PRODUCT OF SIMPLICES AND SETS OF POSITIVE UPPER DENSITY IN $\mathbb{R}^d$

NEIL LYALL ÁKOS MAGYAR

ABSTRACT. We establish that any subset of $\mathbb{R}^d$ of positive upper Banach density necessarily contains an isometric copy of all sufficiently large dilates of any fixed two-dimensional rectangle provided $d \geq 4$.

We further present an extension of this result to configurations that are the product of two non-degenerate simplices; specifically we show that if Δ_{k_1} and Δ_{k_2} are two fixed non-degenerate simplices of $k_1 + 1$ and $k_2 + 1$ points respectively, then any subset of $\mathbb{R}^d$ of positive upper Banach density with $d \geq k_1 + k_2 + 6$ will necessarily contain an isometric copy of all sufficiently large dilates of $\Delta_{k_1} \times \Delta_{k_2}$.

A new direct proof of the fact that any subset of $\mathbb{R}^d$ of positive upper Banach density necessarily contains an isometric copy of all sufficiently large dilates of any fixed non-degenerate simplex of $k + 1$ points provided $d \geq k + 1$, a result originally due to Bourgain, is also presented.

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

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

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 is also given, in fact two different new approaches are presented in Section 3. However, the main purpose of this article is to establish the following new results, namely Theorems 1.2 and 1.3 below.

Theorem 1.2. *Let $\square = \{0, v_1, v_2, v_1 + v_2\} \subseteq \mathbb{R}^2$ with $v_1 \cdot v_2 = 0$ denote a fixed two-dimensional rectangle.*

If $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq 4$, then there exists a threshold $\lambda_0 = \lambda_0(A, \square)$ such that A contains an isometric copy of $\lambda \cdot \square$ for all $\lambda \geq \lambda_0$.

Since $d \geq 4$ we can write $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with $d_1, d_2 \geq 2$. 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 form

$$\{(x, y), (x', y), (x, y'), (x', y')\} \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$$

where $|x - x'| = \lambda|v_1|$ and $|y - y'| = \lambda|v_2|$.

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.

We also establish the following generalization of Theorem 1.2, but with a slight loss in the dimension d.

Theorem 1.3. *Let Δ_{k_1} and Δ_{k_2} be two fixed non-degenerate simplices of dimension k_1 and k_2.*

If $A \subseteq \mathbb{R}^d$ has positive upper Banach density with $d \geq k_1 + k_2 + 6$, then there exists a threshold $\lambda_0 = \lambda_0(A, \Delta_{k_1}, \Delta_{k_2})$ such that A contains an isometric copy of $\lambda \cdot (\Delta_{k_1} \times \Delta_{k_2})$ of the form $\Delta'_{k_1} \times \Delta'_{k_2}$ with each $\Delta'_{k_i} \subseteq \mathbb{R}^{d_i}$ an isometric copy of $\lambda \cdot \Delta_{k_i}$ for all $\lambda \geq \lambda_0$.

It will be clear from the proofs of Theorems 1.3 and 1.2 that if $1 = k_1 < k_2$, then the conclusion of Theorem 1.3 will in fact hold under the weaker hypothesis that $d \geq k_1 + k_2 + 4$.

Note further that if A were a direct product set $B_1 \times B_2 \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ with each $d_i \geq k_i + 1$, then the conclusion of Theorem 1.3 (which contains the conclusion of Theorem 1.2 when each $k_i = 1$) would follow immediately from Theorem 1.1 and under the weaker hypothesis that $d \geq k_1 + k_2 + 2$.

The natural extension of Theorems 1.2 and 1.3 to ℓ-dimensional rectangles and ℓ-fold products of simplices (with $\ell > 2$) also holds, but as the arguments involved in establishing these results are significantly more technical than those needed for Theorems 1.2 and 1.3 we plan to address this in a separate article.

1.3. **Outline of Paper.** Our approach to proving Theorems 1.2 and 1.3 will be to reduce them to quantitative results in the compact setting of $[0,1]^{d_1} \times [0,1]^{d_2}$, namely Propositions 4.1 and 4.2. These reductions are carried out in Section 4.1 with the remainder of Section 4 and the entirety of Sections 5-7 then devoted to establishing Propositions 4.1 and 4.2.

In Section 2 we present a new direct proof of Theorem 1.1 when $k = 1$ and two new proofs of Theorem 1.1, in its full generality, are presented in Section 3. In both cases our novel approach will be to first reduce matters to results for suitably uniformly distributed subsets of $[0,1]^d$.

2. Uniformly Distributed Subsets of $\mathbb{R}^d$ 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 $\mathbb{R}^d$ and prove an (optimal) result, Proposition 2.1 below, on distances in uniformly distributed subsets of $[0,1]^d$. Proposition 2.1 will be critically important in our proof of Proposition 4.1, 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 $A \subseteq \mathbb{R}^d$ has positive upper Banach density and $d \geq 2$, then there exists a threshold $\lambda_0 = \lambda_0(A)$ such that for all $\lambda \geq \lambda_0$ there exist a pair of points*

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

2.1. **Uniform Distribution and Distances.**

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

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

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

Proposition 2.1 (Distances in uniformly distributed sets)**.** *Let $c > 0$, $0 < \lambda \leq \varepsilon \ll \min\{1, c^{-1}\}$ and $d \geq 2$.*

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

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

In fact,

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

where σ denotes the normalized measure on the sphere $\{x \in \mathbb{R}^d : |x| = 1\}$ induced by Lebesgue measure.

Before proving Proposition 2.1 we will first show that when $c = 1$ 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 $A \subseteq \mathbb{R}^d$ 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 $t \in \mathbb{R}^d$

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

(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 $t_0 \in \mathbb{R}^d$.

Combining (i) and (ii) above we see that for any $\lambda \geq \varepsilon^{-4} M_0$, there exist $N \geq \varepsilon^{-4}\lambda$ and $t_0 \in \mathbb{R}^d$ 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 $t \in \mathbb{R}^d$. Consequently, Theorem 2.1 reduces, via a rescaling of $A \cap (t_0 + Q_N)$ to a subset of $[0,1]^d$, to establishing that if $0 < \lambda \leq \varepsilon \ll 1$ and $A \subseteq [0,1]^d$ 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 $t \in \mathbb{R}^d$, 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 $[-\varepsilon^4\lambda, 1 + \varepsilon^4\lambda]^d$ 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

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

and hence that A is $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed. The result therefore follows, provided $d \geq 2$. $\square$

2.3. **Proof of Proposition 2.1.**

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

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

with $d \geq 2$ 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 $f : [0,1]^d \to \mathbb{R}$ we define

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

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

It is an easy, but important, observation that

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

where $\psi_L = L^{-2d}\, 1_{Q_L} * 1_{Q_L}$. Note also that if $A \subseteq [0,1]^d$ with $\alpha = |A| > 0$ and we define

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

then

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

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 $\|f_A\|_{U^1(L)} \le 2\varepsilon$ 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 $c > 0$, $0 < \varepsilon, \lambda \ll \min\{1, c^{-1}\}$ and functions*

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

with $d \ge 2$ we have

$$|T(f_0, f_1)(c\lambda)| \le \prod_{j=0,1} \|f_j\|_{U^1(\varepsilon^4\lambda)} + O(c^{-1/6}\varepsilon^{2/3}).$$

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

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

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 $0 < \lambda \le \varepsilon \ll \min\{1, c^{-1}\}$.

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

$$\begin{aligned} T(f_0, f_1)(c\lambda)^2 &= \Big(\iint f_0(x) f_1(x - c\lambda x_1)\,d\sigma(x_1)\,dx\Big)^2 \\ &\le \Big(\int_{\mathbb{R}^d} |\widehat{f_0}(\xi)||\widehat{f_1}(\xi)||\widehat{\sigma}(c\lambda\xi)|\,d\xi\Big)^2 \\ &\le \prod_{j=0,1}\int_{\mathbb{R}^d} |\widehat{f_j}(\xi)|^2|\widehat{\sigma}(c\lambda\xi)|\,d\xi \end{aligned}$$

where

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

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 [3]) that

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

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

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

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

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

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. A New Proof of Theorem 1.1

In light of the reduction argument presented in Section 2.2 it is clear that in order to prove Theorem 1.1 it would suffice to establish the following result for uniformly distributed subsets of $[0,1]^d$.

Proposition 3.1 (Simplices in uniformly distributed sets)**.** *Let $\Delta_k = \{0, v_1, \dots, v_k\}$ be a fixed non-degenerate k-dimensional simplex with $c_{\Delta_k} = \min_{1\le j\le k} \operatorname{dist}(v_j, \operatorname{span}\{\{v_1,\dots,v_k\}\setminus v_j\})$.*

Let $0<\lambda\le\varepsilon \ll \min\{1, c_{\Delta_k}^{-1}\}$ and $A\subseteq[0,1]^d$ with $d\ge k+1$ and $\alpha=|A|>0$. If A is $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed, then A contains an isometric copy of $\lambda\cdot\Delta_k$ and in fact

$$\iint 1_A(x)1_A(x-\lambda\cdot U(v_1))\cdots 1_A(x-\lambda\cdot U(v_k))\,d\mu(U)\,dx = \alpha^{k+1} + O_k(c_{\Delta_k}^{-1/6}\varepsilon^{2/3}) \tag{9}$$

where μ denotes the Haar measure on $SO(d)$.

Note that Proposition 2.1 is the special case of Proposition 3.1 with $k=1$ and $v_1=1$.

3.1. Proof of Proposition 3.1.

Definition 3.1 (Counting Function for Simplices)**.** For any $0<\lambda\ll 1$ and functions

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

with $d \ge k+1$ we define

$$T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda) = \iint f_0(x)f_1(x-\lambda\cdot U(v_1))\cdots f_k(x-\lambda\cdot U(v_k))\,d\mu(U)\,dx. \tag{10}$$

Proposition 3.1 is an immediate consequence of the following "generalized von-Neumann inequality".

Lemma 3.1 (Generalized von-Neumann for Simplices)**.** *For any $0<\varepsilon,\lambda\ll\min\{1,c_{\Delta_k}^{-1}\}$ and functions*

$$f_0, f_1, \dots, f_k : [0,1]^d \to [-1,1]$$

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)| \le \min_{j=0,1,\dots,k} \|f_j\|_{U^1(\varepsilon^4\lambda)} + O(c_{\Delta_k}^{-1/6}\varepsilon^{2/3}).$$

Indeed, if $A\subseteq[0,1]^d$ with $d\ge k+1$ and $\alpha=|A|>0$, then Lemma 3.1 implies

$$\left|T_{\Delta_k}(1_A,\dots,1_A)(\lambda) - T_{\Delta_k}(\alpha 1_{[0,1]^d},\dots,\alpha 1_{[0,1]^d})(\lambda)\right| \le (2^{k+1}-1)\|f_A\|_{U^1(\varepsilon^4\lambda)} + O_k(c_{\Delta_k}^{-1/6}\varepsilon^{2/3})$$

for any $0<\varepsilon,\lambda\ll\min\{1,c_{\Delta_k}^{-1}\}$. Since $T_{\Delta_k}(\alpha 1_{[0,1]^d},\dots,\alpha 1_{[0,1]^d})(\lambda) = \alpha^{k+1}+O(\lambda)$ it follows that

$$T_{\Delta_k}(1_A,\dots,1_A)(\lambda) = \alpha^{k+1} + O_k(c_{\Delta_k}^{-1/6}\varepsilon^{2/3})$$

provided $0<\lambda\le\varepsilon\ll\min\{1,c_{\Delta_k}^{-1}\}$.

To finish the proof of Proposition 3.1 we are therefore left with the task of proving Lemma 3.1.

Proof of Lemma 3.1. By symmetry it suffices to show that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)| \le \|f_k\|_{U^1(\varepsilon^4\lambda)} + O(c_{\Delta_k}^{-1/6}\varepsilon^{2/3}). \tag{11}$$

As in [1] we start by writing

$$T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda) = \iint\cdots\int f_0(x)f_1(x-\lambda x_1)\cdots f_k(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\cdots d\sigma^{(d-2)}_{x_1}(x_2)\,d\sigma(x_1)\,dx$$

where σ now denotes the normalized measure on the sphere $S^{d-1}(0,|v_1|)$ and $\sigma^{(d-j)}_{x_1,\dots,x_{j-1}}$ denotes, for each $2\le j\le k$, the normalized measure on the spheres

$$S^{d-j}_{x_1,\dots,x_{j-1}} = S^{d-1}(0,|v_j|)\cap S^{d-1}(x_1,|v_j-v_1|)\cap\cdots\cap S^{d-1}(x_{j-1},|v_j-v_{j-1}|) \tag{12}$$

where $S^{d-1}(x,r)=\{x'\in\mathbb{R}^d : |x-x'|=r\}$. Since

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)| \le \iint\cdots\int\left|\int f_k(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\right| d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma^{(d-2)}_{x_1}(x_2)\,d\sigma(x_1)\,dx$$

it follows from an application of Cauchy-Schwarz that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 \leq \int\cdots\iiint\Big|\int f_k(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\Big|^2\,dx \\ d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma^{(d-2)}_{x_1}(x_2)\,d\sigma(x_1). \tag{13}$$

An application of Plancherel therefore shows that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 \leq \int |\widehat{f_k}(\xi)|^2 I(\lambda\,\xi)\,d\xi$$

where

$$I(\xi)=\int\cdots\int\big|\widehat{\sigma^{(d-k)}_{x_1,\dots,x_{j-1}}}(\xi)\big|^2\,d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma^{(d-2)}_{x_1}(x_2)\,d\sigma(x_1). \tag{14}$$

Estimate (11) will follow if we can show that

$$I(\lambda\xi)=I(\lambda\xi)\widehat{\psi}(\varepsilon^4\lambda\xi)+I(\lambda\xi)(1-\widehat{\psi}(\varepsilon^4\lambda\xi))\leq\widehat{\psi}(\varepsilon^4\lambda\xi)+O(c_{\Delta_k}^{-1/3}\varepsilon^{4/3}) \tag{15}$$

since $\|f_k\|_2\leq 1$ and an application of Parseval and appeal to (7) reveals that

$$\int|\widehat{f_k}(\xi)|^2\widehat{\psi}(\varepsilon^4\lambda\xi)\,d\xi=\iint f_k(x)f_k(x-x_1)\psi_{\varepsilon^4\lambda}(x_1)\,dx\,dx_1=\|f_k\|^2_{U^1(\varepsilon^4\lambda)}+O(\varepsilon^4\lambda). \tag{16}$$

To establish (15) we argue as in [1], in particular we use the fact that in addition to being trivially bounded by 1 the Fourier transform of $\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}$ also decays for large ξ in certain directions, specifically

$$\big|\widehat{\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}}(\xi)\big|\leq C\left(r(S^{d-k}_{x_1,\dots,x_{k-1}})\cdot\operatorname{dist}(\xi,\operatorname{span}\{x_1,\dots,x_{k-1}\})\right)^{-(d-k)/2} \tag{17}$$

where $r(S^{d-k}_{x_1,\dots,x_{k-1}})=\operatorname{dist}(v_k,\operatorname{span}\{v_1,\dots,v_{k-1}\})$ denotes the radius of the sphere $S^{d-k}_{x_1,\dots,x_{k-1}}$.

This estimate is a consequence of the well-known asymptotic behavior of the Fourier transform of the measure on the unit sphere $S^{d-k}\subseteq\mathbb{R}^{d-k+1}$ induced by Lebesgue measure, see for example [3].

Together with the trivial uniform bound $I(\xi)\leq 1$, and an appropriate conical decomposition (depending on ξ) of the configuration space over which the integral $I(\xi)$ is defined, this gives

$$I(\xi)\leq\min\{1,C(c_{\Delta_k}|\xi|)^{-(d-k)/2}\}. \tag{18}$$

Combining (18) with the basic bound $|1-\widehat{\psi}(\xi)|\leq\min\{1,C|\xi|\}$ we obtain the uniform bound

$$|1-\widehat{\psi}(\varepsilon^4\lambda\,\xi)|I(\lambda\,\xi)\ll\min\{(\lambda c_{\Delta_k}|\xi|)^{-1/2},\varepsilon^4\lambda|\xi|\}\leq c_{\Delta_k}^{-1/3}\varepsilon^{4/3}$$

from which (15) follows. □

3.2. **A Second New Proof of Theorem 1.1.** In this subsection we present an alternative approach to proving Proposition 3.1 with the slightly worse error bound $O_k(c_{\Delta_k}^{-1/12}\varepsilon^{1/3})$. Specifically, we show that one can in fact establish the following (slightly weaker) generalized von-Neumann inequality for simplices using only Lemma 2.1, namely the generalized von-Neumann inequality for distances.

Lemma 3.2 (Generalized von-Neumann for Simplices II). *For any $0<\lambda\leq\varepsilon\ll\min\{1,c_{\Delta_k}^{-1}\}$ and functions*

$$f_0,f_1,\dots,f_k:[0,1]^d\to[-1,1]$$

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|\leq\sqrt{2\pi}\min_{j=0,1,\dots,k}\|f_j\|^{1/2}_{U^1(\varepsilon^4\lambda)}+O(c_{\Delta_k}^{-1/12}\varepsilon^{1/3}).$$

In the proof below we will make use of the following straightforward observations:

(i) If we let $\Delta_{k-1}=\{0,v_1,\dots,v_{k-1}\}$, then

$$T_{\Delta_k}(f_0,f_1,\dots,f_{k-1},1_{[0,1]^d})(\lambda)=T_{\Delta_{k-1}}(f_0,f_1,\dots,f_{k-1})(\lambda)+O(\lambda). \tag{19}$$

(ii) If we let $\Delta_k'=\{0,v_1',\dots,v_k'\}$ with $v_j'=v_{k-j}-v_k$ for $0\leq j\leq k-1$ and $v_k'=-v_k$, then

$$T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)=T_{\Delta_k'}(f_k,f_{k-1},\dots,f_0)(\lambda). \tag{20}$$

Proof of Lemma 3.2. By symmetry it suffices to show that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 \leq 2\pi\,\|f_k\|_{U^1(\varepsilon^4\lambda)} + O(c_{\Delta_k}^{-1/6}\varepsilon^{2/3}). \tag{21}$$

We initially follow the proof of Lemma 3.1, but after (13) we now proceed differently. Instead of applying Plancherel to the right hand side of

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 \leq \iint\cdots\int\left|\int f_k(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\right|^2 d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx.$$

we now "square out" the right hand side to obtain

$$\iint\cdots\iiint f_k(x-\lambda x_k)f_k(x-\lambda x_{k+1})\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_{k+1})\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\,d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx. \tag{22}$$

If $d=k+1$, then for fixed $x_1,\dots,x_k$ we can use arc-length to parameterize of the circle $S^{d-k}_{x_1,\dots,x_{k-1}}$, with $\theta=0$ and $\theta=2\pi$ corresponding to the point x_k, to write

$$\int f_k(x-\lambda x_{k+1})\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_{k+1}) = \int_0^{2\pi} f_k(x-\lambda x_{k+1}(x_1,\dots,x_k,\theta))\,d\theta. \tag{23}$$

For any fixed $\theta\in[0,2\pi]$ we then define $\Delta_{k+1}(\theta)=\{0,v_1,\dots,v_k,v_{k+1}(\theta)\}$ with $v_{k+1}=v_{k+1}(\theta)$ satisfying $|v_{k+1}|=|v_k|$, $|v_{k+1}-v_j|=|v_k-v_j|$ for all $1\leq j\leq k-1$ and use θ to determine the angle between v_{k+1} and v_k measured from the center of the circle $S^{d-k}_{x_1,\dots,x_{k-1}}$, consequently

$$|v_{k+1}-v_k| = 2\sin(\theta/2)\cdot\operatorname{dist}(v_k,\operatorname{span}\{v_1,\dots,v_{k-1}\}).$$

It follows that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 \leq \int_0^{2\pi} T_{\Delta_{k+1}(\theta)}(1_{[0,1]^d},\dots,1_{[0,1]^d},f_k,f_k)(\lambda)\,d\theta + O(\lambda)$$

and in light of (19) and (20) that

$$\begin{aligned}|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2 &\leq \int_0^{2\pi} T_{\Delta'_{k+1}(\theta)}(f_k,f_k,1_{[0,1]^d},\dots,1_{[0,1]^d})(\lambda)\,d\theta + O(\lambda)\\ &= \int_0^{2\pi} T_{\Delta'_1(\theta)}(f_k,f_k)(\lambda)\,d\theta + O(\lambda)\end{aligned}$$

where

$$T_{\Delta'_1(\theta)}(f_k,f_k)(\lambda) = T(f_k,f_k)(c(\theta)\lambda) := \iint f_k(x)f_k(x-c(\theta)\lambda x_1)\,d\sigma(x_1)\,dx$$

with $c(\theta)=2\sin(\theta/2)\cdot\operatorname{dist}(v_k,\operatorname{span}\{v_1,\dots,v_{k-1}\})$. Lemma 2.1 now implies that

$$|T_{\Delta'_1(\theta)}(f_k,f_k)(\lambda)| \leq \|f_k\|_{U^1(\varepsilon^4\lambda)} + O((\sin(\theta/2))^{-1/6}c_{\Delta_k}^{-1/6}\varepsilon^{2/3})$$

since $c(\theta)\geq 2\sin(\theta/2)\,c_{\Delta_k}$. This completes the proof, when $d=k+1$, as $\int_0^{2\pi}(\sin(\theta/2))^{-1/6}\,d\theta<\infty$, and in fact establishes the result in general, since if $d\geq k+2$, one can define a new non-degenerate simplex

$$\Delta_{d-1} = \{0,v_1,\dots,v_{k-1},v'_k,\dots,v'_{d-2},v'_{d-1}\}$$

with $v'_{d-1}=v_k$ and use the fact that

$$T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda) = T_{\Delta_{d-1}}(f_0,\dots,f_{k-1},1_{[0,1]^d},\dots,1_{[0,1]^d},f_k)(\lambda)+O(\lambda). \qquad\square$$

3.3. **A Direct proof of Lemma 3.2 when $d\geq k+2$.**

We choose to include an additional argument similar to the one presented above that covers the case $d\geq k+2$ directly. Arguments of this nature will be critical important in Section 6.2 when we establish a "relative generalized von-Neumann inequality" for simplices.

If $d\geq k+2$ then in (22), for fixed $x_1,\dots,x_k$, we write

$$\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_{k+1}) = \int_0^{\pi}(\sin\theta)^{d-k-1}\,d\sigma^{(d-k-1)}_{x_1,\dots,x_{k-1},x_k,\theta}(x_{k+1})\,d\theta \tag{24}$$

where $\sigma^{(d-k-1)}_{x_1,\dots,x_{k-1},x_k,\theta}(x_{k+1})$ denotes the normalized measure on the sphere

$$S^{d-k-1}_{x_1,\dots,x_{k-1},x_k,\theta} = S^{d-1}(0,|v_{k+1}|)\cap S^{d-1}(x_1,|v_{k+1}-v_1|)\cap\cdots\cap S^{d-1}(x_k,|v_{k+1}-v_k|) \tag{25}$$

with $v_{k+1}=v_{k+1}(\theta)$ defined such that $|v_{k+1}|=|v_k|$, $|v_{k+1}-v_j|=|v_k-v_j|$ for all $1\leq j\leq k-1$ with θ determining the angle between v_{k+1} and v_k measured from the center of the sphere $S^{d-k}_{x_1,\dots,x_{k-1}}$, consequently

$$|v_{k+1}-v_k|=2\sin(\theta/2)\cdot\operatorname{dist}(v_k,\operatorname{span}\{v_1,\dots,v_{k-1}\}).$$

If we again let $\Delta_{k+1}(\theta)=\{0,v_1,\dots,v_k,v_{k+1}\}$, it follows that

$$|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2\leq\int_0^\pi(\sin\theta)^{d-k-1}T_{\Delta_{k+1}(\theta)}(1_{[0,1]^d},\dots,1_{[0,1]^d},f_k,f_k)(\lambda)\,d\theta+O(\lambda)$$

and in light of (19) and (20) that

$$\begin{aligned}|T_{\Delta_k}(f_0,f_1,\dots,f_k)(\lambda)|^2&\leq\int_0^\pi(\sin\theta)^{d-k-1}T_{\Delta'_{k+1}(\theta)}(f_k,f_k,1_{[0,1]^d},\dots,1_{[0,1]^d})(\lambda)\,d\theta+O(\lambda)\\&=\int_0^\pi(\sin\theta)^{d-k-1}T_{\Delta'_1(\theta)}(f_k,f_k)(\lambda)\,d\theta+O(\lambda)\end{aligned}$$

where again

$$T_{\Delta_1(\theta)}(f_k,f_k)(\lambda)=T(f_k,f_k)(c(\theta)\lambda):=\iint f_k(x)f_k(x-c(\theta)\lambda x_1)\,d\sigma(x_1)\,dx$$

with $c(\theta)=2\sin(\theta/2)\cdot\operatorname{dist}(v_k,\operatorname{span}\{v_1,\dots,v_{k-1}\})$. Lemma 2.1 again implies that

$$|T_{\Delta'_1(\theta)}(f_k,f_k)(\lambda)|\leq\|f_k\|_{U^1(\varepsilon^4\lambda)}+O((\sin(\theta/2))^{-1/6}c_{\Delta_k}^{-1/6}\varepsilon^{2/3})$$

since $c(\theta)\geq2\sin(\theta/2)\,c_{\Delta_k}$ and this completes the proof as $\int_0^\pi(\sin\theta)^{d-k-1}(\sin(\theta/2))^{-1/6}\,d\theta<\infty$. □

4. Proof of Theorems 1.2 and 1.3

We now proceed with the main task, namely the proofs of Theorems 1.2 and 1.3.

4.1. Reducing Theorems 1.2 and 1.3 to quantitative results for subsets of $[0,1]^{d_1}\times[0,1]^{d_2}$.

Proposition 4.1 (Rectangles). *Let $c>0$ and $A\subseteq[0,1]^{d_1}\times[0,1]^{d_2}$ with $d_1,d_2\geq2$ and $\alpha=|A|>0$.*

If $\{\lambda_j\}$ is any sequence in $(0,1)$ with $\lambda_{j+1}<\frac12\lambda_j$ for all $j\geq1$, then there exist $1\leq j\leq J(\alpha,c)$ and a quadruple of points

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

In fact, for $\lambda=\lambda_j$

$$\iiiint 1_A(x,y)1_A(x-\lambda x_1,y)1_A(x,y-c\lambda y_1)1_A(x-\lambda x_1,y-c\lambda y_1)\,d\sigma_1(x_1)\,d\sigma_2(y_1)\,dx\,dy\geq C(\alpha)>0$$

where σ_i denotes, for $i=1,2$, the normalized measure on the unit sphere $S^{d_i-1}\subseteq\mathbb{R}^{d_i}$ centered at the origin induced by the Lebesgue measure on $\mathbb{R}^{d_i}$.

Proposition 4.2 (Product of Simplices). *For $i=1,2$ let $\Delta_{k_i}=\{0,v_1^i,v_2^i,\dots,v_{k_i}^i\}$ be a fixed non-degenerate simplex of dimension k_i. Let $A\subseteq[0,1]^{d_1}\times[0,1]^{d_2}$ with $d_i\geq k_i+3$ and $\alpha=|A|>0$.*

If $\{\lambda_j\}$ is any sequence in $(0,1)$ with $\lambda_{j+1}<\frac12\lambda_j$ for all $j\geq1$, then there exist $1\leq j\leq J(\alpha,\Delta_{k_1},\Delta_{k_2})$ and a product $\Delta'_{k_1}\times\Delta'_{k_2}\subseteq A$ with each $\Delta'_{k_i}\subseteq[0,1]^{d_i}$ an isometric copy of $\lambda_j\cdot\Delta_{k_i}$. In fact, for $\lambda=\lambda_j$

$$\iiiint\prod_{i=0}^{k_1}\prod_{j=0}^{k_2}1_A(x-\lambda\cdot U_1(v_i^1),y-\lambda\cdot U_2(v_j^2))\,d\mu_1(U_1)\,d\mu_2(U_2)\,dx\,dy\geq C(\alpha)>0$$

where $v_0^1=v_0^2=0$ and μ_1 and μ_2 denote the Haar measures on $SO(d_1)$ and $SO(d_2)$ respectively.

The reduction of Theorems 1.2 and 1.3 to these results in the compact setting of $[0,1]^{d_1}\times[0,1]^{d_2}$ is straightforward and precisely the approach taken by Bourgain in [1] to prove Theorem 1.1, but for completeness we supply the details for Theorem 1.2 below.

Proof that Proposition 4.1 implies Theorem 1.2. We may assume that $c := |v_2| \geq |v_1| = 1$.

Arguing indirectly we suppose that $A \subseteq \mathbb{R}^d$ with $d \geq 4$ 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 $J = J(\alpha, c)$ from Proposition 4.1. 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 $[0,1]^d$ and applying Proposition 4.1 leads to a contradiction. $\square$

4.2. **Proof of Propositions 4.1 and 4.2, Part I: A Density Increment Strategy.**

Proposition 4.3 (Dichotomy for Rectangles). *Let $c > 0$ and $B_i \subseteq [0,1]^{d_i}$ with $d_i \geq 2$ 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 $0 < \lambda \leq \varepsilon \ll \beta_1^6\beta_2^6\alpha^{32}\min\{c^{1/4}, c^{-1}\}$, then either*

$$\frac{1}{\beta_1^2\beta_2^2}\iiiint 1_A(x,y)1_A(x-\lambda x_1, y)1_A(x, y-c\lambda y_1)1_A(x-\lambda x_1, y-c\lambda y_1)\,d\sigma_1(x_1)\,d\sigma_2(y_1)\,dx\,dy \geq \frac{1}{2}\alpha^4$$

or there exist cubes $Q_i \subseteq [0,1]^{d_i}$ 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 $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

Proposition 4.4 (Dichotomy for Product of Simplices). *For $i = 1, 2$ let $B_i \subseteq [0,1]^{d_i}$ with $d_i \geq k_i + 3$ and $\beta_i = |B_i| > 0$ and $\Delta_{k_i} = \{v_0^i, v_1^i, v_2^i, \dots, v_{k_i}^i\}$ be a non-degenerate simplex of dimension k_i with $v_0^i = 0$ and*

$$c_{\Delta_{k_i}} = \min_{1 \leq j \leq k_i} \operatorname{dist}(v_j^i, \operatorname{span}\{\{v_1^i, \dots, v_{k_i}^i\} \setminus v_j^i\}).$$

If $A \subseteq B_1 \times B_2$ with $|A| = \alpha\beta_1\beta_2 > 0$ and

$$0 < \lambda \leq \varepsilon \ll_{k_1,k_2} (\beta_1^{k_1+1}\beta_2^{k_2+1}\alpha^{(k_1+1)(k_2+1)})^{16}\min\{(c_{\Delta_{k_1}}c_{\Delta_{k_2}})^2, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}$$

then either

$$\frac{1}{\beta_1^{k_1+1}\beta_2^{k_2+1}}\iiiint \prod_{i=0}^{k_1}\prod_{j=0}^{k_2} 1_A(x - \lambda \cdot U_1(v_i^1), y - \lambda \cdot U_2(v_j^2))\,d\mu_1(U_1)\,d\mu_2(U_2)\,dx\,dy \geq \frac{1}{2}\alpha^{(k_1+1)(k_2+1)}$$

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

$$\frac{|A \cap (B_1' \times B_2')|}{|B_1' \times B_2'|} \geq \alpha + c'\,\alpha^{8(k_1+1)(k_2+1)}.$$

provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

Sections 5 and 6 below are devoted to the proofs of Propositions 4.3 and 4.4. Central to each proof is an appropriate "relative generalized von-Neumann inequality", namely Lemmas 5.1 and 6.1. These relative generalized von-Neumann inequalities in turn imply Corollaries 5.1 and 6.1, which together with Corollaries 5.2 and 6.2 (which are both consequences of an appropriate common "Inverse Theorem", namely Theorem 5.1) immediately imply Propositions 4.3 and 4.4 respectively.

It is important to note that Propositions 4.3 and 4.4 are not in and of themselves sufficient to establish Propositions 4.1 and 4.2. In order to apply a density increment argument one would need that the sets B_1' and B_2' produced by Propositions 4.3 and 4.4, 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 7 we complete the proofs of Proposition 4.1 and 4.2 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 [4] adapted to a sequence of scales.

5. Proof of Proposition 4.3

At the heart of our proof of Proposition 4.3 will be an appropriate "relative generalized von-Neumann inequality for rectangles", namely Lemma 5.1 below. This result, together with a companion "Inverse Theorem" (Theorem 5.1 below) and Proposition 2.1 will ultimately furnish a proof of Proposition 4.3.

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

5.1. A Relative Generalized von-Neumann Inequality for Distances and Rectangles.

Definition 5.1 (A Counting Function for Rectangles)**.** For any $c > 0$, $0 < \lambda \ll 1$ and functions

$$f_{ij} : [0,1]^{d_1} \times [0,1]^{d_2} \to \mathbb{R}$$

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

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

where

$$T_{\square_c}(\lambda) = \iiiint f_{00}(x,y) f_{10}(x - \lambda x_1, y) f_{01}(x, y - c\lambda y_1) f_{11}(x - \lambda x_1, y - c\lambda y_1)\, d\sigma_1(x_1)\, d\sigma_2(y_1)\, dx\, dy \tag{26}$$

Note that if we let

$$\nu(x,y) = \nu_1(x)^{1/2} \nu_2(y)^{1/2} \tag{27}$$

where

$$\nu_1 = \beta_1^{-1} 1_{B_1} \quad \text{and} \quad \nu_2 = \beta_2^{-1} 1_{B_2} \tag{28}$$

then, in light of Proposition 2.1, we have

$$T_{\square_c}(\nu,\nu,\nu,\nu)(\lambda) = T(\nu_1,\nu_1)(\lambda) \cdot T(\nu_2,\nu_2)(c\lambda) = 1 + O(\beta_1^{-2}\beta_2^{-2} c^{-1/6} \varepsilon^{2/3}) \tag{29}$$

for any $0 < \lambda \leq \varepsilon \ll \min\{1, c^{-1}\}$, provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

Definition 5.2 ($\square(L)$-norm)**.** For $0 < L \ll 1$ and functions $f : [0,1]^{d_1} \times [0,1]^{d_2} \to \mathbb{R}$ we define

$$\|f\|^4_{\square(L)} = \int_{[0,1]^{d_1}} \int_{[0,1]^{d_2}} \|f\|^4_{\square(L)(t_1,t_2)}\, dt_2\, dt_1 \tag{30}$$

with

$$\|f\|^4_{\square(L)(t_1,t_2)} = \frac{1}{L^{2(d_1+d_2)}} \iiiint_{\substack{x,x' \in t_1 + Q_{1,L} \\ y,y' \in t_2 + Q_{2,L}}} f(x,y) f(x',y) f(x,y') f(x',y')\; dx'\, dx\, dy'\, dy \tag{31}$$

where $Q_{i,L} = [-L/2, L/2]^{d_i}$ for $i = 1, 2$.

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

$$\iiiint f(x,y) f(x - x_1, y) f(x, y - y_1) f(x - x_1, y - y_1) \psi_{1,L}(x_1) \psi_{2,L}(y_1)\, dx_1\, dx\, dy_1\, dy + O(L) \tag{32}$$

where $\psi_{i,L} = L^{-2d_i}\, 1_{Q_{i,L}} * 1_{Q_{i,L}}$.

In this setting we have the following "generalized von-Neumann inequality" relative to $B_1 \times B_2$.

Lemma 5.1 (Generalized von-Neumann for Rectangles relative to $B_1 \times B_2$)**.** *Let $\nu = \nu_1^{1/2} \otimes \nu_2^{1/2}$ where $\nu_1 = \beta_1^{-1} 1_{B_1}$ and $\nu_2 = \beta_2^{-1} 1_{B_2}$. For any $c > 0$, $0 < \varepsilon, \lambda \ll \min\{1, c^{-1}\}$ and functions*

$$f_{ij} : [0,1]^{d_1} \times [0,1]^{d_2} \to [-1,1]$$

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

$$|T_{\square_c}(f_{00}\nu, f_{10}\nu, f_{01}\nu, f_{11}\nu)(\lambda)| \leq \prod_{i,j \in \{0,1\}} \|f_{ij}\nu\|_{\square(\varepsilon^4\lambda)} + O(\beta_1^{-1}\beta_2^{-1} c^{-1/24} \varepsilon^{1/6}).$$

It is easy to see that Lemma 5.1, combined with Proposition 2.1, gives the following

Corollary 5.1. *Let $c>0$, $0<\alpha,\beta_1,\beta_2\leq 1$ and $0<\lambda\leq\varepsilon\ll\beta_1^6\beta_2^6\alpha^{24}\min\{c^{1/4},c^{-1}\}$.*

If $A\subseteq B_1\times B_2\subseteq[0,1]^{d_1}\times[0,1]^{d_2}$ with $|A|=\alpha\beta_1\beta_2$ and $\|f_A\nu\|_{\square(\varepsilon^4\lambda)}\ll\alpha^4$, then

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

provided B_1 and B_2 are $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

Proof of Corollary 5.1. It follows immediately from Lemma 5.1 that

$$\left|T_{\square_c}(1_A\nu,1_A\nu,1_A\nu,1_A\nu)(\lambda)-\alpha^4T_{\square_c}(\nu,\nu,\nu,\nu)(\lambda)\right|\leq 15\,\|f_A\nu\|_{\square(\varepsilon^4\lambda)}+O(\beta_1^{-1}\beta_2^{-1}c^{-1/24}\varepsilon^{1/6})$$

for any $0<\varepsilon,\lambda\ll\min\{1,c^{-1}\}$, where $f_A=1_A-\alpha 1_{B_1\times B_2}$. The result follows since, as noted in (42), the fact that B_1 and B_2 are $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ allows us to use Proposition 2.1 and conclude that

$$T_{\square_c}(\nu,\nu,\nu,\nu)(\lambda)=1+O(\beta_1^{-2}\beta_2^{-2}c^{-1/6}\varepsilon^{2/3})$$

for any $0<\lambda\leq\varepsilon\ll\min\{1,c^{-1}\}$, as required. □

5.2. **Proof of Lemma 5.1.** The proof of Lemma 5.1 follows from two clever applications of Cauchy-Schwarz combined with the following relative version of Lemma 2.1.

Lemma 5.2 (Relative Version of Lemma 2.1)**.** *Let $B\subseteq[0,1]^d$ with $d\geq 2$ and $\beta=|B|$.*

For any $c>0$, $0<\varepsilon,\lambda\ll\min\{1,c^{-1}\}$ and functions $f_0,f_1:[0,1]^d\to[-1,1]$ we have

$$|T(f_0\nu,f_1\nu)(c\lambda)|\leq\prod_{j\in\{0,1\}}\left(\iint f_j\nu(x)f_j\nu(x-x_1)\psi_{\varepsilon^4\lambda}(x_1)\,dx_1\,dx\right)^{1/2}+O(\beta^{-1}c^{-1/6}\varepsilon^{2/3}).$$

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 5.1 we first observe that

$$|T_{\square_c}(f_{00}\nu,f_{10}\nu,f_{01}\nu,f_{11}\nu)(\lambda)|\leq\iint\left|T(g_0^{x,x_1}\nu_2,g_1^{x,x_1}\nu_2)(c\lambda)\right|\,\nu_1(x)\nu_1(x-\lambda x_1)\,d\sigma_1(x_1)\,dx$$

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

$$\prod_{j\in\{0,1\}}\iint\left|T(h_{0j}^{y,y_1}\nu_1,h_{1j}^{y,y_1}\nu_1)(\lambda)\right|\nu_2(y)\nu_2(y-\lambda y_1)\psi_{2,\varepsilon^4\lambda}(y_1)\,dy_1\,dy+O(\beta_1^{-2}\beta_2^{-2}c^{-1/6}\varepsilon^{2/3})$$

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 5.2 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 $|T_{\square_c}(f_{00}\nu,\dots,f_{11}\nu)(\lambda)|^4$ is majorized by

$$\prod_{i,j\in\{0,1\}}\iiiint h_{ij}^{y,y_1}\nu_1(x)h_{ij}^{y,y_1}\nu_1(x-x_1)\,\nu_2(y)\nu_2(y-\lambda y_1)\psi_{1,\varepsilon^4\lambda}(x_1)\psi_{2,\varepsilon^4\lambda}(y_1)\,dx_1\,dx\,dy_1\,dy+O(\beta_1^{-4}\beta_2^{-4}c^{-1/6}\varepsilon^{2/3})$$

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 (32). □

5.3. **Inverse Theorem for the $\square(L)$-norm.** The final piece in the proof of Proposition 4.3 is the following

Theorem 5.1 (Inverse Theorem)**.** *Let $0<\eta,\beta_1,\beta_2\leq 1$ and B_1 and B_2 be (ε,L)-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ with $0<L\leq\varepsilon\ll\eta^8\,\beta_1^2\beta_2^2$. If $f:[0,1]^{d_1}\times[0,1]^{d_2}\to[-1,1]$ satisfies*

$$\iint f(x,y)\nu_1(x)\nu_2(y)\,dx\,dy=0 \qquad \text{and} \qquad \|f\nu\|_{\square(L)}\geq\eta \tag{33}$$

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 $Q_i\subseteq[0,1]^{d_i}$ of side-length L and sets $B_i'\subseteq B_i\cap Q_i$ such that

$$\frac{1}{L^{d_1+d_2}}\iint_{B_1'\times B_2'} f(x,y)\nu_1(x)\nu_2(y)\,dx\,dy\geq c\,\eta^8. \tag{34}$$

As a consequence of Theorem 5.1 we immediately obtain the following corollary which together with Corollary 5.1 implies Proposition 4.3.

Corollary 5.2. *Let $0<\alpha,\beta_1,\beta_2\leq 1$ and B_1 and B_2 be $(\varepsilon,\varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ with $0<\lambda\leq\varepsilon\ll\beta_1^2\beta_2^2\alpha^{32}$.*

If $A\subseteq B_1\times B_2\subseteq[0,1]^{d_1}\times[0,1]^{d_2}$ 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 $Q_i\subseteq[0,1]^{d_i}$ of side-length $\varepsilon^4\lambda$ and sets B_i' in Q_i for which

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

Proof of Theorem 5.1. If (34) holds for some cubes $Q_i:=t_i+Q_L$ and sets $B_i':=B_i\cap Q_i$, then Theorem 5.1 follows, so we may assume for all $t_1\in[0,1]^{d_1}$ and $t_2\in[0,1]^{d_2}$ that

$$I(t_1,t_2):=\frac{1}{\beta_1\beta_2L^{d_1+d_2}}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x,y)\,dx\,dy\leq c\,\eta^8 \tag{35}$$

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

$$\int \big||B_i\cap(t+Q_L)|-\beta_iL^{d_i}\big|^2\,dt\leq\varepsilon^2L^{2d_i},$$

imply, via an easy averaging argument, that

$$|G_{\eta,\varepsilon}|\geq\frac{\eta^4}{16} \quad \text{where} \quad G_{\eta,\varepsilon}=\left\{(t_1,t_2)\in G_\varepsilon\ :\ \|f\|^4_{\square^2_{B_1,B_2}(L,t_1,t_2)}\geq\frac{\eta^4}{16}\right\} \tag{36}$$

and

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

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 5.1 holds. Indeed, by the pigeonhole principle, we see that given such a pair (t_1,t_2) we may choose $x_1\in[0,1]^{d_1}$ and $y_1\in[0,1]^{d_2}$ so that

$$\left|\frac{1}{\beta_1\beta_2L^{d_1+d_2}}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)f(x_2,y_1)f(x_1,y_2)\,dx_2\,dy_2\right|\geq\frac{\eta^4}{32}. \tag{37}$$

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

$$\left|\frac{1}{\beta_1\beta_2L^{d_1+d_2}}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)g_1(x_2)g_2(y_2)\,dx_2\,dy_2\right|\geq\frac{\eta^4}{2^7},$$

for some functions $g_i:[0,1]^{d_i}\to[0,1]$. 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

$$\left|\frac{1}{\beta_1\beta_2L^{d_1+d_2}}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)1_{U_1}(x_2)1_{V_1}(y_2)\,dx_2\,dy_2\right| \geq \frac{\eta^4}{2^7}. \tag{38}$$

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

$$I_{j,j'} := \frac{1}{\beta_1\beta_2L^{d_1+d_2}}\int_{t_1+Q_L}\int_{t_2+Q_L} f(x_2,y_2)1_{U_j}(x_2)1_{V_{j'}}(y_2)\,dx_2\,dy_2.$$

Note that we know $|I_{1,1}| \geq \eta^4/2^7$ and if $I_{1,1} \geq \eta^4/2^7$ then (34) 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 (34) 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 (35) and (36)

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

While on the other hand

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

by the first assumption of (33), which is a contradiction. This proves the theorem. □

6. Proof of Proposition 4.4

An appropriate "relative generalized von-Neumann inequality" will again be central to our proof of Proposition 4.4, specifically a "relative generalized von-Neumann inequality for product of simplices".

However, the true heart of the argument is in fact the analogous result for *just* simplices, the proof of this "relative generalized von-Neumann inequality for simplices" is necessarily significantly more involved than the analogous relative result for distances (whose proof was essentially identical to the non-relative case) and it is here that our loss in dimension appears.

We fix non-degenerate simplices $\Delta_{k_i} = \{v_0^i, v_1^i, v_2^i, \ldots, v_{k_i}^i\}$ of dimension k_i with $v_0^i = 0$ and let $B_i \subseteq [0,1]^{d_i}$ with $d_i \geq k_i + 3$ and $\beta_i = |B_i| > 0$ denote arbitrary sets, for $i = 1,2$.

In contrast to the proof of Proposition 4.3, we will need to assume that our sets B_1 and B_2 are suitably uniformly distributed, and make use of Proposition 3.1, throughout the proof of Proposition 4.4.

6.1. A Relative Generalized von-Neumann Inequality for Simplices and Products of Simplices.

Definition 6.1 (Counting function for $\Delta_{k_1} \times \Delta_{k_2}$)**.** Let $0 < \lambda \ll 1$.

For functions $f_{ij} : [0,1]^{d_1} \times [0,1]^{d_2} \to \mathbb{R}$ with $(i,j) \in \{0,1,\ldots,k_1\} \times \{0,1,\ldots,k_2\}$ we define

$$T_{\Delta_{k_1},\Delta_{k_2}}(f_{00},\ldots,f_{k_1k_2})(\lambda) = \iiiint \prod_{i=0}^{k_1}\prod_{j=0}^{k_2} f_{ij}(x-\lambda\cdot U_1(v_i^1), y-\lambda\cdot U_2(v_j^2))\,d\mu_1(U_1)\,d\mu_2(U_2)\,dx\,dy \tag{39}$$

Note that if we let

$$\widetilde{\nu}(x,y) = \nu_1(x)^{1/(k_2+1)}\nu_2(y)^{1/(k_1+1)} \tag{40}$$

where

$$\nu_1 = \beta_1^{-1}1_{B_1} \quad \text{and} \quad \nu_2 = \beta_2^{-1}1_{B_2} \tag{41}$$

then $T_{\Delta_{k_1},\Delta_{k_2}}(\widetilde{\nu},\ldots,\widetilde{\nu})(\lambda) = T_{\Delta_{k_1}}(\nu_1,\ldots,\nu_1)(\lambda)\cdot T_{\Delta_{k_2}}(\nu_2,\ldots,\nu_2)(\lambda)$ and in light of Proposition 3.1 we can conclude that

$$T_{\Delta_{k_1},\Delta_{k_2}}(\widetilde{\nu},\ldots,\widetilde{\nu})(\lambda) = 1 + O_{k_1,k_2}(\beta_1^{-k_1-1}\beta_2^{-k_2-1}c_{\Delta_{k_1}}^{-1/6}c_{\Delta_{k_2}}^{-1/6}\varepsilon^{2/3}) \tag{42}$$

for any $0 < \lambda \leq \varepsilon \ll \min\{1, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}$, provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$.

In this setting we have the following "generalized von-Neumann inequality", for which it is essential that our count of product simplices is taken relative to suitably uniformly distributed sets B_1 and B_2.

Lemma 6.1 (Generalized von-Neumann for $\Delta_{k_1} \times \Delta_{k_2}$ relative to $B_1 \times B_2$)**.** *Let*

$$\widetilde{\nu} = \nu_1^{1/(k_2+1)} \otimes \nu_2^{1/(k_1+1)} \quad \text{and} \quad \nu = \nu_1^{1/2} \otimes \nu_2^{1/2}$$

where $\nu_1 = \beta_1^{-1} 1_{B_1}$ and $\nu_2 = \beta_2^{-1} 1_{B_2}$ For any $0 < \lambda \leq \varepsilon \ll \min\{1, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}$ and functions

$$f_{ij} : [0,1]^{d_1} \times [0,1]^{d_2} \to [-1,1]$$

with $(i,j) \in \{0,1,\ldots,k_1\} \times \{0,1,\ldots,k_2\}$ we have

$$|T_{\Delta_{k_1},\Delta_{k_2}}(f_{00}\widetilde{\nu}, \ldots, f_{k_1k_2}\widetilde{\nu})(\lambda)| \leq \min_{\substack{i=0,1,\ldots,k_1\\ j=0,1,\ldots,k_2}} \|f_{ij}\nu\|_{\square(\varepsilon^4\lambda)} + O_{k_1,k_2}(\beta_1^{-k_1-1}\beta_2^{-k_2-1} c_{\Delta_{k_1}}^{-1/8} c_{\Delta_{k_2}}^{-1/8} \varepsilon^{1/16})$$

provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

It is easy to see that Lemma 6.1, combined with Proposition 3.1, gives the following

Corollary 6.1. *Let $0 < \alpha, \beta_1, \beta_2 \leq 1$ and*

$$0 < \lambda \leq \varepsilon \ll_{k_1,k_2} (\beta_1^{k_1+1}\beta_2^{k_2+1}\alpha^{(k_1+1)(k_2+1)})^{16} \min\{(c_{\Delta_{k_1}}c_{\Delta_{k_2}})^2, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}.$$

If $A \subseteq B_1 \times B_2 \subseteq [0,1]^{d_1} \times [0,1]^{d_2}$ with $|A| = \alpha\beta_1\beta_2$ and $\|f_A\nu\|_{\square(\varepsilon^4\lambda)} \ll \alpha^{(k_1+1)(k_2+1)}$, then

$$T_{\Delta_{k_1},\Delta_{k_2}}(1_A\widetilde{\nu}, \ldots, 1_A\widetilde{\nu})(\lambda) \geq \frac{1}{2}\alpha^{(k_1+1)(k_2+1)}$$

provided B_1 and B_2 are $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ respectively.

Proof of Corollary 6.1. It follows immediately from Lemma 6.1 that

$$\begin{aligned}|T_{\Delta_{k_1},\Delta_{k_2}}(1_A\widetilde{\nu}, \ldots, 1_A\widetilde{\nu})(\lambda) - &T_{\Delta_{k_1},\Delta_{k_2}}(\alpha\widetilde{\nu}, \ldots, \alpha\widetilde{\nu})(\lambda)| \\ &\leq (2^{(k_1+1)(k_2+1)} - 1)\|f_A\nu\|_{\square(\varepsilon^4\lambda)} + O_{k_1,k_2}(\beta_1^{-k_1-1}\beta_2^{-k_2-1} c_{\Delta_{k_1}}^{-1/8} c_{\Delta_{k_2}}^{-1/8} \varepsilon^{1/16})\end{aligned}$$

for any $0 < \varepsilon, \lambda \ll \min\{1, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}$, where $f_A = 1_A - \alpha 1_{B_1\times B_2}$ while, as noted in (42), Proposition 3.1 implies that

$$T_{\Delta_{k_1},\Delta_{k_2}}(\alpha\widetilde{\nu}, \ldots, \alpha\widetilde{\nu})(\lambda) = \alpha^{(k_1+1)(k_2+1)}(1 + O_{k_1,k_2}(\beta_1^{-k_1-1}\beta_2^{-k_2-1} c_{\Delta_{k_1}}^{-1/6} c_{\Delta_{k_2}}^{-1/6} \varepsilon^{2/3}))$$

for any $0 < \lambda \leq \varepsilon \ll \min\{1, c_{\Delta_{k_1}}^{-1}, c_{\Delta_{k_2}}^{-1}\}$, as required. □

6.2. **A Relative Version of Lemma 3.1.** Key to the proof of Lemma 6.1 is the following

Lemma 6.2 (Lemma 3.1 relative to uniformly distributed sets)**.** *Let $\Delta_k = \{0, v_1, v_2, \ldots, v_k\}$ be any non-degenerate k-dimensional simplex and $B \subseteq [0,1]^d$ with $d \geq k+3$ be an arbitrary set with $\beta = |B| > 0$.*

If we set $\nu = \beta^{-1}1_B$, then for any $0 < \lambda \leq \varepsilon \ll \min\{c_{\Delta_k}, c_{\Delta_k}^{-1}\}$ and functions $f_0, f_1, \ldots, f_k : [0,1]^d \to [-1,1]$ we have

$$|T_{\Delta_k}(f_0\nu, \ldots, f_k\nu)(\lambda)|^2 \leq \iint f_j\nu(x) f_j\nu(x - x_1)\psi_{\varepsilon^4\lambda}(x_1)\, dx\, dx_1 + O_k(\beta^{-3k-3} c_{\Delta_k}^{-1/2} \varepsilon^{1/4}) \tag{43}$$

for any $0 \leq j \leq k$, provided B is a $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subset of $[0,1]^d$.

Proof. As in the proof of Lemma 3.1 it suffices, by symmetry, to establish (43) for $j = k$. Note also, as in (42) above, that Proposition 3.1 implies

$$T_{\Delta_k}(\nu, \ldots, \nu)(\lambda) = 1 + O_k(\beta^{-k-1} c_{\Delta_k}^{-1/6} \varepsilon^{2/3}), \tag{44}$$

for any non-degenerate k-dimensional simplex Δ_k, provided $0 < \lambda \leq \varepsilon \ll \min\{1, c_{\Delta_k}^{-1}\}$ and B is an $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subset of $[0,1]^d$ with $d \geq k+1$. It is equally easy to see, using Lemma 3.1, that if $1 \leq j \leq k$ and any j of the weights ν are replaced with $1_{[0,1]^d}$ then this modified count will still be asymptotically equal to 1 and will in fact equal $1 + O_k(\beta^{-k-1+j} c_{\Delta_k}^{-1/6} \varepsilon^{2/3})$.

Since

$$|T_{\Delta_k}(f_0\nu,\dots,f_k\nu)(\lambda)| \le \iint\cdots\int \nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-1})\Big|\int f_k\nu(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\Big|$$
$$d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx$$

it follows from an application of Cauchy-Schwarz, facilitated by (44) for the simplex Δ_{k-1}, that

$$|T_{\Delta_k}(f_0\nu,\dots,f_k\nu)(\lambda)|^2 \le \big(1+O_k(\beta^{-k}c_{\Delta_k}^{-1/6}\varepsilon^{2/3})\big)^2\big(\,M(\lambda)+E(\lambda)\,\big)$$

where

$$M(\lambda)=\iint\cdots\int\Big|\int f_k\nu(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\Big|^2 d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx$$

and

$$E(\lambda)=\iint\cdots\int\big[\nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-1})-1(x)\big]\Big|\int f_k\nu(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\Big|^2$$
$$d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx$$

where $1=1_{[0,1]^d}$.

It follows from the proof of Lemma 3.1, specifically the argument from (13) to (16)) that

$$M(\lambda)\le\iint f_k\nu(x_1)f_k\nu(x_2)\psi_{\varepsilon^4\lambda}(x_2-x_1)\,dx_1\,dx_2.$$

We now complete the proof by establishing that $E(\lambda)=O_k(\beta^{-k-3}c_{\Delta_k}^{-1/6}\varepsilon^{1/4})$. Our strategy will be to expand the square in the error term $E(\lambda)$ which will add a new vertex x_{k+1} to the simplex. "Fixing" the distance $|x_{k+1}-x_k|$ leads to an expression which may be viewed as the difference between a weighted and an unweighted average over all isometric copies of a fixed $(k+1)$-dimensional simplex. The reason that this difference is small is that the measure ν behaves suitably random with respect to averages of this type, expressed in (44). To remove the uncontrolled terms f_k one needs another application of Cauchy-Schwarz which leads to simplices of dimension $k+2$ and the requirement $d\ge k+3$ for the underlying dimension of the space.

Writing

$$\nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-1})-1(x)=\sum_{j=0}^{k-1}\big[\nu(x-\lambda x_j)-1(x)\big]\nu(x-\lambda x_{j+1})\cdots\nu(x-\lambda x_{k-1})$$

with the understanding that $x_0=0$, it follows that

$$E(\lambda)=\sum_{j=0}^{k-1}E_j(\lambda)$$

with

$$E_j(\lambda)=\iint\cdots\int\big[\nu(x-\lambda x_j)-1(x)\big]\nu(x-\lambda x_{j+1})\cdots\nu(x-\lambda x_{k-1})\Big|\int f_k\nu(x-\lambda x_k)\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\Big|^2$$
$$d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx.$$

Squaring out we see that

$$E_j(\lambda)=\iint\cdots\iiint\big[\nu(x-\lambda x_j)-1(x)\big]\nu(x-\lambda x_{j+1})\cdots\nu(x-\lambda x_{k-1})f_k\nu(x-\lambda x_k)f_k\nu(x-\lambda x_{k+1})$$
$$d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_{k+1})\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\,d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx.$$

Since $d\ge k+3$ we can follow the argument in Section 3.3 and write

$$\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_{k+1})=\int_0^\pi(\sin\theta_1)^{d-k-1}\,d\sigma^{(d-k-1)}_{x_1,\dots,x_{k-1},x_k,\theta_1}(x_{k+1})\,d\theta_1$$

where $\sigma^{(d-k-1)}_{x_1,\dots,x_{k-1},x_k,\theta_1}(x_{k+1})$ denotes the normalized measure on the sphere

$$S^{d-k-1}_{x_1,\dots,x_{k-1},x_k,\theta_1}=S^{d-1}(0,|v_{k+1}|)\cap S^{d-1}(x_1,|v_{k+1}-v_1|)\cap\cdots\cap S^{d-1}(x_k,|v_{k+1}-v_k|)$$

with $v_{k+1} = v_{k+1}(\theta)$ satisfying $|v_{k+1}| = |v_k|$, $|v_{k+1} - v_j| = |v_k - v_j|$ for all $1 \le j \le k-1$ and θ_1 determining the angle between v_{k+1} and v_k so that $|v_{k+1} - v_k| = 2|v_k|\sin(\theta_1/2)$.

If we now let $\Delta_{k+1}(\theta_1) = \{0, v_1, \dots, v_k, v_{k+1}\}$, then it follows (again using (44)) that

$$E_j(\lambda) = \int_0^\pi (\sin\theta_1)^{d-k-1} T_{\Delta_{k+1}(\theta)}(1,\dots,1,\nu-1,\nu,\dots,\nu,f_k\nu,f_k\nu)(\lambda)\,d\theta_1 + O_k(\beta^{-k-1+j} c_{\Delta_k}^{-1/6}\varepsilon^{2/3})$$

where $T_{\Delta_{k+1}(\theta)}(1,\dots,1,\nu-1,\nu,\dots,\nu,f_k\nu,f_k\nu)(\lambda)$ equals

$$\iint\cdots\iiint \big[\nu(x-\lambda x_j)-1\big]\nu(x-\lambda x_{j+1})\cdots\nu(x-\lambda x_{k-1}) f_k\nu(x-\lambda x_k) f_k\nu(x-\lambda x_{k+1})$$
$$d\sigma^{(d-k-1)}_{x_1,\dots,x_{k-1},x_k,\theta_1}(x_{k+1})\,d\sigma^{(d-k)}_{x_1,\dots,x_{k-1}}(x_k)\,d\sigma^{(d-k+1)}_{x_1,\dots,x_{k-2}}(x_{k-1})\cdots d\sigma(x_1)\,dx.$$

In light of (19) and (20) it suffices to now show that

$$E'_j(\lambda) := \int_0^\pi (\sin\theta_1)^{d-k-1} T_{\Delta'_{k+1-j}(\theta_1)}(f_k\nu,f_k\nu,\nu,\dots\nu,\nu-1)(\lambda)\,d\theta_1 = O(\beta^{-k-1+j}\varepsilon^{1/4})$$

where

$$\Delta'_{k+1-j}(\theta_1) = \{0, v'_1, \dots, v'_{k+1-j}\}$$

with $v'_i = v_{k+1-i} - v_{k+1}$ for $0 \le i \le k$ and $v'_{k+1} = -v_{k+1}$.

Since $|T_{\Delta'_{k+1-j}(\theta_1)}(f_k\nu,f_k\nu,\nu,\dots\nu,\nu-1)(\lambda)|$ is dominated by

$$\iint\cdots\int \nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-j})\Big|\int(\nu-1)(x-\lambda x_{k+1-j})\,d\sigma'^{(d-k-1+j)}_{x_1,\dots,x_{k-j}}(x_{k+1-j})\Big|$$
$$d\sigma'^{(d-k+j)}_{x_1,\dots,x_{k-j-1}}(x_{k-j})\cdots d\sigma'(x_1)\,dx$$

it follows from an application of Cauchy-Schwarz, facilitated by (44) for the simplex $\Delta'_{k-j}(\theta_1)$, that

$$|T_{\Delta'_{k+1-j}(\theta_1)}(f_k\nu,f_k\nu,\nu,\dots,\nu,\nu-1)(\lambda)|^2 \le 2\,I_{\Delta'_{k+1-j}(\theta_1)}(\lambda)$$

where

$$I_{\Delta'_{k+1-j}(\theta_1)}(\lambda) = \iint\cdots\int \nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-j})\Big|\int(\nu-1)(x-\lambda x_{k+1-j})\,d\sigma'^{(d-k-1+j)}_{x_1,\dots,x_{k-j}}(x_{k+1-j})\Big|^2$$
$$d\sigma'^{(d-k+j)}_{x_1,\dots,x_{k-j-1}}(x_{k-j})\cdots d\sigma'(x_1)\,dx.$$

Squaring out we see that $I_{\Delta'_{k+1-j}(\theta_1)}(\lambda)$ equals

$$\iint\cdots\iiint \nu(x)\nu(x-\lambda x_1)\cdots\nu(x-\lambda x_{k-j})(\nu-1)(x-\lambda x_{k+1-j})\,(\nu-1)(x-\lambda x_{k+2-j})$$
$$d\sigma'^{(d-k-1+j)}_{x_1,\dots,x_{k-j}}(x_{k+2-j})\,d\sigma'^{(d-k-1+j)}_{x_1,\dots,x_{k-j}}(x_{k+1-j})\,d\sigma'^{(d-k+j)}_{x_1,\dots,x_{k-j-1}}(x_{k-j})\cdots d\sigma'(x_1)\,dx.$$

Since $d \ge k+3$ we can again argue as above to obtain

$$I_{\Delta'_{k+1-j}(\theta_1)}(\lambda) = \int_0^\pi (\sin\theta_2)^{d-k-2+j}\big[T_1(\lambda) - T_2(\lambda) - T_3(\lambda) + T_4(\lambda)\big]\,d\theta_2$$

where

$$T_1(\lambda) = T_{\Delta'_{k+2-j}(\theta_1,\theta_2)}(\nu,\dots,\nu)(\lambda)$$
$$T_2(\lambda) = T_{\Delta'_{k+2-j}(\theta_1,\theta_2)}(\nu,\dots,\nu,1)(\lambda)$$
$$T_3(\lambda) = T_{\Delta'_{k+2-j}(\theta_1,\theta_2)}(\nu,\dots,1,\nu)(\lambda)$$
$$T_4(\lambda) = T_{\Delta'_{k+2-j}(\theta_1,\theta_2)}(\nu,\dots,\nu,1,1)(\lambda)$$

with

$$\Delta'_{k+2-j}(\theta_1,\theta_2) = \Delta'_{k+1-j}(\theta_1) \cup \{v'_{k+2-j}\}$$

with $v'_{k+2-j} = v'_{k+2-j}(\theta_2)$ satisfying $|v'_{k+2-j}| = |v'_{k+1-j}|$, $|v'_{k+2-j} - v_i| = |v'_{k+1-j} - v_i|$ for all $1 \le i \le k-j$ and θ_2 determining the angle between v'_{k+2-j} and v'_{k+1-j} so that $|v'_{k+2-j} - v'_{k+1-j}| = 2|v_j|\sin(\theta_2/2)$.

We have therefore ultimately established

$$|E'_j(\lambda)|^2 \le C\int_0^\pi\int_0^\pi \big|T_1(\lambda) - T_2(\lambda) - T_3(\lambda) + T_4(\lambda)\big|\,d\theta_2\,d\theta_1$$

for each $0 \le j \le k$. In light of (44) we know that

$$T_i(\lambda) = 1 + O_k(\beta^{-k-3+j} c_{\Delta'_{k+2-j}(\theta_1,\theta_2)}^{-1/6} \varepsilon^{2/3})$$

for $i = 1, \ldots, 4$, and hence

$$E'_j(\lambda) = O_k(\beta^{-k-3+j} \varepsilon^{1/4})$$

provided

$$c_{\Delta'_{k+2-j}(\theta_1,\theta_2)} \ge \varepsilon.$$

The result now follows since the fact that

$$c_{\Delta'_{k+2-j}(\theta_1,\theta_2)} = \min\{c_{\Delta_k}, 2|v_k|\sin(\theta_1/2), 2|v_j|\sin(\theta_2/2)\}$$

and $\varepsilon \ll c_{\Delta_k}$ ensures that

$$|\{(\theta_1,\theta_2) \in [0,\pi]\times[0,\pi] \,:\, c_{\Delta'_{k+2-j}(\theta_1,\theta_2)} \le \varepsilon\}| = O(\varepsilon). \qquad \square$$

6.3. **Proof of Lemma 6.1.** The proof of Lemma 6.1 will follow from two applications of Cauchy-Schwarz combined with Proposition 3.1 and Lemma 6.2. We first observe that if

$$T_{\Delta_{k_1},\Delta_{k_2}}(\lambda) := T_{\Delta_{k_1},\Delta_{k_2}}(f_{00}\widetilde{\nu}, \ldots, f_{k_1k_2}\widetilde{\nu})(\lambda)$$

then

$$T_{\Delta_{k_1},\Delta_{k_2}}(\lambda) = \iint \nu_1(x - \lambda U_1(v_0^1)) \cdots \nu_1(x - \lambda U_1(v_{k_1}^1))\, T_{\Delta_{k_2}}(g_0^{x,U_1}\nu_2, g_1^{x,U_1}, \ldots, g_{k_2}^{x,U_1}\nu_2)(\lambda)\, d\mu_1(U_1)\, dx$$

where

$$g_j^{x,U_1}(y) = f_{0j}(x - \lambda\cdot U_1(v_0^1), y) \cdots f_{k_1j}(x - \lambda\cdot U_1(v_{k_1}^1), y)$$

for each $j = 0, 1, \ldots, k_2$ and that Lemma 6.2 implies

$$\left|T_{\Delta_{k_2}}(g_0^{x,U_1}\nu_2, \ldots, g_{k_2}^{x,U_1}\nu_2)(\lambda)\right|^2 \le \iint g_j^{x,U_1}\nu_2(y) g_j^{x,U_1}\nu_2(y-y_1)\psi_{2,\varepsilon^4\lambda}(y_1)\,dy\,dy_1 + O_{k_2}(\beta^{-3k_2-3} c_{\Delta_{k_2}}^{-1/2}\varepsilon^{1/4})$$

for any $0 \le j \le k_2$. Hence by Cauchy-Schwarz, using (44) for $T_{\Delta_{k_1}}(\nu_1, \ldots, \nu_1)(\lambda)$, and switching the order of integration we obtain that $|T_{\Delta_{k_1},\Delta_{k_2}}(\lambda)|^2$ is majorized by

$$\iint T_{\Delta_{k_1}}(h_{0j}^{y,y_1}\nu_1, \ldots, h_{k_1j}^{y,y_1}\nu_1)(\lambda)\,\nu_2(y)\nu_2(y-y_1)\psi_{2,\varepsilon^4\lambda}(y_1)\,dy\,dy_1 + O_{k_1,k_2}(\beta_1^{-k_1-1}\beta_2^{-3k_2-3} c_{\Delta_{k_1}}^{-1/6} c_{\Delta_{k_2}}^{-1/2}\varepsilon^{1/4})$$

for any $0 \le j \le k_2$ where

$$h_{ij}^{y,y_1}(x) = f_{ij}(x,y) f_{ij}(x, y-y_1)$$

for $i = 0, 1, \ldots, k_1$. A further application of Cauchy-Schwarz (using the fact that $\psi_{2,\varepsilon^4\lambda}$ is L^1-normalized) and appeal to Lemma 6.2 reveals that $|T_{\Delta_{k_1},\Delta_{k_2}}(\lambda)|^4$ is majorized by

$$\iiiint h_{ij}^{y,y_1}\nu_1(x) h_{ij}^{y,y_1}\nu_1(x-x_1)\,\nu_2(y)\nu_2(y-y_1)\psi_{1,\varepsilon^4\lambda}(x_1)\,\psi_{2,\varepsilon^4\lambda}(y_1)\,dx\,dx_1\,dy\,dy_1$$
$$+ O_{k_1,k_2}(\beta_1^{-4k_1-4}\beta_2^{-3k_2-4} c_{\Delta_{k_1}}^{-1/2} c_{\Delta_{k_2}}^{-1/2}\varepsilon^{1/4})$$

for any $0 \le i \le k_1$ and $0 \le j \le k_2$. Since

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

the result follows from (32). $\square$

6.4. **Inverse Theorem Revisited.** We complete this section by noting the following immediate consequence of Theorem 5.1 which together with Corollary 6.1 implies Proposition 4.4.

Corollary 6.2. *Let $0 < \alpha, \beta_1, \beta_2 \le 1$ and B_1 and B_2 be $(\varepsilon, \varepsilon^4\lambda)$-uniformly distributed subsets of $[0,1]^{d_1}$ and $[0,1]^{d_2}$ with $0 < \lambda \le \varepsilon \ll \beta_1^{k_1+1}\beta_2^{k_2+1}\alpha^{8(k_1+1)(k_2+1)}$. If $A \subseteq B_1 \times B_2 \subseteq [0,1]^{d_1} \times [0,1]^{d_2}$ with $|A| = \alpha\beta_1\beta_2$ and*

$$\|f_A\nu\|_{\square(\varepsilon^4\lambda)} \gg \alpha^{(k_1+1)(k_2+1)}$$

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

$$\frac{|A \cap (B'_1 \times B'_2)|}{|B'_1 \times B'_2|} \ge \alpha + c\,\alpha^{8(k_1+1)(k_2+1)}.$$

7. Proof of Proposition 4.1, Part II: Regularization

To complete the proof of Proposition 4.1, as was noted after the Proposition 4.3, 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 4.3 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 4.3 by appealing to a version of Szemerédi's Regularity Lemma [4] adapted to a sequence of scales $\{L_j\}_{1\leq j\leq J}$.

The precise result we need is stated below in Theorem 7.1, but first we state a couple of definitions.

Definition 7.1 (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 $[0,1]^{d_1} \times [0,1]^{d_2}$ 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 7.2 ((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

$$\frac{1}{|Q|}\int_Q \left|\frac{|B\cap(t+Q_L)|}{|Q_L|} - \frac{|B|}{|Q|}\right|^2 dt \leq \varepsilon^2. \tag{45}$$

Theorem 7.1 (Regularity Lemma)**.** *Let $0 < \beta_1, \beta_2, \eta \leq 1$ and $B_i \subseteq [0,1]^{d_i}$ 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 $[0,1]^{d_1} \times [0,1]^{d_2}$ 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 7.1 follows by standard arguments, for completeness we include it in Section 7.1.

An almost immediate consequence of Theorem 7.1 is the following Corollary which, together with Proposition 4.3, provides a complete proof of Proposition 4.1, the easy verification of this we leave to the reader.

Corollary 7.1. *Let $0 < \alpha, \beta_1, \beta_2, \tau, \varepsilon \leq 1$ and $A \subseteq B_1 \times B_2 \subseteq [0,1]^{d_1} \times [0,1]^{d_2}$ 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 7.1 implies Corollary 7.1. Let $\eta = \varepsilon\beta_1\beta_2\tau/3$ and $\mathcal{P} = \mathcal{Q}\cup\mathcal{R}$ be a partition of $[0,1]^{d_1} \times [0,1]^{d_2}$ adapted to the scale L_j that satisfies the conclusions of Theorem 7.1 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 7.1 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 7.1 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 7.1 must hold for at least one cube Q in $\mathcal{U}\setminus\mathcal{S}$. □

7.1. **Proof of Theorem 7.1.** 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 $j=0,1,2,\dots$ we construct partitions $\mathcal{P}^{(j)}$ of $[0,1]^{d_1}\times[0,1]^{d_2}$ into cubes $\mathcal{Q}^{(j)}$ and rectangles $\mathcal{R}^{(j)}$ starting from the trivial partition $\mathcal{P}^{(0)}$ consisting of only one cube $Q=[0,1]^{d_1}\times[0,1]^{d_2}$. 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

$$\sum_{R\in\mathcal{R}^{(j)}}|R|\leq\frac{\eta}{2},\tag{46}$$

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

$$\sum_{Q\in\mathcal{N}^{(j)}}|Q|\leq\frac{\eta}{2}\tag{47}$$

and note that such a partition satisfies the conclusions of Theorem 7.1.

If $[0,1]^{d_1}\times[0,1]^{d_2}\in\mathcal{U}^{(0)}$, then the sets B_1, B_2 are both (ε,L_1)-uniformly distributed and Theorem 7.1 holds. We thus assume that for some $j\geq 0$ we have a partition $\mathcal{P}^{(j)}$ for which (47) 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 $Q_1=t_1+[0,L_j]^{d_1}$ and $L:=L_{j+1}$, we have

$$|E_\eta|\geq\frac{\eta^2}{2}|Q_1|\tag{48}$$

where

$$E_\eta:=\left\{t\in Q_1\ :\ \left|\frac{|B_1\cap(t+Q_L)|}{|Q_L|}-\frac{|B_1\cap Q_1|}{|Q_1|}\right|\geq\frac{\eta}{2}\right\}.\tag{49}$$

Let $m=\lfloor L_j/L_{j+1}\rfloor$ and partition the cube $Q_1'=t_1+[0,(m+1)L]^{d_1}\supseteq Q_1$ into grids of the form $G(s_1)=s_1+\{0,L,\dots,mL\}^{d_1}$ with s_1 running through the cube $t_1+[0,L]^{d_1}$. Since $L<2^{-6}L_j$, by (48) there exist $s_1\in Q_1'$ such that

$$\frac{|G(s_1)\cap E_\eta|}{m^{d_1}}\geq\frac{\eta^2}{4}.\tag{50}$$

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 $(d_1\times d_2)$-dimensional rectangles. After performing this partition of all cubes in $\mathcal{N}^{(j)}$ we obtain the new partition $\mathcal{P}^{(j+1)}$ of $[0,1]^{d_1}\times[0,1]^{d_2}$. 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 (46) 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

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j)}) := \frac{1}{2} \sum_{C^k \in \mathcal{P}^{(j)}} \big((\delta_1^k)^2 + (\delta_2^k)^2\big)\, |C^k|. \tag{51}$$

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

$$\sum_\ell (\delta_1^\ell)^2\, |Q_1^\ell| = (\delta_1)^2\, |Q_1| + \sum_\ell (\delta_1^\ell - \delta_1)^2\, |Q_1^\ell|. \tag{52}$$

Similarly

$$\sum_{\ell'} (\delta_2^{\ell'})^2\, |Q_2^l| = (\delta_2)^2\, |Q_2| + \sum_{\ell'} (\delta_2^{\ell'} - \delta_2)^2\, |Q_2^{\ell'}|. \tag{53}$$

Multiplying equations (52) by $|Q_2|$, (53) by $|Q_1|$, and adding, we get

$$\sum_{\ell,\ell'} \big((\delta_1^\ell)^2 + (\delta_2^{\ell'})^2\big)\, |C^{\ell,\ell'}| = \big((\delta_1)^2 + (\delta_2)^2\big)\, |C| + \sum_{\ell,\ell'} \big((\delta_1^\ell - \delta_1)^2 + (\delta_2^{\ell'} - \delta_2)^2\big)\, |C^{\ell,\ell'}|. \tag{54}$$

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 $\ell \in \{1, \ldots, m\}^{d_1}$ and $\ell' \in \{1, \ldots, m\}^{d_2}$, and into a collection of $(d_1 + d_2)$-dimensional rectangles of small total measure. By (50) there at least $\eta^2 m^{d_1}/4$ values of ℓ for which $|\delta_1^\ell - \delta_1|^2 \geq \eta^2/4$. Thus, as $|Q_1| = L_j^{d_1}$, $|Q_1^\ell| = L_{j+1}^{d_1}$ for all ℓ, and $m = \lfloor L_j/L_{j+1} \rfloor \geq \frac{1}{2} L_j/L_{j+1}$, we have that

$$\sum_{\ell \in \{1,\ldots,m\}^{d_1}} (\delta_1^\ell - \delta_1)^2\, |Q_1^\ell| \geq \frac{\eta^4}{64}\, |Q_1|. \tag{55}$$

By (54) 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 (54) is at least

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j+1)}|_{C^k}) \;\geq\; \frac{1}{2}\, (\delta_1^2 + \delta_2^2)\, |C^k| + \frac{\eta^4}{128}\, |C^k|. \tag{56}$$

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

$$\mathcal{E}(B_1, B_2; \mathcal{P}^{(j+1)}) \;\geq\; \mathcal{E}(B_1, B_2; \mathcal{P}^{(j)}) \;+\; \frac{\eta^5}{256}. \tag{57}$$

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 7.1. $\square$

Acknowledgements. We would like to thank the anonymous referee for useful comments and suggestions which have greatly improved the exposition of this article.

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] E. Stein, *Harmonic Analysis: Real Variable Methods, Orthogonality and Oscillatory Integrals*, Princeton University Press, Princeton, NJ., 1993.
[4] 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