GEOMETRIC SQUARES AND SETS OF POSITIVE UPPER DENSITY IN $\mathbb{R}^4$

NEIL LYALL ÁKOS MAGYAR

Abstract. We establish that any subset of $\mathbb{R}^4$ of positive upper Banach density necessarily contains an isometric copy of all sufficiently large geometric squares. In addition to this we also give a new direct proof of the fact that if $A \subseteq \mathbb{R}^2$ has positive upper Banach density, then its distance set

$$\operatorname{dist}(A) = \{|x - x'| \,:\, x, x' \in A\}$$

contains all large numbers, a result originally due to Katznelson and Weiss [2] and later reproved using Fourier analytic techniques by Bourgain in [1].

1. Introduction

1.1. **Background.** Recall that the *upper Banach density* of a measurable set $A \subseteq \mathbb{R}^d$ is defined by

$$\delta^*(A) = \lim_{N\to\infty} \sup_{t\in\mathbb{R}^d} \frac{|A \cap (t + Q_N)|}{|Q_N|}, \tag{1}$$

where $|\cdot|$ denotes Lebesgue measure on $\mathbb{R}^d$ and Q_N denotes the cube $[-N/2, N/2]^d$.

A result of Katznelson and Weiss [2] states that if $A \subseteq \mathbb{R}^2$ has positive upper Banach density, then its distance set

$$\operatorname{dist}(A) = \{|x - x'| \,:\, x, x' \in A\}$$

contains all large numbers. This result was later reproved using Fourier analytic techniques by Bourgain in [1] where he established the following more general result for arbitrary non-degenerate k-dimensional simplices.

Theorem 1.1 (Bourgain [1])**.** *Let $\Delta_k \subseteq \mathbb{R}^k$ be a fixed non-degenerate k-dimensional simplex.*

If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq k+1$, then there exists a threshold $\lambda_0 = \lambda_0(A, \Delta_k)$ such that A contains an isometric copy of $\lambda \cdot \Delta_k$ for all $\lambda \geq \lambda_0$.

Recall that a set $\Delta_k = \{0, v_1, \ldots, v_k\}$ of $k+1$ points in $\mathbb{R}^k$ is a non-degenerate k-dimensional simplex if the vectors $v_1, \ldots, v_k$ are linearly independent and that a configuration Δ_k' is an isometric copy of $\lambda \cdot \Delta_k$ in $\mathbb{R}^d$ if $\Delta_k' = x + \lambda \cdot U(\Delta_k)$ for some $x \in \mathbb{R}^d$ and $U \in SO(d)$ when $d \geq k+1$.

1.2. **Main Results.** In Section 2 we present a new and direct proof of Theorem 1.1 when $k = 1$, namely a new proof of the aforementioned distance set result of Katznelson and Weiss. A new direct proof of Theorem 1.1 in its full generality can also be established along the same lines, see [3] and [4].

However, the main purpose of this to note is to present a streamlined proof the following result from [3].

Theorem 1.2 (Special case of Theorem 1.2 in [3])**.** *Let $\square = \{0, v_1, v_2, v_1 + v_2\} \subseteq \mathbb{R}^2$ with $v_1 \cdot v_2 = 0$ and $|v_1| = |v_2|$ denote the vertices of a fixed geometric square. If $A \subseteq \mathbb{R}^4$ has positive upper Banach density, then there exists a threshold $\lambda_0 = \lambda_0(A, \square)$ such that A contains an isometric copy of $\lambda \cdot \square$ for all $\lambda \geq \lambda_0$.*

We can of course write $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$ and it is important to note that the isometric copies of $\lambda \cdot \square$, whose existence in A Theorem 1.2 guarantees, will in fact all be of the special "axis-parallel" form

$$\{(x, y), (x', y), (x, y'), (x', y')\} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$$

with $|x - x'| = |y - y'| = \lambda |v_1|$.

2010 *Mathematics Subject Classification.* 11B30.

The first and second authors were partially supported by Simons Foundation Collaboration Grant for Mathematicians 245792 and by Grants NSF-DMS 1600840 and ERC-AdG 321104, respectively.

1.3. **Outline.** Our approach to proving Theorem 1.2 will be to reduce to a quantitative result in the compact setting of $[0,1]^2 \times [0,1]^2$, namely Propositions 3.1 below. This reduction is carried out in Section 3.1 with the remainder of Section 3 and the entirety of Sections 4-5 then devoted to establishing Proposition 3.1.

In Section 2 we present a new direct proof of Theorem 1.1 when $k = 1$, the novelty of our approach will be to first reduce matters to a result for suitably uniformly distributed subsets of $[0,1]^2$.

2. Uniformly Distributed Subsets of $\mathbb{R}^2$ and a New Proof of Theorem 1.1 when $k = 1$

In this section we introduce a precise notion of uniform distribution for subsets of $\mathbb{R}^2$ and prove an (optimal) result, Proposition 2.1 below, on distances in uniformly distributed subsets of $[0,1]^2$. Proposition 2.1 will be critically important in our proof of Proposition 3.1, but as we shall see below it also immediately implies Theorem 1.1 when $k = 1$ and hence provides a new direct proof of the following

Theorem 2.1 (Katznelson and Weiss [2]). *If $A \subseteq \mathbb{R}^2$ has positive upper Banach density, then there exists a threshold $\lambda_0 = \lambda_0(A)$ such that for all $\lambda \geq \lambda_0$ there exist a pair of points*

$$\{x, x'\} \subseteq A \quad \text{with} \quad |x - x'| = \lambda.$$

2.1. Uniform Distribution and Distances.

Definition 2.1 ((ε, L)-uniform distribution). Let $0 < L \leq \varepsilon \ll 1$ and $Q_L = [-L/2, L/2]^2$.

A set $A \subseteq [0,1]^2$ is said to be (ε, L)-uniformly distributed if

$$\int_{[0,1]^2} \left| \frac{|A \cap (t + Q_L)|}{|Q_L|} - |A| \right|^2 dt \leq \varepsilon^2. \tag{2}$$

Proposition 2.1 (Distances in uniformly distributed sets). *Let $c > 0$, $0 < \lambda \leq \varepsilon \ll 1$.*

If $A \subseteq [0,1]^2$ is $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed with $\alpha = |A| > 0$, then there exist a pair of points

$$\{x, x'\} \subseteq A \quad \text{with} \quad |x - x'| = \lambda.$$

In fact,

$$\iint 1_A(x) 1_A(x - \lambda x_1)\, d\sigma(x_1)\, dx = \alpha^2 + O(c^{-1/6}\varepsilon^{2/3}).$$

where σ denotes normalized arc-length measure on the circle $\{x \in \mathbb{R}^2 : |x| = 1\}$.

Before proving Proposition 2.1 we will first show that it immediately implies Theorem 2.1. To the best of our knowledge this observation, which gives a direct proof of Theorem 2.1, is new.

2.2. **Proof that Proposition 2.1 implies Theorem 2.1.** Let $\varepsilon > 0$ and $A \subseteq \mathbb{R}^2$ with $\delta^*(A) > 0$.

The following two facts follow immediately from the definition of upper Banach density, see (1):

(i) There exist $M_0 = M_0(A, \varepsilon)$ such that for all $M \geq M_0$ and all $t \in \mathbb{R}^2$

$$\frac{|A \cap (t + Q_M)|}{|Q_M|} \leq (1 + \varepsilon^4/3)\, \delta^*(A).$$

(ii) There exist arbitrarily large $N \in \mathbb{R}$ such that

$$\frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq (1 - \varepsilon^4/3)\, \delta^*(A)$$

for some $t_0 \in \mathbb{R}^2$.

Combining (i) and (ii) above we see that for any $\lambda \geq \varepsilon^{-4}M_0$, there exist $N \geq \varepsilon^{-4}\lambda$ and $t_0 \in \mathbb{R}^2$ such that

$$\frac{|A \cap (t + Q_{\varepsilon^4\lambda})|}{|Q_{\varepsilon^4\lambda}|} \leq (1 + \varepsilon^4) \frac{|A \cap (t_0 + Q_N)|}{|Q_N|}$$

for all $t \in \mathbb{R}^2$. Consequently, Theorem 2.1 reduces, via a rescaling of $A \cap (t_0 + Q_N)$ to a subset of $[0,1]^2$, to establishing that if $0 < \lambda \leq \varepsilon \ll 1$ and $A \subseteq [0,1]^2$ is measurable with $|A| > 0$ and the property that

$$\frac{|A \cap (t + Q_{\varepsilon^4\lambda})|}{|Q_{\varepsilon^4\lambda}|} \leq (1 + \varepsilon^4)\, |A|$$

for all $t \in \mathbb{R}^2$, then there exist a pair of points $x, x' \in A$ such that $|x - x'| = \lambda$. Now since $A \cap (t + Q_{\varepsilon^4\lambda})$ is only supported in $[-\varepsilon^4\lambda, 1 + \varepsilon^4\lambda]^2$ it follows that

$$|A| = \int_{\mathbb{R}^d} \frac{|A \cap (t + Q_{\varepsilon^4\lambda})|}{|Q_{\varepsilon^4\lambda}|}\, dt = \int_{[0,1]^d} \frac{|A \cap (t + Q_{\varepsilon^4\lambda})|}{|Q_{\varepsilon^4\lambda}|}\, dt + O(\varepsilon^4|A|), \tag{3}$$

from which one can easily deduce that

$$\left|\left\{t \in [0,1]^2 \,:\, \frac{|A \cap (t + Q_{\varepsilon^4\lambda})|}{|Q_{\varepsilon^4\lambda}|} \leq (1 - \varepsilon^2)\,|A|\right\}\right| = O(\varepsilon^2) \tag{4}$$

and hence that A is $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed. The result therefore follows. □

2.3. **Proof of Proposition 2.1.**

Definition 2.2 (Counting Function for Distances)**.** For $0 < \lambda \ll 1$ and functions

$$f_0, f_1 : [0,1]^2 \to \mathbb{R}$$

we define

$$T(f_0, f_1)(\lambda) = \iint f_0(x) f_1(x - \lambda x_1)\, d\sigma(x_1)\, dx. \tag{5}$$

Definition 2.3 ($U^1(L)$-norm)**.** For $0 < L \ll 1$ and functions $f : [0,1]^2 \to \mathbb{R}$ we define

$$\|f\|_{U^1(L)}^2 = \int\limits_{[0,1]^2} \left|\frac{1}{L^d} \int\limits_{t+Q_L} f(x)\, dx\right|^2 dt = \int\limits_{[0,1]^2} \left(\frac{1}{L^{2d}} \iint\limits_{x,x' \in t+Q_L} f(x) f(x')\, dx'\, dx\right) dt \tag{6}$$

where $Q_L = [-L/2, L/2]^2$.

It is an easy, but important, observation that

$$\|f\|_{U^1(L)}^2 = \iint f(x) f(x - x_1) \psi_L(x_1)\, dx_1\, dx + O(L), \tag{7}$$

where $\psi_L = L^{-4}\, 1_{Q_L} * 1_{Q_L}$. Note also that if $A \subseteq [0,1]^d$ with $\alpha = |A| > 0$ and we define

$$f_A := 1_A - \alpha 1_{[0,1]^2}$$

then

$$\int\limits_{[0,1]^2} \left|\frac{1}{L^2} \int\limits_{t+Q_L} f_A(x)\, dx\right|^2 dt = \int\limits_{[0,1]^2} \left|\frac{|A \cap (t + Q_L)|}{|Q_L|} - |A|\right|^2 dt + O(L). \tag{8}$$

Evidently the $U^1(L)$-norm is measuring the mean-square uniform distribution of A on scale L. Specifically if A is (ε, L)-uniformly distributed, then $\|f_A\|_{U^1(L)} \leq 2\varepsilon$ provided $0 < L \ll \varepsilon$.

At the heart of this short proof of Proposition 2.1 is the following "generalized von-Neumann inequality".

Lemma 2.1 (Generalized von-Neumann for Distances)**.** *For any $0 < \varepsilon, \lambda \ll 1$ and functions*

$$f_0, f_1 : [0,1]^2 \to [-1,1]$$

we have

$$|T(f_0, f_1)(c\lambda)| \leq \prod_{j=0,1} \|f_j\|_{U^1(\varepsilon^4\lambda)} + O(\varepsilon^{2/3}).$$

Indeed, if $A \subseteq [0,1]^2$ and $\alpha = |A| > 0$, then Lemma 2.1 implies that

$$\left|T(1_A, 1_A)(\lambda) - T(\alpha 1_{[0,1]^2}, \alpha 1_{[0,1]^2})(\lambda)\right| \leq 3\, \|f_A\|_{U^1(\varepsilon^4\lambda)} + O(\varepsilon^{2/3})$$

for any $0 < \varepsilon, \lambda \ll \min\{1, c^{-1}\}$. Since $T(\alpha 1_{[0,1]^d}, \alpha 1_{[0,1]^d})(\lambda) = \alpha^2 + O(\lambda)$ it follows that

$$T(1_A, 1_A)(\lambda) = \alpha^2 + O(\varepsilon^{2/3})$$

provided $0 < \lambda \leq \varepsilon \ll 1$.

To finish the proof of Proposition 2.1 we are therefore left with the task of proving Lemma 2.1.

Proof of Lemma 2.1. An application of Parseval followed by Cauchy-Schwarz implies that

$$T(f_0,f_1)(\lambda)^2 = \Big(\iint f_0(x)f_1(x-\lambda x_1)\,d\sigma(x_1)\,dx\Big)^2$$
$$\leq \Big(\int_{\mathbb{R}^2} |\widehat{f}_0(\xi)||\widehat{f}_1(\xi)||\widehat{\sigma}(\lambda\xi)|\,d\xi\Big)^2$$
$$\leq \prod_{j=0,1}\int_{\mathbb{R}^d} |\widehat{f}_j(\xi)|^2|\widehat{\sigma}(\lambda\xi)|\,d\xi$$

where

$$\widehat{\mu}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i x\cdot\xi}\,d\mu(x)$$

denotes the Fourier transform of any complex-valued Borel measure $d\mu$ and $\widehat{g}(\xi)$ is the Fourier transform of the measure $d\mu = g\,dx$. Combining the basic fact (see for example [5]) that

$$|\widehat{\sigma}(\xi)| \leq \min\{1, C|\xi|^{-1/2}\}$$

with the simple observation that $|1-\widehat{\psi}(\xi)| \leq \min\{1, C|\xi|\}$ gives

$$|\widehat{\sigma}(\lambda\xi)| = |\widehat{\sigma}(\lambda\xi)|\widehat{\psi}(\varepsilon^4\lambda\xi) + |\widehat{\sigma}(\lambda\xi)|(1-\widehat{\psi}(\varepsilon^4\lambda\xi)) \leq \widehat{\psi}(\varepsilon^4\lambda\xi) + O(\min\{\varepsilon^4\lambda|\xi|, (\lambda|\xi|)^{-1/2}\}).$$

The result now follows, since $\|f_j\|_2^2 \leq 1$,

$$\min\{\varepsilon^4\lambda|\xi|, (\lambda|\xi|)^{-1/2}\} \leq \varepsilon^{4/3}$$

and a further application of Parseval (and appeal to (7)) reveals that

$$\int |\widehat{f}_j(\xi)|^2\widehat{\psi}(\varepsilon^4\lambda\xi)\,d\xi = \iint f_j(x)f_j(x-x_1)\psi_{\varepsilon^4\lambda}(x_1)\,dx_1\,dx = \|f_j\|^2_{U^1(\varepsilon^4\lambda)} + O(\varepsilon^4\lambda). \qquad \square$$

3. Proof of Theorem 1.2

We now proceed with the main task, namely that of proving Theorem 1.2.

3.1. **Reducing Theorem 1.2 to a quantitative result for subsets of $[0,1]^2\times[0,1]^2$.** Arguing indirectly, as in Bourgain's original proof of Theorem 1.1, one can reduce Theorem 1.2 to the following quantitative result in the compact setting of $[0,1]^2\times[0,1]^2$.

Proposition 3.1. *Let $A\subseteq[0,1]^2\times[0,1]^2$ and $\alpha=|A|>0$. If $\{\lambda_j\}$ is any sequence in $(0,1)$ with $\lambda_{j+1}<\frac{1}{2}\lambda_j$ for all $j\geq 1$, then there exist $1\leq j\leq J(\alpha)$ and a quadruple of points*

$$\{(x,y),(x',y),(x,y'),(x',y')\}\subseteq A \quad \textit{with} \quad |x-x'|=|y-y'|=\lambda_j.$$

In fact, for $\lambda=\lambda_j$

$$\iiiint 1_A(x,y)1_A(x-\lambda x_1,y)1_A(x,y-\lambda y_1)1_A(x-\lambda x_1,y-\lambda y_1)\,d\sigma(x_1)\,d\sigma(y_1)\,dx\,dy \geq C(\alpha)>0.$$

The reduction of Theorem 1.2 to this result is straightforward and precisely the approach taken by Bourgain in [1] to prove Theorem 1.1, but for completeness we supply the details below.

Proof that Proposition 3.1 implies Theorem 1.2. We may assume that $|v_1|=1$.

Arguing indirectly we suppose that $A\subseteq\mathbb{R}^4$ is a set with $\delta^*(A)>0$ for which the conclusion of Theorem 1.2 fails to hold, namely that there exist arbitrarily large $\lambda\in\mathbb{R}$ for which A does not contain an isometric copy of $\lambda\cdot\square$.

We now let $0<\alpha<\delta^*(A)$ and set $J=J(\alpha)$ from Proposition 3.1. By our indirect assumption we can choose a sequence $\{\lambda_j\}_{j=1}^J$ with the property that $\lambda_{j+1}<\frac{1}{2}\lambda_j$ for all $1\leq j\leq J-1$ and A does not contain an isometric copy of $\lambda_j\cdot\square$ for each $1\leq j\leq J$. It follows from the definition of upper Banach density that exist $N\in\mathbb{R}$ with $N\gg\lambda_1$ and $t_0\in\mathbb{R}^d$ for which

$$\frac{|A\cap(t_0+Q_N)|}{|Q_N|}\geq\alpha.$$

Rescaling $A\cap(t_0+Q_N)$ to a subset of $[0,1]^4$ and applying Proposition 3.1 leads to a contradiction. $\square$

3.2. **Proof of Proposition 3.1.** We now consider the following "counting function" and uniformity norm.

Definition 3.1 (A Counting Function for "Axix-Parrallel" Squares)**.** For any $0 < \lambda \ll 1$ and functions

$$f_{ij} : [0,1]^2 \times [0,1]^2 \to \mathbb{R}$$

with $i, j \in \{0,1\}$ we define

$$T_\square(\lambda) := T_\square(f_{00}, f_{10}, f_{01}, f_{11})(\lambda)$$

where

$$T_\square(\lambda) = \iiiint f_{00}(x,y) f_{10}(x - \lambda x_1, y) f_{01}(x, y - \lambda y_1) f_{11}(x - \lambda x_1, y - \lambda y_1)\, d\sigma(x_1)\, d\sigma(y_1)\, dx\, dy. \tag{9}$$

Definition 3.2 ($\square(L)$-norm)**.** For $0 < L \ll 1$ and functions $f : [0,1]^2 \times [0,1]^2 \to \mathbb{R}$ we define

$$\|f\|^4_{\square(L)} = \int_{[0,1]^2} \int_{[0,1]^2} \|f\|^4_{\square(L)(t_1,t_2)}\, dt_2\, dt_1 \tag{10}$$

with

$$\|f\|^4_{\square(L)(t_1,t_2)} = \frac{1}{L^8} \iiiint_{\substack{x,x' \in t_1 + Q_L \\ y,y' \in t_2 + Q_L}} f(x,y) f(x',y) f(x,y') f(x',y')\ dx'\, dx\, dy'\, dy \tag{11}$$

where $Q_L = [-L/2, L/2]^2$.

As before it is a straightforward but important observation that $\|f\|^4_{\square(L)}$ equals

$$\iiiint f(x,y) f(x - x_1, y) f(x, y - y_1) f(x - x_1, y - y_1) \psi_L(x_1) \psi_L(y_1)\, dx_1\, dx\, dy_1\, dy + O(L) \tag{12}$$

where $\psi_L = L^{-4}\, 1_{Q_L} * 1_{Q_L}$.

At the heart of our proof of Proposition 3.1 is the following "relative generalized von-Neumann inequality".

Lemma 3.1 (Generalized von-Neumann for "Axis-Parallel" Squares relative to $B_1 \times B_2$)**.** *Let $B_1, B_2 \subseteq [0,1]^2$ with $\beta_i = |B_i| > 0$ for $i = 1, 2$ and set*

$$\nu(x,y) = \nu_1(x)^{1/2} \nu_2(y)^{1/2}$$

where $\nu_1 = \beta_1^{-1} 1_{B_1}$ and $\nu_2 = \beta_2^{-1} 1_{B_2}$. For any $0 < \lambda \le \varepsilon \ll 1$ and functions

$$f_{ij} : [0,1]^2 \times [0,1]^2 \to [-1,1]$$

with $i, j \in \{0,1\}$, we have

$$|T_\square(f_{00}\nu, f_{10}\nu, f_{01}\nu, f_{11}\nu)(\lambda)| \le \prod_{i,j \in \{0,1\}} \|f_{ij}\nu\|_{\square(\varepsilon^4\lambda)} + O(\beta_1^{-1}\beta_2^{-1}\varepsilon^{1/6}).$$

Notice that Proposition 2.1 ensures

$$T_\square(\nu,\nu,\nu,\nu)(\lambda) = T(\nu_1,\nu_1)(\lambda) \cdot T(\nu_2,\nu_2)(\lambda) = 1 + O(\beta_1^{-2}\beta_2^{-2}\varepsilon^{2/3}) \tag{13}$$

for any $0 < \lambda \le \varepsilon \ll 1$, provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^2$. It is then easy to see that Lemma 3.1, combined with observation (13), gives the following

Corollary 3.1. *Let $0 < \alpha, \beta_1, \beta_2 \le 1$, $0 < \lambda \le \varepsilon \ll \beta_1^{12}\beta_2^{12}\alpha^{24}$ and $A \subseteq B_1 \times B_2 \subseteq [0,1]^{d_1} \times [0,1]^{d_2}$ with $d_1, d_2 \ge 2$, $|A| = \alpha\beta_1\beta_2$ and $f_A = 1_A - \alpha 1_{B_1 \times B_2}$. If $\|f_A\nu\|_{\square(\varepsilon^4\lambda)} \ll \alpha^4$, then*

$$T_\square(1_A\nu, 1_A\nu, 1_A\nu, 1_A\nu)(\lambda) \ge \frac{1}{2}\alpha^4$$

provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^2$.

Proof of Corollary 3.1. It follows immediately from Lemma 3.1 that

$$\Big|T_\square(1_A\nu,1_A\nu,1_A\nu,1_A\nu)(\lambda)-\alpha^4 T_\square(\nu,\nu,\nu,\nu)(\lambda)\Big|\leq 15\,\|f_A\nu\|_{\square(\varepsilon^4\lambda)}+O(\beta_1^{-1}\beta_2^{-1}\varepsilon^{1/6})$$

for any $0<\varepsilon,\lambda\ll 1$, where $f_A=1_A-\alpha 1_{B_1\times B_2}$. The result follows since, as noted in (13), the fact that B_1 and B_2 are $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^2$ allows us to use Proposition 2.1 and conclude that

$$T_\square(\nu,\nu,\nu,\nu)(\lambda)=1+O(\beta_1^{-2}\beta_2^{-2}\varepsilon^{2/3})$$

for any $0<\lambda\leq\varepsilon\ll 1$, as required. □

By adaptating existing proofs of inverse theorems for non-localized box-norms one can establish an inverse theorem for the $\square(L)$-norm when $\|f_A\nu\|_{\square(\varepsilon^4\lambda)}\gg\alpha^4$, provided B_1 and B_2 are $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed in $[0,1]^{d_1}$ and $[0,1]^{d_2}$, and establish the following dichotomy

Proposition 3.2 (Dichotomy)**.** *Let $B_i\subseteq[0,1]^2$ and $\beta_i=|B_i|>0$ for $i=1,2$.*

If $A\subseteq B_1\times B_2$ with $|A|=\alpha\beta_1\beta_2>0$ and $0<\lambda\leq\varepsilon\ll\beta_1^6\beta_2^6\alpha^{32}$, then either

$$\frac{1}{\beta_1^2\beta_2^2}\iiiint 1_A(x,y)1_A(x-\lambda x_1,y)1_A(x,y-\lambda y_1)1_A(x-\lambda x_1,y-\lambda y_1)\,d\sigma(x_1)\,d\sigma(y_1)\,dx\,dy\geq\frac{1}{2}\alpha^4$$

or there exist cubes $Q_i\subseteq[0,1]^2$ of side-length $\varepsilon^4\lambda$, sets B_i' in Q_i, and $c>0$ for which

$$\frac{|A\cap(B_1'\times B_2')|}{|B_1'\times B_2'|}\geq\alpha+c\,\alpha^{32}.$$

provided B_1 and B_2 are $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^2$.

It is important to note that Propositions 3.2 is not in and of itself sufficient to establish Proposition 3.1. In order to apply a density increment argument one would need that the sets B_1' and B_2' produced by Proposition 3.2, for which A has increased density on $B_1'\times B_2'$, were (η,L')-uniformly distributed for a sufficiently small η and for L' attached to some of the λ_j's on Q_1 and Q_2 respectively, which they simply may not be. In Section 5 we complete the proof of Proposition 3.1 by showing that we can obtain suitably uniformly distributed sets B_1' and B_2' by appealing to a version of Szemerédi's Regularity Lemma [6] adapted to a sequence of scales.

4. Proof of Proposition 3.2

Throughout this section we fix $B_1,B_2\subseteq[0,1]^2$ to be arbitrary sets with $\beta_i=|B_i|>0$ for $i=1,2$.

4.1. **Proof of Lemma 3.1.** The proof of Lemma 3.1 follows from two clever applications of Cauchy-Schwarz combined with the following relative version of Lemma 2.1.

Lemma 4.1 (Relative Version of Lemma 2.1)**.** *Let $B\subseteq[0,1]^2$ and $\beta=|B|$.*

For any $0<\varepsilon,\lambda\ll 1$ and functions $f_0,f_1:[0,1]^2\to[-1,1]$ we have

$$|T(f_0\nu,f_1\nu)(\lambda)|\leq\prod_{j\in\{0,1\}}\left(\iint f_j\nu(x)f_j\nu(x-x_1)\psi_{\varepsilon^4\lambda}(x_1)\,dx_1\,dx\right)^{1/2}+O(\beta^{-1}\varepsilon^{2/3}).$$

where $\nu=\beta^{-1}1_B$.

Proof. Same as that for Lemma 2.1 above, but noting that $\|f_j\nu\|_2^2\leq\beta^{-1}$ for $j=0,1$. □

To prove Lemma 3.1 we first observe that

$$|T_\square(f_{00}\nu,f_{10}\nu,f_{01}\nu,f_{11}\nu)(\lambda)|\leq\iint|T(g_0^{x,x_1}\nu_2,g_1^{x,x_1}\nu_2)(\lambda)|\;\nu_1(x)\nu_1(x-\lambda x_1)\,d\sigma(x_1)\,dx$$

where

$$g_0^{x,x_1}(y)=f_{00}(x,y)f_{10}(x-\lambda x_1,y)$$
$$g_1^{x,x_1}(y)=f_{01}(x,y)f_{11}(x-\lambda x_1,y).$$

Applying Lemma 4.1 to $T(g_0^{x,x_1}\nu_2, g_1^{x,x_1}\nu_2)(\lambda)$ followed by an application of Cauchy-Schwarz (and switching the order of integration) shows that $|T_\square(f_{00}\nu,\dots,f_{11}\nu)(\lambda)|^2$ is majorized by

$$\prod_{j\in\{0,1\}}\iint \left|T(h_{0j}^{y,y_1}\nu_1, h_{1j}^{y,y_1}\nu_1)(\lambda)\right| \nu_2(y)\nu_2(y-\lambda y_1)\psi_{\varepsilon^4\lambda}(y_1)\,dy_1\,dy + O(\beta_1^{-2}\beta_2^{-2}\varepsilon^{2/3})$$

where

$$h_{0j}^{y,y_1}(x) = f_{0j}(x,y)f_{0j}(x,y-\lambda y_1)$$
$$h_{1j}^{y,y_1}(x) = f_{1j}(x,y)f_{1j}(x,y-\lambda y_1).$$

Applying Lemma 4.1 once more, this time to $T(h_{0j}^{y,y_1}\nu_1, h_{1j}^{y,y_1}\nu_1)(\lambda)$, followed by another application of Cauchy-Schwarz reveals that $|T_\square(f_{00}\nu,\dots,f_{11}\nu)(\lambda)|^4$ is majorized by

$$\prod_{i,j\in\{0,1\}}\iiiint h_{ij}^{y,y_1}\nu_1(x)h_{ij}^{y,y_1}\nu_1(x-x_1)\,\nu_2(y)\nu_2(y-\lambda y_1)\psi_{\varepsilon^4\lambda}(x_1)\psi_{\varepsilon^4\lambda}(y_1)\,dx_1\,dx\,dy_1\,dy + O(\beta_1^{-4}\beta_2^{-4}\varepsilon^{2/3})$$

Since

$$h_{ij}^{y,y_1}\nu_1(x)h_{ij}^{y,y_1}\nu_1(x-x_1)\nu_2(y)\nu_2(y-\lambda y_1) = f_{ij}\nu(x,y)f_{ij}\nu(x-x_1,y)f_{ij}\nu(x,y-y_1)f_{ij}\nu(x-x_1,y-y_1)$$

the result follows in light of observation (12). □

4.2. **Inverse Theorem for the $\square(L)$-norm.** The final piece in the proof of Proposition 3.2 is the following

Theorem 4.1 (Inverse Theorem). *Let $0<\eta,\beta_1,\beta_2\le 1$ and B_1 and B_2 be (ε,L)-uniformly distributed subsets of $[0,1]^2$ with $0<L\le\varepsilon\ll\eta^8\,\beta_1^2\beta_2^2$. If $f:[0,1]^2\times[0,1]^2\to[-1,1]$ satisfies*

$$\iint f(x,y)\nu_1(x)\nu_2(y)\,dx\,dy = 0 \qquad \textit{and} \qquad \|f\nu\|_{\square(L)}\ge\eta \tag{14}$$

with $\nu=\nu_1^{1/2}\otimes\nu_2^{1/2}$ and $\nu_1=\beta_1^{-1}1_{B_1}$ and $\nu_2=\beta_2^{-1}1_{B_2}$, then there exist cubes $Q_i\subseteq[0,1]^2$ of side-length L and sets $B_i'\subseteq B_i\cap Q_i$ such that

$$\frac{1}{L^4}\iint_{B_1'\times B_2'} f(x,y)\nu_1(x)\nu_2(y)\,dx\,dy \ge c\,\eta^8. \tag{15}$$

As a consequence of Theorem 4.1 we immediately obtain the following corollary which together with Corollary 3.1 implies Proposition 3.2.

Corollary 4.1. *Let $0<\alpha,\beta_1,\beta_2\le 1$ and B_1 and B_2 be $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^2$ with $0<\lambda\le\varepsilon\ll\beta_1^2\beta_2^2\alpha^{32}$. If $A\subseteq B_1\times B_2$ with $|A|=\alpha\beta_1\beta_2$ and*

$$\|f_A\nu\|_{\square(\varepsilon^4\lambda)}\gg\alpha^4$$

with $f_A=1_A-\alpha 1_{B_1\times B_2}$, then there exist cubes $Q_i\subseteq[0,1]^2$ of side-length $\varepsilon^4\lambda$ and sets B_i' in Q_i for which

$$\frac{|A\cap(B_1'\times B_2')|}{|B_1'\times B_2'|}\ge\alpha+c\,\alpha^{32}.$$

Proof of Theorem 4.1. If (15) holds for some cubes $Q_i:=t_i+Q_L$ and sets $B_i':=B_i\cap Q_i$, then Theorem 4.1 follows, so we may assume for all $t_1,t_2\in[0,1]^2$ that

$$I(t_1,t_2):=\frac{1}{\beta_1\beta_2L^4}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x,y)\,dx\,dy\le c\,\eta^8 \tag{16}$$

with say $c=2^{-16}$. It is then easy to see that this assumption, together with our assumption on the sets B_i, namely that

$$\int \big||B_i\cap(t+Q_L)|-\beta_iL^2\big|^2\,dt\le\varepsilon^2L^4,$$

imply, via an easy averaging argument, that

$$|G_{\eta,\varepsilon}|\ge\frac{\eta^4}{16} \quad \text{where} \quad G_{\eta,\varepsilon}=\left\{(t_1,t_2)\in G_\varepsilon\;:\;\|f\nu\|^4_{\square(L)(t_1,t_2)}\ge\frac{\eta^4}{16}\right\} \tag{17}$$

and

$$G_\varepsilon = \left\{(t_1,t_2);\ |B_i \cap (t_i+Q_L) - \beta_i L^2| \leq \varepsilon^{1/2}L^2 \text{ for } i=1,2\right\}.$$

We first show that if there exist $(t_1,t_2)\in G_{\eta,\varepsilon}$ for which $|I(t_1,t_2)|\leq \eta^4/2^9$, then Theorem 4.1 holds. Indeed, by the pigeonhole principle, we see that given such a pair (t_1,t_2) we may choose $x_1,y_1\in[0,1]^2$ so that

$$\left|\frac{1}{\beta_1\beta_2L^4}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)f(x_2,y_1)f(x_1,y_2)\,dx_2\,dy_2\right| \geq \frac{\eta^4}{32}. \tag{18}$$

If we now write $f_{y_1}(x_2)=f(x_2,y_1)$, $f_{x_1}(y_2)=f(x_1,y_2)$ and decompose $f_{y_1}=f_{y_1}^+-f_{y_1}^-$ and $f_{x_1}=f_{x_1}^+-f_{x_1}^-$ into their respective positive and negative parts, then it follows that

$$\left|\frac{1}{\beta_1\beta_2L^4}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)g_1(x_2)g_2(y_2)\,dx_2\,dy_2\right| \geq \frac{\eta^4}{2^7},$$

for some functions $g_i:[0,1]^2\to[0,1]$. Writing these functions as an average of indicator functions, namely

$$g_i(x)=\int_0^1 1_{\{g_i(x)\geq s\}}\,ds$$

and appealing again to the pigeonhole principle, we see that we may choose sets U_1 and V_1 so that

$$\left|\frac{1}{\beta_1\beta_2L^4}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)1_{U_1}(x_2)1_{V_1}(y_2)\,dx_2\,dy_2\right| \geq \frac{\eta^4}{2^7}. \tag{19}$$

We now set $U_2=U_1^c$, $V_2=V_1^c$ and define, for $j,j'\in\{1,2\}$, the integrals

$$I_{j,j'}:=\frac{1}{\beta_1\beta_2L^4}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)1_{U_j}(x_2)1_{V_{j'}}(y_2)\,dx_2\,dy_2.$$

Note that we know $|I_{1,1}|\geq\eta^4/2^7$ and if $I_{1,1}\geq\eta^4/2^7$ then (15) holds for the sets $B_1'=B_1\cap(t_1+Q_L)\cap U_1$ and $B_2'=B_2\cap(t_1+Q_L)\cap V_1$. We may therefore assume that $I_{1,1}\leq-\eta^4/2^7$, but this assumption, together with the previous assumption that

$$I(t_1,t_2)=I_{1,1}+I_{1,2}+I_{2,1}+I_{2,2}\geq-\eta^4/2^9$$

immediately implies that $I_{i,j}\geq\eta^4/2^9$ for some $(j,j')\neq(1,1)$ and (15) again follows.

It remains to consider the case when $I(t_1,t_2)\leq-\eta^4/2^9$ for all $(t_1,t_2)\in G_{\eta,\varepsilon}$. Then by (16) and (17)

$$\iint I(t_1,t_2)\,dt_1\,dt_2=\iint_{G_{\eta,\varepsilon}} I(t_1,t_2)\,dt_1\,dt_2+\iint_{G_{\eta,\varepsilon}^c} I(t_1,t_2)\,dt_1\,dt_2\leq-\frac{\eta^4}{2^4}\,\frac{\eta^4}{2^9}\,+\,2\,\frac{\eta^8}{2^{16}}\,\leq\,-\frac{\eta^8}{2^{15}}.$$

While on the other hand

$$\iint I(t_1,t_2)\,dt_1\,dt_2=O(L)$$

by the first assumption of (14), which is a contradiction. This proves the theorem. □

5. Proof of Proposition 3.1, Part II: Regularization

To complete the proof of Proposition 3.1, as was noted after the Proposition 3.2, we need to now produce a pair of new sets B_1'' and B_2'' that are (η,L')-uniformly distributed for a sufficiently small η and for L' attached to some of the λ_j's, but for which A still has increased density on $B_1''\times B_2''$. Proposition 3.2 did produce a pair of sets B_1' and B_2' for which A has increased density on $B_1'\times B_2'$, but these sets are not necessarily uniformly distributed. We will now obtain sets B_1'' and B_2'' with the desired properties from the sets B_1 and B_2 produced by Proposition 3.2 by appealing to a version of Szemerédi's Regularity Lemma [6] adapted to a sequence of scales $\{L_j\}_{1\leq j\leq J}$.

The precise result we need is stated below in Theorem 5.1, but first we state a couple of definitions.

Definition 5.1 (A partition $\mathcal{P}$ being adapted to scale L_j)**.** Let $1=L_0>L_1>L_2>\cdots>0$ be a sequence with the property that $L_{j+1}<\frac{1}{2}L_j$. We say that a partition $\mathcal{P}=\mathcal{Q}\cup\mathcal{R}$ of $[0,1]^2\times[0,1]^2$ into cubes $\mathcal{Q}$ and "rectangles" $\mathcal{R}$ is *adapted to the scale* L_j if each of the cubes in $\mathcal{Q}$ have sidelength L_i for some $0\leq i\leq j$.

Definition 5.2 ((ε, L)-uniform distribution on Q)**.** Let Q be a cube of sidelength L_0 and $0 < L/L_0 \leq \varepsilon \ll 1$.

A set $B \subseteq Q$ is said to be (ε, L)-uniformly distributed on Q if

$$\frac{1}{|Q|}\int_Q \left|\frac{|B\cap(t+Q_L)|}{|Q_L|} - \frac{|B|}{|Q|}\right|^2 dt \leq \varepsilon^2. \tag{20}$$

Theorem 5.1 (Regularity Lemma)**.** *Let $0 < \beta_1, \beta_2, \eta \leq 1$ and $B_i \subseteq [0,1]^2$ with $|B_i| = \beta_i$ for $i = 1, 2$.*

Given any sequence $1 = L_0 > L_1 > \cdots > 0$ with $L_{j+1} < \frac{1}{2}L_j$ there exists $0 \leq j < j' \leq J(\beta_1, \beta_2, \eta)$ and a partition $\mathcal{P} = \mathcal{Q} \cup \mathcal{R}$ of $[0,1]^2 \times [0,1]^2$ adapted to the scale L_j with the following properties:

(i) *For every cube $Q = Q_1 \times Q_2$ in $\mathcal{Q}$ of sidelength L_i with $0 \leq i \leq j-1$, the sets B_1 and B_2 are $(\eta, L_{j'})$-uniformly distributed on the cubes Q_1 and Q_2 respectively.*

(ii) *If $\mathcal{N}$ denotes the collection of cubes in $Q = Q_1 \times Q_2$ in $\mathcal{Q}$ of sidelength L_j for which at least one of the sets B_1 and B_2 is* not *$(\eta, L_{j'})$-uniformly distributed on the cubes Q_1 and Q_2 respectively, then*

$$\sum_{Q\in\mathcal{N}} |Q| + \sum_{R\in\mathcal{R}} |R| \leq \eta.$$

The proof of Theorem 5.1 follows by standard arguments, for completeness we include it in Section 5.1.

An almost immediate consequence of Theorem 5.1 is the following Corollary which, together with Proposition 3.2, provides a complete proof of Proposition 3.1, the easy verification of this we leave to the reader.

Corollary 5.1. *Let $0 < \alpha, \beta_1, \beta_2, \tau, \varepsilon \leq 1$ and $A \subseteq B_1 \times B_2 \subseteq [0,1]^2 \times [0,1]^2$ with $|A| \geq (\alpha+\tau)\beta_1\beta_2$ and $|B_i| = \beta_i$ for $i = 1, 2$. Given any sequence $1 = L_0 > L_1 > \cdots > 0$ with $L_{j+1} < \frac{1}{2}L_j$, there exist $0 \leq j < j' \leq J(\alpha, \beta_1, \beta_2, \tau, \varepsilon)$ and squares Q_1, Q_2 of sidelength L_j such that the sets*

$$B_i' := B_i \cap Q_i$$

with $i = 1, 2$ have the following properties:

(i) $|B_i'| \geq \dfrac{1}{3}\beta_i\tau|Q_i|$.

(ii) *B_i' is $(\varepsilon, L_{j'})$-uniformly distributed on Q_i*

(iii) $\dfrac{|A\cap(B_1'\times B_2')|}{|B_1'\times B_2'|} \geq \alpha + \dfrac{\tau}{3}$.

Proof that Theorem 5.1 implies Corollary 5.1. Let $\eta = \varepsilon\beta_1\beta_2\tau/3$ and $\mathcal{P} = \mathcal{Q} \cup \mathcal{R}$ be a partition of $[0,1]^2 \times [0,1]^2$ adapted to the scale L_j that satisfies the conclusions of Theorem 5.1 for some $0 \leq j < j' \leq J(\beta_1, \beta_2, \eta)$.

Let $B = B_1 \times B_2$ and $\mathcal{U}$ denote the collection of all cubes in $Q = Q_1 \times Q_2$ in $\mathcal{Q}$ of sidelength L_i with $0 \leq i \leq j$ for which B_1 and B_2 are $(\eta, L_{j'})$-uniformly distributed on Q_1 and Q_2 respectively. Note that property (ii) of Corollary 5.1 holds by definition for all cubes Q_1 and Q_2 for which $Q_1 \times Q_2 \in \mathcal{U}$.

If we let $\mathcal{S}$ denote the collection of all cubes Q in $\mathcal{U}$ which are *sparse* in the sense that $|B \cap Q| < \beta\tau|Q|/3$, then property (i) of Corollary 5.1 will hold by definition for all cubes Q_1 and Q_2 with $Q_1 \times Q_2 \in \mathcal{U} \setminus \mathcal{S}$. Finally, it is straightforward to see, using property (ii) of our partition $\mathcal{P}$ (on the size of $\mathcal{N}$ and $\mathcal{R}$) and our assumption on the relative density of A on B, that property (iii) of Corollary 5.1 must hold for at least one cube Q in $\mathcal{U} \setminus \mathcal{S}$. □

5.1. **Proof of Theorem 5.1.** By passing to a subsequence we may assume $L_{j+1} \leq 2^{-(j+6)}\eta L_j$, and in this case we will show that the conclusions of the theorem hold with $j' = j+1$ for some $0 \leq j \leq J(\beta_1, \beta_2, \gamma, \eta)$.

For $j = 0, 1, 2, \ldots$ we construct partitions $\mathcal{P}^{(j)}$ of $[0,1]^2 \times [0,1]^2$ into cubes $\mathcal{Q}^{(j)}$ and rectangles $\mathcal{R}^{(j)}$ starting from the trivial partition $\mathcal{P}^{(0)}$ consisting of only one cube $Q = [0,1]^2 \times [0,1]^2$. The partition $\mathcal{P}^{(j)}$ will consists of two collections of cubes $\mathcal{U}^{(j)}, \mathcal{N}^{(j)}$ and a collection of rectangles $\mathcal{R}^{(j)}$, that is

$$\mathcal{P}^{(j)} = \mathcal{U}^{(j)} \cup \mathcal{N}^{(j)} \cup \mathcal{R}^{(j)}.$$

The collection $\mathcal{R}^{(j)}$ will consist of rectangles $R = R_1 \times R_2$ whose total measure is small, specifically

$$\sum_{R\in\mathcal{R}^{(j)}} |R| \leq \frac{\eta}{2}, \tag{21}$$

while the collection $\mathcal{U}^{(j)}$ will consist of cubes $Q = Q_1 \times Q_2$ of sidelength L_i for some $1 \leq i \leq j$ such that B_1 and B_2 are (η, L_{i+1})-uniformly distributed on Q_1 and Q_2 respectively. Note that the cubes in $\mathcal{U}^{(j)}$ may have different sizes. The remaining collection $\mathcal{N}^{(j)}$ will consist of those cubes Q of sidelength L_j which are not (η, L_{j+1})-uniformly distributed. We will stop the procedure when the total measure of the non-uniform cubes is small enough, specifically when

$$\sum_{Q\in\mathcal{N}^{(j)}} |Q| \leq \frac{\eta}{2} \tag{22}$$

and note that such a partition satisfies the conclusions of Theorem 5.1.

If $[0,1]^2 \times [0,1]^2 \in \mathcal{U}^{(0)}$, then the sets B_1, B_2 are both (ε, L_1)-uniformly distributed and Theorem 5.1 holds. We thus assume that for some $j \geq 0$ we have a partition $\mathcal{P}^{(j)}$ for which (22) does not hold and let $Q = Q_1 \times Q_2$ denote an arbitrary cube in $\mathcal{N}^{(j)}$. By our assumption both cubes have sidelength L_j and B_i is not (η, L_{j+1})-uniformly distributed on Q_i for either $i = 1$ or $i = 2$.

We assume, without loss of generality, that $i = 1$. Averaging show that for $Q_1 = t_1 + [0, L_j]^2$ and $L := L_{j+1}$, we have

$$|E_\eta| \geq \frac{\eta^2}{2}|Q_1| \tag{23}$$

where

$$E_\eta := \left\{ t \in Q_1 \,:\, \left| \frac{|B_1 \cap (t + Q_L)|}{|Q_L|} - \frac{|B_1 \cap Q_1|}{|Q_1|} \right| \geq \frac{\eta}{2} \right\}. \tag{24}$$

Let $m = \lfloor L_j/L_{j+1} \rfloor$ and partition the cube $Q_1' = t_1 + [0, (m+1)L]^2 \supseteq Q_1$ into grids of the form $G(s_1) = s_1 + \{0, L, \ldots, mL\}^2$ with s_1 running through the cube $t_1 + [0, L]^2$. Since $L < 2^{-6}L_j$, by (23) there exist $s_1 \in Q_1'$ such that

$$\frac{|G(s_1) \cap E_\eta|}{m^2} \geq \frac{\eta^2}{4}. \tag{25}$$

Fix such an s_1 and consider the partition of Q_1 into cubes of size $L = L_{j+1}$ and possibly rectangles, defined by the grid $G(s_1)$. Repeat the same partition of the cube Q_2 corresponding to a point s_2 which we can choose arbitrarily from a cube $Q_2' \subseteq Q_2$ of size L. Taking the direct product of these partitions gives a partition of the cube $Q = Q_1 \times Q_2$ into cubes of size $L = L_{j+1}$ and possibly also into some 4-dimensional rectangles. After performing this partition of all cubes in $\mathcal{N}^{(j)}$ we obtain the new partition $\mathcal{P}^{(j+1)}$ of $[0,1]^2 \times [0,1]^2$. The new cubes obtained are then partitioned into classes $\mathcal{U}^{(j+1)}$ and $\mathcal{N}^{(j+1)}$ according to whether they are (η, L_{j+2})-uniform. Note that the cubes in $\mathcal{U}^{(j)}$ and rectangles in $\mathcal{R}^{(j)}$ remain cells of $\mathcal{P}^{(j+1)}$. Note that for each cube $Q \in \mathcal{N}^{(j)}$ the total measure of all the rectangles obtained is at most $16L_{j+1}L_j^{-1}|Q|$, hence summing over all cubes the total measure of the rectangles obtained this way is at most $4L_{j+1}L_j^{-1}$. We adjoin these rectangles to $\mathcal{R}^{(j)}$ to form $\mathcal{R}^{(j+1)}$. Note that this way the total measure of the rectangles is always bounded by

$$\sum_{j=0}^{\infty} \frac{16L_{j+1}}{L_j} \leq \sum_{j=0}^{\infty} 2^{-(j+2)}\eta \leq \frac{\eta}{2},$$

hence (21) holds.

A key notion in regularization arguments is that of the *index* or *energy* of a set with respect to a partition. In our context we define it as follows. Let $\{C_k\}_{k=1}^K$ denote the collection of cells that constitute $\mathcal{P}^{(j)}$. For any given cell $C^k = Q_1^k \times Q_2^k$ in $\mathcal{P}^{(j)}$, where Q_i^k could be either a square or a rectangle, we let δ_i^k denote the relative density of B_i in Q_i^k for $i = 1, 2$, and define the *energy* of (B_1, B_2) with respect to $\mathcal{P}^{(j)}$ by

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j)}) := \frac{1}{2} \sum_{C^k\in\mathcal{P}^{(j)}} \left((\delta_1^k)^2 + (\delta_2^k)^2\right) |C^k|. \tag{26}$$

It is not hard to see that the energy is always at most 1 and is increasing when the partition is refined. To be more precise, we say a partition $\mathcal{P}'$ is a refinement of $\mathcal{P}$ if every cell $C = Q_1 \times Q_2$ of $\mathcal{P}$ is decomposed into cells $C^{\ell,\ell'} = Q_1^\ell \times Q_2^{\ell'}$ of $\mathcal{P}'$ so that cubes (or rectangles) Q_1^ℓ and $Q_2^{\ell'}$ form a partition of Q_1 and Q_2 respectively. Then $|Q_1| = \sum_\ell |Q_1^\ell|$ and $|B_1 \cap Q_1| = \sum_\ell |B_1 \cap Q_1^\ell|$, hence writing δ_1 for the relative density of B_1 on Q_1 and δ_1^ℓ for the relative density of B_1 on Q_1^ℓ one has

$$\sum_\ell (\delta_1^\ell)^2\, |Q_1^\ell| = (\delta_1)^2\, |Q_1| + \sum_\ell (\delta_1^\ell - \delta_1)^2\, |Q_1^\ell|. \tag{27}$$

Similarly

$$\sum_{\ell'} (\delta_2^{\ell'})^2\, |Q_2^l| = (\delta_2)^2\, |Q_2| + \sum_{\ell'} (\delta_2^{\ell'} - \delta_2)^2\, |Q_2^{\ell'}|. \tag{28}$$

Multiplying equations (27) by $|Q_2|$, (28) by $|Q_1|$, and adding, we get

$$\sum_{\ell,\ell'} \big((\delta_1^\ell)^2 + (\delta_2^{\ell'})^2\big)\, |C^{\ell,\ell'}| = \big((\delta_1)^2 + (\delta_2)^2\big)\, |C| + \sum_{\ell,\ell'} \big((\delta_1^\ell - \delta_1)^2 + (\delta_2^{\ell'} - \delta_2)^2\big)\, |C^{\ell,\ell'}|. \tag{29}$$

Going back to our construction we have decomposed each cell $C^k = Q_1 \times Q_2 \in \mathcal{N}^{(j)}$ into cubes of the form $C^{\ell,\ell'} = Q_1^\ell \times Q_2^{\ell'}$ where $Q_1^\ell = s_1 + \ell L + Q_L$, $Q_2^{\ell'} = s_2 + \ell' L + Q_L$ for some $\ell \in \{1,\ldots,m\}^2$ and $\ell' \in \{1,\ldots,m\}^2$, and into a collection of 4-dimensional rectangles of small total measure. By (25) there at least $\eta^2 m^2/4$ values of ℓ for which $|\delta_1^\ell - \delta_1|^2 \geq \eta^2/4$. Thus, as $|Q_1| = L_j^2$, $|Q_1^\ell| = L_{j+1}^2$ for all ℓ, and $m = \lfloor L_j/L_{j+1} \rfloor \geq \frac{1}{2} L_j/L_{j+1}$, we have that

$$\sum_{\ell \in \{1,\ldots,m\}^2} (\delta_1^\ell - \delta_1)^2\, |Q_1^\ell| \geq \frac{\eta^4}{64}\, |Q_1|. \tag{30}$$

By (29) this implies that the energy of (B_1, B_2) with respect to the collection of cells of $\mathcal{P}^{(j+1)}$ contained in $C^k = Q_1 \times Q_2$ given by the left side of (29) is at least

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j+1)}|_{C^k}) \;\geq\; \frac{1}{2}\,(\delta_1^2 + \delta_2^2)\, |C^k| + \frac{\eta^4}{128}\, |C^k|. \tag{31}$$

This holds for all non-uniform cells $C^k \in \mathcal{N}^{(j)}$ and by our assumption that the total measure of $\mathcal{N}^{(j)} \geq \eta/2$ it follows that

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j+1)}) \;\geq\; \mathcal{E}(B_1, B_2; \mathcal{P}^{(j)}) \;+\; \frac{\eta^5}{256}. \tag{32}$$

Thus the procedure must stop in $j \leq 256\,\eta^{-5}$ steps providing a satisfactory partition. As explained above this leads to a cell $C = Q_1 \times Q_2$ satisfying the conclusions of Theorem 5.1. □

References

[1] J. Bourgain, *A Szemerédi type theorem for sets of positive density in R^k*, Israel J. Math. 54 (1986), no. 3, 307–316.

[2] H. Furstenberg, Y. Katznelson and B. Weiss, *Ergodic theory and configurations in sets of positive density*, Israel J. Math. 54 (1986), no. 3, 307–316.

[3] N. Lyall, A. Magyar, *Product of simplices and sets of positive upper density in R^d*, to appear in Math. Proc. Cambridge Philos. Soc.

[4] N. Lyall, A. Magyar, *A new proof of Bourgain's theorem on simplices in* $\mathbb{R}^d$, alpha.math.uga.edu/~ lyall/SimpleBourgain.pdf

[5] E. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality and Oscillatory Integrals*, Princeton University Press, Princeton, NJ., 1993.

[6] J. Komlós, A. Shokoufandeh, M. Simonovits M, E. Szemerédi, *The regularity lemma and its applications in graph theory,* In Theoretical aspects of computer science, pp. 84-112, Springer Berlin Heidelberg., 2002

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA
E-mail address: lyall@math.uga.edu

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA
E-mail address: magyar@math.uga.edu