WEAK HYPERGRAPH REGULARITY AND APPLICATIONS TO GEOMETRIC RAMSEY THEORY

NEIL LYALL ÁKOS MAGYAR

ABSTRACT. Let $\Delta = \Delta_1 \times \ldots \times \Delta_d \subseteq \mathbb{R}^n$, where $\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$ and each with each $\Delta_i \subseteq \mathbb{R}^{n_i}$ a non-degenerate simplex of n_i points. We prove that any set $S \subseteq \mathbb{R}^n$, with $n = n_1 + \cdots + n_d$ of positive upper Banach density necessarily contains an isometric copy of all sufficiently large dilates of the configuration Δ. In particular any such set $S \subseteq \mathbb{R}^{2d}$ contains a d-dimensional cube of side length λ, for all $\lambda \geq \lambda_0(S)$. We also prove analogous results with the underlying space being the integer lattice. The proof is based on a weak hypergraph regularity lemma and an associated counting lemma developed in the context of Euclidean spaces and the integer lattice.

1. INTRODUCTION

1.1. **Existing Results I: Distances and Simplices in Subsets of $\mathbb{R}^n$.** Recall that the *upper Banach density* of a measurable set $S \subseteq \mathbb{R}^n$ is defined by

$$\delta^*(S) = \lim_{N\to\infty} \sup_{t\in\mathbb{R}^n} \frac{|S \cap (t + Q(N))|}{|Q(N)|}, \tag{1.1}$$

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

A result of Katznelson and Weiss [6] states that if $S \subseteq \mathbb{R}^2$ has positive upper Banach density, then its distance set $\{|x - x'| : x, x' \in S\}$ contains all sufficiently large numbers. Note that the distance set of any set of positive Lebesgue measure in $\mathbb{R}^n$ automatically contains all sufficiently small numbers (by the Lebesgue density theorem) and that it is easy to construct a set of positive upper density which does not contain a fixed distance by placing small balls centered on an appropriate square grid.

Theorem A (Katznelson and Weiss [6]). *If $S \subseteq \mathbb{R}^2$ with $\delta^*(S) > 0$, then there exists a threshold $\lambda_0 = \lambda_0(S)$ such that S is guaranteed to contain pairs of points $\{x_1, x_2\}$ with $|x_2 - x_1| = \lambda$ for all $\lambda \geq \lambda_0$.*

This result was later reproved using Fourier analytic techniques by Bourgain in [1] where he established the following more general result for all configurations of n points in $\mathbb{R}^n$ whose affine span is $n-1$ dimensional, namely for all non-degenerate simplices.

Theorem B (Bourgain [1]). *Let $\Delta \subseteq \mathbb{R}^n$ be a non-degenerate simplex of n points. If $S \subseteq \mathbb{R}^n$ with $\delta^*(S) > 0$, then there exists a threshold $\lambda_0 = \lambda_0(S, \Delta)$ such that S contains an isometric copy of $\lambda\Delta$ for all $\lambda \geq \lambda_0$.*

Recall that a finite point configuration Δ' is said to be an isometric copy of $\lambda\Delta$ if there exists a bijection $\phi : \Delta \to \Delta'$ such that $|\phi(v) - \phi(w)| = \lambda |v - w|$ for all $v, w \in \Delta$, i.e. if Δ' is obtained from $\lambda\Delta$ (the dilation of Δ by a factor λ) via a rotation and translation.

Bourgain deduced Theorem B as an immediate consequence of the following stronger quantitative result for measurable subsets of the unit cube of positive measure. In the proposition below, and throughout this article, we shall refer to a decreasing sequence $\{\lambda_j\}_{j=1}^J$ as *lacunary* if $\lambda_{j+1} \leq \lambda_j/2$ for all $1 \leq j < J$.

Proposition B (Bourgain [1]). *Let $\Delta \subseteq \mathbb{R}^n$ be a non-degenerate simplex of n points. For any $0 < \delta \leq 1$ there exists a constant $J = O_\Delta(\delta^{-3n})$ such that if $1 \geq \lambda_1 \geq \cdots \geq \lambda_J$ is any lacunary sequence and $S \subseteq [0,1]^n$ with $|S| \geq \delta$, then there exists $1 \leq j < J$ such that S contains an isometric copy of $\lambda\Delta$ for all $\lambda \in [\lambda_{j+1}, \lambda_j]$.*

2010 *Mathematics Subject Classification.* 11B30.

The first and second authors were partially supported by grants NSF-DMS 1702411 and NSF-DMS 1600840, respectively.

In [12] the authors provided a short direct proof of Theorem B without using Proposition B. It is based on the observation that uniformly distributed sets $S \subseteq \mathbb{R}^d$ contain the expected "number" of isometric copies of dilates $\lambda\Delta$ and that all sets of positive upper density become uniformly distributed at sufficiently large scales. However, for the purposes of this paper it will be important to recall Bourgain's indirect approach.

To see that Proposition B implies Theorem B notice that if Theorem B were not to hold for some set $S \subseteq \mathbb{R}^n$ of upper Banach density $\delta^*(S) > \delta > 0$, then there must exist a lacunary sequence $\lambda_1 \geq \cdots \geq \lambda_J \geq 1$, with J the constant in Proposition B, such that S does not contain an isometric copy of $\lambda_j\Delta$ for any $1 \leq j \leq J$. Taking a sufficiently large cube Q with side length $N \geq \lambda_1$ and $|S \cap Q| \geq \delta|Q|$ and scaling back $Q \to [0,1]^n$ contradicts Proposition B.

We further note that by taking $\lambda_j = 2^{-j}$ in Proposition B we obtain the following "Falconer-type" result for subsets of $[0,1]^n$ of positive Lebesgue measure.

Corollary B. *If $\Delta \subseteq \mathbb{R}^n$ is a non-degenerate simplex of n points, then any $S \subseteq [0,1]^n$ with $|S| > 0$ will necessarily contain an isometric copy of $\lambda\Delta$ for all λ in some interval of length at least* $\exp(-C_\Delta|S|^{-3n})$.

Bourgain further demonstrated in [1] that no result along the lines of Theorem B can hold for configurations that contain any three points in arithmetic progression along a line, specifically showing that for any $n \geq 1$ there are sets of positive upper Banach density in $\mathbb{R}^n$ which do not contain an isometric copy of configurations of the form $\{0, y, 2y\}$ with $|y| = \lambda$ for all sufficiently large λ. This should be contrasted with the following remarkable result of Tamar Ziegler.

Theorem C (Ziegler [25])**.** *Let $\mathcal{F}$ be any configuration of k points in $\mathbb{R}^n$ with $n \geq 2$.*

If $S \subseteq \mathbb{R}^n$ has positive upper density, then there exists a threshold $\lambda_0 = \lambda_0(S, \mathcal{F})$ such that S_ε contains an isometric copy of $\lambda\mathcal{F}$ for all $\lambda \geq \lambda_0$ and any $\varepsilon > 0$, where S_ε denotes the ε-neighborhood of S.

Bourgain's example was later generalized by Graham [9] to establish that the condition that $\varepsilon > 0$ in Theorem C is necessary and cannot be strengthened to $\varepsilon = 0$ for any given *non-spherical* configuration $\mathcal{F}$ in $\mathbb{R}^n$ for any $n \geq 1$, that is for any finite configuration of points that cannot be inscribed in some sphere. We note that the sets constructed by Bourgain and Graham have the property that for any $\varepsilon > 0$ their ε-neighborhoods will contain arbitrarily large cubes and hence trivially satisfy Theorem C with $\lambda_0 = 0$.

It is natural to ask if *any* spherical configuration $\mathcal{F}$, beyond the known example of simplices, has the property that every positive upper Banach density subset of $\mathbb{R}^n$, for some sufficiently large n, contains an isometric copy of $\lambda\mathcal{F}$ for all sufficiently large λ, and even to conjecture that this ought to hold for *all* spherical configurations. The first breakthrough in this direction came in [12] when the authors established this for configurations of four points forming a 2-dimensional rectangle in $\mathbb{R}^4$ and more generally for any configuration that is the direct product of two non-degenerate simplices in $\mathbb{R}^n$ for suitably large n.

The purpose of this article is to present a strengthening of the results in [12] and to extend them to cover configurations with a higher dimensional product structure in both the Euclidean and discrete settings.

1.2. New Results I: Rectangles and Products of Simplices in Subsets of $\mathbb{R}^n$.

The first main result of this article is the following

Theorem 1.1. *Let $\mathcal{R}$ be 2^d points forming the vertices of a fixed d-dimensional rectangle in $\mathbb{R}^{2d}$.*

(i) *If $S \subseteq \mathbb{R}^{2d}$ has positive upper Banach density, then there exists a threshold $\lambda_0 = \lambda_0(S, \mathcal{R})$ such that S contains an isometric copy of $\lambda\mathcal{R}$ for all $\lambda \geq \lambda_0$.*

(ii) *For any $0 < \delta \leq 1$ there exists a constant $c = c(\delta, \mathcal{R}) > 0$ such that any $S \subseteq [0,1]^{2d}$ with $|S| \geq \delta$ is guaranteed to contain an isometric copy of $\lambda\mathcal{R}$ for all λ in some interval of length at least c.*

Moreover, if $\mathcal{R}$ has sidelengths given by $t_1, \ldots, t_d$, then the isometric copies of $\lambda\mathcal{R}$ in both (i) *and* (ii) *above can all be realized in the special form $\{x_{11}, x_{12}\} \times \cdots \times \{x_{d1}, x_{d2}\} \subseteq \mathbb{R}^2 \times \cdots \times \mathbb{R}^2$ with each $|x_{j2} - x_{j1}| = \lambda t_j$.*

The multi-dimensional extension of Szemerédi's theorem on arithmetic progressions in sets of positive density due to Furstenberg and Katznelson [5] implies, and is equivalent to the fact, that there are isometric copies of $\lambda\mathcal{R}$ in S for *arbitrarily* large λ, with sides parallel to the coordinate axis. While Theorem 1.1 states that there is an isometric copy of $\lambda\mathcal{R}$ in S for *every* sufficiently large λ, with sides parallel to given 2-dimensional coordinate subspaces.

A weaker version of Theorem 1.1, with $\mathbb{R}^{2d}$ replaced with $\mathbb{R}^{5d}$, was later established by Durcik and Kovač in [4] using an adaptation of arguments of the second author with Cook and Pramanik in [3]. This approach also makes direct use of the full strength of the multi-dimensional Szemerédi theorem and as such leads to quantitatively weaker results.

Our arguments work for more general patterns where d-dimensional rectangles are replaced with direct products of non-degenerate simplices.

Theorem 1.2. *Let $\Delta = \Delta_1 \times \cdots \times \Delta_d \subseteq \mathbb{R}^n$, where $\mathbb{R}^n = \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_d}$ and each $\Delta_j \subseteq \mathbb{R}^{n_j}$ is a non-degenerate simplex of n_j points.*

(i) *If $S \subseteq \mathbb{R}^n$ has positive upper Banach density, then there exists a threshold $\lambda_0 = \lambda_0(S, \Delta)$ such that S contains an isometric copy of $\lambda\Delta$ for all $\lambda \geq \lambda_0$.*

(ii) *For any $0 < \delta \leq 1$ there exists a constant $c = c(\delta, \Delta) > 0$ such that any $S \subseteq [0,1]^n$ with $|S| \geq \delta$ is guaranteed to contain an isometric copy of $\lambda\Delta$ for all λ in some interval of length at least c.*

Moreover the isometric copies of $\lambda\Delta$ in both (i) *and* (ii) *above can all be realized in the special form $\Delta_1' \times \cdots \times \Delta_d'$ with each $\Delta_j' \subseteq \mathbb{R}^{n_j}$ an isometric copy of $\lambda\Delta_j$.*

Quantitative Remark. A careful analysis of our proof reveals that the constant $c(\delta, \Delta)$ can be taken greater than $W_d(C_\Delta'\delta^{-3n_1\cdots n_d})^{-1}$ where $W_k(m)$ is a tower of exponentials defined by $W_1(m) = \exp(m)$ and $W_{k+1}(m) = \exp(W_k(m))$ for $k \geq 1$.

1.3. **Existing Results II: Distances and Simplices in Subsets of $\mathbb{Z}^n$.** The problem of counting isometric copies of a given non-degenerate simplex in $\mathbb{Z}^n$ (with one vertex fixed) has been extensively studied via its equivalent formulation as the number of ways a quadratic form can be represented as a sum of squares of linear forms, see [11] and [19]. This was exploited by the second author in [16] and [17] to establish analogous results to those described in Section 1.1 above for subsets of the integer lattice $\mathbb{Z}^n$ of positive upper density.

Recall that the *upper Banach density* of a set $S \subseteq \mathbb{Z}^n$ is analogously defined by

$$\delta^*(S) = \lim_{N\to\infty} \sup_{t\in\mathbb{R}^n} \frac{|S \cap (t + Q(N))|}{|Q(N)|}, \tag{1.2}$$

where $|\cdot|$ now denotes counting measure on $\mathbb{Z}^n$ and $Q(N)$ the discrete cube $[-N/2, N/2]^n \cap \mathbb{Z}^n$.

In light of the fact that any pairs of distinct points $\{x_1, x_2\}$ in $\mathbb{Z}^n$ has the property that the square of the distance between them $|x_2 - x_1|^2$ is always a positive integer we introduce the convenient notation

$$\sqrt{\mathbb{N}} := \{\lambda : \lambda > 0 \text{ and } \lambda^2 \in \mathbb{Z}\}.$$

Theorem A′ (Magyar [16])**.** *Let $0 < \delta \leq 1$.*

If $S \subseteq \mathbb{Z}^5$ has upper Banach density at least δ, then there exists an integer $q_0 = q_0(\delta)$ and $\lambda_0 = \lambda_0(S)$ such that S contains pairs of points $\{x_1, x_2\}$ with $|x_2 - x_1| = q_0\lambda$ for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.

Theorem B′ (Magyar [17])**.** *Let $0 < \delta \leq 1$ and $\Delta \subseteq \mathbb{Z}^{2n+3}$ be a non-degenerate simplex of n points.*

(i) *If $S \subseteq \mathbb{Z}^{2n+3}$ has upper Banach density at least δ, then there exists an integer $q_0 = O(\exp(C_\Delta\delta^{-13n}))$ and $\lambda_0 = \lambda_0(S, \Delta)$ such that S contains an isometric copy of $q_0\lambda\Delta$ for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.*

(ii) *If $N \geq \exp(2C_\Delta\delta^{-13n})$, then any $S \subseteq \{1, \ldots, N\}^{2n+3}$ with cardinality $|S| \geq \delta N^{2n+3}$ will necessarily contain an isometric copy of $\lambda\Delta$ for some $\lambda \in \sqrt{\mathbb{N}}$ with $1 \leq \lambda \leq N$.*

Note that the fact that $S \subseteq \mathbb{Z}^n$ could fall entirely into a fixed congruence class of some integer $1 \leq q \leq \delta^{-1/n}$ ensures that the q_0 that appears in Theorems A′ and B′ above must be divisible by the least common multiple of all integers $1 \leq q \leq \delta^{-1/n}$. Indeed if $S = (q\mathbb{Z})^n$ with $1 \leq q \leq \delta^{-1/n}$ then S has upper Banach density at least δ, however the distance between any two points $x, y \in S$ is of the form $|x - y| = q\lambda$ for some $\lambda \in \sqrt{\mathbb{N}}$.

However, in both Theorems A′ and Part (i) of Theorem B′, one can take $q_0 = 1$ if the sets S are assumed to be suitably uniformly distributed on congruence classes of small modulus. This leads via an easy density increment strategy to short new proofs, see [14] for Theorem A′ and Section 8 for Part (i) of Theorem B′.

The original argument in [17] deduced Theorem B′ from the following discrete analogue of Proposition B.

Proposition B′ (Magyar [17]). *Let $\Delta \subseteq \mathbb{Z}^{2n+3}$ be a non-degenerate simplex of n points.*

For any $0 < \delta \leq 1$ there exist constants $J = O_\Delta(\delta^{-3n})$ and $q_0 = O(\exp(C_\Delta \delta^{-13n}))$ such that if $N \geq \lambda_1 \geq \cdots \geq \lambda_J \geq 1$ is any lacunary sequence in $q_0\sqrt{\mathbb{N}}$ and $S \subseteq \{1,\dots,N\}^{2n+3}$ with cardinality $|S| \geq \delta N^{2n+3}$, then S will necessarily contain an isometric copy of $\lambda_j\Delta$ for some $1 \leq j \leq J$.

To see that Proposition B′ implies Theorem B′ notice that if Part (i) of Theorem B′ were not to hold for some set $S \subseteq \mathbb{Z}^{2n+3}$ of upper Banach density $\delta^*(S) > \delta > 0$ with q_0 from Proposition B′, then there must exist a lacunary sequence $\lambda_1 \geq \cdots \geq \lambda_J \geq 1$ in $q_0\sqrt{\mathbb{N}}$, with J the constant from Proposition B′, such that S does not contain an isometric copy of $\lambda_j\Delta$ for any $1 \leq j \leq J$. Since we can find a sufficiently large cube Q with integer side length N that is divisible by q_0 and greater than λ_1 such that $|S \cap Q| \geq \delta|Q|$, this contradicts Proposition B′. Part (ii) of Theorem B′ follows from Proposition B′ by taking $\lambda_j = 2^{J-j}q_0$.

1.4. **New Results II: Rectangles and Products of Simplices in Subsets of $\mathbb{Z}^n$.**

We will also establish the following discrete analogues of Theorem 1.1 and 1.2.

Theorem 1.3. *Let $0 < \delta \leq 1$ and $\mathcal{R}$ be 2^d points forming the vertices of a d-dimensional rectangle in $\mathbb{Z}^{5d}$.*

(i) *If $S \subseteq \mathbb{Z}^{5d}$ has upper Banach density at least δ, then there exist integers $q_0 = q_0(\delta, \mathcal{R})$ and $\lambda_0 = \lambda_0(S, \mathcal{R})$ such that S contains an isometric copy of $q_0\lambda\mathcal{R}$ for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.*

(ii) *There exists a constant $N(\delta, \mathcal{R})$ such that if $N \geq N(\delta, \mathcal{R})$, then any $S \subseteq \{1,\dots,N\}^{5d}$ with cardinality $|S| \geq \delta N^{5d}$ will necessarily contain an isometric copy of $\lambda\mathcal{R}$ for some $\lambda \in \sqrt{\mathbb{N}}$ with $1 \leq \lambda \leq N$.*

If $\mathcal{R}$ has side lengths given by $t_1,\dots,t_d$, then each of the isometric copies in (i) *and* (ii) *above can be realized in the form $\{x_{11}, x_{12}\} \times \cdots \times \{x_{d1}, x_{d2}\} \subseteq \mathbb{Z}^5 \times \cdots \times \mathbb{Z}^5$ with each $|x_{j2} - x_{j1}| = q_0\lambda t_j$ and λt_j, respectively.*

Our arguments again work for more general patterns where d-dimensional rectangles are replaced with direct products of non-degenerate simplices.

Theorem 1.4. *Let $0 < \delta \leq 1$ and $\Delta = \Delta_1 \times \cdots \times \Delta_d \subseteq \mathbb{Z}^n$, where $\mathbb{Z}^n = \mathbb{Z}^{2n_1+3} \times \cdots \times \mathbb{Z}^{2n_d+3}$ and each $\Delta_i \subseteq \mathbb{Z}^{2n_i+3}$ is a non-degenerate simplex of n_i points.*

(i) *If $S \subseteq \mathbb{Z}^n$ has upper Banach density at least δ, then there exist integers $q_0 = q_0(\delta, \Delta)$ and $\lambda_0 = \lambda_0(S, \Delta)$ such that S contains an isometric copy of $q_0\lambda\Delta$ for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.*

(ii) *There exists a constant $N(\delta, \Delta)$ such that if $N \geq N(\delta, \Delta)$, then any $S \subseteq \{1,\dots,N\}^n$ with cardinality $|S| \geq \delta N^n$ will necessarily contain an isometric copy of $\lambda\Delta$ for some $\lambda \in \sqrt{\mathbb{N}}$ with $1 \leq \lambda \leq N$.*

Moreover, each of the isometric copies in (i) *and* (ii) *above can be realized in the special form $\Delta_1' \times \cdots \times \Delta_d'$ with each $\Delta_i' \subseteq \mathbb{Z}^{2n_i+3}$ an isometric copy of $q_0\lambda\Delta_j$ and $\lambda\Delta_j$, respectively.*

Quantitative Remark. A careful analysis of our proof reveals that the constant $q_0(\delta, \Delta)$ (and consequently also $N(\delta, \Delta)$) can be taken less than $W_d(C'_\Delta \delta^{-13n_1\cdots n_d})$ where $W_k(m)$ is a tower of exponentials defined by $W_1(m) = \exp(m)$ and $W_{k+1}(m) = \exp(W_k(m))$ for $k \geq 1$.

1.5. **Notations and Outline.** We will consider the parameters $d, n_1, \dots, n_d$ fixed and will not indicate the dependence on them. Thus we will write $f = O(g)$ if $|f| \leq C(n_1, \dots, n_d)g$. If the implicit constants in our estimates depend on additional parameters $\varepsilon, \delta, K, \dots$ the we will write $f = O_{\varepsilon,\delta,K,\dots}(g)$. We will use the notation $f \ll g$ to indicate that $|f| \leq c\,g$ for some constant $c > 0$ sufficiently small for our purposes.

Given an $\varepsilon > 0$ and a (finite or infinite) sequence $L_0 \geq L_1 \geq \cdots > 0$, we will say that the sequence is *ε-admissible* if $L_j/L_{j+1} \in \mathbb{N}$ and $L_{j+1} \ll \varepsilon^2 L_j$ for all $j \geq 1$. Moreover, if $q \in \mathbb{N}$ is given and $L_j \in \mathbb{N}$ for all $1 \leq j \leq J$, then we will call the sequence $L_0 \geq L_1 \geq \cdots \geq L_J$ *(ε, q)-admissible* if in addition $L_J/q \in \mathbb{N}$. Such sequences of scales will often appear in our statements both in the continuous and the discrete case.

Our proofs are based on a weak hypergraph regularity lemma and an associated counting lemma developed in the context of Euclidean spaces and the integer lattice. In Section 2 we introduce our approach in the model case of finite fields and prove an analogue of Theorem 1.1 in this setting. In Section 3 we review Theorem 1.2 for a single simplex and ultimately establish the base case of our general inductive approach to Theorem 1.2. In Section 4 we address Theorem 1.2 for the direct product of two simplices, this provides a new proof (and strengthening) of the main result of [12] and serves as a gentle preparation for the more

complicated general case which we present in the Section 5. The proof of Theorem 1.4 is outlined in Sections 6 and 7, while a short direct proof of Part (i) of Theorem B′ is presented in Section 8.

2. Model case: vector spaces over finite fields.

In this section we will illustrate our general method by giving a complete proof of Theorem 1.1 in the model setting of $\mathbb{F}_q^n$ where $\mathbb{F}_q$ denotes the finite field of q elements. We do this as the notation and arguments are more transparent in this setting yet many of the main ideas are still present.

We say that two vectors $u, v \in \mathbb{F}_q^n$ are orthogonal, if $x \cdot y = 0$, where "$\cdot$" stands for the usual dot product. A rectangle in $\mathbb{F}_q^n$ is then a set $\mathcal{R} = \{x_1, y_1\} \times \cdots \times \{x_n, y_n\}$ with side vectors $y_i - x_i$ being pairwise orthogonal.

The finite field analogue of Theorem 1.1 is the following

Proposition 2.1. *For any $0 < \delta \leq 1$ there exists an integer $q_0 = q_0(\delta)$ with the following property:*

If $q \geq q_0$ and $t_1, \ldots, t_d \in \mathbb{F}_q^*$, then any $S \subseteq \mathbb{F}_q^{2d}$ with $|S| \geq \delta\, q^{2d}$ will contain points

$$\{x_{11}, x_{12}\} \times \cdots \times \{x_{d1}, x_{d2}\} \subseteq V_1 \times \cdots \times V_d \quad \text{with} \quad |x_{j2} - x_{j1}|^2 = t_j \text{ for } 1 \leq j \leq d$$

where we have written $\mathbb{F}_q^{2d} = V_1 \times \cdots \times V_d$ with $V_j \simeq \mathbb{F}_q^2$ pairwise orthogonal coordinate subspaces.

2.1. **Overview of the proof of Proposition 2.1.** Write $\mathbb{F}_q^{2d} = V_1 \times \ldots \times V_d$ with $V_j \simeq \mathbb{F}_q^2$ pairwise orthogonal coordinate subspaces. For any $\underline{t} := (t_1, \ldots, t_d) \in \mathbb{F}_q^*$ and $S \subseteq \mathbb{F}_q^{2d}$ we define

$$\mathcal{N}_{\underline{t}}(1_S) := \mathbb{E}_{\underline{x}_1 \in V_1^2, \ldots, \underline{x}_d \in V_d^2} \prod_{(\ell_1, \ldots, \ell_d) \in \{1,2\}^d} 1_S(x_{1\ell_1}, \ldots, x_{d\ell_d}) \prod_{j=1}^{d} \sigma_{t_j}(x_{j2} - x_{j1})$$

where we used the shorthand notation $\underline{x}_j := (x_{j1}, x_{j2})$ for each $1 \leq j \leq d$ and the averaging notation:

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

for a finite set $A \neq \emptyset$. We have also used the notation

$$\sigma_t(x) = \begin{cases} q & \text{if } |x|^2 = t \\ 0 & \text{otherwise} \end{cases}$$

for each $t \in \mathbb{F}_q^*$. Note that the function σ_t may be viewed as the discrete analogue of the normalized surface area measure on the sphere of radius $\sqrt{t}$. It is well-known, see [10], that

$$\mathbb{E}_{x \in \mathbb{F}_q^2}\, \sigma_t(x) = 1 + O(q^{-1/2})$$

and for all $\xi \neq 0$ one has

$$\hat{\sigma}_t(\xi) := \mathbb{E}_{x \in \mathbb{F}_q^2}\, \sigma_t(x)\, e^{2\pi i \frac{x \cdot \xi}{q}} = O(q^{-1/2}).$$

Note that if $\mathcal{N}_{\underline{t}}(1_S) > 0$, then this implies that S contains a rectangle of the form $\{x_{11}, x_{12}\} \times \cdots \times \{x_{d1}, x_{d2}\}$ with $x_{j1}, x_{j2} \in V_j$ and $|x_{j2} - x_{j1}|^2 = t_j$ for $1 \leq j \leq d$.

Our approach to Proposition 2.1 in fact establishes the following quantitatively stronger result.

Proposition 2.2. *For any $0 < \varepsilon \leq 1$ there exists an integer $q_0 = q_0(\varepsilon)$ with the following property:*

If $q \geq q_0$, then for any $S \subseteq \mathbb{F}_q^{2d}$ and $t_1, \ldots, t_d \in \mathbb{F}_q^*$ one has

$$\mathcal{N}_{\underline{t}}(1_S) > \left(\frac{|S|}{q^{2d}}\right)^{2^d} - \varepsilon$$

where we have written $\mathbb{F}_q^{2d} = V_1 \times \ldots \times V_d$ with $V_j \simeq \mathbb{F}_q^2$ pairwise orthogonal coordinate subspaces.

A crucial observation in the proof of Proposition 2.2 is that the averages $\mathcal{N}_{\underline{t}}(1_S)$ can be compared to ones which can be easily estimated from below. We define, for any $S \subseteq \mathbb{F}_q^{2d}$, the (unrestricted) count

$$\mathcal{M}(1_S) := \mathbb{E}_{\underline{x}_1 \in V_1^2, \ldots, \underline{x}_d \in V_d^2} \prod_{(\ell_1, \ldots, \ell_d) \in \{1,2\}^d} 1_S(x_{1\ell_1}, \ldots, x_{d\ell_d}).$$

It is easy to see, by carefully applying Cauchy-Schwarz d times to $\mathbb{E}_{x_{11}\in V_1,\ldots,x_{d1}\in V_d} 1_S(x_{11},\ldots,x_{d1})$, that

$$\mathcal{M}(1_S) \geq \left(\frac{|S|}{q^{2d}}\right)^{2^d}. \tag{2.1}$$

Our approach to Proposition 2.2 therefore reduces to establishing that for any $\varepsilon > 0$ one has

$$\mathcal{N}_{\underline{t}}(1_S) \;=\; \mathcal{M}(1_S) + O(\varepsilon) + O_\varepsilon(q^{-1/2}). \tag{2.2}$$

The validity of (2.2) will follow immediately from the $d = k$ case of Proposition 2.3 below. However, before we can state this *counting lemma* we need to introduce some further notation from the theory of hypergraphs, notation that we shall ultimately make use of throughout the paper.

2.2. **Hypergraph Notation and a Counting Lemma.**

In order to streamline our notation we will make use the language of hypergraphs. For $J := \{1,\ldots,d\}$ and $1 \leq k \leq d$, we let $\mathcal{H}_{d,k} = \{e \subseteq J;\ |e| = k\}$ denote the full k-regular hypergraph on the vertex set J. For $K := \{jl;\ j \in J,\, l \in \{1,2\}\}$ we define the projection $\pi : K \to J$ as $\pi(jl) := j$ and use this in turn to define the hypergraph bundle

$$\mathcal{H}^{\underline{2}}_{d,k} := \{e \subseteq K;\ |e| = |\pi(e)| = k\}$$

using the shorthand notation $\underline{2} = (2,2,\ldots,2)$ to indicate that $|\pi^{-1}(j)| = 2$ for all $j \in J$.

Notice when $k = d$ then $\mathcal{H}_{d,d}$ consists of one element, the set $e = \{1,\ldots,d\}$, and

$$\mathcal{H}^{\underline{2}}_{d,d} = \{\,\{1l_1,\ldots,dl_d\};\ (l_1,\ldots,l_d) \in \{1,2\}^d\}.$$

Let $V := \mathbb{F}_q^{2d}$ and $V = V_1 \times \ldots \times V_d$ with $V_j \simeq \mathbb{F}_q^2$ pairwise orthogonal coordinate subspaces. For a given $\underline{x} = (x_{11}, x_{12}, \ldots, x_{d1}, x_{d2}) \in V^2$ with $x_{j1},\, x_{j2} \in V_j$ and a given edge $e = \{1l_1,\ldots,dl_d\}$, we write

$$\underline{x}_e := (x_{1l_1},\ldots,x_{dl_d}).$$

Note that the map $\underline{x} \to \underline{x}_e$ defines a projection $\pi_e : V^2 \to V$. With this notation, we can clearly now write

$$\mathcal{N}_{\underline{t}}(1_S) = \mathbb{E}_{\underline{x}\in V^2} \prod_{e\in\mathcal{H}^{\underline{2}}_{d,d}} 1_S(\underline{x}_e) \prod_{j=1}^{d} \sigma_{t_j}(x_{j2} - x_{j1})$$

$$\mathcal{M}(1_S) = \mathbb{E}_{\underline{x}\in V^2} \prod_{e\in\mathcal{H}^{\underline{2}}_{d,d}} 1_S(\underline{x}_e).$$

Now for any $1 \leq k \leq d$ and any edge $e' \in \mathcal{H}_{d,k}$, i.e. $e' \subseteq \{1,\ldots,d\}$, $|e'| = k$, we let $V_{e'} := \prod_{j\in e'} V_j$. For every $\underline{x} \in V^2$ and $e \in \mathcal{H}^{\underline{2}}_{d,k}$, we define $\underline{x}_e := \pi_e(\underline{x})$ where $\pi_e : V^2 \to V_{\pi(e)}$ is the natural projection map.

Our key *counting lemma*, Proposition 2.3 below, which we will establish by induction on $1 \leq k \leq d$ below, is then the statement that given a family of functions $f_e : V_{\pi(e)} \to [-1,1]$, $e \in \mathcal{H}^{\underline{2}}_{d,k}$, the averages (generalizing those discussed above) which are defined by

$$\mathcal{N}_{\underline{t}}(f_e;\ e \in \mathcal{H}^{\underline{2}}_{d,k}) := \mathbb{E}_{\underline{x}\in V^2} \prod_{e\in\mathcal{H}^{\underline{2}}_{d,k}} f_e(\underline{x}_e) \prod_{j=1}^{d} \sigma_{t_j}(x_{j2} - x_{j1}) \tag{2.3}$$

$$\mathcal{M}(f_e;\, e \in \mathcal{H}^{\underline{2}}_{d,k}) := \mathbb{E}_{\underline{x}\in V^2} \prod_{e\in\mathcal{H}^{\underline{2}}_{d,k}} f_e(\underline{x}_e). \tag{2.4}$$

are approximately equal. Specifically, one has

Proposition 2.3 (Counting Lemma)**.** *Let $1 \leq k \leq d$ and $0 < \varepsilon \leq 1$. For any collection of functions*

$$f_e : V_{\pi(e)} \to [-1,1] \text{ with } e \in \mathcal{H}^{\underline{2}}_{d,k}$$

one has

$$\mathcal{N}_{\underline{t}}(f_e;\ e \in \mathcal{H}^{\underline{2}}_{d,k}) \;=\; \mathcal{M}(f_e;\, e \in \mathcal{H}^{\underline{2}}_{d,k}) + O(\varepsilon) + O_\varepsilon(q^{-1/2}). \tag{2.5}$$

If we apply this Proposition with $d = k$ and $f_e = 1_S$ for all $e \in \mathcal{H}^2_{d,d}$, then Theorem 2.1 clearly follows given the lower bound (2.1).

2.3. **Proof of Proposition 2.3.** We will establish Proposition 2.3 by inducting on $1 \leq k \leq d$.

For $k = 1$ the result follows from the basic observation that if $f_1, f_2 : \mathbb{F}_q^2 \to [-1, 1]$ and let $t \in \mathbb{F}_q^*$, then

$$\begin{aligned} \mathbb{E}_{x_1,x_2\in\mathbb{F}_q^2}\ f_1(x_1)f_2(x_2)\,\sigma_t(x_2-x_1) &= \sum_{\xi\in\mathbb{F}_q^2} \hat{f}_1(\xi)\hat{f}_2(\xi)\hat{\sigma}_t(\xi) \\ &= \hat{f}_1(0)\hat{f}_2(0) + O(q^{-1/2}) \\ &= \mathbb{E}_{x_1,x_2\in\mathbb{F}_q^2}\ f_1(x_1)f_2(x_2) + O(q^{-1/2}) \end{aligned} \tag{2.6}$$

by the properties of the function $\hat{\sigma}$ given above.

To see how this implies Proposition 2.3 for $k = 1$ we note that since $\mathcal{H}^2_{d,1} = \{jl :\ 1 \leq j \leq d,\ 1 \leq l \leq 2\}$ it follows that

$$\begin{aligned} \mathcal{N}_{\underline{t}}(f_e;\ e \in \mathcal{H}^2_{d,1}) &= \prod_{j=1}^{d} \mathbb{E}_{x_{j1},x_{j2}\in\mathbb{F}_q^2}\ f_{j1}(x_{j1})f_{j2}(x_{j2})\,\sigma_t(x_{j2}-x_{j1}) \\ &= \prod_{j=1}^{d} \mathbb{E}_{x_{j1},x_{j2}\in\mathbb{F}_q^2}\ f_{j1}(x_{j1})f_{j2}(x_{j2}) + O(q^{-1/2}) = \mathcal{M}(f_e;\ e \in \mathcal{H}^2_{d,1}) + O(q^{-1/2}). \end{aligned}$$

The induction step has two main ingredients, the first is an estimate of the type which is often referred to as a generalized von-Neumann inequality, namely

Lemma 2.1. *Let $1 \leq k \leq d$. For any collection of functions $f_e : V_{\pi(e)} \to [-1, 1]$ with $e \in \mathcal{H}^2_{d,k}$ one has*

$$\mathcal{N}_{\underline{t}}(f_e;\ e \in \mathcal{H}^2_{d,k}) \leq \min_{e\in\mathcal{H}^2_{d,k}} \|f_e\|_{\square(V_{\pi(e)})} + O(q^{-1/2}) \tag{2.7}$$

where for any $e \in \mathcal{H}^2_{d,k}$ and $f : V_{\pi(e)} \to [-1, 1]$ we define

$$\|f\|^{2^k}_{\square(V_{\pi(e)})} := \mathbb{E}_{\underline{x}\in V^2_{\pi(e)}} \prod_{e\in\mathcal{H}^2_{d,k}} f(\underline{x}_e). \tag{2.8}$$

The corresponding inequality for the multilinear expression $\mathcal{M}(f_e;\ e \in \mathcal{H}^2_{d,k})$, namely the fact that

$$\mathcal{M}(f_e;\ e \in \mathcal{H}^2_{d,k}) \leq \prod_{e\in\mathcal{H}^2_{d,k}} \|f_e\|_{\square(V_{\pi(e)})} \leq \min_{e\in\mathcal{H}^2_{d,k}} \|f_e\|_{\square(V_{\pi(e)})}$$

is well-known and is referred to as the Gowers-Cauchy-Schwarz inequality [8].

The second and main ingredient is an approximate decomposition of a graph to simpler ones, and is essentially the so-called weak (hypergraph) regularity lemma of Frieze and Kannan [7].

We will first introduce this in the case $d = 2$. A bipartite graph with (finite) vertex sets V_1, V_2 is a set $S \subseteq V_1 \times V_2$ and a function $f : V_1 \times V_2 \to \mathbb{R}$ may be viewed as weighted bipartite graph with weights $f(x_1, x_2)$ on the edges (x_1, x_2). If $\mathcal{P}_1$ and $\mathcal{P}_2$ are partitions of V_1 and V_2 respectively then $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2$ is a partition $V_1 \times V_2$ and we let $\mathbb{E}(f|\mathcal{P})$ denote the function that is constant and equal to $\mathbb{E}_{x\in A} f(x)$ on each atom $A = A_1 \times A_2$ of $\mathcal{P}$. The weak regularity lemma states that for any $\varepsilon > 0$ and for any weighted graph $f : V_1 \times V_2 \to [-1, 1]$ there exist partitions $\mathcal{P}_i$ of V_i with $|\mathcal{P}_i| \leq 2^{O(\varepsilon^{-2})}$ for $i = 1, 2$, so that

$$|\mathbb{E}_{x_1\in V_1}\mathbb{E}_{x_2\in V_2}(f - \mathbb{E}(f|\mathcal{P}))(x_1, x_2)\ \mathbf{1}_{U_1}(x_1)\mathbf{1}_{U_2}(x_2)| \leq \varepsilon \tag{2.9}$$

for all $U_1 \subseteq V_1$ and $U_2 \subseteq V_2$. Informally this means that the graph f can be approximated with precision ε with the "low complexity" graph $\mathbb{E}(f, \mathcal{P})$. If we consider the σ-algebras $\mathcal{B}_i$ generated by the partitions $\mathcal{P}_i$ and the σ-algebra $\mathcal{B} = \mathcal{B}_1 \vee \mathcal{B}_2$ generated by $\mathcal{P}_1 \times \mathcal{P}_2$ then we have $\mathbb{E}(f|\mathcal{B})$, the so-called conditional expectation function of f. Moreover it is easy to see, using Cauchy-Schwarz, that estimate (2.9) follows from

$$\|f - \mathbb{E}(f|\mathcal{B}_1 \vee \mathcal{B}_2)\|_{\square(V_1\times V_2)} \leq \varepsilon. \tag{2.10}$$

With this more probabilistic point of view the weak regularity lemma says that the function f can be approximated with precision ε by a low complexity function $\mathbb{E}(f|\mathcal{B}_1 \bigvee \mathcal{B}_2)$, corresponding to σ-algebras $\mathcal{B}_i$ on V_i generated by $O(\varepsilon^{-2})$ sets. This formulation is also referred to as a Koopman- von Neumann type decomposition, see Corollary 6.3 in [23].

We will need a natural extension to k-regular hypergraphs. See [22, 8], and also [2] for extension to sparse hypergraphs. Given an edge $e' \in \mathcal{H}_{d,k}$ of k elements we define its boundary $\partial e' := \{\mathfrak{f}' \in \mathcal{H}_{d,k-1};\ \mathfrak{f}' \subseteq e'\}$. For each $\mathfrak{f}' = e'\backslash\{j\} \in \partial e'$ let $\mathcal{B}'_{\mathfrak{f}}$ be a σ-algebra on $V_{\mathfrak{f}'} := \prod_{j\in\mathfrak{f}'} V_j$ and $\bar{\mathcal{B}}_{\mathfrak{f}'} := \{U \times V_j;\ U \in \mathcal{B}_{\mathfrak{f}'}\}$ denote its pull-back over the space $V_{e'}$. The σ-algebra $\mathcal{B} = \bigvee_{\mathfrak{f}'\in\partial e'} \mathcal{B}_{\mathfrak{f}'}$ is the smallest σ-algebra on $\partial e'$ containing $\bar{\mathcal{B}}_{\mathfrak{f}'}$ for all $\mathfrak{f}' \in \partial e'$. Note that the atoms of $\mathcal{B}$ are of the form $A = \bigcap_{\mathfrak{f}'\in\partial e'} A_{\mathfrak{f}'}$ where $A_{\mathfrak{f}'}$ is an atom of $\bar{\mathcal{B}}_{\mathfrak{f}'}$. We say that the complexity of a σ-algebra $\mathcal{B}_{\mathfrak{f}'}$ is at most m, and write $\mathrm{complex}(\mathcal{B}_{\mathfrak{f}'}) \le m$, if it is generated by m sets.

Lemma 2.2 (Weak hypergraph regularity lemma)**.** *Let $1 \le k \le d$ and $f_e : V_{\pi(e)} \to [-1,1]$ be a given function for each $e \in \mathcal{H}^2_{d,k}$. For any $\varepsilon > 0$ there exists σ-algebras $\mathcal{B}_{\mathfrak{f}'}$ on $V_{\mathfrak{f}'}$ for each $\mathfrak{f}' \in \mathcal{H}_{d,k-1}$ such that*

$$\mathrm{complex}(\mathcal{B}_{\mathfrak{f}'}) = O(\varepsilon^{-2^{k+1}}) \tag{2.11}$$

and

$$\|f_e - \mathbb{E}(f_e|\bigvee_{\mathfrak{f}'\in\partial\pi(e)} \mathcal{B}_{\mathfrak{f}'})\|_{\square(V_{\pi(e)})} \le \varepsilon \quad \text{for all } e \in \mathcal{H}^2_{d,k}. \tag{2.12}$$

The proof of Lemmas 2.1 and 2.2 are presented in Section 2.4 below. We close this subsection by demonstrating how these lemmas can be combined to establish Proposition 2.3.

Proof of Proposition 2.3.

Let $\varepsilon > 0$, $2 \le k \le d$ and assume that the lemma holds for $k-1$. It follows from Lemma 2.2 that there exists σ-algebras $\mathcal{B}_{\mathfrak{f}'}$ of complexity $O(\varepsilon^{-2^{k+1}})$ on $V_{\mathfrak{f}'}$ for each $\mathfrak{f}' \in \mathcal{H}_{d,k-1}$ for which (2.12) holds for all $e \in \mathcal{H}^2_{d,k}$. For each $e \in \mathcal{H}^2_{d,k}$ we let $\bar{f}_e := \mathbb{E}(f_e|\bigvee_{\mathfrak{f}'\in\partial\pi(e)} \mathcal{B}_{\mathfrak{f}'})$ and write $f_e = \bar{f}_e + h_e$. By Lemma 2.1 and multi-linearity we have that

$$\mathcal{N}_{\underline{t}}(f_e;\ e \in \mathcal{H}^2_{d,k}) = \mathcal{N}_{\underline{t}}(\bar{f}_e;\ e \in \mathcal{H}^2_{d,k}) + O(\varepsilon) + O(q^{-1/2}) \tag{2.13}$$

and also by the Gowers-Cauchy-Schwarz inequality

$$\mathcal{M}(f_e;\ e \in \mathcal{H}^2_{d,k}) = \mathcal{M}(\bar{f}_e;\ e \in \mathcal{H}^2_{d,k}) + O(\varepsilon). \tag{2.14}$$

The conditional expectation functions $\bar{f}_e$ are linear combinations of the indicator functions 1_{A_e} of the atoms A_e of the σ-algebras $\mathcal{B}_e := \bigvee_{\mathfrak{f}'\in\partial\pi(e)} \mathcal{B}_{\mathfrak{f}'}$. Since the number of terms in this linear combination is at most $2^{C\varepsilon^{-2^{k+1}}}$, with coefficients at most 1 in modulus, plugging these into the multi-linear expressions $\mathcal{N}_{\underline{t}}(\bar{f}_e;\ e \in \mathcal{H}^2_{d,k})$ and $\mathcal{M}(\bar{f}_e;\ e \in \mathcal{H}^2_{d,k})$ one obtains a linear combination of expressions of the form $\mathcal{N}_{\underline{t}}(1_{A_e};\ e \in \mathcal{H}^2_{d,k})$ and $\mathcal{M}(1_{A_e};\ e \in \mathcal{H}^2_{d,k})$ respectively with each A_e being an atoms of $\mathcal{B}_e$ for all $e \in \mathcal{H}^2_{d,k}$.

The key observation is that these expressions are at level $k-1$ instead of k. Indeed, $1_{A_e} = \prod_{\mathfrak{f}'\in\partial\pi(e)} 1_{A_{e\mathfrak{f}'}}$ where $A_{e\mathfrak{f}'} = A'_{e\mathfrak{f}'} \times V_j$, with $A'_{e\mathfrak{f}'}$ being an atom of $\mathcal{B}_{\mathfrak{f}'}$ when $\mathfrak{f}' = \pi(e)\backslash\{j\}$. If $e = (j_1 l_1, \dots, jl, \dots, j_k l_k)$, let $p_{\mathfrak{f}'}(e) := (j_1 l_1, \dots, j_k l_k) \in \mathcal{H}^2_{d,k-1}$, obtained from e by removing the jl-entry. Then we have $1_{A_{e\mathfrak{f}'}}(\underline{x}_e) = 1_{A'_{e\mathfrak{f}'}}(\underline{x}_{p'_{\mathfrak{f}}(e)})$ since $x_{jl} \in V_j$, and hence

$$1_{A_e}(\underline{x}_e) = \prod_{\mathfrak{f}'\in\partial\pi(e)} 1_{A'_{e\mathfrak{f}'}}(\underline{x}_{p'_{\mathfrak{f}}(e)}).$$

It therefore follows that

$$\mathcal{N}_{\underline{t}}(1_{A_e};\ e\in\mathcal{H}^2_{d,k}) = \mathbb{E}_{\underline{x}\in V^2}\prod_{e\in\mathcal{H}^2_{d,k}}\prod_{\mathfrak{f}'\in\partial\pi(e)} 1_{A'_{e\mathfrak{f}'}}(\underline{x}_{p_{\mathfrak{f}'}(e)})\prod_{j=1}^{d}\sigma_{t_j}(x_{j2}-x_{j1})$$

$$= \mathbb{E}_{\underline{x}\in V^2}\prod_{\mathfrak{f}\in\mathcal{H}^2_{d,k-1}}\underbrace{\prod_{\substack{e\in\mathcal{H}^2_{d,k},\,\mathfrak{f}'\in\partial\pi(e)\\ p_{\mathfrak{f}'}(e)=\mathfrak{f}}} 1_{A'_{e\mathfrak{f}'}}(\underline{x}_{p_{\mathfrak{f}'}(e)})}_{=:g_{\mathfrak{f}}}\prod_{j=1}^{d}\sigma_{t_j}(x_{j2}-x_{j1}) = \mathcal{N}_{\underline{t}}(g_{\mathfrak{f}};\ \mathfrak{f}\in\mathcal{H}^2_{d,k-1})$$

and similarly that

$$\mathcal{M}(1_{A_e};\ e\in\mathcal{H}^2_{d,k}) = \mathcal{M}(g_{\mathfrak{f}};\ \mathfrak{f}\in\mathcal{H}^2_{d,k-1}).$$

It then follows from the induction hypotheses that

$$\mathcal{N}_{\underline{t}}(1_{A_e};\ e\in\mathcal{H}^2_{d,k}) = \mathcal{M}(1_{A_e};\ e\in\mathcal{H}^2_{d,k}) + O(\varepsilon_1) + O_{\varepsilon_1}(q^{-1/2})$$

for any $\varepsilon_1>0$. If we choose $\varepsilon_1 := 2^{-C_1\,\varepsilon^{-2^{k+1}}}$, with $C_1\gg 1$ sufficiently large, then $\varepsilon_1\,2^{C\varepsilon^{-2^{k+1}}} = O(\varepsilon)$ and it follows that

$$\mathcal{N}_{\underline{t}}(\bar{f}_e;\ e\in\mathcal{H}^2_{d,k}) = \mathcal{M}(\bar{f}_e;\ e\in\mathcal{H}^2_{d,k}) + O(\varepsilon) + O_{\varepsilon}(q^{-1/2}).$$

This, together with (2.13) and (2.14), establishes that (2.5) hold for $d=k$ as required. □

2.4. **Proof of Lemmas 2.1 and 2.2.**

Proof of Lemma 2.1. We start by observing the following consequence of (2.6), namely that

$$\left|\mathbb{E}_{x_1,x_2\in\mathbb{F}_q^2} f_1(x_1)f_2(x_2)\sigma_t(x_2-x_1)\right|^2 \le \mathbb{E}_{x_1,x_2\in\mathbb{F}_q^2} f_1(x_1)f_1(x_2) + O(q^{-1/2}) \tag{2.15}$$

for any $f_1,f_2:\mathbb{F}_q^2\to[-1,1]$ and $t\in\mathbb{F}_q^*$.

Now, fix an edge, say $e_0=(11,21,\dots,k1)$. Partition the edges $e\in\mathcal{H}^2_{d,k}$ into three groups; the first group consisting of edges e for which $1\notin\pi(e)$, the second where $11\in e$ and write $e=(11,e')$ with $e'\in\mathcal{H}^2_{d-1,k-1}$ and the third when $12\in e$, using the notation $\mathcal{H}^2_{d-1,k-1} := \{(j_2l_2,\dots,j_kl_k)\}$. Accordingly we can write

$$\mathcal{N}_t(f_e;\ e\in\mathcal{H}^2_{d,k}) = \mathbb{E}_{\underline{x}\in V^2}\prod_{1\notin\pi(e)} f_e(\underline{x}_e)\prod_{e'\in\mathcal{H}^2_{d-1,k-1}} f_{(11,e')}(x_{11},\underline{x}_{e'})\prod_{e'\in\mathcal{H}^2_{d-1,k-1}} f_{(12,e')}(x_{12},\underline{x}_{e'})\prod_{j=1}^{d}\sigma_{t_j}(x_{j2}-x_{j1}). \tag{2.16}$$

If for given $x\in V_1$ and $\underline{x}'=(x_{21},x_{22},\dots,x_{d1},x_{d2})\in V_2^2\times\ldots\times V_d^2$ we define

$$g_1(x,\underline{x}') := \prod_{e'\in\mathcal{H}^2_{d-1,k-1}} f_{(11,e')}(x,\underline{x}_{e'}) \qquad\text{and}\qquad g_2(x,\underline{x}') := \prod_{e'\in\mathcal{H}^2_{d-1,k-1}} f_{(12,e')}(x,\underline{x}_{e'})$$

then we can write

$$\mathcal{N}_t(f_e;\ e\in\mathcal{H}^2_{d,k}) = \mathbb{E}_{x_{21},x_{22},\dots,x_{d1},x_{d2}}\prod_{1\notin\pi(e)} f_e(\underline{x}_e)\prod_{j=2}^{d}\sigma_{t_j}(x_{j2}-x_{j1}) \times \mathbb{E}_{x_{11},x_{12}}\ g_1(x_{11},\underline{x}')g_2(x_{12},\underline{x}')\,\sigma_{t_1}(x_{12}-x_{11}). \tag{2.17}$$

By (2.15) we can estimate the inner sum in (2.17) by the square root of

$$\mathbb{E}_{x_{11},x_{12}}\ g_1(x_{11},\underline{x}')g_1(x_{12},\underline{x}') + O(q^{-1/2}).$$

Thus by Cauchy-Schwarz, and the fact that $f_e:V_{\pi(e)}\to[-1,1]$ for all $e\in\mathcal{H}^2_{d,k}$, we can conclude that

$$\mathcal{N}_t(f_e;\ e\in\mathcal{H}^2_{d,k})^2 \le \mathbb{E}_{x_{11},x_{12},\dots,x_{d1},x_{d2}}\prod_{e'\in\mathcal{H}^2_{d-1,k-1}} f_{(11,e')}(x_{11},\underline{x}_{e'})f_{(11,e')}(x_{12},\underline{x}_{e'})\prod_{j=2}^{d}\sigma_{t_j}(x_{j2}-x_{j2}). \tag{2.18}$$

The expression on the right hand side of the inequality above is similar to that in (2.16) except for the following changes. The functions f_e for $1\notin e$ are eliminated i.e. replaced by 1, as well as the factor σ_{t_1}. The

functions $f_{(12,e')}$, are replaced by $f_{(11,e')}$ for all $e' \in \mathcal{H}^2_{d-1,k-1}$. Repeating the same procedure for $j = 2, \dots, k$ one eliminates all the factors σ_{t_j} for $1 \le j \le k$, moreover all the functions f_e for edges e such that $j \notin \pi(e)$ for some $1 \le j \le k$, which leaves only the edges e so that $\pi(e) = (1, 2, \dots, k)$, moreover for such edges the functions f_e are eventually replaced by $f_{e_0} = f_{11,21,\dots,k1}$. The factors $\sigma_{t_j}(x_{j2} - x_{j1})$ are not changed for $j > k$ however as the function f_{e_0} does not depend on the variables x_{jl} for $j > k$, averaging over these variables gives rise to a factor of $1 + O(q^{-1/2})$. Thus one obtains the following final estimate

$$\mathcal{N}_t(f_e;\ e \in \mathcal{H}^2_{d,k})^{2^k} \le \mathbb{E}_{x_{11},x_{12},\dots,x_{k1},x_{k2}} \prod_{\pi(e)=(1,\dots,k)} f_{e_0}(\underline{x}_e) + O(q^{-1/2}) = \|f_{e_0}\|^{2^k}_{\square(V_{\pi(e_0)})} + O(q^{-1/2}). \tag{2.19}$$

This proves the lemma, as it is clear that the above procedure can be applied to any edge in place of $e_0 = (11, 21, \dots, k1)$. □

Proof of Lemma 2.2. For a function $f_e : V_{\pi(e)} \to [-1, 1]$ and a σ-algebra $\mathcal{B}_{\pi(e)}$ on $V_{\pi(e)}$ define the *energy* of f_e with respect to $\mathcal{B}_{\pi(e)}$ as

$$\mathcal{E}(f_e, \mathcal{B}_{\pi(e)}) := \|\mathbb{E}(f_e|\mathcal{B}_{\pi(e)})\|_2^2 = \mathbb{E}_{x \in V_{\pi(e)}} |\mathbb{E}(f_e|\mathcal{B}_{\pi(e)})(x)|^2,$$

and for a family of functions f_e and σ-algebras $\mathcal{B}_{\pi(e)}$, $e \in \mathcal{H}^2_{d,k}$ its total energy as

$$\mathcal{E}(f_e, \mathcal{B}_{\pi(e)};\ e \in \mathcal{H}^2_{d,k}) := \sum_{e \in \mathcal{H}^2_{d,k}} \mathcal{E}(f_e, \mathcal{B}_{\pi(e)}).$$

We will show that if (2.12) does not hold for a family of σ-algebras $\mathcal{B}_{\pi(e)} = \bigvee_{\mathfrak{f}' \in \partial\pi(e)} \mathcal{B}_{\mathfrak{f}'}$, then the σ-algebras $\mathcal{B}_{\mathfrak{f}'}$ can be refined so that the total energy of the system increases by a quantity depending only on ε. Since the functions f_e are bounded the total energy of the system is $O(1)$, the energy increment process must stop in $O_\varepsilon(1)$ steps, and (2.12) must hold. The idea of this procedure appears already in the proof of Szemerédi's regularity lemma [20], and have been used since in various places [7, 22, 8].

Initially set $\mathcal{B}_{\mathfrak{f}'} := \{\emptyset, V_{\mathfrak{f}'}\}$ and hence $\mathcal{B}_{\pi(e)} = \{\emptyset, V_{\pi(e)}\}$ to be the trivial σ-algebras. Assume that in general (2.12) does not hold for a family of σ-algebras $\mathcal{B}_{\mathfrak{f}'}$, with $\mathfrak{f}' \in \mathcal{H}_{d,k-1}$. Then there exists an edge $e \in \mathcal{H}^2_{d,k}$ so that $\|g_e\|_{\square(V_{\pi(e)})} \ge \varepsilon$, with $g_e := f_e - \mathbb{E}(f_e|\mathcal{B}_{\pi(e)})$. Let $e = (11, \dots, k1)$ for simplicity of notation, hence $\pi(e) = (1, \dots, k)$. Then, with notation $\underline{x}' = (x_{12}, \dots, x_{k2})$, one has

$$\begin{aligned}\varepsilon^{2^k} \le \|g_e\|^{2^k}_{\square(V_{\pi(e)})} &= \mathbb{E}_{x_{11},x_{12},\dots,x_{k1},x_{k2}} \prod_{l_1,\dots,l_k=1,2} g_e(x_{1l_1}, \dots, x_{kl_k}) \\ &\le \mathbb{E}_{x_{12},\dots,x_{k2}} \Big| \mathbb{E}_{x_{11},\dots,x_{k1}} g_e(x_{11}, \dots, x_{k1}) \prod_{j=1}^k h_{j,\underline{x}'}(x_{11}, \dots, x_{j-1\,1}, x_{j+1\,1}, \dots, x_{k1}) \Big|\end{aligned}$$

for some functions $h_{j,\underline{x}'}$ that are bounded by 1 in magnitude. Indeed if and edge $e \ne (11, \dots, k1)$ then x_e does not depend at least one of the variables x_{j1}. Thus there must be an $\underline{x}'$ for which the inner sum in the above expression is at least ε^{2^k}. Fix such an $\underline{x}'$. Decomposing the functions $h_{j,\underline{x}'}$ into their positive and negative parts and then writing them as an average of indicator functions, one obtains that there sets $B_j \subseteq V_{\pi(e)\setminus\{j\}}$ such that

$$\Big| \mathbb{E}_{x_{11},\dots,x_{k1}} g_e(x_{11}, \dots, x_{k1}) \prod_{j=1}^k 1_{B_j}(x_{11}, \dots, x_{j-1\,1}, x_{j+1\,1}, \dots, x_{k1}) \Big| \ \ge\ 2^{-k}\, \varepsilon^{2^k}$$

which can be written more succinctly, using the inner product notation, as

$$\Big| \langle f_e - \mathbb{E}(f_e|\mathcal{B}_{\pi(e)}), \prod_{j=1}^k 1_{B_j} \rangle \Big| \ \ge\ 2^{-k}\, \varepsilon^{2^k}. \tag{2.20}$$

For $\mathfrak{f}' = \partial\pi(e)\setminus\{j\}$ let $\mathcal{B}'_{\mathfrak{f}'}$ be the σ-algebra generated by $\mathcal{B}_{\mathfrak{f}'}$ and the set B_j and let $\mathcal{B}'_{\pi(e)} := \bigvee_{\mathfrak{f}' \in \partial\pi(e)} \mathcal{B}'_{\mathfrak{f}'}$. Since the functions 1_{B_j} are measurable with respect to the σ-algebra $\mathcal{B}'_{\pi(e)}$ for all $1 \le j \le k$, we have that

$$\langle f_e - \mathbb{E}(f_e|\mathcal{B}'_{\pi(e)}), \prod_{j=1}^k 1_{B_j} \rangle = 0 \tag{2.21}$$

and hence, by Cauchy-Schwarz, that

$$\|\mathbb{E}(f_e|\mathcal{B}'_{\pi(e)}) - \mathbb{E}(f_e|\mathcal{B}_{\pi(e)})\|_2^2 = \|\mathbb{E}(f_e|\mathcal{B}'_{\pi(e)})\|_2^2 - \|\mathbb{E}(f_e|\mathcal{B}_{\pi(e)})\|_2^2 \geq 2^{-2k}\,\varepsilon^{2^{k+1}}. \tag{2.22}$$

Note that the first equality above follows from the fact that conditional expectation function $\mathbb{E}(f|\mathcal{B})$ is the orthogonal projection of f to the subspace of $\mathcal{B}$-measurable functions in L^2. This also implies that energy of a function is always increasing when the underlying σ-algebra is refined, and (2.22) tells us that the energy of f_e is increased by at least $c_k\,\varepsilon^{2^{k+1}}$.

For $\mathfrak{f}' \notin \partial\pi(e)$ we set $\mathcal{B}'_{\mathfrak{f}'} := \mathcal{B}_{\mathfrak{f}'}$. Then the total energy of the family f_e with respect to the system $\mathcal{B}'_{\pi(e)} = \bigvee_{\mathfrak{f}'\in\partial\pi(e)} \mathcal{B}'_{\mathfrak{f}'}$, $e \in \mathcal{H}^2_{d,k}$ is also increased by at least $c_k\,\varepsilon^{2^{k+1}}$.

It is clear that the complexity of the σ-algebras $\mathcal{B}_{\mathfrak{f}'}$ are increased by at most 1, hence, as explained above, the lemma follows by applying this energy increment process at most $O(\varepsilon^{-2^{k+1}})$ times. □

3. The base case of an inductive strategy to establish Theorem 1.2

In this section we will ultimately establish the base case of our more general inductive argument. We however start by giving a quick review of the proof of Theorem 1.2 when $d = 1$ (which contains both Theorem B and Corollary B as stated in Section 1.1), namely the case of a single simplex. This was originally addressed in [1] and revisited in [12] and [13].

3.1. A Single Simplex in $\mathbb{R}^n$.

Let $Q \subseteq \mathbb{R}^n$ be a fixed cube and let $l(Q)$ denotes its side length.

Let $\Delta^0 = \{v_1 = 0, v_2, \dots, v_n\} \subseteq \mathbb{R}^n$ be a fixed non-degenerate simplex and define $t_{kl} := v_k \cdot v_l$ for $2 \leq k, l \leq n$ where " $\cdot$ " is the dot product on $\mathbb{R}^n$. Given $\lambda > 0$, a simplex $\Delta = \{x_1 = 0, x_2, \dots, x_n\} \subseteq \mathbb{R}^n$ is isometric to $\lambda\Delta^0$ if and only if $x_k \cdot x_l = \lambda^2 t_{kl}$ for all $2 \leq k, l \leq n$. Thus the configuration space $S_{\lambda\Delta^0}$ of isometric copies of $\lambda\Delta^0$ is a non-singular real variety given by the above equations. Let $\sigma_{\lambda\Delta^0}$ be natural normalized surface area measure on $S_{\lambda\Delta^0}$, described in [1], [12], and [13]. It is clear that the variable x_1 can be replaced by any of the variables x_i by redefining the constants t_{kl}.

For any family of functions $f_1, \dots, f_n : Q \to [-1, 1]$ and $0 < \lambda \ll l(Q)$ we define the multi-linear expression

$$\mathcal{N}^1_{\lambda\Delta^0,Q}(f_1, \dots, f_n) := \fint_{x_1\in Q} \int_{x_2,\dots,x_n} f_1(x_1) \dots f_n(x_n)\, d\sigma_{\lambda\Delta_0}(x_2 - x_1, \dots, x_n - x_1)\, dx_1. \tag{3.1}$$

We note that all of our functions are 1-bounded and both integrals, in fact all integrals in this paper, are normalized. Recall that we are using the normalized integral notation $\fint_A f := \frac{1}{|A|}\int_A f$. Since the normalized measure $\sigma_{\lambda\Delta^0}$ is supported on $S_{\lambda\Delta_0}$ we will not indicate the support of the variables $(x_2, \dots, x_n)$ explicitly.

Note also that if $S \subseteq Q$ is a measurable set and $\mathcal{N}^1_{\lambda\Delta^0,Q}(1_S, \dots, 1_S) > 0$ then S must contain an isometric copy of $\lambda\Delta^0$. The following proposition (with $Q = [0, 1]^n$) is a quantitatively stronger version of Proposition B that appeared in Section 1.1 and hence immediately establishes Theorem 1.2 for $d = 1$.

Proposition 3.1. *For any $0 < \varepsilon \leq 1$ there exists an integer $J = O(\varepsilon^{-2}\log\varepsilon^{-1})$ with the following property: Given any lacunary sequence $l(Q) \geq \lambda_1 \geq \cdots \geq \lambda_J$ and $S \subseteq Q$, there is some $1 \leq j < J$ such that*

$$\mathcal{N}^1_{\lambda\Delta^0,Q}(1_S, \dots, 1_S) > \left(\frac{|S|}{|Q|}\right)^n - \varepsilon \tag{3.2}$$

for all $\lambda \in [\lambda_{j+1}, \lambda_j]$.

Our approach to establishing Proposition 3.1 is to compare the above expressions to simpler ones for which it is easy to obtain lower bounds. Given a scale $0 < \lambda \ll l(Q)$ we define the multi-linear expression

$$\mathcal{M}^1_{\lambda,Q}(f_1, \dots, f_n) := \fint_{t\in Q} \fint_{x_1,x_2,\dots,x_n\in t+Q(\lambda)} f_1(x_1) \dots f_n(x_n)\, dx_1 \dots dx_n\, dt \tag{3.3}$$

where $Q(\lambda) = [-\frac{\lambda}{2}, \frac{\lambda}{2}]^n$ and $t + Q(\lambda)$ is the shift of the cube $Q(\lambda)$ by the vector t. Note that if $S \subseteq Q$ is a set of measure $|S| \geq \delta|Q|$ for some $\delta > 0$, then Hölder implies

$$\mathcal{M}^1_{\lambda,Q}(1_S, \dots, 1_S) = \fint_{t\in Q} \left(\fint_{x\in t+Q(\lambda)} 1_S(x)\, dx\right)^n dt \geq \left(\fint_{t\in Q}\fint_{x\in t+Q(\lambda)} 1_S(x)\, dx\right)^n \geq \delta^n - O(\varepsilon) \tag{3.4}$$

for all scales $0 < \lambda \ll \varepsilon\, l(Q)$.

Recall that for any $\varepsilon > 0$ we call a sequence $L_1 \geq \cdots \geq L_J$ *ε-admissible* if $L_j/L_{j+1} \in \mathbb{N}$ and $L_{j+1} \ll \varepsilon^2 L_j$ for all $1 \leq j < J$. Note that given any lacunary sequence $l(Q) \geq \lambda_1 \geq \cdots \geq \lambda_{J'}$ with $J' \gg (\log \varepsilon^{-1})\, J$, one can always finds an ε-admissible sequence of scales $l(Q) \geq L_1 \geq \cdots \geq L_J$ such that for each $1 \leq j < J$ the interval $[L_{j+1}, L_j]$ contains at least two consecutive elements from the original lacunary sequence.

In light of this observation, and the one above regarding a lower bound for $\mathcal{M}^1_{\lambda,Q}(1_S, \ldots, 1_S)$, our proof of Proposition 3.1 reduces to establishing the following "counting lemma".

Proposition 3.2. *Let $0 < \varepsilon < 1$. There exists an integer $J_1 = O(\varepsilon^{-2})$ such that for any ε-admissible sequence of scales $l(Q) \geq L_1 \geq \cdots \geq L_{J_1}$ and $S \subseteq Q$ there is some $1 \leq j < J_1$ such that*

$$\mathcal{N}^1_{\lambda\Delta^0,Q}(1_S, \ldots, 1_S) = \mathcal{M}^1_{\lambda,Q}(1_S, \ldots, 1_S) + O(\varepsilon) \tag{3.5}$$

for all $\lambda \in [L_{j+1}, L_j]$.

There are two main ingredients in the proof of Proposition 3.2, this will be typical to all of our arguments. The first ingredient is a result which establishes that the our multi-linear forms $\mathcal{N}^1_{\lambda\Delta^0,Q}(f_1, \ldots, f_n)$ are controlled by an appropriate norm which measures the uniformity of distribution of functions $f : Q \to [-1, 1]$ with respect to particular scales L. This is analogous to estimates in additive combinatorics [8], [**?**] which are often referred to as generalized von-Neumann inequalities.

The result below was proved in [12] for $Q = [0, 1]^n$, however a simple scaling of the variables x_i transfers the result to an arbitrary cube Q.

Lemma 3.1 (A Generalized von-Neumann inequality [12])**.** *Let $\varepsilon > 0$, $0 < \lambda \ll l(Q)$, and $0 < L \ll \varepsilon^6 \lambda$.*

For any collections of functions $f_1, \ldots, f_n : Q \to [-1, 1]$ we have

$$|\mathcal{N}^1_{\lambda\Delta^0,Q}(f_1, \ldots, f_n)| \leq \min_{i=1,\ldots,n} \|f_i\|_{U^1_L(Q)} + O(\varepsilon) \tag{3.6}$$

where for any $f : Q \to [-1, 1]$ we define

$$\|f\|^2_{U^1_L(Q)} := \fint_{t \in Q} \Big| \fint_{x \in t + Q(L)} f(x)\, dx \Big|^2 dt \tag{3.7}$$

with $t + Q(L)$ denoting the shift of the cube $Q(L) = [-\frac{L}{2}, \frac{L}{2}]^n$ by the vector t.

The corresponding inequality for the multilinear expression $\mathcal{M}^1_{\lambda,Q}(f_1, \ldots, f_n)$, namely the fact that

$$\mathcal{M}^1_{\lambda,Q}(f_1, \ldots, f_n) \leq \min_{i=1,\ldots,n} \|f_i\|_{U^1_L(Q)} + O(\varepsilon)$$

whenever $0 < L \ll \varepsilon^6 \lambda$ follows easily from Cauchy-Schwarz together with the simple observation that

$$\|f\|_{U^1_L(Q)} \leq \|f\|_{U^1_{L'}(Q)} + O(\varepsilon)$$

whenever $L' \ll \varepsilon L$.

The second key ingredient, proved in [13] and generalized in Lemma 3.3 below, is a Koopman-von Neumann type decomposition of functions where the underlying σ-algebras are generated by cubes of a fixed length. To recall it, let $Q \subseteq \mathbb{R}^n$ be a cube, $L > 0$ be scale that divides $l(Q)$, $Q(L) = [-\frac{L}{2}, \frac{L}{2}]^n$, and $\mathcal{G}_{L,Q}$ denote the collection of cubes $t + Q(L)$ partitioning the cube Q and $\Gamma_{L,Q}$ denote the grids corresponding to the centers of the cubes. By a slightly abuse of notation we also write $\mathcal{G}_{L,Q}$ for the σ-algebra generated by the grid. Recall that the conditional expectation function $\mathbb{E}(f|\mathcal{G}_{L,Q})$ is constant and equal to $\fint_A f$ on each cube $A \in \mathcal{G}_{L,Q}$.

Lemma 3.2 (A Koopman-von Neumann type decomposition [12])**.** *Let $0 < \varepsilon \leq 1$ and $Q \subseteq \mathbb{R}^n$ be a cube.*

There exists an integer $\bar{J}_1 = O(\varepsilon^{-2})$ such that for any ε-admissible sequence $l(Q) \geq L_1 \geq \cdots \geq L_{\bar{J}_1}$ and function $f : Q \to [-1, 1]$ there is some $1 \leq j < \bar{J}_1$ such that

$$\|f - \mathbb{E}(f|\mathcal{G}_{L_j,Q})\|_{U^1_{L_{j+1}}(Q)} \leq \varepsilon \tag{3.8}$$

Proof of Proposition 3.2. Let $\mathcal{G}_{L_j,Q}$ be the grid obtained from Lemma 3.2 for the functions $f = 1_S$ for some fixed $\varepsilon > 0$. Let $\bar{f} := \mathbb{E}(f|\mathcal{G}_{L_j,Q})$, then by (3.6) and multi-linearity, we have

$$\mathcal{N}^1_{\lambda\Delta^0,Q}(f, \ldots, f) = \mathcal{N}^1_{\lambda\Delta^0,Q}(\bar{f}, \ldots, \bar{f}) + O(\varepsilon),$$

and also

$$\mathcal{M}^1_{\lambda,Q}(f,\dots,f) = \mathcal{M}^1_{\lambda,Q}(\bar{f},\dots,\bar{f}) + O(\varepsilon)$$

provided for $\varepsilon^{-6}L_{j+1} \ll \lambda$. Thus in showing (6.4) one can replace the functions f with $\bar{f}$. If we make the additional assumption that $\lambda \ll \varepsilon L_j$ then it is easy to see, using the fact that the function $\bar{f}$ is constant on the cubes $Q_t(L_j) \in \mathcal{G}_{L_j,Q}$, that

$$\mathcal{N}^1_{\lambda\Delta^0,Q}(\bar{f},\dots,\bar{f}) = \mathcal{M}^1_{\lambda,Q}(\bar{f},\dots,\bar{f}) + O(\varepsilon).$$

Since the condition $\varepsilon^{-6}L_{j+1} \ll \lambda \ll \varepsilon L_j$ can be replaced with $L_{j+1} \ll \lambda \ll L_j$ if one passes to a subsequence of scales, for example $L'_j = L_{5j}$, this completes the proof of Proposition 3.2. □

3.2. **The base case of a general inductive strategy.**

In this section, as preparation to handle the case of products of simplices, we prove a parametric version of Proposition 3.2, namely Proposition 3.3 below, which will serve as the base case for later inductive arguments.

Let $Q = Q_1 \times \cdots \times Q_d$ with $Q_i \subseteq \mathbb{R}^{n_i}$ be cubes of equal side length $l(Q)$. Let L be a scale dividing $l(Q)$ and for each $\underline{t} = (t_1,\dots,t_d) \in \Gamma_{L,Q}$ let $Q_{\underline{t}}(L) = \underline{t} + Q(L)$ and $Q_{t_i}(L) = t_i + Q_i(L)$. Note that $Q_{\underline{t}}(L) = Q_{t_1}(L) \times \cdots \times Q_{t_d}(L)$. Here $Q(L) = [-\frac{L}{2},\frac{L}{2}]^n$ and $Q_i(L) = [-\frac{L}{2},\frac{L}{2}]^{n_i}$ for each $1 \le i \le d$.

Let $\Delta^0_i = \{v^i_1,\dots,v^i_{n_i}\} \subseteq \mathbb{R}^{n_i}$ be a non-degenerate simplex for each $1 \le i \le d$.

Proposition 3.3 (Parametric Counting Lemma on $\mathbb{R}^n$ for Simplices)**.**

Let $0 < \varepsilon \le 1$ and $R \ge 1$. There exists an integer $J_1 = J_1(\varepsilon,R) = O(R\,\varepsilon^{-4})$ such that for any ε-admissible sequence of scales $L_0 \ge L_1 \ge \cdots \ge L_{J_1}$ with the property that L_0 divides $l(Q)$ and collection of functions

$$f^{i,r}_{k,\underline{t}} : Q_{t_i}(L_0) \to [-1,1] \quad \text{with} \quad 1 \le i \le d,\ 1 \le k \le n_i,\ 1 \le r \le R \text{ and } \underline{t} \in \Gamma_{L_0,Q}$$

there exists $1 \le j < J_1$ and a set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| \le \varepsilon|\Gamma_{L_0,Q}|$ such that

$$\mathcal{N}^1_{\lambda\Delta^0_i,Q_{t_i}(L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) = \mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) + O(\varepsilon) \tag{3.9}$$

for all $\lambda \in [L_{j+1},L_j]$ and $\underline{t} \notin T_\varepsilon$ uniformly in $1 \le i \le d$ and $1 \le r \le R$.

The proof of Proposition 3.3 will follow from Lemma 3.1 and the following generalization of Lemma 3.2 in which we simultaneously consider a family of functions supported on the subcubes in a partition of an original cube Q.

Lemma 3.3 (A simultaneous Koopman-von Neumann type decomposition)**.**

Let $0 < \varepsilon \le 1$, $m \ge 1$, and $Q \subseteq \mathbb{R}^n$ be a cube. There exists an integer $\bar{J}_1 = O(m\varepsilon^{-3})$ such that for any ε-admissible sequence $L_0 \ge L_1 \ge \cdots \ge L_{\bar{J}_1}$ with the property that L_0 divides $l(Q)$, and collection of functions

$$f_{1,t},\dots,f_{m,t} : Q_t(L_0) \to [-1,1]$$

defined for each $t \in \Gamma_{L_0,Q}$, there is some $1 \le j < \bar{J}_1$ and a set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| \le \varepsilon|\Gamma_{L_0,Q}|$ such that

$$\|f_{i,t} - \mathbb{E}(f_{i,t}|\mathcal{G}_{L_j,Q_t(L_0)})\|_{U^1_{L_{j+1}}(Q_t(L_0))} \le \varepsilon \tag{3.10}$$

for all $1 \le i \le m$ and $t \notin T_\varepsilon$.

Proof of Proposition 3.3. Fix $1 \le i \le d$. For $1 \le k \le n_i$ and $\underline{t} = (t_1,\dots,t_d) \in \Gamma_{L_0,Q}$, we will abuse notation and write

$$f^{i,r}_{k,\underline{t}}(x_1,\dots,x_d) := f^{i,r}_{k,\underline{t}}(x_i)$$

for $(x_1,\dots,x_d) \in Q_{\underline{t}}(L_0)$.

If we apply Lemma 3.3 to the family of functions $f^{i,r}_{k,\underline{t}}$ on $Q_{\underline{t}}(L_0)$ for $1 \le i \le d$, $1 \le k \le n_i$, and $1 \le r \le R$, with $m = (n_1 + \ldots + n_d)R$, then this produces a grid $\mathcal{G}_{L_j,Q}$ for some $1 \le j \le \bar{J}_1 = O(\varepsilon^{-3}R)$, and a set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| \le \varepsilon|\Gamma_{L_0,Q}|$, such that

$$\|f^{i,r}_{k,\underline{t}} - \mathbb{E}(f^{i,r}_{k,\underline{t}}|\mathcal{G}_{L_j,Q})\|_{U^1_{L_{j+1}}(Q_{\underline{t}}(L_0))} \le \varepsilon$$

uniformly for $1 \le i \le d$, $1 \le k \le n_i$ and $1 \le r \le R$ for $t \notin T_\varepsilon$.

Since $f^{i,r}_{k,\underline{t}}(x_1,\dots,x_d) = f^{i,r}_{k,\underline{t}}(x_i)$ for $(x_1,\dots,x_d) \in Q_{\underline{t}}(L_0)$ it is easy to see that

$$\|f^{i,r}_{k,\underline{t}} - \mathbb{E}(f^{i,r}_{k,\underline{t}}|\mathcal{G}_{L_j,Q})\|_{U^1_{L_{j+1}}(Q_{\underline{t}}(L_0))} = \|f^{i,r}_{k,\underline{t}} - \mathbb{E}(f^{i,r}_{k,\underline{t}}|\mathcal{G}_{L_j,Q_i})\|_{U^1_{L_{j+1}}(Q_{t_i}(L_0))}.$$

Let $\bar{f}^{i,r}_{k,\underline{t}} := \mathbb{E}(f^{i,r}_{k,\underline{t}}|\mathcal{G}_{L_j,Q_i})$, then by Lemma 3.1, one has

$$\mathcal{N}^1_{\lambda\Delta^0_i,Q_{t_i}(L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) = \mathcal{N}^1_{\lambda\Delta^0_i,Q_{t_i}(L_0)}(\bar{f}^{i,r}_{1,\underline{t}},\dots,\bar{f}^{i,r}_{n_i,\underline{t}}) + O(\varepsilon),$$

and

$$\mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) = \mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(\bar{f}^{i,r}_{1,\underline{t}},\dots,\bar{f}^{i,r}_{n_i,\underline{t}}) + O(\varepsilon)$$

for all $\underline{t} \notin T_\varepsilon$ provided $\varepsilon^{-6}L_{j+1} \ll \lambda$. Finally, if we also have $\lambda \ll \varepsilon L_j$ then it is easy to see that

$$\mathcal{N}^1_{\lambda\Delta^0_i,Q_{t_i}(L_0)}(\bar{f}^{i,r}_{1,\underline{t}},\dots,\bar{f}^{i,r}_{n_i,\underline{t}}) = \mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(\bar{f}^{i,r}_{1,\underline{t}},\dots,\bar{f}^{i,r}_{n_i,\underline{t}}) + O(\varepsilon)$$

as the functions $\bar{f}^{i,r}_{k,\underline{t}}$ are constant on cubes $Q_{t_i}(L_j)$ of $\mathcal{G}_{L_j,Q_i}$, which are of size $L_j \ll \varepsilon L_0$.

Passing first to a subsequence of scales, for example $L'_j = L_{5j}$, the condition $\varepsilon^{-6}L_{j+1} \ll \lambda \ll \varepsilon L_j$ can be replaced with $L_{j+1} \ll \lambda \ll L_j$ so this completes the proof of the Proposition. □

We conclude this section with a sketch of the proof of Lemma 3.3. These arguments are standard, see for example the proof of Lemma 3.2 given in [12].

Proof of Lemma 3.3. First we make an observation about the $U^1_L(Q)$-norm. Suppose $0 < L' \ll \varepsilon^2 L$ with L' dividing L. If $s \in \Gamma_{L',Q}$ and $t \in Q_s(L')$ then $|t-s| = O(L')$ and hence

$$\fint_{x\in Q_t(L)} g(x)\,dx = \fint_{x\in Q_s(L)} g(x)\,dx + O(L'/L)$$

for any function $g : Q \to [-1,1]$. Moreover, since the cube $Q_s(L)$ is partitioned into the smaller cubes $Q_t(L')$, we have by Cauchy-Schwarz

$$\Big|\fint_{x\in Q_s(L)} g(x)\,dx\Big|^2 \le \mathbb{E}_{t\in\Gamma_{L',Q_s(L)}}\Big|\fint_{x\in Q_t(L')} g(x)\,dx\Big|^2.$$

From these observations it is easy to see that

$$\|g\|^2_{U^1_L(Q)} = \fint_{t\in Q}\Big|\fint_{x\in Q_t(L)} g(x)\,dx\Big|^2 dt \le \mathbb{E}_{t\in\Gamma_{L',Q}}\Big|\fint_{x\in Q_t(L')} g(x)\,dx\Big|^2 + O(L'/L)$$

and we note that the right side of the above expression is $\|\mathbb{E}(g|\mathcal{G}_{L',Q})\|^2_{L^2(Q)}$ since the conditional expectation function $\mathbb{E}(g|\mathcal{G}_{L',Q})$ is constant and equal to $\fint_{x\in Q_t(L')} g(x)\,dx$ on the cubes $Q_t(L')$.

Suppose that (3.10) does not hold for some $1 \le i \le m$ for every t in some set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| > \varepsilon\,|\Gamma_{L_0,Q}|$. If we apply the above observation to $g := f_{i,t} - \mathbb{E}(f_{i,t}|\mathcal{G}_{L_j,Q_t(L_0)})$, for every $t \in T_\varepsilon$, we obtain by orthogonality that

$$\sum_{i=1}^m \|\mathbb{E}(f_{i,t}|\mathcal{G}_{L_{j+2},Q_t(L_0)})\|^2_{L^2(Q_t(L_0))} \ge \sum_{i=1}^m \|\mathbb{E}(f_{i,t}|\mathcal{G}_{L_j,Q_t(L_0)})\|^2_{L^2(Q_t(L_0))} + c\varepsilon^2$$

for some constant $c > 0$.

If we now define $f_i : Q \to [-1,1]$ such that $f_i|_{(Q_t(L_0))} = f_{i,t}$, for $1 \le i \le m$, average over $t \in \Gamma_{L_0,Q}$, and use the fact $\|f_i\|^2_{L^2(Q)} = \mathbb{E}_{t\in\Gamma_{L_0,Q}}\|f_{i,t}\|^2_{L^2(Q_t(L_0))}$, we obtain

$$\sum_{i=1}^m \|\mathbb{E}(f_i|\mathcal{G}_{L_{j+2},Q})\|^2_{L^2(Q)} \ge \sum_{i=1}^m \|\mathbb{E}(f_i|\mathcal{G}_{L_j,Q})\|^2_{L^2(Q)} + c\varepsilon^3. \tag{3.11}$$

It is clear that the sums in the above expressions are bounded by m for all $j \ge 1$, thus (3.11) cannot hold for some $1 \le j \le \bar{J}_1$ for $\bar{J}_1 := C\,m\,\varepsilon^{-3}$. This implies that (3.10) must hold for some $1 \le j \le \bar{J}_1$, for all $1 \le i \le m$ and all $t \notin T_\varepsilon$ for a set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| \le \varepsilon\,|\Gamma_{L_0,Q}|$. □

4. Product of two simplices in $\mathbb{R}^n$

Although not strictly necessary, we discuss in this section the special case $d = 2$ of Theorem 1.2. This already gives an improvement of the main results of [12], but more importantly serves as a gentle preparation for the more complicated general case, presented in the Section 5, which involve both a plethora of different scales and the hypergraph bundle notation introduced in Section 2.2.

4.1. Proof of Theorem 1.2 with $d = 2$.

Let $Q = Q_1 \times Q_2$ with $Q_1 \subseteq \mathbb{R}^{n_1}$ and $Q_2 \subseteq \mathbb{R}^{n_2}$ be cubes of equal side length $l(Q)$ and $\Delta^0 = \Delta_1^0 \times \Delta_2^0$ with $\Delta_1^0 = \{v_{11}, \dots, v_{1n_1}\} \subseteq \mathbb{R}^{n_1}$ and $\Delta_2^0 = \{v_{11}, \dots, v_{2n_2}\} \subseteq \mathbb{R}^{n_2}$ two non-degenerate simplices.

In order to "count" configurations of the form $\Delta = \Delta_1 \times \Delta_2 \subseteq \mathbb{R}^{n_1+n_2}$ with Δ_1 and Δ_2 isometric copies of $\lambda\Delta_1^0$ and $\lambda\Delta_2^0$ respectively for some $0 < \lambda \ll l(Q)$ in a set $S \subseteq Q$ we introduce the multi-linear expression

$$\mathcal{N}^2_{\lambda\Delta^0,Q}(\{f_{kl}\}) := \fint_{x_{11}\in Q_1} \fint_{x_{21}\in Q_2} \int_{x_{12},\dots,x_{1n_1}} \int_{x_{22},\dots,x_{2n_2}} \prod_{k=1}^{n_1}\prod_{l=1}^{n_2} f_{kl}(x_{1k}, x_{2l}) \\ d\sigma_{\lambda\Delta_1^0}(x_{12} - x_{11}, \dots, x_{1n_1} - x_{11})\, d\sigma_{\lambda\Delta_2^0}(x_{22} - x_{21}, \dots, x_{2n_2} - x_{21})\, dx_{21}\, dx_{11}$$

for any family of functions $f_{kl} : Q_1 \times Q_2 \to [-1, 1]$ with $1 \le k \le n_1$ and $1 \le l \le n_2$.

Indeed, if $f_{kl} = 1_S$ for all $1 \le k \le n_1$ and $1 \le l \le n_2$ then the above expression is 0 unless there exists a configuration $\Delta \subseteq S$ of the form $\Delta_1 \times \Delta_2$ with Δ_1 and Δ_2 isometric copies of $\lambda\Delta_1^0$ and $\lambda\Delta_2^0$ respectively.

The short argument presented in Section 1.1 demonstrating how both Theorem B and Corollary B follow from Proposition B, and hence from Proposition 3.1, applies equally well to each of our main theorems. This reduces our main theorems to analogous quantitative results involving an arbitrary lacunary sequence of scales. In the case $d = 2$ of Theorem 1.2 this stronger quantitative result takes the following form:

Proposition 4.1. *For any $0 < \varepsilon \ll 1$ there exists an integer $J = O(\exp(C\varepsilon^{-13}))$ with the following property: Given any lacunary sequence $l(Q) \ge \lambda_1 \ge \cdots \ge \lambda_J$ and $S \subseteq Q$, there is some $1 \le j < J$ such that*

$$\mathcal{N}^2_{\lambda\Delta^0,Q}(\{1_S\}) > \left(\frac{|S|}{|Q|}\right)^{n_1 n_2} - \varepsilon \tag{4.1}$$

for all $\lambda \in [\lambda_{j+1}, \lambda_j]$.

Our approach to establishing Proposition 4.1 is again to compare the above expressions to simpler ones for which it is easy to obtain lower bounds. For any $0 < \lambda \ll l(Q)$ and family of functions $f_{kl} : Q_1 \times Q_2 \to [-1, 1]$ with $1 \le k \le n_1$ and $1 \le l \le n_2$ we consider

$$\mathcal{M}^2_{\lambda,Q}(\{f_{kl}\}) := \fint_{\underline{t}\in Q} \fint_{\underline{x}_1\in(t_1+Q_1(\lambda))^{n_1}} \fint_{\underline{x}_2\in(t_2+Q_2(\lambda))^{n_2}} \prod_{k=1}^{n_1}\prod_{l=2}^{n_2} f_{kl}(x_{1k}, x_{2l})\, d\underline{x}_2\, d\underline{x}_1\, d\underline{t}$$

where $\underline{t} = (t_1, t_2) \in Q_1 \times Q_2$, $\underline{x}_i = (x_{i1}, \dots, x_{in_i})$ and $Q_i(\lambda) = [-\frac{\lambda}{2}, \frac{\lambda}{2}]^{n_i}$ for $i = 1, 2$.

Note that if $S \subseteq Q$ is a set of measure $|S| \ge \delta|Q|$ for some $\delta > 0$, then careful applications of Hölder's inequality give

$$\mathcal{M}^2_{\lambda,Q}(\{1_S\}) \ge \fint_{\underline{t}\in Q} \left(\fint_{(x_1,x_2)\in\underline{t}+Q(\lambda)} 1_S(x_1, x_2)\, dx_1 dx_2\right)^{n_1 n_2} d\underline{t} \ge \delta^{n_1 n_2} - O(\varepsilon)$$

for all scales $0 < \lambda \ll \varepsilon\, l(Q)$.

In light of the observation above, and the discussion preceding Proposition 3.2, we see that Proposition 4.1, and hence Theorem 1.2 when $d = 2$, will follows as a consequence of the following

Proposition 4.2. *Let $0 < \varepsilon \ll 1$. There exists an integer $J_2 = O(\exp(C\varepsilon^{-12}))$ such that for any ε-admissible sequence of scales $l(Q) \ge L_1 \ge \cdots \ge L_{J_2}$ and $S \subseteq Q$ there is some $1 \le j < J_2$ such that*

$$\mathcal{N}^2_{\lambda\Delta^0,Q}(\{1_S\}) = \mathcal{M}^2_{\lambda,Q}(\{1_S\}) + O(\varepsilon) \tag{4.2}$$

for all $\lambda \in [L_{j+1}, L_j]$.

There are again two main ingredients in the proof of Proposition 4.2. The first establishes that the our multi-linear forms $\mathcal{N}^2_{\lambda\Delta^0,Q}(\{f_{kl}\})$ are controlled by an appropriate box-type norm attached to a scale L.

Let $Q = Q_1 \times Q_2$ be a cube. For any scale $0 < L \ll l(Q)$ and function $f : Q \to \mathbb{R}$ we define its local box norm at scale L to be

$$\|f\|^4_{\square_L(Q_1\times Q_2)} := \fint_{\underline{t}\in Q} \|f\|^4_{\square(\underline{t}+Q(L))}\, d\underline{t} \tag{4.3}$$

where $Q(L) = [-\frac{L}{2}, \frac{L}{2}]^{n_1+n_2}$ and

$$\|f\|^4_{\square(\widetilde{Q})} := \fint_{x_{11},x_{12}\in\widetilde{Q}_1} \fint_{x_{21},x_{22}\in\widetilde{Q}_2} f(x_{11},x_{21})f(x_{12},x_{21})f(x_{11},x_{22})f(x_{12},x_{22})\, dx_{11}\dots dx_{22} \tag{4.4}$$

for any cube $\widetilde{Q} \subseteq Q$ of the form $\widetilde{Q} = \widetilde{Q}_1 \times \widetilde{Q}_2$ with $\widetilde{Q}_j \subseteq Q_j$ for $j = 1, 2$.

Lemma 4.1 (A Generalized von-Neumann inequality [12])**.** *Let $\varepsilon > 0$, $0 < \lambda \ll l(Q)$, and $0 < L \ll \varepsilon^{24}\lambda$.*

For any collections of functions $f_{kl} : Q_1 \times Q_2 \to [-1, 1]$ with $1 \le k \le n_1$ and $1 \le l \le n_2$ we have both

$$|\mathcal{N}^2_{\lambda\Delta^0,Q}(\{f_{kl}\})| \le \min_{1\le k\le n_1,\, 1\le l\le n_2} \|f_{kl}\|_{\square_L(Q_1\times Q_2)} + O(\varepsilon) \tag{4.5}$$

$$|\mathcal{M}^2_{\lambda,Q}(\{f_{kl}\})| \le \min_{1\le k\le n_1,\, 1\le l\le n_2} \|f_{kl}\|_{\square_L(Q_1\times Q_2)}. \tag{4.6}$$

The result above was essentially proved in [12] for the multi-linear forms $\mathcal{N}^2_{\lambda\Delta^0,Q}$ when $Q = [0, 1]^{n_1+n_2}$, however a simple scaling argument transfers the result to an arbitrary cube Q. For completeness we include its short proof in Section 4.2 below.

The second and main ingredient is an analogue of a weak form of Szemerédi's regularity lemma due to Frieze and Kannan [7]. The more probabilistic formulation, we will use below, can be found for example in [21], [22], and [23], and is also sometimes referred to as a Koopman-von Neumann type decomposition.

For any cube $Q \subseteq \mathbb{R}^n$ and scale $L > 0$ that divides $l(Q)$ we will let $Q(L) = [-\frac{L}{2}, \frac{L}{2}]^n$ and $\mathcal{G}_{L,Q}$ denote the collection of cubes $Q_{\underline{t}}(L) = \underline{t} + Q(L)$ partitioning the cube Q and let $\Gamma_{L,Q}$ denote grid corresponding to the centers of these cubes. We will say that a finite σ-algebra $\mathcal{B}$ on Q is of *scale* L if it contains $\mathcal{G}_{L,Q}$ and for simplicity of notation will write $\mathcal{B}_{\underline{t}}$ for $\mathcal{B}|_{Q_{\underline{t}}(L)}$.

Recall that if we have two σ-algebras $\mathcal{B}_1$ on a cube Q_1 and $\mathcal{B}_2$ on Q_2 then by $\mathcal{B}_1 \vee \mathcal{B}_2$ we mean the σ-algebra on $Q = Q_1 \times Q_2$ generated by the sets $B_1 \times B_2$ with $B_1 \in \mathcal{B}_1$ and $B_2 \in \mathcal{B}_2$. Recall also that we say the complexity of a σ-algebra $\mathcal{B}$ is at most m, and write $\mathrm{complex}(\mathcal{B}) \le m$, if it is generated by m sets.

Lemma 4.2 (Weak regularity lemma in $\mathbb{R}^n$)**.**

Let $0 < \varepsilon \ll 1$ and $Q = Q_1 \times Q_2$ with $Q_1 \subseteq \mathbb{R}^{n_1}$ and $Q_2 \subseteq \mathbb{R}^{n_2}$ be cubes of equal side length $l(Q)$.

There exists an integer $\bar{J}_2 = O(\varepsilon^{-12})$ such that for any ε^4-admissible sequence $l(Q) \ge L_1 \ge \cdots \ge L_{\bar{J}_2}$ and function $f : Q \to [-1, 1]$ there is some $1 \le j \le \bar{J}_2$ and a σ-algebra $\mathcal{B}$ of scale L_j on Q such that

$$\|f - \mathbb{E}(f|\mathcal{B})\|_{\square_{L_{j+1}}(Q_1\times Q_2)} \le \varepsilon \tag{4.7}$$

which has the additional local structure that for each $\underline{t} = (t_1, t_2) \in \Gamma_{L_j,Q}$ there exist σ-algebras $\mathcal{B}_{1,\underline{t}}$ on $Q_{t_1}(L_j)$ and $\mathcal{B}_{2,\underline{t}}$ on $Q_{t_2}(L_j)$ with $\mathrm{complex}(\mathcal{B}_{i,\underline{t}}) = O(j)$ for $i = 1, 2$ such that $\mathcal{B}_{\underline{t}} = \mathcal{B}_{1,\underline{t}} \vee \mathcal{B}_{2,\underline{t}}$.

Comparing the above statement to Lemma 2.2 for $d = 2$, i.e to the weak regularity lemma, note that the σ-algebra $\mathcal{B}$ of scale L_j has a direct product structure only locally, inside each cube $Q_{\underline{t}}(L_j)$. Moreover this product structure varies with $\underline{t} \in \Gamma_{L_j,Q}$, however the "local complexity" remains uniformly bounded.

Assuming for now the validity of Lemmas 4.1 and 4.2 we prove Proposition 4.2. We will make crucial use of Proposition 3.3, namely our parametric counting lemma on $\mathbb{R}^n$ for simplices.

Proof of Proposition 4.2. Let $0 < \varepsilon \ll 1$, $\varepsilon_1 := \exp(-C_1\varepsilon^{-12})$ for some $C_1 \gg 1$, and $\{L_j\}_{j\ge1}$ be an ε_1-admissible sequence of scales. Set $R = \varepsilon\,\varepsilon_1^{-1}$ and $J_1(\varepsilon_1, R)$ be the parameter appearing in Proposition 3.3, noting that $J_1(\varepsilon_1, R) = O(\varepsilon_1^{-5})$.

For $L \in \{L_j\}_{j\geq 1}$ write $\mathrm{index}(L) = j$ if $L = L_j$. We now choose a subsequence $\{L'_j\} \subseteq \{L_j\}$ so that $L'_1 = L_1$ and $\mathrm{index}(L'_{j+1}) \geq \mathrm{index}(L'_j) + J_1(\varepsilon_1, R) + 2$. Applying Lemma 4.2, with $f_{kl} = f := 1_S$ for all $1 \leq k \leq n_1$ and $1 \leq l \leq n_2$, guarantees the existence of a σ-algebra $\mathcal{B}$ of scale L'_j on Q such that

$$\|f - \mathbb{E}(f|\mathcal{B})\|_{\square_{L'_{j+1}}(Q_1\times Q_2)} \leq \varepsilon \tag{4.8}$$

for some $1 \leq j \leq C\varepsilon^{-12}$. Moreover, we know that $\mathcal{B}$ has the additional local structure that for each $\underline{t} = (t_1, t_2) \in \Gamma_{L'_j,Q}$ there exist σ-algebras $\mathcal{B}_{1,\underline{t}}$ on $Q_{t_1}(L'_j)$ and $\mathcal{B}_{2,\underline{t}}$ on $Q_{t_2}(L'_j)$ with $\mathrm{complex}(\mathcal{B}_{i,\underline{t}}) = O(\varepsilon^{-12})$ for $i = 1, 2$ such that $\mathcal{B}_{\underline{t}} = \mathcal{B}_{1,\underline{t}} \vee \mathcal{B}_{2,\underline{t}}$. Thus, if we let $R_{1,\underline{t}}$ and $R_{2,\underline{t}}$ denote the number of atoms in $\mathcal{B}_{1,\underline{t}}$ and $\mathcal{B}_{2,\underline{t}}$ respectively, then we can assume, by formally adding the empty set to these collections of atoms if necessary, that $R_{1,\underline{t}} = R_{2,\underline{t}} = R' := \exp(C\varepsilon^{-12})$ for all $\underline{t} \in \Gamma_{L'_j,Q}$.

If we let $\bar{f} := \mathbb{E}(f|\mathcal{B}_1 \vee \mathcal{B}_2)$, then by Lemma 4.1 and multi-linearity we have

$$\mathcal{N}^2_{\lambda\Delta^0,Q}(\{f\}) = \mathcal{N}^2_{\lambda\Delta^0,Q}(\{\bar{f}\}) + O(\varepsilon) \quad \text{and} \quad \mathcal{M}^2_{\lambda,Q}(\{f\}) = \mathcal{M}^2_{\lambda,Q}(\{\bar{f}\}) + O(\varepsilon) \tag{4.9}$$

provided for $\varepsilon^{-24}L'_{j+1} \ll \lambda$. For a given $\underline{t} \in \Gamma_{Q,L'_j}$ write $\bar{f}_{\underline{t}}$ for the restriction of $\bar{f}$ to the cube $Q_{\underline{t}}(L'_j)$. By localization, one then has

$$\mathcal{N}^2_{\lambda\Delta^0,Q}(\{\bar{f}\}) = \mathbb{E}_{\underline{t}\in\Gamma_{L'_j,Q}}\, \mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) + O(\varepsilon), \tag{4.10}$$

and

$$\mathcal{M}^2_{\lambda,Q}(\{\bar{f}\}) = \mathbb{E}_{\underline{t}\in\Gamma_{L'_j,Q}}\, \mathcal{M}^2_{\lambda,Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) + O(\varepsilon) \tag{4.11}$$

provided one also insists that $\lambda \ll \varepsilon\, L'_j$.

For given $\underline{t} \in \Gamma_{L'_j,Q}$, the functions $\bar{f}_{\underline{t}}(x_1, x_2)$ are linear combinations of functions of the form $1_{A^{r_1}_{1,\underline{t}}}(x_1)1_{A^{r_2}_{2,\underline{t}}}(x_2)$, where $\{A^{r_1}_{1,\underline{t}}\}_{1\leq r_1\leq R'}$ and $\{A^{r_2}_{2,\underline{t}}\}_{1\leq r_2\leq R'}$ are the collections of the atoms of the σ-algebras $\mathcal{B}_{1,\underline{t}}$ and $\mathcal{B}_{2,\underline{t}}$ defined on the cubes $Q_{\underline{t}_1}(L'_j)$ and $Q_{\underline{t}_2}(L'_j)$. Thus for each $\underline{t} \in \Gamma_{L'_j,Q}$ one has

$$\bar{f}_{\underline{t}} = \sum_{r_1=1}^{R'}\sum_{r_2=1}^{R'} \alpha_{\underline{r},\underline{t}} 1_{A^{r_1}_{1,\underline{t}}} \times 1_{A^{r_2}_{2,\underline{t}}}$$

where $\underline{r} = (r_1, r_2)$. Plugging these linear expansions into the multi-linear expressions in above one obtains

$$\mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) = \sum_{\underline{r}=\{\underline{r}_{kl}\}_{kl}} \alpha_{\underline{r},\underline{t}}\, \mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(\{1_{A^{r_{1,kl}}_{1,\underline{t}}} \times 1_{A^{r_{2,kl}}_{2,\underline{t}}}\})$$

using the notations $\underline{r}_{kl} = (r_{1,kl}, r_{2,kl})$, $\alpha_{\underline{r},\underline{t}} = \prod_{kl} \alpha_{\underline{r}_{kl},\underline{t}}$. Notice that the product

$$\prod_{k=1}^{n_1}\prod_{l=1}^{n_2} 1_{A^{r_{1,kl}}_{1,\underline{t}}}(x_{1k})1_{A^{r_{2,kl}}_{2,\underline{t}}}(x_{2l})$$

is nonzero only if $A^{r_{1,kl}}_{1,\underline{t}} = A^{r_{1,k}}_{1,\underline{t}}$, that is if $r_{1,kl} = r_{1,k}$ for all $1 \leq l \leq n_2$, as the atoms $A^r_{1,\underline{t}}$ are all disjoint. Similarly, one has that $r_{2,kl} = r_{2,l}$ for all $1 \leq k \leq n_1$. Thus, in fact

$$\mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) = \sum_{\underline{r}=\{\underline{r}_{kl}\}_{kl}} \alpha_{\underline{r},\underline{t}}\, \mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(\{1_{A^{r_{1,k}}_{1,\underline{t}}} \times 1_{A^{r_{2,l}}_{2,\underline{t}}}\}) \tag{4.12}$$

and similarly

$$\mathcal{M}^2_{\lambda,Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) = \sum_{\underline{r}=\{\underline{r}_{kl}\}_{kl}} \alpha_{\underline{r},\underline{t}}\, \mathcal{M}^2_{\lambda,Q_{\underline{t}}(L'_j)}(\{1_{A^{r_{1,k}}_{1,\underline{t}}} \times 1_{A^{r_{2,l}}_{2,\underline{t}}}\}). \tag{4.13}$$

Note, that indices $\underline{r}$ are running through the index set $[1, R']^{n_1} \times [1, R']^{n_2}$ of size at most R if $C_1 \gg 1$.

The key observation is that

$$\mathcal{N}^2_{\lambda\Delta^0,Q_{\underline{t}}(L'_j)}(1_{A^{r_{1,k}}_{1,\underline{t}}} \times 1_{A^{r_{2,l}}_{2,\underline{t}}}) = \mathcal{N}^1_{\lambda\Delta^0_1,Q_{t_1}(L'_j)}(1_{A^{r_{1,1}}_{1,\underline{t}}}, \ldots, 1_{A^{r_{1,n_1}}_{1,\underline{t}}})\, \mathcal{N}^1_{\lambda\Delta^0_2,Q_{t_2}(L'_j)}(1_{A^{r_{2,1}}_{2,\underline{t}}}, \ldots, 1_{A^{r_{2,n_2}}_{2,\underline{t}}}) \tag{4.14}$$

and

$$\mathcal{M}^2_{\lambda,Q_{\underline{t}}(L'_j)}(1_{A^{r_{1,k}}_{1,\underline{t}}} \times 1_{A^{r_{2,l}}_{2,\underline{t}}}) = \mathcal{M}^1_{\lambda,Q_{t_1}(L'_j)}(1_{A^{r_{1,1}}_{1,\underline{t}}}, \ldots, 1_{A^{r_{1,n_1}}_{1,\underline{t}}})\, \mathcal{M}^1_{\lambda,Q_{t_2}(L'_j)}(1_{A^{r_{2,1}}_{2,\underline{t}}}, \ldots, 1_{A^{r_{2,n_2}}_{2,\underline{t}}}). \tag{4.15}$$

Let $\underline{r} = \{(r_{1,k}, r_{2,l})\}_{kl}$ and $g^{1,\underline{r}}_{k,\underline{t}} := 1_{A^{r_{1,k}}_{1,\underline{t}}}$, $g^{2,\underline{r}}_{l,\underline{t}} := 1_{A^{r_{2,l}}_{2,\underline{t}}}$. Writing $j' := \text{index}(L'_j)$ and $J' := \text{index}(L'_{j+1})$, one may apply Proposition 3.3 for the families of functions $g^{1,\underline{r}}_{k,\underline{t}}$, $g^{2,\underline{r}}_{l,\underline{t}}$, where $1 \le k \le n_1$, $1 \le l \le n_2$ and $\underline{r} = (r_{1,k}, r_{2,l})_{kl} \in [1, R']^{n_1} \times [1, R']^{n_2}$, with respect to the ε_1-admissible sequence of scales

$$L_{j'+1} \ge L_{j'+2} \ge \cdots \ge L_{J'-1}.$$

This is possible as $J' - j' = J_1(\varepsilon_1, R)$. Then there is a scale L_j with $j' \le j < J'$ so that

$$\mathcal{N}^1_{\lambda\Delta^0_1, Q_{t_1}(L'_j)}(g^{1,\underline{r}}_{1,\underline{t}}, \dots, g^{1,\underline{r}}_{n_1\underline{t}}) = \mathcal{M}^1_{\lambda, Q_{t_1}(L'_j)}(g^{1,\underline{r}}_{1,\underline{t}}, \dots, g^{1,\underline{r}}_{n_1,\underline{t}}) + O(\varepsilon_1) \tag{4.16}$$

and

$$\mathcal{N}^1_{\lambda\Delta^0_2, Q_{t_2}(L'_j)}(g^{2,\underline{r}}_{1,\underline{t}}, \dots, g^{2,\underline{r}}_{n_2,\underline{t}}) = \mathcal{M}^1_{\lambda, Q_{t_2}(L'_j)}(g^{2,\underline{r}}_{1,\underline{t}}, \dots, g^{2,\underline{r}}_{n_2,\underline{t}}) + O(\varepsilon_1), \tag{4.17}$$

for all $\lambda \in [L_{j+1}, L_j]$ uniformly in $\underline{r} = \{(r_{1,k}, r_{2,l})\}_{kl}$ and $\underline{t} \notin T_{\varepsilon_1} \subseteq \Gamma_{L'_j, Q}$, for a set of size $|T_{\varepsilon_1}| \le \varepsilon_1 |\Gamma_{L'_j, Q}|$. Then, by (4.14)-(4.15) and (4.12)-(4.13), we have

$$\mathcal{N}^2_{\lambda\Delta^0, Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) = \mathcal{M}^2_{\lambda, Q_{\underline{t}}(L'_j)}(\{\bar{f}_{\underline{t}}\}) + O(\varepsilon) \tag{4.18}$$

for $\underline{t} \notin T_{\varepsilon_1}$, as $|\alpha_{\underline{r},\underline{t}}| \le 1$ and $R\varepsilon_1 \le \varepsilon$. Finally, since $|T_{\varepsilon_1}| \le \varepsilon_1 |\Gamma_{L'_j, Q}|$, by averaging in $\underline{t} \in \Gamma_{L'_j, Q}$, one has

$$\mathcal{N}^2_{\lambda\Delta^0, Q}(\{\bar{f}\}) = \mathcal{M}^2_{\lambda, Q}(\{\bar{f}\}) + O(\varepsilon)$$

using (4.10)-(4.11) and the Proposition follows by (4.9) with an index $1 \le j < J_2 = O(\varepsilon^{-12}\varepsilon_1^{-5})$. □

4.2. **Proof of Lemmas 4.1 and 4.2.**

Proof of Lemma 4.1. First we note that if $\chi_L := L^{-n} 1_{[-L/2, L/2]^n}$ and $\psi_L := \chi_L * \chi_L$, then

$$\psi_L(x_2 - x_1) = \int_t \chi_L(x_1 - t)\chi_L(x_2 - t)\, dt$$

and hence for any function $f : Q \to [-1, 1]$, with $Q \subseteq \mathbb{R}^n$ being a cube of side length $l(Q)$, one has

$$\|f\|^2_{U^1_L(Q)} = \fint_{x_1 \in Q} \int_{x_2} f(x_1) f(x_2) \psi_L(x_2 - x_1)\, dx_1 dx_2 + O(L/l(Q)).$$

Write $\underline{x}' := (x_{21}, \dots, x_{2n_2})$ and let $g_{k,\underline{x}'}(x) := \prod_{l=1}^{n_2} f_{kl}(x, x_{2l})$. Then one may write

$$\mathcal{N}^2_{\lambda\Delta^0, Q}(\{f_{kl}\}) = \fint_{x_{21} \in Q_2} \int_{x_{22}, \dots, x_{2n_2}} \mathcal{N}^1_{\lambda\Delta^0_1, Q_1}(g_{1,\underline{x}'}, \dots, g_{n_1,\underline{x}'})\, d\sigma_{\lambda\Delta^0_2}(x_{22} - x_{21}, \dots, x_{2n_2} - x_{21})\, dx_{21}.$$

Using estimate (3.6), the above observation, and Cauchy-Schwarz one has

$$|\mathcal{N}^2_{\lambda\Delta^0, Q}(\{f_{kl}\})|^2 \le \fint_{x_{11} \in Q_1} \int_{x_{12}} \psi_L(x_{12} - x_{11})\, \mathcal{N}^1_{\lambda\Delta^0_2, Q_2}(h_{1,x_{11},x_{12}}, \dots, h_{n_2,x_{11},x_{12}})\, dx_{11} dx_{12} + O(\varepsilon^4)$$

provided $0 < \lambda \ll l(Q)$ and $0 < L \ll \varepsilon^{24}\lambda$ where $h_{l,x_{11},x_{12}}(x) = f_{1l}(x_{11}, x) f_{1l}(x_{12}, x)$ for $1 \le l \le n_2$. Applying the same procedure again ultimately gives

$$|\mathcal{N}^2_{\lambda\Delta^0, Q}(\{f_{kl}\})|^4 \le \|f_{11}\|^4_{\square_L(Q_1 \times Q_2)} + O(\varepsilon^4).$$

The same estimate can of course be given for any function f_{kl} in place of f_{11}. This establishes (4.5). Estimate (4.6) is established similarly. □

Proof of Lemma 4.2. For each $\underline{t} = (t_1, t_2) \in \Gamma_{L_1, Q}$ we will let $\mathcal{B}_{1,\underline{t}}(L_1) := \{\emptyset, Q_{t_1}(L_1)\}$ and $\mathcal{B}_{2,\underline{t}}(L_1) := \{\emptyset, Q_{t_2}(L_1)\}$, in other words the trivial σ-algebras on $Q_{t_1}(L_1)$ and $Q_{t_2}(L_1)$ respectively. If (4.8) holds with $\mathcal{B}(L_1) = \mathcal{G}_{L_1, Q}$, noting that $\mathcal{B}_{\underline{t}}(L_1) := \mathcal{B}_{1,\underline{t}}(L_1) \vee \mathcal{B}_{2,\underline{t}}(L_1)$ in this case, then we are done.

We now assume that we have developed, for each $\underline{t} = (t_1, t_2) \in \Gamma_{L_j, Q}$, σ-algebras $\mathcal{B}_{1,\underline{t}}(L_j)$ on $Q_{t_1}(L_j)$ and $\mathcal{B}_{2,\underline{t}}(L_j)$ on $Q_{t_2}(L_j)$ with $\text{complex}(\mathcal{B}_{i,\underline{t}}(L_j)) \le j$ for $i = 1, 2$. Let $\mathcal{B}(L_j)$ be the σ-algebra such that $\mathcal{B}_{\underline{t}}(L_j) = \mathcal{B}_{1,\underline{t}}(L_j) \vee \mathcal{B}_{2,\underline{t}}(L_j)$ for all $\underline{t} \in \Gamma_{L_j, Q}$ and assume that (4.8) does not hold, namely that

$$\|g\|_{\square_{L_{j+1}}(Q)} \ge \varepsilon$$

where $g := f - \mathbb{E}(f|\mathcal{B}(L_j))$. By the definition of the local box norm this means that

$$\fint_{\underline{t}\in Q} \|g\|^4_{\square(\underline{t}+Q(L_{j+1}))}\,dt \geq \varepsilon^4$$

and hence, as $L_{j+2} \ll \varepsilon^4 L_{j+1}$, it is easy to see that

$$\mathbb{E}_{\underline{s}\in\Gamma_{L_{j+2},Q}}\ \|g\|^4_{\square(\underline{s}+Q(L_{j+2}))} \geq \varepsilon^4/2.$$

This implies that there is a set $S \subseteq \Gamma_{L_{j+2},Q}$ of size $|S| \geq (\varepsilon^4/4)|\Gamma_{L_{j+2},Q}|$ such that for all $\underline{s} = (s_1, s_2) \in S$, one has that $\|g\|^4_{\square(Q_{\underline{s}}(L_{j+2}))} \geq \varepsilon^4/4$. It therefore follows, as is well-known see for example [12] or [23], that there exist sets $B_{1,\underline{s}} \subseteq Q_{s_1}(L_{j+2})$ and $B_{2,\underline{s}} \subseteq Q_{s_2}(L_{j+2})$ such that

$$\fint_{x_1\in Q_{s_1}(L_{j+2})} \fint_{x_2\in Q_{s_2}(L_{j+2})} g(x_1,x_2)\,1_{B_{1,\underline{s}}}(x_1)1_{B_{2,\underline{s}}}(x_2)\,dx_1\,dx_2 \geq \varepsilon^4/16. \tag{4.19}$$

For a given $\underline{s} \in \Gamma_{L_{j+2},Q}$ there is a unique $\underline{t} = \underline{t}(\underline{s})$ such that $Q_{\underline{s}}(L_{j+2}) \subseteq Q_{\underline{t}}(L_j)$. Let $\mathcal{B}'_{1,\underline{s}}(L_{j+2}) := \mathcal{B}_{1,\underline{t}}(L_j)|_{Q_{s_1}(L_{j+2})}$ and $\mathcal{B}'_{2,\underline{s}}(L_{j+2}) := \mathcal{B}_{2,\underline{t}}(L_j)|_{Q_{s_2}(L_{j+2})}$ noting that $\text{complex}(\mathcal{B}'_{i,\underline{s}}(L_{j+2})) \leq j$ for $i = 1, 2$, as the complexity of a σ-algebra does not increase when restricted to a set. If, for $i = 1, 2$, we let $\mathcal{B}_{i,\underline{s}}(L_{j+2})$ denote the σ-algebra generated by $\mathcal{B}'_{i,\underline{s}}(L_{j+2})$ and the set $B_{i,\underline{s}}$ if $\underline{s} \in S$ and let $\mathcal{B}_{i,\underline{s}}(L_{j+2}) := \mathcal{B}'_{i,\underline{s}}(L_{j+2})$ otherwise, then clearly $\text{complex}(\mathcal{B}_{i,\underline{s}}(L_{j+2}))$ is at most $j+1$. We now define $\mathcal{B}(L_{j+2})$ to the the sigma algebra of scale L_{j+2} with the property that $\mathcal{B}_{\underline{s}}(L_{j+2}) = \mathcal{B}_{1,\underline{s}}(L_{j+2}) \vee \mathcal{B}_{2,\underline{s}}(L_{j+2})$ for all $\underline{s} \in \Gamma_{L_{j+2},Q}$.

Using the inner product notation $\langle f, g\rangle_Q = \fint_Q f(x)g(x)\,dx$ we can rewrite (4.19) as

$$\langle f - \mathbb{E}(f|\mathcal{B}(L_j))\,,\, 1_{B_{1,\underline{s}}} \times 1_{B_{2,\underline{s}}}\rangle_{Q_{\underline{s}}(L_{j+2})} \geq \varepsilon^4/16$$

for all $\underline{s} \in S$. Since the function $1_{B_{1,\underline{s}}} \times 1_{B_{2,\underline{s}}}$ is measurable with respect to $\mathcal{B}(L_{j+2})$ one clearly has

$$\langle f - \mathbb{E}(f|\mathcal{B}(L_{j+2}))\,,\, 1_{B_{1,\underline{s}}} \times 1_{B_{2,\underline{s}}}\rangle_{Q_{\underline{s}}(L_{j+2})} = 0$$

and hence

$$\langle \mathbb{E}(f|\mathcal{B}(L_{j+2})) - \mathbb{E}(f|\mathcal{B}(L_j))\,,\, 1_{B_{1,\underline{s}}} \times 1_{B_{2,\underline{s}}}\rangle_{Q_{\underline{t}}(L_{j+2})} \geq \varepsilon^4/16.$$

It then follows from Cauchy-Schwarz and orthogonality that

$$\|\mathbb{E}(f|\mathcal{B}(L_{j+2}))\|^2_{L^2(Q_{\underline{s}}(L_{j+2}))} - \|\mathbb{E}(f|\mathcal{B}_1(L_j))\|^2_{L^2(Q_{\underline{s}}(L_{j+2}))} \geq \varepsilon^8/256.$$

Since $|S| \geq (\varepsilon^4/4)|\Gamma_{L_{j+2},Q}|$ averaging over all $\underline{s} \in \Gamma_{L_{j+2},Q}$ gives

$$\|\mathbb{E}(f|\mathcal{B}(L_{j+2}))\|^2_{L^2(Q)} \geq \|\mathbb{E}(f|\mathcal{B}(L_j))\|^2_{L^2(Q)} + \varepsilon^{12}/2^{10}.$$

Trivially both sides are at most 1 thus the process must stop at a step $j = O(\varepsilon^{-12})$ where (4.8) holds for a σ-algebra of "local complexity" at most j. This proves the Lemma. □

5. Proof of Theorem 1.2: The general case.

After these preparations we will now consider the general case of Theorem 1.2. Let $Q = Q_1 \times \cdots \times Q_d \subseteq \mathbb{R}^n$ with $Q_i \subseteq \mathbb{R}^{n_i}$ cubes of equal side length $l(Q)$ and $\Delta^0 = \Delta^0_1 \times \cdots \times \Delta^0_d$ with each $\Delta_i \subseteq \mathbb{R}^{n_i}$ a non-degenerate simplex of n_i points for $1 \leq i \leq d$.

We will use a generalized version of the hypergraph terminology introduced in Section 2. In particular, for a vertex set $I = \{1, 2, \ldots, d\}$ and set $K = \{il;\ 1 \leq i \leq d, 1 \leq l \leq n_i\}$ we will let $\pi : K \to I$ denote the projection defined by $\pi(il) := i$. As before we will let $\mathcal{H}_{d,k} := \{e \subseteq I;\ |e| = k\}$ denote the complete k-regular hypergraph with vertex set I, and for the multi-index $\underline{n} = (n_1, \ldots, n_d)$ define the hypergraph bundle

$$\mathcal{H}^{\underline{n}}_{d,k} := \{e \subseteq K;\ |e| = |\pi(e)| = k\}$$

noting that $|\pi^{-1}(i)| = n_i$ for all $i \in I$.

In order to parameterize the vertices of direct products of simplices, i.e. sets of the form $\Delta = \Delta_1 \times \cdots \times \Delta_d$ with $\Delta_i \subseteq Q_i$, we consider points $\underline{x} = (\underline{x}_1, \ldots, \underline{x}_d)$ with $\underline{x}_i = (x_{i1}, \ldots, x_{in_i}) \in Q_i^{n_i}$ for each $i \in I$. Now for any $1 \leq k \leq d$ and any edge $e' \in \mathcal{H}_{d,k}$ we will write $Q_{e'} := \prod_{i\in e'} Q_i$, and for every $\underline{x} \in Q_1^{n_1} \times \cdots \times Q_d^{n_d}$ and $e \in \mathcal{H}^{\underline{n}}_{d,k}$ we define $\underline{x}_e := \pi_e(\underline{x})$, where $\pi_e : Q_1^{n_1} \times \cdots \times Q_d^{n_d} \to Q_{\pi(e)}$ is the natural projection map.

Writing $\Delta_i = \{x_{i1}, \dots, x_{in_i}\}$ we have that $\Delta_1 \times \cdots \times \Delta_d = \{\underline{x}_e : \ e \in \mathcal{H}^n_{d,d}\}$ since every edge $\underline{x}_e$ is of the form $(x_{1l_1}, \dots, x_{dl_d})$. We can therefore identify points $\underline{x}$ with configurations of the form $\Delta_1 \times \cdots \times \Delta_d$.

For any $0 < \lambda \ll l(Q)$ the measures $d\sigma_{\lambda\Delta_i^0}$, introduced in Section 3.1, are supported on points $(y_2, \dots, y_{n_i})$ for which the simplex $\Delta_i = \{0, y_2, \dots, y_{n_i}\}$ is isometric to $\lambda\Delta_i^0$. For simplicity of notation we will write

$$\int_{\underline{x}_i} f(\underline{x}_i)\, d\sigma_i^\lambda(\underline{x}_i) := \fint_{x_{i1} \in Q_i} \int_{x_{i2}, \dots, x_{in_i}} f(\underline{x}_i)\, d\sigma_{\lambda\Delta_i^0}(x_{i2} - x_{i1}, \dots, x_{in_i} - x_{i1})\, dx_{i1}$$

Note that the support of the measure $d\sigma_i^\lambda$ is the set of points $\underline{x}_i$ so that the simplex $\Delta_i := \{x_{i1}, \dots, x_{in_i}\}$ is isometric to $\lambda\Delta_i^0$ and $x_{i1} \in Q_i$, moreover the measure is normalized. Thus if $S \subseteq Q$ is a set then the density of configurations Δ in S of the form $\Delta = \Delta_1 \times \ldots \times \Delta_d$ with each $\Delta_i \subseteq Q_i$ an isometric copy of $\lambda\Delta_i^0$ is given by the expression

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) := \int_{\underline{x}_1} \cdots \int_{\underline{x}_d} \prod_{e \in \mathcal{H}^n_{d,d}} 1_S(\underline{x}_e)\ d\sigma_1^\lambda(\underline{x}_1) \dots d\sigma_d^\lambda(\underline{x}_d). \tag{5.1}$$

The proof of Theorem 1.2 reduces to establishing the following stronger quantitative result.

Proposition 5.1. *For any $0 < \varepsilon \ll 1$ there exists an integer $J_d = J_d(\varepsilon)$ with the following property:*

Given any lacunary sequence $l(Q) \geq \lambda_1 \geq \cdots \geq \lambda_{J_d}$ and $S \subseteq Q$, there is some $1 \leq j < J_d$ such that

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) > \left(\frac{|S|}{|Q|}\right)^{n_1 \cdots n_d} - \varepsilon \tag{5.2}$$

for all $\lambda \in [\lambda_{j+1}, \lambda_j]$.

Quantitative Remark. A careful analysis of our proof reveals that there is a choice of $J_d(\varepsilon)$ which is less than $W_d(\log(C_\Delta \varepsilon^{-3}))$, where $W_k(m)$ is again the tower-exponential function defined by $W_1(m) = \exp(m)$ and $W_{k+1}(m) = \exp(W_k(m))$ for $k \geq 1$.

For any $0 < \lambda \ll l(Q)$ and set $S \subseteq Q$ we define the expression:

$$\mathcal{M}^d_{\lambda,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) := \fint_{\underline{t} \in Q} \mathcal{M}^d_{\underline{t}+Q(\lambda)}(1_S\,;\, e \in \mathcal{H}^n_{d,d})\, d\underline{t} \tag{5.3}$$

where $Q(\lambda) = [-\frac{\lambda}{2}, \frac{\lambda}{2}]^n$ and

$$\mathcal{M}^d_{\widetilde{Q}}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) := \fint_{\underline{x}_1 \in \widetilde{Q}_1^{n_1}} \cdots \fint_{\underline{x}_d \in \widetilde{Q}_d^{n_d}} \prod_{e \in \mathcal{H}^n_{d,d}} 1_S(\underline{x}_e)\, d\underline{x}_1 \dots d\underline{x}_d \tag{5.4}$$

for any cube $\widetilde{Q} \subseteq Q$ of the form $\widetilde{Q} = \widetilde{Q}_1 \times \cdots \times \widetilde{Q}_d$ with $\widetilde{Q}_i \subseteq Q_i$ for $1 \leq i \leq d$. Note that if $S \subseteq Q$ is a set of measure $|S| \geq \delta|Q|$ for some $\delta > 0$, then careful applications of Hölder's inequality give

$$\mathcal{M}^d_{\lambda,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) \geq \fint_{\underline{t} \in Q} \left(\fint_{(x_1, \dots, x_d) \in \underline{t}+Q(\lambda)} 1_S(x_1, \dots, x_d)\, dx_1 \dots dx_d \right)^{n_1 \cdots n_d} d\underline{t} \ \geq \delta^{n_1 \cdots n_d} - O(\varepsilon)$$

for all scales $0 < \lambda \ll \varepsilon\, l(Q)$.

In light of the discussion above, and that preceding Proposition 3.2, we see that Proposition 5.1, and hence Theorem 1.2 in general, will follows as a consequence of the following

Proposition 5.2. *Let $0 < \varepsilon \ll 1$. There exists an integer $J_d = J_d(\varepsilon)$ such that for any ε-admissible sequence of scales $l(Q) \geq L_1 \geq \cdots \geq L_{J_d}$ and $S \subseteq Q$ there is some $1 \leq j < J_d$ such that*

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) = \mathcal{M}^d_{\lambda,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) + O(\varepsilon) \tag{5.5}$$

for all $\lambda \in [L_{j+1}, L_j]$.

The validity of Proposition 5.2 will follow immediately from the $d = k$ case of Proposition 5.3 below.

5.1. **Reduction of Proposition 5.2 to a more general "local" counting lemma.**

For any given $1 \leq k \leq d$ and collection of functions $f_e : Q_{\pi(e)} \to [-1,1]$ with $e \in \mathcal{H}^n_{d,k}$ we define the following multi-linear expressions

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e; e \in \mathcal{H}^n_{d,k}) := \int_{\underline{x}_1} \cdots \int_{\underline{x}_d} \prod_{e \in \mathcal{H}^n_{d,k}} f_e(\underline{x}_e)\ d\sigma^\lambda_1(\underline{x}_1) \ldots . d\sigma^\lambda_d(\underline{x}_d) \tag{5.6}$$

and

$$\mathcal{M}^d_{\lambda,Q}(f_e\,;\, e \in \mathcal{H}^n_{d,k}) := \fint_{\underline{t} \in Q} \mathcal{M}^d_{\underline{t}+Q(\lambda)}(f_e\,;\, e \in \mathcal{H}^n_{d,k})\, d\underline{t} \tag{5.7}$$

where $Q(\lambda) = [-\frac{\lambda}{2}, \frac{\lambda}{2}]^n$ and

$$\mathcal{M}^d_{\widetilde{Q}}(f_e\,;\, e \in \mathcal{H}^n_{d,k}) := \fint_{\underline{x}_1 \in \widetilde{Q}^{n_1}_1} \cdots \fint_{\underline{x}_d \in \widetilde{Q}^{n_d}_d} \prod_{e \in \mathcal{H}^n_{d,k}} f_e(\underline{x}_e)\, d\underline{x}_1 \ldots d\underline{x}_d \tag{5.8}$$

for any cube $\widetilde{Q} \subseteq Q$ of the form $\widetilde{Q} = \widetilde{Q}_1 \times \cdots \times \widetilde{Q}_d$ with $\widetilde{Q}_i \subseteq Q_i$ for $1 \leq i \leq d$.

Our strategy to proving Proposition 5.2 is the same as illustrated in the finite field settings, that is we would like to compare averages $\mathcal{N}_{\lambda\Delta^0,Q}(f_e; e \in \mathcal{H}^n_{d,k})$ to those of $\mathcal{M}^d_{\lambda,Q}(f_e\,;\, e \in \mathcal{H}^n_{d,k})$, at certain scales $\lambda \in [L_{j+1}, L_j]$, inductively for $1 \leq k \leq d$. However in the Euclidean case, an extra complication emerges due to the fact the (hypergraph) regularity lemma, the analogue of Lemma 2.2, does not produce σ-algebras $\mathcal{B}_{\mathfrak{f}}$, for $\mathfrak{f} \in \mathcal{H}^n_{d,k-1}$, on the cubes $Q_{\mathfrak{f}}$. In a similar manner to the case for $d = 2$ discussed in the previous section, we will only obtain σ-algebras "local" on cubes $Q_{\underline{t}_{\mathfrak{f}}}(L_0)$ at some scale $L_0 > 0$. This will have the effect that the functions f_e will be replaced by a family of functions $f_{e,\underline{t}}$, where $\underline{t}$ runs through a grid $\Gamma_{L_0,Q}$.

To be more precise, let $L > 0$ be a scale dividing the side-length $l(Q)$. For $\underline{t} \in \Gamma_{L,Q}$ and $e' \in \mathcal{H}_{d,k}$ we will use $\underline{t}_{e'}$ to denote the projection of $\underline{t}$ onto $Q_{e'}$ and $Q_{\underline{t}_{e'}}(L) := \underline{t}_{e'} + Q_{e'}(L)$ to denote the projection of the cube $Q_{\underline{t}}(L)$ centered at $\underline{t}$ onto $Q_{e'}$. It is then easy to see that for any $\varepsilon > 0$ we have

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{t} \in \Gamma_{L,Q}}\, \mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L)}(f_{e,\underline{t}}\,; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{5.9}$$

and

$$\mathcal{M}^d_{\lambda,Q}(f_e; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{t} \in \Gamma_{L,Q}}\, \mathcal{M}^d_{\lambda,Q_{\underline{t}}(L)}(f_{e,\underline{t}}\,; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{5.10}$$

provided $0 < \lambda \ll \varepsilon L$ where $f_{e,\underline{t}}$ denotes the restriction of a function f_e to the cube $Q_{\underline{t}}(L)$.

At this point the proof of Proposition 5.2 reduces to showing that the expressions in (7.8) and (7.9) only differ by $O(\varepsilon)$ at some scales $\lambda \in [L_{j+1}, L_j]$, given an ε-admissible sequence $L_0 \geq L_1 \geq \cdots \geq L_J$, for any collection of bounded functions $f_{e,\underline{t}}$, $e \in \mathcal{H}^n_{d,k}$, $\underline{t} \in \Gamma_{L_0,Q}$. Indeed, our crucial result will be the following

Proposition 5.3 (Local Counting Lemma)**.** *Let $0 < \varepsilon \ll 1$ and $M \geq 1$. There exists an integer $J_k = J_k(\varepsilon, M)$ such that for any ε-admissible sequence of scales $L_0 \geq L_1 \geq \cdots \geq L_{J_k}$ with the property that L_0 divides $l(Q)$, and collection of functions*

$$f^m_{e,\underline{t}} : Q_{t_{\pi(e)}}(L_0) :\to [-1,1] \quad \textit{with} \quad e \in \mathcal{H}^n_{d,k},\ 1 \leq m \leq M \ \textit{and}\ \underline{t} \in \Gamma_{L_0,Q}$$

there exists $1 \leq j < J_k$ and a set $T_\varepsilon \subseteq \Gamma_{L_0,Q}$ of size $|T_\varepsilon| \leq \varepsilon |\Gamma_{L_0,Q}|$ such that

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathcal{M}_{\lambda,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{5.11}$$

for all $\lambda \in [L_{j+1}, L_j]$ and $\underline{t} \notin T_\varepsilon$ uniformly in $e \in \mathcal{H}^n_{d,k}$ and $1 \leq m \leq M$.

5.2. **Proof of Proposition 5.3.**

We will prove Proposition 5.3 by induction on $1 \leq k \leq d$. For $k = 1$ this is basically Proposition 3.3.

Indeed, in this case for a given $\underline{t} = (t_1, \ldots, t_d) \in \Gamma_{L_0,Q}$ and edge $e \in \mathcal{H}^n_{d,1} = \{il : 1 \leq i \leq d,\ 1 \leq l \leq n_i\}$ we have that $f^m_{e,\underline{t}}(\underline{x}_e) = f^m_{il,\underline{t}}(x_{il})$ with $x_{il} \in Q_{t_i}(L_0)$ and hence both

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,1}) = \prod_{i=1}^d \mathcal{N}^1_{\lambda\Delta^0_i,Q_{t_i}(L_0)}(f^m_{i1,\underline{t}}, \ldots, f^m_{in_i,\underline{t}})$$

$$\mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}};\, e\in\mathcal{H}^{\underline{n}}_{d,1}) = \prod_{i=1}^{d}\mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(f^m_{i1,\underline{t}},\dots,f^m_{in_i,\underline{t}}).$$

By Proposition 3.3 there exists an $1\le j<J_1=O(M\varepsilon^{-4})$ and an exceptional set $T_\varepsilon\subseteq\Gamma_{L_0,Q}$ of size $|T_\varepsilon|\le\varepsilon|\Gamma_{L_0,Q}|$, such that uniformly for $\underline{t}\notin T_\varepsilon$ and for $1\le i\le d$, one has

$$\mathcal{N}^1_{\lambda\Delta^0,Q_{t_i}(L_0)}(f^m_{i1,\underline{t}},\dots,f^m_{in_i,\underline{t}}) = \mathcal{M}^1_{\lambda,Q_{t_i}(L_0)}(f^m_{i1,\underline{t}},\dots,f^m_{in_i,\underline{t}})+O(\varepsilon)$$

hence

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}};\, e\in\mathcal{H}^{\underline{n}}_{d,1}) = \mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}};\, e\in\mathcal{H}^{\underline{n}}_{d,1})+O(\varepsilon)$$

as the all factors are trivially bounded by 1 in magnitude. This implies (5.11) for $k=1$.

For the induction step we again need two main ingredients. The first establishes that the our multi-linear forms $\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e;\, e\in\mathcal{H}^{\underline{n}}_{d,k})$ are controlled by an appropriate box-type norm attached to a scale L.

Let $Q=Q_1\times\cdots\times Q_d$ and $1\le k\le d$. For any scale $0<L\ll l(Q)$ and function $f:Q_{e'}\to[-1,1]$ with $e'\in\mathcal{H}_{d,k}$ we define its local box norm at scale L by

$$\|f\|^{2^k}_{\square_L(Q_{e'})} := \fint_{\underline{s}\in Q_{e'}}\|f\|^{2^k}_{\square(\underline{s}+Q(L))}\,d\underline{s} \tag{5.12}$$

where

$$\|f\|^{2^k}_{\square(\widetilde{Q})} := \fint_{x_{11},x_{12}\in\widetilde{Q}_1}\cdots\fint_{x_{k1},x_{k2}\in\widetilde{Q}_k}\prod_{(\ell_1,\dots,\ell_k)\in\{1,2\}^k} f(x_{1\ell_1},\dots,x_{k\ell_k})\,dx_{11}\,dx_{12}\dots\,dx_{k1}\,dx_{k2} \tag{5.13}$$

for any cube $\widetilde{Q}$ of the form $\widetilde{Q}=\widetilde{Q}_1\times\cdots\times\widetilde{Q}_k$.

Lemma 5.1 (Generalized von-Neumann inequality)**.** *Let $\varepsilon>0$, $0<\lambda\ll l(Q)$ and $0<L\ll(\varepsilon^{2^k})^6\lambda$.*

For any $1\le k\le d$ and collection of functions $f_e:Q_{\pi(e)}\to[-1,1]$ with $e\in\mathcal{H}^{\underline{n}}_{d,k}$ we have both

$$|\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e;\, e\in\mathcal{H}^{\underline{n}}_{d,k})|\le\min_{e\in\mathcal{H}^{\underline{n}}_{d,k}}\|f_e\|_{\square_L(Q_{\pi(e)})}+O(\varepsilon) \tag{5.14}$$

$$|\mathcal{M}^d_{\lambda,Q}(f_e;\, e\in\mathcal{H}^{\underline{n}}_{d,k})|\le\min_{e\in\mathcal{H}^{\underline{n}}_{d,k}}\|f_e\|_{\square_L(Q_{\pi(e)})}. \tag{5.15}$$

The crucial ingredient is the following analogue of the weak hypergraph regularity lemma.

Lemma 5.2 (Parametric weak hypergraph regularity lemma for $\mathbb{R}^n$)**.** *Let $0<\varepsilon\ll1$, $M\ge1$, and $1\le k\le d$.*

There exists $\bar{J}_k=O(M\varepsilon^{-2^{k+3}})$ such that for any ε^{2^k}-admissible sequence $L_0\ge L_1\ge\cdots\ge L_{\bar{J}_k}$ with the property that L_0 divides $l(Q)$ and collection of functions

$$f^m_{e,\underline{t}}:Q_{\underline{t}_{\pi(e)}}(L_0)\to[-1,1]\quad\text{with}\quad e\in\mathcal{H}^{\underline{n}}_{d,k},\ 1\le m\le M,\ \text{and}\ \underline{t}\in\Gamma_{L_0,Q}$$

there is some $1\le j<\bar{J}_k$ and σ-algebras $\mathcal{B}_{e',\underline{t}}$ of scale L_j on $Q_{\underline{t}_{e'}}(L_0)$ for each $\underline{t}\in\Gamma_{L_0,Q}$ and $e'\in\mathcal{H}_{d,k}$ such that

$$\|f^m_{e,\underline{t}}-\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}})\|_{\square_{L_{j+1}}(Q_{\underline{t}_{\pi(e)}}(L_0))}\le\varepsilon \tag{5.16}$$

uniformly for all $t\notin T_\varepsilon$, $e\in\mathcal{H}^{\underline{n}}_{d,k}$, and $1\le m\le M$, where $T_\varepsilon\subseteq\Gamma_{L_0,Q}$ with $|T_\varepsilon|\le\varepsilon|\Gamma_{L_0,Q}|$.

Moreover, the σ-algebras $\mathcal{B}_{e',\underline{t}}$ have the additional local structure that the exist σ-algebras $\mathcal{B}_{e',\mathfrak{f}',\underline{s}}$ on $Q_{\underline{s}_{\mathfrak{f}'}}(L_j)$ with $\operatorname{complex}(\mathcal{B}_{e',\mathfrak{f}',\underline{s}})=O(j)$ for each $\underline{s}\in\Gamma_{L_j,Q}$, $e'\in\mathcal{H}_{d,k}$, and $\mathfrak{f}'\in\partial e'$ such that if $\underline{s}\in Q_{\underline{t}}(L_0)$, then

$$\mathcal{B}_{e',\underline{t}}\big|_{Q_{\underline{s}_{e'}}(L_j)} = \bigvee_{\mathfrak{f}'\in\partial e'}\mathcal{B}_{e',\mathfrak{f}',\underline{s}}. \tag{5.17}$$

Lemma 5.2 is the parametric and simultaneous version of the extension of Lemma 3.7 to the product of d simplices. The difference is that in the general case one has to deal with a parametric family of functions $f^m_{e,\underline{t}}$ as $\underline{t}$ is running through a grid $\Gamma_{L_0,Q}$. The essential new content of Lemma 5.2 is that one can develop σ-algebras $\mathcal{B}_{e',\underline{t}}$ on the cubes $Q_{\underline{t}}(L_0)$ with respect to the family of functions $f^m_{e,\underline{t}}$ such that the local structure described above and (5.16) hold simultaneously for almost all $\underline{t}\in\Gamma_{L_0,Q}$.

Proof of Proposition 5.3. Assume the Proposition holds for $k-1$.

Let $\varepsilon > 0$, $\varepsilon_1 := \exp(-C_1\varepsilon^{-2^{k+3}})$ for some large constant $C_1 = C_1(n,k,d) \gg 1$, and $\{L_j\}_{j\geq 1}$ be an ε_1-admissible sequence of scales. Set $F(\varepsilon) := J_{k-1}(\varepsilon_1, M)$ with $M = \varepsilon\,\varepsilon_1^{-1}$.

For $L \in \{L_j\}_{j\geq 1}$ we again write $\mathrm{index}(L) = j$ if $L = L_j$. We now choose a subsequence $\{L'_j\} \subseteq \{L_j\}$ so that $L'_0 = L_0$ and $\mathrm{index}(L'_{j+1}) \geq \mathrm{index}(L'_j) + F(\varepsilon) + 2$. Lemma 5.2 then guarantees the existence of σ-algebras $\mathcal{B}_{e',\underline{t}}$ of scale L'_j on $Q_{\underline{t}_{e'}}(L_0)$ for each $\underline{t} \in \Gamma_{L_0,Q}$ and $e' \in \mathcal{H}_{d,k}$, with the local structure described above, such that

$$\|f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}})\|_{\square_{L'_{j+1}}(Q_{\underline{t}_{\pi(e)}}(L_0))} \leq \varepsilon \tag{5.18}$$

uniformly for all $t \notin T'_\varepsilon$, $e \in \mathcal{H}^n_{d,k}$, and $1 \leq m \leq M$, for some $1 \leq j < \bar{J}_k(\varepsilon, M) = O(M\varepsilon^{-2^{k+3}})$, where $T'_\varepsilon \subseteq \Gamma_{L_0,Q}$ with $|T'_\varepsilon| \leq \varepsilon|\Gamma_{L_0,Q}|$. Let $\bar{f}^m_{e,\underline{t}} := \mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}})$ for $\underline{t} \in \Gamma_{L_0,Q}$ and $e \in \mathcal{H}^n_{d,k}$. If $t \notin T'_\varepsilon$, then by (5.14), (5.15), and (5.16) we have both

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(\bar{f}^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{5.19}$$

$$\mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}(f^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}(\bar{f}^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon). \tag{5.20}$$

provided $(\varepsilon^{-2^k})^6 L'_{j+1} \ll \lambda$. For given $\underline{s} \in \Gamma_{L'_j,Q_{\underline{t}}(L_0)}$ one may write $\bar{f}^m_{e,\underline{s}}$ for the restriction of $\bar{f}^m_{e,\underline{t}}$ on the cube $Q_{\underline{s}}(L'_j) \subseteq Q_{\underline{t}}(L_0)$, as $\underline{s}$ uniquely determines $\underline{t}$. By localization, provided $\lambda \ll \varepsilon L'_j$, we then have both

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}(\bar{f}^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{s}\in\Gamma_{L'_j,Q_{\underline{t}}(L_0)}}\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{e,\underline{s}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon), \tag{5.21}$$

$$\mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}(\bar{f}^m_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{s}\in\Gamma_{L'_j,Q_{\underline{t}}(L_0)}}\mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{e,\underline{s}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon). \tag{5.22}$$

For a fixed cube $Q_{\underline{s}}(L'_j)$ we have that

$$\bar{f}^m_{e,\underline{s}} = \sum_{r_e=1}^{R_{e,\underline{s}}} \alpha_{\underline{s},r_e,m}\ 1_{A^{r_e}_{\pi(e),\underline{s}}} \tag{5.23}$$

where $\{A^{r_e}_{\pi(e),\underline{s}}\}_{1\leq r\leq R_{e,\underline{s}}}$ is the family of atoms of the σ-algebra $\mathcal{B}_{\pi(e),\underline{t}}$ restricted to the cube $Q_{\underline{s}}(L'_j)$. Note that $|\alpha_{\underline{s},r_e}| \leq 1$ and $|R_{e,\underline{s}}| = O(\exp(C\varepsilon^{-2^{k+3}}))$. By adding the empty set to the collection of atoms one may assume $|R_{e,\underline{s}}| = R := \exp(C\varepsilon^{-2^{k+3}})$ for all $e \in \mathcal{H}^n_{d,k}$ and $\underline{s} \in \Gamma_{L'_j,Q}$. Then, by multi-linearity, using the notations $\underline{r} = (r_e)_{e\in\mathcal{H}^n_{d,k}}$ and $\alpha_{\underline{r},\underline{s}} = \prod_e \alpha_{\underline{s},r_e}$, one has both

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{\underline{s},e};\ e \in \mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{s},\underline{r},m}\ \mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(1_{A^{r_e}_{\pi(e),\underline{s}}};\ e \in \mathcal{H}^n_{d,k}) \tag{5.24}$$

$$\mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{\underline{s},e};\ e \in \mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{s},\underline{r},m}\ \mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}(1_{A^{r_e}_{\pi(e),\underline{s}}};\ e \in \mathcal{H}^n_{d,k}). \tag{5.25}$$

The key observation is that these expressions in the sum above are all at level $k-1$ instead of k. To see this let $e = (i_1l_1, \ldots, i_ml_m, \ldots, i_kl_k)$ so $e' = \pi(e) = (i_1, \ldots, i_m, \ldots, i_k)$. If $\mathfrak{f}' = e'\backslash\{i_m\}$ then recall that the edge $p_{\mathfrak{f}'}(e) = (i_1l_1, \ldots, i_kl_k) \in \mathcal{H}^n_{d,k-1}$ is obtained from e by removing the i_ml_m-entry. Thus, for any atom $A_{e',\underline{s}}$ of $\mathcal{B}_{\underline{s},e'}(L'_j)$ we have by (5.17), that

$$1_{A_{e',\underline{s}}}(\underline{x}_e) = \prod_{\mathfrak{f}'\in\partial e'} 1_{A_{e',\mathfrak{f}',\underline{s}},}(\underline{x}_{p_{\mathfrak{f}'}(e)}) \tag{5.26}$$

where $A_{e',\mathfrak{f}',\underline{s}}$ is an atom of the σ-algebra $\mathcal{B}_{e',\mathfrak{f}',\underline{s}}$. Thus

$$\prod_{e\in\mathcal{H}^n_{d,k}} 1_{A^{r_e}_{\pi(e),\underline{s}}}(\underline{x}_e) = \prod_{\mathfrak{f}\in\mathcal{H}^n_{d,k-1}} \prod_{\substack{e\in\mathcal{H}^n_{d,k},\mathfrak{f}'\in\partial\pi(e)\\ p_{\mathfrak{f}'}(e)=f}} 1_{A^{r_e}_{\pi(e),\mathfrak{f}',\underline{s}}}(\underline{x}_{\mathfrak{f}}) = \prod_{\mathfrak{f}\in\mathcal{H}^n_{d,k-1}} g^{\underline{r}}_{\mathfrak{f},\underline{s}}(\underline{x}_{\mathfrak{f}}). \tag{5.27}$$

It follows that

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(1_{A^{r_e}_{\pi(e),\underline{s}}};\ e \in \mathcal{H}^n_{d,k}) = \mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\ \mathfrak{f} \in \mathcal{H}^n_{d,k-1}) \tag{5.28}$$

and hence that

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{e,\underline{s}};\, e\in\mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{s},\underline{r},m}\, \mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f}\in\mathcal{H}^n_{d,k-1}) \tag{5.29}$$

and similarly

$$\mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}(\bar{f}^m_{e,\underline{s}};\, e\in\mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{r},\underline{s},m}\, \mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f}\in\mathcal{H}^n_{d,k-1}). \tag{5.30}$$

Note that number of index vectors $\underline{r}=(r_e)_{e\in\mathcal{H}^n_{d,k}}$ is R^D with $D:=|\mathcal{H}^n_{d,k}|$ and hence $R^D\le M$ if $C_1\gg 1$.

Writing $j':=\mathrm{index}(L'_j)$ and $J':=\mathrm{index}(L'_{j+1})$ it then follows from our inductive hypothesis functions, applied with respect to the ε_1-admissible sequence of scales

$$L_{j'+1}\ge L_{j'+2}\ge\cdots\ge L_{J'-1}$$

which is possible as $J'-j'\gg J_{k-1}(\varepsilon_1,R^D)$, that there is a scale L_j with $j'\le j<J'$ so that

$$\mathcal{N}_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}\,(g^{\underline{r}}_{\underline{s},\mathfrak{f}};\, \mathfrak{f}\in\mathcal{H}^n_{d,k-1}) = \mathcal{M}_{\lambda,Q_{\underline{s}}(L'_j)}\,(g^{\underline{r}}_{\underline{s},\mathfrak{f}};\, \mathfrak{f}\in\mathcal{H}^n_{d,k-1}) + O(\varepsilon_1) \tag{5.31}$$

for all $\lambda\in[L_{j+1},L_j]$ uniformly in $\underline{r}$ for $\underline{s}\notin S_{\varepsilon_1}$, where $S_{\varepsilon_1}\subseteq\Gamma_{L'_j,Q}$ is a set of size $|S_{\varepsilon_1}|\le\varepsilon_1|\Gamma_{L'_j,Q}|$.

Since the cubes $Q_{\underline{t}}(L_0)$ form a partition of Q as $\underline{t}$ runs through the grid $\Gamma_{L_0,Q}$ the relative density of the set S_{ε_1} can substantially increase only of a few cubes $Q_{\underline{t}}(L_0)$. Indeed, it is easy to see that $|T''_{\varepsilon_1}|\le\varepsilon_1^{1/2}|\Gamma_{L_0,Q}|$ for the set

$$T''_{\varepsilon_1}:=\{\underline{t}\in\Gamma_{L_0,Q}:\ |S_{\varepsilon_1}\cap Q_{\underline{t}}(L_0)|\ge\varepsilon_1^{1/2}\,|\Gamma_{L'_j,Q}\cap Q_{\underline{t}}(L_0)|\}.$$

We claim that (5.11) holds for $\lambda\in[L_{j+1},L_j]$ uniformly in $t\notin T_\varepsilon:=T'_\varepsilon\cup T''_{\varepsilon_1}$, $e\in\mathcal{H}^n_{d,k}$, and $1\le m\le M$. Indeed, from (7.17), (7.18), and (5.31) and the fact that $|\alpha_{\underline{s},\underline{r}}|\le 1$, it follows

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{s}}(L'_j)}\,(\bar{f}_{e,\underline{s}};\, e\in\mathcal{H}^n_{d,k}) = \mathcal{M}^d_{\lambda,Q_{\underline{s}}(L'_j)}\,(\bar{f}_{e,\underline{s}};\, e\in\mathcal{H}^n_{d,k}) + O(\varepsilon)$$

for $\underline{s}\notin S_{\varepsilon_1}\cap Q_{\underline{t}}(L_0)$ since $R^D\varepsilon_1\ll\varepsilon$. Finally, the fact that $\underline{t}\notin T''_{\varepsilon_1}$ together with localization, namely (5.21) and (5.22), ensures that averaging over $\Gamma_{L'_j,Q_{\underline{t}}(L_0)}$ gives

$$\mathcal{N}^d_{\lambda\Delta^0,Q_{\underline{t}}(L_0)}\,(\bar{f}_{e,\underline{t}};\, e\in\mathcal{H}^n_{d,k}) = \mathcal{M}^d_{\lambda,Q_{\underline{t}}(L_0)}\,(\bar{f}_{e,\underline{t}};\, e\in\mathcal{H}^n_{d,k}) + O(\varepsilon) + O(\varepsilon_1^{1/2})$$

which in light of (5.19), (5.20), and the fact that $\varepsilon_1\ll\varepsilon^2$ complete the proof. □

5.3. **Proof of Lemmas 5.1 and 5.2.**

Proof of Lemma 5.1. The argument is similar to that of Lemma 2.1. Fix an edge, say $e_0=(11,12,\dots,1k)$, and partition the edges $e\in\mathcal{H}^n_{d,k}$ in to as follows. Let $\mathcal{H}_0$ be the set of those edges e for which $1\notin\pi(e)$, and for $l=1,\dots,n_1$ let $\mathcal{H}_l$ denote the collection of edges of the form $e=(1l,j_2l_2,\dots,j_kl_k)$, in other words $e\in\mathcal{H}_l$ if $e=(1l,e')$ for some edge $e'=(j_2l_2,\dots,j_kl_k)\in\mathcal{H}^n_{d-1,k-1}$. Accordingly write

$$\prod_{e\in\mathcal{H}^n_{d,k}} f_e(\underline{x}_e) = \prod_{e\in\mathcal{H}_0} f_e(\underline{x}_e)\prod_{l=1}^{n_1}\prod_{e'\in\mathcal{H}^n_{d-1,k-1}} f_{1l,e'}(x_{1l},\underline{x}_{e'}).$$

For $x\in Q_1$ and $\underline{x}'=(\underline{x}_2,\dots,\underline{x}_d)$ with $\underline{x}_i\in Q_i^{n_i}$, define

$$g_l(x,\underline{x}') := \prod_{e'\in\mathcal{H}^n_{d-1,k-1}} f_{1l,e'}(x_{1l},\underline{x}_{e'}) \tag{5.32}$$

Then one may write

$$\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e;\, e\in\mathcal{H}^n_{d,k}) = \fint_{\underline{x}_2}\dots\fint_{\underline{x}_d}\prod_{e\in\mathcal{H}_0} f_e(\underline{x}_e)\left(\fint_{\underline{x}_1}\prod_{l=1}^{n_1} g_l(x_{1l},\underline{x}')\,d\sigma_1^\lambda(\underline{x}_1)\right)d\sigma_d^\lambda(\underline{x}_d)\dots d\sigma_2^\lambda(\underline{x}_2). \tag{5.33}$$

For the inner integrals we have, using (3.6), the estimate

$$\left(\fint_{\underline{x}_1}\prod_{l=1}^{n_1} g_l(x_{1l},\underline{x}')\,d\sigma_1^\lambda\right)^2 \le \|g_1\|^2_{U^1_L(Q)} + O(\varepsilon^{2^k}) = \fint_{y_{11}}\int_{y_{12}} g_1(y_{11})g_1(y_{12})\psi^1_L(y_{12}-y_{11})\,dy_{11}\,dy_{12} + O(\varepsilon^{2^k}).$$

provided $0 < L \ll (\varepsilon^{2^k})^6\lambda$, where as in the proof of Lemma 4.1 we use the notation

$$\psi^i_L(y_2 - y_1) = \int_t \chi^i_L(y_1 - t)\chi^i_L(y_2 - t)\, dt$$

with $\chi^i_L := L^{-n_i}1_{[-L/2,L/2]^{n_i}}$ for $1 \le i \le k$. By Cauchy-Schwarz we then have

$$\Big|\mathcal{N}^d_{\lambda\Delta^0,Q}(f_e;\, e \in \mathcal{H}^n_{d,k}\Big|^2 \le \int_{\underline{y}_1} \fint_{\underline{x}_2} \dots \fint_{\underline{x}_d} \prod_{e' \in \mathcal{H}^n_{d-1,k-1}} f_{11,e'}(x_{11}, \underline{x}_{e'}) f_{11,e'}(x_{12}, \underline{x}_{e'})\, d\sigma^\lambda_d \dots d\sigma^\lambda_2\, d\omega^1_L(\underline{y}_1) + O(\varepsilon^{2^k})$$

where $d\omega^i_L(\underline{y}_i) = |Q_i|^{-1}\psi^i_L(y_{i2} - y_{i1})\, dy_{i1}\, dy_{i2}$ with $\underline{y}_i = (y_{i1}, y_{i2}) \in Q^2_i$ for $1 \le i \le k$.

The expression we have obtained above is similar to the one in (5.6) except for the following changes. The variable $\underline{x}_1 \in Q^{n_1}_1$ is replaced by $\underline{y}_1 \in Q^2_1$ and the measure $d\sigma^\lambda_1$ by $d\omega^1_L$. The functions $f_{1l,e'}$ are replaced by $f_{11,e'}$, for $1 \le l \le n_1$, while the functions f_e for all $e \in \mathcal{H}^n_{d,k}$ such that $1 \notin \pi(e)$ are eliminated, that is replaced by 1. Repeating the same procedure for $i = 2, \dots, k$ replaces all variables $\underline{x}_i$ with variables $\underline{y}_i$ as well as the measures $d\sigma^\lambda_i$ with $d\omega^i_L$. The procedure eliminates all functions f_e when e is an edge such that $i \notin \pi(e)$ for some $1 \le i \le k$; for the remaining edges, when $\pi(e) = (1, \dots, k)$, it replaces the functions f_e with $f_{e_0} = f_{11,21,\dots,1k}$. For $k < i$ the variables $\underline{x}_i$ and the measures $d\sigma^\lambda_i$ are not changed, however integrating in these variables will have no contribution as the measures are normalized. Thus one obtains the following final estimate

$$\Big|\mathcal{N}_{\lambda\Delta^0,Q}(f_e;\, e \in \mathcal{H}^n_{d,k}\Big|^{2^k} \le \frac{1}{|Q_1|}\int_{\underline{y}_1} \dots \frac{1}{|Q_k|}\int_{\underline{y}_k} \prod_{e \in \mathcal{H}^2_{k,k}} f_{e_0}(\underline{y}_e) \prod_{i=1}^k \psi^i_L(y_{i2} - y_{i1})\, dy_{i1}\, dy_{i2} + O(\varepsilon^{2^k}) \tag{5.34}$$

noting that these integrals are not normalized. Thus, one may write the expression in (5.34), using a change of variables $y_{i1} := y_{i1} - t_i$, $y_{i2} := y_{i2} - t_i$, as

$$\frac{1}{|Q_1|}\int_{t_1} \fint_{\underline{y}_1 \in t_1 + Q_1} \dots \frac{1}{|Q_k|}\int_{t_k} \fint_{\underline{y}_k \in t_k + Q_k} \prod_{e \in \mathcal{H}^2_{k,k}} f_{e_0}(\underline{y}_e)\, d\underline{y}_1 \dots d\underline{y}_k\, d\underline{t} = \|f_{e_0}\|^{2^k}_{\square_L(Q_{\pi(e_0)})} + O(\varepsilon^{2^k}) \tag{5.35}$$

where the last equality follows from the facts that the function f_{e_0} is supported on the cube $Q_{\pi(e_0)}$ and hence the integration in $\underline{t}$ is restricted to the cube $Q + Q(L)$, giving rise an error of $O(L/l(Q))$. Estimate (5.14) follows from (5.34) and (5.35) noting that the above procedure can be applied to any $e \in \mathcal{H}^n_{d,k}$ in place of e_0. Estimate (5.15) is established similarly. □

Proof of Lemma 5.2. For $j = 0$ we set $\mathcal{B}_{e',\underline{t}}(L_0) := \{Q_{\underline{t}}(L_0), \emptyset\}$ and $\mathcal{B}_{e',\mathfrak{f}',\underline{s}}(L_0) := \{Q_{\underline{s}_{\mathfrak{f}'}}(L_0), \emptyset\}$ for $e' \in \mathcal{H}_{d,k}$, $\mathfrak{f}' \in \partial e'$, and $\underline{t}, \underline{s} \in \Gamma_{L_0,Q}$. We will develop σ-algebras $\mathcal{B}_{e',\underline{t}}(L_j)$ of scale L_j such that (5.17) holds with $\text{complex}(\mathcal{B}_{e',\mathfrak{f}',\underline{s}}(L_j)) \le j$.

We define the *total energy* of a family of functions $f^m_{e,\underline{t}}$ with respect to a family of σ-algebras $\mathcal{B}_{e',\underline{t}}(L_j)$ as

$$\mathcal{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j)) := \mathbb{E}_{\underline{t} \in \Gamma_{L_0,Q}} \sum_{m=1}^M \sum_{e \in \mathcal{H}^n_{d,k}} \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}}(L_j))\|^2_{L^2(Q_{\underline{t}_{\pi(e)}}(L_0))}. \tag{5.36}$$

Since $|f^m_{e,\underline{t}}| \le 1$ for all e, m, and $\underline{t}$ it follows that the total energy is bounded by $M \cdot |\mathcal{H}^n_{d,k}| = O(M)$. Our strategy will be to show that if (5.16) does not hold then there exist a family of σ-algebras $\mathcal{B}_{e',\underline{t}}(L_{j+2})$ such that the total energy of the family of functions $f^m_{e,\underline{t}}$ is increased by at least $c_k\varepsilon^{2^{k+3}}$ with respect to this new family of σ-algebras, and at the same time ensuring that (5.17) remains valid with $\text{complex}(\mathcal{B}_{e',\mathfrak{f}',\underline{s}}(L_{j+2})) \le j + 2$. This iterative process must stop at some $j = O(M\,\varepsilon^{-2^{k+3}})$ proving the Lemma.

Assume that we have developed σ-algebras $\mathcal{B}_{e',\underline{t}}(L_j)$ and $\mathcal{B}_{e',\mathfrak{f}',\underline{s}}(L_j)$ of scale L_j such that (5.17) holds with $\text{complex}(\mathcal{B}_{e',\mathfrak{f}',\underline{s}}(L_j)) \le j$. If (5.16) does not hold then $|T_\varepsilon| \ge \varepsilon|\Gamma_{L_0,Q}|$ for the set

$$T_\varepsilon := \{\underline{t} \in \Gamma_{L_0,Q} \,:\, \|f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}}(L_j))\|_{\square_{L_{j+1}}(Q_{\underline{t}_{\pi(e)}}(L_0))} \ge \varepsilon \ \text{ for some } e \in \mathcal{H}^n_{d,k} \text{ and } 1 \le m \le M\}.$$

Fix $\underline{t} \in T_\varepsilon$ and let $e \in \mathcal{H}^n_{d,k}$ and $1 \le m \le M$ be such that

$$\|f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}}(L_j))\|_{\square_{L_{j+1}}(Q_{\underline{t}_{\pi(e)}}(L_0))} \ge \varepsilon$$

and write $e' := \pi(e)$. Consider the partition of the cube $Q_{\underline{t}_{e'}}(L_0)$ into small cubes $Q_{\underline{s}_{e'}}(L_{j+2})$ where $\underline{s}_{e'} \in \Gamma_{L_{j+2},Q_{e'}} \cap Q_{\underline{t}_{e'}}(L_0)$. By the localization properties of the $\square_{L_{j+1}}(Q)$-norm, and the fact that $L_{j+2} \ll \varepsilon^{2^k} L_{j+1}$ we have that

$$\|f\|^{2^k}_{\square_{L_{j+1}}(Q_{\underline{t}_{e'}}(L_0))} \leq \mathbb{E}_{\underline{s}_{e'} \in \Gamma_{L_{j+2},Q_{\underline{t}_{e'}}(L_0)}} \|f\|^{2^k}_{\square(Q_{\underline{s}_{e'}}(L_{j+2}))} + \frac{\varepsilon^{2^k}}{2}$$

for any function $f : Q_{\underline{t}_{e'}}(L_0) \to [-1,1]$. Thus there exists a set $S_{\varepsilon,e,\underline{t}} \subseteq \Gamma_{L_{j+2},Q_{\underline{t}_{e'}}(L_0)}$ of size

$$|S_{\varepsilon,e,\underline{t}}| \geq \frac{\varepsilon^{2^k}}{4} |\Gamma_{L_{j+2},Q_{\underline{t}_{e'}}(L_0)}|$$

such that

$$\|f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}} | \mathcal{B}_{e',\underline{t}}(L_j))\|^{2^k}_{\square(Q_{\underline{s}_{e'}}(L_{j+2})} \geq \frac{\varepsilon^{2^k}}{4} \tag{5.37}$$

for all $\underline{s}_{e'} \in S_{\varepsilon,e,\underline{t}}$.

For a given cube Q and functions $f, g : Q \to \mathbb{R}$, define the normalized inner product of f and g as

$$\langle f, g \rangle_Q := \fint_Q f(x) g(x)\, dx.$$

Then by the well-known property of the $\square$-norm, see for example [23] or the proof of Lemma 2.2, it follows from (5.37) that there exits sets

$$B_{\mathfrak{f}',\underline{s}_{e'},\underline{t}} \subseteq Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$$

for $\mathfrak{f}' \in \partial e'$ such that

$$\Big\langle f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}} | \mathcal{B}_{e',\underline{t}}(L_j)) \,,\, \prod_{\mathfrak{f}' \in \partial e'} 1_{B_{\mathfrak{f}',\underline{s}_{e'},\underline{t}}} \Big\rangle_{Q_{\underline{s}_{e'}}(L_{j+2})} \geq \frac{\varepsilon^{2^k}}{2^{k+2}}. \tag{5.38}$$

If $\underline{s} \in \Gamma_{L_{j+2},Q}$ then there is a unique $\underline{t} = \underline{t}(\underline{s}) \in \Gamma_{L_0,Q}$ such that $\underline{s} \in Q_{\underline{t}}(L_0)$. If $\underline{t} \in T_\varepsilon$ and $\underline{s}_{e'} \in S_{\varepsilon,e,\underline{t}}$ then we define the σ-algebras $\mathcal{B}_{\mathfrak{f}',e',\underline{s}}(L_{j+2})$ on $Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$ as follows. Write $B_{\mathfrak{f}',e',\underline{s}} = B_{\mathfrak{f}',\underline{s}_{e'},\underline{t}}$ where $\underline{t} = \underline{t}(\underline{s})$ and let $\mathcal{B}_{\mathfrak{f}',e',\underline{s}}(L_{j+2})$ be the σ-algebra generated by the set $B_{\mathfrak{f}',e',\underline{s}}$ and the σ-algebra $\mathcal{B}_{\mathfrak{f}',e',\underline{s}'}(L_j)$ restricted to $Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$ where $\underline{s}' \in \Gamma_{L_j,Q}$ is the unique element so that $\underline{s} \in Q_{\underline{s}'}(L_j)$. Note that that the complexity of the σ-algebra $\mathcal{B}_{\mathfrak{f}',e',\underline{s}}(L_{j+2})$ is at most one larger then the complexity of the σ-algebra $\mathcal{B}_{\mathfrak{f}',e',\underline{s}'}(L_j)$ as restricting a σ-algebra to a set does not increase its complexity. If $\underline{t} = \underline{t}(\underline{s}) \notin T_\varepsilon$ or $\underline{s}_{e'} \notin S_{\varepsilon,e,\underline{t}}$ then let $\mathcal{B}_{\mathfrak{f}',e',\underline{s}}(L_{j+2})$ be simply the restriction of $\mathcal{B}_{\mathfrak{f}',e',\underline{s}'}(L_j)$ to the cube $Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$, or equivalently define the sets $B_{\mathfrak{f}',e',\underline{s}} := Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$. Finally, let

$$\mathcal{B}_{e',\underline{s}}(L_{j+2}) := \bigvee_{\mathfrak{f}' \in \partial e'} \mathcal{B}_{\mathfrak{f}',e',\underline{s}}(L_{j+2}) \tag{5.39}$$

be the corresponding σ-algebra on the cube $Q_{\underline{s}_{e'}}(L_{j+2})$.

Since the cubes $Q_{\underline{s}_{e'}}(L_{j+2})$ partition the cube $Q_{\underline{t}_{e'}}(L_0)$ as $\underline{s}_{e'}$ runs through the grid $\Gamma_{L_{j+2},Q_{e'}} \cap Q_{\underline{t}_{e'}}(L_0)$, these σ-algebras define a σ-algebra $\mathcal{B}_{e',\underline{t}}(L_{j+2})$ on $Q_{\underline{t}_{e'}}(L_0)$, such that its restriction to the cubes $Q_{\underline{s}_{e'}}(L_{j+2})$ is equal to the σ-algebras $\mathcal{B}_{e',\underline{s}}(L_{j+2})$.

Since the function $\prod_{\mathfrak{f}' \in \partial e'} 1_{B_{\mathfrak{f}',e',\underline{s}}}$ is measurable with respect to the σ-algebra $\mathcal{B}_{e',\underline{t}}(L_{j+2})$ restricted to the cube $Q_{\underline{s}_{e'}}(L_{j+2})$ one clearly has

$$\langle\, f^m_{e,\underline{t}} - \mathbb{E}(f^m_{e,\underline{t}} | \mathcal{B}_{e',\underline{t}}(L_{j+2})),\, \prod_{\mathfrak{f}' \in \partial e'} 1_{B_{\mathfrak{f}',e',\underline{s}}} \rangle_{Q_{\underline{s}_{e'}}(L_{j+2})} = 0. \tag{5.40}$$

and hence, by (5.38), that

$$\langle\, \mathbb{E}(f^m_{e,\underline{t}} | \mathcal{B}_{e',\underline{t}}(L_{j+2})) - \mathbb{E}(f^m_{e,\underline{t}} | \mathcal{B}_{e',\underline{t}}(L_j)),\, \prod_{\mathfrak{f}' \in \partial e'} 1_{B_{\mathfrak{f}',e',\underline{s}}} \rangle_{Q_{\underline{s}_{e'}}(L_{j+2})} \geq \frac{\varepsilon^{2^k}}{2^{k+2}}. \tag{5.41}$$

It then follows from Cauchy-Schwarz and orthogonality, using the fact that the σ-algebra $\mathcal{B}_{e',\underline{t}}(L_{j+2}))$ is a refinement of $\mathcal{B}_{e',\underline{t}}(L_{j+2})$, that

$$\begin{aligned}(5.42)\qquad \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_{j+2}))-\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j))\|^2_{L^2(Q_{\underline{s}_{e'}}(L_{j+2}))}&\\ = \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_{j+2}))\|^2_{L^2(Q_{\underline{s}_{e'}}(L_{j+2}))} - \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j))\|^2_{L^2(Q_{\underline{s}_{e'}}(L_{j+2}))}&\\ \geq \Big(\frac{\varepsilon^{2^k}}{2^{k+2}}\Big)^2&\end{aligned}$$

for $\underline{s}_{e'} \in S_{\varepsilon,e,\underline{t}}$. Since $|S_{\varepsilon,e,\underline{t}}| \geq \dfrac{\varepsilon^{2^k}}{4}|\Gamma_{L_{j+2},Q_{\underline{t}_{e'}}}(L_0)|$ averaging over $\underline{s}_{e'} \in \Gamma_{L_{j+2},Q_{\underline{t}_{e'}}}(L_0)$ implies

$$(5.43)\qquad \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_{j+2}))\|^2_{L^2(Q_{\underline{t}_{e'}}(L_0))} \geq \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j))\|^2_{L^2(Q_{\underline{t}_{e'}}(L_0))} + \frac{\varepsilon^{2^{k+2}}}{2^{2k+6}}.$$

At this point we have shown that if $\underline{t} \in T_\varepsilon$ then there exists an edge $e \in \mathcal{H}^n_{d,k}$, $1 \leq m \leq M$, and σ-algebras $\mathcal{B}_{e',\underline{t}}(L_{j+2}))$ of scale L_{j+2} on $Q_{\underline{t}_{e'}}(L_0)$, with $e' = \pi(e)$, such that (5.43) holds.

For all $e'' \in \mathcal{H}_{d,k}$ with $e'' \neq e'$ let $\mathcal{B}_{\mathfrak{f}',e'',\underline{s}}(L_{j+2})$ be the restriction of the σ-algebra $\mathcal{B}_{\mathfrak{f}',e'',\underline{s}'}(L_j)$ to the cube $Q_{\underline{s}_{\mathfrak{f}'}}(L_{j+2})$, where $\underline{s}'$ is such that $\underline{s} \in Q_{\underline{s}'}(L_j)$. By (5.39) this implies that $\mathcal{B}_{e'',\underline{s}}(L_{j+2})$ is also the restriction of $\mathcal{B}_{e'',\underline{s}'}(L_j)$ to the cube $Q_{\underline{s}_{e''}}(L_{j+2})$, and hence the σ-algebra $\mathcal{B}_{e'',\underline{t}}(L_{j+2})$ is generated by the grid $\mathcal{G}_{L_{j+2},Q_{\underline{t}_{e''}}(L_0)}$ and the σ-algebra $\mathcal{B}_{e'',\underline{t}}(L_j)$.

We have therefore defined a family of the σ-algebras $\mathcal{B}_{e',\underline{t}}(L_{j+2})$ for $e' \in \mathcal{H}_{d,k}$, satisfying

$$\sum_{m=1}^{M}\sum_{e\in\mathcal{H}^n_{d,k}} \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}}(L_{j+2}))\|^2_{L^2(Q_{\underline{t}_{\pi(e)}}(L_0))} \geq \sum_{m=1}^{M}\sum_{e'\in\mathcal{H}^n_{d,k}} \|\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}}(L_j))\|^2_{L^2(Q_{\underline{t}_{\pi(e)}}(L_0))} + \frac{\varepsilon^{2^{k+2}}}{2^{2k+6}}.$$

Using the fact that $|T_\varepsilon| \geq \varepsilon|\Gamma_{L_0,Q}|$ and averaging over $\underline{t} \in \Gamma_{L_0,Q}$ it follows using the notations of (5.36) that

$$\mathcal{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_{j+2})) \geq \mathcal{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j)) + \frac{\varepsilon^{2^{k+3}}}{2^{2k+6}}.$$

As the total energy $\mathcal{E}(f^m_{e,\underline{t}}|\mathcal{B}_{e',\underline{t}}(L_j))$ is bounded by $O(M)$, the process must stop at a step $j = O(M\,\varepsilon^{-2^{k+3}})$ where (5.16) holds for a σ-algebra of "local complexity" at most j, completing the proof of Lemma 5.2. □

6. The base case of an inductive strategy to establish Theorem 1.4

In this section we will ultimately establish the base case of our more general inductive argument. We will however start by giving a (new) proof of Theorem B′, namely the case $d = 1$ of Theorem 1.4.

6.1. **A Single Simplex in $\mathbb{Z}^n$.** Let $\Delta^0 = \{v_1 = 0, v_2, \ldots, v_{n_1}\}$ be a fixed non-degenerate simplex of n_1 points in $\mathbb{Z}^n$ with $n = 2n_1 + 3$ and define $t_{kl} := v_k \cdot v_l$ for $2 \leq k, l \leq n_1$. Recall, see [17], that a simplex $\Delta = \{m_1 = 0, \ldots, m_{n_1}\} \subseteq \mathbb{Z}^n$ is isometric to $\lambda\Delta^0$ if and only if $m_k \cdot m_l = \lambda^2 t_{kl}$ for all $2 \leq k, l \leq n_1$.

For any positive integer q and $\lambda \in q\sqrt{\mathbb{N}}$ we define $S_{\lambda\Delta^0,q}(m_2, \ldots, m_{n_1}) : \mathbb{Z}^{n(n_1-1)} \to \{0, 1\}$ be the function whose value is 1 if $m_k \cdot m_l = \lambda^2 t_{kl}$ with both m_k and m_l in $(q\mathbb{Z})^n$ for all $2 \leq k, l \leq n_1$ and is equal to 0 otherwise. It is a well-known fact in number theory, see [11] or [17], that for $n \geq 2n_1 + 1$ we have that

$$\sum_{m_2,\ldots,m_{n_1}} S_{\lambda\Delta^0,q}(m_2, \ldots, m_{n_1}) = \rho(\Delta^0)\,(\lambda/q)^{(n-n_1)(n_1-1)}(1 + O(\lambda^{-\tau}))$$

for some absolute constant $\tau > 0$ and some constant $\rho(\Delta^0) > 0$, the so-called singular series, which can be interpreted as the product of the densities of the solutions of the above system of equations among the p-adics and among the reals. Thus if we define

$$\sigma_{\lambda\Delta^0,q} := \rho(\Delta^0)^{-1}(\lambda/q)^{-(n-n_1)(n_1-1)}S_{\lambda\Delta^0,q}$$

then $\sigma_{\lambda\Delta^0,q}$ is normalized in so much that

$$\sum_{m_2,\ldots,m_{n_1}} \sigma_{\lambda\Delta^0,q}(m_2, \ldots, m_{n_1}) = 1 + O(\lambda^{-\tau})$$

for some absolute constant $\tau > 0$.

Let $Q \subseteq \mathbb{Z}^n$ be a fixed cube and let $l(Q)$ denotes its side length. For any family of functions

$$f_1, \dots, f_{n_1} : Q \to [-1, 1]$$

and $0 < \lambda \ll l(Q)$ we define the following two multi-linear expressions

$$\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1, \dots, f_{n_1}) := \mathbb{E}_{m_1 \in Q} \sum_{m_2,\dots,m_{n_1}} f_1(m_1) \dots f_{n_1}(m_{n_1})\, \sigma_{\lambda\Delta^0,q}(m_2 - m_1, \dots, m_{n_1} - m_1) \tag{6.1}$$

and

$$\mathcal{M}^1_{\lambda,q,Q}(f_1, \dots, f_{n_1}) := \mathbb{E}_{t \in Q}\, \mathbb{E}_{m_1,\dots,m_{n_1} \in t+Q(q,\lambda)}\ f_1(m_1) \dots f_{n_1}(m_{n_1}) \tag{6.2}$$

where $Q(q, \lambda) := [-\frac{\lambda}{2}, \frac{\lambda}{2}]^n \cap (q\mathbb{Z})^n$. Note that if $S \subseteq Q$ and $\mathcal{N}^1_{\lambda\Delta^0,q,Q}(1_S, \dots, 1_S) > 0$ then S must contain an isometric copy of $\lambda\Delta^0$, while if $|S| \geq \delta|Q|$ for some $\delta > 0$ then as before Hölder implies that

$$\mathcal{M}^1_{\lambda,q,Q}(1_S, \dots, 1_S) \geq \delta^n - O(\varepsilon) \tag{6.3}$$

for all scales $\lambda \in q\sqrt{\mathbb{N}}$ with $0 < \lambda \ll \varepsilon\, l(Q)$.

Recall that for any $0 < \varepsilon \ll 1$ and positive integer q we call a sequence $L_1 \geq \dots \geq L_J$ (ε, q)-*admissible* if $L_j/L_{j+1} \in \mathbb{N}$ and $L_{j+1} \ll \varepsilon^2 L_j$ for all $1 \leq j < J$ and $L_J/q \in \mathbb{N}$. Note that if $\lambda_1 \geq \dots \geq \lambda_{J'} \geq 1$ is any lacunary sequence in $q\sqrt{\mathbb{N}}$ with $J' \gg (\log \varepsilon^{-1})\, J + \log q$, one can always finds an (ε, q)-admissible sequence of scales $L_1 \geq \dots \geq L_J$ with the property that for each $1 \leq j < J$ the interval $[L_{j+1}, L_j]$ contains at least two consecutive elements from the original lacunary sequence.

In light of these observations we see that the following "counting lemma" ultimately establishes a quantitatively stronger version of Proposition B′ that appeared in Section 1.3 and hence immediately establishes Theorem 1.4 for $d = 1$.

Proposition 6.1. *Let $0 < \varepsilon \ll 1$ and $q_j := q_1(\varepsilon)^j$ for $j \geq 1$ with $q_1(\varepsilon) := \operatorname{lcm}\{1 \leq q \leq C\varepsilon^{-10}\}$.*

There exists $J_1 = O(\varepsilon^{-2})$ such that for any (ε, q_{J_1})-admissible sequence of scales $l(Q) \geq L_1 \geq \dots \geq L_{J_1}$ and $S \subseteq Q$ there is some $1 \leq j < J_1$ such that

$$\mathcal{N}^1_{\lambda\Delta^0,q_j,Q}(1_S, \dots, 1_S) = \mathcal{M}^1_{\lambda,q_j,Q}(1_S, \dots, 1_S) + O(\varepsilon) \tag{6.4}$$

for all $\lambda \in q_j\sqrt{\mathbb{N}}$ with $L_{j+1} \leq \lambda \leq L_j$.

As in the continuous setting the proof of Proposition 6.1 has two main ingredients, namely Lemmas 6.1 and 6.2 below. In these lemmas, and for the remainder Sections 6 and 7, we will continue to use the notation

$$q_1(\varepsilon) := \operatorname{lcm}\{1 \leq q \leq C\varepsilon^{-10}\}$$

for any given $\varepsilon > 0$.

Lemma 6.1 (A Generalized von Neumann inequality)**.**

Let $0 < \varepsilon \ll 1$, $q, q' \in \mathbb{N}$ with $qq_1(\varepsilon)|q'$, and $\lambda \in q\sqrt{\mathbb{N}}$ with $\lambda \ll l(Q)$ and $1 \ll L \ll \varepsilon^{10}\lambda$. For any collection of functions $f_1, \dots, f_{n_1} : Q \to [-1, 1]$ we have

$$|\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1, \dots, f_{n_1})| \leq \min_{1 \leq i \leq n_1} \|f_i\|_{U^1_{q',L}(Q)} + O(\varepsilon) \tag{6.5}$$

where for any function $f : Q \to [-1, 1]$ we define

$$\|f\|_{U^1_{q,L}(Q)} := \Big(\frac{1}{|Q|} \sum_{t \in Q} |f * \chi_{q,L}(t)|^2\Big)^{1/2} \tag{6.6}$$

with $\chi_{q,L}$ denoting the normalized characteristic function of the cubes $Q(q, L) := [-\frac{L}{2}, \frac{L}{2}]^n \cap (q\mathbb{Z})^n$.

For any cube $Q \subseteq \mathbb{Z}^n$ of side length $l(Q)$ and $q, L \in \mathbb{N}$ satisfying $q \ll L$ with L dividing $l(Q)$, we shall now partition Q into cubic grids $Q_t(q, L) = t + ((q\mathbb{Z})^n \cap Q(L))$, with $Q(L) = [-\frac{L}{2}, \frac{L}{2}]^n$ as usual. These grids form the atoms of a σ-algebra $\mathcal{G}_{q,L,Q}$. Note that if $q|q'$ and $L'|L$ then $\mathcal{G}_{q,L,Q} \subseteq \mathcal{G}_{q',L',Q}$.

Lemma 6.2 (A Koopman-von Neumann type decomposition)**.**

Let $0 < \varepsilon \ll 1$ and $q_j := q_1(\varepsilon)^j$ for all $j \geq 1$. There exists an integer $\bar{J}_1 = O(\varepsilon^{-2})$ such that any $(\varepsilon, q_{\bar{J}_1})$-admissible sequence of scales $l(Q) \geq L_1 \geq \cdots \geq L_{\bar{J}_1}$ and function $f : Q \to [-1,1]$ there is some $1 \leq j < \bar{J}_1$ such that

$$\|f - \mathbb{E}(f|\mathcal{G}_{q_j,L_j,Q})\|_{U^1_{q_{j+1},L_{j+1}}(Q)} \leq \varepsilon. \tag{6.7}$$

The reduction of Proposition 6.1 to these two lemmas is essentially identical to the analogous argument in the continuous setting as presented at the end of Section 3.1, we choose to omit the details.

Proof of Lemma 6.1. We will rely on some prior exponential sum estimates, specifically Propositions 4.2 and 4.4 in [17]. First we deal with the case $n_1 \geq 3$. By the change of variables $m_1 := m_1,\ m_i := m_i - m_1$ for $2 \leq i \leq n_1$, one may write

$$\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1,\ldots,f_{n_1}) := \mathbb{E}_{m_1 \in Q_N} \sum_{m_2,\ldots,m_{n_1}} f_1(m_1) f_2(m_1+m_2)\cdots f_{n_1}(m_1+m_{n_1})\, \sigma_{\lambda\Delta^0,q}(m_2,\ldots,m_{n_1}).$$

We now write

$$\sigma_{\lambda\Delta^0,q}(m_2,\ldots,m_{n_1}) = \sigma_{\lambda\Delta^{0\prime},q}(m_2,\ldots,m_{n_1-1})\, \sigma^{m_2,\ldots,m_{n_1-1}}_{\lambda,q}(m_{n_1})$$

where $\Delta^{0\prime} = \{v_1 = 0, v_2, \ldots, v_{n_1-1}\}$ and for each $m_2,\ldots,m_{n_1-1} \in (q\mathbb{Z})^n$ we are using $\sigma^{m_2,\ldots,m_{n_1-1}}_{\lambda,q}(m)$ denote the (essentially) normalized indicator function of the subset of $(q\mathbb{Z})^n$ that contains m if and only if $m \cdot m_k = \lambda^2 t_{kn_1}$ for all $2 \leq k \leq n_1$.

Using the fact that $|f_i| \leq 1$, together with Cauchy-Schwarz and Plancherel, one can then easily see that

$$|\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1,\ldots,f_{n_1})|^2 \leq |Q|^{-1} \int_{\xi \in \mathbb{T}^n} |\widehat{f}_{n_1}(\xi)|^2 H_{\lambda,q}(\xi)\, d\xi \tag{6.8}$$

with

$$H_{\lambda,q}(\xi) = \sum_{m_2,\ldots,m_{n_1}} \sigma_{\lambda\Delta^{0\prime},q}(m_2,\ldots,m_{n_1-1})\, |\widehat{\sigma^{m_2,\ldots,m_{n_1-1}}_{\lambda,q}}(\xi)|^2.$$

It then follows by Propositions 4.2 and 4.4 in [17], with $\delta = \varepsilon^4$ and after rescaling by q, that in addition to being non-negative and uniformly bounded in ξ we in fact have

$$H_{\lambda,q}(\xi) = O(\varepsilon) \quad \text{whenever} \quad \left|q\xi - \frac{l}{q_1(\varepsilon)}\right| \geq \frac{q}{\varepsilon^4 \lambda}, \tag{6.9}$$

for all $l \in \mathbb{Z}^n$.

We note that the expression $H_{\lambda,q}(\xi)$ may be interpreted as the Fourier transform of the indicator function of the set of integer points on a certain variety, and estimate (6.9) indicates that this concentrates near rational points of small denominator. It is this crucial fact from number theory which makes results like Theorem B′ possible.

Since

$$\widehat{\chi}_{q,L}(\xi) = \frac{q^n}{L^n} \sum_{m \in [-\frac{L}{2},\frac{L}{2})^n,\ q|m} e^{-2\pi i m\cdot\xi}$$

it is easy to see that $\widehat{\chi}_{q,L}(l/q) = 1$ for all $l \in \mathbb{Z}^n$ and that there exists some absolute constant $C > 0$ such that

$$0 \leq 1 - \widehat{\chi}_{q,L}(\xi)^2 \leq C\, L\, |\xi - l/q| \tag{6.10}$$

for all $\xi \in \mathbb{T}^n$ and $l \in \mathbb{Z}^n$. It is then easy to see using our assumption that $qq_1(\varepsilon)|q'$ that

$$0 \leq H_{\lambda,q}(\xi)(1 - \widehat{\chi}_{q',L}(\xi)^2) \leq C\varepsilon \tag{6.11}$$

for some constant $C > 0$ uniformly in $\xi \in \mathbb{T}^n$ provided $L \ll \varepsilon^5 \lambda$. Substituting inequality (6.7) into (6.8), we obtain

$$\begin{aligned} |\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1,\ldots,f_{n_1})|^2 &\leq |Q|^{-1} \left(\int |\hat{f}_{n_1}(\xi)|^2 H_\lambda(\xi) \widehat{\chi}_{q',L}(\xi)^2\, d\xi + \int |\hat{f}_{n_1}(\xi)|^2 H_\lambda(\xi)(1 - \widehat{\chi}_{q',L}(\xi)^2)\, d\xi \right) \\ &\leq \|f_{n_1}\|^2_{U^1_{q',L}(Q)} + O(\varepsilon) \end{aligned}$$

provided $L \ll \varepsilon^5\lambda$. This proves Lemma 6.1 for $k \geq 3$, as it is clear that by re-indexing the above estimate holds for any of the functions f_i in place of f_{n_1}. For $n_1 = 2$ an easy modification of arguments in [14], specifically the proof of Lemma 3 therein, establishes that

$$|\mathcal{N}^1_{\lambda\Delta^0,q,Q}(f_1,f_2)|^2 \leq \|f_i\|^2_{U^1_{q',L}(Q)} + O(\varepsilon)$$

for $i = 1, 2$ provided $L \ll \varepsilon^5\lambda$. □

Proof of Lemma 6.2. Let $q, L \in \mathbb{N}$ such that $L|N$, $q|L$. The "modulo q" grids $Q_t(q,L) = t + Q(q,L)$ partition the cube Q with t running through the set $\Gamma_{q,L,Q} = \{1,\dots,q\}^n + \Gamma_{L,Q}$, where $\Gamma_{L,Q}$ denote the centers of the "integer" grids $t + Q(L)$ in an initial partition of Q. Let q', L' be positive integers so that $q|q'$, $L'|L$ and $L' \ll \varepsilon^2 L$. If $s \in \Gamma_{q',L',Q}$ and $t \in Q_s(q',L')$ then $|t-s| = O(L')$ and hence

$$\mathbb{E}_{x\in Q_t(q,L)}g(x) = \mathbb{E}_{x\in Q_s(q,L)}g(x) + O(L'/L)$$

for any function $g : Q \to [-1,1]$. Moreover, since the cube $Q_s(q,L)$ is partitioned into the smaller cubes $Q_t(q',L')$, we have by Cauchy-Schwarz

$$|\mathbb{E}_{x\in Q_s(q,L)}\, g(x)|^2 \leq \mathbb{E}_{t\in\Gamma_{q',L',Q_s(q,L)}}|\mathbb{E}_{x\in Q_t(q',L')}g(x)|^2.$$

From this it is easy to see that

$$\|g\|^2_{U^1_{q,L}(Q)} = \mathbb{E}_{t\in Q}|\mathbb{E}_{x\in Q_t(q,L)}g(x)|^2 \leq \mathbb{E}_{t\in\Gamma_{q',L',Q}}\ |\mathbb{E}_{x\in Q_t(q',L')}g(x)|^2 + O(L'/L)$$

and we note that the right side of the above expression is $\|\mathbb{E}(g|\mathcal{G}_{q',L',Q})\|^2_{L^2(Q)}$ since the conditional expectation function $\mathbb{E}(g|\mathcal{G}_{q',L',Q})$ is constant and equal to $\mathbb{E}_{x\in Q_t(q',L')}g(x)$ on the cubes $Q_t(q',L')$.

Now suppose (6.7) does not hold for some $j \geq 1$, that is

$$\|f - \mathbb{E}(f|\mathcal{G}_{q_j,L_j,Q})\|^2_{U^1_{q_{j+1},L_{j+1}}(Q)} \geq \varepsilon^2.$$

Since $L_{j+2} \ll \varepsilon^2 L_{j+1}$, $L_{j+2}|L_j$, and $q_{j+1}|q_{j+2}$ we can apply the above observations to $g := f - \mathbb{E}(f|\mathcal{G}_{q_j,L_j,Q})$ and obtain, by orthogonality, that

$$\|\mathbb{E}(f|\mathcal{G}_{q_{j+2},L_{j+2},Q})\|^2_{L^2(Q)} \geq \|\mathbb{E}(f|\mathcal{G}_{q_j,L_j,Q})\|^2_{L^2(Q)} + c\varepsilon^2 \tag{6.12}$$

for some constant $c > 0$. Since the above expressions are clearly bounded by 1, the above procedure must stop in $O(\varepsilon^{-2})$ steps at which (6.7) must hold for some $1 \leq j \leq \bar{J}_1(\varepsilon)$ with $\bar{J}_1(\varepsilon) = O(\varepsilon^{-2})$. □

6.2. **The base case of our general inductive strategy.**

Let $Q = Q_1 \times \ldots \times Q_d$ with $Q_i \subseteq \mathbb{Z}^{2n_i+3}$ be cubes of equal side length $l(Q)$ and $\Delta^0_i \subseteq \mathbb{Z}^{2n_i+3}$ be a non-degenerate simplex of n_i points for $1 \leq i \leq d$.

We note that for any $q_0 \in \mathbb{N}$ and scale L_0 dividing $l(Q)$ if $\underline{t} = (t_1,\dots,t_d) \in \Gamma_{q_0,L_0,Q}$, then the corresponding grids $Q_{\underline{t}}(q_0,L_0)$ in the partition of Q take the form $Q_{\underline{t}}(q_0,L_0) = Q_{t_1}(q_0,L_0) \times \cdots \times Q_{t_d}(q_0,L_0)$.

As in the continuous setting we will ultimately need a parametric version of Proposition 6.1, namely Proposition 6.2 below.

Proposition 6.2 (Parametric Counting Lemma on $\mathbb{Z}^n$ for Simplices)**.** *Let $0 < \varepsilon \leq 1$ and $R \geq 1$.*

There exists an integer $J_1 = J_1(\varepsilon,R) = O(R\,\varepsilon^{-4})$ such that for any (ε, q_{J_1})-admissible sequence of scales $L_0 \geq L_1 \geq \cdots \geq L_{J_1}$ with L_0 dividing $l(Q)$ and $q_j := q_0q_1(\varepsilon)^j$ for $0 \leq j \leq J_1$ with $q_0 \in \mathbb{N}$, and collection of functions

$$f^{i,r}_{k,\underline{t}} : Q_{t_i}(q_0,L_0) \to [-1,1] \quad \text{with} \quad 1 \leq i \leq d,\ 1 \leq k \leq n_i,\ 1 \leq r \leq R \text{ and } \underline{t} \in \Gamma_{q_0,L_0,Q}$$

there exists $1 \leq j < J_1$ and a set $T_\varepsilon \subseteq \Gamma_{q_0,L_0,Q}$ of size $|T_\varepsilon| \leq \varepsilon|\Gamma_{q_0,L_0,Q}|$ such that

$$\mathcal{N}^1_{\lambda\Delta^0_i,q_j,Q_{t_i}(q_0,L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) = \mathcal{M}^1_{\lambda,q_j,Q_{t_i}(q_0,L_0)}(f^{i,r}_{1,\underline{t}},\dots,f^{i,r}_{n_i,\underline{t}}) + O(\varepsilon) \tag{6.13}$$

for all $\lambda \in q_j\sqrt{\mathbb{N}}$ with $L_{j+1} \leq \lambda \leq L_j$ and $\underline{t} \notin T_\varepsilon$ uniformly in $1 \leq i \leq d$ and $1 \leq r \leq R$.

This proposition follows, as the analogous result did in the continuous setting, from Lemma 6.1 and the follow parametric version of Lemma 6.2.

Lemma 6.3 (A simultaneous Koopman-von Neumann type decomposition)**.**

Let $0 < \varepsilon \ll 1$, $m \geq 1$, and $Q \subseteq \mathbb{Z}^n$ be a cube. There exists an integer $\bar{J}_1 = O(m\varepsilon^{-3})$ such that for any $(\varepsilon, q_{\bar{J}_1})$-admissible sequence $L_0 \geq L_1 \geq \cdots \geq L_{\bar{J}_1}$ with L_0 dividing $l(Q)$ and $q_j := q_0 q_1 (\varepsilon)^j$ for $0 \leq j \leq \bar{J}_1$ with $q_0 \in \mathbb{N}$, and collection of functions

$$f_{1,t}, \dots f_{m,t} : Q_t(q_0, L_0) \to [-1, 1]$$

defined for each $t \in \Gamma_{q_0,L_0,Q}$, there is some $1 \leq j < \bar{J}_1$ and a set $T_\varepsilon \subseteq \Gamma_{q_0,L_0,Q}$ of size $|T_\varepsilon| \leq \varepsilon |\Gamma_{q_0,L_0,Q}|$ such that

$$\|f_{i,t} - \mathbb{E}(f_{i,t}|\mathcal{G}_{q_j,L_j,Q_t(q_0,L_0)}\|_{U^1_{q_{j+1},L_{j+1}}(Q_t(q_0,L_0))} \leq \varepsilon \tag{6.14}$$

for all $1 \leq i \leq m$ and $t \notin T_\varepsilon$.

Lemma 6.3 above is of course the discrete analogue of Lemma 3.2. Since the proofs of Proposition 6.2 and Lemma 6.3 are almost identical to the arguments presented in Section 3.2 we choose to omit these details.

7. Proof of Theorem 1.4: The general case

After the preparations in Section 6 we can proceed very similarly as in Section 5 to prove our main result in the discrete case, namely Theorem 1.4. The main difference will be that given $0 < \varepsilon \ll 1$ and $1 \leq k \leq d$, we construct a positive integer $q_k(\varepsilon)$ and assume that all our sequences of scales will be $(\varepsilon, q_k(\varepsilon))$-admissible. The cubes $Q_{\underline{t}}(L)$ will be naturally now be replaced by the grids $Q_{\underline{t}}(q, L)$ of the form that already appear in Section 6 where we always assume $q|L$.

Let $\Delta^0 = \Delta^0_1 \times \ldots \times \Delta^0_d$ with each $\Delta^0_i \subseteq \mathbb{Z}^{2n_i+3}$ a non-degenerate simplex of n_i points for $1 \leq i \leq d$ and $Q = Q_1 \times \ldots \times Q_d \subseteq \mathbb{Z}^n$ with $Q_i \subseteq \mathbb{Z}^{2n_i+3}$ cubes of equal side length $l(Q)$ (taken much larger than the diameter of Δ^0). We will use the same parameterizations in terms of hypergraph bundles $\mathcal{H}^n_{d,k}$ and corresponding notations as in Section 5 to count the configurations $\Delta = \Delta_1 \times \ldots \times \Delta_d \subseteq Q$ with each $\Delta_i \subseteq Q_i$ an isometric copy of $\lambda \Delta^0_i$ for some $\lambda \in \sqrt{\mathbb{N}}$.

Given any positive integer q and $\lambda \in q\sqrt{\mathbb{N}}$ we will make use of the notation

$$\sum_{\underline{x}_i} f(\underline{x}_i)\, \sigma^i_{\lambda,q}(\underline{x}_i) := \mathbb{E}_{x_{i1} \in Q_i} \sum_{x_{i2},\dots,x_{in_i}} f(\underline{x}_i)\, \sigma_{\lambda\Delta^0_i,q}(x_{i2} - x_{i1}, \dots, x_{in_i} - x_{i1})\, dx_{i1} \tag{7.1}$$

with $\sigma_{\lambda\Delta^0_i,q}$ as defined in the previous section and $\underline{x}_i = (x_{i1}, \dots, x_{in_i}) \in Q_i^{n_i}$.

Note that if $S \subseteq Q$ then the density of configurations Δ in S, of the form $\Delta = \Delta_1 \times \ldots \times \Delta_d$ with each $\Delta_i \subseteq Q_i$ an isometric copy of $\lambda\Delta^0_i$ for some $\lambda \in q\sqrt{\mathbb{N}}$ is given by the expression

$$\mathcal{N}^d_{\lambda\Delta^0,q,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) := \sum_{\underline{x}_1} \cdots \sum_{\underline{x}_d} \prod_{e \in \mathcal{H}^n_{d,d}} 1_S(\underline{x}_e)\ \sigma^1_{\lambda,q}(\underline{x}_1) \dots \sigma^d_{\lambda,q}(\underline{x}_d). \tag{7.2}$$

More generally, for any given $1 \leq k \leq d$ and a family of functions $f_e : Q_{\pi(e)} \to [-1, 1]$ with $e \in \mathcal{H}^n_{d,k}$ we define the multi-linear expression

$$\mathcal{N}^d_{\lambda\Delta^0,q,Q}(f_e; e \in \mathcal{H}^n_{d,k}) := \sum_{\underline{x}_1} \cdots \sum_{\underline{x}_d} \prod_{e \in \mathcal{H}^n_{d,k}} f_e(\underline{x}_e)\ \sigma^1_{\lambda,q}(\underline{x}_1) \dots . \sigma^d_{\lambda,q}(\underline{x}_d). \tag{7.3}$$

as well as

$$\mathcal{M}^d_{\lambda,q,Q}(f_e; e \in \mathcal{H}^n_{d,k}) := \mathbb{E}_{\underline{t} \in Q}\ \mathcal{M}^d_{\underline{t}+Q(q,L)}\,(f_e; e \in \mathcal{H}^n_{d,k}) \tag{7.4}$$

where $Q(q, L) = Q_1(q, L) \times \cdots \times Q_d(q, L)$ with each $Q_i(q, L) = (q\mathbb{Z} \cap [-\frac{L}{2}, \frac{L}{2}])^{2n_i+3}$ and

$$\mathcal{M}^d_{\widetilde{Q}}(f_e; e \in \mathcal{H}^n_{d,k}) := \mathbb{E}_{\underline{x}_1 \in \widetilde{Q}_1^{n_1}} \cdots\ \mathbb{E}_{\underline{x}_d \in \widetilde{Q}_d^{n_d}} \prod_{e \in \mathcal{H}^n_{d,k}} f_e(\underline{x}_e) \tag{7.5}$$

for any cube $\widetilde{Q} \subseteq Q$ of the form $\widetilde{Q} = \widetilde{Q}_1 \times \cdots \times \widetilde{Q}_d$ with $\widetilde{Q}_i \subseteq Q_i$ for $1 \leq i \leq d$.

We note that it is easy to show, as in the continuous, that if $S \subseteq Q$ with $|S| \geq \delta |Q|$ for some $\delta > 0$ then

$$\mathcal{M}^d_{\lambda,q,Q}(1_S; e \in \mathcal{H}^n_{d,d}) \geq \delta^{n_1 \cdots n_d} - O(\varepsilon) \tag{7.6}$$

for all scales $\lambda \in q\sqrt{\mathbb{N}}$ with $0 < \lambda \ll \varepsilon\, l(Q)$. In light of this observation and the discussion preceding Proposition 6.1 the proof of Theorem 1.4 reduces, as it did in the continuous setting, to the following

Proposition 7.1. *Let $0 < \varepsilon \ll 1$. There exist positive integers $J_d = J_d(\varepsilon)$ and $q_d(\varepsilon)$ such that for any $(\varepsilon, q_d(\varepsilon)^{J_d})$-admissible sequence of scales $l(Q) \geq L_1 \geq \cdots \geq L_{J_1}$ and $S \subseteq Q$ there is some $1 \leq j < J_d$ such that*

$$\mathcal{N}^d_{\lambda\Delta^0,q_j,Q}(1_S\,;\, e \in \mathcal{H}^n_{d,d}) = \mathcal{M}^d_{\lambda,q_j,Q}(1_S; e \in \mathcal{H}^n_{d,d}) + O(\varepsilon), \tag{7.7}$$

for all $\lambda \in q_j\sqrt{\mathbb{N}}$ with $L_{j+1} \leq \lambda \leq L_j$ with $q_j := q_d(\varepsilon)^j$.

Quantitative Remark. A careful analysis of our proof reveals that there exist choices of $J_d(\varepsilon)$ and $q_d(\varepsilon)$ which are less than $W_d(\log(C_\Delta \varepsilon^{-3}))$ and $W_d(C_\Delta \varepsilon^{-13})$ respectively where $W_k(m)$ is again the tower-exponential function defined by $W_1(m) = \exp(m)$ and $W_{k+1}(m) = \exp(W_k(m))$ for $k \geq 1$.

The proof of Proposition 7.1 follows along the same lines as the analogous result in the continuous setting. As before we will compare the averages $\mathcal{N}^d_{\lambda\Delta^0,q,Q}(f_e; e \in \mathcal{H}^n_{d,k})$ to those of $\mathcal{M}^d_{\lambda,q,Q}(f_e; e \in \mathcal{H}^n_{d,k})$, at certain scales q and $\lambda \in q\sqrt{\mathbb{N}}$ with with $L_{j+1} \leq \lambda \leq L_j$, inductively for $1 \leq k \leq d$. As the arguments closely follow those given in Section 5 we will be brief and emphasize mainly just the additional features.

7.1. Reduction of Proposition 7.1 to a more general "local" counting lemma.

For any given $1 \leq k \leq d$ and a family of functions $f_e : Q_{\pi(e)} \to [-1,1]$ with $e \in \mathcal{H}^n_{d,k}$ it is easy to see that for any $\varepsilon > 0$, scale $L_0 > 0$ dividing the side-length $l(Q)$, and $q_0 | q$ we have

$$\mathcal{N}^d_{\lambda\Delta^0,q,Q}(f_e; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{t} \in \Gamma_{q_0,L_0,Q}}\, \mathcal{N}^d_{\lambda\Delta^0,q,Q_{\underline{t}}(q_0,L_0)}(f_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{7.8}$$

and

$$\mathcal{M}^d_{\lambda,q,Q}(f_e; e \in \mathcal{H}^n_{d,k}) = \mathbb{E}_{\underline{t} \in \Gamma_{L,Q}}\, \mathcal{M}^d_{\lambda,q,Q_{\underline{t}}(q_0,L_0)}(f_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{7.9}$$

provided $0 < \lambda \ll \varepsilon L_0$ where $f_{e,\underline{t}}$ denotes the restriction of a function f_e to the cube $Q_{\underline{t}}(q_0, L_0)$.

Thus the proof of Proposition 7.1 reduces to showing that the expressions in (7.8) and (7.9) only differ by $O(\varepsilon)$ for all scales $\lambda \in q\sqrt{\mathbb{N}}$ with $L_{j+1} \leq \lambda \leq L_j$, given an (ε, q)-admissible sequence $L_0 \geq L_1 \geq \cdots \geq L_J$, for any collection of bounded functions $f_{e,\underline{t}}$, $e \in \mathcal{H}^n_{d,k}$, $\underline{t} \in \Gamma_{q_0,L_0,Q}$. Indeed, our crucial result will be the following

Proposition 7.2 (Local Counting Lemma in $\mathbb{Z}^n$)**.** *Let $0 < \varepsilon \ll 1$ and $q_0, M \in \mathbb{N}$.*

There exist positive integers $J_k = J_k(\varepsilon, M)$ and $q_k(\varepsilon)$ such that for any (ε, q_{J_d})-admissible sequence of scales $L_0 \geq L_1 \geq \cdots \geq L_{J_1}$ with L_0 dividing $l(Q)$ and $q_j := q_0\, q_k(\varepsilon)^j$ for $j \geq 1$, and collection of functions

$$f^m_{e,\underline{t}} : Q_{t_{\pi(e)}}(q_0, L_0) :\to [-1,1] \quad \text{with} \quad e \in \mathcal{H}^n_{d,k},\ 1 \leq m \leq M \text{ and } \underline{t} \in \Gamma_{q_0,L_0,Q}$$

there exists $1 \leq j < J_k$ and a set $T_\varepsilon \subseteq \Gamma_{q_0,L_0,Q}$ of size $|T_\varepsilon| \leq \varepsilon |\Gamma_{q_0,L_0,Q}|$ such that

$$\mathcal{N}^d_{\lambda\Delta^0,q_j,Q_{\underline{t}}(q_0,L_0)}(f_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) = \mathcal{M}^d_{\lambda,q_j,Q_{\underline{t}}(q_0,L_0)}(f_{e,\underline{t}}; e \in \mathcal{H}^n_{d,k}) + O(\varepsilon) \tag{7.10}$$

for all $\lambda \in q_j\sqrt{\mathbb{N}}$ with $L_{j+1} \leq \lambda \leq L_j$ and $\underline{t} \notin T_\varepsilon$ uniformly in $e \in \mathcal{H}^n_{d,k}$ and $1 \leq m \leq M$.

Note that if $k = d$, $L_0 = l(Q)$, $q_0 = M = 1$, then $|\Gamma_{q_0,L_0,Q}| = 1$, and moreover if $f_{e,\underline{t}} = 1_S$ for all $e \in \mathcal{H}^n_{d,k}$ for a set $S \subseteq Q$, then Proposition 7.2 reduces to precisely Proposition 7.1. In fact, Proposition 7.2 is a parametric, multi-linear and simultaneous extension of Proposition 7.1 which we need in the induction step, i.e. when going from level $k-1$ to level k.

7.2. Proof of Proposition 7.2.

We will prove Proposition 7.2 by induction on $1 \leq k \leq d$.

For $k = 1$ this is basically Proposition 6.2, exactly as it was in the base case of the proof of Proposition 5.3.

For the induction step we will again need two main ingredients. The first establishes that the our multi-linear forms $\mathcal{N}^d_{\lambda\Delta^0,q,Q}(f_e; e \in \mathcal{H}^n_{d,k})$ are controlled by a box-type norm attached to scales q' and L.

Let $Q = Q_1 \times \ldots \times Q_d$ with $Q_i \subseteq \mathbb{Z}^{2n_i+3}$ be cubes of equal side length $l(Q)$ and $1 \leq k \leq d$. For any scale $0 < L \ll l(Q)$ and function $f : Q_{e'} \to [-1,1]$ with $e' \in \mathcal{H}_{d,k}$ we define its local box norm at scales q' and L by

$$\|f\|^{2^k}_{\square_{q',L}(Q_{e'})} := \mathbb{E}_{\underline{s} \in Q_{e'}} \|f\|^{2^k}_{\square(Q_{\underline{s}}(q',L))} \tag{7.11}$$

where

$$\|f\|_{\square(\widetilde{Q})}^{2^k} := \mathbb{E}_{x_{11},x_{12}\in\widetilde{Q}_1}\cdots\ \mathbb{E}_{x_{k1},x_{k2}\in\widetilde{Q}_k}\prod_{(\ell_1,\ldots,\ell_k)\in\{1,2\}^k} f(x_{1\ell_1},\ldots,x_{k\ell_k}) \tag{7.12}$$

for any cube $\widetilde{Q}$ of the form $\widetilde{Q}=\widetilde{Q}_1\times\cdots\times\widetilde{Q}_k$. We note that (7.4) and (7.5) are special cases of (7.11) and (7.12) with $k=d$, $\underline{n}=(2,\ldots,2)$, and $f_e=f$ for all $e\in\mathcal{H}^{\underline{n}}_{d,d}$.

Lemma 7.1 (A Generalized von-Neumann inequality on $\mathbb{Z}^n$)**.** *Let $1\leq k\leq d$.*

Let $0<\varepsilon\ll 1$, $q,q'\in\mathbb{N}$ with $qq_1(\varepsilon)|q'$, and $\lambda\in q\sqrt{\mathbb{N}}$ with $\lambda\ll l(Q)$ and $1\ll L\ll(\varepsilon^{2^k})^{10}\lambda$. For any collection of functions $f_e:Q_{\pi(e)}\to[-1,1]$ with $e\in\mathcal{H}^{\underline{n}}_{d,k}$ we have both

$$|\mathcal{N}^d_{\lambda\Delta^0,q,Q}(f_e;e\in\mathcal{H}^{\underline{n}}_{d,k})|\leq\min_{e\in\mathcal{H}^{\underline{n}}_{d,k}}\|f_e\|_{\square_{q',L'}(Q_{\pi(e)})}+O(\varepsilon) \tag{7.13}$$

and

$$|\mathcal{M}^d_{\lambda,q,Q}(f_e;e\in\mathcal{H}^{\underline{n}}_{d,k})|\leq\min_{e\in\mathcal{H}^{\underline{n}}_{d,k}}\|f_e\|_{\square_{q',L'}(Q_{\pi(e)})}. \tag{7.14}$$

The proof of inequalities (7.13) and (7.14) follow exactly as in the continuous case, see Lemma 5.1, using Lemma 6.1 in place of Lemma 3.1. We omit the details.

The crucial ingredient is again a parametric weak hypergraph regularity lemma, i.e. Lemma 5.2 adapted to the discrete settings. The proof is essentially the same as in the continuous case, with exception that the $\square_{L_j}$-norms are replaced by $\square_{q_j,L_j}$-norms where $q_j=q_0q^j$ is a given sequence of positive integers and $L_0\geq L_1\geq\cdots\geq L_J$ is an (ε,q_J)-admissible sequence of scales. To state it we say that a σ-algebra $\mathcal{B}$ on a cube Q is of *scale* (q,L) if it is refinement of the grid $\mathcal{G}_{q,L,Q}$, i.e. if its atoms partition each cube $Q_{\underline{t}}(q,L)$ of the grid. We will always assume that $q|L$ and $L|l(Q)$. Recall also that we say the complexity of a σ-algebra $\mathcal{B}$ is at most m, and write $\mathrm{complex}(\mathcal{B})\leq m$, if it is generated by m sets.

Lemma 7.2 (Parametric weak hypergraph regularity lemma for $\mathbb{Z}^n$)**.**

Let $0<\varepsilon\ll 1$, $1\leq k\leq d$, $q_0,q,L_0,M\in\mathbb{N}$, and let $q_j:=q_0q^j$ for $j\geq 1$. There exists $\bar{J}_k=O(M\varepsilon^{-2^{k+3}})$ such that for any $(\varepsilon^{2^k},q_{\bar{J}_k})$-admissible sequence $L_0\geq L_1\geq\cdots\geq L_{\bar{J}_k}$ with the property that L_0 divides $l(Q)$ and collection of functions

$$f^m_{e,\underline{t}}:Q_{\underline{t}_{\pi(e)}}(q_0,L_0)\to[-1,1]\quad\text{with}\quad e\in\mathcal{H}^{\underline{n}}_{d,k},\ 1\leq m\leq M,\ \text{and}\ \underline{t}\in\Gamma_{q_0,L_0,Q}$$

there is some $1\leq j<\bar{J}_k$ and σ-algebras $\mathcal{B}_{e',\underline{t}}$ of scale (q_j,L_j) on $Q_{\underline{t}_{e'}}(q_0,L_0)$ for each $\underline{t}\in\Gamma_{q_0,L_0,Q}$ and $e'\in\mathcal{H}_{d,k}$ such that

$$\|f^m_{e,\underline{t}}-\mathbb{E}(f^m_{e,\underline{t}}|\mathcal{B}_{\pi(e),\underline{t}})\|_{\square_{q_{j+1},L_{j+1}}(Q_{\underline{t}_{\pi(e)}}(L_0))}\leq\varepsilon \tag{7.15}$$

uniformly for all $t\notin T_\varepsilon$, $e\in\mathcal{H}^{\underline{n}}_{d,k}$, and $1\leq m\leq M$, where $T_\varepsilon\subseteq\Gamma_{q_0,L_0,Q}$ with $|T_\varepsilon|\leq\varepsilon|\Gamma_{q_0,L_0,Q}|$.

Moreover, the σ-algebras $\mathcal{B}_{e',\underline{t}}$ have the additional local structure that the exist σ-algebras $\mathcal{B}_{e',\mathfrak{f}',\underline{s}}$ on $Q_{\underline{s}_{\mathfrak{f}'}}(q_j,L_j)$ with $\mathrm{complex}(\mathcal{B}_{e',\mathfrak{f}',\underline{s}})=O(j)$ for each $\underline{s}\in\Gamma_{q_j,L_j,Q}$, $e'\in\mathcal{H}_{d,k}$, and $\mathfrak{f}'\in\partial e'$ such that if $\underline{s}\in Q_{\underline{t}}(q_0,L_0)$, then

$$\mathcal{B}_{e',\underline{t}}\big|_{Q_{\underline{s}_{e'}}(q_j,L_j)}=\bigvee_{\mathfrak{f}'\in\partial e'}\mathcal{B}_{e',\mathfrak{f}',\underline{s}}. \tag{7.16}$$

The proof of Lemma 7.2 follows exactly as the corresponding proof of Lemma 5.2 in the continuous setting, so we will omit the details. We will however provide some details of how one deduces Proposition 7.2, from Lemmas 7.1 and 7.2. The arguments are again very similar to those in the continuous setting, however one needs to make a careful choice of the integers $q_k(\varepsilon)$, appearing in the statement of the Proposition.

Proof of Proposition 7.2. Let $2\leq k\leq d$ and assume that the lemma holds for $k-1$.

Let $0<\varepsilon\ll 1$ and $\varepsilon_1:=\exp(-C_1\varepsilon^{-2^{k+3}})$ for some large constant $C_1=C_1(n,k,d)\gg 1$.

We then define $q_k(\varepsilon):=q_{k-1}(\varepsilon_1)$ recalling that $q_1(\varepsilon):=\mathrm{lcm}\{1\leq q\leq C\varepsilon^{-10}\}$ and note that it is easy to see by induction that $q_k(\varepsilon)|q_k(\varepsilon')$ for $0<\varepsilon'\leq\varepsilon$ and $q_{k-1}(\varepsilon)|q_k(\varepsilon)$. We further define the function $F(\varepsilon):=J_{k-1}(\varepsilon_1,M)$ with $M=\varepsilon\,\varepsilon_1^{-1}$ and recall that $q_j:=q_0\,q_k(\varepsilon)^j$ for $j\geq 1$.

We now proceed exactly as in the proof of Proposition 5.3 but with $\{L_j\}_{j\geq 1}$ being a $(\varepsilon_1, q_{\widetilde{J}})$-admissible sequence of scales, with $\widetilde{J} \gg F(\varepsilon)\,\bar{J}_k(\varepsilon, M)$. We again choose a subsequence $\{L'_j\} \subseteq \{L_j\}$ so that $L'_0 = L_0$ and $\mathrm{index}(L'_{j+1}) \geq \mathrm{index}(L'_j) + F(\varepsilon) + 2$, but also now set $q'_j = q_{j'}$, where $j' := \mathrm{index}(L'_j)$. Lemma 7.2 then guarantees the existence of σ-algebras $\mathcal{B}_{e',\underline{t}}$ of scale (q'_j, L'_j) on $Q_{\underline{t}_{e'}}(q_0, L_0)$ for each $\underline{t} \in \Gamma_{q_0,L_0,Q}$ and $e' \in \mathcal{H}^n_{d,k}$, with the local structure described above, such that (7.15) holds uniformly for all $t \notin T'_\varepsilon$, $e \in \mathcal{H}^n_{d,k}$, and $1 \leq m \leq M$, for some $1 \leq j < \bar{J}_k(\varepsilon, M) = O(M\varepsilon^{-2^{k+3}})$, where $T'_\varepsilon \subseteq \Gamma_{q_0,L_0,Q}$ with $|T'_\varepsilon| \leq \varepsilon|\Gamma_{q_0,L_0,Q}|$.

Arguing as in the proof of Proposition 5.3 we can conclude from this that for each $j' \leq l < J'$ we have

$$\mathcal{N}^d_{\lambda\Delta^0,q_l,Q_{\underline{s}}(q'_j,L'_j)}(f^m_{e,\underline{s}};\, e \in \mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{s},\underline{r},m}\; \mathcal{N}^d_{\lambda\Delta^0,q_l,Q_{\underline{s}}(q'_j,L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f} \in \mathcal{H}^n_{d,k-1}) + O(\varepsilon) \tag{7.17}$$

and

$$\mathcal{M}^d_{\lambda,q_l,Q_{\underline{s}}(q'_j,L'_j)}(f^m_{e,\underline{s}};\, e \in \mathcal{H}^n_{d,k}) = \sum_{\underline{r}} \alpha_{\underline{r},\underline{s},m}\; \mathcal{M}^d_{\lambda,q_l,Q_{\underline{s}}(q'_j,L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f} \in \mathcal{H}^n_{d,k-1}) + O(\varepsilon) \tag{7.18}$$

provided $(\varepsilon^{-2^k})^{10} L'_{j+1} \ll \lambda$ with $\lambda \in q_l\sqrt{\mathbb{N}}$, where each $|\alpha_{\underline{s},r_e}| \leq 1$ and number of index vectors $\underline{r} = (r_e)_{e\in\mathcal{H}^n_{d,k}}$ is R^D with $D := |\mathcal{H}^n_{d,k}|$ and hence $R^D \leq M$ if $C_1 \gg 1$.

By induction, we apply Proposition 7.2 to the sequence of scales $L'_j = L_{j'} \geq L_{j'+1} \geq \cdots \geq L_{J'} = L'_{j+1}$ with $\varepsilon_1 > 0$ and for $q_l := q'_j\, q_k(\varepsilon)^{l-j'} = q_{j'}\, q_{k-1}(\varepsilon_1)^{l-j'}$ where $j' \leq l \leq J'$ with respect to the family of functions $g^{\underline{r}}_{\underline{s},\mathfrak{f}} : Q_{\underline{s}_\mathfrak{f}}(q'_j, L'_j) \to [-1,1]$. This is possible as $J' - j' \gg J_{k-1}(\varepsilon_1, R^D)$ and our sequence of scales is $(\varepsilon_1, q_{J'})$-admissible. Thus there exists an index $j' \leq l < J'$ such that for all $\lambda \in q_l\sqrt{\mathbb{N}}$ with $L_{l+1} \leq \lambda \leq L_l$ we have

$$\mathcal{N}^d_{\lambda\Delta^0,q_l,Q_{\underline{s}}(q'_j,L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f} \in \mathcal{H}^n_{d,k-1}) = \mathcal{M}^d_{\lambda,q_l,Q_{\underline{s}}(q'_j,L'_j)}\,(g^{\underline{r}}_{\mathfrak{f},\underline{s}};\, \mathfrak{f} \in \mathcal{H}^n_{d,k-1}) + O(\varepsilon_1) \tag{7.19}$$

uniformly in $\underline{r}$ for $\underline{s} \notin S_{\varepsilon_1}$, where $S_{\varepsilon_1} \subseteq \Gamma_{q'_j,L'_j,Q}$ is a set of size $|S_{\varepsilon_1}| \leq \varepsilon_1|\Gamma_{q'_j,L'_j,Q}|$.

The remainder of the proof follows as just as it did for Proposition 5.3. □

8. Appendix: A short direct proof of Part (i) of Theorem B′

We conclude by providing a short direct proof of Part (i) of Theorem B′, namely the following

Theorem 8.1 (Magyar [17])**.** *Let $0 < \delta \leq 1$ and $\Delta \subseteq \mathbb{Z}^{2k+3}$ be a non-degenerate simplex of k points.*

If $S \subseteq \mathbb{Z}^{2k+3}$ has upper Banach density at least δ, then there exists an integer $q_0 = q_0(\delta)$ and $\lambda_0 = \lambda_0(S,\Delta)$ such that S contains an isometric copy of $q_0\lambda\Delta$ for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \geq \lambda_0$.

For any $\varepsilon > 0$ we define

$$q_\varepsilon := \mathrm{lcm}\{1 \leq q \leq C\varepsilon^{-10}\}$$

with $C > 0$ a (sufficiently) large absolute constant. Following [14] we further define $S \subseteq \mathbb{Z}^n$ to be *ε-uniformly distributed (modulo q_ε)* if its *relative* upper Banach density on any residue class modulo q_ε never exceeds $(1+\varepsilon^2)$ times its density on $\mathbb{Z}^n$, namely if

$$\delta^*(S \,|\, s + (q_\varepsilon\mathbb{Z})^d) \leq (1+\varepsilon^2)\,\delta^*(S)$$

for all $s \in \{1,\ldots,q_\varepsilon\}^d$. It turns out that this notion is closely related to the $U^1_{q,L}(Q)$-norm introduced in Section 6. Recall that for any cube $Q \subseteq \mathbb{Z}^n$ and function $f : Q \to [-1,1]$ we define

$$\|f\|_{U^1_{q,L}(Q)} := \Big(\frac{1}{|Q|}\sum_{t\in Q} |f * \chi_{q,L}(t)|^2\Big)^{1/2} \tag{8.1}$$

with $\chi_{q,L}$ denoting the normalized characteristic function of the cubes $Q(q,L) := [-\frac{L}{2}, \frac{L}{2}]^n \cap (q\mathbb{Z})^n$. Note that the $U^1_{q,L}(Q)$-norm measures the mean square oscillation of a function with respect to cubic grids of size L and gap q.

The following observation from [14] (specifically Lemmas 1 and 2) is key to our short proof of Theorem 8.1.

Lemma 8.1. *Let $\varepsilon > 0$. If $S \subseteq \mathbb{Z}^n$ be ε-uniformly distributed with $\delta := \delta^*(S) > 0$, then there exists an integer $L = L(S, \varepsilon) > 0$ and cubes Q of arbitrarily large side length $l(Q)$ with $l(Q) \gg \varepsilon^{-4}L$ such that*

$$\|1_S - \delta 1_Q\|_{U^1_{q_\varepsilon, L}(Q)} = O(\varepsilon).$$

Let $\Delta^0 = \{v_1 = 0, v_2, \dots, v_k\}$ be a fixed non-degenerate simplex of k points in $\mathbb{Z}^n$ with $n = 2k + 3$ and define $t_{ij} := v_i \cdot v_j$ for $2 \leq i, j \leq k$. We now define a function which counts isometric copies of $\lambda\Delta^0$.

Recall, see [17], that a simplex $\Delta = \{m_1 = 0, \dots, m_k\} \subseteq \mathbb{Z}^n$ is isometric to $\lambda\Delta^0$ if and only if $m_i \cdot m_j = \lambda^2 t_{ij}$ for all $2 \leq i, j \leq k$. For any $\lambda \in \sqrt{\mathbb{N}}$ we define $S_{\lambda\Delta^0}(m_2, \dots, m_k) : \mathbb{Z}^{n(k-1)} \to \{0, 1\}$ be the function whose value is 1 if $m_i \cdot m_j = \lambda^2 t_{ij}$ for all $2 \leq i, j \leq k$ and is equal to 0 otherwise. It is a well-known fact in number theory, see [11] or [17], that for $n \geq 2k + 1$ we have that

$$\sum_{m_2, \dots, m_k} S_{\lambda\Delta^0}(m_2, \dots, m_k) = \rho(\Delta^0)\, \lambda^{(n-k)(k-1)}(1 + O(\lambda^{-\tau}))$$

for some absolute constant $\tau > 0$ and constant $\rho(\Delta^0) > 0$, the so-called singular series, which can be interpreted as the product of the densities of the solutions of the above system of equations among the p-adics and among the reals. Thus if we define

$$\sigma_{\lambda\Delta^0} := \rho(\Delta^0)^{-1} \lambda^{-(n-k)(k-1)} S_{\lambda\Delta^0}$$

then $\sigma_{\lambda\Delta^0}$ is normalized in so much that

$$\sum_{m_2, \dots, m_k} \sigma_{\lambda\Delta^0}(m_2, \dots, m_k) = 1 + O(\lambda^{-\tau})$$

for some absolute constant $\tau > 0$.

Let $Q \subseteq \mathbb{Z}^n$ be a fixed cube and let $l(Q)$ denotes its side length. For any family of functions

$$f_1, \dots, f_k : Q \to [-1, 1]$$

and $0 < \lambda \ll l(Q)$ we define

$$\mathcal{N}^1_{\lambda\Delta^0, Q}(f_1, \dots, f_k) := \mathbb{E}_{m_1 \in Q} \sum_{m_2, \dots, m_k} f_1(m_1) \dots f_k(m_k)\, \sigma_{\lambda\Delta^0}(m_2 - m_1, \dots, m_k - m_1). \tag{8.2}$$

It is clear that if $f_1 = \dots = f_k = 1_S$ restricted to Q, then the above expression is a normalized count of the isometric copies of $\lambda\Delta^0$ in $S \cap Q$. Thus, Theorem 8.1 will follow from Lemma 8.1 and the following special case (with $q = 1$) of Lemma 6.1.

Lemma 8.2 (A Generalized von Neumann inequality)**.** *Let $0 < \varepsilon \ll 1$.*

If $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \ll l(Q)$ and $1 \ll L \ll \varepsilon^{10}\lambda$ then for any collection of functions $f_1, \dots, f_k : Q \to [-1, 1]$ we have

$$|\mathcal{N}^1_{\lambda\Delta^0, Q}(f_1, \dots, f_k)| \leq \min_{1 \leq j \leq k} \|f_j\|_{U^1_{q_\varepsilon, L}(Q)} + O(\varepsilon). \tag{8.3}$$

This compares with the purely number theoretic fact that the number of simplices $\Delta = \{v_1 = 0, v_2, \dots, v_k\} \subseteq \mathbb{Z}^n$ isometric to $\lambda\Delta^0$ is asymptotic to $\rho(\Delta^0)\, \lambda^{(n-k)(k-1)}$. Thus, under the same conditions as in Lemma 8.2, we have

$$\mathcal{N}^1_{\lambda\Delta^0, Q}(1_Q, \dots, 1_Q) = 1 + O(\lambda^{-\tau}) + O(\varepsilon) \tag{8.4}$$

provided one also has $\lambda \ll \varepsilon l(Q)$.

Proof of Theorem 8.1. Let $0 < \varepsilon \ll \delta^k$ and $S \subseteq \mathbb{Z}^n$ be a set of upper Banach density δ.

We assume first that S is ε-uniformly distributed. Select a scale $L = L(\varepsilon, S)$ and a sufficiently large cube Q so that the conclusion of Lemma 8.1 holds. For a given $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda \ll \varepsilon l(Q)$ and $L \ll \varepsilon^{10}\lambda$ write $1_S = \delta 1_Q + g$ and substitute this decomposition into the multi-linear expression $\mathcal{N}^1_{\lambda\Delta^0, Q}(1_S, \dots, 1_S)$. Then by Lemma 8.2 and (8.3)-(8.4), we have that

$$\mathcal{N}^1_{\lambda\Delta^0, Q}(1_S, \dots, 1_S) \geq \delta^k - O(\varepsilon) \tag{8.5}$$

and we can conclude that S must contain an isometric copy of $\lambda\Delta^0$.

If S is not ε-uniformly distributed, then its upper Banach density is increased to at least $\delta_1 := (1+\varepsilon^2)\delta$ when restricted to a residue class $s+(q_\varepsilon\mathbb{Z})^n$. Identify $s+(q_\varepsilon\mathbb{Z})^n$ with $\mathbb{Z}^n$ and simultaneously the set $S|_{s+(q_\varepsilon\mathbb{Z})^n}$ with a set $S_1 \subseteq \mathbb{Z}^n$, via the map $y \to q_\varepsilon^{-1}(y-s)$. Note that if S_1 is ε-uniformly distributed then it contains an isometric copy of $\lambda\Delta^0$ for all sufficiently large $\lambda \in \sqrt{\mathbb{N}}$ and hence S contains an isometric copy of $q_\varepsilon\lambda\Delta^0$.

Repeating the above procedure one arrives to a set $S_j = q_\varepsilon^{-j}(S-s_j) \subseteq \mathbb{Z}^n$ for some $s_j \in \mathbb{Z}^n$ in $j = O(\log \varepsilon^{-1})$ steps which contains an isometric copy of $\lambda\Delta^0$ for all sufficiently large $\lambda \in \sqrt{\mathbb{N}}$. □

References

[1] J. Bourgain, *A Szemerédi type theorem for sets of positive density in $\mathbb{R}^k$*, Israel J. Math. 54 (1986), no. 3, 307–316.
[2] D. Conlon, J. Fox, Y. Zhao, *A relative Szemeredi theorem*, Geometric and Functional Analysis 25.3 (2015): 733-762.
[3] B. Cook, Á. Magyar, M. Pramanik, *A Roth-type theorem for dense subsets of $\mathbb{R}^d$*, Bull. Lond. Math. Soc. 49 (2017), no. 4, 676-689.
[4] P. Durcik, V. Kovač, *Boxes, extended boxes, and sets of positive upper density in the Euclidean space*, arXiv 1809.08692
[5] H. Furstenberg, Y. Katznelson, *An ergodic Szemerédi theorem for commuting trnasformations*, J. Analyse Math. 31 (1978), 275-291
[6] H. Furstenberg, Y. Katznelson and B. Weiss, *Ergodic theory and configurations in sets of positive density*, Mathematics of Ramsey theory, 184–198, Algorithms Combin., 5, Springer, Berlin, 1990.
[7] A. Frieze, R. Kannan, *The regularity lemma and approximation schemes for dense problems*, In Foundations of Computer Science (1996) Proc. 37th Annual Symp. IEEE., 12-20
[8] W. T. Gowers. *Hypergraph regularity and the multidimensional Szemerdi theorem* Annals of Mathematics (2007), 897-946.
[9] R.L. Graham. *Recent trends in Euclidean Ramsey theory*, Discrete Mathematics 136, no. 1-3 (1994), 119-127.
[10] A. Iosevich and M. Rudnev, *Erdős distance problem in vector spaces over finite fields*, Trans. Amer. Math. Soc. 359, no. 12 (2007), 6127-6142.
[11] Y.Kitaoka, *Siegel modular forms and representation by quadratic forms* Lectures on Mathe- matics and Physics, Tata Institute of Fundamental Research, Springer-Verlag, (1986)
[12] N. Lyall and Á. Magyar, *Product of simplices and sets of positive upper density in $\mathbb{R}^d$*, Math. Proc. of the Cambridge Philos. Soc. 165. no. 1. (2018), 25-51
[13] N. Lyall and Á. Magyar, *Distance Graphs and sets of positive upper density in $\mathbb{R}^d$*, to appear in Anal. PDE
[14] N. Lyall and Á. Magyar, *Distances and trees in dense subsets of $\mathbb{Z}^d$*, arXiv 1509.09298
[15] N. Lyall, Á. Magyar, H. Parshall, *Spherical configurations over finite fields*, to appear in Amer. J. Math.
[16] Á. Magyar, *Distance sets of large sets of integer points*, Israel J. Math., v (2008) pp.
[17] Á. Magyar, *k-point configurations in sets of positive density of Z^n*, Duke Math. J., v 146/1, (2009) pp. 1-34.
[18] H. Parshall, *Simplices over finite fields*, Proc. Amer. Math. Soc. 145.6 (2017), 2323-2334.
[19] C. L. Siegel, *On the theory of indefinite quadratic forms*, Ann. of Math. (2) 45 (1944), 577 622
[20] E., Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. 27 (1975), 199-245.
[21] T. Tao, *Szemerédi's regularity lemma revisited*, arXiv preprint math/0504472 (2005)
[22] T. Tao, *A variant of the hypergraph removal lemma*, Journal of Combinatorial Theory, Series A 113.7 (2006): 1257-1280
[23] T. Tao, *The ergodic and combinatorial approaches to Szemerdi's theorem*, Additive combinatorics, 145–193, CRM Proc. Lecture Notes, 43, Amer. Math. Soc., Providence, RI, 2007.
[24] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, 105. Cambridge University Press, Cambridge, 2006. xviii+512 pp.
[25] T. Ziegler, *Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$*, J. Anal. Math. 99 (2006), 249-266.

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: lyall@math.uga.edu

E-mail address: magyar@math.uga.edu