On an Argument of Shkredov on Two-Dimensional Corners

Michael T. Lacey
School of Mathematics
Georgia Institute of Technology
Atlanta GA 30332
`lacey@math.gatech.edu`

William McClain
School of Mathematics
Georgia Institute of Technology
Atlanta GA 30332
`bill@math.gatech.edu`

Submitted: October 25, 2005; Accepted: January 19, 2007; Published: March 13, 2007

Abstract

Let $\mathbb{F}_2^n$ be the finite field of cardinality 2^n. For all large n, any subset $A \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$ of cardinality

$$|A| \gtrsim 4^n \frac{\log \log n}{\log n},$$

must contain three points $\{(x,y), (x+d,y), (x,y+d)\}$ for $x, y, d \in \mathbb{F}_2^n$ and $d \neq 0$. Our argument is an elaboration of an argument of Shkredov [14], building upon the finite field analog of Ben Green [10]. The interest in our result is in the exponent on $\log n$, which is larger than has been obtained previously.

1 Main Theorem

We are interested in extensions of Roth's theorem on arithmetic progressions in dense sets of integers to a two-dimensional, finite-field setting. Specifically, for a finite group G define the quantity $r_\angle(G)$ to be the cardinality of the largest subset of $G \times G$ containing no *corner*. A *corner* is triple of points of the form $\{(x,y), (x+d,y), (x,y+d)\}$ with $d \neq 0$.

While this concept is most interesting in the context of the groups $G = \mathbb{Z}_N$, it already makes sense–and is substantial–in the context of finite fields. In this paper, we only consider the case of $G = \mathbb{F}_2^n$. Here and throughout this paper we write $N = 2^n = |\mathbb{F}_2^n|$.

Theorem 1.1. $r_\angle(\mathbb{F}_2^n) \ll N^2 \frac{\log\log\log N}{\log\log N}$.

This bound is an improvement, in the setting of $\mathbb{F}_2^n$, of the bounds provided by Shkredov [14, 15], and as simplified by Ben Green [10, 11]. The main point is that we elaborate on the 'Density Increment' procedure, obtaining a density increment on a set which is the intersection of sublattices in two distinct sets of bases.

Our theorem is an example of the quantitative bounds on questions of arithmetic combinatorics. We refer the reader to the papers of Gowers [6], and surveys by T. Tao [17] and Ben Green [8, 10] for more history.

Erdős and Graham raised the question of quantitative bounds for $r_\angle$, and this question was raised again by Gowers [6]. Ajtai and Szemerédi [1] first proved that $r_\angle(\mathbb{Z}_N) = o(N^2)$. Furstenberg and Katznelson [4, 5] gave a far reaching extension, though their method of proof does not in and of itself permit explicit bounds. Solymosi [16], and V. Vu [18] provided such bounds, although of a weak nature.

I. Shkredov [13, 14] provided the first 'reasonable' bounds. We are using his ingenious argument, as explained and simplified by Ben Green [10] in the finite field setting. In particular Green showed that one could achieve an estimate in which $r_\angle(N)/N^2$ decreased like $(\log\log N)^{-c}$ where c could be taken to be $1/21$. We find an additional extension of this argument, and sharpen some inequalities to obtain our Theorem.

We comment that 'the finite field thesis' holds that questions of this type should first be studied in the context of finite fields. This is because one can implement many of the tools of analysis, e. g. convolution and Fourier transform, in that setting. In addition, one has the powerful concept of being able to pass to appropriate affine subspaces. Moreover, there are a range of methods that one can use to 'lift' the finite field argument to $\mathbb{Z}_N$. See papers by Bourgain [2], Green and Tao [9] and Shkredov [15] for more information.

Acknowledgment. One of us (M.L.) was supported in part by the NSF and the Guggenheim Foundation. One of us (W.M.) was supported in part by a VIGRE grant awarded to Georgia Institute of Technology. Both authors are grateful for the hospitality of the University of British Columbia. While there, we benefited greatly from conversations with Izabella Łaba and Jozef Solymosi. Ben Green and I. Shkredov pointed out an error in an earlier version of this paper. We also thank the referee for additional corrections, which lead to strengthening of the main result.

1.1 Outline of Proof

Ben Green has provided a comprehensive outline of the method of proof [10], so we will be somewhat brief.

Let $A \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$. It is natural to count the expected number of corners that A has. This expectation will be approximately what it should be if the 'Box norms' of A are small. These norms, we use three of them, are explicitly given in (2.8)—(2.10), and are a two dimensional analog of the 'combinatorial square' norms that play such a prominent role in the proof of Roth's Theorem [9, 10, 12]. It is therefore a certain measure of 'uniformity.'

An important difference between our approach and the previous ones is that we emphasize the role of three distinct coordinates in the problem. Two of these are the obvious X and Y coordinates, given by the canonical basis e_1 and e_2 respectively. The third is the *diagonal* coordinate D, given by $\mathrm{e}_1 + \mathrm{e}_2$. In this way, our argument resembles the ergodic theoretic arguments, and in particular that of Conze and Lesigne [3] who describe the characteristic factor for the ergodic averages[1]

$$N^{-1} \sum_{n=1}^{N} f_1(T_1^n) f_2(T_2^n)$$

where T_1, T_2 are commuting measure preserving transformations.

Two of the three box norms employ this diagonal coordinate. Either being large is an 'obstruction to uniformity.' (See [17].) But, in contrast to the case of Roth's theorem, this obstruction to uniformity has no clear arithmetic consequence. It does imply, however, that the set A has an increased density on a sublattice. That is, there is a subsets $X, Y, D \subset \mathbb{F}_2^n$, of relatively large density, for which A has a larger density on the *intersection of* $X \times Y$ *and*

[1] The argument for two commuting transformations in this paper is complete. We are grateful to Bryna Kra for pointing out this reference to us and to Emmaunel Lesigne for providing us with a copy of this paper.

$X \overset{\text{diag}}{\times} D$, where in the latter product, we are taking the product in the coordinate system $(\mathrm{e}_1, \mathrm{e}_1 + \mathrm{e}_2)$. See Lemma 3.4. Forming the intersection of two sublattices in this way is the main new contribution of this paper.

It was a significant insight of Shkredov [14] that (1) one could, after an additional argument, assume that X, Y and D had an arithmetic structure, namely that it was *uniform* in the sense of (2.2), see Lemma 3.8; and (2) the Box norm could still be used as an 'obstruction to uniformity' if X, Y and D were uniform. This is the content of the Lemma 3.1.

If there is an obstruction to uniformity, an increment in the density of A can be found. One can find the obstruction to uniformity only a finite number of times, else the density of A would exceed one. Thus at some point, there would be no obstruction to uniformity, and so A would have a corner. All of the details of the proof are below.

2 Preliminaries and Definitions

We will let $H \subset \mathbb{F}_2^n$ denote a subspace. The dimension of this subspace will decrease in the course of the proof. $X, D, Y \subset H$ denote subsets. We adopt the notations of probability and expectation with respect to the counting measure on H. We view X as a subset of the first coordinate associated to basis element e_1; Y as a subset of the second coordinate, associated to basis element e_2; and D as a subset of the 'diagonal' coordinate associated to $\mathrm{e}_1 + \mathrm{e}_2$. We will be working with subsets of

$$S \overset{\text{def}}{=} X \times Y \cap X \overset{\text{diag}}{\times} D. \tag{2.1}$$

It is to be emphasized that the products are taken in the first instance in the e_1 and e_2 coordinates; the second in the e_1 and $\mathrm{e}_1 + \mathrm{e}_2$ coordinates.

Write *the density of X in H* as

$$\delta_X := \mathbb{P}(X \mid H) = \tfrac{|X|}{|H|},$$

and similarly for δ_Y and δ_D. In the iterative procedure used in this proof, these densities will be decreasing, somewhat rapidly. Throughout this paper we will not only be concerned with the density of the subsets X and D, but also with how 'uniformly distributed' they are in H. A quantification of this quality comes in the following definition:

$$\|X\|_{\text{Uni}} := \sup_{\xi \neq 0} \tfrac{|\widehat{X}(\xi)|}{|H|}. \tag{2.2}$$

This definition only makes sense relative to a subspace H. If $\|X\|_{\text{Uni}} \le \eta$ then we say that X is *η–uniform*. Here $\widehat{X}$ represents the the Fourier transform of X which is defined as follows, in which $\omega = -1$ is a second root of unity:

$$\widehat{g}(\xi) := \sum_{x \in H} g(x) \omega^{x \cdot \xi}.$$

It is immediate from the definition that a translate of a uniform set by an element of H is again uniform. After the deletion of a small subset, a uniform set is again uniform. Let $E \subset X$, we have

$$\|X - E\|_{\text{Uni}} \le \|X\|_{\text{Uni}} + \mathbb{P}(E \mid X). \tag{2.3}$$

This proposition will be used repeatedly.

Proposition 2.4. *Let $X \subset H$ and denote the density of X in H by δ_X. Then*

$$\big[\mathbb{E}_{d\in H}\big|\mathbb{E}_{y\in H}X(d-y)G(y) - \delta_X\mathbb{E}_{y\in H}G(y)\big|^2\big]^{1/2} \le \|X\|_{\mathrm{Uni}}\big[\mathbb{E}_y G(y)^2\big]^{1/2} \tag{2.5}$$

for any function G.

Notice that we are comparing a convolution to it's zero Fourier mode. The Proposition follows from Plancherel, and the definition of uniformity. A form of this inequality that we will use several times is this: For any two sets $A, B \subset H$,

$$\mathbb{E}_{x\in X}\big|\mathbb{E}_{\in Y}A(x+y)B(y) - \mathbb{P}(A)\mathbb{P}(B)\big|^2 \le \min\{\|A\|_{\mathrm{Uni}}^{1/2}\mathbb{P}(B)\,,\ \|B\|_{\mathrm{Uni}}^{1/2}\mathbb{P}(A)\}\,. \tag{2.6}$$

That is, only uniformity in one coordinate is required.

Now, S is as in (2.1), and let $A \subset S$. Write *the density of A* as

$$\delta := \mathbb{P}(A \mid S)\,.$$

This quantity will increase in the iterative procedure used in the proof. We define the *balanced function* of A to be the function supported on S as

$$f(x,y) := A(x,y) - \delta S\,.$$

Our standing assumption will be

$$\|X\|_{\mathrm{Uni}}\,,\ \|Y\|_{\mathrm{Uni}}\,,\ \|D\|_{\mathrm{Uni}} \le \upsilon\,, \qquad \upsilon \stackrel{\mathrm{def}}{=} (\delta\delta_X\delta_Y\delta_D)^C \tag{2.7}$$

where C is a large constant which we need not specify exactly, as its precise value only influences implied constants in our main Theorem. In the proofs of Lemmas, we will use the notation υ' for a fixed, but unimportant, function of υ, that tends to zero as υ does.

Further, for a function $f : S \longrightarrow \mathbb{C}$, define the following norm

$$\|f\|_{\Box} \stackrel{\mathrm{def}}{=} \delta_D^{-4}\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y}} f(x,y)f(x',y)f(x,y')f(x',y') \tag{2.8}$$

where we use the standard basis $(\mathrm{e}_1, \mathrm{e}_2)$. This norm averages 'cross correlations' of f over all boxes in $X \times Y$. When f is the balanced function of A, the norm being 'large' is an obstacle to A having the correct number of corners.

We use two additional norms. In the $(\mathrm{e}_1, \mathrm{e}_1 + \mathrm{e}_2)$ coordinate system,

$$\|f\|_{\Box,X}^4 := \delta_Y^{-4}\mathbb{E}_{\substack{x,x'\in X\\ d,d'\in D}} f(x,d)f(x',d)f(x,d')f(x',d')\,. \tag{2.9}$$

Similarly, with respect to the $(\mathrm{e}_2, \mathrm{e}_1 + \mathrm{e}_2)$ coordinate system, define

$$\|f\|_{\Box,Y}^4 := \delta_X^{-4}\mathbb{E}_{\substack{y,y'\in Y\\ d,d'\in D}} f(y,d)f(y',d)f(y,d')f(y',d')\,. \tag{2.10}$$

The leading normalizations in these definitions are initially confusing, but chosen so that the norms are essentially bounded from above by the L^∞ norm of f. The point of these next propositions is that the quantities introduced above behave as they should, under the assumption of uniformity. In particular, (2.16) and (2.17) justify the normalizations in the definition of the Box norms.

Proposition 2.11. *Let $X, Y, D \subset H$ be as above, and let S be as in (2.1). Assuming (2.7) we have*

$$\mathbb{E}_{\substack{x\in X\\ y\in Y}} S = \delta_D + O(\upsilon'); \tag{2.12}$$

$$\mathbb{E}_{\substack{x\in X\\ d\in D}} S = \delta_Y + O(\upsilon'); \tag{2.13}$$

$$\mathbb{E}_{\substack{d\in D\\ y\in Y}} S = \delta_X + O(\upsilon'); \tag{2.14}$$

$$\mathbb{E}_{\substack{x\in X\\ y\in Y\\ s\in H}} S(x,y)S(x+s,y)S(x,y+s) = \delta_X\delta_Y\delta_D^2 + O(\upsilon'); \tag{2.15}$$

$$\mathbb{E}_{\substack{x,x'\in X\\ d,d'\in D}} S(x,d)S(x',d)S(x,d')S(x',d') = \delta_Y^4 + O(\upsilon'); \tag{2.16}$$

$$\mathbb{E}_{\substack{y,y'\in Y\\ d,d'\in D}} S(y,d)S(y',d)S(y,d')S(y',d') = \delta_X^4 + O(\upsilon'). \tag{2.17}$$

In (2.16) we are using the $(\mathrm{e}_1, \mathrm{e}_1+\mathrm{e}_2)$ *coordinate systems, while in the (2.17) we are using the* $(\mathrm{e}_2, \mathrm{e}_1+\mathrm{e}_2)$ *coordinate system.*

Proof. For (2.12), observe that

$$\delta_X\delta_Y \mathbb{E}_{\substack{x\in X\\ y\in Y}} S = \mathbb{E}_{x,y\in H} X(x)Y(y)D(x+y) = \delta_X\delta_Y\delta_D + O(\delta_X^{1/2}\delta_Y^{1/2}\upsilon)\,.$$

This equality only requires uniformity in one coordinate. See (2.6). The equalities (2.13) and (2.14) are corollaries, after a change of variables.

To see (2.15), we apply Lemma 2.19, with $f = X$ and $g = Y$. Using the notation Φ in (2.20), we have

$$\begin{aligned}
&\mathbb{E}_{x,y,s\in H} S(x,y)S(x+s,y)S(x,y+s)\\
&\qquad = \mathbb{E}_{x,y,s\in H} X(x)X(x+s)D(x+y)D(s)Y(y)Y(y+s)\\
&\qquad = \mathbb{E}_{x,y\in H} X(x)Y(y)D(x+y)\Phi(x+y) + O(\upsilon')\\
&\qquad = \delta_X\delta_Y \mathbb{E}_{x,y\in H} D(x+y)\Phi(x+y) + O(\upsilon')\\
&\qquad = \delta_X\delta_Y \mathbb{E}_{x\in H} D(x)\Phi(x) + O(\upsilon')
\end{aligned}$$

It remains to estimate the last expectation, which we view as an inner product. Observe that $\widehat{\Phi}(0) = |H|\delta_D\delta_X\delta_Y$. And, by Plancherel,

$$\begin{aligned}
\bigl|\mathbb{E}_{x\in H} D(x)\Phi(x) - \delta_D^2\delta_X\delta_Y\bigr| &= |H|^{-2}\Bigl|\sum_{\alpha\neq 0}\widehat{D}(\alpha)\widehat{\Phi}(\alpha)\Bigr|\\
&\le \|D\|_{\mathrm{Uni}}|H|^{-1}\sum_{\alpha\neq 0}|\widehat{\Phi}(\alpha)|\\
&\le \upsilon|H|^{-2}\delta_D\sum_{\alpha\neq 0}|\widehat{X}(\alpha)\widehat{Y}(\alpha)|\\
&\le \upsilon\delta_D\delta_X^{1/2}\delta_Y^{1/2}\,.
\end{aligned}$$

Concerning (2.16), we use a similar proof to the one above. We follow the notation in

Lemma 2.19, and its proof. Set

$$\Psi(x,x') \stackrel{\text{def}}{=} \mathbb{E}_d Y(x+d)D(d)Y(x'+d)\,,$$
$$\Phi(x) \stackrel{\text{def}}{=} \frac{\delta_D}{|H|^2}\sum_{\alpha\in H}\widehat{Y}(\alpha)^2\omega^{\alpha\cdot x}\,.$$

Lemma 2.19 implies that Ψ is well approximated by Φ. Hence, we can estimate

$$\begin{aligned}
\delta_X^2\delta_D^2\cdot(2.16) &= \mathbb{E}_{\substack{x,x'\in H\\ d,d'\in H}} X(x)X(x')D(d)D(d')Y(x+d)Y(x'+d)Y(x+d')Y(x'+d')\\
&= \mathbb{E}_{x,x'\in H}X(x)X(x')\Psi(x,x')^2\\
&= \mathbb{E}_{x,x'\in H}X(x)X(x')\Phi(x+x')^2 + O(\upsilon')\\
&= \delta_X^2\mathbb{E}_{x\in H}\Phi(x+x')^2 + O(\upsilon')\\
&= \delta_X^2\delta_D^2|H|^{-4}\sum_{\alpha\in H}\widehat{Y}(\alpha)^4 + O(\upsilon')\\
&= \delta_X^2\delta_D^2\delta_Y^4 + O(\upsilon')\,.
\end{aligned}$$

Here, we have used Lemma 2.19, uniformity in X, Plancherel, and uniformity in Y. The second equality (2.17) follows from the first. This completes the proof.

□

Remark 2.18. There is a second way to see (2.15), which we only briefly indicate, since the method of proof is not self contained. Consider $S = X\times Y \cap X \overset{\text{diag}}{\times} D$ as a subset of $X\times Y$, and let Δ be it's balanced function. Namely $\Delta = S - \mathbb{P}(S \mid X\times Y)X\times Y$. One can then define the Box norm, as does Shkredov

$$\|\Delta\|_{\text{RectBox}}^4 \stackrel{\text{def}}{=} \mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y}}\Delta(x,y)\Delta(x',y)\Delta(x,y')\Delta(x',y')$$

It follows from the proof of (2.16), that we have $\|\Delta\|_{\text{RectBox}} \lesssim \upsilon'$. Shkredov [14] showed that under this assumption, and uniformity in X and Y, that the set S has nearly the expected number of point in it. That is the content of his 'Generalized von Neumann Lemma.'

Lemma 2.19. *Let D be uniform, and let f,g be two functions on H. Define*

$$\Phi(x) \stackrel{\text{def}}{=} \frac{\delta_D}{|H|^2}\sum_{\alpha\in H}\widehat{f}(\alpha)\widehat{g}(\alpha)\omega^{\alpha\cdot x}\,. \tag{2.20}$$

We have the inequality

$$\begin{aligned}
&\Big[\mathbb{E}_{x,y\in H}\big|\mathbb{E}_s f(x+s)D(s)g(y+s) - \Phi(x+y)\big|^2\Big]^{1/2}\\
&\qquad\lesssim \|D\|_{\text{Uni}}\big[\mathbb{E}_x f(x)^2\big]^{1/2}\cdot\big[\mathbb{E}_y g(y)^2\big]^{1/2}
\end{aligned} \tag{2.21}$$

Proof. Consider

$$\Psi(x,y) = \mathbb{E}_s f(x+s)D(s)g(y+s)$$

as a function on $H \times H$. Expanding f in dual variable α and g in dual variable β we have

$$\begin{aligned}\Psi(x,y) &= |H|^{-2} \sum_{\alpha,\beta\in H} \widehat{f}(\alpha)\widehat{g}(\beta)\omega^{\alpha\cdot x+\beta\cdot y}\mathbb{E}_s D(s)\omega^{(\alpha+\beta)\cdot s} \\ &= |H|^{-2} \sum_{\alpha,\beta\in H} \widehat{f}(\alpha)\widehat{g}(\beta)\omega^{\alpha\cdot x+\beta\cdot y}\frac{\widehat{D}(\alpha+\beta)}{|H|}.\end{aligned}$$

This shows that $\widehat{\Psi}(\alpha,\beta) = \widehat{f}(\alpha)\widehat{g}(\beta)|H|^{-1}\widehat{D}(\alpha+\beta)$. Clearly, Φ of the Lemma consists of the reconstruction of Ψ from those Fourier coefficients (α,β) for which $\alpha+\beta=0$. And by Plancherel, the Lemma follows from

$$\begin{aligned}\sum_{\alpha+\beta\neq 0} |\widehat{\Psi}(\alpha,\beta)|^2 &\leq \|D\|_{\text{Uni}}^2 \sum_{\alpha,\beta} |\widehat{f}(\alpha)\widehat{g}(\beta)|^2 \\ &= |H|^4 \cdot \|D\|_{\text{Uni}}^2 \cdot \mathbb{E}_x f(x)^2 \cdot \mathbb{E}_y g(y)^2 \,.\end{aligned}$$

□

3 Primary Lemmata

Our first lemma is a generalized von Neumann estimate, a term coined by Green and Tao [7, 9, 10]. It gives us sufficient conditions from which to conclude that A has a corner.

Lemma 3.1 (Generalized von Neumann)**.** *Suppose that $A \subset S$ with $\mathbb{P}(A \mid S) = \delta$; (2.7) holds; and we have the two inequalities*

$$\delta_X \delta_Y \delta_D \delta^2 N > C \,, \tag{3.2}$$

$$\max\{\|f\|_{\square}\,,\ \|f\|_{\square,X}\,,\ \|f\|_{\square,Y}\} \leq \kappa\delta^{5/4}\,. \tag{3.3}$$

Then A has a corner.

Here, and throughout this paper, C represents a large absolute constant. The exact value of C does not impact the qualitative nature of our estimate, so we do not seek to specify an optimal value for it. Also $0 < c, \kappa, \kappa' < 1$ are small fixed constants, which plays a role similar to C.

Our second lemma tells us that if the conditions of our first lemma are not satisfied then we can find a sublattice on which A has increased density.

Lemma 3.4 (Density Increment)**.** *For $0 < \kappa$ there is a constant $0 < \kappa' < 1$ for which the following holds. Suppose that $A \subset S = X \times Y \cap X \overset{\text{diag}}{\times} D$ with $\mathbb{P}(A \mid S) = \delta$, that f is the balanced function of A on S, and that*

$$\max\{\|f\|_{\square}\,,\ \|f\|_{\square,X}\,,\ \|f\|_{\square,Y}\} > \kappa\delta^{5/4}\,.$$

Then there exists $X' \subset X$, $Y' \subset Y$, $D' \subset D$ such that three conditions hold.

$$\text{either } X' = X, \text{ or } Y' = Y, \text{ or } D' = D; \tag{3.5}$$

$$\mathbb{P}(A \mid S') \geq \delta + \kappa'\delta^2\,, \qquad S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'\,; \tag{3.6}$$

$$\mathbb{P}(X' \mid X)\,,\ \mathbb{P}(Y' \mid Y)\,,\ \mathbb{P}(D' \mid D) \geq \kappa'\delta^2. \tag{3.7}$$

We need only refine two of the three sets X, Y and D above. Note that with uniformity in coordinate that is *not* refined, we then have that the set $S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'$ has about the expected number of points in it.

Our third lemma is a modification of one from a note of Ben Green [10]. It tells us that we can find a *uniform* sublattice on which A has increased density which is important since it is a required premise in applying the generalized von Neumann Lemma.

Lemma 3.8 (Uniformizing a Sublattice)**.** *Suppose that X, Y, D are as above. In addition*

1. *X, Y, and D satisfy (2.7);*
2. *$X' \subset X$, $Y' \subset Y$, $D' \subset D$, with $\mathbb{P}(X' \mid X) \geq c\delta^2$ and similarly for Y and D;*
3. *Either $X' = X$, $Y' = Y$ or $D' = D$;*
4. *$S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D'$;*
5. *$\mathbb{P}(A \mid S') = \delta + c\delta^2$;*
6. *$\dim(H) > C[\delta^4(\upsilon'')^2]^{-1}$, where $0 < \upsilon'' < 1$ is fixed.*

Then there exists $X'' \subset X'$, $Y'' \subset Y$, $D'' \subset D'$ and H', H'', translates of the same *subspace $H_0 \leq H$, so that*

$$\|X''\|_{\text{Uni}}\,, \ \|Y''\|_{\text{Uni}}\,, \ \|D''\|_{\text{Uni}} \leq \upsilon'', \tag{3.9}$$

$$\mathbb{P}(A \mid S'') \geq \delta + \tfrac{c}{2}\delta^2\,, \qquad S'' = X'' \times Y'' \cap X'' \overset{\text{diag}}{\times} D''\,, \tag{3.10}$$

$$\dim(H_0) \geq \dim(H) - C[\delta^4(\upsilon'')^2]^{-1}\,, \tag{3.11}$$

$$\mathbb{P}(X'' \mid H') \geq \kappa\delta^2\mathbb{P}(X' \mid H)\,. \tag{3.12}$$

An inequality similar to the last one also holds for Y'' and D''. In particular $\mathbb{P}(D'' \mid H' + H'') \geq \kappa\delta^2\mathbb{P}(D' \mid H)$

It is to be emphasized that H' and H'' are translates of the same subspace of $H_0 < H$, therefore $H' + H''$ is also a translate of H_0. Thus, after a joint translation of A, X, Y and D, we can assume that H' and H'' are in fact the same subspace H. It is this translation that is used in the iteration of the proof.

4 Proof of Theorem 1.1

Combining Lemmas 3.1 through 3.8 of the previous section yields the proof of Theorem 1.1. Since the proof is by recursion, we describe the conditional loop needed for the proof.

Proof. Initialize $X \leftarrow \mathbb{F}_2^n$, $Y \leftarrow \mathbb{F}_2^n$, $D \leftarrow \mathbb{F}_2^n$, $S \leftarrow \mathbb{F}_2^n \times \mathbb{F}_2^n$, $H \leftarrow \mathbb{F}_2^n$. Likewise $\delta_X, \delta_Y, \delta_D \leftarrow 1$. Fix a set A_0 with density δ_0 in $\mathbb{F}_2^n \times \mathbb{F}_2^n$. Initialize $A \leftarrow A_0$ and $\delta \leftarrow \mathbb{P}(A \mid S)$. Now we will iteratively apply the following steps:

1. If $\max\{\|f\|_{\square}\,, \|f\|_{\square,X}\,, \|f\|_{\square,Y}\} > \kappa\delta^{5/4}$, apply Lemma 3.4.

2. If X', Y' or D' is not $\upsilon = (\delta\delta_{X'}\delta_{Y'}\delta_D)^C$ uniform, apply Lemma 3.8. Suppose these sets are as in the Lemma: subsets $X'' \subset X'$, $Y'' \subset Y'$, $D'' \subset D'$ and affine subspaces $H', H'' \subset H$ containing X'', Y'', D''. After joint translation of X'', Y'', D'', A and H', H'', we can assume that $H' = H''$ and are subspaces of H.

3. Update variables:

$$X \leftarrow X'', \quad Y \leftarrow Y'', \quad D \leftarrow D'', \quad H \leftarrow H',$$
$$\delta_X \leftarrow \mathbb{P}(X'' \mid H'), \quad \delta_Y \leftarrow \mathbb{P}(Y'' \mid H'), \quad \delta_D \leftarrow \mathbb{P}(D'' \mid H'),$$
$$S \leftarrow X \times Y \cap X \overset{\text{diag}}{\times} D, \quad \delta \leftarrow \mathbb{P}(A \mid S).$$

4. Observe that the density of the incremented A on the set S has increased by at least $\kappa\delta_0^2$. Also, the incremented densities, δ_X, δ_Y, and δ_D have decreased by no more than $(\kappa\delta_0)^C$.

Once this loop stops, Lemma 3.1 applies—provided that the initial dimension is large enough—and we conclude that A has a corner. This iteration must stop in $\lesssim \delta_0^{-1}$ iterates, else the density of A on the sublattice would exceed one. Thus we need to be able to apply Lemmas 3.4 and 3.8 $\lesssim \delta_0^{-1}$ times. In order to do that, both X and H must be sufficiently large at each stage of the loop.

This requirement places several lower bounds on $N = 2^n$. The most stringent of these comes from the loss of dimensions in (3.11). Note that before the loop terminates, we can have δ_X as small as

$$\delta_X \geq (\kappa\delta_0)^{(\kappa\delta_0)^{-1}}.$$

In order to apply Lemma 3.8 at that stage, we need

$$N > 2^{(C\delta_0)^{-C\delta_0^{-1}}}.$$

From this condition we get the bound stated in the Theorem. □

5 Proof of Generalized von Neumann Lemma

Proof of Lemma 3.1. Define

$$\mathrm{T}(f, g, h) = \mathbb{E}_{x,s,y \in H} f(x, y) g(x + s, y) h(x, y + s). \tag{5.1}$$

Thus, $\mathrm{T}(A, A, A)$ is the expected number of corners in A. We show that this quantity is at least a fixed small multiple of $\delta_X^2\delta_Y^2\delta_D^2\delta^3$. By assumption (3.2), $C\delta_X^2\delta_Y^2\delta_D^2\delta^3 N^3 > \delta\delta_X\delta_Y\delta_D N^2$. The left hand side is the expected number of corners in A, while the right is the number of trivial corners in A—that is the number of points in A. Thus A is seen to have a corner.

Throughout the proof, it is convenient to make the substitution $s \to x + y + s$ in the expression for $\mathrm{T}(f, g, h)$, thus

$$\mathrm{T}(f, g, h) = \mathbb{E}_{x,y,s} f(x, y) g(y + s, y) h(x, x + s).$$

We are of course using the fact that we work in a field of characteristic two, but there is a similar substitution for any field.

We make the substitution $A = f + \delta S$ to get

$$\mathrm{T}(A,A,A) = \delta^3\,\mathrm{T}(S,S,S) \tag{5.2}$$
$$+\,\delta^2\,\mathrm{T}(f,S,S) + \delta^2\,\mathrm{T}(S,f,S) + \delta^2\,\mathrm{T}(S,S,f) \tag{5.3}$$
$$+\,\delta\,\mathrm{T}(f,f,S) + \delta\,\mathrm{T}(S,f,f) + \delta\,\mathrm{T}(f,S,f) \tag{5.4}$$
$$+\,\mathrm{T}(f,f,f)\,. \tag{5.5}$$

We have grouped the terms according to the number of f's that appear.

The main term is the right hand side of (5.2). Using (2.15), we see that $\delta^3\,\mathrm{T}(S,S,S) \geq \frac{1}{2}\delta_X^2\delta_Y^2\delta_D^2\delta^3$. (Observe the difference in the normalizations on the expectations in (2.15).)

All three terms in (5.3) are approximately zero, but we have to use uniformity to see this. For instance, we appeal to (2.5) and (2.7) to see that

$$\begin{aligned}\delta^2\,\mathrm{T}(S,f,S) &= \delta^2\mathbb{E}_{x,y,s\in H}X(x)Y(x+s)D(x+y)f(y+s,y)\\ &= \delta^2\delta_X\mathbb{E}_{x,y,s\in H}Y(x+s)D(x+y)f(y+s,y) + O(\upsilon')\\ &= \delta^2\delta_X\delta_Y\delta_D\mathbb{E}_{y,s\in H}f(y,s) + O(\upsilon')\\ &= O(\upsilon')\,.\end{aligned} \tag{5.6}$$

In this line and below, υ' is an unimportant function of υ which tends to zero.

The three terms in (5.4) are all controlled by appeal to Lemma 5.8. For instance, we have using the assumption about the maximal size of box norms (3.3)

$$|\delta\,\mathrm{T}(f,S,f)| \leq \upsilon' + \delta(\delta_X\delta_Y\delta_D)^2\|f\|_{\Box}\|f\|_{\Box,X} \leq \kappa\delta^{7/2}(\delta_X\delta_Y\delta_D)^2\,.$$

The other two terms admit a similar bound.

We bound the final term $\mathrm{T}(f,f,f)$, with the inequality Lemma 5.8:

$$\bigl|\mathrm{T}(f,f,f)\bigr| \leq \upsilon' + \delta^{1/2}(\delta_X\delta_Y\delta_D)^2\|f\|_{\Box,X}\|f\|_{\Box,Y} \leq \kappa(\delta_X\delta_Y\delta_D)^2\delta^3\,. \tag{5.7}$$

Using the hypothesis on the Box norm, (3.3), and the equation (5.6) will prove the Lemma. □

Lemma 5.8. *For $f_j \in \{f, S\}$ we have the estimate*

$$|\mathrm{T}(f_0,f_1,f_2)| \leq \upsilon' + (\delta_X\delta_Y\delta_D)^2 \cdot \begin{cases}\|f_0\|_2\cdot\|f_1\|_{\Box,Y}\cdot\|f_2\|_{\Box,X}\\ \|f_0\|_{\Box}\cdot\|f_1\|_{\Box,Y}\cdot\|f_2\|_2\\ \|f_0\|_{\Box}\cdot\|f_1\|_2\cdot\|f_2\|_{\Box,X}\end{cases} \tag{5.9}$$

Here, $\|f\|_2 = \delta^{1/2}$ while $\|S\|_2 = 1$.

Proof. We prove an instance of the claimed inequalities:

$$|\mathrm{T}(f_0,f_1,f_2)| \leq \upsilon' + \|f_0\|_2(\delta_X\delta_Y\delta_D)^2\|f_1\|_{\Box,Y}\cdot\|f_2\|_{\Box,X}\,. \tag{5.10}$$

By a change of basis, this inequality implies the other two.

Apply Cauchy Schwartz once, in the variables (x,y), to get

$$|\mathrm{T}(f_0,f_1,f_2)| \leq \bigl(\mathbb{E}_{x,y}f_0(x,y)^2\bigr)^{1/2}\cdot\mathrm{U}^{1/2}$$

The first term on the right is no more than $\|f_0\|_2(\delta_X\delta_Y\delta_D)^{1/2}$. As for the second term, it is

$$\begin{aligned}\mathrm{U} &\stackrel{\text{def}}{=} \mathbb{E}_{x,y}D(x+y)\big|\mathbb{E}_s f_1(y+s,y)f_2(x,x+s)\big|^2\\ &= \mathbb{E}_{y,s,s'}\big\{\mathbb{E}_x D(x+y)f_1(y+s,y)f_1(y+s',y)\big\}\cdot\{f_2(x,x+s)f_2(x,x+s')\}\end{aligned}$$

Note that in the definition of U, we have inserted the term $D(x+y)$, which arises from $A(x,y)$.[2] We apply Proposition 2.4 to replace $D(x+y)$ by δ_D and then use Cauchy Schwartz again in the variables x, x' to get

$$\begin{aligned}\mathrm{U} &= \delta_D(\mathrm{U}_1\cdot\mathrm{U}_2)^{1/2}+O(\upsilon');\\ \mathrm{U}_1 &\stackrel{\text{def}}{=} \mathbb{E}_{y,y'\in H}\big|\mathbb{E}_{s\in H}f_1(y+s,y)f_1(y'+s,y')\big|^2=\delta_X^4\delta_Y^2\delta_D^2\|f_1\|_{\Box,Y}^4\,;\\ \mathrm{U}_2 &\stackrel{\text{def}}{=} \mathbb{E}_{x,x'\in H}|\mathbb{E}_{s\in H}f_2(x,x+s)f_2(x',x'+s)|^2=\delta_X^2\delta_Y^4\delta_D^2\|f_2\|_{\Box,X}^4\,.\end{aligned}$$

A change of variables permits the identification of U_1 and U_2. This completes the proof of (5.7).

□

6 Proof of Density Increment Lemma

Proof of Lemma 3.4. We prove this assertion: Fix $c<1$. There is a constant κ so that the following holds. Assume that $\|f\|_\Box \ge c\delta^{5/4}$, and show that there are subsets $X'\subset X$, $Y'\subset Y$ so that

$$\begin{gathered}\mathbb{P}(X'\mid X)\,,\,\mathbb{P}(Y'\mid Y)\ge\kappa\delta^2\,,\\ \mathbb{P}(A\mid X'\times Y'\cap X'\overset{\text{diag}}{\times}D)\ge\delta+\kappa\delta^2\,.\end{gathered}$$

We emphasize that the box norm we use is the one given in (2.8). And we will *not* refine the diagonal coordinate. This is one instance of the Lemma, which by a change of coordinates, this will prove the Lemma as stated.

We can assume that the fibers above points $x\in X$, and $y\in Y$, typically behave as expected. Namely, we assume that

$$\begin{aligned}\mathbb{E}_{x\in X}\big|\mathbb{E}_{y\in Y}f(x,y)\big|^2&\le\kappa^3\delta^2\delta_D^2\,,\\ \mathbb{E}_{y\in Y}\big|\mathbb{E}_{x\in X}f(x,y)\big|^2&\le\kappa^3\delta^2\delta_D^2\,.\end{aligned}\tag{6.1}$$

for otherwise, we can apply Lemma 6.25 to conclude the Lemma.

For a point $(x,y)\in A$, consider $N_x\stackrel{\text{def}}{=}\{y'\mid(x,y')\in A\}$, $N_y\stackrel{\text{def}}{=}\{x'\mid(x',y)\in A\}$. These are the neighbors of x and of y, respectively. We need to consider points for which these sets are about as big as they should be. Set

$$\begin{aligned}X''&=\{x\in X\mid\mathbb{P}(N_x|Y)\ge\kappa\delta^5\delta_D\,,\,\big|\mathbb{E}_{y\in Y}f(x,y)\big|<\kappa\delta\delta_D\}\,,\\ Y''&=\{y\in Y\mid\mathbb{P}(N_y|X)\ge\kappa\delta^5\delta_D\,,\,\big|\mathbb{E}_{x\in X}f(x,y)\big|<\kappa\delta\delta_D\}\,.\end{aligned}$$

It is clear that these sets are most of X and Y respectively. In particular, in view of (6.1) we have

$$\mathbb{P}(X''\mid X)\,,\,\mathbb{P}(Y''\mid Y)\ge1-\delta^2\,.\tag{6.2}$$

[2]Without this term, we would not get the right power of δ_D in our estimates.

Clearly, we can assume that

$$\mathbb{E}_{\substack{x\in X''\\ y\in Y''}} f(x,y) < \kappa\delta^2\delta_D \tag{6.3}$$

for otherwise we already proved the Lemma. But, it is also the case that we can assume

$$-\kappa\delta^4\delta_D < \mathbb{E}_{\substack{x\in X''\\ y\in Y''}} f(x,y) \tag{6.4}$$

for if this inequality fails, we apply Lemma 6.26 to conclude the proof of the Lemma.

We further note that we have

$$\mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} f(x,y)f(x',y)f(x,y')f(x',y') \geq \tfrac{c^4}{2}\delta^5\,. \tag{6.5}$$

If we were taking the expectation over X and Y, this would follow from the assumption that $\|f\|_{\Box} \geq c\delta^{5/4}$. As we are taking the expectation over X'' and Y'', we need to show that taking the expectation over the complement of X'' we get an appropriate upper bound, namely

$$\Big|\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)f(x,y)f(x',y)f(x,y')f(x',y')\Big| \leq \upsilon' + \kappa\delta^5\delta_D^4\,. \tag{6.6}$$

Three similar inequalities hold, so using the assumption that $\|f\|_{\Box} \geq c\delta^{5/4}$, taking $0 < \kappa < \frac{c^4}{8}$ we will see that (6.5) holds.

To see (6.6), first observe that by definition of X'',

$$\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)N_x(y)D(x+y')D(x'+y)D(x'+y') \leq \upsilon' + \kappa\delta^5\delta_D^4\,. \tag{6.7}$$

This does not prove (6.6) since $f(x,y)$ is not supported on $N_x(y)$. But, we also have the similar inequality

$$\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)D(x+y)N_x(y')D(x'+y)D(x'+y')\upsilon' + \kappa\delta^5\delta_D^4\,. \tag{6.8}$$

Of course $f = \delta S - A$, so we can estimate

$$\begin{aligned}
\delta\Big|&\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)D(x+y)f(x,y')f(x',y)f(x',y')\Big| \\
&\leq \upsilon' + \kappa\delta^5\delta_D^4 + \delta^2\Big|\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)D(x+y)D(x+y')f(x',y)f(x',y')\Big| \\
&\leq \upsilon' + \kappa\delta^5\delta_D^4 + \delta^2\delta_D^2\Big|\mathbb{E}_{\substack{x,x'\in X\\ y,y'\in Y''}} \mathbf{1}_{X-X''}(x)f(x',y)f(x',y')\Big| \\
&\leq \upsilon' + \kappa\delta^5\delta_D^4 + \delta^4\delta_D^2\mathbb{E}_{x'\in X}\Big|\mathbb{E}_{y\in Y''}f(x',y)\Big|^2 \\
&\leq \upsilon' + \kappa\delta^5\delta_D^4 + \kappa\delta^6\delta_D^4\,.
\end{aligned}$$

This completes the proof of (6.6).

To find the subset on which A has increased density, we consider any subset of the form $N_y \times N_x \cap N_y \overset{\text{diag}}{\times} D$, where $y \in Y''$ and $x \in X''$. Clearly we are interested in the largest increase in density, for which we estimate

$$\sup_{\substack{x\in X'',y\in Y''\\(x,y)\in A}} \frac{|A\cap\{N_y\times N_x\cap S''\}|}{|N_y\times N_x\cap S''|}\ge \frac{\mathrm{Q}''(A,A,A,A)}{\mathrm{Q}''(A,A,A,S)}\,, \tag{6.9}$$

$$S''\overset{\text{def}}{=} X''\times Y''\cap X''\overset{\text{diag}}{\times} D\,,$$

$$\mathrm{Q}''(f_0,f_1,f_2,f_3)\overset{\text{def}}{=}\mathbb{E}_{\substack{x,x'\in X''\\y,y''\in Y''}} f_0(x,y)f_1(x',y)f_2(x,y')f_3(x',y')\,.$$

Here, we note that for $(x,y)\in X''\times Y''$ we have $|A\cap\{N_y\times N_x\cap S''\}|>0$, so that we are not dividing by zero in (6.9). By definition, we have

$$\mathbb{P}(N_y\times N_x\mid X\times Y)\ge k^2\delta^{10}\delta_D^2\,.$$

And by uniformity, we have

$$\begin{aligned}\mathbb{P}(N_y\times N_x\cap S'') &= \mathbb{E}_{x',y'\in H}N_y(x')N_x(y')D(x'+y')\\ &\ge \upsilon'+k^2\delta^{10}\delta_D^3\delta_X\delta_Y>0\,.\end{aligned}$$

In the last inequality, υ' is a function of the uniformity constant.

There is a gain of regularity in passing to the expectations on the right hand side of (6.9). Expand $A=f+\delta S$ in the last place in $\mathrm{Q}(A,A,A,A)$ to see that

$$\frac{\mathrm{Q}''(A,A,A,A)}{\mathrm{Q}''(A,A,A,S)}\ge\delta+\frac{\mathrm{Q}''(A,A,A,f)}{\mathrm{Q}''(A,A,A,S)}$$

and so we should show that the last fraction is at least $c'\delta^2$, which we do by showing that

$$\mathrm{Q}''(A,A,A,f)\ge c'\delta^5\delta_D^4\,, \tag{6.10}$$

$$0<\mathrm{Q}''(A,A,A,S)<10\delta^3\delta_D^4\,. \tag{6.11}$$

This we will do, assuming one more condition. If this last condition fails, we will get a density increment of δ^2.

Define

$$\alpha_X^2\overset{\text{def}}{=}\delta\delta_D^2\mathbb{E}_{x\in X''}\big|\mathbb{E}_{y\in Y''}f(x,y)\big|^2\,,\qquad \alpha_Y^2\overset{\text{def}}{=}\delta\delta_D^2\mathbb{E}_{y\in Y''}\big|\mathbb{E}_{x\in X''}f(x,y)\big|^2\,,$$

By (6.1) and the definitions of X'' and Y'', one can see that these two quantities are at most $\kappa\delta^3\delta_D^4$, and hence are only a small fraction of the major term in the considerations below.

Using $A=\delta S+f$, $\mathrm{Q}''(A,A,A,S)$ has the expansion

$$\mathrm{Q}''(A,A,A,S)=\delta^3\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}D(x+y)D(x+y')D(x'+y)D(x'+y') \tag{6.12}$$

$$+3\delta^2\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}f(x,y)D(x+y')D(x'+y)D(x'+y') \tag{6.13}$$

$$+\delta\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}f(x,y)f(x,y')D(x'+y)D(x'+y') \tag{6.14}$$

$$+\delta\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}D(x+y)f(x,y')f(x',y)D(x'+y') \tag{6.15}$$

$$+\delta\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}f(x,y)D(x+y')f(x',y)D(x'+y') \tag{6.16}$$

$$+\mathbb{E}_{\substack{x,x'\in X''\\y,y'\in Y''}}f(x,y)f(x,y')f(x',y)D(x'+y')\,. \tag{6.17}$$

Clearly, the uniformity in D is relevant. The right hand side of (6.12) is $\delta^3\delta_D^4$ plus a term controlled by uniformity; the term in (6.13) is, by (6.4), at least $-3\kappa\delta^6$, plus a term controlled by uniformity; (6.14) is α_X^2, plus a term controlled by uniformity; (6.15), by (6.3) and (6.4), obeys the inequality

$$(6.15) < \upsilon' + \delta\delta_D^2\big|\mathbb{E}_{x\in X, y\in Y} f(x,y)\big|^2 < \upsilon' + \kappa^2\delta^3\delta_D^4\,;$$

(6.16) is approximately in α_Y^2; while the last term (6.17) is not one that admits an obvious control, and we write

$$(6.17) = \upsilon' + \Delta\,, \qquad \Delta \stackrel{\text{def}}{=} \delta_D \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} f(x,y)f(x,y')f(x',y)\,. \tag{6.18}$$

We can assume that $|\Delta| \le \kappa\delta^3\delta_D^4$, otherwise we apply Lemma 6.31 to finish the proof of the Lemma, getting a density increment of the order of δ^2. This proves (6.11).

The expression $\mathrm{Q}(A,A,A,f)$ admits a very similar expansion.

$$\begin{aligned}
\mathrm{Q}''(A,A,A,f) &= \delta^3 \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} D(x+y)D(x+y')D(x'+y)f(x',y') && (6.19)\\
&\quad + \mathrm{Q}_2 + \mathrm{Q}_3 && (6.20)\\
&\quad + \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} f(x,y)f(x,y')f(x',y)f(x',y')\,, && (6.21)\\
\mathrm{Q}_2 &\stackrel{\text{def}}{=} \delta^2\big\{\mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} f(x,y)D(x+y')D(x'+y)f(x',y') && (6.22)\\
&\quad + \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} D(x+y)f(x,y')D(x'+y)f(x',y')\\
&\quad + \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} D(x+y)D(x+y')f(x',y)f(x',y')\big\}\,,\\
\mathrm{Q}_3 &\stackrel{\text{def}}{=} \delta\{\mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} D(x+y)f(x,y')f(x',y)f(x',y') && (6.23)\\
&\quad + \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} f(x,y)D(x+y')f(x',y)f(x',y')\\
&\quad + \mathbb{E}_{\substack{x,x'\in X''\\ y,y'\in Y''}} D(x+y)f(x,y')f(x',y)f(x',y')\}\,.
\end{aligned}$$

The right hand side of (6.19) is greater than $-\kappa\delta^3\delta_D^4$, plus a term controlled by uniformity by (6.3) and (6.4); the term in (6.21) is the box norm, which by (6.5) is at least $c'\delta^5\delta_D^4$; here of course, $c' > 0$ is fixed in advance, and we take κ much smaller than c'; the terms which make up the definition of Q_2 all involve two f's, and after taking uniformity into account each individual term is positive, and so can be ignored as we are obtaining a lower bound for $\mathrm{Q}(A,A,A,f)$; the three terms in the definition of Q_3 are all of the form $\upsilon' + \Delta$ where Δ is defined in (6.18). In particular, we have already assumed $|\Delta| \le \kappa\delta^3\delta_D^4$. For $0 < \kappa$ sufficiently small, this proves (6.10), and so the proof of this Lemma.

□

We use the following simple variant of the Paley Zygmund inequality. It states, in particular, that a random variable, bounded in L^∞ norm by one, with mean zero, and standard deviation σ, must be at least a constant multiple of σ on a set of probability proportional to the variance σ^2.

Proposition 6.24. *Let $1 < p < \infty$. Then there is a $c > 0$ so that for all $1 \leq p < \infty$, and random variables Z with $-1 \leq Z \leq 1$, $\mathbb{E}Z = 0$, and $\mathbb{E}|Z|^p = \sigma^p$. Then, $\mathbb{P}(Z > c\sigma^p) \geq c_p \sigma^p$.*

Proof. In fact, we can take $c = \frac{1}{5}$. We assume that the conclusion is false and seek a contradiction. Since $\mathbb{E}Z = 0$ we have

$$\begin{aligned} -\mathbb{E}Z\mathbf{1}_{\{Z<0\}} &= \mathbb{E}Z\mathbf{1}_{\{Z>0\}} \\ &\leq \mathbb{P}(Z > c\sigma^p) + \mathbb{E}Z\mathbf{1}_{\{0<Z<c\sigma^p\}} \\ &\leq 2c\sigma^p \,. \end{aligned}$$

With this, and the fact that $p \geq 1$ while Z is bounded by 1, we can now estimate

$$\begin{aligned} \sigma^p = \mathbb{E}|Z|^p &= \mathbb{E}|Z|^p\mathbf{1}_{\{Z<0\}} + \mathbb{E}|Z|^p\mathbf{1}_{\{Z>0\}} \\ &\leq 2\mathbb{E}Z\mathbf{1}_{\{Z>0\}} \leq 4c\sigma^p \,. \end{aligned}$$

This is a contradiction. □

Lemma 6.25. *If it is the case that*

$$\mathbb{E}_{x\in X}\left|\mathbb{E}_{y\in Y} f(x,y)\right|^2 \geq c\delta^2\delta_D^2$$

then there is a set $X' \subset X$ for which $\mathbb{P}(X' \mid X) \geq \frac{c}{12}\delta^2$ on which we have

$$\mathbb{P}(A \mid X' \times Y \cap X' \overset{\text{diag}}{\times} D) \geq \delta + \tfrac{c}{12}\delta^2 \,.$$

Proof. The set of $x \in X$ for which

$$2\delta_D \leq \left|\mathbb{E}_{y\in Y} f(x,y)\right| \leq \left|\mathbb{E}_{y\in Y} D(x+y)\right|$$

has probability that is controlled by uniformity, and hence is negligible. Thus, the Lemma follows immediately from the Paley Zygmund inequality. □

Lemma 6.26. *Suppose that there is a sublattice $S'' = X'' \times Y'' \cap X'' \overset{\text{diag}}{\times} D$ with*

$$\mathbb{P}(X'' \mid X)\,,\, \mathbb{P}(Y'' \mid Y) \geq 1 - \lambda\,, \tag{6.27}$$

$$\mathbb{P}(A \mid S'') \leq \delta - \tau\,. \tag{6.28}$$

Then, there is a sublattice $S' = X' \times Y' \cap X' \overset{\text{diag}}{\times} D$ on which

$$\mathbb{P}(X' \mid X)\,,\, \mathbb{P}(Y' \mid Y) \geq \tfrac{1}{2}\tau\,, \tag{6.29}$$

$$\mathbb{P}(A \mid S') \geq \delta + \kappa\tau\lambda^{-1}\,. \tag{6.30}$$

Proof. Notice that the density of A on $S - S''$ must be strictly larger than δ. Namely,

$$\begin{aligned} \mathbb{P}(A \mid S - S'') &\geq \frac{\delta[1 - \mathbb{P}(S'' \mid S)] + \tau\mathbb{P}(S'' \mid S)}{1 - \mathbb{P}(S'' \mid S)} \\ &\geq \delta + \tau\frac{\mathbb{P}(S'' \mid S)}{1 - \mathbb{P}(S'' \mid S)} \\ &\geq \delta + \kappa\tau\lambda^{-1}\,. \end{aligned}$$

But, clearly $S - S''$ is a union of three sublattices, and on one of these three, A must have density at least $\delta + \kappa\tau\lambda^{-1}$. We must have $\mathbb{P}(X'' \mid X) \leq 1 - \frac{1}{2}\tau$, otherwise we contradict (6.28). This finishes the proof. □

Lemma 6.31. *Fix $c>0$. If it is the case that*

$$\bigl|\mathbb{E}_{x,y} f(x,y)\mathbb{E}_{x'} f(x',y)\,\mathbb{E}_{y'} f(x,y')\bigr| \geq c\delta^3\delta_D^3 \tag{6.32}$$

Then, there is a constant $c'=c'(c)$, $X'\subset X$, $Y'\subset Y$ with

$$\mathbb{P}(X' \mid X)\,,\ \mathbb{P}(Y' \mid Y) \geq c'\delta^2\,, \tag{6.33}$$

$$\mathbb{P}(A \mid X'\times Y' \cap X' \overset{\text{diag}}{\times} D) \geq \delta + c'\delta^2\,. \tag{6.34}$$

Proof. By uniformity in D, we have

$$\mathbb{P}\bigl(\bigl|\mathbb{E}_y f(x,y)\bigr| > 2\delta_D\bigr) \leq \upsilon'\,.$$

That is, the effective L^∞ bound on $\mathbb{E}_y f(x,y)$ is $2\delta_D$.

Hölders inequality and the assumption (6.32) implies

$$\begin{aligned} c\delta^3\delta_D^3 &\leq \bigl[\mathbb{E}_{x,y}|f(x,y)|^3\bigr]^{1/3}\cdot\Bigl[\mathbb{E}_{x,y}D(x+y)\bigl|\mathbb{E}_y f(x,y)\bigr|^{3/2}\cdot\bigl|\mathbb{E}_x f(x,y)\bigr|^{3/2}\Bigr]^{2/3} \\ &\leq 2\delta^{1/3}\delta_D^{1/3}\Bigl[\mathbb{E}_{x,y}D(x+y)\bigl|\mathbb{E}_{y'} f(x,y')\bigr|^{3/2}\cdot\bigl|\mathbb{E}_{x'} f(x',y)\bigr|^{3/2}\Bigr]^{2/3} \\ &\leq \upsilon' + 2\delta^{1/3}\delta_D\Bigl[\mathbb{E}_x\bigl|\mathbb{E}_{y'} f(x,y')\bigr|^{3/2}\cdot\mathbb{E}_y\bigl|\mathbb{E}_{x'} f(x',y)\bigr|^{3/2}\Bigr]^{2/3}. \end{aligned}$$

Here it is essential that we insert the term $D(x+y)$ when we apply Hölders inequality. Note that uniformity in D is then used to obtain the full power of δ_D in the last line.

Thus, we must have e.g.

$$\mathbb{E}_y\bigl|\mathbb{E}_x f(x,y)\bigr|^{3/2} \geq c''\delta^2\delta_D^{3/2}$$

Then, the conclusion of our Lemma follows from the Paley Zygmund inequality.

□

7 Proof of Lemma 3.8

We include a proof of this Lemma as we are requiring a more than is claimed in e.g. Ben Green's survey [10]. In particular, we claim that all three sets X,Y,D can be uniformized.

Let us indicate the central way that uniformity is used in this proof. See [10, Lemma 3.4(1)].

Proposition 7.1. *For any subset $X\subset H$, there is a partition of H into two affine subspaces H' and H'' for which*

$$\tfrac{1}{2}\bigl\{\mathbb{P}(X \mid H')^2 + \mathbb{P}(X \mid H'')^2\bigr\} \geq \mathbb{P}(X \mid H)^2 + \tfrac{1}{8}\|X\|_{\text{Uni}}^2\,. \tag{7.2}$$

The following more technical Lemma describes a key inductive procedure in the proof.

Lemma 7.3. *Let $0<t,u<1$ be positive parameters. Let $U\subset H$. Suppose that $\dim(H)\geq 10(tu^2)^{-1}$. Then, there is a partition $\mathcal{P}$ of H, so that writing $\mathcal{P}=\mathcal{U}\cup\mathcal{N}$,*

1. *All $H'\in\mathcal{P}$ have dimension $\dim(H')\geq\dim(H)-4(tu^2)^{-1}$;*
2. *$\|U\|_{\text{Uni}}\leq u$ for all $H'\in\mathcal{U}$;*

3. For $\delta_U = \mathbb{P}(U \mid H)$,

$$\mathbb{P}\Big(\bigcup\{H' \mid H' \in \mathcal{N}\} \mid H\Big) \le t\delta_U \,.$$

Remark 7.4. We will not use this Lemma as stated, but rather the more complicated variant that follows. We include this statement and proof for clarity's sake. Some of the notation above is taken from Ben Green's note [10]. In particular, $\mathcal{U}$ is for 'uniform', and $\mathcal{N}$ is for 'non–uniform.'

Proof. The proof is an inductive procedure, though we will not define the collections of hyperspaces $\mathcal{U}$ and $\mathcal{N}$ until the conclusion of the iteration. Initialize variables

$$\mathcal{Q} \leftarrow \{H\}.$$

Also initialize a counter $m \leftarrow 0$. Given $\mathcal{Q}$, define $\mathcal{R}$ to be those $H' \in \mathcal{Q}$ for which $\|U \cap H'\|_{\mathrm{Uni}} \ge u$. `WHILE`

$$\mathbb{P}\Big(\bigcup\{H' \mid H' \in \mathcal{R}\}\Big) \ge t\delta_U$$

update $m \leftarrow m+1$. And for each $H' \in \mathcal{R}$ apply Lemma 7.3 to H'. Thus, $H' = H'_1 \cup H'_2$ for which

$$\tfrac{1}{2}\big\{\mathbb{P}(U \mid H'_1)^2 + \mathbb{P}(U \mid H'_2)^2\big\} \ge \mathbb{P}(U \mid H')^2 + \tfrac{1}{4}u^2 \,. \tag{7.5}$$

Update $\mathcal{Q} \leftarrow \mathcal{Q} - \mathcal{R} \cup \{H'_1, H'_2 \mid H' \in \mathcal{R}\}$. When each $H' \in \mathcal{R}$ has been so split, set $\mathcal{P}_m \leftarrow \mathcal{Q}$. The WHILE loop then repeats.

Once the WHILE loop stops, return the value of m, the sequence of partitions $\mathcal{P}_1, \ldots, \mathcal{P}_m$, and $\mathcal{Q} = \mathcal{P}_m$. Define

$$\mathcal{U} \stackrel{\mathrm{def}}{=} \{H' \in \mathcal{Q} \mid \|X \cap H'\|_{\mathrm{Uni}} \le u\} \,,$$

and $\mathcal{N} \stackrel{\mathrm{def}}{=} \mathcal{Q} - \mathcal{U}$. The subspaces in $\mathcal{Q}$ have dimension at least $\dim(H) - m$.

Once the `WHILE` loop stops, it is clear that the conclusions (2) and (3) of the Lemma hold. We concern ourselves with the first conclusion about the dimension of the subspaces involved. Observe that at each iterate of the `WHILE` loop, the dimensions of the partition can only lose one dimension. It suffices to estimate the number of iterates of the Loop, which is the counter m.

Consider the quantities

$$S_j = \sum_{H' \in \mathcal{P}_j} \mathbb{P}(U \mid H')^2 \mathbb{P}(H' \mid H) = \mathbb{E}\, \mathbb{E}(U \mid \mathcal{P}_j)^2 \,,$$

which are the mean square densities of U relative to the partitions $\mathcal{P}_j$. (Here we are relying on the usual notations for conditional second moments.) Obviously these quantities are less than $\delta_U = \mathbb{P}(U \mid H)$. Each H' which is split in (7.5), the mean square density in H' is increased by $\frac{1}{4}u^2$. And, this takes place, at each iterate of the loop, in a portion of the whole space H that is at least probability $t\delta_U$. From this, we see that $S_{j+1} \ge S_j + \frac{1}{4}t\delta_U u^2$. Consequently, $mt\delta_U u^2 \le S_m \le \delta_U$. This proves (1), and so the proof is complete. □

The more technical statement that we need is as follows. Whereas in the first Lemma, we have a single subset $U \subset H$ and construct a 'good' partition, in this statement we have three subsets $U_1, U_2, U_3 \subset H$ and construct a single 'good' partition in the product space $H \times H$.

Lemma 7.6. *Let $0 < t, u < 1$ be positive parameters. Let $U_1, U_2, U_3 \subset H$. Suppose that $\dim(H) \geq 10(tu^2)^{-1}$. Then, there is a partition $\mathcal{P}$ of $H \times H$, so that writing $\mathcal{P} = \mathcal{U} \cup \mathcal{N}_1 \cup \mathcal{N}_2 \cup \mathcal{N}_3$, we have the following.*

1. *For all $V_1 \times V_2 \in \mathcal{P}$,*
$$\dim(V_1) = \dim(V_2) \geq \dim(H) - 4(tu^2)^{-1} .$$
2. *Moreover, V_1, V_2 as affine subspaces of H, are translates of each other.*
3. *For all $V_1 \times V_2 \in \mathcal{U}$, $\max_{j=1,2} \|U_j \cap V_j\|_{\mathrm{Uni}} \leq u$ and $\|U_3 \cap (V_1 + V_2)\|_{\mathrm{Uni}} \leq u$.*
4. $$\mathbb{P}\Big(\bigcup\{V_j \mid V_1 \times V_2 \in \mathcal{N}_j\} \mid H\Big) \leq t\delta_{U_j} \quad j = 1, 2 \,,$$
$$\mathbb{P}\Big(\bigcup\{V_1 + V_2 \mid V_1 \times V_2 \in \mathcal{N}_3\} \mid H\Big) \leq t\delta_{U_3} \,.$$

Remark 7.7. If we did not insist on the second conclusion, that the relevant subspaces be translates of one another, we could simply take a product of partitions arising from Lemma 7.3. Due to the prominent role of the diagonals in our problem, this is an essential point for us.

Proof. We describe the recursive procedure. Initialize variables
$$\mathcal{Q} \leftarrow \{H \times H\}.$$
and three counters $m_j \leftarrow 0$, $j = 1, 2, 3$. The recursive procedure we describe will return the partition we need, as well as some auxiliary data that we need to prove the Lemma.

Given $\mathcal{Q}$, define $\mathcal{R}_j$, $j = 1, 2, 3$ by
$$\mathcal{R}_j = \{V_1 \times V_2 \in \mathcal{Q} \mid \|U_j \cap V_j\|_{\mathrm{Uni}} \geq u\} \,, \qquad j = 1, 2$$
$$\mathcal{R}_3 = \{V_1 \times V_2 \in \mathcal{Q} \mid \|U_3 \cap (V_1 + V_2)\|_{\mathrm{Uni}} \geq u\} \,.$$

For $j = 1, 2$
$$\textsf{IF} \qquad \mathbb{P}\Big(\bigcup\{V_j \mid V_1 \times V_2 \in \mathcal{R}_j\}\Big) \geq t\delta_{U_j} \tag{7.8}$$
$\textsf{THEN}$ update $m_j \leftarrow m_j + 1$. $\textsf{WHILE}$ $\mathcal{R}_j \neq \emptyset$

- Apply Lemma 7.3 to V_j. Thus, $V_j = V_j' \cup V_j''$ for which
$$\tfrac{1}{2}\big\{\mathbb{P}(U_j \mid V_j')^2 + \mathbb{P}(U_j \mid V_j'')^2\big\} \geq \mathbb{P}(U_j \mid V_j)^2 + \|U_j \cap V_j\|_{\mathrm{Uni}}^2 \,. \tag{7.9}$$
- (This point is a departure from the previous proof.) Let k be the other index.[3] Since V_k is a translate of V_j, we can choose translates V_k' and V_k'' of V_j' which also partition V_k.
- At this point, let us observe that we will have $V_1'+V_2'$, $V_1'+V_2''$, $V_1''+V_2'$, $V_1''+V_2'' \subset V_1+V_2$. And indeed, all of these subspaces are affine translates of V_1', hence these four subspaces are made of two equal pairs, which partition $V_1 + V_2$.

[3]That is, if $j = 1$, then $k = 2$. Notice that we are enforcing a partition on V_k that does not necessarily have anything to do with increasing conditional variances.

- Update

$$\mathcal{Q} \leftarrow (\mathcal{Q} - \{V_1 \times V_2\}) \cup \{V_1' \times V_2', V_1'' \times V_2', V_1' \times V_2'', V_1'' \times V_2''\}$$
$$\mathcal{R}_j \leftarrow \mathcal{R}_j - \{V_1 \times V_2\}.$$

When the WHILE loop stops, define $\mathcal{P}_{m_1+m_2+m_3} \overset{\text{def}}{=} \mathcal{Q}$ and $\sigma_j(m_j) = m_1 + m_2 + m_3$.

We now describe the procedure as applied to the set U_3. That the products in $\mathcal{Q}$ are products of translates of the same subspace plays a critical role in this formulation. In particular, for $V_1 \times V_2 \in \mathcal{P}$, $V_1 + V_2$ is a translate of V_1 (and V_2).

$$\text{IF} \qquad \mathbb{P}\Big(\bigcup\{V_1 + V_2 \mid V_1 \times V_2 \in \mathcal{R}_3\}\Big) \geq t\delta_{U_3} \tag{7.10}$$

THEN update $m_3 \leftarrow m_3 + 1$. For each $V_1 \times V_2 \in \mathcal{R}_3$

- Apply Lemma 7.3 to $W = V_1 + V_2$. Thus, $W = W' \cup W''$ for which

$$\tfrac{1}{2}\big\{\mathbb{P}(U_3 \mid W')^2 + \mathbb{P}(U_3 \mid W'')^2\big\} \geq \mathbb{P}(U_3 \mid W)^2 + \|U_3 \cap W\|_{\text{Uni}}^2 . \tag{7.11}$$

- Since the spaces V_j are translates of W, we can choose translates V_j' and V_j'' of W' which also partition V_j, $j = 1, 2.$.

- Update

$$\mathcal{Q} \leftarrow (\mathcal{Q} - \{V_1 \times V_2\}) \cup \{V_1' \times V_2', V_1'' \times V_2', V_1' \times V_2'', V_1'' \times V_2''\}$$
$$\mathcal{R}_3 \leftarrow \mathcal{R}_3 - \{V_1 \times V_2\}.$$

Repeat these steps until $\mathcal{R}_3$ is exhausted. Then, define $\mathcal{P}_{m_1+m_2+m_3} \overset{\text{def}}{=} \mathcal{Q}$ and $\sigma_3(m_j) = m_1 + m_2 + m_3$.

Iteratively apply the three conditionals, two in (7.8) and one in (7.10). STOP when all three conditionals fail. Return the values of m_j, the sequence of partitions of $\mathcal{P}_j$ for $1 \leq j \leq m_1 + m_2$, the 'partition times' $\{\sigma_j(n) \mid 1 \leq n \leq m_j\}$ and collection $\mathcal{Q}$.

Define

$$\mathcal{N}_j \overset{\text{def}}{=} \{V_1 \times V_2 \in \mathcal{Q} \mid \|U_j \cap V_j\|_{\text{Uni}} \geq u\}\,, \qquad j = 1, 2$$

and similarly define $\mathcal{N}_3$. Set $\mathcal{U} = \mathcal{Q} - \mathcal{N}_1 - \mathcal{N}_2 - \mathcal{N}_3$.

All subspaces chosen in this way have dimension at least equal to $\dim(H) - m_1 - m_2 - m_3$. We need only provide upper bounds on the m_j, as all the other claims of the Lemma are evident from the construction.

We claim that $m_j \leq (\delta t u^2)^{-1}$, and for this, we can use the previous proof, with one additional fact. Let π be a finite partition of a probability space, and suppose that π' refines π. Then, for any random variable Z we have

$$\mathbb{E}(\mathbb{E}(Z \mid \pi)^2) \leq \mathbb{E}(\mathbb{E}(Z \mid \pi')^2)$$

Here, we are using a standard notation for conditional expectation given π. This is a simple martingale fact. Indeed, let $Y = \mathbb{E}(Z \mid \pi') - \mathbb{E}(Z \mid \pi)$, and observe that $\mathbb{E}\, Y \cdot \mathbb{E}(Z \mid \pi) = 0$.

The random variable in question is U_j. Set

$$S_{j,n} = \mathbb{E}\,\mathbb{E}(U_j \mid \mathcal{P}_n)^2\,, \qquad 1 \leq n \leq m \overset{\text{def}}{=} m_1 + m_2 + m_3\,.$$

We have just seen that the $S_{j,n}$ are increasing in $1 \le n \le m$. They are obviously at most δ_{U_j}. And by construction, and in particular using (7.5) and (7.11), at each time at which the corresponding conditional is satisfied, we increase these numbers by a definite amount. At each iterate of the loop, in a portion of the whole space H that is at least probability $t\delta_{U_j}$, where this uniformity constant is at least u. From this, we see that

$$\delta_{U_j} \ge S_{j,\sigma_j(\ell)} \ge S_{j,\sigma_j(\ell-1)} + tu^2\delta_{U_j}\,, \qquad 1 \le \ell \le m_j\,,\ j = 1,2,3\,.$$

Therefore, $m_j \le (tu^2)^{-1}$. The proof is complete. $\square$

Proof of Lemma 3.8. Let us assume that e. g. $D' = D$. Apply Lemma 7.6 to the sets $X' = U_1$, $Y' = U_2$, $D = D' = U_3$, with $t = \frac{c}{16}\delta^2$ and $u = \upsilon''$. Let $\mathcal{P}$ be the partition of $H \times H$ that this Lemma gives us.

Define two subsets $\mathcal{E}_1$ and $\mathcal{E}_2$ of $\mathcal{P}$ by

$$\mathcal{E}_1 \overset{\text{def}}{=} \{V_1 \times V_2 \in \mathcal{P} \mid \mathbb{P}(X' \cap V_j \mid V_j) \le t\}\,.$$

We define $\mathcal{E}_2$ similarly, with X' replaced by Y', and $\mathcal{E}_3$ with X' replaced by D. These are the 'empty' portions of the partition which nearly avoid X' or Y' entirely.

Consider

$$X_0 \overset{\text{def}}{=} \bigcup \{X' \cap V_1 \mid V_1 \times V_2 \in \mathcal{E}_1 \cup \mathcal{N}_1\}$$

and similarly define Y_0. Then by construction, $\mathbb{P}(X_0 \cap X') \le 2t = \frac{c}{8}\delta^2$. Observe that uniformity in D then implies that

$$\mathbb{E}_{x,y\in H} X_0(x) D(x+y) Y'(y) = \upsilon' + \mathbb{P}(X_0 \mid H)\delta_D\delta_{Y'}\,.$$

That is, we can assume

$$\mathbb{P}(X_0 \times Y' \cap X_0 \overset{\text{diag}}{\times} D \mid X' \times Y' \cap X' \overset{\text{diag}}{\times} D) \le \tfrac{c}{4}\delta^2\,.$$

The same inequality holds with X_0 replaced by X and Y by Y_0. The import of this is the inequality

$$\mathbb{P}(A \mid X_1 \times Y_1 \cap X_1 \overset{\text{diag}}{\times} D) \ge \delta + \tfrac{c}{2}\delta^2\,, \qquad X_1 \overset{\text{def}}{=} X - X_0\,,\ Y_1 \overset{\text{def}}{=} Y - Y_0\,.$$

Let $V_1 \times V_2 \in \mathcal{P} - \mathcal{E}_1 - \mathcal{E}_2 - \mathcal{N}_1 - \mathcal{N}_2$. In particular, $X \cap V_1$ is a uniform set, obeying $\|X \cap V_1\|_{\mathrm{Uni}} \le \upsilon''$. As a consequence, we have

$$\mathbb{E}_{x\in V_1, y\in V_2} X(x)V_1(x)D(x+y)Y(y)V_2(y) = \nu + \mathbb{P}(X \mid V_1)\mathbb{P}(D \mid V_1+V_2)\mathbb{P}(Y \mid V_2)\,.$$

Here, ν is a function of υ'' that need not concern us. Set

$$D_0 \overset{\text{def}}{=} \{D \cap V_1 + V_2 \mid V_1 \times V_2 \in \mathcal{E}_3 \cup \mathcal{N}_3\}.$$

It follows that

$$\mathbb{P}(X_1 \times Y_1 \cap X_1 \overset{\text{diag}}{\times} D_0 \mid X_1 \times Y_1 \cap X_1 \overset{\text{diag}}{\times} D) \le \tfrac{c}{4}\delta^2.$$

Then, for $D_1 = D - D_0$, we have

$$\mathbb{P}(A \mid X_1 \times Y_1 \cap X_1 \overset{\text{diag}}{\times} D_1) \ge \delta + \tfrac{c}{4}\delta^2.$$

From this, we see that we must have some $H' \times H'' \in \mathcal{U}$ for which

$$\mathbb{P}(A \mid H' \times H'' \cap X_1 \times Y_1 \cap X_1 \overset{\text{diag}}{\times} D_1) \geq \delta + \frac{c}{4}\delta^2 .$$

We take $X'' = X_1 \cap V_1$, $Y'' = Y_1 \cap V_2$, $D'' = D_1 \cap H' + H''$. These are all uniform subsets, satisfying (3.9); the second conclusion (3.10) is the inequality above; the lower bound on the dimension of H' and H'' follows from Lemma 7.6; and the final conclusion follows from the fact that the element of the partition that we chose, $H' \times H''$ is not in the collection $\mathcal{E}_1 \cup \mathcal{E}_2 \cup \mathcal{E}_3$.

□

References

[1] M. Ajtai and E. Szemerédi, *Sets of lattice points that form no squares*, Stud. Sci. Math. Hungar. **9** (1974), 9–11 (1975).MR0369299 (51 #5534) ↑1

[2] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984.MR1726234 (2001h:11132) ↑2

[3] Jean-Pierre Conze and Emmanuel Lesigne, *Théorèmes ergodiques pour des mesures diagonales*, Bull. Soc. Math. France **112** (1984), no. 2, 143–175 (French, with English summary). MR **788966 (86i:**28019) ↑2

[4] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168.MR833409 (87m:28007) ↑1

[5] ———, *A density version of the Hales-Jewett theorem*, J. Anal. Math. **57** (1991), 64–119.MR1191743 (94f:28020) ↑1

[6] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.MR1844079 (2002k:11014) ↑1

[7] Ben Green and Terence Tao, *The primes contain arbitrarily long arithmetic progressions*, available at `arXiv:math.NT/0404188`. ↑7

[8] Ben Green, *Long arithmetic progressions of primes*, available at `arXiv:math.NT/0508063`. ↑1

[9] Ben Green and Terence Tao, *An inverse theorem for the Gowers U^3 norm*, available at `arXiv:math.NT/0503014`. ↑2, 7

[10] Ben Green, *Finite field models in additive combinatorics*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 1–27. MR **2187732 (2006j:**11030) ↑1, 2, 7, 8, 16, 17

[11] ———, *An Argument of Shkredov in the Finite Field Setting*, available at `http://www.dpmms.cam.ac.uk/~bjg23/`. ↑1

[12] Roy Meshulam, *On subsets of finite abelian groups with no* 3*-term arithmetic progressions*, J. Combin. Theory Ser. A **71** (1995), no. 1, 168–172.MR1335785 (96g:20033) ↑2

[13] I. D. Shkredov, *On a problem of Gowers*, Dokl. Akad. Nauk **400** (2005), no. 2, 169–172 (Russian).MR2157918 ↑2

[14] ———, *On a problem of Gowers*, Izv. Ross. Akad. Nauk Ser. Mat. **70** (2006), no. 2, 179–221 (Russian, with Russian summary). MR 2223244 ↑1, 2, 3, 6

[15] I.D. Shkredov, *On a Generalization of Szemerédi's Theorem*, available at `arXiv:math.NT/0503639`. ↑1, 2

[16] J. Solymosi, *A note on a queston of Erdős and Graham*, Combin. Probab. Comput. **13** (2004), no. 2, 263–267.MR2047239 (2004m:11012) ↑1

[17] Terence Tao, *Obstructions to uniformity, and arithmetic patterns in the primes*, available at `arXiv:math.NT/0505402`. ↑1, 2

[18] V. H. Vu, *On a question of Gowers*, Ann. Comb. **6** (2002), no. 2, 229–233.MR1955522 (2003k:11013) ↑1