SIMPLEX–CENTER CONFIGURATIONS IN DENSE SUBSETS OF EUCLIDEAN SPACES AND THE INTEGER LATTICE

Abstract. We obtain density Ramsey theorems for the configuration consisting of the vertices of a simplex Δ together with their barycenter. We prove that any subset $A \subseteq \mathbb{R}^n$ of positive upper density contains an isometric copy of all sufficiently large dilates of Δ together with its barycenter. Since this configuration is non-spherical such results are not possible with respect the quadratic Euclidean metric, hence consider general metrics ρ defined by a positive-definite, convex, non-singular homogeneous forms of even degree at least four. We prove the analogues result in the discrete setting, for subsets A of the integer lattice $\mathbb{Z}^n$, under some natural and necessary congruence restrictions on the scales λ at which the set A can contain an isometric copy of the simplex.

1. Introduction.

In the 1970's Erdős at al. [5] initiated the study of geometric or Euclidean Ramsey theory. They studied finite point configurations F such that for every r-coloring of a sufficiently large dimensional Euclidean space there is at least one color class which contains an isometric copy of F. Later Bourgain [2] has obtained a natural density analogue for non-degenerate simplices $\Delta = \{w_1, \ldots, w_m\}$. He showed that any positive density set $A \subseteq \mathbb{R}^m$ necessarily contains an isometric copy of its dilates $\lambda\Delta$, for sufficiently large scales $\lambda \geq \lambda(A, \Delta)$. An example of Graham [8] shows that in both the coloring and in the density case positive results are only possible when the configuration is spherical, i.e. can be inscribed in a sphere.

However, it was shown in [3] that such results are still possible for certain non-spherical configurations if one changes the underlying metric of the ambient space. In particular it was shown that a positive density set $A \subseteq \mathbb{R}^n$ contains a three term progression $x, x+y, x+2y$ with $|y|_p = \lambda$ for all $\lambda \geq \lambda(A)$, for $1 < p < \infty,\ p \neq 2$, at least in large dimensions $n \geq n(p)$. Here $|y|_p = (\sum_{i=1}^n |y_i|^p)^{1/p}$ denotes the ℓ^p-metric of the ambient space. Our first first result is an extension to configurations consisting of the vertices of an m-simplex $\Delta = \{w_1, \ldots, w_m\}$ together with its barycenter with respect to metrics generated by certain positive, convex, homogeneous forms. To be more precise, let

$$\Delta_o = (w_1, \ldots, w_m), \qquad \sum_{i=1}^m w_i = 0, \tag{1.1}$$

be a fixed simplex in $\mathbb{R}^n$ centered at the origin. We will denote by

$$E := \binom{m}{2}, \tag{1.2}$$

the set of its edges.

Let $Q \in \mathbb{Z}[x_1, \ldots, x_n]$ be homogeneous of even degree $r \geq 4$. Throughout we assume that Q is positive-definite and convex, so that

$$\rho(x) := Q(x)^{1/r}$$

is a norm, thus it defines a metric in on the underlying Euclidean space $\mathbb{R}^n$. We will also assume that Q is non-singular in the sense that

$$\nabla Q(x) = 0 \quad (x \in \mathbb{C}^n) \qquad \Longrightarrow \qquad x = 0. \tag{1.3}$$

Date: August 2026.

and define

$$\ell'_{ij} := \rho(w_i - w_j) \quad \text{and} \quad \ell_{ij} = Q(w_i - w_j) \qquad (1 \le i < j \le m). \tag{1.4}$$

Note that $\ell'_{ij} = \ell_{ij}^{1/r}$ is the length of the edge of the simplex connecting the points w_i and w_j, however it will be more convenient to work with the quantities ℓ_{ij}.

We refer to a configuration

$$z, \quad z + v_1, \dots, z + v_m, \qquad \sum_{i=1}^{m} v_i = 0, \tag{1.5}$$

as a simplex–center pattern, note that z is the barycenter of the other m points. Prescribing

$$Q(v_i - v_j) = \lambda^r \ell_{ij} \qquad (i < j) \tag{1.6}$$

makes the simplex $\Delta(z) = \{z + v_1, \dots, z + v_m\}$ ρ-isometric to $\lambda\Delta$. This pattern is genuinely non-spherical in the ordinary Euclidean sense as the barycenter z lies in the interior of the convex hull of the vertices.

Write

$$H_m^n := \left\{ (v_1, \dots, v_m) \in (\mathbb{R}^n)^m : \sum_{i=1}^{m} v_i = 0 \right\}, \qquad D := (m-1)n, \tag{1.7}$$

and introduce the edge map

$$\Phi : H_m^n \longrightarrow \mathbb{R}^E, \qquad \Phi(v) := \big(Q(v_i - v_j)\big)_{i<j}. \tag{1.8}$$

We define the notion of a Q-regular simplex,

Definition 1.1 (Regular simplex)**.** The centered simplex Δ_o is *Q-regular* if

$$\operatorname{rank}\left(Jac\,\Phi(\Delta_o)\right) = E. \tag{1.9}$$

This means that the map Φ has maximal rank at the simplex $\Delta = (w_1, \dots, w_m) \in (\mathbb{R}^n)^m$. Bourgain's theorem [2] treats simplices which are non-degenerate in the sense that their vertices are affinely independent. For the standard Euclidean metric and for triangles this the rank condition is equivalent to the affine independency of the vertices, but for metrics arising from higher degree forms the latter is not sufficient for all barycentric configurations, see Section 2.

1.1. **The Euclidean setting.**

Recall the upper Banach density of a measurable set $A \subset \mathbb{R}^n$

$$\overline{d}(A) := \limsup_{N\to\infty} \sup_{x\in\mathbb{R}^n} \frac{|A \cap (x + [0,N]^n)|}{N^n}.$$

Our first result is the following.

Theorem 1.2. *Let Q satisfy the preceding hypotheses and let Δ be an affinely independent, Q-regular simplex. Suppose*

$$n > E(r-1)2^{r-1}. \tag{1.10}$$

If $A \subset \mathbb{R}^n$ is measurable and $\overline{d}(A) > 0$, there is $\lambda_0 = \lambda_0(A, Q, \Delta)$ such that, for every $\lambda \ge \lambda_0$, the set A contains a ρ-isometric copy of $\lambda\Delta$ together with its barycenter, i.e. a configuration of the form (1.5) *satisfying* (1.6).

Note that the exponential dependence of the dimension n on the degree r is a consequence of the generality of the metric $\rho = Q^{1/r}$ for diagonal forms e.g. for the l^r-metric the dimension depending only quadratically on r using the best available results on the Waring problem [14].

The proof of Theorem 1.2 follows broadly the approach of [3]. It is based on an oscillatory representation of Leray measure of the configuration space and a decomposition of the range of integration into a small, an intermediate and a large region. Correspondingly the counting function of isometric copies of a scaled simplex in the set A is decomposed in to three parts; a main part, an intermediate part, and a high frequency error term. The main term can be estimated below via standard results from additive combinatorics. The error term is controlled by the Gowers U^3-uniformity norm of the phase function e^{itQ}, whose approximating Riemann sum is an exponential sumcwhich is estimated via Weyl-differencing [1]. The intermediate terms are treated via estimates from time-frequency analysis due to Muscalu, Tao and Thiele [11].

1.2. **The discrete setting.** Our main results concern subsets of the integer lattice $\mathbb{Z}^n$. We prove an analogue of Theorem 1.2, as well as a quantitative result about the existence of a single ρ-isometric copy in a finite $A \subseteq [N]^n$ of density $\delta > 0$. Both results can be viewed as extensions of the main results of [10], from simplices to simplex-center configurations with respect to general metrics ρ, generated by positive-definite, convex, non-singular forms.

Here we assume that $\Delta \subseteq \mathbb{Z}^n$ and $Q : \mathbb{Z}^n \to \mathbb{Z}$ is an integral form. This restricts the "levels" $L = \lambda^r$ to rational numbers with a fixed denominator depending only on Δ_o. Indeed, a simplex $\Delta = \{v_1, \ldots, v_m\}$ is ρ-isometric to $\lambda\Delta_o$ if $Q(v_i - v_j) = L\ell_{iJ}$ for all $1 \le i < j \le m$; we will only consider levels $L \in \mathbb{N}$. Another restriction on the levels is due to the existence of grids. Fixing a density $\delta > 0$ the grid $A = (q\mathbb{Z})^n$ has density at least $\delta > 0$ if $q \le \delta^{-1/n}$, as $q^r \mid Q(v_i - v_j)$ for $v_i, v_j \in A$. Thus a discrete analogue of Theorem 1.2 can only hold for those levels L which are integer multiples of $q_0(\delta) := lcm\{1 \le q \le \delta^{-1/n}\}^r$, as it have to be valid for all such grids.

In order to find integer solutions $v = (v_1, \ldots, v_m) \in (\mathbb{Z}^n)^m$ of the diophantine system $Q(v_i - v_j) = L\ell_{ij}$, $(1 \le i < j \le m)$ one needs make some assumptions on the map $\Phi : (\mathbb{Z}^n)^m \to \mathbb{Z}^{m(m-1)/2}$. In order to find solutions $v = (v_1, \ldots, v_m) \in A^m$ for any set $A \subseteq \mathbb{Z}^n$ of positive upper density one needs to have an abundance of solutions in $v \in (\mathbb{Z}^n)^m$, and the solutions must be uniformly distributed in some sense. The standard analytic tool which can provide such information is the Hardy-Littlewood circle method. Indeed, if the rank of the map Φ is sufficiently large and it also satisfies certain local conditions then it provides an asymptotic formula for the number of integer solutions to the system $\Phi(v) = Ll$ [1], at least for sufficiently large scales L which are multiples of a fixed integer $q = q(\ell)$.

Recall that the Birch rank $\mathcal{B}(\Phi)$of the map Φ is defined by

$$\mathcal{B}(\Phi) := nm - \dim_{\mathbb{C}} V^*_\Phi \quad \text{where} \quad V^*_\Phi := \{v \in (\mathbb{C}^n)^m : \operatorname{rank}(Jac\,\Phi(v)) < E\} \tag{1.11}$$

is the so-called singular locus of the map Φ. The following explicit estimate provides a lower bound for the Birch rank of the system Φ, exploiting the non-singularity of the form Q.

Proposition 1.3 (Rank inherited from Q)**.** *If Q is non-singular, then*

$$\mathcal{B}(\Phi) \ge n - E + 1. \tag{1.12}$$

The proof of Proposition 1.3 is will be given Section 4. As corollary in dimensions n satisfying

$$n - E \ge E(E+1)(r-1)2^{r-1}, \tag{1.13}$$

one has the lower bound

$$\mathcal{B}(\Phi) > E(E+1)(r-1)2^{r-1}. \tag{1.14}$$

This rank condition allows the use of the Birch's circle method for the diophantine system $\Phi(v) = L\ell$ which is crucial for our main result in the discrete settings.

One also needs some local conditions satisfied by the system Φ. Let $\Lambda = H^n_m \cap \mathbb{Z}^{nm}$. For a prime p, define

$$\sigma_p(L) := \lim_{s\to\infty} p^{-s(D-E)} \# \{v \in \Lambda/p^s\Lambda : \; \Phi(v) \equiv L\ell \pmod{p^s}\} \tag{1.15}$$

when the limit exists, and let $\mathfrak{S}(L) := \prod_{p \text{ prime}} \sigma_p(L)$.

It is not hard to show via Hensel's lemma, that $\sigma_p(L) > 0$ if the equation $\Phi(v) = 0$ has a non-singular solution $v = v^{(p)}$ in $(\mathbb{Q}^n_p)^m$ ($\mathbb{Q}_p$ being the field of p-adic numbers), and if L is divisible with a sufficiently large power of p.

Lemma 1.4. *Let $Q : \mathbb{Z}^n \to \mathbb{Z}$ be a positive-definite, non-singular form and let Δ be a Q-regular simplex. Assume that the dimension n satisfies* (1.14)*. If for all primes p there exists $v^{(p)} \in (\mathbb{Q}^n_p)^m$ such that*

$$\Phi(v^{(p)}) = 0, \quad rank\,(Jac\,\Phi(v^{(p)})) = E, \tag{1.16}$$

then one has

$$\beta_* := \inf_{q\geq 1} \beta_q(0) > 0, \tag{1.17}$$

for the quantities

$$\beta_q(0) := q^{-(D-E)}\, \# \{v \in \Lambda/q\Lambda : \Phi(v) \equiv 0 \pmod{q}\}\,, \tag{1.18}$$

and there exist $q_o \in \mathbb{N}$ and $c, C > 0$ such that

$$c \leq \mathfrak{S}(L) \leq C, \quad \text{whenever } q_o \mid L. \tag{1.19}$$

The proof of Lemma 1.4 follows from standard tools p-adic analysis, we postpone its proof to an appendix.

Remark 1.5. Let us remark that condition (1.16) automatically holds for all but finitely many primes. Indeed, for every sufficiently large prime of good reduction, the Birch-rank estimate gives $p^{D-E}(1 + o(1))$ solutions of $\Phi(v) = 0$, $v \in (\mathbb{F}^n_p)^m$, while the singular solutions form a lower-dimensional variety. Hence a non-singular solution exists modulo p, and by Hensel's lemma it can be lifted to $(\mathbb{Z}^n_p)^m$. Thus (1.16) needs to be verified only at finitely many exceptional primes.

Our first result in the discrete settings is the following,

Theorem 1.6. *Let $Q : \mathbb{Z}^n \to \mathbb{Z}$ be a positive-definite, convex, non-singular form of degree $r > 2$, such that the associated map Φ satisfies the local condition* (1.16)*, and let Δ be a Q-regular simplex. If*

$$n - E \geq E(E + 1)(r - 1)2^{r-1}. \tag{1.20}$$

then the following holds for any set $A \subset \mathbb{Z}^n$ has positive upper Banach density $\delta > 0$.

There are positive integers $q = q(\delta, Q, \Delta)$ and $L_0 = L_0(A, Q, \Delta)$ such that, for every

$$L \geq L_0, \quad q \mid L,$$

the set A contains a barycentric configuration $\Delta(z) = \{z, z + v_1, \ldots, z + v_m\}$ satisfying the equations

$$\sum_{i=1}^m v_i = 0, \qquad Q(v_i - v_j) = L\ell_{ij} \quad (1 \leq i < j \leq m). \tag{1.21}$$

The main conclusion of Theorem 1.6 is that the set A contains a simplex Δ together with its barycenter that is ρ-isometric to $\lambda\Delta_0$ for all $\lambda = L^{1/r}$, for all sufficiently large L divisible by a fixed modulus q, which depends only on the initial data δ, Q and Δ_o. We will refer to such levels L, and corresponding and scales $\lambda = L^{1/r}$ as *admissible* scales. This compares to the main result of [10], the crucial extra conclusion being that side vectors $v_1, \ldots, v_m$ satisfy the linear relation $v_1 + \ldots + v_m = 0$, and the isometry is defined with respect to metrics generated by positive-definite convex forms.

The proof of Theorem 1.6 is based on a combination of elements of the the proof in the Euclidean case and arithmetic inputs from the Hardy-Littlewood circle method. If $A \subseteq [N]^n$ then the count of barycentric configurations $\Delta(z) \subseteq A$ isometric $\lambda\Delta_o$ is given by the quantity

$$\mathcal{N}_\lambda(A, \Delta_o) := \sum_{z \in \mathbb{Z}^n} \sum_{v \in H_m^n} f(z) f(z + v_1) \ldots f(z + v_m)\, \mathbf{1}_{\{\Phi(v) = L\ell\}} \tag{1.22}$$

Our approach starts standard representation of the kernel

$$\mathbf{1}_{\{\Phi(v) = L\ell\}} = \int_{\mathbb{T}^E} e(\alpha(\Phi(v) - L\ell)\, d\alpha, \qquad (e(\beta) := e^{2\pi i \beta}),$$

and then a standard major/minor arcs decomposition of the phase variables α corresponding to the level L. The contribution of the minor arcs can be reduced after a change of variables, via three applications of the Cauchy-Schwarz inequality to estimating the Gowers U^3 uniformity norm of the phase functions $e(\alpha Q(s))$, similarly as in the Euclidean setting. The U^3-norm is bounded by the U^{r-1}-norm which can be estimated via Weyl-differencing, already implicit in [1].

The majors arcs are further decomposed into a small and an intermediate boxes. The integral over the small range provides the main term. The crucial difference from the Euclidean case that the corresponding kernel contains a local arithmetic factor $\chi(v) = \mathbf{1}_{\{\Phi(v) \equiv 0 \pmod{q}\}}$. The lower bound for the main term then is proved using the local condition (1.17) together with the finite multidimensional Szemerédi theorem. This is different from the continuous case and does not provide effective quantitative lower bounds.

Finally, the intermediate terms are handled by Poisson summation and Villarroya's transference theorem [13], which provides a bridge between estimates for continuous multilinear multipliers and their discrete analogues. This leads to an estimate for the total sum of the intermediate terms across a family of scales $\lambda_1 \ll \ldots \ll \lambda_J$ by a quantity independent of the number of scales J; a standard indirect argument then finishes the proof.

Theorem 1.6 states that every sufficiently large admissible level L occurs, but its use of multi-dimensional Szemerédi theorem makes it ineffective for the smallest admissible level L.

Our next result provides an effective bound on N such that every $A \subseteq [N]^d$ of density $\delta > 0$ necessarily contains a barycentric configuration $\Delta(z) = \{z, z + v_1, \ldots, z + v_m\}$, such the simplex $\Delta = \{\{z, z + v_1, \ldots, z + v_m\}$ is ρ-isometric to $\lambda\Delta_o$ for some scale $\lambda > 0$. We will refer to such a barycentric configuration $\Delta(z)$ *ρ-similar similar copy* of Δ_o.

Theorem 1.7. *Let $\delta > 0$ and let $\Delta_o \subseteq \mathbb{Z}^n$ be a fixed simplex. Assume the form Q, the associated map Φ, the simplex Δ_o and the dimension n satisfy the hypotheses of Theorem 1.6.*

Then there exists a constant $C = C(n, m, Q, \Delta_o)$ such that if $N \geq N(\delta) := \exp\exp(C\,\delta^{-C})$, then every set $A \subseteq [N]^n$ of size $|A| \geq \delta N^n$ contains a barycentric configuration $\Delta(z)$ which is a ρ-similar copy of Δ_0.

Note that when the metric ρ is the standard Euclidean metric generated by $Q(x) = |x|^2$, this is an ordinary Euclidean similar copy with dilation λ, and z is the barycenter of its vertices. For a general form Q, the terminology does not assert that a linear map carries Δ to $\lambda\Delta_o$.

2. Further barycentric configurations

The simplex-center configuration is a special case of general barycentric configurations $\Delta_{k,l}$ consisting of a simplex of k-points together with the barycenters of all of its l-faces. Much of the main ingredients of our proof extends to these configurations; a lower bound for the main term using a quantitative version of Szemerédi's theorem or its multi-dimensional extension and error term estimate via estimating Gowers norms. The missing ingredient is the applicability of a result in time frequency analysis that would provide a uniform bound on the sum of the intermediate terms over the scales λ_j $(1 \leq j \leq J)$, independent of J. However, good quantitative bounds maybe provided via the multi-scale orthogonality methods of Durcik-Kovač [4] both in the Euclidean and discrete setting. We plan to address general barycentric configurations in future work.

3. Regularity with respect to the form Q.

Put

$$g_{ij} := \nabla Q(a_i - a_j), \qquad g_{ji} := -g_{ij}. \tag{3.1}$$

For $h = (h_1, \dots, h_m) \in H^n_m$,

$$\big(Jac\,\Phi(\Delta)h\big)_{ij} = g_{ij} \cdot (h_i - h_j). \tag{3.2}$$

Proposition 3.1 (Q-regularity criterion)**.** *The simplex Δ is Q-regular if and only if the only symmetric edge weights $\omega_{ij} = \omega_{ji}$ satisfying*

$$\sum_{j \neq i} \omega_{ij} g_{ij} = 0 \qquad (1 \leq i \leq m) \tag{3.3}$$

are $\omega_{ij} = 0$.

Proof. Let $R_Q(\Delta)$ be the $E \times mn$ matrix whose ij-row has g_{ij} in the i-block, $-g_{ij}$ in the j-block, and zero elsewhere. A linear dependence among its rows is exactly (3.3). Simultaneous translations lie in the kernel of $D\Phi$, and every variation is the sum of a centered variation and a translation. Restriction to H^n_m therefore does not change the rank. Hence

$$\operatorname{rank} Jac\,\Phi(\Delta) = E \iff \det\big(R_Q(\Delta) R_Q(\Delta)^T\big) \neq 0 \iff \text{(3.3) has only the zero solution.}$$

$\square$

Proposition 3.2 (Sufficient conditions)**.** *The following statements hold.*

1. *If Q is a non-singular quadratic form, affine independence of the vertices implies Q-regularity.*
2. *If $m = 3$ and the unit sphere of $\rho = Q^{1/r}$ is smooth and strictly convex, every (n0n-degenerate) triangle is Q-regular.*
3. *Regularity follows if the vertices can be ordered $i_1, \dots, i_m$ so that, for every s, the vectors*

$$\{\nabla Q(a_{i_s} - a_{i_t}) : t > s\}$$

are linearly independent.

Proof. For a quadratic form, ∇Q is an invertible linear map, and the usual simplex rigidity argument applies. For a smooth strictly convex norm, the projective Gauss map is injective. Thus the two conormals at a vertex of a non-collinear triangle are independent.

Under (3), apply the equilibrium equation at i_1 to eliminate every edge from i_1 to a later vertex, and repeat at $i_2, i_3, \dots$. This eliminates every edge. $\square$

However, affine independence alone is not sufficient for higher simplices dimensional simplices, even for diagonal forms satisfying all the hypotheses on Q.

Proposition 3.3 (A non-singular diagonal counterexample)**.** *Let $r \geq 4$ be even, put $p = r-1$ and $c = 2^{-1/p}$, and take*

$$Q(x) = \sum_{\nu=1}^{n} b_\nu x_\nu^r, \qquad b_\nu > 0.$$

In the first three coordinates set

$$a_0 = 0, \quad a_1 = (1, -c, -c), \quad a_2 = (-c, 1, -c), \quad a_3 = (-c, -c, 1).$$

These four vertices are affinely independent but not Q-regular. Translating all four vertices by their centroid gives an example in H_4^n.

Proof. The matrix with columns a_1, a_2, a_3 has eigenvalues $1+c, 1+c, 1-2c$, all nonzero. Define

$$\omega_{0i} = 1 \quad (1 \leq i \leq 3), \qquad \omega_{ij} = -\frac{1}{2(1+c)^p} \quad (1 \leq i < j \leq 3).$$

Since $c^p = 1/2$, the equilibrium equation at a_0 holds coordinatewise. At a_1, after suppressing the invertible factors rb_ν, it is

$$(1, -\tfrac{1}{2}, -\tfrac{1}{2}) - \frac{1}{2}\big((1,-1,0) + (1,0,-1)\big) = 0.$$

At the other two vertices the same coordinate-wise identity holds after factoring out the corresponding nonzero coefficients rb_ν. Notice that Q is non-singular, since $\nabla Q(x) = r(b_\nu x_\nu^p)_\nu$ vanishes only at the origin. $\square$

4. Leray measure and the basic decomposition

We will only use regularity near the fixed simplex Δ_o as the configuration space $S_{\Phi,v} := \{v \in H_m^n : \Phi(v) = \lambda^r \ell\}$ may have singuarl point. Let $\chi_\Delta, \chi_0 \in C_c^\infty(H_m^n)$ be smooth cut-off functions such that

- χ_{Δ_o} is supported in a sufficiently small regular neighborhood of Δ_o and is positive at Δ_o;
- χ_0 is supported near the origin and is positive on a smaller neighborhood of the origin;
- $\operatorname{supp} \chi_0 \cap \Phi^{-1}(\ell) = \emptyset$.

Set $\chi = \chi_\Delta + \chi_0$. The normalized localized Leray measure maybe written formally as, see Section 7.

$$d\sigma_\lambda(v) : = \lambda^{rE-D} \chi(v/\lambda) \prod_{i<j} \delta\big(Q(v_i - v_j) - \lambda^r \ell_{ij}\big)\, dv_H \tag{4.1}$$

$$= \lambda^{rE-D} \chi(v/\lambda) \int_{\mathbb{R}^E} e\big(\sum_{i<j} \alpha_{ij}(Q(v_i - v_j) - \lambda^r \ell_{ij})\big) \prod_{i<j} d\alpha_{ij} \tag{4.2}$$

It is positive, nonzero, and supported near $\lambda\Delta_o$, while the χ_0-part vanishes on the level set $S_{\Phi,v}$. Its normalization follows from the scaling properties,

$$dv_H(\lambda u) = \lambda^D du_H, \qquad \delta^{(E)}\big(\lambda^r(\Phi(u) - \ell)\big) = \lambda^{-rE} \delta^{(E)}(\Phi(u) - \ell). \tag{4.3}$$

Choose an even function $\psi \in C_c^\infty(\mathbb{R}^E)$, $\psi \geq 0$, with $\int \psi = 1$, positive near both 0 and $-\ell$. Define

$$\omega_\lambda(v) := \lambda^{-D} \chi(v/\lambda) \psi\big(\Phi(v)/\lambda^r - \ell\big), \tag{4.4}$$

$$\omega_\lambda^\varepsilon(v) := \lambda^{-D} \varepsilon^{-E} \chi(v/\lambda) \psi\left(\frac{\Phi(v)/\lambda^r - \ell}{\varepsilon}\right). \tag{4.5}$$

The kernel $\omega_\lambda^\varepsilon$ converges to σ_λ in local coordinates as $\varepsilon \to 0$. Let

$$c_\varepsilon := \frac{\int_{H_m^n} \omega_\lambda^\varepsilon(v)\, dv_H}{\int_{H_m^n} \omega_\lambda(v)\, dv_H}. \tag{4.6}$$

Scaling shows that c_ε is independent of λ, and it is easy to see in local coordinates local coordinates that it stays between two positive constants for small ε. Let

$$k_\lambda^\varepsilon := \omega_\lambda^\varepsilon - c_\varepsilon \omega_\lambda, \qquad e_\lambda^\varepsilon := \sigma_\lambda - \omega_\lambda^\varepsilon. \tag{4.7}$$

Then the correctly signed decomposition is

$$\sigma_\lambda = c_\varepsilon \omega_\lambda + k_\lambda^\varepsilon + e_\lambda^\varepsilon, \qquad \int k_\lambda^\varepsilon = 0. \tag{4.8}$$

For $0 \leq f \leq 1$, supported in a cube of side N, define

$$\mathcal{N}_\lambda(f) := \int_{\mathbb{R}^n} \int_{H_m^n} f(z) \prod_{i=1}^m f(z+v_i)\, d\sigma_\lambda(v)\, dz. \tag{4.9}$$

Thus

$$\mathcal{N}_\lambda(f) = c_\varepsilon \mathcal{M}_\lambda(f) + \mathcal{K}_\lambda^\varepsilon(f) + \mathcal{E}_\lambda^\varepsilon(f). \tag{4.10}$$

5. Proof of Theorem 1.2

In this section we carry out the three main elements of the proof, a lower bound for the main term, en error estimate, and estimate for the total contribution of the intermediate terms for the decomposition (4.10). We follow [3], the new feature being an "edge-separation" lemma which makes it possible to reduce the error estimate which to that of estimating U^3 uniformity norm of the phase function $e(tQ)$.

5.1. Lower bound for the main term.

Lemma 5.1. *For every $m \geq 2$ and $\delta > 0$, there is a constant $c = c_{m,n}(\delta) > 0$ with the following property. If $A \subset [0,N]^n$ has measure at least δN^n and $1 \leq \lambda \leq \frac{c}{10n} N$, then*

$$\int_z \int_{\substack{v \in H_m^n \\ |v_i| \leq c\lambda}} \mathbf{1}_A(z) \prod_{i=1}^m \mathbf{1}_A(z+v_i)\, dv_H\, dz \geq c\, N^n \lambda^D. \tag{5.1}$$

Moreover, one may take $c = c_{m,n}(\delta) = \exp\left(-C_{m,n} \log^7(2/\delta)\right)$ for $m \geq 3$ and $c_{m,n}(\delta) = \exp\left(-C_{m,n} \log^9(2/\delta)\right)$ for $m = 2$, with some constant $C_{m,n} > 0$.

Proof. Partition the large cube into boxes of side comparable with λ, on at least $\delta/2$ proportion of the boxes the relative density of the set A is at least $\delta/2$, let's call these boxes *dense*. Discretize each dense box by a fine grid, apply the lower bound for the number of solutions to the linear, translation invariant equation

$$x_1 + \ldots + x_m - mz = 0 \tag{5.2}$$

and average over translations, grid frames, and mesh sizes. By the usual Varnavides averaging one has that a $c(\delta) > 0$ proportion of all solutions $(x_1, \ldots, x_m, z)$ to equation (5.2) are in A^{m+1}. Thickening the selected grid points inside Lebesgue-density cells and then averaging over the mesh offset removes the discretization, see [3]. Averaging over the dense boxes using the best current bounds [12, 9] $c_{m,n}(\delta) = \exp\left(-C_{m,n} \log^7(2/\delta)\right)$ shows (5.1). $\square$

Since ψ is positive near $-\ell$, the kernel $\omega_\lambda(v)$ is bounded below by a constant times λ^{-D} when all v_i are sufficiently small compared with λ. Therefore

$$\mathcal{M}_\lambda(\mathbf{1}_A) \geq c(\delta, m, Q, \Delta) N^n \qquad (1 \ll \lambda \ll_\delta N). \tag{5.3}$$

5.2. **Pair isolation and the edge transform.** In order to apply the Cauchy-Schwarz inequality to estimate the error term via the U^3-uniformity norm one needs to find an edge $(v_a - v_b)/2 = s$ such that in the s-variable the phase

$$P_t(v) := \sum_{i<j} t_{ij} Q(v_i - v_j). \tag{5.4}$$

is non-vanishing for $t = (t_{ij})_{i<j} \in \mathbb{R}^E \backslash \{0\}$.

Let $1 \leq a < b \leq m$. If all v_j, $j \neq a, b$, are fixed, the condition $\sum_i v_i = 0$ fixes $v_a + v_b$. Write

$$w := -\sum_{j \neq a,b} v_j, \qquad v_a = \frac{w}{2} + s, \qquad v_b = \frac{w}{2} - s. \tag{5.5}$$

Extend the variable $t = (t_{ij})_{i<j}$ symmetrically to $t = (t_{ij})_{i \neq j}$ via $t_{ji} := t_{ij}$ if $i < j$. As a polynomial in s, the degree-r part of P_t is:

$$C_{ab}(t) Q(s), \qquad C_{ab}(t) := 2^r t_{ab} + \sum_{j \neq a,b} (t_{aj} + t_{bj}). \tag{5.6}$$

Lemma 5.2. *We have,*

$$\max_{a<b} |C_{ab}(t)| \geq c_{m,r} |t|. \tag{5.7}$$

Proof. If we set $t_{ji} := t_{ij}$ for $i < j$, then we may write

$$P_t(v) := \frac{1}{2} \sum_{i \neq j} t_{ij} Q(v_i - v_j),$$

Let's define $C_{ba}(t) := C_{ab}(t)$ where $C_{ab}(t)$ is defined by (5.6). Consider the linear map: $C_r : t \to (C_{ab}(t))_{a \neq b})$ mapping $\mathbb{R}^{m(m-1}$ to itself. We calculate the eigenvalues of this map. Assume $C_r(t) = \gamma\, t$ for some eigenevalue γ and eigenvector $t \neq 0$. Write,

$$\tau_a := \sum_{b \neq a} t_{ab}, \quad \tau := \sum_a \tau_a$$

Then, from (5.6) we have for all $a \neq b$

$$(2^r - 2 - \gamma) t_{ab} + \tau_a + \tau_b = 0.$$

Summing for $b \neq a$ gives

$$(2^r + m - 4 - \gamma) \tau_a + \tau = 0,$$

for all $a \in [m]$. Summing for a, we get

$$(2^r + 2m - 4 - \gamma) \tau = 0.$$

Then either $\gamma = 2^r + 2m - 4$ or $\tau = 0$. If $\tau = 0$ the either $\gamma = 2^r + m - 4$ or $\tau_a = 0$ for all $a \in [m]$. In the latter case either $\gamma = 2^r - 2$ or $t_{ab} = 0$ for all $a, b \in [m]$ and hence $t = 0$. Thus the eigen-values are: $2^r - 2$, $2^r + m - 4$ and $2^r + 2m - 4$. The operator C_r is symmetric an its eigenvalues are at least $2^r - 2$ thus $\|C_r(t)\|_2 \geq (2^r - 2)\|t\|_2$. Hence

$$\max_{a<b} |C_{ab}(t)| \geq c_{m,r} |t|, \quad \text{with} \quad c_{m,r} = \frac{2^r - 2}{\sqrt{m(m-1)}}.$$

□

If g is compactly supported, and $B \subseteq \mathbb{R}^n$ is a box of size P, the normalized local norm of g in the box B, is defined by

$$\|g\|_{U^s(B)}^{2^s} := P^{-n(s+1)} \int_{\mathbb{R}^{n(s+1)}} \prod_{\omega \in \{0,1\}^s} \mathcal{C}^{|\omega|} g(x + \omega \cdot h)\, \mathbf{1}_B(x + \omega \cdot h)\, dx\, dh, \tag{5.8}$$

where $\mathbf{1}_B$ is the indicator function of the box B.

Lemma 5.3 (Pair U^3 estimate)**.** *Let $F_0, F_1, F_2 : B \to \mathbb{C}$ bounded in magnitude by 1, where B is box of size P.*

$$P^{-2n} \left| \int \int F_0(x) F_1(x+s) F_2(x-s) g(s)\, dx\, ds \right| \ll \|g\|_{U^3([-4P,4P]^n)},$$

with an implicit constant depending only on the dimension n.

Proof. Applying Cauchy–Schwarz inequality successively together with shifting the variables $x+s \to x$ after the first, and $x+2s \to x$ after the second application eliminates the factors F_0, F_1, F_2. After a third application of Cauchy–Schwarz the eighth power is bounded by a constant times the integral of

$$P^{-4n} \int \prod_{\omega \in \{0,1\}^3} \mathcal{C}^{|\omega|} g(s + \omega \cdot h)\, \mathbf{1}_{[-4P,4P]^n}(s + \omega \cdot h)\, dh\, ds$$

which is the eighth power localized $U^3([-4P, 4P]^n)$-norm of the function g. ☐

5.3. **Gowers norm estimates for polynomial phase functions.** We will need to estimates the Gowers norms of certain phase functions. Such estimates are not directly available in the Euclidean settings for general polynomial phases Q, thus we first estimate the more standard discrete Gowers norms of phase functions $e(\theta Q(x))$. Fortunately such estimates are known for phase functions $g(x) = e(\alpha Q(x) + R(x))$ where *deg* $R < r$ and α belongs to the so-called minor arcs, in fact if the form Q has sufficiently large rank these are provide by provided by Birch [1]. To be more precise, let

$$s := r - 1, \qquad p := 2^{r-1}, \qquad \kappa := \frac{n}{(r-1)2^{r-1}}. \tag{5.9}$$

If $g : \mathbb{Z}^n \to \mathbb{C}$ is supported in a box of side $O(P)$, define the normalized Gowers local norm by

$$\|g\|_{U_P^s}^p := P^{-nr} \sum_{x, h_1, \ldots, h_s \in \mathbb{Z}^n} \prod_{\omega \in \{0,1\}^s} \mathcal{C}^{|\omega|} g(x + \omega \cdot h), \tag{5.10}$$

where g is extended by zero outside the box.

We recall Birch's minor arcs estimate for a single non-singular form of degree r. For $0 < \theta < 1$, let recall the major/minor arcs at level θ.

$$\mathfrak{M}_P(\theta) := \bigcup_{1 \le q \le P^{(r-1)\theta}} \bigcup_{\substack{a \in \mathbb{Z} \,(\mathrm{mod}\ q) \\ (a,q)=1}} \left\{ \alpha \in \mathbb{T} : 2|q\alpha - a| < P^{-r+(r-1)\theta} \right\}, \qquad \mathfrak{m}_P(\theta) := \mathbb{T} \setminus \mathfrak{M}_P(\theta). \tag{5.11}$$

Lemma 5.4 (Birch's minor-arc estimate for Gowers norms)**.** *Let $Q \in \mathbb{Z}[x_1, \ldots, x_n]$ be a homogeneous non-singular form of degree r and let $W \in C_c^\infty(\mathbb{R}^n)$. If R is any polynomial of degree less than r, then, uniformly in all coefficients of R, one has*

$$\|W(\cdot/P) e\big(\alpha Q + R_{<r}\big)\|_{U_P^{r-1}} \ll_{Q,W} P^{-n\theta/2^{r-1}} \qquad \big(\alpha \in \mathfrak{m}_P(\theta)\big). \tag{5.12}$$

Consequently, if $P^{-r} \le |\alpha| < P^{-r/2}$, where α is represented in $[-1/2, 1/2]$, then

$$\|W(\cdot/P) e\big(\alpha Q + R_{<r}\big)\|_{U_P^{r-1}} \ll_{\varepsilon,Q,W} \big(P^r|\alpha|\big)^{-\kappa}. \tag{5.13}$$

Remark 5.5. Since all weights occurring here are smooth and compactly supported, the factor P^{ε} that is appearing in in Lemma 4.4 may be omitted. Indeed, after the final differencing step, Poisson summation in the remaining variable replaces the usual reciprocal sums by a rapidly decreasing periodic kernel. The resulting dyadic sums are uniformly summable. Consequently, for every $\kappa_0 < \kappa$,

$$\big\|W(\cdot/\lambda)e(\theta Q + R_{<r})\big\|_{U_\lambda^{r-1}} \ll_{\kappa_0,W} H^{-\kappa_0}, \qquad \theta \notin \mathfrak{M}_1(\lambda; H),$$

uniformly in the coefficients of $R_{<r}$ and throughout the range of H used below.

Proof. Let Γ_Q be the symmetric r-linear form associated with Q and put

$$\Psi_j(\mathbf{h}) := (-1)^{r-1} r!\, \Gamma_Q(e_j, h_1, \dots, h_{r-1}), \qquad \mathbf{h} = (h_1, \dots, h_{r-1}).$$

For

$$g(x) := W(x/P)e\big(\alpha Q(x) + R_{<r}(x)\big),$$

expansion of the p-th power in (5.10) gives

$$\|g\|^p_{U_P^{r-1}} = P^{-nr} \sum_{\mathbf{h}} \sum_{x} A_{\mathbf{h}}(x) e\left(\alpha \sum_{j=1}^{n} x_j \Psi_j(\mathbf{h}) + C_{\alpha,R}(\mathbf{h})\right), \tag{5.14}$$

where

$$A_{\mathbf{h}}(x) := \prod_{\omega \in \{0,1\}^{r-1}} \mathcal{C}^{|\omega|} W\left(\frac{x + \omega \cdot h}{P}\right).$$

Here $\Delta_h F(x) := F(x) - F(x+h)$, and we used the fact that

$$\Delta_{h_1} \cdots \Delta_{h_{r-1}} \big(\alpha Q + R_{<r}\big)(x) = \alpha \sum_{j=1}^{n} x_j \Psi_j(\mathbf{h}) + C_{\alpha,R}(\mathbf{h}).$$

In particular, all lower-degree terms are independent of the base variable x and hence disappear after taking the absolute value of the inner sum. The support of $A_{\mathbf{h}}$ forces $|h_i| \ll_W P$. The support of $A_{\mathbf{h}}$ is forces $|h_i| \ll P$ its rapidly decaying Fourier transform yields,

$$\|g\|^p_{U_P^{r-1}} \ll_{W,C} \mathcal{B}_P(\alpha), \qquad \mathcal{B}_P(\alpha) := P^{-nr} \sum_{|\mathbf{h}| \ll_W P} \prod_{j=1}^{n} \big(1 + P\|\alpha \Psi_j(\mathbf{h})\|_{\mathbb{R}/\mathbb{Z}}\big)^{-C}. \tag{5.15}$$

The quantity $\mathcal{B}_P(\alpha)$ compares to the terminal quantity on the right of Birch's Lemma 2.1, equation (2), after division by P^{np}, with the difference that the quantity in Lemma 2.1 in [1] corresponds to $C = 1$. The argument proving Birch's minor-arc estimate (Lemmas 2.2–3.3 and Lemma 4.3 of [1]), starting from this terminal quantity, applies without any changes. Indeed, the exponential sum enters that argument only through Lemma 2.1, which places its normalized p-th power below $\mathcal{B}_P(\alpha)$. Starting instead from $\mathcal{B}_P(\alpha)^{1/p}$ gives the same alternative: either

$$\mathcal{B}_P(\alpha)^{1/p} \ll_\varepsilon P^{-K\theta}, \tag{5.16}$$

or $\alpha \in \mathfrak{M}_P(\theta)$, or the forms Ψ_i lie on Birch's singular locus. Since Q is non-singular, Birch's parameter K, defined by

$$\dim V_Q^* = n - 2^{r-1} K,$$

equals $n/2^{r-1}$, and the singular alternative is absent. Combining (5.15) and (5.16) proves (5.12).

Finally, Birch's "small-α" corollary following Lemma 4.3 applies to $\mathcal{B}_P(\alpha)^{1/p}$. Indeed, choose θ_0 such that

$$P^r |\alpha| = P^{(r-1)\theta_0}$$

and approach $\theta_)$ value of θ from below. Since $P^{-r} \le |\alpha| < P^{-r/2}$, one has $2(r-1)\theta < r$, and the minor arcs estimate (5.16) applies. Thus

$$P^{-K\theta} = \big(P^r|\alpha|\big)^{-K/(r-1)} = \big(P^r|\alpha|\big)^{-\kappa},$$

which gives (5.13). □

Lemma 5.6 (Transfer to Euclidean Gowers norms)**.** *Let $r \ge 4$ and fix $\varphi \in C_c^\infty(\mathbb{R}^n)$. Suppose*

$$P_t(y) = c_t Q(y) + R_t(y), \qquad \deg R_t < r, \tag{5.17}$$

where, for $|t| \ge 1$,

$$|c_t| \asymp |t| \quad \text{and the coefficients of } R_t \text{ are } O(1+|t|).$$

Let $W \in C_c\infty(\mathbb{R}^n)$ satisfying $\|W\|_{C^M} \le C_W$. Then,

$$\|W\, e(P_t)\|_{U^3(\mathbb{R}^n)} \ll_{Q,W} (1+|t|)^{-\kappa}, \tag{5.18}$$

with implicit constant depending only on the support and finitely many derivatives of the weight function W.

Proof. The range $|t| \le 1$ is trivial. Put $T := 2 + |t|$, choose

$$A > 1 + \frac{n}{r-1}, \qquad P := \lceil T^A \rceil, \tag{5.19}$$

and sample the continuous function on the mesh $P^{-1}\mathbb{Z}^n$. Since the integrand defining the p-th power of the $U^{r-1}(\mathbb{R}^n)$-norm has fixed compact support and Lipschitz norm $O(T)$, the Riemann-sum comparison gives

$$\|We(P_t)\|^p_{U^{r-1}(\mathbb{R}^n)} = \|W\,(\cdot/P)e\big(P_t(\cdot/P)\big)\|^p_{U_P^{r-1}} + O(TP^{-1}). \tag{5.20}$$

On the lattice,

$$P_t(x/P) = \alpha Q(x) + R_{t,P}(x), \qquad \alpha := \frac{c_t}{P^r}, \qquad \deg R_{t,P} < r.$$

Moreover,

$$P^r|\alpha| = |c_t| \asymp T, \qquad P^{-r} \le |\alpha| \le P^{-r/2}$$

for T sufficiently large. Hence (5.13) gives

$$\|W(\cdot/P)e\big(P_t(\cdot/P)\big)\|_{U_P^{r-1}} \ll_{Q,W} T^{-\kappa}. \tag{5.21}$$

After taking the p-th root in (5.20), the error is

$$O\big((TP^{-1})^{1/p}\big) = O\big(T^{-(A-1)/2^{r-1}}\big) = o(T^{-\kappa})$$

by (5.19). Combining this with (5.21) proves the corresponding $U^{r-1}(\mathbb{R}^n)$ estimate. Since $r \ge 4$, using monotonicity of the local Gowers norms gives $U^3 \lesssim U^{r-1}$. This proves (5.18). □

5.4. **The high-frequency term.** Fourier inversion, followed by the changes of variables $v = \lambda u$ and $t = \lambda^r \alpha$, gives

$$\mathcal{E}^\varepsilon_\lambda(f) = \int_{\mathbb{R}^E} \big(1 - \widehat{\psi}(\varepsilon t)\big) e(-t\cdot \ell)\, \mathcal{T}_{\lambda,t}(f)\, dt,$$

where

$$\mathcal{T}_{\lambda,t}(f) := \int_{\mathbb{R}^n} \int_{H^n_m} \chi(u) f(z) \prod_{i=1}^m f(z+\lambda u_i)\, e\,\big(P_t(u)\big)\, du_H\, dz$$

and

$$P_t(u) := t \cdot \Phi(u) = \sum_{i<j} t_{ij} Q(u_i - u_j).$$

Fixing the remaining simplex variables and using Theorem 5.2, choose $a < b$ with $|C_{ab}(t)| \gtrsim |t|$. Then Theorems 5.3 and 5.4 give

$$|T_t(f)| \lesssim N^n(1+|t|)^{-\kappa}. \tag{5.22}$$

Since $\widehat{\psi}(0) = 1$ and ψ is even,

$$|1-\widehat{\psi}(\varepsilon t)| \lesssim \min(1, \varepsilon^2|t|^2). \tag{5.23}$$

Thus, whenever $\kappa > E$, there is $c > 0$ such that

$$|\mathcal{E}^{\varepsilon}_{\lambda}(f)| \lesssim \varepsilon^c N^n. \tag{5.24}$$

One may take any $c < \min(2, \kappa - E)$.

5.5. **The intermediate terms.** The kernels of the intermediate terms are of the form,

$$k^{\varepsilon}_{\lambda_j}(v) = \lambda_j^{-D} k^{\varepsilon}_1(v/\lambda_j), \qquad \int_{H^n_m} k^{\varepsilon}_1 = 0, \tag{5.25}$$

with k^{ε}_1 smooth and compactly supported. We assume the scales λ_j are *lacunary* i.e. $\lambda_{j+1} \geq 2\lambda_j$.

Proposition 5.7. *If $\lambda_1 < \cdots < \lambda_J$ is lacunary and f is supported in a cube of side N, with $|f| \leq 1$, then*

$$\sum_{j=1}^{J} |\mathcal{K}^{\varepsilon}_{\lambda_j}(f)|^2 \leq CN^{2n}, \tag{5.26}$$

with a constant $C = C(\varepsilon, m, Q, \Delta)$ independent of J and the sequence.

Proof. The argument follows closely that of [3, Section 5]. Applying Cauchy–Schwarz in the variable eliminates $f(z)$. Expanding the square and summing in j gives a $2m$-linear form with kernel

$$K^{\varepsilon}_J(v, v') = \sum_{j=1}^{J} k^{\varepsilon}_{\lambda_j}(v)\overline{k^{\varepsilon}_{\lambda_j}(v')}. \tag{5.27}$$

On the Fourier hyperplane

$$\Gamma = \left\{ (\xi_1, \ldots, \xi_m, \xi'_1, \ldots, \xi'_m) : \sum_i (\xi_i + \xi'_i) = 0 \right\},$$

the singular subspace is

$$\Gamma' = \{(\theta, \ldots, \theta, -\theta, \ldots, -\theta) : \theta \in \mathbb{R}^n\}. \tag{5.28}$$

To see this subspace directly, use $v_m = -\sum_{i<m} v_i$ as coordinates on H^n_m. The kernel-frequency map is

$$L(\xi, \xi') = \big((\xi_i - \xi_m)_{i<m}, (\xi'_i - \xi'_m)_{i<m}\big) \in (H^n_m)^* \times (H^n_m)^*. \tag{5.29}$$

The multiplier is singular only when $L(\xi, \xi') = 0$. On Γ this means that all unprimed frequencies equal some θ, all primed frequencies equal some θ', and $m(\theta + \theta') = 0$; hence $\theta' = -\theta$ and the singular set is exactly (5.28). It has integral rank one and is a graph over every frequency block. In the notation of the integral-rank theorem there are $2m$ input blocks, each of dimension n, while the singular space has dimension n. Thus its rank is $1 < m = (2m)/2$, and the block-graph nondegeneracy is exactly the hypothesis in the rank-one integral case of [6, Section 4.4]; see also [3, 11].

The symbol verification is short. In kernel coordinates,

$$m_J(\eta,\eta') = \sum_{j=1}^{J} \widehat{k}_1^{\varepsilon}(\lambda_j\eta)\overline{\widehat{k}_1^{\varepsilon}(\lambda_j\eta')}.$$

Put $\varrho = |(\eta,\eta')|$. For $\lambda_j\varrho \le 1$, the identity $\widehat{k}_1^{\varepsilon}(0) = 0$ controls the zeroth derivative, while the geometric series in λ_j controls every positive derivative. For $\lambda_j\varrho > 1$, Schwartz decay gives the corresponding reverse geometric series. Consequently, for every derivative required by the multiplier theorem,

$$|\partial^\beta m_J(\xi)| \le C_{\beta,\varepsilon}\,\mathrm{dist}(\xi,\Gamma')^{-|\beta|}, \tag{5.30}$$

uniformly in J. Apply the theorem with all $2m$ input exponents equal to $2m$. Their L^{2m} norms contribute N^n, and the initial Cauchy–Schwarz factor contributes another N^n, proving (5.26). □

Proof of Theorem 1.2. Suppose arbitrarily large scales are missing and select a finite lacunary sequence $\lambda_1 < \cdots < \lambda_J$ of missing scales. Upper Banach density provides a cube of side $N \gg \lambda_J$ in which A has density bounded below by δ. Replace A by its intersection with this cube and put $f = \mathbf{1}_A$.

At every missing scale $\mathcal{N}_{\lambda_j}(f) = 0$. Choose ε so small that (5.24) is at most half the lower bound (5.3). Since $c_\varepsilon \asymp 1$, (4.10) gives

$$|\mathcal{K}^{\varepsilon}_{\lambda_j}(f)| \ge c_{m,n}(\delta)N^n \qquad (1 \le j \le J).$$

Therefore

$$J\,c_{m,n}(\delta)^2 N^{2n} \le \sum_{j=1}^{J} |\mathcal{K}^{\varepsilon}_{\lambda_j}(f)|^2 \lesssim_\varepsilon N^{2n},$$

contradicting Theorem 5.7 for large J. □

6. The discrete settings

Here we describe the proofs of our main results in the setting of the integer lattice $\mathbb{Z}^n$. Fix a level L and write

$$P := L^{1/r}$$

for its physical scale. We insert the same cut-off function χ as in the Euclidean case, and define

$$\sigma_L^{\mathbb{Z}}(v) := P^{rE-D}\chi(v/P)\prod_{i<j}\mathbf{1}_{\{Q(v_i-v_j)=L\ell_{ij}\}}, \qquad v \in H^n_{m,\mathbb{Z}}. \tag{6.1}$$

Note that $\sigma_L^{\mathbb{Z}}$ is normalized if the system Φ has sufficiently large rank [1], which we verify in Section 6.1.

For f supported in a box of side N, the normalized lattice count is

$$\mathcal{N}_L^{\mathbb{Z}}(f) := N^{-n}\sum_{z\in\mathbb{Z}^n}\sum_{v\in H^n_{m,\mathbb{Z}}} f(z)\prod_{i=1}^{m} f(z+v_i)\,\sigma_L^{\mathbb{Z}}(v). \tag{6.2}$$

Fourier inversion gives the representation

$$\sigma_L^{\mathbb{Z}}(v) = P^{rE-D}\chi(v/P)\int_{\mathbb{T}^E} e\big(\alpha\cdot(\Phi(v)-L\ell)\big)\,d\alpha, \tag{6.3}$$

6.1. **The rank of the map Φ.** We prove the lower bound for the rank of the system Φ by exploiting the non-singularity of the form Q. This will be needed to apply the circle method to the Diophantine system $\Phi(v) = L\ell$.

Proof of Theorem 1.3. For $0 \neq t \in \mathbb{C}^E$, let $P_t = t \cdot \Phi$. By Theorem 5.2 we can choose $a < b$ with $C_{ab}(t) \neq 0$ and use (5.5). Fixing all edges v_i for $i \neq a, b$ and setting $s = (v_a - v_b)/2$ and $u = (v_i)_{i \neq a,b}$, we have

$$P_t(v) = P_t(s, u) = C_{ab}(t)Q(s) + \text{terms of lower degree in } s.$$

Let

$$F_j^u(s) := \partial_{s_j} P_t(u, s), \qquad 1 \leq j \leq n.$$

These are polynomials of degree at most $r - 1$ in s, and their homogeneous parts of degree $r - 1$ are

$$C_{ab}(t)\partial_j Q(s), \qquad 1 \leq j \leq n.$$

Since Q is non-singular, the equations

$$\partial_1 Q(s) = \cdots = \partial_n Q(s) = 0$$

have no nonzero common solution in $\mathbb{C}^n$.

We claim that, for every fixed u, the common zero set

$$\{s \in \mathbb{C}^n : F_1^u(s) = \cdots = F_n^u(s) = 0\}$$

is finite. Indeed, homogenize the polynomials F_j^u in the variables (s, w). On the hyperplane at infinity $w = 0$, the homogenized equations reduce to

$$C_{ab}(t)\partial_j Q(s) = 0, \qquad 1 \leq j \leq n,$$

which have no projective solution $[s : 0]$, however positive-dimensional affine common zero set would have a projective closure meeting the hyperplane at infinity.

Consequently, every fiber of the projection

$$\{(u, s) : \nabla_s P_t(u, s) = 0\} \longrightarrow \mathbb{C}^{D-n}, \qquad (u, s) \longmapsto u,$$

is zero-dimensional. The fiber-dimension theorem therefore gives

$$\dim\{(u, s) : \nabla_s P_t(u, s) = 0\} \leq D - n.$$

Since

$$V_{P_t}^* = \{(u, s) : \nabla_{u,s} P_t(u, s) = 0\} \subseteq \{(u, s) : \nabla_s P_t(u, s) = 0\},$$

we conclude that

$$\dim V_{P_t}^* \leq D - n.$$

Consider the "incidence" variety,

$$\mathcal{V}_\Phi := \{(v, [t]) \in H_{m,\mathbb{C}}^n \times \mathbb{P}^{E-1} : \nabla_v (t \cdot \Phi)(v) = 0\},$$

where $[t]$ denotes the projective line through t. Every fiber over $[t]$ has dimension at most $D - n$, so $\dim \mathcal{V}_\Phi \leq D - n + E - 1$. Its projection to $H_{m,\mathbb{C}}^n$ is $V_\Phi^* := \{v \in H_{m,\mathbb{C}}^n : \ rank\, Jac_\Phi(v) < E\}$. Indeed, if $v \in V_\Phi^*$ then there exists $t \in \mathbb{C}^E,\ t \neq 0$ such that $\nabla_v(t \cdot \Phi) = 0$. Therefore $\dim V_\Phi^* \leq D - n + E - 1$, which implies $\mathcal{B}(\Phi) \geq n - E + 1$. $\square$

6.2. **The circle-method decomposition.** Fix a physical scale P, or equivalently a level $L = P^r$. As usual in the circle method, we decompose the range of integration first into major/minor arcs depending on a parameter $\varepsilon = \varepsilon(\delta)$. We say that α is in a major arc at height R if there exists $1 \le q \le R$ such that $\|q\alpha\| \le R/L$, where $\|\beta\| = |\{\beta\}|_\infty$ and $\{\beta\} \in [-\frac{1}{2}, \frac{1}{2}]^E$ is the fractional part of $\beta \in \mathbb{R}^E$. We denote the union of these major arcs by $\mathfrak{M}_{P,R}$ and use heights $\varepsilon^{-1/2} \le R \le P$. We enlarge the major arcs at height $R = \varepsilon^{-1/2}$ by taking $q = q(\varepsilon) := \mathrm{lcm}\{1 \le q \le \varepsilon^{-1/2}\}$ and cubes of side comparable to $\varepsilon^{-1/2}/L$ centered at $a/q \in \mathbb{T}^E$, with $a \in [q]^E$; denote their union by $\mathfrak{M}'_{\varepsilon,P}$. We also assume throughout this section that $q \mid L$. Then we have the decomposition

$$\sigma_L^{\mathbb{Z}} = c_\varepsilon \omega_{q,L} + k_{q,L}^\varepsilon + e_{q,L}^\varepsilon, \tag{6.4}$$

where

$$\Psi_{P,\tau}(\alpha) := \sum_{k \in \mathbb{Z}^E} \widehat{\psi}\big(\tau P^r(\alpha - k)\big), \qquad \tau > 0,$$

is the periodization of the Euclidean multiplier. Thus the following torus representations are exact:

$$\omega_{q,L}(v) := P^{rE-D}\chi(v/P) \int_{\mathbb{T}^E} \sum_{a \in [q]^E} e\big(\alpha \cdot (\Phi(v) - L\ell)\big)\, \Psi_{P,1}(\alpha - a/q)\, d\alpha,$$

$$\omega_{q,L}^\varepsilon(v) := P^{rE-D}\chi(v/P) \int_{\mathbb{T}^E} \sum_{a \in [q]^E} e\big(\alpha \cdot (\Phi(v) - L\ell)\big)\, \Psi_{P,\varepsilon}(\alpha - a/q)\, d\alpha. \tag{6.5}$$

Set

$$k_{q,L}^\varepsilon := \omega_{q,L}^\varepsilon - c_\varepsilon \omega_{q,L}, \qquad e_{q,L}^\varepsilon := \sigma_L^{\mathbb{Z}} - \omega_{q,L}^\varepsilon.$$

Changing variables $\alpha := \alpha - a/q$ one obtains

$$\omega_{q,L}(v) = \sum_{a \in [q]^E} e\bigg(\frac{a \cdot (\Phi(v) - L\ell)}{q}\bigg)\omega_P(v) = q^E\, \mathbf{1}_{\{\Phi(v) \equiv 0 \pmod q\}}\, \omega_P(v) =: \chi_q(v)\omega_P(v), \tag{6.6}$$

assuming $v \in H_{m,\mathbb{Z}}^n$ and $q = q(\varepsilon) \mid L$. Thus the extra arithmetic factor in the discrete case is

$$\chi_q(v) := q^E \mathbf{1}_{\{\Phi(v) \equiv 0 \pmod q\}}, \tag{6.7}$$

By scaling, we have that

$$\omega_{q,L}(v) = \chi_q(v)\omega_P(v), \qquad k_{q,L}^\varepsilon(v) = \chi_q(v) k_P^\varepsilon(v), \tag{6.8}$$

ω_P and k_P^ε being the kernels in (4.7).

For $0 \le f \le 1$ supported in $[1, N]^n$, write

$$\mathcal{M}_{q,L}^{\mathbb{Z}}(f) := N^{-n} \sum_{z \in \mathbb{Z}^n} \sum_{v \in H_{m,\mathbb{Z}}^n} f(z) \prod_{i=1}^m f(z + v_i)\omega_{q,L}(v), \tag{6.9}$$

and define $\mathcal{K}_{q,L}^{\varepsilon,\mathbb{Z}}$ and $\mathcal{E}_{q,L}^{\varepsilon,\mathbb{Z}}$ analogously using $k_{q,L}^\varepsilon$ and $e_{q,L}^\varepsilon$. Thus (6.4) gives

$$\mathcal{N}_L^{\mathbb{Z}}(f) = c_\varepsilon \mathcal{M}_{q,L}^{\mathbb{Z}}(f) + \mathcal{K}_{q,L}^{\varepsilon,\mathbb{Z}}(f) + \mathcal{E}_{q,L}^{\varepsilon,\mathbb{Z}}(f). \tag{6.10}$$

Proposition 6.1 (Main term lower bound.)**.** *Assume* (1.17)*. For every $\delta > 0$ there are $c = c(\delta, m, n, Q, \Delta) > 0$, $\eta = \eta(\delta, m, n, Q, \Delta) > 0$, and $C = C(\delta, m, n, Q, \Delta) < \infty$ such that, if $A \subset [1, N]^n \cap \mathbb{Z}^n$ has density at least δ, $q \mid L$, $P = L^{1/r}$, and*

$$Cq \le P \le \eta N,$$

then

$$\mathcal{M}_{q,L}^{\mathbb{Z}}(\mathbf{1}_A) \ge c(\delta, m, n, Q, \Delta) > 0, \tag{6.11}$$

We remark that it is crucial in our argument that the lower bound is independent of q, L, and N.

Proof. Let $s = m-1$ and identify $H^n_{m,\mathbb{Z}}$ with $(\mathbb{Z}^n)^s$ by writing

$$R = (r_1, \ldots, r_s, r_m), \qquad r_m = -\sum_{j=1}^{s} r_j. \tag{6.12}$$

Choose $\theta > 0$ so small that

$$\omega_P(v) \ge c_0\, P^{-D}, \quad \text{whenever } \max_i |v_i|_\infty \le \theta P. \tag{6.13}$$

This is possible because χ_0 is positive near the origin and ψ is positive near $-\ell$.

We embed A in the finite group

$$\Gamma_N := (\mathbb{Z}/(4N+1)\mathbb{Z})^n.$$

Its density in Γ_N is at least $\delta_0 := 5^{-n}\delta$. By the finite multidimensional Szemerédi theorem [7] there exists $K = K(\delta_0, m)$ such that every subset of $[K]^s$ of density at least $\delta_0/2$ contains a nontrivial homothetic copy

$$a + tP_m \subset [K]^s, \qquad 1 \le t \le K, \tag{6.14}$$

where $P_m = \{0, e_1, \ldots, e_{m-1}, -\sum_{i=1}^{m-1} e_i\}$.

Let $\mathcal{R}_q(P)$ be the set of integral frames (6.12) such that

$$|r_j|_\infty \le \frac{\theta P}{sK} \quad (1 \le j \le s), \qquad \Phi(R) \equiv 0 \pmod q. \tag{6.15}$$

Each frame $v \in H^n_{m,(Z/qZ)}$ such that $\Phi \equiv 0 \pmod q$ has $\gg (P/(Kq))^D$ such lifts if $P \ge Cq$. Thus it follows from (1.18)–(1.17) that

$$|\mathcal{R}_q(P)| \gtrsim_{\delta,m,n,Q,\Delta} \beta_* P^D q^{-E}. \tag{6.16}$$

Fix $R \in \mathcal{R}_q(P)$. For $x \in \Gamma_N$, define the coefficient set

$$B_{x,R} := \left\{ u \in [K]^s : x + \sum_{j=1}^{s} u_j r_j \in A \right\}.$$

Translation invariance of Γ_N gives

$$\mathbb{E}_{x\in\Gamma_N} \frac{|B_{x,R}|}{K^s} = \frac{|A|}{|\Gamma_N|} \ge \delta_0.$$

Consequently, $\gg_{\delta,n} |\Gamma_N|$ choices of x have $|B_{x,R}| \ge (\delta_0/2)K^s$. For each such x, choose a copy (6.14) and put

$$z = x + \sum_{j=1}^{s} a_j r_j, \qquad v_j = t r_j \ (1 \le j \le s), \qquad v_m = -t \sum_{j=1}^{s} r_j.$$

Then $z, z+v_1, \ldots, z+v_m$ all belong to A. Moreover, homogeneity gives

$$\Phi(v) = t^r \Phi(R) \equiv 0 \pmod q, \tag{6.17}$$

and (6.15) gives $\max_i |v_i|_\infty \le \theta P$.

The map $(R, x, a, t) \mapsto (z, v)$ has multiplicity at most K^{s+1}: after t and a are fixed, $R = v/t$ and $x = z - \sum_j a_j r_j$ are fixed. Taking η sufficiently small that $\theta\eta < 1$ also prevents wraparound in

Γ_N. Thus every modular configuration counted is an actual integer configuration. Summing over $R \in \mathcal{R}_q(P)$ and using (6.16), we obtain

$$\sum_{z} \sum_{\substack{v \in H^n_{m,\mathbb{Z}}, |v|_\infty \le \theta P \\ \Phi(v) \equiv 0 \pmod q}} \mathbf{1}_A(z) \prod_{i=1}^{m} \mathbf{1}_A(z+v_i) \gtrsim_{\delta,m,n,Q,\Delta} \beta_* N^n P^D q^{-E}. \tag{6.18}$$

Finally, (6.7), (6.13), and (6.18) give

$$\mathcal{M}^{\mathbb{Z}}_{q,L}(\mathbf{1}_A) \gtrsim N^{-n} q^E P^{-D} \big(\beta_* N^n P^D q^{-E}\big) \gtrsim_{\delta,m,n,Q,\Delta} \beta_* > 0,$$

which proves (6.11). □

For a smooth cutoff W on H^n_m, introduce the normalized oscillatory form

$$\Lambda_{\alpha,P}(f) := N^{-n} P^{-D} \sum_{z} \sum_{v \in H^n_{m,\mathbb{Z}}} W(v/P) f_0(z) \prod_{i=1}^{m} f_i(z+v_i) e(\alpha \cdot \Phi(v)).$$

Remark 6.2. We remark that for $m = 2$, i.e. for three-term progressions $x-y, x, x+y$, one can use Roth's theorem, which provides a quantitative bound. This simplifies much of the discussion in Section 7 and leads to a quick double-exponential bound for Theorem 7.2. Thus it would be desirable to obtain a quantitative lower bound for $m \ge 3$ as well, bypassing the multidimensional Szemerédi theorem.

Lemma 6.3. *For $M \ge 2$, define*

$$\mathfrak{M}_C(P; M) := \bigcup_{1 \le q \le M} \left\{ \alpha \in \mathbb{T}^E : \|q\, C_r \alpha\|_{\mathbb{T}^E,\infty} \le M P^{-r} \right\}.$$

Suppose that

$$E(E+1) < \kappa_0 < \kappa, \quad \textit{with} \quad \kappa = \frac{n}{(r-1)2^{r-1}}.$$

Let

$$\gamma := \frac{\kappa_0}{E} - (E+1) > 0.$$

Then, for every fixed $M_0 \ge 2$,

$$P^{rE} \int_{\mathbb{T}^E \setminus \mathfrak{M}_C(P;M_0)} |\Lambda_{\alpha,P}(f)|\, d\alpha \ll M_0^{-\gamma}, \tag{6.19}$$

uniformly for $|f_i| \le 1$.

Proof. For $H \ge 2$, put

$$\mathfrak{M}_1(P; H) := \bigcup_{1 \le q \le H} \left\{ \theta \in \mathbb{T} : \|q\theta\|_{\mathbb{T}} \le H P^{-r} \right\}.$$

Taking $H = P^{(r-1)\vartheta}$, these are, up to harmless constant changes, the major arcs $\mathfrak{M}_P(\vartheta)$ of Theorem 5.4. The standard dyadically "pruned" consequence of that lemma therefore gives, for every $\kappa_0 < \kappa$,

$$\|W_0(\cdot/P) e(\theta Q + R_{<r})\|_{U_P^{r-1}} \ll H^{-\kappa_0}, \quad \text{for} \quad \theta \notin \mathfrak{M}_1(P; H), \tag{6.20}$$

uniformly in the coefficients of $R_{<r}$ and for the uniformly smooth family of weights W_0 arising below. The remainder in (6.20) is uniform under the dyadic summation used below.

Fix an edge $a < b$ and all vertices other than v_a, v_b. Using

$$w = -\sum_{j \ne a,b} v_j, \qquad v_b = -u, \qquad v_a = w + u,$$

we have

$$\alpha \cdot \Phi(v) = (C_r\alpha)_{ab} Q(u) + R_{\alpha,v'}(u), \qquad \deg R_{\alpha,v'} < r.$$

The Cauchy–Schwarz argument of Theorem 5.3, followed by $U^3 \lesssim U^{r-1}$, gives

$$|\Lambda_{\alpha,P}(f)| \lesssim \mathbb{E}_{v'} \left\| W_{v'}(\cdot/P) e\big((C_r\alpha)_{ab} Q + R_{\alpha,v'}\big) \right\|_{U_P^{r-1}}. \tag{6.21}$$

Suppose that every coordinate $\beta_{ab} := (C_r\alpha)_{ab}$ belongs to $\mathfrak{M}_1(P;H)$. Multiplying the corresponding E denominators gives a common denominator $q \le H^E$ satisfying

$$\|q\, C_r \alpha\|_{\mathbb{T}^E} \le H^E P^{-r}.$$

Taking $H = M^{1/E}$, we conclude that

$$\alpha \notin \mathfrak{M}_C(P;M) \quad \Longrightarrow \quad (C_r\alpha)_{ab} \notin \mathfrak{M}_1(P;M^{1/E})$$

for at least one edge $a < b$. Hence (6.20)–(6.21) give

$$|\Lambda_{\alpha,P}(f)| \ll M^{-\kappa_0/E}, \qquad \text{for} \quad \alpha \notin \mathfrak{M}_C(P;M). \tag{6.22}$$

Since C_r is an integral non-singular matrix, each map $qC_r : \mathbb{T}^E \to \mathbb{T}^E$ is a surjective torus endomorphism. Its pullback preserves normalized Haar measure, and consequently

$$|\mathfrak{M}_C(P;M)| \lesssim \sum_{q \le M} (MP^{-r})^E \lesssim M^{E+1} P^{-rE}. \tag{6.23}$$

Thus, on the dyadic shell

$$\mathfrak{M}_C(P;2M) \setminus \mathfrak{M}_C(P;M),$$

we obtain

$$P^{rE} \int_{\mathfrak{M}_C(P;2M)\setminus \mathfrak{M}_C(P;M)} |\Lambda_{\alpha,P}(f)|\, d\alpha \lesssim M^{E+1-\kappa_0/E} = M^{-\gamma}. \tag{6.24}$$

Finally, simultaneous Dirichlet approximation gives

$$\mathfrak{M}_C(P;M_*) = \mathbb{T}^E \qquad \text{for} \qquad M_* \asymp P^{rE/(E+1)}.$$

Indeed, taking $H = P^{r/(E+1)}$, one obtains

$$q \le H^E = M_*, \qquad \|qC_r\alpha\|_\infty \le H^{-1} = M_* P^{-r}.$$

Summing (6.24) over $M = M_0, 2M_0, \dots, M_*$ proves (6.19). ☐

We can now estimate the high-frequency error term $\mathcal{E}^{\varepsilon,\mathbb{Z}}_{q,L}(f)$.

Proposition 6.4. *Assume* (1.13). *For every $\varepsilon > 0$, the error term in* (6.4) *satisfies*

$$|\mathcal{E}^{\varepsilon,\mathbb{Z}}_{q,L}(f)| \le C\varepsilon^c \tag{6.25}$$

uniformly for $|f| \le 1$, provided $q = q(\varepsilon)$, $q \mid L$, and $P_0(\varepsilon) \le P \ll N$, where $P_0(\varepsilon) \le \varepsilon^{-C}$ and $C = C_{n,m} > 0$. Moreover, the common modulus $q = q(\varepsilon)$ and the lower scale threshold $P_0(\varepsilon)$ may be chosen so that

$$\log q(\varepsilon) + \log P_0(\varepsilon) \le C\varepsilon^{-C}. \tag{6.26}$$

Proof. Choose M_0 to be a sufficiently large fixed power of ε^{-1}. By Theorem 6.3, the contribution outside $\mathfrak{M}_C(P;M_0)$ is $O(\varepsilon^c)$. Enlarge the common denominator $q(\varepsilon)$ by $|\det C_r|$ and by the denominators up to M_0, so that these vector major arcs are absorbed into the major term of (6.4). Since $P^{rE-D}\sum_v$ in (6.3) equals $P^{rE}(P^{-D}\sum_v)$, (6.19) gives (6.25) once $P \ge P_0(\varepsilon)$. The construction of M_0, $q(\varepsilon)$, and the finitely many small-scale cutoffs gives (6.26). ☐

Next, we deal with the sum of the intermediate terms. The key tool we use is Villaroya's transference argument for multi-linear multiplier operators.

Proposition 6.5. *For a lacunary sequence of levels L_j and physical scales $P_j = L_j^{1/r}$, $1 \le j \le J$,*

$$\sum_j |\mathcal{K}^{\varepsilon,\mathbb{Z}}_{q,L_j}(f)|^2 \lesssim_{m,Q,\Delta} \varepsilon^{-C} q^{D+2E}, \tag{6.27}$$

with a constant $C = C_{n,m}$, uniformly in the number of levels J and in $|f| \le 1$.

Proof. Expanding χ_q in its finite Fourier series on $H^n_{m,\mathbb{Z}}/qH^n_{m,\mathbb{Z}}$, each mode merely modulates the external functions. Thus it suffices to prove the estimate without χ_q, accepting an extra factor q^{D+2E} via a simple application of Cauchy-Schwarz.

After the same Cauchy–Schwarz step as in Theorem 5.7 as in the Euclidean case, the sampled kernel is the restriction to the lattice of K^ε_J in (5.27). Poisson summation writes its torus multiplier as a sum over

$$h \in (H^n_{m,\mathbb{Z}})^* \times (H^n_{m,\mathbb{Z}})^* \simeq \mathbb{Z}^{2D}$$

of the Euclidean multiplier. Insert fixed smooth cutoffs in every input frequency on a fundamental cube. The map L in (5.29) is surjective onto the two kernel-frequency spaces, so every fixed h translation can be absorbed into modulations of the external input functions.

For the finitely many bounded vectors h, view the scalar $2m$-linear form as a $(2m-1)$-linear operator

$$\prod_{i=1}^{2m-1} L^{2m} \longrightarrow L^{2m/(2m-1)}.$$

The continuous integral-rank MTT estimate [11] then transfers to the corresponding sequence spaces by Villarroya's multi-linear transference theorem [13, Theorem 3.3]. Note that the proof there is written in one dimension for notational simplicity, but the paper explicitly notes, and it is easy to see from its proof that it is unchanged for multi-linear operators on $\mathbb{R}^n$ and $\mathbb{Z}^n$.

For a large h, the cutoff multiplier $m_{J,h}$ obeys, for every fixed derivative order and every M,

$$\|\partial^\beta m_{J,h}\|_\infty \le C_{\beta,M,\varepsilon}(1+|h|)^{-M}, \tag{6.28}$$

uniformly in J. Sufficiently many derivatives give

$$\|\check{m}_{J,h}\|_{L^1} \le C_{M,\varepsilon}(1+|h|)^{-M}. \tag{6.29}$$

The physical-space kernel here lives on the full product of input variables. The resulting multi-linear norm has the same decay and is summable over h. This proves (6.27). ☐

We can now give the proof of Theorem 1.6.

Proof of Theorem 1.6. Let $c(\delta)$ be the q-independent constant in (6.11), and let

$$c_0 := \inf_{0<\varepsilon\le\varepsilon_0} c_\varepsilon > 0$$

for a sufficiently small fixed ε_0. Choose $\varepsilon = \varepsilon(\delta) \le \varepsilon_0$ so that the bound in (6.25) is at most $c_0 c(\delta)/2$, and then take the common denominator $q = q(\varepsilon)$ in (6.4). The uniform local conditions (1.19) and (1.17) control, respectively, the exact arithmetic normalization and the lower bound for the main term uniformly over the indicated levels and moduli.

Suppose there are arbitrarily large admissible levels, divisible by q, for which no configuration exists. Select a finite lacunary sequence $L_1 < \cdots < L_J$ of such levels, write $P_j = L_j^{1/r}$, and require

$$P_1 \ge \max\{Cq(\varepsilon), P_0(\varepsilon)\}.$$

Here C is the constant in Theorem 6.1. Choose a box of side $N \geq \eta^{-1}P_J$ on which A has density comparable to δ. Such a box may be chosen arbitrarily large by the definition of upper Banach density. Translate the box to $[1, N]^n$ and let f be the indicator of the corresponding translate of A inside that box. The exact count vanishes at every L_j. By Theorems 6.1 and 6.4 and (6.10),

$$|\mathcal{K}^{\varepsilon,\mathbb{Z}}_{q,L_j}(f)| \geq c_\varepsilon c(\delta) - |\mathcal{E}^{\varepsilon,\mathbb{Z}}_{q,L_j}(f)| \geq \frac{1}{2}c_0 c(\delta) =: c'(\delta) > 0. \quad (1 \leq j \leq J).$$

Thus the left side of (6.27) is at least $J(c'(\delta))^2$, contradicting Theorem 6.5 for large J. All sufficiently large admissible levels divisible by q therefore occur. □

7. A quantitative finitary variant

We prove a quantitative result about the existence of barycentric configuration in a finite set $A \subseteq [N]^n$ of density δ, which is Q-similar to a given simplex Δ_o. The proof will combine our basic decomposition with a density increment method. Doing this over a family of scales is technically involved, mainly due to the inverse results for Roth type affine configurations where one does not have good control of the scales at which the density of the underlying set increases. We handle this by assigning a large family scales to each density, which remains closed under the induction on the densities.

Definition 7.1 (Q-similar simplex–center copy)**.** Let $L \in \mathbb{N}$. A collection

$$z, \quad z + v_1, \ldots, z + v_m \in \mathbb{Z}^n$$

is a *Q-similar copy of Δ_o together with its center, at level L*, if

$$\sum_{i=1}^m v_i = 0, \qquad Q(v_i - v_j) = L\ell_{ij} \quad (i < j). \tag{7.1}$$

The corresponding physical scale is $P = L^{1/r}$. The terminology refers to the edge equations in (7.1); for a general form Q it does not assert that a linear map carries Δ_o to $(v_1, \ldots, v_m)$.

Theorem 7.2. *Assume the hypotheses of Theorem 1.6. There is a constant $C = C(m, n, Q, \Delta)$ such that, for every $0 < \delta \leq 1$ and every scale $R_0 \geq 1$, there are a finite set of levels $\mathcal{T}_{\delta,R_0} \subset \mathbb{N}$ and an integer $N_*(\delta, R_0)$ satisfying*

$$N_*(\delta, R_0) \leq \exp\bigl(\exp(C\delta^{-C})(1 + \log R_0)\bigr) = (eR_0)^{\exp(C\delta^{-C})}, \tag{7.2}$$

with the following property. If $N \geq N_*(\delta, R_0)$ and

$$A \subset [1, N]^n \cap \mathbb{Z}^n, \qquad |A| \geq \delta N^n,$$

then A contains a Q-similar copy of Δ_o together with its center at some level $L \in \mathcal{T}_{\delta,R_0}$. Moreover,

$$R_0 \leq L^{1/r} \leq N_*(\delta, R_0) \qquad (L \in \mathcal{T}_{\delta,R_0}). \tag{7.3}$$

7.1. **A quantitative block dichotomy.** Let us first record the three local inputs. For functions $f_0, \ldots, f_m$ we use the multi-linear versions of the forms introduced in Section 6, replacing the "diagonal" product by $f_0(z)\prod_{i=1}^m f_i(z + v_i)$.

Fix throughout

$$\theta := \frac{1}{n+1}. \tag{7.4}$$

Choose constants $c_*, C_* > 0$ once and for all, with c_* sufficiently small and C_* sufficiently large, and set

$$\kappa_\rho := c_*\rho^{C_*} \qquad (0 < \rho \leq 1). \tag{7.5}$$

These constants are set to be below the density increments obtained from Theorem 7.4.

To make the conditioning cells precise, put $M_H := \lfloor H/q \rfloor$ and, for every $a \in (\mathbb{Z}/q\mathbb{Z})^n$, partition $a + q\mathbb{Z}^n$ into scalar grids

$$C = a + q\big(b + [1, M_H]^n\big), \qquad b \in M_H\mathbb{Z}^n.$$

Let $\mathcal{C}_{H,q}$ be the cells contained in $[1,N]^n$ and discard the boundary cells, whose union has size $O(HN^{n-1})$. Write

$$d_A(C) := \frac{|A \cap C|}{|C|}, \qquad h_{H,q} := \sum_{C \in \mathcal{C}_{H,q}} d_A(C)\mathbf{1}_C.$$

If φ is a fixed nonnegative Schwartz function with $\int \varphi = 1$ and $\widehat{\varphi}$ supported in a sufficiently small cube, let

$$\varphi_{H,q}(x) := q^n H^{-n}\mathbf{1}_{q\mathbb{Z}^n}(x)\varphi(x/H), \qquad F_{H,q} := \mathbf{1}_A * \varphi_{H,q}. \tag{7.6}$$

If $T = H/q \geq 2$, then

$$\psi_T(y) := \varphi_{H,q}(qy) = T^{-n}\varphi(y/T), \qquad \sum_{y \in \mathbb{Z}^n} \psi_T(y) = 1,$$

the last identity following from Poisson summation and the Fourier support of φ. Thus convolution by ψ_T is a positive mean-preserving operator on each normalized residue-class grid. If $f = \mathbf{1}_A$ and $f_a(y) = f(a + qy)$, then

$$F_{H,q}(a + qy) = f_a * \psi_T(y), \qquad 0 \leq F_{H,q} \leq 1.$$

Choosing φ as a positive majorant and absorbing its rapidly decaying tails, one has $F_{H,q} \gtrsim h_{H,q}$ away from the boundary. For later use, put

$$\mathcal{S}_q := \{u \in H^n_{m,\mathbb{Z}}/qH^n_{m,\mathbb{Z}} : \Phi(u) \equiv 0 \pmod{q}\}.$$

The following lemma is similar in spirit to that of [10, Lemma 3.1]

Lemma 7.3 (Asymmetrically smoothed coarse main term)**.** *Let $A \subset [1,N]^n$ have density at least ρ, and put $t_0 = m + 1$. There are effective constants $c^*_\rho, C^*_\rho > 0$ such that the following holds. Let $P = L^{1/r}$, let $q \mid L$, and choose*

$$P \geq C^*_\rho q, \qquad H = C^*_\rho P, \qquad H \leq c^*_\rho N.$$

If $d_A(C) \leq \rho + \kappa_\rho$ for every $C \in \mathcal{C}_{H,q}$, then

$$\mathcal{M}^{\mathbb{Z}}_{q,L}(\mathbf{1}_A, F_{H,q}, \ldots, F_{H,q}) \geq c\,\beta_q(0)\rho^{t_0}, \tag{7.7}$$

where $c > 0$ depends only on the initial configuration Δ_o. Moreover,

$$\log C^*_\rho + \log(c^*_\rho)^{-1} \leq \exp(C\rho^{-C}). \tag{7.8}$$

Proof. For each $u \in \mathcal{S}_q$, split the base point into residue classes modulo q. The shifts in the support of ω_P have size $O(P)$ and hence are $O_\rho(H)$. The discarded boundary contributes $O(H/N)$, which is absorbed by the choice of c^*_ρ. By our assumption, positivity of $F_{H,q}$, and the standard Varnavides averaging give a lower bound $c\rho^{t_0}$ for (7.7) with $h_{H,q}$. Only the affine relations

$$z, \quad z + v_1, \ldots, z + v_{m-1}, \quad z - v_1 - \cdots - v_{m-1}$$

are used; no independence of the vectors v_i is required.

There are $|\mathcal{S}_q| = \beta_q(0)q^{D-E}$ vectors $v = (v_1, \ldots, v_m)$ satisfying $\Phi(v) \equiv 0 \pmod{q}$. Their density q^{-E} cancels the factor q^E of χ_q in (6.7), while P^{-D} cancels the number of lifts. This gives the factor $\beta_q(0)$ in (7.7); compare (6.16)–(6.18) and [10]. $\square$

For a function G on $[1,N]^n$ and $a \in (\mathbb{Z}/q\mathbb{Z})^n$, set $G_a(y) := G(a+qy)$ and extend G_a by zero. We use the standard local U^2 norm at scale S and, crucially, retain the average over all residue classes:

$$\|G_a\|_{U^2(S)}^4 := \mathbb{E}_x \mathbb{E}_{h,k\in[-S,S]^n} G_a(x)\overline{G_a(x+h)G_a(x+k)}G_a(x+h+k),$$
$$\mathcal{U}_{S,q}(G)^4 := \mathbb{E}_{a \bmod q} \|G_a\|_{U^2(S)}^4. \tag{7.9}$$

Here all averages have the usual normalized finite-box interpretation; changing the boundary convention changes the quantities by $O(qS/N)$.

Lemma 7.4 (Local U^2 increment on a scalar grid)**.** *Let $A \subset [1,N]^n$ have density $\bar\rho := |A|/N^n \geq \rho$, let $H/q, S \geq 2$, and suppose that*

$$\mathcal{U}_{S,q}(\mathbf{1}_A - F_{H,q}) \geq \eta. \tag{7.10}$$

Then there are integers $s, M \geq 1$ and a translate b such that

$$b + qs[1,M]^n \subset [1,N]^n,$$

$$s \leq C_\eta S^{1-\theta}, \qquad M \geq c_\eta S^\theta, \qquad sM \leq C_\eta S, \tag{7.11}$$

and

$$\frac{|A \cap (b+qs[1,M]^n)|}{M^n} \geq \bar\rho + c\eta^C - C_\eta \frac{H+qS}{N}. \tag{7.12}$$

In particular, if $H + qS \leq c_\eta N$, then, after adjusting the constants, the right side is at least $\bar\rho + c\eta^C \geq \rho + c\eta^C$. The constants may be chosen so that $\log C_\eta + \log c_\eta^{-1} \ll \eta^{-C}$.

Proof. Put $f = \mathbf{1}_A$, $T = H/q$, and

$$\Omega_a := \{y \in \mathbb{Z}^n : a + qy \in [1,N]^n\}, \qquad U_a := \|(f - F_{H,q})_a\|_{U^2(S)}.$$

The hypothesis says $\mathbb{E}_a U_a^4 \geq \eta^4$. Since $U_a \leq 1$, thresholding shows that $U_a \geq \eta/2$ on a set of residue classes of measure $\gg \eta^4$. Apply the standard Fourier–Dirichlet U^2 inverse argument with the individual norm U_a in those fibres, and choose any admissible grid with the same bounds in the remaining fibres. Averaging the resulting lower bounds (and enlarging the absolute exponent C) gives scalar grids $\mathcal{R}_a = s_a[1,M_a]^n$ satisfying (7.11). For these grids put

$$\mathcal{T}_a := \{t \in \mathbb{Z}^n : \mathcal{R}_a + t \subset \Omega_a\}, \qquad \mathbb{E}_{a,t} := \mathbb{E}_{a \bmod q}\mathbb{E}_{t\in\mathcal{T}_a}.$$

Then

$$\mathbb{E}_{a,t}\left|\mathbb{E}_{x\in\mathcal{R}_a+t}\big(f_a - f_a * \psi_T\big)(x)\right| \geq c\eta^C - O_\eta(qS/N), \tag{7.13}$$

The usual proof permits the choices s_a, M_a to depend on a; this causes no difficulty because the residue average has not been discarded.

Set

$$d_a(t) := \mathbb{E}_{x\in\mathcal{R}_a+t} f_a(x) \quad (t \in \mathbb{Z}^n), \qquad K_T d_a := d_a * \psi_T.$$

Here f_a is extended by zero, so d_a and hence $K_T d_a$ are defined for every t, including those near the boundary of $\mathcal{T}_a$. Since convolution commutes with averaging over a translated grid,

$$\mathbb{E}_{x\in\mathcal{R}_a+t}(f_a * \psi_T)(x) = K_T d_a(t).$$

Thus the inner average in (7.13) is $\alpha_a(t) := d_a(t) - K_T d_a(t)$. Averaging over all residue classes and all translates gives

$$\mathbb{E}_{a,t} d_a(t) = \bar\rho + O_\eta(qS/N), \qquad \mathbb{E}_{a,t} K_T u_a(t) = \mathbb{E}_{a,t} u_a(t) + O_\eta(H/N) \quad (|u_a| \leq 1).$$

The second identity holds for every bounded family $u_a : \mathbb{Z}^n \to \mathbb{C}$ and follows by translating $\mathcal{T}_a$ under the Markov kernel; only a boundary layer of physical width $O(H)$ is lost. Consequently (7.13) implies

$$\mathbb{E}_{a,t}(\alpha_a(t))_+ \geq c\eta^C - O_\eta((H+qS)/N). \tag{7.14}$$

On the other hand, positivity and constant preservation of K_T give the pointwise inequality

$$(d_a - K_T d_a)_+ \leq (d_a - \bar{\rho})_+ + K_T\big((\bar{\rho} - d_a)_+\big).$$

If $\Delta := \sup_{a,t}(d_a(t) - \bar{\rho})$, averaging this inequality and using the preceding mean identities yields

$$\mathbb{E}_{a,t}(\alpha_a(t))_+ \leq 2\Delta + O_\eta((H+qS)/N).$$

Together with (7.14), this gives $\Delta \geq c\eta^C - C_\eta(H+qS)/N$. Selecting the corresponding a,t and returning to physical coordinates gives $b + qs_a[1, M_a]^n$, with $b = a + qt$, and proves (7.12). ☐

The other input is the explicit form of Theorem 6.5: for every physically lacunary sequence of levels,

$$\sum_j |\mathcal{K}^{\varepsilon,\mathbb{Z}}_{q,L_j}(f)|^2 \lesssim \varepsilon^{-C} q^{D+2E}, \tag{7.15}$$

uniformly in the number of levels and in $|f| \leq 1$.

Proposition 7.5 (Quantitative block dichotomy)**.** *For every $0 < \rho \leq 1$ there are an integer q_ρ, a block length J_ρ, and effective constants $b_\rho, C_\rho > 0$ and $0 < c_\rho \leq 1$ such that*

$$\log q_\rho \leq \rho^{-C}, \qquad J_\rho \leq \exp(\rho^{-C}), \tag{7.16}$$

and the following holds. For every $X \geq 2$ one can choose levels $L_1 < \cdots < L_{J_\rho}$, divisible by q_ρ, whose physical scales $P_j = L_j^{1/r}$ satisfy

$$b_\rho q_\rho X^{1/\theta} 4^{j-1} \leq P_j \leq 2 b_\rho q_\rho X^{1/\theta} 4^{j-1}. \tag{7.17}$$

If $A \subset [1,N]^n$ has density at least ρ and

$$P_{J_\rho} \leq c_\rho N, \tag{7.18}$$

then either

1. *A contains a simplex–center copy at one of the levels L_j; or*
2. *for some j there are integers $s, M \geq 1$ and a translate a such that*

$$a + q_\rho s[1,M]^n \subset [1,N]^n,$$

$$s \leq C_\rho (P_j/q_\rho)^{1-\theta}, \qquad M \geq X, \tag{7.19}$$

and

$$\frac{|A \cap (a + q_\rho s[1,M]^n)|}{M^n} \geq \rho + \kappa_\rho. \tag{7.20}$$

Moreover,

$$\log b_\rho + \log c_\rho^{-1} + \log C_\rho \leq \exp(C\rho^{-C}). \tag{7.21}$$

Proof. Take $\varepsilon_\rho = c_0\, \rho^{C_0}$, with $c_0 > 0$ small and C_0 large enough that Theorem 6.4 gives

$$|\mathcal{E}^{\varepsilon_\rho,\mathbb{Z}}_{q(\varepsilon_\rho),L}(\mathbf{1}_A)| \leq \frac{c}{8}\beta_* \rho^{t_0}, \qquad t_0 = m+1. \tag{7.22}$$

Fix also

$$\eta_\rho := c_1 \rho^{t_0}, \tag{7.23}$$

where $c_1 > 0$ is sufficiently small. The constants in (7.5) are chosen so that κ_ρ is below the increment $c\eta_\rho^C$ supplied by Theorem 7.4. Choose the common denominator $q_\rho = q(\varepsilon_\rho)$ in (6.4). The construction in Theorem 6.4 gives $\log q_\rho \le \rho^{-C}$, and its lower scale threshold is absorbed into b_ρ. Choose the fit constant in the proposition so that

$$c_\rho \le \min\left\{\frac{c_\rho^*}{C_\rho^*}, \frac{c_{\eta_\rho}}{C_\rho^*+1}\right\}. \tag{7.24}$$

Shrink c_ρ further, if necessary, to meet the upper-scale hypotheses in Theorem 6.4. Thus $P_j \le c_\rho N$ controls both the $C_\rho^* P_j/N$ and P_j/N boundary errors. Choose b_ρ sufficiently large to absorb the lower threshold $P_0(\varepsilon_\rho)$, the condition $P_1 \ge C_\rho^* q_\rho$, the constants in Theorem 7.4, the nesting inequality (7.32) below, and the requirement $M \ge X$.

For each j, let $H_j = C_\rho^* P_j$ and $F_j = F_{H_j,q_\rho}$. If, for some j, a cell $C \in \mathcal{C}_{H_j,q_\rho}$ has density at least $\rho + \kappa_\rho$, then, since $C = b + q_\rho[1, M_{H_j}]^n$, it gives the second alternative with $s = 1$. Our choice of b_ρ ensures $M_{H_j} \ge X$. Otherwise Theorem 7.3 applies for every j. If the first alternative fails, every level L_j is missing, and for each j we write (6.10) as

$$0 = \mathcal{B}_j + \mathcal{D}_j + \mathcal{K}_j + \mathcal{E}_j, \tag{7.25}$$

where

$$\begin{aligned}
\mathcal{B}_j &:= c_{\varepsilon_\rho}\mathcal{M}^{\mathbb{Z}}_{q_\rho,L_j}(\mathbf{1}_A, F_j, \dots, F_j),\\
\mathcal{D}_j &:= c_{\varepsilon_\rho}\left[\mathcal{M}^{\mathbb{Z}}_{q_\rho,L_j}(\mathbf{1}_A, \dots, \mathbf{1}_A) - \mathcal{M}^{\mathbb{Z}}_{q_\rho,L_j}(\mathbf{1}_A, F_j, \dots, F_j)\right].
\end{aligned}$$

Thus

$$\mathcal{B}_j \ge c\,\beta_{q_\rho}(0)\rho^{t_0}. \tag{7.26}$$

Telescoping the m replacements in $\mathcal{D}_j$ and decomposing both the frame and the base point into residue classes modulo q_ρ, a typical term has the normalized form

$$\beta_{q_\rho}(0)\mathbb{E}_{u\in\mathcal{S}_{q_\rho}}\mathbb{E}_{a \bmod q_\rho}\Lambda_{u,a}(F_0, \dots, \mathbf{1}_A - F_j, \dots, F_m), \tag{7.27}$$

where $\Lambda_{u,a}$ is an unweighted affine form on a q_ρ-grid at scale P_j/q_ρ. Thus the height q_ρ^E of χ_{q_ρ} has canceled against the q_ρ^{-E} density of null residue frames v satisfying $\Phi(v) \equiv 0 \pmod{q_\rho}$.

The affine system satisfying $v_1 + \cdots + v_m = 0$ has complexity one. Put

$$S_j := \lfloor P_j/q_\rho \rfloor.$$

The localized generalized von Neumann inequality, with all residue-class averages in (7.27) retained, therefore gives the following estimate. Indeed, for fixed u, translating a by the residue of the balanced slot merely permutes the residue classes, so the u-average does not alter the displayed norm:

$$\mathbb{E}_{a \bmod q_\rho}\|(\mathbf{1}_A - F_j)_a\|_{U^2(S_j)} \gtrsim \rho^{t_0}$$

whenever

$$|\mathcal{D}_j| > \frac{c}{4}\beta_{q_\rho}(0)\rho^{t_0}, \tag{7.28}$$

By Jensen's inequality this implies

$$\mathcal{U}_{S_j,q_\rho}(\mathbf{1}_A - F_j) \ge \eta_\rho.$$

Now Theorem 7.4 gives the second alternative. The choice of c_ρ absorbs the error in (7.12), while the choice of b_ρ ensures $M \ge X$ and gives the bound for s in (7.19).

If neither alternative occurs, then (7.26), (7.22), and (7.25) imply

$$|\mathcal{K}_j| \ge c'\beta_*\rho^{t_0} \qquad (1 \le j \le J_\rho).$$

On the other hand, (7.15) gives

$$\sum_{j=1}^{J_\rho} |\mathcal{K}_j|^2 \lesssim \varepsilon_\rho^{-C} q_\rho^{D+2E}.$$

It is therefore enough to take

$$J_\rho > C\varepsilon_\rho^{-C} q_\rho^{D+2E} (c'\beta_* \rho^{t_0})^{-2}. \tag{7.29}$$

This choice satisfies $J_\rho \le \exp(\rho^{-C})$.

Finally, with $Y_j = b_\rho q_\rho X^{1/\theta} 4^{j-1}$, choose

$$L_j = q_\rho \left\lceil Y_j^r / q_\rho \right\rceil.$$

With this fixed choice of b_ρ, these levels satisfy (7.17); in particular their physical scales are lacunary. Since the logarithmic width of the block is $O(J_\rho)$, all constants obey (7.21). This proves the proposition. □

7.2. **Construction of the finite tree of scales.**

Proof of Theorem 7.2. We first define some fixed parameters independently of the set which will allow "enough room" to define a family of scales for our induction on the density arguments. Let $\rho_0 = \delta$ and define

$$\rho_{h+1} := \rho_h + \kappa_{\rho_h}. \tag{7.30}$$

Let h_* be the first index for which $\rho_{h_*} > 1$. Since $\kappa_\rho = c_* \rho^{C_*}$,

$$h_* \le \delta^{-C}. \tag{7.31}$$

Write $q_h = q_{\rho_h}$ and $J_h = J_{\rho_h}$.

The constants b_ρ in Theorem 7.5 are chosen large enough that, whenever P satisfies the lower bound in (7.17), we have

$$q_\rho \left\lceil C_\rho (P/q_\rho)^{1-\theta} \right\rceil X \le P, \tag{7.32}$$

which follows from $P \ge b_\rho q_\rho X^{1/\theta}$. Let

$$B_\rho := 2c_\rho^{-1} b_\rho q_\rho 4^{J_\rho - 1} + 2. \tag{7.33}$$

Then (7.16) and (7.21) give

$$\log B_\rho \le \exp(C\rho^{-C}). \tag{7.34}$$

Set

$$X_{h_*} := \max\{2, R_0\}. \tag{7.35}$$

For $h = h_* - 1, h_* - 2, \dots, 0$, define

$$X_h := B_{\rho_h} X_{h+1}^{1/\theta}, \tag{7.36}$$

and then apply Theorem 7.5 with $X = X_{h+1}$ to choose a block of lacunary scales

$$L_{h,1} < \cdots < L_{h,J_h}, \qquad P_{h,j} := L_{h,j}^{1/r}.$$

The explicit upper bound in (7.17) and the definition of B_{ρ_h} give

$$P_{h,J_h} \le c_{\rho_h} X_h, \tag{7.37}$$

while (7.32) gives, for

$$s_{h,j}^{\max} := \left\lceil C_{\rho_h} (P_{h,j}/q_h)^{1-\theta} \right\rceil, \tag{7.38}$$

one has the relation

$$q_h s_{h,j}^{\max} X_{h+1} \le P_{h,j} \qquad (0 \le h < h_*,\ 1 \le j \le J_h). \tag{7.39}$$

This is precisely the reason for choosing the parent block at physical scale comparable to $q_h X_{h+1}^{1/\theta}$.

We now build the "decision" tree. Let $\mathcal{V}_0$ consist of a single root $\varnothing$, and put $g_\varnothing = 1$. Having constructed the finite set $\mathcal{V}_h$ of nodes of generation h, attach to every $\nu \in \mathcal{V}_h$ one child (ν, j, s) for every

$$1 \le j \le J_h, \qquad 1 \le s \le s_{h,j}^{\max},$$

and define

$$g_{(\nu,j,s)} := g_\nu q_h s. \tag{7.40}$$

Translations and side lengths do not label formal children: they affect the location and available room of an actual increment box, but not the level of a copy lifted back to the original coordinates.

Define the predetermined root-level family

$$\mathcal{T}_{\delta,R_0} := \{g_\nu^r L_{h,j} : 0 \le h < h_*,\ \nu \in \mathcal{V}_h,\ 1 \le j \le J_h\}. \tag{7.41}$$

This family is finite and depends only on δ, R_0 and the fixed configuration, not on A.

Let $A \subset [1, N]^n$ have density at least δ, where $N \ge X_0$, and suppose for a contradiction that A avoids every level in (7.41). We follow a path down the already constructed tree. At an actual node $\nu \in \mathcal{V}_h$ we keep an affine scalar embedding

$$\iota_\nu(x) = b_\nu + g_\nu x$$

and a normalized set

$$A_\nu := \{x \in [1, N_\nu]^n : \iota_\nu(x) \in A\}$$

satisfying

$$|A_\nu| \ge \rho_h N_\nu^n, \qquad N_\nu \ge X_h. \tag{7.42}$$

At the root take $b_\varnothing = 0$, $g_\varnothing = 1$, and $N_\varnothing = N$.

If A_ν contained a copy at a direct level $L_{h,j}$, homogeneity would give

$$Q\big(g_\nu(v_i - v_j)\big) = g_\nu^r L_{h,j} \ell_{ij},$$

so A would contain a copy at the level $g_\nu^r L_{h,j}$ in (7.41). Hence all levels in the block are missing from A_ν. The fit condition follows from (7.37) and (7.42), so Theorem 7.5 supplies, for some j, s, a grid

$$a + q_h s[1, M]^n \subset [1, N_\nu]^n, \qquad M \ge X_{h+1},$$

on which the normalized density is at least ρ_{h+1}. Select the formal child $\nu' = (\nu, j, s)$ and put

$$b_{\nu'} := b_\nu + g_\nu a, \qquad g_{\nu'} := g_\nu q_h s, \qquad N_{\nu'} := M.$$

The corresponding set $A_{\nu'}$ satisfies (7.42) at generation $h+1$. Notice that the full M-box is retained; no passage to a subbox of exactly side X_{h+1} is needed.

After h_* such forward steps, (7.42) would give a set of density at least $\rho_{h_*} > 1$, a contradiction. Thus A contains a copy at a level in (7.41).

It remains to bound the scales. Along a child edge generated by (j, s), (7.39) gives

$$g_{\nu'} X_{h+1} = g_\nu q_h s X_{h+1} \le g_\nu P_{h,j}. \tag{7.43}$$

Since (7.37) also gives $P_{h,j} \le X_h$, induction down the formal tree yields the exact invariant

$$g_\nu X_h \le X_0 \qquad (\nu \in \mathcal{V}_h). \tag{7.44}$$

Hence every physical test scale $g_\nu P_{h,j}$ in (7.41) is at most X_0. On the other hand, (7.39) gives $P_{h,j} \ge X_{h+1} \ge X_{h_*} \ge R_0$, while $g_\nu \ge 1$. Therefore $g_\nu P_{h,j} \ge R_0$ for every test level.

Write $x_h = \log X_h$ and $d_\theta = 1/\theta = n+1$. From (7.36) and (7.34),

$$x_h \le d_\theta x_{h+1} + \exp(C\rho_h^{-C}). \tag{7.45}$$

Since $h_* \le \delta^{-C}$ and $\rho_h \ge \delta$, iteration gives

$$x_0 \le \exp(C\delta^{-C})(1+\log R_0).$$

Taking $N_*(\delta, R_0) = \lceil X_0 \rceil$ proves (7.2) and (7.3), and completes the proof of Theorem 7.2. Taking $R_0 = 1$ proves Theorem 1.7. □

Remark 7.6. The key observation is that the refinement parameter s can make the tree very broad, but it does not enter its depth or the maximum-scale recurrence (7.45). The construction of the tree is necessary because we do not have a direct quantitative lower bound, independent of the modulus q, for the main term in the discrete case when $m \ge 3$, owing to the congruence restriction $\Phi(v) \equiv 0 \pmod{q}$. Such estimates may be possible and would make the quantitative argument substantially simpler and more similar to the Euclidean case.

References

[1] B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. Ser. A **265** (1961/62), 245–263.
[2] J. Bourgain, *A Szemerédi type theorem for sets of positive density in* $\mathbb{R}^k$, Israel J. Math. **54** (1986), 307–316.
[3] B. Cook, A. Magyar, and 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 and V. Kovač, *A Szemerédi-type theorem for subsets of the unit cube*, Anal. PDE **15** (2022), no. 2, 507–549.
[5] P. Erdös, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, E. G. Straus, *Euclidean Ramsey theorems. I.* Journal of Combinatorial Theory, Series A, **14** no, 3, (1973) 341-363.
[6] C. Demeter, M. Pramanik, and C. Thiele, *Multilinear singular operators with fractional rank*, Pacific J. Math. **246** (2010), no. 2, 293–324.
[7] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291.
[8] R. L. Graham *Recent trends in Euclidean Ramsey theory*, Discrete Mathematics, **136** no. 1-3 (1994), 119-127.
[9] T. Kościuszko, *Invariant Equations in Many Variables*, Electron. J. Combin., **32**, no. 3 (2025) 3-24
[10] A. Magyar, *k-point configurations in sets of positive density of* $\mathbb{Z}^n$, Duke Math. J. **146** (2009), 1–34.
[11] C. Muscalu, T. Tao, and C. Thiele, *Multi-linear operators given by singular multipliers*, J. Amer. Math. Soc. **15** (2002), no. 2, 469–496.
[12] R. Raghavan *Improved Bounds for 3-Progressions*, arXiv.2603.27045 (2026)
[13] P. Villarroya, *On boundedness of discrete multilinear singular integral operators*, J. Math. Anal. Appl. **382** (2011), no. 2, 534–548.
[14] T. D. Wooley, *Nested efficient congruencing and relatives of Vinogradov's mean value theorem*, Proceedings of the London Mathematical Society 118, no. 4 (2019): 942-1016.