ON POLYNOMIAL CONFIGURATIONS IN FRACTAL SETS

K. HENRIOT, I. ŁABA, M. PRAMANIK

ABSTRACT. We show that subsets of $\mathbb{R}^n$ of large enough Hausdorff and Fourier dimension contain polynomial patterns of the form

$$(x,\, x + A_1 y,\, \dots,\, x + A_{k-1} y,\, x + A_k y + Q(y) e_n), \quad x \in \mathbb{R}^n,\ y \in \mathbb{R}^m,$$

where A_i are real $n \times m$ matrices, Q is a real polynomial in m variables and $e_n = (0, \dots, 0, 1)$.

1. Introduction

In this work we investigate the presence of point configurations in subsets of $\mathbb{R}^n$ which are large in a certain sense. When E is a subset of $\mathbb{R}^n$ of positive Lebesgue measure, a consequence of the Lebesgue density theorem is that E contains a similar copy of any finite set. A more difficult result of Bourgain [9] states that sets of positive upper density in $\mathbb{R}^n$ contain, up to isometry, all large enough dilates of the set of vertices of any fixed non-degenerate $(n-1)$-dimensional simplex. In a different setting, Roth's theorem [37] in additive combinatorics states that subsets of $\mathbb{Z}$ of positive upper density contain non-trivial three-term arithmetic progressions.

When a subset $E \subset \mathbb{R}$ is only supposed to have a positive Hausdorff dimension, a direct analogue of Roth's theorem is impossible. Indeed Keleti [26] has constructed a set of full dimension in $[0,1]$ not containing the vertices of any non-degenerate parallelogram, and in particular not containing any non-trivial three-term arithmetic progression. Maga [30] has since extended this construction to dimensions $n \geqslant 2$. The work of Łaba and Pramanik [28] and its multidimensional extension by Chan et al. [10] circumvent these obstructions under additional assumptions on the set E, which we now describe.

When E is a compact subset of $\mathbb{R}^n$, Frostman's lemma [46, Chapter 8] essentially states that its Hausdorff dimension is equal to

$$\dim_{\mathcal{H}} E = \sup\{\alpha \in [0,n) \,:\, \exists\, \mu \in \mathcal{M}(E) \,:\, \sup_{x \in \mathbb{R}^n,\, r > 0} \mu\big[B(x,r)\big] r^{-\alpha} < \infty\},$$

where $\mathcal{M}(E)$ is the space of probability measures supported on E. On the other hand, the Fourier dimension of E is

$$\dim_{\mathcal{F}} E = \sup\{\beta \in [0,n) \,:\, \exists\, \mu \in \mathcal{M}(E) \,:\, \sup_{\xi \in \mathbb{R}^n} |\widehat{\mu}(\xi)| (1 + |\xi|)^{-\beta/2} < \infty\}.$$

It is well-known that we have $\dim_{\mathcal{F}}(E) \leqslant \dim_{\mathcal{H}}(E)$ for every compact set E, with strict inequality in many instances, and we call E a Salem set when equality holds. There are various known constructions of Salem sets [7,8,22,24,25,28,39], several of which [11,27] also produce sets with prescribed Hausdorff and Fourier dimensions $0 < \beta \leqslant \alpha < n$.

In a very abstract setting, one may ask whether it is possible to find translation-invariant patterns of the form

$$\Phi(x, y) = (x, x + \varphi_1(y), \dots, x + \varphi_k(y)) \tag{1.1}$$

in the product set $E \times \cdots \times E$, where the $\varphi_j : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$ are certain shift functions. When $n + m \geqslant (k+1)n$, the map Φ considered is often a submersion of an open subset of $\mathbb{R}^{n+m}$ onto $\mathbb{R}^{(k+1)n}$, and then one can find a pattern of the desired kind in E via the implicit function theorem. A natural restriction is therefore to assume that $m < kn$ in this multidimensional setting. Chan et al. [10] studied the case where the maps $\varphi_j(y) = A_j y$ are linear for matrices $A_j \in \mathbb{R}^{n \times m}$, generalizing the study of Łaba and Pramanik for three-term arithmetic progressions, under the following technical assumption.

Definition 1.1. *Let $n, k, m \geqslant 1$ and suppose that $m = (k-r)n + n'$ with $1 \leqslant r < k$ and $0 \leqslant n' < n$. We say that the system of matrices $A_1, \dots, A_k \in \mathbb{R}^{n \times m}$ is non-degenerate when*

$$\operatorname{rk} \begin{bmatrix} A_{j_1}^\intercal & \dots & A_{j_{k-r+1}}^\intercal \\ I_{n \times n} & \dots & I_{n \times n} \end{bmatrix} = (k - r + 1)n,$$

for every set of indices $\{j_1, \dots, j_{k-r+1}\} \subset \{0, \dots, k\}$, with the convention that $A_0 = 0_{n \times n}$.

Requirements similar to the above arise when analysing linear patterns by ordinary Fourier analysis in additive combinatorics [38], and there is a close link with the modern definition of linear systems of complexity one [16]. The main result of Chan et al. [10] gives a fractal analogue of the multidimensional Szemerédi theorem [15] for non-degenerate linear systems, when the Frostman measure has both dimensional and Fourier decay. We only state it in the case where n divides m for simplicity.

Theorem 1.2 (Chan, Łaba and Pramanik)**.** *Let $n, k, m \geqslant 1$, $D \geqslant 1$ and $\alpha, \beta \in (0, n)$. Suppose that E is a compact subset of $\mathbb{R}^n$ and μ is a probability measure supported on E such that*[1]

$$\mu\big[B(x, r)\big] \leqslant D r^\alpha \quad \text{and} \quad |\widehat{\mu}(\xi)| \leqslant D(n - \alpha)^{-D}(1 + |\xi|)^{-\beta/2}$$

[1]In fact, this theorem was proved in [10] under the more restrictive condition $|\widehat{\mu}(\xi)| \leqslant D(1 + |\xi|)^{-\beta/2}$ for a fixed constant D. However, by examining the proof there, one can see that the constant $D = D_\alpha$ may be allowed to grow polynomially in $n - \alpha$, as was the case in the original argument of Łaba and Pramanik [28].

for all $x \in \mathbb{R}^n$, $r > 0$ and $\xi \in \mathbb{R}^n$. Suppose that $(A_1, \dots, A_k)$ is a non-degenerate system of $n \times m$ matrices in the sense of Definition 1.1. Assume finally that $m = (k-r)n$ with $1 \leqslant r < k$ and, for some $\varepsilon \in (0,1)$,

$$\left\lceil \frac{k}{2} \right\rceil n \leqslant m < kn, \qquad \frac{2(kn-m)}{k+1} + \varepsilon \leqslant \beta < n, \qquad n - c_{n,k,m,\varepsilon,D,(A_i)} \leqslant \alpha < n,$$

for a sufficiently small constant $c_{n,k,m,\varepsilon,D,(A_i)} > 0$. Then, for every collection of strict subspaces $V_1, \dots, V_q$ of $\mathbb{R}^{n+m}$, there exists $(x,y) \in \mathbb{R}^{n+m} \smallsetminus V_1 \cup \dots V_q$ such that

$$(x, x + A_1 y, \dots, x + A_k y) \in E^{k+1}.$$

Note that the Hausdorff dimension α is required to be large enough with respect to the constants involved in the dimensional and Fourier decay bounds for the Frostman measure. A construction due to Shmerkin [41] shows that the dependence of α on the constants cannot be removed.

In practice, Salem set constructions provide a family of fractal sets indexed by α, and it is often possible to verify the conditions of Theorem 1.2 for α close to n ; this was done in a number of cases in [28]. The requirement of Fourier decay of the measure μ serves as an analogue of the notion of pseudorandomness in additive combinatorics [44], under which we expect a set to contain the same density of patterns as a random set of same size.

In this work we consider a class of polynomial patterns, which generalizes that of Theorem 1.2. We aim to obtain results similar in spirit to the Furstenberg-Sarközy theorem [14, 40] in additive combinatorics, which finds patterns of the form $(x, x+y^2)$ in dense subsets of $\mathbb{Z}$. A deep generalization of this result is the multidimensional polynomial Szemerédi theorem in ergodic theory [4, 6] (see also [5, Section 6.3]), which handles patterns of the form (1.1) where each shift function φ_j is an integer polynomial vector with zero constant term. By contrast, the class of patterns we study includes only one polynomial term, which should satisfy certain non-degeneracy conditions. We are also forced to work with a dimension $n \geqslant 2$, and all these limitations are due to the inherent difficulty in analyzing polynomial patterns through Fourier analysis. On the other hand, we are able to relax the Fourier decay condition on the fractal measure needed in Theorem 1.2.

Theorem 1.3. *Let $n, m, k \geqslant 2$, $D \geqslant 1$ and $\alpha, \beta \in (0, n)$. Suppose that E is a compact subset of $\mathbb{R}^n$ and μ is a probability measure supported on E such that*

$$\mu\big[B(x,r)\big] \leqslant D r^{\alpha} \quad \text{and} \quad |\widehat{\mu}(\xi)| \leqslant D(n-\alpha)^{-D}(1+|\xi|)^{-\beta/2}$$

for all $x \in \mathbb{R}^n$, $r > 0$ and $\xi \in \mathbb{R}^n$. Suppose that $(A_1, \dots, A_k)$ is a non-degenerate system of real $n \times m$ matrices in the sense of Definition 1.1. Let Q be a real polynomial in m

variables such that $Q(0) = 0$ *and the Hessian of* Q *does not vanish at zero. Assume furthermore that, for a constant* $\beta_0 \in (0, n)$,

$$(k-1)n < m < kn, \qquad \beta_0 \leqslant \beta < n, \qquad n - c_{\beta_0,n,k,m,D,(A_i),Q} < \alpha < n,$$

for a sufficiently small constant $c_{\beta_0,n,k,m,D,(A_i),Q} > 0$. *Then, for every collection* $V_1, \dots, V_q$ *of strict subspaces of* $\mathbb{R}^{m+n}$, *there exists* $(x, y) \in \mathbb{R}^{n+m} \smallsetminus (V_1 \cup \dots \cup V_q)$ *such that*

$$(x,\, x + A_1 y, \dots,\, x + A_{k-1}y,\, x + A_k y + Q(y)e_n) \in E^{k+1}, \tag{1.2}$$

where $e_n = (0, \dots, 0, 1)$.

Our argument follows broadly the transference strategy devised by Łaba and Pramanik [28], and its extension by Chan and these two authors [10]. However, the case of polynomial configurations requires a more delicate treatment of the singular integrals arising in the analysis. The weaker condition on β is obtained by exploiting restriction estimates for fractal measures due to Mitsis [33] and Mockenhaupt [34]. A more detailed outline of our strategy can be found in Section 3. By the method of this paper, one can also obtain an analogue of Theorem 1.2 with the same relaxed condition on the exponent β, and we state this version precisely in Section 9.

For concreteness' sake, we highlight the lowest dimensional situation handled by Theorem 1.3. When $k = n = 2$ and $m = 3$, this theorem allows us to detect patterns of the form

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + A_1 \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + A_2 \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 \\ Q(y_1, y_2, y_3) \end{bmatrix},$$

for matrices $A_1, A_2 \in \mathbb{R}^{2\times 3}$ of full rank such that $A_1 - A_2$ has full rank, and for a non-degenerate quadratic form Q in three variables. We may additionally impose that $y_1, y_2, y_3 \in \mathbb{R} \smallsetminus \{0\}$ by setting $V_i = \{(x, y) \in \mathbb{R}^5 : y_i = 0\}$ in Theorem 1.3. For example, when $A_1 = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right]$, $A_2 = \left[\begin{smallmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{smallmatrix}\right]$ and $Q(y) = |y|^2$, we can detect the configuration

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}, \begin{bmatrix} x_1 + y_3 \\ x_2 + y_1 + y_1^2 + y_2^2 + y_3^2 \end{bmatrix}$$

with $y_1, y_2, y_3 \in \mathbb{R} \smallsetminus \{0\}$. However, we cannot detect the configuration

$$(x, x + y, x + y^2), \quad x \in \mathbb{R},\, y \in \mathbb{R} \smallsetminus \{0\},$$

for then we have $n = m = 1$ and $k = 2$, and the condition $m > (k-1)n$ is not satisfied.

Note also that, in the statement of Theorem 1.3, one may add a linear term in variables $y_1, \dots, y_m$ to the polynomial Q without affecting the assumptions on it. This allows for some flexibility in satisfying the matrix non-degeneracy conditions of Definition 1.1, since one may alter the last line of A_k at will. For example, the degenerate system of matrices $\left[\begin{smallmatrix}1&0&0\\0&1&0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0&0&1\\0&0&0\end{smallmatrix}\right]$ and the polynomial $Q(y) = |y|^2$ give rise to the configuration

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}, \quad \begin{bmatrix} x_1 + y_3 \\ x_2 + |y|^2 \end{bmatrix}.$$

Rewriting $|y|^2 = y_1 + y_2 + y_3 + Q_1(y)$, we see that Q_1 still has non-degenerate Hessian at zero and the configuration is now associated to the system of matrices $\left[\begin{smallmatrix}1&0&0\\0&1&0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0&0&1\\1&1&1\end{smallmatrix}\right]$, which is easily seen to be non-degenerate. One possible explanation for this curious phenomenon is that, by comparison with the setting of Theorem 1.2, we have an extra variable at our disposition, since $m > (k-1)n \geqslant \lceil k/2 \rceil n$.

Finally, we note that there is a large body of literature on configurations in fractal sets where Fourier decay assumptions are not required. Here, the focus is often on finding a large variety (in a specified quantitative sense) of certain types of configurations. A well-known conjecture of Falconer [46, Chapter 9] states that when a compact subset E of $\mathbb{R}^n$ has Hausdorff dimension at least $n/2$, its set of distances $\Delta(E) = \{|x-y|,\ x, y \in E\}$ must have positive Lebesgue measure. This can be phrased in terms of E containing configurations $\{x, y\}$ with $|x-y| = d$ for all $d \in \Delta(E)$, where $\Delta(E)$ is "large." Wolff [45] and Erdoğan [12, 13] proved that the distance set $\Delta(E)$ has positive Lebesgue measure for $\dim_{\mathcal{H}} E > \frac{n}{2} + \frac{1}{3}$, and Mattila and Sjölin [32] showed that it contains an open interval for $\dim_{\mathcal{H}} E > \frac{n+1}{2}$. More recently, Orponen [36] proved using very different methods that $\Delta(E)$ has upper box dimension 1 if E is s-Ahlfors-David regular with $s \geqslant 1$. There is a rich literature generalizing these results to other classes of configurations, such as triangles [19], simplices [17, 20], or sequences of vectors with prescribed consecutive lengths [3, 21].

In a sense, the configurations studied in these references enjoy a greater degree of directional freedom, which ensures that they are not avoided by sets of full Hausdorff dimension. By contrast, a Fourier decay assumption is necessary to locate 3-term progressions in a fractal set of full Hausdoff dimension (as mentioned earlier), and in light of recent work of Mathé [31], it is likely that a similar assumption is needed to find polynomial patterns of the form (1.2). It is, however, possible that our non-degeneracy assumptions are not optimal, or that special cases of our results could be proved without Fourier decay assumptions[2]. Loosely speaking, we would expect that configurations with more degrees of freedom are less likely to require Fourier conditions, but the specifics are far from understood and we do not feel that we have sufficient data to attempt to make a conjecture in this direction.

[2]After this article was first submitted for publication, a result of this type was indeed proved by Iosevich and Liu [23].

Acknowledgements. This work was supported by NSERC Discovery grants 22R80520 and 22R82900.

2. Notation

We define the following standard spaces of complex-valued functions and measures:

$$\begin{aligned}\mathcal{C}(\mathbb{R}^d) &= \{\text{continuous functions on } \mathbb{R}^d\},\\ \mathcal{S}(\mathbb{R}^d) &= \{\text{Schwartz functions on } \mathbb{R}^d\},\\ \mathcal{C}_c^\infty(\mathbb{R}^d) &= \{\text{smooth compactly-supported functions on } \mathbb{R}^d\},\\ \mathcal{C}_{c,+}^\infty(\mathbb{R}^d) &= \{\text{non-negative smooth compactly-supported functions on } \mathbb{R}^d\},\\ \mathcal{M}^+(\mathbb{R}^d) &= \{\text{finite non-negative Borelian measures on } \mathbb{R}^d\}.\end{aligned}$$

Similar notation is employed for functions on $\mathbb{T}^d$. We write $e(x) = e^{2i\pi x}$ for $x \in \mathbb{R}$. We let $\mathcal{L}$ denote either the Lebesgue measure on $\mathbb{R}^d$ or the normalized Haar measure on $\mathbb{T}^d$. We let $\mathrm{d}\sigma$ denote generically the Euclidean surface measure on a submanifold of $\mathbb{R}^d$. When f is a function on an abelian group G and t is an element of G, we denote the t-shift of f by $T^t f(x) = f(x+t)$. When A is a matrix we denote its transpose by A^T. We also write $[n] = \{1,\dots,n\}$ for an integer n and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

3. Broad scheme

In this section we introduce the basic objects that we will work with in this paper. We also state the intermediate propositions corresponding to the main steps of our argument, and we derive Theorem 1.3 from them at the outset.

We fix a compact set $E \subset \mathbb{R}^n$ and a probability measure μ supported on E. For technical reasons, we suppose that $E \subset \left[-\frac{1}{16}, \frac{1}{16}\right]^n$. We fix two exponents $0 < \beta \leqslant \alpha < n$, as well as two constants $D, D_\alpha \geqslant 1$, where the subscript in the second constant indicates that it is allowed to vary with α. We assume that the measure μ verifies the following dimensional and Fourier decay conditions:

$$\mu\big[B(x,r)\big] \leqslant D r^\alpha \qquad (x \in \mathbb{R}^n, r > 0), \tag{3.1}$$

$$|\widehat{\mu}(\xi)| \leqslant D_\alpha (1+|\xi|)^{-\beta/2} \qquad (\xi \in \mathbb{R}^n). \tag{3.2}$$

We suppose that the second constant involved blows up (if at all) at most polynomially as α tends to n:

$$D_\alpha \lesssim (n-\alpha)^{-O(1)}. \tag{3.3}$$

We also let $k \geqslant 3$ and we consider smooth functions $\varphi_1, \ldots, \varphi_k : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$, where Ω is an open neighborhood of zero. We are interested in locating the pattern

$$\Phi(x, y) = (x, x + \varphi_1(y), \ldots, x + \varphi_k(y)) \tag{3.4}$$

in E^{k+1}. While this abstract notation is sometimes useful, in practice we work with the maps

$$(\varphi_1(y), \ldots, \varphi_k(y)) = (A_1 y, \ldots,\ A_{k-1} y,\ A_k y + Q(y) e_n), \tag{3.5}$$

where $(A_1, \ldots, A_k)$ is a non-degenerate system of $n \times m$ matrices in the sense of Definition 1.1 and $Q \in \mathbb{R}[y_1, \ldots, y_m]$ is such that $Q(0) = 0$ and the Hessian of Q does not vanish at zero. We also fix a smooth cutoff $\psi \in \mathcal{C}^\infty_{c,+}(\mathbb{R}^m)$ supported on Ω such that $\psi \geqslant 1$ on a small box $[-c, c]^m$ and the Hessian of Q is bounded away from zero on the support of ψ. This cutoff is used in Definition 3.2 below. We take the opportunity here to state an equivalent form of Definition 1.1 when $m \geqslant (k-1)n$.

Definition 3.1. *If $m \geqslant (k-1)n$, we say that the system of matrices $(A_i)_{1 \leqslant i \leqslant k}$ with $A_i \in \mathbb{R}^{n \times m}$ is non-degenerate when, for every $1 \leqslant j \leqslant k$ and writing $[k] = \{i_1, \ldots, i_{k-1}, j\}$, the matrices*

$$[A_1^\intercal \ \ldots \ \widehat{A_j^\intercal} \ \ldots \ A_k^\intercal], \qquad [(A_{i_1}^\intercal - A_j^\intercal) \ \ldots \ (A_{i_{k-1}}^\intercal - A_j^\intercal)]$$

(where the hat indicates omission) have rank $(k-1)n$.

We also state a few notational conventions applied throughout the article. When $(A_1, \ldots, A_k)$ is a system of $n \times m$ matrices, we define the $kn \times m$ matrix $\mathbf{A}$ by $\mathbf{A}^\intercal = [A_1^\intercal \ \ldots \ A_k^\intercal]$. Unless mentioned otherwise, we allow every implicit or explicit constant in the article to depend on the integers n, k, m, the constant D, the matrices A_i and the polynomial Q, and the cutoff function ψ. This convention is already in effect in the propositions stated later in this section.

We start by defining a multilinear form which plays a central role in our argument.

Definition 3.2 (Configuration form)**.** *For functions $f_0, \ldots, f_k \in \mathcal{S}(\mathbb{R}^n)$, we let*

$$\Lambda(f_0, \ldots, f_k) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f_0(x) f_1(x + \varphi_1(y)) \cdots f_k(x + \varphi_k(y)) \mathrm{d}x \, \psi(y) \mathrm{d}y.$$

In Section 4, we show that the multilinear form has the following convenient Fourier expression.

Proposition 3.3. *For measurable functions $F_0, \ldots, F_k$ on $\mathbb{R}^n$ and K on $\mathbb{R}^{nk}$, we let*

$$\Lambda^*(F_0, \ldots, F_k; K) = \int_{(\mathbb{R}^n)^k} F_0(-\xi_1 - \cdots - \xi_k) F_1(\xi_1) \cdots F_k(\xi_k) K(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}, \tag{3.6}$$

whenever the integral is absolutely convergent or the integrand is non-negative. For every $f_0, \dots, f_k \in \mathcal{S}(\mathbb{R}^n)$*, we have*

$$\Lambda(f_0, \dots, f_k) = \Lambda^*(\widehat{f_0}, \dots, \widehat{f_k}; J),$$

where J is the oscillatory integral of Definition 4.1.

We may extend the configuration operator to measures, whenever we have absolute convergence of the dual form.

Definition 3.4 (Configuration form for measures)**.** *When* $\lambda_0, \dots, \lambda_k \in \mathcal{M}^+(\mathbb{R}^n)$ *are such that* $\Lambda^*(\,|\widehat{\lambda}_0|, \dots, |\widehat{\lambda}_k|; |J|\,) < \infty$*, we define*

$$\Lambda(\lambda_0, \dots, \lambda_k) = \Lambda^*(\widehat{\lambda}_0, \dots, \widehat{\lambda}_k; J).$$

When $\lambda_j \in \mathcal{S}(\mathbb{R}^n)$*, this is compatible with Definition 3.2 by Proposition 3.3.*

The next step, carried out in Section 5, is to obtain bounds for the dual multilinear form evaluated at the Fourier-Stieljes transform of the fractal measure μ. Such bounds hold only in certain ranges of α, β and under certain restrictions on n, k, m.

Proposition 3.5. *Let* $\beta_0 \in (0, n)$ *and suppose that for a constant* $c > 0$ *small enough with respect to* n, k, m*,*

$$(k-1)n < m < kn, \quad \beta_0 \leqslant \beta < n, \quad n - c\beta_0 \leqslant \alpha < n. \tag{3.7}$$

Then

$$\Lambda^*(\,|\widehat{\mu}|, \dots, |\widehat{\mu}|; |J|\,) \lesssim_{\beta_0} (n-\alpha)^{-O(1)}. \tag{3.8}$$

Recalling Definition 3.4, we see that $\Lambda(\mu, \dots, \mu)$ is well-defined under the conditions (3.7). In practice, we will need slight variants of Proposition 3.5, which are discussed in Section 5. In the same section, we obtain singular integral bounds for bounded functions of compact support.

Proposition 3.6. *Suppose that* $m > (k-1)n$*. Then there exists* $\varepsilon \in (0,1)$ *depending at most on* n, k, m *such that the following holds. For functions* $f_0, \dots, f_k \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ *with support in* $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$*,*

$$|\Lambda(f_0, \dots, f_k)| \lesssim \prod_{0 \leqslant j \leqslant k} \|\widehat{f_j}\|_\infty^\varepsilon \cdot \|f_j\|_\infty^{1-\varepsilon}.$$

In Section 6, we construct a measure detecting polynomial configurations, by exploiting the finiteness of the singular integral in (3.8) and the uniform decay of the fractal measure.

Proposition 3.7. *Let* $\beta_0 \in (0, n)$ *and suppose that* (3.7) *holds. Then there exists a measure* $\nu \in \mathcal{M}^+(\mathbb{R}^{n+m})$ *such that*

- $\|\nu\| = \Lambda(\mu, \dots, \mu)$,
- *ν is supported on the set of $(x, y) \in \mathbb{R}^n \times \Omega$ such that:* $(x, x + \varphi_1(y), \dots, x + \varphi_k(y)) \in E^{k+1}$,
- *$\nu(H) = 0$ for every hyperplane $H < \mathbb{R}^{n+m}$.*

In Section 7, we show how to obtain a positive mass of polynomial configurations in sets of positive density, through the singular integral bound of Proposition 3.6 and the arithmetic regularity lemma from additive combinatorics.

Proposition 3.8. *Suppose that $m > (k-1)n$. Then, uniformly for every function $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that* $\operatorname{Supp} f \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^n$, $0 \leqslant f \leqslant 1$ *and* $\int f = \tau \in (0, 1]$, *we have*

$$\Lambda(f, \dots, f) \gtrsim_\tau 1.$$

In Section 8, we show how to obtain a positive mass of configurations by a transference argument, by which the fractal measure μ is replaced by a mollified version of itself which is absolutely continuous with bounded density, allowing us to invoke Proposition 3.8.

Proposition 3.9. *Let $\beta_0 \in (0, n)$ and suppose that*

$$(k-1)n < m < kn, \quad \beta_0 \leqslant \beta < n, \quad n - c(\beta_0) \leqslant \alpha < n,$$

for a sufficiently small constant $c(\beta_0) > 0$. Then

$$\Lambda(\mu, \dots, \mu) > 0.$$

At this stage we have stated all the necessary ingredients to prove the main theorem.

Proof of Theorem 1.3. We may assume that $E \subset \left[-\frac{1}{16}, \frac{1}{16}\right]^n$ after a translation and dilation, which does not affect the assumptions on $\mu, (A_i), Q$ except for the introduction of constant factors in bounds. By Proposition 3.7 , there exists a measure $\nu \in \mathcal{M}^+(\mathbb{R}^{n+m})$ with mass $\Lambda(\mu, \dots, \mu)$ supported on

$$X = \{(x, y) \in \mathbb{R}^n \times \Omega \,:\, (x, x + A_1 y, \dots, x + A_{k-1} y, x + A_k y + Q(y) e_n) \in E^{k+1}\},$$

and such that $\nu(V_i) = 0$ for every collection of hyperplanes $V_1, \dots, V_q$ of $\mathbb{R}^{n+m}$. We have therefore proven the result if we can show that $\|\nu\| = \Lambda(\mu, \dots, \mu) > 0$, for then $\nu(X \smallsetminus (V_1 \cup \dots \cup V_q)) > 0$ and the set $X \smallsetminus (V_1 \cup \dots \cup V_q)$ cannot be empty. We may apply Proposition 3.9 to obtain precisely this conclusion when α is close enough to n with respect to β_0 (and the other implicit parameters $n, k, m, D, \mathbf{A}, Q$). $\square$

To conclude this outline, we comment briefly on the role that the Fourier decay hypothesis plays in our argument. Using the restriction theory of fractals, the assumption (3.2) is used together with the ball condition (3.1) in Appendix B to deduce that $\|\widehat{\mu}\|_{2+\varepsilon} < \infty$ for an arbitrary $\varepsilon > 0$, provided that α is close enough to n (depending on ε). The Hausdorff dimension condition (3.1) alone does yield information on the average Fourier decay growth of μ, via the energy formula [46, Chapter 8], but this type of estimate seems to be insufficient to establish the boundedness of the singular integrals we encounter. Section 5 on singular integral bounds and Section 7 on absolutely continuous estimates only use the Fourier moment bound above. On the other hand, the estimation of degenerate configurations in Section 6 and the transference argument of Section 8 exploit in an essential way the assumption of uniform Fourier decay.

4. Counting operators and Fourier expressions

In this section we describe the various types of pattern-counting operators and singular integrals that arise in trying to detect translation-invariant patterns in the fractal set of the introduction. First, we define an oscillatory integral which arises naturally in the Fourier expression of the configuration form in Definition 3.2.

Definition 4.1 (Oscillatory integral)**.** *For $\boldsymbol{\xi} \in (\mathbb{R}^n)^k$ and $\theta \in \mathbb{R}^m$ we define*

$$J_\theta(\boldsymbol{\xi}) = \int_{\mathbb{R}^m} e\big[(\theta + \mathbf{A}^\intercal \boldsymbol{\xi}) \cdot y + \xi_{kn} Q(y)\big] \psi(y) \mathrm{d}y, \qquad J = J_0.$$

We now derive the dual expression of the configuration form announced in Section 3.

Proof of Proposition 3.3. By inserting the Fourier expansions of $f_1, \dots, f_k$ and by Fubini, we have

$$\begin{aligned}\Lambda(f_0, \dots, f_k) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f_0(x) f_1(x + \varphi_1(y)) \cdots f_k(x + \varphi_k(y)) \mathrm{d}x \psi(y) \mathrm{d}y \\ &= \int_{(\mathbb{R}^n)^k} \widehat{f}_1(\xi_1) \cdots \widehat{f}(\xi_k) \\ &\qquad \int_{\mathbb{R}^n} f_0(x) e\big[(\xi_1 + \cdots + \xi_k) \cdot x\big] \mathrm{d}x \\ &\qquad \int_{\mathbb{R}^m} e\big[\xi_1 \cdot \varphi_1(y) + \cdots + \xi_k \cdot \varphi_k(y)\big] \psi(y) \mathrm{d}y \; \mathrm{d}\xi_1 \dots \mathrm{d}\xi_k.\end{aligned}$$

Recalling Definition 4.1 and the choice (3.5), we deduce that

$$\Lambda(f_0, \dots, f_k) = \int_{(\mathbb{R}^n)^k} \widehat{f}_0(-\xi_1 - \cdots - \xi_k) \widehat{f}_1(\xi_1) \cdots \widehat{f}_k(\xi_k) J(\boldsymbol{\xi}) \mathrm{d}\xi_1 \dots \mathrm{d}\xi_k.$$

□

We single out a useful bound for the configuration operator, typically used when the λ_i are either the measure μ or a mollified version of it.

Proposition 4.2. *For measures $\lambda_0, \dots, \lambda_k \in \mathcal{M}^+(\mathbb{R}^n)$, we have*

$$|\Lambda(\lambda_0, \dots, \lambda_k)| \leqslant \prod_{j=0}^{k} \|\widehat{\lambda}_j\|_\infty^\varepsilon \cdot \Lambda^*\big(|\widehat{\lambda}_0|^{1-\varepsilon}, \dots, |\widehat{\lambda}_k|^{1-\varepsilon}; |J| \big),$$

where the left-hand side is absolutely convergent if the right-hand side is finite.

Proof. This follows from Definition 3.4 and the successive bounds

$$\begin{aligned}
&|\Lambda^*(\widehat{\lambda}_0, \dots, \widehat{\lambda}_k; J)| \\
&\leqslant \int_{(\mathbb{R}^n)^k} |\widehat{\lambda}_0(\xi_1 + \cdots + \xi_k)||\widehat{\lambda}_1(\xi_1)| \cdots |\widehat{\lambda}_k(\xi_k)||J(\boldsymbol{\xi})|\mathrm{d}\boldsymbol{\xi} \\
&\leqslant \prod_{j=0}^{k} \|\widehat{\lambda}_j\|_\infty^\varepsilon \cdot \int_{(\mathbb{R}^n)^k} |\widehat{\lambda}_0(\xi_1 + \cdots + \xi_k)|^{1-\varepsilon}|\widehat{\lambda}_1(\xi_1)|^{1-\varepsilon} \cdots |\widehat{\lambda}_k(\xi_k)|^{1-\varepsilon}|J(\boldsymbol{\xi})|\mathrm{d}\boldsymbol{\xi}.
\end{aligned}$$

□

In some instances we will need a slightly more general multilinear form, as follows.

Definition 4.3 (Smoothed configuration form)**.** *For functions $f_0, \dots, f_k \in \mathcal{S}(\mathbb{R}^n)$ and $F \in \mathcal{S}(\mathbb{R}^{n+m})$, let*

$$\Lambda(f_0, \dots, f_k; F) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} F(x,y) f_0(x) f_1(x + \varphi_1(y)) \cdots f_k(x + \varphi_k(y)) \mathrm{d}x \, \psi(y) \mathrm{d}y. \tag{4.1}$$

Proposition 4.4. *For functions $f_0, \dots, f_k \in \mathcal{S}(\mathbb{R}^n)$ and $F \in \mathcal{S}(\mathbb{R}^{n+m})$, we have*

$$\Lambda(f_0, \dots, f_k; F) = \int_{\mathbb{R}^n \times \mathbb{R}^m} \widehat{F}(\kappa, \theta) \int_{(\mathbb{R}^n)^k} \widehat{f}_0(-\kappa - \xi_1 - \cdots - \xi_k) \prod_{j=1}^{k} \widehat{f}_j(\xi_j) J_\theta(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\kappa \mathrm{d}\theta.$$

Proof. By inserting the Fourier expansions of $F, f_1, \dots, f_k$ and by Fubini, we obtain

$$\begin{aligned}
&\Lambda(f_0, \dots, f_k; F) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} F(x,y) f_0(x) f_1(x + \varphi_1(y)) \cdots f_k(x + \varphi_k(y)) \mathrm{d}x \psi(y) \mathrm{d}y \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^m} \widehat{F}(\kappa, \theta) \int_{(\mathbb{R}^n)^k} \widehat{f}_1(\xi_1) \cdots \widehat{f}(\xi_k) \\
&\qquad \int_{\mathbb{R}^n} f_0(x) e\big[(\kappa + \xi_1 + \cdots + \xi_k) \cdot x\big] \mathrm{d}x \\
&\qquad \int_{\mathbb{R}^m} e\big[\theta \cdot y + \xi_1 \cdot \varphi_1(y) + \cdots + \xi_k \cdot \varphi_k(y)\big] \psi(y) \mathrm{d}y \ \mathrm{d}\xi_1 \dots \mathrm{d}\xi_k \mathrm{d}\kappa \mathrm{d}\theta \\
&= \int_{\mathbb{R}^n \times \mathbb{R}^m} \widehat{F}(\kappa, \theta) \int_{(\mathbb{R}^n)^k} \widehat{f}_0(-\kappa - \xi_1 - \cdots - \xi_k) \widehat{f}_1(\xi_1) \cdots \widehat{f}_k(\xi_k) J_\theta(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\kappa \mathrm{d}\theta.
\end{aligned}$$

□

5. Bounding the singular integral

This section is devoted to the central task of bounding the singular integral (3.6), when the kernel K involved is the oscillatory integral J_θ from Definition 4.1. We will rely crucially on the following decay estimate.

Proposition 5.1. *Assuming that the neighborhood Ω of zero has been chosen small enough, we have*

$$|J_\theta(\boldsymbol{\xi})| \lesssim (1 + |\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta|)^{-m/2} \qquad (\boldsymbol{\xi} \in (\mathbb{R}^n)^k,\ \theta \in \mathbb{R}^m). \tag{5.1}$$

Proof. By Definition 4.1, we have $J_\theta(\boldsymbol{\xi}) = I(\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta, \xi_{kn})$, where

$$I(\gamma, \gamma_{m+1}) = \int_{\mathbb{R}^m} e(\gamma \cdot y + \gamma_{m+1} Q(y))\psi(y)\mathrm{d}x.$$

Consider the hypersurface $S = \{(y, Q(y)) \,:\, y \in \mathrm{Supp}(\psi)\}$ of $\mathbb{R}^{m+1}$, then our assumptions on Q mean that S has non-zero Gaussian curvature. Observe that I is the Fourier transform of $\widetilde{\psi}\mathrm{d}\sigma_S$, where σ_S is the surface measure on S and $\widetilde{\psi}$ is a smooth function with same support as ψ. Therefore it satisfies the decay estimate [42, Chapter VIII]

$$|I(\gamma, \gamma_{m+1})| \lesssim (1 + |\gamma| + |\gamma_{m+1}|)^{-m/2}$$

uniformly in $(\gamma, \gamma_{m+1}) \in \mathbb{R}^{m+1}$, which concludes the proof. □

The main result of this section is a bound on the singular integral for functions in L^s, for a range of s depending on n, m, k. In practice we will apply the proposition below when s is close to 2, which requires the parameter m' to be larger than $(k-1)n$, and when the functions F_i are powers of $|\widehat{\mu}|$ or bounded functions supported on $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$.

Proposition 5.2. *Let $1 + \frac{1}{k} < s < k+1$ and $m' > 0$, and write*

$$K_{\theta,m'}(\boldsymbol{\xi}) = (1 + |\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta|)^{-m'/2} \qquad (\boldsymbol{\xi} \in (\mathbb{R}^n)^k, \theta \in \mathbb{R}^m).$$

Let $F_0, \dots, F_k$ be non-negative measurable functions on $\mathbb{R}^n$. Provided that

$$m' > 2kn - \frac{2(k+1)}{s}n, \tag{5.2}$$

we have, uniformly in $\theta \in \mathbb{R}^m$,

$$\Lambda^*(F_0, \dots, F_k; K_{\theta,m'}) \lesssim_{s,m'} \|F_0\|_s \cdots \|F_k\|_s.$$

The first step towards the proof of this proposition is to bound moments of the kernels $K_{\theta,m'}$ on certain subspaces. Consider the $k+1$ linear maps $(\mathbb{R}^n)^k \to \mathbb{R}^n$ given by

$$\boldsymbol{\xi} \mapsto -(\xi_1 + \cdots + \xi_k) =: \xi_0, \qquad \boldsymbol{\xi} \mapsto \xi_j \qquad (1 \leqslant j \leqslant k).$$

For every $0 \leqslant j \leqslant k$ and $\eta \in \mathbb{R}^n$, the set $\{\boldsymbol{\xi} \in (\mathbb{R}^n)^k \,:\, \xi_j = \eta\}$ is an affine subspace of $(\mathbb{R}^n)^k$ of dimension $(k-1)n$. Recall that $\mathbf{A}^\mathsf{T} : \mathbb{R}^{nk} \to \mathbb{R}^m$, so that in the regime

$m \geqslant (k-1)n$ we expect $(1+|\mathbf{A}^\mathsf{T}\cdot|)^{-1}$ to have bounded moments of order $q > (k-1)n$ on each of the subspaces $\{\xi_j = \eta\}$, under reasonable non-degeneracy conditions on the matrix $\mathbf{A}$. As the next lemma shows, what is needed is precisely the content of Definition 3.1.

Proposition 5.3. *Let $0 \leqslant j \leqslant k$ and suppose that $m \geqslant (k-1)n$. Then for $q > (k-1)n$ we have, uniformly in $\eta \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^m$,*

$$\int_{\xi_j=\eta} (1+|\mathbf{A}^\mathsf{T}\boldsymbol{\xi}+\theta|)^{-q}\,\mathrm{d}\sigma(\boldsymbol{\xi}) \lesssim_q 1.$$

Proof. First note that the assumptions of Definition 3.1 mean that $\mathbf{A}^\mathsf{T}$ is injective on $\{\boldsymbol{\xi} : \xi_j = 0\}$ for $0 \leqslant j \leqslant k$. To see that, observe that the conditions

$$\mathbf{A}^\mathsf{T}\boldsymbol{\xi} = 0,\ \xi_j = 0 \quad \Rightarrow \quad \boldsymbol{\xi} = 0 \qquad\qquad (0 \leqslant j \leqslant k)$$

can be put in matrix form

$$\begin{bmatrix} A_1^\mathsf{T} & \dots & A_j^\mathsf{T} & \dots & A_k^\mathsf{T} \\ 0 & \dots & I_{n\times n} & \dots & 0 \end{bmatrix} \boldsymbol{\xi} = 0 \quad \Rightarrow \quad \boldsymbol{\xi} = 0 \qquad (1 \leqslant j \leqslant k),$$

$$\begin{bmatrix} A_1^\mathsf{T} & \dots & A_k^\mathsf{T} \\ I_{n\times n} & \dots & I_{n\times n} \end{bmatrix} \boldsymbol{\xi} = 0 \quad \Rightarrow \quad \boldsymbol{\xi} = 0.$$

Since $m+n \geqslant kn$, the $(m+n)\times kn$ matrices above have empty kernel if and only if they have rank kn, a set of conditions which is easily seen to be equivalent to that of Definition 3.1.

Now let

$$I = \int_{\xi_j=\eta} (1+|\mathbf{A}^\mathsf{T}\boldsymbol{\xi}+\theta|)^{-q}\mathrm{d}\sigma(\xi).$$

We parametrize the affine subspace $\{\xi_j = \eta\}$ by $\boldsymbol{\xi} = \mathbf{R}\boldsymbol{\xi}' + \boldsymbol{\xi}_\eta$, where $\boldsymbol{\xi}'$ runs over $(\mathbb{R}^n)^k$, $\boldsymbol{\xi}_\eta \in (\mathbb{R}^n)^k$ is picked such that $(\boldsymbol{\xi}_\eta)_j = \eta$, and $\mathbf{R} \in O(\mathbb{R}^{kn})$ is a rotation mapping the subspace $\mathbb{R}^{(k-1)n}$ to $\{\xi_j = 0\}$. We obtain

$$I = \int_{\mathbb{R}^{(k-1)n}} \Big(1+|\mathbf{A}^\mathsf{T}\mathbf{R}\boldsymbol{\xi}'+\mathbf{A}^\mathsf{T}\boldsymbol{\xi}_\eta+\theta|\Big)^{-q}\mathrm{d}\boldsymbol{\xi}',$$

and we write $\mathbf{B} = \mathbf{A}^\mathsf{T}\mathbf{R} \in \mathbb{R}^{m\times kn}$, which is injective on $\mathbb{R}^{(k-1)n}$. Consider the orthogonal decomposition $\mathbf{A}^\mathsf{T}\boldsymbol{\xi}_\eta + \theta = \mathbf{B}\boldsymbol{\xi}_{\eta,\theta} + \gamma$ with $\boldsymbol{\xi}_{\eta,\theta} \in \mathbb{R}^{(k-1)n}$ and $\gamma \in (\mathbf{B}(\mathbb{R}^{(k-1)n}))^\perp$, and observe that by Pythagoras and injectivity,

$$|\mathbf{B}\boldsymbol{\xi}'+\mathbf{A}^\mathsf{T}\boldsymbol{\xi}_\eta+\theta| = |\mathbf{B}(\boldsymbol{\xi}'+\boldsymbol{\xi}_{\eta,\theta})+\gamma| \geqslant |\mathbf{B}(\boldsymbol{\xi}'+\boldsymbol{\xi}_{\eta,\theta})| \gtrsim |\boldsymbol{\xi}'+\boldsymbol{\xi}_{\eta,\theta}|.$$

Via the change of variables $\boldsymbol{\xi}' \leftarrow \boldsymbol{\xi}' + \boldsymbol{\xi}_{\eta,\theta}$, we deduce that

$$I \lesssim \int_{\mathbb{R}^{(k-1)n}} \big(1 + |\boldsymbol{\xi}'|\big)^{-q} \mathrm{d}\boldsymbol{\xi}',$$

which is bounded for $q > (k-1)n$, uniformly in $\eta \in \mathbb{R}^n$. $\square$

Proposition 5.4. *Let $F_0, \dots, F_k$ be non-negative measurable functions on $\mathbb{R}^n$. Let $\tau \in (0,1)$ and $p, p' \in (1, +\infty)$ be parameters with $\frac{1}{p} + \frac{1}{p'} = 1$. Let $H \geqslant 0$ be a parameter and suppose that K is a non-negative measurable function on $\mathbb{R}^{nk}$ such that*

$$\int_{\xi_j = \eta} K(\boldsymbol{\xi})^{p'} \mathrm{d}\sigma(\boldsymbol{\xi}) \leqslant H \qquad (\eta \in \mathbb{R}^n,\ 0 \leqslant j \leqslant k).$$

Then

$$\begin{aligned}&\Lambda^*(F_0, \dots, F_k; K) \\ &\leqslant H^{1/p'} \prod_{j=0}^{k} \left(\int_{\mathbb{R}^n} F_j(\eta)^{\tau p (k+1)/k} \mathrm{d}\eta \right)^{\frac{k}{k+1}\frac{1}{p}} \left(\int_{\mathbb{R}^n} F_j(\eta)^{(1-\tau)p'(k+1)} \mathrm{d}\eta \right)^{\frac{1}{k+1}\frac{1}{p'}}.\end{aligned}$$

Proof. We write $I = \Lambda^*(F_0, \dots, F_k; K)$. By a first application of Hölder:

$$\begin{aligned} I &= \int_{(\mathbb{R}^n)^k} \prod_{j=0}^{k} F_j(\xi_j)^{\tau + (1-\tau)} K(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi} \\ &\leqslant \left[\int_{(\mathbb{R}^n)^k} \left(\prod_{j=0}^{k} F_j(\xi_j) \right)^{\tau p} \mathrm{d}\boldsymbol{\xi} \right]^{\frac{1}{p}} \times \left[\int_{(\mathbb{R}^n)^k} \left(\prod_{j=0}^{k} F_j(\xi_j) \right)^{(1-\tau)p'} K(\boldsymbol{\xi})^{p'} \mathrm{d}\boldsymbol{\xi} \right]^{\frac{1}{p'}} \\ &=: (I_1)^{\frac{1}{p}} \times (I_2)^{\frac{1}{p'}}. \end{aligned} \tag{5.3}$$

We can rewrite I_1 as follows:

$$I_1 = \int_{(\mathbb{R}^n)^k} \prod_{j=0}^{k} F_j(\xi_j)^{\tau p} \mathrm{d}\boldsymbol{\xi} = \int_{(\mathbb{R}^n)^k} \prod_{i=0}^{k} \left[\prod_{\substack{0 \leqslant j \leqslant k \\ j \neq i}} F_j(\xi_j)^{\tau p} \right]^{\frac{1}{k}} \mathrm{d}\boldsymbol{\xi}.$$

By Hölder, we can then reduce to integrals involving each only k of the ξ_j's:

$$I_1 \leqslant \prod_{i=0}^{k} \left[\int_{(\mathbb{R}^n)^k} \prod_{\substack{0 \leqslant j \leqslant k \\ j \neq i}} F_j(\xi_j)^{\tau p (k+1)/k} \mathrm{d}\boldsymbol{\xi} \right]^{\frac{1}{k+1}}.$$

Recall that $\xi_0 = \xi_1 + \cdots + \xi_k$, so that after appropriate changes of variables, each inner integral splits and we have

$$\begin{aligned} I_1 &\leqslant \prod_{i=0}^{k} \Bigg[\prod_{\substack{0 \leqslant j \leqslant k \\ j \neq i}} \int_{\mathbb{R}^n} F_j(\eta)^{\tau p(k+1)/k} \mathrm{d}\eta \Bigg]^{\frac{1}{k+1}} \\ &= \prod_{j=0}^{k} \Big(\int_{\mathbb{R}^n} F_j(\eta)^{\tau p(k+1)/k} \mathrm{d}\eta \Big)^{\frac{k}{k+1}}. \end{aligned} \tag{5.4}$$

To treat the integral I_2, we separate variables by Hölder, and then integrate along slices [35]:

$$\begin{aligned} I_2 &= \int_{(\mathbb{R}^n)^k} \prod_{j=0}^{k} F_j(\xi_j)^{(1-\tau)p'} K(\xi)^{p'} \mathrm{d}\boldsymbol{\xi} \\ &\leqslant \prod_{j=0}^{k} \Big[\int_{(\mathbb{R}^n)^k} F_j(\xi_j)^{(1-\tau)p'(k+1)} K(\boldsymbol{\xi})^{p'} \mathrm{d}\boldsymbol{\xi} \Big]^{\frac{1}{k+1}} \\ &= \prod_{j=0}^{k} \Big[\int_{\eta \in \mathbb{R}^n} F_j(\eta)^{(1-\tau)p'(k+1)} \Big(\int_{\xi_j = \eta} K(\boldsymbol{\xi})^{p'} \mathrm{d}\sigma(\boldsymbol{\xi}) \Big) \mathrm{d}\eta \Big]^{\frac{1}{k+1}}. \end{aligned}$$

Inside each inner integral we use the fiber moment condition, so that eventually

$$I_2 \leqslant H \prod_{j=0}^{k} \Big(\int_{\mathbb{R}^n} F_j(\eta)^{(1-\tau)p'(k+1)} \mathrm{d}\eta \Big)^{\frac{1}{k+1}}. \tag{5.5}$$

The proof is finished upon inserting (5.4) and (5.5) into (5.3). □

It remains to determine the parameters (τ, p) in Proposition 5.4 that lead to a bound involving a single L^s norm.

Corollary 5.5. *Suppose that* $1 + \frac{1}{k} < s < k + 1$*. Then there exists unique parameters* $\tau \in (0, 1)$ *and* $p \in (1, \infty)$ *depending on* k *and* s *such that*

$$s = \frac{k+1}{k} p\tau = (k+1)p'(1-\tau), \tag{5.6}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$*, and for such* (τ, p) *we have*

$$\frac{k+1}{s} = \frac{k}{p} + \frac{1}{p'}, \tag{5.7}$$

$$\frac{1}{p'} = \frac{1}{k-1}\Big(k - \frac{k+1}{s}\Big). \tag{5.8}$$

Proof. Starting from (5.6), and dividing by $\frac{k+1}{k}p$ in the first identity and by $\frac{k+1}{k}pp'$ in the second, we obtain the equivalent identities

$$\tau = \frac{k}{k+1}\frac{s}{p} \quad \text{and} \quad \Big(\frac{k}{p}+\frac{1}{p'}\Big)\tau = \frac{k}{p}. \tag{5.9}$$

Inserting the left-hand expression of τ in the right-hand identity, we deduce the relation (5.7). This is easily solved in p, p' and one finds that

$$\frac{1}{p} = \frac{1}{k-1}\Big(\frac{k+1}{s}-1\Big), \qquad \frac{1}{p'} = \frac{1}{k-1}\Big(k-\frac{k+1}{s}\Big),$$

which in particular recovers (5.8). It can be checked that $\frac{1}{p} \in (0,1)$ under the given conditions on s. Inserting this value of $\frac{1}{p}$ in the first identity of (5.9), we find that

$$\tau = \frac{k}{k-1}\Big(1-\frac{s}{k+1}\Big),$$

which again lies in $(0,1)$ for the given range of s. $\square$

Proof of Proposition 5.2. Apply Proposition 5.4 with $K(\boldsymbol{\xi}) = (1+|\mathbf{A}^{\mathsf{T}}\boldsymbol{\xi}+\theta|)^{-m'/2}$, and the choice of parameters (τ, p) from Proposition 5.5. By (5.7), this gives

$$\begin{aligned}|\lambda^*(F_0,\dots,F_k;K)| &\leqslant H^{1/p'}\prod_{j=0}^{k}(\|F_j\|_s^s)^{\frac{1}{k+1}\left(\frac{k}{p}+\frac{1}{p'}\right)} \\ &= H^{1/p'}\prod_{j=0}^{k}\|F_j\|_s,\end{aligned}$$

where $H = \max_j \sup_{\eta,\theta}\int_{\xi_j=\eta}(1+|\mathbf{A}^{\mathsf{T}}\boldsymbol{\xi}+\theta|)^{-p'm'/2}\mathrm{d}\sigma(\boldsymbol{\xi})$. Via Proposition 5.3 and (5.8), we have $H \lesssim_{s,m'} 1$ provided that

$$m' > \frac{2(k-1)n}{p'} = 2\Big(k-\frac{k+1}{s}\Big)n.$$

$\square$

From Proposition 5.2, we now derive useful bounds on the dual form Λ^*, which are needed to develop the results of Sections 6–8. In the course of the proof, we refer to a restriction estimate from Appendix B, which states essentially that $\widehat{\mu}$ is in $L^{2+\varepsilon}$ when β remains bounded away from zero and α is close enough to n. Recall the notation $T^\kappa f = f(\kappa + \cdot)$ from Section 2.

Proposition 5.6. *Let $\beta_0 \in (0,n)$ and suppose that for a constant $c>0$ small enough with respect to n,k,m,*

$$(k-1)n < m < kn, \quad \beta_0 \leqslant \beta < n, \quad n - c\beta_0 \leqslant \alpha < n.$$

Then there exists $\varepsilon \in (0,1)$ depending at most on n, k, m such that

$$\sup_{(\kappa,\theta)\in\mathbb{R}^n\times\mathbb{R}^m} \Lambda^*(\, T^\kappa|\widehat{\mu}|^{1-\varepsilon}, |\widehat{\mu}|^{1-\varepsilon}, \dots, |\widehat{\mu}|^{1-\varepsilon}; |J_\theta|^{1-\varepsilon}\,) < \infty,$$

$$\Lambda^*(\, |\widehat{\mu}|^{1-\varepsilon}, \dots, |\widehat{\mu}|^{1-\varepsilon}; |J|\,) \lesssim_{\beta_0} (n-\alpha)^{-O(1)}.$$

Proof. Let $\varepsilon, \delta \in (0,1)$ be parameters. Recalling the majoration (5.1), we apply Proposition 5.2 to $F_0 = T^\kappa|\widehat{\mu}|^{1-\varepsilon}$ and $F_i = |\widehat{\mu}|^{1-\varepsilon}$ for $i \geqslant 1$, with parameters $m' = (1-\varepsilon)m$ and $s = \frac{2+\delta}{1-\varepsilon}$. The condition (5.2) is fulfilled when $m > (k-1)n$ and ε, δ are small enough with respect to n, k, m. We obtain, uniformly in $\kappa \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^m$,

$$\Lambda^*(T^\kappa|\widehat{\mu}|^{1-\varepsilon}, |\widehat{\mu}|^{1-\varepsilon}, \dots, |\widehat{\mu}|^{1-\varepsilon}; |J_\theta|^{1-\varepsilon}) \lesssim_{\varepsilon,s} \| |\widehat{\mu}|^{1-\varepsilon} \|_s^{k+1} = \|\widehat{\mu}\|_{2+\delta}^{(1-\varepsilon)(k+1)},$$

By Proposition B.3 and (3.3), we conclude that

$$\Lambda^*(T^\kappa|\widehat{\mu}|^{1-\varepsilon}, |\widehat{\mu}|^{1-\varepsilon}, \dots, |\widehat{\mu}|^{1-\varepsilon}; |J_\theta|^{1-\varepsilon}) \lesssim_{\varepsilon,\delta,\beta_0} (n-\alpha)^{-O(1)},$$

and the second bound follows since $|J| \lesssim |J|^{1-\varepsilon}$. □

Proof of Proposition 3.6. Let $\varepsilon \in (0,1)$ be a parameter. By Proposition 4.2 and (5.1), we have

$$|\Lambda(f_0, \dots, f_k)| \leqslant \textstyle\prod_{0\leqslant j\leqslant k} \|\widehat{f}_j\|_\infty^\varepsilon \cdot \Lambda^*\big(|\widehat{f}_0|^{1-\varepsilon}, \dots, |\widehat{f}_k|^{1-\varepsilon}; (1+|\mathbf{A}^\intercal \cdot|)^{-m/2}\big). \tag{5.10}$$

For ε small enough with respect to n, k, m, we may apply Proposition 5.2 with $s = \frac{2}{1-\varepsilon}$ and $m' = m$, together with Plancherel:

$$\begin{aligned}
\Lambda^*(|\widehat{f}_0|^{1-\varepsilon}, \dots, |\widehat{f}_k|^{1-\varepsilon}; (1+|\mathbf{A}^\intercal \cdot|)^{-m/2}) &\lesssim \textstyle\prod_{j=0}^k \| |\widehat{f}_j|^{1-\varepsilon} \|_{2/(1-\varepsilon)} \\
&= \textstyle\prod_{j=0}^k \|\widehat{f}_j\|_2^{1-\varepsilon} \\
&= \textstyle\prod_{j=0}^k \|f_j\|_2^{1-\varepsilon} \\
&\leqslant \textstyle\prod_{j=0}^k \|f_j\|_\infty^{1-\varepsilon},
\end{aligned}$$

where we used the assumption $\mathrm{Supp}(f_j) \subset \big[-\frac{1}{2}, \frac{1}{2}\big]^n$ in the last line. Inserting this bound in (5.10) finishes the proof. □

6. The configuration measure

In this section, we aim to construct the measure $\nu \in \mathcal{M}^+(\mathbb{R}^{n+m})$ specified in Proposition 3.7. We make extensive use of the singular integral bounds derived in the previous section. Our treatment is similar to that of Chan et al. [10], but we work in a more abstract setting. We assume throughout this section that the dimensionality conditions (3.7) are met, so that singular integral bounds are available.

We start with the proper definition of ν, which is the content of the next proposition (recall Definition 3.2 and Proposition 3.3). We define an extra shift function $\varphi_0 = 0$ for notational convenience.

Proposition 6.1. *Define the functional ν at $F \in \mathcal{S}(\mathbb{R}^{n+m})$ by*

$$\langle \nu, F \rangle = \lim_{\varepsilon \to 0} \Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon; F),$$

where $\mu_\varepsilon = \mu * \phi_\varepsilon$ for an approximate identity ϕ_ε with $\phi \in \mathcal{C}^\infty_{c,+}(\mathbb{R}^n)$. Then ν is well-defined and we have, for every $F \in \mathcal{S}(\mathbb{R}^{n+m})$,

$$\langle \nu, F \rangle = \Lambda^*(\widehat{\mu}, \dots, \widehat{\mu}; \widehat{F}),$$
$$|\langle \nu, F \rangle| \leqslant \|F\|_\infty \Lambda(\mu, \dots, \mu),$$

where the integrals defined by the right-hand side expressions converge absolutely. Therefore ν extends by density to a positive bounded linear operator on $\mathcal{C}_c(\mathbb{R}^{n+m})$.

Proof. By Proposition 4.4, we have

$$\begin{aligned}&\Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon; F)\\ &= \int_{\mathbb{R}^{n+m}} \widehat{F}(\kappa, \theta) \int_{(\mathbb{R}^n)^k} \widehat{\mu}(-\kappa - \xi_1 - \dots - \xi_k) \prod_{j=1}^k \widehat{\mu}(\xi_j) J_\theta(\boldsymbol{\xi}) h_\varepsilon(\boldsymbol{\xi}, \kappa) \mathrm{d}\boldsymbol{\xi} \mathrm{d}\kappa \mathrm{d}\theta\end{aligned}$$

where $h_\varepsilon(\boldsymbol{\xi}, \kappa) = \widehat{\phi}(-\varepsilon(\kappa + \xi_1 + \dots + \xi_k)) \prod_{j=1}^k \widehat{\phi}(\varepsilon \xi_j)$. Since h_ε is bounded by one in absolute value and tends to 1 pointwise as $\varepsilon \to 0$, the limit of $\Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon; F)$ as $\varepsilon \to 0$ exists and equals $\Lambda^*(\widehat{\mu}, \dots, \widehat{\mu}; \widehat{F})$ by dominated convergence, since we have uniform boundedness of

$$\begin{aligned}&\int_{\mathbb{R}^{n+m}} |\widehat{F}(\kappa, \theta)| \int_{(\mathbb{R}^n)^k} |\widehat{\mu}(\kappa + \xi_1 + \dots + \xi_k)| \prod_{j=1}^k |\widehat{\mu}(\xi_j)| |J_\theta(\boldsymbol{\xi})| \mathrm{d}\boldsymbol{\xi} \mathrm{d}\kappa \mathrm{d}\theta\\ &\leqslant \sup_{(\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^m} \Lambda^*(\,|T^\kappa \widehat{\mu}|, |\widehat{\mu}|, \dots, |\widehat{\mu}|; |J_\theta|\,) \times \int_{\mathbb{R}^{n+m}} |\widehat{F}(\kappa, \theta)| \mathrm{d}\kappa \mathrm{d}\theta < \infty,\end{aligned}$$

via Proposition 5.6 and the majorations $|J_\theta| \lesssim |J_\theta|^{1-\varepsilon}$, $|\widehat{\mu}| \leqslant |\widehat{\mu}|^{1-\varepsilon}$. Recalling Definitions 3.2 and 4.3, and using the positivity of μ_ε, we have also

$$|\langle \nu, F \rangle| = \lim_{\varepsilon \to 0} |\Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon; F)| \leqslant \|F\|_\infty \varlimsup_{\varepsilon \to 0} \Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon).$$

By Fourier inversion (Proposition 3.3) and another instance of the dominated convergence theorem, exploiting the finiteness of $\Lambda^*(|\widehat{\mu}|, \dots, |\widehat{\mu}|; |J|)$ provided by Proposition 5.6, we obtain

$$|\langle \nu, F \rangle| \leqslant \|F\|_\infty \varlimsup_{\varepsilon \to 0} \Lambda^*(\widehat{\mu}_\varepsilon, \dots, \widehat{\mu}_\varepsilon; J) = \|F\|_\infty \Lambda^*(\widehat{\mu}, \dots, \widehat{\mu}; J).$$

This last quantity equals $\|F\|_\infty \Lambda(\mu, \dots, \mu)$ by Definition 3.4. $\square$

Proposition 6.2. *When defined, the measure ν of Proposition 6.1 is supported on the compact set*

$$X = \{(x,y) \in E \times \operatorname{Supp}\psi \,:\, (x, x+\varphi_1(y), \dots, x+\varphi_k(y)) \in E^{k+1}\}.$$

Proof. We can rewrite $X = (E \times \operatorname{Supp}\psi) \cap \Phi^{-1}(E^{k+1})$, where Φ is the smooth map defined by (3.4), so that X is closed and bounded, and therefore compact. Since its complement X^c is open, it is enough to show that $\langle \nu, F\rangle = 0$ for every $F \in \mathcal{C}^\infty_{c,+}(\mathbb{R}^{n+m})$ such that $\operatorname{Supp} F \subset X^c$. By compactness we know that there exists $c > 0$ such that

$$\max_{0 \leqslant j \leqslant k} d(x+\varphi_j(y), E) \geqslant c > 0 \quad \text{for all} \quad (x,y) \in \operatorname{Supp} F \cap (\mathbb{R}^n \times \operatorname{Supp}\psi).$$

On the other hand,

$$\langle \nu, F\rangle = \lim_{\varepsilon \to 0} \int_{\operatorname{Supp} F \cap (\mathbb{R}^n \times \operatorname{Supp}\psi)} F(x,y) \prod_{j=0}^{k} \mu_\varepsilon(x+\varphi_j(y)) \mathrm{d}x\, \psi(y) \mathrm{d}y.$$

For ε small enough, since μ_ε is supported on $E + B(0, C\varepsilon)$ for a certain $C > 0$, the integrand above is always zero. $\square$

Proposition 6.3. *We have $\|\nu\| = \Lambda(\mu, \dots, \mu)$.*

Proof. Consider the compact set X from Proposition 6.2, and the larger compact set

$$Y = \{(x,y) \in \mathbb{R}^n \times \operatorname{Supp}\psi \,:\, d(x+\varphi_j(y), E) \leqslant 1 \quad \text{for } 0 \leqslant j \leqslant k\}.$$

Pick a smoothed ball indicator $F \in \mathcal{C}^\infty_{c,+}(\mathbb{R}^{n+m})$ such that $F = 1$ on Y. Since ν is supported on $X \subset Y$, we have

$$\nu(\mathbb{R}^{n+m}) = \langle \nu, F\rangle = \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n \times \operatorname{Supp}\psi} F(x,y) \prod_{j=0}^{k} \mu_\varepsilon(x+\varphi_j(y)) \mathrm{d}x \psi(y) \mathrm{d}y.$$

Since $(x,y) \mapsto \prod_{j=0}^k \mu_\varepsilon(x+\varphi_j(y))$ is supported on Y for ε small enough, we have therefore

$$\nu(\mathbb{R}^{n+m}) = \lim_{\varepsilon \to 0} \Lambda(\mu_\varepsilon, \dots, \mu_\varepsilon).$$

By the same reasoning as in the end of the proof of Proposition 6.1, using again the bound $\Lambda^*(\,|\widehat{\mu}|, \dots, |\widehat{\mu}|; |J|\,) < \infty$ provided by Proposition 5.6, we find eventually that $\|\nu\| = \Lambda(\mu, \dots, \mu)$. $\square$

We now turn to the last expected feature of the configuration measure ν, which is that it has zero mass on any hyperplane.

Proposition 6.4. *We have $\nu(H) = 0$ for every hyperplane H of $\mathbb{R}^{n+m}$.*

Proof. Consider a hyperplane $H < \mathbb{R}^{n+m}$ and a rotation $R \in O_{n+m}(\mathbb{R})$ such that $H = R(\mathbb{R}^{n+m-1} \times \{0\})$. Consider parameters $L \geqslant 1$ and $\delta \in (0,1]$. We consider a Schwartz function F_δ of the form

$$F_\delta \circ R = \chi\Big(\frac{\cdot}{L}\Big)\,\Xi\Big(\frac{\cdot}{\delta}\Big), \tag{6.1}$$

where $\chi \in \mathcal{S}(\mathbb{R}^{n+m-1})$ and $\Xi \in \mathcal{S}(\mathbb{R})$ are non-negative such that $\chi \geqslant 1$ on $[-1,1]^{n+m-1}$ and $\Xi(0) \geqslant 1$. Writing $H_L = R(\,[-L,L]^{n+m-1} \times \{0\}\,)$, we have therefore $\nu(H_L) \leqslant \langle \nu, F_\delta \rangle$, and it is enough to show that $\langle \nu, F_\delta\rangle$ tends to 0 as $\delta \to 0$, for every fixed $L \geqslant 1$. By Proposition 6.1, and writing $\gamma = (\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^m$, we have

$$\langle \nu, F_\delta \rangle = \int\limits_{\mathbb{R}^{n+m}} \int\limits_{(\mathbb{R}^n)^k} \widehat{F}_\delta(\gamma)\widehat{\mu}(-\kappa - \xi_1 - \cdots - \xi_k) \prod_{j=1}^{k} \widehat{\mu}(\xi_j) J_\theta(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}\mathrm{d}\gamma. \tag{6.2}$$

We assume that χ, Ξ have been chosen so that their Fourier transforms are supported on centered balls of radius 1, which is certainly possible. Recalling (6.1), we have therefore, for every $(u,v) \in \mathbb{R}^{n+m-1} \times \mathbb{R}$,

$$|\widehat{F}_\delta \circ R(u,v)| = |\widehat{F_\delta \circ R}(u,v)| \lesssim L^{n+m-1} \cdot 1_{|u| \leqslant L^{-1}} \cdot \delta \cdot 1_{|v| \leqslant \delta^{-1}}. \tag{6.3}$$

We next show how to obtain some uniform γ-decay from the other factor in the integrand of (6.2). By (3.2) and (5.1), and since $\beta \leqslant n \leqslant m$, we have

$$\begin{aligned}
&|\widehat{\mu}(\kappa + \xi_1 + \cdots + \xi_k)| \textstyle\prod_{j=1}^k |\widehat{\mu}(\xi_j)||J_\theta(\boldsymbol{\xi})| \\
&\lesssim_\alpha (1 + |\kappa + \xi_1 + \cdots + \xi_k|)^{-\frac{\beta}{2}} \textstyle\prod_{j=1}^k (1+|\xi_j|)^{-\frac{\beta}{2}} (1 + |\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta|)^{-\frac{m}{2}} \\
&\lesssim_\alpha (1 + |\kappa + \xi_1 + \cdots + \xi_k| + |\xi_1| + \cdots + |\xi_k| + |\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta|)^{-\frac{\beta}{2}}.
\end{aligned}$$

Using the above in cunjunction with the triangle inequality and the decompositions $\theta = (\mathbf{A}^\mathsf{T}\boldsymbol{\xi} + \theta) - \sum_{j=1}^k A_j^\mathsf{T}\xi_j$ and $\kappa = (\kappa + \xi_1 + \cdots + \xi_k) - \sum_{j=1}^k \xi_j$, we deduce that

$$\begin{aligned}
&|\widehat{\mu}(\kappa + \xi_1 + \cdots + \xi_k)| \textstyle\prod_{j=1}^k |\widehat{\mu}(\xi_j)||J_\theta(\boldsymbol{\xi})| \\
&\lesssim_\alpha (1 + |\kappa| + |\theta|)^{-\frac{\beta}{2}} \\
&\asymp (1 + |\gamma|)^{-\frac{\beta}{2}}.
\end{aligned} \tag{6.4}$$

Let $\varepsilon \in (0,1)$ be the small parameter in the statement of the proposition. At this point we have two parametrizations $\gamma = (\kappa,\theta) = R(u,v)$ with $(\kappa,\theta) \in \mathbb{R}^n \times \mathbb{R}^m$, $(u,v) \in \mathbb{R}^{n+m-1} \times \mathbb{R}$. By integrating in (u,v)-coordinates in (6.2), and bounding $\widehat{F}_\delta(\gamma)$ via (6.3), we obtain

$$|\langle \nu, F_\delta\rangle| \lesssim \int_{\mathbb{R}^{n+m-1}\times\mathbb{R}} 1_{|u|\leqslant L^{-1}} \cdot L^{n+m-1} \cdot \delta \cdot 1_{|v|\leqslant\delta^{-1}} \left[\int_{(\mathbb{R}^n)^k} |\widehat{\mu}(\kappa+\xi_1+\cdots+\xi_k)| \prod_{j=1}^{k} |\widehat{\mu}(\xi_j)||J_\theta(\boldsymbol{\xi})|\mathrm{d}\boldsymbol{\xi}\right]\mathrm{d}u\mathrm{d}v$$

By pulling out an ε-th power of the inner integrand and using (6.4), we infer that

$$\begin{aligned}
&|\langle \nu, F_\delta\rangle| \\
&\lesssim_\alpha \int_{\mathbb{R}^{n+m-1}} L^{n+m-1}\cdot 1_{|u|\leqslant L^{-1}} \int_{\mathbb{R}} \delta\cdot 1_{|v|\leqslant\delta^{-1}}\cdot(1+|(u,v)|)^{-\frac{\varepsilon\beta}{2}} \\
&\qquad \left[\int_{(\mathbb{R}^n)^k} |\widehat{\mu}(\kappa+\xi_1+\cdots+\xi_k)|^{1-\varepsilon} \prod_{j=1}^{k} |\widehat{\mu}(\xi_j)|^{1-\varepsilon}|J_\theta(\boldsymbol{\xi})|^{1-\varepsilon}\mathrm{d}\boldsymbol{\xi}\right]\mathrm{d}u\mathrm{d}v \\
&\lesssim \sup_{(\kappa,\theta)\in\mathbb{R}^n\times\mathbb{R}^m} \Lambda^*(\,|T^\kappa\widehat{\mu}|^{1-\varepsilon}, |\widehat{\mu}|^{1-\varepsilon},\ldots,|\widehat{\mu}|^{1-\varepsilon}; |J_\theta|^{1-\varepsilon}\,)\times\delta\int_{|v|\leqslant\delta^{-1}}(1+|v|)^{-\frac{\varepsilon\beta}{2}}\mathrm{d}v.
\end{aligned}$$

The supremum above is finite by Proposition 5.6, and for ε small enough, the last factor is bounded by $\delta^{\varepsilon\beta/2}$. Therefore $\langle\nu, F_\delta\rangle \to 0$ as $\delta\to 0$, as was to be shown. □

Proof of Proposition 3.7. It suffices to combine Propositions 6.1– 6.4, recalling that we assumed (3.7) in this section. □

7. Absolutely continuous estimates

In this section we verify that absolutely continuous estimates are available when the shifts in (3.4) are given by polynomial vectors and the singular integral converges. We work with the notation of abstract shift functions.

The strategy, as in the regularity proof of Roth's theorem [43], is to use the U^2 arithmetic regularity lemma to decompose a non-negative bounded function into an almost-periodic component, an L^2 error, and a part which is Fourier-small. The precise version of the regularity lemma that we need is found in Appendix A. To neglect the contribution of Fourier-small functions, we use the fact that the counting operator is controlled by the Fourier L^∞ norm for bounded functions, in the sense of Proposition 3.6. To show that the pattern count for almost-periodic functions is high, we need uniform lower bounds for certain Bohr sets of almost-periods, the proof of which will occupy subsequent parts of this section. We define a Bohr set of $\mathbb{T}^n$ of frequency set $\Gamma\subset\mathbb{Z}^n$, radius $\delta\in\left(0,\frac{1}{2}\right]$ and dimension $d=|\Gamma|<\infty$ by

$$B = B(\Gamma,\delta) = \{x\in\mathbb{T}^n : \|\xi\cdot x\|\leqslant\delta\quad\forall\,\xi\in\Gamma\}. \tag{7.1}$$

We first prove the following conditional version of Proposition 3.8.

Proposition 7.1. *Suppose that $m > (k-1)n$ and, uniformly for every Bohr set B of $\mathbb{T}^n$ of dimension d and radius $\delta > 0$,*

$$\mathcal{L}\left\{ y \in [-c,c]^m \, : \, \varphi_1(y), \dots, \varphi_k(y) \in B \right\} \gtrsim_{d,\delta} 1.$$

Then for every function $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ supported on $\left[-\frac{1}{8}, \frac{1}{8}\right]^n$ such that $0 \leqslant f \leqslant 1$ and $\int f = \tau \in (0,1]$, we have

$$\Lambda(f, \dots, f) \gtrsim_\tau 1.$$

Proof. We let $\kappa : (0,1]^3 \to (0,1]$ be a decay function and $\varepsilon \in (0,1]$ be a parameter, both to be determined later. Write the decomposition of Proposition A.2 with respect to ε, κ as $f = f_1 + f_2 + f_3 = g + f_3$. Note that $f_1, g \geqslant 0$ and f_1, f_2, f_3, g are supported in $\left[-\frac{1}{4}, \frac{1}{4}\right]^n$ and uniformly bounded by 2 in absolute value. Expanding $f = g + f_3$ by multilinearity, and using Proposition 3.6 together with the Fourier bound on f_3 in (A.5), we obtain

$$\begin{aligned} \Lambda(f, \dots, f) &= \Lambda(g, \dots, g) + O\big(\textstyle\sum \Lambda(*, \dots, f_3, \dots, *) \big) \\ &= \Lambda(g, \dots, g) + O(\kappa(\varepsilon, d^{-1}, \delta)^{\varepsilon'}), \end{aligned} \tag{7.2}$$

for a certain $\varepsilon' \in (0,1)$ depending at most on n, k, m. Recall that we assumed that ψ is at least 1 on a box $[-c,c]^m$ in Section 3, and let

$$E = \left\{ y \in [-c,c]^m \, : \, \varphi_1(y), \dots, \varphi_k(y) \in B \right\}, \tag{7.3}$$

where B is the Bohr set of Proposition A.2. For reasons that shall be clear later, we first restrict integration to the set E, using the non-negativity of g:

$$\Lambda(g, \dots, g) \geqslant \int_E \left(\int_{\mathbb{R}^n} g \cdot T^{\varphi_1(y)} g \cdots T^{\varphi_k(y)} g \, \mathrm{d}\mathcal{L} \right) \mathrm{d}y.$$

Next, we focus on the decomposition $g = f_1 + f_2$ and exploit the L^2 bound on f_2 in (A.5) by Cauchy-Schwarz in the inner integral:

$$\begin{aligned} \Lambda(g, \dots, g) &\geqslant \int_E \left(\int_{\mathbb{R}^n} g \cdot T^{\varphi_1(y)} g \cdots T^{\varphi_k(y)} g \, \mathrm{d}\mathcal{L} \right) \mathrm{d}y \\ &\geqslant \int_E \left(\int_{\mathbb{R}^n} f_1 \cdot T^{\varphi_1(y)} f_1 \cdots T^{\varphi_k(y)} f_1 \, \mathrm{d}\mathcal{L} - \sum \int_{\mathbb{R}^n} * \cdots T^{\varphi_j(y)} f_2 \cdots * \, \mathrm{d}\mathcal{L} \right) \mathrm{d}y \\ &\geqslant \int_E \left(\int_{\mathbb{R}^n} f_1 \cdot T^{\varphi_1(y)} f_1 \cdots T^{\varphi_k(y)} f_1 \, \mathrm{d}\mathcal{L} - O(\varepsilon) \right) \mathrm{d}y. \end{aligned}$$

Finally, we use the almost-periodicity estimate for f_1 in (A.5) and the definition (7.3) of E to replace the shifts of f_1 by itself:

$$\Lambda(g,\dots,g) \geqslant \int_E \Big(\int_{[-\frac{1}{2},\frac{1}{2}]^n} f_1^{k+1} \mathrm{d}\mathcal{L} - O(\varepsilon) \Big) \mathrm{d}y$$

By nesting of $L^p\big(\big[-\frac{1}{2},\frac{1}{2}\big]^n\big)$ norms and non-negativity of f_1, we infer that

$$\begin{aligned}\Lambda(g,\dots,g) &\geqslant \int_E \Big(\Big(\int_{[-\frac{1}{2},\frac{1}{2}]^n} f_1 \mathrm{d}\mathcal{L} \Big)^{k+1} - O(\varepsilon) \Big) \mathrm{d}y \\ &= \mathcal{L}(E)\cdot(\tau^{k+1} - O(\varepsilon)).\end{aligned}$$

Choosing $\varepsilon = c\tau^{k+1}$ with a small $c > 0$, and recalling (7.2) and the assumption on E, we obtain

$$\Lambda(f,\dots,f) \geqslant c(\delta, d^{-1})\tau^{k+1} - O(\kappa(c\tau^{k+1}, d^{-1}, \delta)^{\varepsilon'}).$$

Choosing $\kappa(\varepsilon, d^{-1}, \delta) = c' \cdot \big(c(\delta, d^{-1})\varepsilon\big)^{1/\varepsilon'}$, we obtain

$$\Lambda(f,\dots,f) \geqslant \tfrac{1}{2} c(\delta, d^{-1})\tau^{k+1} \gtrsim_\tau 1,$$

recalling that $d, \delta^{-1} \lesssim_{\varepsilon,\kappa} 1 \lesssim_\tau 1$. □

It remains to determine a lower bound on the measure of the intersection of preimages of a Bohr set by the shift functions. This can be done when the shift functions are polynomial vectors, by reduction to a known diophantine approximation problem, and in fact there will be a series of intermediate reductions. We let d denote the L^∞ metric on $\mathbb{R}^n$ or $\mathbb{R}$ and we define

$$\|x\|_{\mathbb{T}^n} = d(x, \mathbb{Z}^n) = \max_{1\leqslant i\leqslant n} d(x_i, \mathbb{Z})$$

for $x \in \mathbb{R}^n$. In all subsequent propositions in this section we also liberate the letters n, k, m from their usual meaning, and we indicate the dependencies of implicit constants in all parameters. Our objective is to prove the following statement.

Proposition 7.2. *Let $t, m, n, \ell, d \geqslant 1$. Let $Q_1,\dots,Q_t : \mathbb{R}^m \to \mathbb{R}^n$ be polynomial vectors with components of degree at most ℓ, and such that $Q_i(0) = 0$ for all $i \in [t]$. For $\xi_1,\dots,\xi_d \in \mathbb{R}^n$, we have*

$$\mathcal{L}\big\{\, y \in [-c,c]^m \,:\, \|Q_i(y)\cdot\xi_j\|_{\mathbb{T}} < \varepsilon \quad \forall (i,j) \in [t]\times[d] \,\big\} \gtrsim_{\varepsilon,\ell,m,t,d,n} 1.$$

Our first reduction is to a finite system of conditions on monomials modulo one.

Proposition 7.3. *Let $\ell, m \geqslant 1$ and $X = \{0,\dots,\ell\}^m \smallsetminus \{0\}$. For every $I \in X$, let $d_I \in \mathbb{N}_0$ and $\xi_I \in \mathbb{R}^{d_I}$. We have*[3]

$$\mathcal{L}\big\{\, y \in [-c,c]^m \;:\; \|y^I \xi_I\|_{\mathbb{T}^{d_I}} \leqslant \varepsilon \quad \forall I \in X \,\big\} \gtrsim_{\varepsilon,\ell,m,(d_I)} 1.$$

Proof that Proposition 7.3 implies Proposition 7.2.

We let $X = \{0,\dots,\ell\}^m \smallsetminus \{0\}$ and we write $Q_i = \sum_{k\in[n]} Q_{ik} e_k$, $Q_{ik} = \sum_{I\in X} a_I^{(ik)} y^I$. For every $I \in X$ we define $d_I = t+d+n$ and $\xi_I = (a_I^{(ik)}\xi_{jk})_{(i,j,k)} \in \mathbb{T}^{t+d+n}$, to make the following observation:

$$\begin{aligned}
&\|Q_i(y)\cdot\xi_j\|_{\mathbb{T}} \leqslant \varepsilon && \forall (i,j) \in [t]\times[d] \\
\Leftrightarrow\;& \|\textstyle\sum_{k\in[n]}\sum_{I\in X} a_I^{(ik)} y^I \xi_{jk}\|_{\mathbb{T}} \leqslant \varepsilon && \forall (i,j) \in [t]\times[d] \\
\Leftarrow\;& \|y^I a_I^{(ik)}\xi_{jk}\|_{\mathbb{T}} \leqslant \frac{\varepsilon}{n\ell^m} && \forall (i,j,k) \in [t]\times[d]\times[n],\ I\in X \\
\Leftrightarrow\;& \|y^I \xi_I\|_{\mathbb{T}^{d_I}} \leqslant \frac{\varepsilon}{n\ell^m} && \forall I \in X.
\end{aligned}$$

Applying Proposition 7.3 with $\varepsilon \leftarrow \varepsilon/n\ell^m$ and (d_I,ξ_I) as above, we find a lower bound on the quantity under study which depends only on ε,ℓ,m,t,d,n. ☐

Our second reduction consists in a straightforward induction which reduces the dimension of the problem to 1.

Proposition 7.4. *Let $\ell \geqslant 1$ and $d_1,\dots,d_\ell \in \mathbb{N}_0$, $\xi_1 \in \mathbb{R}^{d_1},\dots,\xi_\ell \in \mathbb{R}^{d_\ell}$. We have*

$$\mathcal{L}\big\{\, y \in [-c,c] \;:\; \|y^j \xi_j\|_{\mathbb{T}^{d_j}} \leqslant \varepsilon \quad \forall j \in [\ell] \,\big\} \gtrsim_{\varepsilon,\ell,(d_i)} 1.$$

Proof that Proposition 7.4 implies Proposition 7.3.

We induct on $m \geqslant 1$, the case $m = 1$ being precisely Proposition 7.4. Assume that we have proven the estimate for dimensions less than or equal to m, and write a tuple $I \in \{0,\dots,\ell\}^{m+1} \smallsetminus \{0\}$ as $I = (J, i_{m+1})$ with $J \in \{0,\dots,\ell\}^m$ and $i_{m+1} \geqslant 0$. We distinguish the conditions involving y_{m+1} or not by Fubini:

$$\begin{aligned}
&\mathcal{L}\big\{\, y \in [-c,c]^{m+1} \;:\; \|y^I\xi_I\|_{\mathbb{T}^{d_I}} \leqslant \varepsilon \quad \forall\, I \in X \,\big\} \\
&= \int_{[-c,c]^{m+1}} 1\Big[\, \|y^J y_{m+1}^{i_{m+1}} \xi_I\|_{\mathbb{T}^{d_I}} \leqslant \varepsilon \quad \forall\,(J,i_{m+1}) = I \in X \,\Big] \mathrm{d}y_1 \dots \mathrm{d}y_m \mathrm{d}y_{m+1} \\
&= \int_{[-c,c]^m} 1\Big[\, \|y^J\xi_I\|_{\mathbb{T}^{d_I}} \leqslant \varepsilon \quad \forall (J,0) = I \in X \,\Big] \\
&\qquad \int_{[-c,c]} 1\Big[\, \|y_{m+1}^{i_{m+1}} \cdot y^J \xi_I\|_{\mathbb{T}^{d_I}} \leqslant \varepsilon \quad \forall (J,i_{m+1}) = I \in X \,:\, i_{m+1} \geqslant 1 \,\Big] \mathrm{d}y_{m+1}\, \mathrm{d}y_1 \dots \mathrm{d}y_m.
\end{aligned}$$

[3] Here and in the sequel we set $\mathbb{R}^0 = \{0\}$ and $\|0\|_{\mathbb{T}^0} = 0$ so that the conditions involving a space $\mathbb{R}^0$ are void.

By first applying the induction hypothesis with $m = 1$ at fixed $y_1, \dots, y_m$, and then by applying another instance of the induction hypothesis, we find that this quantity is indeed bounded from below by a positive constant depending only on ε, ℓ, m and (d_I). □

Our final reduction is a simple discretization argument which reduces the problem to the following known diophantine approximation estimate [29, Proposition B.2] (see also [18, Proposition A.2], [2, Chapter 7]).

Proposition 7.5. *Let $\ell \geqslant 1$ and $d_1, \dots, d_\ell \in \mathbb{N}_0$. Let $\alpha_1 \in \mathbb{R}^{d_1}, \dots, \alpha_\ell \in \mathbb{R}^{d_\ell}$ and $N \geqslant 1$. We have*

$$N^{-1} \# \big\{ |n| \leqslant N \,:\, \| n^j \alpha_j \|_{\mathbb{T}^{d_j}} \leqslant \varepsilon \quad \forall j \in [\ell] \big\} \gtrsim_{\varepsilon, \ell, (d_j)} 1.$$

Proof that Proposition 7.5 implies Proposition 7.4.

Consider a scale $N \geqslant 1$ going to infinity. Write each $|y| \leqslant c$ as $y = \frac{n+u}{N}$ with $n \in \mathbb{Z}$ and $u \in \big(-\frac{1}{2}, \frac{1}{2} \big]$, so that $y^j = n^j / N^j + O_\ell(1/N)$ for every $j \in [\ell]$. For N large enough with respect to (ξ_j), ε and ℓ, we have therefore

$$\| y^j \xi_j \|_{\mathbb{T}} \leqslant \varepsilon \quad \Leftarrow \quad \Big\| n^j \frac{\xi_j}{N^j} \Big\|_{\mathbb{T}} \leqslant \frac{\varepsilon}{2}.$$

This yields:

$$\begin{aligned}
&\mathcal{L}\big\{ \, y \in [-c, c] \,:\, \| y^j \xi_j \|_{\mathbb{T}^{d_j}} \leqslant \varepsilon \quad \forall j \in [\ell] \, \big\} \\
&\geqslant \sum_{|n| \leqslant cN/2} \mathcal{L}\big\{ \, y = \tfrac{n+u}{N} \,:\, |u| \leqslant \tfrac{1}{2}, \ \Big\| n^j \frac{\xi_j}{N^j} \Big\|_{\mathbb{T}^{d_j}} \leqslant \varepsilon/2 \quad \forall j \in [\ell] \, \big\} \\
&\geqslant N^{-1} \# \big\{ |n| \leqslant cN/2 \,:\, \Big\| n^j \frac{\xi_j}{N^j} \Big\|_{\mathbb{T}^{d_j}} \leqslant \varepsilon \quad \forall j \in [\ell] \, \big\}.
\end{aligned}$$

Applying Proposition 7.5 concludes the proof. □

To conclude this section we may now derive the absolutely continuous estimates stated in Section 3.

Proof of Proposition 3.8. It suffices to combine Propositions 7.1 and 7.2, recalling the shape (3.5) of our shift functions. □

8. The transference argument

This section is concerned with proving that $\Lambda(\mu, \dots, \mu) > 0$, by the transference argument of Łaba and Pramanik [28] exploiting the pseudorandomness of the fractal measure μ as $\alpha \to n$. We start by recalling the decomposition of Chan et al. [10, Section 6] of the fractal measure μ into a bounded smooth part (a mollified version of μ) and a Fourier-small part (the difference with the first part). This is the part of the argument where one lets α tend to n in a certain sense, and then the Fourier tail exhibits very strong, exponential-type decay in $n-\alpha$.

Proposition 8.1. *There exists a constant $C_1 > 0$ depending at most on n, D and a decomposition $\mu = \mu_1 + \mu_2$, where $\mu_1 = f \mathrm{d}\mathcal{L}$, $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, $0 \leqslant f \leqslant C_1$, $\int f = 1$, $\operatorname{Supp} f \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^n$, $|\widehat{\mu}_i| \leqslant 2|\widehat{\mu}|$ for $i \in \{1,2\}$ and*

$$\|\widehat{\mu}_2\|_\infty \lesssim (n-\alpha)^{-O(1)} e^{-\frac{\beta}{2+\beta}\cdot\frac{1}{n-\alpha}}.$$

Proof. Let $L \geqslant 1$ be a parameter. Consider a cutoff $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\int \phi = 1$, $\operatorname{Supp} \phi \subset B(0, \frac{1}{16})$ and $0 \leqslant \phi \leqslant C_0$, for a certain $C_0 = C_0(n) > 0$, and define $\phi_L = L^n \phi(L\,\cdot)$. Let $f = \mu * \phi_L$ and consider the decomposition $\mu = \mu_1 + \mu_2$ with $\mu_1 = f \mathrm{d}\mathcal{L}$ and $\mu_2 = \mu - \mu_1$. We can already infer that $f \geqslant 0$, $\int f = 1$, $|\widehat{\mu}_i| \leqslant 2|\widehat{\mu}|$ for $i = 1, 2$ and $\operatorname{Supp} \mu_1 \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^n$, since we assumed that $E \subset \left[-\frac{1}{16}, \frac{1}{16}\right]^n$ in Section 3.

Next, we show that f is bounded. Since ϕ_L has support in $B(0, \frac{1}{16L})$, by (3.1) we have

$$\begin{aligned} f(x) &= \int_{B(x, \frac{1}{16L})} \phi_L(x-y) \mathrm{d}\mu(y) \\ &\leqslant \|\phi_L\|_\infty \cdot \mu\big[B(x, \tfrac{1}{16L})\big] \\ &\leqslant C_0 D L^{n-\alpha}. \end{aligned}$$

Choosing $L = e^{\frac{1}{n-\alpha}}$, we deduce that

$$\|f\|_\infty \leqslant C_0 D e =: C_1.$$

Finally, we bound the Fourier transform of $\widehat{\mu}_2$. Observe that, for every $\xi \in \mathbb{R}^n$,

$$\widehat{\mu}_2(\xi) = \widehat{\mu}(\xi)\Big(1 - \widehat{\phi}\Big(\frac{\xi}{L}\Big)\Big). \tag{8.1}$$

Since $\int \phi = 1$, we always have $|1 - \widehat{\phi}(\frac{\xi}{L})| \leqslant 2$. On the other hand, since ϕ has support in $B(0, \frac{1}{16})$, we have

$$\Big|1 - \widehat{\phi}\Big(\frac{\xi}{L}\Big)\Big| = \left| \int_{B(0,\frac{1}{16})} \phi(x)\Big(1 - e\Big(\frac{\xi \cdot x}{L}\Big)\Big) \mathrm{d}x \right| \lesssim \frac{|\xi|}{L}.$$

By inserting these two last bounds in (8.1), we obtain

$$|\widehat{\mu}_2(\xi)| \lesssim \min\Big(1, \frac{|\xi|}{L}\Big) |\widehat{\mu}(\xi)|.$$

Consequently, by (3.2) and (3.3) we have

$$|\widehat{\mu}_2(\xi)| \lesssim (n-\alpha)^{-O(1)} \min\Big(1, \frac{|\xi|}{L}\Big) \min(1, |\xi|^{-\beta/2}).$$

By considering separately the ranges $|\xi| \geqslant L^{2/(2+\beta)}$ and $|\xi| \leqslant L^{2/(2+\beta)}$, we find that

$$|\widehat{\mu}_2(\xi)| \lesssim (n-\alpha)^{-O(1)} L^{-\frac{\beta}{2+\beta}}.$$

Recalling our choice of L, we have

$$|\widehat{\mu}_2(\xi)| \lesssim (n-\alpha)^{-O(1)} e^{-\frac{\beta}{2+\beta}\cdot\frac{1}{n-\alpha}}.$$

□

We now establish the positivity of $\Lambda(\mu,\dots,\mu)$, using the previous decomposition, with the main contribution from the absolutely continuous part estimated by Proposition 3.8, and the other contributions bounded away by Proposition 5.6.

Proof of Proposition 3.9. We consider the decomposition $\mu = \mu_1 + \mu_2$ from Proposition 8.1, and expand by multilinearity in

$$\Lambda(\mu,\dots,\mu) = C_1^{-(k+1)} \Lambda(\mu_1/C_1,\dots,\mu_1/C_1) + O\big(\textstyle\sum \Lambda(*,\dots,\mu_2,\dots,*) \big),$$

where the sum is over $2^{k+1}-1$ terms and the stars denote measures equal to either μ_1 or μ_2. By Proposition 3.8, we deduce that for a certain constant $c>0$, we have

$$\Lambda(\mu,\dots,\mu) \geqslant c - O\big(\textstyle\sum \Lambda(*,\dots,\mu_2,\dots,*) \big).$$

By Proposition 4.2, we have furthermore, for any $\varepsilon \in (0,1)$,

$$\Lambda(\mu,\dots,\mu) \geqslant c - O\big(\|\widehat{\mu}_2\|_\infty^\varepsilon \, \Lambda^*(|\widehat{\mu}|^{1-\varepsilon},\dots,|\widehat{\mu}|^{1-\varepsilon}; |J|) \big).$$

By taking ε to be that appearing in Proposition 5.6, and inserting the Fourier bound on μ_2 from Proposition 8.1, we find that

$$\Lambda(\mu,\dots,\mu) \geqslant c - O_{\beta_0}\Big((n-\alpha)^{-O(1)} e^{-\varepsilon\cdot\frac{\beta_0}{2+\beta_0}\cdot\frac{1}{n-\alpha}} \Big),$$

where we used the monotonicity of $x/(2+x)$. This can be made positive for $\alpha \geqslant n - c(\beta_0,\varepsilon)$ with $c(\beta_0,\varepsilon) > 0$ small enough. □

9. Revisiting the linear case

In this section we indicate how the method of this article may be modified to obtain the following extension of Theorem 1.2, which allows for any positive exponent of Fourier decay for the fractal measure. For simplicity we only treat the case where n divides m, which already covers all the geometric applications discussed in [10].

Theorem 9.1. *Let $n,k,m \geqslant 1$, $D \geqslant 1$ and $\alpha,\beta \in (0,n)$. Suppose that E is a compact subset of $\mathbb{R}^n$ and μ is a probability measure supported on E such that*

$$\mu\big[B(x,r)\big] \leqslant Dr^\alpha \quad \text{and} \quad |\widehat{\mu}(\xi)| \leqslant D(n-\alpha)^{-D}(1+|\xi|)^{-\beta/2}$$

for all $x \in \mathbb{R}^n$, $r > 0$ and $\xi \in \mathbb{R}^n$. Suppose that $(A_1, \ldots, A_k)$ is a non-degenerate system of $n \times m$ matrices in the sense of Definition 1.1. Assume finally that $m = (k-r)n$ with $1 \leqslant r < k$ and, for some $\beta_0 \in (0, n)$,

$$\frac{k-1}{2} n < m < kn, \qquad \beta_0 \leqslant \beta < n, \qquad n - c_{n,k,m,\beta_0,D,(A_i)} \leqslant \alpha < n,$$

for a sufficiently small constant $c_{n,k,m,\beta_0,D,(A_i)} > 0$. Then, for every collection of strict subspaces $V_1, \ldots, V_q$ of $\mathbb{R}^{n+m}$, there exists $(x, y) \in \mathbb{R}^{n+m} \smallsetminus V_1 \cup \ldots V_q$ such that

$$(x, x + A_1 y, \ldots, x + A_k y) \in E^{k+1}.$$

Note that the condition on m is equivalent to that of Theorem 1.2. We only sketch the proof of Theorem 9.1, since it follows by a straightforward adaption of the methods of this paper, with the only difference lying in the treatment of the singular integral.

We start by stating a slight generalization of Hölder's inequality that was already used (for $\ell = k+1$, $r = k$) in the proof of Proposition 5.4. We write $\binom{[\ell]}{r}$ for the set of subsets of $[\ell]$ of size r.

Proposition 9.2. *Let $(X, \mathfrak{M}, \lambda)$ be a measured space and let $1 \leqslant r \leqslant \ell$. For measurable functions $F_1, \ldots, F_\ell : X \to \mathbb{C}$, we have*

$$\int_X \prod_{j \in [\ell]} |F_j| \, \mathrm{d}\lambda \leqslant \prod_{S \in \binom{[\ell]}{r}} \left[\int_X \prod_{j \in S} |F_j|^{\ell/r} \mathrm{d}\lambda \right]^{1/\binom{\ell}{r}}.$$

Proof. First observe that, for arbitrary real numbers $a_1, \ldots a_\ell \geqslant 0$, we have

$$\prod_{j \in [\ell]} a_j = \prod_{S \in \binom{[\ell]}{r}} \left(\prod_{j \in S} a_j \right)^{1/\binom{\ell-1}{r-1}}.$$

Next, let $I = \int_X \prod_{j \in [\ell]} |F_j| \, \mathrm{d}\lambda$ and apply Hölder's inequality in

$$\begin{aligned} I &= \int_X \prod_{S \in \binom{[\ell]}{r}} \left(\prod_{j \in S} |F_j| \right)^{1/\binom{\ell-1}{r-1}} \mathrm{d}\lambda \\ &\leqslant \prod_{S \in \binom{[\ell]}{r}} \left[\int_X \left(\prod_{j \in S} |F_j| \right)^{\binom{\ell}{r}/\binom{\ell-1}{r-1}} \mathrm{d}\lambda \right]^{1/\binom{\ell}{r}}. \end{aligned}$$

A quick computation shows that $\binom{\ell}{r}/\binom{\ell-1}{r-1} = \ell/r$, which concludes the proof. $\square$

We now place ourselves under the assumptions of Theorem 9.1, and in particular we assume that the matrices $A_1, \ldots, A_k$ are non-degenerate in the sense of Definition 1.1. We also write $A_0 = 0_{n \times n}$ throughout. This matches the framework of this paper except that now $Q = 0$.

We fix a smooth cutoff $\psi \in \mathcal{C}^\infty_{c,+}(\mathbb{R}^n)$ which is at least 1 on a box $[-c,c]^n$. We define the oscillatory integral

$$J(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e(\mathbf{A}^\mathsf{T}\boldsymbol{\xi}\cdot y)\psi(y)\mathrm{d}y = \widehat{\psi}(-\mathbf{A}^\mathsf{T}\boldsymbol{\xi}). \tag{9.1}$$

The counting operators are now defined by[4]

$$\Lambda(f_0,\dots,f_k) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} f_0(x)f_1(x+A_1y)\cdots f_k(x+A_ky)\mathrm{d}x\ \psi(y)\mathrm{d}y,$$
$$\Lambda^*(F_0,\dots,F_k;J) = \int_{(\mathbb{R}^n)^k} F_0(-\xi_1-\cdots-\xi_k)F_1(\xi_1)\cdots F_k(\xi_k)J(\boldsymbol{\xi})\mathrm{d}\boldsymbol{\xi},$$

for functions $f_i, F_i \in \mathcal{S}(\mathbb{R}^n)$, and we have $\Lambda(f_0,\dots,f_k) = \Lambda^*(\widehat{f_0},\dots,\widehat{f_k};J)$ as before.

Since we assumed that $\psi \in \mathcal{C}^\infty_c(\mathbb{R}^m)$, it follows from (9.1) that

$$|J(\boldsymbol{\xi})| \lesssim_N (1+|\mathbf{A}^\mathsf{T}\boldsymbol{\xi}|)^{-N} \tag{9.2}$$

for every $N>0$. Via some matricial considerations (as in [10, Lemma 3.2]), it can be checked that Definition 1.1 is equivalent to the requirement that $\mathbf{A}^\mathsf{T} : \mathbb{R}^{kn} \to \mathbb{R}^{kn-rn}$ is injective on each subspace of the form

$$\{\boldsymbol{\xi} \in (\mathbb{R}^n)^k \,:\, (\xi_j)_{j\in S} = \boldsymbol{\eta}\},$$

where S is a subset of $\{0,\dots,k\}$ of size r and $\boldsymbol{\eta} \in \mathbb{R}^{rn}$, and we wrote $\xi_0 = -(\xi_1+\cdots+\xi_k)$ as before. Now consider an arbitrary subset S of $\{0,\dots,k\}$ of size r. By (9.2) one quickly deduces that

$$\int_{(\xi_j)_{j\in S}=\boldsymbol{\eta}} |J(\boldsymbol{\xi})|^q \mathrm{d}\sigma(\boldsymbol{\xi}) \lesssim_q 1 \qquad (q>0,\ \boldsymbol{\eta}\in(\mathbb{R}^n)^r), \tag{9.3}$$

in the same manner as in the proof of Proposition 5.3.

In our linear setting one may naturally obtain a better range of m for which the multilinear form Λ^* is controlled by L^s norms. The next proposition demonstrates this, and it is applicable to our problem only when when $\frac{k+1}{r} > 2$, or equivalently $m = (k-r)n > \frac{k-1}{2}n$.

Proposition 9.3. *We have*

$$|\Lambda^*(F_0,\dots,F_k;J)| \lesssim \|F_0\|_{(k+1)/r}\cdots\|F_k\|_{(k+1)/r}.$$

[4]In fact, one could work without cutoff functions in the y-variable, as was done in [10], which simplifies the estimates somewhat. Here we keep smooth cutoffs to stay closer to the framework of the article.

Proof. Write $I = \Lambda^*(F_0, \dots, F_k; J)$ and $[0,k] = \{0, \dots, k\}$ for the purpose of this proof. By Proposition 9.2, we have

$$I \leqslant \int_{(\mathbb{R}^n)^k} \prod_{j=0}^{k} \big(F_j(\xi_j) \cdot |J(\boldsymbol{\xi})|^{\frac{1}{k+1}} \big) \mathrm{d}\boldsymbol{\xi}$$

$$\leqslant \prod_{S \in \binom{[0,k]}{r}} \left[\int_{(\mathbb{R}^n)^k} \prod_{j \in S} F_j(\xi_j)^{\frac{k+1}{r}} |J(\boldsymbol{\xi})|^{\frac{1}{r}} \mathrm{d}\boldsymbol{\xi} \right]^{1/\binom{k+1}{r}}.$$

Integrating along slices, and invoking (9.3), we obtain

$$I \leqslant \prod_{S \in \binom{[0,k]}{r}} \left[\int_{(\mathbb{R}^n)^r} \prod_{j \in S} F_j(\eta_j)^{\frac{k+1}{r}} \bigg(\int_{(\xi_j)_{j\in S} = \boldsymbol{\eta}} |J(\boldsymbol{\xi})|^{\frac{1}{r}} \mathrm{d}\sigma(\boldsymbol{\xi}) \bigg) \mathrm{d}\boldsymbol{\eta} \right]^{1/\binom{k+1}{r}}$$

$$\lesssim \prod_{S \in \binom{[0,k]}{r}} \left[\int_{(\mathbb{R}^n)^r} \prod_{j \in S} F_j(\eta_j)^{\frac{k+1}{r}} \mathrm{d}\boldsymbol{\eta} \right]^{1/\binom{k+1}{r}}.$$

Therefore each inner integral splits and we have

$$I \lesssim \prod_{S \in \binom{[0,k]}{r}} \left[\prod_{j \in S} \int_{\mathbb{R}^n} F_j(\eta)^{\frac{k+1}{r}} \mathrm{d}\boldsymbol{\eta} \right]^{1/\binom{k+1}{r}}$$

$$= \prod_{j \in [0,k]} \left[\int_{\mathbb{R}^n} F_j(\eta)^{\frac{k+1}{r}} \mathrm{d}\boldsymbol{\eta} \right]^{\binom{k}{r-1}/\binom{k+1}{r}}.$$

Since $\binom{k+1}{r}/\binom{k}{r-1} = (k+1)/r$, it follows that $I \leqslant \prod_{j \in [0,k]} \|F_j\|_{\frac{k+1}{r}}$, as was to be shown. ☐

With Proposition 9.3 in hand, it is a simple matter to adapt the rest of the argument in this paper. In fact, one would need a slight variant of that proposition involving a shift θ, as in the case of Proposition 5.2. From such a proposition one may deduce the natural analogues of Propositions 5.6 and 3.6, which will impose the same conditions on α and β, and a distinct condition $m > \frac{k-1}{2} n$ on m. With these singular integral bounds in hand, the arguments of Sections 6–8 go through essentially unchanged, and one obtains Theorem 9.1 by the process described at the end of Section 3.

Appendix A. The arithmetic regularity lemma

In this section, we derive a version of the U^2 arithmetic regularity lemma following Tao's argument [43], with minor twists to accomodate functions defined over $\mathbb{R}^n$ instead of $\mathbb{T}^n$. This set of ideas itself originates in a paper of Bourgain [9], albeit in a rather different language. We include the complete proof since the exact result we need is not stated in a convenient form in the literature.

We defined a Bohr set of $\mathbb{T}^n$ of frequency set $\Gamma \subset \mathbb{Z}^n$, radius $\delta \in \left(0, \frac{1}{2}\right]$ and dimension $d = |\Gamma| < \infty$ by (7.1). We define the dilate of a Bohr set B of frequency set Γ and radius δ by a factor $\rho \in (0,1]$ as $B(\Gamma,\delta)_\rho = B(\Gamma, \rho\delta)$. Note that $B(\Gamma,\delta) = \phi^{-1}(2\delta \cdot Q)$ for the cube $Q = \left[-\frac{1}{2}, \frac{1}{2}\right]^n$ and the morphism $\phi : \mathbb{R}^n \to \mathbb{T}^d$, $x \mapsto (\xi \cdot x)_{\xi \in \Gamma}$. We can find a cube covering of the form $Q \subset \bigcup_{t \in T}(t + \delta \cdot Q)$ with $|T| = \lceil 1/\delta \rceil^d \leqslant (2/\delta)^d$, and therefore

$$1 = |\phi^{-1}(Q)| \leqslant \sum_{t \in T} |\phi^{-1}(t + \delta \cdot Q)|.$$

By the pigeonhole principle, there exists $t \in T$ such that $|\phi^{-1}(t + \delta \cdot Q)| \geqslant (\delta/2)^d$, and since $\phi^{-1}(t + \delta \cdot Q) - \phi^{-1}(t + \delta \cdot Q) \subset B$, we deduce that

$$|B| = |B(\Gamma,\delta)| \geqslant (\delta/2)^d \quad \text{for all } \delta \in (0, \tfrac{1}{2}]. \tag{A.1}$$

Now consider the tent function $\Delta(x) = (1 - |x|)^+$ on $\mathbb{R}$, which is 1-Lipschitz, bounded by 1 everywhere, and bounded from below by 1/2 on $\left[-\frac{1}{2}, \frac{1}{2}\right]$. For any Bohr set B, we define functions $\phi_B, \nu_B : \mathbb{T}^n \to \mathbb{C}$ by

$$\phi_B(x) = \Delta\Big(\frac{1}{\delta} \sup_{\xi \in \Gamma} \|\xi \cdot x\|\Big), \quad \nu_B = \frac{\phi_B}{\int \phi_B},$$

so that $\int \nu_B = 1$ and $\frac{1}{2} 1_{B_{1/2}} \leqslant \nu_B \leqslant 1_B$. The function ν_B is essentially a smoothed normalized indicator function of the Bohr set B, and its most important properties are summarized in the following proposition.

Proposition A.1. *For any Bohr set B of frequency set $\Gamma \subset \mathbb{Z}^n$ and radius $\delta \in (0, \frac{1}{2}]$, we have*

$$\|\nu_B\|_\infty \lesssim (\delta/4)^{-d}, \tag{A.2}$$

$$\|T^t \nu_B - \nu_B\|_\infty \lesssim (\delta/4)^{-d} \rho \qquad \textit{for } t \in B_\rho,\ \rho \in (0,1], \tag{A.3}$$

$$\widehat{\nu_B}(\xi) = 1 + O(\delta) \qquad \textit{for } \xi \in \Gamma. \tag{A.4}$$

Proof. Note that $\int \phi_B \geqslant \frac{1}{2}|B_{1/2}| \geqslant \frac{1}{2}(\delta/4)^d$ by (A.1), which implies the first estimate. For every $x, t \in \mathbb{T}^n$, we also have

$$|\nu_B(x+t) - \nu_B(x)| \leqslant 2(\delta/4)^{-d} \Big| \Delta\Big(\frac{1}{\delta} \sup_{\xi \in \Gamma} \|\xi \cdot (x+t)\|\Big) - \Delta\Big(\frac{1}{\delta} \sup_{\xi \in \Gamma} \|\xi \cdot x\|\Big) \Big|.$$

When $t \in B_\rho$, we have $\|\xi \cdot t\| \leqslant \rho\delta$ for every $\xi \in \Gamma$, and therefore $|\nu_B(x+t) - \nu_B(x)| \lesssim (\delta/4)^{-d}\rho$ since Δ is 1-Lipschitz, and we have established the second estimate. To obtain the third, consider $\xi \in \Gamma$, and observe that since ν_B is supported on B and $\|\xi \cdot x\| \leqslant \delta$

for $x \in B$, we have

$$\widehat{\nu}_B(\xi) = \int_B \nu_B(x) e(-\xi \cdot x) \mathrm{d}x = (1 + O(\delta)) \cdot \int \nu_B = 1 + O(\delta).$$

$\square$

Proposition A.2. *Let $\varepsilon \in (0,1]$ be a parameter and let $\kappa : (0,1]^3 \to (0,1]$ be a decay function. Suppose that $f \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ is such that $0 \leqslant f \leqslant 1$ and $\operatorname{Supp} f \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^n$. Then there exists a decomposition $f = f_1 + f_2 + f_3$ with $f_i \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, $\operatorname{Supp} f_i \subset \left[-\frac{1}{4}, \frac{1}{4}\right]^n$, $\|f_i\|_\infty \leqslant 1$, $f_1 \geqslant 0$, $f_1 + f_2 \geqslant 0$, $\int f_1 = \int f$ as well as a Bohr set B of dimension $d \lesssim_{\varepsilon,\kappa} 1$ and radius $\delta \gtrsim_{\varepsilon,\kappa} 1$ such that*

$$\text{(A.5)} \qquad \|T^t f_1 - f_1\|_\infty \leqslant \varepsilon \quad \forall t \in B, \qquad \|f_2\|_2 \leqslant \varepsilon, \qquad \|\widehat{f_3}\|_{L^\infty(\mathbb{R}^n)} \leqslant \kappa(\varepsilon, d^{-1}, \delta).$$

Proof. We initially consider f as defined on the torus $\mathbb{T}^n$, by identification with its 1-periodization from the cube $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$. Consider sequences of positive real numbers

$$\tfrac{1}{2} \geqslant \delta_0 \geqslant \delta_1 \geqslant \ldots \geqslant \delta_i \geqslant \ldots \quad \text{and} \quad 1 \geqslant \eta_1 \geqslant \ldots \geqslant \eta_i \geqslant \ldots$$

to be determined later. We define sequences of frequency sets Γ_i and Bohr sets B_i of dimension d_i, and measures ν_i inductively for $i \geqslant 0$ by

$$\text{(A.6)} \qquad \Gamma_{i+1} = \Gamma_i \bigcup \{\, |\widehat{f}| \geqslant \eta_{i+1} \,\} \bigcup \bigcup_{j=0}^{i} \{\, |\widehat{\nu}_j| \geqslant \eta_{i+1} \,\},$$

$$B_{i+1} = B\big(\Gamma_{i+1}, \delta_{i+1}\big), \qquad \nu_{i+1} = \nu_{B_{i+1}}.$$

We initialize with $\Gamma_0 = \{e_1, \ldots, e_n\}$, $\delta_0 \leqslant \frac{1}{8}$, $B_0 = B(\Gamma_0, \delta_0)$, $\nu_0 = \nu_{B_0}$, so that $d_0 = n$ and by the definition (7.1) of Bohr sets, we have $B_i \subset \left[-\frac{1}{8}, \frac{1}{8}\right]^n$ for all i. Note that, by Tchebychev, we also have a dimension bound

$$d_{i+1} \leqslant d_i + \frac{\|\widehat{f}\|_2^2}{\eta_{i+1}^2} + \sum_{j=0}^{i} \frac{\|\widehat{\nu}_j\|_2^2}{\eta_{i+1}^2}.$$

By Plancherel and the bound (A.2), it follows that

$$\text{(A.7)} \qquad d_i \lesssim_{\delta_0, \ldots, \delta_{i-1}, d_{i-1}, \eta_i} 1 \qquad (i \geqslant 1).$$

We start by finding a piece of the Fourier expansion of f which is small in L^2. To this end observe that

$$\sum_{i=0}^{k} \sum_{\Gamma_{i+2} \smallsetminus \Gamma_i} |\widehat{f}|^2 \leqslant 2\|\widehat{f}\|_2^2 = 2\|f\|_2^2 \leqslant 2.$$

By Tchebychev's bound, it follows that

$$\#\Big\{0 \leqslant i \leqslant k \,:\, \sum_{\Gamma_{i+2} \smallsetminus \Gamma_i} |\widehat{f}|^2 \geqslant \frac{\varepsilon^2}{2}\Big\} \leqslant \frac{4}{\varepsilon^2}.$$

Choosing $k = \lceil 4/\varepsilon^2 \rceil$, we obtain the existence of an index $0 \leqslant i \leqslant k$ such that

$$\sum_{\Gamma_{i+2} \smallsetminus \Gamma_i} |\widehat{f}|^2 \leqslant \frac{\varepsilon^2}{2}. \tag{A.8}$$

We now decompose f into three pieces $f_1, f_2, f_3 : \mathbb{T}^n \to \mathbb{C}$ defined by

$$f = f * \nu_i + (f * \nu_{i+1} - f * \nu_i) + (f - f * \nu_{i+1}) = f_1 + f_2 + f_3.$$

Since f takes values in $[0, 1]$ and $\int \nu_i = 1$, the functions f_1, f_2, f_3 take values in $[-1, 1]$ by simple convolution bounds. It is also clear that f_1 and $f_1 + f_2$ are non-negative and $\int f_1 = \int f$.

Let us first analyze the L^2-small piece. By Plancherel and (A.8), we have

$$\begin{aligned} \|f * \nu_{i+1} - f * \nu_i\|_2^2 &= \sum_{m \in \mathbb{Z}^n} |\widehat{f}(m)|^2 |\widehat{\nu_{i+1}}(m) - \widehat{\nu_i}(m)|^2 \\ &\leqslant \frac{\varepsilon^2}{2} + \sum_{m \in \Gamma_i \cup (\mathbb{Z}^n \smallsetminus \Gamma_{i+2})} |\widehat{f}(m)|^2 |\widehat{\nu_{i+1}}(m) - \widehat{\nu_i}(m)|^2 \end{aligned} \tag{A.9}$$

For $m \in \Gamma_i \subset \Gamma_{i+1}$, by (A.4) we have

$$|\widehat{\nu_{i+1}}(m) - \widehat{\nu_i}(m)| \lesssim \delta_{i+1} + \delta_i.$$

For $m \notin \Gamma_{i+2}$ the definition (A.6) of Γ_{i+2} implies that $|\widehat{\nu_i}(m)| \leqslant \eta_{i+2}$ and $|\widehat{\nu_{i+1}}(m)| \leqslant \eta_{i+2}$. Inserting these bounds into (A.9), we obtain

$$\|f * \nu_{i+1} - f * \nu_i\|_2^2 \leqslant \frac{\varepsilon^2}{2} + O(\delta_i + \delta_{i+1} + \eta_{i+2}) \leqslant \varepsilon^2, \tag{A.10}$$

provided that $\delta_j, \eta_j \leqslant c\varepsilon^2$ for all j.

Next, let us focus on the almost-periodic piece. Introducing a parameter $\rho_i \in (0, 1]$, we deduce from (A.3) that for $t \in B_{\rho_i}$, we have

$$\begin{aligned} \|T^t f * \nu_i - f * \nu_i\|_\infty &\leqslant \|f\|_1 \|T^t \nu_i - \nu_i\|_\infty \\ &\lesssim_n \delta_i^{-d_i} \rho_i \\ &\leqslant \varepsilon, \end{aligned} \tag{A.11}$$

choosing $\rho_i = c_n \varepsilon \delta_i^{d_i}$. We write $\widetilde{\delta}_i = \rho_i \delta_i$, and from (A.7) we see that $\widetilde{\delta}_i$ depends at most on $n, \varepsilon, \delta_0, \ldots, \delta_i, \eta_1, \ldots, \eta_i$.

Finally, we consider the Fourier-small piece. By Fourier inversion,

$$\|(f - f * \nu_{i+1})^\wedge\|_{\ell^\infty(\mathbb{Z}^n)} = \sup_{m \in \mathbb{Z}^n} |\widehat{f}(m)||1 - \widehat{\nu}_{i+1}(m)|.$$

For $m \in \Gamma_{i+1}$, we have $|1 - \widehat{\nu}_{i+1}(m)| \lesssim \delta_{i+1}$ by (A.4), while for $m \notin \Gamma_{i+1}$, the definition (A.6) of Γ_{i+1} shows that $|\widehat{f}(m)| \leqslant \eta_{i+1}$. Therefore

$$\|(f - f * \nu_{i+1})^\wedge\|_{\ell^\infty(\mathbb{Z}^n)} \lesssim \delta_{i+1} + \eta_{i+1} \leqslant c\kappa(\varepsilon, d_i^{-1}, \widetilde{\delta}_i), \tag{A.12}$$

for a small constant $c > 0$ provided that we choose the δ_j, η_j recursively satisfying

$$\max(\delta_{i+1}, \eta_{i+1}) = c \min(\kappa(\varepsilon, d_i^{-1}, \widetilde{\delta}_i), \varepsilon^2).$$

At this stage we have obtained the desired bounds (A.5) over $\mathbb{T}^n$ and for a Bohr set $\widetilde{B}_i = B_i(\Gamma_i, \widetilde{\delta}_i)$, and from (A.7) and the construction of the δ_i it follows that

$$d_i \lesssim_{\varepsilon,\kappa} 1 \quad \text{and} \quad \delta_i \gtrsim_{\varepsilon,\kappa} 1.$$

To finish the proof we now consider the functions f_1, f_2, f_3 as functions on $\mathbb{R}^n$ supported on $\left[-\frac{1}{2}, \frac{1}{2}\right]^n$. Since f and the Bohr sets measures ν_i are supported on $\left[-\frac{1}{8}, \frac{1}{8}\right]^n$, the convolutions $f * \nu_i$ over $\mathbb{T}^n$ may be readily interpreted as convolutions over $\mathbb{R}^n$, and the functions f_i are supported on $\left[-\frac{1}{4}, \frac{1}{4}\right]^n$. The properties (A.10) and (A.11) are readily viewed as holding over $\mathbb{R}^n$, thus we only need to verify that f_3 has the appropriate Fourier decay at real frequencies. We claim that since f_3 has support in $\left[-\frac{1}{4}, \frac{1}{4}\right]^n$, we have $\|\widehat{f_3}\|_{L^\infty(\mathbb{R}^n)} \lesssim \|\widehat{f_3}\|_{\ell^\infty(\mathbb{Z}^n)}$ and by taking the constant c in (A.12) small enough, we obtain the desired Fourier decay estimate. To prove this claim, consider a smooth bump function χ equal to 1 on $\left[-\frac{1}{4}, \frac{1}{4}\right]^n$. For $\xi \in \mathbb{R}^n$, expanding f as a Fourier series yields

$$\begin{aligned}
\widehat{f_3}(\xi) &= \int_{[-\frac{1}{4}, \frac{1}{4}]^n} f_3(x)\chi(x)e(-\xi \cdot x)\mathrm{d}x \\
&= \sum_{k \in \mathbb{Z}^n} \widehat{f_3}(k) \int_{\mathbb{R}^n} \chi(x)e((k - \xi) \cdot x)\mathrm{d}x \\
&= \sum_{k \in \mathbb{Z}^n} \widehat{f_3}(k)\widehat{\chi}(\xi - k).
\end{aligned}$$

Using the smoothness of χ, it follows that, uniformly in $\xi \in \mathbb{R}^n$,

$$|\widehat{f_3}(\xi)| \lesssim \|\widehat{f_3}\|_{\ell^\infty(\mathbb{Z}^n)} \sum_{k \in \mathbb{Z}^n} (1 + |\xi - k|)^{-(n+1)} \lesssim \|\widehat{f_3}\|_{\ell^\infty(\mathbb{Z}^n)}.$$

□

Appendix B. Uniform restriction estimates for fractal measures

In this section we obtain restriction estimates for fractal measures satisfying dimensionality and Fourier decay conditions, with uniformity in all the parameters involved.

Throughout this section, we liberate μ, α, β from their usual meaning, and we track dependencies on all parameters such as the dimension n. To facilitate our quoting of the literature, we first recall the functional equivalences in Tomas' T^*T argument [46, Chapter 7].

Fact B.1. *Suppose that $\mu \in \mathcal{M}^+(\mathbb{R}^n)$ and $p \in (1, +\infty]$, and that p' is given by $\frac{1}{p} + \frac{1}{p'} = 1$. Let $R > 0$. The following statements are equivalent:*

$$\|\widehat{f}\|_{L^2(\mathrm{d}\mu)} \leqslant R\|f\|_{L^{p'}(\mathbb{R}^n)} \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^n), \tag{B.1}$$

$$\|\widehat{g\mathrm{d}\mu}\|_{L^p(\mathbb{R}^n)} \leqslant R\|g\|_{L^2(\mathrm{d}\mu)} \qquad \forall\, g \in L^2(\mathrm{d}\mu). \tag{B.2}$$

We now fix two exponents $0 < \beta \leqslant \alpha \leqslant n$ and two constants $A, B \geqslant 1$, an we restrict our attention to probability measures μ on $\mathbb{R}^n$ satisfying

$$\mu\big[B(x,r)\big] \leqslant Ar^\alpha \qquad (x \in \mathbb{R}^n, r > 0), \tag{B.3}$$

$$|\widehat{\mu}(\xi)| \leqslant B(1 + |\xi|)^{-\beta/2} \qquad (\xi \in \mathbb{R}^n). \tag{B.4}$$

We define the critical exponent

$$p_0 = 2 + \frac{4(n-\alpha)}{\beta}, \tag{B.5}$$

so that the Mitsis-Mockenhaupt restriction theorem [33, 34] states that each of the inequalities in Fact B.1 holds for $p > p_0$, for a certain constant $R = R(A, B, \alpha, \beta, p, n)$. We wish to use (B.2) with $g \equiv 1$ and $p = 2 + \delta$ with a fixed small $\delta > 0$, which is possible when α is close enough to n by (B.5), but to be useful this requires some uniformity in α. The constants in [33, 34] can be given explicit expressions in terms of the parameters involved, and in fact one could likely adapt the version of Mockenhaupt's argument in [28, Proposition 4.1], to relax the condition $\beta > 2/3$ there to $\beta > 0$. We provide instead a direct derivation from the estimate of Bak and Seeger [1], which includes explicit constants.

Proposition B.2. *Let $\beta_0 \in (0, n)$. There exists $C_{n,\beta_0} > 0$ such that, when $\beta \geqslant \beta_0$, the estimate (B.1) holds for $p \geqslant p_0$ with $R = C_{n,\beta_0} \max(A, B)^{p_0/2p}$.*

Proof. Apply [1, Eq. (1.5)], replacing $a \leftarrow \alpha$, $b \leftarrow \beta/2$, $d \leftarrow n$, $p \leftarrow p'$, so that $q = \frac{2p}{p_0}$; and note that α, β belong to the compact interval $[\beta_0, n]$. Since $q \geqslant 2$ for $p \geqslant p_0$, by nesting of $L^s(\mathrm{d}\mu)$ norms this yields

$$\begin{aligned}
\|\widehat{f}\|_{L^2(\mathrm{d}\mu)} &\leqslant \|\widehat{f}\|_{L^q(\mathrm{d}\mu)} \\
&\leqslant (C_{n,\beta_0})^{\frac{2}{q}} A^{\frac{1}{q}\cdot\frac{2}{p_0}} B^{\frac{1}{q}\left(1-\frac{2}{p_0}\right)} \|f\|_{L^{p'}(\mathbb{R}^n)} \\
&\leqslant C_{n,\beta_0} \max(A, B)^{\frac{p_0}{2p}} \|f\|_{L^{p'}(\mathbb{R}^n)}.
\end{aligned}$$

□

Alternatively, one may choose to track down the dependencies on constants in Mitsis' simpler argument [33], which would lead to a similar estimate for the constant R in (B.1), upto a harmless (for our argument) factor $(p-p_0)^{-1}$. Via Proposition B.2, it is now possible to bound the moments of $\widehat{\mu}$ of order slightly larger than 2 when α is close enough to n, with only a moderate dependency of constants on α.

Proposition B.3. *Let $\delta \in (0,1)$ and $\beta_0 \in (0,n)$. Suppose that μ is a probability measure satisfying (3.1) and (3.2). Then, uniformly for $n - \frac{\delta\beta_0}{4} \leqslant \alpha < n$ and $\beta_0 \leqslant \beta < n$, we have*

$$\|\widehat{\mu}\|_{2+\delta} \lesssim_{\beta_0,n} D_\alpha^{1/2}.$$

Proof. We consider the exponent $p = 2+\delta$. Recalling (B.5), we have $p \geqslant p_0$ in the stated range of α. We can therefore invoke Proposition B.2 with $A = D \asymp 1$ and $B = D_\alpha$, so that the extension inequality (B.2) holds for $g \equiv 1$ with $R \lesssim_{\beta_0,n} D_\alpha^{1/2}$. □

References

1. J.-G. Bak and A. Seeger, *Extensions of the Stein-Tomas theorem*, Math. Res. Lett. **18** (2011), no. 4, 767–781.
2. R. C. Baker, *Diophantine inequalities*, The Clarendon Press, Oxford University Press, New York, 1986.
3. M. Bennett, A. Iosevich, and K. Taylor, *Finite chains inside thin subsets of $\mathbb{R}^d$*, to appear in Anal. PDE, http://arxiv.org/abs/1409.2581.
4. V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753.
5. V. Bergelson, A. Leibman, and E. Lesigne, *Intersective polynomials and the polynomial Szemerédi theorem*, Adv. Math. **219** (2008), no. 1, 369–388.
6. V. Bergelson and R. McCutcheon, *An ergodic IP polynomial Szemerédi theorem*, Mem. Amer. Math. Soc. **146** (2000), no. 695, viii+106.
7. C. Bluhm, *Random recursive construction of Salem sets*, Ark. Mat. **34** (1996), no. 1, 51–63.
8. ———, *On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets*, Ark. Mat. **36** (1998), no. 2, 307–316.
9. J. Bourgain, *A Szemerédi type theorem for sets of positive density in $\mathbf{R}^k$*, Israel J. Math. **54** (1986), no. 3, 307–316.
10. V. Chan, I. Łaba, and M. Pramanik, *Finite configurations in sparse sets*, to appear in J. Anal. Math., http://arxiv.org/abs/1307.1174.
11. X. Chen, *Sets of Salem type and sharpness of the L^2 Fourier restriction theorem*, to appear in Trans. Amer. Math. Soc., http://arxiv.org/abs/1305.5584.
12. M. B. Erdoğan, *A bilinear Fourier extension theorem and applications to the distance set problem*, Int. Math. Res. Not. (2005), no. 23, 1411–1425.
13. ———, *On Falconer's distance set conjecture*, Rev. Mat. Iberoam. **22** (2006), no. 2, 649–662.

14. H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.
15. H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979).
16. W. T. Gowers and J. Wolf, *The true complexity of a system of linear equations*, Proc. Lond. Math. Soc. (3) **100** (2010), no. 1, 155–176.
17. L. Grafakos, A. Greenleaf, A. Iosevich, and E. Palsson, *Multilinear generalized Radon transforms and point configurations*, Forum Math. **27** (2015), no. 4, 2323–2360.
18. B. Green and T. Tao, *New bounds for Szemerédi's theorem. II. A new bound for $r_4(N)$*, Analytic number theory, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.
19. A. Greenleaf and A. Iosevich, *On triangles determined by subsets of the Euclidean plane, the associated bilinear operators and applications to discrete geometry*, Anal. PDE **5** (2012), no. 2, 397–409.
20. A. Greenleaf, A. Iosevich, B. Liu, and E. Palsson, *A group-theoretic viewpoint on Erdős-Falconer problems and the Mattila integral*, Preprint (2014), http://arxiv.org/abs/1306.3598.
21. A. Greenleaf, A. Iosevich, and M. Pramanik, *On necklaces inside thin subsets of $\mathbb{R}^d$*, to appear in Math. Res. Lett., http://arxiv.org/abs/1409.2588.
22. K. Hambrooke, *Explicit Salem sets and applications to metrical Diophantine approximation*, Preprint (2014), http://www.math.rochester.edu/people/faculty/khambroo/research/index.html.
23. A. Iosevich and B. Liu, *Equilateral triangles in subsets of $\mathbb{R}^d$ of large Hausdorff dimension*, Preprint (2016), http://arxiv.org/abs/1603.01907.
24. J.-P. Kahane, *Some random series of functions*, second ed., Cambridge University Press, Cambridge, 1985.
25. R. Kaufman, *On the theorem of Jarník and Besicovitch*, Acta Arith. **39** (1981), no. 3, 265–267.
26. 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** (1998/99), no. 2, 843–844.
27. T. W. Körner, *Hausdorff and Fourier dimension*, Studia Math. **206** (2011), no. 1, 37–50.
28. I. Łaba and M. Pramanik, *Arithmetic progressions in sets of fractional dimension*, Geom. Funct. Anal. **19** (2009), no. 2, 429–456.
29. N. Lyall and A. Magyar, *Simultaneous polynomial recurrence*, Bull. Lond. Math. Soc. **43** (2011), no. 4, 765–785.
30. P. Maga, *Full dimensional sets without given patterns*, Real Anal. Exchange **36** (2010/11), no. 1, 79–90.
31. A. Mathé, *Sets of large dimension not containing polynomial configurations*, Preprint (2012), http://arxiv.org/abs/1201.0548.
32. P. Mattila and P. Sjölin, *Regularity of distance measures and sets*, Math. Nachr. **204** (1999), 157–162.
33. T. Mitsis, *A Stein-Tomas restriction theorem for general measures*, Publ. Math. Debrecen **60** (2002), no. 1-2, 89–99.
34. G. Mockenhaupt, *Salem sets and restriction properties of Fourier transforms*, Geom. Funct. Anal. **10** (2000), no. 6, 1579–1587.
35. L. Nicolaescu, *The coarea formula*, Expository note (2011), http://www3.nd.edu/~lnicolae/vita.html.
36. T. Orponen, *On the distance sets of AD-regular sets*, Preprint (2015), http://arxiv.org/abs/1509.06675.

37. K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.
38. ———, *On certain sets of integers. II*, J. London Math. Soc. **29** (1954), 20–26.
39. R. Salem, *On singular monotonic functions whose spectrum has a given Hausdorff dimension*, Ark. Mat. **1** (1951), 353–365.
40. A. Sárközy, *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), no. 1–2, 125–149.
41. P. Shmerkin, *Salem sets with no arithmetic progressions*, Preprint, http://arxiv.org/abs/1510.07596.
42. E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton, NJ, 1993.
43. T. Tao, *A proof of Roth's theorem*, Blog post (2014), https://terrytao.wordpress.com/2014/04/24/a-proof-of-roths-theorem/.
44. T. Tao and V. Vu, *Additive combinatorics*, Cambridge University Press, Cambridge, 2006.
45. T. Wolff, *Decay of circular means of Fourier transforms of measures*, Internat. Math. Res. Notices (1999), no. 10, 547–567.
46. T. H. Wolff, *Lectures on harmonic analysis*, American Mathematical Society, Providence, RI, 2003.

Kevin Henriot

Department of mathematics, University of British Columbia, Room 121, 1984 Mathematics Road, Vancouver BC V6T 1Z2, Canada

Email address: khenriot@math.ubc.ca

Izabella Łaba

Department of mathematics, University of British Columbia, Room 121, 1984 Mathematics Road, Vancouver BC V6T 1Z2, Canada

Email address: ilaba@math.ubc.ca

Malabika Pramanik

Department of mathematics, University of British Columbia, Room 121, 1984 Mathematics Road, Vancouver BC V6T 1Z2, Canada

Email address: malabika@math.ubc.ca