(x,y),$$

which we can write as

$$\mathbb{E}_{x,y,h_1,h_2}\Delta_{(0,-h_1),(-h_2,h_2)}f_1(x,y)\mu'(x,y,h_1,h_2),$$

where

$$\mu'(x,y,h_1,h_2) := \mathbb{E}_z\Delta_{2h_2}B(y+z)\Delta_{-2h_1}B(y-z)\Delta_{-2h_1,-h_2}C(x+y-z)\Delta_{h_2}C(x+y+z)$$
$$\Delta_{-2h_1,-2h_2}D(2x+y-z)D(2x+y+z)\Delta_{(-h_2,h_2)}\Phi(x,y+z)$$
$$\Delta_{(h_1,-2h_1)}\Phi(x+z,y-z).$$

The quantity (27) and the weight $\mu'$ can be analyzed in the same manner as the corresponding quantity and weight in the proof of (16), so that

$$|\Lambda(T,f_1,f_2,f_3)|^4 \leq \alpha^8\beta^8\gamma^{10}\delta^8\rho^8\|f_1\|_{\star_2}^4 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

Finally, for (18), we make a change of variables and apply the Cauchy–Schwarz inequality to bound $|\Lambda(T,T,f_2,f_3)|^2$ by $\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T)$ times

$$\mathbb{E}_{x,y}B(y)C(x+y)\Phi(x,y)|\mathbb{E}_z B(y+z)C(x+y-z)D(2x+y-2z)D(2x+y-z)$$
$$\Phi(x-z,y+z)f_2(x-z,y+2z)|^2.$$

Expanding the square in the second quantity and making a change of variables yields

$$\mathbb{E}_{x,y,z,h_1}B(y)C(x+y-z)\Phi(x+z,y-2z)\Delta_{h_1}B(y-z)\Delta_{-h_1}C(x+y-2z)$$
$$\Delta_{-2h_1}D(2x+y-2z)\Delta_{-h_1}D(2x+y-z)\Delta_{(-h_1,h_1)}\Phi(x,y-z)\Delta_{(-h_1,2h_1)}f_2(x,y),$$

which can be written as

$$\mathbb{E}_{x,y,h_1}\Delta_{(-h_1,2h_1)}f_2(x,y)\mu''(x,y,h_1),$$

where

$$\mu''(x,y,h_1) := \mathbb{E}_z B(y)C(x+y-z)\Phi(x+z,y-2z)\Delta_{h_1}B(y-z)\Delta_{-h_1}C(x+y-2z)$$
$$\Delta_{-2h_1}D(2x+y-2z)\Delta_{-h_1}D(2x+y-z)\Delta_{(-h_1,h_1)}\Phi(x,y-z).$$

This weight can again be analyzed in the same manner as the weights appearing in the proof of (16), giving

$$|\Lambda(T,T,f_2,f_3)|^2 \leq \alpha\beta^3\gamma^3\delta^4\rho^3\|f_2\|_{\star_3}^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

and completing the proof of the lemma. □



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

## 6. Obtaining a density increment

As was mentioned in § 2, one ingredient in the proof of Theorem 2.5 is an inverse theorem for the $U^2(\Phi(x,\cdot))$-norm localized to pseudorandom sets. Applying the standard, nonlocalized inverse theorem for the $U^2(\Phi(x,\cdot))$-norm would yield a density increment that gets weaker as $T$ becomes sparser, which is inadequate to close the density-increment iteration. To get a strong enough localized version of this inverse theorem, we will need to use the transference principle. The particular instance of it required is an immediate consequence of the dense model lemma from [Zha14], which appears as Lemma 3.3 in that paper.

LEMMA 6.1 (Dense model lemma for the $U^s$-norm on subspaces). *For every natural number $s$, there exists a constant $c_s > 0$ such that the following holds. Let $\varepsilon > 0$, $V \leq \mathbb{F}_p^n$ be a subspace, and $f, \nu : V \to [0,\infty)$ be functions satisfying:*

(i) $0 \leq f \leq \nu$;
(ii) $\mathbb{E}_{x\in V} f(x) \leq 1$; *and*
(iii) $\|\nu - 1\|_{U^s(V)} \leq \exp(-\varepsilon^{-c_s})$.

*Then there exists a $\tilde{f} : V \to [0,1]$ such that $\mathbb{E}_{x\in V} f(x) = \mathbb{E}_{x\in V} \tilde{f}(x)$ and $\|f - \tilde{f}\|_{U^s(V)} \leq \varepsilon$.*

One can see that this lemma is a consequence of Zhao's lemma by using the Gowers–Cauchy–Schwarz inequality to translate between his $(s,\varepsilon)$-discrepancy pair condition and our $U^s$-uniformity condition.

In the course of the proof of Theorem 2.5, we will encounter various averages of linear forms that turn out to be controlled by certain degree 1 and 2 directional uniformity norms. Because of this, we will also need to obtain a density increment when these norms of the balanced function $g_S = S - \sigma$ are large. The first two subsections of this section are devoted to proving that this is possible.

### 6.1 Results on degree 1 norms

We first show that the relevant fibers of any set of the form (11) typically have close to their average density, provided that $A, B, C, D$, and $\Phi$ are sufficiently pseudorandom.

LEMMA 6.2. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ has density $\alpha\rho$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ and takes the form*

$$\Phi = \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\varepsilon > 0$ and assume that*

$$\|A - \alpha\|_{U^5(\mathbb{F}_p^n)}, \quad \|B - \beta\|_{U^5(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^5(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^5(\mathbb{F}_p^n)}, \quad \|\Phi - \alpha\rho\|_{U^2(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < \varepsilon.$$

*Define $T$ by (11). Then*

$$\mathbb{P}(x \in A : |\mu_{\mathbb{F}_p^n}(T(x,\cdot)) - \beta\gamma\delta\rho| > \varepsilon') \ll \frac{\varepsilon^{1/8}}{(\varepsilon')^2\alpha},$$

$$\mathbb{P}(y \in B : |\mu_{\mathbb{F}_p^n}(T(\cdot,y)) - \alpha\gamma\delta\rho| > \varepsilon') \ll \frac{\varepsilon^{1/8}}{(\varepsilon')^2\beta},$$

$$\mathbb{P}(z \in C : |\mu_{\mathbb{F}_p^n}(T(\cdot,z-\cdot)) - \alpha\beta\delta\rho| > \varepsilon') \ll \frac{\varepsilon^{1/8}}{(\varepsilon')^2\gamma},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*and*

$$\mathbb{P}(w \in D : |\mu_{\mathbb{F}_p^n}(T(\cdot, w - 2\cdot)) - \alpha\beta\gamma\rho| > \varepsilon') \ll \frac{\varepsilon^{1/8}}{(\varepsilon')^2\delta}$$

*for any* $\varepsilon' > 0$.

*Proof.* From Lemma 5.1, we have $\mathbb{E}_{x\in A}\mu_{\mathbb{F}_p^n}(T(x,\cdot)) = \beta\gamma\delta\rho + O(\varepsilon^{1/8}/\alpha)$. For the second moment, we argue as in § 5 to estimate

$$\mathbb{E}_x A(x)|\mathbb{E}_y B(y)C(x+y)D(2x+y)\Phi(x,y)|^2 = \alpha\beta^2\gamma^2\delta^2\rho^2 + O(\varepsilon^{1/8})$$

using Lemmas 3.5 and 3.6, giving $\mathbb{E}_{x\in A}\mu_{\mathbb{F}_p^n}(T(x,\cdot))^2 = \beta^2\gamma^2\delta^2\rho^2 + O(\varepsilon^{1/8}/\alpha)$. It now follows from Markov's inequality that

$$\mathbb{P}(x \in A : |\mu_{\mathbb{F}_p^n}(T(x,\cdot)) - \beta\gamma\delta\rho| > \varepsilon') \ll \frac{\varepsilon^{1/8}}{(\varepsilon')^2\alpha}$$

for all $\varepsilon' > 0$. The three other estimates are proved analogously. □

Now we can obtain a density increment when the degree 1 uniformity norms controlled by $\|\cdot\|_{\star_1}$ and $\|\cdot\|_{\star_2}$ are large. The proof is essentially an averaging argument, like the proof of the analogous Lemma 3.1 in [Gre05a]. The most substantial new feature, which will arise many times in this section, is that we must now verify that the set on which we claim to have found a density increment actually has close to the correct density in $\mathbb{F}_p^n \times \mathbb{F}_p^n$. In the case of corners, any product set $A \times B$ trivially has density equal to the product of the densities of $A$ and $B$. This is not, in general, the case for sets of the form (11) unless further assumptions are made about $A, B, C, D$, and $\Phi$.

Lemma 6.3. *There exist absolute constants* $0 < c_1 < 1 < c_2$ *such that the following holds. Let* $d$ *be a nonnegative integer, and set* $\rho := p^{-d}$. *Suppose that* $A, B, C, D \subset \mathbb{F}_p^n$ *have densities* $\alpha, \beta, \gamma, \delta$, *respectively, and that* $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ *takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each* $V_x$ *is a subspace of* $\mathbb{F}_p^n$ *of codimension* $d$. *Let* $\tau > 0$ *and* $\varepsilon \le c_1(\tau\alpha\beta\gamma\delta\rho)^{c_2}$, *and assume that*

$$\|A - \alpha\|_{U^5(\mathbb{F}_p^n)}, \quad \|B - \beta\|_{U^5(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^5(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^5(\mathbb{F}_p^n)}, \quad \|\Phi_0 - \rho\|_{U^2(\mathbb{F}_p^n\times\mathbb{F}_p^n)} < \varepsilon.$$

*Define* $T$ *by* (11), *and let* $S \subset T$ *have density* $\sigma$ *in* $T$. *Suppose that*

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \ge \tau\alpha\beta^2\gamma^2\delta^2\rho^2, \tag{28}$$

$$\mathbb{E}_{x,y,h}\Delta_{(h,0)}g_S(x,y) \ge \tau\alpha^2\beta\gamma^2\delta^2\rho^2, \tag{29}$$

*or*

$$\mathbb{E}_{x,y,h}\Delta_{(-h,h)}g_S(x,y) \ge \tau\alpha^2\beta^2\gamma\delta^2\rho^2. \tag{30}$$

*Then* $S$ *has density at least* $\sigma + \Omega(\tau^{O(1)})$ *on some subset* $T'$ *of* $T$ *of the form*

$$T' = \{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : A'(x)B'(y)C'(x+y)D(2x+y)\Phi(x,y) = 1\},$$

*where the densities of* $A', B', C' \subset \mathbb{F}_p^n$ *are* $\Omega(\tau^{O(1)}\alpha)$, $\Omega(\tau^{O(1)}\beta)$, *and* $\Omega(\tau^{O(1)}\gamma)$, *respectively.*

*Proof.* The assumption (28) can be written as

$$\mathbb{E}_{x\in A}|\mathbb{E}_y g_S(x,y)|^2 \ge \tau\beta^2\gamma^2\delta^2\rho^2.$$

Since $\mu_{\mathbb{F}_p^n}(T(x,\cdot)) > 2\beta\gamma\delta\rho$ for at most a $O(\varepsilon^{1/8}/\alpha^3\beta^2\gamma^2\delta^2\rho^2)$-proportion of $x \in A$ by Lemma 6.2, it follows that, as long as $c_2$ is sufficiently large, there exists a subset $A_0 \subset A$ of relative density



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

at least $\tau/4$ for which

$$|\mathbb{E}_y g_S(x,y)| \geq \frac{\sqrt{\tau}\beta\gamma\delta\rho}{4}$$

for all $x \in A_0$, provided that $c_1$ is small enough and $c_2$ is large enough. Note that $\mathbb{E}_y g_S(x,y)$ is a real number, and thus is either positive or negative. There must therefore exist a subset $A_1 \subset A_0$ of density at least $1/2$ in $A_0$ such that either

$$\mathbb{E}_y g_S(x,y) \geq \frac{\sqrt{\tau}\beta\gamma\delta\rho}{4}$$

for every $x \in A_1$ or

$$\mathbb{E}_y g_S(x,y) \leq -\frac{\sqrt{\tau}\beta\gamma\delta\rho}{4}$$

for every $x \in A_1$.

In the first case, setting $A' = A_1$ and $\alpha' = \mu_{\mathbb{F}_p^n}(A')$, we have

$$|\mathbb{E}_{x\in A'}\mathbb{E}_y g_S(x,y) - (\mathbb{E}_{x\in A'}\mathbb{E}_y S(x,y) - \sigma\beta\gamma\delta\rho)| < \frac{\sqrt{\tau}\beta\gamma\delta\rho}{8}$$

by Lemma 6.2 whenever $c_1$ is small enough and $c_2$ is large enough, so that

$$\mathbb{E}_{x,y} A'(x)S(x,y) > \left(\sigma + \frac{\sqrt{\tau}}{8}\right)\alpha'\beta\gamma\delta\rho.$$

Since $\mathbb{E}_{x,y}A'(x)B(y)C(x+y)D(2x+y)\Phi(x,y) = \alpha'\beta\gamma\delta\rho + O(\varepsilon^{1/16}/\alpha)$ by Lemma 6.2, the conclusion of the lemma follows when $c_1$ is small enough and $c_2$ is large enough by taking $B' = B$.

In the second case, setting $A' = A \setminus A_1$ and $\alpha' = \mu_{\mathbb{F}_p^n}(A')$, we use the fact that $\mathbb{E}_{x,y} g_S(x,y) = 0$ to deduce that

$$\mathbb{E}_{x\in A'}\mathbb{E}_y g_S(x,y) \geq \frac{\tau^{3/2}\beta\gamma\delta\rho}{16}$$

whenever $c_1$ is small enough and $c_2$ is large enough. Note that $\mathbb{E}_{x,y}A'(x)B(y)C(x+y)D(2x+y)\Phi(x,y) = \alpha'\beta\gamma\delta\rho + O(\varepsilon^{1/16}/\alpha)$ in this case as well by Lemma 6.2, so the conclusion of the lemma will follow as long as $\alpha' = \Omega(\tau^{O(1)}\alpha)$. But this also follows from Lemma 6.2, for we have

$$\tau\alpha(\beta\gamma\delta\rho) + O(\varepsilon^{1/16}) \ll |\mathbb{E}_x A_1(x) g_S(x,y)| = |\mathbb{E}_x A'(x) g_S(x,y)| = \alpha'(\beta\gamma\delta\rho) + O(\alpha'\varepsilon^{1/16}/\alpha).$$

The proof of the lemma starting from the assumptions (29) and (30) is essentially identical, but using the second and third probability estimates from Lemma 6.2, respectively. □

### 6.2 Results on degree 2 norms

To obtain a density increment when the relevant degree 2 directional uniformity norms of $g_S$ are large, we first need to show that certain degree 2 'inner products' are controlled by the degree 1 directional uniformity norms studied in the previous subsection.

Lemma 6.4. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\varepsilon > 0$ and assume that*

$$\|A - \alpha\|_{U^8(\mathbb{F}_p^n)}, \quad \|B - \beta\|_{U^8(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^8(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^8(\mathbb{F}_p^n)}, \quad \|\Phi - \alpha\rho\|_{U^4(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < \varepsilon.$$

*Define $T$ by (11), and let $S \subset T$ and $\tau > 0$.*



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*If*

$$\left|\mathbb{E}_{x,y,h,k}\prod_{\omega\in\{0,1\}^2} g_\omega((x,y)+\omega\cdot((0,h),(0,k)))\right| \geq \tau\alpha\beta^4\gamma^4\delta^4\rho^3, \tag{31}$$

*where $g_\omega$ equals $T$ or $g_S$ for all $\omega\in\{0,1\}^2$, at least one $g_\omega$ equals $g_S$, and at least one $g_\omega$ equals $T$, then*

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \geq \tau^2\alpha\beta^2\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right).$$

*If*

$$\left|\mathbb{E}_{x,y,h,k}\prod_{\omega\in\{0,1\}^2} g_\omega((x,y)+\omega\cdot((h,0),(0,k)))\right| \geq \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4, \tag{32}$$

*where $g_\omega$ equals $T$ or $g_S$ for all $\omega\in\{0,1\}^2$, at least one $g_\omega$ equals $g_S$, and at least one $g_\omega$ equals $T$, then*

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \geq \tau^2\alpha\beta^2\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right),$$

*or*

$$\mathbb{E}_{x,y,h}\Delta_{(h,0)}g_S(x,y) \geq \tau^2\alpha^2\beta\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right).$$

*If*

$$\left|\mathbb{E}_{x,y,h,k}\prod_{\omega\in\{0,1\}^2} g_\omega((x,y)+\omega\cdot((0,h),(-k,k)))\right| \geq \tau\alpha^2\beta^4\gamma^2\delta^4\rho^4, \tag{33}$$

*where $g_\omega$ equals $T$ or $g_S$ for all $\omega\in\{0,1\}^2$, at least one $g_\omega$ equals $g_S$, and at least one $g_\omega$ equals $T$, then*

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \geq \tau^2\alpha\beta^2\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right),$$

*or*

$$\mathbb{E}_{x,y,h}\Delta_{(-h,h)}g_S(x,y) \geq \tau^2\alpha^2\beta^2\gamma\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right).$$

*Proof.* We rewrite the various assumptions that (31), (32), and (33) hold when at least two of the $g_\omega$ equal $T$ as

$$\begin{aligned}
|\mathbb{E}_{x,y,h}g_S(x,y)g_S(x,y+h)\mu_1(x,y,h)| &\geq \tau\alpha\beta^4\gamma^4\delta^4\rho^3,\\
|\mathbb{E}_{x,y}g_S(x,y)\mu_2(x,y)| &\geq \tau\alpha\beta^4\gamma^4\delta^4\rho^3,\\
|\mathbb{E}_{x,y,k}g_S(x,y)g_S(x,y+k)\mu_3(x,y,k)| &\geq \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4,\\
|\mathbb{E}_{x,y,h}g_S(x,y)g_S(x+h,y)\mu_4(x,y,h)| &\geq \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4,\\
|\mathbb{E}_{x,y}g_S(x,y)\mu_5(x,y)| &\geq \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4,\\
|\mathbb{E}_{x,y,h}g_S(x,y)g_S(x,y+h)\mu_6(x,y,h)| &\geq \tau\alpha^2\beta^4\gamma^2\delta^4\rho^4,\\
|\mathbb{E}_{x,y,k}g_S(x,y)g_S(x-k,y+k)\mu_7(x,y,k)| &\geq \tau\alpha^2\beta^4\gamma^2\delta^4\rho^4,
\end{aligned}$$

and

$$|\mathbb{E}_{x,y}g_S(x,y)\mu_8(x,y)| \geq \tau\alpha^2\beta^4\gamma^2\delta^4\rho^4,$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

where

$$\begin{aligned}
\mu_1(x,y,h) &= \mathbb{E}_k T(x,y+k)T(x,y+h+k),\\
\mu_2(x,y) &= \mathbb{E}_{h,k} T(x,y+h)T(x,y+k)T(x,y+h+k),\\
\mu_3(x,y,k) &= \mathbb{E}_h T(x+h,y)T(x+h,y+k),\\
\mu_4(x,y,h) &= \mathbb{E}_k T(x,y+k)T(x+h,y+k),\\
\mu_5(x,y) &= \mathbb{E}_{h,k} T(x+h,y)T(x,y+k)T(x+h,y+k),\\
\mu_6(x,y,h) &= \mathbb{E}_k T(x-k,y+k)T(x-k,y+h+k),\\
\mu_7(x,y,k) &= \mathbb{E}_h T(x,y+h)T(x-k,y+h+k),
\end{aligned}$$

and

$$\mu_8(x,y) = \mathbb{E}_{h,k} T(x,y+h)T(x-k,y+k)T(x-k,y+h+k).$$

Using the definition (11) of $T$ and arguing as in § 5 using Lemmas 3.5, 3.6, and 4.1 gives that each of $\mu_1,\dots,\mu_8$ is typically very close to its average value on its support. Precisely, we have the estimates

$$\begin{aligned}
\mathbb{E}_{x,y,h} A(x)\Phi(x,h-u)|\mu_1(x,y,h) - \beta^2\gamma^2\delta^2\rho|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y} A(x)\Phi(x,y)|\mu_2(x,y) - \beta^3\gamma^3\delta^3\rho^2|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y,k} B(y)B(y+k)|\mu_3(x,y,k) - \alpha\gamma^2\delta^2\rho^2|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y,h} A(x)A(x+h)|\mu_4(x,y,h) - \beta\gamma^2\delta^2\rho^2|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y} A(x)B(y)|\mu_5(x,y) - \beta^2\gamma^3\delta^3\rho^3|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y,h} C(x+y)C(x+y+h)|\mu_6(x,y,h) - \alpha\beta^2\delta^2\rho^2|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},\\
\mathbb{E}_{x,y,k} A(x)A(x-k)|\mu_7(x,y,k) - \beta^2\gamma\delta^2\rho^2|^2 &\ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},
\end{aligned}$$

and

$$\mathbb{E}_{x,y} A(x)C(x+y)|\mu_8(x,y) - \alpha\beta^3\gamma\delta^3\rho^3|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

which imply that

$$|\mathbb{E}_{x,y,h} g_S(x,y)g_S(x,y+h)\mu_1(x,y,h)| = \beta^2\gamma^2\delta^2\rho\mathbb{E}_{x,y,h} g_S(x,y)g_S(x,y+h) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

$$|\mathbb{E}_{x,y} g_S(x,y)\mu_2(x,y)| = \beta^3\gamma^3\delta^3\rho^2|\mathbb{E}_{x,y} g_S(x,y)| + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

$$|\mathbb{E}_{x,y,k} g_S(x,y) g_S(x,y+k) \mu_3(x,y,k)| = \alpha\gamma^2\delta^2\rho^2 \mathbb{E}_{x,y,k} g_S(x,y) g_S(x,y+k) + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg),$$

$$|\mathbb{E}_{x,y,h} g_S(x,y) g_S(x+h,y) \mu_4(x,y,h)| = \beta\gamma^2\delta^2\rho^2 \mathbb{E}_{x,y,h} g_S(x,y) g_S(x+h,y) + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg),$$

$$|\mathbb{E}_{x,y} g_S(x,y) \mu_5(x,y)| = \beta^2\gamma^3\delta^3\rho^3 |\mathbb{E}_{x,y} g_S(x,y)| + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg),$$

$$|\mathbb{E}_{x,y,h} g_S(x,y) g_S(x,y+h) \mu_6(x,y,h)| = \alpha\beta^2\delta^2\rho^2 \mathbb{E}_{x,y,h} g_S(x,y) g_S(x,y+h) + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg),$$

$$|\mathbb{E}_{x,y,k} g_S(x,y) g_S(x-k,y+k) \mu_7(x,y,k)| = \beta^2\gamma\delta^2\rho^2 \mathbb{E}_{x,y,k} g_S(x,y) g_S(x-k,y+k) + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg),$$

and

$$|\mathbb{E}_{x,y} g_S(x,y) \mu_8(x,y)| = \alpha\beta^3\gamma\delta^3\rho^3 |\mathbb{E}_{x,y} g_S(x,y)| + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg).$$

Since $|\mathbb{E}_{x,y} g_S(x,y)|^2 \le \alpha \mathbb{E}_{x,y,h} g_S(x,y) g_S(x,y+h)$ by an application of the Cauchy–Schwarz inequality, the conclusion of the lemma easily follows starting from any one of the assumptions (31), (32), or (33) when at least two of the $g_\omega$ equal $T$.

To prove the lemma when only one $g_\omega$ in (31), (32), or (33) equals $T$, we will apply the Cauchy–Schwarz inequality once, and then argue analogously. By making a change of variables, we may start from the assumption that

$$|\mathbb{E}_{x,y,h,k} g_S(x,y) g_S(x,y+h) g_S(x,y+k) T(x,y+h+k)| \ge \tau\alpha\beta^4\gamma^4\delta^4\rho^3,$$
$$|\mathbb{E}_{x,y,h,k} g_S(x,y) g_S(x+h,y) g_S(x,y+k) T(x+h,y+k)| \ge \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4,$$

or

$$|\mathbb{E}_{x,y,h,k} g_S(x,y) g_S(x,y+h) g_S(x-k,y+k) T(x-k,y+h+k)| \ge \tau\alpha^2\beta^4\gamma^2\delta^4\rho^4.$$

We apply the Cauchy–Schwarz inequality in each of these three cases to get that

$$\mathbb{E}_{x,y,h} |\mathbb{E}_k g_S(x,y+k) T(x,y+h+k)|^2 \Delta_h B(y) \Delta_h C(x+y) \Delta_h D(2x+y) \Phi(x,y) \tag{34}$$

times $\mathbb{E}_{x,y,h} \Delta_{(0,h)} T(x,y)$ is at least $\tau^2\alpha^2\beta^8\gamma^8\delta^8\rho^6$,

$$\mathbb{E}_{x,y,h} |\mathbb{E}_k g_S(x,y+k) T(x+h,y+k)|^2 B(y) \Delta_h C(x+y) \Delta_{2h} D(2x+y) \Delta_{(h,0)} \Phi(x,y) \tag{35}$$

times $\mathbb{E}_{x,y,h} \Delta_{(h,0)} T(x,y)$ is at least $\tau^2\alpha^4\beta^4\gamma^8\delta^8\rho^8$, or

$$\mathbb{E}_{x,y,h} |\mathbb{E}_k g_S(x-k,y+k) T(x-k,y+h+k)|^2 A(x) \Delta_h B(y) \Delta_h D(2x+y) \Delta_{(0,h)} \Phi(x,y) \tag{36}$$

times $\mathbb{E}_{x,y,h} \Delta_{(0,h)} T(x,y)$ is at least $\tau^2\alpha^4\beta^8\gamma^4\delta^8\rho^8$.

We have

$$\mathbb{E}_{x,y,h} \Delta_{(0,h)} T(x,y) = \alpha\beta^2\gamma^2\delta^2\rho^2 + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg)$$

and

$$\mathbb{E}_{x,y,h} \Delta_{(h,0)} T(x,y) = \alpha^2\beta\gamma^2\delta^2\rho^2 + O\bigg(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\bigg)$$

by Lemmas 3.5, 3.6, and 4.1. Expanding the square and making a change of variables, (34) equals

$$\mathbb{E}_{x,y,l} \Delta_{(0,\ell)} g_S(x,y) \mu_9(x,y,\ell),$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

where

$$\mu_9(x,y,\ell) := \mathbb{E}_{h,k}\Delta_h B(y-k)\Delta_\ell B(y+h)\Delta_h C(x+y-k)\Delta_\ell C(x+y+h)$$
$$\Delta_h D(2x+y-k)\Delta_\ell D(2x+y+h)\Phi(x,y-k)\Phi(x,y+h)$$

(35) equals

$$\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}g_S(x,y)\mu_{10}(x,y,\ell),$$

where

$$\mu_{10}(x,y,\ell) := \mathbb{E}_{h,k}B(y-k)\Delta_h C(x+y-k)\Delta_\ell C(x+y+h)$$
$$\Delta_{2h}D(2x+y-k)\Delta_\ell D(2x+y+2h)\Delta_{(h,0)}\Phi(x,y-k)\Delta_{(0,\ell)}\Phi(x+h,y)$$

and (36) equals

$$\mathbb{E}_{x,y,\ell}\Delta_{(-\ell,\ell)}g_S(x,y)\mu_{11}(x,y,\ell),$$

where

$$\mu_{11}(x,y,\ell) := \mathbb{E}_{h,k}\Delta_h B(y-k)\Delta_\ell B(y+h)C(x+y+h)$$
$$\Delta_h D(2x+y+k)\Delta_{-\ell}D(2x+y+h)\Delta_{(0,h)}\Phi(x+k,y-k)\Delta_{(-\ell,\ell)}\Phi(x,y+h).$$

Analogously to the weights $\mu_1,\dots,\mu_8$, Lemmas 3.5, 3.6, and 4.1 give the estimates

$$\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}\Phi(x,y)|\mu_9(x,y,\ell)-\beta^4\gamma^4\delta^4\rho^2|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

$$\mathbb{E}_{x,y,\ell}A(x)\Delta_\ell B(y)|\mu_{10}(x,y,\ell)-\alpha\beta\gamma^4\delta^4\rho^4|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

and

$$\mathbb{E}_{x,y,\ell}\Delta_{-\ell}A(x)|\mu_{11}(x,y,\ell)-\alpha\beta^4\gamma\delta^4\rho^4|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

from which it follows that

$$|\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}g_S(x,y)\mu_9(x,y,\ell)| = \beta^4\gamma^4\delta^4\rho^2\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}g_S(x,y)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

$$|\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}g_S(x,y)\mu_{10}(x,y,\ell)| = \alpha\beta\gamma^4\delta^4\rho^4\mathbb{E}_{x,y,\ell}\Delta_{(0,\ell)}g_S(x,y)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

and

$$|\mathbb{E}_{x,y,\ell}\Delta_{(-\ell,\ell)}g_S(x,y)\mu_{11}(x,y,\ell)| = \alpha\beta^4\gamma\delta^4\rho^4\mathbb{E}_{x,y,\ell}\Delta_{(-\ell,\ell)}g_S(x,y)+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

This completes the proof of the lemma. □

Now we are almost ready to prove our desired density-increment result for the localized degree 2 directional uniformity norms controlled by $\|\cdot\|_{\star_1}$.

Lemma 6.5. *There exist absolute constants $0<c_1<1<c_2,c_3$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A,B,C,D\subset\mathbb{F}_p^n$ have densities*



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*$\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\tau > 0$ and*

$$\varepsilon \leq c_1(\tau\alpha\beta\gamma\delta\rho)^{c_2} \exp(-(32/\tau)^{c_3}),$$

*and assume that*

$$\|A - \alpha\|_{U^8(\mathbb{F}_p^n)}, \quad \|B - \beta\|_{U^8(\mathbb{F}_p^n)}, \quad \|C - \gamma\|_{U^8(\mathbb{F}_p^n)}, \quad \|D - \delta\|_{U^8(\mathbb{F}_p^n)}, \quad \|\Phi - \alpha\rho\|_{U^4(\mathbb{F}_p^n \times \mathbb{F}_p^n)} < \varepsilon.$$

*Define $T$ by (11), and let $S \subset T$ have density $\sigma$ in $T$. Suppose that*

$$\mathbb{E}_{x,y,h,k} \Delta_{(0,h),(0,k)} g_S(x, y) \geq \tau\alpha\beta^4\gamma^4\delta^4\rho^3 \tag{37}$$

*or*

$$\mathbb{E}_{x,y,h,k} \Delta_{(h,0),(0,k)} g_S(x, y) \geq \tau\alpha^2\beta^2\gamma^4\delta^4\rho^4. \tag{38}$$

*Then $S$ has density at least $\sigma + \Omega(\tau^{O(1)})$ on some subset $T'$ of $T$ of the form*

$$T' = \{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : A'(x)B'(y)C(x + y)D(2x + y)\Phi'(x, y) = 1\},$$

*where the densities of $A', B' \subset \mathbb{F}_p^n$ are both $\Omega((\sigma\tau\alpha\beta\gamma\delta\rho)^{O(1)})$, and $\Phi'$ is of the form*

$$\Phi' = \{(x, y) \in A' \times \mathbb{F}_p^n : y \in u' + V'_x\},$$

*where each $V'_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d + 1$.*

To prove this lemma starting from the assumption (37), we will apply a localized $U^2(\Phi(x, \cdot))$-norm inverse theorem for many fixed $x$, which will produce a density increment on a set of the form

$$\{(x, y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : A'(x)B(y)C(x + y)D(2x + y)\Psi(x, y) = 1\}, \tag{39}$$

where

$$\Psi = \{(x, y) \in A' \times \mathbb{F}_p^n : y \in u'_x + V'_x\}.$$

This is not yet what the conclusion of the lemma promises, since $u'_x$ varies with $x$. To show that we can select a fixed $u'$ (at the cost of shrinking the size of $A'$ a bit) we will need Lemma 6.6.

The reader may wonder why we cannot just run the density-increment argument on sets of the form (39) and skip having to prove Lemma 6.6. The issue with this hypothetical proof is that $U^s(\mathbb{F}_p^n \times \mathbb{F}_p^n)$-uniformity of a set of the form $\Psi$ is not strong enough to guarantee that the analogue of Lemma 4.1 is true, regardless of how large $s$ is taken to be. Thus, such sets are not as amenable to a Shkredov-like pseudorandomization procedure.

LEMMA 6.6. *There exist absolute constants $0 < c_1 < 1 < c_2$ such that the following holds. Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Let $A \subset \mathbb{F}_p^n$ have density $\alpha$, $\Psi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ take the form*

$$\Psi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u_x + V_x\},$$

*where $u_x \in \mathbb{F}_p^n$ and each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$, and $K \subset A \times \mathbb{F}_p^n$ satisfy*

$$|\mathbb{E}_y K(x, y) - \kappa|, \quad |\mathbb{E}_y K(x, y)H(y) - \kappa\rho| < \varepsilon$$

*for every $x \in A$ and every affine subspace $H$ of $\mathbb{F}_p^n$ of codimension $d$. Let $\tau > 0$, assume that $\varepsilon \leq c_1(\tau\alpha\beta\gamma\delta\rho)^{c_2}$, and suppose that $S \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ has density at least $\sigma + \tau$ in $K \cap \Psi$,*



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*where $\sigma > 0$. Then there exists a $u \in \mathbb{F}_p^n$ and a subset $A' \subset A$ such that $S \cap T'$ has density at least $\sigma + \tau/4$ in $T'$, where*

$$T' := A' \times \mathbb{F}_p^n \cap K \cap \Phi',$$
$$\Phi' := \{(x, y) \in A' \times \mathbb{F}_p^n : y \in u + V_x\},$$

*and $\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(A') \geq \alpha\rho\tau/2$.*

*Proof.* Define

$$\Psi_u := \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\}$$

and

$$A_u := \{x \in A : u_x - u \in V_x\}$$

for every $u \in \mathbb{F}_p^n$, and set $\alpha_u := \mu_{\mathbb{F}_p^n}(A_u)$, so that $\mathbb{E}_u \alpha_u = \alpha\rho$. By our assumption on $K$, we have

$$\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(K) = \alpha\kappa + O(\varepsilon)$$

and

$$\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(K \cap \Psi_u) = \alpha_u\kappa\rho + O(\varepsilon)$$

for all $u \in \mathbb{F}_p^n$.

Note that

$$\mathbb{E}_u \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(S \cap K \cap \Psi_u) \geq \rho(\sigma + \tau)\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(K) = (\sigma + \tau)\alpha\kappa\rho^2 + O(\varepsilon),$$

and set

$$G(u) := \frac{\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(S \cap K \cap \Psi_u)}{\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(K \cap \Psi_u)}$$

for each $u \in \mathbb{F}_p^n$. Then we have

$$\begin{aligned}(\sigma + \tau)\alpha\kappa\rho^2 + O(\varepsilon) &\leq \kappa\rho\mathbb{E}_u\alpha_u G(u) \\ &< \frac{\kappa\rho}{p^n}\left(\left(\sigma + \frac{\tau}{4}\right) \sum_{\substack{u \in \mathbb{F}_p^n \\ G(u) < \sigma + \tau/4}} \alpha_u + \sum_{\substack{u \in \mathbb{F}_p^n \\ G(u) \geq \sigma + \tau/4}} \alpha_u\right) \\ &< \kappa\rho\left(\left(\sigma + \frac{\tau}{4}\right)(\alpha\rho - \eta) + \eta\right),\end{aligned}$$

where

$$\eta := \frac{1}{p^n} \sum_{\substack{u \in \mathbb{F}_p^n \\ G(u) \geq \sigma + \tau/4}} \alpha_u,$$

so that

$$\eta > \frac{5\tau\alpha\rho}{8}$$

when $c_1$ is small enough and $c_2$ is large enough. The contribution to $\eta$ coming from $u$ for which $\alpha_u < \tau\alpha\rho/2$ is obviously at most $\tau\alpha\rho/2$, which implies that

$$\frac{1}{p^n} \sum_{\substack{u \in \mathbb{F}_p^n \\ G(u) \geq \sigma + \tau/4 \\ \alpha_u \geq \tau\alpha\rho/2}} \alpha_u > \frac{\tau\alpha\rho}{8},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

which is clearly positive. We thus conclude that there must exist a $u \in \mathbb{F}_p^n$ for which $\alpha_u = \mu_{\mathbb{F}_p^n}(A_u) \geq \tau\alpha\rho/2$ and

$$\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(S \cap K \cap \Psi_u) \geq \left(\sigma + \frac{\tau}{4}\right)\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(K \cap \Psi_u).$$

The conclusion of the lemma now follows by taking $A' = A_u$ and $\Phi' = \Psi_u \cap (A' \times \mathbb{F}_p^n)$. □

Now we can prove Lemma 6.5.

*Proof of Lemma 6.5.* First assume that (37) holds. By writing $S = g_S + \sigma T$, we see that $\mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}S(x,y)$ equals

$$\sigma^4\mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}T(x,y) + \mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}g_S(x,y)$$

plus 14 terms of the form

$$\mathbb{E}_{x,y,h,k}\prod_{\omega\in\{0,1\}^2} g_\omega((x,y) + \omega\cdot((0,h),(0,k))), \tag{40}$$

where at least one $g_\omega$ equals $g_S$ and at least one other equals $\sigma T$, and

$$\sigma^4\mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}T(x,y) + \mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}g_S(x,y) \geq (\sigma^4+\tau)\alpha\beta^4\gamma^4\delta^4\rho^3 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

If one of the terms (40) has absolute value larger than $\tau\alpha\beta^4\gamma^4\delta^4\rho^3/32$, then combining Lemmas 6.4 and 6.3 produces the desired density increment. Thus, we may proceed under the assumption that all have size at most $\tau\alpha\beta^4\gamma^4\delta^4\rho^3/32$, so that

$$\mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}S(x,y) \geq \left(\sigma^4 + \frac{\tau}{2}\right)\alpha\beta^4\gamma^4\delta^4\rho^3.$$

For each $x \in \mathbb{F}_p^n$, set $\Phi_x(y) := \Phi(x,y)$, $T_x(y) := T(x,y)$, and $S_x(y) = S(x,y)$. Lemmas 3.5, 3.6, and 4.1 tell us that

$$\mathbb{E}_x A(x)|\mathbb{E}_{y,h,k}\Delta_{h,k}T_x(y) - \beta^4\gamma^4\delta^4\rho^3| \ll \varepsilon^{\Omega(1)},$$
$$\mathbb{E}_x A(x)\|T_x - \beta\gamma\delta\|_{U^2(\Phi_x)}^4 \ll \varepsilon^{\Omega(1)},$$

and that

$$\mathbb{E}_x A(x)|\mathbb{E}_{y\in\Phi_x}S_x(y) - \sigma\beta\gamma\delta|^2 < \frac{\tau^4}{64}\alpha\beta^2\gamma^2\delta^2\rho^2,$$

or else Lemma 6.3 will again give the desired density increment. It follows that there exists a subset $A_0 \subset A$ of density $\gg \tau$ in $A$ such that

$$\|S_x\|_{U^2(\Phi_x)}^4 \geq \left(\sigma^4 + \frac{\tau}{4}\right)\beta^4\gamma^4\delta^4, \quad \|T_x - \beta\gamma\delta\|_{U^2(\Phi_x)} \ll \varepsilon^{\Omega(1)},$$

and

$$|\mathbb{E}_{y\in\Phi_x}S_x(y) - \sigma\beta\gamma\delta| < \frac{\tau\beta\gamma\delta}{64}$$

for every $x \in A_0$. Setting

$$f_x(y) := \frac{1}{\beta\gamma\delta}S_x(y) \quad \text{and} \quad \nu_x := \frac{1}{\beta\gamma\delta}T_x,$$

for all $x \in A_0$ we then have $\|f_x\|_{U^2(\Phi_x)}^4 \geq \sigma^4 + \tau/4$, $0 \leq f_x \leq \nu_x$, $\mathbb{E}_y f_x(y) \leq 1$, and

$$\|\nu_x - 1\|_{U^2(\Phi_x)} \ll \frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}} \leq \exp(-(32/\tau)^{c_3}),$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

provided that $c_1$ is small enough and $c_2$ is large enough. Thus, as long as $c_3$ is sufficiently large, Lemma 6.1 tells us that there exists a function $\tilde{f}_x : \Phi_x \to [0,1]$ such that $\mathbb{E}_{y\in\Phi_x} f_x(y) = \mathbb{E}_{y\in\Phi_x}\tilde{f}_x(y)$ and $\|f_x - \tilde{f}_x\|_{U^2(\Phi_x)}^4 \le \tau/32$. As a consequence, since $\|f_x\|_{U^2(\Phi_x)}^4 \ge \sigma^4 + \tau/4$, we must have

$$\|\tilde{f}_x\|_{U^2(\Phi_x)}^4 \ge \sigma^4 + \frac{\tau}{8}$$

as well. Set $\tilde{\sigma}_x := \mathbb{E}_y f_x(y) = \mathbb{E}_y \tilde{f}_x(y)$ and let $v_x \in \Phi_x$, so that

$$\|\tilde{f}_x\|_{U^2(\Phi_x)}^4 = \tilde{\sigma}_x^4 + \sum_{0\neq\xi\in\widehat{\Phi_x - v_x}} |\widehat{(\tilde{f}_x - \tilde{\sigma})}(\xi)|^4.$$

Since $|\tilde{\sigma}_x - \sigma| < \tau/64$, it follows that there exists a nonzero $\xi_x \in \widehat{\Phi_x - v_x}$ such that

$$|\mathbb{E}_{y\in\Phi_x}(\tilde{f}_x - \sigma)(y) e_p(\xi_x \cdot y)| \ge \frac{\tau}{16},$$

where we have crucially used that $\tilde{f}_x$ is 1-bounded. As $\|f_x - \tilde{f}_x\|_{U^2(\Phi_x)}^4 \le \tau/32$, it therefore follows that

$$|\mathbb{E}_{y\in\Phi_x}(f_x - \sigma)(y) e_p(\xi_x \cdot y)| \ge \frac{\tau}{32}$$

for every $x \in A_0$.

Extend $x \mapsto \xi_x$ from $A_0$ to $A$ by picking a nonzero $\xi_x \in \widehat{\Phi_x - v_x}$ arbitrarily for all $x \in A \setminus A_0$. We now split the average over $y \in \Phi_x$ above into an average of averages over cosets of $\langle \xi_x \rangle^\perp$ in $\Phi_x$ and average over all of $A$ to get that

$$\mathbb{E}_{x\in A}\mathbb{E}_{t\in\mathbb{F}_p}\left|\mathbb{E}_{\substack{y\in\Phi_x \\ \xi_x\cdot y = t}}(f_x - \sigma)(y)\right| \gg \tau,$$

and use the fact that

$$\mathbb{E}_{x\in A}\mathbb{E}_{t\in\mathbb{F}_p}\mathbb{E}_{\substack{y\in\Phi_x \\ \xi_x\cdot y = t}}(f_x - \sigma)(y) = \mathbb{E}_{x\in A}\mathbb{E}_{y\in\Phi_x}(f_x - \sigma)(y) = 0$$

to deduce that

$$\mathbb{E}_{x\in A}\mathbb{E}_{t\in\mathbb{F}_p}\max\left(0, \mathbb{E}_{\substack{y\in\Phi_x \\ \xi_x\cdot y = t}}(f_x - \sigma)(y)\right) \gg \tau.$$

By applying the pigeonhole principle in the $x$ and $t$ variables, it follows that there exists a subset $A_1 \subset A$ of density $\gg \tau$ in $A$ and, for each $x \in A_1$, an element $t_x \in \mathbb{F}_p$ for which

$$\mathbb{E}_{\substack{y\in\Phi_x \\ \xi_x\cdot y = t_x}}(f_x - \sigma)(y) \gg \tau.$$

Thus, recalling the definition of $f_x$, we have

$$\mathbb{E}_{x\in A_1}\mathbb{E}_{\substack{y\in\Phi_x \\ \xi_x\cdot y = t_x}} S(x,y) \ge (\sigma + \Omega(\tau))\beta\gamma\delta. \tag{41}$$

Define $\phi : \mathbb{F}_p^n \to \mathbb{F}_p^n$ by taking $\phi(x) = \xi_x$ for all $x \in A$ and $\phi(x)$ to be an arbitrary element of $(\Phi_x - v_x)^\perp \setminus \{0\}$ for all $x \in \mathbb{F}_p^n \setminus A$, and similarly extend $x \mapsto t_x$ from $A_1$ to $A$ by taking $t_x$ to be an arbitrary element of $\mathbb{F}_p$ for which

$$\mathbb{E}_{\substack{y\in\Phi_x \\ \phi(x)\cdot y = t_x}} S(x,y) \ge \mathbb{E}_{y\in\Phi_x} S(x,y)$$

for all $x \in A \setminus A_1$. Such an element must exist by the pigeonhole principle. Set

$$\Psi := \{(x,y) \in \Phi : \phi(x)\cdot y = t_x\},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

so that $\operatorname{codim}\{y \in \mathbb{F}_p^n : \Psi(x,y) = 1\} = d+1$ for every $x \in A$, and $\alpha_1 := |A_1|/p^n$. Then (41) can be rewritten as

$$\mathbb{E}_{x,y} S(x,y)A_1(x)\Psi(x,y) \geq (\sigma + \Omega(\tau))\frac{\alpha_1\beta\gamma\delta\rho}{p}.$$

It remains to check that the density $\mathbb{E}_{x,y}A_1(x)B(y)C(x+y)D(2x+y)\Psi(x,y)$ is close to $\alpha_1\beta\gamma\delta\rho/p$, so that we indeed have the desired density increment. But by Lemmas 3.5 and 3.6, we have

$$\mathbb{E}_y B(y)C(x+y)D(2x+y)\Psi(x,y) = \frac{\beta\gamma\delta\rho}{p} + O(\varepsilon^{1/8})$$

for all but a $O(\sqrt{\varepsilon})$-proportion of $x \in A$, from which it follows that

$$\mathbb{E}_{x,y}A_1(x)B(y)C(x+y)D(2x+y)\Psi(x,y) = \frac{\alpha_1\beta\gamma\delta\rho}{p} + O(\varepsilon^{1/8}).$$

Thus, $S$ has density at least $\sigma + \Omega(\tau)$ on

$$Q := \{(x,y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : A_1(x)B(y)C(x+y)D(2x+y)\Psi(x,y) = 1\},$$

provided that $c_2$ is large enough. The conclusion of the lemma now follows from Lemma 6.6

Now suppose that (38) holds. By writing $S = g_S + \sigma T$ and arguing as in the first case, we may proceed under the assumption that

$$\mathbb{E}_{x,x',y,y'} S(x,y)S(x,y')S(x',y)S(x',y') \geq \left(\sigma^4 + \frac{\tau}{2}\right)\alpha^2\beta^2\gamma^4\delta^4\rho^4.$$

We will first show that either

$$\mathbb{E}_{x,y} S(x,y')S(x',y)C(x+y)D(2x+y)\Phi(x,y) = (\sigma^2 + O(\tau^2))\alpha\beta\gamma^3\delta^3\rho^3$$

for almost every pair $(x',y') \in S$, or else we can deduce the desired density increment using Lemma 6.3.

Consider the average

$$\mathbb{E}_{x,x',y,y'} S(x,y')S(x',y)S(x',y')C(x+y)D(2x+y)\Phi(x,y).$$

Using that $S = g_S + \sigma T$, the above can be written as

$$\sigma^3\mathbb{E}_{x,x',y,y'} T(x,y')T(x',y)T(x',y')C(x+y)D(2x+y)\Phi(x,y) \tag{42}$$

plus seven other terms of the form

$$\mathbb{E}_{x,x',y,y'} g_0(x,y')g_1(x',y)g_2(x',y')C(x+y)D(2x+y)\Phi(x,y), \tag{43}$$

where $g_0, g_1$, and $g_2$ all equal $g_S$ or $\sigma T$ and at least one $g_i$ equals $g_S$. By Lemmas 3.5, 3.6, and 4.1, the quantity (42) equals $\sigma^3\alpha^2\beta^2\gamma^4\delta^4\rho^4 + O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$. Suppose that $k$ of the functions $g_0, g_1$, and $g_2$ in (43) equal $\sigma T$. By a similar argument to those used to prove Lemma 6.4, if any term of the form (43) has size at least $\tau^4\sigma^k\alpha^2\beta^2\gamma^4\delta^4\rho^4$, then

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \geq \tau^8\alpha\beta^2\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)$$

or

$$\mathbb{E}_{x,y,h}\Delta_{(h,0)}g_S(x,y) \geq \tau^8\alpha^2\beta\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

so that the desired density increment follows from Lemma 6.3. Thus, we may proceed under the assumption that

$$\mathbb{E}_{(x',y')\in S}\mathbb{E}_{x,y} S(x,y')S(x',y)C(x+y)D(2x+y)\Phi(x,y) = \sigma^2\alpha\beta\gamma^3\delta^3\rho^3 + O(\tau^4\alpha\beta\gamma^3\delta^3\rho^3).$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Now consider the average

$$\mathbb{E}_{x',y'} S(x',y') |\mathbb{E}_{x,y} S(x,y') S(x',y) C(x+y) D(2x+y) \Phi(x,y)|^2.$$

Using that $S = g_S + \sigma T$, the above can be written as

$$\sigma^5 \mathbb{E}_{x',y'} T(x',y') |\mathbb{E}_{x,y} T(x,y') T(x',y) C(x+y) D(2x+y) \Phi(x,y)|^2, \tag{44}$$

plus 31 other terms of the form

$$\begin{aligned}\mathbb{E}_{x,y,x',y',w,z} g_0(x',y') g_1(x,y') g_2(x',y) g_3(z,y') g_4(x',w) C(x+y) C(z+w) \\ \cdot D(2x+y) D(2z+w) \Phi(x,y) \Phi(z,w),\end{aligned} \tag{45}$$

where $g_0, g_1, g_2, g_3$, and $g_4$ all equal $g_S$ or $\sigma T$ and at least one $g_i$ equals $g_S$.

By Lemmas 3.5, 3.6, and 4.1, the quantity (44) equals $\sigma^5 \alpha^3 \beta^3 \gamma^7 \delta^7 \rho^7 + O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$. Suppose that $k$ of the functions $g_0, \dots, g_4$ in (45) equal $\sigma T$. Analogously to the situation for the first moment, if any of the terms of the form (45) has size at least $\tau^8 \sigma^k \alpha^3 \beta^3 \gamma^7 \delta^7 \rho^7$, then we will be able to deduce the desired density increment. The most involved case is when $g_0 = \cdots = g_4 = g_S$. All other cases can be handled using a simpler version of the argument we are about to carry out.

Thus, consider this most involved case, i.e. that

$$\begin{aligned}\mathbb{E}_{x,y,x',y',w,z} g_S(x',y') g_S(x,y') g_S(x',y) g_S(z,y') g_S(x',w) C(x+y) C(z+w) \\ \cdot D(2x+y) D(2z+w) \Phi(x,y) \Phi(z,w)\end{aligned}$$

has size at least $\tau^8 \alpha^3 \beta^3 \gamma^7 \delta^7 \rho^7$. Applying the Cauchy–Schwarz inequality in the variables $x', y', y$, and $w$ gives that

$$\mathbb{E}_{x',y',y,w} T(x',y') T(x',y) T(x',w)$$

times

$$\begin{aligned}\mathbb{E}_{x',y',y,w} (|\mathbb{E}_{x,z} g_S(x,y') g_S(z,y') C(x+y) C(z+w) D(2x+w) D(2z+w) \Phi(x,y) \Phi(z,w)|^2 \\ \cdot B(y) B(w) C(x'+y') C(x'+y) C(x'+w) \\ \cdot D(2x'+y') D(2x'+y) D(2x'+w) \Phi(x',y') \Phi(x',y) \Phi(x',w))\end{aligned}$$

is at least $\tau^8 \alpha^6 \beta^6 \gamma^{14} \delta^{14} \rho^{14}$. By Lemmas 3.5, 3.6, and 4.1, $\mathbb{E}_{x',y',y,w} T(x',y') T(x',y) T(x',w) = \alpha \beta^3 \gamma^3 \delta^3 \rho^3 + O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$. Expanding the square in the average above, this means that

$$\begin{aligned}\mathbb{E}_{x',y',y,w,x,z,u,v} (g_S(x,y') g_S(z,y') g_S(u,y') g_S(v,y') \\ B(y) B(w) C(x'+y') C(x'+y) C(x'+w) \\ C(x+y) C(z+w) C(u+y) C(v+w) \\ D(2x+w) D(2z+w) D(2u+w) D(2v+w) \\ D(2x'+y') D(2x'+y) D(2x'+w) \\ \Phi(x,y) \Phi(z,w) \Phi(u,y) \Phi(v,w) \Phi(x',y') \Phi(x',y) \Phi(x',w))\end{aligned}$$

is $\gg \tau^8 \alpha^5 \beta^3 \gamma^{11} \delta^{11} \rho^{11}$, provided that $c_1$ is small enough and $c_2$ is large enough. We can write the above as

$$\mathbb{E}_{y',x,z,u,v} g_S(x,y') g_S(z,y') g_S(u,y') g_S(v,y') \mu(y',x,z,u,v),$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

where

$$\begin{aligned}\mu(y',x,z,u,v) := \mathbb{E}_{x',y,w}(&B(y)B(w)C(x'+y')C(x'+y)C(x'+w)\\ &C(x+y)C(z+w)C(u+y)C(v+w)\\ &D(2x+w)D(2z+w)D(2u+w)D(2v+w)\\ &D(2x'+y')D(2x'+y)D(2x'+w)\\ &\Phi(x,y)\Phi(z,w)\Phi(u,y)\Phi(v,w)\Phi(x',y')\Phi(x',y)\Phi(x',w)).\end{aligned}$$

Yet more applications of Lemmas 3.5, 3.6, and 4.1 give the estimate

$$\mathbb{E}_{y',x,z,u,v}A(x)A(z)A(u)A(w)|\mu(y',x,z,u,v)-\alpha\beta^2\gamma^7\delta^7\rho^7|^2 \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

so that

$$\mathbb{E}_{y',x,z,u,v}g_S(x,y')g_S(z,y')g_S(u,y')g_S(v,y') \gg \tau^8\alpha^4\beta\gamma^4\delta^4\rho^4.$$

It follows from one more application of the Cauchy–Schwarz inequality in the variables $y', z, u$, and $v$ and a similar analysis to above that

$$\mathbb{E}_{x,y,h}\Delta_{(h,0)}g_S(x,y') \gg \tau^{16}\alpha^2\beta\gamma^2\delta^2\rho^2,$$

which, combined with Lemma 6.3, gives the desired density increment.

Thus, we may also proceed under the assumption that

$$\mathbb{E}_{(x',y')\in S}|\mathbb{E}_{x,y}S(x,y')S(x',y)C(x+y)D(2x+y)\Phi(x,y)-\sigma^2\alpha\beta\gamma^3\delta^3\rho^3|^2 \ll \tau^4\alpha^2\beta^2\gamma^6\delta^6\rho^6.$$

By Markov's inequality, we therefore have

$$\mathbb{E}_{x,y}S(x,y')S(x',y)C(x+y)D(2x+y)\Phi(x,y) = (\sigma^2+O(\tau^2))\alpha\beta\gamma^3\delta^3\rho^3 \tag{46}$$

for all but a $O(\tau^2)$-proportion of $(x',y') \in S$. As a consequence, there exists a pair $(x',y') \in S$ for which both (46) holds and

$$\mathbb{E}_{x,y}S(x,y)S(x,y')S(x',y) \geq \left(\sigma^3+\frac{\tau}{4}\right)\alpha\beta\gamma^3\delta^3\rho^3,$$

which together imply that

$$\frac{|S\cap T'|}{|T'|} \geq \sigma+\Omega(\tau)$$

when we take $A'(x) = S(x,y')$, $B'(y) = S(x',y)$, $C' = C$, $D' = D$, and $\Phi' = \Phi$ in the definition of $T'$. □

### 6.3 More preliminaries for $\|\cdot\|_{\star_1}$

We will similarly need that certain $\|\cdot\|_{\star_1}$-inner products are controlled by the degree 1 and 2 directional uniformity norms appearing in the previous subsections. The proof of the lemma below is similar to the proof of Lemma 6.4, but with an extra application of the Cauchy–Schwarz inequality in some cases.

Lemma 6.7. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\varepsilon > 0$, and assume that*

$$\|A-\alpha\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|B-\beta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|C-\gamma\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|D-\delta\|_{U^{10}(\mathbb{F}_p^n)}, \quad \|\Phi-\alpha\rho\|_{U^8(\mathbb{F}_p^n\times\mathbb{F}_p^n)} < \varepsilon.$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*Define $T$ by (11), and let $S \subset T$ and $\tau > 0$. If*

$$\left|\mathbb{E}_{x,y,h_1,h_2,h_3} \prod_{\omega\in\{0,1\}^3} g_\omega((x,y)+\omega\cdot((0,h_1),(0,h_2),(h_3,0)))\right| \geq \tau\alpha^2\beta^4\gamma^8\delta^8\rho^6,$$

*where at least one $g_\omega$ equals $T$ and another equals $g_S$, then*

$$\mathbb{E}_{x,y,h}\Delta_{(0,h)}g_S(x,y) \geq \tau^8\alpha\beta^2\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right),$$

$$\mathbb{E}_{x,y,h}\Delta_{(h,0)}g_S(x,y) \geq \tau^8\alpha^2\beta\gamma^2\delta^2\rho^2 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right),$$

$$\mathbb{E}_{x,y,h,k}\Delta_{(0,h),(0,k)}g_S(x,y) \geq \tau^8\alpha\beta^4\gamma^4\delta^4\rho^3 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right),$$

*or*

$$\mathbb{E}_{x,y,h,k}\Delta_{(h,0),(0,k)}g_S(x,y) \geq \tau^8\alpha^2\beta^2\gamma^4\delta^4\rho^4 + O\left(\frac{\varepsilon^{\Omega(1)}}{(\alpha\beta\gamma\delta\rho)^{O(1)}}\right).$$

Our final preliminary lemma says that, for almost every $(x, x+h_3) \in A^2$, the function $\Delta_{(h_3,0)}S(x,\cdot)$ is supported on a Fourier uniform subset of the affine subspace $\{y \in \mathbb{F}_p^n : \Delta_{(h_3,0)}\Phi(x,y) = 1\}$.

LEMMA 6.8. *Let $d$ be a nonnegative integer, and set $\rho := p^{-d}$. Suppose that $A, B, C, D \subset \mathbb{F}_p^n$ have densities $\alpha, \beta, \gamma, \delta$, respectively, and that $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ takes the form*

$$\Phi = \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension $d$. Let $\varepsilon > 0$, and assume that*

$$\|A-\alpha\|_{U^4(\mathbb{F}_p^n)}, \quad \|B-\beta\|_{U^4(\mathbb{F}_p^n)}, \quad \|C-\gamma\|_{U^4(\mathbb{F}_p^n)}, \quad \|D-\delta\|_{U^4(\mathbb{F}_p^n)}, \quad \|\Phi-\alpha\rho\|_{U^4(\mathbb{F}_p^n\times\mathbb{F}_p^n)} < \varepsilon.$$

*Setting*

$$R_{x,h}(y) := B(y)\Delta_h C(x+y)\Delta_{2h}D(2x+y)$$

*and*

$$\Phi_{x,h} := \Delta_{(h,0)}\Phi,$$

*then the probability*

$$\mathbb{P}\left((x,x+h) \in A^2 : \operatorname{codim}\{y \in \mathbb{F}_p^n : \Phi_{x,h}(y) = 1\} \neq 2d \text{ or } \|R_{x,h}\Phi_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h})} \geq \frac{\varepsilon^{1/32}}{\rho^{3/2}}\right)$$

*is $\ll \varepsilon^{\Omega(1)}/\rho^{O(1)}$.*

*Proof.* By Lemmas 3.6 and 4.1, we have that $\|R_{x,y} - \beta\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)} \geq \varepsilon^{1/8}$ or $\operatorname{codim}\{y \in \mathbb{F}_p^n : \Phi_{x,h}(y) = 1\} \neq 2d$ for at most a $O(\varepsilon^{\Omega(1)}/\rho^{O(1)})$-proportion of pairs $(x, x+h) \in A^2$. For all of these typical pairs $(x,h)$, we have

$$\|R_{x,h}\Phi_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h})}^4 = \rho^{-6}\mathbb{E}_{y,k,\ell}\Delta_{k,\ell}[R_{x,h} - \beta\gamma^2\delta^2](y)\Phi_{x,h}(y)\Phi_{x,h}(y+k)\Phi_{x,h}(y+\ell),$$

which is at most

$$\sum_{\xi,\eta,\nu\in(\Phi_{x,h}-u)^\perp} |\mathbb{E}_{y,k,\ell}[R_{x,h} - \beta\gamma^2\delta^2](y)e_p(\xi\cdot y)[R_{x,h} - \beta\gamma^2\delta^2](y+k)e_p(\eta\cdot(y+k))$$
$$[R_{x,h} - \beta\gamma^2\delta^2](y+\ell)e_p(\nu\cdot(y+\ell))[R_{x,h} - \beta\gamma^2\delta^2](y+k+\ell)|$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

by inserting the identity

$$\Phi_{x,h}(z) = \frac{1}{p^{2d}} \sum_{\xi \in (\Phi_{x,h}-u)^{\perp}} e_p(\xi \cdot [z-u])$$

for every $z \in \mathbb{F}_p^n$. For each fixed triple $(\xi, \eta, \nu)$, the interior average above is bounded by $\|R_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)}$ by the Gowers–Cauchy–Schwarz inequality. Since there are $\rho^{-6}$ possible triples to sum over, we must have

$$\|R_{x,h}\Phi_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h})}^4 \leq \rho^{-6}\|R_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)} < \varepsilon^{1/8}/\rho^6,$$

from which the conclusion of the lemma follows. □

### 6.4 Proof of Theorem 2.5

Now we can finally finish the proof of Theorem 2.5.

*Proof of Theorem 2.5.* First assume that

$$\|g_S\|_{\star_1} \geq \tau\alpha^{1/4}\beta^{1/2}\gamma\delta\rho^{3/4}.$$

Analogously to the proof of Lemma 6.5, by writing $S = g_S + \sigma T$, we see that $\|S\|_{\star_1}^8$ equals

$$\sigma^8\|T\|_{\star_1}^8 + \|g_S\|_{\star_1}^8$$

plus 62 terms of the form

$$\mathbb{E}_{x,y,h_1,h_2,h_3} \prod_{\omega \in \{0,1\}^3} g_\omega((x,y) + \omega \cdot ((0,h_1),(0,h_2),(h_3,0))), \tag{47}$$

where at least one $g_\omega$ equals $\sigma T$ and at least one other equals $g_S$. Note that

$$\|T\|_{\star_1}^8 = \alpha^2\beta^4\gamma^8\delta^8\rho^6 + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right)$$

by Lemmas 3.5, 3.6, and 4.1. If any of the terms (47) have absolute value at least $\frac{1}{128}\tau^8\alpha^2\beta^4\gamma^8\delta^8\rho^6$, then combining Lemma 6.7 with Lemma 6.3 or Lemma 6.5 produces the desired density increment. We may thus proceed under the assumption that these terms are all small, so that

$$\mathbb{E}_{x,y,h_1,h_2,h_3}\Delta_{(0,h_1),(0,h_2),(h_3,0)}S(x,y) \geq \left(\sigma^8 + \frac{\tau^8}{2}\right)\alpha^2\beta^4\gamma^8\delta^8\rho^6.$$

For each pair $(x, x+h) \in A^2$, let $R_{x,h}$ and $\Phi_{x,h}$ be as in Lemma 6.8, and set $S_{x,h}(y) := \Delta_{(h,0)}S(x,y)$. By Markov's inequality, either

$$\mathbb{P}\left((x,x+h) \in A \times A : |\mathbb{E}_y S_{x,h}(y) - \sigma^2\beta\gamma^2\delta^2\rho^2| \geq \frac{\tau^8}{64}\beta\gamma^2\delta^2\rho^2\right) < \frac{\tau^8}{4}, \tag{48}$$

or else $\mathbb{E}_{x,x+h\in A}|\mathbb{E}_y S_{x,h}(y) - \sigma^2\beta\gamma^2\delta^2\rho^2|^2$ is $\gg \tau^{16}\beta^2\gamma^4\delta^4\rho^4$, in which case a combination of Lemmas 6.3, 6.4, and 6.5 will produce the desired density increment. We may thus proceed under the assumption that (48) holds.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

By (48) and Lemma 6.8, there exists a subset $A_0 \subset \{(x,h) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : x, x+h \in A\}$ of density $\gg \tau^{O(1)}\alpha^2$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ such that

$$\|S_{x,h}\|_{U^2(\Phi_{x,h})}^4 \geq \left(\sigma^8 + \frac{\tau^8}{4}\right)\beta^4\gamma^8\delta^8,$$

$$|\mathbb{E}_y S_{x,h}(y) - \sigma^2\beta\gamma^2\delta^2\rho^2| < \frac{\tau^8}{64}\beta\gamma^2\delta^2\rho^2,$$

$$\|R_{x,h}\Phi_{x,h} - \beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h})} < \frac{\varepsilon^{1/32}}{\rho^{3/2}},$$

and $\operatorname{codim}\{y \in \mathbb{F}_p^n : \Phi_{x,h}(y) = 1\} = 2d$ for all $(x,h) \in A_0$. Note that $S_{x,h}$ is supported on $R_{x,h}\Phi_{x,h}$, and, setting

$$f_{x,h} := \frac{1}{\beta\gamma^2\delta^2}S_{x,h} \quad \text{and} \quad \nu_{x,h} := \frac{1}{\beta\gamma^2\delta^2}R_{x,h}\Phi_{x,h},$$

we have $\|f_{x,h}\|_{U^2(\Phi_{x,h})}^4 \geq \sigma^8 + \tau^8/4$, $0 \leq f_{x,h} \leq \nu_{x,h}$, $\mathbb{E}f_{x,h} \leq 1$, $|\mathbb{E}f_{x,h} - \sigma^2| < \tau^8/64$, and $\|\nu_{x,h} - 1\|_{U^2(\Phi_{x,h})} < \varepsilon^{1/32}/\beta\gamma^2\delta^2\rho^{3/2} < \exp(-(64/\tau)^{c_3})$, provided that $c_1$ is small enough and $c_2$ is large enough. Applying Lemma 6.1 on the affine subspace $\Phi_{x,h}$ yields a function $\tilde{f}_{x,h} : \Phi_{x,h} \to [0,1]$ such that

$$\mathbb{E}\tilde{f}_{x,h} = \mathbb{E}f_{x,h} \quad \text{and} \quad \|\tilde{f}_{x,h} - f_{x,h}\|_{U^2(\Phi_{x,h})} \leq \frac{\tau^8}{64},$$

provided that $c_3$ is large enough. We must then also have

$$\|\tilde{f}_{x,h}\|_{U^2(\Phi_{x,h})}^4 \geq \sigma^8 + \frac{\tau^8}{16}.$$

Arguing as in the proof of the first part of Lemma 6.5, it follows that, for every pair $(x,h) \in A_0$, there exists a nonzero $\xi_{x,h} \in \widehat{\Phi_{x,h} - u}$ such that

$$|\mathbb{E}_{y \in \Phi_{x,h}}(\tilde{f}_{x,h} - \sigma^2)(y)e_p(\xi_{x,h} \cdot y)| \geq \frac{\tau^8}{32},$$

so that

$$|\mathbb{E}_{y \in \Phi_{x,h}}(f_{x,h} - \sigma^2)(y)e_p(\xi_{x,h} \cdot y)| \geq \frac{\tau^8}{64}$$

as well.

Extend $(x,h) \mapsto \xi_{x,h}$ from $A_0$ to the set $\{(x,h) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : x, x+h \in A\}$ by picking a nonzero $\xi_{x,h} \in \widehat{\Phi_{x,h} - u}$ arbitrarily for all pairs outside of $A_0$. We split the average over $y \in \Phi_{x,h}$ up into an average of averages over cosets of $\langle \xi_{x,h} \rangle^\perp$ and average over all pairs $(x,h)$ such that $x, x+h \in A$ to get that

$$\mathbb{E}_{x,x+h \in A}\mathbb{E}_{t \in \mathbb{F}_p}\left|\mathbb{E}_{\substack{y \in \Phi_{x,h} \\ \xi_{x,h} \cdot y = t}}(f_{x,h} - \sigma^2)(y)\right| \gg \tau^c \tag{49}$$

for some absolute constant $c > 0$. Note that

$$\begin{aligned}\mathbb{E}_{x,x+h \in A}\mathbb{E}_{t \in \mathbb{F}_p}\mathbb{E}_{\substack{y \in \Phi_{x,h} \\ \xi_{x,h} \cdot y = t}}(f_{x,h} - \sigma^2)(y) &= \frac{1}{\alpha^2\beta\gamma^2\delta^2\rho^2}\mathbb{E}_{x,y,h}S(x,y)S(x+h,y) - \sigma^2 \\ &= \frac{1}{\alpha^2\beta\gamma^2\delta^2\rho^2}\mathbb{E}_{x,y,h}g_S(x,y)g_S(x+h,y),\end{aligned}$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

so that, if it were the case that

$$\left|\mathbb{E}_{x,x+h\in A}\mathbb{E}_{t\in\mathbb{F}_p}\mathbb{E}_{\substack{y\in\Phi_{x,h}\\ \xi_{x,h}\cdot y=t}}(f_{x,h}-\sigma^2)(y)\right| \gg \tau^{2c},$$

then we would be able to deduce the desired density increment from Lemma 6.3. Thus, we may proceed under the assumption

$$\left|\mathbb{E}_{x,x+h\in A}\mathbb{E}_{t\in\mathbb{F}_p}\mathbb{E}_{\substack{y\in\Phi_{x,h}\\ \xi_{x,h}\cdot y=t}}(f_{x,h}-\sigma^2)(y)\right| \ll \tau^{2c},$$

which we can combine with (49), a change of variables, and an application of the pigeonhole principle in the $t$ variable to deduce that

$$\mathbb{E}_{x,h\in A}\max\left(0,\mathbb{E}_{\substack{y\in\Phi_{x,h-x}\\ \xi_{x,h-x}\cdot y=t}}(f_{x,h-x}-\sigma^2)(y)\right) \gg \tau^c \tag{50}$$

for some fixed $t\in\mathbb{F}_p$. Set

$$\Psi_h := \{(x,y)\in\mathbb{F}_p^n\times\mathbb{F}_p^n : y\in\Phi_{x,h-x} \text{ and } \xi_{x,h-x}\cdot y = t\}. \tag{51}$$

To deduce the desired density increment by applying the pigeonhole principle to (50), we will have to show that almost every set of the form

$$\{(x,y)\in T : S(h,y)A_1(x,h-x)\Psi_h(x,y)=1\}$$

has close to the 'correct' density. Most of the remainder of our argument for $\|\cdot\|_{\star_1}$ is devoted to this task.

Set $Q_{x,h}(y) := S(h,y)C(x+y)D(2x+y)$. We start by showing that either $\|Q_{x,h}-\sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})}$ is small for almost every pair $(x,h)\in A\times A$, or else we can deduce a density increment from Lemmas 6.3 and 6.5. Note first that

$$\begin{aligned}\mathbb{E}_{x,y,h}A(x)A(h)\Phi_{x,h-x}(y)Q_{x,h}(y) &= \mathbb{E}_{x,y,h}S(h,y)A(x)C(x+y)D(2x+y)\Phi(x,y)\\ &= \mathbb{E}_{h,y}S(h,y)\mu(y),\end{aligned}$$

where $\mu(y) := \mathbb{E}_x A(x)C(x+y)D(2x+y)\Phi(x,y)$. By Lemmas 3.5, 3.6, and 4.1, we have

$$\mathbb{P}(y\in\mathbb{F}_p^n : |\mu(y)-\alpha\gamma\delta\rho|>\varepsilon) \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$

so that $|\mathbb{E}_{x,h,y}A(x)A(h)\Phi_{x,h-x}(y)Q_{x,h}(y)-\sigma\alpha^2\beta\gamma^2\delta^2\rho^2| < \varepsilon^{\Omega(1)}/\rho^{O(1)}$. Similarly, the average $\mathbb{E}_{x,h}A(x)A(h)|\mathbb{E}_y Q_{x,h}(y)\Phi_{x,h-x}(y)-\sigma\beta\gamma^2\delta^2\rho^2|^2$ equals

$$\mathbb{E}_{x,y,h,k}A(x)A(h)\Phi_{x,h-x}(y)\Phi_{x,h-x}(y+k)Q_{x,h}(y)Q_{x,h}(y+k)-\sigma^2\alpha^2\beta^2\gamma^4\delta^4\rho^4+O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right),$$

and $\mathbb{E}_{x,y,h,k}A(x)A(h)\Phi_{x,h-x}(y)\Phi_{x,h-x}(y+k)Q_{x,h}(y)Q_{x,h}(y+k)$ equals

$$\mathbb{E}_{y,h,k}S(h,y)S(h,y+k)\mu'(y,k),$$

where

$$\mu'(y,k) = \mathbb{E}_x A(x)C(x+y)C(x+y+k)D(2x+y)D(2x+y+k)\Phi(x,y)\Phi(x,y+k).$$

By Lemmas 3.5, 3.6, and 4.1, we have

$$\mathbb{P}((y,k)\in\mathbb{F}_p^n\times\mathbb{F}_p^n : |\mu'(y,k)-\alpha\gamma^2\delta^2\rho^2|>\varepsilon) \ll \frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}},$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

so that

$$\mathbb{E}_{x,y,h,k} A(x)A(h)\Phi_{x,h-x}(y)\Phi_{x,h-x}(y+k)Q_{x,h}(y)Q_{x,h}(y+k)$$

equals

$$\alpha\gamma^2\delta^2\rho^2\mathbb{E}_{y,h,k}S(h,y)S(h,y+k) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

Thus, if

$$\mathbb{E}_{x,h}A(x)A(h)|\mathbb{E}_y Q_{x,h}(y)\Phi_{x,h-x}(y) - \sigma\beta\gamma^2\delta^2\rho^2|^2 \gg \tau^{4c'}\sigma^2\alpha^2\beta^2\gamma^4\delta^4\rho^4$$

for some $c' > 0$ (to be chosen later), and $c_1$ is small enough and $c_2$ is large enough, then

$$\mathbb{E}_{x,y,h}S(x,y)S(x,y+h) \geq (\sigma^2 + \Omega(\tau^{4c'}))\alpha\beta^2\gamma^2\delta^2\rho^2,$$

in which case the desired density increment follows from Lemma 6.3. We may thus proceed under the assumption that

$$\mathbb{E}_{x,h}A(x)A(h)|\mathbb{E}_y Q_{x,h}(y)\Phi_{x,h-x}(y) - \sigma\beta\gamma^2\delta^2\rho^2|^2 \ll \tau^{4c'}\sigma^2\alpha^2\beta^2\gamma^4\delta^4\rho^4,$$

so that, by Markov's inequality, we have

$$\mathbb{P}((x,h) \in A \times A : |\mathbb{E}_y Q_{x,h}(y)\Phi_{x,h-x}(y) - \sigma\beta\gamma^2\delta^2\rho^2| \geq \tau^{2c'}\sigma\beta\gamma^2\delta^2\rho^2) \ll \tau^{2c'}.$$

It follows that $\mathbb{E}_{x,h}A(x)A(h)\|Q_{x,h} - \sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})}^4$ equals

$$\mathbb{E}_{\substack{x,h\in\mathbb{F}_p^n \\ y,y+k,y+\ell\in\Phi_{x,h-x}}} A(x)A(h)\Delta_{k,\ell}Q_{x,h}(y) - \sigma^4\alpha^2\beta^4\gamma^8\delta^8 + O(\tau^{2c'}\alpha^2\beta^4\gamma^8\delta^8).$$

By Lemmas 3.5 and 3.6,

$$\mathbb{E}_{\substack{x,h\in\mathbb{F}_p^n \\ y,y+k,y+\ell\in\Phi_{x,h-x}}} A(x)A(h)\Delta_{k,\ell}Q_{x,h}(y) = \frac{\alpha\gamma^4\delta^4}{\rho^3}\mathbb{E}_{x,y,k,\ell}\Delta_{(0,k),(0,\ell)}S(x,y) + O\left(\frac{\varepsilon^{\Omega(1)}}{\rho^{O(1)}}\right).$$

If

$$|\mathbb{E}_{x,y,k,\ell}\Delta_{(0,k),(0,\ell)}S(x,y) - \sigma^4\alpha\beta^4\gamma^4\delta^4\rho^3| \gg \tau^{4c'}\alpha\beta^4\gamma^4\delta^4\rho^3,$$

then we could deduce the desired density increment from Lemma 6.5. Thus, we may proceed under the assumption that this inequality does not hold, which implies that $\mathbb{E}_{x,h}A(x)A(h)\|Q_{x,h} - \sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})}^4 \ll \tau^{4c'}\alpha^2\beta^4\gamma^8\delta^8$. It then follows from Markov's inequality that

$$\mathbb{P}((x,h) \in A \times A : \|Q_{x,h} - \sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})} \gg \tau^{2c'}\beta\gamma^2\delta^2) \ll \tau^{2c'}. \tag{52}$$

Next, we will show that, for typical $(x,h) \in A \times A$, the average size of $S(x,y)S(h,y)\Psi_h(x,y)$ is not very large. Certainly,

$$\begin{aligned}\mathbb{E}_y S(x,y)S(h,y)\Psi_h(x,y) &\leq \mathbb{E}_y T(x,y)T(h,y)\Psi_h(x,y)\\ &= \mathbb{E}_y B(y)C(x+y)C(h+y)D(2x+y)D(2h+y)\Psi_h(x,y)\end{aligned}$$

for every $(x,h) \in A \times A$. Setting

$$F(x,h,y) := B(y)C(x+y)C(h+y)D(2x+y)D(2h+y),$$

Lemma 3.6 says that

$$\mathbb{P}((x,h) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : \|F(x,h,\cdot) - \beta\gamma^2\delta^2\|_{U^2(\mathbb{F}_p^n)} > \varepsilon^{1/8}) \ll \sqrt{\varepsilon}.$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Thus, by Lemma 3.5, as long as $c_2$ is large enough, we have

$$\mathbb{P}\bigg((x,h) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : |\mathbb{E}_y S(x,y)S(h,y)\Psi_h(x,y)| > 100\frac{\beta\gamma^2\delta^2\rho^2}{p}\bigg) \ll \sqrt{c_1}(\sigma\tau\alpha\beta\gamma\delta\rho)^{16c}, \quad (53)$$

say.

Now, setting $c' = 8c$, the contribution to the left-hand side of (50) coming from pairs $(x,h) \in A \times A$ for which $\|Q_{x,h} - \sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})} \gg \tau^{16c}\beta\gamma^2\delta^2$ or $|\mathbb{E}_y S(x,y)S(h,y)\Psi_h(x,y)| > 100\beta\gamma^2\delta^2\rho^2/p$ is $\ll \tau^{16c}$. Thus, by the pigeonhole principle, there exists an $h \in A$ and a subset $A' \subset A$ of density $\alpha' \gg \tau^{O(1)}\alpha$ in $\mathbb{F}_p^n$ such that

$$\mathbb{E}_{x\in A'}\mathbb{E}_{(x,y)\in\Psi_h}(f_{x,h} - \sigma^2)(y) \gg \tau^c$$

and $\|Q_{x,h} - \sigma\beta\gamma^2\delta^2\|_{U^2(\Phi_{x,h-x})} \ll \tau^{16c}\beta\gamma^2\delta^2$ for every $x \in A'$. Recalling the definition of $f_{x,h}$, the above displayed equation says that

$$\mathbb{E}_{x,y} S(x,y)A'(x)S(h,y)\Psi_h(x,y) \geq (\sigma + \Omega(\tau^c))\frac{\alpha'\sigma\beta\gamma^2\delta^2\rho^2}{p}.$$

We now use Lemma 6.6 to find a subset $A'' \subset A'$ of density $\alpha'' \gg \alpha'\rho\tau^{O(1)}$ and a $u' \in \mathbb{F}_p^n$ such that

$$\mathbb{E}_{x,y} S(x,y)A'(x)S(h,y)\Phi_h'(x,y) \geq (\sigma + \Omega(\tau^c))\frac{\alpha''\sigma\beta\gamma^2\delta^2\rho^2}{p},$$

where

$$\Phi_h' = \{(x,y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : y \in \Phi_{x,h-x} \text{ and } \xi_{x,h}\cdot(y-u') = 0\}.$$

This gives the conclusion of the lemma.

The proofs of the two remaining cases are very similar to those appearing in earlier subsections, so we will be more brief in our arguments. Next, assume that

$$\|g_S\|_{\star_2} \geq \tau\alpha^{1/2}\beta\gamma^{1/2}\delta\rho.$$

As in the second part of the proof of Lemma 6.5, we may proceed under the assumption that

$$\mathbb{E}_{x,y,x',y'} S(x,y)S(x,y'-x)S(-x',x+y+x')S(-x',x'+y') \geq \bigg(\sigma^4 + \frac{\tau^4}{2}\bigg)\alpha^2\beta^4\gamma^2\delta^4\rho^4,$$

and use it to show that either

$$\mathbb{E}_{x,y} B(y)D(2x+y)\Phi(x,y)S(x,y'-x)S(-x',x+y+x') = (\sigma^2 + O(\tau^5))\alpha\beta^3\gamma\delta^3\rho^3$$

for almost every pair $(x',y')$ for which $(-x',x'+y') \in S$, or else the desired density increment follows from Lemma 6.3. By Lemmas 3.5, 3.6, and 4.1 and the Cauchy–Schwarz inequality, either

$$\begin{aligned}&\mathbb{E}_{\substack{x',y'\\(-x',x'+y')\in S}} \mathbb{E}_{x,y} B(y)D(2x+y)\Phi(x,y)S(x,y'-x)S(-x',x+y+x')\\ &\quad = \sigma^2\alpha\beta^3\gamma\delta^3\rho^3 + O(\tau^{16}\alpha\beta^3\gamma\delta^3\rho^3)\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}_{\substack{x',y'\\(-x',x'+y')\in S}} |\mathbb{E}_{x,y} B(y)D(2x+y)\Phi(x,y)S(x,y'-x)S(-x',x+y+x') - \sigma^2\alpha\beta^3\gamma\delta^3\rho^3|^2\\ &\quad \ll \tau^{32}\alpha^2\beta^6\gamma^2\delta^6\rho^6\end{aligned}$$

or else we have the desired density increment from Lemma 6.3. It then follows from Markov's inequality that

$$\mathbb{E}_{x,y} B(y)D(2x+y)\Phi(x,y)S(x,y'-x)S(-x',x+y+x') = (\sigma^2 + O(\tau^5))\alpha\beta^3\gamma\delta^3\rho^3$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

for all but a $\tau/4$-proportion of pairs $(x', y')$ for which $(-x', x' + y') \in S$, which means that we can find such a pair for which we also have

$$\mathbb{E}_{x,y} S(x, y)S(x, y' - x)S(-x', x + y + x') \geq \left(\sigma^4 + \frac{\tau^4}{2}\right)\alpha^2\beta^4\gamma^2\delta^4\rho^4,$$

so that we get the desired density increment by taking $A' = S(x, y' - x)$, $B' = B$, $C' = S(-x', x + y + x')$, $D' = D$, and $\Phi' = \Phi$ in the definition of $T'$.

Now assume that

$$\|g_S\|_{\star_3} \geq \tau\alpha\beta\gamma\delta^{1/2}\rho,$$

which means

$$\mathbb{E}_{z \in D} |\mathbb{E}_{2x+y=z} g_S(x, y)|^2 \geq \tau^2\alpha^2\beta^2\gamma^2\rho^2.$$

By Lemma 6.2 and the pigeonhole principle, there exists a subset $D_0 \subset D$ of density at least $\tau^2/2$ for which

$$|\mathbb{E}_{2x+y=z} g_S(x, y)| \geq \frac{\tau\alpha\beta\gamma\rho}{4}$$

whenever $z \in D_0$. As in the proof of Lemma 6.3, there is a subset $D_1 \subset D_0$ of density at least $1/2$ in $D_0$ such that either $\mathbb{E}_{2x+y=z} g_S(x, y) \geq \tau\alpha\beta\gamma\rho/4$ or $\mathbb{E}_{2x+y=z} g_S(x, y) \leq -\tau\alpha\beta\gamma\rho/4$ for all $z \in D_1$. As in the proof of Lemma 6.3, we can simply take $D' = D_1$ and $D' = D \setminus D_1$ in these cases, respectively, since, by Lemma 6.2, the fibers $\{(x, y) \in T : 2x + y = z\}$ typically have very close to their average size. □

## 7. Pseudorandomization

This section begins with yet more preliminaries. To prove Lemma 2.6, we will need a result of Cohen and Tal, which says that, for any finite set of polynomials in $\mathbb{F}_p[x_1, \ldots, x_m]$, one can find a partition of $\mathbb{F}_p^m$ into affine subspaces of relatively large dimension on which all of the polynomials are constant.

THEOREM 7.1 (Cohen and Tal [Theorem 3.6, CT15] specialized to prime fields). *Let $d, m$, and $t$ be natural numbers. There exists a positive integer $m'$ satisfying*

$$m' \gg \frac{m^{1/(d-1)!}}{t^e}$$

*such that, for any polynomials $P_1, \ldots, P_t \in \mathbb{F}_p[x_1, \ldots, x_m]$ of degree at most $d$, there is a partition of $\mathbb{F}_p^m$ into affine subspaces of dimension $m'$ such that $P_1, \ldots, P_t$ are constant on each affine subspace.*

To see that this formulation of Cohen and Tal's theorem is equivalent to Theorem 3.6 of [CT15], note that one can make all of the affine subspaces in the partitions produced by their theorem have the same dimension by simply partitioning each subspace not of the minimum possible dimension into more subspaces.

We will also need a 'bilinear' version of Cohen and Tal's result, which we will use to find partitions of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ into product spaces of the form $(u + V) \times (w + V)$ on which polynomials in two sets of variables $x_1, \ldots, x_m, y_1, \ldots, y_m$ are constant.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Corollary 7.2. *Let $d$ and $m$ be natural numbers. There exists a positive integer $m'$ satisfying*

$$m' \gg \frac{m^{1/(d-1)!^2}}{d^e}$$

*such that, for any polynomial $R \in \mathbb{F}_p[x_1, \dots, x_m, y_1, \dots, y_m]$ of degree at most $d$, there is a partition of $\mathbb{F}_p^m \times \mathbb{F}_p^m$ into products of affine subspaces of the form*

$$(u + V) \times (w + V),$$

*with each $\dim V = m'$, such that $R$ is constant on each product of affine subspaces.*

*Proof.* Write

$$R(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{d} P_i(\mathbf{x}) Q_i(\mathbf{y}),$$

where $P_0, \dots, P_d, Q_0, \dots, Q_d \in \mathbb{F}_p[z_1, \dots, z_m]$ satisfy $\deg P_i \leq i$ and $\deg Q_i \leq d - i$ for all $0 \leq i \leq d$. Applying Theorem 7.1 to $P_0, \dots, P_d$ gives us a positive integer $m_0' \gg m^{1/(d-1)!}/(d+1)^e$ and a partition

$$\mathbb{F}_p^m = \coprod_{j \in J} (u_j + V_j)$$

of $\mathbb{F}_p^m$, with $\dim V_j = m_0'$ for each $j \in J$, such that $P_0, \dots, P_d$ are all constant on each affine subspace $u_j + V_j$.

Now write

$$\mathbb{F}_p^m \times \mathbb{F}_p^m = \coprod_{j \in J} (u_j + V_j) \times \mathbb{F}_p^m = \coprod_{j \in J} \coprod_{w + V_j \in \mathbb{F}_p^m / V_j} (u_j + V_j) \times (w + V_j).$$

Since restricting a polynomial to a subspace cannot increase its degree, for each $j \in J$, we can apply Theorem 7.1 on each affine subspace $w + V_j$ to the polynomial

$$Q_{j,w}(\mathbf{y}) := \sum_{i=0}^{d} P_i(u_j + V_j) Q_i(\mathbf{y})$$

to get that there exists a positive integer $m'$ satisfying

$$m' \gg \frac{(m_0')^{1/(d-1)!}}{(d+1)^e} \gg \frac{m^{1/(d-1)!^2}}{d^e}$$

and a partition

$$w + V_j = \coprod_{k \in K_{j,w}} (w_k + V_k'),$$

with each subspace $V_k' \leq V_j$ having $\dim V_k' = m'$, such that $Q_{j,w}$ is constant on each $w_k + V_k'$.

Thus, we can write

$$\mathbb{F}_p^m \times \mathbb{F}_p^m = \coprod_{j \in J} \coprod_{w + V_j \in \mathbb{F}_p^m / V_j} \coprod_{k \in K_{j,w}} (u_j + V_j) \times (w_k + V_k'),$$

where $R$ is constant on each product $(u_j + V_j) \times (w_k + V_k')$. Since $V_k' \leq V_j$ we can further refine this partition to one of the desired form

$$\mathbb{F}_p^m \times \mathbb{F}_p^m = \coprod_{j \in J} \coprod_{w + V_j \in \mathbb{F}_p^m / V_j} \coprod_{k \in K_{j,w}} \coprod_{u + V_k' \in V_j / V_k'} (u + V_k') \times (w_k + V_k'),$$

on each part of which $R$ is still constant. □



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Finally, we will recall the recent quantitative inverse theorem of Gowers and Milićević for the $U^s$-norms on vector spaces over finite fields.

Theorem 7.3 (Gowers and Milićević [Theorem 7, GM20]). *Let $s$ be a natural number and assume that $p \geq s$. There exist constants $c_s, c'_{s,p} > 0$ so that, if $f : \mathbb{F}_p^m \to \mathbb{C}$ is a 1-bounded function satisfying*

$$\|f\|_{U^s} \geq \delta,$$

*then there exists a polynomial $P \in \mathbb{F}_p[x_1, \ldots, x_m]$ of degree $\deg P \leq s-1$ such that*

$$|\mathbb{E}_x f(x) e_p(P(x))| \gg_{s,p} \frac{1}{\exp^{c_s}(c'_{s,p}/\delta)}.$$

### 7.1 Proof of Lemma 2.6

Our proof of Lemma 2.6 is modeled after the corresponding pseudorandomization arguments in [Shk06a, Shk06b] and [Gre05a]. As was said in the outline, some new features arise from our desire for $B', C', D'$, and $\Phi'$ to be uniform with respect to $U^s$-norms of degree greater than 2 and from $\Phi$'s particular structure as a union of affine subspaces of the second factor of $\mathbb{F}_p^n \times \mathbb{F}_p^n$. The first of these two complications can be overcome by using Theorem 7.3 and then Theorem 7.1 or Corollary 7.2. We will discuss how $\Phi$'s structure influences our proof shortly.

The proof of Lemma 2.6 proceeds via an energy-increment argument. Each step of the energy-increment iteration will produce a partition $\mathscr{C}_j = (\mathcal{C}_{i,j})_{i \in I_j}$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ into cells $\mathcal{C}_{i,j}$ of the form

$$\mathcal{C}_{i,j} = (u_{i,j} + V_{i,j}) \times (w_{i,j} + V_{i,j})$$

for some subspace $V_{i,j} \leq \mathbb{F}_p^n$. For each cell $\mathcal{C} = (u+V) \times (w+V)$ in a partition $\mathscr{C}$, we set

- $B_{\mathcal{C}} := B \cap (w+V)$,
- $C_{\mathcal{C}} := C \cap (u+w+V)$,
- $D_{\mathcal{C}} := D \cap (2u+w+V)$, and
- $\Phi_{\mathcal{C}} := \Phi \cap \mathcal{C}$,

and, correspondingly,

- $\beta(\mathcal{C}) := \mu_{w+V}(B_{\mathcal{C}})$,
- $\gamma(\mathcal{C}) := \mu_{u+w+V}(C_{\mathcal{C}})$,
- $\delta(\mathcal{C}) := \mu_{2u+w+V}(D_{\mathcal{C}})$, and
- $\phi(\mathcal{C}) := \mu_{\mathcal{C}}(\Phi_{\mathcal{C}})$,

so that

$$T \cap \mathcal{C} = \{(x,y) \in \mathcal{C} : B_{\mathcal{C}}(y) C_{\mathcal{C}}(x+y) D_{\mathcal{C}}(2x+y) \Phi_{\mathcal{C}}(x,y) = 1\}.$$

Analogously to the pseudorandomization procedure for corners, we will show that if $\|B_{\mathcal{C}} - \beta(\mathcal{C})\|_{U^8(\mathbb{F}_p^n)}$, $\|C_{\mathcal{C}} - \gamma(\mathcal{C})\|_{U^8(\mathbb{F}_p^n)}$, or $\|D_{\mathcal{C}} - \delta(\mathcal{C})\|_{U^8(\mathbb{F}_p^n)}$ is large for a substantial portion of cells $\mathcal{C}$ in a partition $\mathscr{C}$, then there exists a refinement $\mathscr{C}'$ of $\mathscr{C}$ which has substantially larger energy. But even if $\Phi_{\mathcal{C}}$ is a union of affine subspaces of the same codimension $d$ for most cells in the partition, this may not be the case for the $\Phi_{\mathcal{C}'}$ corresponding to the cells $\mathcal{C}'$ in the refinement $\mathscr{C}'$. The codimensions of the affine subspaces $\{y \in w' + V' : \Phi_{\mathcal{C}'}(x,y) = 1\}$ can range from 0 to $d$, so before even considering how to obtain a pseudorandom $\Phi'$, we have already found an obstacle to even getting a set of the same general form as $\Phi'$.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

To get around this issue, we will pseudorandomize each of the sets

$$\Phi_{\mathcal{C}}^{\leq i} := \{(x,y) \in \Phi_{\mathcal{C}} : \mathbb{E}_{z \in w+V} \Phi(x,z) \geq p^{-i}\} \tag{54}$$

for $0 \leq i \leq d$, instead of just $\Phi_{\mathcal{C}} = \Phi_{\mathcal{C}}^{\leq d}$ itself. The definition (54) of $\Phi_{\mathcal{C}}^{\leq i}$ selects all of the subspaces of the second factor of $\mathcal{C}$ comprising $\Phi_{\mathcal{C}}$ that have codimension at most $i$. This will pseudorandomize each of

$$\Phi_{\mathcal{C}}^{i} := \{(x,y) \in \Phi_{\mathcal{C}} : \mathbb{E}_{z \in w+V} \Phi(x,z) = p^{-i}\} = \Phi_{\mathcal{C}}^{\leq i} \setminus \Phi_{\mathcal{C}}^{\leq i-1}$$

as well. At the end of the proof of Lemma 2.6, we will use an averaging argument to choose $\Phi'$ to be some suitable $\Phi_{\mathcal{C}}^{i}$.

For each $0 \leq i \leq d$, set $\phi^{\leq i}(\mathcal{C}) := \mu_{\mathcal{C}}(\Phi_{\mathcal{C}}^{\leq i})$. We define the energy $E(\mathscr{C})$ of a partition $\mathscr{C}$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ by

$$E(\mathscr{C}) := \frac{1}{4+d} \sum_{\mathcal{C} \in \mathscr{C}} \left( \beta(\mathcal{C})^2 + \gamma(\mathcal{C})^2 + \delta(\mathcal{C})^2 + \sum_{i=0}^{d} \phi^{\leq i}(\mathcal{C})^2 \right) \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}).$$

Note that the energy of any partition is bounded above by 1. We will now prove a couple of lemmas concerning this energy.

LEMMA 7.4. *Let $m' > m$ and $d$ be nonnegative integers, $A, B, C, D \subset \mathbb{F}_p^n$, $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ be of the form*

$$\Phi = \{(x,y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension between 0 and $d$, $\mathscr{C}$ be a partition of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ with each cell $\mathcal{C}$ taking the form*

$$\mathcal{C} = (u_{\mathcal{C}} + V_{\mathcal{C}}) \times (w_{\mathcal{C}} + V_{\mathcal{C}})$$

*for some subspace $V_{\mathcal{C}} \leq \mathbb{F}_p^n$ of codimension $m$, and suppose that $\mathscr{C}'$ is a refinement of $\mathscr{C}$ with each cell $\mathcal{C}'$ taking the form*

$$\mathcal{C}' = (u'_{\mathcal{C}'} + V'_{\mathcal{C}'}) \times (w'_{\mathcal{C}'} + V'_{\mathcal{C}'})$$

*for some subspace $V'_{\mathcal{C}'}$ of codimension $m'$. Then*

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \beta(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \sum_{\mathcal{C} \in \mathscr{C}} \beta(\mathcal{C})^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}),$$

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \gamma(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \sum_{\mathcal{C} \in \mathscr{C}} \gamma(\mathcal{C})^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}),$$

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \delta(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \sum_{\mathcal{C} \in \mathscr{C}} \delta(\mathcal{C})^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}),$$

*and*

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \phi^{\leq i}(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \sum_{\mathcal{C} \in \mathscr{C}} \phi^{\leq i}(\mathcal{C})^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C})$$

*for every $0 \leq i \leq d$.*

*Proof.* Note that it suffices to prove the result with the sum over $\mathcal{C} \in \mathscr{C}$ restricted to a single cell $\mathcal{C}_0$ and the sum over $\mathcal{C}' \in \mathscr{C}'$ restricted to all cells contained in $\mathcal{C}_0$, since one can then just sum over $\mathcal{C}_0$ in $\mathscr{C}$ to get the desired result. Thus, we may assume without loss of generality that $\mathscr{C}$ is the trivial partition $\{\mathbb{F}_p^n \times \mathbb{F}_p^n\}$.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Let $\beta, \gamma$, and $\delta$ denote the densities of $B, C$, and $D$, respectively, in $\mathbb{F}_p^n$. Since $\mathbb{F}_p^n \times B$ has density $\beta$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$, we have

$$\beta = \sum_{\mathcal{C}' \in \mathscr{C}'} \beta(\mathcal{C}') \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}'),$$

so that, by the Cauchy–Schwarz inequality,

$$\begin{aligned} \beta^2 &\leq \Bigg( \sum_{\mathcal{C}' \in \mathscr{C}'} \beta(\mathcal{C}')^2 \Bigg) \Bigg( \sum_{\mathcal{C}' \in \mathscr{C}'} \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}')^2 \Bigg) \\ &= \sum_{\mathcal{C}' \in \mathscr{C}'} \beta(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}'), \end{aligned}$$

as desired, since

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}')^2 = \frac{p^{2n}}{p^{2n-2m'}} \cdot \frac{1}{p^{4m'}} = \frac{1}{p^{2m'}} = \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}')$$

for all cells $\mathcal{C}'$ of $\mathscr{C}'$. Similarly, since $\{(x,y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : x + y \in C\}$ has density $\gamma$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$ and $\{(x,y) \in \mathbb{F}_p^n \times \mathbb{F}_p^n : 2x + y \in D\}$ has density $\delta$ in $\mathbb{F}_p^n \times \mathbb{F}_p^n$, we have

$$\gamma = \sum_{\mathcal{C}' \in \mathscr{C}'} \gamma(\mathcal{C}') \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}')$$

and

$$\delta = \sum_{\mathcal{C}' \in \mathscr{C}'} \delta(\mathcal{C}') \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}'),$$

it follows again from the Cauchy–Schwarz inequality that

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \gamma(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \gamma^2$$

and

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \delta(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq \delta^2.$$

The argument for the $\phi^{\leq i}$ is also essentially identical, but with one small difference. For ease of notation, set $\Phi^{\leq i} := \Phi^{\leq i}_{\mathbb{F}_p^n \times \mathbb{F}_p^n}$ and $\phi^{\leq i} := \phi^{\leq i}(\mathbb{F}_p^n \times \mathbb{F}_p^n)$ for all $0 \leq i \leq d$. Then we actually have

$$\phi^{\leq i} \leq \sum_{\mathcal{C}' \in \mathscr{C}'} \phi^{\leq i}(\mathcal{C}') \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}'),$$

instead of equality, since $\Phi^{\leq i} \cap \mathcal{C}' \subset \Phi^{\leq i}_{\mathcal{C}'}$ (which is why we run the energy-increment argument with the $\Phi^{\leq i}$, instead of the $\Phi^i$). It therefore follows yet again from the Cauchy–Schwarz inequality that

$$\sum_{\mathcal{C}' \in \mathscr{C}'} \phi^{\leq i}(\mathcal{C}')^2 \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C}') \geq (\phi^{\leq i})^2$$

for all $0 \leq i \leq d$. □



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

LEMMA 7.5. *Let $m$ and $d$ be nonnegative integers, $A, B, C, D \subset \mathbb{F}_p^n$, $\Phi \subset \mathbb{F}_p^n \times \mathbb{F}_p^n$ be of the form*

$$\Phi = \{(x, y) \in A \times \mathbb{F}_p^n : y \in u + V_x\},$$

*where each $V_x$ is a subspace of $\mathbb{F}_p^n$ of codimension between 0 and $d$, and $\mathscr{C}$ be a partition of $\mathbb{F}_p^n \times \mathbb{F}_p^n$ with each cell $\mathcal{C}$ taking the form*

$$\mathcal{C} = (u_{\mathcal{C}} + V_{\mathcal{C}}) \times (w_{\mathcal{C}} + V_{\mathcal{C}})$$

*for some subspace $V_{\mathcal{C}} \leq \mathbb{F}_p^n$ of dimension $m$. There exists a positive integer*

$$m' \gg m^{1/(8!)^2}$$

*and positive integers $c, c_p' > 0$, such that the following holds.*

*Let $\mathcal{C} \in \mathscr{C}$.*

(i) *If $\|B_{\mathcal{C}} - \beta(\mathcal{C})\|_{U^{10}(w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$, then there exists a partition $\mathscr{C}_{\mathcal{C}}'$ of $\mathcal{C}$ with each cell $\mathcal{C}'$ taking the form*

$$\mathcal{C}' = (u_{\mathcal{C}'}' + V_{\mathcal{C}'}') \times (w_{\mathcal{C}'}' + V_{\mathcal{C}'}'),$$

*with each $\dim V_{C'}' = m'$, such that*

$$\sum_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}'} \beta(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \beta(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c_p'/\varepsilon)^2}\right).$$

(ii) *If $\|C_{\mathcal{C}} - \gamma(\mathcal{C})\|_{U^{10}(u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$, then there exists a partition $\mathscr{C}_{\mathcal{C}}'$ of $\mathcal{C}$ with each cell $\mathcal{C}'$ taking the form*

$$\mathcal{C}' = (u_{\mathcal{C}'}' + V_{\mathcal{C}'}') \times (w_{\mathcal{C}'}' + V_{\mathcal{C}'}'),$$

*with each $\dim V_{C'}' = m'$, such that*

$$\sum_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}'} \gamma(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \gamma(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c_p'/\varepsilon)^2}\right).$$

(iii) *If $\|D_{\mathcal{C}} - \delta(\mathcal{C})\|_{U^{10}(2u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$, then there exists a partition $\mathscr{C}_{\mathcal{C}}'$ of $\mathcal{C}$ with each cell $\mathcal{C}'$ taking the form*

$$\mathcal{C}' = (u_{\mathcal{C}'}' + V_{\mathcal{C}'}') \times (w_{\mathcal{C}'}' + V_{\mathcal{C}'}'),$$

*with each $\dim V_{C'}' = m'$, such that*

$$\sum_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}'} \delta(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \delta(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c_p'/\varepsilon)^2}\right).$$

(iv) *Let $0 \leq i \leq d$. If $\|\Phi_{\mathcal{C}}^{\leq i} - \phi^{\leq i}(\mathcal{C})\|_{U^8(\mathcal{C})} \geq \varepsilon$, then there exists a partition $\mathscr{C}_{\mathcal{C}}'$ of $\mathcal{C}$ with each cell $\mathcal{C}'$ taking the form*

$$\mathcal{C}' = (u_{\mathcal{C}'}' + V_{\mathcal{C}'}') \times (w_{\mathcal{C}'}' + V_{\mathcal{C}'}'),$$

*with each $\dim V_{C'}' = m'$, such that*

$$\sum_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}'} \phi^{\leq i}(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \phi^{\leq i}(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c_p'/\varepsilon)^2}\right).$$

*Proof.* Let $c$ and $c_p'$ denote the constants $c_{10}$ and $c_{10,p}'$, respectively, from Theorem 7.3, and $m'$ denote the smaller of the two minimum values of $m'$ appearing in Theorem 7.1 when we take $m$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

as in this lemma, $d = 9$, and $t = 1$ and $m'$ appearing in Corollary 7.2 when we take $m$ as in this lemma and $d = 7$.

First assume that $\|B_{\mathcal{C}} - \beta(\mathcal{C})\|_{U^{10}(w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$. Then applying Theorem 7.3 with $s = 10$ yields a polynomial $P \in \mathbb{F}_p[x_1, \dots, x_m]$ of degree at most 9 such that

$$|\mathbb{E}_{x \in w_{\mathcal{C}}+V_{\mathcal{C}}}(B_{\mathcal{C}} - \beta(\mathcal{C}))(x)e_p(P(x))| \gg \frac{1}{\exp^c(c'_p/\varepsilon)}.$$

By Theorem 7.1, there exists a partition $(w_i + V_i)_{i \in I}$ of $w_{\mathcal{C}} + V_{\mathcal{C}}$ into affine subspaces of $w_{\mathcal{C}} + V_{\mathcal{C}}$ of dimension $m'$, on each of which $P$ is constant. Thus, by the triangle inequality,

$$\mathbb{E}_{i \in I}|\mathbb{E}_{x \in w_i+V_i}(B_{\mathcal{C}} - \beta(\mathcal{C}))(x)| \gg \frac{1}{\exp^c(c'_p/\varepsilon)},$$

so that, by the Cauchy–Schwarz inequality,

$$\mathbb{E}_{i \in I}|\mathbb{E}_{x \in w_i+V_i}(B_{\mathcal{C}} - \beta(\mathcal{C}))(x)|^2 \gg \frac{1}{\exp^c(c'_p/\varepsilon)^2}.$$

Expanding the square, this means that

$$\mathbb{E}_{i \in I}|\mathbb{E}_{x \in w_i+V_i}B_{\mathcal{C}}(x)|^2 \geq \beta(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\right). \tag{55}$$

Now we partition the whole cell of interest $\mathcal{C}$ by writing

$$\mathcal{C} = (u_{\mathcal{C}} + V_{\mathcal{C}}) \times \coprod_{i \in I}(w_i + V_i) = \coprod_{i \in I}(u_{\mathcal{C}} + V_{\mathcal{C}}) \times (w_i + V_i),$$

and, for each $i \in I$, use that $V_i \leq V_{\mathcal{C}}$ to partition $u_{\mathcal{C}} + V_{\mathcal{C}}$ into cosets of $V_i$ to get

$$\mathcal{C} = \coprod_{i \in I} \coprod_{u'+V_i \in V_{\mathcal{C}}/V_i} (u_{\mathcal{C}} + u' + V_i) \times (w_i + V_i) =: \coprod_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}} \mathcal{C}'.$$

Since $\mu_{\mathcal{C}}(\mathcal{C}') = |\mathscr{C}'|^{-1}$ for each $\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}$, (55) reads

$$\sum_{\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}} \beta(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \beta(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\right).$$

The arguments for $C_{\mathcal{C}}$ and $D_{\mathcal{C}}$ are again analogous, but we include them for the sake of completeness. Next, assume that $\|C_{\mathcal{C}} - \gamma(\mathcal{C})\|_{U^{10}(u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$. Applying Theorem 7.3 yields a polynomial $P \in \mathbb{F}_p[x_1, \dots, x_m]$ of degree at most 9 such that

$$|\mathbb{E}_{x \in u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}}}(C_{\mathcal{C}} - \gamma(\mathcal{C}))(x)e_p(P(x))| \gg \frac{1}{\exp^c(c'_p/\varepsilon)}.$$

By Theorem 7.1, there exists a partition $(v_i + V_i)_{i \in I}$ of $u_{\mathcal{C}} + w_{\mathcal{C}} + V_{\mathcal{C}}$ into affine subspaces of $u_{\mathcal{C}} + w_{\mathcal{C}} + V_{\mathcal{C}}$ of dimension $m'$ on each of which $P$ is constant. Thus, by the Cauchy–Schwarz inequality again,

$$\mathbb{E}_{i \in I}|\mathbb{E}_{x \in v_i+V_i}C_{\mathcal{C}}(x)|^2 \geq \gamma(\mathcal{C})^2 + \Omega\left(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\right).$$

Now we partition the whole cell $\mathcal{C}$ by writing

$$\mathcal{C} = \coprod_{i \in I} \coprod_{v'+V_i \in V_{\mathcal{C}}/V_i} (v_i - w_{\mathcal{C}} + v' + V_i) \times (w_{\mathcal{C}} - v' + V_i) =: \coprod_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}} \mathcal{C}'.$$



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Since $(v_i - w_{\mathcal{C}} + v' + V_i) + (w_{\mathcal{C}} - v' + V_i) = v_i + V_i$, we conclude that

$$\sum_{\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}} \gamma(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \gamma(\mathcal{C})^2 + \Omega\bigg(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\bigg).$$

Now assume that $\|D_{\mathcal{C}} - \delta(\mathcal{C})\|_{U^{10}(2u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}})} \geq \varepsilon$. Applying Theorem 7.3 yields a polynomial $P \in \mathbb{F}_p[x_1, \dots, x_m]$ of degree at most 9 such that

$$|\mathbb{E}_{x \in 2u_{\mathcal{C}}+w_{\mathcal{C}}+V_{\mathcal{C}}}(D_{\mathcal{C}} - \delta(\mathcal{C}))(x) e_p(P(x))| \gg \frac{1}{\exp^c(c'_p/\varepsilon)}.$$

By Theorem 7.1, there exists a partition $(v_i + V_i)_{i \in I}$ of $2u_{\mathcal{C}} + w_{\mathcal{C}} + V_{\mathcal{C}}$ into affine subspaces of $2u_{\mathcal{C}} + w_{\mathcal{C}} + V_{\mathcal{C}}$ of dimension $m'$ on each of which $P$ is constant. Thus, by the Cauchy–Schwarz inequality yet again,

$$\mathbb{E}_{i \in I} |\mathbb{E}_{x \in v_i + V_i} D_{\mathcal{C}}(x)|^2 \geq \delta(\mathcal{C})^2 + \Omega\bigg(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\bigg).$$

Now we partition the whole cell $\mathcal{C}$ by writing

$$\mathcal{C} = \coprod_{i \in I} \coprod_{v' + V_i \in V_{\mathcal{C}}/V_i} (u_{\mathcal{C}} + v' + V_i) \times (v_i - 2u_{\mathcal{C}} - 2v' + V_i) =: \coprod_{\mathcal{C}' \in \mathscr{C}_{\mathcal{C}}} \mathcal{C}'.$$

Since $(2u_{\mathcal{C}} + 2v' + V_i) + (v_i - 2u_{\mathcal{C}} - 2v' + V_i) = v_i + V_i$, we conclude that

$$\sum_{\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}} \delta(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \delta(\mathcal{C})^2 + \Omega\bigg(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\bigg).$$

Finally, suppose that $\|\Phi_{\mathcal{C}}^{\leq i} - \phi^{\leq i}(\mathcal{C})\|_{U^8(\mathcal{C})} \geq \varepsilon$ for some $0 \leq i \leq d$. Theorem 7.3 then says that there exists a polynomial $R \in \mathbb{F}_p[x_1, \dots, x_m, y_1, \dots, y_m]$ of degree at most 7 such that

$$|\mathbb{E}_{(x,y) \in \mathcal{C}}(\Phi_{\mathcal{C}}^{\leq i} - \phi^{\leq i}(\mathcal{C}))(x,y) e_p(R(x,y))| \gg \frac{1}{\exp^c(c'_p/\varepsilon)}.$$

By Corollary 7.2, there exists a partition $\mathscr{C}'_{\mathcal{C}}$ of $\mathcal{C}$ into affine subspaces of the form $(u + V) \times (w + V)$ with $\dim V = m'$, on each of which $R$ is constant. Thus, by the Cauchy–Schwarz inequality, we have

$$\mathbb{E}_{\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}} |\mathbb{E}_{(x,y) \in \mathcal{C}'} \Phi_{\mathcal{C}}^{\leq i}(x,y)|^2 \geq \phi^{\leq i}(\mathcal{C})^2 + \Omega\bigg(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\bigg).$$

Since

$$\mathbb{E}_{(x,y) \in \mathcal{C}'} \Phi_{\mathcal{C}}^{\leq i}(x,y) \leq \mathbb{E}_{(x,y) \in \mathcal{C}'} \Phi_{\mathcal{C}'}^{\leq i}(x,y) = \phi^{\leq i}(\mathcal{C}'),$$

the conclusion

$$\sum_{\mathcal{C}' \in \mathscr{C}'_{\mathcal{C}}} \phi^{\leq i}(\mathcal{C}')^2 \mu_{\mathcal{C}}(\mathcal{C}') \geq \phi^{\leq i}(\mathcal{C})^2 + \Omega\bigg(\frac{1}{\exp^c(c'_p/\varepsilon)^2}\bigg)$$

now follows. □

Now we can prove Lemma 2.6.

*Proof of Lemma 2.6.* We proceed via an energy-increment argument, as described at the beginning of the subsection. A cell $\mathcal{C} = (u + V) \times (w + V)$ in a partition $\mathscr{C}_j$ is said to be *expired* if $\beta(\mathcal{C}), \gamma(\mathcal{C}), \delta(\mathcal{C})$, or $\phi^{\leq d}(\mathcal{C})$ is less than $\tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/4$, and a nonexpired cell $\mathcal{C}$ is said to be



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

*uniform* if

$$\|B_{\mathcal{C}} - \beta(\mathcal{C})\|_{U^{10}(w+V)}, \quad \|C_{\mathcal{C}} - \gamma(\mathcal{C})\|_{U^{10}(u+w+V)}, \quad \|D_{\mathcal{C}} - \delta(\mathcal{C})\|_{U^{10}(2u+w+V)} < \varepsilon$$

and

$$\|\Phi_{\mathcal{C}}^{\leq i} - \phi^{\leq i}\|_{U^8(\mathcal{C})} < \varepsilon$$

for all $0 \leq i \leq d$. We will denote the subset of expired cells of $\mathscr{C}_j$ by $\mathscr{E}_j$, the subset of uniform cells by $\mathscr{U}_j$, and the subset of nonexpired, nonuniform cells by $\mathscr{N}_j$, so that $\mathscr{E}_j, \mathscr{U}_j$, and $\mathscr{N}_j$ partition $\mathscr{C}_j$. For any subset $K \subset I_j$, we define $\eta(K)$ to be the measure of all cells indexed by $K$:

$$\eta(K) := \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}\Big(\coprod_{k \in K} \mathcal{C}_{k,j}\Big).$$

Finally, we define a sequence of integers $(m_j)_{j=0}^{\infty}$ by setting $m_0 = n$ and, for every $j > 0$, $m_j$ to be the minimum of the value of $m'$ appearing in Theorem 7.1 when we take $m = m_{j-1}$, $d = 9$, and $t = 1$ and of $m'$ appearing in Corollary 7.2 when we take $m = m_{j-1}$ and $d = 7$, so that

$$m_j \geq c_1 n^{c_2^j}$$

for some absolute constants $0 < c_1, c_2 < 1$.

Set $\mathscr{C}_0$ to be the trivial partition $\{\mathbb{F}_p^n \times \mathbb{F}_p^n\}$ of $\mathbb{F}_p^n \times \mathbb{F}_p^n$. Letting $c = c_{10}$ and $c' = c'_{10,p}$ be as in Lemma 7.5, then, as long as $\eta(\mathscr{N}_j) \geq \tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/2$ and $m_{j+1} \geq 1$, there exists a refinement $\mathscr{C}_{j+1}$ of $\mathscr{C}_j$ such that:

(i) $\dim V_{i,j+1} \geq m_{j+1}$ for every $i \in I_{j+1}$ and $\dim V_{i,j+1} = m_{j+1}$ whenever $\mathcal{C}_{i,j+1} \in \mathscr{N}_{j+1}$; and
(ii) $E(\mathscr{C}_{j+1}) \geq E(\mathscr{C}_j) + \Omega(\tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/d \exp^c(c'_p/\varepsilon)^2)$.

Indeed, suppose that $\eta(\mathscr{N}_j) \geq \tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/2$. Each cell $\mathcal{C}$ in $\mathscr{N}_j$ must be of dimension $m_j \times m_j$. By Lemmas 7.4 and 7.5, there exists a partition $\mathscr{C}_{k,j+1} = (\mathcal{C}_{k,j+1})_{k \in K_{\mathcal{C}}}$ of each $\mathcal{C} \in \mathscr{N}_j$ such that

$$\frac{1}{4+d} \sum_{k \in K_{\mathcal{C}}} \Big(\beta(\mathcal{C}_{k,j+1})^2 + \gamma(\mathcal{C}_{k,j+1})^2 + \delta(\mathcal{C}_{k,j+1})^2 + \sum_{i=0}^{d} \phi^{\leq i}(\mathcal{C}_{k,j+1})^2\Big)\mu_{\mathcal{C}}(\mathcal{C}_{k,j+1})$$

is at least

$$\frac{\beta(\mathcal{C})^2 + \gamma(\mathcal{C})^2 + \delta(\mathcal{C})^2 + \sum_{i=0}^{d} \phi^{\leq i}(\mathcal{C})^2}{4+d} + \Omega\Big(\frac{1}{d \exp^c(c'/\varepsilon)^2}\Big)$$

and each $\mathcal{C}_{k,j+1}$ is of the form $(u' + V') \times (w' + V')$ with $\dim V' = m_{j+1} \geq 1$. Taking

$$\mathscr{C}_{j+1} := \{\mathcal{C}_{k,j+1} : k \in K_{\mathcal{C}}, \mathcal{C} \in \mathscr{N}_j\} \cup \mathscr{E}_j \cup \mathscr{U}_j,$$

we see that multiplying both sides of the above by $\mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(\mathcal{C})$ and summing over $\mathcal{C} \in \mathscr{N}_j$ yields

$$E(\mathscr{C}_{j+1}) \geq E(\mathscr{C}_j) + \Omega\Big(\frac{\tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)}{d \exp^c(c'/\varepsilon)^2}\Big).$$

Since $E(\mathscr{C}) \leq 1$ for all partitions $\mathscr{C}$, this iteration must terminate for some

$$j = j_0 \ll d \exp^c(c'/\varepsilon)^2/\tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T),$$

at which point either $\eta(\mathscr{N}_{j_0}) < \tau \mu_{\mathbb{F}_p^n \times \mathbb{F}_p^n}(T)/2$ or $m_{j_0+1} < 1$. Assuming that $n \geq c_1^{-c_2^{-(j_0+1)}}$ ensures that the latter case cannot occur.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

Since $\eta(\mathscr{E}_j) < \tau\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T)/4$, we have

$$\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}\Big(S\cap\bigcup_{\mathcal{C}\in\mathscr{U}_{j_0}}\mathcal{C}\Big)\geq\Big(\sigma+\frac{\tau}{4}\Big)\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T).$$

This certainly implies that

$$\sum_{\mathcal{C}\in\mathscr{U}_{j_0}}\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(S\cap\mathcal{C})\geq\Big(\sigma+\frac{\tau}{4}\Big)\sum_{\mathcal{C}\in\mathscr{U}_{j_0}}\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T\cap\mathcal{C}),$$

so that, by the pigeonhole principle, there exists a cell $\mathcal{C}_0$ in $\mathscr{U}_{j_0}$ for which

$$\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(S\cap\mathcal{C}_0)\geq\Big(\sigma+\frac{\tau}{4}\Big)\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T\cap\mathcal{C}_0).$$

Since $\Phi_{\mathcal{C}_0}=\coprod_{i=0}^d\Phi_{\mathcal{C}_0}^i$, another application of the pigeonhole principle tells us that there exists a $0\leq i\leq d$ for which we also have the density increment

$$\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(S\cap\mathcal{C}_0\cap\Phi_{\mathcal{C}_0}^i)\geq\Big(\sigma+\frac{\tau}{4}\Big)\mu_{\mathbb{F}_p^n\times\mathbb{F}_p^n}(T\cap\mathcal{C}_0\cap\Phi_{\mathcal{C}_0}^i).$$

As noted at the beginning of this subsection,

$$\|\Phi_{\mathcal{C}_0}^i-\phi^i(\mathcal{C}_0)\|_{U^8(\mathcal{C}_0)}=\|\Phi_{\mathcal{C}_0}^{\leq i}-\phi^{\leq i}(\mathcal{C}_0)+\phi^{\leq i-1}(\mathcal{C}_0)-\Phi_{\mathcal{C}_0}^{\leq i-1}\|_{U^8(\mathcal{C}_0)}<2\varepsilon,$$

so that the conclusion of the lemma now follows. □

## 8. The density-increment argument

Now we can finally prove Theorem 1.2 by iterating Lemma 2.7.

*Proof of Theorem 1.2.* Suppose that $S\subset\mathbb{F}_p^n\times\mathbb{F}_p^n$ has density $\sigma$ and contains no nontrivial L-shaped configurations. Set $S_0:=S$, $n_0:=n$, $d_0=0$, $\varepsilon_0:=1$, $A_0=B_0=C_0=D_0=\mathbb{F}_p^n$, and $\Phi_0:=\mathbb{F}_p^n\times\mathbb{F}_p^n$. Applying Lemma 2.7 repeatedly produces sequences of $S_i$, $n_i$, $d_i$, $\varepsilon_i$, $A_i$, $B_i$, $C_i$, $D_i$, and $\Phi_i$, with $A_i,B_i,C_i,D_i\subset\mathbb{F}_p^{n_i}$ and $\Phi_i\subset A_i\times\mathbb{F}_p^{n_i}$ of the form

$$\Phi_i=\{(x,y)\in A_i\times\mathbb{F}_p^{n_i}:y\in u+V_x\},$$

where each $V_x\leq\mathbb{F}_p^{n_i}$ is a subspace of codimension $d_i$, such that, on setting

$$T_i:=\{(x,y)\in\mathbb{F}_p^{n_i}\times\mathbb{F}_p^{n_i}:B_i(y)C_i(x+y)D_i(2x+y)\Phi_i(x,y)=1\},$$

$\alpha_i:=\mu_{\mathbb{F}_p^{n_i}}(A_i)$, $\beta_i:=\mu_{\mathbb{F}_p^{n_i}}(B_i)$, $\gamma_i:=\mu_{\mathbb{F}_p^{n_i}}(C_i)$, $\delta_i:=\mu_{\mathbb{F}_p^{n_i}}(D_i)$, and $\rho_i:=\mu_{\mathbb{F}_p^{n_i}\times\mathbb{F}_p^{n_i}}(\Phi_i)/\alpha_i=p^{-d_i}$, we have, for each $i\geq1$, that:

(i) $S_i\subset T_i$ has density $\sigma_i$ in $T_i$, where $\sigma_i\geq\sigma_{i-1}+\Omega(\sigma^{O(1)})$;
(ii) $n_i\gg n_{i-1}^{c_1^{O(\exp^c(c'/\varepsilon_i)/(\sigma\alpha_{i-1}\beta_{i-1}\gamma_{i-1}\delta_{i-1}\rho_{i-1})^{O(1)})}}$;
(iii) $\varepsilon_i\leq(\sigma\alpha_i\beta_i\gamma_i\delta_i\rho_i)^{O(1)}\exp(-(64/\sigma)^{O(1)})$;
(iv) $\alpha_i,\beta_i,\gamma_i,\delta_i\gg(\sigma\alpha_{i-1}\beta_{i-1}\gamma_{i-1}\delta_{i-1}\rho_{i-1})^{O(1)}$;
(v) $d_i\leq d_{i-1}+1$;
(vi) $\|A_i-\alpha_i\|_{U^{10}(\mathbb{F}_p^{n_i})}$, $\|B_i-\beta_i\|_{U^{10}(\mathbb{F}_p^{n_i})}$, $\|C_i-\gamma_i\|_{U^{10}(\mathbb{F}_p^{n_i})}$, $\|D_i-\delta_i\|_{U^{10}(\mathbb{F}_p^{n_i})}$, $\|\Phi_i-\alpha_i\rho_i\|_{U^8(\mathbb{F}_p^{n_i}\times\mathbb{F}_p^{n_i})}<\varepsilon_i$; and
(vii) $S_i$ contains no nontrivial L-shaped configurations;



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

provided that

$$n_{i-1} \geq \exp^2\left(O\left(\frac{\exp^c(c'/\varepsilon_i)}{(\sigma\alpha_{i-1}\beta_{i-1}\gamma_{i-1}\delta_{i-1}\rho_{i-1})^{c_4}}\right)\right). \tag{56}$$

Since no set can have density larger than 1, the lower bound (56) must fail for some $i = i_0 + 1 \ll \sigma^{-O(1)}$. Thus, there exists an absolute constant $c'' > 1$ such that

$$n_{i_0} \ll \exp^{c''}(O(1/\sigma^{O(1)}))$$

while, on the other hand,

$$n_{i_0} \gg n^{c_1^{O(\exp^{c''}(O(1/\sigma^{O(1)})))}}.$$

Comparing the upper and lower bounds for $n_{i_0}$ and taking the $c''$-fold iterated logarithm of both sides yields the bound in Theorem 1.2. □

## Acknowledgements

The author thanks Ben Green and Freddie Manners for helpful discussions, and Ben Green, Noah Kravitz, Terry Tao, and the anonymous referees for useful comments on earlier drafts. The author was supported by the NSF Mathematical Sciences Postdoctoral Research Fellowship Program under Grant No. DMS-1903038 and the Oswald Veblen fund, and also gratefully acknowledges the support and hospitality of the Hausdorff Institute for Mathematics, where the bulk of this paper was written.

## Conflicts of Interest

None.

## References

Aus13a T. Austin, *Partial difference equations over compact Abelian groups, I: modules of solutions*, Preprint (2013), arXiv:1305.7269.

Aus13b T. Austin, *Partial difference equations over compact Abelian groups, II: step-polynomial solutions*, Preprint (2013), arXiv:1309.3577.

CT15 G. Cohen and A. Tal, *Two structural results for low degree polynomials and applications*, in *Approximation, randomization, and combinatorial optimization. Algorithms and techniques*, LIPICS - Leibniz International Proceedings in Informatics, vol. 40 (Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2015), 680–709; MR 3441991.

FK78 H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Anal. Math. **34** (1978), 275–291 (1979); MR 531279.

Gow98a W. T. Gowers, *Fourier analysis and Szemerédi's theorem*, in *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, Documenta Mathematica, **Extra Volume ICM** (Deutsche Mathematiker-Vereinigung, 1998), 617–629; MR 1660658.

Gow98b W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Funct. Anal. **8** (1998), 529–551; MR 1631259.

Gow01a W. T. Gowers, *Arithmetic progressions in sparse sets*, in *Current developments in mathematics, 2000* (International Press, Somerville, MA, 2001), 149–196; MR 1882535.

Gow01b W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588; MR 1844079.

Gow07 W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), 897–946; MR 2373376.



https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press

GM20 W. T. Gowers and L. Milićević, *An inverse theorem for Freiman multi-homomorphisms*, Preprint (2020), arXiv:2002.11667.

Gra97 R. L. Graham, *Euclidean Ramsey theory*, in *Handbook of discrete and computational geometry*, Discrete Mathematics and Its Applications (CRC Press, Boca Raton, FL, 1997), 153–166; MR 1730164.

Gre05a B. Green, *An argument of Shkredov in the finite field setting*, Preprint (2005), http://people.maths.ox.ac.uk/greenbj/papers/corners.pdf.

Gre05b B. Green, *Finite field models in additive combinatorics*, in *Surveys in combinatorics 2005*, London Mathematical Society Lecture Note series, vol. 327 (Cambridge University Press, Cambridge, 2005), 1–27; MR 2187732.

GT10 B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), 1753–1850; MR 2680398.

Man18 F. Manners, *Quantitative bounds in the inverse theorem for the Gowers $U^{s+1}$-norms over cyclic groups*, Preprint (2018), arXiv:1811.00718.

NRS06 B. Nagle, V. Rödl and M. Schacht, *The counting lemma for regular $k$-uniform hypergraphs*, Random Structures Algorithms **28** (2006), 113–179; MR 2198495.

Pre15 S. Prendiville, *Matrix progressions in multidimensional sets of integers*, Mathematika **61** (2015), 14–48; MR 3333959.

RS04 V. Rödl and J. Skokan, *Regularity lemma for $k$-uniform hypergraphs*, Random Structures Algorithms **25** (2004), 1–42; MR 2069663.

Shk06a I. D. Shkredov, *On a generalization of Szemerédi's theorem*, Proc. Lond. Math. Soc. (3) **93** (2006), 723–760; MR 2266965.

Shk06b I. D. Shkredov, *On a problem of Gowers*, Izv. Ross. Akad. Nauk Ser. Mat. **70** (2006), 179–221; MR 2223244.

Shk09 I. D. Shkredov, *On an inverse theorem for $\mathbf{U}^3(\square)$-norm*, in *Modern problems of mathematics and mechanics*, Mathematics, Dynamical Systems, vol. 4, 2 (2009), 55–127.

Tao06 T. Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A **113** (2006), 1257–1280; MR 2259060.

Tao12 T. Tao, *Higher order Fourier analysis*, Graduate Studies in Mathematics, vol. 142 (American Mathematical Society, Providence, RI, 2012); MR 2931680.

Zha14 Y. Zhao, *An arithmetic transference proof of a relative Szemerédi theorem*, Math. Proc. Cambridge Philos. Soc. **156** (2014), 255–261; MR 3177868.

Sarah Peluse speluse@umich.edu

Department of Mathematics, University of Michigan, East Hall, Church Street, Ann Arbor, MI 48109, USA





https://doi.org/10.1112/S0010437X2300756X Published online by Cambridge University Press