

# A Roth-type theorem for dense subsets of $\mathbb{R}^d$

Brian Cook, Ákos Magyar and Malabika Pramanik

## ABSTRACT

Let  $1 < p < \infty$ ,  $p \neq 2$ . We prove that if  $d \geq d_p$  is sufficiently large, and  $A \subseteq \mathbb{R}^d$  is a measurable set of positive upper density then there exists  $\lambda_0 = \lambda_0(A)$  such that for all  $\lambda \geq \lambda_0$  there are  $x, y \in \mathbb{R}^d$  such that  $\{x, x + y, x + 2y\} \subseteq A$  and  $\|y\|_p = \lambda$ , where  $\|y\|_p = (\sum_i |y_i|^p)^{1/p}$  is the  $l^p(\mathbb{R}^d)$ -norm of a point  $y = (y_1, \dots, y_d) \in \mathbb{R}^d$ . This means that dense subsets of  $\mathbb{R}^d$  contain 3-term progressions of all sufficiently large gaps when the gap size is measured in the  $l^p$ -metric. This statement is known to be false in the Euclidean  $l^2$ -metric as well as in the  $l^1$  and  $\ell^\infty$ -metrics. One of the goals of this note is to understand this phenomenon. A distinctive feature of the proof is the use of multilinear singular integral operators, widely studied in classical time-frequency analysis, in the estimation of forms counting configurations.

## 1. Introduction

Here we are concerned with certain density theorems in Ramsey theory over  $\mathbb{R}^d$ . A basic and representative result of the type we are interested in states that, with  $d \geq 2$ , a set  $A \subseteq \mathbb{R}^d$  of positive upper Banach density contains all large distances, that is, for every sufficiently large  $\lambda \geq \lambda_0(A)$  there are points  $x, x + y \in A$  such that  $\|y\|_2 = \lambda$ . Recall that the positive upper Banach density of  $A$  is defined as

$$\bar{\delta}(A) := \lim_{N \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \frac{|A \cap (x + [0, N]^d)|}{N^d}.$$

The quoted result was obtained independently, along with various generalizations, by a number of authors, for example Furstenberg, Katznelson and Weiss [5], Falconer and Marstrand [4], and Bourgain [1].

Alternatively, the above result shows that for any two-point configuration  $X \in \mathbb{R}^d$  we are guaranteed the existence, up to congruence, of *all sufficiently large* dilates of  $X$  inside of  $A$ . The term configuration simply refers to a finite point set. From this point of view it is a natural question to ask whether similar statements exist that involve configurations with a greater number of points. If one looks for some large dilate of a given configuration, such results are well known in the discrete regime of the integer lattice, under suitable assumptions of largeness on the underlying set. These results can often be translated to existence of configurations in the Euclidean setting as well. For instance, Roth's theorem [10] in the integers states that a subset of  $\mathbb{Z}$  of positive upper density contains a three-term arithmetic progression  $\{x, x + y, x + 2y\}$  and it easily implies that a measurable set  $A \subseteq \mathbb{R}$  of positive upper density contains a three-term progression whose gap size can be arbitrarily large. Results ensuring all sufficiently large dilates of a configuration in a set of positive Banach density are typically more difficult. Bourgain [1] shows that if  $X$  is any non-degenerate  $k$ -point simplex in  $\mathbb{R}^d$ ,  $d \geq k \geq 2$  (that is,  $X$  spans a

---

Received 24 June 2016; revised 27 January 2017; published online 29 May 2017.

2010 *Mathematics Subject Classification* 05D10, 42B20 (primary).

The first author is supported by NSF grant DMS1147523 and by the Fields Institute. The second author is supported by grant ERC-AdG. 321104 and by NSF grant DMS 1600840. The third is supported by an NSERC Discovery grant.

$(k-1)$ -dimensional space), then any subset of  $\mathbb{R}^d$  of positive upper Banach density contains a congruent copy of  $\lambda X$  for all sufficiently large  $\lambda$ .

On the other hand, a simple example given in [1] shows that there is a set  $A \subseteq \mathbb{R}^d$  in any dimension  $d \geq 1$ , such that the gap lengths of all 3-progressions in  $A$  do not contain all sufficiently large numbers. In other words, the result of [1] is false for the degenerate configuration  $X = \{0, e_1, 2e_1\}$ , where  $e_1$  is the canonical unit vector in the  $x_1$ -direction. More precisely, the counterexample provided in [1] is the set  $A$  of points  $x \in \mathbb{R}^d$  such that  $||x||_2^2 - m \leq \frac{1}{10}$  for some  $m \in \mathbb{N}$ . The parallelogram identity  $2||y||_2^2 = ||x||_2^2 + ||x+2y||_2^2 - 2||x+y||_2^2$  then dictates that  $||y||_2^2 - \frac{\ell}{2} \leq \frac{4}{10}$  (for some  $\ell \in \mathbb{N}$ ) for any progression  $\{x, x+y, x+2y\} \subseteq A$ . Thus the squares of the gap lengths are restricted to lie close to the half-integers, and therefore cannot realize all sufficiently large numbers.

It is interesting to observe that while Bourgain's counterexample prevents an existence theorem for three-term arithmetic progressions of all sufficiently large *Euclidean* gaps, it does not exclude the validity of such a result when the gaps are measured using some other metric on  $\mathbb{R}^d$  that does not obey the parallelogram law. The aim of this note is to prove that such results do indeed exist for the  $l^p$  metrics  $||y||_p := (\sum_{i=1}^d |y_i|^p)^{1/p}$  for all  $1 < p < \infty$ ,  $p \neq 2$ . In this sense, a counterexample as described above is more the exception rather than the rule.

Variations of our arguments also work for other metrics given by specific classes of positive homogeneous polynomials of degree at least 4 and those generated by symmetric convex bodies with special structure. Results of the first type were obtained in the finite field setting by the first two authors [2]. Also, the arguments here can be applied to obtain similar results for certain other degenerate point configurations. We hope to pursue these extensions elsewhere.

## 2. Main results

**THEOREM 2.1.** *Let  $1 < p < \infty$ ,  $p \neq 2$ . Then there exists a constant  $d_p \geq 2$  such that for  $d \geq d_p$  the following holds. Any measurable set  $A \subseteq \mathbb{R}^d$  of positive upper Banach density contains a three-term arithmetic progression  $\{x, x+y, x+2y\} \subseteq A$  with gap  $||y||_p = \lambda$  for all sufficiently large  $\lambda \geq \lambda(A)$ .*

**REMARKS.** (a) The result is sharp in the range of  $p$ . Easy variants of the example in [1] show that Theorem 2.1 and in fact even the two-point results of [1, 4, 5] cannot be true for  $p = 1$  and  $p = \infty$ . Indeed, if  $A = \mathbb{Z}^d + \epsilon_0[-1, 1]^d$  for some small  $\epsilon_0 > 0$ , then on one hand  $A$  is of positive upper Banach density. On the other hand, if  $x, x+y \in A$  for some  $y \neq 0$ , then both  $||y||_\infty$  and  $||y||_1$  are restricted to lie within distance  $O(\epsilon_0)$  from some positive integer.

(b) Indeed, counterexamples similar to [1] and the above can be constructed for norms given by a symmetric, convex body, a non-trivial part of whose boundary is either flat or coincides with an  $l^2$ -sphere. An appropriate formulation of a positive result for a general norm, and indeed the measuring failure of the parallelogram law for such norms, remains an interesting open question.

(c) We do not know whether the  $p$ -dependence of the dimensional threshold  $d_p$  stated in the theorem is an artifact of our proof. In our analysis,  $d_p$  grows without bound as  $p \nearrow \infty$ , while other implicit constants involved in the proof blow up near  $p = 1$  and  $p = 2$ . See, in particular, Proposition 2.2 and Lemma 4.3. It would be of interest to determine whether Theorem 2.1 holds for all  $d \geq 2$  for the specified values of  $p$ .

(d) Since three collinear points do not lie on an  $l^p$ -sphere for any  $p \in (1, \infty)$ , Theorem 2.1 shows that a result of the type considered by Graham in [7] is, in general, false for an  $l^p$ -sphere if  $p \neq 1, 2, \infty$ . Thus any connection between Ramsey-like properties and the notion of sphericity appears to be a purely  $l^2$  phenomenon.

### 2.1. Overview of proof

We describe below the main elements of the proof. Details will be provided in subsequent sections.

Our main observation is a stronger finitary version of Theorem 2.1 for bounded measurable sets.

**THEOREM 2.2.** *Let  $1 < p < \infty$ ,  $p \neq 2$ . Let  $d \geq d_p$ ,  $\delta > 0$  and let  $N \geq N(\delta)$  be sufficiently large. Then for any measurable set  $A \subseteq [0, N]^d$  of measure  $|A| \geq \delta N^d$  the following holds.*

*For any lacunary sequence  $1 < \lambda_1 < \dots < \lambda_J \ll N$  with  $\lambda_{j+1} \geq 2\lambda_j$  and  $J \geq J(\delta)$ , there exists a three-term arithmetic progression  $\{x, x+y, x+2y\} \subseteq A$  such that  $\|y\|_p = \lambda_j$  for some  $1 \leq j \leq J$ ,*

*Proof of Theorem 2.1.* Theorem 2.2 implies Theorem 2.1. Indeed, assume that Theorem 2.1 does not hold. Then there exists an infinite sequence  $\{\lambda_j\}_{j=1}^\infty \subseteq \mathbb{N}$  such that  $\lambda_j \neq \|y\|_p$  for any  $j$  and any  $y$  which is the gap of a 3-progression contained in  $A$ . Without loss of generality, the sequence may be assumed to be lacunary, that is,  $\lambda_{j+1} \geq 2\lambda_j$  for all  $j$ . After setting  $\delta = \overline{\delta}(A)/2$ , fix any sufficiently large  $J = J(\delta)$  and any sufficiently large box  $B_N$  of size  $N = N(\delta, \lambda_J)$  on which the density of  $A$  is  $|A|/N^d \geq \delta$ . By translation invariance we may assume  $B_N = [0, N]^d$ , contradicting Theorem 2.2.  $\square$

**2.1.1. A counting function and its variants.** For the rest of the paper we fix a finite exponent  $p > 1$ ,  $p \neq 2$ , and for simplicity of notation write  $|y| = \|y\|_p$ . We start by counting three-term arithmetic progressions  $P = \{x, x+y, x+2y\}$  contained in  $A$  via a positive measure  $\sigma_\lambda$  supported on the  $l^p$ -sphere  $S_\lambda = \{y \in \mathbb{R}^d; |y| = \lambda\}$ . Let  $f := \mathbf{1}_A$  be the indicator function of a measurable set  $A \subseteq [0, N]^d$ . As is standard in enumerating configurations, we introduce the counting function

$$\mathcal{N}_\lambda(f) := \int_{\mathbb{R}^d} \int_{S_\lambda} f(x)f(x+y)f(x+2y) d\sigma_\lambda(y) dx. \quad (2.1)$$

Clearly, if  $\mathcal{N}_\lambda(f) > 0$  then  $A$  must contain a 3-progression  $x, x+y, x+2y$  with  $|y| = \lambda$ . We will define the measure  $\sigma_\lambda$  via the oscillatory integral

$$\sigma_\lambda(y) := \lambda^{-d+p} \int_{\mathbb{R}} e^{it(|y|^p - \lambda^p)t} dt. \quad (2.2)$$

It is well known (see [8], Chapter 2) that the above oscillatory integral defines an absolutely continuous measure with respect to the surface area measure on  $S_\lambda$  whose density function is  $|\nabla Q(y)|^{-1}$  with  $Q(y) = |y|^p$ . The normalizing factor  $\lambda^{-d+p}$  is inserted to ensure that  $\sigma_\lambda(S_\lambda) = \sigma_1(S_1) > 0$ , which is independent of  $\lambda$ .

Let  $\psi$  be a Schwarz function such that  $0 \leq \psi \leq 1$ ,  $\psi(0) = 1$  and  $\widehat{\psi} \geq 0$  is compactly supported. Define the quantity

$$\omega_\lambda(y) := \lambda^{-d+p} \int_{\mathbb{R}} e^{it(|y|^p - \lambda^p)t} \psi(\lambda^p t) dt. \quad (2.3)$$

Note that by scaling

$$\omega_\lambda(y) = \lambda^{-d} \omega(y/\lambda) = \lambda^{-d} \widehat{\psi}(|y/\lambda|^p - 1), \quad (2.4)$$

hence  $\omega_\lambda$  is compactly supported on  $B(0; C\lambda)$  with

$$\int \omega_\lambda(y) dy = \int \omega(y) dy = C_\omega > 0.$$

When the subscript is omitted it should be assumed that  $\lambda = 1$ . Also define

$$\mathcal{M}_\lambda(f) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)f(x+2y)\omega_\lambda(y)dydx. \quad (2.5)$$

The first step is to show that this quantity is large.

**PROPOSITION 2.1.** *Let  $0 < \delta \leq 1$  and let  $A \subseteq [0, N]^d$  be such that  $|A| \geq \delta N^d$ . Then there exists a constant  $c(\delta) > 0$  depending only on  $\delta$  such that for  $0 < \lambda \ll N$ ,*

$$\mathcal{M}_\lambda(\mathbf{1}_A) \geq c(\delta) N^d. \quad (2.6)$$

As we will see in Section 3, the proof of this proposition is in essence Roth's theorem adapted to the Euclidean setting (see [1]).

Next, we define a variant of  $\mathcal{M}_\lambda$  indexed by a small  $\varepsilon > 0$  which is a good approximation to  $\mathcal{N}_\lambda(f)$ . Let

$$\omega_\lambda^\varepsilon(y) := \lambda^{-d+p} \int_{\mathbb{R}} e^{it(|y|^p - \lambda^p)} \psi(\varepsilon \lambda^p t) dt. \quad (2.7)$$

It is easy to see that

$$\omega_\lambda^\varepsilon(y) = \lambda^{-d} \varepsilon^{-1} \widehat{\psi} \left( \frac{|y/\lambda|^p - 1}{\varepsilon} \right) = \lambda^{-d} \omega^\varepsilon(y/\lambda). \quad (2.8)$$

Define

$$\mathcal{M}_\lambda^\varepsilon(f) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)f(x+y)f(x+2y)\omega_\lambda^\varepsilon(y)dx dy. \quad (2.9)$$

We establish the error estimate

**PROPOSITION 2.2.** *Let  $f : [0, N]^d \rightarrow [-1, 1]$  and let  $0 < \varepsilon < 1$ . Then there exist constants  $\gamma = \gamma_p > 0$  and  $C_{p,d} > 0$  both independent of  $\lambda$  such that for  $0 < \lambda \ll N$ ,*

$$|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\varepsilon(f)| \leq C_{p,d} \varepsilon^{d\gamma_p - 1} N^d. \quad (2.10)$$

*In particular,  $\gamma_p$  is independent of  $d$ . The constant  $C_{p,d} \nearrow \infty$  as  $p \rightarrow 1$  or  $2$ , while  $\gamma_p \rightarrow 0$  as  $p \nearrow \infty$ .*

The proof of Proposition 2.2 is based on two facts that may be of independent interest. The first is an inequality showing that the so-called  $U^3$ -uniformity norm of Gowers [6] controls expressions like  $\mathcal{N}_\lambda(f)$ . Let us recall the definition of the  $U^3$  norm for a compactly supported bounded measurable function  $g$ :

$$\|g\|_{U^3(\mathbb{R}^d)}^8 = \int_{(x,y) \in \mathbb{R}^d \times \mathbb{R}^{3d}} \left( \prod_{\nu \in \{0,1\}^3} \bar{g}^\nu(x + \nu_1 y_1 + \nu_2 y_2 + \nu_3 y_3) \right) dx dy, \quad (2.11)$$

where  $\nu_1, \nu_2, \nu_3$  can take the values 0 or 1,  $\bar{g}^\nu = \bar{g}$  if  $\nu_1 + \nu_2 + \nu_3$  is odd and  $\bar{g}^\nu = g$  otherwise.

**LEMMA 2.3.** *Let  $f : [0, N]^d \rightarrow [-1, 1]$  and let  $0 < \varepsilon < 1$ . Then for  $0 < \lambda \ll N$  one has*

$$|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\varepsilon(f)| \lesssim N^d \|\sigma - \omega^\varepsilon\|_{U^3}. \quad (2.12)$$

Although it is not *a priori* clear how to define the  $U^3$ -norm of the measure  $\sigma$  defined in (2.2), we note that  $\omega^\varepsilon \rightarrow \sigma$  weakly as  $\varepsilon \rightarrow 0$ . To prove (2.12), we first establish that  $\{\omega^\eta\}$  is a Cauchy sequence with respect to the  $U^3$ -norm and then define  $\|\sigma - \omega^\varepsilon\|_{U^3} := \lim_{\eta \rightarrow 0} \|\omega^\eta - \omega^\varepsilon\|_{U^3}$ .

We also show that

LEMMA 2.4. *Let  $0 < \varepsilon < 1$ ,  $p > 1$ ,  $p \neq 2$ . If  $d > 8r_p$  with  $r_p = \max(p+1, 2p-1)$ , then one has*

$$\|\sigma - \omega^\varepsilon\|_{U^3} \leq C_{p,d} \varepsilon^{d/8r_p-1}. \quad (2.13)$$

where the constant  $C_{p,d}$  has the same behaviour as described in Proposition 2.2.

Let us note in passing that (2.12) and (2.13) yield (2.11) with  $\gamma = 1/(8r_p)$ . The proof of Lemma 2.4 uses in an essential way that on  $\mathbb{R}^d$  the norm is defined by the expression  $|y|^p = \sum_i |y_i|^p$  for some  $p > 1$ ,  $p \neq 2$ . In particular, a pivotal role is played by the fact that, for  $p \neq 1, 2$  and  $x, x+y, x+2y \in \mathbb{R}_+^d$ ,

$$|x|^p + |x+2y|^p - 2|x+y|^p = \sum_{i=1}^d [x_i^p + (x_i+2y_i)^p - 2(x_i+y_i)^p]$$

does not vanish identically. It is worth remarking that this part of the argument fails both for the standard Euclidean norm and the  $l^1$ -norm. Indeed, estimate (2.13) does not hold for either the  $l^2$  or the  $l^1$ -metric.

2.1.2. *Multilinear Calderón–Zygmund singular integral operators.* The final ingredient in the proof of Theorem 2.2 is an estimate given in [9] for certain multilinear operators similar to the bilinear Hilbert transform. In order to describe the form in which we need this estimate, let us fix  $0 < \varepsilon \ll 1$  and define the constant  $c_1(\varepsilon)$  as follows,

$$c_1(\varepsilon) \int_{\mathbb{R}^d} \omega(y) dy = \int_{\mathbb{R}^d} \omega^\varepsilon(y) dy. \quad (2.14)$$

We see below in Lemma 4.1 that  $c_1(\varepsilon) \approx 1$  is bounded by two positive constants depending only on the dimension  $d$ . Write  $k^\varepsilon(y) := \omega_\lambda^\varepsilon(y) - c_1(\varepsilon) \omega_{\lambda_j}(y)$ , and

$$\mathcal{E}_\lambda(f) := \mathcal{M}_\lambda^\varepsilon(f) - c_1(\varepsilon) \mathcal{M}_\lambda(f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) f(x+y) f(x+2y) k^\varepsilon(y) dy dx, \quad (2.15)$$

so that by (2.14) one has the cancellation property

$$\int_{\mathbb{R}^d} k^\varepsilon(y) dy = \int_{\mathbb{R}^d} (\omega_\lambda^\varepsilon(y) - c_1(\varepsilon) \omega_\lambda(y)) dy = \int_{\mathbb{R}^d} (\omega^\varepsilon(y) - c_1(\varepsilon) \omega(y)) dy = 0. \quad (2.16)$$

The key estimate concerning the operator  $\mathcal{E}_\lambda$  is the following:

PROPOSITION 2.5. *Suppose that  $\{\lambda_j : 1 \leq j \leq J\}$  is a lacunary sequence (finite or infinite) with  $\lambda_{j+1} \geq 2\lambda_j$  for all  $j$ . Then for any  $f : [0, N]^d \rightarrow [-1, 1]$ ,*

$$\sum_{j=1}^J |\mathcal{E}_{\lambda_j}(f)|^2 \leq C_\epsilon N^d \|f\|_4^4 \leq C_\epsilon N^{2d},$$

where the constant  $C_\epsilon$  depends only on the quantity  $\epsilon$  used to define  $\mathcal{E}_\lambda$  and, in particular, is independent of  $f$  and the number  $J$  of elements in the lacunary sequence.

We provide details of this result in Section 5.

*Proof of Theorem 2.2.* Assuming Propositions 2.1, 2.2 and 2.5 for now, the proof proceeds by contradiction. Assume that there exist arbitrarily large  $N$ , a measurable set  $A \subseteq [0, N]^d$  with  $|A| \geq \delta N^d$ , and a sequence of non-admissible progression gaps  $\lambda_1 < \lambda_2 < \dots < \lambda_J \ll N$  for some  $J \geq J(\delta)$ , such that  $\mathcal{N}_{\lambda_j}(f) = 0$  for  $f = 1_A$ . The sequence may be chosen to be lacunary,

and  $J$  may be assumed to be arbitrarily large as well, by choosing  $N$  large enough. Thus, for  $1 \leq j \leq J$ ,

$$0 = \mathcal{N}_{\lambda_j}(f) = c_1(\epsilon)\mathcal{M}_{\lambda_j}(f) + [\mathcal{N}_{\lambda_j}(f) - \mathcal{M}_{\lambda_j}^\epsilon(f)] + \mathcal{E}_{\lambda_j}(f).$$

In view of this, Propositions 2.1, 2.2, and recalling that  $c_1(\epsilon) \approx 1$ , we find that for some sufficiently small  $\epsilon$  depending on  $p, d$  and  $\delta$ , the inequality

$$\frac{c(\delta)}{2}N^d \leq (c(\delta) - C_{p,d}\epsilon^{\gamma p-1})N^d \leq |c_1(\epsilon)\mathcal{M}_{\lambda_j}(f) + [\mathcal{N}_{\lambda_j}(f) - \mathcal{M}_{\lambda_j}^\epsilon(f)]| = |\mathcal{E}_{\lambda_j}(f)|$$

holds for every  $1 \leq j \leq J$ . Squaring both sides and summing over all  $j \leq J$  yields, after an application of Proposition 2.5 with  $f = 1_A$ ,

$$c(\delta)JN^{2d} \leq \sum_{j=1}^J |\mathcal{E}_{\lambda_j}(f)|^2 \leq C_\epsilon N^{2d}.$$

This implies that  $J \leq C_{p,d,\delta}$ , contradicting the hypothesis that  $J$  can be chosen arbitrarily large.  $\square$

### 3. The main term

We now set about proving the main propositions leading up to the theorem. In this section we prove Proposition 2.1 via an application of Roth's theorem on compact abelian groups (see [1], Appendix A, Theorem 3). The compact group of interest is of course the  $d$ -dimensional torus  $\Pi^d$ .

*Proof of Proposition 2.1.* Fix  $\delta \in (0, 1]$  and  $\lambda \ll N$  for a sufficiently large  $N$ , and also let  $A$  be given with the stated properties. Consider a real-valued function  $f : [-N, N]^d \rightarrow [0, 1]$  with  $\int f \geq \delta N^d$ . The proposition, as phrased, is proven from what follows with the choice  $f = \mathbf{1}_A$ .

Equipartition the cube  $[-N, N]^d$  in the natural way into disjoint boxes with sides parallel to the coordinate axes and length  $\ell = c\lambda$  by choosing a sufficiently small number  $c > 0$  so that  $N/\ell$  is a positive integer. Enumerate the boxes by  $\{B_i : 1 \leq i \leq L\}$ . Each box  $B_i$  may then be identified with its leftmost endpoint  $x_i$ , so that  $B_i = [0, \ell]^d + x_i$ .

Split  $f$  into pieces restricted to each box. More precisely, define  $g_i(x) : [0, \ell]^d \rightarrow [0, 1]$  by  $g_i(x) = \mathbf{1}_{[0, \ell]^d}(x + x_i)f(x + x_i)$ , so that  $f = \sum_{i=1}^L g_i$  on  $[-N, N]^d$ . The non-negativity of  $f$  implies the bound

$$\mathcal{M}_\lambda(f) \geq \sum_{i=1}^L \iint_{x, y \in \mathbb{R}^d} g_i(x)g_i(x+y)g_i(x+2y)\omega_\lambda(y)dydx. \quad (3.1)$$

Recall from (2.4) that  $\omega_\lambda(y) = \lambda^{-d} \hat{\psi}(|y/\lambda|^p - 1)$ ; hence we may choose  $\psi$  such that  $\hat{\psi}(|y/\lambda|^p - 1) \geq 1/10$  for  $y \in [-\ell, \ell]^d$ . Then (3.1) yields

$$\mathcal{M}_\lambda(f) \geq \frac{\lambda^{-d}}{10} \sum_{i=1}^L \iint_{x, y \in \mathbb{R}^d} g_i(x)g_i(x+y)g_i(x+2y)dydx. \quad (3.2)$$

Now identify  $\Pi^d$  with the cube  $[-\frac{1}{2}, \frac{1}{2}]^d$ . After a change of variable  $(x, y) \mapsto 10\ell(x, y)$ , each summand on the right hand side of (3.2) may be written as

$$\begin{aligned} (10\ell)^{2d} \iint_{x, y \in \mathbb{R}^d} g_i(10\ell x)g_i(10\ell(x+y))g_i(10\ell(x+2y))dydx \\ = (10\ell)^{2d} \iint_{x, y \in \Pi^d} g_i(10\ell x)g_i(10\ell(x+y))g_i(10\ell(x+2y))dydx. \end{aligned}$$

Note that the support assumptions on the  $g_i$  dictate that the integrand is supported on  $[-\frac{1}{10}, \frac{1}{10}]^{2d} \subset \Pi^{2d}$ , as indicated in the last step. If we also have that

$$\int_{\Pi^d} g_i(10\ell x) dx \geq \eta > 0 \quad (3.3)$$

for some index  $i$  then Roth's theorem on compact abelian groups would imply that for such an index

$$\iint_{\Pi^d \times \Pi^d} g_i(10\ell x) g_i(10\ell(x+y)) g_i(10\ell(x+2y)) dx dy \geq c_0(d, \eta),$$

where  $c_0(d, \eta) > 0$  is a constant uniform in  $i$  but depending only on  $d$  and  $\eta$ . We prove below in Lemma 3.1 that (3.3) holds with  $\eta = \delta(10)^{-d}/2$  for at least  $\delta L/2$  indices  $i$ . Summing over all these indices in (3.2) then leads to the bound

$$\mathcal{M}_\lambda(f) \geq c_0(d, \eta) \lambda^{-d} \frac{\delta}{2} L(10\ell)^{2d} = c(d, \delta),$$

as claimed.  $\square$

LEMMA 3.1. *The relation (3.3) holds with  $\eta = \delta(10)^{-d}/2$  for at least  $\delta L/2$  indices  $i$ .*

*Proof.* This is a simple pigeonholing argument. Let  $I$  denote the number of indices  $i$  for which the integral inequality in (3.3) holds. After a scaling change of variable, this is the same set of indices  $i$  for which  $\int_{[0, \ell]^d} g_i(x) dx \geq (\delta/2)\ell^d$ . By our hypothesis on  $f$ ,

$$\int_{\mathbb{R}^d} f(x) dx = \sum_{i=1}^L \int g_i(x) dx \geq \delta N^d. \quad (3.4)$$

On the other hand,  $0 \leq g_i \leq 1$  is supported on  $[0, \ell]^d$ , so  $\|g_i\|_1 \leq \ell^d$  for trivial reasons. This leads to the estimate

$$\sum_{i=1}^L \int_{\mathbb{R}^d} g_i(x) dx \leq (\delta/2)N^d + I\ell^d. \quad (3.5)$$

for each  $1 \leq i \leq L$ . Combining the lower bound in (3.4) with the upper bound in (3.5) and recalling that  $L = N^d/\ell^d$  leads to the claimed statement.  $\square$

#### 4. Error estimates

As indicated in the introduction, our main objective here is to prove Proposition 2.2, which is a direct consequence of Lemma 2.3 and 2.4. Before turning our attention to the proof of these lemmas, let us first make the simple but important observation that  $\int \omega^\varepsilon(y) dy \approx 1$  uniformly for sufficiently small  $\varepsilon > 0$ . Let  $\nu_p := |\{y; |y|_p \leq 1\}|$ , then by homogeneity  $|\{y; |y|_p \leq \eta\}| = \nu_p \eta^d$ . This fact was used in the proof of Theorem 2.2 in order to bound from below the main term  $c_1(\varepsilon)\mathcal{M}_{\lambda_j}(f)$ .

LEMMA 4.1. *There exist constants  $0 < c_1 < c_2$  depending only on the function  $\psi$  and the parameter  $d$  and  $p$  such that*

$$c_1 \leq \int_{\mathbb{R}^d} \omega^\varepsilon(y) dy \leq c_2, \quad (4.1)$$

uniformly for  $0 < \varepsilon < \frac{1}{10d}$ .

*Proof.* By the definition of the function  $\psi$  we have  $|\hat{\psi}(t)| \geq c$  for  $|t| \leq \tau$  for some constants  $c > 0$  and  $0 < \tau \leq 1$ . Then by (2.7) one has

$$\int \omega_1^\varepsilon(y) dy \geq \frac{c}{\varepsilon} |\{y; 1 - \tau\varepsilon \leq |y|^p \leq 1 + \tau\varepsilon\}| = \frac{c\nu_p}{\varepsilon} ((1 + \tau\varepsilon)^{d/p} - (1 - \tau\varepsilon)^{d/p}) \geq c_1,$$

uniformly for  $0 < \varepsilon \leq 1/10d$  as  $p > 1$ . Similarly, as  $0 \leq \hat{\psi} \leq 1$  and  $\hat{\psi}$  is supported on  $(-2, 2)$ ,

$$\int \omega^\varepsilon(y) dy \leq \frac{1}{\varepsilon} |\{y; 1 - 2\varepsilon \leq |y|^p \leq 1 + 2\varepsilon\}| = \frac{\nu_p}{\varepsilon} ((1 + 2\varepsilon)^{d/p} - (1 - 2\varepsilon)^{d/p}) \leq c_2. \quad \square$$

Since  $\omega^\varepsilon(y)$  is invariant under reflections to the coordinate hyperplanes, we have that

$$\int \chi_+(y) \omega^\varepsilon(y) dy = 2^{-d} \int \omega^\varepsilon(y) dy,$$

and the above lemma holds for  $\chi_+(y) \omega^\varepsilon(y)$  as well.

To prove Lemma 2.3 we need the following result.

LEMMA 4.2. Let  $f : [-N, N]^d \rightarrow [-1, 1]$  and  $g : [0, \lambda]^d \rightarrow [-1, 1]$  be given functions. Then

$$\int_{x, y \in \mathbb{R}^d} f(x) f(x+y) f(x+2y) g(y) dx dy \lesssim N^d \lambda^{d/2} \|g\|_{U^3}, \quad (4.2)$$

where the implicit constant depends only on  $d$ .

*Proof.* The proof involves several changes of variables and successive applications of the Cauchy–Schwarz inequality. Set

$$T = \int_{x, y \in \mathbb{R}^d} f(x) f(x+y) f(x+2y) g(y) dx dy.$$

Applying the Cauchy–Schwarz inequality in the  $x$  integration to get

$$T^2 \leq N^d \int_x \int_y \int_{y'} f(x+y) f(x+2y) f(x+y') f(x+2y') g(y) g(y') dx dy dy'.$$

Use the substitution  $y' = y + h$  followed by the substitution  $x \rightarrow x - y$ , and define

$$\Delta_h F(x) = F(x+h) \overline{F(x)} \quad (4.3)$$

for a generic complex valued function  $F$ . Then one may write

$$T^2 \leq N^d \int_x \int_y \int_h \Delta_h f(x) \Delta_{2h} f(x+y) \Delta_h g(y) dx dy dh.$$

The integrals in  $y, h$  may be restricted to a region with  $|y|, |h| \lesssim \lambda$  due to the support of  $g$ . Then another application of Cauchy–Schwarz in the  $x$  and  $h$  integration gives

$$T^4 \lesssim N^{3d} \lambda^d \int_{x, h, y, y'} \Delta_{2h} f(x+y) \Delta_{2h} f(x+y') \Delta_h g(y) \Delta_h g(y') dx dy dh dy'.$$

Again use the substitutions  $y' = y + k$  and  $x \rightarrow x - y$  in turn to get

$$T^4 \lesssim N^{3d} \lambda^d \int_{x, y, k, h} \Delta_{2h} f(x) \Delta_{2h} f(x+k) \Delta_h g(y) \Delta_h g(y+k) dx dy dh dk.$$

One final application of the Cauchy–Schwarz inequality in  $x$  and  $h$  and  $k$  integration gives

$$T^8 \lesssim N^{7d} \lambda^{4d} \int_{x, y, h, k, y'} 1_{[0, N]^d}(x) \Delta_h g(y) \Delta_h g(y') \Delta_h g(y+k) \Delta_h g(y'+k) dx dy dy' dh dk.$$

The  $x$  integration may be carried out, and applying the substitution  $y' \rightarrow y + l$  gives the final form

$$T^8 \lesssim N^{8d} \lambda^{4d} \int_{y,h,k,l} \Delta_{h,k,l} f(y) dy dh dk dl, \quad (4.4)$$

where  $\Delta_{h,k,l}$  is well defined as the composition of the operators  $\Delta_h$ ,  $\Delta_k$ , and  $\Delta_l$ . The integral is easily verified to be  $\|g\|_{U^3}^8$ , which completes the proof.  $\square$

*Proof of Lemma 2.3.* As indicated in the introduction, the right-hand side of (2.12) is to be interpreted as

$$\|(\sigma - \omega^\epsilon)\|_{U^3} := \lim_{\eta \rightarrow 0} \|(\omega^\eta - \omega^\epsilon)\|_{U^3}.$$

Since the integral representation of

$$\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\epsilon(f) = \lim_{\eta \rightarrow 0} \mathcal{M}_\lambda^\eta(f) - \mathcal{M}_\lambda^\epsilon(f)$$

is of the form (4.2), we may apply Lemma 4.2 with  $g(y) = (\omega_\lambda^\eta(y) - \omega_\lambda^\epsilon(y))$  to get

$$|\mathcal{N}_\lambda(f) - \mathcal{M}_\lambda^\epsilon(f)| \lesssim N^d \lambda^{d/2} \lim_{\eta \rightarrow 0} \|\chi_+(\omega_\lambda^\eta - \omega_\lambda^\epsilon)\|_{U^3}.$$

Since  $\omega_\lambda^\epsilon$  is a rescaled version of  $\omega^\epsilon := \omega_1^\epsilon$ , scaling properties of the  $U^3$  norm imply that

$$\|(\omega_\lambda^\eta - \omega_\lambda^\epsilon)\|_{U^3} = \lambda^{-d} \lambda^{d/2} \|(\omega^\eta - \omega^\epsilon)\|_{U^3},$$

which leads to the claimed upper bound.  $\square$

Next we turn to the proof of Lemma 2.4. In what follows we assume that  $\lambda$  and  $N \gg \lambda$  are fixed, and  $f$  is the characteristic function of a set  $A \subset [-N, N]^d$  with measure  $\delta N^d$ . First we need an estimate for one-dimensional scalar oscillatory integrals of the following type.

**LEMMA 4.3.** *Let  $1 < p < \infty$ ,  $p \neq 2$ . For a smooth cut-off function  $\phi$  on  $\mathbb{R}$ , let  $\phi_+$  be its restriction to the positive real numbers and define the integral*

$$I(t) := \int_{y,h,k,l \in \mathbb{R}} \Delta_{h,k,l}(\phi_+(y) e^{it|y|^p}) dy dh dk dl, \quad (4.5)$$

where  $\Delta_{h,k,l}$  is defined via iterated compositions of  $\Delta_h$  as described in (4.3) and the proof of Lemma 4.2. Then there exists a constant  $r = r(p) > 0$  such that

$$|I(t)| \leq C_p |t|^{-1/r}, \quad \text{for } |t| \geq 1. \quad (4.6)$$

One may take  $r(p) = p + 1$  for  $1 < p < 2$ , and  $r(p) = 2p - 1$  for  $p > 2$ . The constant  $C_p$  is finite in the indicated range of  $p$ , and tends to infinity as  $p \rightarrow 1$  or  $2$ .

*Proof.* Replacing  $y + h$  by a new variable  $y'$ , the integral  $I(t)$  may be rewritten as

$$I(t) = \int |I_{k,l}(t)|^2 dk dl,$$

where

$$I_{k,l}(t) = \int \Delta_{k,l} \phi_+(y) e^{it\psi_{k,l}(y)} dy,$$

with

$$\Delta_{k,l}\phi_+(y) = \phi_+(y)\phi_+(y+k)\phi_+(y+l)\phi_+(y+k+l),$$

$$\psi_{k,l}(y) = y^p + (y+k+l)^p - (y+k)^p - (y+l)^p.$$

The reason we can write  $\psi_{k,l}(y)$  in this form is that  $y, y+k, y+l, y+k+l$  are all positive on the support of  $\Delta_{k,l}\phi_+$ .

It is clear that  $I_{k,l}(\eta)$  is uniformly bounded, hence  $I(t)$  receives small contribution from regions where at least one of the integration variables  $k, l$  is small. For a small parameter  $0 < \eta < 1$  to be chosen later, we may therefore write

$$I(t) = \int_{|k|, |l| \geq \eta} |I_{k,l}(t)|^2 dk dl + O(\eta).$$

We now estimate the integral  $I_{k,l}(t)$  for fixed  $k, l$  assuming  $|k|, |l| \geq \eta$ . Introducing a smooth partition of unity, we have  $I_{k,l}(t) = I'_{k,l}(t) + J_{k,l}(t)$ , where the domain of integration of  $I'_{k,l}(t)$  ranges over those  $y$  for which at least one of the quantities  $y, y+k, y+l, y+k+l$  is  $O(\eta)$ . Thus,  $I'_{k,l}(t) = O(\eta)$ .

For  $J_{k,l}(t)$  one may write, using Taylor's remainder formula

$$\psi_{k,l}(y) = klp(p-1) \int_{[0,1]^2} (y+uk+sl)^{p-2} du ds.$$

Therefore, its derivative is given by

$$\psi'_{k,l}(y) = klp(p-1)(p-2) \int_{[0,1]^2} (y+uk+sl)^{p-3} du ds.$$

By our assumptions, we have that

$$\eta \lesssim y+uk+sl \lesssim 1,$$

uniformly for  $0 \leq u, s \leq 1$ , and also that  $|k|, |l| \gtrsim \eta$ . Thus for  $p > 1, p \neq 2$ ,

$$|\psi'_{k,l}(y)| \gtrsim \eta^{p-1},$$

with an implicit constant independent of  $k$  and  $l$ . Then, writing  $\psi = \psi_{k,l}$  and  $\chi$  for the amplitude, integration by parts yields

$$\begin{aligned} J_{k,l}(t) &= \int_{\mathbb{R}^d} \frac{d}{dy} \left( e^{it\psi(y)} \right) \frac{\chi(y)}{it\psi'(y)} dy = -\frac{1}{it} \int_{\mathbb{R}^d} e^{it\psi(y)} \frac{d}{dy} \left( \frac{\chi(y)}{\psi'(y)} \right) dy \\ &= \frac{1}{it} \int_{\mathbb{R}^d} e^{it\psi(y)} \left( \frac{\chi'}{\psi'} + \frac{\chi\psi''}{\psi'^2} \right) dy. \end{aligned}$$

Here we have used the support properties of  $\chi$  in the form  $\|\chi\|_\infty = O(1)$  and  $\|\chi'\|_\infty = O(\eta^{-1})$ . Therefore,

$$|J_{k,l}(t)| \lesssim |t|^{-1} (\eta^{-p} + \eta^{-2p+2}) \leq |t|^{-1} \eta^{-r'(p)},$$

with  $r'(p) = \max(p, 2p-2) > 0$ . This implies that  $I_{k,l}(t) = O(\eta) + O(|t|^{-1} \eta^{-r'_p})$ , choosing  $\eta := |t|^{-1/(r_p+1)}$  yields

$$|I(t)| \lesssim |t|^{-1/r_p}, \quad (4.7)$$

with  $r(p) = r'(p) + 1 = \max(p+1, 2p-1)$  and the lemma follows.  $\square$

*Proof of Lemma 2.2.* By the triangle inequality

$$\|\sigma - \omega^\varepsilon\|_{U^3(\mathbb{R}^d)} \leq \sum_{i=1}^{2^d} \|\chi_i(\sigma - \omega^\varepsilon)\|_{U^3(\mathbb{R}^d)},$$

where  $\chi_i$  is the indicator function of the  $i$ th orthant of  $\mathbb{R}^d$ . Since both  $\sigma$  and  $\omega^\varepsilon$  are invariant under reflections about the coordinate hyperplanes, it suffices to estimate  $\|\chi_+(\sigma - \omega^\varepsilon)\|_{U^3}$ , where  $\chi_+$  is the indicator function of the positive orthant. Recalling that  $\omega^\eta - \omega^\varepsilon$  is compactly supported, say on  $[-C, C]^d$ , let us fix a smooth cutoff function  $\Phi(y) = \phi^{\otimes d}(y)$ , where  $\phi$  is a smooth bump function supported on an interval of the form  $[-2C, 2C]$  and identically 1 on the middle half of it. Then,

$$\|\chi_+(\omega^\eta - \omega^\varepsilon)\|_{U^3} = \left\| \Phi_+(y) \int_t (\psi(\eta t)) - \psi(\varepsilon t) e^{i(|y|^p-1)t} dt \right\|_{U^3(dy)}, \quad (4.8)$$

with  $\Phi_+(y) = \chi_+(y)\Phi(y)$ . Applying Minkowski's inequality to the right-hand side of the last equation (4.8) this is further estimated by

$$\int_t |\psi(\eta t) - \psi(\varepsilon t)| \left\| \Phi_+(y) e^{it|y|^p} \right\|_{U^3(\mathbb{R}^d)} dt. \quad (4.9)$$

Note that as  $\Phi_+(y) e^{it|y|^p} = \prod_{i=1}^d \phi_+(y_i) e^{it y_i^p}$ , we have

$$\left\| \Phi_+(y) e^{it|y|^p} \right\|_{U^3(\mathbb{R}^d)} = \left\| \phi_+(y) e^{it y^p} \right\|_{U^3(\mathbb{R})}^d,$$

where the one-dimensional integrals

$$\left\| \phi_+(y) e^{it y^p} \right\|_{U^3(\mathbb{R})}^8 = \int_{y,h,k,l \in \mathbb{R}} \left( \Delta_{h,k,l} \phi_+(y) e^{it|y|^p} \right) dy dh dk dl$$

are estimated in Lemma 4.3. Thus, we have for  $|t| \geq 1$

$$\left\| \Phi_+(y) e^{it|y|^p} \right\|_{U^3(y)} \lesssim |t|^{-d/8r},$$

with  $r = r(p)$  given in (4.6). Inserting this bound into (4.9), we complete the estimation as follows:

$$\begin{aligned} \int_t |\psi(\eta t) - \psi(\varepsilon t)| |t|^{-d/8r} dt &\leq \int_t [|\psi(\eta t)| + |\psi(\varepsilon t)|] |t|^{-d/8r} dt \\ &\lesssim \eta^{d/8r-1} + \varepsilon^{d/8r-1} \lesssim \epsilon^{d/8r-1}, \end{aligned}$$

which is bounded uniformly in  $\eta$  provided that  $\eta \ll \epsilon$  and  $d > 8r$ .  $\square$

## 5. A result from time-frequency analysis

Here we will prove Proposition 2.5 by using the main result of [9]. The necessary verifications of the hypotheses of [9] will be done subsequently.

*Proof of Proposition 2.5.* By the Cauchy–Schwarz inequality and support restrictions on  $f$ ,

$$\begin{aligned} |\mathcal{E}_{\lambda_j}(f)|^2 &\lesssim N^d \iiint_{(\mathbb{R}^d)^3} f(x+y)f(x+z)f(x+2y)f(x+2z) k_j^\varepsilon(y) k_j^\varepsilon(z) dy dz dx \\ &= N^d \iiint_{(\mathbb{R}^d)^3} f(x)f(x+z-y)f(x+y)f(x+2z-y) k_j^\varepsilon(y) k_j^\varepsilon(z) dy dz dx. \end{aligned} \quad (5.1)$$

Summing (5.1) for  $1 \leq j \leq J$  one has

$$\sum_{j=1}^J |\mathcal{E}_{\lambda_j}(f)|^2 \lesssim N^d \int_{\mathbb{R}^d} f(x) \mathcal{K}_J^\varepsilon(f, f, f)(x) dx, \quad (5.2)$$

where the trilinear operator  $\mathcal{K}_J^\varepsilon$  on the right side is given by

$$\mathcal{K}_J^\varepsilon(f_1, f_2, f_3)(x) := \int \int_{(\mathbb{R}^d)^2} f_1(x+z-y) f_2(x+y) f_3(x+2z-y) K_J^\varepsilon(y, z) dy dz, \quad (5.3)$$

with integration kernel

$$K_J^\varepsilon(y, z) = \sum_{j=1}^J k_j^\varepsilon(y) k_j^\varepsilon(z), \quad \text{where} \quad k_j^\varepsilon(y) = \omega_{\lambda_j}(y) - c_1(\varepsilon) \omega_{\lambda_j}(y). \quad (5.4)$$

We justify in Lemma 5.1 below that this is a Calderón–Zygmund kernel. Lebesgue mapping properties of associated multilinear operators of the type given in (5.3) were studied by Muscalu, Tao and Thiele [9]. Indeed, by Theorem 1.1 in [9] one has the estimate

$$\|\mathcal{K}_J^\varepsilon(f, f, f)\|_{L^{4/3}(\mathbb{R}^d)} \leq C_\varepsilon \|f\|_{L^4(\mathbb{R}^d)}^3, \quad (5.5)$$

with a constant  $C_\varepsilon > 0$  independent of  $J$  and the sequence  $\lambda_1 < \dots < \lambda_J$ . Then

$$\begin{aligned} \sum_{j=1}^J |\mathcal{E}_{\lambda_j}(f)|^2 &\lesssim N^d \int_{\mathbb{R}^d} f(x) \mathcal{K}_J^\varepsilon(f, f, f)(x) dx \\ &\leq C_\varepsilon N^d \|\mathcal{K}_J^\varepsilon(f, f, f)\|_{L^{4/3}(\mathbb{R}^d)} \|f\|_{L^4(\mathbb{R}^d)} \leq C_\varepsilon N^d \|f\|_4^4, \end{aligned} \quad (5.6)$$

as claimed.  $\square$

In preparation for Lemma 5.1, let us begin by rewriting (5.6) as

$$\mathcal{K}_J^\varepsilon(f_1, f_2, f_3)(x) := \iiint_{(\mathbb{R}^d)^3} e^{ix \cdot (\xi_1 + \xi_2 + \xi_3)} \widehat{f}_1(\xi_1) \widehat{f}_2(\xi_2) \widehat{f}_3(\xi_3) m(\xi_1, \xi_2, \xi_3) d\xi_1 d\xi_2 d\xi_3, \quad (5.7)$$

where

$$\begin{aligned} m(\xi_1, \xi_2, \xi_3) &= \iint_{(\mathbb{R}^d)^2} K_J^\varepsilon(y, z) e^{-iy \cdot (\xi_1 - \xi_2 + \xi_3)} e^{ix \cdot (\xi_1 + 2\xi_3)} dy dz \\ &= \widehat{K_J^\varepsilon}(-\xi_1 + \xi_2 - \xi_3, \xi_1 + 2\xi_3). \end{aligned} \quad (5.8)$$

Set

$$\Gamma = \{(\xi_1, \xi_2, \xi_3, \xi_4) \in (\mathbb{R}^d)^4 : \xi_1 + \xi_2 + \xi_3 + \xi_4 = 0\}, \quad \text{and} \quad (5.9)$$

$$\Gamma' = \{(\xi_1, \xi_2, \xi_3, \xi_4) \in \Gamma : \xi_1 - \xi_2 + \xi_3 = 0, \xi_1 + 2\xi_3 = 0\}, \quad (5.10)$$

so that  $\dim(\Gamma) = 3d$  and  $\dim(\Gamma') = d$ . By taking the Fourier transform one may re-express  $\mathcal{K}_J^\varepsilon$  as a multiplier, namely,

$$\widehat{\mathcal{K}_J^\varepsilon}(f_1, f_2, f_3)(-\xi_4) := \int_{\Gamma} \widehat{f}_1(\xi_1) \widehat{f}_2(\xi_2) \widehat{f}_3(\xi_3) m(\xi_1, \xi_2, \xi_3, \xi_4) d\xi_1 d\xi_2 d\xi_3, \quad (5.11)$$

identifying the operator  $\mathcal{K}_J^\varepsilon$  with the ones studied in [9]. The only difference is that the functions  $f_i$  are now defined on  $\mathbb{R}^d$  instead of on  $\mathbb{R}$ ; however, this does not affect the arguments given there. To clarify,  $\Gamma'$  is a graph over  $d$  of the canonical variables, and  $\mathcal{K}_J^\varepsilon$  generates a 4-linear form, so in the notation of [3],  $k = d$  and  $n = 4$ . Thus, the rank of the operator as described in this paper is  $m = k/d = 1$ , an integer  $< 2 = n/2$ . As has been pointed out in page

295 of [3], the higher-dimensional adaptation of the main result of [9] in this setting of integral rank is fairly straightforward, provided certain requirements on the multiplier is met. Thus, in order to apply the main result of [9], one needs to establish certain growth and differentiability properties of the multiplier  $m(\xi)$ . This is the goal of the following lemma.

LEMMA 5.1. *Given an integration kernel  $K_J$  as in (5.4), with the summands obeying the cancellation condition (2.16), let  $m(\xi)$  be the associated multiplier defined in (5.8), where  $\xi = (\xi_1, \dots, \xi_3) \in \mathbb{R}^{3d}$  is a coordinate system identifying the subspace  $\Gamma$ . Then for any multi-index  $\alpha \in \mathbb{Z}_+^{3d}$  one has the estimate*

$$|\partial_\xi^\alpha m(\xi)| \leq C_{\alpha, \varepsilon} (\text{dist}(\xi, \Gamma'))^{-|\alpha|}. \quad (5.12)$$

Here  $\Gamma$  and  $\Gamma'$  are as in (5.9) and (5.10).

REMARK. The crucial point here is that the constant  $C_{\alpha, \varepsilon}$  is independent of  $J$  and the lacunary sequence  $\lambda_1 < \dots < \lambda_J$ . Also, the above estimate is needed just up to some fixed finite order. Once this is established, our main result Theorem 2.2 follows as explained at the end of Section 2.

*Proof.* Since  $m$  is essentially  $\widehat{K}_J^\varepsilon$  composed with a linear transformation, we study the latter function in detail. The relation (5.4) implies that

$$\widehat{K}_J^\varepsilon(\eta, \zeta) = \sum_{j=1}^J \widehat{k}_j^\varepsilon(\eta) \widehat{k}_j^\varepsilon(\zeta)$$

and

$$\widehat{k}_j^\varepsilon(\eta) = \int_{y \in \mathbb{R}^d} e^{iy \cdot \eta} (\omega_{\lambda_j}^\varepsilon(y) - c_1(\varepsilon) \omega_{\lambda_j}(y)) dy = \widehat{\omega}^\varepsilon(\lambda_j \eta) - c_1(\varepsilon) \widehat{\omega}(\lambda_j \eta).$$

Let us recall the definitions  $\omega_1^\varepsilon(y) = \varepsilon^{-1} \widehat{\psi}(|y|^p - 1)/\varepsilon$ , where  $\omega(y) = \omega_1^1(y)$  with  $\widehat{\psi}$  is a compactly supported smooth function. Therefore for all multi-indices  $\alpha$

$$\left| \partial_\eta^\alpha \widehat{\omega}_1^\varepsilon(\eta) \right| \leq C_\alpha. \quad (5.13)$$

Integrating by parts  $k$  times in the integral expression for  $\partial_\eta^\alpha \omega_1^\varepsilon$  we also obtain

$$\left| \partial_\eta^\alpha \widehat{\omega}_1^\varepsilon(\eta) \right| \leq C_{\alpha, k} \varepsilon^{-|\alpha|} (1 + |\eta|)^{-k}. \quad (5.14)$$

Thus for any  $k \in \mathbb{N}$

$$\left| \partial_\eta^\alpha \widehat{k}_j^\varepsilon(\eta) \right| \leq C_{\alpha, k} \varepsilon^{-|\alpha|} (1 + |\lambda_j \eta|)^{-k}. \quad (5.15)$$

The cancellation property (2.16) gives that  $\widehat{k}_j^\varepsilon(0) = 0$ , which leads to the additional pointwise estimate

$$\left| \widehat{k}_j^\varepsilon(\eta) \right| \leq C \lambda_j |\eta|. \quad (5.16)$$

This implies that  $\widehat{K}_J^\varepsilon(0, \zeta) = \widehat{K}_J^\varepsilon(\eta, 0) = 0$  and

$$\left| \widehat{K}_J^\varepsilon(\eta, \zeta) \right| \lesssim \sum_{j \leq J} \min \left( \lambda_j |\eta|, \frac{1}{\lambda_j |\eta|} \right) \lesssim 1 \quad (5.17)$$

as the sequence  $\mu_j := \lambda_j |\eta|$  is lacunary and  $\|\widehat{k}_j^\varepsilon\|_\infty \lesssim 1$ . This shows that the multiplier  $m(\xi)$  defined in (5.11) is bounded. To estimate its partial derivatives we apply (5.15) with

$k = |\alpha| + |\beta| + 1$  to write

$$\begin{aligned} \left| \partial_\eta^\alpha \partial_\zeta^\beta \widehat{K}_J^\varepsilon(\eta, \zeta) \right| &\leq C_{\alpha, \beta, \varepsilon} \sum_{j \leq J} \lambda_j^{|\alpha| + |\beta|} \min \left( (1 + |\lambda_j \eta|)^{-|\alpha| - |\beta| - 1}, (1 + |\lambda_j \zeta|)^{-|\alpha| - |\beta| - 1} \right) \\ &\leq C_{\alpha, \beta, \varepsilon} \min \left( |\eta|^{-|\alpha| - |\beta|}, |\zeta|^{-|\alpha| - |\beta|} \right) \leq C_{\alpha, \beta, \varepsilon} (|\eta| + |\zeta|)^{-|\alpha| - |\beta|}. \end{aligned}$$

Here we have used the fact that

$$\sum_{j \leq J} \mu_j^k (1 + \mu_j)^{-k-1} \leq \sum_{j: \mu_j \leq 1} \mu_j + \sum_{j: \mu_j \geq 1} \mu_j^{-1} \leq C,$$

for the lacunary sequences  $\mu_j := \lambda_j |\eta|$  and  $\mu_j := \lambda_j |\zeta|$  with  $k = |\alpha| + |\beta| \geq 1$ . By (5.8) this leads to the estimate

$$|\partial_\xi^\alpha m(\xi)| \leq C_{\alpha, \varepsilon} \left( |-\xi_1 + \xi_2 - \xi_3|^{-|\alpha|} + |\xi_1 + 2\xi_3|^{-|\alpha|} \right) \leq C'_{\alpha, \varepsilon} \text{dist}(\xi, \Gamma')^{-|\alpha|}.$$

for  $\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^{3d} \simeq \Gamma$  and any multi-index  $\alpha \in \mathbb{Z}_+^{3d}$ . This proves Lemma 5.1.  $\square$

### References

1. J. BOURGAIN, ‘A Szemerédi type theorem for sets of positive density in  $\mathbb{R}^k$ ’, *Israel J. Math.* 54 (1986) 307–316.
2. B. COOK and Á. MAGYAR, ‘On restricted arithmetic progressions over finite fields’, *Online J. Anal. Comb.* 7 (2012) 1–10.
3. C. DEMETER, M. PRAMANIK and C. THIELE, ‘Multilinear singular operators with fractional rank’, *Pacific J. Math.* 246 (2010) 293–324.
4. K. FALCONER and J. MARSTRAND, ‘Plane sets with positive density at infinity contain all large distances’, *Bull. London Math. Soc.* 18 (1986) 471–474.
5. H. FURSTENBERG, Y. KATZNELSON and B. WEISS, *Ergodic theory and configurations in sets of positive density*, Mathematics of Ramsey Theory (Springer, Berlin/Heidelberg, 1990) 184–198.
6. W. T. GOWERS, ‘A new proof of Szemerédi’s theorem for arithmetic progressions of length four’, *Geom. Funct. Anal.* 8 (1998) 529–551.
7. R. L. GRAHAM, ‘Recent trends in Euclidean Ramsey theory’, *Discrete Math.* 136 (1994) 119–127.
8. L. HÖRMANDER, *The analysis of linear partial differential operators Vol. I: Distribution theory and Fourier analysis* (Springer, Berlin-Heidelberg, 1983).
9. C. MUSCALU, T. TAO and C. THIELE, ‘Multi-linear operators given by singular multipliers’, *J. Amer. Math. Soc.* 15 (2002) 469–496.
10. K. F. ROTH, ‘On certain sets of integers’, *J. London Math. Soc.* 28 (1953) 104–109.

Brian Cook  
The Fields Institute for Research in  
Mathematical Sciences  
222 College Street  
Toronto, ON  
Canada M5T 3J1

[bcook@fields.utoronto.ca](mailto:bcook@fields.utoronto.ca)

Ákos Magyar  
Department of Mathematics  
University of Georgia  
452 Boyd Graduate Studies  
Athens, GA 30602  
USA

[magyar@math.uga.edu](mailto:magyar@math.uga.edu)

Malabika Pramanik  
Department of Mathematics  
University of British Columbia  
Room 121, 1884 Mathematics Road  
Vancouver, BC  
Canada V6T 1Z4

[malabika@math.ubc.ca](mailto:malabika@math.ubc.ca)