

## Boxes, extended boxes and sets of positive upper density in the Euclidean space

BY POLONA DURCIK

*California Institute of Technology, 1200 E California Blvd,  
Pasadena CA 91125, U.S.A.*

*Current address: Chapman University, One University Drive, Orange,  
CA 92866, U.S.A.*

*e-mails: [durcik@chapman.edu](mailto:durcik@chapman.edu), [durcik@caltech.edu](mailto:durcik@caltech.edu)*

AND VJEKOSLAV KOVAČ

*Department of Mathematics, Faculty of Science, University of Zagreb,  
Bijenička cesta 30, 10000 Zagreb, Croatia*

*e-mail: [vjekovac@math.hr](mailto:vjekovac@math.hr)*

*(Received 11 March 2019; revised 27 August 2020)*

### *Abstract*

We prove that sets with positive upper Banach density in sufficiently large dimensions contain congruent copies of all sufficiently large dilates of three specific higher-dimensional patterns. These patterns are:  $2^n$  vertices of a fixed  $n$ -dimensional rectangular box, the same vertices extended with  $n$  points completing three-term arithmetic progressions, and the same vertices extended with  $n$  points completing three-point corners. Our results provide common generalizations of several Euclidean density theorems from the literature.

2020 Mathematics Subject Classification: Primary: 05D10, 28A75; Secondary: 11B30, 42B20

---

### 1. Introduction

Euclidean Ramsey theory typically seeks for a given pattern, such as vertices of a square, an arithmetic progression, etc., in a single partition class determined by an arbitrary (or only measurable) coloring of the Euclidean space. Stronger results than the mere coloring theorems are the so-called density theorems, which establish existence of the pattern inside an arbitrary measurable subset of positive density. The appropriate notion of density for this purpose is the *upper Banach density*, defined as

$$\bar{\delta}(A) := \limsup_{N \rightarrow \infty} \sup_{x \in \mathbb{R}^d} \frac{|A \cap (x + [0, N]^d)|}{N^d} \quad (1.1)$$

for any measurable  $A \subseteq \mathbb{R}^d$ . Here and in what follows,  $|B|$  denotes the Lebesgue measure of a measurable set  $B \subseteq \mathbb{R}^d$ .

An interesting class of density results tries to find congruent copies of all sufficiently large dilates of a given pattern. There is always a critical dimension  $d_{\min}$  below which positive statements cannot hold. Since many dimension-related issues are still unresolved, one is often content with proving the claim when  $d$  is sufficiently large. An initial result of this type starts with the simplest possible pattern, a pair of points in  $\mathbb{R}^d$  for  $d \geq 2$ , and it was established independently by Bourgain [1], Falconer and Marstrand [8], and Furstenberg, Katznelson and Weiss [10]. Moreover, Bourgain [1] generalised it to non-degenerate  $k$ -point patterns, also viewed as vertices of  $(k - 1)$ -dimensional simplices, in  $\mathbb{R}^d$  for  $d \geq k$ .

More recently, Lyall and Magyar [13] initiated the consideration of product-type patterns. They proved that, for fixed  $a_1, a_2 > 0$ , a positive density subset of  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ ,  $d_1, d_2 \geq 2$ , contains vertices of a rectangle,

$$(x_1, x_2), \quad (x_1, x_2 + s_2), \quad (x_1 + s_1, x_2), \quad (x_1 + s_1, x_2 + s_2),$$

with  $x_1, s_1 \in \mathbb{R}^{d_1}$ ,  $x_2, s_2 \in \mathbb{R}^{d_2}$ ,  $\|s_1\|_{\ell^2} = \lambda a_1$ , and  $\|s_2\|_{\ell^2} = \lambda a_2$ , for all sufficiently large  $\lambda > 0$ . We write  $\|v\|_{\ell^2}$  for the Euclidean norm of a vector  $v = (v_1, v_2, \dots, v_d) \in \mathbb{R}^d$ , since later we will also consider more general  $\ell^p$ -norms,  $1 < p < \infty$ , defined as

$$\|v\|_{\ell^p} := \left( \sum_{i=1}^d |v_i|^p \right)^{1/p}.$$

The particular case  $a_1 = a_2 = 1$  corresponds to the search for squares. In the same paper the authors proceed to Cartesian products of two general non-degenerate simplices.

As our first result, we establish a different generalization, replacing (vertices of) rectangles with (vertices of) higher-dimensional rectangular solids. Let  $d_1, d_2, \dots, d_n$  be positive integers. In what follows, a *box* will be a pattern consisting of  $2^n$  points in  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$  of the form

$$(x_1 + k_1 s_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n), \quad k_1, k_2, \dots, k_n \in \{0, 1\} \tag{1.2}$$

for any  $x_j, s_j \in \mathbb{R}^{d_j}$ ,  $s_j \neq \mathbf{0}$ ,  $j = 1, 2, \dots, n$ .

**THEOREM 1.** *Fix numbers  $a_1, a_2, \dots, a_n > 0$ . For any positive integers  $d_1, d_2, \dots, d_n \geq 5$  and any measurable set  $A \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$  with  $\bar{\delta}(A) > 0$  one can find  $\lambda_0 > 0$  with the property that for any real number  $\lambda \geq \lambda_0$  the set  $A$  contains a box (1.2) with  $x_j, s_j \in \mathbb{R}^{d_j}$  and  $\|s_j\|_{\ell^2} = \lambda a_j$ ,  $j = 1, 2, \dots, n$ .*

The possibility of generalising the aforementioned results of Lyall and Magyar to  $n$ -fold products with  $n \geq 3$  was announced by the same authors in [13]. The present paper and the more recent preprint by Lyall and Magyar [14] achieve this goal for boxes independently of each other and using quite different approaches. In fact, [14] proves a sharp variant of Theorem 1 above, in which the assumptions  $d_j \geq 5$  are relaxed to  $d_j \geq 2$ . It is clearly necessary to assume  $d_j \geq 2$ : if we had  $d_1 = 1$ , then the set of all points with the first coordinate from

$$\bigcup_{k \in \mathbb{Z}} [(k - 1/10)a_1, (k + 1/10)a_1]$$

would be a counterexample, since it would contain no boxes associated with half-integer values of  $\lambda$ .

The approach pursued in this paper is in the spirit of the paper by Cook, Magyar and Pramanik [2], and the same method will allow us to handle certain enlarged patterns we are about to discuss.

Bourgain [1] also constructed a measurable set  $A \in \mathbb{R}^d$  with  $\bar{\delta}(A) > 0$  such that lengths  $\|s\|_{\ell^2}$  of gaps  $s$  for all 3-term arithmetic progressions

$$x, x + s, x + 2s$$

inside  $A$  omit an unbounded set of positive values. This prevents us from having the most obvious candidate for a density theorem for 3-term arithmetic progressions. On the other hand, Cook, Magyar, and Pramanik [2] showed that the corresponding density theorem still holds if one is allowed to measure sizes of gaps  $s$  in the  $\ell^p$ -norms for  $1 < p < \infty, p \neq 2$ .

Our second result is a common generalisation of Theorem 1 above and [2, theorem 2.1]. Consider  $n$  additional points in  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$ ,

$$(x_1 + 2s_1, x_2, \dots, x_n), \quad (x_1, x_2 + 2s_2, \dots, x_n), \quad \dots, \quad (x_1, x_2, \dots, x_n + 2s_n) \quad (1.3)$$

for given  $x_j, s_j \in \mathbb{R}^{d_j}, s_j \neq \mathbf{0}, j = 1, 2, \dots, n$ . The union of (1.2) and (1.3) will be called a *3AP-extended box*: it has a 3-term arithmetic progression attached to each edge coming from a fixed vertex of the box. From the aforementioned observation of Bourgain we know that an analogue of Theorem 1 for the 3AP-extended boxes is not possible, so one has to give up on the Euclidean norm.

**THEOREM 2.** *Fix numbers  $a_1, a_2, \dots, a_n > 0$  and an exponent  $1 < p < \infty, p \neq 2$ . There exists a dimensional threshold  $d_{\min}$  such that for any positive integers  $d_1, d_2, \dots, d_n \geq d_{\min}$  and any measurable set  $A \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n}$  with  $\bar{\delta}(A) > 0$  one can find  $\lambda_0 > 0$  with the property that for any real number  $\lambda \geq \lambda_0$  the set  $A$  contains a 3AP-extended box (1.2)∪(1.3) with  $x_j, s_j \in \mathbb{R}^{d_j}$  and  $\|s_j\|_{\ell^p} = \lambda a_j, j = 1, 2, \dots, n$ .*

The pattern consisting of points (1.2) and (1.3) can also be viewed as a subset of the grid

$$(x_1 + k_1s_1, x_2 + k_2s_2, \dots, x_n + k_ns_n), \quad k_1, k_2, \dots, k_n \in \{0, 1, 2\},$$

consisting of  $3^n$  points. At the moment we are not able to prove a result analogous to Theorem 2 for this grid. Larger grids bring further complications: one should first handle longer arithmetic progressions and it is known that additional restrictions on the values of  $p$  are needed; see the remarks in [4].

The same approach will enable a further generalisation of Theorems 1 and 2. The present authors and Rimanić [4] have raised the generality of the result by Cook, Magyar and Pramanik [2] from 3-term arithmetic progressions to *corners*, which are triples of points in  $\mathbb{R}^d \times \mathbb{R}^d$  of the form

$$(x, y), (x + s, y), (x, y + s)$$

for  $x, y, s \in \mathbb{R}^d, s \neq \mathbf{0}$ . A *corner-extended box* will be a pattern in  $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n})^2$  consisting of  $2^n$  points forming a box,

$$(x_1 + k_1s_1, x_2 + k_2s_2, \dots, x_n + k_ns_n, y_1, y_2, \dots, y_n), \quad k_1, k_2, \dots, k_n \in \{0, 1\}, \quad (1.4)$$

and  $n$  additional points completing corners with  $n$  of its edges,

$$(x_1, x_2, \dots, x_n, y_1 + s_1, y_2, \dots, y_n), \quad \dots, \quad (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n + s_n), \tag{1.5}$$

where  $x_j, y_j, s_j \in \mathbb{R}^{d_j}, s_j \neq \mathbf{0}, j = 1, 2, \dots, n$ . The following result is a common generalisation of Theorem 1 above and [4, theorem 1.2].

**THEOREM 3.** *Fix numbers  $a_1, a_2, \dots, a_n > 0$  and an exponent  $1 < p < \infty, p \neq 2$ . There exists a dimensional threshold  $d_{\min}$  such that for any positive integers  $d_1, d_2, \dots, d_n \geq d_{\min}$  and any measurable set  $A \subseteq (\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n})^2$  with  $\bar{\delta}(A) > 0$  one can find  $\lambda_0 > 0$  with the property that for any real number  $\lambda \geq \lambda_0$  the set  $A$  contains a corner-extended box (1.4)  $\cup$  (1.5) with  $x_j, y_j, s_j \in \mathbb{R}^{d_j}$  and  $\|s_j\|_{\ell^p} = \lambda a_j, j = 1, 2, \dots, n$ .*

Theorem 3 implies Theorem 2, as can be seen by considering the skew projections  $(x_j, y_j) \mapsto y_j - x_j$ ; see [4] for details. Consequently, it is still necessary to assume  $p \neq 2$ , while the endpoint cases  $p = 1$  and  $p = \infty$  clearly do not allow any nontrivial results; see the comments in [2].

Let us emphasise that the dimensional threshold  $d_{\min}$  in Theorems 2 and 3 depends on the exponent  $p$ . Careful analysis of the arguments below can give  $d_{\min} = O(p)$  for each fixed  $n$ , but we do not even have counterexamples to the possibility  $d_{\min} = 2$ . The optimal value of  $d_{\min}$  is still far from understood, even in the case  $n = 1$  studied in [2] and [4].

Lyll and Magyar [12] also worked on the Euclidean embedding of all large dilates of a fixed distance graph. Their results do not include Theorem 1, since the boxes (or even rectangles) are simultaneously “too rigid” and “too degenerate;” compare with the definition of a *proper  $k$ -degenerate distance graph* from [12]. They also clearly do not overlap with Theorems 2 and 3, simply because these theorems fail in the Euclidean metric.

As we have already mentioned, our method of approach is based on the paper by Cook, Magyar, and Pramanik [2]. This method reduces Theorems 1–3 to boundedness of certain multilinear singular integral operators. In order to obtain bounds for these operators we invoke the main result from the recent paper by Thiele and one of the present authors [7], which in turn uses techniques gradually developed in a series of papers including [3–6, 11].

In Section 2 we list the main ingredients of the proofs in the form of several propositions and we explain how they imply the three theorems. Section 3 establishes the propositions that belong to the combinatorial part of the proof, by either invoking [2], or performing necessary modifications. Section 4 establishes the propositions dealing with singular integral operators, which constitute the analytical part of the proof.

### 2. Scheme of the proofs

We have already explained how Theorem 2 can be derived from Theorem 3, so in this section we give outlines of proofs of Theorems 1 and 3. In complete analogy with the steps from [2], they will be reduced to Propositions 4–6 below.

If  $A, B: \mathcal{D} \rightarrow [0, \infty)$  are two functions or functionals for which there exists a finite constant  $C$  depending on a set of parameters  $P$  such that  $A(x) \leq CB(x)$  for each  $x \in \mathcal{D}$ , then we write

$$A(x) \lesssim_P B(x).$$

If both  $A(x) \lesssim_P B(x)$  and  $B(x) \lesssim_P A(x)$ , then we write

$$A(x) \sim_P B(x).$$

The parameters in  $P$  that are understood throughout the text will be omitted from this notation. In particular, it will always be understood that all constants implicit in the notation  $\lesssim$  and  $\sim$  depend on the fixed vector of positive numbers

$$\mathbf{a} := (a_1, a_2, \dots, a_n),$$

which determines the shape of the patterns (the aspect ratios of the boxes), and the exponent  $p$ , which is relevant to the proof of Theorem 3 only.

Characteristic function (i.e. the indicator function) of a set  $E$  will be written as  $\mathbb{1}_E$ . Let us write  $g_t$  for an  $L^1$ -normalised dilate of a function  $g: \mathbb{R}^d \rightarrow \mathbb{C}$  by a factor  $t > 0$ , i.e.,

$$g_t(s) := t^{-d} g(t^{-1}s) \quad (2.1)$$

for each  $s \in \mathbb{R}^d$ . The Fourier transform of  $L^1$  functions is normalised as

$$\widehat{g}(\xi) := \int_{\mathbb{R}^d} g(s) e^{-2\pi i s \cdot \xi} ds,$$

where  $s \cdot \xi$  stands the standard scalar product of vectors  $s$  and  $\xi$  in  $\mathbb{R}^d$ . If  $\sigma$  is a measure on Borel subsets of  $\mathbb{R}^d$ , then we define its dilate by  $t > 0$  as another measure  $\sigma_t$  given as

$$\sigma_t(E) := \sigma(t^{-1}E) \quad (2.2)$$

for each Borel set  $E \subseteq \mathbb{R}^d$ . A consequence of a linear change of variables is

$$\int_{\mathbb{R}^d} f(s) d\sigma_t(s) = \int_{\mathbb{R}^d} f(ts) d\sigma(s)$$

for any measurable function  $f: \mathbb{R}^d \rightarrow \mathbb{C}$  such that the above integrals exist. Notation (2.1) and (2.2) is mutually consistent when  $\sigma$  is absolutely continuous with respect to the Lebesgue measure with density  $g$ . Occasionally we will need an  $L^p$ -normalised dilate of  $g: \mathbb{R}^d \rightarrow \mathbb{C}$  by  $t > 0$ , for a more general exponent  $1 \leq p < \infty$ , which will be denoted  $D_t^p g$  and defined as

$$D_t^p g(s) := t^{-d/p} g(t^{-1}s). \quad (2.3)$$

Let us fix an exponent  $1 < p < \infty$ ; it will simply be  $p = 2$  in relation with Theorem 1, while the proof of Theorem 3 will assume  $p \neq 2$ . We introduce a measure  $\sigma^{d,p}$  on Borel subsets of  $\mathbb{R}^d$  in the Dirac  $\delta$  notation as

$$\sigma^{d,p}(s) := \delta(1 - \|s\|_{\ell^p}^p)$$

or, less formally and abusing the integral representation for the Fourier transform, as

$$\sigma^{d,p}(s) = \int_{\mathbb{R}} e^{-2\pi i u(1 - \|s\|_{\ell^p}^p)} du.$$

Its dilate  $\sigma_\lambda^{d,p}$  by  $\lambda > 0$  is clearly supported on the  $C^1$  surface  $\{s \in \mathbb{R}^d : \|s\|_{\ell^p} = \lambda\}$ . Let us also fix a Schwartz function  $\psi: \mathbb{R} \rightarrow [0, 1]$  such that  $\widehat{\psi} \geq 0$ ,  $\widehat{\psi}$  is supported in  $[-4, 4]$ ,  $\widehat{\psi}(1) > 0$ , and  $\psi(0) = 1$ . For instance, we can take a  $C^\infty$  function  $\rho: \mathbb{R} \rightarrow [0, \infty)$  such

that  $\rho > 0$  on  $[-1, 1]$ ,  $\rho = 0$  outside  $[-2, 2]$ , and  $\rho$  has integral 1; then we can simply set  $\psi = |\widehat{\rho}|^2$ . Any constants implicit in the notation  $\lesssim$  and  $\sim$  will also be understood to depend on  $\psi$ . Furthermore, for any  $\varepsilon > 0$  we introduce a function  $\omega^{d,p,\varepsilon} : \mathbb{R}^d \rightarrow \mathbb{C}$  by the formula

$$\omega^{d,p,\varepsilon}(s) := \int_{\mathbb{R}} e^{-2\pi i u(1-\|s\|_{\ell^p}^p)} \psi(\varepsilon u) \, du = \varepsilon^{-1} \widehat{\psi}(\varepsilon^{-1}(1 - \|s\|_{\ell^p}^p)).$$

It was shown in [2, lemma 4.1] that

$$\int_{\mathbb{R}^d} \omega^{d,p,\varepsilon}(s) \, ds \sim_{d,p} 1$$

for all  $0 < \varepsilon < 1/10d$ . Thus, for such  $\varepsilon$  we set

$$c(d, p, \varepsilon) := \frac{\int_{\mathbb{R}^d} \omega^{d,p,\varepsilon}(s) \, ds}{\int_{\mathbb{R}^d} \omega^{d,p,1}(s) \, ds} \sim_{d,p} 1 \tag{2.4}$$

and then  $k^{d,p,\varepsilon} : \mathbb{R}^d \rightarrow \mathbb{R}$  defined by

$$k^{d,p,\varepsilon}(s) := \omega^{d,p,\varepsilon}(s) - c(d, p, \varepsilon)\omega^{d,p,1}(s) \tag{2.5}$$

has integral equal to 0.

We introduce the number  $D = d_1 + \dots + d_n$ , so that

$$\mathbb{R}^D \cong \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_n}.$$

Throughout the proofs we will use the shorthand notation

$$\mathbf{x} := (x_1, \dots, x_n), \quad \mathbf{y} := (y_1, \dots, y_n), \quad \mathbf{s} := (s_1, \dots, s_n)$$

and we view  $\mathbf{x}$ ,  $\mathbf{y}$ , and  $\mathbf{s}$  as vectors from  $\mathbb{R}^D$ . It will also be convenient to adopt some derived notation, such as

$$d\mathbf{x} := dx_1 \dots dx_n, \quad d\sigma_{\lambda \mathbf{a}}^p(\mathbf{s}) := d\sigma_{\lambda a_1}^{d_1,p}(s_1) \dots d\sigma_{\lambda a_n}^{d_n,p}(s_n).$$

In the same spirit we define

$$\omega_{\lambda \mathbf{a}}^{p,\varepsilon}(\mathbf{s}) := \omega_{\lambda a_1}^{d_1,p,\varepsilon}(s_1) \dots \omega_{\lambda a_n}^{d_n,p,\varepsilon}(s_n).$$

Let us also write  $\mathbf{k} = (k_1, \dots, k_n) \in \{0, 1\}^n$  and denote

$$(\mathcal{F}f)(\mathbf{x}, \mathbf{s}) := \prod_{\mathbf{k} \in \{0,1\}^n} f(x_1 + k_1 s_1, \dots, x_n + k_n s_n)$$

for a function  $f : \mathbb{R}^D \rightarrow [0, 1]$  and

$$(\widetilde{\mathcal{F}}f)(\mathbf{x}, \mathbf{y}, \mathbf{s}) := \left( \prod_{\mathbf{k} \in \{0,1\}^n} f(x_1 + k_1 s_1, \dots, x_n + k_n s_n, y_1, \dots, y_n) \right) f(x_1, \dots, x_n, y_1 + s_1, \dots, y_n) \dots f(x_1, \dots, x_n, y_1, \dots, y_n + s_n)$$

for a function  $f : (\mathbb{R}^D)^2 \rightarrow [0, 1]$ .

The most important objects are the *pattern-counting forms*, defined as follows. For a “scale”  $\lambda > 0$  we set

$$\mathcal{N}_\lambda^p(f) := \int_{(\mathbb{R}^D)^2} (\mathcal{F}f)(\mathbf{x}, \mathbf{s}) \, d\sigma_{\lambda\mathbf{a}}^p(\mathbf{s}) \, d\mathbf{x}$$

and

$$\tilde{\mathcal{N}}_\lambda^p(f) := \int_{(\mathbb{R}^D)^3} (\tilde{\mathcal{F}}f)(\mathbf{x}, \mathbf{y}, \mathbf{s}) \, d\sigma_{\lambda\mathbf{a}}^p(\mathbf{s}) \, d\mathbf{x} \, d\mathbf{y}.$$

The name comes from the fact that if  $\mathcal{N}_\lambda^p(\mathbb{1}_A) > 0$  (resp.  $\tilde{\mathcal{N}}_\lambda^p(\mathbb{1}_A) > 0$ ), then  $A$  contains a box (1.2) (resp. a corner-extended box (1.4)∪(1.5)) with  $\|s_j\|_{\ell^p} = \lambda a_j, j = 1, 2, \dots, n$ . We will also need their smoothed versions, defined for  $\varepsilon > 0$  as

$$\mathcal{M}_\lambda^{p,\varepsilon}(f) := \int_{(\mathbb{R}^D)^2} (\mathcal{F}f)(\mathbf{x}, \mathbf{s}) \omega_{\lambda\mathbf{a}}^{p,\varepsilon}(\mathbf{s}) \, d\mathbf{s} \, d\mathbf{x}$$

and

$$\tilde{\mathcal{M}}_\lambda^{p,\varepsilon}(f) := \int_{(\mathbb{R}^D)^3} (\tilde{\mathcal{F}}f)(\mathbf{x}, \mathbf{y}, \mathbf{s}) \omega_{\lambda\mathbf{a}}^{p,\varepsilon}(\mathbf{s}) \, d\mathbf{s} \, d\mathbf{x} \, d\mathbf{y}.$$

By the standard approximation of identity arguments,

$$\lim_{\varepsilon \rightarrow 0^+} \mathcal{M}_\lambda^{p,\varepsilon}(f) = \mathcal{N}_\lambda^p(f), \quad \lim_{\varepsilon \rightarrow 0^+} \tilde{\mathcal{M}}_\lambda^{p,\varepsilon}(f) = \tilde{\mathcal{N}}_\lambda^p(f) \tag{2.6}$$

for functions  $f$  as above. Finally, we denote

$$\mathcal{E}_\lambda^{p,\varepsilon}(f) := \mathcal{M}_\lambda^{p,\varepsilon}(f) - b(p, \varepsilon) \mathcal{M}_\lambda^{p,1}(f)$$

and

$$\tilde{\mathcal{E}}_\lambda^{p,\varepsilon}(f) := \tilde{\mathcal{M}}_\lambda^{p,\varepsilon}(f) - b(p, \varepsilon) \tilde{\mathcal{M}}_\lambda^{p,1}(f),$$

where we recall that the numbers  $c(d_j, p, \varepsilon)$  come from (2.4) and use the shorthand notation

$$b(p, \varepsilon) := c(d_1, p, \varepsilon) \cdots c(d_n, p, \varepsilon).$$

Here are the three main propositions needed in the proofs.

**PROPOSITION 4.** *Suppose that  $1 < p < \infty$  and that  $\delta, \lambda, N$  are real numbers such that  $0 < \delta \leq 1$  and  $0 < \lambda \leq N$ .*

(a) *If  $f : \mathbb{R}^D \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^D$  and satisfying  $\int_{[0, N]^D} f \geq \delta N^D$ , then*

$$\mathcal{M}_\lambda^{p,1}(f) \gtrsim_{D,\delta} N^D.$$

(b) *If  $f : \mathbb{R}^{2D} \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^{2D}$  and satisfying  $\int_{[0, N]^{2D}} f \geq \delta N^{2D}$ , then*

$$\tilde{\mathcal{M}}_\lambda^{p,1}(f) \gtrsim_{D,\delta} N^{2D}.$$

PROPOSITION 5. *Suppose that  $\varepsilon, \lambda, N$  are real numbers such that  $0 < \varepsilon < 1$  and  $0 < \lambda \leq N$ .*

(a) *If  $d_j \geq 5$  for  $j = 1, 2, \dots, n$  and if  $f : \mathbb{R}^D \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^D$ , then*

$$|\mathcal{N}_\lambda^2(f) - \mathcal{M}_\lambda^{2,\varepsilon}(f)| \lesssim_D \varepsilon^{1/4} N^D.$$

(b) *Additionally, take  $1 < p < \infty, p \neq 2$ . If each  $d_j$  is sufficiently large for  $j = 1, 2, \dots, n$  and if  $f : \mathbb{R}^{2D} \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^{2D}$ , then*

$$|\tilde{\mathcal{N}}_\lambda^p(f) - \tilde{\mathcal{M}}_\lambda^{p,\varepsilon}(f)| \lesssim_D \varepsilon N^{2D}.$$

PROPOSITION 6. *Suppose that  $1 < p < \infty, 0 < \varepsilon < 1/10D$ , and that  $\lambda_1, \lambda_2, \dots, \lambda_M$  are positive numbers such that  $\lambda_{m+1}/\lambda_m \geq 2$  for  $m = 1, 2, \dots, M - 1$ .*

(a) *If  $f : \mathbb{R}^D \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^D$ , then*

$$\sum_{m=1}^M |\mathcal{E}_{\lambda_m}^{p,\varepsilon}(f)| \lesssim_{D,\varepsilon} N^D.$$

(b) *If  $f : \mathbb{R}^{2D} \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^{2D}$ , then*

$$\left( \sum_{m=1}^M |\tilde{\mathcal{E}}_{\lambda_m}^{p,\varepsilon}(f)|^2 \right)^{1/2} \lesssim_{D,\varepsilon} N^{2D}.$$

Proofs of Propositions 4–6 are postponed to the later sections. Now we show how they imply Theorems 1 and 3.

*Proof of Theorem 1.* We argue by contradiction and suppose that there is a set  $A$  with strictly positive upper Banach density  $\bar{\delta}(A)$  for which the claim does not hold: there exists a sequence  $(\lambda_m)_{m=1}^\infty$  such that  $\lim_{m \rightarrow \infty} \lambda_m = \infty$  and that, for each  $m$ , the set  $A$  contains no boxes (1.2) with  $\|s_j\|_{\ell^2} = \lambda_m a_j, j = 1, 2, \dots, n$ . By omitting some terms we can achieve  $\lambda_{m+1}/\lambda_m \geq 2$  for each  $m$ . Fix an arbitrary positive integer  $M$ . For  $\delta := \bar{\delta}(A)/2 > 0$ , by the definition of the upper Banach density (1.1), we can find  $N \geq \lambda_M$  and  $\mathbf{x} \in \mathbb{R}^D$  such that

$$|A \cap (\mathbf{x} + [0, N]^D)| \geq \delta N^D.$$

Denote  $A' := (-\mathbf{x} + A) \cap [0, N]^D$  and  $f = \mathbb{1}_{A'}$ , so that  $A'$  is now a measurable subset of  $[0, N]^D$  satisfying

$$\int_{[0, N]^D} f(\mathbf{x}) \, d\mathbf{x} = |A'| \geq \delta N^D \tag{2.7}$$

and it still has no boxes (1.2) of the previously described sizes determined by  $\lambda$ . Consequently,

$$\mathcal{N}_{\lambda_m}^2(f) = 0 \tag{2.8}$$

for  $m = 1, 2, \dots, M$ . Because of condition (2.7), we can apply part (a) of Proposition 4 and get

$$\mathcal{M}_{\lambda_m}^{2,1}(f) \gtrsim_{D,\delta} N^D, \tag{2.9}$$

again for each  $m = 1, 2, \dots, M$ . Moreover, (2.8), Proposition 5(a) and (2.9) together give

$$\mathcal{M}_{\lambda_m}^{2,\varepsilon}(f) = |\mathcal{N}_{\lambda_m}^2(f) - \mathcal{M}_{\lambda_m}^{2,\varepsilon}(f)| \lesssim_D \varepsilon^{1/4} N^D \lesssim_{D,\delta} \varepsilon^{1/4} \mathcal{M}_{\lambda_m}^{2,1}(f). \tag{2.10}$$

By (2.10) and (2.4), for a sufficiently small  $\varepsilon$  depending on the dimensions and  $\delta$ , we have

$$\mathcal{M}_{\lambda_m}^{2,\varepsilon}(f) \leq \frac{1}{2} b(2, \varepsilon) \mathcal{M}_{\lambda_m}^{2,1}(f),$$

so,

$$|\mathcal{E}_{\lambda_m}^{2,\varepsilon}(f)| = b(2, \varepsilon) \mathcal{M}_{\lambda_m}^{2,1}(f) - \mathcal{M}_{\lambda_m}^{2,\varepsilon}(f) \geq \frac{1}{2} b(2, \varepsilon) \mathcal{M}_{\lambda_m}^{2,1}(f).$$

By (2.4) and (2.9) again, we conclude

$$|\mathcal{E}_{\lambda_m}^{2,\varepsilon}(f)| \gtrsim_{D,\delta} N^D \tag{2.11}$$

for each  $m = 1, 2, \dots, M$ . Summing the lower bound (2.11) in  $m$  gives

$$\sum_{m=1}^M |\mathcal{E}_{\lambda_m}^{2,\varepsilon}(f)| \gtrsim_{D,\delta} M N^D. \tag{2.12}$$

Finally, combining (2.12) with Proposition 6 yields  $M \lesssim_{D,\delta} 1$ , which contradicts the fact that  $M$  could have been chosen arbitrarily large.

*Proof of Theorem 3.* The same outline also applies here. The only difference is that Proposition 5(b) only holds for sufficiently large dimensions  $d_j$  depending on  $p$ . The reader can also consult the corresponding proofs of [2, theorem 2.2] and [4, theorem 1.2].

### 3. Combinatorial results

The following lemma is needed in the proof of Proposition 4 in the same way in which Bourgain’s version of Roth’s theorem for compact abelian groups [1] is needed in the analogous proposition in [2].

LEMMA 7. *Suppose  $0 < \delta \leq 1$ .*

(a) *If  $f: \mathbb{R}^D \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^D$  and satisfying  $\int_{[0,N]^D} f \geq \delta N^D$ , then*

$$\int_{([0,1]^D)^2} (\mathcal{F}f)(\mathbf{x}, \mathbf{s}) \, ds \, dx \gtrsim_{D,\delta} 1.$$

(b) *If  $f: \mathbb{R}^{2D} \rightarrow [0, 1]$  is a measurable function supported in  $[0, N]^{2D}$  and satisfying  $\int_{[0,N]^{2D}} f \geq \delta N^{2D}$ , then*

$$\int_{([0,1]^D)^3} (\tilde{\mathcal{F}}f)(\mathbf{x}, \mathbf{y}, \mathbf{s}) \, ds \, dx \, dy \gtrsim_{D,\delta} 1.$$

*Proof of Lemma 7.* Both parts of the lemma are shown using multidimensional Szemerédi’s theorem of Furstenberg and Katznelson [9]. By this result, for any dimension  $n$  and any number  $0 < \beta \leq 1$  there exists a positive integer  $m_{n,\beta}$  such that for each positive integer  $m \geq m_{n,\beta}$  one has the following:

- (i) each subset  $S \subseteq \{0, 1, \dots, m - 1\}^n$  of cardinality at least  $\beta m^n$  contains (vertices of) an  $n$ -dimensional cube,

$$(i_1 + k_1l, i_2 + k_2l, \dots, i_n + k_nl), \quad k_1, k_2, \dots, k_n \in \{0, 1\}$$

for some  $i_1, i_2, \dots, i_n, l \in \mathbb{Z}$  with  $l \neq 0$ ;

- (ii) each subset  $S \subseteq \{0, 1, \dots, m - 1\}^{2n}$  of cardinality at least  $\beta m^{2n}$  contains a  $2n$ -dimensional corner-extended cube,

$$(i_1 + k_1l, i_2 + k_2l, \dots, i_n + k_nl, j_1, j_2, \dots, j_n), \quad k_1, k_2, \dots, k_n \in \{0, 1\},$$

$$(i_1, i_2, \dots, i_n, j_1 + l, j_2, \dots, j_n), \quad \dots, \quad (i_1, i_2, \dots, i_n, j_1, j_2, \dots, j_n + l)$$

for some  $i_1, i_2, \dots, i_n, j_1, j_2, \dots, j_n, l \in \mathbb{Z}$  with  $l \neq 0$ .

Then one applies the averaging trick of Varnavides [16], in the same way it was done in the proof of lemma 3.2 in [4] for the particular case of the three-point corners.

Indeed, multidimensional Szemerédi’s theorem applies to any finite pattern on the integer lattice, not only to boxes and corner-extended boxes, so Lemma 7 can be generalized easily. The reasons why we restrict our attention to very special patterns lie in the rigidity of other auxiliary results, most notably Proposition 6 above and Theorem 10 from Section 4.

*Proof of Proposition 4.* The proposition is shown by cutting the Euclidean space into cubes, the scaled copies of  $([0, 1]^D)^2$  or  $([0, 1]^D)^3$ , and applying Lemma 7 on each of them. For details the reader can consult the proof of proposition 2.1 in [2].

Now we turn to the proof of the second proposition. We will need the Euclidean version of the notion of the Gowers norms, so let us begin by setting

$$(\Delta_h g)(s) := g(s)\overline{g(s+h)}$$

for  $s, h \in \mathbb{R}^d$  and a function  $g: \mathbb{R}^d \rightarrow \mathbb{C}$ . If such  $g$  is also measurable, then its *Gowers uniformity norm of degree  $k$*  is defined as

$$\|g\|_{U^k(\mathbb{R}^d)} := \left( \int_{(\mathbb{R}^d)^{k+1}} (\Delta_{h_k} \cdots \Delta_{h_1} g)(s) \, ds \, dh_1 \cdots dh_k \right)^{2^{-k}}$$

$$= \left( \int_{(\mathbb{R}^d)^{k-1}} \left| \int_{\mathbb{R}^d} (\Delta_{h_{k-1}} \cdots \Delta_{h_1} g)(s) \, ds \right|^2 dh_1 \cdots dh_{k-1} \right)^{2^{-k}}. \quad (3.1)$$

We will only need the norms  $\|\cdot\|_{U^2(\mathbb{R}^d)}$  and  $\|\cdot\|_{U^3(\mathbb{R}^d)}$ . The Gowers norms scale properly with respect to the  $L^1$ -normalised dilations of the function. In particular,

$$\|g_t\|_{U^2(\mathbb{R}^d)} = t^{-d/4} \|g\|_{U^2(\mathbb{R}^d)}, \quad \|g_t\|_{U^3(\mathbb{R}^d)} = t^{-d/2} \|g\|_{U^3(\mathbb{R}^d)}, \quad (3.2)$$

as is shown by an easy change of variables of integration.

LEMMA 8. Suppose that  $\lambda$  and  $N$  are real numbers such that  $0 < \lambda \leq N$ .

(a) If  $f_1, f_2: \mathbb{R}^d \rightarrow [0, 1]$  are measurable functions supported in  $[0, N]^d$ , and  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  is a measurable function supported in  $[-3\lambda, 3\lambda]^d$ , then

$$\left| \int_{(\mathbb{R}^d)^2} f_1(x) f_2(x + s) g(s) \, ds \, dx \right| \lesssim_d N^d \lambda^{d/4} \|g\|_{U^2(\mathbb{R}^d)}. \tag{3.3}$$

(b) If  $f_1, f_2, f_3: \mathbb{R}^{2d} \rightarrow [0, 1]$  are measurable functions supported in  $[0, N]^{2d}$  and  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  is a measurable function supported in  $[-3\lambda, 3\lambda]^d$ , then

$$\left| \int_{(\mathbb{R}^d)^3} f_1(x, y) f_2(x + s, y) f_3(x, y + s) g(s) \, ds \, dx \, dy \right| \lesssim_d N^{2d} \lambda^{d/2} \|g\|_{U^3(\mathbb{R}^d)}. \tag{3.4}$$

*Proof of Lemma 8.* This lemma will be shown in a way similar to the proof of lemma 4.2 in [2].

Proof of (a). The left-hand side of (3.3) can be rewritten as

$$\left| \int_{[0, N]^d} f_1(x) \left( \int_{\mathbb{R}^d} f_2(x + s) g(s) \, ds \right) \, dx \right|,$$

so the Cauchy–Schwarz inequality in  $x$  bounds it by

$$N^{d/2} \left( \int_{(\mathbb{R}^d)^3} f_2(x + s) f_2(x + s') g(s) g(s') \, ds \, ds' \, dx \right)^{1/2}.$$

Substituting  $y = x + s$ ,  $h = s' - s$ , applying the Cauchy–Schwarz inequality again, and recognising the  $U^2$ -norm from (3.1), we get

$$N^{d/2} \|g\|_{U^2(\mathbb{R}^d)} \left( \int_{[-6\lambda, 6\lambda]^d} \left( \int_{[0, N]^d} f_2(y) f_2(y + h) \, dy \right)^2 \, dh \right)^{1/4}.$$

The expression within the outer parentheses is clearly at most a constant times  $N^{2d} \lambda^d$ .

Proof of (b). Using the Cauchy–Schwarz inequality we first bound the left-hand side of (3.4) by

$$N^d \left( \int_{(\mathbb{R}^d)^4} f_2(x + s, y) f_2(x + s', y) f_3(x, y + s) f_3(x, y + s') g(s) g(s') \, ds \, ds' \, dx \, dy \right)^{1/2}.$$

Then we substitute  $h = s' - s$  and perform the shift  $x \mapsto x - s$  to transform the integral inside parentheses into

$$\int_{(\mathbb{R}^d)^4} f_2(x, y) f_2(x + h, y) f_3(x - s, y + s) f_3(x - s, y + s + h) g(s) g(s + h) \, dx \, dy \, ds \, dh.$$

Another application of the Cauchy–Schwarz inequality controls the left-hand side of (3.4) with

$$N^{3d/2} \lambda^{d/4} \left( \int_{(\mathbb{R}^d)^5} f_3(x - s, y + s) f_3(x - s', y + s') f_3(x - s, y + s + h) \right. \\ \left. \times f_3(x - s', y + s' + h) g(s) g(s') g(s + h) g(s' + h) \, ds \, ds' \, dx \, dy \, dh \right)^{1/4}.$$

It remains to substitute  $h' = s' - s$ , shift  $x \mapsto x + s$ ,  $y \mapsto y - s$ , use the Cauchy–Schwarz inequality one more time, and finally recognise  $\|g\|_{U^3(\mathbb{R}^d)}$  from (3.1).

LEMMA 9. *Suppose that  $\eta$  and  $\varepsilon$  are real numbers such that  $0 < \eta < \varepsilon < 1$ .*

(a) *For  $d \geq 5$  we have*

$$\|\omega^{d,2,\eta} - \omega^{d,2,\varepsilon}\|_{U^2(\mathbb{R}^d)} \lesssim_d \varepsilon^{1/4}.$$

(b) *For  $1 < p < \infty$ ,  $p \neq 2$  and sufficiently large  $d$  depending on  $p$  we have*

$$\|\omega^{d,p,\eta} - \omega^{d,p,\varepsilon}\|_{U^3(\mathbb{R}^d)} \lesssim_{p,d} \varepsilon.$$

*Proof of Lemma 9.* Part (b) was already established in [2, lemma 2.4].

We will show part (a) using the same lines of proof, but we need to use to our advantage the fact that we only need the  $U^2$ -norm and get more concrete decay in  $\varepsilon$  as  $\varepsilon \rightarrow 0^+$ . Let  $\varphi: \mathbb{R} \rightarrow [0, \infty)$  be a compactly supported  $C^\infty$  function that is constantly equal to 1 on  $[-3, 3]$  and set

$$\Phi = \underbrace{\varphi \otimes \cdots \otimes \varphi}_d.$$

All constants are assumed to depend on  $\varphi$  without further mention. Observe that  $\omega^{d,2,\eta}$  and  $\omega^{d,2,\varepsilon}$  are supported on  $[-3, 3]^d$ , so

$$\omega^{d,2,\eta}(s) - \omega^{d,2,\varepsilon}(s) = \int_{\mathbb{R}} (\psi(\eta u) - \psi(\varepsilon u)) \Phi(s) e^{2\pi i u(\|s\|_2^2 - 1)} du.$$

Integral version of the triangle inequality for the Gowers norm and the tensor product splitting of the exponential give

$$\|\omega^{d,2,\eta} - \omega^{d,2,\varepsilon}\|_{U^2(\mathbb{R}^d)} \leq \int_{\mathbb{R}} |\psi(\eta u) - \psi(\varepsilon u)| \|\varphi(s) e^{2\pi i u s^2}\|_{U_s^2(\mathbb{R})}^d du. \tag{3.5}$$

By definition of the Gowers norm (3.1),

$$\|\varphi(s) e^{2\pi i u s^2}\|_{U_s^2(\mathbb{R})}^4 = \int_{[-3,3]} \left| \int_{\mathbb{R}} \varphi(s) \varphi(s+h) e^{-4\pi i u h s} ds \right|^2 dh,$$

which is certainly bounded by a constant, for each  $u \in \mathbb{R}$ . However, for  $|u| \geq 1$  we get a better estimate by splitting the outer domain of integration into  $|h| \leq |u|^{-1}$  and  $|u|^{-1} < |h| \leq 3$ . The first part of the integral is clearly at most a constant times  $|u|^{-1}$ . Integration by parts in the second part gives

$$\left| \int_{\mathbb{R}} \varphi(s) \varphi(s+h) e^{-4\pi i u h s} u h ds \right| \lesssim 1,$$

so that

$$\int_{\{|u|^{-1} < |h| \leq 3\}} \left| \int_{\mathbb{R}} \varphi(s) \varphi(s+h) e^{-4\pi i u h s} ds \right|^2 dh \lesssim |u|^{-2} \int_{\{|u|^{-1} < |h| \leq 3\}} h^{-2} dh \lesssim |u|^{-1}.$$

From these we conclude

$$\|\varphi(s) e^{2\pi i u s^2}\|_{U_s^2(\mathbb{R})} \lesssim \min\{1, |u|^{-1/4}\}. \tag{3.6}$$

Now we combine (3.5) and (3.6) into a single estimate

$$\|\omega^{d,2,\eta} - \omega^{d,2,\varepsilon}\|_{U^2(\mathbb{R}^d)} \lesssim \int_{\mathbb{R}} |\psi(\eta u) - \psi(\varepsilon u)| \min\{1, |u|^{-d/4}\} du.$$

This time we split the domain of integration into three parts:  $|u| \leq 1$ ,  $1 < |u| \leq \varepsilon^{-1}$ , and  $|u| > \varepsilon^{-1}$ . We bound the corresponding integrals respectively as

$$\int_{\{|u| \leq 1\}} \|\psi'\|_{L^\infty(\mathbb{R})} \varepsilon |u| du \lesssim \varepsilon \lesssim \varepsilon^{1/4},$$

$$\int_{\{1 < |u| \leq \varepsilon^{-1}\}} \|\psi'\|_{L^\infty(\mathbb{R})} \varepsilon |u|^{1-d/4} du \lesssim_d \varepsilon^{d/4-1} \lesssim \varepsilon^{1/4},$$

and

$$\int_{\{|u| > \varepsilon^{-1}\}} 2\|\psi\|_{L^\infty(\mathbb{R})} |u|^{-d/4} du \lesssim_d \varepsilon^{d/4-1} \lesssim \varepsilon^{1/4}.$$

In the last display we needed  $d > 4$  for the convergence of the improper integral and also to have  $d/4 - 1 \geq 1/4 > 0$ . This allows us to conclude the desired inequality.

*Proof of Proposition 5.* Both parts of the proposition are shown in exactly the same way, using Lemmata 8 and 9, so we only elaborate on the proof of part (a).

Because of (2.6) it is enough to bound the difference

$$\mathcal{M}_\lambda^{2,\eta}(f) - \mathcal{M}_\lambda^{2,\varepsilon}(f) \tag{3.7}$$

for all  $0 < \eta < \varepsilon < 1$ , with a constant independent of  $\eta$ . The difference of the corresponding cutoff functions can be expanded as

$$\omega_{\lambda \mathbf{a}}^{p,\eta}(\mathbf{s}) - \omega_{\lambda \mathbf{a}}^{p,\varepsilon}(\mathbf{s}) = \sum_{j=1}^n \tau_\lambda^{(j)}(\mathbf{s}),$$

where

$$\tau^{(j)}(\mathbf{s}) := \left( \prod_{i=1}^{j-1} \omega_{a_i}^{d_i,p,\eta}(s_i) \right) \left( \omega_{a_j}^{d_j,p,\eta}(s_j) - \omega_{a_j}^{d_j,p,\varepsilon}(s_j) \right) \left( \prod_{i=j+1}^n \omega_{a_i}^{d_i,p,\varepsilon}(s_i) \right).$$

This decomposes (3.7) into  $n$  pieces,  $\sum_{j=1}^n \mathcal{P}_\lambda^{(j)}(f)$ , where

$$\mathcal{P}_\lambda^{(j)}(f) := \int_{(\mathbb{R}^D)^2} (\mathcal{F}f)(\mathbf{x}, \mathbf{s}) \tau_\lambda^{(j)}(\mathbf{s}) ds dx.$$

Without loss of generality we will estimate the piece  $\mathcal{P}_\lambda^{(1)}(f)$ . Part Lemma 9(a) and (3.2) give

$$\|\tau_\lambda^{(j)}(\mathbf{s})\|_{U_{s_1}^2(\mathbb{R}^{d_1})} \| \cdot \|_{L_{s_2, \dots, s_n}^\infty(\mathbb{R}^{D-d_1})} \lesssim_D \lambda^{-d_1/4} \varepsilon^{1/4} \lambda^{-(d_2+\dots+d_n)} = \varepsilon^{1/4} \lambda^{-D+3d_1/4}.$$

Then we observe that  $(\mathcal{F}f)(\mathbf{x}, \mathbf{s})$  can, for fixed  $x_2, \dots, x_n, s_2, \dots, s_n$ , be written in the form  $f_1(x_1) f_2(x_1 + s_1)$  from Lemma 8, so we obtain

$$\left| \int_{(\mathbb{R}^{d_1})^2} (\mathcal{F}f)(\mathbf{x}, \mathbf{s}) \tau_\lambda^{(j)}(\mathbf{s}) ds_1 dx_1 \right| \leq N^{d_1} \lambda^{d_1/4} \varepsilon^{1/4} \lambda^{-D+3d_1/4} = \varepsilon^{1/4} N^{d_1} \lambda^{-D+d_1}.$$

Integrating in  $x_j \in [0, N]^{d_j}$  and  $s_j \in [-3\lambda, 3\lambda]^{d_j}$ ,  $j = 1, 2, \dots, n$  we finally get

$$|\mathcal{P}_\lambda^{(1)}(f)| \lesssim_D \varepsilon^{1/4} N^{d_1} \lambda^{-D+d_1} N^{D-d_1} \lambda^{D-d_1} = \varepsilon^{1/4} N^D,$$

which completes the proof.

4. Analytical results

The main ingredient in the proof of Proposition 6 is an estimate for multilinear singular integral forms. We formulate it as a separate theorem.

THEOREM 10.

- (a) Suppose that  $K : \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_n} \rightarrow \mathbb{C}$  is a bounded compactly supported function and that its Fourier transform satisfies the standard symbol estimates (cf. [15]),

$$|\widehat{K}(\xi)| \leq C_\kappa \|\xi\|_{\ell^2}^{-|\kappa|} \tag{4.1}$$

for any multi-index  $\kappa$ . Then we have the inequality

$$\left| \int_{(\mathbb{R}^D)^2} K(\mathbf{s}) \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) \, ds \, dx \right| \lesssim \prod_{\mathbf{k} \in \{0,1\}^n} \|F_{\mathbf{k}}\|_{L^{2^n}} \tag{4.2}$$

with the implicit constant depending only on  $(C_\kappa)_\kappa$  and the dimensions  $d_i$ .

- (b) Suppose that  $K : \mathbb{R}^{d_1} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \dots \times \mathbb{R}^{d_n} \rightarrow \mathbb{C}$  is a bounded compactly supported function such that its Fourier transform satisfies the standard symbol estimates (4.1) for any multi-index  $\kappa$ . Then we have the inequality

$$\left| \int_{\mathbb{R}^{3D+d_1}} K(s_1, s'_1, s_2, \dots, s_n) \left( \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \right) \right. \\ \left. \times F_{\tilde{\mathbf{k}}}(x_1 + l_1 s_1 + l_2 s'_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1 + l_3 s_1 + l_4 s'_1, y_2, \dots, y_n) \right) \\ \times ds_1 ds'_1 ds_2 \dots ds_n \, dx \, dy \Big| \lesssim \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \|F_{\tilde{\mathbf{k}}}\|_{L^{2^{n+1}}},$$

with the implicit constant depending only on  $(C_\kappa)_\kappa$  and the dimensions  $d_i$ .

Note that the implicit constants in both inequalities claimed by Theorem 10 depend only on the implicit constants from the symbol estimates (4.1) and the dimensions, the latter being regarded as fixed throughout the text. Also observe that part (b) of the theorem specialised to  $n = 1$  coincides with [4, theorem 1.3], the main analytic result of that paper.

Once Theorem 10 is established, it is easy to complete the proof of Proposition 6. Let us elaborate on that argument and postpone the proof of the theorem to the second half of this section.

*Proof of Proposition 6.* For the proof of part (a), note that the cutoff function appearing in  $\mathcal{E}_{\lambda_m}^{p,\varepsilon}(f)$  is

$$\omega_{\lambda_m \mathbf{a}}^{p,\varepsilon}(\mathbf{s}) - b(p, \varepsilon) \omega_{\lambda_m \mathbf{a}}^{p,1}(\mathbf{s}) = \prod_{j=1}^n \omega_{\lambda_m a_j}^{d_j, p, \varepsilon}(s_j) - \prod_{j=1}^n c(d_j, p, \varepsilon) \omega_{\lambda_m a_j}^{d_j, p, 1}(s_j).$$

Recalling the introduction of  $k^{d,p,\varepsilon}$  in (2.5), we can rewrite it as

$$\sum_{j=1}^n \left( \prod_{i=1}^{j-1} \omega_{\lambda_m a_i}^{d_i,p,\varepsilon}(s_i) \right) k_{\lambda_m a_j}^{d_j,p,\varepsilon}(s_j) \left( \prod_{i=j+1}^n c(d_i, p, \varepsilon) \omega_{\lambda_m a_i}^{d_i,p,1}(s_i) \right). \tag{4.3}$$

Note that each summand in (4.3) is of the form

$$\varphi_{\lambda_m}^{(1)}(s_1) \varphi_{\lambda_m}^{(2)}(s_2) \dots \varphi_{\lambda_m}^{(n)}(s_n), \tag{4.4}$$

where  $\varphi^{(j)}$ ,  $j = 1, 2, \dots, n$ , are  $C^1$  functions and one of them has integral equal to 0, while the others are nonnegative. Take arbitrary signs  $\alpha_m \in \{-1, 1\}$ ,  $m = 1, 2, \dots, M$ . By a standard computation (see [2]) the kernels

$$K(\mathbf{s}) := \sum_{m=1}^M \alpha_m \varphi_{\lambda_m}^{(1)}(s_1) \varphi_{\lambda_m}^{(2)}(s_2) \dots \varphi_{\lambda_m}^{(n)}(s_n)$$

satisfy the conditions from Theorem 10, with constants  $C_\kappa$  independent of the numbers  $M, \lambda_1, \dots, \lambda_M$  and signs  $\alpha_1, \dots, \alpha_M$ , but we allow dependencies on the dimensions (i.e.  $D$ ), on  $\varepsilon$ , on the numbers  $a_1, \dots, a_n$ , and on the exponent  $p$ . Applying part (a) of Theorem 10 to those kernels and  $F_{\mathbf{k}} = f$ , we obtain

$$\left| \sum_{m=1}^M \alpha_m \mathcal{E}_{\lambda_m}^{p,\varepsilon}(f) \right| \lesssim_{D,\varepsilon} \|f\|_{L^{2^n}}^{2^n} \leq N^D.$$

It remains to choose the signs  $\alpha_m$  appropriately, so that the left hand side becomes  $\sum_{m=1}^M |\mathcal{E}_{\lambda_m}^{p,\varepsilon}(f)|$ .

In the proof of part (b) we begin with the same splitting (4.3) into summands of the form (4.4). Since the notation has become symmetric in  $j$ , without loss of generality we can suppose  $\int_{\mathbb{R}^{d_1}} \varphi^{(1)} = 0$ , i.e. the cancellation comes from the variable  $s_1$ . Gathering inside parentheses all factors containing that variable, the corresponding part of  $\tilde{\mathcal{E}}_{\lambda_m}^{p,\varepsilon}(f)$  can be rewritten as

$$\begin{aligned} & \int_{\mathbb{R}^{3D-d_1}} \left( \int_{\mathbb{R}^{d_1}} \left( \prod_{(k_2, \dots, k_n) \in \{0,1\}^{n-1}} f(x_1 + s_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1, \dots, y_n) \right) \right. \\ & \quad \times f(x_1, \dots, x_n, y_1 + s_1, \dots, y_n) \varphi_{\lambda_m}^{(1)}(s_1) \, ds_1 \Big) \\ & \quad \times \left( \prod_{(k_2, \dots, k_n) \in \{0,1\}^{n-1}} f(x_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1, \dots, y_n) \right) \\ & \quad \times f(x_1, \dots, x_n, y_1, y_2 + s_2, \dots, y_n) \cdots f(x_1, \dots, x_n, y_1, y_2, \dots, y_n + s_n) \\ & \quad \times \varphi_{\lambda_m}^{(2)}(s_2) \dots \varphi_{\lambda_m}^{(n)}(s_n) \, ds_2 \cdots ds_n \, d\mathbf{x} \, d\mathbf{y}. \end{aligned}$$

By the Cauchy–Schwarz inequality its square is bounded with

$$\mathcal{A}_{\lambda_m}(f) \mathcal{B}_{\lambda_m}(f), \tag{4.5}$$

where

$$\mathcal{A}_\lambda(f) := \int_{\mathbb{R}^{3D+d_1}} \left( \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \right) \times f_{\tilde{\mathbf{k}}}(x_1 + l_1s_1 + l_2s'_1, x_2 + k_2s_2, \dots, x_n + k_ns_n, y_1 + l_3s_1 + l_4s'_1, y_2, \dots, y_n) \times \varphi_\lambda^{(1)}(s_1)\varphi_\lambda^{(1)}(s'_1)\varphi_\lambda^{(2)}(s_2) \dots \varphi_\lambda^{(n)}(s_n) ds_1 ds'_1 ds_2 \dots ds_n dx dy$$

and

$$\mathcal{B}_\lambda(f) := \int_{\mathbb{R}^{3D-d_1}} \left( \prod_{(k_2, \dots, k_n) \in \{0,1\}^{n-1}} \mathbb{1}_{[0,N]^{2D}}(x_1, x_2 + k_2s_2, \dots, x_n + k_ns_n, y_1, \dots, y_n) \right) \times \mathbb{1}_{[0,N]^{2D}}(x_1, \dots, x_n, y_1, y_2 + s_2, \dots, y_n) \dots \mathbb{1}_{[0,N]^{2D}}(x_1, \dots, x_n, y_1, y_2, \dots, y_n + s_n) \times \varphi_\lambda^{(2)}(s_2) \dots \varphi_\lambda^{(n)}(s_n) ds_2 \dots ds_n dx dy.$$

The functions  $f_{\tilde{\mathbf{k}}}$  appearing in the definition of  $\mathcal{A}_\lambda(f)$  are all supported in  $[0, N]^{2D}$  and taking values in  $[0, 1]$ ; some of them are equal to  $f$ , while the others are artificially inserted into the expression as  $\mathbb{1}_{[0,2N]^{2D}}$ . Clearly,

$$\mathcal{B}_{\lambda_m}(f) \leq \left( \prod_{j=2}^n \int_{\mathbb{R}^{d_j}} \varphi_{\lambda_m}^{(j)} \right) \left( \prod_{j=1}^n \int_{\mathbb{R}^{d_j}} \mathbb{1}_{[0,N]^{d_j}} \right)^2 \lesssim_{D,\varepsilon} N^{2D}. \tag{4.6}$$

On the other hand, Theorem 10(b) applied with the kernel

$$K(\mathbf{s}) := \sum_{m=1}^M \varphi_{\lambda_m}^{(1)}(s_1)\varphi_{\lambda_m}^{(1)}(s'_1)\varphi_{\lambda_m}^{(2)}(s_2) \dots \varphi_{\lambda_m}^{(n)}(s_n)$$

and the functions  $F_{\tilde{\mathbf{k}}} = f_{\tilde{\mathbf{k}}}$  yields

$$\left| \sum_{m=1}^M \mathcal{A}_{\lambda_m}(f) \right| \lesssim_{D,\varepsilon} N^{2D}. \tag{4.7}$$

Summing the products (4.5) in  $m = 1, 2, \dots, M$ , from (4.6) and (4.7) we finally conclude

$$\sum_{m=1}^M |\tilde{\mathcal{E}}_{\lambda_m}^{p,\varepsilon}(f)|^2 \lesssim_{D,\varepsilon} N^{4D},$$

as claimed.

We finalise the paper with the proof of the remaining analytical result. As we will soon see, the case when all dimensions  $d_i$  are equal will be an easy consequence of the main result from the paper by Thiele and one of the present authors [7]. We will spend just a slight additional effort to reduce the case of possibly different dimensions  $d_i$  to the very same result. An alternative to this addition could be considering appropriate lower-dimensional sections of the sets and functions appearing throughout the paper.

*Proof of Theorem 10.* Proof of (a). First we consider the case of equal dimensions, i.e.  $d_1 = d_2 = \dots = d_n$ , and we write them simply as  $d$ . Relabelling  $x_i$  to  $x_i^0$  in the left-hand side of (4.2) and changing variables  $x_i^0 + s_i = x_i^1$  for  $i = 1, \dots, n$ , we see that we need to show

$$\left| \int_{\mathbb{R}^{2dn}} \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(\Pi_{\mathbf{k}}x) K(\Pi x) \, dx \right| \lesssim_{(C_{\kappa})_{\kappa}} \prod_{\mathbf{k} \in \{0,1\}^n} \|F_{\mathbf{k}}\|_{L^{2^n}(\mathbb{R}^{nd})},$$

where  $x = (x_1^0, \dots, x_n^0, x_1^1, \dots, x_n^1) \in (\mathbb{R}^d)^{2n}$  and  $\Pi_{\mathbf{k}}, \Pi : (\mathbb{R}^d)^{2n} \rightarrow (\mathbb{R}^d)^n$  are linear operators given by

$$\Pi_{\mathbf{k}}x := (x_1^{k_1}, \dots, x_n^{k_n}), \quad \Pi x := (x_1^1 - x_1^0, \dots, x_n^1 - x_n^0).$$

This estimate can be recognised as one of the singular Brascamp–Lieb inequalities from the main theorem of [7], which establishes the claim in the case of equal dimensions.

The general case of different dimensions in (4.2) will be deduced from the case of equal dimensions as follows. By approximating  $K$  with smooth compactly supported functions in  $L^1$  and applying Hölder’s inequality, we may assume that  $K$  is a smooth compactly supported function on  $\mathbb{R}^D \cong \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_n}$ , satisfying the symbol estimates (4.1). We set

$$d := \max_{1 \leq i \leq n} d_i + 1.$$

(We have added 1 for technical reasons, so that  $d - d_i > 0$  for all  $1 \leq i \leq n$ .) First we define a new function  $\tilde{K}$  on  $\mathbb{R}^{nd}$ , whose integration in a certain direction gives  $K$ . To achieve this take a smooth compactly supported function  $\varphi$  on  $\mathbb{R}^{nd-D} \cong \mathbb{R}^{d-d_1} \times \dots \times \mathbb{R}^{d-d_n}$  satisfying  $\varphi(\mathbf{0}) = 1$ . Then we define  $\tilde{K}$  by setting

$$\widehat{\tilde{K}}(\xi, \tilde{\xi}) := \widehat{K}(\xi) \varphi\left(\frac{\tilde{\xi}}{\|\xi\|}\right) \tag{4.8}$$

for  $\mathbf{0} \neq \xi \in \mathbb{R}^D, \tilde{\xi} \in \mathbb{R}^{nd-D}$ . Observe that we have

$$\widehat{\tilde{K}}(\xi, \mathbf{0}) = \widehat{K}(\xi)$$

or, passing to the spatial side,

$$K(\mathbf{s}) = \int_{\mathbb{R}^{nd-D}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) \, d\tilde{\mathbf{s}}.$$

Moreover, the function  $\widehat{\tilde{K}}$  satisfies the symbol estimates

$$|\widehat{\tilde{K}}(\xi, \tilde{\xi})| \leq \tilde{C}_{\kappa} \|(\xi, \tilde{\xi})\|_{\ell^2}^{-|\kappa|}$$

for all multi-indices  $\kappa$  and all  $(\xi, \tilde{\xi}) \neq \mathbf{0}$ , with  $\tilde{C}_{\kappa}$  depending only on  $C_{\kappa}$ .

Assuming the estimate (4.2) in the case of equal dimensions, let us plug in the kernel  $\tilde{K}$  and the functions  $\tilde{F}_{\mathbf{k}} : (\mathbb{R}^d)^n \rightarrow \mathbb{C}$  defined by

$$\tilde{F}_{\mathbf{k}}(\mathbf{z}, \tilde{\mathbf{z}}) := F_{\mathbf{k}}(\mathbf{z})(D_{\lambda}^{2^n} \varphi_1 \otimes \dots \otimes D_{\lambda}^{2^n} \varphi_n)(\tilde{\mathbf{z}}),$$

where  $\mathbf{z} \in \mathbb{R}^D$ ,  $\tilde{\mathbf{z}} \in \mathbb{R}^{nd-D}$ ,  $F_{\mathbf{k}} : \mathbb{R}^D \rightarrow \mathbb{C}$ ,  $\varphi_i$  is a Schwartz function on  $\mathbb{R}^{d-d_i}$  and  $D_\lambda^{2n} \varphi_i$  is defined by (2.3). Then the form in question becomes

$$\int_{\mathbb{R}^{2nd}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) \left( \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) \times D_\lambda^{2n} \varphi_1(\tilde{x}_1 + k_1 \tilde{s}_1) \cdots D_\lambda^{2n} \varphi_n(\tilde{x}_n + k_n \tilde{s}_n) \right) ds d\tilde{s} dx d\tilde{x}, \tag{4.9}$$

where  $\tilde{\mathbf{s}} = (\tilde{s}_1, \dots, \tilde{s}_n)$ ,  $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_n) \in \mathbb{R}^{nd-D}$ . From the case of equal dimensions we know that the last display is bounded by

$$\prod_{\mathbf{k} \in \{0,1\}^n} \|\tilde{F}_{\mathbf{k}}\|_{L^{2n}(\mathbb{R}^{nd})} = \|\varphi_1 \otimes \cdots \otimes \varphi_n\|_{L^{2n}(\mathbb{R}^{nd-D})} \prod_{\mathbf{k} \in \{0,1\}^n} \|F_{\mathbf{k}}\|_{L^{2n}(\mathbb{R}^D)} \tag{4.10}$$

times a constant depending only on  $(C_\kappa)_\kappa$ . On the other hand, (4.9) equals

$$\int_{\mathbb{R}^{2nd}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) \left( \prod_{i=1}^n \lambda^{-(d-d_i)} \varphi_i(\lambda^{-1} \tilde{x}_i)^{2^{n-1}} \varphi_i(\lambda^{-1}(\tilde{x}_i + \tilde{s}_i))^{2^{n-1}} \right) \times \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) ds d\tilde{s} dx d\tilde{x}. \tag{4.11}$$

Integrating in  $\tilde{\mathbf{x}}$  we obtain

$$\int_{\mathbb{R}^{nd+D}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) \Phi(\lambda^{-1} \tilde{\mathbf{s}}) \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) ds d\tilde{s} dx, \tag{4.12}$$

where we have set

$$\Phi := (\varphi_1^{2^{n-1}} * \tilde{\varphi}_1^{2^{n-1}}) \otimes \cdots \otimes (\varphi_n^{2^{n-1}} * \tilde{\varphi}_n^{2^{n-1}})$$

and  $\tilde{\varphi}_i(s) := \varphi_i(-s)$ . Integrating in  $\tilde{\mathbf{s}}$  and taking the limit as  $\lambda \rightarrow \infty$ , (4.12) becomes, up to a constant,

$$\int_{(\mathbb{R}^D)^2} K(\mathbf{s}) \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) ds dx \tag{4.13}$$

and we know that it is bounded by (4.10), as desired.

To justify passage to the limit we observe that the difference of (4.12) and (4.13) equals

$$\int_{(\mathbb{R}^D)^2} \left( \int_{\mathbb{R}^{nd-D}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) (\Phi(\lambda^{-1} \tilde{\mathbf{s}}) - 1) d\tilde{\mathbf{s}} \right) \prod_{\mathbf{k} \in \{0,1\}^n} F_{\mathbf{k}}(x_1 + k_1 s_1, \dots, x_n + k_n s_n) ds dx,$$

which tends to zero as  $\lambda \rightarrow \infty$ . Indeed, this follows by applying Hölder’s inequality in  $\mathbf{x}$ , which bounds the last display by

$$\left( \int_{\mathbb{R}^D} \left| \int_{\mathbb{R}^{nd-D}} \tilde{K}(\mathbf{s}, \tilde{\mathbf{s}}) (\Phi(\lambda^{-1} \tilde{\mathbf{s}}) - 1) d\tilde{\mathbf{s}} \right| ds \right) \prod_{\mathbf{k} \in \{0,1\}^n} \|F_{\mathbf{k}}\|_{L^{2n}(\mathbb{R}^D)}.$$

We note that the expression in the bracket tends to zero as  $\lambda \rightarrow \infty$ , as desired.

Proof of (b). Note that it suffices to show the bound

$$\left| \int_{\mathbb{R}^{2D+2d_1}} K(\mathbf{s}, s'_1) \left( \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \right) \right. \\ \left. \times F_{\tilde{\mathbf{k}}}(x_1 + l_1 s_1 + l_2 s'_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1 + l_3 s_1 + l_4 s'_1) \right) \\ \times ds ds'_1 d\mathbf{x} dy_1 \Big| \lesssim_{(C_\kappa)_k} \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \|F_{\tilde{\mathbf{k}}}\|_{L^{2^{n+1}}(\mathbb{R}^{(n+1)d})} \quad (4.14)$$

for functions  $F_{\tilde{\mathbf{k}}} : \mathbb{R}^{D+d_1} \rightarrow \mathbb{C}$ , where  $\mathbf{s} = (s_1, \dots, s_n) \in \mathbb{R}^D$ . Indeed, part (b) of Theorem 10 then follows by Fubini, applying the estimate (4.14) and Hölder’s inequality in  $y_2, \dots, y_n$ .

To show (4.14) we again first consider the case of equal dimensions,  $d_1 = \dots = d_n$ . Once again, we write them simply as  $d$ . Changing variables  $x_1 + y_1 + s_1 = x_{n+1}^0$ ,  $x_1 + y_1 + s'_1 = x_{n+1}^1$ , we obtain

$$\int_{\mathbb{R}^{2d(n+1)}} K(x_{n+1}^0 - x_1 - y_1, s_2, \dots, s_n, x_{n+1}^1 - x_1 - y_1) \left( \prod_{k=(k_2, \dots, k_n) \in \{0,1\}^{n-1}} \right. \\ \times F_{(k, e_1)}(x_{n+1}^0 - y_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1) \\ \times F_{(k, e_2)}(x_{n+1}^1 - y_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1) \\ \times F_{(k, e_3)}(x_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, x_{n+1}^0 - x_1) \\ \left. \times F_{(k, e_4)}(x_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, x_{n+1}^1 - x_1) \right) \\ \times dx_{n+1}^0 dx_{n+1}^1 ds_2 \cdots ds_n d\mathbf{x} dy_1,$$

where  $e_i$  are standard unit vectors in  $\mathbb{R}^4$ . Shearing the functions  $F_{(k, e_i)}$  we see that it suffices to show estimate an estimate for the form

$$\int_{\mathbb{R}^{2d(n+1)}} K(x_{n+1}^0 - x_1 - y_1, s_2, \dots, s_n, x_{n+1}^1 - x_1 - y_1) \left( \prod_{k=(k_2, \dots, k_n) \in \{0,1\}^{n-1}} \right. \\ \times F_{(k, e_1)}(x_{n+1}^0, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1) \\ \times F_{(k, e_2)}(x_{n+1}^1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, y_1) \\ \times F_{(k, e_3)}(x_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, x_{n+1}^0) \\ \left. \times F_{(k, e_4)}(x_1, x_2 + k_2 s_2, \dots, x_n + k_n s_n, x_{n+1}^1) \right) dx_{n+1}^0 dx_{n+1}^1 ds_2 \cdots ds_n d\mathbf{x} dy_1.$$

We relabel  $y_1$  into  $x_1^1$ ,  $x_i$  into  $x_i^0$  for  $1 \leq i \leq n$ , and change variables  $x_i^0 + s_i = x_i^1$  for  $2 \leq i \leq n$ . Then we see that it suffices to show the estimate

$$\left| \int_{\mathbb{R}^{2d(n+1)}} \prod_{\mathbf{k} \in \{0,1\}^{n+1}} F_{\mathbf{k}}(\Pi_{\mathbf{k}} x) K(\Pi x) dx \right| \lesssim_{(C_\kappa)_k} \prod_{\mathbf{k} \in \{0,1\}^n} \|F_{\mathbf{k}}\|_{L^{2^{n+1}}(\mathbb{R}^{(n+1)d})}$$

for  $x = (x_1^0, \dots, x_{n+1}^0, x_1^1, \dots, x_{n+1}^1) \in (\mathbb{R}^d)^{n+1}$  and linear operators  $\Pi_{\mathbf{k}}, \Pi : (\mathbb{R}^d)^{2(n+1)} \rightarrow (\mathbb{R}^d)^{n+1}$  given by

$$\begin{aligned} \Pi_{\mathbf{k}}x &:= (x_1^{k_1}, \dots, x_{n+1}^{k_{n+1}}), \\ \Pi x &:= (x_{n+1}^0 - x_1^1 - x_1^0, x_2^1 - x_2^0, \dots, x_n^1 - x_n^0, x_{n+1}^1 - x_1^1 - x_1^0). \end{aligned}$$

This estimate again follows from the main result in [7].

To finish the proof of (4.14) it remains to deduce the case of different dimensions from the case of equal dimensions. This follows similarly as in part (a) and we only sketch the necessary modifications. Let  $d$  be defined as in the proof of part (a) of this theorem. Assuming the estimate (4.14) in the case of equal dimensions, let us plug in the kernel  $\tilde{K}$  defined on  $(\mathbb{R}^d)^{n+1}$  as in (4.8) and the functions  $F_{\tilde{\mathbf{k}}} : (\mathbb{R}^d)^{n+1} \rightarrow \mathbb{C}$  given by

$$\tilde{F}_{\tilde{\mathbf{k}}}(\mathbf{z}, \tilde{\mathbf{z}}) := F_{\tilde{\mathbf{k}}}(\mathbf{z})(D_\lambda^{2^{n+1}} \varphi_1 \otimes \dots \otimes D_\lambda^{2^{n+1}} \varphi_n \otimes D_\lambda^{2^{n+1}} \varphi_1)(\tilde{\mathbf{z}}),$$

where  $\mathbf{z} \in \mathbb{R}^{D+d_1}$ ,  $\tilde{\mathbf{z}} \in \mathbb{R}^{(n+1)d-D-d_1}$ ,  $F_{\tilde{\mathbf{k}}} : \mathbb{R}^{D+d_1} \rightarrow \mathbb{C}$ ,  $\varphi_i$  is a Schwartz function on  $\mathbb{R}^{d-d_i}$  for  $1 \leq i \leq n$ , and  $D_\lambda^{2^{n+1}} \varphi_i$  was defined in (2.3). Then the form in (4.14) becomes, analogously to the display (4.11) in part (a),

$$\begin{aligned} &\int_{\mathbb{R}^{2d(n+1)}} \tilde{K}(\mathbf{s}, s'_1) \lambda^{-(d-d_1)} \varphi_1(\lambda^{-1}\tilde{x}_1)^{2^n} \varphi_1(\lambda^{-1}(\tilde{x}_1 + \tilde{s}_1))^{2^{n-1}} \varphi_1(\lambda^{-1}(\tilde{x}_1 + \tilde{s}'_1))^{2^{n-1}} \\ &\quad \times \left( \prod_{i=2}^n \lambda^{-(d-d_i)} \varphi_i(\lambda^{-1}\tilde{x}_i)^{2^n} \varphi_i(\lambda^{-1}(\tilde{x}_i + \tilde{s}_i))^{2^n} \right) \\ &\quad \times \lambda^{-(d-d_1)} \varphi_1(\lambda^{-1}\tilde{y}_1)^{2^n} \varphi_1(\lambda^{-1}(\tilde{y}_1 + \tilde{s}_1))^{2^{n-1}} \varphi_1(\lambda^{-1}(\tilde{y}_1 + \tilde{s}'_1))^{2^{n-1}} \\ &\quad \times \left( \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \right) \\ &\quad \times F_{\tilde{\mathbf{k}}}(x_1 + l_1s_1 + l_2s'_1, x_2 + k_2s_2, \dots, x_n + k_ns_n, y_1 + l_3s_1 + l_4s'_1) \\ &\quad \times ds ds' d\tilde{s} d\tilde{s}' d\mathbf{x} d\tilde{\mathbf{x}} dy_1 d\tilde{y}_1, \end{aligned} \tag{4.15}$$

where  $\tilde{\mathbf{s}} = (\tilde{s}_1, \dots, \tilde{s}_n) \in \mathbb{R}^{nd-D}$ ,  $\tilde{s}'_1 \in \mathbb{R}^{d-d_1}$ ,  $\tilde{y}_1 \in \mathbb{R}^{d-d_1}$ . From the case of equal dimensions we know that it is bounded by a constant times

$$\prod_{\tilde{\mathbf{k}}} \| \tilde{F}_{\tilde{\mathbf{k}}} \|_{L^{2^{n+1}}(\mathbb{R}^{(n+1)d})} = \| \varphi_1 \otimes \dots \otimes \varphi_n \otimes \varphi_1 \|_{L^{2^{n+1}}(\mathbb{R}^{(n+1)d-D-d_1})} \prod_{\tilde{\mathbf{k}}} \| F_{\tilde{\mathbf{k}}} \|_{L^{2^{n+1}}(\mathbb{R}^{D+d_1})}.$$

On the other hand, integrating in  $\tilde{\mathbf{x}}$  and  $\tilde{y}_1$  gives that the form (4.15) equals

$$\begin{aligned} &\int_{\mathbb{R}^{2d(n+1)}} \tilde{K}(\mathbf{s}, s'_1) \Phi_1(\lambda^{-1}\tilde{s}_1, \lambda^{-1}\tilde{s}'_1)^2 \Phi(\lambda^{-1}(\tilde{s}_2, \dots, \tilde{s}_n)) \\ &\quad \times \left( \prod_{\substack{\tilde{\mathbf{k}}=(k_2, \dots, k_n, l_1, l_2, l_3, l_4) \in \{0,1\}^{n+3} \\ l_1+l_2+l_3+l_4=1}} \right) \\ &\quad \times F_{\tilde{\mathbf{k}}}(x_1 + l_1s_1 + l_2s'_1, x_2 + k_2s_2, \dots, x_n + k_ns_n, y_1 + l_3s_1 + l_4s'_1) \\ &\quad \times ds ds' d\tilde{s} d\tilde{s}' d\mathbf{x} dy_1, \end{aligned}$$

where

$$\Phi_1(\tilde{s}_1, \tilde{s}'_1) := \int_{\mathbb{R}^{d-d_1}} \varphi_i(u)^{2^n} \varphi_i(u + \tilde{s}_1)^{2^{n-1}} \varphi_1(u + \tilde{s}'_1)^{2^{n-1}} du$$

and

$$\Phi := (\varphi_2^{2^n} * \tilde{\varphi}_2^{2^n}) \otimes \cdots \otimes (\varphi_n^{2^n} * \tilde{\varphi}_n^{2^n}),$$

where  $\tilde{\varphi}_i$  is defined as in (a). Taking the limit as  $\lambda \rightarrow \infty$  and integrating in  $\tilde{\mathfrak{s}}, \tilde{\mathfrak{s}}'_1$ , similarly as in the proof of part (a), we recover the form on the left-hand side of (4.14).

*Acknowledgements.* The authors thank Christoph Thiele for inspiring discussions aided by the bilateral DAAD-MZO grant *Multilinear singular integrals and applications*. The authors are also grateful to the anonymous referee for useful suggestions. V. K. was supported in part by the Croatian Science Foundation under the project UIP-2017-05-4129 (MUNHANAP).

#### REFERENCES

- [1] J. BOURGAIN. A Szemerédi type theorem for sets of positive density in  $\mathbb{R}^k$ . *Israel J. Math.* **54** (1986), no. 3, 307–316.
- [2] B. COOK, Á. MAGYAR and M. PRAMANIK. A Roth type theorem for dense subsets of  $\mathbb{R}^d$ . *Bull. London Math. Soc.* **49** (2017), no. 4, 676–689.
- [3] P. DURCIK. An  $L^4$  estimate for a singular entangled quadrilinear form. *Math. Res. Lett.* **22** (2015), no. 5, 1317–1332.
- [4] P. DURCIK, V. KOVAČ and L. RIMANIĆ. On side lengths of corners in positive density subsets of the Euclidean space. *Int. Math. Res. Not.* (2018), no. 22, 6844–6869.
- [5] P. DURCIK, V. KOVAČ, K. A. ŠKREB and C. THIELE. Norm-variation of ergodic averages with respect to two commuting transformations. *Ergodic Theory Dynam. Systems* **39** (2019), no. 3, 658–688.
- [6] P. DURCIK, V. KOVAČ, and C. THIELE. Power-type cancellation for the simplex Hilbert transform. *J. Anal. Math.* **139** (2019), 67–82.
- [7] P. DURCIK and C. THIELE. Singular Brascamp–Lieb inequalities with cubical structure. *Bull. London Math. Soc.* **52** (2020), no. 2, 283–298.
- [8] K. J. FALCONER and J. M. MARSTRAND. Plane sets with positive density at infinity contain all large distances. *Bull. London Math. Soc.* **18** (1986), no. 5, 471–474.
- [9] H. FURSTENBERG and Y. KATZNELSON. An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.* **38** (1978), no. 1, 275–291.
- [10] H. FURSTENBERG, Y. KATZNELSON and B. WEISS. Ergodic theory and configurations in sets of positive density. Mathematics of Ramsey theory, pp. 184–198. *Algorithms Combin.* **5** (Springer, Berlin, 1990).
- [11] V. KOVAČ. Boundedness of the twisted paraproduct. *Rev. Mat. Iberoam.* **28** (2012), no. 4, 1143–1164.
- [12] N. LYALL and Á. MAGYAR. Distance graphs and sets of positive upper density in  $\mathbb{R}^d$ . *Anal. PDE.* **13** (2020), no. 3, 685–700.
- [13] N. LYALL and Á. MAGYAR. Product of simplices and sets of positive upper density in  $\mathbb{R}^d$ . *Math. Proc. Camb. Phil. Soc.* **165** (2018), no. 1, 25–51.
- [14] N. LYALL and Á. MAGYAR. Weak hypergraph regularity and applications to geometric Ramsey theory. Preprint (2019).
- [15] E. M. STEIN. Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. *Princeton Math. Ser.* **43** (Princeton University Press, Princeton, 1993).
- [16] P. VARNAVIDES. On certain sets of positive density. *J. London Math. Soc.* **34** (1959), 358–360.

