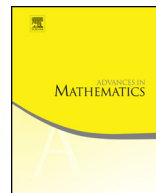




ELSEVIER

Contents lists available at ScienceDirect

Advances in Mathematics

journal homepage: www.elsevier.com/locate/aim

Maximal subspace averages

Francesco Di Plinio^{a,1}, Ioannis Parissis^{b,c,*,2}^a *Dipartimento di Matematica e Applicazioni “R. Caccioppoli”, Università di Napoli “Federico II”, Via Cintia, Monte S. Angelo 80126 Napoli, Italy*^b *Departamento de Matemáticas, Universidad del País Vasco, Aptdo. 644, 48080 Bilbao, Spain*^c *Ikerbasque, Basque Foundation for Science, Bilbao, Spain*

ARTICLE INFO

Article history:

Received 8 July 2021

Received in revised form 27

September 2022

Accepted 7 October 2022

Available online 18 October 2022

Communicated by C. Fefferman

MSC:

primary 42B25

secondary 42B20

Keywords:

Directional operators

Nikodym sets

Keakeya problem

Zygmund’s conjecture

ABSTRACT

We consider maximal operators associated to singular averages along finite subsets Σ of the Grassmannian $\text{Gr}(d, n)$ of d -dimensional subspaces of $\mathbb{R}^n$. The well studied $d = 1$ case corresponds to the directional maximal function with respect to arbitrary finite subsets of $\text{Gr}(1, n) = \mathbb{S}^{n-1}$. We provide a systematic study of all cases $1 \leq d < n$ and prove essentially sharp $L^2(\mathbb{R}^n)$ bounds for the maximal subspace averaging operator in terms of the cardinality of Σ , with no assumption on the structure of Σ . In the codimension 1 case, that is $n = d + 1$, we prove the precise critical weak $(2, 2)$ -bound.

Drawing on the analogy between maximal subspace averages and (d, n) -Nikodym maximal averages, we also formulate the appropriate maximal Nikodym conjecture for general $1 < d < n$ by providing examples that determine the critical L^p -space for the (d, n) -Nikodym problem. Unlike the $d = 1$ case, the maximal Keakeya and Nikodym problems are shown not to be equivalent when $d > 1$. In this context, we prove the best possible $L^2(\mathbb{R}^n)$ -bound for the (d, n) -Nikodym maximal function for all combinations of dimension and codimension. Our estimates rely on Fourier analytic almost orthogonality principles, combined with polynomial partitioning, but we

* Corresponding author.

E-mail addresses: francesco.diplinio@unina.it (F. Di Plinio), ioannis.parissis@ehu.es (I. Parissis).¹ F. Di Plinio has been partially supported by the National Science Foundation under the grants NSF-DMS-2000510, NSF-DMS-2054863.² I. Parissis is partially supported by the project PGC2018-094528-B-I00 (AEI/FEDER, UE) with acronym “IHAIP”, project PID2021-122156NB-I00/AEI/10.13039/501100011033 funded by Agencia Estatal de Investigación with acronym “HAMIP”, grant IT1615-22 of the Basque Government and IKERBASQUE.

also use spatial analysis based on the precise calculation of intersections of d -dimensional plates in $\mathbb{R}^n$.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

For $\sigma \in \text{Gr}(d, n)$, the Grassmannian of d -dimensional subspaces of $\mathbb{R}^n$, the scale s subspace average of $f \in \mathcal{C}(\mathbb{R}^n)$ is

$$\langle f \rangle_{s, \sigma}(x) := \int_{sB_n \cap \sigma} f(x - y) \frac{dy}{s^d}, \quad x \in \mathbb{R}^n,$$

where $B_n \subset \mathbb{R}^n$ is the n -dimensional unit ball centered at the origin, sB_n is its concentric dilate, and $dy = d\mathcal{L}^d(y)$ denotes the Lebesgue measure on $\sigma \in \text{Gr}(d, n)$. Fubini's theorem ensures that, up to the dimensional constant $\mathcal{L}^d(B_d)$, the map $f \mapsto \langle f \rangle_{s, \sigma}$ preserves the $L^1(\mathbb{R}^n)$ -norm of f and contracts all $L^p(\mathbb{R}^n)$ -norms for $1 < p \leq \infty$.

The general concern of this article is the L^p -behavior of the positive maps

$$f \mapsto \langle f \rangle_{s(\cdot), \sigma(\cdot)}(\cdot)$$

corresponding to a measurable choice of $\sigma \in \text{Gr}(d, n)$ and scale $s > 0$ depending on the point $x \in \mathbb{R}^n$. When $d = n$, these maps are pointwise controlled by the standard Hardy-Littlewood maximal operator for any pair of choice functions σ, s . The singular cases $d < n$ give rise to a family of nontrivial problems of intrinsic relevance within the theory of differentiation of integrals, and possessing applications to singular and oscillatory operators, geometric measure theory and partial differential equations.

A central example is the classical question, attributed to Zygmund, of characterizing the class of planar vector fields that differentiate $L^2(\mathbb{R}^2)$ functions. The corresponding singular integral variant of Zygmund's question, usually attributed to Stein, asks whether Lipschitz vector-fields of directions allow for a weak $(2, 2)$ bound for the corresponding directional Hilbert transform, after suitable truncation.

The case of choice functions $\sigma(\cdot)$ whose range is a *finite* subset $\Sigma \subset \text{Gr}(d, n)$ is also of particular importance. The averaging operator is of maximal nature, that is for $s \in S \subset (0, \infty)$

$$M_{\Sigma, \{s\}} f := \sup_{\sigma \in \Sigma} \langle |f| \rangle_{s, \sigma}(\cdot), \quad M_{\Sigma, S} f := \sup_{s \in S} M_{\Sigma, \{s\}} f.$$

The study of one-dimensional directional averages in $\mathbb{R}^n$, corresponding to $\text{Gr}(1, n)$, is connected to the problem of determining the Hausdorff dimension of Kakeya sets in $\mathbb{R}^n$. There is a classification of such questions for all $d < n$, with $\text{Gr}(d, n)$ corresponding to the

problem of studying the existence of (d, n) Besicovitch sets. For the study of directional averages as above in relation to the Kakeya-Besicovitch category of problems, it suffices to consider finite subsets Σ which are uniformly distributed at some distinct scale δ . More specifically, the range Σ of $\sigma(\cdot)$ is a maximal δ -net in $\text{Gr}(d, n)$. This in turn leads to seeking for L^p -bounds for the corresponding maximal directional averages as a function of δ .

1.1. Thin (d, n) -averages for arbitrary finite $\Sigma \subset \text{Gr}(d, n)$

The case of more general maximal directional averages, where σ takes values in a finite but *arbitrary* subset $\Sigma \subset \text{Gr}(d, n)$, is in general much harder as there is no distinct scale in the set of directions, and any suitable method must make up for the lack of uniform density in Σ . This obstruction is already present in the *single scale problem*. Our first main result is an essentially complete description of the sharp $L^2(\mathbb{R}^n)$ -bounds in the whole range of dimensions and codimensions $1 \leq d < n$.

Theorem A. *For all $N > 0$, $1 \leq d < n$ there holds*

$$\sup_{s>0} \sup_{\substack{\Sigma \subset \text{Gr}(d,n) \\ \#\Sigma \leq N}} \|M_{\Sigma, \{s\}}\|_{L^2(\mathbb{R}^n)} \lesssim N^{\frac{(n-d-1)}{2(n-d)}} (\log N)^{\frac{n-d}{2}} \quad (1.1)$$

with an implicit numerical constant depending only upon d, n . This bound is sharp in terms of N when $n = d + 1$ and sharp up to the logarithmic factor in general.

For the special case $d = 1$ and $n = 3$, we obtain an improved version of Theorem A where the corrective logarithmic term appearing in the right hand side of (1.1) is replaced by an arbitrarily iterated logarithm of N . The precise statement is given in Theorem F of Section 5.

Before proceeding with the description of our second group of results, some remarks are in order. The case $d = 1, n = 2$ is due to Katz, [18]. It should be noted that two-dimensional versions of the theorem above are somewhat related to the resolution of the Kakeya conjecture in two dimensions and to earlier results of Córdoba [6,8] and Strömberg [29]. We note here that Theorem A recovers the sharp result of Katz, [18], for single scale maximal directional averages on $L^2(\mathbb{R}^2)$, so necessarily $d = 1$, in fact with a new and independent proof. Note that Theorem A is a single-scale but with $\Sigma \subset \text{Gr}(d, n)$ *arbitrary*. The $d = 1, n = 3$ improved result of Theorem F is itself an amelioration of [12, Theorem B]. The proof technique for Theorem F combines the polynomial partitioning ideas of [12] with a new algebraic almost-orthogonality result, Theorem E, which is of independent interest and may be seen as a higher-dimensional analogue of the well-known Alfonseca-Soria-Vargas almost-orthogonality principle [2].

1.2. Sharp critical estimates for multiscale $(d, d+1)$ -averages

In the cases of maximal directional averages in arbitrary dimension and codimension $1 = n - d$, we give a final theorem that provides the best possible L^p -bounds, $p \geq 2$, for the corresponding maximal averages with respect to arbitrary finite $\Sigma \subset \text{Gr}(d, d+1)$. It should be noted that in this case we can actually tackle the multiscale problem in a sharp fashion. An important feature shared by the codimension 1 problems is that the critical integrability space for the maximal function $M_{\Sigma, S}$ is $L^2(\mathbb{R}^n)$.

Theorem B. *Let $n = d + 1$ and $\Sigma \subset \text{Gr}(d, d+1)$ be a finite set. Then*

$$\|M_{\Sigma, (0, \infty)} : L^2(\mathbb{R}^n) \rightarrow L^{2, \infty}(\mathbb{R}^n)\| \lesssim (\log \#\Sigma)^{\frac{1}{2}}$$

and

$$\|M_{\Sigma, (0, \infty)} : L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)\| \lesssim \begin{cases} (\log \#\Sigma)^{\frac{1}{p}}, & p > 2, \\ \log \#\Sigma, & p = 2. \end{cases}$$

These bounds are best possible in terms of the dependence on the cardinality $\#\Sigma$.

For the sharpness of the obtained bounds, we send to Proposition 5.4. The main tool in the proof of the upper bounds is a directional Carleson embedding theorem for suitable Carleson sequences indexed by δ -plates, see Theorem I in Section 7. More specifically, Theorem B is obtained by an application of Theorem I to the adjoint of the (linearized) maximal operator $M_{\Sigma, S}$. More general applications of the directional Carleson theorem are revealed by couplings with time-frequency analysis. Along these lines, as a representative sample of the scope of Theorem I, we derive from it a sharp estimate for the Rubio de Francia square function associated to N well-distributed conical sectors in $\mathbb{R}^n$, see Theorem J.

1.3. L^2 -estimates for the (d, n) -Nikodym maximal operator

With precise definitions and discussion to come later, we define the (d, n) -Nikodym maximal function on $\mathbb{R}^n$ by

$$\mathcal{N}_\delta f(x) := \sup_{\sigma \in \text{Gr}(d, n)} \int_{x+T_\delta(\sigma)} |f(y)| \, dy, \quad x \in \mathbb{R}^n, \quad (1.2)$$

where $T_\delta(\sigma)$ is a δ -neighborhood of $\sigma \cap B_n(1)$; see (2.3). This is the maximal operator coupled with the dimensional analysis of (d, n) -Nikodym sets for general $1 \leq d < n$. As discussed in §6, suitable $L^p(\mathbb{R}^n)$ -bounds for $\mathcal{N}_\delta$ imply corresponding lower bounds for the Hausdorff dimension of (d, n) -Nikodym sets. Our main result here is the following

Theorem C. Let $1 \leq d < n$ and $\delta > 0$. There holds

$$\|\mathcal{N}_\delta : L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)\| \lesssim \begin{cases} \delta^{-\frac{n-d-1}{2}}, & 1 \leq d < n-1, \\ \sqrt{\log(\delta^{-1})}, & d = n-1, \end{cases}$$

and

$$\|\mathcal{N}_\delta : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)\| \lesssim \begin{cases} \delta^{-\frac{n-d-1}{2}} \sqrt{\log \delta^{-1}}, & 1 \leq d < n-1, \\ \log(\delta^{-1}), & d = n-1. \end{cases}$$

The weak $(2,2)$ bound is best possible while the strong $(2,2)$ bound is best possible up to the logarithmic factor $\sqrt{\log(\delta^{-1})}$.

1.4. Motivation and background

The study of maximal directional averages has a long history, motivated for instance by the Kakeya, Zygmund and Stein conjectures mentioned above. Maximal averages along subsets of $\text{Gr}(d, n)$ are also quantitatively connected with the behavior of the ball and polygon multipliers and the convergence of Fourier series in higher dimensions, as well as to square functions formed by frequency projections to the corresponding dual subsets of the frequency domain. This last theme is explored in detail in [1] in dimension $n = 2$. As mentioned above, we present an application of this type, for arbitrary dimension n , in §7.8 below.

The study of $\text{Gr}(d, n)$ -averages can be classified according to the structure of $\Sigma \subset \text{Gr}(d, n)$. This classification is more established in the case $d = 1$, where $\Sigma \subset \mathbb{S}^{n-1}$ and three particular cases of interest arise.

First, lacunary sets [7,21,27,28] are the only infinite subsets of $\mathbb{S}^{n-1}$ that give rise to bounded directional maximal functions. This characterization is due to Bateman in $n = 2$, and albeit in a weaker form, to Parcet and Rogers [24], and relies ultimately on the existence of Kakeya and Nikodym sets. Second, sharp $L^p(\mathbb{R}^n)$ -bounds for maximal averages along δ -uniform sets of directions $\Sigma_\delta \subset \mathbb{S}^{n-1}$ are the subject of the maximal Nikodym conjecture, whose lower bounds tells us *a fortiori* that $\|\mathcal{M}_{\Sigma_\delta, \{1\}}\|_{L^p(\mathbb{R}^n)}$ cannot be independent of δ for any $p \in (1, \infty)$. In fact, the maximal Nikodym conjecture is formally weaker than statements involving $\mathcal{M}_{\Sigma_\delta, \{1\}}$, dealing with averages along δ -tubes instead of *thin*, δ -uniformly spaced averages of the form $\mathcal{M}_{\Sigma_\delta, \{1\}}$.

In two dimensions the sharp bounds for $\mathcal{M}_{\Sigma_\delta, S}$ are known, see for example [6,8,29]. In higher dimensions $n > 3$ the Kakeya conjecture is open but several partial and very significant results are available; see for example [16] and the references therein. Best possible $L^2(\mathbb{R}^n)$ bounds for multiscale averages along uniformly distributed sets in $\mathbb{S}^{n-1}$ recently appeared in [19].

The problem of sharp L^p -bounds for $\mathcal{M}_{\Sigma, S}$ when $\Sigma \subset \mathbb{S}^{n-1}$ is instead arbitrary is also fully solved when $n = 2$ [17,18]. However, in particular when $n > 2$, it is in general

much harder than the δ -net case, as there is no distinct scale on the set of directions, or alternatively, there is no fixed density of directions. This lack of structure does not allow for e.g. the use of induction on scales, and new tools are necessary. Recently, we proved in [12] essentially sharp $L^2(\mathbb{R}^n)$ -bounds for $M_{\Sigma, \{1\}}$ for arbitrary finite $\Sigma \subset \mathbb{S}^{n-1}$ or $\Sigma \subset Z$ with Z algebraic submanifold of $\mathbb{S}^{n-1}$, via the polynomial method.

Moving away from the restriction $d = 1$, we introduce in this paper a family of problems related to averages with respect to $\text{Gr}(d, n)$ for general $1 \leq d < n$. Such Radon-type transforms have been studied in several forms in the literature, for example for the (d, n) -Kakeya maximal function

$$\mathcal{K}f(L) := \sup_{x \in \mathbb{R}^n} \int_{L+x} f(y) \, dy, \quad L \in \text{Gr}(d, n),$$

where f is a suitable function in $\mathbb{R}^n$. In analogy to the case $d = 1$, suitable $L^p(\mathbb{R}^n) \rightarrow L^q(\text{Gr}(d, n))$ bounds for the (d, n) -Kakeya operator, or for corresponding d -plane transforms, relate to the existence and dimension of (d, n) -Besicovitch sets; see [20, §24], [22, 23], and the discussion in Section 6. Our point of view is different, unifying the study of d -plane averages for all $1 \leq d < n$ in the form of *thin* subspace averages $M_{\Sigma, \{s\}}$, possibly at different scales s , as operators acting on $L^p(\mathbb{R}^n)$. As in the case $d = 1$ the structure of the subset $\Sigma \subset \text{Gr}(d, n)$ under consideration is of paramount importance. In this light, our setup is new and, for example, there is currently no definition of lacunary subsets of $\text{Gr}(d, n)$. In this paper we thoroughly study the cases of arbitrary $\Sigma \subset \text{Gr}(d, n)$. The case of uniformly distributed subsets $\Sigma_\delta \subset \text{Gr}(d, n)$ is also important and is implicit in the study of the (d, n) -Kakeya and Nikodym maximal operators in Section 6 and more precisely in the statements of Proposition 6.2 and Theorem C. There is again a critical integrability space $L^{p(d, n)}(\mathbb{R}^n)$ for $M_{\Sigma_\delta, \{1\}}$ relating to the problem of existence and dimension of (d, n) -Nikodym sets.

An important difference is that for $d > 1$ the (d, n) -Kakeya conjecture and the (d, n) -Nikodym conjecture appear to be independent of each other and so are the corresponding critical exponents $p(d, n)$. This is in stark contrast with the case $d = 1$ where the maximal Kakeya and maximal Nikodym conjectures are equivalent; see [30]. For example a well known result of Falconer, [13], implies the there are no (d, n) Besicovitch sets for $d > n/2$ while another result of Falconer, [14], shows that there exist (d, n) -Nikodym sets for all $1 \leq d < n$. This difference is also reflected in the fact that, unlike the case $d = 1$, the possibility of δ -free bounds for the (d, n) -Kakeya operator $\mathcal{K}$, defined in (6.1) below, is not excluded. An instance of this is contained in the statement of Proposition 6.2. On the other hand this is not the case for the (d, n) -Nikodym maximal operator, nor for $M_{\Sigma, \{s\}}$, as exhibited in Theorem C, in accordance to the previously mentioned result of Falconer. A more general study of d -plane averages of the form $M_{\Sigma, S}$ and of (d, n) -Nikodym maximal operators for $1 < d < n$ is motivated by these connections and nuances. We introduce the maximal multiscale problem with $S = (0, \infty)$ and generic $\Sigma \subset \text{Gr}(d, n)$ and manage to fully resolve the problem in the codimension $1 = n - d$

case, in the form of Theorem B above. At the same time we formulate a general (d, n) -Nikodym conjecture and discuss the critical integrability index for general d . From that point of view Theorem C stated above describes a sharp but subcritical estimate for the (d, n) -Nikodym function, and corresponding maximal conjecture.

The investigations in the current paper lead to several natural questions concerning the $L^p(\mathbb{R}^n)$ -bounds for d -subspace averages $M_{\Sigma, S}$, and especially the study of such L^p -norms close to the critical exponent $p = p(d, n)$ is particularly interesting, and consistently hard. The current paper addresses in particular all the $L^2(\mathbb{R}^n)$ -bounds for such operators in a sharp fashion, whether $L^2(\mathbb{R}^n)$ happens to be subcritical as in the case $n > 2$, $d < n - 1$, or critical as in the case $d = n - 1$ in any dimension.

1.5. Methodology

In this paper we employ a mix of geometric, Fourier analytic, and polynomial methods. The latter technique for the study of directional maximal operators along arbitrary sets of directions was introduced in [12]. Using a polynomial partition we divide the set of directions into subsets (cells) of controlled cardinality and such that the boundary of these subsets is an algebraic variety of controlled degree. The properties of this partition allow us to prove an almost orthogonality principle via Fourier methods, as the algebraic nature of the boundary of the cells (wall) yields suitable overlap estimates for the relevant Fourier multipliers. Using this scheme we prove a general almost orthogonality principle for single scale directional averages defined with respect to an arbitrary set of directions. In the current paper we apply this principle to yield essentially sharp bounds for maximal $d = 1$ -dimensional averages given by arbitrary directions on the sphere. The application of the polynomial partitioning scheme to the case of d -dimensional averages in $\mathbb{R}^n$ will require a suitable polynomial partition on the Grassmannian $\text{Gr}(d, n)$ and will be taken up in a future work. However, for general codimension $n - d$ we present an alternative argument that recovers almost sharp L^2 -bounds for d -dimensional single-scale averages given by *arbitrary* subsets $\Sigma \subset \text{Gr}(d, n)$. Indeed this approach misses the conjectured sharp L^2 -bound, which is polynomial in $\#\Sigma$, by a logarithmic factor in $\#\Sigma$.

In two cases, we employ a different point of view in order to obtain sharp $L^2(\mathbb{R}^n)$ -bounds via space analysis. This is particular efficient when proving $L^2(\mathbb{R}^n)$ -bounds for maximal directional d -plane averages with $n = d + 1$. It is important to note that for this codimension-1 case, the space $L^2(\mathbb{R}^n)$ is critical and we do prove the best possible bound in Theorem B, in fact even for the multiscale maximal function. The approach, inspired by the works of Katz [17] and Bateman [4], is via a TT^* -argument on the adjoint of the linearized maximal operator and an appeal to a suitable directional Carleson embedding theorem. In [1] this method was elaborated into a directional Carleson embedding theorem for suitable directional Carleson sequences. Here we suitably adapt the geometric part of the argument, resulting in a corresponding directional Carleson embedding for sequences indexed by $(n - 1)$ -dimensional plates in $\mathbb{R}^n$ and satisfying a Carleson condition adjusted to the geometry of such plates.

A second TT^* instance appears in the proof of sharp $L^2(\mathbb{R}^n)$ -bounds for the $\text{Gr}(d, n)$ -Nikodym maximal operator. Here we are able to exploit specific structure of the nets Σ_δ and prove explicit estimates for the volume of the intersections of such δ -plates in all combinations of dimension and codimension. These volume estimates and the TT^* argument yield the sharp bound for the (d, n) -Nikodym maximal operator of Theorem C.

1.6. Notation

The purpose of this paragraph is to provide easy reference for a few central definitions, in particular for the several maximal and averaging operators that appear throughout the paper.

- The notation $B_k(z)$ is reserved for the unit ball in $\mathbb{R}^k$ centered at $z \in \mathbb{R}^k$, and we write $B_k := B_k(0)$.
- If σ is a subspace of $\mathbb{R}^n$, we denote by Π_σ the corresponding orthogonal projection. For $v \in \mathbb{S}^{n-1}$ we abuse notation and write Π_v instead of $\Pi_{\text{span}\{v\}}$.
- The notation M is reserved for the Hardy-Littlewood maximal operator on the corresponding $\mathbb{R}^n$.
- For $\Sigma \subset \text{Gr}(d, n)$ and $S \subset (0, \infty)$, $M_{\Sigma, S}$ stands for the maximal averaging operator with respect to *thin* plates $\sigma \cap B_n(s)$ with $\sigma \in \Sigma$ and choice of scale $s \in S$. In symbols,

$$M_{\Sigma, S}f(x) := \sup_{\substack{s \in S \\ \sigma \in \Sigma}} \langle |f| \rangle_{\sigma, s}(x), \quad x \in \mathbb{R}^n.$$

The case $S = \{s\}$ for some $s > 0$ is the single-scale case and will appear in several places below. When $S = \{1\}$ we simplify the notation to $M_\Sigma f := M_{\Sigma, \{1\}}f$.

- The smooth, compactly Fourier supported version of $M_{\Sigma, S}$, denoted by $A_{\Sigma, S}$, and defined in (2.4), will be used throughout the paper.
- Also, given $\delta > 0$ we consider the (d, n) -Nikodym maximal operator $\mathcal{N}_\delta$ which is a maximal –with respect to σ – average along plates $x + \sigma \cap B_n(1)$, $x \in \mathbb{R}^n$, oriented along any $\sigma \in \text{Gr}(d, n)$ and having thickness δ in the $\sigma^\perp$ -directions, see (1.2).

All of the above operators are functions defined on $\mathbb{R}^n$ and we will be proving $L^p(\mathbb{R}^n) \rightarrow L^q(\mathbb{R}^n)$ operator norm-bounds.

1.7. Structure of the article

In Section 2, we collect a few definitions related to the Grassmannian and its distance, and develop a technical subspace switch lemma for the Fourier version of our averages. Section 3 uses the switch lemma to give a new and simple proof of the L^2 -almost-orthogonality principle of [2] for maximal directional operators in the plane. The

argument of Section 3 serves as a model for the more complex Section 4, where an algebraic almost orthogonality principle in arbitrary dimension, of similar flavor, is proved. Section 5 contains the proof of Theorems A and F. In Section 6 we discuss the Nikodym analogue of the maximal function M_Σ , formulating the relevant maximal conjecture and proving the L^2 case. Finally, Section 7 is dedicated to the full solution of the codimension 1 case via a subspace Carleson embedding theorem and to the application of the latter to the Rubio de Francia estimate for conical cutoffs in $\mathbb{R}^n$.

2. Grassmannian, Fourier averages and switch lemmas

This section contains a few definitions and technical lemmas that will be used throughout the paper.

2.1. Grassmannian

We write $\text{Gr}(d, n)$ for the Grassmannian of d -dimensional subspaces of $\mathbb{R}^n$. If $O(n)$ stands for the orthogonal group on $\mathbb{R}^n$ then

$$\text{Gr}(d, n) = O(n) \setminus [O(d) \otimes O(n - d)], \quad (2.1)$$

identifying each subspace $\sigma \in \text{Gr}(d, n)$ with the orthogonal map sending the first d canonical vectors onto an orthonormal basis of σ . In particular $\text{Gr}(d, n)$ is a smooth algebraic variety of dimension $d(n - d)$. Equipped with the metric

$$d(\sigma, \tau) := \sup_{v \in \mathbb{S}^{n-1}} |\sigma v - \tau v|, \quad \sigma, \tau \in \text{Gr}(d, n),$$

the Grassmanian $\text{Gr}(d, n)$ can be viewed as a compact metric space. For $\delta > 0$ and $\sigma \in \text{Gr}(d, n)$ we denote by $B_\delta(\sigma) := \{\tau \in \text{Gr}(d, n) : d(\sigma, \tau) < \delta\}$ the open δ -ball centered at $\sigma \in \text{Gr}(d, n)$.

In analogy with the classical Kakeya-Nikodym directional maximal functions, we will consider below the maximal subspace averages along δ -separated subsets $\Sigma \subset \text{Gr}(d, n)$. We say $\Sigma \subset \text{Gr}(d, n)$ is δ -separated if $\{B_\delta(\sigma) : \sigma \in \Sigma\}$ is a collection of pairwise disjoint sets. We will need the following lemma concerning the cardinality of the subset of a δ -separated set Σ consisting of subspaces which are δ -approximately orthogonal to some $\xi \in \mathbb{R}^n$. We will see that these belong to a δ -neighborhood of the Grassmanian hyperplane

$$H_\xi(d) := \{\tau \in \text{Gr}(d, n) : \Pi_\tau \xi = 0\}. \quad (2.2)$$

Notice that $H_\xi(d)$ is linearly isomorphic to $\text{Gr}(d, n - 1)$.

Lemma 2.2. *Let $\xi \in \mathbb{R}^n \setminus \{0\}$, $\Sigma \subset \text{Gr}(d, n)$, $\delta > 0$ and $\Sigma_\xi := \left\{ \sigma \in \Sigma : \frac{|\Pi_\sigma \xi|}{|\xi|} < \frac{\delta}{4} \right\}$. Then*

1. The set Σ_ξ is contained in the $\frac{\delta}{3}$ -neighborhood of $H_\xi(d)$.
2. If Σ is a δ -separated set, then $\#\Sigma_\xi \lesssim \delta^{-d(n-d-1)}$.

Proof. We first prove claim 2 assuming claim 1. Indeed, assuming 1, for each $\sigma \in \Sigma_\xi$ we may pick $a_\sigma \in H_\xi(d)$ with $d(a_\sigma, \sigma) < \frac{\delta}{3}$. By the triangle inequality $\Sigma \ni \sigma, \tau, \sigma \neq \tau \implies d(a_\sigma, a_\tau) \geq \frac{\delta}{3}$. Therefore, the set $\{a_\sigma : \sigma \in \Sigma_\xi\} \subset H_\xi(d)$ has at most $\sim \delta^{-d(n-d-1)}$ elements by dimensionality of $H_\xi(d)$. This completes the proof of the lemma up to establishing the claim.

We now prove claim 1. By rotation invariance of the statement it suffices to prove the claim for $\xi = e_1$. Let $\sigma \in \Sigma_\xi$. Let $u = |\Pi_\sigma e_1|$. If $u = 0$ there is nothing to prove, which means we may work with $0 < u < \delta/4$. Pick an orthonormal basis $\{b_1, \dots, b_d\}$ of σ with $b_1 = \Pi_\sigma e_1 / |\Pi_\sigma e_1|$. Then $b_2, \dots, b_d \in e_1^\perp \cap b_1^\perp$. Let

$$c_1 := \frac{b_1 - ue_1}{|b_1 - ue_1|}.$$

Then $a_\sigma := \text{span}\{c_1, b_2, \dots, b_d\} \in H_\xi(d)$ and as $|c_1 - b_1| < \frac{\delta}{3}$, we have shown that $d(a_\sigma, \sigma) < \frac{\delta}{3}$ as claimed. $\square$

2.3. Fourier averages and switch lemmas

If $0 \leq \delta \ll 1$ and $\sigma \in \text{Gr}(d, n)$ then

$$T_\delta(\sigma) := T_\delta^1(\sigma) := \{\xi \in \mathbb{R}^n : |\Pi_\sigma \xi| \leq 1, |\Pi_{\sigma^\perp} \xi| < \delta\} \quad (2.3)$$

will stand for the unit scale d -dimensional δ -plate oriented along σ . In general, we think of $T_\delta(\sigma)$ as a slight fattening of the unit ball B_d on σ . Throughout the paper a few slightly different versions of this fattening will be employed depending on the problem being considered.

Let $\phi_d \in \mathcal{S}(\mathbb{R}^d)$ be a real valued even function with support in $2^{-8}B_d$, and $\|\phi_d\|_{L^1(\mathbb{R}^d)} = 1$. For $\sigma \in \text{Gr}(d, n)$ and $s > 0$ we define the smooth subspace averages and maximal averages of $f \in \mathcal{S}(\mathbb{R}^d)$ by

$$A_{\sigma, s} f(x) := \int_{\mathbb{R}^n} \widehat{f}(\xi) \phi_d(s \Pi_\sigma \xi) e^{ix \cdot \xi} d\xi, \quad A_{\Sigma, S} f(x) := \sup_{\substack{s \in S \\ \sigma \in \Sigma}} |A_{\sigma, s} f(x)|, \quad x \in \mathbb{R}^n. \quad (2.4)$$

When $\sigma = \mathbb{R}v$ for some $v \in \mathbb{S}^{n-1}$, we write $A_{v, s}$ in place of $A_{\sigma, s}$. The easily verified fact

$$\|M_{\Sigma, S}\|_{L^2(\mathbb{R}^n)} = \sup_{k \in \mathbb{Z}} \|M_{\Sigma, 2^k S}\|_{L^2(\mathbb{R}^n)} \sim \sup_{k \in \mathbb{Z}} \|A_{\Sigma, 2^k S}\|_{L^2(\mathbb{R}^n)} = \|A_{\Sigma, S}\|_{L^2(\mathbb{R}^n)} \quad (2.5)$$

will be used in what follows, often in combination with (2.10) below.

Fix a radial function $\Phi \in \mathcal{S}(\mathbb{R}^n)$ with $\mathbf{1}_{B_n} \leq \Phi \leq \mathbf{1}_{2B_n}$ and for $\delta > 0$ introduce the low-high splitting

$$A_{\sigma,s}f = A_{\sigma,s}^{>\delta}f + A_{\sigma,s}^{<\delta}f, \quad \widehat{A_{\sigma,s}^{>\delta}f}(\xi) := \widehat{A_{\sigma,s}f}(\xi)\Phi(4s\delta\xi), \quad \xi \in \mathbb{R}^n. \quad (2.6)$$

The notation is motivated by the fact that $A_{\sigma,s}^{>\delta}$ is a smooth average at spatial scale s in the directions of σ , and scale $s\delta$ in the directions orthogonal to σ , and consequently containing frequency scales at most $1/(s\delta)$. With this in mind, the parameter $0 < \delta \leq 1$ in the following lemma measures the eccentricity of the plates with long directions oriented along the subspace τ appearing in the averages on the right hand side.

Lemma 2.4. *Let $s > 0$, $\sigma \neq \tau \in \text{Gr}(d, n)$ and $\delta \in [d(\sigma, \tau), 1]$ be given. Then*

$$|A_{\sigma,s}^{>\delta}f(x)| \lesssim \sum_{k=0}^{\infty} 2^{-kn} \int_{x+2^k s T_{\delta}(\tau)} |f|, \quad (2.7)$$

$$|A_{\sigma,s}^{>\delta}f(x) - A_{\tau,s}^{>\delta}f(x)| \lesssim \frac{d(\sigma, \tau)}{\delta} \sum_{k=0}^{\infty} 2^{-kn} \int_{x+2^k s T_{\delta}(\tau)} |f|. \quad (2.8)$$

Proof. The case $\sigma = \tau$ of (2.7) follows easily by the Schwartz decay of the smooth function ϕ_d used in the definition of $A_{\sigma,s}$. This means that the general case of (2.7) is an immediate consequence of (2.8), which we now prove. By isotropic scaling and rotational invariance it suffices to treat the case $s = 1$, $\tau = \text{span}\{e_1, \dots, e_d\}$. Let

$$\zeta(\xi, y) := \Pi_{\tau}\xi + y(\Pi_{\sigma} - \Pi_{\tau})\xi, \quad y \in [0, 1].$$

The Fourier transform of the integral kernel K of $A_{\sigma,1}^{>\delta} - A_{\tau,1}^{>\delta}$ is given by

$$\widehat{K}(\xi) = \Phi(4\delta\xi) \int_0^1 \nabla \phi_d(\zeta(\xi, y)) \cdot [(\Pi_{\sigma} - \Pi_{\tau})\xi] dy$$

and is supported in a 1-neighborhood of $\sigma^{\perp} \cap 2\delta^{-1}B_n$, which has measure $\sim \delta^{d-n}$. It also satisfies for each multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$

$$|\partial^{\alpha} \widehat{K}(\xi)| \lesssim d(\sigma, \tau) \delta^{\alpha_{d+1} + \dots + \alpha_n - 1}. \quad (2.9)$$

This estimate is obtained via repeated use of the Leibniz rule and the following considerations:

$$|\text{supp } \widehat{K}| \lesssim \delta^{d-n}, \quad \sup_{y \in [0,1]} \sup_{d+1 \leq j \leq n} |\partial_j \zeta(\xi, y)| \lesssim d(\sigma, \tau) \leq \delta, \quad \|\Pi_{\sigma} - \Pi_{\tau}\| = d(\sigma, \tau).$$

Integration by parts then readily yields

$$|K(x)| \lesssim \frac{d(\sigma, \tau)}{\delta^{n-d+1}} [1 + |\Pi_{\tau}x|]^{-4n} \left[1 + \frac{|\Pi_{\tau^{\perp}}x|}{\delta} \right]^{-4n}, \quad x \in \mathbb{R}^n,$$

which in turn implies (2.8). The proof is thus complete. $\square$

Remark 2.5. If M is the standard maximal function in $\mathbb{R}^n$, a simple averaging argument leads from (2.7) and (2.8) to

$$|A_{\sigma,s}^{>\delta} f(x)| + \frac{\delta}{d(\sigma, \tau)} |A_{\sigma,s}^{>\delta} f(x) - A_{\tau,s}^{>\delta} f(x)| \lesssim \sum_{k=0}^{\infty} 2^{-kn} M[\langle |f| \rangle_{\tau, 2^k s}](x), \quad x \in \mathbb{R}^n. \quad (2.10)$$

The previous lemma suggests that the low frequencies $|\xi| \lesssim (s\delta)^{-1}$ of the averages $A_{\sigma,s}$ may be approximated by plates oriented along δ -nearby subspaces. The next lemma records the support of the complementary high-frequency components. For $\sigma \in \text{Gr}(d, n)$ and $\delta > 0$, consider the two-sheeted cone

$$\Gamma_{\sigma, \delta} := \left\{ \xi \in \mathbb{R}^n \setminus \{0\} : \frac{|\Pi_{\sigma} \xi|}{|\xi|} < 2^{-2} \delta \right\}$$

and, abusing notation, denote also the corresponding Fourier restriction by $\Gamma_{\sigma, \delta}$, namely we write $(\Gamma_{\sigma, \delta} f)^{\wedge}(\xi) := \mathbf{1}_{\Gamma_{\sigma, \delta}}(\xi) \hat{f}(\xi)$.

Lemma 2.6. *Let $s > 0, \sigma \in \text{Gr}(d, n)$. Then $A_{\sigma,s}^{<\delta} f = A_{\sigma,s}^{<\delta} \Gamma_{\sigma, \delta} f$.*

Proof. Suppose ξ belongs to the Fourier support of $A_{\sigma,s}^{<\delta} f$. Then $|\xi| \geq 4(\delta s)^{-1}$, and $|\Pi_{\sigma} \xi| < s^{-1} \leq \delta |\xi|/4$. The latter means $\xi \in \Gamma_{\sigma, \delta}$. $\square$

3. Almost orthogonality for directions in the plane

In this section, we present a simple proof of the L^2 -almost-orthogonality principle of [2] for maximal directional operators in the plane. Our argument uses Fourier analysis and overlap estimates instead of geometric considerations and TT^* -type arguments. Although this result is known, we include here a new argument that serves as a prelude to the more technical algebraic almost orthogonality in arbitrary dimension devised in the next section. The statement is as follows.

Theorem D. *There is an absolute constant C such that the following holds. Let $S \subset (0, \infty)$, $U = \{u_1, \dots, u_{N+1}\} \subset \mathbb{S}^1$ be a set of directions ordered counterclockwise. For each $j = 1, \dots, N$ let $V_j \subset \mathbb{S}^1$ be a set of directions contained in the cone bordered by u_j, u_{j+1} and let $V = \bigcup \{V_j : 1 \leq j \leq N\}$. Then,*

$$\|A_{V,S}\|_{L^2(\mathbb{R}^2)} \leq C \|A_{U,S}\|_{L^2(\mathbb{R}^2)} + \max_{1 \leq j \leq N} \|A_{V_j,S}\|_{L^2(\mathbb{R}^2)}.$$

A simple induction argument using the leftmost and middle directions as elements of the dividing set U , together with the control from (2.5), bootstrap Theorem D to recover the following sharp bound from [17, 2].

Corollary D.1. *Let $V \subset \mathbb{S}^1$ be a finite set. Then $\|M_{V,(0,\infty)}\|_{L^2(\mathbb{R}^2)} \lesssim \log(\#V)$.*

Remark 3.1. The sharpness of the estimate in Corollary D.1 follows by a variation of an example employed in [9, Proposition 3.3] for the directional Hilbert transform. In particular one considers the action of $M_{\Sigma_\delta,(0,\infty)}$ on the function $f(x) := |x|^{-1} \mathbf{1}_{\{1 \leq |x| \lesssim N\}}$, where Σ_δ is a $\delta = N^{-1}$ -net on $\mathbb{S}^1$. We omit the details.

Proof of Theorem D. In this proof the constant $C > 0$ is absolute and may vary at each occurrence. For each $1 \leq j \leq N$, let Γ_j be the two-sheeted cone bordered by the supporting lines to $u_j^\perp, u_{j+1}^\perp$, and denote also by Γ_j be the corresponding Fourier restriction. The cones $\{\Gamma_j : 1 \leq j \leq n\}$ are pairwise disjoint, so that

$$\sum_{j=1}^N \|\Gamma_j f\|_2^2 \leq \|f\|_2^2. \quad (3.1)$$

For $v \in V_j$, let $u(v) = \arg \min \{d(v, u_j), d(v, u_{j+1})\}$ and $\delta(v) = d(v, u(v))$. By assumption, the direction $v_j^\perp \in \mathbb{S}^1$ lies between $u_j^\perp, u_{j+1}^\perp$ and thus in Γ_j . The sense of the definitions above is that the whole cone $\Gamma_{v,\delta(v)} := \{\xi : |\xi \cdot v| < \delta_v |\xi|\}$ is contained in Γ_j , namely

$$\Gamma_{v,\delta(v)} \subset \Gamma_j, \quad v \in V_j. \quad (3.2)$$

With this choice, Lemma 2.6 and (3.2) tell us that

$$\left| A_{v,s}^{<\delta(v)} f \right| = \left| A_{v,s}^{<\delta(v)} \Gamma_j f \right| \leq \left| A_{v,s}^{>\delta(v)} \Gamma_j f \right| + |A_{v,s} \Gamma_j f|.$$

Applying Remark 2.5 twice, we obtain

$$\begin{aligned} |A_{v,s} f| &\leq \left| A_{v,s}^{>\delta(v)} f \right| + \left| A_{v,s}^{<\delta(v)} f \right| \leq \left| A_{v,s}^{>\delta(v)} f \right| + \left| A_{v,s}^{>\delta(v)} \Gamma_j f \right| + |A_{v,s} \Gamma_j f| \\ &\leq |A_{v,s} \Gamma_j f| + C \sum_{k=0}^{\infty} 2^{-k} \left(M[\langle |f| \rangle_{u(v), 2^k s}] + M[\langle |\Gamma_j f| \rangle_{u(v), 2^k s}] \right), \end{aligned}$$

where we remember that M denotes the Hardy-Littlewood maximal operator. Taking supremum over $v \in V_j, s \in S, 1 \leq j \leq N$, and subsequently taking L^2 norms,

$$\begin{aligned} \|A_{V,S} f\|_2 &\leq \left\| \max_{1 \leq j \leq N} \{A_{V_j,S} \Gamma_j f\} \right\|_2 \\ &\quad + C \sup_k \|M \circ M_{U, 2^k S} f\|_2 + C \sup_k \left\| \sup_{1 \leq j \leq N} \{M \circ M_{U, 2^k S} [\Gamma_j f]\} \right\|_2. \end{aligned} \quad (3.3)$$

Using the observation in (2.5)

$$\sup_k \|M \circ M_{U,2^k S}\|_{2 \rightarrow 2} \leq C \sup_k \|M_{U,2^k S}\|_{2 \rightarrow 2} \leq C \|A_{U,S}\|_{2 \rightarrow 2}, \quad (3.4)$$

and similarly using the orthogonality in (3.1) and a square function argument

$$\begin{aligned} & \left\| \sup_{1 \leq j \leq N} \{M \circ M_{U,2^k S}[\Gamma_j f]\} \right\|_2 \\ & \leq C \|A_{U,S}\|_{2 \rightarrow 2} \sqrt{\sum_{1 \leq j \leq N} \|\Gamma_j f\|_2^2} \leq C \|A_{U,S}\|_{2 \rightarrow 2} \|f\|_2 \end{aligned} \quad (3.5)$$

and

$$\begin{aligned} & \left\| \max_{1 \leq j \leq N} \{A_{V_j,S} \Gamma_j f\} \right\|_2 \\ & \leq \left(\max_{1 \leq j \leq N} \|A_{V_j,S}\|_{2 \rightarrow 2} \right) \sqrt{\sum_{1 \leq j \leq N} \|\Gamma_j f\|_2^2} \leq \left(\max_{1 \leq j \leq N} \|A_{V_j,S}\|_{2 \rightarrow 2} \right) \|f\|_2. \end{aligned} \quad (3.6)$$

Inserting (3.4), (3.5) and (3.6) into estimate (3.3) completes the proof of the theorem. $\square$

4. Algebraic almost orthogonality in general codimension

This section contains an analogue of Theorem D in higher dimensions, where subsets of $\text{Gr}(d, n)$ are partitioned by algebraic sets. The almost orthogonality result thus obtained may then be employed in combination with polynomial partitioning to obtain a sharpening of the recent result of [12] concerning averages along arbitrary directions in $\mathbb{R}^3$. For simplicity, we restrict ourselves to the case $d = 1$ below and identify $\text{Gr}(1, n)$ with $\mathbb{S}^{n-1}$ in the obvious way. Higher d analogues and the related polynomial partitioning theorems on $\text{Gr}(d, n)$ for $d > 1$ are the object of a forthcoming companion paper.

Definition 1. Let $p \in \mathbb{R}[x_1, \dots, x_n]$ be a degree D polynomial and $Z(p) := \{x \in \mathbb{R}^n : p(x) = 0\}$ be the corresponding zero set. The associated set of *cells* $\mathbf{C}(p)$ is the set of connected components of $\mathbb{S}^{n-1} \setminus Z(p)$. Then $\mathbf{C}(p)$ has $\lesssim D^{n-1}$ elements, see e.g. [3].

Theorem E. Let $p \in \mathbb{R}[x_1, \dots, x_n]$ be a degree D polynomial. For every finite set $\Sigma \subset \mathbb{S}^{n-1}$ and $S \subset (0, \infty)$,

$$\|A_{\Sigma,S}\|_{L^2(\mathbb{R}^n)} \lesssim \sup_{\substack{U \subset \mathbb{S}^{n-1} \cap Z(p) \\ \#U \leq \#\Sigma}} \|A_{U,S}\|_{L^2(\mathbb{R}^n)} + D^{\frac{(n-2)}{2}} \sup_{C \in \mathbf{C}(p)} \|A_{\Sigma \cap C,S}\|_{L^2(\mathbb{R}^n)}.$$

The gain in Theorem E is that the set $Z := \mathbb{S}^{n-1} \cap Z(p)$ is a $(n-2)$ -dimensional algebraic variety of controlled degree and $A_{U,S}$ is better behaved when $U \subset Z$.

4.1. Proof of Theorem E

The strategy is similar to the one we used for Theorem D but with a few twists. We keep using the notation $Z = \mathbb{S}^{n-1} \cap Z(p)$ and the distance $\mathbf{d}$ used here is between elements of $\mathbb{S}^{n-1}$. First of all we fix $\mathbf{C} \in \mathbf{C}(P)$ and $\sigma \in \mathbf{C}$. We set

$$u(\sigma) := \arg \min \{ \mathbf{d}(\sigma, u) : u \in Z \}, \quad \delta(\sigma) := \frac{|\mathbf{d}(\sigma, u(\sigma))|}{4},$$

and further introduce

$$\Gamma_{\mathbf{C}} := \bigcup_{v \in \mathbf{C} \cap \Sigma} \Gamma_{\sigma, \delta(\sigma)}, \quad \widehat{\Gamma_{\mathbf{C}} f}(\xi) := \widehat{f}(\xi) \mathbf{1}_{\Gamma_{\mathbf{C}}}(\xi), \quad \xi \in \mathbb{R}^n.$$

Note that we are again conflating the set $\Gamma_{\mathbf{C}}$ with the corresponding Fourier restriction. We also define the sets

$$U := \bigcup_{\mathbf{C} \in \mathbf{C}(P)} U(\mathbf{C}), \quad U(\mathbf{C}) := \{u(\sigma) : \sigma \in \mathbf{C}\}.$$

Clearly we have that $U \subset Z$ and $\#U \leq \#\Sigma$. An immediate though important geometric observation is made in the following lemma. For $\xi \in \mathbb{R}^n \setminus \{0\}$, recall the notation $H_{\xi}(1) = \{\tau \in \mathbb{S}^{n-1} : \Pi_{\tau} \xi = 0\}$.

Lemma 4.2. *Let $\mathbf{C} \in \mathbf{C}(P)$ and $\xi \in \Gamma_{\mathbf{C}}$. Then $\mathbf{C} \cap H_{\xi}(1) \neq \emptyset$.*

Proof. As $\Gamma_{\mathbf{C}}$ is a union of cones, it suffices to work with $\xi \in \Gamma_{\mathbf{C}} \cap \mathbb{S}^{n-1}$. By definition of $\Gamma_{\mathbf{C}}$ we may find $\sigma \in \mathbf{C}$ such that $|\Pi_{\sigma} \xi| < \delta(\sigma)$. The first claim of Lemma 2.2 tells us that there exists $\tau \in H_{\xi}(1)$ with $\mathbf{d}(\sigma, \tau) < \delta(\sigma)/3$. However the set $\{\tau \in \mathbb{S}^{n-1} : \mathbf{d}(\sigma, \tau) < \delta(\sigma)\}$ is contained in $\mathbf{C}$ by the definition of $\delta(\sigma)$. $\square$

Lemma 4.2 is the main cog in the proof of the following square function estimate.

Lemma 4.3. $\sum_{\mathbf{C} \in \mathbf{C}(P)} \|\Gamma_{\mathbf{C}} f\|_2^2 \lesssim D^{n-2} \|f\|_2^2.$

Proof. It suffices to show that for almost every $\xi \in \mathbb{S}^{n-1}$

$$\#\mathbf{C}(\xi) \lesssim D^{n-2}, \quad \mathbf{C}(\xi) := \{\mathbf{C} \in \mathbf{C}(P) : \xi \in \Gamma_{\mathbf{C}}\}.$$

Lemma 4.2 tells us that if $\mathbf{C} \in \mathbf{C}(\xi)$ then $\mathbf{C} \cap H_{\xi}(1)$ is a connected component of $H_{\xi}(1) \setminus Z$. As $\dim H_{\xi}(1) = n-2$ this set has $\lesssim D^{n-2}$ connected components, see e.g. [26, Chapter 4] or [3]. $\square$

We begin the main argument. Fixing $C \in \mathbf{C}(p)$, $\sigma \in \mathbf{C}$, we apply Lemma 2.6 and the definition of Γ_C to obtain

$$\left| A_{\sigma,s}^{<\delta(\sigma)} f \right| = \left| A_{\sigma,s}^{<\delta(\sigma)} \Gamma_C f \right| \leq CM[A_{\sigma,s} \Gamma_C f],$$

which together with Lemma 2.4 yields

$$|A_{\sigma,s} f| \leq \left| A_{\sigma,s}^{>\delta(\sigma)} f \right| + \left| A_{\sigma,s}^{<\delta(\sigma)} f \right| \leq CM \circ M_{u(\sigma),s} f + CM[A_{v,s} \Gamma_C f].$$

We first take supremum over $s \in S$, $\sigma \in \mathbf{C}$ and obtain

$$A_{\Sigma \cap C, S} f \leq CM \circ M_{U(C), S} f + CM \circ A_{\Sigma \cap C}(\Gamma_C f).$$

Subsequently taking supremum over $C \in \mathbf{C}(P)$ leads to

$$A_{\Sigma, S} f \lesssim M \circ M_U f + \left(\sum_{C \in \mathbf{C}(P)} |M \circ A_{\Sigma \cap C}(\Gamma_C f)|^2 \right)^{\frac{1}{2}}. \quad (4.1)$$

The estimate of Theorem E then follows easily from (4.1) and the square function estimate of Lemma 4.3.

5. Sharp or nearly sharp bounds for maximal subspace averages

We now focus on the single scale maximal operator $M_{\Sigma, \{1\}} =: M_{\Sigma}$ when Σ is a generic finite subset of $\text{Gr}(d, n)$. The majority of this section is in fact dedicated to the proof of Theorem A. However, we first detail the announced sharpening of the case $n = 3$, $d = 1$. In this case, Theorem A tells us that

$$\sup_{\substack{\Sigma \subset \mathbb{S}^2 \\ \#\Sigma \leq N}} \|M_{\Sigma}\|_{L^2(\mathbb{R}^3)} \lesssim N^{\frac{1}{4}} \log N$$

which is sharp up to the logarithmic factor and recovers the bound from [10]. The following more precise estimate was previously proved in [12]: for any positive integer k we have

$$\sup_{\substack{\Sigma \subset \mathbb{S}^2 \\ \#\Sigma \leq N}} \|M_{\Sigma}\|_{L^2(\mathbb{R}^3)} \lesssim_k N^{\frac{1}{4}} (\log N)^{\frac{1}{2}} \log^{[k]} N,$$

where $\log^{[1]} N := \log(2 + N)$, $\log^{[k]} N := \log(2 + \log^{[k-1]} N)$. For this special case of dimension and codimension we improve the result of the theorems above by exploiting the algebraic almost orthogonality principle of §4. This is the content of the following theorem.

Theorem F. For all $N > 0$ and every positive integer k there holds

$$\sup_{\substack{\Sigma \subset \mathbb{S}^2 \\ \#\Sigma \leq N}} \|M_\Sigma\|_{L^2(\mathbb{R}^3)} \lesssim_k N^{\frac{1}{4}} \log^{[k]} N,$$

with implicit constant depending only on k . This bound is sharp in terms of N up to the iterated logarithmic factor.

We will prove the estimates of both Theorems A and F for the corresponding norm-equivalent Fourier operator $A_{\Sigma, \{1\}} =: A_\Sigma$.

5.1. Clusters on $\text{Gr}(d, n)$

We begin with a definition related to $H_\xi(d)$ from (2.2). We say that Σ is a δ -cluster with top $\xi \in \mathbb{S}^{n-1}$ if Σ is a finite set contained in the δ -neighborhood of $H_\xi(d)$. Somewhat dual to δ -clusters are the cone cutoffs

$$\Gamma_{\sigma, \delta} = \{\eta \in \mathbb{R}^n \setminus \{0\} : |\Pi_\sigma \eta| < 2^{-4} \delta |\eta|\}, \quad \Gamma_{\sigma, \delta} f(x) = \int_{\Gamma_{\Sigma, \delta}} \widehat{f}(\eta) e^{ix \cdot \eta} d\eta.$$

We summarize two key steps of our proof in the estimate of the next lemma.

Lemma 5.2. Suppose $f \in \mathcal{S}(\mathbb{R}^d)$ with

$$\text{supp } \widehat{f} \subset \text{Ann}(\delta) := \{\eta \in \mathbb{R}^n : 2^{-4} < \delta |\eta| < 2^{-2}\}.$$

Then

1. For all $\sigma \in \text{Gr}(d, n)$ and $s \geq 1$, we have $A_{\sigma, s} f = A_{\sigma, s} \Gamma_{\sigma, \delta} f$.
2. If Σ is a δ -cluster, we have the estimate

$$\|A_{\Sigma, \{1\}} f\|_2 \lesssim \left(\sup_{\substack{\Sigma' \subset \text{Gr}(d, n-1) \\ \#\Sigma' \leq \#\Sigma}} \|A_{\Sigma', \{1\}}\|_{L^2(\mathbb{R}^{n-1})} \right) \|f\|_2.$$

Proof. First of all we dispense with the support claim. As ϕ_d in the definition of $A_{\sigma, s}$ is supported in $2^{-8} B_d$ we have that $\widehat{A_{\sigma, s} f}(\eta) = 0$ unless $|\Pi_\sigma \eta| < 2^{-8} s^{-1}$ and $\delta |\eta| > 2^{-4}$, whence the claim.

We then move to the proof of the estimate. Suppose that Σ is a δ -cluster with top $\xi \in \mathbb{S}^{n-1}$. For $\sigma \in \Sigma$ let $\tau(\sigma) \in H_\xi(d)$ be such that $d(\sigma, \tau) < \delta$. As $\widehat{f}(\eta) = 0$ for $\delta |\eta| \geq 2^{-2}$, we may insert the $\Phi(2^2 \delta \cdot)$ cutoff for free and $A_{\sigma, 1} f = A_{\sigma, 1}^{>\delta} f$. Then using (2.7) from Lemma 2.4 and Remark 2.5 returns

$$A_{\sigma,1}f \lesssim \sum_{k=0}^{\infty} 2^{-kn} M(\langle f \rangle_{\tau(\sigma),2^k})$$

which coupled with the norm equivalences of (2.5) tells us that

$$\|A_{\Sigma,\{1\}}f\|_2 \lesssim \left(\sup_{\substack{\Sigma' \subset H_{\xi}(d) \\ \#\Sigma' \leq \#\Sigma}} \|A_{\Sigma',\{1\}}\|_{L^2(\mathbb{R}^n)} \right) \|f\|_2.$$

The conclusion of the lemma then follows from recalling that $H_{\xi}(d) \cong \text{Gr}(d, n-1)$ and applying Fubini's theorem in the ξ direction. $\square$

5.3. Counterexamples yielding sharpness of Theorem A

In this subsection we discuss the sharpness of the estimate in Theorem A. In fact we will prove the more general proposition below.

Proposition 5.4. *Let $n \geq 2$ and $1 < d \leq n-1$. There holds*

$$\sup_{\substack{\Sigma \subset \text{Gr}(d,n) \\ \#\Sigma \leq N}} \|M_{\Sigma}\|_{L^p(\mathbb{R}^n)} \gtrsim \begin{cases} N^{\frac{n-d+1-p}{p(n-d)}}, & 1 < p < n-d+1, \\ (\log N)^{\frac{1}{p}}, & p \geq n-d+1. \end{cases}$$

Proof of Proposition 5.4 for $d=1$. This case is classical and well understood but we include it here as it is instructive for case of $2 \leq d < n$ that follows. Indeed we just need to consider f to be the indicator function of the unit ball in $\mathbb{R}^n$ and Σ be a $\varepsilon\delta$ -net in the unit sphere with ε a numerical constant to be chosen momentarily. Clearly $\#\Sigma \asymp_{\varepsilon} \delta^{-(n-1)}$ and note that for $x \in \mathbb{R}^n$ and $v(x) := x/|x|$ we have that

$$A_{v(x),\delta^{-1}}^{>\delta} f(x) \gtrsim \delta, \quad x \in B_n(\delta^{-1}) \setminus B_n(0,1).$$

On the other hand for every $x \in \mathbb{R}^n$ there exists $v \in \Sigma$ with $\text{dist}(v, v(x)) \leq \varepsilon\delta$ so we get by Lemma (2.4)

$$A_{v,\delta^{-1}}^{>\delta} f(x) \geq c\delta - C\varepsilon \sum_{k=0}^{\infty} 2^{-kn} \int_{x+2^k\delta^{-1}T_{\delta}} |f| \gtrsim \delta$$

if ε is chosen to be sufficiently small. This readily implies

$$\|M_{\Sigma,\{\delta^{-1}\}}\|_{L^p(\mathbb{R}^n)} \gtrsim \|A_{\Sigma,\{\delta^{-1}\}}\|_{L^p(\mathbb{R}^n)} \gtrsim \delta^{-\frac{n-p}{p}}, \quad 1 < p < n.$$

Since $\#\Sigma = \delta^{-(n-1)} = N$ this proves the proposition when $d=1$ and $p < n$.

For the case $d = 1$, $p \geq n$ the lower bound of the order $(\log N)^{1/p}$ follows by the Besicovitch construction of a Kakeya set as in [15]. Briefly, one constructs a set $K_\delta = \bigcup T_\delta \subset \mathbb{R}^2$, the δ -neighborhood of a Kakeya set, which is a union of $\delta \times 1$ tubes pointing along the directions Σ of a δ -net on $\mathbb{S}^1$, and so that $|K_\delta| \lesssim (\log N)^{-1}$. Then it is simple to check that

$$M_{\Sigma, \{3\}} \mathbf{1}_{K_\delta} \gtrsim 1 \quad \text{on the set} \quad K_\delta^* := \bigcup T_\delta^* \setminus K_\delta,$$

where T_δ^* is the tube with the same center and direction as T_δ and 3 times the length. It follows that $|K_\delta^*| \approx 1$ and so

$$\|M_{\Sigma, 1}\|_{L^p(\mathbb{R}^2)} \gtrsim (\log N)^{\frac{1}{p}}.$$

For general $n \geq 2$ we just test the maximal operator on the tensor product of the Kakeya-type set above with a unit cube in $\mathbb{R}^{n-2}$, namely $K_\delta \times [-1/2, 1/2]^{n-2}$ and the lower bound $\|M_{\Sigma, 1}\|_{L^p(\mathbb{R}^n)} \gtrsim (\log N)^{\frac{1}{p}}$ follows. $\square$

Proof of Proposition 5.4 for $1 < d < n$. In general we consider (d, n) plates of scale $M := N^{1/(n-d)}$ and thickness 0 as follows. If $\sigma \in \text{Gr}(d, n)$, then

$$T_0^M(\sigma) := \{x \in \mathbb{R}^n : |x| < M, \Pi_{\sigma^\perp} x = 0\}$$

is the scale M plate oriented along σ . Let $\omega \in \text{Gr}(d-1, n)$ that will remain fixed throughout the argument below. Then we define

$$E = E(\omega) := \{\eta \in \text{Gr}(1, n) : \Pi_\omega v = 0 \ \forall v \in \eta\} \sim \text{Gr}(1, n-d+1) \sim \mathbb{S}^{n-d}.$$

Notice that $\text{span}\{\omega, \eta\} \in \text{Gr}(d, n)$ for all $\eta \in E$. We may then write every $x \in \mathbb{R}^n$ as $x = \Pi_\omega x + \rho\eta$ for $\eta \in E \sim \mathbb{S}^{n-d}$ and $\rho \geq 0$. We use the notation $x = (x_\omega, \rho\eta)$ accordingly where $x_\omega = \Pi_\omega x$. We also set

$$C_M = C_M(\omega) := \{x = (x_\omega, \rho\eta) \in \mathbb{R}^n : |x_\omega| \leq M, |\rho| \leq 1\}.$$

Lemma 5.5. *Let $\omega \in \text{Gr}(d-1, n)$ and $x = (x_\omega, \rho\eta)$ with $|x_\omega| \leq \frac{M}{2}$, $2^{-8}M \leq \rho \leq 2^{-7}M$. Suppose $v \in E$ with $|v - \eta| < \frac{2^{-8}}{M}$. Let $\sigma = \text{span}(\omega, \eta) \in \text{Gr}(d, n)$. Then*

$$|[x + T_0^M(\sigma)] \cap C_M| \geq 2^{-10} M^{d-1}.$$

Proof. By a rotation we may assume $\omega = \{e_1, \dots, e_{d-1}\}$. Choose β so that $\{\eta, \beta\}$ is an orthonormal basis of $\text{span}\{\eta, v\}$. Then from the assumption of $|v - \eta| < 2^{-8}M^{-1}$ we obtain that

$$v = (\cos \theta)\eta + (\sin \theta)\beta, \quad |\theta| \leq \frac{2^{-7}}{M}$$

so that $x = (x_\omega, \rho\eta)$ and write the generic point on the plate $x + T_0^M(\sigma)$ as

$$y = (x_\omega + t_1 e_1 + \dots + t_{d-1} e_{d-1}, (\rho + t_d \cos \theta)\eta + t_d (\sin \theta)\beta), \quad t_1, \dots, t_d \in (-M, M).$$

Note that

$$\sup_j |t_j| < \frac{M}{2(d-1)} \implies |y_\omega| \leq |x_\omega| + (d-1) \sup_j |t_j| < M.$$

Also,

$$t_d \in I := \left[-\frac{\rho}{\cos \theta} + \frac{1}{8}, -\frac{\rho}{\cos \theta} + \frac{1}{4} \right] \implies 2^{-10}M \leq |t_d| \leq 2^{-6}M$$

so that

$$|(\rho + t_d \cos \theta)\eta + t_d (\sin \theta)\beta| \leq \frac{1}{4} + 2^{-6}M |\sin \theta| \leq \frac{1}{2}.$$

It follows that $y \in C_M$ in the range $|t_1|, \dots, |t_{d-1}| < \frac{M}{2(d-1)}, t_d \in I$ which is a set of measure $\geq 2^{-10}M^{d-1}$. $\square$

We continue with the proof of Proposition 5.4. Let E_M be a $2^{-18}M$ -net in $E \sim \mathbb{S}^{n-d}$ and consider the set

$$\Sigma_M := \{\sigma = \text{span}(\omega, v) : v \in E_M\} \subset \text{Gr}(d, n)$$

which has $\sim M^{n-d} = N$ elements. The maximal function

$$M_{\Sigma_M, \{M\}} f(x) = \sup_{\sigma \in \Sigma_M} \frac{1}{M^d} \int_{x+T_0^M(\sigma)} |f|$$

then satisfies

$$\begin{aligned} M_{\Sigma_M, \{M\}} \mathbf{1}_{C_M}(x) &\gtrsim \frac{M^{d-1}}{M^d} \sim M^{-1}, \\ x \in U_M &:= \{(x_\omega, \rho\eta) : |x_\omega| \leq 2^{-1}M, 2^{-8}M \leq \rho \leq 2^{-7}M\}, \end{aligned}$$

whence

$$\|M_{\Sigma_M, \{M\}}\|_{L^p(\mathbb{R}^n)} \geq \frac{\|M_{\Sigma_M, \{M\}} \mathbf{1}_{C_M}\|_p}{|C_M|^{\frac{1}{p}}} \gtrsim M^{-1+\frac{1-d}{p}} |U_M|^{\frac{1}{p}} \gtrsim M^{\frac{n-d+1}{p}-1} = N^{\frac{n-d+1-p}{p(n-d)}}.$$

This proves the desired bound for $1 < d < n$ and $p < n - d + 1$.

For the case $p \geq n - d + 1$ we repeat the proof above but replacing the ball with a Kakeya set. More precisely, fixing $\omega \in \text{Gr}(d-1, n)$ we let $K_\delta \subset \mathbb{R}^2 \subseteq \omega^\perp$ be a δ -neighborhood of a Kakeya set as before, namely K_δ is a union of $\delta \times 1$ tubes pointing along a δ -net $\Sigma_\delta \subset \mathbb{S}^1$. Note that this is always possible as $\omega^\perp$ always contains a copy of $\mathbb{R}^2$. Remembering that $x_\omega = \Pi_\omega x$ we modify the definition of C_N as follows

$$C_\delta = C_\delta(\omega) := \{x \in \mathbb{R}^n : |x_\omega| \leq 1, x - x_\omega \in K_\delta\}.$$

It is then easy to check that

$$\|M_{\Sigma_M, \{M\}}\|_{L^p(\mathbb{R}^n)} \geq \frac{\|M_{\Sigma_M, \{M\}} \mathbf{1}_{C_\delta}\|_p}{|C_\delta|^{\frac{1}{p}}} \gtrsim (\log \delta^{-1})^{\frac{1}{p}} \approx (\log \#\Sigma_\delta)^{\frac{1}{p}}$$

as desired. $\square$

5.6. Proof of the upper bound in Theorem A

We seek an inductive estimate for

$$K_{d,n} := \sup_{N>0} \frac{1}{N^{\frac{n-d-1}{n-d}} (\log N)^{\alpha(d,n)}} \sup_{\substack{\Sigma \subset \text{Gr}(d,n) \\ \#\Sigma \leq N}} \|A_\Sigma\|_{L^2(\mathbb{R}^n)}^2,$$

where $\alpha(d, n)$ will be determined along the induction argument. The first step in the reduction is a classical use of the Chang-Wilson-Wolff inequality. For similar applications in the setting of directional singular integrals see for example [9, 11].

Lemma 5.7. *There holds $K_{d,n} \lesssim_n Q_{d,n}$, where*

$$Q_{d,n} := \sup_{N>0} \frac{1}{N^{\frac{n-d-1}{n-d}} (\log N)^{\alpha(d,n)-1}} \sup_{0<\delta<1} \sup_{\substack{\Sigma \subset \text{Gr}(d,n) \\ \#\Sigma \leq N}} \sup_{\substack{\|f\|_2=1 \\ \text{supp } \hat{f} \subset \text{Ann}(\delta)}} \|A_\Sigma f\|_{L^2(\mathbb{R}^n)}^2,$$

with $\text{Ann}(\delta)$ as defined in Lemma 5.2.

The induction parameter is n , while d is kept fixed along the induction. The seed for the induction is the base case $n = d + 1$.

Lemma 5.8. *If $\alpha(d, d+1) = 1$, then $Q_{d,d+1} \lesssim_d 1$.*

The proof of Lemma 5.8 is given at the end of this section. Notice that the lemma, together with Lemma 5.7, implies $K_{d,d+1} \lesssim 1$ with the choice $\alpha(d, d+1) = 1$. Fix now $N > 0$, $0 < \delta < 1$, $\Sigma \subset \text{Gr}(d, n)$ with $\#\Sigma \leq N$, $f \in \mathcal{S}(\mathbb{R}^n)$ with $\|f\|_2 = 1$ and $\text{supp } \hat{f} \subset \text{Ann}(\delta)$. Using a greedy selection algorithm, we may achieve

$$\Sigma = \Sigma_0 \cup \bigcup_{j=1}^{\Theta} \Sigma_j,$$

where $0 \leq \Theta \leq N^d$, each Σ_j is a δ -cluster and Σ_0 has the property that

$$\sum_{\sigma \in \Sigma_0} \mathbf{1}_{\Gamma_{\sigma, \delta}}(\eta) \leq N^{\frac{n-d-1}{n-d}} \quad \forall \eta \in \mathbb{R}^n \setminus \{0\}. \quad (5.1)$$

Say $\xi \in \mathbb{S}^{n-1}$ is bad for Σ' if the set $\{\sigma \in \Sigma' : \xi \in \Gamma_{\sigma, \delta}\}$ has cardinality $> N^{\frac{n-d-1}{n-d}}$. The first step in the greedy selection algorithm is to initialize and set $\Sigma' = \Sigma$. If some bad ξ exists for Σ' , set $\Sigma_1 := \{\sigma \in \Sigma' : \xi \in \Gamma_{\sigma, \delta}\}$. Notice that by Lemma 2.2 the set Σ_1 is a δ -cluster. Now set $\Sigma' := \Sigma' \setminus \Sigma_1$ and repeat. The algorithm terminates when no bad ξ exists for Σ' , in which case set $\Sigma_0 = \Sigma'$. Notice that (5.1) then holds by construction. Cardinality considerations tell us that the algorithm terminates after $\Theta \leq N^d$ steps. Using the first claim of Lemma 5.2, we get at once

$$\|A_{\Sigma_0} f\|_2^2 \leq \sum_{\sigma \in \Sigma_0} \|A_{\Sigma_0} \Gamma_{\sigma, \delta} f\|_2^2 \lesssim \sum_{\sigma \in \Sigma_0} \mathbf{1}_{\Gamma_{\sigma, \delta}} \leq N^{\frac{n-d-1}{n-d}}. \quad (5.2)$$

Suppose that the cluster Σ_j has $\sim 2^k N^{\frac{n-d-1}{n-d}}$ elements for some $1 \leq 2^k \leq N^{\frac{1}{n-d}}$. Then the second claim of Lemma 5.2 tells us that

$$\|A_{\Sigma_j} f\|_2^2 \lesssim K_{d, n-1} 2^k N^{\frac{n-d-2}{n-d-1}} N^{\frac{n-d-2}{n-d}} (\log N)^{\alpha(d, n-1)}$$

Notice that $\#\{j : \Sigma_j \sim 2^k N^{\frac{n-d-1}{n-d}}\} \leq 2^{-k} N^{\frac{1}{n-d}}$, so that

$$\begin{aligned} \sum_{j=1}^{\Theta} \|A_{\Sigma_j} f\|_2^2 &\lesssim K_{d, n-1} N^{\frac{n-d-1}{n-d}} (\log N)^{\alpha(d, n-1)} \sum_{k=1}^{\infty} 2^{k \frac{n-d-2}{n-d-1} - k} \\ &\lesssim K_{d, n-1} N^{\frac{n-d-1}{n-d}} (\log N)^{\alpha(n-1)}. \end{aligned} \quad (5.3)$$

Combining (5.2) and (5.3) we see that $Q_{d, n} \lesssim K_{d, n-1}$, so that $K_{d, n} \lesssim K_{d, n-1}$ provided that $\alpha(d, n) = \alpha(d, n-1) + 1$. Induction completes the proof with $\alpha(d, n) = n - d$.

5.9. Proof of Lemma 5.8

We have $n = d + 1$ throughout the proof of the lemma. We use below the slight enlargement of $\text{Ann}(\delta)$ from the statement of Lemma 5.2

$$\text{Ann}^+(\delta) := \{\eta \in \mathbb{R}^n : 2^{-5} < \delta|\eta| < 2^{-1}\} \supset \text{Ann}(\delta).$$

Before the proof proper we carefully reshuffle the conclusion of Lemma 2.4. We say that $m \in \mathcal{S}(\mathbb{R}^n)$ is δ -adapted to $\tau = \text{span}\{e_1, \dots, e_d\}$ if

- a. $\text{supp } m \subset O_{\tau, \delta} := \{\xi \in \text{Ann}^+(\delta) : |\Pi_\tau \xi| \leq 2\delta|\xi|\},$
- b. $\|\partial^\alpha m\|_\infty \leq \delta^{\alpha_{d+1}}$ for all multi-indices $\alpha = (\alpha_1, \dots, \alpha_{d+1})$ of order $|\alpha| \leq 100d$.

Now if $\tau \in \text{Gr}(d, d+1)$ is generic, we say that m is δ -adapted to τ if $m \circ \tau$ is δ -adapted to $\text{span}\{e_1, \dots, e_d\}$, where τ also stands for the rotation mapping $\text{span}\{e_1, \dots, e_d\}$ to τ and e_{d+1} to $\tau^\perp$. A typical example of function δ -adapted to τ is

$$m(\xi) := \mu(\Pi_\tau \xi) \Psi(\delta \xi), \quad \xi \in \mathbb{R}^{d+1},$$

where $\Psi \in \mathcal{S}(\mathbb{R}^{d+1})$ is supported in $\text{Ann}(1)$ and $\mu \in \mathcal{S}(\mathbb{R}^d)$ is supported on a small ball near the origin and suitably normalized. We denote by $\mathfrak{m}(\tau, \delta)$ the class of multipliers which are δ -adapted to τ and define

$$A_\tau^\delta f(x) := \sup_{m \in \mathfrak{m}(\tau, \delta)} \left| \int_{\mathbb{R}^{d+1}} \widehat{f}(\xi) m(\xi) e^{ix \cdot \xi} d\xi \right|, \quad x \in \mathbb{R}^{d+1}.$$

A repetition of the proof of (2.7) of Lemma 2.4 tells us that $A_\tau^\delta f$ is pointwise bounded by the right hand side of (2.7) for $s = 1$. This together with frequency support considerations yield

$$\|A_\tau^\delta f\|_2 \lesssim \|O_{\tau, \delta} f\|_2 \quad (5.4)$$

uniformly in $\tau \in \text{Gr}(d, d+1)$ and $\delta > 0$, where we have denoted by $O_{\tau, \delta}$ the frequency cutoff to the corresponding conical sector. Now, if Σ_δ is $2^{-10}\delta$ -net in $\text{Gr}(d, d+1)$, we have

$$\sum_{\tau \in \Sigma_\delta} \|O_{\tau, \delta} f\|_2^2 \lesssim \|f\|_2^2 \quad (5.5)$$

as the projection on $\mathbb{S}^d$ of $O_{\tau, \delta}$ lies in a $\sim \delta$ -neighborhood of τ . A square function argument combining (5.4) with (5.5) then yields

$$\left\| \sup_{\tau \in \Sigma_\delta} A_\tau^\delta f \right\|_2 \lesssim \|f\|_2. \quad (5.6)$$

We are ready to complete the proof proper of this lemma. Fix $\delta > 0$ and let $f \in L^2(\mathbb{R}^{d+1})$ with $\text{supp } \widehat{f} \subset \text{Ann}(\delta)$. Read from the proof of Lemma 2.4, cf. (2.9) in particular, that

$$\sup_{\substack{\sigma \in \text{Gr}(d, d+1) \\ d(\sigma, \tau) < \delta}} |A_{\sigma, 1} f| \lesssim A_\tau^\delta f \quad (5.7)$$

by means of a suitable insertion of a Littlewood-Paley cutoff equal to 1 on $\text{Ann}(\delta)$ and supported on $\text{Ann}^+(\delta)$. As Σ_δ is a $2^{-10}\delta$ -net, an application of (5.7) followed by (5.6) returns

$$\left\| \sup_{\sigma \in \text{Gr}(d, d+1)} |A_{\sigma, 1} f| \right\|_{L^2(\mathbb{R}^{d+1})} \lesssim \left\| \sup_{\tau \in \Sigma_\delta} A_\tau^\delta f \right\|_2 \lesssim \|f\|_2$$

which is what is claimed in the lemma.

5.10. Polynomial partition for the proof of Theorem F

We plan to apply the algebraic almost orthogonality principle of §4 in order to prove Theorem F. The first order of business is to feed the almost orthogonality result of Theorem E with a suitable polynomial partition of the set $\Sigma \subseteq \mathbb{S}^2$ with $\#\Sigma \leq N^2$, tailored to the problem in hand.

Proposition 5.11. *Let $\Sigma \subset \mathbb{S}^2$ be a finite set with $\#\Sigma \leq N$ and let $\delta > 0$. For each integer $D \geq 2^3$ there exists a partition*

$$\Sigma = \Sigma_\circ \cup \Sigma_\times$$

satisfying the following properties:

1. *There exist $O(1)$ transverse complete intersections $W_j \subset \mathbb{S}^2$ of dimension 1 and degree $O(D)$ such that*

$$\sup_{\sigma \in \Sigma_\times} \inf_j \text{dist}(\sigma, W_j) < \delta.$$

2. *There exist $O(D^2)$ disjoint connected open subsets $C \in \vec{C}$ of $\mathbb{S}^2$ with the property that*

$$\Sigma_\circ = \bigcup_{C \in \vec{C}} \Sigma_C, \quad \Sigma_C := \Sigma \cap C, \quad \#\Sigma_C \leq \frac{N}{D^2}.$$

The proposition above is a consequence of the more general polynomial partitioning result of [12, Proposition 2.10]. We also refer the reader to [12, §2.9] for the definition of a *transverse complete intersection* and further background on polynomial partitioning results.

The partition of Proposition 5.11 is not directly applicable as an input for Theorem E as the set $\Sigma_\times$ also contains directions *close* to the algebraic variety $Z := \cup_j W_j$ instead of just directions *on* the Z . This is easily remedied by a soft approximation argument. Indeed as the conclusion of Theorem F is an $L^2(\mathbb{R}^3)$ -operator norm bound, we can work with functions $f \in L^2(\mathbb{R}^3)$ such that $\text{supp}(\hat{f}) \subseteq B_3(R)$ for some $R > 0$, as long as we prove bounds independent of R . Now if we choose $\delta \ll R^{-1}$ in Proposition 5.11, we will have that $\|A_\sigma f\|_{L^2(\mathbb{R}^3)} \lesssim \|M_{\tau(\sigma)} f\|_{L^2(\mathbb{R}^3)}$ for all $\sigma \in \Sigma_\times$, where $\tau(\sigma)$ denotes a direction on Z such that $\text{dist}(\sigma, \tau(\sigma)) < \delta$, whose existence is guaranteed from point 1. of Proposition 5.11. This remark allows us to assume that $\Sigma_\times \subset Z = \cup_j W_j$.

With the remark above taken as understood, the almost orthogonality principle of Theorem F reduces the proof of Theorem F to a recursive *cellular* term, and a *wall* term which is the $L^2(\mathbb{R}^3)$ -norm of $A_{\Sigma_\times}$. The latter operator is a single scale, maximal directional average along directions on a one-dimensional algebraic subvariety of $\mathbb{S}^2$. It follows from [12, Theorem D] that

$$\sup_{\substack{\Sigma' \subset \cup_j W_j \\ \#\Sigma' \leq N}} \|A_{\Sigma'}\|_{L^2(\mathbb{R}^3)} \lesssim D^{\frac{1}{2}} (\log N)^{\frac{3}{2}}. \quad (5.8)$$

5.12. The proof of Theorem F

We are seeking an estimate for

$$K_N := \sup_{\substack{\Sigma \subset \mathbb{S}^2 \\ \#\Sigma \leq N}} \|A_\Sigma\|_{L^2(\mathbb{R}^n)}^2.$$

Combining Theorem E with the polynomial partition of Proposition 5.11, the subsequent remarks, and (5.8) we can estimate for any $D \gtrsim 1$

$$K_N \leq K_1 D (\log N)^3 + K_2 D K_{\frac{N}{D^2}}$$

for numerical constants $K_1, K_2 > 0$ and D a sufficiently large degree, to be chosen momentarily. Indeed choosing $D := \sqrt{N}/(\log N)^3$ yields

$$\frac{K_N}{\sqrt{N}} \leq K_1 + K_2 \frac{K_{(\log N)^6}}{\sqrt{(\log N)^6}}$$

which readily implies the estimate in statement of Theorem F by recursion.

6.akeya and Nikodym maximal operators

In this section we digress a bit in order to discuss two maximal operators that appear naturally in the context of this paper, namely the *akeya* and *Nikodym* maximal operators. These operators have been studied extensively in the case $d = 1$ in relation to the maximal *akeya* conjecture and the maximal *Nikodym* conjecture, which are equivalent when $d = 1$; see [30]. For $d > 1$, the corresponding *akeya* maximal function on $\text{Gr}(d, n)$ has been studied in relation to the existence and dimension of (d, n) -Besicovitch sets. We send the interested reader to [20] for general background on the topic, and will focus below on just a few notions central to our discussion.

We briefly recall some elementary properties of the Haar measure on $\text{Gr}(d, n)$. Let $1 \leq d < n$ be fixed parameters and $d\sigma$ be the quotient Haar measure on $\text{Gr}(d, n)$ seen as a quotient group as in (2.1), that is

$$\int_{\text{Gr}(d,n)} f(\sigma) d\sigma = \int_{O(n)} f(\omega) d\omega$$

where $O(n)$ is equipped with its left-invariant Haar probability measure. When $A \subset \text{Gr}(d, n)$, the notation $|A|$ stands for the $d\sigma$ -measure of A . As the measure $d\sigma$ is the unique probability measure on $\text{Gr}(d, n)$ which is left invariant under the action of $O(n)$, it coincides with the normalized Riemannian volume form on $\text{Gr}(d, n)$. This implies that, if

$$\mathbf{B}_\delta(\sigma) := \{\tau \in \text{Gr}(d, n) : d(\sigma, \tau) < \delta\},$$

the measure $|\mathbf{B}_\delta(\sigma)|$ is independent of σ and $|\mathbf{B}_\delta(\sigma)| \sim \delta^{d(n-d)}$.

6.1. The *Kekeya* maximal operator

Recall from (2.3) the notation $T_\delta(\sigma)$ for the δ -plate oriented along $\sigma \in \text{Gr}(d, n)$. Consider the dual maximal operator acting on $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and its tailed version

$$\mathcal{K}_\delta f(\sigma) = \sup_{x \in \mathbb{R}^n} \int_{x+T_\delta(\sigma)} |f|, \quad \widetilde{\mathcal{K}}_\delta f(\sigma) = \sup_{x \in \mathbb{R}^n} \sum_{k=0}^{\infty} 2^{-2kn} \int_{x+2^k T_\delta(\sigma)} |f|, \quad \sigma \in \text{Gr}(d, n). \quad (6.1)$$

Definition 2. Let $1 \leq d < n$. A Borel set $A \subset \mathbb{R}^n$ is said to be a (d, n) -set if for every $\sigma \in \text{Gr}(d, n)$ there is $y \in \mathbb{R}^n$ such that $B_n(y, 1) \cap (\sigma + y) \subset A$. If $|A| = 0$ then A is called a (d, n) -Besicovitch set.

It is well known that for $d = 1$ there exist $(1, n)$ -Besicovitch or *Kekeya* sets and the *Kekeya conjecture* states that the Hausdorff dimension of *Kekeya* sets should be at least n . The *maximal Kekeya* conjecture is the statement that for all $\delta, \varepsilon > 0$ we should have the estimate

$$\|\mathcal{K}_\delta\|_{L^n(\text{Gr}(1,n))} \lesssim_\varepsilon \delta^{-\varepsilon} \|f\|_{L^n(\mathbb{R}^n)}. \quad (6.2)$$

When $1 < d < n$, it is conjectured that no (d, n) -Besicovitch sets exist. For $d > n/2$ this conjecture was verified by Falconer, [13]. The range of non-existence of (d, n) -Besicovitch sets was extended by Bourgain in [5] and further by Oberlin in [23]. The reader is also referred to [23] for further results and lower bounds on the Hausdorff dimensions of (d, n) -Besicovitch for general d . The connection with the *Kekeya* maximal function is revealed by the following well known implication; see [23, p. 3].

Proposition 6.2. Let $1 \leq d < n$ and suppose that there exists $1 \leq p < \infty$ and $\varepsilon > 0$ such that for every $\delta > 0$ we have the bound

$$\|\mathcal{K}_\delta f\|_{L^p(\mathrm{Gr}(d,n))} \lesssim_\varepsilon \delta^{-\frac{\varepsilon}{p}} \|f\|_{L^p(B_n)}.$$

Then every (d, n) -set has Hausdorff dimension at least $n - \varepsilon$.

In the context of the current paper the case $1 < d < n$ is the most relevant. The estimate of the following proposition is a natural consequence of the methods of this paper and we include as an illustration of how these methods can be applied in the context of the (d, n) -Kakeya maximal function. Our estimate below recovers the well known result of Falconer [13]: there are no (d, n) -Besicovitch sets when $d > n/2$. By Proposition 6.2, it also yields that a (d, n) -Besicovitch set necessarily has full Hausdorff dimension when $d = n/2$.

Proposition 6.3. $\|\mathcal{K}_\delta : L^2(\mathbb{R}^n) \rightarrow L^2(\mathrm{Gr}(d, n))\| \lesssim \begin{cases} 1 & 2d > n \\ \log \delta & 2d = n \\ \delta^{d-\frac{n}{2}} & 2d < n. \end{cases}$

Proof. By an isotropic rescaling of the input function f , the norm equivalence

$$\|\mathcal{K}_\delta : L^p(\mathbb{R}^n) \rightarrow L^q(\mathrm{Gr}(d, n))\| \sim_{p,q} \|\widetilde{\mathcal{K}}_\delta : L^p(\mathbb{R}^n) \rightarrow L^q(\mathrm{Gr}(d, n))\| \quad \forall 1 \leq p, q \leq \infty$$

holds. By standard arguments and Lemma 2.4, if $\tau \in \mathrm{Gr}(d, n)$ and $\sigma \in B_{2^{-10}\delta}(\tau)$ we have the pointwise estimates

$$\mathcal{K}_\delta f(\sigma) \lesssim \sup_{x \in \mathbb{R}^n} |A_{\sigma,1}^{\delta} f(x)| \lesssim \sup_{x \in \mathbb{R}^n} A_\tau^\delta f(x) \lesssim \widetilde{\mathcal{K}}_\delta f(\tau) \lesssim \delta^{-\frac{(n-d)}{p}} \|f\|_p. \quad (6.3)$$

The proof of the Proposition relies on an inductive estimate for

$$]Q_\delta := \|\mathcal{K}_\delta : L^2(\mathbb{R}^n) \rightarrow L^2(\mathrm{Gr}(d, n))\|,$$

having defined the Fourier localized operator

$$\mathcal{K}_\delta f(\sigma) := \sup_{x \in \mathbb{R}^n} |A_{\sigma,1}^{\delta} f(x)|, \quad \sigma \in \mathrm{Gr}(d, n).$$

Let Φ be as in (2.6) and set $f_\delta := \mathcal{F}^{-1}[\widehat{f}(\cdot)(1 - \Phi(2^4\delta\cdot))]$. Then

$$A_{\sigma,1}^{\delta} f = A_{\sigma,1}^{>4\delta} f + A_{\sigma,1}^{\delta} f_\delta, \quad \sigma \in \mathrm{Gr}(d, n)$$

which tells us immediately that

$$Q_\delta \leq Q_{4\delta} + \sup_{\|f\|_2=1} \|\mathcal{K}_\delta f_\delta\|_2. \quad (6.4)$$

The important fact to notice here is that due to the support of f_δ being contained in $\text{Ann}^+(\delta)$, we have

$$\sigma \in B_{h\delta}(\tau) \implies A_{\sigma,1}^{>\delta} f_\delta = A_{\sigma,1}^{>\delta} O_{\tau,\delta} f$$

where $O_{\tau,\delta}$ is the frequency cutoff to the set $O_{\tau,\delta} := \{\xi \in \text{Ann}^+(\delta) : |\Pi_\tau \xi| \leq 2\delta|\xi|\}$. Notice that by Lemma 2.2 we have for every δ -separated set $\Sigma_\delta \subset \text{Gr}(d, n)$ that

$$\sum_{\tau \in \Sigma_\delta} \|O_{\tau,\delta} f\|_2^2 \lesssim \delta^{-d(n-d-1)} \|f\|_2^2 \quad (6.5)$$

Therefore if Σ_δ is a maximal $2^{-11}\delta$ -net in $\text{Gr}(d, n)$, relying on (6.3) to pass to the second line and on the overlap estimate (6.5)

$$\begin{aligned} \|K_\delta f_\delta\|_2^2 &\leq \delta^{d(n-d)} \sum_{\tau \in \Sigma_\delta} \sup_{\sigma \in B_{2^{-10}\delta}(\tau)} \sup_{x \in \mathbb{R}^n} |A_{\sigma,1}^{>\delta} O_{\tau,\delta} f(x)|^2 \\ &\lesssim \delta^{(d-1)(n-d)} \sum_{\tau \in \Sigma_\delta} \|O_{\tau,\delta} f\|_2^2 \lesssim \delta^{2d-n} \end{aligned} \quad (6.6)$$

Combining (6.6) with (6.4) yields the recursion

$$Q_\delta \leq Q_{4\delta} + \Theta \delta^{2d-n},$$

for some dimensional constant Θ . This proves the proposition via easy induction. $\square$

6.4. The Nikodym maximal operator

For $1 \leq d < n$, $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, consider the maximal δ -plate averages

$$\mathcal{N}_\delta f(x) := \sup_{\sigma \in \text{Gr}(d,n)} \sup_{x \in T_\delta(\sigma)} |f|, \quad x \in \mathbb{R}^n.$$

The role of $\text{Gr}(d, n)$ is kept implicit in the notation of the Nikodym maximal function $\mathcal{N}_\delta$. The study of the Nikodym maximal operator is motivated by the question of existence and dimension of (d, n) -Nikodym sets as defined below.

Definition 3. We say that $A \subset \mathbb{R}^n$ is a (d, n) -Nikodym set if $|A| = 0$ and for every $x \in A$ there exists an affine d -plane $\sigma + y$, with $\sigma \in \text{Gr}(d, n)$ and $y \in \mathbb{R}^n$, such that

$$x \in \sigma + y \quad \text{and} \quad B_n(y, 1) \cap (\sigma + y) \subset E.$$

For $d = 1$ it is easy to see that the Keakeya conjecture would imply that all $(1, n)$ -Nikodym sets have Hausdorff dimension at least n . Again for the case $d = 1$ the maximal Keakeya conjecture (6.2) is equivalent to the statement that for all $\delta, \varepsilon > 0$ we have

$$\|\mathcal{N}_\delta f\|_{L^n(\mathbb{R}^n)} \lesssim_\varepsilon \delta^{-\varepsilon} \|f\|_{L^n(\mathbb{R}^n)}, \quad d = 1.$$

On the other hand, while for $d > n/2$ there are no (d, n) -Besicovitch sets, it is known that for all $1 \leq d < n$ there exist (d, n) -Nikodym sets; see [14, Corollary 6.6]. It is thus natural to focus on lower bounds for the Hausdorff dimension of such Nikodym sets. In order to formulate the relevant conjecture on the maximal level we briefly explore below the relation of Hausdorff dimension with $L^p(\mathbb{R}^n)$ -bounds for $\mathcal{N}_\delta$. The proof is similar to [20, Theorem 22.9].

Proposition 6.5. *Suppose that there exists $1 \leq p < \infty$, $1 \leq q < \infty$ and $\varepsilon > 0$ such that the following estimate holds. For all $\delta, \varepsilon > 0$ there holds*

$$\|\mathcal{N}_\delta f : L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)\| \lesssim_\varepsilon \delta^{-\frac{\varepsilon}{p}}.$$

Then every (d, n) -Nikodym set has Hausdorff dimension at least $n - \varepsilon$.

For general $1 \leq d < n$ the counterexample presented in Proposition 5.4 for the maximal operator $M_{\Sigma, \{1\}}$ can be easily modified for the Nikodym maximal function with $\delta^{-1} = N^{1/(n-d)}$, showing that the following conjecture would yield the best possible quantification of the range of boundedness.

Conjecture G. *Let $\mathcal{N}_\delta$ denote the (d, n) -Nikodym maximal function for $1 \leq d < n$. There holds*

$$\|\mathcal{N}_\delta : L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)\| \lesssim \begin{cases} \delta^{-\frac{n-d+1-p}{p}}, & 1 < p < n-d+1, \\ (\log \delta^{-1})^{\alpha(n,d,p)}, & p \geq n-d+1, \end{cases}$$

for some exponent $\alpha(n, d, p) > 0$.

The conjecture above suggests that the critical exponent for the (d, n) -Nikodym problem is $p_o := n - d + 1$. We verify the conjecture in the case $d = n - 1$, so that $p_o = 2$, in Section 7 below; in fact we prove the same logarithmic dependence for a more general operator given with respect to $(n - 1)$ -plates in $\mathbb{R}^n$ of arbitrary eccentricity. A range of off-diagonal estimates can be conjectured by interpolating the estimate of Conjecture G with the trivial $L^1 \rightarrow L^\infty$ estimate.

Conjecture H. *Let $\mathcal{N}_\delta$ denote the (d, n) -Nikodym maximal function for $1 \leq d < n$. For $1 \leq p \leq n - d + 1$, $q = p'(n - d)$, and all $\varepsilon > 0$*

$$\|\mathcal{N}_\delta f : L^p(\mathbb{R}^n) \rightarrow L^q(\mathbb{R}^n)\| \lesssim \begin{cases} \delta^{-\left(\frac{n-d+1-p}{p}\right)}, & 1 \leq p < n-d+1, \\ (\log \delta^{-1})^{\beta(n,d,p)}, & p = n-d+1, \end{cases}$$

for some exponent $\beta(n, d, p) > 0$.

For the case of general $1 < d < n$, we include here the proof of the best possible $L^2(\mathbb{R}^n)$ -estimate for the (d, n) -Nikodym maximal operator, namely the proof of Theorem C as stated in the introduction of the paper. Unlike the previous parts of this paper, the proof of Theorem C, as well as the one yielding the sharp critical bound for $d = n - 1$ in Section 7 below, rely on geometric rather than Fourier analytic considerations, together with TT^* types of arguments. In particular we will rely on precise estimate for pairwise intersections of δ -neighborhoods of elements of $\text{Gr}(d, n)$. When $n \geq 2d$ pairwise intersections of elements of $\text{Gr}(d, n)$ may have any dimension between 0 and $d - 1$, requiring a corresponding classification of the volume estimates; these estimates are stated and proved in Lemma 6.9 below. Similar estimates have appeared in [25] for $d = 2$.

6.6. Pairwise intersections in $\text{Gr}(d, n)$ and volume estimates

For $\sigma \in \text{Gr}(d, n)$ it is convenient to denote $\mathbb{S}^\sigma := \mathbb{S}^n \cap \sigma$. Let $\sigma, \tau \in \text{Gr}(d, n)$ and suppose that $\eta = \sigma \cap \tau \in \text{Gr}(m, n)$ with $0 \leq m < d$. We define the angles

$$0 < \theta_{m+1}(\sigma, \tau) \leq \cdots \leq \theta_d(\sigma, \tau)$$

inductively as follows. Initially set $\sigma_{m+1} = \sigma \cap \eta^\perp$, $\tau_{m+1} = \tau \cap \eta^\perp$. For $m + 1 \leq j < d$ and σ_j, τ_j have been defined let

$$(s_j, t_j) := \arg \min \{ \arccos(st) : s \in \mathbb{S}^{\sigma_j}, t \in \mathbb{S}^{\tau_j} \}, \quad \theta_j := \arccos(s_j t_j),$$

$$\sigma_{j+1} := \sigma_j \cap s_j^\perp, \quad \tau_{j+1} := \tau_j \cap t_j^\perp,$$

and repeat with $j + 1$ in place of j . The algorithm stops when $j = d$.

Remark 6.7. Notice that $\theta_{m+1} > 0$, as there is no pair $(s, t) \in \mathbb{S}^{\sigma_{m+1}} \times \mathbb{S}^{\tau_{m+1}}$ with s and t colinear, and that clearly θ_j is nondecreasing. Also notice that if $s_1 = t_1, \dots, s_m = t_m$ is an orthonormal basis of $\sigma \cap \tau$, then $\{s_1, \dots, s_d\}$ and $\{t_1, \dots, t_d\}$ are orthonormal bases of σ, τ respectively. In addition, the construction of $s_{m+1}, \dots, s_d, t_{m+1}, \dots, t_d$ yields a further orthogonality property; namely if $\pi_j = \text{span}\{s_j, t_j\}$ for $m + 1 \leq j \leq d$ then

$$1 \leq j < k \leq d \implies xy = 0 \quad \forall x \in \pi_j, y \in \pi_k.$$

This is obvious if $j \leq m$. Otherwise, if $t \in \tau_k$ is such that $s_j t \neq 0$ for some $j < k$, we may write $t = (\cos \phi) \tilde{t} + (\sin \phi) s_j$ for some $\phi \neq 0$ and $\tilde{t} \in \tau_k$. Thus for any $s \in \sigma_k$ as $ss_j = 0$ we have

$$st = (\cos \phi) s \tilde{t} + (\sin \phi) ss_j = (\cos \phi) s \tilde{t} < s \tilde{t}$$

thus no $s \in \sigma_k$ exists such that (s, t) is a minimizer. It follows that t_k is orthogonal to s_j whence $xy = 0$ for all $x \in \pi_j, y \in \pi_k$. The angles and bases

$$0 = \theta_1 = \cdots = \theta_m, \quad 0 < \theta_{m+1} \leq \cdots \leq \theta_d \leq \frac{\pi}{2},$$

$$S = \{s_1, \dots, s_d\}, \quad T = \{t_1, \dots, t_d\},$$

are respectively called *canonical angles and bases* and may be obtained, respectively, as inverse cosines of the eigenvalues and eigenvectors of the $2d \times 2d$ matrix M of inner products between elements of arbitrary orthonormal bases of σ and τ . Finally we stress that $\theta_d(\sigma, \tau) \sim d(\sigma, \tau)$, where the latter refers to the $\text{Gr}(d, n)$ distance.

Remark 6.8. Given $\sigma, \tau \in \text{Gr}(d, n)$, suppose $\sigma \cap \tau \in \text{Gr}(m, n)$, so that $\zeta = \text{span}\{\sigma, \tau\} \in \text{Gr}(2d - m, n)$. Let $S = \{s_1, \dots, s_d\}$ and $T = \{t_1, \dots, t_d\}$ be the principal bases of σ, τ respectively that we constructed above. For each $j = m + 1, \dots, d$ let z_j be the unit vector so that

$$t_j = (\cos \theta_j)s_j + (\sin \theta_j)z_j.$$

Notice that z_j belongs to π_j and is thus orthogonal to s_k for all $k \neq j$ by Remark 6.7 and to s_j by construction. It follows that

$$Z = \{s_1, \dots, s_d, z_{m+1}, \dots, z_d\}$$

is an orthonormal basis of ζ .

Lemma 6.9. Let $\sigma, \tau \in \text{Gr}(d, n)$ and suppose that $\eta = \sigma \cap \tau \in \text{Gr}(m, n)$. Let $a, b \in \mathbb{R}^n$, $P(\sigma) = a + T_\delta(\sigma)$, $P(\tau) = b + T_\delta(\tau)$ be δ -plates. Then

$$|P(\sigma) \cap P(\tau)| \lesssim_{n,d} \delta^{n-m} \left(\prod_{j=m+1}^d \max\{\delta, \theta_j(\sigma, \tau)\} \right)^{-1}.$$

Proof. Let $\{s_1, \dots, s_d\}$ and $\{t_1, \dots, t_d\}$ be the orthonormal bases of σ, τ respectively we obtained with the principal angle construction. Let $\{s_{d+1}, \dots, s_n\}$ also be a basis of $\sigma^\perp$. Pick any point $p \in P(\sigma) \cap P(\tau)$. Choose coordinates $y = (y_1, \dots, y_j) \in \mathbb{R}^n$ so that p is the origin and $y_j = y \cdot s_j$. We claim that $P(\sigma) \cap P(\tau)$ is contained in the intersection of the n bands

$$B_j := \begin{cases} \{y \in \mathbb{R}^n : |y_j| < 3\}, & j = 1, \dots, m, \\ \left\{ y \in \mathbb{R}^n : |y_j| < \frac{C\delta}{\max\{\delta, \theta_j\}} \right\}, & j = m + 1, \dots, d, \\ \{y \in \mathbb{R}^n : |y_j| < 3\delta\}, & j = d + 1, \dots, n. \end{cases}$$

The claim readily yields the conclusion of the lemma and is completely obvious for $j = 1, \dots, m$ and $j = d + 1, \dots, n$, because for those values of j one has $P(\sigma) \subset B_j$ as well. It is also obvious for $j \in \{m + 1, \dots, d\}$ if $\max\{\delta, \theta_j\} = \delta$, thus we fix now

$j \in \{m+1, \dots, d\}$ and arguing by contradiction, suppose that there exists $y \in P(\sigma) \cap P(\tau)$ with $|y_j| \geq \frac{C\delta}{\theta_j}$. Let z_j be as in Remark 6.8. It follows that

$$|y \cdot z_j| \leq 2\delta.$$

However simple geometry shows that

$$\tan \theta_j = \frac{|y \cdot z_j|}{|y_j|} \leq \frac{2\delta}{\frac{C\delta}{\theta_j}} = \frac{2\theta_j}{C},$$

a contradiction if C is large enough. $\square$

In the next lemma we describe how to construct the essentially minimal dilate of a δ -plate $P(\sigma)$ that covers a nearby plate $P(\tau)$, depending on the principal angles of σ, τ . We need some notation first

Let $\sigma, \tau \in \text{Gr}(d, n)$ and suppose that $\sigma \cap \tau \in \text{Gr}(m, n)$. Let S, T and Z be the respective orthonormal bases of σ, τ and $\zeta = \text{span}(\sigma, \tau)$ constructed in Remark 6.8. Then the plates $T_\delta(\sigma), T_\delta(\tau)$ may be described by

$$\begin{aligned} T_\delta(\sigma) &= \left\{ x : \max_{\{1, \dots, d\}} |x \cdot s_j| < 1, \max_{m+1 \leq j \leq d} |x \cdot z_j| < \delta, \max_{w \in \zeta^\perp} |x \cdot w| < \delta \right\}, \\ T_\delta(\tau) &= \left\{ x : \max_{\{1, \dots, m\}} |x \cdot s_j| < 1, \max_{\{m+1, \dots, d\}} |x \cdot [(\cos \theta_j)s_j + (\sin \theta_j)z_j]| < 1, \right. \\ &\quad \left. \max_{w \in \tau^\perp} |x \cdot w| < \delta \right\}. \end{aligned}$$

We define the dilation $T_\delta^{+\tau}(\sigma)$ of $T_\delta(\sigma)$ by

$$T_\delta^{+\tau}(\sigma) := \left\{ x : \max_{\{1, \dots, d\}} |x \cdot s_j| < 3, \max_{m+1 \leq j \leq d} |x \cdot z_j| < 3 \max\{\delta, \theta_j\}, \max_{w \in \zeta^\perp} |x \cdot w| < 3\delta \right\}.$$

Lemma 6.10. *Let $a, b \in \mathbb{R}^n$, $P(\sigma) = a + T_\delta(\sigma)$, $P(\tau) = b + T_\delta(\tau)$ be δ -plates with $P(\sigma) \cap P(\tau) \neq \emptyset$. Then*

$$P(\tau) \subset a + T_\delta^{+\tau}(\sigma).$$

Proof. By translation invariance, we may assume $a = 0$. Then $P(\tau)$ is contained in the moderate dilate

$$\begin{aligned} 3T_\delta(\tau) &:= \left\{ x : \max_{\{1, \dots, m\}} |x \cdot s_j| < 3, \max_{\{m+1, \dots, d\}} |x \cdot [(\cos \theta_j)s_j + (\sin \theta_j)z_j]| < 3, \right. \\ &\quad \left. \max_{w \in \tau^\perp} |x \cdot w| < 3\delta \right\}. \end{aligned}$$

Simple geometry tells us that when $j \in \{m+1, \dots, d\}$

$$x \in 3T_\delta(\tau) \implies |x \cdot [(\cos \theta_j)s_j + (\sin \theta_j)z_j]| \leq 3 \implies |x \cdot z_j| \leq 3 \sin \theta_j \leq 3\theta_j.$$

On the other hand it is obvious that $|x \cdot s_j| < 3$ for all $1 \leq j \leq d$ when $x \in 3T_\delta(\tau)$. Therefore $3T_\delta(\tau) \subset T_\delta^{+\tau}(\sigma)$, and the proof is complete. $\square$

6.11. The proof of Theorem C

As noted in the discussion leading to the formulation of Conjecture G, the bound of Theorem C is best possible so it suffices to prove the upper bound. We begin by linearizing the maximal operator $\mathcal{N}_\delta$ as follows. For $\Sigma \subset \text{Gr}(d, n)$ that will remain fixed throughout the proof we let $\mathcal{T}_\Sigma$ denote the collection of all δ -plates of the form $a_T + T_\delta(\sigma_T)$ for $a_T \in \mathbb{R}^n$ and $\sigma_T \in \Sigma$. Now given $f \in \mathcal{S}(\mathbb{R}^n)$ and $\mathcal{T} \subset \mathcal{T}_\Sigma$ we consider the linear operator

$$\mathcal{N}_\mathcal{T}f(x) := \sum_{T \in \mathcal{T}} \left(\int_T f \right) \mathbf{1}_{F_T}(x)$$

where $F_T \subset T$ for every $T \in \mathcal{T}$ and the collection $\{F_T\}_{T \in \mathcal{T}}$ is pairwise disjoint. Denoting by $\mathcal{N}_\mathcal{T}^*$ the adjoint of $\mathcal{N}_\mathcal{T}$ we have that

$$\begin{aligned} \|\mathcal{N}_\delta : L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)\| &\approx \sup_{\mathcal{T} \subset \mathcal{T}_\Sigma} \|\mathcal{N}_\mathcal{T} : L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)\| \\ &\approx \sup_{\mathcal{T} \subset \mathcal{T}_\Sigma} \sup_{E \subset \mathbb{R}^n} \frac{\|\mathcal{N}_\mathcal{T}^* \mathbf{1}_E\|_2}{|E|^{\frac{1}{2}}}, \end{aligned}$$

where the supremum in $\mathcal{T} \subset \mathcal{T}_\Sigma$ can be taken over finite collections $\mathcal{T}$. As the adjoint operator has the form

$$\mathcal{N}_\mathcal{T}^*g := \sum_{T \in \mathcal{T}} \left(\frac{1}{|T|} \int_{F_T} g \right) \mathbf{1}_T(x)$$

we readily see that

$$\mathcal{N}_\mathcal{T}^* \mathbf{1}_E = \sum_{T \in \mathcal{T}} |F_T \cap E| \frac{\mathbf{1}_T}{|T|} =: \sum_{T \in \mathcal{T}} |E_T| \frac{\mathbf{1}_T}{|T|}$$

with $E_T := F_T \cap E$ pairwise disjoint and $E_T \subset T$. Note that the collection $\{|E_T|\}_{T \in \mathcal{T}}$ is a Carleson sequence: for any open set $U \subset \mathbb{R}^n$ we have

$$\sum_{\substack{T \in \mathcal{T} \\ T \subset U}} |E_T| \leq |U \cap E|.$$

We now expand the square of the L^2 -norm as follows

$$\|\mathcal{N}_{\mathcal{T}}^* \mathbf{1}_E\|_2^2 = \sum_{T \in \mathcal{T}} |E_T| \sum_{\substack{T' \in \mathcal{T} \\ T \cap T' \neq \emptyset}} |E_{T'}| \frac{|T \cap T'|}{|T||T'|}.$$

For fixed $T \in \mathcal{T}$ and $0 \leq k \leq \log \delta^{-1}$ we define

$$\mathcal{T}_k(T) := \{T' \in \mathcal{T} : T \cap T' \neq \emptyset, 2^k \delta \leq \text{dist}(T, T') \leq 2^{k+1} \delta\}.$$

This notation allows us to write

$$\|\mathcal{N}_{\mathcal{T}}^* \mathbf{1}_E\|_2^2 = \sum_{k=0}^{\log \delta^{-1}} \sum_{T \in \mathcal{T}} \frac{|E_T|}{|T|} \sum_{T' \in \mathcal{T}_k(T)} |E_{T'}| \frac{|T \cap T'|}{|T'|}. \quad (6.7)$$

Now for fixed k, T and $T' \in \mathcal{T}_k(T)$ we use Lemma 6.9 to estimate for every $m \in \{0, \dots, d-1\}$

$$\frac{|T \cap T'|}{|T'|} \leq \delta^{n-m} \frac{1}{\delta^{d-m-1} \text{dist}(\sigma_T, \sigma_{T'})} \frac{1}{|T'|} \approx \frac{\delta^{n-d+1}}{2^k \delta} \frac{1}{\delta^{n-d}} = 2^{-k}.$$

Then for every fixed k we will have by Lemma 6.10 and the Carleson property of the sequence $\{|E_T|\}_{T \in \mathcal{T}}$

$$\sum_{T' \in \mathcal{T}_k} |E_{T'}| \leq \left| \bigcup_{T' \in \mathcal{T}_k} T' \right| \leq |T_\delta^+(\sigma_T)|.$$

Since $\theta_j(\sigma_T, \sigma_{T'}) \leq \theta_d(\sigma_T, \sigma_{T'}) \approx 2^k \delta$ and $2^k \delta \geq \delta$ we can use the definition of $T_\delta^+(\sigma_T)$ to estimate

$$|T_\delta^+(\sigma_T)| \lesssim (2^k \delta)^{n-d}.$$

Using the estimate in the last two displays and the calculation in (6.7) we gather

$$\|\mathcal{N}_{\mathcal{T}}^* \mathbf{1}_E\|_2^2 \lesssim \sum_{k=0}^{\log \delta^{-1}} \sum_{T \in \mathcal{T}} \frac{|E_T|}{|T|} (2^k \delta)^{n-d} 2^{-k} \lesssim \begin{cases} \delta^{-(n-d-1)} |E|, & 1 \leq d < n-1, \\ \log(\delta^{-1}) |E|, & d = n-1, \end{cases}$$

which proves the desired weak $(2, 2)$ norm-estimate. The corresponding strong $(2, 2)$ bound with an additional $\sqrt{\log \delta^{-1}}$ -term follows from the corresponding weak-type $(2, 2)$ bound and a well known interpolation argument of Strömberg, [29]; see also [17, p. 77–78] for the details of this argument.

7. The sharp bound for codimension one maximal operators

In this section, we prove Theorem B as a consequence of a more general directional Carleson embedding theorem, Theorem I below. Therefore, we work with fixed codimension 1, so that $n = d + 1 \geq 2$ throughout. Recall that our goal is to prove a sharp estimate in terms of the cardinality parameter N for the maximal d -subspace averaging operator $M_{\Sigma, (0, \infty)}$ when $\Sigma \subset \text{Gr}(d, d + 1)$ is a finite set with $\#\Sigma = N$.

Our setup is more conveniently described by taking advantage of the isometric isomorphism

$$\sigma \in \Sigma \subset \text{Gr}(d, d + 1) \mapsto v = \sigma^\perp \subset \text{Gr}(1, d + 1) \sim \mathbb{S}^d$$

By finite splitting and rotational invariance, we may work under the assumption that $V \subset \mathbb{S}^d$ is contained in a small neighborhood of e_n . This choice of coordinate system is conveniently exploited by modifying slightly our definition of d -plate (2.3), as follows. For a d -dimensional axis-parallel cube $I \subset e_n^\perp$ with center $(c_I, 0) \in \mathbb{R}^n$ and sidelength ℓ_I , for $v \in V$, an interval $K \subset \mathbb{R}$ with $|K| \leq \ell_I$, let

$$p(I, t, v) := \{y \in v^\perp + (c_I, t) : \Pi_{e_n^\perp} y \in I\}, \quad P(I, K, v) := \bigcup_{t \in K} p(I, t, v).$$

The set $Q = P(I, K, v)$ is a d -plate with orientation $v_Q = v$, basis $I_Q = I$, scale $s_Q = \ell_I$, height $K_Q = K$ and center $c_Q = (c_{I_Q}, c_{K_Q})$, where c_{K_Q} is the center of the interval $K \subset \mathbb{R}$. Each set $p(I_Q, t, v_Q)$ with $t \in K$ is referred to as the t -slice of Q .

Remark 7.1. The plate Q is the shearing of an axis-parallel box whose short side is oriented along e_n . To compare with (2.3), observe that if $\delta = \ell_{K_Q}/\ell_{I_Q}$, then Q and $Q' := c_Q + s_Q T_\delta(v_Q^\perp)$ are comparable, that is $Q \subset CQ'$, $Q' \subset CQ$ for a suitably chosen dimensional constant C .

We will work with different special collections of d -plates which we define below.

Definition 4. Let $V \subset \mathbb{S}^d$ be a finite set of directions, $\delta > 0$ be a small parameter.

- The collection of all d -plates in $\mathbb{R}^n$ with orientation along $v \in V$ will be denoted by $\mathcal{P}_v$, and $\mathcal{P}_V := \bigcup_{v \in V} \mathcal{P}_v$.
- For $Q = P(I, K, v) \in \mathcal{P}_V$, write $\text{af}_Q := v^\perp + c_Q$ and call af_Q the *plane* of Q .
- A d -plate $Q(I, K, v)$ will be called a (d, δ) -plate if $\ell(K_Q) = \delta > 0$, namely if it is a d -plate with fixed thickness $\delta > 0$. The subcollection of those (d, δ) -plates belonging to $\mathcal{P}_v$ is referred to by $\mathcal{P}_{v, \delta}$ and $\mathcal{P}_{V, \delta} := \bigcup_v \mathcal{P}_{v, \delta}$.
- Given a dyadic grid $\mathcal{D}$ in $\mathbb{R}^d$, special subcollections of $\mathcal{P}_v, \mathcal{P}_V$ are produced by defining $\mathcal{D}_v := \{Q \in \mathcal{P}_v : I_Q \in \mathcal{D}\}$, and $\mathcal{D}_V := \bigcup_{v \in V} \mathcal{D}_v$. The special subcollection of $\mathcal{D}_V$ consisting of (d, δ) -plates will be denoted by $\mathcal{D}_{V, \delta} := \bigcup_{v \in V} \mathcal{D}_{v, \delta}$.

- For a generic collection $\mathcal{Q} \subseteq \mathcal{P}_V$, set $\mathcal{Q}_v := \{Q \in \mathcal{Q} : v_Q = v\}$. This yields $\mathcal{Q} = \bigcup_{v \in V} \mathcal{Q}_v$.
- A partial order on $\mathcal{D}_V$ is defined as follows. If $Q, R \in \mathcal{D}_V$, say $Q \leq R$ if $Q \cap R \neq \emptyset$ and $I_Q \subseteq I_R$.
- If $\mathcal{Q} \subset \mathcal{P}_V$ we will use the notation $\text{sh}(\mathcal{Q}) := \bigcup_{Q \in \mathcal{Q}} Q$ for the *shadow* of the collection.

With these definitions we introduce below maximal operators defined with respect to collection of plates. For any $\mathcal{L} \subseteq \mathcal{P}_V$, set

$$M_{\mathcal{L}}f(x) := \sup_{L \in \mathcal{L}} \left(\int_L |f| \, dy \right) \mathbf{1}_L(x), \quad x \in \mathbb{R}^n.$$

Note that in general $M_{\mathcal{L}}$ is a *directional operator* as the plates have variable orientation. In the special case that $\mathcal{L} = \mathcal{L}_v \subseteq \mathcal{P}_v$ for some fixed $v \in V$ then $M_{\mathcal{L}_v}$ is pointwise bounded by the strong maximal function in a suitable coordinate system. Another special case of interest occurs when $\mathcal{L} \subseteq \mathcal{P}_{v,\delta}$ for fixed $v \in V$ and $\delta > 0$, in which case $M_{\mathcal{L}}$ is a one-parameter operator and satisfies weak $(1, 1)$ bounds uniformly in v and δ . Note that the weak $(1, 1)$ -bound persists if $\mathcal{L} \subset \mathcal{P}_v$ is a collection of fixed eccentricity: in that case the operators $M_{\mathcal{L}}$ are again of weak-type $(1, 1)$ uniformly in $v \in V$ and the eccentricity of the collection.

Carleson sequences. Directional Carleson sequences of positive numbers $\{a_Q\}_{Q \in \mathcal{D}_V}$ are introduced in the string of definitions that follow.

Definition 5. Let $\mathcal{L} \subset \mathcal{P}_V$ be a collection of d -plates and let $v \in V$ be a fixed direction. The collection $\mathcal{L}$ is *subordinate to* $\mathcal{T} \subset \mathcal{P}_v$ if for every $L \in \mathcal{L}$ there exists $T \in \mathcal{T}$ such that $L \subseteq T$.

We stress that $\mathcal{T} \subset \mathcal{P}_v$ in the definition above *only contains plates with fixed orientation* v .

Definition 6. Let $a = \{a_Q\}_{Q \in \mathcal{D}_V}$ be a sequence of positive numbers. The sequence a is an $(L^\infty\text{-normalized})$ *Carleson sequence* if for every $\mathcal{L} \subset \mathcal{D}_V$ which is subordinate to some $\mathcal{T} \subset \mathcal{P}_v$ for some fixed $v \in V$ we have

$$\sum_{L \in \mathcal{L}} a_L \leq |\text{sh}(\mathcal{T})|, \quad \text{mass}_a(\mathcal{Q}) := \sum_{Q \in \mathcal{D}_V} a_Q < \infty.$$

For $\mathcal{Q} \subset \mathcal{D}_V$ and a Carleson sequence $a = \{a_Q\}_{Q \in \mathcal{D}_V}$ define the *balayage*

$$T_{\mathcal{Q}}(a)(x) := \sum_{Q \in \mathcal{Q}} a_Q \frac{\mathbf{1}_Q(x)}{|Q|}, \quad x \in \mathbb{R}^n. \quad (7.1)$$

It follows from the definition above that if a is a Carleson sequence and $\mathcal{T} \subset \mathcal{P}_v$ for some fixed $v \in V$ then $\mathcal{T}$ is subordinate to itself and thus $\text{mass}_a(\mathcal{T}) \leq |\text{sh}(\mathcal{T})|$.

An L^2 -Carleson embedding theorem for d -plates. Here we describe and prove the main result of this section, a directional Carleson embedding theorem for d -plates in $\mathbb{R}^n$. In order to state it we also introduce for any $\mathcal{Q} \subseteq \mathcal{P}_V$ the notation

$$\widehat{\mathcal{Q}} := \bigcup_{1 \leq s \leq 100n} \bigcup_{Q \in \mathcal{Q}} (1+s)Q.$$

Theorem I. *Let $V \subset \mathbb{S}^d$ be a finite set of directions and $\mathcal{Q} \subseteq \mathcal{D}_V$ be a collection of d -plates in $\mathbb{R}^n$. We assume that the operators $\{M_{\mathcal{Q}_v} : v \in V\}$ satisfy*

$$\sup_{v \in V} \|M_{\mathcal{Q}_v} : L^1(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mathbb{R}^n)\| \lesssim_n 1.$$

If $a = \{a_Q\}_{Q \in \mathcal{D}_{V,\delta}}$ is a Carleson sequence then

$$\|T_{\mathcal{Q}}(a)\|_{L^2(\mathbb{R}^n)} \lesssim_n (\log \#V)^{\frac{1}{2}} \text{mass}_a(\mathcal{Q})^{\frac{1}{2}}$$

with implicit constant depending only upon dimension.

The proof of Theorem I begins with some reductions that simplify and highlight the main argument. First, for any $\mathcal{Q} \subseteq \mathcal{D}_V$ we expand the square of the L^2 -norm in the statement of the theorem as

$$\begin{aligned} \frac{1}{2} \|T_{\mathcal{Q}}(a)\|_{L^2(\mathbb{R}^n)}^2 &\leq \sum_{R \in \mathcal{Q}} a_R \int_R \sum_{\mathcal{Q} \ni Q \leq R} a_Q \frac{1_Q}{|Q|} \\ &\leq \mu(\log \#V) \text{mass}_a(\mathcal{Q}) + \mu(\#V) \sum_{k > \mu(\log \#V)} (k+1) \sup_{v \in V} \sum_{R \in \mathcal{Q}_{v,k}} a_R \end{aligned} \quad (7.2)$$

where $\mu > 0$ is a numerical constant to be chosen later and

$$\mathcal{Q}_{v,k} := \left\{ R \in \mathcal{Q}_v : \mu k \leq \int_R \sum_{\mathcal{Q} \ni Q \leq R} a_Q \frac{1_Q}{|Q|} < \mu(k+1) \right\}.$$

Thus the proof reduces to proving a suitable estimate for $\text{mass}_a(\mathcal{Q}_{v,k})$ for every fixed $v \in V$ and every $k > \mu(\log \#V)$. The next remark encapsulates some simple but important geometric observations that are at the heart of the argument.

Remark 7.2. Fix $R = P(I_R, K_R, v_R) \in \mathcal{P}_V$ and consider a d -plate $\mathcal{P}_V \ni Q = P(I_Q, K_Q, v_Q) \leq R$. Note that if $v_Q \neq v_R$ then $\text{af}_Q \cap \text{af}_R$ is a $(d-1)$ -dimensional affine subspace and let $\text{lin}_{Q,R}$ be the subspace parallel to $\text{af}_Q \cap \text{af}_R$. As v_Q, v_R lie in a small

neighborhood of $e_n^\perp$, the subspace $\Pi_{e_n^\perp} \text{lin}_{Q,R}$ has dimension $d-1$ as well and is a codimension 1 subspace of $e_n^\perp$. We may thus pick an orthonormal basis $G_{Q,R} := (g_1, \dots, g_d)$ for $e_n^\perp$ such that $(g_1, \dots, g_{d-1})$ is an orthonormal basis of $\Pi_{e_n^\perp} \text{lin}_{Q,R}$.

Let $\widehat{I}_{Q,R} \subset$ be the smallest d -dimensional cube in the coordinates $G_{Q,R}$ that contains I_Q . Then

$$\widehat{I}_{Q,R} \subset e_n^\perp, \quad I_Q \subseteq \widehat{I}_{Q,R} \subseteq dI_Q,$$

where the dilation is taken with respect to the center of I_Q . Furthermore, defining

$$\widehat{Q} := P(\widehat{I}_{Q,R}, K_Q, v_Q)$$

then $Q \subseteq \widehat{Q}$ and $|\widehat{Q}| \asymp_n |Q|$. In case $v_Q = v_R$ we just set $\widehat{Q} := Q$ for the sake of having a general definition. Finally, setting $\widehat{I}_R := dI_R$ and $\widehat{R} := P(\widehat{I}_R, K_R, v_R)$ yields

$$v_{\widehat{R}} = v_R, \quad \widehat{R} \supseteq R, \quad I_{\widehat{R}} \supseteq I_{\widehat{Q}} \quad \forall Q \leq R.$$

As $v_Q = v_{\widehat{Q}}$ and $\text{af}_{\widehat{Q}} = \text{af}_Q$, the plate $\widehat{Q}$ is a rotation and af_Q -tangential dilation of Q with respect to the line $\{c_Q + tv_Q : t \in v_Q\}$. Also, our construction yields that one of the $(d-1)$ -dimensional edges of $\widehat{Q}$ lies on an affine copy of $\text{af}_Q \cap \text{af}_R$. Note also that $\widehat{I}_{Q,R}$ depends both on Q and R ; we will however many times suppress the R -dependence as R will be fixed and just write $\widehat{I}_Q$ in place of $\widehat{I}_{Q,R}$.

With these definitions and conventions in hand, we state a geometric slicing lemma that will be important for the proof of Theorem I.

Lemma 7.3. *Let $R = P(I_R, K_R, v_R) \in \mathcal{D}_v$ and $\mathcal{D}_V \ni Q = P(I_Q, K_Q, v_Q) \leq R$, and $\widehat{R}, \widehat{Q}$ be as in Remark 7.2. Let $K \subset \mathbb{R}$ be an interval with $\Pi_{e_n}(\widehat{Q}) \not\subseteq 3K$ and $K_R \subseteq K$. Then*

$$\max_{a \in \mathbb{R}} \frac{|\widehat{Q} \cap p(I_{\widehat{R}}, a, v_{\widehat{R}})|}{|I_{\widehat{R}}|} \lesssim_n \frac{1}{|I_{\widehat{R}}||K|} \int_{I_{\widehat{R}} \times 3K} \mathbf{1}_{\widehat{Q}}$$

with implicit constant depending only upon dimension.

Proof. By composing d shearing transformations, we reduce to the case of $v_{\widehat{R}} = e_n$. In this case the slices may simply be described by $p(I_{\widehat{R}}, a, v_{\widehat{R}}) = I_{\widehat{R}} \times \{a\}$. The conclusion is immediate if $v_Q = v_{\widehat{R}} = e_n$ since then the slices of Q by planes perpendicular to e_n have all constant measure. Thus we assume that $e_n = v_{\widehat{R}} \neq v_Q$. Then, $\text{af}_Q \cap \text{af}_R$ is an affine space of dimension $d-1$ parallel to the subspace $\text{lin}_{Q,R}$. As $v_{\widehat{R}} = e_n$, we have $\text{lin}_{Q,R} = \Pi_{e_n^\perp} \text{lin}_{Q,R}$, in other words $\text{lin}_{Q,R} \subset e_n^\perp$. Setting $H(t) := t + (\text{lin}_{Q,R})^\perp$ for $t \in \text{af}_R \cap \text{af}_Q \cap \widehat{Q}$, we have for any $a \in \mathbb{R}$ that

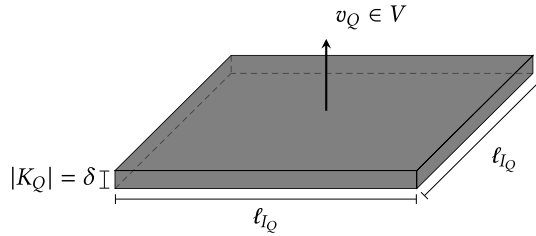


Fig. 7.1. A plate $Q = P(I_Q, K_Q, v_Q)$ in $\mathbb{R}^3$ perpendicular to $v_Q \in V$.

$$|\widehat{Q} \cap (I_{\widehat{R}} \times \{a\})| = \int_{\text{af}_R \cap \text{af}_Q \cap \widehat{Q}} |H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a\})| dt.$$

Now observe that for each $t \in \text{af}_R \cap \text{af}_Q \cap \widehat{Q}$ the set $H(t) \cap \widehat{Q}$ is a two-dimensional parallelogram lying on $H(t)$ with long side perpendicular to v_Q and short side of length $|K_Q|$ parallel to e_n (Fig. 7.1). Our assumptions yield

$$\Pi_{e_n}(\widehat{Q}) \cap K \neq \emptyset, \quad \Pi_{e_n}(\widehat{Q}) \not\subseteq 3K;$$

a two-dimensional calculation then reveals that there exists a set $A \subset 3K \setminus K$ with $|A| \geq |K|/3$ such that for each $a' \in A$ and all $t \in \text{af}_R \cap \text{af}_Q \cap \widehat{Q}$

$$|H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a'\})| \approx \max_{a \in \mathbb{R}} |H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a\})|;$$

the implicit constant in the estimate above is independent everything and in particular this estimate holds uniformly in t . This clearly implies that

$$\max_{a \in \mathbb{R}} |H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a\})| \leq \int_A |H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a'\})| da' \lesssim \int_{3K} |\widehat{Q} \cap H(t) \cap (I_{\widehat{R}} \times \{a'\})| da'.$$

The conclusion of the lemma readily follows by noticing that for every $a \in \mathbb{R}$ the quantity $|H(t) \cap \widehat{Q} \cap (I_{\widehat{R}} \times \{a\})|$ is independent of t and integrating for $t \in \text{af}_R \cap \text{af}_Q \cap \widehat{Q}$. $\square$

We now return to the estimate for $\text{mass}_a(\mathcal{Q}_{v,k})$. Letting $R \in \mathcal{Q}_{v,k}$ we note that for every $\mathcal{Q} \ni Q \leq R$ we have $\widehat{Q} \supseteq Q$ and $|\widehat{Q}| \approx_n |Q|$ so that

$$\mu k \leq \int_R \sum_{\mathcal{Q} \ni Q \leq R} a_Q \frac{1_Q}{|Q|} \lesssim_n \int_R \sum_{\mathcal{Q} \ni Q \leq R} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|} \lesssim_n \int_{\widehat{R}} \sum_{\mathcal{Q} \ni Q \leq R} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|}$$

since $R \subseteq \widehat{R}$ and $|R| \simeq_n |\widehat{R}|$. For a collection of plates $\mathcal{L} \subseteq \mathcal{D}_V$ we define

$$B_R^{\mathcal{L}} := \int_{\widehat{R}} \sum_{\substack{Q \in \mathcal{L} \\ Q \leq R}} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|}.$$

Then the considerations above imply that

$$\mathcal{Q}_{v,k} \subseteq \widehat{\mathcal{Q}}_{v,k} := \{Q \in \mathcal{Q}_v : B_R^Q > c_n \mu k\} \quad (7.3)$$

for some dimensional constant $c_n > 0$ and μ to be chosen.

The proof of Theorem I relies upon the estimate for $\text{sh}(\widehat{\mathcal{Q}}_{v,k})$ contained in the following lemma.

Lemma 7.4. *Let $\delta > 0$, $V \subset \mathbb{S}^d$ be a finite set of directions, and $\mathcal{Q} \subseteq \mathcal{D}_V$ satisfy the assumptions of Theorem I. For $v \in V$ and $k \geq 1$ we define $\widehat{\mathcal{Q}}_{v,k}$ as in (7.3) above, with μ a sufficiently large dimensional constant. Then*

$$|\text{sh}(\widehat{\mathcal{Q}}_{v,k})| \lesssim_n 2^{-k} \text{mass}_a(\mathcal{Q}).$$

With Lemma 7.4 in our disposal we can complete the proof of the main result of this paragraph.

Proof of Theorem I. Let $\mathcal{Q} \subseteq \mathcal{D}_V$ be a collection of plates such that the operators $\{M_{\mathcal{Q}_v} : v \in V\}$ are of weak-type $(1, 1)$, uniformly in $v \in V$, with weak $(1, 1)$ bound depending only on the dimension. By (7.2) combined with (7.3) and the estimate of Lemma 7.4 we have that

$$\|T_{\mathcal{Q}}(a)\|_{L^2(\mathbb{R}^n)}^2 \lesssim \mu \left((\log \#V) + (\#V) \sum_{k \geq \mu(\log \#V)} k 2^{-k} \right) \text{mass}_a(\mathcal{L}) \lesssim \mu(\log \#V) \text{mass}_a(\mathcal{L})$$

so that

$$\|T_{\mathcal{Q}}(a)\|_{L^2(\mathbb{R}^n)} \leq C_n \mu^{\frac{1}{2}} (\log \#V)^{\frac{1}{2}} \text{mass}_a(\mathcal{Q})^{\frac{1}{2}}$$

with μ as in the assumption of Lemma 7.4. Note that this means that μ can be chosen to be a dimensional constant and this completes the proof of the theorem. $\square$

It remains to prove Lemma 7.4 which follows by an iterative application of the lemma below, as in the proof of [1, Lemma 2.21].

Lemma 7.5. *Let $V \subset \mathbb{S}^d$ be a finite set of directions and $\mathcal{L} \subseteq \mathcal{R} \subseteq \mathcal{D}_V$ such that for every $L \in \mathcal{L}$ there exists $R \in \mathcal{R}$ with $L \leq R$. Furthermore we assume that*

$$\sup_{v \in V} \|M_{\widehat{\mathcal{R}}_v} : L^1(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mathbb{R}^n)\| \lesssim_n 1.$$

Fix some $v \in V$ and let μ be a sufficiently large dimensional constant. There exists $\mathcal{L}_1 \subset \mathcal{L}$ such that

$$(i) \text{ mass}_a(\mathcal{L}_1) \leq \frac{1}{2} \text{mass}_a(\mathcal{L});$$

(ii) For all plates $R \in \mathcal{R}_v$ such that $B_R^{\mathcal{L}} > \mu$ we have

$$B_R^{\mathcal{L}} \leq \mu + B_R^{\mathcal{L}_1}.$$

Proof. Without loss of generality we can assume that $v = e_n$. Let $\mathcal{R}'_v$ denote the collection of $R \in \mathcal{R}_v$ with $B_R^{\mathcal{L}} > \mu$. For the plates $R \in \mathcal{R}'_v$ and plates $Q \in \mathcal{L}$ with $Q \leq R$ we define $\widehat{R}, \widehat{Q}$ as in Remark 7.2. We remember also the notation $\widehat{R} = P(\widehat{I}_R, K_R, v_R)$ and $Q = P(\widehat{I}_Q, K_Q, v_Q)$; since we will always consider the case $Q \leq R$ we have that $\widehat{I}_Q \subseteq \widehat{I}_R$.

Given some interval $K \subset \mathbb{R}$ and $R \in \mathcal{R}'_v$ we define the collections

$$\begin{aligned} \mathcal{B}_{R,K}^{\text{in}} &:= \{Q \in \mathcal{L} : Q \leq R, \quad \pi_{e_n}(\widehat{Q}) \subseteq 3K\}, \\ \mathcal{B}_{R,K}^{\text{out}} &:= \{Q \in \mathcal{L} : Q \leq R, \quad \pi_{e_n}(\widehat{Q}) \not\subseteq 3K\}. \end{aligned}$$

We define

$$B_{R,K}^{\text{in}} := \sum_{Q \in \mathcal{B}_{R,K}^{\text{in}}} a_Q \frac{|Q' \cap (\widehat{I}_R \times K)|}{|\widehat{Q}| |\widehat{I}_R \times K|}, \quad B_{R,K}^{\text{out}} := \sum_{Q \in \mathcal{B}_{R,K}^{\text{out}}} a_Q \frac{|\widehat{Q} \cap (\widehat{I}_R \times K)|}{|\widehat{Q}| |\widehat{I}_R \times K|},$$

and note that for any K we have the splitting

$$B_R^{\mathcal{L}} = \int_{\widehat{R}} \sum_{Q \in \mathcal{B}_{R,K}^{\text{in}}} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|} + \int_{\widehat{R}} \sum_{Q \in \mathcal{B}_{R,K}^{\text{out}}} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|}.$$

Easy case: Let $\mathcal{R}_1$ be the collection of those $R \in \mathcal{R}'_v$ such that $B_{R,K_R}^{\text{out}} \leq \mu$. Then we have for $R \in \mathcal{R}_1$ that

$$B_R^{\mathcal{L}} \leq \mu + B_{R,K_R}^{\text{in}} = \mu + B_R^{\mathcal{L}'_1}, \quad \mathcal{L}'_1 := \bigcup_{\rho \in \mathcal{R}_1} \mathcal{B}_{\rho,K_\rho}^{\text{in}}.$$

Noting that for all $Q \in \mathcal{L}'_1$ we have that $Q \subseteq \widehat{Q} \subseteq \widehat{I}_\rho \times 3K_\rho$ for some $\rho \in \mathcal{R}_1$, we get

$$\text{mass}_{a,1}(\mathcal{L}'_1) := \sum_{Q \in \mathcal{L}'_1} a_Q \leq \left| \bigcup_{\rho \in \mathcal{R}_1} \bigcup_{Q \in \mathcal{B}_{\rho,K_\rho}^{\text{in}}} Q \right| \leq \left| \bigcup_{\rho \in \mathcal{R}_1} (\widehat{I}_\rho \times 3K_\rho) \right|.$$

Furthermore for every $\rho \in \mathcal{R}_1$ we have

$$\mu < B_\rho^{\mathcal{L}} = \int_{\widehat{\rho}} \sum_{\substack{Q \leq \rho \\ Q \in \mathcal{L}}} a_Q \frac{1_{\widehat{Q}}}{|\widehat{Q}|}.$$

As all the $\rho \in \mathcal{R}_1$ have fixed orientation our assumption entails that the maximal operator $M_{\widehat{\mathcal{R}}_v} = M_{\widehat{\mathcal{R}}_{e_n}}$ is of weak type $(1, 1)$ with constant depending only on the dimension. Since

$$\bigcup_{\rho \in \mathcal{R}_1} (\hat{I}_\rho \times 3K_\rho) \subseteq \left\{ M_{\hat{\mathcal{R}}_v} \left(\sum_{Q \in \mathcal{L}} a_Q \frac{\mathbf{1}_{\hat{Q}}}{|\hat{Q}|} \right) \gtrsim_n \mu \right\},$$

we conclude that $\text{mass}_a(\mathcal{L}'_1) \leq \text{mass}_a(\mathcal{L})/4$ if $\mu > 0$ is chosen to be a sufficiently large dimensional constant.

The main case: Here we consider $R \in \mathcal{R}_2 := \mathcal{R}'_v \setminus \mathcal{R}_1$. Let us write again $\hat{R} = \hat{I}_R \times K_R$ and consider the intervals J of the form $J = 3^\ell K_R$ for $\ell \geq 0$ such that $B_{R,J}^{\text{out}} > \mu$. Since $B_{R,K_R}^{\text{out}} > \mu$ for $R \in \mathcal{R}_2$ the maximal such J which we call J_R will contain K_R and $B_{R,3J_R}^{\text{out}} \leq \mu$.

By Lemma 7.3 we have for each $a \in 3J_R$ and $Q \in \mathcal{B}_{R,J_R}^{\text{out}}$

$$|\hat{Q} \cap (\hat{I}_R \times \{a\})| \lesssim \frac{1}{|3J_R|} \int_{\hat{I}_R \times 3J_R} \mathbf{1}_{\hat{Q}}.$$

We can then calculate

$$\begin{aligned} \sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}}} a_Q \frac{|\hat{Q} \cap (\hat{I}_R \times \{a\})|}{|\hat{Q}| |\hat{I}_R \times \{a\}|} &\lesssim \sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}}} a_Q \frac{|\hat{Q} \cap (\hat{I}_R \times 3J_R)|}{|\hat{Q}| |\hat{I}_R \times 3J_R|} \\ &\lesssim B_{R,3J_R}^{\text{out}} + \int_{(\hat{I}_R \times 3J_R)} \sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}} \setminus \mathcal{B}_{R,3J_R}^{\text{out}}} a_Q \frac{\mathbf{1}_{\hat{Q}}}{|\hat{Q}|} \\ &\leq \mu + \sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}} \setminus \mathcal{B}_{R,3J_R}^{\text{out}}} a_Q \frac{|\hat{Q} \cap (\hat{I}_R \times 3J_R)|}{|\hat{Q}| |\hat{I}_R \times 3J_R|}, \end{aligned}$$

in passing to the last line we used the maximality of J_R . Now all the plates $\hat{Q}$ appearing in the sum of the right hand side in the estimate above are contained in $\hat{I}_R \times 9J_R$ and so the sum of the second summand above is estimated by a dimensional constant $c_n > 1$ so that

$$\sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}}} a_Q \frac{|\hat{Q} \cap (\hat{I}_R \times \{a\})|}{|\hat{Q}| |\hat{I}_R \times \{a\}|} \lesssim_n \mu + c_n \lesssim_n \mu$$

if μ is sufficiently large depending only upon dimension. Since $J_R \supseteq K_R$ we can integrate for $a \in K_R = \hat{K}_R$ to conclude that

$$\sum_{Q \in \mathcal{B}_{R,J_R}^{\text{out}}} a_Q \frac{|\hat{Q} \cap \hat{R}|}{|\hat{Q}| |\hat{R}|} \leq \kappa_n \mu$$

for some dimensional constant $\kappa_n > 1$. This shows that for μ sufficiently large depending upon dimension we have

$$B_R^{\mathcal{L}} \leq \kappa_n \mu + \sum_{Q \in B_{R,J_R}^{\text{in}}} a_Q \frac{|\widehat{Q} \cap \widehat{R}|}{|\widehat{Q}||\widehat{R}|} = \kappa_n \mu + B_R^{\mathcal{L}'_2}, \quad \mathcal{L}'_2 := \bigcup_{\rho \in \mathcal{R}_2} \mathcal{B}_{\rho,J_\rho}^{\text{in}},$$

with

$$\sum_{Q \in \mathcal{L}'_2} a_Q \leq \left| \bigcup_{\rho \in \mathcal{R}_2} \widehat{I}_\rho \times 3J_\rho \right|. \quad (7.4)$$

By the previous estimates we have that for each $\rho \in \mathcal{R}_2$

$$\begin{aligned} \mu < B_{\rho,J_\rho}^{\text{out}} &= \sum_{Q \in \mathcal{B}_{\rho,J_\rho}^{\text{out}}} a_Q \frac{|\widehat{Q} \cap (\widehat{I}_\rho \times J_\rho)|}{|\widehat{Q}||\widehat{I}_\rho \times J_\rho|} = \int_{J_\rho} \sum_{Q \in \mathcal{B}_{\rho,J_\rho}^{\text{out}}} a_Q \frac{|\widehat{Q} \cap (\widehat{I}_\rho \times \{a\})|}{|\widehat{Q}||\widehat{I}_\rho \times \{a\}|} da \\ &\leq \frac{\kappa_n \mu}{|J_\rho|} |\{a \in J_\rho : \psi_\rho(a) > \mu/2\}| + \mu/2 \end{aligned}$$

with

$$\psi_\rho(a) := \sum_{Q \in \mathcal{B}_{\rho,J_\rho}^{\text{out}}} a_Q \frac{|\widehat{Q} \cap (\widehat{I}_\rho \times \{a\})|}{|\widehat{Q}||\widehat{I}_\rho \times \{a\}|}.$$

Thus there exists a set $J'_\rho \subseteq J_\rho$ with $|J'_\rho| \gtrsim_n |J_\rho|$ so that $\psi_\rho(a) > \mu/2$ for $a \in J'_\rho$. Now note that for $a \in J'_\rho$

$$\frac{\mu}{2} \leq \psi_\rho(a) \leq - \int_{\widehat{I}_\rho \times \{a\}} \sum_{Q \in \mathcal{B}_{\rho,K_\rho}^{\text{out}}} a_Q \frac{\mathbf{1}_{\widehat{Q}}}{|\widehat{Q}|} \leq \inf_{x \in \widehat{I}_\rho} M_{v^\perp} \left(\sum_{\substack{Q \in \mathcal{L} \\ Q \leq \rho}} a_Q \frac{\mathbf{1}_{\widehat{Q}}}{|\widehat{Q}|} \right)(x, a).$$

In the estimate above we write $M_{v^\perp}$ for the maximal function

$$M_{v^\perp} f(x) := \sup_{s>0} \int_{Q_{v^\perp}(0,s)} |f(x+t)| dt, \quad x \in \mathbb{R}^n,$$

where $Q_{v^\perp}(0,s)$ denotes the cube in $v^\perp \simeq \mathbb{R}^d$ with sidelength $s > 0$ and centered at $0 \in v^\perp$. Note that for $v \in \mathbb{S}^d$ the operator $M_{v^\perp}$ is of weak-type $(1,1)$, uniformly in v . Thus

$$\bigcup_{\rho \in \mathcal{R}_2} (\widehat{I}_\rho \times J'_\rho) \subseteq S := \left\{ z \in \mathbb{R}^n : M_{v^\perp} \left(\sum_{Q \in \mathcal{R}} a_Q \frac{\mathbf{1}_{c_2 Q}}{|c_2 Q|} \right)(z) > \mu/2 \right\}$$

On the other hand since $J'_\rho \subseteq J_R$ with $|J'_\rho| \gtrsim |J_\rho|$ we readily see that

$$\begin{aligned} \bigcup_{\rho \in \mathcal{R}_2} (\widehat{I}_\rho \times J_\rho) &\subseteq \left\{ z \in \mathbb{R}^n : M_{e_n} (\mathbf{1}_{\bigcup_{\rho \in \mathcal{R}_2} \widehat{I}_\rho \times J'_\rho})(z) \gtrsim_n 1 \right\} \\ &\subseteq \{ z \in \mathbb{R}^n : M_{e_n} (\mathbf{1}_S)(z) \gtrsim_n 1 \}. \end{aligned}$$

Combining (7.4) with the weak type $(1, 1)$ inequalities for M_{e_n} and $M_{v^\perp} = M_{e_n^\perp}$ (since we assume that $v = e_n$) and choosing μ to be a sufficiently large dimensional constant

$$\text{mass}_{a,1}(\mathcal{L}'_2) = \sum_{Q \in \mathcal{L}'_2} a_Q \leq \frac{C}{\lambda} \sum_{Q \in \mathcal{L}} \leq \frac{1}{4} \sum_{Q \in \mathcal{L}} a_Q = \text{mass}_{a,1}(\mathcal{L})/4.$$

Now we set $\mathcal{L}' := \mathcal{L}'_1 \cup \mathcal{L}'_2$ and the proof is complete. $\square$

7.6. Application to a maximal function estimate

As an immediate application of the directional Carleson embedding theorem for plates we describe below a sharp theorem for maximal averages along codimension 1 plates. Let $d = n - 1$ and consider $\sigma \in \text{Gr}(d, n) = \text{Gr}(d, d + 1)$. We remember that the codimension 1 averages at scale $s > 0$ of a function $f \in \mathcal{S}(\mathbb{R}^{d+1})$ can be given in the form

$$\langle f \rangle_{s,\sigma}(x) := \int_{B_{d+1}(s) \cap \sigma} f(x - y) \frac{dy}{s^d}, \quad x \in \mathbb{R}^{d+1},$$

where $B_{d+1}(s)$ denotes the ball of radius s and centered at $0 \in \mathbb{R}^{d+1}$. Given a finite subset $\Sigma \subset \text{Gr}(d, d + 1)$ we are interested in the corresponding maximal averaging operator along codimension 1 plates given by Σ

$$M_\Sigma f(x) := \sup_{s > 0} \sup_{\sigma \in \Sigma} \langle |f| \rangle_{\sigma,s}(x), \quad x \in \mathbb{R}^{d+1}.$$

As a consequence of the directional Carleson embedding theorem we obtain the sharp bounds for M_Σ for arbitrary finite $\Sigma \subset \text{Gr}(d, d + 1)$.

Proof of Theorem B. We write $n = d + 1$ throughout the proof. It suffices to prove the weak-type $(2, 2)$ estimate. Indeed the L^p -estimate will then follow by interpolation between the $L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n) \rightarrow L^\infty(\mathbb{R}^n)$ bounds. Furthermore the strong-type $(2, 2)$ estimate follows by the corresponding weak-type estimate with an additional $\sqrt{\log \#\Sigma}$ -loss by the well known interpolation argument of Strömberg, [29] as in the proof of Theorem C.

For the $L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)$ note that it suffices to prove the weak-type $(2, 2)$ estimate with the same dependence on $\#\Sigma$ for the closely related dyadic maximal operator

$$M_{V,\delta} f(x) := \sup_{\substack{Q \in \mathcal{D}_{V,\delta} \\ Q \ni x}} \int_Q |f|, \quad x \in \mathbb{R}^n,$$

for $V \subset \mathbb{S}^d$ finite and fixed $\delta > 0$, and with a bound independent of δ . This operator can be linearized as in the proof of Theorem C in the form

$$T_{\mathcal{Q}}f(x) := \sum_{Q \in \mathcal{Q}} \left(\int_Q |f| \right) \mathbf{1}_{F_Q}(x), \quad x \in \mathbb{R}^n$$

where $\mathcal{Q} \subset \mathcal{D}_{V,\delta}$ a finite collection of (d, δ) plates and $\{F_Q\}_{Q \in \mathcal{Q}}$ a pairwise disjoint collection of measurable sets with $F_Q \subseteq Q$ for every $Q \in \mathcal{Q}$. Denoting by $T_{\mathcal{Q}}^*$ the adjoint of $T_{\mathcal{Q}}$ we have that

$$\|M_{V,\delta}\|_{L^2(\mathbb{R}^n) \rightarrow L^{2,\infty}(\mathbb{R}^n)} = \sup_{\mathcal{Q} \subset \mathcal{D}_{V,\delta}} \sup_{0 < |E| < \infty} \frac{\|T_{\mathcal{Q}}^*(\mathbf{1}_E)\|_{L^2(\mathbb{R}^n)}}{|E|^{\frac{1}{2}}}$$

where

$$T_{\mathcal{Q}}^*(\mathbf{1}_E)(x) = \sum_{Q \in \mathcal{Q}} |F_Q \cap E| \frac{\mathbf{1}_Q}{|Q|}.$$

Clearly $a = \{a_Q\}_{Q \in \mathcal{Q}} = \{|F_Q \cap E|\}_{Q \in \mathcal{Q}}$ is a Carleson sequence in the sense of Definition 6 so the required estimate for $T_{\mathcal{Q}}^*(\mathbf{1}_E)$ follows by a straightforward application of Theorem 1.

The fact that these estimates are best possible follows by considering A to be a Kakeya collection of $\delta \times 1$ -tubes in $\mathbb{R}^2$ and taking $A' := A \times [-1, 1]^{n-2}$. Now for each tube in A we can consider a $1 \times \delta^{n-2}$ plate that contains the tube and is perpendicular to the copy of $\mathbb{R}^2$ that contains A . Calculating the averages of $\mathbf{1}_{A'}$ with respect to these plates yields the sharpness of the weak-type $(2, 2)$ estimate and the sharpness of the strong (p, p) estimate for $p > 2$. Note that the numerology here is $\#\Sigma = \#V \simeq 1/\delta$. The optimality of the strong $(2, 2)$ -estimate follows similarly by considering a function in $\mathbb{R}^2$ that yields the sharpness of the 2-dimensional results and extending them in $\mathbb{R}^n$ by taking a tensor product with a smooth bump in $\mathbb{R}^{n-2}$; see also Remark 3.1. $\square$

Remark 7.7. We stress an important switch in our point of view when proving estimates for the maximal operator $M_{V,\delta}$ above, compared to say the corresponding estimates for the Nikodym operator $\mathcal{N}_{\delta}$ in §6. Indeed although these two operators appear to be quite similar, in the case of $M_{V,\delta}$ we are interested in proving estimates for *arbitrary finite subsets* $V \subset \mathbb{S}^d \approx \text{Gr}(d, d+1)$. Thus our δ -fattening of the thin plates $B_n(s) \cap \sigma$ for $\sigma \in \text{Gr}(d, d+1)$ is purely qualitative, it is there just to allow us to use δ -plates which have positive measure in $\mathbb{R}^{d+1}$ and are more amenable to geometric arguments. These estimates are to be δ -free as we use a limiting argument in order to recover thin plates as $\delta \rightarrow 0^+$. Necessarily, for this argument the cardinality $\#V$ and δ are completely independent of each other. This is in contrast to the geometric setup underlying the definition of the Nikodym operator $\mathcal{N}_{\delta}$ where the implicit subset of $\text{Gr}(d, d+1)$ is a δ -net and thus has cardinality $\sim \delta^{-d}$, namely the thickness of the plates and the set of essentially directions present in $\mathcal{N}_{\delta}$ are intimately connected.

7.8. Application to a conical frequency square function

We describe below a square function estimate in the spirit of Rubio de Francia, given with respect to conical frequency projections in $\mathbb{R}^{d+1}$. Let $\{B_k\}_{k=1}^N$ be a collection of open balls in $\mathbb{R}^{d+1}$ whose centers v_k lie on $\mathbb{S}^d$, and is *well-distributed* in the sense that

$$\sum_{k=1}^N \mathbf{1}_{3B_k} \lesssim 1$$

Here, as usual, $3B_k$ is a threefold dilation of the Euclidean ball B_k with respect to its center. For each k , let

$$\phi_k \in \mathcal{S}(\mathbb{R}^{d+1}), \quad \mathbf{1}_{B_k} \leq \phi_k \leq \mathbf{1}_{3B_k}$$

and define the conical frequency projection

$$S_k f(x) := \int_{\mathbb{S}^d} \int_0^\infty \hat{f}(r\xi') \phi_k(\xi') e^{ix \cdot r\xi'} r^d dr d\sigma_d(\xi'), \quad x \in \mathbb{R}^{d+1}.$$

The d -plate Carleson embedding Theorem I may be used to deduce a square function estimate for the projections S_k with sharp dependence on the parameter N .

Theorem J. For $2 \leq p < 4$,

$$\left\| \left(\sum_k |S_k f|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^{d+1})} \lesssim_{p,d} (\log N)^{\frac{1}{2} - \frac{1}{p}} \|f\|_{L^p(\mathbb{R}^{d+1})}.$$

Furthermore, the restricted $L^4(\mathbb{R}^n)$ -type estimate

$$\left\| \left(\sum_k |S_k(f\mathbf{1}_E)|^2 \right)^{\frac{1}{2}} \right\|_{L^4(\mathbb{R}^{d+1})} \lesssim_d (\log N)^{\frac{1}{4}} \|f\|_{L^\infty(\mathbb{R}^{d+1})} |E|^{\frac{1}{4}}$$

holds for all bounded measurable sets $E \subset \mathbb{R}^{d+1}$. These bounds are best possible up to the implicit numerical constants.

For the optimality in the estimate of Theorem J we send to [1, Section 8], noting that the two-dimension bound becomes a lower bound in $\mathbb{R}^{d+1}$ for general d by taking balls $\{B_k\}_{k=1}^N$ having centers lying on a copy of $\mathbb{S}^1$ and functions f which are suitable tensor products.

The remainder of this section contains the proof of the upper bounds in Theorem J. Below, ℓ_N^2 stands for the Euclidean norm on $\mathbb{C}^N$. The first step consists of the radial decoupling

$$\|S_k f\|_{L^p(\mathbb{R}^{d+1}; \ell_N^2)} \lesssim_{p,d} \|S_k Q_m f\|_{L^p(\mathbb{R}^{d+1}; \ell_N^2 \otimes \ell_{\mathbb{Z}}^2)} \quad (7.5)$$

where $p \in [2, \infty)$ and $\{Q_m : m \in \mathbb{Z}\}$ are Fourier multiplier operators whose associated multipliers are radial and a partition of unity of $\mathbb{R}^{d+1} \setminus \{0\}$ subordinated to the finitely overlapping cover $A_m = \{\xi \in \mathbb{R}^{d+1} : 2^{-m-1} < |\xi| < 2^{-m+1}\}, m \in \mathbb{Z}$. The proof is a simple application of the weighted norm inequality

$$\|g\|_{L^2(w)} \lesssim \|Q_m g\|_{L^2(\widetilde{M}w; \ell_{\mathbb{Z}}^2)}$$

where $\widetilde{M}$ stands for the third iterate of the standard $(d+1)$ -dimensional Hardy-Littlewood maximal operator see [1, Lemma 5.6] for details.

Note that $S_k Q_m$ is supported in the frequency tube

$$\omega_{k,m} := \left\{ \xi \in A_m : \frac{\xi}{|\xi|} \in 3B_k \right\}$$

whose center line is through the center v_k of B_k and whose spatial dual is the *plate*

$$R_{m,k}^0 := \{x \in \mathbb{R}^{d+1} : |\Pi_{v_k} x| < 2^m \delta_k, \Pi_{v_k^\perp} x < 2^m\}$$

of *eccentricity* δ_k , the radius of B_k and sidelength 2^m . Let $\mathcal{R}_{m,k}$ be a $\lesssim_d$ 1-overlapping cover of $\mathbb{R}^{d+1}$ by translates R of $R_{m,k}^0$ and $t \in \mathbf{T}_{m,k}$ be the collection of *tiles* $t = (R_t, \omega_t)$ with $\omega_t = \omega_{m,k}$ and $R \in \mathcal{R}_{m,k}$. The space-frequency projection on t is represented by the intrinsic coefficient

$$a_t(f) = \sup_{\phi \in \Phi_t^M} |\langle f, \phi \rangle|^2$$

where $\Phi_t^M \subset \mathcal{S}(\mathbb{R}^{d+1})$ is the class of functions whose frequency support is contained in ω_t and are uniformly spatially adapted to R_t in the sense that

$$\sqrt{|R_t|} |\phi| \leq \mathbf{1}_{R_t} + \sum_{k=0}^{\infty} 2^{-Mk} \mathbf{1}_{2^{k+1}R_t \setminus 2^k R_t}$$

uniformly over $\phi \in \Phi_t^M$. A standard space-frequency discretization, see e.g. [1, Sect. 5] the right-hand side of (7.5) is pointwise bounded by the discretized square function associated to the coefficients a_t , namely

$$\|S_k Q_m f\|_{\ell_N^2 \otimes \ell_{\mathbb{Z}}^2} \lesssim \Delta f := \left(\sum_{t \in \mathbf{T}} a_t(f) \frac{\mathbf{1}_{R_t}}{|R_t|} \right)^{\frac{1}{2}}$$

where $\mathbf{T}_k = \bigcup_{m \in \mathbb{Z}} \mathbf{T}_{m,k}$ and $\mathbf{T} = \bigcup_{1 \leq k \leq N} \mathbf{T}_k$. Theorem J is thus reduced to the corresponding bounds for the discretized square function Δf . In fact, by standard

restricted-type interpolation, it suffices to prove the restricted type estimate that follows.

Proposition 7.9. $\left\| \Delta(f\mathbf{1}_E)^2 \right\|_{L^2(\mathbb{R}^{d+1})} \lesssim_d (\log N)^{\frac{1}{2}} \|f\|_{L^\infty(\mathbb{R}^{d+1})} |E|^{\frac{1}{2}}.$

Proof. By a limiting argument, it suffices to replace the universe of tiles $\mathbf{T}$ in the definition of Δ by a finite subcollection, which we still call $\mathbf{T}$. We still denote by $\mathbf{T}_{m,k}, \mathbf{T}_k$ the subcollections of $\mathbf{T}$ with sidelength and frequency parameters m, k . Recall that $n = d + 1$. By linearity, we may restrict to the case $\|f\|_\infty = 1$. By finite splitting of the cones, we can assume all v_k lie in a small neighborhood of e_n as specified in Theorem I. Using the well known 3^d -grid lemma and finite splitting of $\mathbf{T}$, we may find a dyadic grid $\mathcal{D}$ on $e_n^\perp$ such that for all $t \in \mathbf{T}_{m,k}$, $\Pi_{e_n^\perp} R_t \subset I_t$ for some $I_t \in \mathcal{D}$ with $\ell_{I_t} = 2^{m+3}$. Let also $\mathcal{K}^j$, $j = 0, 1, 2$ be a system of three shifted dyadic grids on $\mathbb{R}$ and define

$$\mathcal{Q}_{m,k} := \{P(I, K, v_k) : I \in \mathcal{D}, \ell_{I_t} = 2^{m+3}, K \in \mathcal{K}^0 \cup \mathcal{K}^1 \cup \mathcal{K}^2, \ell_K \in [2^{m+3}\delta_k, 2^{m+4}\delta_k]\},$$

$$\mathcal{Q}_k := \bigcup_{m \in \mathbb{Z}} \mathcal{Q}_{m,k}, \quad \mathcal{Q} := \bigcup_{1 \leq k \leq N} \mathcal{Q}_k.$$

Notice that for each $t \in \mathbf{T}_{m,k}$ there exists at least 1 and at most 3 elements $Q \in \mathcal{Q}_{m,k}$ with $R_t \subset Q$ and $|Q| \lesssim |R_t|$. Thus, setting

$$\mathbf{T}(Q) := \{t \in \mathbf{T}_{m,k} : R_t \subset Q\}, \quad a_Q := \sum_{t \in \mathbf{T}(Q)} a_t(f\mathbf{1}_E), \quad Q \in \mathcal{Q}_{m,k}$$

leads to the pointwise estimate $\Delta(f\mathbf{1}_E)^2 \lesssim T_Q(a)$, with $T_Q(a)$ in the form (7.1).

Proposition 7.9 may then be obtained by an application of Theorem I to the union $\mathcal{Q}$ of the N collections $\mathcal{Q}_k$. Notice that the plates of $\mathcal{Q}_k$ have fixed eccentricity δ_k and thus obey the weak (1,1) assumption of that theorem. We must then compute the directional Carleson norm of the sequence $\{a_Q : Q \in \mathcal{Q}\}$. Firstly, from the finite frequency overlap of the Fourier supports of any collection $\{\phi_t \in \Phi_t^M : t \in \mathbf{T}\}$ and the spatial localization of the collection $\{\phi_t \in \Phi_t : t \in \mathbf{T} : \omega_t = \omega_{m,k}\}$ to a finitely overlapping collection $\mathcal{R}_{m,k}$, we gather that

$$\sum_{t \in \mathbf{T}} |\langle g, \phi_t \rangle|^2 \lesssim \|g\|_2^2 \quad (7.6)$$

whence

$$\text{mass}_a(\mathcal{Q}) \leq 2 \sum_{t \in \mathbf{T}} a_t(f\mathbf{1}_E) \lesssim |E|.$$

The details of this estimate are similar to those of [1, Lemma 4.4].

We turn to the verification of the directional Carleson sequence property. That is for each fixed k , let $\mathcal{T} \subset \mathcal{P}_{v_k}$, we need to prove that

$$\sum_{L \in \mathcal{L}} a_L \lesssim |\text{sh}(\mathcal{T})|$$

whenever $\mathcal{L}$ is a collection subordinated to $\mathcal{T}$. If M stands locally for the maximal averaging operator over all plates $\mathcal{P}_{v_k}$, which is of type $(2, 2)$ say, define the enlargement of $\text{sh}(\mathcal{T})$

$$U := \{M\mathbf{1}_{\text{sh}(\mathcal{T})} > 2^{-10}\}$$

so that $|U| \lesssim |\text{sh}(\mathcal{T})|$. Then

$$\sum_{L \in \mathcal{L}} a_L \leq \sum_{L \in \mathcal{L}} \sum_{t \in \mathbf{T}(Q)} a_t(f\mathbf{1}_{E \cap U}) + \sum_{L \in \mathcal{L}} \sum_{t \in \mathbf{T}(L)} a_t(f\mathbf{1}_{E \cap U^c}).$$

The local part $f\mathbf{1}_{E \cap U}$ is then dealt with using (7.6) as follows:

$$\sum_{t \in \mathbf{T}(L)} a_t(f\mathbf{1}_{E \cap U}) \lesssim \|f\mathbf{1}_{E \cap U}\|_2^2 \lesssim |U| \lesssim |\text{sh}(\mathcal{T})|.$$

To estimate the nonlocal part split $T \in \mathcal{T}$ into the union of collections $\mathbf{T}(u)$, saying $T \in \mathbf{T}(u)$ if u is the least integer such that $2^{u+1}T \cap U^c \neq \emptyset$. A suitable version of Journé's lemma [1, Lemma 4.7] yields

$$\sum_{T \in \mathbf{T}(u)} |T| \lesssim 2^u |\text{sh}(\mathcal{T})|$$

therefore the estimate for the non-local part follows from bounding uniformly in $T \in \mathcal{T}(u)$

$$\sum_{\substack{L \in \mathcal{L} \\ L \subset T}} \sum_{t \in \mathbf{T}(L)} a_t(f\mathbf{1}_{E \cap U^c}) \lesssim 2^{-10du} |T|. \quad (7.7)$$

To prove the latter estimate, write

$$\chi(x) := \left(1 + \frac{|\Pi_{v_k^\perp} x|^2}{\ell_{I_T}} + \frac{|\Pi_{v_k} x|^2}{\ell_{K_T}}\right)^{-100d}$$

for the rapidly decaying function adapted to the plate T . Note that $\phi_t \in \Phi_t^M$ with $t \in \mathbf{T}(L)$ is adapted to R_t and thus to the slight enlargement L . When $L \subset T$, $\tilde{\phi}_t := c\phi_t\chi^{-1} \in \Phi_t^{M/2}$ if $M > 2^{200d}$ and $c > 0$ is suitably chosen; in particular we have used that the frequency support of $\tilde{\phi}_t$ is the same as that of ϕ_t as χ is the inverse of a polynomial.

Let $h = f\mathbf{1}_{E \cap U^c}$ and $T^r = 2^{u+r+1}T \setminus 2^{u+r}T$. As $2^u T \cap U^c = \emptyset$, we may write $h = \sum_{r=0}^\infty h\mathbf{1}_{T^r}$. Then for suitable choice of $\phi_t \in \Phi_t^M$ and applying (7.6) to pass to the second line

$$\begin{aligned}
\sum_{\substack{L \in \mathcal{L} \\ L \subset T}} \sum_{t \in \mathbf{T}(L)} a_t(h\mathbf{1}_{T^r}) &\leq \sum_{\substack{L \in \mathcal{L} \\ L \subset T}} \sum_{t \in \mathbf{T}(L)} |\langle h\mathbf{1}_{T^r}, \phi_t \rangle|^2 \lesssim \sum_{\substack{L \in \mathcal{L} \\ L \subset T}} \sum_{t \in \mathbf{T}(L)} |\langle h\mathbf{1}_{T^r} \chi, \tilde{\phi}_t \rangle|^2 \\
&\lesssim \|h\mathbf{1}_{T^r} \chi\|_2^2 \leq \|\chi\mathbf{1}_{T^r}\|_\infty^2 |T^r| \\
&\lesssim 2^{-100d(u+r)} \times (2^{rd}|T|) \lesssim 2^{-99d(u+r)} |T|
\end{aligned}$$

Summing up over $r \geq 0$ yields the bound (7.7), and completes the proof of the Proposition. $\square$

References

- [1] N. Accomazzo, F.D. Plinio, P. Hagelstein, I. Parissis, L. Roncal, Directional square functions, *Anal. PDE* (2022), in press, available at [arXiv:2004.06509](https://arxiv.org/abs/2004.06509).
- [2] A. Alfonseca, F. Soria, A. Vargas, A remark on maximal operators along directions in $\mathbb{R}^2$, *Math. Res. Lett.* 10 (1) (2003) 41–49. MR1960122.
- [3] S. Barone, S. Basu, Refined bounds on the number of connected components of sign conditions on a variety, *Discrete Comput. Geom.* 47 (3) (2012) 577–597. MR2891249.
- [4] M. Bateman, Maximal averages along a planar vector field depending on one variable, *Trans. Am. Math. Soc.* 365 (8) (2013) 4063–4079. MR3055689.
- [5] J. Bourgain, Besicovitch type maximal operators and applications to Fourier analysis, *Geom. Funct. Anal.* 1 (2) (1991) 147–187. MR1097257.
- [6] A. Córdoba, The multiplier problem for the polygon, *Ann. Math.* (2) 105 (3) (1977) 581–588. MR0438022.
- [7] A. Córdoba, R. Fefferman, On differentiation of integrals, *Proc. Natl. Acad. Sci. USA* 74 (6) (1977) 2211–2213. MR0476977.
- [8] A. Córdoba, The Keakey maximal function and the spherical summation multipliers, *Am. J. Math.* 99 (1) (1977) 1–22. MR0447949.
- [9] C. Demeter, Singular integrals along N directions in $\mathbb{R}^2$, *Proc. Am. Math. Soc.* 138 (12) (2010) 4433–4442. MR2680067.
- [10] C. Demeter, L^2 bounds for a Keakey-type maximal operator in $\mathbb{R}^3$, *Bull. Lond. Math. Soc.* 44 (4) (2012) 716–728. MR2967239.
- [11] F. Di Plinio, S. Guo, C. Thiele, P. Zorin-Kranich, Square functions for bi-Lipschitz maps and directional operators, *J. Funct. Anal.* 275 (8) (2018) 2015–2058. MR3841536.
- [12] F. Di Plinio, I. Parissis, Maximal directional operators along algebraic varieties, *Am. J. Math.* 143 (5) (2021) 1463–1503. MR4334401.
- [13] K.J. Falconer, Continuity properties of k -plane integrals and Besicovitch sets, *Math. Proc. Camb. Philos. Soc.* 87 (2) (1980) 221–226. MR553579.
- [14] K.J. Falconer, Sets with prescribed projections and Nikodým sets, *Proc. Lond. Math. Soc.* (3) 53 (1) (1986) 48–64. MR842156.
- [15] C. Fefferman, The multiplier problem for the ball, *Ann. Math.* (2) 94 (1971) 330–336. MR0296602.
- [16] J. Hickman, K.M. Rogers, R. Zhang, Improved bounds for the Keakey maximal conjecture in higher dimensions, *Am. J. Math.* (2022), in press, available at [arXiv:1908.05589](https://arxiv.org/abs/1908.05589).
- [17] N.H. Katz, Maximal operators over arbitrary sets of directions, *Duke Math. J.* 97 (1) (1999) 67–79. MR1681088.
- [18] N.H. Katz, Remarks on maximal operators over arbitrary sets of directions, *Bull. Lond. Math. Soc.* 31 (6) (1999) 700–710. MR1711029.
- [19] J. Kim, Almost-orthogonality principles for certain directional maximal functions, *J. Geom. Anal.* 31 (7) (2021) 7320–7332. MR4289260.
- [20] P. Mattila, *Fourier Analysis and Hausdorff Dimension*, Cambridge Studies in Advanced Mathematics, vol. 150, Cambridge University Press, Cambridge, 2015. MR3617376.
- [21] A. Nagel, E.M. Stein, S. Wainger, Differentiation in lacunary directions, *Proc. Natl. Acad. Sci. USA* 75 (3) (1978) 1060–1062. MR0466470.
- [22] R. Oberlin, Bounds for Keakey-type maximal operators associated with k -planes, *Math. Res. Lett.* 14 (1) (2007) 87–97. MR2289622.
- [23] R. Oberlin, The (d,k) Keakey problem and estimates for the X-ray transform, ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)—The University of Wisconsin - Madison, 2007.

- [24] J. Parcet, K.M. Rogers, Directional maximal operators and lacunarity in higher dimensions, *Am. J. Math.* 137 (6) (2015) 1535–1557. MR3432267.
- [25] K.M. Rogers, On a planar variant of the Kakeya problem, *Math. Res. Lett.* 13 (2–3) (2006) 199–213. MR2231112.
- [26] A. Sheffer, *Polynomial Methods and Incidence Theory*, Cambridge Studies in Advanced Mathematics, vol. 197, Cambridge University Press, Cambridge, 2022. MR4394303.
- [27] P. Sjögren, P. Sjölin, Littlewood-Paley decompositions and Fourier multipliers with singularities on certain sets, *Ann. Inst. Fourier (Grenoble)* 31 (1) (1981) 157–175, vii. MR613033.
- [28] J.-O. Strömberg, Weak estimates on maximal functions with rectangles in certain directions, *Ark. Mat.* 15 (2) (1977) 229–240. MR0487260.
- [29] J.-O. Strömberg, Maximal functions associated to rectangles with uniformly distributed directions, *Ann. Math. (2)* 107 (2) (1978) 399–402. MR0481883.
- [30] T. Tao, The Bochner-Riesz conjecture implies the restriction conjecture, *Duke Math. J.* 96 (2) (1999) 363–375. MR1666558.