

Finite configurations in sparse sets

Vincent Chan, Izabella Łaba, Malabika Pramanik

July 2, 2013

Abstract

Let $E \subseteq \mathbb{R}^n$ be a closed set of Hausdorff dimension α . For $m \geq n$, let $\{B_1, \dots, B_k\}$ be $n \times (m - n)$ matrices. We prove that if the system of matrices B_j is non-degenerate in a suitable sense, α is sufficiently close to n , and if E supports a probability measure obeying appropriate dimensionality and Fourier decay conditions, then for a range of m depending on n and k , the set E contains a translate of a non-trivial k -point configuration $\{B_1 y, \dots, B_k y\}$. As a consequence, we are able to establish existence of certain geometric configurations in Salem sets (such as parallelograms in \mathbb{R}^n and isosceles right triangles in \mathbb{R}^2). This can be viewed as a multidimensional analogue of the result of [25] on 3-term arithmetic progressions in subsets of \mathbb{R} .

Contents

1	Introduction	2
1.1	Notation and definitions	4
1.2	The main result	5
1.3	Examples	5
1.4	Outline of proof	8
1.5	Acknowledgement	10
2	A functional multilinear form for counting configurations	10
3	Extending the multilinear form to measures	11
4	Counting geometric configurations in sparse sets	15
4.1	Existence of candidate ν	15
4.2	Proof of Proposition 4.1(a)	18
4.3	Proof of Proposition 4.1(b)	18
4.4	Proof of Proposition 4.1(c)	19
5	Absolutely continuous estimates	23
5.1	Preliminaries	23
5.2	Almost periodic functions	25
5.3	Ubiquity of almost periodic functions	27

5.4	Proof of Proposition 5.1	30
5.5	Quantitative Szemerédi bounds fail for general \mathbb{A}	32
6	Proof of the main theorem	32
7	Examples	34
A	Approximate Identity	38
B	Supplementary facts from §5	41

Keywords and phrases: Finite point configurations, systems of linear equations, Salem sets, Hausdorff dimension, Fourier dimension

2010 Mathematics Subject Classification: 28A78, 42A32, 42A38, 42A45, 11B25

1 Introduction

This paper is a contribution to the study of Szemerédi-type problems in continuous settings in Euclidean spaces. Specifically, given a class of subsets of \mathbb{R}^n that are “large” in a certain sense, one may ask whether every set in this class must contain certain geometric configurations. The precise meaning of this will vary. For instance, given a fixed set $F \subset \mathbb{R}^n$ (usually discrete, but not necessarily finite), one could ask if there is a geometrically similar copy of F contained in every set E in a given class; or one could ask if every such E contains solutions to a given system of linear equations. This could be viewed as continuous analogues of Szemerédi’s theorem on arithmetic progressions in sets of integers of positive upper asymptotic density [32], or of its multidimensional variants [9].

It is an easy consequence of the Lebesgue density theorem that any set $E \subset \mathbb{R}^n$ of positive Lebesgue measure contains a similar copy of any finite set F . A famous conjecture of Erdős [6] states that for any infinite set of real numbers F , there exists a set E of positive measure which does not contain any non-trivial affine copy of F . Falconer [8] shows using a Cantor-like construction that if $F \subseteq \mathbb{R}$ contains a slowly decaying sequence $\{x_n\}$ such that $x_n \searrow 0$ and $\liminf x_{i+1}/x_i = 1$, there exists a closed set $E \subseteq \mathbb{R}$ of positive measure which does not contain any affine copies of F . For other classes of negative examples and related results, see [5], [16], [23], [24]. For faster decaying sequences, such as the geometric sequence $\{2^{-i}\}$, the question remains open.

We will focus on the case when the configuration is finite, but E has Lebesgue measure 0, making a trivial resolution of the problem impossible. Instead, the assumptions on the size of E will be given in terms of its Hausdorff dimension $\dim_H(E)$. Recall that by Frostman’s lemma, we have

$$\dim_H(E) = \sup \left\{ \alpha \in [0, n] : \exists \mu \in \mathcal{M}(E) \text{ with } \sup_{\varepsilon > 0} \frac{\mu(B(x, \varepsilon))}{\varepsilon^\alpha} < \infty \right\}, \quad (1.1)$$

where $\mathcal{M}(E)$ is the set of probability measures supported on E .

Consider the first non-trivial case where F consists of three points. Then we have the following result in \mathbb{R} .

Theorem 1.1 (Keleti, [21]). *For a given distinct triple of points $\{x, y, z\}$, there exists a compact set in \mathbb{R} with Hausdorff dimension 1 which does not contain any similar copy of $\{x, y, z\}$.*

Keleti also proves the existence of sets $E \subset \mathbb{R}$ with $\dim_H(E) = 1$ that avoid all “one-dimensional parallelograms” $\{x, x + y, x + z, x + y + z\}$, with $y, z \neq 0$ [20], or all similar copies of any 3-point configuration from a given sequence $\{(x_i, y_i, z_i)\}_{i=1}^\infty$ [21].

In higher dimensions, there is a range of results of this type. For instance, we have the following.

Theorem 1.2 (Maga, [26]). *For distinct points $x, y, z \in \mathbb{R}^2$, there exists a compact set in \mathbb{R}^2 with Hausdorff dimension 2 which does not contain any similar copy of $\{x, y, z\}$.*

Further examples will be given in Section 1.3.

Additive combinatorics suggests that sets E that are “pseudorandom” in an appropriate sense should be better behaved with regard to Szemerédi-type phenomena than generic sets of the same size. For example, it is well known that there are sets $A \subset \{1, 2, \dots, N\}$ of cardinality at least $C_\varepsilon N^{1-\varepsilon}$ for all ε such that A does not contain any non-trivial 3-term arithmetic progression (see [31], [1]). On the other hand, there are Szemerédi-type results available for sets of integers of zero asymptotic density if additional randomness or pseudorandomness conditions are assumed, see e.g. [22], [12]. Different types of pseudorandomness conditions have been used in the literature, depending on the context and especially on the type of configurations being sought. In the particular case of 3-term arithmetic progressions, the appropriate conditions take a Fourier analytic form [29]; a more recent (and much deeper) result is that Fourier analytic conditions can also be used to control more general finite configurations of “true complexity” [10].

In the continuous setting, this leads us to considering the Fourier dimension of a set $E \subset \mathbb{R}^n$, defined as

$$\dim_F(E) = \sup\{\beta \in [0, n] : \exists \mu \in \mathcal{M}(E) \text{ with } \sup_{\xi \in \mathbb{R}^n} |\widehat{\mu}(\xi)| (1 + |\xi|)^{\beta/2} < \infty\}. \quad (1.2)$$

Here, $\widehat{\mu}(\xi) = \int e^{-2\pi i x \cdot \xi} d\mu(x)$. It is well-known that $\dim_F(E) \leq \dim_H(E)$ for all $E \subseteq \mathbb{R}^n$. Strict inequality is possible and common, for instance the middle-thirds Cantor set has Fourier dimension 0 and Hausdorff dimension $\log 2 / \log 3$. When $\dim_F(E) = \dim_H(E)$, we say that E is a *Salem set*. Most of the known constructions of Salem sets are probabilistic, see e.g. [30], [18], [2], [3], [25]; deterministic examples are in [17], [19].

It turns out that a suitable combination of Hausdorff and Fourier dimensionality conditions does indeed force the presence of three term progressions in subsets of \mathbb{R} .

Theorem 1.3 (Łaba, Pramanik [25]). *Suppose $E \subseteq [0, 1]$ is a closed set which supports a probability measure μ with the following properties:*

- (i) $\mu([x, x + \varepsilon]) \leq C_1 \varepsilon^\alpha$ for all $0 < \varepsilon \leq 1$,

(ii) $|\widehat{\mu}(\xi)| \leq C_2(1 - \alpha)^{-B}|\xi|^{-\beta/2}$ for all $\xi \neq 0$,

where $0 < \alpha < 1$ and $2/3 < \beta \leq 1$. If $\alpha > 1 - \varepsilon_0(C_1, C_2, B, \beta)$, then E contains a 3-term arithmetic progression.

While Theorem 1.3 is stated and proved in [25] only for arithmetic progressions, the same proof works for any fixed 3-point configuration $\{x, y, z\}$. The assumptions (i), (ii) are not equivalent to the requirement that E be a Salem set, but many constructions of Salem sets produce also a measure that satisfies these conditions; this is discussed in [25] in more detail.

The present paper may be seen as an extension of Theorem 1.3 to \mathbb{R}^n , and to a larger class of patterns, by showing that certain point configurations are realized by all sets in \mathbb{R}^n supporting a measure that satisfies a ball condition of type (1.1) and a Fourier decay condition of type (1.2).

1.1 Notation and definitions

Definition 1.4. Fix integers $n \geq 2$, $k \geq 3$ and $m \geq n$. Suppose B_1, \dots, B_k are $n \times (m - n)$ matrices.

- (a) We say E contains a k -point \mathbb{B} -configuration if there exists $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^{m-n} \setminus \{0\}$ such that $\{x + B_j y\}_{j=1}^k \subseteq E$.
- (b) Given any finite collection of subspaces $V_1, \dots, V_q \subseteq \mathbb{R}^{m-n}$ with $\dim(V_i) < m - n$, we say that E contains a non-trivial k -point \mathbb{B} -configuration with respect to (V_1, \dots, V_q) if there exists $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^{m-n} \setminus \bigcup_{i=1}^q V_i$ such that $\{x + B_j y\}_{j=1}^k \subseteq E$.

For both of these definitions, we will drop the k from the notation if there is no confusion.

Let A_1, \dots, A_k be $n \times m$ matrices. For any set of distinct indices $J = \{j_1, \dots, j_s\} \subseteq \{1, \dots, k\}$, define the $ns \times m$ matrix \mathbb{A}_J by

$$\mathbb{A}_J^t = (A_{j_1}^t \ \cdots \ A_{j_s}^t).$$

We shall use $\mathbb{A} = \mathbb{A}_{\{1, \dots, k\}}$.

Let r be the unique positive integer such that

$$n(r - 1) < nk - m \leq nr. \tag{1.3}$$

If we have $nk - m$ components and account for $r - 1$ groups of size n , we are left with $nk - m - n(r - 1)$. This quantity is useful in the main theorem, and bulky to constantly use. We shall denote

$$n' = nk - m - n(r - 1). \tag{1.4}$$

Notice if $nk - m$ is a multiple of n , then $n' = n$, and in general, $0 < n' \leq n$.

Definition 1.5. We say that $\{A_1, \dots, A_k\}$ is non-degenerate if for any $J \subseteq \{1, \dots, k\}$ with $\#(J) = k - r$ and any $j \in \{1, \dots, k\} \setminus J$, the $m \times m$ matrix

$$(\mathbb{A}_J^t \ \widetilde{A}_j^t)$$

is non-singular for any choice of \widetilde{A}_j a submatrix of A_j . Observe that for the above matrix to be $m \times m$, we require \widetilde{A}_j to consist of $n - n'$ rows.

1.2 The main result

Theorem 1.6. *Suppose*

$$n \left\lceil \frac{k+1}{2} \right\rceil \leq m < nk \quad (1.5)$$

and

$$\frac{2(nk-m)}{k} < \beta < n. \quad (1.6)$$

Let $\{B_1, \dots, B_k\}$ be a collection of $n \times (m-n)$ matrices such that $A_j = (I_{n \times n} \ B_j)$ is non-degenerate in the sense of Definition 1.5, where $I_{n \times n}$ is the $n \times n$ identity matrix. Then for any constant C , there exists a positive number $\epsilon_0 = \epsilon_0(C, n, k, m, \mathbb{B}) \ll 1$ with the following property. Suppose the set $E \subseteq \mathbb{R}^n$ with $|E| = 0$ supports a positive, finite, Radon measure μ with the two conditions:

(a) (ball condition) $\sup_{\substack{x \in E \\ 0 < r < 1}} \frac{\mu(B(x;r))}{r^\alpha} \leq C$ if $n - \epsilon_0 < \alpha < n$,

(b) (Fourier decay) $\sup_{\xi \in \mathbb{R}^n} |\widehat{\mu}(\xi)|(1 + |\xi|)^{\beta/2} \leq C$.

Then:

(i) E contains a k -point \mathbb{B} -configuration in the sense of Definition 1.4 (a).

(ii) Moreover, for any finite collection of subspaces $V_1, \dots, V_q \subseteq \mathbb{R}^{m-n}$ with $\dim(V_i) < m-n$, E contains a non-trivial k -point \mathbb{B} -configuration with respect to (V_1, \dots, V_q) in the sense of Definition 1.4 (b).

Note that (a) implies E has Hausdorff dimension at least α by Frostman's Lemma, and (b) implies that E has Fourier dimension at least β .

The existence and constructions of measures on \mathbb{R} that satisfy (a), (b) are discussed in detail in [25]. In higher dimensions, it should be possible to generalize the construction in [25, Section 6] to produce examples in \mathbb{R}^n ; alternatively, it is easy to check that if $\mu = \tilde{\mu}(dr) \times \sigma(d\omega)$ is a product measure in radial coordinates (r, ω) , where $\tilde{\mu}$ is a Salem measure on $[0, 1]$ as in [25, Section 6] and σ is the Lebesgue measure on S^{n-1} , then μ satisfies the conditions (a), (b) of Theorem 1.6.

1.3 Examples

We now give a few examples of geometric configurations covered by Theorem 1.6. For proofs and further discussion, see Section 7.

Corollary 1.7. *Let a, b, c be three distinct points in the plane. Suppose that $E \subset \mathbb{R}^2$ satisfies the assumptions of Theorem 1.6 with ϵ_0 small enough depending on C and on the configuration a, b, c . Then E must contain three distinct points x, y, z such that the triangle Δxyz is a similar (possibly rotated) copy of the triangle Δabc .*

Theorem 1.2 shows that Corollary 1.7 fails without the assumption (b), even if E has Hausdorff dimension 2.

In dimensions $n \geq 2$, one may also consider the following modified question: given a set $E \subset \mathbb{R}^n$, how large can $\dim_H(E)$ be if E does not contain a triple of points forming a particular angle θ ? For ease, we say $E \subseteq \mathbb{R}^n$ *contains the angle* θ if there exist distinct points $x, y, z \in E$ such that the angle between the vectors $y - x$ and $z - x$ is θ , and write $\angle\theta \in E$. Define

$$C(n, \theta) = \sup\{s : \exists E \subseteq \mathbb{R}^n \text{ compact with } \dim_H(E) = s, \angle\theta \notin E\}.$$

Harangi, Keleti, Kiss, Maga, Máthé, Matilla, and Strenner [15] give upper bounds on $C(n, \theta)$ (which they show is tight for $\theta = 0, \pi$), and Máthé [28] provides lower bounds. Their results are summarized below.

θ	lower bound on $C(n, \theta)$	upper bound on $C(n, \theta)$
$0, \pi$	$n - 1$	$n - 1$
$\pi/2$	$n/2$	$\lfloor (n + 1)/2 \rfloor$
$\cos^2 \theta \in \mathbb{Q}$	$n/4$	$n - 1$
other θ	$n/8$	$n - 1$

Corollary 1.7 shows that if θ is given, then any set $E \subset \mathbb{R}^2$ as in Theorem 1.6 must not only contain θ , but in fact θ can be realized as the angle at the apex of a non-degenerate isosceles triangle with vertices in E . This also answers, for such sets, a question posed by Maga [26]:

Question 1.1. *If $E \subseteq \mathbb{R}^2$ is compact with $\dim_H(E) = 2$, must E contain the vertices of an isosceles triangle?*

A different point of view is adopted in [13], where the following question is considered. For $E \subset \mathbb{R}^2$, let $T_2(E) = E^3 / \sim$, where $(a, b, c) \sim (a', b', c')$ if and only if the triangles $\triangle abc$ and $\triangle a'b'c'$ are congruent. Observe that $T_2(\mathbb{R})$ can be parametrized as a 3-dimensional space, e.g. by one angle and the sidelengths of the two sides adjacent to it. What can we say about the size of $T_2(E)$ if the dimension of E is given?

Theorem 1.8 (Greenleaf and Iosevich, [13]). *Let $E \subset \mathbb{R}^2$ be a compact set with $\dim_H(E) > 7/4$. Then $T_2(E)$ has positive 3-dimensional measure.*

Theorem 1.8 does not (and, in light of Theorem 1.2, could not) guarantee the existence of a triangle similar to any particular triangle $\triangle abc$ given in advance. It does, however, ensure that the set of triangles spanned by points of E is large. Our Theorem 1.6 does not provide this type of results, due to the dependence of ε_0 on C and on the choice of configurations. While it might be possible to keep track of this dependence with more effort, Theorem 1.8 is simpler and holds for a much larger class of sets. For extensions of Theorem 1.8 to finite configurations in higher dimensions, see e.g. [7], [11], [14].

In yet another direction, Bourgain [4] proved that if $E \subset \mathbb{R}^n$ has positive upper density (with respect to the Lebesgue measure), and if Δ is a non-degenerate $(k - 1)$ -dimensional simplex (i.e. a set of k points in general position) with $k \leq n$, then E contains a translated

and rotated copy of $\lambda\Delta$ for all λ sufficiently large. A discrete analogue of this result was proved more recently by Magyar [27]. In dimension 2, Bourgain’s result says that there is a $\lambda_0 > 0$ such that for any $\lambda > \lambda_0$, E contains two points x, y with $|x - y| = \lambda$. Similarly, if $E \subset \mathbb{R}^3$ has positive upper density and a, b, c are three distinct and non-collinear points in \mathbb{R}^3 , then for all λ large enough, E contains translated and rotated copies of Δabc rescaled by λ . However, Bourgain’s result does not apply to configurations of $n + 1$ or more points in \mathbb{R}^n , such as triangles in \mathbb{R}^2 .

A related consequence of Theorem 1.6 is the following.

Corollary 1.9. *Let a, b, c be three distinct colinear points in \mathbb{R}^n . Suppose that $E \subset \mathbb{R}^n$ satisfies the assumptions of Theorem 1.6 with ε_0 small enough. Then E must contain three distinct points x, y, z that form a similar image of the triple a, b, c .*

For $n = 2$, this is a special case of Corollary 1.9. While Theorem 1.6 does not seem to allow for similar images of general triangles in dimensions $n \geq 3$, it does cover the case of 3 colinear points in any dimension. Note in particular that this includes 3-term arithmetic progressions $\{x, x + y, x + 2y\}$ with $y \neq 0$, which is a “degenerate” configuration in the sense of Bourgain [4].

We now turn to parallelograms in dimensions 2 and higher.

Corollary 1.10. *Suppose that $E \subset \mathbb{R}^n$ satisfies the assumptions of Theorem 1.6, with ε_0 small enough depending on C . Then E contains a parallelogram $\{x, x + y, x + z, x + y + z\}$, where the four points are all distinct.*

Again, this should be compared to a result of Maga, which shows that the result is false without the Fourier decay assumption.

Theorem 1.11 (Maga, [26]). *There exists a compact set in \mathbb{R}^n with Hausdorff dimension n which does not contain any parallelogram $\{x, x + y, x + z, x + y + z\}$, with $y, z \neq 0$.*

We end with a polynomial example.

Corollary 1.12. *Let a_1, \dots, a_6 be distinct numbers, all greater than 1. Suppose that $E \subset \mathbb{R}^3$ satisfies the assumptions of Theorem 1.6, with ε_0 small enough depending on C and a_i . Then E contains a configuration of the form*

$$x, \quad x + B_2y, \quad x + B_3y, \quad x + B_4y, \tag{1.7}$$

for some $x \in \mathbb{R}^3$ and $y \in \mathbb{R}^6$ with $B_i y \neq 0$ for $i = 2, 3, 4$, where

$$B_2 = \begin{pmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_6 \\ a_1^2 & \dots & a_6^2 \end{pmatrix}, \quad B_3 = \begin{pmatrix} a_1^3 & \dots & a_6^3 \\ a_1^4 & \dots & a_6^4 \\ a_1^5 & \dots & a_6^5 \end{pmatrix}, \quad B_4 = \begin{pmatrix} a_1^6 & \dots & a_6^6 \\ a_1^7 & \dots & a_6^7 \\ a_1^8 & \dots & a_6^8 \end{pmatrix}.$$

This is a non-trivial result in the following sense. Since Vandermonde matrices are non-singular, the set of 6 vectors

$$\begin{pmatrix} 1 \\ a_1 \\ \vdots \\ a_1^5 \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} 1 \\ a_6 \\ \vdots \\ a_6^5 \end{pmatrix}$$

forms a basis for \mathbb{R}^6 . It follows that if $a, b, c \in E$, there is a unique $y \in \mathbb{R}^6$ such that $b = a + B_2 y$ and $c = a + B_3 y$. This also determines uniquely the point $a + B_4 y$, which might or might not be in E . Our result asserts that, under the conditions of Corollary 1.12, we may choose x and y so that in fact all 4 points in (1.7) lie in E .

1.4 Outline of proof

We introduce the following multilinear form:

$$\Lambda(f_1, \dots, f_k) = \int_{\mathbb{R}^m} \prod_{j=1}^k f_j(A_j \vec{x}) \, d\vec{x}. \quad (1.8)$$

If $f_j = f$ for all j , we write $\Lambda(f)$ instead of $\Lambda(f, \dots, f)$. The use of multilinear forms formally similar to (1.8) is common in the literature on Szemerédi-type problems.

Our strategy, roughly following that of [25], is to define an analogue of (1.8) for measures via its Fourier-analytic representation (1.9) below, prove that this analogue can be used to count the k -point configurations we seek, and obtain lower bounds on it that imply the existence of such configurations. A key feature in the present work is the multidimensional geometry of the problem, determined by the system of matrices A_j . While this issue is almost nonexistent in [25], here it will play a major role at every stage of the proof and will account for most of the difficulties, both technical and conceptual.

The multilinear form (1.8) is initially defined for $f_j \in C_c^\infty(\mathbb{R}^n)$. We prove in Proposition 2.1 that for such functions we have

$$\Lambda(f_1, \dots, f_k) = C \int_S \prod_{j=1}^k \widehat{f}_j(\xi_j) \, d\sigma(\xi_1, \dots, \xi_k), \quad (1.9)$$

where σ is the Lebesgue measure on the subspace

$$S = \left\{ \xi = (\xi_1, \dots, \xi_k) \in (\mathbb{R}^n)^k : \sum_{j=1}^k A_j^t \xi_j = \vec{0} \right\} \quad (1.10)$$

and $C = C(\mathbb{A})$ is a constant only depending on the matrices A_j . The importance of (1.9) for us is twofold. First, it allows us to use Fourier bounds on f_j to control the size of $\Lambda(f_1, \dots, f_k)$, a fact that we will use repeatedly in the paper. Second, unlike (1.8), (1.9) makes sense for more general (possibly singular) measures μ_j instead of C_c^∞ functions, provided that their Fourier transforms are well enough behaved so that the integral

$$\Lambda^*(\widehat{\mu}_1, \dots, \widehat{\mu}_k) = C \int_S \prod_{j=1}^k \widehat{\mu}_j(\xi_j) \, d\sigma(\xi_1, \dots, \xi_k), \quad (1.11)$$

converges.

We prove in Proposition 3.1 that the integral in (1.11) does indeed converge, provided that μ_j and A_j obey the assumptions of Theorem 1.6. While the full strength of these assumptions is not needed at this point, we would like to emphasize two key conditions.

First, we require pointwise decay of $\widehat{\mu}_j$. Second, in the absence of additional assumptions on the matrices A_j , the decay of $\widehat{\mu}_j(\xi_j)$ in the ξ_j variables would not necessarily translate into decay in any direction on the subspace S in (1.10). The nondegeneracy conditions in Definition 1.5 ensure that S is in “general position” relative to the subspaces $\{\xi_j = 0\}$ of the full configuration space $(\mathbb{R}^n)^k$ along which the functions $\widehat{\mu}_j(\xi_j)$ do not decay. This puts us in the best possible geometric case with regard to the convergence of (1.11), and allows us to complete the proof of the proposition. Similar geometric issues (with somewhat different details) arise later in the proof as well, notably in the proofs of Propositions 4.1 and 5.1.

Let $\Lambda^*(\widehat{\mu})$ be the multilinear form thus defined, with $\mu_1 = \cdots = \mu_k = \mu$. We claim that a lower bound of the form

$$\Lambda^*(\widehat{\mu}) > 0 \tag{1.12}$$

implies the existence of the k -point configurations we seek. If μ were absolutely continuous with density f , this would follow trivially from Proposition 2.1, since (1.12) would be equivalent to a bound $\Lambda(f) > 0$, and $\Lambda(f)$ has a direct interpretation in terms of such configurations. For singular measures, however, (1.8) need not make sense.

We therefore proceed less directly, following the same route as in [25]. Namely, we prove in Proposition 4.1 that there exists a non-negative, finite, Radon measure $\nu = \nu(\mu)$ on $[0, 1]^m$ such that

- $\nu(\mathbb{R}^m) = \Lambda^*(\widehat{\mu})$.
- $\text{supp } \nu \subseteq \{x \in \mathbb{R}^m : A_1 x, \dots, A_k x \in \text{supp } \mu\}$.
- For any subspace $V \subseteq \mathbb{R}^m$ with $\dim V < m$, $\nu(V) = 0$.

The last condition implies that the ν -measure of the set of “degenerate” configurations is 0. It follows that (1.12) indeed implies the existence of desired configurations in $\text{supp } \mu$.

It remains to prove (1.12). Following the strategy of [25] again, we decompose μ as $\mu = \mu_1 + \mu_2$, where μ_1 is absolutely continuous with bounded density, and μ_2 is singular but obeys good Fourier bounds. We then have

$$\Lambda^*(\widehat{\mu}) = \Lambda^*(\widehat{\mu}_1) + \Lambda(\widehat{\mu}_2, \widehat{\mu}_1, \dots, \widehat{\mu}_1) + \cdots + \Lambda(\widehat{\mu}_2). \tag{1.13}$$

We will treat the first term on the right side of (1.13) as the main term, and the remaining terms as error terms.

To bound $\Lambda(\mu_1)$ from below, we need a quantitative Szemerédi-type estimate from below for bounded functions. In Proposition 5.1, we prove a bound

$$\Lambda(f) \geq c(\delta, M) \tag{1.14}$$

for all functions $f : [0, 1]^n \rightarrow \mathbb{R}$ such that $0 \leq f \leq M$ and $\int f \geq \delta$. To this end, we will modify the “quantitative ergodic” proof of Varnavides’ Theorem given in [33]. While in [25] this proof could be reused almost verbatim, more substantial changes are needed in the multidimensional case. This will take up the bulk of Section 5.

The proof of Theorem 1.6 is completed in Section 6. We first carry out the decomposition $\mu = \mu_1 + \mu_2$ as described earlier, and prove the required bounds on the density of μ_1 and the Fourier transform of μ_2 . The first term on the right side of (1.13) is bounded from below by

a constant $c > 0$ using (1.14), and the sum of the remaining terms can be shown to be less than c in absolute value, using Proposition 3.1 again and the Fourier estimates on μ_2 .

The assumptions of Theorem 1.6 are sufficient for the entire proof to go through. However, many of our intermediate results hold under weaker conditions. We indicate this explicitly in the statements of the results in question, for clarity and possible use in future work.

1.5 Acknowledgement

The second and third authors were supported by NSERC Discovery Grants.

2 A functional multilinear form for counting configurations

We will consider the following multilinear form, initially defined for $f_j \in C_c^\infty(\mathbb{R}^n)$:

$$\Lambda(f_1, \dots, f_k) = \int_{\mathbb{R}^m} \prod_{j=1}^k f_j(A_j \vec{x}) d\vec{x}. \quad (2.1)$$

If $f_j = f$ for all j , we write $\Lambda(f)$ instead of $\Lambda(f, \dots, f)$. Clearly if $\Lambda(f) \neq 0$, then the support of f contains configurations of the form $\{A_j \vec{x} : 1 \leq j \leq k\}$ for a set of \vec{x} of positive measure. We will rewrite $\Lambda(f_1, \dots, f_k)$ in a form that will allow us to extend it to measures, not just functions.

Proposition 2.1. *For $f_j \in C_c^\infty(\mathbb{R}^n)$, $\Lambda(f_1, \dots, f_k)$ defined in (2.1) admits the representation*

$$\Lambda(f_1, \dots, f_k) = C \int_S \prod_{j=1}^k \widehat{f}_j(\xi_j) d\sigma(\xi_1, \dots, \xi_k),$$

where σ is the Lebesgue measure on the subspace

$$S = \left\{ \xi = (\xi_1, \dots, \xi_k) \in (\mathbb{R}^n)^k : \sum_{j=1}^k A_j^t \xi_j = \vec{0} \right\} \quad (2.2)$$

and $C = C(\mathbb{A})$ is a constant only depending on the matrices A_j .

Proof. For $\Phi \in \mathcal{S}(\mathbb{R}^m)$ with $\Phi(0) = 1$, the dominated convergence theorem gives

$$\Lambda(f_1, \dots, f_k) = \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^m} \left(\prod_{j=1}^k f_j(A_j \vec{x}) \right) \Phi(\vec{x}\varepsilon) d\vec{x}.$$

Applying the Fourier inversion formula in \mathbb{R}^n , $g(y) = \int_{\mathbb{R}^n} e^{2\pi i y \cdot \xi} \widehat{g}(\xi) d\xi$ where $\widehat{g}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} g(x) dx$, we get

$$\Lambda(f_1, \dots, f_k) = \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^m} \prod_{j=1}^k \left[\int_{\mathbb{R}^n} e^{2\pi i A_j \vec{x} \cdot \xi_j} \widehat{f}_j(\xi_j) d\xi_j \right] \Phi(\vec{x}\varepsilon) d\vec{x}$$

$$\begin{aligned}
&= \lim_{\varepsilon \rightarrow 0^+} \int_{\vec{\xi} = (\xi_1, \dots, \xi_k) \in (\mathbb{R}^n)^k} \prod_{j=1}^k \widehat{f}_j(\xi_j) \left[\int_{\mathbb{R}^m} e^{2\pi i \vec{x} \cdot \mathbb{A}^t \xi} \Phi(\vec{x} \varepsilon) d\vec{x} \right] d\vec{\xi} \\
&= \lim_{\varepsilon \rightarrow 0^+} \int_{\vec{\xi} \in (\mathbb{R}^n)^k} \prod_{j=1}^k \widehat{f}_j(\xi_j) \frac{1}{\varepsilon^p} \widehat{\Phi} \left(\frac{\mathbb{A}^t \xi}{\varepsilon} \right) d\vec{\xi} \\
&= C_{\mathbb{A}^t} \int_S \prod_{j=1}^k \widehat{f}_j(\xi_j) d\sigma(\xi),
\end{aligned}$$

for some constant $C_{\mathbb{A}^t}$. This last step follows from Proposition A.3, with $p = m$, $d = nk$, $V = S$, $P = \mathbb{A}^t$, and $F = \prod \widehat{f}_j$. \square

Proposition 2.2. *Let $g \in \mathcal{S}(\mathbb{R}^m)$, $f_j \in C_c^\infty(\mathbb{R}^n)$. Then the integral*

$$\Theta(g; f_1, \dots, f_k) := \int_{\mathbb{R}^m} g(\vec{x}) \prod_{j=1}^k f_j(A_j \vec{x}) d\vec{x}$$

is absolutely convergent, and admits the representation

$$\Theta(g; f_1, \dots, f_k) = \int_{(\mathbb{R}^n)^k} \widehat{g}(-\mathbb{A}^t \vec{\xi}) \prod_{j=1}^k \widehat{f}_j(\vec{\xi}_j) d\vec{\xi}.$$

Proof. By Fourier inversion,

$$\Theta(g; f_1, \dots, f_k) = \int_{\mathbb{R}^m} g(\vec{x}) \prod_{j=1}^k \left[\int_{\mathbb{R}^n} e^{2\pi i A_j \vec{x} \cdot \vec{\xi}_j} \widehat{f}_j(\vec{\xi}_j) d\vec{\xi}_j \right] d\vec{x},$$

which is absolutely convergent since $\widehat{f}_j, g \in \mathcal{S}(\mathbb{R}^m)$. Then by Fubini's Theorem,

$$\begin{aligned}
\Theta(g; f_1, \dots, f_k) &= \int_{(\mathbb{R}^n)^k} \prod_{j=1}^k \widehat{f}_j(\vec{\xi}_j) \left[\int_{\mathbb{R}^m} g(x) e^{2\pi i \vec{x} \cdot \mathbb{A}^t \vec{\xi}} d\vec{x} \right] d\vec{\xi} \\
&= \int_{(\mathbb{R}^n)^k} \widehat{g}(-\mathbb{A}^t \vec{\xi}) \prod_{j=1}^k \widehat{f}_j(\vec{\xi}_j) d\vec{\xi},
\end{aligned}$$

where the last line follows by the definition of the Fourier transform. \square

3 Extending the multilinear form to measures

With S as in (2.2), denote $S^\perp = \{\tau \in (\mathbb{R}^n)^k : \langle \tau, \xi \rangle = 0 \text{ for all } \xi \in S\}$ and fix a $\tau \in S^\perp$. We will use the variable

$$\eta = (\eta_1, \dots, \eta_k) = \xi + \tau = (\xi_1, \dots, \xi_k) + (\tau_1, \dots, \tau_k) \in S + \tau, \quad (3.1)$$

where $\xi \in S$, and $\eta_j, \xi_j, \tau_j \in \mathbb{R}^n$. Define

$$\Lambda_\tau^*(g_1, \dots, g_k) = \int_{S+\tau} \prod_{j=1}^k g_j(\eta_j) d\sigma(\eta), \quad (3.2)$$

initially defined for $g_j \in C_c(\mathbb{R}^n)$ so that the integral is absolutely convergent. If $g_j = g$ for all j , we write $\Lambda_\tau^*(g)$ instead of $\Lambda_\tau^*(g, \dots, g)$. We will use Λ^* in place of Λ_0^* . Our next proposition shows that we may extend this multilinear form to continuous functions with appropriate decay.

In applications, g_j will be $\widehat{\mu}$, with μ as in Theorem 1.6. While defining the multilinear form Λ for measures requires only the use of Λ_0^* (so that the integration is on S as in Proposition 2.1), the proof of our main result will rely crucially on estimates on Λ_τ^* uniform in τ .

Proposition 3.1. *Let $\{A_1, \dots, A_k\}$ be a non-degenerate collection of $n \times m$ matrices in the sense of Definition 1.5. Assume that $nk/2 < m < nk$ and $g_1, \dots, g_k : \mathbb{R}^n \rightarrow \mathbb{C}$ are continuous functions satisfying*

$$|g_j(\kappa)| \leq M(1 + |\kappa|)^{-\beta/2}, \quad \kappa \in \mathbb{R}^n, \quad (3.3)$$

for some $\beta > 2(nk - m)/k$. Then the integral defining $\Lambda_\tau^* = \Lambda_\tau^*(g_1, \dots, g_k)$ is absolutely convergent for every $\tau \in S^\perp$. Indeed,

$$\sup_{\tau \in S^\perp} \Lambda_\tau^*(|g_1|, \dots, |g_k|) \leq C$$

where C depends only on n, k, m, M , and \mathbb{A} .

Since $S + \tau$ for $\tau \in S^\perp$ includes all possible translates of S , Proposition 3.1 gives a uniform upper bound on the integral of $\prod g_j$ over all affine copies of S .

The proof of this proposition is based on a lemma which requires some additional notation. Let $0_{a \times b}$ denote the $a \times b$ matrix consisting of 0's; we will also use 0 when the size is evident. Let $I_{n \times n}$ denote the $n \times n$ identity matrix; we will also use I_n for short if the context is clear. Define the $nk \times n$ matrix E_j ($1 \leq j \leq k$) by

$$E_j^t = (0_{n \times n} \cdots I_{n \times n} \cdots 0_{n \times n}),$$

where $I_{n \times n}$ is in the j th block. For any $\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n) \in \{0, 1\}^n$, we denote $I_{n \times n}(\vec{\varepsilon}) = \text{diag}(\varepsilon_1, \dots, \varepsilon_n)$. For a subset $J' \subseteq \{1, \dots, n\}$ and index $j \in \{1, \dots, k\}$, define the $nk \times n$ matrix

$$E_j(J')^t = (0_{n \times n} \cdots I_{n \times n}(\varepsilon) \cdots 0_{n \times n}),$$

where $I_{n \times n}(\varepsilon)$ is in the j th block and $\varepsilon_i = 1$ if and only if $i \in J'$. Finally, for $\xi \in S$ we define

$$\xi_j(J') = E_j(J')\xi.$$

Lemma 3.2. *Let $\{A_1, \dots, A_k\}$ be a non-degenerate collection of $n \times m$ matrices in the sense of Definition 1.5. Let $J = \{j_1, \dots, j_r\} \subseteq \{1, \dots, k\}$ and $J' \subseteq \{1, \dots, n\}$ be collections of distinct indices with $\#(J') = n'$, defined by (1.4). Then the projection of $(\xi_{j_1}, \dots, \xi_{j_r}, \xi_{j_r}(J'))$ on S is a coordinate system on S as defined in (2.2). In particular, there exists a constant $C = C(J, J', j, \mathbb{A})$ such that $d\sigma(\xi) = C d\xi_{j_1} \cdots d\xi_{j_{r-1}} d\xi_{j_r}(J')$, where $\sigma(\xi)$ is the Lebesgue measure on S .*

Proof. It suffices to prove

$$S' = \{\xi \in S : \xi_{j_1} = \xi_{j_2} = \cdots = \xi_{j_{r-1}} = \xi_{j_r}(J') = 0\} \implies \dim S' = 0. \quad (3.4)$$

We will examine $S' = \{\xi \in S : \xi_1 = \xi_2 = \cdots = \xi_{r-1} = \xi_r(J') = 0\}$; the other cases are similar. S' is the subspace defined by

$$\left\{ \xi \in (\mathbb{R}^n)^k : \begin{pmatrix} A_1^t & A_2^t & \cdots & A_{r-1}^t & A_r^t & A_{r+1}^t & \cdots & A_k^t \\ I_n & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & I_n & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & I_n & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & I_n(\vec{1} - \vec{\varepsilon}) & 0 & \cdots & 0 \end{pmatrix} \xi = 0 \right\},$$

where $\varepsilon_i = 1$ if and only if $i \in J'$, and $\vec{1} = (1, \dots, 1)$. For $\dim S' = 0$, we need the kernel of the above $(m+nr) \times nk$ matrix to have dimension 0, so we need the rank of said matrix to be nk . Notice I_n is of rank n and we have $r-1$ of these, and $I(\vec{1} - \vec{\varepsilon})$ has rank $[nk - m - n(r-1)]$, so it suffices for

$$\text{rank} \left((A_r(\vec{\varepsilon}))^t \quad A_{r+1}^t \quad \cdots \quad A_k^t \right) = nk - nr.$$

A similar condition holds if we examine $S' = \{\xi \in S : \xi_{i_1} = \xi_{i_2} = \cdots = \xi_{i_{r-1}} = \xi_{i_r}(J') = 0\}$ in general, and upon taking the transpose we arrive at the sufficient condition

$$\text{rank} \begin{pmatrix} \mathbb{A}_I \\ A_{i_{k-(r-1)}}(\vec{\varepsilon}) \end{pmatrix} = m \quad (3.5)$$

for $I = \{i_1, \dots, i_{k-r}\} \subseteq \{1, \dots, k\}$ a set of distinct indices, and $\vec{\varepsilon} \cdot \vec{1} = n - n'$. Notice this means the matrix is of full rank. (3.5) follows from the non-degeneracy assumption in Definition 1.5. \square

Proof of Proposition 3.1. Fix $\tau \in S^\perp$. We use Sym_k to denote the symmetric group on k elements. For a permutation $\theta \in \text{Sym}_k$, we define the region

$$\Omega_\theta = \{\eta \in S + \tau : |\eta_{\theta(1)}| \leq |\eta_{\theta(2)}| \leq \cdots \leq |\eta_{\theta(k)}|\}$$

so that

$$S + \tau = \bigcup_{\theta \in \text{Sym}_k} \Omega_\theta.$$

For simplicity and without loss of generality, we will examine the case of $\theta = \text{id}$ and write Ω_{id} as Ω ; the other cases are analogous. It then suffices to show the convergence of the integral

$$I = \int_\Omega \prod_{j=1}^k |g_j(\eta_j)| \, d\sigma \lesssim \int_\Omega \prod_{j=1}^k (1 + |\eta_j|)^{-\beta/2} \, d\sigma.$$

Let $L = \prod_{j=1}^{r-1} (1 + |\eta_j|)^{-\beta/2}$. By Hölder's inequality over $k - (r-1)$ terms,

$$I \lesssim \int_\Omega \prod_{j=1}^k (1 + |\eta_j|)^{-\beta/2} \, d\sigma$$

$$\begin{aligned}
&= \int_{\Omega} L \prod_{i=r}^k (1 + |\eta_i|)^{-\beta/2} d\sigma \\
&= \int_{\Omega} \prod_{i=r}^k [L^{1/(k-(r-1))} (1 + |\eta_i|)^{-\beta/2}] d\sigma \\
&\leq \prod_{i=r}^k \left[\int_{|\eta_1| \leq \dots \leq |\eta_{r-1}| \leq |\eta_i|}^{S+\tau} L (1 + |\eta_i|)^{-\frac{\beta}{2}(k-(r-1))} d\sigma \right]^{1/(k-(r-1))} \\
&\leq \int_{|\eta_1| \leq \dots \leq |\eta_r|}^{S+\tau} \prod_{j=1}^{r-1} (1 + |\eta_j|)^{-\beta/2} (1 + |\eta_r|)^{-\frac{\beta}{2}(k-(r-1))} d\sigma.
\end{aligned}$$

In the last inequality, we see that for each i , the integral is the same with a different dummy variable, so we collect the terms under the single index $i = r$. Recalling our decomposition of η into ξ and τ as per (3.1) and since τ is fixed, we arrive at

$$I \lesssim \int_{|\xi_1 + \tau_1| \leq \dots \leq |\xi_r + \tau_r|}^S \prod_{j=1}^{r-1} (1 + |\xi_j + \tau_j|)^{-\beta/2} (1 + |\xi_r + \tau_r|)^{-\frac{\beta}{2}(k-(r-1))} d\sigma(\xi).$$

Suppose $\eta_r = (\eta_{r,1}, \dots, \eta_{r,n})$, and for any permutation $\pi \in \text{Sym}_n$, let

$$\eta_r^\pi = (\eta_{r,\pi(1)}, \dots, \eta_{r,\pi(n)}).$$

As in the beginning of this section, we may partition our current region of integration into a finite number of regions of the form $\{|\eta_{r,\pi(1)}| \geq \dots \geq |\eta_{r,\pi(n)}|\}$ for $\pi \in \text{Sym}_n$. Then $|\eta_r| \leq n|\eta_{r,\pi(1)}|$, so that $|\eta_r| \sim |\eta_r^\pi|$.

By Lemma 3.2, $d\sigma(\xi) = C d\xi_1 \dots d\xi_{r-1} d\xi_r^\pi$, hence

$$\begin{aligned}
I &\lesssim \int_{|\xi_1 + \tau_1| \leq \dots \leq |\xi_r + \tau_r|}^S \prod_{j=1}^{r-1} (1 + |\xi_j + \tau_j|)^{-\beta/2} (1 + |\xi_r^\pi + \tau_r^\pi|)^{-\frac{\beta}{2}(k-(r-1))} d\xi_1 \dots d\xi_{r-1} d\xi_r^\pi \\
&\lesssim \int_{\mathbb{R}^{n'}} \prod_{j=1}^{r-1} \left[\int_{\substack{\xi_j + \tau_j \in \mathbb{R}^n \\ |\xi_j + \tau_j| \leq |\xi_r^\pi + \tau_r^\pi|}} (1 + |\xi_j + \tau_j|)^{-\beta/2} d\xi_j \right] (1 + |\xi_r^\pi + \tau_r^\pi|)^{-\frac{\beta}{2}(k-(r-1))} d\xi_r^\pi.
\end{aligned}$$

Translating ξ_j by τ_j ,

$$\begin{aligned}
I &\lesssim \int_{\mathbb{R}^{n'}} \prod_{j=1}^{r-1} \left[\int_{\substack{\xi_j \in \mathbb{R}^n \\ |\xi_j| \leq |\xi_r^\pi|}} (1 + |\xi_j|)^{-\beta/2} d\xi_j \right] (1 + |\xi_r^\pi|)^{-\frac{\beta}{2}(k-(r-1))} d\xi_r^\pi \\
&\lesssim \int_{\mathbb{R}^{n'}} \left[\int_0^{|\xi_r^\pi|} (1 + \rho)^{-\beta/2+n-1} d\rho \right]^{r-1} (1 + |\xi_r^\pi|)^{-\frac{\beta}{2}(k-(r-1))} d\xi_r^\pi \\
&\lesssim \int_{\mathbb{R}^{n'}} (1 + |\xi_r^\pi|)^{(n-\beta/2)(r-1) - \frac{\beta}{2}(k-(r-1))} d\xi_r^\pi \\
&= \int_{\mathbb{R}^{n'}} (1 + |\xi_r^\pi|)^{n(r-1) - \beta k/2} d\xi_r^\pi,
\end{aligned}$$

where the Jacobian in making the spherical change of coordinates $\rho = |\xi_j|$ above is independent of τ_j . This last expression is finite (with a bound independent of τ) when

$$\beta k/2 - n(r-1) > n' = nk - m - n(r-1),$$

which holds since $2(nk - m)/k < \beta < n$ and $m > nk/2$. □

4 Counting geometric configurations in sparse sets

In this section, we will show that the multilinear form Λ^* defined in (3.2) is effective in counting non-trivial configurations supported on appropriate sparse sets.

Proposition 4.1. *Suppose $nk/2 < m < nk$ and $2(nk - m)/k < \beta < n$. Let $\{A_1, \dots, A_k\}$ be a collection of $n \times m$ matrices that are non-degenerate in the sense of Definition 1.5. Let μ be a positive, finite, Radon measure μ with*

$$\sup_{\xi \in \mathbb{R}^n} |\widehat{\mu}(\xi)| (1 + |\xi|)^{\beta/2} \leq C. \quad (4.1)$$

Then there exists a non-negative, finite, Radon measure $\nu = \nu(\mu)$ on $[0, 1]^m$ such that

- (a) $\nu(\mathbb{R}^m) = \Lambda^*(\widehat{\mu})$.
- (b) $\text{supp } \nu \subseteq \{x \in \mathbb{R}^m : A_1 x, \dots, A_k x \in \text{supp } \mu\}$.
- (c) For any subspace $V \subseteq \mathbb{R}^m$ with $\dim V < m$, $\nu(V) = 0$.

4.1 Existence of candidate ν

Fix a non-negative $\phi \in \mathcal{S}(\mathbb{R}^m)$ with $\int \phi = 1$ and let $\phi_\varepsilon(y) = \varepsilon^{-n} \phi(\varepsilon^{-1}y)$. Let $\mu_\varepsilon = \mu * \phi_\varepsilon$. Notice $\widehat{\phi} \in \mathcal{S}(\mathbb{R}^m)$ since $\phi \in \mathcal{S}(\mathbb{R}^m)$, so

$$|\widehat{\mu}_\varepsilon(\xi)| = |\widehat{\mu}(\xi) \widehat{\phi}(\varepsilon\xi)| \leq C(1 + |\xi|)^{-\beta/2} \quad (4.2)$$

with $C = \|\widehat{\phi}\|_\infty$ independent of ε . Furthermore, $\widehat{\phi}(\varepsilon\xi) \rightarrow \widehat{\phi}(0) = \int \phi = 1$ as $\varepsilon \rightarrow 0$, hence

$$\widehat{\mu}_\varepsilon(\xi) \rightarrow \widehat{\mu}(\xi) \text{ pointwise as } \varepsilon \rightarrow 0. \quad (4.3)$$

We prove that the multilinear form Λ_τ^* satisfies a weak continuity property, in the following sense:

Lemma 4.2. $\Lambda_\tau^*(\widehat{\mu}_\varepsilon) \rightarrow \Lambda_\tau^*(\widehat{\mu})$ as $\varepsilon \rightarrow 0$ for every fixed $\tau \in S^\perp$.

Proof. Fix $\tau \in S^\perp$. By definition,

$$\Lambda_\tau^*(\widehat{\mu}_\varepsilon) = \int_{S+\tau} \prod_{j=1}^k \widehat{\mu}_\varepsilon(\xi_j) d\xi.$$

By (4.2), $|\widehat{\mu}_\varepsilon(\eta)| \leq C(1 + |\eta|)^{-\beta/2} =: g(\eta)$ uniformly in ε , so

$$\prod_{j=1}^k |\widehat{\mu}_\varepsilon(\xi_j)| \leq C \prod_{j=1}^k g(\xi_j).$$

By Proposition 3.1, $\Lambda_\tau^*(g)$ is finite, so by (4.3) and the dominated convergence theorem,

$$\Lambda_\tau^*(\widehat{\mu}_\varepsilon) \rightarrow \Lambda_\tau^*(\widehat{\mu}).$$

□

For $F \in C([0, 1]^m)$ such that $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$, define the linear functional ν by

$$\langle \nu, F \rangle = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} F(\vec{x}) \prod_{j=1}^k \mu_\varepsilon(A_j \vec{x}) d\vec{x}. \quad (4.4)$$

We will prove in Lemma 4.3 below that the limit exists and extends as a bounded linear functional on $C([0, 1]^m)$. Clearly, $\langle \nu, F \rangle \geq 0$ if $F \geq 0$. By the Riesz representation theorem, there exists a non-negative, finite, Radon measure ν that identifies this linear functional; namely $\langle \nu, F \rangle = \int F d\nu$.

Lemma 4.3. *There exists a non-negative, bounded, linear functional ν on $C([0, 1]^m)$, that is,*

$$|\langle \nu, F \rangle| \leq C \|F\|_\infty \quad (4.5)$$

for some positive constant C independent of $F \in C([0, 1]^m)$, which agrees with (4.4) if $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$.

Proof. Assume the limit (4.4) exists for $F \in C([0, 1]^m)$, $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$. Then

$$\begin{aligned} |\langle \nu, F \rangle| &\leq \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} |F(\vec{x})| \prod_{j=1}^k \mu_\varepsilon(A_j \vec{x}) d\vec{x} \\ &\leq \|F\|_\infty \lim_{\varepsilon \rightarrow 0} \Lambda(\mu_\varepsilon) \\ &= \|F\|_\infty \lim_{\varepsilon \rightarrow 0} \Lambda^*(\widehat{\mu}_\varepsilon) \\ &\leq C \|F\|_\infty, \end{aligned}$$

where the last line follows by Proposition 3.1, with a constant C independent of ε . Thus, (4.5) holds.

It remains to prove that $\langle \nu, F \rangle$ is well-defined. We will prove this by showing that the limit in (4.4) exists for $F \in C([0, 1]^m)$ such that $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$ and use density arguments to extend the functional to all of $C([0, 1]^m)$. Applying Corollary 2.2 with $g = F$, $f_1 = \dots = f_k = \mu_\varepsilon$, we obtain

$$\langle \nu, F \rangle = \lim_{\varepsilon \rightarrow 0} \Theta(F; \mu_\varepsilon, \dots, \mu_\varepsilon) = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^{nk}} \widehat{F}(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}_\varepsilon(\xi_j) d\xi. \quad (4.6)$$

By (4.3),

$$\widehat{F}(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}_\varepsilon(\xi_j) \rightarrow \widehat{F}(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}(\xi_j)$$

pointwise, and by (4.2),

$$\left| \widehat{F}(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}_\varepsilon(\xi_j) \right| \leq C |\widehat{F}(-\mathbb{A}^t \xi)| \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2}.$$

Existence of the limit in (4.6) will follow from the dominated convergence theorem, if we prove $|\widehat{F}(\mathbb{A}^t \xi)| \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2} \in L^1(\mathbb{R}^{nk})$. To this end, let $g(t) = (1 + |t|)^{\beta/2}$ for $t \in \mathbb{R}^n$. Then

$$\begin{aligned} \int_{\mathbb{R}^{nk}} |\widehat{F}(\mathbb{A}^t \xi)| \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2} d\xi &= \int_{\mathbb{R}^m} |\widehat{F}(\kappa)| \int_{\mathbb{A}^t \xi = \kappa} \prod_{j=1}^k g(\xi_j) d\sigma(\xi) d\kappa \\ &= \int_{\mathbb{R}^m} |\widehat{F}(\kappa)| \Lambda_{\tau(\kappa)}^*(g) d\kappa \\ &\leq C \int_{\mathbb{R}^m} |\widehat{F}(\kappa)| d\kappa < \infty. \end{aligned}$$

Here $\tau(\kappa)$ is the unique vector in S^\perp such that $\{\mathbb{A}^t \xi = \kappa\} = S + \tau(\kappa)$. We have used Proposition 3.1 to bound $\Lambda_{\tau(\kappa)}^*$ in the last displayed inequality above, and used the fact that $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$ to deduce that $\widehat{F} \in L^1(\mathbb{R}^m)$. By the dominated convergence theorem, the limit in (4.4) exists.

To extend ν to all of $C([0, 1]^m)$, fix $F \in C([0, 1]^m)$. Extend F to $\widetilde{F} \in C_c(\mathbb{R}^m)$ so that $F = \widetilde{F}$ on $[0, 1]^m$. We will reuse F to mean \widetilde{F} for convenience. Get a sequence of functions $F_n \in C^\infty([0, 1]^m)$ with $\widehat{F}_n \in \mathcal{S}(\mathbb{R}^m)$ such that $\|F - F_n\|_\infty \rightarrow 0$. By the preceding proof,

$$|\langle \nu, F_n - F_m \rangle| \leq C \|F_n - F_m\|_\infty \rightarrow 0$$

as $n, m \rightarrow \infty$ since F_n is Cauchy in sup norm. Thus the sequence of scalars $\langle \nu, F_n \rangle$ is Cauchy and hence converges. Define

$$\langle \nu, F \rangle = \lim_{n \rightarrow \infty} \langle \nu, F_n \rangle.$$

Clearly,

$$|\langle \nu, F \rangle| = \lim_{n \rightarrow \infty} |\langle \nu, F_n \rangle| \leq C \lim_{n \rightarrow \infty} \|F_n\|_\infty = C \|F\|_\infty.$$

□

The proof of Lemma 4.3 yields the following corollary which will be used later in the sequel:

Corollary 4.4. *For $F \in C([0, 1]^m)$ with $\widehat{F} \in \mathcal{S}(\mathbb{R}^m)$,*

$$\langle \nu, F \rangle = \int_{\mathbb{R}^{nk}} \widehat{F}(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}(\xi_j) d\xi.$$

4.2 Proof of Proposition 4.1(a)

Proof. We have

$$\nu(\mathbb{R}^m) = \langle \nu, 1 \rangle = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} \prod_{j=1}^k \mu_\varepsilon(A_j x) dx = \lim_{\varepsilon \rightarrow 0} \Lambda(\mu_\varepsilon) = \lim_{\varepsilon \rightarrow 0} \Lambda^*(\widehat{\mu}_\varepsilon) = \Lambda^*(\widehat{\mu})$$

by Lemma 4.2, with $\tau = 0$. □

4.3 Proof of Proposition 4.1(b)

Proof. Define

$$X := \{x \in \mathbb{R}^m : A_1 x, \dots, A_k x \in \text{supp } \mu\}.$$

Since $\text{supp } \mu$ is closed, X is closed. Let F be any continuous function on \mathbb{R}^m with $\text{supp } F$ disjoint from X , then $\text{dist}(\text{supp } F, X) > 0$. In order to prove that ν is supported on X , we aim to show that $\langle \nu, F \rangle = 0$. To this end, let us define

$$\begin{aligned} X_N &:= \{\vec{x} \in \mathbb{R}^m : \text{dist}(A_j \vec{x}, \text{supp } (\mu)) \leq 1/N \text{ for every } 1 \leq j \leq k\} \\ &= \bigcap_{j=1}^k \{\vec{x} \in \mathbb{R}^m : \text{dist}(A_j \vec{x}, \text{supp } (\mu)) \leq 1/N\}. \end{aligned}$$

Then $X \subseteq X_N$ for every N , and $X = \bigcap_{N=1}^{\infty} X_N$. Furthermore,

$$X_N^c = \bigcup_{j=1}^k \{\vec{x} \in \mathbb{R}^m : \text{dist}(A_j \vec{x}, \text{supp } (\mu)) > 1/N\}$$

is an open set for every $N \geq 1$, with

$$\text{supp } (F) \subseteq X^c = \bigcup_{N=1}^{\infty} X_N^c.$$

Introducing a smooth partition of unity subordinate to $\{X_N^c\}_N$, we can write $F = \sum_N F_N$ where each $F_N \in C_c^\infty(\mathbb{R}^m)$ with $\text{supp } (F_N) \subseteq X_N^c$. Note that since $\text{supp } (F)$ is a compact subset of X^c , it follows from the definition of a partition of unity that the infinite sum above is in fact a finite sum, so there is no issue of convergence.

Let $\mu_\varepsilon^{A_j}(\vec{x}) := \mu_\varepsilon(A_j \vec{x})$. To compute

$$\begin{aligned} \langle \nu, F_N \rangle &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} F_N(\vec{x}) \prod_{j=1}^k \mu_\varepsilon(A_j \vec{x}) d\vec{x} \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^m} F_N(\vec{x}) \prod_{j=1}^k \mu_\varepsilon^{A_j}(\vec{x}) d\vec{x} \end{aligned}$$

for a fixed $N \geq 1$, we observe that

$$\text{supp } (\mu_\varepsilon^{A_j}) \subseteq \{\vec{x} \in \mathbb{R}^m : \text{dist}(A_j \vec{x}, \text{supp } (\mu)) \leq \varepsilon\}$$

$$\subseteq \{\vec{x} \in \mathbb{R}^m : \text{dist}(A_j \vec{x}, \text{supp}(\mu)) \leq 1/N\}$$

if $\varepsilon \leq 1/N$. Thus the product $\prod_{j=1}^k \mu_\varepsilon^{A_j}(\vec{x})$ is supported on X_N , whereas F_N is supported on X_N^c . This implies

$$\int_{\mathbb{R}^m} F_N(\vec{x}) \prod_{j=1}^k \mu_\varepsilon^{A_j}(\vec{x}) d\vec{x} = 0$$

for all $\varepsilon \leq 1/N$, so that $\langle \nu, F_N \rangle = 0$ for every $N \geq 1$. Therefore, $\langle \nu, F \rangle = 0$ as claimed. \square

4.4 Proof of Proposition 4.1(c)

Proof. It suffices to prove the proposition for $\dim V = v = m - 1$, since smaller subspaces have even less measure. Let P_V denote the projection onto V . Fix $v_0 \in V$ and define

$$V_{\delta, \gamma} = \{x \in \mathbb{R}^m : |v_0 - P_V x| \leq \gamma, \text{dist}(x, V) = |P_{V^\perp} x| \leq \delta\}.$$

It suffices to prove $\nu(V_\delta) \rightarrow 0$ as $\delta \rightarrow 0$. If ϕ_δ is any smooth function with

$$\phi_\delta = \begin{cases} 1 & \text{on } V_{\delta, 1}, \\ 0 & \text{on } \mathbb{R}^m \setminus V_{\delta, 2}, \end{cases} \quad (4.7)$$

then $\nu(V_\delta) \leq \int \phi_\delta d\nu = \langle \nu, \phi_\delta \rangle$, so we aim to show that $\langle \nu, \phi_\delta \rangle \rightarrow 0$ as $\delta \rightarrow 0$.

Fix bases $\{\mathbf{a}_1, \dots, \mathbf{a}_{m-1}\}$ and $\{\mathbf{b}\}$ for V and V^\perp respectively, such that $\{\mathbf{a}_1, \dots, \mathbf{a}_{m-1}, \mathbf{b}\}$ forms an orthonormal basis of \mathbb{R}^m . Thus, for any $x \in \mathbb{R}^m$, there is a unique decomposition

$$x = u + w, \text{ with } u = \sum_{j=1}^{m-1} a_j \mathbf{a}_j \in V, w = b \mathbf{b} \in V^\perp, \quad (4.8)$$

$$\text{where } \vec{a} = (a_1, \dots, a_{m-1})^t \in \mathbb{R}^{m-1}, b \in \mathbb{R}.$$

Without loss of generality, we may assume ϕ_δ as in (4.7) to be variable-separated as

$$\phi_\delta(x) = \phi_V(\vec{a}) \phi_{V^\perp}(\delta^{-1} b), \quad (4.9)$$

where $\phi_V \in C_c^\infty(\mathbb{R}^{m-1})$ is supported on $\{\vec{a} : |\sum_{j=1}^{m-1} a_j \mathbf{a}_j - v_0| \leq 2\}$ and $\phi_{V^\perp} \in C_c^\infty(\mathbb{R})$ is supported on $\{b : |b| \leq 2\}$.

By Corollary 4.4,

$$\langle \nu, \phi_\delta \rangle = \int_{\mathbb{R}^{nk}} \widehat{\phi}_\delta(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}(\xi_j) d\xi. \quad (4.10)$$

We will show that this integral tends to 0 as $\delta \rightarrow 0$. The estimation of this integral relies on an orthogonal decomposition of \mathbb{R}^{nk} into specific subspaces, which we now describe. Let

$$W = \{\xi \in \mathbb{R}^{nk} : \mathbb{A}^t \xi \cdot x = 0 \text{ for all } x \in V\}. \quad (4.11)$$

Then S is clearly a subspace of W , as $\mathbb{A}^t \xi = 0$ if $\xi \in S$. It is also not difficult to see that $\dim W = nk - v = nk - (m - 1)$. The proof of this has been relegated to Lemma 4.5 below. A consequence of this fact is that

$$\dim W \cap S^\perp = 1 \quad (4.12)$$

since S is $(nk - m)$ -dimensional. Now write $\xi \in \mathbb{R}^{nk}$ as

$$\xi = \zeta + \eta + \lambda, \text{ where } \zeta \in S, \eta \in W \cap S^\perp, \lambda \in W^\perp,$$

so that $d\xi = d\sigma_{S+\eta+\lambda}(\xi) d\eta d\lambda$. Here $d\sigma_{S+\eta+\lambda}$ denotes the surface measure on $S + \eta + \lambda$, as defined in Definition A.1. We will soon show, in Lemmas 4.6 and 4.7 below, that the two factors of the integrand in (4.10) obey the size estimates:

$$\left| \prod_{j=1}^k \widehat{\mu}(\xi_j) \right| \lesssim (1 + |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-\varepsilon} \prod_{j=1}^k g(\xi_j), \quad (4.13)$$

and

$$|\widehat{\phi}_\delta(-\mathbb{A}^t \xi)| \leq \delta C_M (1 + |\lambda|)^{-M} (1 + \delta |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-M} \quad (4.14)$$

for any $M \geq 1$. Here $g(\xi_j) = (1 + |\xi_j|)^{-\beta/2+\varepsilon}$, where $\varepsilon > 0$ is chosen sufficiently small so that

$$\beta - 2\varepsilon > 2(nk - m)/k. \quad (4.15)$$

Notice this is possible since $\beta > 2(nk - m)/k$.

Assuming (4.13) and (4.14) temporarily, the estimation of (4.10) proceeds as follows.

$$\begin{aligned} |\langle \nu, \phi_\delta \rangle| &\leq \int_{\mathbb{R}^{nk}} \left| \widehat{\phi}_\delta(-\mathbb{A}^t \xi) \prod_{j=1}^k \widehat{\mu}(\xi_j) \right| d\xi \\ &\lesssim \int_{W^\perp} \left[\int_{W \cap S^\perp} \left[\int_{S+\eta+\lambda} \prod_{j=1}^k g(\xi_j) d\sigma_{S+\eta+\lambda}(\xi) \right] J(\eta, \lambda) d\eta \right] (1 + |\lambda|)^{-M} d\lambda, \end{aligned}$$

where

$$J(\eta, \lambda) = \delta C_M (1 + |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-\varepsilon} (1 + \delta |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-M}.$$

We claim that

$$\int_{S+\eta+\lambda} \prod_{j=1}^k g(\xi_j) d\sigma(\xi) \leq C \quad (4.16)$$

and

$$\sup_{\lambda \in W^\perp} \int_{W \cap S^\perp} J(\eta, \lambda) d\eta \leq C_M \delta^{\varepsilon/2}. \quad (4.17)$$

These two estimates yield, for $M \geq \dim(W^\perp) + 1$,

$$\begin{aligned} |\langle \nu, \phi_\delta \rangle| &\lesssim C_M \delta^{\varepsilon/2} \int_{W^\perp} (1 + |\lambda|)^{-M} d\lambda \\ &\lesssim C_M \delta^{\varepsilon/2} \rightarrow 0 \end{aligned}$$

as $\delta \rightarrow 0$, as required.

It remains to establish the estimates in (4.16) and (4.17). For the former, we observe that the left hand side of the inequality is $\Lambda_{\eta+\lambda}^*(g)$, so the desired conclusion follows from Proposition 3.1 and our choice (4.15) of ε .

To prove (4.17), we recall (4.12) so we may parametrize $\eta = sw_0$ for some fixed unit vector $w_0 \in W \cap S^\perp \setminus \{0\}$, with $d\eta = ds$. To confirm that $J(\eta, \lambda)$ has decay in η , we need to verify that $\mathbb{A}^t w_0 \cdot \mathbf{b} \neq 0$. Indeed, if $\mathbb{A}^t w_0 \cdot \mathbf{b} = 0$, then $\mathbb{A}^t w_0 \in V$ since $\mathbf{b} \in V^\perp$. Since w_0 also lies in W given by (4.11), this implies $\mathbb{A}^t w_0 \cdot \mathbb{A}^t w_0 = 0$, so that $\mathbb{A}^t w_0 = 0$. But the last equation says $w_0 \in S$, whereas $w_0 \in S^\perp$ by assumption. This forces $w_0 = 0$, a contradiction to $\|w_0\| = 1$.

We now set $c_0 := \mathbb{A}^t w_0 \cdot \mathbf{b}$ which is nonzero by the discussion in the preceding paragraph. Making a linear change of variable $t = sc_0 + \mathbb{A}^t \lambda \cdot \mathbf{b}$, with Jacobian $ds = dt/c_0$, we proceed to estimate the integral in (4.17) by partitioning the region of integration as follows,

$$\begin{aligned}
& \int_{W \cap S^\perp} J(\eta, \lambda) d\eta \\
&= \delta C_M \int_{\mathbb{R}} (1 + |t|)^{-\varepsilon} (1 + \delta|t|)^{-M} \frac{dt}{c_0} \\
&= \delta C_M \int_{|t| \leq \delta^{-1/2}} (1 + |t|)^{-\varepsilon} (1 + \delta|t|)^{-M} dt \\
&\quad + \delta C_M \int_{|t| > \delta^{-1/2}} (1 + |t|)^{-\varepsilon} (1 + \delta|t|)^{-M} dt \\
&\lesssim \delta C_M \int_{|t| \leq \delta^{-1/2}} 1 dt + \delta^{\varepsilon/2} \delta C_M \int_{|t| > \delta^{-1/2}} (1 + \delta|t|)^{-M} dt \\
&\lesssim \delta^{1/2} C_M + \delta^{\varepsilon/2} \delta C_M \int_{\mathbb{R}} (1 + \delta|\eta|)^{-M} d\eta \\
&\approx \delta^{1/2} C_M + \delta^{\varepsilon/2} C_M \lesssim \delta^{\varepsilon/2} C_M.
\end{aligned}$$

This completes the proof of (4.17) and hence the proof of the proposition. \square

Now we prove the three lemmas required earlier for this proof.

Lemma 4.5. *Define W as in (4.11). Then $\dim W = nk - v = nk - (m - 1)$.*

Proof. As before, let P_V denote the projection onto V . By (4.11),

$$\begin{aligned}
W &= \{\xi \in \mathbb{R}^{nk} : \mathbb{A}^t \xi \cdot x = 0 \text{ for all } x \in V\} \\
&= \{\xi \in \mathbb{R}^{nk} : \mathbb{A}^t \xi \cdot P_V x = 0 \text{ for all } x \in \mathbb{R}^m\} \\
&= \{\xi \in \mathbb{R}^{nk} : P_V^t \mathbb{A}^t \xi \cdot x = 0 \text{ for all } x \in \mathbb{R}^m\} \\
&= \{\xi \in \mathbb{R}^{nk} : P_V^t \mathbb{A}^t \xi\} \\
&= \mathcal{N}(P_V^t \mathbb{A}^t),
\end{aligned}$$

Writing \mathbb{R}^m as $V \oplus V^\perp$, the dimension of $\mathbb{A}(V)$ must be equal to $\dim V = m - 1$ as \mathbb{A} is of full rank and hence an isomorphism from \mathbb{R}^m to the range of \mathbb{A} . Then $\mathbb{A}P_V$ is an isomorphism from V to the range of $\mathbb{A}P_V$, so $\text{rank}(P_V^t \mathbb{A}^t) = \dim(V) = m - 1$, and thus $\dim W = nk - (m - 1)$. \square

Lemma 4.6. *With $\mu, \xi, \eta, \lambda, \beta$ defined as in the proof of Proposition 4.1(c), we have*

$$\left| \prod_{j=1}^k \widehat{\mu}(\xi_j) \right| \lesssim (1 + |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-\varepsilon} \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2+\varepsilon},$$

for any $\varepsilon > 0$.

Proof. The decay condition (4.1) on $\widehat{\mu}$ gives

$$\left| \prod_{j=1}^k \widehat{\mu}(\xi_j) \right| \leq C \prod_{j=1}^k (1 + |\xi_j|)^{-\varepsilon} \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2+\varepsilon},$$

for any $\varepsilon > 0$. We have

$$|\eta + \lambda| \leq |\xi| \leq k \max_{1 \leq j \leq k} |\xi_j|,$$

and by Cauchy-Schwarz

$$|\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}| = |(\eta + \lambda) \cdot \mathbb{A}\mathbf{b}| \leq |\eta + \lambda| |\mathbb{A}\mathbf{b}|.$$

Since $\mathbb{A}\mathbf{b}$ is fixed, $|\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}| \lesssim |\eta + \lambda|$, and so

$$\begin{aligned} \left| \prod_{j=1}^k \widehat{\mu}(\xi_j) \right| &\lesssim (1 + |\eta + \lambda|)^{-\varepsilon} \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2+\varepsilon} \\ &\lesssim (1 + |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-\varepsilon} \prod_{j=1}^k (1 + |\xi_j|)^{-\beta/2+\varepsilon}. \end{aligned}$$

□

Lemma 4.7. *With $\phi_\delta, \xi, \zeta, \eta, \lambda, \beta$ defined as in the proof of Proposition 4.1(c), we have*

$$|\widehat{\phi}_\delta(-\mathbb{A}^t\xi)| \leq \delta C_M (1 + |\lambda|)^{-M} (1 + \delta |\mathbb{A}^t(\eta + \lambda) \cdot \mathbf{b}|)^{-M},$$

for any $M \in \mathbb{R}$.

Proof. Since $\zeta \in S$ implies $\mathbb{A}^t\zeta = 0$, we have

$$\begin{aligned} \widehat{\phi}_\delta(-\mathbb{A}^t\xi) &= \widehat{\phi}_\delta(-\mathbb{A}^t(\zeta + \eta + \lambda)) \\ &= \widehat{\phi}_\delta(-\mathbb{A}^t(\eta + \lambda)) \\ &= \iint \phi_V(\vec{a}) \phi_{V^\perp}(\delta^{-1}b) e^{2\pi i \mathbb{A}^t(\eta + \lambda) \cdot (u+w)} du dw. \end{aligned}$$

By definition, $\eta \in W$ and $u \in V$ give $\mathbb{A}^t\eta \cdot u = 0$, and so

$$\widehat{\phi}_\delta(-\mathbb{A}^t\xi) = \left[\int_V \phi_V(\vec{a}) e^{2\pi i \mathbb{A}^t\lambda \cdot A_0 \vec{a}} du \right] \left[\int_{V^\perp} \phi_{V^\perp}(\delta^{-1}b) e^{2\pi i b \mathbb{A}^t(\eta + \lambda) \cdot b} dw \right]$$

$$= \left[\int_{\mathbb{R}^{m-1}} \phi_V(\vec{a}) e^{2\pi i A_0^t \mathbb{A}^t \lambda \cdot \vec{a}} d\vec{a} \right] \left[\int_{\mathbb{R}} \phi_{V^\perp}(\delta^{-1}b) e^{2\pi i b \mathbb{A}^t (\eta + \lambda) \cdot \mathbf{b}} db \right],$$

where we have used $u = \sum_{i=1}^{m-1} a_i \mathbf{a}_i = A_0 \vec{a}$ for some matrix A_0 , $w = b \mathbf{b}$, and $du dw = da db$.

The first factor is by definition $\widehat{\phi}_V(-A_0^t \mathbb{A}^t \lambda)$. Since $\widehat{\phi}_V \in \mathcal{S}(\mathbb{R}^{m-1})$, for every $M \in \mathbb{R}$ we have

$$|\widehat{\phi}_V(-A_0^t \mathbb{A}^t \lambda)| \leq C_M (1 + |A_0^t \mathbb{A}^t \lambda|)^{-M}.$$

We claim $|A_0^t \mathbb{A}^t \lambda| \gtrsim |\lambda|$ for all $\lambda \in W^\perp$. Since $A_0^t \mathbb{A}^t$ is linear, it suffices to prove $A_0^t \mathbb{A}^t \lambda \neq 0$ for any $\lambda \in W^\perp$. If $\lambda \in W^\perp$, then by definition of W^\perp there exists $x \in V \setminus \{0\}$ such that $(\mathbb{A}^t \lambda, x) \neq 0$. Then $x = A_0 \vec{a}$ for some $\vec{a} \neq 0$, so

$$(A_0^t \mathbb{A}^t \lambda, \vec{a}) = (\mathbb{A}^t \lambda, A_0^t \vec{a}) \neq 0,$$

and hence $A_0^t \mathbb{A}^t \lambda \neq 0$. Then for every $M \in \mathbb{R}$,

$$|\widehat{\phi}_V(-A_0^t \mathbb{A}^t \lambda)| \leq C_M (1 + |\lambda|)^{-M}.$$

The second factor is, upon scaling, $\delta \widehat{\phi}_{V^\perp}(-\delta \mathbb{A}^t (\eta + \lambda) \cdot \mathbf{b})$. As $\widehat{\phi}_{V^\perp} \in \mathcal{S}(\mathbb{R})$, for every $M \in \mathbb{R}$ we have

$$|\delta \widehat{\phi}_{V^\perp}(-\delta \mathbb{A}^t (\eta + \lambda) \cdot \mathbf{b})| \leq \delta (1 + \delta |\mathbb{A}^t (\eta + \lambda) \cdot \mathbf{b}|)^{-M},$$

completing the proof. □

5 Absolutely continuous estimates

In this section, we will restrict to the case when A_j is of the form $A_j = (I_{n \times n} B_j)$, where B_j are $n \times (m - n)$ matrices. Set $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^{m-n}$, so that our configurations are of the form $\{x + B_1 y, \dots, x + B_k y\}$. With r defined as in (1.3), we will also assume $k - 1 \geq 2r$, or equivalently, $n \lceil (k + 1)/2 \rceil \leq m$, the first inequality of condition (1.5) of the main result.

Proposition 5.1. *For every $\delta, M > 0$, there exists a constant $c(\delta, M) > 0$ with the following property: for every function $f : [0, 1]^n \rightarrow \mathbb{R}$, $0 \leq f \leq M$, $\int f \geq \delta$, we have $\Lambda(f) \geq c(\delta, M)$.*

We will proceed as in the proof of Varnavides' Theorem given in [33]. The strategy will be to decompose $f = g + b$ into a “good” function g which is the major contribution and a “bad” function b whose contribution is negligible. This will be made precise in the following subsection.

5.1 Preliminaries

Proposition 5.2. *Let f be as in Proposition 5.1. Suppose $f = g + b$ where*

$$\|g\|_\infty, \|b\|_\infty \leq M; \quad \|g\|_1, \|b\|_1 = \delta.$$

Then

$$\Lambda(f) = \Lambda(g) + O(C(M, \delta) \|\widehat{b}\|_\infty).$$

Proof. We use the decomposition $f = g + b$ and the linearity of Λ to decompose $\Lambda(f)$ into 2^k pieces. The main piece will be $\Lambda(g)$ and the remaining pieces which constitute the error term have at least one copy of b . By the hypothesis and Hölder's inequality,

$$\|g\|_2^2, \|b\|_2^2 \leq M\delta.$$

We will apply Lemma 5.3 below to estimate each of the $2^k - 1$ summands in the error term, arriving at an upper bound of $(2^k - 1)\|\widehat{b}\|_\infty(M\delta)^r$. \square

We now prove the lemma required for the previous proposition.

Lemma 5.3. *Let f_j be as in Proposition 5.1. Assume moreover that $k-1 \geq 2r$ and $\|f_j\|_1 \leq 1$ for $1 \leq j \leq k$. Then*

$$|\Lambda(f_1, \dots, f_k)| \leq M\|\widehat{f}_k\|_\infty \|f_r\|_1^{1/2} \|f_{2r}\|_1^{1/2} \prod_{\substack{j=1 \\ j \neq r}}^{2r-1} \|\widehat{f}_j\|_2.$$

We have a similar bound for permutations of f_1, \dots, f_k .

Proof. Let us recall the Fourier representation of Λ from Proposition 2.1, which gives

$$|\Lambda(f_1, \dots, f_k)| \leq \int_S \prod_{j=1}^k |\widehat{f}_j(\xi_j)| d\sigma(\xi).$$

Since $\|\widehat{f}_j\|_\infty \leq \|f_j\|_1 \leq 1$ for each j , reducing the number of factors in the product that appears in the last integrand only makes the integral larger. We use the hypothesis $k-1 \geq 2r$ to drop $(k-1-2r)$ of these factors and split the remaining $2r$ into two groups and apply the Cauchy-Schwarz inequality. Executing these steps leads to

$$\begin{aligned} |\Lambda(f_1, \dots, f_k)| &\leq \int_S \prod_{j=1}^k |\widehat{f}_j(\xi_j)| d\sigma(\xi) \\ &\leq \|\widehat{f}_k\|_\infty \int_S \prod_{j=1}^r |\widehat{f}_j(\xi_j)| \prod_{j=r+1}^{2r} |\widehat{f}_j(\xi_j)| d\sigma(\xi) \\ &\leq \|\widehat{f}_k\|_\infty \left(\int_S \prod_{j=1}^r |\widehat{f}_j(\xi_j)|^2 d\sigma(\xi) \right)^{1/2} \left(\int_S \prod_{j=r+1}^{2r} |\widehat{f}_j(\xi_j)|^2 d\sigma(\xi) \right)^{1/2}. \end{aligned}$$

Both of the above integrals are estimated in the same way; we will focus only on the first. If $\xi_j = (\xi_{j,1}, \dots, \xi_{j,n})$, let $\xi'_j = (\xi_{j,1}, \dots, \xi_{j,n'})$ with n' as defined in Definition 1.4; notice this is the same as ξ_j^{id} as defined in Proposition 3.1. By Lemma 5.4 below, $\|\widehat{f}_r\|_{L^2_{\xi'_r}} \leq M\|f_r\|_1$, and so by Lemma 3.2,

$$\int_S \prod_{j=1}^r |\widehat{f}_j(\xi_j)|^2 d\sigma(\xi)$$

$$\begin{aligned}
&= \int_S \prod_{j=1}^r |\widehat{f}_j(\xi_j)|^2 d\xi'_r d\xi_1 \cdots d\xi_{r-1} \\
&\leq \int_{\mathbb{R}^{n(r-1)}} M \|f_r\|_1 \prod_{j=1}^{r-1} |\widehat{f}_j(\xi_j)|^2 d\xi_1 \cdots d\xi_{r-1} \\
&= M \|f_r\|_1 \prod_{j=1}^{r-1} \|\widehat{f}_j\|_2^2.
\end{aligned}$$

The result follows. \square

Lemma 5.4. *Let $1 \leq j \leq k$. If $\xi_j = (\xi_{j,1}, \dots, \xi_{j,n})$, we denote $\xi'_j = (\xi_{j,1}, \dots, \xi_{j,n'})$ and $\xi''_j = (\xi_{j,n'+1}, \dots, \xi_{j,k})$. Suppose*

(a) $|f_j| \leq M$,

(b) $\text{supp } f_j \subseteq [0, 1]^n$.

Then

$$\|\widehat{f}_j\|_{L^2_{\xi'_j}}^2 \leq M \|f_j\|_1$$

uniformly for all ξ''_j .

Proof. Fix ξ''_j , and let $F(x', \xi''_j) = \int f_j(x', x'') e^{-2\pi i \xi''_j \cdot x''} dx''$, where x', x'' are the dual variables to ξ'_j, ξ''_j respectively. Then (a) and (b) give $|F(x', \xi''_j)| \leq M$ for all x' , and we calculate

$$\int |F(x', \xi''_j)| dx' \leq \iint |f_j(x', x'')| dx' dx'' = \|f_j\|_1.$$

By Hölder's inequality, $\|F\|_{L^2_{x'}}^2 \leq \|F\|_\infty \|F\|_{L^1_{x'}} \leq M \|f_j\|_1$. Now,

$$\widehat{f}_j(\xi'_j, \xi''_j) = \int F(x', \xi''_j) e^{-2\pi i \xi'_j \cdot x'} dx',$$

which is the Fourier transform of F in x' . Therefore by Plancherel's theorem in the x' variables, $\|\widehat{f}_j\|_{L^2_{\xi'_j}}^2 = \|F\|_{L^2_{x'}}^2 \leq M \|f_j\|_1$. \square

5.2 Almost periodic functions

In light of Proposition 5.2, our next goal will be to identify a large class of “good” functions g for which we can bound $\Lambda(g)$ from below. It turns out that almost periodic functions, defined analogously to [33], can be used for this purpose.

Definition 5.5. 1. A character is a function $\chi : [0, 1]^n \rightarrow \mathbb{C}$ of the form $\chi(x) = e^{2\pi i v \cdot x}$ for some $v \in \mathbb{Z}^n$.

2. If $K \in \mathbb{N}$, then a K -quasiperiodic function is a function f of the form $\sum_{\ell=1}^K c_\ell \chi_\ell$ where each χ_ℓ are characters (not necessarily distinct), and c_ℓ are scalars with $|c_\ell| \leq 1$.

3. If $\sigma > 0$, then $f : [0, 1]^n \rightarrow \mathbb{C}$ is (σ, K) -almost periodic if there exists a K -quasiperiodic function f_{QP} such that $\|f - f_{QP}\|_{L^2([0,1]^n)} \leq \sigma$. We call f_{QP} a K -quasiperiodic function approximating f within σ .

Lemma 5.6. Let $K \in \mathbb{N}$, $M > 0$, $0 < \delta < 1$, and

$$0 < \sigma \leq \frac{\delta^k}{4kM^{k-1}}. \quad (5.1)$$

Then there exists $c(K, \delta, M) > 0$ such that for any non-negative (σ, K) -almost periodic function f bounded by M and obeying $\int f \geq \delta$,

$$\Lambda(f) \geq c(K, \delta, M).$$

Proof. Our goal is to bound $\Lambda(f)$ from below by a multiple of $\|f\|_1^k$, which is known to be at least as large as δ^k . We will achieve this by approximating each factor in the integral defining Λ by f , on a reasonably large set with acceptable error terms.

To this end, let f_{QP} be a K -quasiperiodic function approximating f within σ , say $f_{QP}(x) = \sum_{\ell=1}^K c_\ell e^{2\pi i v_\ell \cdot x}$ and $\|f - f_{QP}\|_2 \leq \sigma$. Let $\varepsilon > 0$ be a small constant to be fixed later, and define

$$C_\varepsilon = \{y \in \mathbb{R}^{m-n} : \|A_j^t v_\ell \cdot y\| \leq \varepsilon, \text{ for all } 1 \leq j \leq k, 1 \leq \ell \leq K\}, \quad (5.2)$$

where $\|t\|$ denotes the distance of $t \in \mathbb{R}$ to the nearest integer. We shall prove in Corollary B.2 that

$$|C_\varepsilon| \geq c(\varepsilon, K) > 0 \quad (5.3)$$

for some $c(\varepsilon, K)$ possibly depending on k , m and n but is independent of f . Let T^a be the shift map $T^a f(x) := f(x + a)$. For $y \in C_\varepsilon$ and any $x \in [0, 1]^n$,

$$\begin{aligned} |T^{B_j y} f_{QP}(x) - f_{QP}(x)| &= \left| \sum_{\ell=1}^K c_\ell e^{2\pi i v_\ell \cdot x} (1 - e^{2\pi i v_\ell \cdot A_j y}) \right| \\ &\leq \sum_{\ell=1}^K |1 - e^{2\pi i A_j^t v_\ell \cdot y}| \\ &\leq \sum_{\ell=1}^K |A_j^t v_\ell \cdot y| \\ &\leq K\varepsilon. \end{aligned} \quad (5.4)$$

The bound above leads to the following estimate:

$$\begin{aligned} \|T^{B_j y} f - f\|_{L_x^1} &\leq \|T^{B_j y} f - T^{B_j y} f_{QP}\|_{L_x^2} + \|T^{B_j y} f_{QP} - f_{QP}\|_{L_x^\infty} + \|f_{QP} - f\|_{L_x^2} \\ &= 2\|f_{QP} - f\|_{L_x^2} + \|T^{B_j y} f_{QP} - f_{QP}\|_{L_x^\infty} \\ &\leq 2\sigma + K\varepsilon \\ &\leq \frac{\delta^k}{2kM^{k-1}} + K\varepsilon. \end{aligned}$$

In the sequence of inequalities above, we have used Hölder's inequality in the first step, triangle inequality in the second step, norm-invariance of the shift operator in the third step, and (5.4) in the last step.

We now choose $\varepsilon = \delta^k/(4kM^{k-1})$, so that

$$\|T^{B_j y} f - f\|_{L_x^1} \leq \frac{3\delta^k}{4kM^{k-1}}. \quad (5.5)$$

For every $1 \leq j \leq k$, the bound $\|T^{B_j y} f\|_{L_x^\infty} = \|f\|_\infty \leq M$ holds trivially, so by Lemma B.3 with $C = M$, $f_j = f$, $g_j = T^{B_j y} f$, $R = k$, $p = 1$, and $\kappa = (3\delta^k)/(4kM^{k-1})$, we have

$$\left\| \prod_{j=1}^k T^{B_j y} f - f^k \right\|_{L_x^1} \leq kM^{k-1} \frac{3\delta^k}{4kM^{k-1}} = \frac{3\delta^k}{4}$$

using (5.5). On the other hand, the bounded non-negativity of f , the hypothesis $\int f \geq \delta$, and Hölder's inequality lead to

$$\|f^k\|_1 \geq \|f\|_1^k \geq \delta^k,$$

and so for $y \in C_\varepsilon$,

$$\left\| \prod_{j=1}^k T^{B_j y} f \right\|_{L_x^1} \geq \|f^k\|_1 - \left\| \prod_{j=1}^k T^{B_j y} f - f^k \right\|_{L_x^1} \geq \frac{\delta^k}{4}. \quad (5.6)$$

We now combine 5.3, the positivity of f and the above to obtain

$$\begin{aligned} \Lambda(f) &= \int_{\mathbb{R}^{m-n}} \int_{\mathbb{R}^n} \prod_{j=1}^k f_j(x + B_j y) \, dx \, dy \\ &= \int_{\mathbb{R}^{m-n}} \left\| \prod_{j=1}^k T^{B_j y} f \right\|_{L_x^1} \, dy \\ &\geq \int_{C_\varepsilon} \left\| \prod_{j=1}^k T^{B_j y} f \right\|_{L_x^1} \, dy \\ &\geq \delta^k c(\varepsilon, K)/4 = c(K, \delta, M). \end{aligned}$$

□

5.3 Ubiquity of almost periodic functions

To make use of Lemma 5.6, we will approximate a general function f by an almost periodic function. In the following sequence of lemmas, we construct an increasingly larger family of σ -algebras with the property that any function measurable with respect to these will be almost periodic. We do this by an iterative random mechanism, the building block of which is summarized in the next result.

Lemma 5.7. *Let $0 < \varepsilon \ll 1$ and let χ be a character. Viewing \mathbb{C} as \mathbb{R}^2 , partition the complex plane $\mathbb{C} = \bigcup_{Q \in \mathbb{Q}_\varepsilon} Q$ into squares of side-length ε with corners lying in the lattice $\varepsilon\mathbb{Z}^2$. For $\omega \in [0, 1]^2$, define $\mathcal{B}_{\varepsilon, \chi, \omega}$ to be the σ -algebra generated by the atoms*

$$\{\chi^{-1}(Q + \varepsilon\omega) : Q \in \mathbb{Q}_\varepsilon\}.$$

There exists ω such that

1. $\|\chi - \mathbb{E}(\chi | \mathcal{B}_{\varepsilon, \chi, \omega})\|_\infty \leq C\varepsilon$.
2. *For every $\sigma > 0$ and $M < 0$, there exists $K = K(\sigma, \varepsilon, M)$ such that every function f which is measurable with respect to $\mathcal{B}_{\varepsilon, \chi, \omega}$ with $\|f\|_\infty \leq M$ is (σ, K) -almost periodic.*

Proof. (1) follows from definition of $\mathcal{B}_{\varepsilon, \chi, \omega}$, for any ω .

To prove (2), it suffices to prove: for each integer $\ell > \ell_0$ for some sufficiently large ℓ_0 , there exists a set $\Omega_\ell \subseteq [0, 1]^2$, $|\Omega_\ell| > 1 - C2^{-\ell}\varepsilon$ with the following property. For $\omega \in \Omega_\ell$, there exists $K = K(\sigma, \varepsilon, M)$ such that every function f which is measurable with respect to $\mathcal{B}_{\varepsilon, \chi, \omega}$ with $\|f\|_\infty \leq M$ is (σ, K) -almost periodic, for $\sigma = 2^{-\ell}$.

Indeed, taking $\Omega = \bigcap_{\ell > \ell_0} \Omega_\ell$, we have

$$|\Omega| > 1 - C \sum_{\ell > \ell_0} 2^{-\ell}\varepsilon \geq 1 - C2^{-\ell_0}\varepsilon$$

and so we may find $\omega \in \Omega$. Then if $\sigma > 0$, get $\ell > \ell_0$ with $2^{-\ell} \leq \sigma$, and by the above, there exists $K = K(2^{-\ell}, \varepsilon, M)$ such that every function f which is measurable with respect to $\mathcal{B}_{\varepsilon, \chi, \omega}$ with $\|f\|_\infty \leq M$ is $(2^{-\ell}, K)$ -almost periodic.

Fix $\sigma = 2^{-\ell}$. We prove in Lemma B.4 that $\mathcal{B}_{\varepsilon, \chi, \omega}$ has at most $C\varepsilon^{-1}$ atoms, which we use to reduce the proof of Lemma 5.7 to a simpler form. Namely, it suffices to prove that for f an indicator function of one of those atoms, say

$$f(x) = f_{Q, \omega}(x) := 1_{\chi^{-1}(Q + \varepsilon\omega)}(x) = 1_Q(\chi(x) - \varepsilon\omega),$$

there exists a $C(\sigma, \varepsilon)$ -quasiperiodic function $g_{Q, \omega}$ such that $\|f_{Q, \omega} - g_{Q, \omega}\|_2 \leq C^{-1}\sigma\varepsilon$ with probability $1 - C\sigma\varepsilon^{-1}$. (Here and below, $C(\sigma, \varepsilon)$ will denote a constant which may change from line to line, but always depends only on M, σ, ε , in particular remains independent of f .) Indeed, if this were the case, then any measurable f may be written as

$$f(x) = \sum_{i \in I} c_i f_{Q_i, \omega}(x)$$

where $f_{Q_i, \omega}(x) = 1_{Q_i}(\chi(x) - \varepsilon\omega)(x)$, $Q_i \in \mathbb{Q}_\varepsilon$ are distinct, and $\#I \leq C\varepsilon^{-1}$. Notice $\|f\|_\infty \leq M$ and the fact that $f_{Q_i, \omega}$ have disjoint support means $|c_i| \leq M$ for $i \in I$. Letting $g_{Q_i, \omega}$ be a $C(\sigma/M, \varepsilon)$ -quasiperiodic function approximating $f_{Q_i, \omega}$ to within $C^{-1}M^{-1}\sigma\varepsilon$, we have that $g = \sum_{i \in I} c_i g_{Q_i, \omega}$ is $C(\sigma, \varepsilon)$ -quasiperiodic (repeating $g_{Q_i, \omega}$ at most M times if necessary) and

$$\|f - g\|_2 \leq \sum_{i \in I} |c_i| \|f_{Q_i, \omega} - g_{Q_i, \omega}\|_2 \leq \sigma.$$

Thus we restrict to the case when f is the indicator function of one of those atoms, which we denote f_ω for ease.

$$f(x) = f_\omega(x) := 1_Q(\chi(x) - \varepsilon\omega).$$

Let $0 \leq h \leq 1$ be a continuous function on $[-2, 2]^2$ which is a good L^2 -approximation for 1_Q ; precisely,

$$\|1_Q - h\|_{L^2([-2, 2]^2)}^2 < \frac{\sigma^3 \varepsilon^3}{10CM^2}. \quad (5.7)$$

By the Weierstrass Approximation Theorem, there exists a polynomial P such that

$$\|h - P\|_{L^\infty([-2, 2]^2)} < \left(\frac{\sigma^3 \varepsilon^3}{10CM^2} \right)^{1/2}. \quad (5.8)$$

Let $h_\omega(x) := h(\chi(x) - \varepsilon\omega)$, and $g_\omega(x) = P(\chi(x) - \varepsilon\omega)$; notice g_ω can be written as a linear combination of at most $C(\sigma, \varepsilon)$ characters, with coefficients at most $C(\sigma, \varepsilon)$. Repeating characters if necessary, we may reduce the coefficients to be less than 1 and so g_ω is $C(\sigma, \varepsilon)$ -quasiperiodic. It should also be noted that $\|g_\omega\|_\infty \leq \|g_\omega - h_\omega\|_\infty + \|h_\omega\|_\infty \leq 2$.

It remains to show $\|f_\omega - g_\omega\|_2 \leq C^{-1}M^{-1}\sigma\varepsilon$ with probability at least $1 - C\sigma\varepsilon^{-1}$. Define

$$F(\omega) = \|f_\omega - g_\omega\|_2^2 = \int_{[0, 1]^n} |f_\omega(x) - g_\omega(x)|^2 dx.$$

By an application of Cauchy-Schwarz inequality on the integrand, combined with (5.7) and Tonelli's Theorem, we obtain

$$\begin{aligned} \|F\|_{L^1_\omega} &= \int_{[0, 1]^2} \int_{[0, 1]^n} |f_\omega(x) - g_\omega(x)|^2 dx d\omega \\ &\leq 2 \int_{[0, 1]^2} \int_{[0, 1]^n} [|f_\omega(x) - h_\omega(x)|^2 + |h_\omega(x) - g_\omega(x)|^2] dx d\omega \\ &< \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \int_{[0, 1]^2} |h_\omega(x) - g_\omega(x)|^2 d\omega dx \\ &< \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \int_{[0, 1]^2} |h(\chi(x) - \varepsilon\omega) - P(\chi(x) - \varepsilon\omega)|^2 d\omega dx. \end{aligned}$$

Using the change of variables $\omega' = \chi(x) - \varepsilon\omega$ for fixed $x \in [0, 1]^n$, we have $d\omega' = \varepsilon^2 d\omega$. Notice ω' belongs to $[0, \varepsilon]^2$ shifted by $\chi(x)$, so is contained in $[-2, 2]^2$. By (5.8),

$$\begin{aligned} \|F\|_1 &\leq \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \int_{\omega' \in [-2, 2]^2} |h(\omega') - P(\omega')|^2 \frac{d\omega'}{\varepsilon^2} dx \\ &\leq \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \int_{\omega' \in [-2, 2]^2} |h(\omega') - P(\omega')|^2 \frac{d\omega'}{\varepsilon^2} dx \\ &= \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \frac{1}{\varepsilon^2} \|h - P\|_{L^2([-2, 2]^2)}^2 dx \\ &\leq \frac{\sigma^3 \varepsilon^3}{2CM^2} + \int_{[0, 1]^n} \frac{1}{\varepsilon^2} \left(\frac{\sigma^3 \varepsilon^3}{10CM^2} \right) dx < \frac{1}{\varepsilon^2} \left(\frac{\sigma^3 \varepsilon^3}{CM^2} \right). \end{aligned}$$

By Markov's inequality,

$$\begin{aligned} |\{\omega \in [0, 1]^2 : \|f_\omega - g_\omega\|_2^2 > C^{-2}M^{-2}\sigma^2\varepsilon^2\}| &= |\{\omega \in [0, 1]^2 : F(\omega) > C^{-2}M^{-2}\sigma^2\varepsilon^2\}| \\ &\leq \frac{C^2M^2\|F\|_1}{\sigma^2\varepsilon^2} \leq \frac{C^2M^2\frac{1}{\varepsilon^2}\left(\frac{\sigma^3\varepsilon^3}{CM^2}\right)}{\sigma^2\varepsilon^2} = C\sigma\varepsilon^{-1}. \end{aligned}$$

Thus,

$$|\{\omega \in [0, 1]^2 : \|f_\omega - g_\omega\|_2^2 \leq C^{-1}M^{-1}\sigma\varepsilon\}| \geq 1 - C\sigma\varepsilon^{-1},$$

as required. \square

The above proof also gives the following result.

Corollary 5.8. *Let $0 < \varepsilon \ll 1$ and let χ be a character. Then the σ -algebra $\mathcal{B}_{\varepsilon, \chi, \omega}$ described in the statement of Lemma 5.7 can be chosen to have the additional property that for every atom $\chi^{-1}(Q + \varepsilon\omega)$, $Q \in \mathbb{Q}_\varepsilon$, there exists a K -quasiperiodic function $g_{Q, \omega}$ that obeys $\|g_{Q, \omega}(\cdot) - 1_Q(\chi(\cdot) - \varepsilon\omega)\|_2 < \sigma$ for every $\sigma > 0$, and in addition $\|g_{Q, \omega}\|_\infty \leq 2$.*

We can concatenate the σ -algebras from Lemma 5.7. If $\mathcal{B}_1, \dots, \mathcal{B}_R$ are σ -algebras, denote by $\mathcal{B}_1 \vee \dots \vee \mathcal{B}_R$ the smallest σ -algebra which contains all of them.

Corollary 5.9. *Let $0 < \varepsilon_1, \dots, \varepsilon_R \ll 1$ and let χ_1, \dots, χ_R be characters. Let $\mathcal{B}_{\varepsilon_1, \chi_1}, \dots, \mathcal{B}_{\varepsilon_R, \chi_R}$ be the σ -algebras arising from Lemma 5.7. Then for every $\sigma > 0$, there exists $K = K(R, \sigma, \varepsilon_1, \dots, \varepsilon_R)$ such that every function f which is measurable with respect to $\mathcal{B}_{\varepsilon_1, \chi_1} \vee \dots \vee \mathcal{B}_{\varepsilon_R, \chi_R}$ with $\|f\|_\infty \leq M$ is (σ, K) -almost periodic.*

Proof. Since there are at most $C(R, \varepsilon_1, \dots, \varepsilon_R)$ atoms in $\mathcal{B}_{\varepsilon_1, \chi_1} \vee \dots \vee \mathcal{B}_{\varepsilon_R, \chi_R}$, it suffices to prove the claim in the case when f is the indicator function of a single atom. Then f is the product of R indicator functions f_1, \dots, f_R , where f_j is the indicator function of an atom from $\mathcal{B}_{\varepsilon_j, \chi_j}$. Let g_j be a $K(\sigma/(R2^{R-1}), \varepsilon_j)$ -quasiperiodic function approximating f_j to within $\sigma/(R2^{R-1})$ as provided in Corollary 5.9; notice $\|g_j\|_\infty \leq 2$. Then $g = \prod g_j$ is a K -quasiperiodic function where $K = \prod K(\sigma/(R2^{R-1}), \varepsilon_j)$ depends only on $R, \sigma, \varepsilon_1, \dots, \varepsilon_R$. Finally, by Lemma B.3 with $C = 2$, $p = 2$, and $\kappa = \sigma/(R2^{R-1})$, we have

$$\left\| \prod_{j=1}^R f_j - \prod_{j=1}^R g_j \right\|_2 \leq \sigma.$$

\square

5.4 Proof of Proposition 5.1

We will need two more auxiliary results, analogous to [33, Lemma 2.10 and 2.11].

Lemma 5.10. *Let b be a function bounded by M with $\|\widehat{b}\|_\infty \geq \sigma > 0$. Then there exists $0 < \varepsilon \ll \sigma$, a character χ , and an associated σ -algebra $\mathcal{B}_{\varepsilon, \chi}$ (as defined earlier in this section) such that*

$$\|\mathbb{E}(b|\mathcal{B}_{\varepsilon, \chi})\|_2 \geq C^{-1}\sigma.$$

Proof. Since $\|\widehat{b}\|_\infty \geq \sigma$, there exists a character χ such that

$$\left| \int_{[0,1]^n} b(x)\chi(x) dx \right| \geq \frac{\sigma}{2}. \quad (5.9)$$

On the other hand, the σ -algebra $\mathcal{B}_{\varepsilon,\chi}$ is generated by the atoms $\{\chi^{-1}(Q + \varepsilon\omega); Q \in \mathbb{Q}_\varepsilon\}$ for some ω in the unit square. On each atom, χ can vary by at most $C\varepsilon$, hence

$$\|\chi - \mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})\|_\infty \leq C\varepsilon. \quad (5.10)$$

Since b is bounded by M , (5.9) and (5.10) yield

$$\int_{[0,1]^n} b(x)\mathbb{E}(\chi|\mathcal{B}_{\varepsilon,\chi})(x) dx \geq \frac{\sigma}{2} - MC\varepsilon.$$

Conditional expectation being self-adjoint, the inequality above may be rewritten as

$$\int_{[0,1]^n} \mathbb{E}(b|\mathcal{B}_{\varepsilon,\chi})(x)\chi(x) dx \geq \frac{\sigma}{2} - MC\varepsilon.$$

Recalling that χ is bounded above by 1, the desired result now follows by choosing ε sufficiently small relative to σ and M , and applying Cauchy-Schwarz inequality to the integral on the left. \square

Lemma 5.11. *Let $F : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be an arbitrary function, let $0 < \delta \leq 1$, and let $f \geq 0$ be a function bounded by M with $\int f \geq \delta$. Let σ satisfy (5.1). Then there exists a K with $0 < K \leq C(F, \delta)$ and a decomposition $f = g + b$ where $g \geq 0$ is a bounded (σ, K) -almost periodic function with $\int g \geq \delta$, and b obeys the bound*

$$\|\widehat{b}\|_\infty \leq F(\delta, K). \quad (5.11)$$

The proof of Lemma 5.11 is exactly identical to [33, 2.11].

Proof of Proposition 5.1. Let $F : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a function to be chosen later. Decompose $f = g + b$ as in Lemma 5.11. By Lemma 5.6,

$$\Lambda(g) \geq c(K, \delta, M).$$

By Proposition 5.2, (5.11), and the above inequality,

$$\Lambda(f) \geq c(K, \delta, M) + O(C(M, \delta)F(\delta, K)).$$

By choosing F sufficiently small and since $K \leq C(F, \delta)$, we get

$$\Lambda(f) \geq c(\delta, M)$$

as required. \square

5.5 Quantitative Szemerédi bounds fail for general \mathbb{A}

At the beginning of this section, we restricted to the case when A_j is of the form $A_j = (I_{n \times n} B_j)$ where B_j are $n \times (m - n)$ matrices. The reason for this is that generic A_i will not provide a lower bound on Λ when $\int f = \delta$, even when satisfying the non-degeneracy condition.

In the case $n = 2, k = 3, m = 4$, consider the function $f = \mathbb{1}_{B((0,1),\delta)}$, the indicator of the ball centered at $(0, 1)$ with radius $\delta \leq 1/3$. Define

$$\begin{aligned} A_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \\ A_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ A_3 &= \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

It is clear that (3.5) holds for these matrices. In the integral defining Λ , we consider the conditions for $\vec{x} = (x_1, \dots, x_4)$ to be in the support of $\prod_{i=1}^3 f(A_i \vec{x})$. The first term of the product gives $f(A_1 \vec{x}) = f(x_1, x_2)$, and so in particular, $|x_2 - 1| < \delta$ which implies

$$|x_2| > 1 - \delta. \tag{5.12}$$

Similarly, considering the second term yields in particular

$$|x_3| < \delta, \tag{5.13}$$

while the third term gives

$$|x_2 + x_3| < \delta. \tag{5.14}$$

On the other hand, (5.12) and (5.13) give

$$|x_2 + x_3| \geq |x_2| - |x_3| > 1 - 2\delta \geq \delta.$$

Then the support of $\prod_{i=1}^3 f(A_i \vec{x})$ is empty, and $\Lambda = 0$.

6 Proof of the main theorem

The preceding section gave a quantitative lower bound on the Λ quantity in the case of absolutely continuous measures with bounded density. This suggests the strategy of decomposing the measure μ as $\mu = \mu_1 + \mu_2$ where μ_1 is absolutely continuous with bounded density, and μ_2 gives negligible contribution. In light of the Fourier form of Λ , the key property of μ_2 here will be having good bounds on the Fourier transform.

Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ be a non-negative function supported on $B(0, 1)$ with $\int \phi = 1$. For any positive integer N , define $\phi_N(x) = N^n \phi(Nx)$. Let $N \gg 1$ be a large constant to be determined later, and let

$$\mu_1(x) = \mu * \phi_N(x).$$

Clearly, $\mu_1 \geq 0$ is a C^∞ function of compact support with $\int d\mu_1 = 1$. Since ϕ_N is supported on $B(0, N^{-1})$,

$$\begin{aligned} |\mu_1(x)| &\leq \int_{B(x, N^{-1})} |\phi_N(x-y)| d\mu(y) \\ &= \int_{B(x, N^{-1})} N^n |\phi(N(x-y))| d\mu(y) \\ &\leq CN^n \mu(B(x, N^{-1})) \\ &\leq CN^{n-\alpha} \end{aligned}$$

where the last inequality follows by the ball condition (a). Then $|\mu_1(x)| \leq M = Ce$ if $N = e^{1/(n-\alpha)}$, which tends to infinity as $\alpha \rightarrow n^-$.

Focusing now on μ_2 , we will prove that

$$|\widehat{\mu}_2(\xi)| \lesssim N^{-\frac{\varepsilon\beta}{2}} (1 + |\xi|)^{-\frac{\beta}{2}(1-\varepsilon)} \quad (6.1)$$

for some constant $\varepsilon > 0$ to be chosen later. Since $\int \phi = 1$ and $\phi \in \mathcal{S}(\mathbb{R}^n)$,

$$|1 - \widehat{\phi}(\xi)| = |\widehat{\phi}(0) - \widehat{\phi}(\xi)| = \left| \int_0^1 \frac{d}{dt} \widehat{\phi}(t\xi) dt \right| = \int_0^1 |\xi \cdot \nabla \phi| dt \leq C|\xi|.$$

In particular, defining $\mu_2 = \mu - \mu_1$ we have

$$|\widehat{\mu}_2(\xi)| \lesssim |\widehat{\mu}(\xi)| \min(1, |\xi|N^{-1}).$$

Notice if $|\xi| \geq N$, then

$$\begin{aligned} |\widehat{\mu}_2(\xi)| &\leq |\widehat{\mu}(\xi)| \\ &\lesssim (1 + |\xi|)^{-\beta/2} \\ &= (1 + |\xi|)^{-\varepsilon\beta/2} (1 + |\xi|)^{-\beta/2(1-\varepsilon)} \\ &\lesssim N^{-\varepsilon\beta/2} (1 + |\xi|)^{-\beta/2(1-\varepsilon)}. \end{aligned}$$

On the other hand, if $|\xi| < N$, then we still have

$$\begin{aligned} |\widehat{\mu}_2(\xi)| &\leq |\widehat{\mu}(\xi)| \\ &\lesssim (1 + |\xi|)^{-\beta/2} |\xi| N^{-1} \\ &= (1 + |\xi|)^{-\beta/2} |\xi|^{\varepsilon\beta/2} |\xi|^{1-\varepsilon\beta/2} N^{-1} \\ &\lesssim N^{-\varepsilon\beta/2} (1 + |\xi|)^{-\beta/2(1-\varepsilon)}. \end{aligned}$$

Now, decompose

$$\Lambda^*(\widehat{\mu}) = \Lambda^*(\widehat{\mu}_1) + \Lambda(\widehat{\mu}_2, \widehat{\mu}_1, \dots, \widehat{\mu}_1) + \dots + \Lambda(\widehat{\mu}_2).$$

By Proposition 5.1, $\Lambda^*(\widehat{\mu}_1) = \Lambda(\mu_1) > c(\delta, M)$. It remains to show the Λ^* quantities containing at least one copy of μ_2 are negligible relative to $c(\delta, Ce)$. These quantities can be

written as $\Lambda^*(g_1, \dots, g_k)$ where for each $1 \leq j \leq k$, g_j is either $\widehat{\mu}_1$ or $\widehat{\mu}_2$ and at least one g_j is $\widehat{\mu}_2$. Without loss of generality, suppose $g_1 = \widehat{\mu}_2$, so that

$$|g_1(\eta_1)| \lesssim N^{-\varepsilon\beta/2}(1 + |\eta_j|)^{-\beta/2(1-\varepsilon)}$$

by the above estimate on $\widehat{\mu}_2$. For $j \geq 2$, we have

$$|g_j(\eta_j)| \lesssim (1 + |\eta_j|)^{-\beta/2(1-\varepsilon)}$$

by the above estimate on $\widehat{\mu}_2$ and the general Fourier decay condition (b) on μ . Then

$$\begin{aligned} \Lambda^*(g_1, \dots, g_k) &= \int_S \prod_{j=1}^k g_j(\eta_j) d\sigma \\ &\leq N^{-\varepsilon\beta/2} \int_S \prod_{j=1}^k (1 + |\eta_j|)^{-\beta/2(1-\varepsilon)} d\sigma. \end{aligned}$$

Since $\beta > 2(nk - m)/k$, we may choose $\varepsilon > 0$ so that $\beta' = \beta(1 - \varepsilon) > 2(nk - m)/k$. Then by Proposition 3.1 with β' in place of β , the integral above is bounded by a constant independent of N . Then we may choose N sufficiently large that $\Lambda^*(g_1, \dots, g_k) \leq 2^{-k}c(\delta, M)$, and so

$$\Lambda^*(\widehat{\mu}) \geq 2^{-k}c(\delta, M).$$

7 Examples

For a fixed choice of $n \geq 1$ and $k \geq 3$, let $m = n\lceil(k+1)/2\rceil$, the smallest value allowed by (1.5). Non-degeneracy in this case will be the condition

$$\text{rank} \begin{pmatrix} A_{i_1} \\ \vdots \\ A_{i_{m/n}} \end{pmatrix} = \text{rank} \begin{pmatrix} I_{n \times n} & B_{i_1} \\ \vdots & \vdots \\ I_{n \times n} & B_{i_{m/n}} \end{pmatrix} = \lceil m, \quad (7.0)$$

for $i_1, \dots, i_{\lceil(k+1)/2\rceil} \in \{1, \dots, k\}$ distinct. Reducing,

$$\text{rank} \begin{pmatrix} I_{n \times n} & B_{i_1} \\ 0_{n \times n} & B_{i_2} - B_{i_1} \\ \vdots & \vdots \\ 0_{n \times n} & B_{i_{m/n}} - B_{i_1} \end{pmatrix} = m.$$

Since $I_{n \times n}$ is of rank n , it suffices for

$$\text{rank} \begin{pmatrix} B_{i_2} - B_{i_1} \\ \vdots \\ B_{i_{m/n}} - B_{i_1} \end{pmatrix} = m - n, \quad (7.1)$$

for $i_1, \dots, i_{m/n} \in \{1, \dots, k\}$ distinct. Notice that while it is necessary to check (7.1) for every choice of m/n indices $i_1, \dots, i_{m/n}$, we do not need to check for permutations of the indices, any permutation suffices.

Example 7.1 (Triangles). We now prove the claim in Corollary 1.7 that if a, b, c are three distinct points in the plane, then any set $E \subset \mathbb{R}^2$ obeying the assumptions of Theorem 1.6 with ε_0 small enough (depending on C and on a, b, c) must contain a similar copy of the triangle $\triangle abc$. Note that our proof allows for degenerate triangles where a, b, c are colinear.

Let θ be the angle between the line segments \overline{ab} and \overline{ac} , measured counter-clockwise, and let $\lambda = \frac{|c-a|}{|b-a|}$. Permuting the points a, b, c if necessary, we may assume without loss of generality that $\theta \in (0, \pi]$. Then it suffices to prove that E contains a configuration of the form

$$x, \quad x + y, \quad x + \lambda y_\theta, \tag{7.2}$$

where y_θ is the vector y rotated by an angle θ counter-clockwise, for some $x, y \in \mathbb{R}^2$ with $y \neq 0$.

Fix $n = 2$, $k = 3$, and $m = 4$. Let $B_1 = 0_{2 \times 2}$. By (7.1), non-degeneracy means that

$$\text{rank}(B_j) = 2, \quad \text{rank}(B_3 - B_2) = 2,$$

for $j = 2, 3$. With $\theta \in (0, \pi]$ and $\lambda > 0$ as above, let

$$B_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B_3 = \begin{pmatrix} \lambda \cos \theta & -\lambda \sin \theta \\ \lambda \sin \theta & \lambda \cos \theta \end{pmatrix}.$$

It is easy to check that non-degeneracy holds, and this collection of matrices corresponds to configurations of the form (7.2). Letting $V = \{0\}$, Theorem 1.6 asserts that any set $E \subset \mathbb{R}^2$ obeying its assumptions with ε_0 small enough must contain such a configuration, non-degenerate in the sense that $y \neq 0$. This proves Corollary 1.7.

Example 7.2 (Colinear triples). We prove that if a, b, c are three distinct colinear points in \mathbb{R}^n , then any set $E \subset \mathbb{R}^n$ obeying the assumptions of Theorem 1.6 with ε_0 small enough (depending on C and on a, b, c) must contain a non-degenerate similar copy of $\{a, b, c\}$.

Without loss of generality, suppose $|c - a| > |b - a|$. Let $\lambda = \frac{|c-a|}{|b-a|} > 1$. Then it suffices to prove that E contains a configuration of the form

$$x, \quad x + y, \quad x + \lambda y, \tag{7.3}$$

for some $x, y \in \mathbb{R}^n$ with $y \neq 0$.

Fix a positive integer n , $k = 3$, and $m = 2n$. Let $B_1 = 0_{n \times n}$, $B_2 = I_{n \times n}$, $B_3 = \lambda I_{n \times n}$. Similarly to Example 7.1, this system of matrices produces configurations of the form (7.3), and the non-degeneracy condition (7.1) becomes

$$\text{rank}(B_j) = n, \quad \text{rank}(B_3 - B_2) = n,$$

for $j = 2, 3$, which is easy to check for B_j as above. Applying Theorem 1.6 with $V = \{0\}$ as before, we get the desired conclusion.

Example 7.3 (Parallelograms). We now prove Corollary 1.10. Fix $n \geq 1$, $k = 4$, and $m = 3n$. Let $B_1 = 0_{n \times 2n}$; (7.1) tells us non-degeneracy will be the condition

$$\text{rank} \begin{pmatrix} B_{i_1} \\ B_{i_2} \end{pmatrix} = 2n, \quad \text{rank} \begin{pmatrix} B_2 - B_4 \\ B_3 - B_4 \end{pmatrix} = 2n$$

for $i_1, i_2 \in \{2, 3, 4\}$ distinct.

Let

$$\begin{aligned} B_2 &= \begin{pmatrix} I_{n \times n} & 0_{n \times n} \end{pmatrix}, \\ B_3 &= \begin{pmatrix} 0_{n \times n} & I_{n \times n} \end{pmatrix}, \\ B_4 &= B_2 + B_3 = \begin{pmatrix} I_{n \times n} & I_{n \times n} \end{pmatrix}. \end{aligned}$$

Non-degeneracy clearly holds, and this collection of matrices corresponds to configurations of the form

$$x, \quad x + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad x + \begin{pmatrix} y_{n+1} \\ \vdots \\ y_{2n} \end{pmatrix}, \quad x + \begin{pmatrix} y_1 + y_{n+1} \\ \vdots \\ y_n + y_{2n} \end{pmatrix}, \quad (7.4)$$

for some $x \in \mathbb{R}^n$ and $y_1, \dots, y_{2n} \in \mathbb{R}$. Geometrically, such configurations describe 2-dimensional parallelograms. To exclude degenerate cases where the parallelogram becomes a line segment, we define the ‘‘exceptional’’ subspaces

$$\begin{aligned} V_1 &= \{y \in \mathbb{R}^{2n} : y_1 = \dots = y_n = 0\}, \\ V_2 &= \{y \in \mathbb{R}^{2n} : y_{n+1} = \dots = y_{2n} = 0\}, \\ V_3 &= \{y \in \mathbb{R}^{2n} : y_1 + y_{n+1} = 0, \dots, y_n + y_{2n} = 0\}, \\ V_4 &= \{y \in \mathbb{R}^{2n} : y_1 - y_{n+1} = 0, \dots, y_n - y_{2n} = 0\} \end{aligned}$$

Then Theorem 1.6 provides for the existence of parallelograms with y not in V_1, V_2, V_3, V_4 , so that the four points in (7.4) are all distinct.

Example 7.4 (Polynomial configurations). Finally, we prove Corollary 1.12. We will in fact prove a stronger statement, namely that the result in Corollary 1.12 holds in \mathbb{R}^n for all $n \geq 3$, with (1.7) replaced by 4-point configurations defined below in Corollary 7.1.

As in Example 7.3, fix $n \geq 1$, $k = 4$, and $m = 3n$, and let $B_1 = 0_{n \times 2n}$. We will use a Vandermonde-style matrix for the remaining B_i . To make the notation less cumbersome, for a function

$$\begin{aligned} g &: \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R} \\ &(i, j) \mapsto g(i, j), \end{aligned}$$

we denote by $(g(i, j))_{a \times b}$ the $a \times b$ matrix whose entry in the i th row and j th column is given by $g(i, j)$.

Corollary 7.1. *Let $a_1, \dots, a_{2n} > 1$ be distinct real numbers, and let $\eta, d \in \mathbb{N}$. Consider the following matrices:*

$$B_2 = (a_j^{\eta+(i-1)d})_{n \times 2n}, \quad B_3 = (a_j^{\eta+(n+i-1)d})_{n \times 2n}, \quad B_4 = (a_j^{\eta+(2n+i-1)d})_{n \times 2n}.$$

Suppose that $E \subset \mathbb{R}^n$ obeys the assumptions of Theorem 1.6, with ϵ_0 small enough depending on C and a_i . Then E contains a configuration of the form

$$x, \quad x + B_2 y, \quad x + B_3 y, \quad x + B_4 y \quad (7.5)$$

for some $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^{2n}$ with $B_i y \neq 0$ for $i = 2, 3, 4$.

The proof of Corollary 7.1 will rely on two short lemmas.

Lemma 7.2. *Suppose $0 \leq \eta_1 < \eta_2 < \dots < \eta_t$ are integers. Then for any choice of constants c_1, c_2, \dots, c_t that are not all zero, the polynomial*

$$P(x) = \sum_{i=1}^t c_i x^{\eta_i}$$

has fewer than t distinct positive roots.

Proof. We prove this with induction. For $t = 1$, it is clear that $c_1 x^{\eta_1}$ cannot have a positive root since $c_1 \neq 0$, so the base case is satisfied. We make the inductive hypothesis that the lemma holds for t , and check $t + 1$. Suppose to the contrary that there exist constants c_1, c_2, \dots, c_{t+1} , not all zero, such that the polynomial

$$P(x) = \sum_{i=1}^{t+1} c_i x^{\eta_i}$$

has at least $t + 1$ distinct positive roots. But then

$$x^{-\eta_1} P(x) = c_1 + c_2 x^{\eta_2 - \eta_1} + \dots + c_{t+1} x^{\eta_{t+1} - \eta_1},$$

so by Rolle's Theorem, the following polynomial has at least t distinct positive roots:

$$\begin{aligned} P_1(x) &:= \frac{d}{dx}(x^{-\eta_1} P(x)) \\ &= c_2(\eta_2 - \eta_1)x^{\eta_2 - \eta_1 - 1} + c_3(\eta_3 - \eta_1)x^{\eta_3 - \eta_1 - 1} + \dots + c_{t+1}(\eta_{t+1} - \eta_1)x^{\eta_{t+1} - \eta_1 - 1} \\ &= \sum_{i=1}^t c_{i+1}(\eta_{i+1} - \eta_1)x^{\eta_{i+1} - \eta_1 - 1} \end{aligned}$$

Since η_i were strictly increasing integers, $c_{i+1}(\eta_{i+1} - \eta_1)$ are not all zero, and $\eta_{i+1} - \eta_1 - 1 \geq 0$ are strictly increasing integers. This contradicts the induction hypothesis and completes the proof. \square

Lemma 7.3. *If*

$$A = (a_j^{\eta_i})_{t \times s}$$

where a_1, a_2, \dots, a_s are distinct, positive real numbers and $0 \leq \eta_1 < \eta_2 < \dots < \eta_t$ are integers, then A has full rank.

Proof. Without loss of generality, $t \leq s$. It suffices to show the following submatrix has full rank:

$$A_t = (a_j^{\eta_i})_{t \times t}.$$

This holds if and only if $\det A_t \neq 0$. If to the contrary $\det A_t = 0$, then we can find constants c_1, c_2, \dots, c_t that are not all zero such that $\sum_{i=1}^t c_i R_i = \vec{0}$, where R_i is the i th row of A_t ; considering the k th position this says $\sum_{i=1}^t c_i a_j^{\eta_i} = 0$ for $1 \leq j \leq t$. That is, the polynomial

$$P(x) = \sum_{i=1}^t c_i x^{\eta_i}$$

has at least the t distinct positive roots $x = a_j$ for $1 \leq j \leq t$. This contradicts Lemma 7.2, so we must have $\det A_t \neq 0$ and hence A has full rank. \square

Proof of Corollary 7.1. By Lemma 7.3,

$$\text{rank} \begin{pmatrix} B_{i_1} \\ B_{i_2} \end{pmatrix} = 2n$$

for $i_1, i_2 \in \{2, 3, 4\}$ distinct. It remains to check

$$\text{rank} \begin{pmatrix} B_2 - B_4 \\ B_3 - B_4 \end{pmatrix} = \text{rank} \begin{pmatrix} (a_j^{\eta+(i-1)d} - a_j^{\eta+(2n+i-1)d})_{n \times 2n} \\ (a_j^{\eta+(n+i-1)d} - a_j^{\eta+(2n+i-1)d})_{n \times 2n} \end{pmatrix} = 2n. \quad (7.6)$$

For constants c_1, \dots, c_{2n} , consider the polynomial

$$\begin{aligned} Q_{c_1, \dots, c_n}(x) &= c_1(x^\eta - x^{\eta+2nd}) + c_2(x^{\eta+d} - x^{\eta+(2n+1)d}) + \dots + c_n(x^{\eta+(n-1)d} - x^{\eta+(3n-1)d}) \\ &\quad + c_{n+1}(x^{\eta+nd} - x^{\eta+2nd}) + \dots + c_{2n}(x^{\eta+(2n-1)d} - x^{\eta+(3n-1)d}) \end{aligned} \quad (7.7)$$

If (7.6) fails to hold, then as in the proof of Lemma 7.3, there are constants c_1, \dots, c_{2n} not all 0 whose corresponding polynomial $Q(x) := Q_{c_1, \dots, c_n}(x)$ has at least the $2n$ distinct roots a_1, \dots, a_{2n} , all of which are larger than 1. We may simplify $Q(x)$ as

$$\begin{aligned} Q(x) &= c_1 x^\eta (1 - x^{2nd}) + c_2 x^{\eta+d} (1 - x^{2nd}) + \dots + c_n x^{\eta+(n-1)d} (1 - x^{2nd}) \\ &\quad + c_{n+1} x^{\eta+nd} (1 - x^{nd}) + \dots + c_{2n} x^{\eta+(2n-1)d} (1 - x^{nd}) \\ &= (1 - x^{nd}) [c_1 x^\eta (1 + x^{nd}) + c_2 x^{\eta+d} (1 + x^{nd}) + \dots + c_n x^{\eta+(n-1)d} (1 + x^{nd}) \\ &\quad + c_{n+1} x^{\eta+nd} + \dots + c_{2n} x^{\eta+(2n-1)d}] \\ &= (1 - x^{nd}) P(x), \end{aligned}$$

where

$$P(x) = c_1 x^\eta + c_2 x^{\eta+d} + \dots + c_n x^{\eta+(n-1)d} + (c_1 + c_{n+1}) x^{\eta+nd} + \dots + (c_n + c_{2n}) x^{\eta+(2n-1)d}.$$

The roots of $Q(x)$ which are larger than 1 coincide with the roots of $P(x)$. Notice that not all of the coefficients of $P(x)$ are 0, since not all of c_1, \dots, c_{2n} are 0. Then by Lemma 7.2, $P(x)$ has fewer than $2n$ positive roots, a contradiction. Thus, (7.6) holds and $\{A_1, \dots, A_k\}$ is non-degenerate. The result follows by applying Theorem 1.6, with $V_i = \{y \in \mathbb{R}^{2n} : B_i y = 0\}$ for $i = 2, 3, 4$. \square

A Approximate Identity

For a $p \times d$ ($p \leq d$) matrix P of full rank p , define

$$V = \{\xi \in \mathbb{R}^d : P\xi = 0\} = \mathcal{N}(P),$$

and $v = \dim(V) = d - p$.

Definition A.1. Fix an orthonormal basis $\{\vec{\alpha}_1, \dots, \vec{\alpha}_v\}$ of V . The surface measure $d\sigma$ on V is defined as follows:

$$\int_V F d\sigma = \int_{\mathbb{R}^v} F(x_1\vec{\alpha}_1 + \dots + x_v\vec{\alpha}_v) dx_1 \cdots dx_v$$

for every $F \in C_c(\mathbb{R}^d)$.

Note that this definition is independent of the choice of basis. Indeed, if $\{\vec{\beta}_1, \dots, \vec{\beta}_v\}$ is another orthonormal basis of V , then the mapping $(x_1, \dots, x_v) \mapsto (y_1, \dots, y_v)$ given by $\sum x_j\vec{\alpha}_j = \sum y_j\vec{\beta}_j$ is a linear isometry, hence given by an orthogonal matrix which has determinant 1. Then $d\vec{x} = d\vec{y}$ and hence

$$\int_{\mathbb{R}^v} F(x_1\vec{\alpha}_1 + \dots + x_v\vec{\alpha}_v) dx_1 \cdots dx_v = \int_{\mathbb{R}^v} F(x_1\vec{\alpha}_1 + \dots + x_v\vec{\alpha}_v) dx_1 \cdots dx_v.$$

Lemma A.2. For any $g \in C_c(\mathbb{R}^d)$ and any $\Psi \in \mathcal{S}(\mathbb{R}^p)$ with $\Psi(0) \neq 0$,

$$\lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} g(y_1, y_2) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) dy_1 dy_2 = \Psi(0) \int_{\mathbb{R}^{d-p}} g(y_1, 0) dy_1.$$

Here, $y = (y_1, y_2) \in \mathbb{R}^d$, with $y_1 \in \mathbb{R}^{d-p}$ and $y_2 \in \mathbb{R}^p$.

Proof. Fix $\kappa > 0$. Our goal is to show

$$\left| \int_{\mathbb{R}^d} g(y_1, y_2) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) dy_1 dy_2 - \Psi(0) \int_{\mathbb{R}^{d-p}} g(y_1, 0) dy_1 \right| < \kappa$$

for all $\varepsilon > 0$ sufficiently small. Then for every $\varepsilon > 0$, we have by definition of $\Psi(0)$

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} g(y_1, y_2) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) dy_1 dy_2 - \Psi(0) \int_{\mathbb{R}^{d-p}} g(y_1, 0) dy_1 \right| \\ &= \left| \int_{\mathbb{R}^d} g(y_1, y_2) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) dy_1 dy_2 - \int_{\mathbb{R}^p} \int_{\mathbb{R}^{d-p}} g(y_1, 0) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) dy_1 dy_2 \right| \\ &\leq \iint |g(y_1, y_2) - g(y_1, 0)| \frac{1}{\varepsilon^p} \left| \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) \right| dy_1 dy_2. \end{aligned}$$

We now partition our region of integration into where $|y_2|$ is small and where it is large. By uniform continuity of g on its compact support K , we may choose $\eta > 0$ sufficiently small so that

$$\sup_{y_1 \in K} |g(y_1, y_2) - g(y_1, 0)| < \frac{\kappa}{2|\Psi(0)|} \tag{A.1}$$

if $|y_2| \leq \eta$. On this region,

$$\begin{aligned} & \iint_{|y_2| \leq \eta} |g(y_1, y_2) - g(y_1, 0)| \frac{1}{\varepsilon^p} \left| \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) \right| dy_1 dy_2 \\ &< \iint \frac{\kappa}{2|\Psi(0)|} \frac{1}{\varepsilon^p} \left| \widehat{\Psi} \left(\frac{y_2}{\varepsilon} \right) \right| dy_1 dy_2 = \frac{\kappa}{2}. \end{aligned}$$

Since $\widehat{\Psi}$ is integrable, we can make the tail integral as small as we would like. In particular, there exists $\varepsilon > 0$ sufficiently small relative to η so that

$$\int_{|y_2| > \eta/\varepsilon} |\widehat{\Phi}(y_2)| dy_2 < \frac{\kappa}{4\|g\|_\infty \text{diam}(K)}. \quad (\text{A.2})$$

Then for this ε ,

$$\begin{aligned} & \iint_{|y_2| > \eta, y \in K} |g(y_1, y_2) - g(y_1, 0)| \frac{1}{\varepsilon^p} \left| \widehat{\Psi}\left(\frac{y_2}{\varepsilon}\right) \right| dy_1 dy_2 \\ & \leq 2\|g\|_\infty \text{diam}(K) \int_{|y_2| > \eta/\varepsilon} |\widehat{\Psi}(y_2)| dy_2 < \frac{\kappa}{2}. \end{aligned}$$

As this inequality holds for every $\kappa > 0$, the result follows. \square

Proposition A.3. *For any P as above, there exists a constant $C_P > 0$ with the property that for any $\Phi \in \mathcal{S}(\mathbb{R}^p)$ with $\Phi(0) = 1$, the limit*

$$\lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} F(\xi) \frac{1}{\varepsilon^p} \widehat{\Phi}\left(\frac{P\xi}{\varepsilon}\right) d\xi$$

exists and equals $C_P \int_V F d\sigma$.

Proof. Fix $\Phi \in \mathcal{S}(\mathbb{R}^p)$ with $\Phi(0) = 1$. Let $\{\vec{\alpha}_1, \dots, \vec{\alpha}_v\}$ be an orthonormal basis of V . Extend this to an orthonormal basis of \mathbb{R}^d , say $\{\vec{\alpha}_1, \dots, \vec{\alpha}_v, \vec{\alpha}_{v+1}, \dots, \vec{\alpha}_d\}$. Given any function $F \in C_c(\mathbb{R}^d)$, we define $G_F : \mathbb{R}^d \rightarrow \mathbb{R}$ as follows:

$$G(x_1, \dots, x_d) = F\left(\sum_{j=1}^d x_j \vec{\alpha}_j\right).$$

Notice $G_F \in C_c(\mathbb{R}^d)$ as well. Then by definition,

$$\begin{aligned} \int_V F d\sigma &= \int_{\mathbb{R}^v} F\left(\sum_{j=1}^v x_j \vec{\alpha}_j\right) dx_1 \cdots dx_v \\ &= \int_{\mathbb{R}^v} G_F(x_1, \dots, x_v, 0, \dots, 0) dx_1 \cdots dx_v \\ &= \frac{1}{\Psi(0)} \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} G_F(x_1, \dots, x_v, x_{v+1}, \dots, x_d) \frac{1}{\varepsilon^p} \widehat{\Psi}\left(\frac{x_{d+1}}{\varepsilon}, \dots, \frac{x_d}{\varepsilon}\right) d\vec{x} \end{aligned}$$

for any $\Psi \in \mathcal{S}(\mathbb{R}^d)$ with $\Psi(0) \neq 0$, by Lemma A.2. By the definition of G_F , this gives

$$\int_V F d\sigma = \frac{1}{\Psi(0)} \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} F(\xi) \frac{1}{\varepsilon^p} \widehat{\Psi}\left(\frac{x_{v+1}}{\varepsilon}, \dots, \frac{x_d}{\varepsilon}\right) d\vec{x}, \quad (\text{A.3})$$

for any $\Psi \in \mathcal{S}(\mathbb{R}^d)$ with $\Psi(0) \neq 0$, where we denote $\xi = \left(\sum_{j=1}^d x_j \vec{\alpha}_j\right)$. Now, let Q be the $p \times p$ matrix defined by

$$Q \begin{pmatrix} x_{v+1} \\ \vdots \\ x_d \end{pmatrix} = \sum_{j=v+1}^d x_j P \vec{\alpha}_j.$$

Since P is of full rank and acting on basis vectors, Q is non-singular. Recall $V = \{\xi : P\xi = 0\}$ and $\{\vec{\alpha}_1, \dots, \vec{\alpha}_v\}$ is a basis for V , so $P\vec{\alpha}_j = 0$ for $1 \leq j \leq v$. Then

$$\begin{pmatrix} x_{v+1} \\ \vdots \\ x_d \end{pmatrix} = Q^{-1} \sum_{j=v+1}^d x_j P\vec{\alpha}_j = Q^{-1} \sum_{j=1}^d x_j P\vec{\alpha}_j = Q^{-1} P\xi. \quad (\text{A.4})$$

Define $\Psi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\widehat{\Psi}(\xi) = \widehat{\Phi}(Q\xi).$$

Then by (A.3) and (A.4),

$$\begin{aligned} \int_V F \, d\sigma &= \frac{1}{\Psi(0)} \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} F(\xi) \frac{1}{\varepsilon^p} \widehat{\Psi} \left(\frac{Q^{-1}P\xi}{\varepsilon} \right) \, d\vec{x} \\ &= \frac{1}{\Psi(0)} \lim_{\varepsilon \rightarrow 0^+} \int_{\mathbb{R}^d} F(\xi) \frac{1}{\varepsilon^p} \widehat{\Phi} \left(\frac{P\xi}{\varepsilon} \right) \, d\vec{x}. \end{aligned}$$

Finally,

$$\begin{aligned} \Psi(0) &= \int_{\mathbb{R}^p} \widehat{\Psi}(\xi) \, d\xi \\ &= \int_{\mathbb{R}^p} \widehat{\Phi}(Q\xi) \, d\xi \\ &= \frac{1}{|Q|} \int_{\mathbb{R}^p} \widehat{\Phi}(\xi) \, d\xi \\ &= \frac{1}{|Q|} \Phi(0) = \frac{1}{|Q|} \end{aligned}$$

and so the result follows with $C_P = |Q|$. Note that Q is also independent of the choice of basis, and is a function only of P . \square

B Supplementary facts from §5

Lemma B.1. *Given $0 < \varepsilon < 1$ and any integer $K \geq 1$, there exists a positive constant $c'(\varepsilon, K)$ such that*

$$|\{t \in [0, 1] : \|tv_\ell\| \leq \varepsilon \text{ for all } 1 \leq \ell \leq K\}| \geq c'(\varepsilon, K),$$

for any choice of $v_1, \dots, v_K \in \mathbb{Z}$.

Proof. Clearly it suffices to prove the lemma for the case when $\varepsilon \leq 1$. Let N be the unique integer that $N^{-1} < \varepsilon \leq (N-1)^{-1}$, and consider the partition of the unit cube $[0, 1]^K$ into N^K disjoint cubes of side length N^{-1} . That is, the vertices of the cubes are at points of the form $N^{-1}\mathbb{Z}^K \bmod 1$. Define

$$X_Q = \{t \in [0, 1] : (tv_1, \dots, tv_K) \bmod 1 \in Q\}.$$

Then since $|[0, 1]| = 1$, there must exist a cube Q such that

$$|X_Q| \geq \frac{1}{N^K} \geq \left(\frac{\varepsilon}{2}\right)^K,$$

where we have used that $\varepsilon \leq (n-1)^{-1}$. For $t \in X_Q$,

$$(tv_1, \dots, tv_K) \bmod 1 \in Q - Q \subseteq [-n^{-1}, n^{-1}]^K \subseteq [-\varepsilon, \varepsilon]^K,$$

and so $\|tv_\ell\| \leq \varepsilon$ for every $1 \leq \ell \leq K$. Notice

$$|X_Q - X_Q| \geq |X_Q| \geq \left(\frac{\varepsilon}{2}\right)^K,$$

and so by symmetry

$$|\{t \in [0, 1] : \|tv_\ell\| \leq \varepsilon \text{ for all } 1 \leq \ell \leq K\}| \geq \frac{1}{2} \left(\frac{\varepsilon}{2}\right)^K.$$

□

Corollary B.2. *Given $0 < \varepsilon < 1$, and integers $k, K, m, n \in \mathbb{N}$, $m > n$, there exists a positive constant c depending on all of these quantities, for which the set*

$$C_\varepsilon = \{y \in \mathbb{R}^{m-n} : \|A_j^t v_\ell \cdot y\| \leq \varepsilon, \text{ for all } 1 \leq j \leq k, 1 \leq \ell \leq K\}.$$

defined as in Lemma 5.6 obeys the size estimate

$$|C_\varepsilon| \geq c$$

for any choice of matrices $\{A_j\}$ and vectors $\{v_\ell\}$.

Proof. Let $A_j^t v_\ell(i)$ denote the i th component of $A_j^t v_\ell$. Let

$$D_i = \{y_i \in [0, 1] : \|A_j^t v_\ell(i) y_i\| \leq \varepsilon / (m-n) \text{ for all } 1 \leq j \leq k, 1 \leq \ell \leq K\}.$$

By Lemma B.1,

$$|D_i| \geq c'(\varepsilon / (m-n), K)^k$$

for $1 \leq i \leq m-n$. If $y = (y_1, \dots, y_{m-n}) \in \prod_{i=1}^{m-n} D_i$, then

$$\|A_j^t v_\ell \cdot y\| = \left\| \sum_{i=1}^{m-n} A_j^t v_\ell(i) y_i \right\| \leq (m-n) \frac{\varepsilon}{m-n} = \varepsilon,$$

and so

$$C_\varepsilon \supseteq D_1 \times \dots \times D_{m-n}.$$

Therefore,

$$|C_\varepsilon| \geq c'(\varepsilon / (m-n), K)^{k(m-n)} = c(\varepsilon, K).$$

□

Lemma B.3. Suppose $\|f_j\|_\infty, \|g_j\|_\infty \leq C$ and $\|f_j - g_j\|_p \leq \kappa$ for $1 \leq j \leq R$, for some $1 \leq p \leq \infty$. Then

$$\left\| \prod_{j=1}^R f_j - \prod_{j=1}^R g_j \right\|_p \leq RC^{R-1}\kappa.$$

Proof.

$$\begin{aligned} \left\| \prod_{j=1}^R f_j - \prod_{j=1}^R g_j \right\|_p &\leq \left\| f_1 \cdot \prod_{j=2}^R f_j - g_1 \cdot \prod_{j=2}^R f_j \right\|_p \\ &\quad + \left\| g_1 \cdot f_2 \cdot \prod_{j=3}^R f_j - g_1 \cdot g_2 \cdot \prod_{j=3}^R f_j \right\|_p \\ &\quad + \cdots + \left\| \prod_{j=1}^{R-1} g_j \cdot f_R - \prod_{j=1}^{R-1} g_j \cdot g_R \right\|_p \\ &\leq \left(\prod_{j=2}^R \|f_j\|_\infty \right) \|f_1 - g_1\|_p \\ &\quad + \|g_1\|_\infty \left(\prod_{j=3}^R \|f_j\|_\infty \right) \|f_2 - g_2\|_p \\ &\quad + \cdots + \left(\prod_{j=1}^{R-1} \|g_j\|_\infty \right) \|f_R - g_R\|_p \\ &\leq RC^{R-1}\kappa. \end{aligned}$$

□

Lemma B.4. For a fixed $\varepsilon, \chi, \omega$, define $\mathcal{B}_{\varepsilon, \chi, \omega}$ as in Lemma 5.7. Then $\mathcal{B}_{\varepsilon, \chi, \omega}$ has at most $\pi 4\sqrt{2}/\varepsilon$ atoms.

Proof. Let N be the number of atoms of $\mathcal{B}_{\varepsilon, \chi, \omega}$. Recall that the atoms are of the form $\chi^{-1}(Q + \varepsilon\omega)$ for $Q \in \mathbb{Q}_\varepsilon$. Since the image of χ lies in the unit circle \mathbb{S}^1 , the preimage of $(Q + \varepsilon\omega)$ under χ is non-empty if and only if $(Q + \varepsilon\omega) \cap \mathbb{S}^1 \neq \emptyset$. Since $\text{diam}(Q + \varepsilon\omega) = \sqrt{2}\varepsilon$, this holds if and only if $(Q + \varepsilon\omega)$ lies in the $(\sqrt{2}\varepsilon)$ -thickened unit circle,

$$\mathbb{S}_{\sqrt{2}\varepsilon}^1 = \{x \in \mathbb{C} : \text{dist}(x, \mathbb{S}^1) \leq \sqrt{2}\varepsilon\}.$$

It is easy to calculate $|\mathbb{S}_{\sqrt{2}\varepsilon}^1| = \pi 4\sqrt{2}\varepsilon$ and $|Q + \varepsilon\omega| = \varepsilon^2$ for every $Q \in \mathbb{Q}_\varepsilon$, so since the squares $Q + \varepsilon\omega$ are disjoint,

$$N \leq \frac{|\mathbb{S}_{\sqrt{2}\varepsilon}^1|}{|Q + \varepsilon\omega|} = \frac{\pi 4\sqrt{2}}{\varepsilon}$$

as claimed. □

References

- [1] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U. S. A.*, 32:331–332, 1946. 3
- [2] C. Bluhm. Random recursive construction of Salem sets. *Ark. Mat.*, 34(1):51–63, 1996. 3
- [3] C. Bluhm. On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets. *Ark. Mat.*, 36(2):307–316, 1998. 3
- [4] J. Bourgain. A szemerédi type theorem for sets of positive density in \mathbb{R}^k . *Israel J. Math.*, 54:307–316, 1986. 6, 7
- [5] J. Bourgain. Construction of sets of positive measure not containing an affine image of a given infinite structure. *Israel J. Math.*, 60(3):333–344, 1987. 2
- [6] P. Erdős. Remarks on some problems in number theory. *Math. Balkanica*, 4:197–202, 1974. Papers presented at the Fifth Balkan Mathematical Congress (Belgrade, 1974). 2
- [7] B. Erdoan, D. Hart, and A. Iosevich. Multiparameter projection theorems with applications to sums-products and finite point configurations in the euclidean setting. In D. Bilyk, L. De Carli, A. Petukhov, A. M. Stokolos, and B. D. Wick, editors, *Recent Advances in Harmonic Analysis and Applications*, volume 25 of *Springer Proceedings in Mathematics & Statistics*, pages 93–103. Springer New York, 2013. 6
- [8] K. J. Falconer. On a problem of Erdős on sequences and measurable sets. *Proc. Amer. Math. Soc.*, 90(1):77–78, 1984. 2
- [9] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.*, 34:275–291 (1979), 1978. 2
- [10] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc. (3)*, 100(1):155–176, 2010. 3
- [11] L. Grafakos, A. Greenleaf, A. Iosevich, and E. Palsson. Multilinear generalized Radon transforms and point configurations. *ArXiv e-prints*, April 2012. 6
- [12] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008. 3
- [13] A. Greenleaf and A. Iosevich. On three point configurations determined by subsets of the Euclidean plane, the associated bilinear operator and applications to discrete geometry. *ArXiv e-prints*, September 2010. 6
- [14] A. Greenleaf, A. Iosevich, B. Liu, and E. Palsson. A group-theoretic viewpoint on Erdos-Falconer problems and the Mattila integral. *ArXiv e-prints*, June 2013. 6
- [15] V. Harangi, T. Keleti, G. Kiss, P. Maga, A. Máthé, P. Mattila, and B. Strenner. How large dimension guarantees a given angle? *ArXiv e-prints*, January 2011. 6

- [16] P. D. Humke and M. Laczkovich. A visit to the Erdős problem. *Proc. Amer. Math. Soc.*, 126(3):819–822, 1998. 2
- [17] J.-P. Kahane. Sur certains ensembles de Salem. *Acta Math. Acad. Sci. Hungar.*, 21:87–89, 1970. 3
- [18] J.-P. Kahane. *Some random series of functions*, volume 5 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 1985. 3
- [19] R. Kaufman. On the theorem of Jarník and Besicovitch. *Acta Arith.*, 39(3):265–267, 1981. 3
- [20] T. Keleti. A 1-dimensional subset of the reals that intersects each of its translates in at most a single point. *Real Anal. Exchange*, 24(2):843–844, 1998/99. 3
- [21] T. Keleti. Construction of one-dimensional subsets of the reals not containing similar copies of given patterns. *Anal. PDE*, 1(1):29–33, 2008. 3
- [22] Y. Kohayakawa, T. Łuczak, and V. Rödl. Arithmetic progressions of length three in subsets of a random set. *Acta Arith.*, 75(2):133–163, 1996. 3
- [23] M. N. Kolountzakis. Infinite patterns that can be avoided by measure. *Bull. London Math. Soc.*, 29(4):415–424, 1997. 2
- [24] P. Komjáth. Large sets not containing images of a given sequence. *Canad. Math. Bull.*, 26(1):41–43, 1983. 2
- [25] I. Łaba and M. Pramanik. Arithmetic progressions in sets of fractional dimension. *Geom. Funct. Anal.*, 19(2):429–456, 2009. 1, 3, 4, 5, 8, 9
- [26] P. Maga. Full dimensional sets without given patterns. *Real Anal. Exchange*, 36:79–90, 2010. 3, 6, 7
- [27] A. Magyar. k -point configurations in sets of positive density of \mathbb{Z}^n . *Duke Math. J.*, 146:1–34, 2009. 7
- [28] A. Máthé. Sets of large dimension not containing polynomial configurations. *ArXiv e-prints*, January 2012. 6
- [29] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953. 3
- [30] R. Salem. On singular monotonic functions whose spectrum has a given Hausdorff dimension. *Ark. Mat.*, 1:353–365, 1951. 3
- [31] R. Salem and D. C. Spencer. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U. S. A.*, 28:561–563, 1942. 3
- [32] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975. Collection of articles in memory of Juriĭ Vladimirovič Linnik. 2

- [33] T. Tao. Arithmetic progressions and the primes. *Collect. Math.*, (Vol. Extra):37–88, 2006. 9, 23, 25, 30, 31

Vincent Chan
University of British Columbia
1984 Mathematics Road
Vancouver BC, Canada V6T 1Z2
Email: vchan@math.ubc.ca

Izabella Łaba
University of British Columbia
1984 Mathematics Road
Vancouver BC, Canada V6T 1Z2
Email: ilaba@math.ubc.ca

Malabika Pramanik
University of British Columbia, Vancouver
1984 Mathematics Road
Vancouver BC, Canada V6T 1Z2
Email: malabika@math.ubc.ca