



PROJECT MUSE®

Maximal directional operators along algebraic varieties

Francesco Di Plinio, Ioannis Parissis

American Journal of Mathematics, Volume 143, Number 5, October 2021,
pp. 1463-1503 (Article)

Published by Johns Hopkins University Press

DOI: <https://doi.org/10.1353/ajm.2021.0037>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/813903/summary>

MAXIMAL DIRECTIONAL OPERATORS ALONG ALGEBRAIC VARIETIES

By FRANCESCO DI PLINIO and IOANNIS PARISSIS

Abstract. We establish the sharp growth order, up to epsilon losses, of the L^2 -norm of the maximal directional averaging operator along a finite subset V of a polynomial variety of arbitrary dimension m , in terms of cardinality. This is an extension of the works by Córdoba, for one-dimensional manifolds, Katz for the circle in two dimensions, and Demeter for the 2-sphere. For the case of directions on the two-dimensional sphere we improve by a factor of $\sqrt{\log N}$ on the best known bound, due to Demeter, and we obtain a sharp estimate for our model operator. Our results imply new L^2 -estimates for Keakeya type maximal functions with tubes pointing along polynomial directions. Our proof technique is novel and in particular incorporates an iterated scheme of polynomial partitioning on varieties adapted to directional operators, in the vein of Guth, Guth-Katz, and Zahl.

1. Main results, motivation, background and techniques. We are interested in maximal directional averaging operators, defined with respect to a given set of directions. Our focus is on the higher dimensional setting where the directions are distributed on algebraic varieties in $\mathbb{R}^n$, of any given codimension.

1.1. Main results. More precisely, let $n \geq 2$ and define the directional averages of a smooth function f on $\mathbb{R}^n$ by

$$\langle f \rangle_v(x) := \int_{-1}^1 f(x - tv) dt, \quad x, v \in \mathbb{R}^n,$$

and consider the maximal averaging operator, at unit scale, associated to a set of directions $V \subset \mathbb{R}^n$

$$(1.1) \quad M_V f(x) := \sup_{v \in V} \langle |f| \rangle_v(x).$$

It is customary to study the *single scale operator* (1.1) for sets of directions V which are normalized to live in a fixed but arbitrary annular region excluding the origin. To state our results, we will use the unit annulus

$$\mathcal{A}_n(1) := \{v \in \mathbb{R}^n : 1 \leq |v| < 2\}.$$

Manuscript received October 3, 2018.

Research of the first author supported in part by NSF grants DMS-1650810, DMS-1800628, and DMS-2000510; research of the second author supported in part by the project PGC2018-094528-B-I00 (AEI/FEDER, UE) with acronym “IHAIP”, grant T1247-19 of the Basque Government and IKERBASQUE.

American Journal of Mathematics 143 (2021), 1463–1503. © 2021 by Johns Hopkins University Press.

The first main result of this article is the sharp bound in terms of the cardinality parameter N , up to arbitrarily small losses, for the maximal operator norm

$$\sup \left\{ \|M_V\|_{L^2(\mathbb{R}^n)} : V \subset Z_m \cap \mathcal{A}_n(1), \#V \leq N^m \right\}$$

when $Z_m \subseteq \mathbb{R}^n$ is a real algebraic variety of a fixed dimension $1 \leq m \leq n-1$. We send to Section 4 for the precise definition of a real algebraic variety of dimension m ; here, we restrict ourselves to mentioning a prototypical example. If $D \geq 1$, we say that Z_m belongs to the class $\mathcal{Z}_{m,n}^\times(D)$ if

$$Z_m = \{x \in \mathbb{R}^n : P_1(x) = \cdots = P_{n-m}(x) = 0\}$$

where $P_1, \dots, P_{n-m}$ are polynomials in n real variables of degree at most D , and the tangent space to Z_m is m -dimensional at all points $x \in Z_m$, in the sense of (2.10) below. For the class $\mathcal{Z}_{m,n}^\times(D)$ we are able to obtain a uniform bound on the operator norm. In particular, polynomial graphs over m variables in $\mathbb{R}^n$, of degree at most D ,

$$Z_m = \{(y, h(y)) : y \in \mathbb{R}^m\}, \quad h: \mathbb{R}^m \rightarrow \mathbb{R}^{n-m} \text{ polynomial of degree at most } D,$$

belong to the class $\mathcal{Z}_{m,n}^\times(D)$.

THEOREM A. *Let $n \geq 2$ and $1 \leq m \leq n-1$. Let $Z_m \subseteq \mathbb{R}^n$ be a real algebraic variety of dimension m . Then for all $\eta > 0$ there is a constant $\Theta = \Theta(Z_m, \eta)$ such that*

$$(1.2) \quad \sup_{\substack{V \subset Z_m \cap \mathcal{A}_n(1) \\ \#V \leq N^m}} \|M_V\|_{L^2(\mathbb{R}^n)} \leq \Theta N^{\frac{m-1}{2} + \eta}.$$

The constant $\Theta = \Theta(Z_m, \eta)$ depends on η and on explicit algebraic properties of the variety Z_m .

Furthermore, if $D \geq 1$ and $\mathcal{Z}_{m,n}^\times(D)$ is the class of real algebraic varieties defined above, we have the uniform bound

$$(1.3) \quad \sup_{Z_m \in \mathcal{Z}_{m,n}^\times(D)} \sup_{\substack{V \subset Z_m \cap \mathcal{A}_n(1) \\ \#V \leq N^m}} \|M_V\|_{L^2(\mathbb{R}^n)} \leq \Theta N^{\frac{m-1}{2} + \eta}$$

for all $\eta > 0$; the constant $\Theta = \Theta(m, n, D, \eta)$ depends only on the dimension parameters m, n , on the degree D , and on $\eta > 0$.

As we will see in the subsequent section, the constant $\Theta(Z_m, \eta)$ that appears in the statement of Theorem A depends on certain notions of *degree* and *count* of the variety Z_m . We refer the reader to §4 for precise definitions and further discussion.

Let $M_{V,r}$ be the analogue of M_V of (1.1) obtained from the directional averages at scale $r > 0$. Scaling shows that $\|M_{V,r}\|_{L^p(\mathbb{R}^n)} = \|M_V\|_{L^p(\mathbb{R}^n)}$. When f is the

indicator of the unit ball in $\mathbb{R}^n$, there holds

$$\mathbf{M}_{V,N}f(x) \gtrsim \frac{1}{|x|}, \quad \frac{N}{2} < |x| < N,$$

if $V \subset \mathbb{S}^{n-1}$ is a $\frac{c}{N}$ -net with $c > 0$ sufficiently small; we then gather that (1.2), (1.3) are sharp up to the η -correction for the codimension one case, $m = n - 1$. For general codimensions we can adjust the example above by taking $V \subset \mathbb{S}^m \subset \mathbb{S}^{n-1}$ and $f := f_1 \otimes f_2 : \mathbb{R}^{m+1} \times \mathbb{R}^{n-(m+1)} \rightarrow \mathbb{R}$ where f_1 is the indicator of the unit ball in $\mathbb{R}^{m+1}$ and f_2 is a smooth bump function in $\mathbb{R}^{n-(m+1)}$ which is identically 1 on the unit ball of $\mathbb{R}^{n-(m+1)}$. This modification also proves the sharpness of (1.2), (1.3), up to the η -correction for general codimensions. In fact, a logarithmic correction is necessary in the case $m = 1$; see [12, Proposition 1.3].

All cases of Theorem A are new, except for the case $m = 1$, which dates back to the work of Córdoba [13] and Barrionuevo [5], and the case of $Z = \mathbb{S}^2$, which is the main result of [19] by Demeter. In the latter case, we prove a more precise result improving on the logarithmic correction of [19]. To describe this improvement, we introduce the following iterated logarithmic function; for integers $k, N \geq 1$ we set

$$\log^{[1]} N := \log(2 + N), \quad \log^{[k]} N := \log(2 + \log^{[k-1]} N).$$

THEOREM B. *For every $k \geq 1$ there exists a constant $\Theta_k > 1$ such that*

$$\sup_{\substack{V \subset \mathbb{S}^2 \\ \#V \leq N^2}} \|\mathbf{M}_V\|_{L^2(\mathbb{R}^3)} \leq \Theta_k N^{\frac{1}{2}} \sqrt{\log N} \log^{[k]} N.$$

As customary, we prove our estimates for a Fourier analogue of $\mathbf{M}_V$, namely the maximal directional multiplier operator $\mathbf{A}_V$ defined in (2.2) below, which dominates $\mathbf{M}_V$ pointwise on the cone of positive functions. If f has frequency support in the annulus $\{\xi \in \mathbb{R}^n : s < |\xi| < 2s\}$, our proof yields the stronger estimate

$$\sup_{\substack{V \subset \mathbb{S}^2 \\ \#V \leq N^2}} \|\mathbf{A}_V f\|_{L^2(\mathbb{R}^3)} \lesssim_k N^{\frac{1}{2}} \log^{[k]} N \|f\|_{L^2(\mathbb{R}^3)}$$

uniformly over $s > 0$. Modifying the example discussed after the statement of Theorem A, namely testing the operator norm on a suitable frequency cutoff and modulation of the indicator of the unit ball in $\mathbb{R}^3$, reveals that the N -dependence of the latter estimate is sharp modulo the iterated logarithmic correction; see [32] for details.

Theorem A yields $L^2(\mathbb{R}^n)$ -bounds for the Nikodym maximal function

$$\mathbf{M}_{Z_m, \delta} f(x) = \sup_{x \in T \in \mathcal{T}(Z_m, \delta)} \frac{1}{|T|} \int_T |f|$$

where $\mathcal{T}(Z_m, \delta)$ is the collection of tubes T with length 1 and $n - 1$ -dimensional cross section of width $0 < \delta \ll 1$, where the long side oriented along some direction $v \in Z_m \cap \mathcal{A}_n(1)$.

THEOREM C. *Let $Z_m \subseteq \mathbb{R}^n$ be a real algebraic variety of dimension $1 \leq m \leq n - 1$. Then for all $\eta > 0$ there is a constant $\Theta = \Theta(Z_m, \eta)$ independent of $0 < \delta \ll 1$, such that*

$$\|M_{Z_m, \delta}\|_{L^2(\mathbb{R}^n)} \leq \Theta \delta^{-\frac{(m-1)}{2} - \eta}.$$

Proof. Let $\delta > 0$ and $V \subset Z_m$ be a δ -net with $\#V \lesssim \delta^{-m}$. A well-known reduction [13, 19] yields that

$$\|M_{Z_m, \delta}\|_{L^2(\mathbb{R}^n)} \lesssim \|M_V\|_{L^2(\mathbb{R}^n)}$$

so that the claim follows by an application of Theorem A. $\square$

The exponent in Theorem C is in general optimal up to the arbitrarily small η -loss. Indeed, notice that the example provided after the statement of Theorem A with $N \sim 1/\delta$ also applies to the bounds of Theorem C above. This is because the Nikodym maximal operator is defined with respect to δ -tubes; one then can reduce the supremum in the definition of the Nikodym maximal function to be taken over a subset of δ -tubes pointing along a δ -net on the unit sphere; see [19, p. 718] for the details of this argument.

This result appears to be new in all cases except for $m = n - 1$, which is a classical estimate of Córdoba [12], and $m = 1$ [13, Theorem B], also due to Córdoba.

The converse direction of the implication leading from Theorem A to Theorem C, namely obtaining sharp bounds for the *thin* averages of (1.1) from sharp bounds for *thick* averages over tubes, as the ones involved in the definition of the Nikodym maximal operator, is not feasible as the averages (1.1) are more singular; see also [19]. An exception to this heuristic is, in two dimensions, the case of uniformly distributed directions on the circle $\mathbb{S}^1$. In this setting, the difference between thin and thick averages is a square function which, at least in $L^2(\mathbb{R}^2)$, can be easily treated by overlap considerations; see also the discussion below in §1.3.

1.2. Motivation. In addition to the intrinsic relevance of Theorems A, B and C to the realm of classical differentiation theory, this paper finds motivation within a more general program, aimed at understanding directional maximal and singular integral operators in higher dimensions, *under no particular structural assumptions*. This is in contrast with previous results in the literature, for instance [5, 6, 13], which are obtained for sets of directions complying with some type of geometric configuration. In the papers cited above, the structure is that of uniform distribution of the discrete set on the sphere. On the other hand, in e.g., [37] and references therein, the set of directions is allowed to be infinite but is assumed to be

lacunary. In both cases, these conditions allow for an effective splitting into subsets which are either of controlled cardinality, as in the case of uniform distribution, or exhibit self-similar behavior as in the case of lacunary directions. Both families of results mentioned above, as well as their two-dimensional counterparts as in [1, 2, 3], show that a very effective way to handle directional operators is the divide and conquer technique.

For arbitrary sets of directions in higher dimensions this technique presents a new challenge: an efficient partitioning of a non-uniformly distributed, not naturally ordered set of directions on $\mathbb{S}^2$, for instance, is far from obvious. Inspired by recent developments in harmonic analysis, most notably the improvements on the restriction problem by Guth [26, 27], our proofs are based on a polynomial partitioning scheme adapted to maximal directional operators. We establish a novel strategy for handling these operators in a very general setting: the directions lie on an algebraic variety of arbitrary codimension in $\mathbb{R}^n$, and essentially no other structure is assumed. To exemplify the intrinsic gain that is brought by this perspective, we mention Theorem B, which is about directions on the sphere $\mathbb{S}^2$ and where no *a-priori* relevant algebraic structure is present. The improvement over [19] is obtained by partitioning our directions into connected components of the complement of the zero set of a polynomial P on the sphere, each containing a roughly constant number of points, and with favorable overlap properties—not too many cells intersecting each given hyperplane—and using the algebraic structure, see the dimension 1 estimate of Theorem D, to handle the contribution of those directions falling on the partitioning zero set.

We believe that our methods can be furthered to obtain L^p -estimates for more general singular and maximal averages in higher dimensions, when there is no structure in the set of directions. The natural point of view within the polynomial method is that of systematically attacking the cases of directions lying on algebraic varieties in $\mathbb{R}^n$ of arbitrary codimension. As an example of relevant open question, it is worth noting that no nontrivial bound is currently available for the *multi-scale directional maximal operator* along directions lying on a polynomial subvariety of $\mathbb{S}^{n-1}$, without any additional structure.

1.3. Background. The study of directional maximal and singular integrals initially arose as a natural companion to questions on Kakeya-type maximal operators, Bochner-Riesz multipliers, and the conjectures of Zygmund and Stein for corresponding objects defined along 2-dimensional vector fields: we provide a short overview with particular focus on the L^2 -theory. In [12] Córdoba showed that the maximal average defined with respect to rectangles in the plane, of fixed eccentricity δ , is bounded on $L^2(\mathbb{R}^2)$ with an operator norm of the order $(\log \delta^{-1})^{\frac{1}{2}}$; the bound above is best possible as revealed by a counterexample constructed by means of the Kakeya set. We note here that, in two dimensions, this operator is essentially equivalent to a maximal averaging operator along uniformly δ -spaced directions in

$\mathbb{S}^1$, their difference being that of an easy to treat square function; it is also of some importance to highlight that this comparison is not as readily available in higher dimensions, nor in the case of more general algebraic varieties in place of the sphere. Soon after the result of [12], Strömberg, [41], showed that the $L^2(\mathbb{R}^2) \rightarrow L^{2,\infty}(\mathbb{R}^2)$ norm of the maximal average with respect to rectangles pointing in a uniformly distributed set of N directions in $\mathbb{S}^1$, without any restriction on their eccentricity, is of the order $(\log N)^{\frac{1}{2}}$, and this is best possible. The numerology here is $\delta = N^{-1}$. These results were generalized by Katz in [29, 30], where the author proved the same sharp bound for the $L^2(\mathbb{R}^2) \rightarrow L^{2,\infty}(\mathbb{R}^2)$ bound for the operator norm of the maximal average with respect to an *arbitrary* set of N directions in $\mathbb{S}^1$. We note here that L^2 is the critical L^p -space for the Kakeya, or Nikodym maximal operators in $\mathbb{R}^2$, while in $\mathbb{R}^n$ the critical exponent is $p = n$; see [42].

Parallel to the results above was the investigation of the maximal averages in the plane given by a set of directions which is *infinite*, but possesses certain arithmetic-geometric structure. Collectively Strömberg, [40], R. Fefferman and Córdoba, [14], Nagel, Stein, and Wainger, [35], and Sjögren and Sjölin, [39], proved that such maximal operators are bounded on $L^p(\mathbb{R}^2)$ whenever the set of directions is a lacunary set of finite order. The striking result of Bateman, [8], characterized lacunary sets of finite order in the plane as the only sets that give rise to bounded directional maximal operators on $L^p(\mathbb{R}^2)$ for some (equivalently any) $p \in (1, \infty)$.

In the higher (ambient space $\mathbb{R}^n$, $n \geq 2$) dimensional setting, the picture is far from complete. We note that some special cases of higher-dimensional lacunary sets of directions were introduced by Nagel, Stein, and Wainger in [35], and by Carbery in [10], where the authors showed the boundedness of the corresponding maximal operator. However, it was only recently that Parcet and Rogers, [37], gave a general definition of lacunary sets of directions in any dimension and proved the boundedness of maximal operators along such directions in $L^p(\mathbb{R}^n)$, for all $p \in (1, \infty)$ and all $n \geq 2$. Recently in [22], the authors of the present paper obtained the best possible bound for the Hilbert transform along sets of directions in $\mathbb{S}^2$ that are lacunary in the sense of [37].

For averaging operators with respect to arbitrary sets of directions in dimensions $n \geq 2$ Demeter has studied the case of maximal directional averages along sets $V \subset \mathbb{S}^2$ and showed L^2 -bounds of the order $N^{\frac{1}{2}}(\log N)$ for the case that $\#V = N^2$. This result is improved by a $\sqrt{\log N}$, modulo iterated logarithmic losses, in Theorem B of the present paper. Also relevant for us is the result of Córdoba, [13], for equispaced directions lying on a smooth parametrizable curve in $\mathbb{S}^{n-1}$, proving L^2 -bounds of the order $(\log N)^2$, for the corresponding multiscale maximal operator. This bound was improved by Barrionuevo in [5] to the best possible order $\log N$. Finally, Barrionuevo in [6], has proved almost optimal L^2 -bounds in arbitrary dimension for the case of N uniformly distributed directions in $\mathbb{S}^{n-1}$.

1.4. Techniques. The bulk of the paper is devoted to the proofs of Theorems A and B. These proofs have a similar coarse structure. A finite set of directions V lying on an algebraic variety $Z \subset \mathbb{R}^n$ of dimension m (the sphere in Theorem B) is partitioned by the zero set of a polynomial P into cells C , each containing at most a fixed portion of the original points, plus points lying exactly on (or sufficiently close to) the zero set of P . These lie on an algebraic variety of dimension $\mu = m - 1$ and their contribution is estimated by induction, when $\mu > 1$, or by direct methods, when $\mu = 1$; see for example Theorem D.

For the cellular part, the estimate is based on the observation that the directions contributing at a given frequency point ξ are localized on a fattening of the hyperplane perpendicular to ξ . If, at first approximation, we ignore the complications brought by the fattening, we can estimate how many of the $V \cap C$ operators overlap at ξ by counting, via Milnor-Thom-type theorems, in how many components the variety $\xi^\perp \cap Z$ is split by the zero set of P . When $Z = \mathbb{S}^2$ for instance, $\xi^\perp \cap Z$ is a circle, and the number of zeros of P on the circle is controlled by the degree of P . In the realistic case, we are dealing with fat hyperplanes and the additional term consisting of those cells staying close to a $(n - 1)$ -dimensional hyperplane is handled by comparison with directions lying on an algebraic subvariety of the hyperplane: an approximate projection of Z . This term is dealt via induction on the ambient space dimension n .

Polynomial partition on subvarieties of $\mathbb{R}^n$ is in general challenging and an optimal procedure is not yet completely understood, especially in codimension 2 and higher: see the articles [7, 24, 34, 44] and references therein for recent developments. Motivated by the study of the Fourier restriction operator in higher dimensions, Guth [27] introduced a polynomial partitioning scheme based on *transverse complete intersections* (TCI), namely algebraic varieties of the classes $\mathcal{Z}_{m,n}^\times(D)$ defined in Section 4, which in effect allows one to partition densities supported on m -varieties as if they were on $\mathbb{R}^m$, by keeping some uncertainty in the partitioning polynomials. Similar procedures have been used in [36]. We adapt the scheme of [27] to our context of directional operators, the final result in this sense being Proposition 2.10. The first main additional difficulty, compared to [27, 36], is that we are partitioning points and not densities. We tackle this issue by replacing points from the initial polynomial wall with nearby points, sitting on a nicer wall of the TCI-type; this replacement can be made harmless for our quantitative estimates. The second is that the approximate projection procedure hinted at above does not preserve TCIs. We remedy this by covering an arbitrary algebraic variety of dimension m with an arbitrarily small neighborhood of a union of a controlled number of TCIs of dimension no more than m ; see Proposition 4.5.

Interestingly, the recent article by Katz and Rogers [31] solves a conjecture of Guth [27] concerning the maximal number of δ -separated δ -tubes contained in the δ -neighborhood of an algebraic subvariety of $\mathbb{R}^n$. While their estimate is somewhat dual to that of Theorem C, none of the two can be promptly reduced to the other

one. However, remarkably, the proof techniques of the main result of [31] share some aspects with our arguments of Section 4: in particular, the authors of [31] appeal to Tarski's quantifier elimination principle within the category of semialgebraic sets, combined with Gromov's complexity estimate. In contrast, inspired by arguments in [34] by Matoušek and Patakova, we construct (approximate) projections of algebraic varieties by explicitly computing elimination ideals of the original variety, via Gröbner bases and the related quantifier elimination theorem; see [15, 16] and references therein. This technique allows us to work within the class of algebraic sets.

1.5. Structure. Section 2 contains some preliminary analytic tools and a version of polynomial partition on m -dimensional polynomial subvarieties $\mathbb{R}^n$ adapted to our problem, and involving directional averages; see Proposition 2.10. The proof of Theorem B is given in Section 3. Section 4 contains several definitions and tools from algebraic geometry. The proof of Proposition 2.10, which will find use in the proof of Theorem A, can also be found in Section 4. Finally, the proof of Theorem A is given in the final Section 5.

Acknowledgments. The authors want to express their gratitude to the reviewers for their careful reading and suggestions which contributed to improvements in the presentation.

2. Preliminaries.

2.1. Notation and recurring definitions. By $\Theta_{\alpha_1, \dots, \alpha_j}$ we denote a positive constant, possibly depending on the parameters $\alpha_1, \dots, \alpha_j$ only, which may differ at each occurrence. We also write

$$\begin{aligned} A \lesssim_{\alpha_1, \dots, \alpha_j} B &\iff A \leq \Theta_{\alpha_1, \dots, \alpha_j} B, \\ A \sim_{\alpha_1, \dots, \alpha_j} B &\iff A \lesssim_{\alpha_1, \dots, \alpha_j} B, \quad B \lesssim_{\alpha_1, \dots, \alpha_j} A. \end{aligned}$$

For a bounded operator $T : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$ we use the shorthand notation $\|T\|_{L^2(\mathbb{R}^n)} := \|T : L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)\|$. If $v \in V \subset \mathbb{R}^n \setminus \{0\}$ we write $v' := v/|v|$ and $V' := \{v' : v \in V\} \subset \mathbb{S}^{n-1}$. We use the notation

$$\text{dist}(U, V) := \sup_{u \in U} \inf_{v \in V} |u - v|$$

for the (asymmetric) distance between $U, V \subset \mathbb{R}^n$. We will be working with (frequency and directional) annuli: for $R \geq 1$,

$$\mathcal{A}_n(R) := \{\xi \in \mathbb{R}^n : R^{-1} \leq |\xi| < 2R\}.$$

We define the frequency bands

$$(2.1) \quad R_{\xi,s} := \{\eta \in \mathcal{A}_n(1) : |\xi \cdot \eta| < s|\eta|\}, \quad \xi \in \mathbb{R}^n \setminus \{0\}, \quad s > 0.$$

Let $A_{v,s}$ be the operator defined in (2.2) below and S_1 be a smooth frequency cutoff adapted to the annulus $\mathcal{A}_n(1)$. Abusing notation we will write $\widehat{S_1}$ for the (smooth) Fourier multiplier of the operator S_1 ; we will be using repeatedly that the Fourier support of $A_{v,s} \circ S_1$ is contained in the (fat) band $R_{v,s}$. For $R \subset \mathbb{R}^n$, we write

$$f_R(x) := \int_R \widehat{f}(\xi) e^{ix \cdot \xi} d\xi.$$

to indicate the rough frequency restriction to R .

2.2. Single scale averages. Our tools will be largely Fourier analytic in nature: in fact, we will work with maximal operators obtained by replacing the rough averages in (1.1) with the smooth directional Fourier multipliers

$$(2.2) \quad A_{V,s} f := \sup_{v \in V} |A_{v,s} f|, \quad A_{v,s} f(x) := \int_{\mathbb{R}^n} \widehat{f}(\xi) \psi\left(\frac{\xi \cdot v}{s}\right) e^{ix \cdot \xi} d\xi$$

with $x \in \mathbb{R}^n$ and $s > 0$. Throughout, $\psi : \mathbb{R} \rightarrow \mathbb{R}$ will be the Fourier transform of an even nonnegative Schwartz function $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ with the following properties:

$$\text{supp } \psi \subset [-2^{-10}, 2^{-10}], \quad \Psi(t) \geq \mathbf{1}_{(-1,1)}(t).$$

This smooth function will remain fixed throughout the paper so we suppress any implicit dependence on the choice of ψ . We are chiefly interested in the case $s = 1$, where we suppress the subscript and simply write A_v, A_V in place of $A_{v,1}, A_{V,1}$. By scaling f , we see that

$$(2.3) \quad \|A_{V,s}\|_{L^p(\mathbb{R}^n)} = \|A_V\|_{L^p(\mathbb{R}^n)} \quad \forall s > 0, \quad 0 < p \leq \infty.$$

This scaling invariance is destroyed when restricting the operators $A_{V,s}$ to act on functions with frequency support in $\mathcal{A}_n(1)$, and for this reason we work with a more general scale parameter below; see for instance Proposition 2.5.

With our choice of ψ , we have the following comparison principle; the upper bound is immediate by pointwise comparison for $f \geq 0$, while the lower bound follows from (2.8) below, and scaling.

LEMMA 2.3. *Let $V \subset \mathbb{R}^n$. Then*

$$\|M_V\|_{L^2(\mathbb{R}^n)} \sim_n \|A_V\|_{L^2(\mathbb{R}^n)}.$$

2.4. Decoupling of frequency annuli and localization. Let S_1 be a smooth Littlewood-Paley radial cutoff to the annulus $\mathcal{A}_n(1)$. Using the Chang-Wilson-Wolff inequality as in [18, 22, 21, 25], and scaling, we are able to reduce our maximal estimates to functions with frequency support in $\mathcal{A}_n(1)$, as recorded in the following proposition.

PROPOSITION 2.5. *Let $V \subset \mathbb{R}^n$ be a finite subset. Then*

$$\|A_V\|_{L^2(\mathbb{R}^n)} \lesssim_n \sqrt{\log \#V} \sup_{s>0} \|A_{V,s} \circ S_1\|_{L^2(\mathbb{R}^n)}.$$

The main advantage of the reduction of Proposition 2.5 is that, in addition to being Fourier localized, the averages (2.2) also enjoy a localization property with respect to the direction v , in the sense of the next lemma. Below we denote by M the usual Hardy-Littlewood maximal function in $\mathbb{R}^n$

$$Mf(x) := \sup_{r>0} \frac{1}{r^n} \int_{|t| \leq r} |f(x-t)| dt, \quad x \in \mathbb{R}^n.$$

LEMMA 2.6. *Let $s > 0$ and $0 < a < 2^3$. Let $U, V \subset \mathcal{A}_n(2)$ have the property that*

$$(2.4) \quad \text{dist}(U', V') < as.$$

where $U' := \{u/|u| : u \in U\}$ and similarly for V . Then

$$\|A_{U,s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \max \left(\|A_V\|_{L^2(\mathbb{R}^n)}, \|M\|_{L^2(\mathbb{R}^n)} \right) \approx \|A_V\|_{L^2(\mathbb{R}^n)}.$$

Proof. Below all the implicit constants in the almost inequality sign may depend on n only, and may differ at each occurrence. The approximate equality in the conclusion is obvious so it suffices to show the estimate

$$\|A_{U,s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \max \left(\|A_V\|_{L^2(\mathbb{R}^n)}, \|M\|_{L^2(\mathbb{R}^n)} \right).$$

Note also the trivial estimate

$$(2.5) \quad \sup_{s \geq 1} \|A_{U,s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \|M\|_{L^2(\mathbb{R}^n)}.$$

Indeed the operator on the left-hand side has symbol

$$m(\xi) = \psi \left(\frac{\xi \cdot v}{s} \right) \phi(|\xi|)$$

where ϕ is smooth and supported on an annulus $|\xi| \approx 1$. As $|\xi| \approx 1 \leq s$ on $\text{supp } m$ we get that $|\partial_\xi^\alpha m(\xi)| \lesssim |\xi|^{-|\alpha|}$ for $0 \leq |\alpha| \leq 10n$, whence the estimate (2.5).

For the rest of the proof we assume that $s \leq 1$; it is convenient to work with the scaled rough averages and maximal operators

$$\langle f \rangle_{v,s}(x) := \int_{s|t|<1} f(x-tv) \frac{sdt}{2}, \quad \mathbf{M}_{V,s}f(x) := \sup_{v \in V} \langle |f| \rangle_{v,s}(x);$$

note the somewhat nonstandard use of space scales $1/s$ and that the same scaling property (2.3) holds for $\mathbf{M}_V$ in place of $\mathbf{A}_V$.

As $V \subset \mathcal{A}_n(2)$ and $V' \subset \mathbb{S}^{n-1}$, we have the estimates

$$(2.6) \quad \|\mathbf{M}_{V',\sigma}\|_{L^2(\mathbb{R}^n)} \lesssim \|\mathbf{M}_{V,\frac{\sigma}{2}}\|_{L^2(\mathbb{R}^n)} \lesssim \|\mathbf{A}_{V,\frac{\sigma}{2}}\|_{L^2(\mathbb{R}^n)}$$

uniformly over $\sigma > 0$, where the last inequality is obtained by Lemma 2.3. By virtue of (2.6) it suffices to prove

$$(2.7) \quad \|A_{U,s} \circ S_1 f\|_{L^2(\mathbb{R}^n)} \lesssim \sup_{\sigma < s} \|\mathbf{M}_{V',\sigma}(S_1 f)\|_{L^2(\mathbb{R}^n)}.$$

We write $g := S_1 f$ below for simplicity. Let $v(u) \in V$ be chosen such that $|v(u)' - u'| < as$. We claim the estimate

$$(2.8) \quad |A_{u,s}g| \lesssim \sum_{k \geq 1} 2^{-kn} \mathbf{M}(\langle |g| \rangle_{v(u)', 2^{-k}s}).$$

We first show how (2.8) implies (2.7):

$$\begin{aligned} \|A_{U,s}g\|_{L^2(\mathbb{R}^n)} &\lesssim \left\| \sup_{v' \in V'} \sum_{k \geq 1} 2^{-kn} \mathbf{M}(\langle |g| \rangle_{v', 2^{-k}s}) \right\|_{L^2(\mathbb{R}^n)} \\ &\lesssim \sum_{k \geq 1} 2^{-kn} \|\mathbf{M} \circ \mathbf{M}_{V', 2^{-k}s}g\|_{L^2(\mathbb{R}^n)} \\ &\lesssim \sup_{k \geq 1} \|\mathbf{M}_{V', 2^{-k}s}g\|_{L^2(\mathbb{R}^n)} \end{aligned}$$

which is (2.7); note that (2.8) has been used in the first inequality.

In order to prove (2.8) we may assume, by rotation invariance, that $v(u)' = (1, 0, \dots, 0)$. Then

$$(2.9) \quad |u_1| \lesssim 1, \quad \frac{1}{s} \sup_{2 \leq j \leq n} |u_j| \leq \frac{4}{s} \sup_{2 \leq j \leq n} |u'_j| \leq 4 \min\left(\frac{1}{s}, a\right).$$

Let us write

$$(A_{u,s}S_1 f)^\wedge =: m_{u,s} \widehat{f}$$

for the multiplier of $A_{u,s}S_1$. Now note that under our assumptions on the parameter a there exists a dimensional constant $c > 1$ such that the set $R_{v,cs}$ contains the

support of $m_{u,s}$. Relying on (2.9), for every multiindex α we obtain

$$|\partial_\xi^\alpha m_{u,s}(\xi)| = \left| \partial_\xi^\alpha [\widehat{S_1}(\xi)\psi(s^{-1}u \cdot \xi)] \right| \lesssim s^{-\alpha_1} \min(1, s^{-|\alpha|+\alpha_1}) \approx s^{-\alpha_1}$$

since $s \leq 1$. This readily implies that $K_{u,s} := \widehat{m_{u,s}}$ satisfies the estimate

$$|K_{u,s}(x_1, \dots, x_n)| \lesssim \frac{s}{(1+s|x_1|)^{10n}} \left(1 + \sum_{j=2}^n |x_j|\right)^{-10n} \lesssim \sum_{k \geq 1} 2^{-kn} \widetilde{\mathbf{1}}_{B_k}(x),$$

where B_k is the tube centered at the origin and of sidelengths $s^{-1}2^k$ along e_1 and 2^k along $e_2, \dots, e_n$ and $\widetilde{\mathbf{1}}_{B_k} = \mathbf{1}_{B_k}/|B_k|$. Hence

$$\begin{aligned} |A_{u,s}g| &= |g * K_{u,s}| \leq |g| * |K_{u,s}| \\ &\lesssim \sum_{k \geq 1} 2^{-kn} |g| * \widetilde{\mathbf{1}}_{B_k} \lesssim \sum_{k \geq 1} 2^{-kn} \mathbf{M}(\langle |g| \rangle_{v(u)', 2^{-k}s}) \end{aligned}$$

where we used the inequality $|g| * \widetilde{\mathbf{1}}_{B_k} \lesssim \mathbf{M}(\langle |g| \rangle_{v(u)', 2^{-k}s})$. We have reached (2.8) and the proof is thus complete. $\square$

Lemma 2.6 tells us that it suffices to study small scales when f has annular frequency support.

COROLLARY 2.7. *There holds*

$$\sup_{s > 2^{-5}} \|A_{\mathcal{A}_n(2),s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n 1.$$

Proof. Let $s > 2^{-5}$. Let V be a 2^{-10} -net in $\mathbb{S}^{n-1}$. Then $U = \mathcal{A}_n(2)$ and V satisfy the assumption (2.4) of Lemma 2.6 with $a = 2^{-4}$. Hence

$$\|A_{\mathcal{A}_n(2),s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \|A_V\|_{L^2(\mathbb{R}^n)} \lesssim_n 1$$

as V has cardinality Θ_n and for every v the averaging operators $f \mapsto A_v f$ are bounded on $L^2(\mathbb{R}^n)$. Note that an alternative proof of the statement of the corollary follows by the kernel estimates established in the proof of Lemma 2.6, applied for $s \gtrsim 1$. $\square$

COROLLARY 2.8. *Let $0 < s < 2^{-5}$, $n \geq 3$, and let $Z \subset \mathcal{A}_n(1)$ have the property that*

$$\text{dist}(Z, \xi^\perp) \leq as$$

for some $a \leq 2^3$. Denote by $\Pi_\xi Z$ the orthogonal projection of Z on $\xi^\perp \equiv \mathbb{R}^{n-1}$. Then

$$\sup_{\substack{U \subset Z \\ \#U \leq N}} \|A_{U,s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \sup_{\substack{V \subset \Pi_\xi Z \\ \#V \leq N}} \|A_V\|_{L^2(\mathbb{R}^{n-1})}.$$

Proof. Let $U \subset Z$ have $\#U \leq N$. The set $V := \Pi_\xi U$ is contained in $\Pi_\xi Z \subset \mathbb{R}^{n-1}$ and has $\#V \leq N$. Then U and V satisfy the assumption (2.4) of Lemma 2.6. We obtain that

$$\|A_{U,s} \circ S_1\|_{L^2(\mathbb{R}^n)} \lesssim_n \|A_V\|_{L^2(\mathbb{R}^n)} \lesssim_n \|A_V\|_{L^2(\mathbb{R}^{n-1})}$$

where the first inequality is an application of Lemma 2.6, and the second follows from Fubini's theorem applied on the slices $E_b := \{x \in \mathbb{R}^n : x \cdot \xi = b\}$, $b \in \mathbb{R}$. The proof is complete. $\square$

2.9. Polynomial partition of directions. We write $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and denote by $\mathbb{R}[x_1, \dots, x_n]$ the ring of n -variable polynomials with real coefficients.

A *transverse complete intersection* Z in $\mathbb{R}^n$ of dimension $m \in \{1, \dots, n-1\}$ is the common zero set of polynomials $P_1, \dots, P_{n-m} \in \mathbb{R}[x_1, \dots, x_n]$:

$$Z = \mathbf{Z}(P_1, \dots, P_{n-m}) = \{x \in \mathbb{R}^n : P_1(x) = \dots = P_{n-m}(x) = 0\},$$

with the property that

$$(2.10) \quad \nabla P_1(x) \wedge \dots \wedge \nabla P_{n-m}(x) \neq 0$$

holds for all $x \in Z$. In particular transverse complete intersections are m -dimensional smooth submanifolds of $\mathbb{R}^n$. We will in general use the shorthand notation (TCI) for transverse complete intersections.

The following proposition, which is fundamental for the analysis in the current paper, is a form of polynomial partitioning on manifold, inspired by [26, 27] (see [44, Theorem 2.3] for a different approach) and adapted to the structure of our problem. It partitions a finite set of directions V lying on an m -dimensional TCI into a boundary component $V_\times$ of points lying arbitrarily close to a controlled number of $(m-1)$ -dimensional TCIs, and a cellular component $V_\circ$. The component $V_\circ$ is itself partitioned into connected components of the complement of zero sets of boundedly many polynomials of controlled degree, each containing at most a fraction of the original number of points. The proof of this proposition together with a detailed presentation and analysis of relevant tools from algebraic geometry is contained in Section 4.

PROPOSITION 2.10. *Let $Z = \mathbf{Z}(P_1, \dots, P_{n-m}) \subset \mathbb{R}^n$ be a transverse complete intersection of degree D . Let $V \subset Z$ be a finite point set with $\#V \leq N^m$. For each*

integer E with $E^{m-1} \geq D^n$ and $\delta > 0$ we may perform a partition

$$V = V_{\circ} \cup V_{\times}$$

with the following properties.

- (1) There exist $\leq \Theta_{m,n} D^n$ transverse complete intersections

$$W_j = \mathbf{Z}(P_1, \dots, P_{n-m}, Q_j)$$

of dimension $m-1$ and degree $\leq \Theta_{m,n} E$ such that

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

- (2) There exist $\leq \Theta_{m,n} D^n E^m$ disjoint connected subsets $C \in \vec{C}$ of Z with the property that

$$V_{\circ} = \bigcup_{C \in \vec{C}} V_C, \quad V_C := V \cap C, \quad \#V_C \leq \left(\frac{N}{E}\right)^m,$$

and such that for almost every $\xi \in \mathbb{R}^n$ and every $a \in \mathbb{R}$, the intersection $C \cap \{x \in \mathbb{R}^n : \xi \cdot x = a\}$ is nontrivial for at most $\leq \Theta_{n,m} D^{2n} E^{m-1}$ elements of $\vec{C}$.

3. Directional averages along arbitrary directions on $\mathbb{S}^2$. This section is dedicated to the proof of Theorem B. Given a set of directions in $\mathbb{S}^2$ we will apply polynomial partitioning, in the form of Proposition 2.10, to reduce the problem of estimating the corresponding maximal directional average into two parts. The first one, corresponding to the cells of the partition, will be handled inductively. The second term corresponds loosely to the zero set of the polynomial partitioning, which in this case is an algebraic variety of dimension 1. In particular, the application of Proposition 2.10 of the previous section allows us to work with a nice class of algebraic varieties, namely transverse complete intersections.

3.1. Directional averages along one dimensional varieties. In the theorem below we work on the particular case of algebraic varieties of dimension one, contained in $\mathbb{S}^2$. We note here that for a single smooth parametrizable curve in $\mathbb{S}^{n-1}$ an argument of Córdoba from [13, p. 223], or of Christ, Duoandikoetxea, and Rubio de Francia from [11], could be used to yield a bound of the form $D \log N$. However these arguments do not seem to directly provide the degree dependence $D^{\frac{1}{2}}$ in the case of a general one-dimensional algebraic variety which might have several connected components. This degree dependence is crucial for our application in the case of averages along directions in $\mathbb{S}^2$, forcing us to modify the argument from [13] in order to get the correct behavior $D^{\frac{1}{2}}$ in terms of the degree, at the price of a slightly worse bound in terms of the cardinality of the set of directions. As we will need to apply the result below for a relatively high degree D this trade-off

is in our favor. We only insist on this variation for the case of algebraic varieties of dimension 1 contained in $\mathbb{S}^2$ as for the higher codimensional case the result of Córdoba is sufficient for our purposes; see Subsection 5.3.

THEOREM D. *Let $Z \subset \mathbb{S}^2$ be a transverse complete intersection of dimension 1, given by a nonzero polynomial P with $\deg P = D$, namely*

$$Z = \{x \in \mathbb{R}^3 : P(x) = P_{\text{sph}}(x) = 0\}$$

where $P_{\text{sph}}(x) := x_1^2 + x_2^2 + x_3^2 - 1$ and

$$\nabla P(x) \wedge x \neq 0 \quad \forall x \in Z.$$

Then

$$\sup_{\substack{V \subset Z \\ \#V=N}} \|A_V f\|_{L^2(\mathbb{R}^3)} \lesssim D^{\frac{1}{2}} (\log N)^{\frac{3}{2}} \|f\|_{L^2(\mathbb{R}^3)}.$$

Furthermore we have the following single-annulus estimate

$$(3.1) \quad \sup_{s>0} \sup_{\substack{V \subset Z \\ \#V=N}} \|A_{V,s} \circ S_1 f\|_{L^2(\mathbb{R}^3)} \lesssim D^{\frac{1}{2}} \log N \|S_1 f\|_{L^2(\mathbb{R}^3)}.$$

Proof. We fix a transverse complete intersection Z as in the statement of the theorem and let $V \subset Z$ be a set of directions with $\#V = N$. By an application of the Chang-Wilson-Wolff inequality of Proposition 2.5, the proof of the lemma will follow from the single annulus estimate (3.1) which we prove in what follows. By Corollary 2.7, it suffices to treat the case $0 < s < 2^{-5}$. The proof is divided in steps.

Decomposition into good and bad components. The variety Z can be decomposed into connected components $Z = \cup_j Z_j$. It is important to note here that the number of connected components of Z is at most $O(D^2)$; see for example [4] and the references therein.

Let $\Omega \subset \mathbb{S}^2$ be a $2^{-10}s$ -net, that is a set of $2^{-10}s$ -separated vectors $\xi \in \mathbb{S}^2$ such that the collection of caps

$$\omega_{\xi,s} = \{v \in \mathbb{S}^2 : |v - \xi| < s\}, \quad \xi \in \Omega,$$

is a finitely overlapping cover of $\mathbb{S}^2$. A *cluster* $k_\xi \subset \{Z_j\}_j$ with *top* $\xi \in \Omega$ is a subset $\{Z_{j_k}\}_k \subseteq \{Z_j\}_j$ with $Z_{j_k} \subseteq R_{\xi,3s}$. We say that a cluster is *bad* if its interior contains (in the sense of set inclusion) more than D distinct components from $\{Z_j\}_j$. By a greedy selection algorithm we can identify a set $\Omega_{\text{bad}} \subseteq \Omega$ and a corresponding set $Z_{\text{bad}} \subseteq \{Z_j\}$ of at most $O(D)$ bad clusters. Setting $\Omega_{\text{good}} := \Omega \setminus \Omega_{\text{bad}}$ we then know that for every $\xi \in \Omega_{\text{good}}$ the band $R_{\xi,3s}$ contains at most D connected components from $\{Z_j\}_j$. We define V_{good} and V_{bad} to be those directions

from W contained in Z_{good} , and Z_{bad} , respectively. We also write V_ξ for those directions contained in a cluster k_ξ for some $\xi \in \Omega$.

Estimate for the bad component directions V_{bad} . We first estimate the operator corresponding to directions in V_{bad} on a single frequency annulus. We have

$$(3.2) \quad \|A_{V_{\text{bad}},s} \circ S_1 f\|_{L^2(\mathbb{R}^3)} \leq \left(\sum_{\xi \in \Omega_{\text{bad}}} \|A_{V_\xi,s} \circ S_1 f\|_{L^2(\mathbb{R}^3)}^2 \right)^{\frac{1}{2}}.$$

We note that for $\xi \in \Omega_{\text{bad}}$ the set $\bigcup \{Z_j : Z_j \in k_\xi\}$ is contained in a frequency band $R_{\xi,3s}$ so we can approximate the directions in V_ξ for $\xi \in \Omega_{\text{bad}}$ by directions lying on $\xi^\perp \cap \mathbb{S}^2 \equiv \mathbb{S}^1$. Note that the sets $\{V_\xi\}_{\xi \in \Omega_{\text{bad}}}$ are pairwise disjoint, by our greedy selection algorithm, and each has cardinality $\leq N$. More precisely, we apply Corollary 2.8 with $\xi \in \Omega_{\text{bad}}$ to obtain

$$(3.3) \quad \|A_{V_\xi,s} \circ S_1 f\|_{L^2(\mathbb{R}^3)} \lesssim \left(\sup_{\substack{U \subset \xi^\perp \cap \mathbb{S}^2 \\ \#U \leq N}} \|A_U\|_{L^2(\mathbb{R}^2)} \right) \|S_1 f\|_{L^2(\mathbb{R}^3)} \\ \lesssim \sqrt{\log N} \|S_1 f\|_{L^2(\mathbb{R}^3)}.$$

The last inequality follows by an application of the two-dimensional single-scale result of Katz [30]. Alternative proofs of this two-dimensional result have since been given in [18, 20]. Combining (3.2) with (3.3) we gather that

$$\|A_{V_{\text{bad}},s} \circ S_1 f\|_{L^2(\mathbb{R}^3)} \lesssim \left(\sum_{\xi \in \Omega_{\text{bad}}} \log N \right)^{\frac{1}{2}} \|S_1 f\|_{L^2(\mathbb{R}^3)} \lesssim D^{\frac{1}{2}} \sqrt{\log N} \|S_1 f\|_{L^2(\mathbb{R}^3)}$$

since $\#\Omega_{\text{bad}} \leq D$. This gives the desired estimate for the bad part.

Estimate for V_{good} . We now move to the estimation of V_{good} which relies on overlap considerations. Let Z_{good} consist of some connected components $\{Z_j\}_{j \in J}$ and let V_j denote the directions of V contained in Z_j . We write

$$V_j = \{v_{j,1}, \dots, v_{j,n_j}\}$$

where, for fixed $j \in J$, the directions are sorted in consecutive order.

In this proof, we need a rough analogue of $A_{v,s}$. Let $I \subset \mathbb{R}$ be any interval and for $g \in L^2(\mathbb{R}^n)$ define

$$B_{v,I}g(x) := \int_{\mathbb{R}^n} \widehat{g}(\beta) \mathbf{1}_I(\beta \cdot v) e^{ix \cdot \beta} d\beta, \quad B_{V,I}g := \sup_{v \in V} |B_{v,I}g|.$$

We also define $I_{a,\sigma} := (a - \sigma, a + \sigma)$. By an averaging argument we have that

$$\|A_{V_{\text{good}},s} \circ S_1 f\|_{L^2(\mathbb{R}^n)} \lesssim \|\psi\|_{\text{BV}} \sup_{0 < \sigma \leq s} \sup_{|a| < s} \|B_{V_{\text{good}},I_{a,\sigma}}(S_1 f)\|_{L^2(\mathbb{R}^n)}.$$

Therefore, it suffices to estimate the $L^2(\mathbb{R}^3)$ -norm of the operator $B_{V_{\text{good}},I_{a,\sigma}} \circ S_1$ uniformly over $0 < \sigma \leq s$. This can be done as follows. Set

$$L_{\beta,a,\sigma} := \{v \in \mathbb{S}^2 : |\beta \cdot v - a| < \sigma\} = \{v \in \mathbb{S}^2 : \beta \cdot v \in I_{a,\sigma}\}, \quad \beta \in \mathcal{A}_n(1).$$

Then

$$\begin{aligned} \|B_{V_{\text{good}},I_{a,\sigma}} \circ S_1\|_{L^2(\mathbb{R}^3)} &\leq \sup_{\beta \in \mathcal{A}_n(1)} \left(\sum_{j \in J} \sum_{\ell=1}^{n_j-1} |\mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,\ell+1}) - \mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,\ell})|^2 \right)^{\frac{1}{2}} \\ &\quad + \sup_{\beta \in \mathcal{A}_n(1)} \left(\sum_{j \in J} |\mathbf{1}_{L_{\beta,a,\sigma}}(v_{j+1,1}) - \mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,n_j})|^2 \right)^{\frac{1}{2}} \\ &\quad + \|B_{\tilde{V}_{\text{good}},I_{a,\sigma}} \circ S_1\|_{L^2(\mathbb{R}^3)} \end{aligned}$$

where $\tilde{V}_{\text{good}} \subset V_{\text{good}}$ satisfies $\#\tilde{V}_{\text{good}} \leq \#V_{\text{good}}/2$.

We now estimate the square functions in the first two summands. For this fix some $\beta \in \mathcal{A}_n(1)$ and pick $\xi \in \Omega$ such that $\beta' \in \omega_{\xi,s}$. Here we remember that $\beta' = \beta/|\beta|$ and note that ξ remains fixed as long as β is fixed. For the square function in the first summand note that if $\mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,\ell+1}) - \mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,\ell}) \neq 0$ then the piece of Z_j between $v_{j,\ell}$ and $v_{j,\ell+1}$, which is connected, must cross one of the hyperplanes $\{v : \beta \cdot v = a \pm \sigma\}$. Overall the variety Z can cross these hyperplanes at most $O(D)$ times so the first summand above is $O(D^{\frac{1}{2}})$.

For the square function in the second summand we note that $\mathbf{1}_{L_{\beta,a,\sigma}}(v_{j+1,1}) - \mathbf{1}_{L_{\beta,a,\sigma}}(v_{j,n_j}) \neq 0$ implies one of the two following possibilities

$$\begin{aligned} \text{either} \quad & |\beta \cdot v_{j+1,1} - a| < \sigma \quad \text{and} \quad |\beta \cdot v_{j,n_j} - a| \geq \sigma, \\ \text{or} \quad & |\beta \cdot v_{j+1,1} - a| \geq \sigma \quad \text{and} \quad |\beta \cdot v_{j,n_j} - a| < \sigma. \end{aligned}$$

In either one of the cases above there is at least one component $Z_* \in Z_{\text{good}}$ and $\beta' \in \mathbb{S}^2$ such that $|\beta' \cdot v - a| \leq s$ for some $v \in Z_*$, where we crucially use that $|a|, \sigma \leq s$ and that $\beta \in \mathcal{A}(1)$. This implies that $Z_* \cap R_{\xi,3s} \neq \emptyset$ using the triangle inequality and the facts that $|a| \leq s$ and $\beta' \in \omega_{\xi,s}$. Thus a bound for the second summand can be calculated by counting how many of the components $Z_* \in Z_{\text{good}}$ can intersect $R_{\xi,3s}$. As the components $Z_* \in Z_{\text{good}}$, at most D of them can be contained in the interior of $R_{\xi,3s}$. On the other hand the components Z_* that satisfy $Z_* \not\subset R_{\xi,3s}$ and $Z_* \cap R_{\xi,3s} \neq \emptyset$ can be counted by the number of connected components of the set $\{\eta \in \mathbb{S}^2 : \eta \cdot \xi = \pm 3s\} \setminus Z$ which is $O(D)$. This follows from the fact that the set $\{\eta \in \mathbb{S}^2 : \eta \cdot \xi = \pm 3s\}$ is a smooth algebraic variety of dimension 1 and an

application of the Milnor-Thom theorem; see for example [4] and the references therein.

We have thus proved the estimate

$$\left\| B_{V_{\text{good}}, I_a, \sigma} \circ S_1 \right\|_{L^2(\mathbb{R}^3)} \leq CD^{\frac{1}{2}} + \left\| B_{\tilde{V}_{\text{good}}, I_a, \sigma} \circ S_1 \right\|_{L^2(\mathbb{R}^3)}.$$

This recursive inequality can be iterated on V_{good} yielding the bound $D^{\frac{1}{2}} \log N$ for the good part so the proof is complete. $\square$

3.2. Reduction to single annulus and polynomial partition. Theorem B is reduced via Proposition 2.5 and Corollary 2.7 to the following single annulus bound. As mentioned after the statement of Theorem B in §1, standard counterexamples show that the growth rate in (3.4) is sharp up to the iterated logarithmic term.

PROPOSITION 3.3. *Let $V \subset \mathbb{S}^2$ be a finite subset with $\#V = N^2$. Then for all $k \geq 1$*

$$(3.4) \quad \sup_{s>0} \left\| A_{V,s} \circ S_1 \right\|_{L^2(\mathbb{R}^3)} \lesssim_k N^{\frac{1}{2}} \log^{[k]}(N).$$

As anticipated, the proof uses polynomial partitioning on $\mathbb{S}^2$. We will prove the proposition by showing inductively that

$$K_N := \sup_{s>0} \sup_{\substack{V \subset \mathbb{S}^2 \\ \#V = N^2}} \left\| A_{V,s} \circ S_1 \right\|_{L^2(\mathbb{R}^3)}^2$$

fulfills (3.4). More precisely, for every positive integer $E > 1$ we will prove the recursive estimate

$$(3.5) \quad K_N \leq \Theta E \left[(\log N)^3 + K_{\frac{N}{E}} \right]$$

where the constant $\Theta > 0$ is independent of N and E . We show at the end of the proof how (3.5) implies the proposition.

Proof of (3.5). By Corollary 2.7 it suffices to argue for $0 < s < 2^{-5}$. Throughout, $\Theta > 0$ will be some absolute constant which may vary from line to line. To prove (3.5) we fix some $V \subset \mathbb{S}^2$ with cardinality at most N^2 . Applying the polynomial partition on $\mathbb{S}^2 = \mathbf{Z}(P_{\text{sph}}) \subset \mathbb{R}^3$ from Proposition 2.10, so that $D = 2$, with the choice $\delta = 2^{-10}s$, we obtain a decomposition

$$(3.6) \quad V = V_{\circ} \cup V_{\times}$$

with the following properties. First, there exist $\leq \Theta$ transverse complete intersections $W_j = \mathbf{Z}(P_{\text{sph}}, Q_j)$ of dimension 1 and degree $\leq \Theta E$ such that

$$(3.7) \quad \sup_{v \in V_\times} \inf_j \text{dist}(v, W_j) < 2^{-10} s.$$

Secondly, there exist $\leq \Theta E^2$ disjoint connected subsets of S^2 , denoted by $C \in \vec{\mathcal{C}}$, with the property that

$$(3.8) \quad V_\circ = \bigcup_{C \in \vec{\mathcal{C}}} V_C, \quad V_C := V_\circ \cap C, \quad \#V_C \leq \left(\frac{N}{E}\right)^2,$$

and such that for almost every $\xi \in \mathbb{R}^2$ and every $a \in \mathbb{R}$, the intersection $C \cap \{x \in \mathbb{R}^2 : \xi \cdot x = a\}$ is nontrivial for $\leq \Theta E$ elements of $\vec{\mathcal{C}}$.

Remark 3.4. For those readers who have not looked ahead at the proof of Proposition 2.10, we provide an explanation of how the bounded crossing property

$$(3.9) \quad \#\{C \in \vec{\mathcal{C}} : C \cap \{x \in \mathbb{R}^2 : \xi \cdot x = a\} \neq \emptyset\} \leq \Theta E$$

is obtained. This explanation is fully rigorous if the points of V initially lie on a coordinate neighborhood of $\mathbb{S}^2$ (say the $1/10$ neighborhood of the north pole), and is made fully rigorous in a more general setup in the proof of Proposition 2.10. The cells $C \in \vec{\mathcal{C}}$ are the connected components of $\mathbb{S}^2 \setminus Z$, where Z is an algebraic subvariety of $\mathbb{S}^2$ obtained as the zero set $Z = \mathbf{Z}(P, P_{\text{sph}})$ with P a polynomial of degree $\leq \Theta E$. The intersection of a generic hyperplane with $\mathbb{S}^2 \setminus Z$ can have at most $\leq \Theta E$ connected components (as many as the common zeros of P, P_{sph} on the hyperplane), and the number of these connected components coincides with the left-hand side of the last display.

Controlling $V_\times$. In view of (3.7) we may find subsets $V_j \subset W_j$ with $\#V_j \leq N^2$ so that for each $v \in V_\times$ there exists j and $w(v) \in V_j$ with $|v - w(v)| < 2^{-10} s$. Hence using Lemma 2.6, directions from $V_\times$ can be effectively approximated with directions from $V_j \subset W_j$. More precisely, we have that

$$(3.10) \quad \begin{aligned} \|A_{V_\times, s} \circ S_1 f\|_{L^2(\mathbb{R}^3)}^2 &\leq \Theta \left(\sup_j \|A_{V_j}\|_{L^2(\mathbb{R}^3)}^2 \right) \|S_1 f\|_2^2 \\ &\leq \Theta E (\log N)^3 \|S_1 f\|_{L^2(\mathbb{R}^3)}^2 \end{aligned}$$

applying Lemma 2.6 for the first inequality and Theorem D for each subset V_j of the transverse complete intersection W_j to finish.

Splitting V_o into good and bad clusters. First of all, let $\xi \in \Omega$ be a $2^{-10}s$ -net on $\mathbb{S}^2$ chosen so that the bounded crossing property (3.9) holds with $a = \pm 3s$, for all $\xi \in \Omega$. We define $k_\xi \subset \vec{\mathcal{C}}$ to be a cluster of cells with top $\xi \in \Omega$ if

$$C \subset R_{\xi,3s} \quad \forall C \in k_\xi.$$

We say that a cluster k_ξ is bad if it contains more than E cells. As $\#\vec{\mathcal{C}} \leq \Theta E^2$, by an iterative selection algorithm, we may construct $\Omega_{\text{bad}} \subset \Omega$ with $\#\Omega_{\text{bad}} \leq \Theta E$ such that setting

$$\vec{\mathcal{C}}_{\text{good}} := \vec{\mathcal{C}} \setminus \bigcup_{\xi \in \Omega_{\text{bad}}} k_\xi$$

there holds

$$\#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \subset R_{\xi,3s}\right\} \leq E \quad \forall \xi \in \Omega.$$

Accordingly we define

$$V_\xi := \bigcup_{C \in k_\xi} V_C, \quad V_{o,\text{bad}} := \bigcup_{\xi \in \Omega_{\text{bad}}} V_\xi, \quad V_{o,\text{good}} := \bigcup_{C \in \vec{\mathcal{C}}_{\text{good}}} V_C.$$

Controlling $V_{o,\text{bad}}$. Each of the $\leq \Theta E$ bad clusters will be controlled by a lower dimensional estimate. By construction, for each $v \in V_\xi$ we may find $u(v) \in W_\xi := \mathbb{S}^2 \cap \xi^\perp$ with $|v - u(v)| \leq 3s$. Define $U_\xi = \{u(v) : v \in V_\xi\} \subset W_\xi$, and note that $\#U_\xi \leq N^2$. Using Lemma 2.6 again we get

$$\begin{aligned} (3.11) \quad \|A_{V_{o,\text{bad}},s} \circ S_1 f\|_{L^2(\mathbb{R}^3)}^2 &\leq \Theta E \left(\sup_{\xi \in \Omega_{\text{bad}}} \|A_{U_\xi}\|_{L^2(\mathbb{R}^3)}^2 \right) \|S_1 f\|_2^2 \\ &\leq \Theta E (\log N)^3 \|S_1 f\|_2^2 \end{aligned}$$

where we used the $D = 1$ case of Theorem D on each $U_\xi \subset W_\xi$.

Controlling $V_{o,\text{good}}$. It is convenient to write $F := S_1 f$ below. For $C \in \vec{\mathcal{C}}_{\text{good}}$ we define

$$(3.12) \quad R_C := \bigcup_{v \in V_C} R_{v,s}.$$

We rely on the observation below (2.1) to gather that $A_{v,s}F = A_{v,s}F_{R_C}$ whenever $v \in C$, where we remember that F_R denotes the rough frequency restriction of F to some $R \subset \mathbb{R}^3$. Then we compute the overlap of the bands $\{R_C : C \in \vec{\mathcal{C}}_{\text{good}}\}$ at a generic $\beta \in \mathbb{R}^3$. By homogeneity it is enough to take $\beta \in \mathbb{S}^2$. Choose $\xi \in \Omega$ with

$|\xi - \beta| < 2^{-10}s$. Then the overlap of the frequency supports R_C at β is bounded by

$$(3.13) \quad \#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap R_{\beta,s} \neq \emptyset\right\} \leq \#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap R_{\xi,2s} \neq \emptyset\right\}.$$

As by construction of $\mathcal{C}_{\text{good}}$

$$\#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \subset R_{\xi,3s}\right\} \leq E$$

we are left with counting

$$\begin{aligned} & \#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap R_{\xi,2s} \neq \emptyset, C \not\subset R_{\xi,3s}\right\} \\ & \leq \#\left\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap \{x \in \mathbb{R}^2 : \xi \cdot x = \pm 3s\} \neq \emptyset\right\} \leq \Theta E \end{aligned}$$

where (3.9) has been taken into account. As each V_C contains at most $(N/E)^2$ directions, see (3.8), and $F_{R_C} = S_1 F_{R_C}$, we note that

$$\|A_{V_C,s} F_{R_C}\|_{L^2(\mathbb{R}^3)}^2 \leq K_{\frac{N}{E}} \|F_{R_C}\|_{L^2(\mathbb{R}^3)}^2.$$

Therefore

$$\begin{aligned} (3.14) \quad & \|A_{V_{\circ,\text{good}},s} F\|_{L^2(\mathbb{R}^3)}^2 \leq \sum_{C \in \vec{\mathcal{C}}_{\text{good}}} \|A_{V_C,s} F\|_{L^2(\mathbb{R}^3)}^2 \\ & = \sum_{C \in \vec{\mathcal{C}}_{\text{good}}} \|A_{V_C,s} F_{R_C}\|_{L^2(\mathbb{R}^3)}^2 \\ & \leq K_{\frac{N}{E}} \sum_{C \in \vec{\mathcal{C}}_{\text{good}}} \|F_{R_C}\|_{L^2(\mathbb{R}^3)}^2 \\ & \leq K_{\frac{N}{E}} \left(\sup_{\beta \in \mathbb{S}^2} \sum_{C \in \vec{\mathcal{C}}_{\text{good}}} \mathbf{1}_{R_C}(\beta) \right) \|F\|_{L^2(\mathbb{R}^3)}^2 \\ & \leq \Theta E K_{\frac{N}{E}} \|F\|_{L^2(\mathbb{R}^3)}^2. \end{aligned}$$

Collecting (3.6), (3.8) with estimates (3.10), (3.11), and (3.14) completes the proof of (3.5). $\square$

Proof that (3.5) implies Proposition 3.3. Observe that an application of (3.5) with $E \simeq N/(\log N)^3$ gives the recursive estimate

$$\frac{K_N}{N} \leq \Theta + \Theta \frac{K_{(\log N)^3}}{(\log N)^3}$$

for some numerical constant $\Theta > 0$. Iterating we get for any integer $k \geq 1$ that

$$\frac{K_N}{N} \lesssim \Theta^k + \Theta^k \frac{K_{(\log^{[k]} N)^3}}{(\log^{[k]} N)^3}.$$

However for any integer $M \geq 1$ we have the trivial estimate $K_M \lesssim M^2$ which together with the recursive estimate above gives

$$K_N \lesssim \Theta^k N + \Theta^k (\log^{[k]} N)^3 N$$

and completes the proof. $\square$

4. Notions from algebraic geometry, projection varieties, and polynomial partition. In this section, we turn to setting up the algebraic geometry definitions and tools needed in the proof of Theorem A, beyond the notion of transverse complete intersection already introduced in Subsection 2.9. We will also prove the Proposition 2.10, namely the version of polynomial partition used throughout the paper. A short summary that can be used in parallel with reading the proof of Theorem A is the following:

§4.1 definition and basic notions concerning m -dimensional real algebraic varieties in $\mathbb{R}^n$;

§4.3 approximation of m -dimensional real algebraic varieties with transverse complete intersections; approximate projection of a transverse complete intersection on a lower dimensional subspace;

§4.7 approximate polynomial partitioning in $\mathbb{R}^n$ and proof of Proposition 2.10.

4.1. Real algebraic varieties and dimension. For us a real algebraic variety in $\mathbb{R}^n$ is a set

$$\mathbf{Z}(I) := \{x \in \mathbb{R}^n : P(x) = 0 \quad \forall P \in I\}$$

where $I \subset \mathbb{R}[x_1, \dots, x_n]$ is an ideal. We possibly write $\mathbf{Z}(I) = \mathbf{Z}(P_1, \dots, P_J)$ if $\{P_1, \dots, P_J\} \subset \mathbb{R}[x_1, \dots, x_n]$ is a basis (finite generating set) for the ideal I .

Count and degree. Following the approach of [38, Section 4], we write $\text{ct}\mathbf{Z}(I)$ to denote the *count* of $\mathbf{Z}(I)$, that is least cardinality of a basis of the ideal I , and $\text{deg}\mathbf{Z}(I)$ to denote the least integer D such that there exist polynomials $P_1, \dots, P_J$ with $\text{deg} P_j \leq D$ for all j , that generate the ideal I .

Dimension. The dimension $\dim Z$ of a real algebraic variety $Z = \mathbf{Z}(I)$ is defined as in [9, Def. 2.8.11]; see also [43]. Let $\{P_1, \dots, P_J\} \subset \mathbb{R}[x_1, \dots, x_n]$ be a basis for the ideal I . If $\mathbf{Z}(I)$ has dimension $1 \leq m \leq n$, we say that $x \in \mathbf{Z}(I)$ is a *smooth point* if

$$\text{rank}(\mathbf{D}\vec{P})(x) = n - m,$$

where $D\vec{P}$ is the Jacobian matrix of the vector valued function $\vec{P}: \mathbb{R}^n \rightarrow \mathbb{R}^J$, namely the $J \times n$ matrix whose j th row is the vector $\nabla P_j = (\partial_{x_1} P_j, \dots, \partial_{x_n} P_j)$. We say that $Z = \mathbf{Z}(I)$ is a *smooth algebraic variety* of dimension $1 \leq m \leq n$ if every point $x \in Z$ is a smooth point in dimension m . The set of smooth points of an m -dimensional algebraic variety in $\mathbb{R}^n$ will be denoted by $Z^{\circ m}$. A special important case of smooth m -dimensional algebraic varieties in $\mathbb{R}^n$ are the transverse complete intersections $Z = \mathbf{Z}(P_1, \dots, P_{n-m})$ defined in §2.9.

We conclude our discussion of dimension of algebraic varieties by recalling the following standard result; see [38, Theorem 4.8] for a proof.

LEMMA 4.2. (smooth components) *Fix positive integers J, D and $m \leq n$. There exists a constant $\Theta = \Theta_{J,D,m,n}$ such that the following holds. Let $Z = \mathbf{Z}(I) \subset \mathbb{R}^n$ be a real algebraic variety with*

$$\text{ct } Z = J, \quad \deg Z = D, \quad \dim Z = m.$$

Then $Z' := Z \setminus Z^{\circ m}$ is a real algebraic variety with $\text{ct } W \leq \Theta, \deg W \leq \Theta$, and $\dim W < m$.

Classes of algebraic varieties. Finally, it will be convenient to us to introduce pieces of notation for the classes of algebraic varieties involved in our theorems, namely:

◊ $\mathcal{Z}_{m,n}(D, J)$ is the collection of real algebraic varieties Z in $\mathbb{R}^n$ with $\dim Z = m, \deg Z \leq D, \text{ct } Z \leq J$.

◊ $\mathcal{Z}_{m,n}^{\times}(D)$ is the collection of transverse complete intersections Z in $\mathbb{R}^n$ with $\dim Z = m, \deg Z \leq D$. This is the collection appearing in the quantitative estimate (1.3) of Theorem A.

4.3. Approximation by and projection of transverse complete intersections. The ultimate aim of this subsection is to show that any real algebraic variety of dimension m can be approximated in a suitable sense by a controlled number of transverse complete intersections of dimension up to m and of controlled degree.

We begin with detailing a continuity property of nonsingular zero sets, which are semialgebraic sets; see [9]. It is convenient to introduce some pieces of notation before the statement. Given $1 \leq c \leq n$ and polynomials $\{P_j : 1 \leq j \leq c\} \subset \mathbb{R}[x_1, \dots, x_n]$ we write $\vec{P} = (P_1, \dots, P_c): \mathbb{R}^n \rightarrow \mathbb{R}^c$. Let τ be an indexing of the $\binom{n}{c}$ choices of $c \times c$ minors $D_{\tau} \vec{P}$ of $D\vec{P}$ and for each τ we write $\Delta_{\tau} \vec{P} := \det D_{\tau} \vec{P}$, which is a polynomial of degree at most D^c . In the statements below the notion of distance $\text{dist}(U, V)$ is as in §2.1.

LEMMA 4.4. *Let $P_1, \dots, P_c \in \mathbb{R}[x_1, \dots, x_n]$ and $Z = \mathbf{Z}(P_1, \dots, P_c)$. For $\rho, R > 0$, let*

$$U_{\rho,R} := \left\{ x \in Z : |x| < R, \max_{\tau} |\Delta_{\tau} \vec{P}| \geq \rho \right\}.$$

For $\vec{\alpha} = (\alpha_1, \dots, \alpha_c) \in \mathbb{R}^c$, $\vec{\beta} = (\beta_1, \dots, \beta_{n-1}) \in \mathbb{R}^{n-1}$ define the algebraic varieties

$$V_{\vec{\alpha}} := \mathbf{Z}(P_1 + \alpha_1, \dots, P_c + \alpha_c), \quad W_{\vec{\beta}} := \mathbf{Z}(P_{1,\vec{\beta}}, P_2, \dots, P_c),$$

where

$$P_{1,\vec{\beta}}(x_1, \dots, x_n) := P_1(x_1 + \beta_1 x_n, \dots, x_{n-1} + \beta_{n-1} x_n, x_n).$$

Then for all $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, \rho, R, \vec{P}) > 0$ such that

$$(4.1) \quad |\vec{\alpha}| + |\vec{\beta}| < \delta \implies \text{dist}(U_{\rho,R}, V_{\vec{\alpha}}) + \text{dist}(U_{\rho,R}, W_{\vec{\beta}}) < \varepsilon.$$

In addition, if Z is a transverse complete intersection, so are $V_{\vec{\alpha}}$ and $W_{\vec{\beta}}$.

Proof. We prove the approximation claim for $W_{\vec{\beta}}$, a similar but simpler proof works for $V_{\vec{\alpha}}$. Pick a point $\bar{x} \in U_{\rho,R}$. By symmetry, suppose that the $c \times c$ matrix $A := \{\partial_{x_j} P_k(\bar{x}) : 1 \leq j, k \leq c\}$ is invertible and $|\det A| \geq \rho$. Let $F = F(x, \vec{\beta}) : \mathbb{R}^n \times \mathbb{R}^{n-1} \rightarrow \mathbb{R}^c$ be the vector valued function with $F_1(x, \vec{\beta}) := R_{1,\vec{\beta}}(x)$ and $F_k(x) := P_k(x)$ for $2 \leq k \leq c$. Then A is a $c \times c$ invertible minor of $DF(\bar{x}, 0)$. By the implicit function theorem, for $1 \leq j \leq c$ we find smooth functions $h_j = h_j(x_{c+1}, \dots, x_n, \vec{\beta})$ defined for $|\vec{\beta}| + \sup_{j>c} |x_j - \bar{x}_j|$ sufficiently small such that

$$h_j(\bar{x}_{c+1}, \dots, \bar{x}_n, \vec{0}_{\mathbb{R}^{n-1}}) = \bar{x}_j, \\ R_{1,\vec{\beta}}(h_1(x_{c+1}, \dots, x_n, \vec{\beta}), \dots, h_c(x_{c+1}, \dots, x_n, \vec{\beta}), x_{c+1}, \dots, x_n) = 0.$$

Then (4.1) follows from the above display and standard estimates on the L^∞ -norms of the derivatives of h_j , obtained from the lower bound on $|\det A|$. The claim that if Z is a transverse complete intersection so are $V_{\vec{\alpha}}$, $W_{\vec{\beta}}$, is actually a byproduct of the application of the implicit function theorem above. $\square$

Using Lemma 4.4 in combination with the smooth component Lemma 4.2 we obtain the following approximation result by transverse complete intersections.

PROPOSITION 4.5. *Let $Z = \mathbf{Z}(I)$ be an algebraic variety in $\mathbb{R}^n$ of dimension $0 \leq m \leq n$, degree D , and $\{P_1, \dots, P_J\}$ be a generating set for I . Let $\varepsilon > 0$, $R \geq 1$. Then there exists a set $Z_0 \subseteq Z$ with $\#Z_0 \leq \Theta$, and collections of transverse complete intersections $\mathcal{Z}_\mu(Z) := \{Z_{\mu,\vartheta} : 1 \leq \vartheta \leq \Theta\} \subset \mathcal{Z}_{\mu,n}^\times(\Theta)$, such that*

$$\text{dist} \left(Z \cap \mathcal{A}_n(R), Z_0 \cup \bigcup_{\mu=1}^m \bigcup_{\vartheta=1}^{\Theta} Z_{\mu,\vartheta} \right) < \varepsilon.$$

The constant $\Theta = \Theta_{m,n,J,D}$ depends only on the indicated parameters and can be explicitly computed.

Proof. We prove the proposition by induction on the dimension m . If Z has dimension zero, then $Z \cap \mathcal{A}_n(R)$ is a finite set with $\#Z \leq \Theta$ points, whence the claim holds with $Z_0 = Z$ and empty collections $\mathcal{Z}_\mu$. If Z has dimension n , there is nothing to prove.

We now deal with the inductive step for which some preliminary notation is needed. Let $1 \leq m \leq n-1$, Z and $\{P_1, \dots, P_J\}$ be as in the statement of the lemma. Let $c = n - m$ and σ be a choice of increasing c -tuple of indices $\{j_1, \dots, j_c\} \subset \{1, \dots, J\}$; there are $\binom{J}{c}$ choices. Let $\vec{P}_\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^c$ be the vector-valued function with $(\vec{P}_\sigma)_k := P_{j_k}$ and $\Delta_\tau \vec{P}_\sigma = \det D_\tau \vec{P}_\sigma$, as in the statement of Lemma 4.4. Calling

$$Z' := Z \cap \mathbf{Z}(\{\Delta_\tau \vec{P}_\sigma : \sigma, \tau\}),$$

that is Z' is the complement in Z of the set of smooth points Z^{om} , we gather from Lemma 4.2 that $Z' \in \mathcal{Z}_{\bar{\mu}, n}(\Theta, \Theta)$ for some $0 \leq \bar{\mu} < m$. We now partition

$$Z \cap \mathcal{A}_n(R) = U \cup W$$

where W is the intersection of $Z \cap \mathcal{A}_n(R)$ with the $\varepsilon/2$ neighborhood of Z' . By the inductive assumption applied to $Z' \cap \mathcal{A}_n(R)$ with $\varepsilon/2$ in place of ε , we may find Z_0 with $\#Z_0 \leq \Theta$ and families $\mathcal{Z}_\mu(Z') \subset \mathcal{Z}_{\mu, n}(\Theta)$ for $1 \leq \mu \leq \bar{\mu}$ each with at most Θ elements so that

$$(4.2) \quad \text{dist} \left(W \cap \mathcal{A}_n(R), Z_0 \cup \bigcup_{\mu=1}^m \bigcup_{\vartheta=1}^{\Theta} Z_{\mu, \vartheta} \right) < \varepsilon.$$

We are left with treating the U -component. Notice that, as we have excised an $\varepsilon/2$ -neighborhood of $\mathbf{Z}(\{\Delta_\tau \vec{P}_\sigma : \sigma, \tau\})$, there exists $\rho > 0$ such that

$$U \subset \bigcup_{\sigma} U^\sigma, \quad U^\sigma := \left\{ x \in \mathbf{Z}(\vec{P}_\sigma) : |x| \leq 2R, \max_{\tau} \min_{x \in U} |\Delta_\tau \vec{P}_\sigma(x)| \geq \rho \right\}.$$

Now we may apply Lemma 4.4 to $\vec{P} = \vec{P}_\sigma$, with U^σ in place of $U_{\rho, 2R}$ and obtain that

$$\text{dist}(U_\sigma, V_{\sigma, \vec{\alpha}}) < \varepsilon, \quad V_{\sigma, \vec{\alpha}} := \mathbf{Z}(P_{\sigma(1)} + \alpha_1, \dots, P_{\sigma(c)} + \alpha_c),$$

provided $|\vec{\alpha}| < \delta(\varepsilon, \rho)$ is sufficiently small. We also note that a recursive application of [27, Lemma 5.1] tells us that the variety $V_{\sigma, \vec{\alpha}}$ is a transverse complete intersection in $\mathcal{Z}_{m, n}^\times(D)$ for almost all $\vec{\alpha} \in \mathbb{R}^c$. Choosing one such $\vec{\alpha}$ for each σ , we obtain

$$(4.3) \quad \text{dist} \left(U, \bigcup_{\sigma} V_{\sigma, \vec{\alpha}} \right) < \varepsilon.$$

Combining (4.2) with (4.3) completes the inductive step. $\square$

The next lemma will be used in the course of the proof of Theorem A to control the contribution to our maximal averaging operator of (a piece of) a m -dimensional transverse complete intersection lying s -close to the $n - 1$ dimensional hyperplane $x_n = 0$, by a m -dimensional algebraic variety *lying on* the hyperplane $x_n = 0$.

LEMMA 4.6. (approximate projection of a transverse complete intersection)
Let

$$Z = \mathbf{Z}(P_1, \dots, P_{n-m})$$

be a transverse complete intersection in $\mathbb{R}^n$ with $\deg Z = D$. Let

$$U := Z \cap \{x \in \mathbb{R}^n : 1 \leq |x| < 2, |x_n| < s\}$$

for some $0 < s < \frac{1}{2}$. There exists an algebraic variety W with the following properties:

- (i) $W \subset e_n^\perp := \{x \in \mathbb{R}^n : x_n = 0\}$;
- (ii) $\dim W \leq m$;
- (iii) $\text{ct } W + \deg W \leq \Theta_{n,m,D}$;
- (iv) $\text{dist}(U, W) < 2s$.

Proof. We identify $e_n^\perp$ with $\mathbb{R}^{n-1}$. The proof includes a technical reduction which we first include in our assumptions, postponing its justification until the end of the argument.

Claim. The ideal $I \subset \mathbb{R}[x_1, \dots, x_n]$ generated by $P_1, \dots, P_{n-m}$ contains a polynomial f of degree d where the monomial x_n^d appears with nonzero coefficient, for some $d \geq 1$.

We proceed assuming the claim. Define

$$\widehat{Z} := \{z \in \mathbb{C}^n : P(z) = 0 \ \forall P \in I\}.$$

Observe that $\widehat{Z}$ is a complex variety of dimension m which is invariant under complex conjugation and $Z = \widehat{Z} \cap \mathbb{R}^n$. Let $\Pi: \mathbb{C}^n \rightarrow \mathbb{C}^{n-1}$ indicate the canonical projection erasing the last coordinate. By virtue of our claim we may apply [17, Theorem 1.68] (see also [34, Theorem 3.3]) and obtain that

$$\widehat{W} := \Pi \widehat{Z} = \{z \in \mathbb{C}^{n-1} : P(z) = 0 \ \forall P \in I'\},$$

where I' is the first elimination ideal of I , namely $I' = I \cap \mathbb{R}[x_1, \dots, x_{n-1}]$. Such a $\widehat{W}$ is a complex algebraic variety of dimension m . We may find a generating set of for I' by arguing as follows. According to the results of Dubé [23] and Latyshev [33], we may pick a Gröbner basis for I , which we denote $\{Q_j : 1 \leq j \leq J\}$, $Q_j \in \mathbb{R}[x_1, \dots, x_n]$, whose degree $\Delta := \max \deg Q_j$, and cardinality J , is bounded by a constant depending only on D, n, m . Using the elimination theorem [16, Chap. 2,

Theorem 3], a (Gröbner) basis for I' is given by those Q_j which do not depend on the coordinate x_n : this completes the claim about the generating set of I' . For definitions and constructive algorithms leading to Gröbner bases we send to the monographs [15, 16] and references therein.

At this point we denote $W := \widehat{W} \cap \mathbb{R}^{n-1}$ and W satisfies (i) by construction. It is immediate to verify that W contains the projection of Z on the last $n-1$ coordinates, whence the claim (iv) follows; in fact the stronger $\text{dist}(U, W) < s$ holds. Furthermore, the fact that W is a real algebraic variety of dimension $\leq m$, namely claim (ii), follows by a straightforward application of [34, Lemma 2.1]. As $W = \mathbf{Z}(I')$, the upper bound of claim (iii) is a consequence of the above construction of a generating set for I' . This last assertion completes the proof of the lemma up to the verification of the preliminary claim, which follows next.

We make sure that the claim holds by perturbing Z slightly. For $\delta_1, \dots, \delta_{n-1} \in \mathbb{R}$ consider the polynomial

$$f(x) := P_1(x_1 + \delta_1 x_n, \dots, x_{n-1} + \delta_{n-1} x_n, x_n).$$

As noticed in [34, Lemma 3.4], for all $\delta > 0$ there exist $\delta_1, \dots, \delta_{n-1}$ with $\max |\delta_j| < \delta$ such that f satisfies the condition required in the claim. We take advantage of Lemma 4.4 and obtain that for a suitable choice of δ_j

$$\tilde{Z} := \mathbf{Z}(f, P_2, \dots, P_{n-m})$$

is a transverse complete intersection satisfying the claim and $\text{dist}(U, \tilde{Z}) < 2^{-10}s$. In particular we may find a set $\tilde{U} \subset \tilde{Z} \cap \{x \in \mathbb{R}^n : 1 - 2^{-8} \leq |x| < 2 + 2^{-8}, |x_n| < (1 + 2^{-8})s\}$ such that $\text{dist}(U, \tilde{U}) < 2^{-10}s$. Applying the above proof to $\tilde{U}$ in place of U with the slightly different value of s completes the reduction and therefore the proof of the Lemma. $\square$

4.7. Approximate polynomial partitioning on $\mathbb{R}^n$ and proof of Proposition 2.10. We now begin the proof of Proposition 2.10. One of the main tools is the following approximate polynomial partitioning theorem on $\mathbb{R}^n$ which can be seen as the adaptation to our setting of the Guth-Katz polynomial partitioning theorem [28] and of its hypersurface refinement due to Zahl [43]. These results are appealed to in the proof.

LEMMA 4.8. (Approximate polynomial partitioning on $\mathbb{R}^n$) *Let $V \subset \mathbb{R}^n$ be a finite point set with $\#V \leq N^n$ and $\delta > 0$. There exists a polynomial $Q \in \mathbb{R}[x_1, \dots, x_n]$ of degree $\leq \Theta_n E$ and $\varepsilon_0 > 0$ such that*

(1) *Each of the $\leq \Theta_n E^n$ connected components of $\mathbb{R}^n \setminus \mathbf{Z}(Q)$ contains at most at most $(N/E)^n$ points of V .*

(2) *If $|\varepsilon| < \varepsilon_0$, then $\text{dist}(V \cap \mathbf{Z}(Q), \mathbf{Z}(Q + \varepsilon)) < \delta$.*

Proof. We use the polynomial partitioning theorem of Guth and Katz, [28, Theorem 4.1], with the improvement of Zahl from [44, Corollary 2.3]. The theorem of Guth and Katz yields the existence of $Q \in \mathbb{R}[x_1, \dots, x_n]$ of degree $\leq \Theta_n E$ such that each of the $\Theta_n E^n$ connected components of $\mathbb{R}^n \setminus \mathbf{Z}(Q)$ contains at most $(N/E)^n$ points. Now the improvement of Zahl implies that we can assume that every irreducible component $\tilde{Q}$ of Q satisfies $\dim(\tilde{Q}) = n - 1$ and $\nabla \tilde{Q}$ does not vanish identically on $\mathbf{Z}(\tilde{Q})$. In particular the latter condition implies that $\mathbf{Z}(Q)$ is $(n - 1)$ -dimensional and smooth points are dense in $\mathbf{Z}(Q)$. Therefore for all $v \in V \cap \mathbf{Z}(Q)$ we may find a smooth point $w(v) \in \mathbf{Z}(Q)$ with $|v - w(v)| < 2^{-1}\delta$. By Lemma 4.4 and a compactness argument there exists $\varepsilon_0 > 0$ such that for all $|\varepsilon| < \varepsilon_0$

$$\sup_{v \in V \cap \mathbf{Z}(Q)} \text{dist}(w(v), \mathbf{Z}(Q + \varepsilon)) < \frac{\delta}{2}.$$

The above estimate together with the construction of $w(v)$ prove the second claim. $\square$

The proof proper of Proposition 2.10 is articulated into several steps.

Step 1. (reduction to coordinate patches) This step is an adaptation of [27, Subsection 5.3] by Guth. For an element g of the Grassmanian $\Lambda^m \mathbb{R}^n$ define the polynomial

$$P_g(x) := \nabla P_1(x) \wedge \cdots \wedge \nabla P_{n-m}(x) \wedge g$$

which has degree $\leq \Theta_{n,m} D$. It is proved in [27, Lemma 5.6] that one may choose a $\frac{1}{100}$ -net $G \subset \Lambda^m \mathbb{R}^n$ so that, for every $g \in G$, we have that

$$W_g := \mathbf{Z}(P_1, \dots, P_{n-m}, P_g)$$

is a transverse complete intersection of dimension $m - 1$ and degree $\leq \Theta_{n,m} D$. Before moving further we operate a first excision from V , namely we set

$$V_G := \left\{ v \in V : \inf_{g \in G} \text{dist}(v, W_g) < \delta \right\}.$$

Clearly, if $v \in V \setminus V_G$, then v belongs to a connected component $O \in \mathcal{O}$ of the set $Z \setminus \cup \{Z_g : g \in G\}$. Notice that there are at most $\Theta_{n,m} D^n$ elements in $\mathcal{O}$. We thus partition

$$V \setminus V_G = \bigcup_{O \in \mathcal{O}} V_O, \quad V_O := V \cap O.$$

Step 2. (partition of each coordinate patch) Fix one such connected component $O \in \mathcal{O}$. Proceeding exactly like in [27, Lemma 5.6] we notice that O is the graph of a Lipschitz map $h : O' \rightarrow \mathbb{R}^{n-m}$ with Lipschitz constant $< \frac{1}{10}$, where O' is an

open connected subset of $\mathbb{R}^m$. By rotational invariance we may identify $\mathbb{R}^m \equiv \text{span}\{e_1, \dots, e_m\}$ and write

$$O := \{(y_1, \dots, y_m, h(y_1, \dots, y_m)) : (y_1, \dots, y_m) \in O'\}.$$

Let $\Pi : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\Pi(x_1, \dots, x_n) := (x_1, \dots, x_m)$; polynomials of $\mathbb{R}[x_1, \dots, x_m]$ will be identified with a corresponding polynomial in $\mathbb{R}[x_1, \dots, x_n]$ by precomposing with Π without explicit mention.

Observe that Π is injective on O and set $V'_O := \Pi V_O \subset O' \subset \mathbb{R}^m$. We may apply Lemma 4.8 to the set V'_O , with m in place of n and $\delta/2$ in place of δ , in order to find a polynomial $Q_O \in \mathbb{R}[x_1, \dots, x_m]$ with the following properties: Q_O partitions O' into $\leq \Theta_m E^m$ cells C' , the connected components of $O' \setminus \mathbf{Z}(Q_O)$, with the property that each C' contains at most $(N/E)^m$ points of V'_O ; and $V'_O \cap \mathbf{Z}(Q_O)$ is contained in the $(\delta/2)$ -neighborhood of $\mathbf{Z}(Q_O + \varepsilon)$ in $\mathbb{R}^m$ for all sufficiently small ε . By [27, Lemma 5.1] we may choose ε such that the above property holds and

$$(4.4) \quad W_O := \mathbf{Z}(P_1, \dots, P_m, Q_O + \varepsilon)$$

is a transverse complete intersection. It follows that the $\leq \Theta_m E^m$ cells

$$C := \{(y_1, \dots, y_m, h(y_1, \dots, y_m)) : (y_1, \dots, y_m) \in C'\} \subset O,$$

which we group into $C \in \vec{C}_O$, are the connected components of

$$O \setminus \mathbf{Z}(P_1, \dots, P_{n-m}, Q_O)$$

and the sets

$$(4.5) \quad V_C := V_O \cap C$$

have at most $(N/E)^m$ elements. It remains to take care of those points $v \in V_O$ such that $\Pi v \in \mathbf{Z}(Q_O)$. We call this set $V_{O, \times}$. By the above, we may pick $w' = (w_1, \dots, w_m) \in \mathbf{Z}(Q_O + \varepsilon)$ such that $|\Pi v - w| < \delta/2$. Notice that the point $w = (w', h(w'))$ belongs to W_O defined in (4.4) and satisfies $|v - w| \leq |\Pi v - w'| + |h(\Pi v) - h(w')| < \delta$ using the Lipschitz constant of h . We have proved that

$$\sup_{v \in V_{\times, O}} \text{dist}(v, W_O) < \delta.$$

Step 3. (definition of $V_{\times}$) By the constructions in Steps 1 and 2, the set

$$V_{\times} := V_G \cup \bigcup_{O \in \mathcal{O}} V_{\times, O}$$

satisfies the required claim, and the collection of transverse complete intersections W_j of the statement is obtained by putting together the $\leq \Theta_{m,n}$ transverse complete intersections $\{W_g : g \in G\}$ of Step 1 with the $\leq \Theta_{m,n} D^n$ transverse complete intersections $\{W_O : O \in \mathcal{O}\}$.

Step 4. (definition of $V_\circ$ and counting of the hyperplane crossings) We define

$$\vec{C} := \bigcup \{\vec{C}_O : O \in \mathcal{O}\}.$$

By the construction in Step 2, the sets V_C defined in (4.5) exhaust $V \setminus V_\times$ and satisfy the cardinality requirement. We are left with counting how many of the intersections $\{C \cap \{x \in \mathbb{R}^n : \xi \cdot x = a\} : C \in \vec{C}\}$ are nontrivial. As we have at most $\leq \Theta_m D^n$ elements $O \in \mathcal{O}$, it suffices to show that for almost every $\xi \in \mathbb{R}^n$, $a \in \mathbb{R}$ the intersection $\{C \cap \{x \in \mathbb{R}^n : \xi \cdot x = a\}$ is nontrivial for at most $\leq \Theta_{n,m} D^n E^{m-1}$ connected components $C \in \vec{C}_O$.

Fixing one such O , we may go back to the coordinate system where O is a graph over the first m variables. Consider $P_{\xi,a}(x_1, \dots, x_m) = \Pi \xi \cdot x - a$ as an element of $\mathbb{R}[x_1, \dots, x_m]$; we prove the claim for all $\xi \in \mathbb{R}^n$ such that $Z_\xi := \mathbf{Z}(P_1, \dots, P_{n-m}, P_{\xi,a})$ is a transverse complete intersection of dimension $m-1$, as these ξ are a set whose complement is a null set in $\mathbb{R}^n$.

We first count how many $C \in \vec{C}_O$ satisfy the condition

$$C \cap \mathbf{Z}(P_{\xi,a}) \neq \emptyset, \quad \partial C \cap \mathbf{Z}(P_{\xi,a}) = \emptyset.$$

This number is controlled by the number of connected components of the set $\mathbf{Z}(P_{\xi,a}) \setminus Z$, which is controlled by $\Theta_n D^{n-1}$: to see this apply [38, Theorem 4.11], which is a reformulation of the main result by Barone and Basu [4]. Now we count how many $C \in \vec{C}_O$ satisfy instead the condition

$$C \cap \mathbf{Z}(P_{\xi,a}) \neq \emptyset, \quad \partial C \cap \mathbf{Z}(P_{\xi,a}) \neq \emptyset.$$

This number is controlled by the number of connected components of the set

$$\mathbf{Z}(P_{\xi,a}, P_1, \dots, P_m) \setminus \mathbf{Z}(Q_O)$$

which is controlled by $\Theta_{n,m} D^n E^{m-1}$: to see this apply again [4] in the form of [38, Theorem 4.11], with $U = \mathbf{Z}(P_{\xi,a}, P_1, \dots, P_m)$, which has dimension $m-1$, and $W = \mathbf{Z}(Q_O)$, which has degree E . Putting together the two counting arguments above completes the last claim of the second point of Proposition 2.10 and therefore the proof of the proposition.

5. Proof of Theorem A.

5.1. Main line of proof. Both statements in the theorem will be a consequence of the following uniform estimate: for all $1 \leq m < n$, $D, J \geq 1$, and $\eta > 0$, there exists a constant $\Theta = \Theta_{m,n,D,J,\eta}$ such that

$$(5.1) \quad \begin{aligned} K_{m,n,D,J}(N) &:= \sup_{s>0} \sup_{Z \in \mathcal{Z}_{m,n}(D,J)} \sup_{\substack{V \subset Z \cap \mathcal{A}_n(\frac{3}{2}) \\ \#V \leq N^m}} \|A_{V,s} \circ S_1\|_{L^2(\mathbb{R}^n)}^2 \\ &\leq \Theta N^{m-1+\eta}, \end{aligned}$$

where $\mathcal{Z}_{m,n}(D, J)$ is the class of real algebraic varieties introduced in Subsection 4.1. Notice the slightly enlarged annulus, which is for technical reasons. This reduction was made using Proposition 2.5 to insert the annular cutoff. The proof of (5.1) will be inductive, and the constants we will induct on are the following; for $R \in \{1, 2\}$, $N \geq 1$

$$\begin{aligned} K_{m,n,D}^\times(N, R) &:= \sup_{s>0} \sup_{Z \in \mathcal{Z}_{m,n}^\times(D)} \sup_{\substack{V \subset Z \cap \mathcal{A}_n(R) \\ \#V \leq N^m}} \|A_{V,s} \circ S_1\|_{L^2(\mathbb{R}^n)}^2, \\ K_{m,n,D}^\times(N) &:= K_{m,n,D}^\times(N, 1). \end{aligned}$$

The approximation Proposition 4.2 allows us to relate $K_{m,n,D,J}(N)$ to $K_{\mu,n,\Theta}^\times(N)$ for $1 \leq \mu \leq m$ and some suitable constant $\Theta = \Theta_{m,n,J,D}$ and ultimately induct on the latter quantity only.

LEMMA 5.2. *There exists a constant $\Theta = \Theta_{m,n,J,D}$ such that*

$$K_{m,n,D,J}(N) \leq \Theta \log N \sup_{1 \leq \mu \leq m} K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}} \right).$$

Proof. Throughout the constant $\Theta = \Theta_{m,n,D,J}$ is understood to depend only on the indicated parameters and may vary from line to line. For every positive integer M the inequality

$$(5.2) \quad K_{m,n,D}^\times(M, 2) \lesssim K_{m,n,D}^\times(M)$$

is obvious by rescaling. So it suffices to prove that

$$K_{m,n,D,J}(N) \leq \Theta \log N \sup_{1 \leq \mu \leq m} K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}}, 2 \right).$$

To do so, by virtue of Corollary 2.7, it is enough to estimate operator norms on the left-hand sides of (5.1) for $0 < s < 2^{-5}$. Fix such an s and $Z \in \mathcal{Z}_{m,n}(D, J)$. Notice that by a straightforward application of Proposition 4.5, we may find that there exist $\Theta = \Theta_{m,n,D,J}$ and a set $Z_0 \subset Z$ with $\#Z_0 \leq \Theta$ and collections of transverse

complete intersections $\mathcal{Z}_\mu(Z) := \{Z_{\mu,\vartheta} : 1 \leq \vartheta \leq \Theta\} \subset \mathcal{Z}_{\mu,n}^\times(\Theta)$ such that

$$\text{dist} \left(Z \cap \mathcal{A}_n(3/2), Z_0 \cup \bigcup_{\mu=1}^m \bigcup_{\vartheta=1}^{\Theta} Z_{\mu,\vartheta} \right) < 2^{-10} s.$$

Applying Lemma 2.6 we have

$$\begin{aligned} \sup_{\substack{V \subset Z \cap \mathcal{A}_n(3/2) \\ \#V \leq N^m}} \|A_{V,s} \circ S_1\|_{L^2(\mathbb{R}^n)}^2 &\leq \Theta \sup_{1 \leq \mu \leq m} \sup_{1 \leq \vartheta \leq \Theta} \sup_{\substack{V \subset Z_{\mu,\vartheta} \cap \mathcal{A}_n(2) \\ \#V \leq N^m}} \|A_V\|_{L^2(\mathbb{R}^n)}^2 \\ &\leq \Theta \log N \left(\sup_{1 \leq \mu \leq m} K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}}, 2 \right) \right) \end{aligned}$$

where in the last step we applied Proposition 2.5. The claim follows by taking supremum over s and $Z \in \mathcal{Z}_{m,n}(D, J)$. $\square$

We now show how Lemma 5.2 reduces (5.1) to the following statement; here $1 \leq m < n$.

THE $(\text{TCI})_{(m,n)}$ ESTIMATE. *For all $D \geq 1$ and for all $\eta > 0$ there exists a constant $\Upsilon_{m,n,D,\eta}$ such that*

$$(\text{TCI})_{(m,n)} \quad K_{m,n,D}^\times(N) \leq \Upsilon_{m,n,D,\eta} N^{m-1+\eta}.$$

Proof that the $(\text{TCI})_{(\mu,n)}$ estimate for all $1 \leq \mu \leq m$ implies (5.1). Fix $D, J \geq 1$. Applying Lemma 5.2, we may find a constant $\Theta = \Theta_{m,n,D,J}$ such that

$$\begin{aligned} K_{m,n,D,J}(N) &\leq \Theta \log N \left(\sup_{1 \leq \mu \leq m} K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}} \right) \right) \\ &\leq \frac{\Theta N^{\frac{\eta}{2}}}{\eta} \left(\sup_{1 \leq \mu \leq m} K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}} \right) \right). \end{aligned}$$

Using the estimate $(\text{TCI})_{(m,n)}$ for $\frac{\eta}{2} > 0$, we have that

$$K_{m,n,\Theta}^\times(N) \leq \Upsilon_{m,n,\Theta,\frac{\eta}{2}} N^{m-1+\frac{\eta}{2}}$$

while using the $(\text{TCI})_{(\mu,n)}$ estimate with $(m-\mu)/m$ in place of η when $\mu < m$, we have that

$$K_{\mu,n,\Theta}^\times \left(N^{\frac{m}{\mu}} \right) \leq \Upsilon_{\mu,n,\Theta,(m-\mu)/m} N^{m-1}.$$

Putting together the last three displays completes the proof of the implication. $\square$

Summarizing, we have reduced Theorem A to proving $(\text{TCI})_{(m,n)}$ for all pairs (m,n) with $1 \leq m < n$. This will be done by the induction scheme

$$(m-1, n) \wedge \bigwedge_{\mu=1}^m (\mu, n-1) \implies (m, n), \quad 1 < m < n-1,$$

while using the two base cases $(1, n)$ and $(n-1, n)$ as seeds of our induction. The main inductive step is summarized in the following estimate which is a consequence of polynomial partitioning.

THE $(\text{PART})_{(m,n)}$ ESTIMATE. *Let $1 < m < n-1$ and $D \geq 1$. For all integers E with $E^m \geq D^n$, there holds*

$$\begin{aligned} K_{m,n,D}^\times(N) &\leq \Theta_{m,n} D^{2n} \left[E^{m-1} K_{m,n,D}^\times \left(\frac{N}{E} \right) \right. \\ (\text{PART})_{(m,n)} \quad &+ (\log N) K_{m-1,n,\Theta_{m,n}E}^\times \left(N^{\frac{m}{m-1}} \right) \\ &\left. + E(\log N)^2 \sup_{1 \leq \mu \leq m} K_{\mu,n-1,\Theta_{m,n,D}}^\times \left(N^{\frac{m}{\mu}} \right) \right]. \end{aligned}$$

We detail the two base cases in Subsection 5.3 and 5.5 and prove the partitioning estimate $(\text{PART})_{(m,n)}$ in the final Subsection 5.6. The main line of proof of Theorem A will be complete once we establish the implication

$$(\text{PART})_{(m,n)} \wedge (\text{TCI})_{(m-1,n)} \wedge \bigwedge_{\mu=1}^m (\text{TCI})_{(\mu,n-1)} \implies (\text{TCI})_{(m,n)}$$

for all $1 < m < n-1$.

Proof of the implication. Fix $D \geq 1, \eta > 0$. The goal is to find a uniform in N bound for

$$\tau(N) := \frac{K_{m,n,D}^\times(N)}{N^{m-1+\eta}}.$$

For each $E \geq 1$ we may use the induction assumptions $(\text{TCI})_{(m-1,n)}$, $(\text{TCI})_{(\mu,n-1)}$ for $1 \leq \mu \leq m$ to get the estimates

$$\begin{aligned} \frac{K_{m-1,n,\Theta_{m,n}E}^\times(N^{\frac{m}{m-1}})}{N^{m-1}} &\leq \Upsilon_{m-1,n,\Theta_{m,n}E,\frac{1}{m}}, \\ \frac{K_{\mu,n-1,\Theta_{m,n,D}}^\times(N^{\frac{m}{\mu}})}{N^{m-1}} &\leq \Upsilon_{\mu,n-1,\Theta_{m,n,D},\frac{m-\mu}{m}}, \quad 1 \leq \mu \leq m-1, \\ \frac{K_{m,n-1,\Theta_{m,n,D}}^\times(N)}{N^{m-1}} &\leq \Upsilon_{m,n-1,\Theta_{m,n,D},\frac{\eta}{2}} N^{\frac{\eta}{2}}. \end{aligned}$$

By virtue of these estimates and of the elementary bound $\log N \leq \frac{4}{\eta} N^{\frac{\eta}{4}}$, dividing $(\text{PART})_{(m,n)}$ by $N^{m-1+\eta}$ yields that for all $E^m > D^n$

$$(5.3) \quad \begin{aligned} \tau(N) &\leq \frac{\Theta_{m,n} D^{2n}}{E^\eta} \tau\left(\frac{N}{E}\right) \\ &+ \frac{\Theta_{m,n} D^{2n}}{\eta^2} \left[\Upsilon_{m-1,n,\Theta_{m,n}E,\frac{1}{m}} \right. \\ &\quad \left. + \Upsilon_{m,n-1,\Theta_{m,n},D,\frac{\eta}{2}} + \sup_{1 \leq \mu \leq m-1} \Upsilon_{\mu,n-1,\Theta_{m,n},D,\frac{m-\mu}{m}} \right]. \end{aligned}$$

Now choosing

$$\begin{aligned} E = E(m,n,D,\eta) &:= \left\lceil (2\Theta_{m,n} D^{2n})^{\frac{1}{\eta}} \right\rceil, \\ \Upsilon_{m,n,D,\eta} &:= \frac{2\Theta_{m,n} D^{2n}}{\eta^2} \left[\Upsilon_{m-1,n,\Theta_{m,n}E,\frac{1}{m}} + \Upsilon_{m,n-1,\Theta_{m,n},D,\frac{\eta}{2}} \right. \\ &\quad \left. + \sup_{1 \leq \mu \leq m-1} \Upsilon_{\mu,n-1,\Theta_{m,n},D,\frac{m-\mu}{m}} \right], \end{aligned}$$

estimate (5.3) becomes

$$(5.4) \quad \tau(N) - \frac{1}{2} \tau\left(\frac{N}{E}\right) \leq \frac{\Upsilon_{m,n,D,\eta}}{2},$$

and induction on N of (5.4) yields the claimed $(\text{TCI})_{(m,n)}$ for this choice of D, η . $\square$

5.3. The $(\text{TCI})_{(1,n)}$ estimate. To prove the $(\text{TCI})_{(1,n)}$ estimate, we fix a transverse complete intersection $Z = \mathbf{Z}(P_1, \dots, P_{n-1}) \in \mathcal{Z}_{1,n}^\times(D)$. Note that in particular Z is the union of at most D^n smooth connected curves $\mathcal{V}_j = \{\gamma_j(t) : t \in I_j\} \subset \mathbb{R}^n$ where $I_j \subset \mathbb{R}$ are parameter intervals. Now for each $\xi \in \mathbb{R}^n$, $a \in \mathbb{R}$ the function $t \mapsto \xi \cdot \gamma_j(t) - a$ may change sign at most $O(D_j)$ times with $\sum_j D_j \leq D$. The latter property allows us to appeal to an estimate of Córdoba from [13] which can be summarized in the following calculation. If $\mathcal{V}_j \cap V = \{v_{j,1}, \dots, v_{j,n_j}\}$ where for each j the directions are in consecutive order then

$$\begin{aligned} \|A_{Z \cap V, s} f\|_{L^2(\mathbb{R}^n)} &\leq \|A_{Z \cap \tilde{V}, s} f\|_{L^2(\mathbb{R}^n)} \\ &+ \left(\sum_{j=1}^{D^n} \sum_{\ell=1}^{n_j-1} \|(A_{v_{j,\ell+1}, s} - A_{v_{j,\ell}, s}) f\|_{L^2(\mathbb{R}^n)}^2 \right)^{\frac{1}{2}} \\ &+ \left(\sum_{j=1}^{D^n} \|(A_{v_{j+1,1}, s} - A_{v_{j,n_j}, s}) f\|_{L^2(\mathbb{R}^n)}^2 \right)^{\frac{1}{2}} \end{aligned}$$

where $\#\tilde{V} \leq \#V/2$. The first square function is estimated as in the proof of the good part in Theorem D by a constant multiple of $\|\psi\|_{\text{BV}} D^{\frac{1}{2}} \|f\|_{L^2(\mathbb{R}^n)}$, using the D_j crossing property of each $\mathcal{V}_j$; here we remember that ψ is the smooth function used to define $A_{v,s}$. The square function of the second summand is estimated just by triangle inequality and is bounded by a constant multiple of $D^{\frac{n}{2}} \|f\|_{L^2(\mathbb{R}^n)}$. Thus induction on $\#V$ yields the bound $\|A_{Z \cap V, s}\|_{L^2(\mathbb{R}^n)} \lesssim_n D^{\frac{n}{2}} \log(\#V \cap Z)$. Note that the argument above is essentially identical to the one used for the good part in the proof of Theorem D; here however we can afford implicit losses depending on D . The estimate described above is summarized in the following proposition.

PROPOSITION 5.4. *Let $\mathcal{V} = \{\gamma(t) : t \in I\} \subset \mathbb{R}^n$ be the image of a smooth curve γ with the property that for each $\xi \in \mathbb{R}^n$, $a \in \mathbb{R}$, the function $t \mapsto \xi \cdot \gamma(t) - a$ changes sign at most D times. Then*

$$\sup_{\substack{V \subset \mathcal{V} \cap \mathcal{A}_n(1) \\ \#V \leq N}} \|A_V\|_{L^2(\mathbb{R}^n)} \leq \Theta_n D^{\frac{n}{2}} \log N.$$

In consequence of this proposition we have

$$K_{1,n,D}^\times(N) \leq \Theta_n D^{2n} (\log N)^2 \leq \frac{\Theta_n D^{2n}}{\eta^2} N^\eta$$

for all $\eta > 0$, which complies with the claimed $(\text{TCI})_{(1,n)}$ estimate.

5.5. The $(\text{TCI})_{(n-1,n)}$ estimate. For the codimension 1 case, we first observe that whenever $V \subset \mathcal{A}_n(1)$, denoting $V' = \{v' : v \in V\} \subset \mathbb{S}^{n-1}$ we have

$$\|M_V\|_{L^2(\mathbb{R}^n)} \leq 2 \|M_{V'}\|_{L^2(\mathbb{R}^n)}.$$

By virtue of the above display and of Lemma 2.3, we obtain that

$$(5.5) \quad K_{n-1,n,D}^\times(N) \leq \Theta \sup_{0 < s < 2^{-5}} \sup_{\substack{V \subset \mathbb{S}^{n-1} \\ \#V \leq N^{n-1}}} \|A_{V,s} \circ S_1\|_{L^2(\mathbb{R}^n)}^2.$$

We know, respectively from [30] and from Theorem B with $k = 1$ (say), that the supremum in the right-hand side of (5.5) is bounded by $\Theta \log N \leq \frac{\Theta}{\eta} N^\eta$ when $n = 2$ and by $\Theta N \log N \leq \frac{\Theta}{\eta} \Theta N^{1+\eta}$ when $n = 3$. These observations deal with the cases $n = 2, 3$ of $(\text{TCI})_{(n-1,n)}$.

We now estimate (likely not optimally) that the right-hand side of (5.5) is controlled by $\Theta_n N^{n-2} (\log N)^{n-2}$ when $n \geq 4$, completing the proof of the $(\text{TCI})_{(n-1,n)}$ estimate in all cases. Fix $0 < s < 2^{-5}$ and take a $2^{-10}s$ -net in Ω and for each $\xi \in \Omega$ we define the cluster $V_\xi = V \cap R_{\xi,3s}$. We say that a cluster is bad if $\#V_\xi \geq N^{n-2}$. By a greedy selection algorithm we can identify a set $\Omega_{\text{bad}} \subset \Omega$ having at most N elements, and disjoint sets $W_\xi \subset V_\xi$ with the property that,

defining,

$$V_{\text{bad}} := \bigcup_{\xi \in \Omega_{\text{bad}}} V_{\xi} = \bigcup_{\xi \in \Omega_{\text{bad}}} W_{\xi}, \quad V_{\text{good}} := V \setminus \bigcup_{\xi \in \Omega_{\text{bad}}} V_{\xi},$$

there holds

$$\sup_{\xi \in \Omega} \#(V_{\text{good}}) \cap R_{\xi, s} \leq N^{n-2}.$$

An easy overlap estimate on the Fourier support of the multipliers $\{A_{v, s} : v \in V_{\text{good}}\}$ then leads to the bound

$$\|A_{V_{\text{good}}, s} \circ S_1 f\|_{L^2(\mathbb{R}^n)}^2 \leq \Theta_n N^{n-2} \|f\|_2^2.$$

For the directions in $V \setminus V_{\text{good}}$ we use an $n-2$ dimensional estimate for each W_{ξ} , as W_{ξ} is $3s$ -close to a set $U_{\xi} \subset \mathbb{S}^{n-1} \cap \xi^{\perp} \equiv \mathbb{S}^{n-2}$ with $\#W_{\xi} = \#U_{\xi}$. Arguing inductively with the help of Proposition 2.5, we know

$$\|A_{U_{\xi}}\|_{L^2(\mathbb{R}^{n-1})}^2 \leq \Theta_n (\#U_{\xi})^{\frac{n-3}{n-2}} (\log N)^{n-2},$$

one more log factor than the inductive estimate we assumed for $A_{U_{\xi}, s} \circ S_1$. In fact, applying Corollary 2.8,

$$\begin{aligned} & \|A_{V_{\text{bad}}, s} \circ S_1 f\|_{L^2(\mathbb{R}^n)}^2 \\ & \leq \Theta_n \sum_{\xi \in \Omega_{\text{bad}}} \|A_{W_{\xi}, s} \circ S_1 f\|_{L^2(\mathbb{R}^n)}^2 \\ & \leq \Theta_n \left(\sum_{\xi \in \Omega_{\text{bad}}} \|A_{U_{\xi}}\|_{L^2(\mathbb{R}^{n-1})}^2 \right) \|S_1 f\|_2^2 \\ & \leq \Theta_n (\log N)^{n-2} \left(\sum_{\xi \in \Omega_{\text{bad}}} (\#U_{\xi})^{\frac{n-3}{n-2}} \right) \|S_1 f\|_2^2 \\ & \leq \Theta_n N^{n-3} (\log N)^{n-2} \left(\sum_{k=1}^{\log N} 2^{k \frac{n-3}{n-2}} \#\{\xi \in \Omega_{\text{bad}} : \#U_{\xi} \sim 2^k N^{n-2}\} \right) \|S_1 f\|_2^2 \\ & \leq \Theta_n (N \log N)^{n-2} \|S_1 f\|_2^2 \end{aligned}$$

as $\#\{\xi \in \Omega_{\text{bad}} : 2^{k-1} N^{n-2} < \#U_{\xi} \leq 2^k N^{n-2}\} \leq 2^{-k} N$.

Note that if we apply the approach above for directions on $\mathbb{S}^2$ we get the bound $N^{\frac{1}{2}} (\log N)^{\frac{1}{2}}$ for the single annulus estimate, which has been proved in [19] with similar ideas. This is worse than the bound proved in Proposition 3.3 where the factor of $\sqrt{\log N}$ is replaced by an arbitrary iterated logarithm.

5.6. Proof of the (PART)_(m,n) estimate. Let $D \geq 1$, $E^m \geq D^n$ and

$$Z = \mathbf{Z}(P_1, \dots, P_{n-m}) \in \mathcal{Z}_{m,n}^\times(D)$$

be a TCI. By Corollary 2.7 it is enough to argue for $0 < s < 2^{-5}$. Fix a subset $V \subset Z \cap \mathcal{A}_n(1)$ with $\#V \leq N^m$. To estimate the maximal operator $A_{V,s} \circ S_1$, we will partition V using Proposition 2.10, choosing $\delta = 2^{-10}s$. We obtain the decomposition

$$V = V_\circ \cup V_\times$$

with the following properties. To begin with, we may find $\leq \Theta_{m,n} D^n$ transverse complete intersections $W_j = \mathbf{Z}(P_1, \dots, P_{n-m}, Q_j)$ of dimension $m-1$ and degree $\leq \Theta_{m,n} E$ such that

$$(5.6) \quad \sup_{v \in V_\times} \inf_j \text{dist}(v, W_j) < 2^{-10}s.$$

Moreover, there exist $\leq \Theta_{m,n} D^n E^m$ disjoint connected subsets of Z , denoted by $\mathbf{C} \in \tilde{\mathcal{C}}$, with the property that

$$(5.7) \quad V_\circ = \bigcup_{\mathbf{C} \in \tilde{\mathcal{C}}} V_{\mathbf{C}}, \quad V_{\mathbf{C}} := V_\circ \cap \mathbf{C}, \quad \#V_{\mathbf{C}} \leq \left(\frac{N}{E}\right)^m,$$

and such that for almost every $\xi \in \mathbb{R}^n$ and every $a \in \mathbb{R}$, there holds

$$(5.8) \quad \#\{\mathbf{C} \in \tilde{\mathcal{C}} : \mathbf{C} \cap \{x \in \mathbb{R}^n : \xi \cdot x = a\} \neq \emptyset\} \leq \Theta_{m,n} D^{2n} E^{m-1}.$$

Controlling $V_\times$. In a similar way to what was done in the proof of Theorem B, we approximate directions from $V_\times$ with directions from the $\leq \Theta_{m,n} D^n$ transverse complete intersections W_j which are lower-dimensional. To wit, we use (5.6) to find subsets $V_j \subset W_j$ with $\#V_j \leq N^m$ so that for each $v \in V_\times$ there exists j and $w(v) \in V_j$ with $|v - w(v)| < 2^{-10}s$. In particular $V_j \subset \mathcal{A}_n(2)$. Applying the approximation of Lemma 2.6, followed by Proposition 2.5, we deduce that

$$(5.9) \quad \begin{aligned} \|A_{V_\times, s} \circ S_1 f\|_2^2 &\leq \Theta_{m,n} D^n \left(\sup_j \|A_{V_j}\|_{L^2(\mathbb{R}^n)}^2 \right) \|S_1 f\|_2^2 \\ &\leq \Theta_{m,n} D^n K_{m-1, n, \Theta_{m,n} E}^\times (N^{\frac{m}{m-1}}, 2) \log N \|S_1 f\|_2^2 \\ &\leq \Theta_{m,n} D^n K_{m-1, n, \Theta_{m,n} E}^\times (N^{\frac{m}{m-1}}) \log N \|S_1 f\|_2^2 \end{aligned}$$

as each $W_j \in \mathcal{Z}_{m-1, n}^\times(\Theta_{m,n} E)$, and using inequality (5.2) in the last step. This completes the control of the $V_\times$ component of V .

Splitting V_o into good and bad clusters. Using that $\vec{C}$ is a finite set, we may choose a $2^{-10}s$ -net Ω on $\mathbb{S}^{n-1}$ such that the bounded crossing property (5.8) holds with $a = \pm 3s$ for all $\xi \in \Omega$. For $\xi \in \Omega$ we define $k_\xi \subset \vec{C}$ to be a cluster of cells with top $\xi \in \Omega$ if

$$C \subset R_{\xi, 3s} \quad \forall C \in k_\xi.$$

We say that a cluster k_ξ is bad if it contains more than E^{m-1} cells. As $\#\vec{C} \leq \Theta_{m,n} D^n E^m$, by an iterative selection algorithm, we may construct $\Omega_{\text{bad}} \subset \Omega$ with $\#\Omega_{\text{bad}} \leq \Theta_{m,n} D^n E$ such that setting

$$\vec{C}_{\text{good}} := \vec{C} \setminus \bigcup_{\xi \in \Omega_{\text{bad}}} k_\xi,$$

there holds

$$\#\left\{C \in \vec{C}_{\text{good}} : C \subset R_{\xi, 3s}\right\} \leq E^{m-1} \quad \forall \xi \in \Omega.$$

We split accordingly

$$V_\xi := \bigcup_{C \in k_\xi} V_C, \quad V_{o, \text{bad}} := \bigcup_{\xi \in \Omega_{\text{bad}}} V_\xi, \quad V_{o, \text{good}} := \bigcup_{C \in \vec{C}_{\text{good}}} V_C.$$

Controlling $V_{o, \text{bad}}$. We still aim to control the contribution of the $\leq \Theta_{m,n} D^n E$ bad clusters with a lower dimensional estimate, replacing the $\leq N^m$ vectors V_ξ with a set U_ξ consisting of $\leq N^m$ vectors from the $(n-1)$ -dimensional hyperplane $\xi^\perp$. This step is more difficult than the analogous one from the proof of Theorem B because we need to make sure that the obtained U_ξ is contained in an m -dimensional algebraic variety on $\xi^\perp$ of controlled degree and count. This is not in general possible by simply taking projections and using Lemma 2.8 as the projection of an algebraic variety is in general a semi-algebraic set [9].

We tackle this difficulty by applying the approximate projection Lemma 4.6, after a rotation taking ξ to e_n , with choice of U given by $U := Z \cap \mathcal{A}_n(1) \cap R_{\xi, 3s}$. Notice that the latter set contains V_ξ . Then Lemma 4.6 yields the existence of an algebraic variety $W_\xi \in \mathcal{Z}_{m, n-1}(\Theta_{m,n,D}, \Theta_{m,n,D})$ contained in $\xi^\perp$ such that for every $v \in V_\xi$ we may find $u(v) \in W_\xi$ with $|v - u(v)| \leq 6s$. Define $U_\xi = \{u(v) : v \in V_\xi\} \subset W_\xi \cap \mathcal{A}_{n-1}(3/2)$, and observe that $\#U_\xi \leq N^m$. Using that there are at most $\Theta_{m,n} D^n E$ bad clusters together with the approximation Lemma 2.6 in the first step, Fubini's theorem in the second, Proposition 2.5 in the third, we obtain

the chain of inequalities

$$\begin{aligned}
 & \|A_{V_{\circ, \text{bad}}, s} \circ S_1 f\|_2^2 \\
 & \leq \Theta_{m,n} D^n E \left(\sup_{\xi \in \Omega_{\text{bad}}} \|A_{U_\xi}\|_{L^2(\mathbb{R}^n)}^2 \right) \|S_1 f\|_2^2 \\
 & \leq \Theta_{m,n} D^n E \left(\sup_{\xi \in \Omega_{\text{bad}}} \|A_{U_\xi}\|_{L^2(\mathbb{R}^{n-1})}^2 \right) \|S_1 f\|_2^2 \\
 (5.10) \quad & \leq \Theta_{m,n} D^n E \log N \left(\sup_{\xi \in \Omega_{\text{bad}}} \sup_{\sigma > 0} \|A_{U_{\xi, \sigma}} \circ S_1\|_{L^2(\mathbb{R}^{n-1})}^2 \right) \|S_1 f\|_2^2 \\
 & \leq \Theta_{m,n} D^n E \log N \cdot K_{m,n-1, \Theta_{m,n,D}, \Theta_{m,n,D}}(N) \|S_1 f\|_2^2 \\
 & \leq \Theta_{m,n} D^n E (\log N)^2 \left(\sup_{1 \leq \mu \leq m} K_{\mu, n-1, \Theta_{m,n,D}}^\times \left(N^{\frac{m}{\mu}} \right) \right) \|S_1 f\|_2^2,
 \end{aligned}$$

where we used Lemma 5.2 in the very last step. This completes the treatment of $V_{\circ, \text{bad}}$.

Controlling $V_{\circ, \text{good}}$. This estimate is not much different to the one appearing in the corresponding proof of Theorem B. We keep when possible the same notations. For $C \in \vec{\mathcal{C}}_{\text{good}}$ we define R_C as in (3.12) and again notice that $A_{v,s} f = A_{v,s} f_{R_C}$ whenever $v \in C$. The next step is the estimation of the overlap of the sets $\{R_C : C \in \vec{\mathcal{C}}_{\text{good}}\}$ at a generic $\beta \in \mathbb{R}^n$. By homogeneity in the definition of $R_{v,s}$ it is enough to take $\beta \in \mathbb{S}^{n-1}$ and approximate by $\xi \in \Omega$ with $|\xi - \beta| < 2^{-10}s$. Then the overlap of the multipliers 1_{R_C} at β is bounded by the same right-hand side of (3.13). The removal of the bad clusters when constructing $\mathcal{C}_{\text{good}}$ ensured the property

$$\#\{C \in \vec{\mathcal{C}}_{\text{good}} : C \subset R_{\xi, 3s}\} \leq E^{m-1}$$

so we are left with counting

$$\begin{aligned}
 & \#\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap R_{\xi, 2s} \neq \emptyset, C \not\subset R_{\xi, 3s}\} \\
 & \leq \#\{C \in \vec{\mathcal{C}}_{\text{good}} : C \cap \{x \in \mathbb{R}^2 : \xi \cdot x = \pm 3s\} \neq \emptyset\} \leq \Theta_{m,n} D^{2n} E^{m-1},
 \end{aligned}$$

where (5.8) has been taken into account. Therefore arguing exactly like in (3.14) and using that V_C contains at most $(N/E)^m$ directions, see (5.7), we obtain

$$(5.11) \quad \|A_{V_{\circ, \text{good}}, s} f\|_2^2 \leq \Theta_{n,m} D^{2n} E^{m-1} K_{m,n,D}^\times (N/E) \|f\|_2^2.$$

The sought after $(\text{PART})_{(m,n)}$ estimate for these values of m, n, D, N, E follows by collecting (5.9), (5.10), and (5.11).

DEPARTMENT OF MATHEMATICS, WASHINGTON UNIVERSITY IN ST. LOUIS,
1 BROOKINGS DR, ST. LOUIS, MO 63130

E-mail: francesco.diplinio@wustl.edu

DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DEL PAÍS VASCO, APTDO. 644,
48080 BILBAO, SPAIN AND IKERBASQUE, BASQUE FOUNDATION FOR SCIENCE, BIL-
BAO, SPAIN

E-mail: ioannis.parissis@ehu.es

REFERENCES

- [1] A. Alfonseca, Strong type inequalities and an almost-orthogonality principle for families of maximal operators along directions in $\mathbb{R}^2$, *J. London Math. Soc. (2)* **67** (2003), no. 1, 208–218.
- [2] A. Alfonseca, F. Soria, and A. Vargas, An almost-orthogonality principle in L^2 for directional maximal functions, *Harmonic Analysis at Mount Holyoke (South Hadley, MA, 2001)*, *Contemp. Math.*, vol. 320, Amer. Math. Soc., Providence, RI, 2003, pp. 1–7.
- [3] ———, A remark on maximal operators along directions in $\mathbb{R}^2$, *Math. Res. Lett.* **10** (2003), no. 1, 41–49.
- [4] S. Barone and S. Basu, Refined bounds on the number of connected components of sign conditions on a variety, *Discrete Comput. Geom.* **47** (2012), no. 3, 577–597.
- [5] J. Barrionuevo, Averages along uniformly distributed directions on a curve, *Proc. Amer. Math. Soc.* **119** (1993), no. 3, 823–827.
- [6] ———, Estimates for some Kakeya-type maximal operators, *Trans. Amer. Math. Soc.* **335** (1993), no. 2, 667–682.
- [7] S. Basu and M. Sombra, Polynomial partitioning on varieties of codimension two and point-hypersurface incidences in four dimensions, *Discrete Comput. Geom.* **55** (2016), no. 1, 158–184.
- [8] M. Bateman, Kakeya sets and directional maximal operators in the plane, *Duke Math. J.* **147** (2009), no. 1, 55–77.
- [9] J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry. Translated from the 1987 French original. Revised by the authors*, *Ergeb. Math. Grenzgeb. (3)*, vol. 36, Springer-Verlag, Berlin, 1998.
- [10] A. Carbery, Differentiation in lacunary directions and an extension of the Marcinkiewicz multiplier theorem, *Ann. Inst. Fourier (Grenoble)* **38** (1988), no. 1, 157–168.
- [11] M. Christ, J. Duoandikoetxea, and J. L. Rubio de Francia, Maximal operators related to the Radon transform and the Calderón-Zygmund method of rotations, *Duke Math. J.* **53** (1986), no. 1, 189–209.
- [12] A. Córdoba, The Kakeya maximal function and the spherical summation multipliers, *Amer. J. Math.* **99** (1977), no. 1, 1–22.
- [13] A. Córdoba, Geometric Fourier analysis, *Ann. Inst. Fourier (Grenoble)* **32** (1982), no. 3, vii, 215–226.
- [14] A. Córdoba and R. Fefferman, On differentiation of integrals, *Proc. Nat. Acad. Sci. U.S.A.* **74** (1977), no. 6, 2211–2213.
- [15] D. A. Cox, J. Little, and D. O’Shea, *Using Algebraic Geometry*, 2nd ed., *Grad. Texts in Math.*, vol. 185, Springer-Verlag, Berlin, 2005.
- [16] ———, *Ideals, Varieties, and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 4th ed., *Undergrad. Texts Math.*, Springer-Verlag, Cham, 2015.
- [17] W. Decker and G. Pfister, *A First Course in Computational Algebraic Geometry*, *AIMS Libr. Ser.*, Cambridge University Press, Cambridge, 2013.
- [18] C. Demeter, Singular integrals along N directions in $\mathbb{R}^2$, *Proc. Amer. Math. Soc.* **138** (2010), no. 12, 4433–4442.
- [19] ———, L^2 bounds for a Kakeya-type maximal operator in $\mathbb{R}^3$, *Bull. Lond. Math. Soc.* **44** (2012), no. 4, 716–728.

- [20] F. Di Plinio, S. Guo, C. Thiele, and P. Zorin-Kranich, Square functions for bi-Lipschitz maps and directional operators, *J. Funct. Anal.* **275** (2018), no. 8, 2015–2058.
- [21] F. Di Plinio and I. Parissis, A sharp estimate for the Hilbert transform along finite order lacunary sets of directions, *Israel J. Math.* **227** (2018), no. 1, 189–214.
- [22] Francesco Di Plinio and Ioannis Parissis, On the maximal directional Hilbert transform in three dimensions, *Int. Math. Res. Not. IMRN* **2020** (2020), no. 14, 4324–4356.
- [23] T. W. Dubé, The structure of polynomial ideals and Gröbner bases, *SIAM J. Comput.* **19** (1990), no. 4, 750–775.
- [24] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl, A semi-algebraic version of Zarankiewicz’s problem, *J. Eur. Math. Soc. (JEMS)* **19** (2017), no. 6, 1785–1810.
- [25] L. Grafakos, P. Honzik, and A. Seeger, On maximal functions for Mikhlin-Hörmander multipliers, *Adv. Math.* **204** (2006), no. 2, 363–378.
- [26] L. Guth, A restriction estimate using polynomial partitioning, *J. Amer. Math. Soc.* **29** (2016), no. 2, 371–413.
- [27] ———, Restriction estimates using polynomial partitioning II, *Acta Math.* **221** (2018), no. 1, 81–142.
- [28] L. Guth and N. H. Katz, On the Erdős distinct distances problem in the plane, *Ann. of Math. (2)* **181** (2015), no. 1, 155–190.
- [29] N. H. Katz, Maximal operators over arbitrary sets of directions, *Duke Math. J.* **97** (1999), no. 1, 67–79.
- [30] ———, Remarks on maximal operators over arbitrary sets of directions, *Bull. London Math. Soc.* **31** (1999), no. 6, 700–710.
- [31] N. H. Katz and K. M. Rogers, On the polynomial Wolff axioms, *Geom. Funct. Anal.* **28** (2018), no. 6, 1706–1716.
- [32] M. T. Lacey and X. Li, Maximal theorems for the directional Hilbert transform on the plane, *Trans. Amer. Math. Soc.* **358** (2006), no. 9, 4099–4117.
- [33] V. N. Latyshev, A combinatorial complexity of Gröbner bases. Algebra, 13, *J. Math. Sci. (New York)* **102** (2000), no. 3, 4134–4138.
- [34] J. Matoušek and Z. Patáková, Multilevel polynomial partitions and simplified range searching, *Discrete Comput. Geom.* **54** (2015), no. 1, 22–41.
- [35] A. Nagel, E. M. Stein, and S. Wainger, Differentiation in lacunary directions, *Proc. Nat. Acad. Sci. U.S.A.* **75** (1978), no. 3, 1060–1062.
- [36] Y. Ou and H. Wang, A cone restriction estimate using polynomial partitioning, preprint, <https://arxiv.org/abs/1704.05485>.
- [37] J. Parcet and K. M. Rogers, Directional maximal operators and lacunarity in higher dimensions, *Amer. J. Math.* **137** (2015), no. 6, 1535–1557.
- [38] A. Sheffer, Polynomial methods and incidence theory, book in progress, <http://faculty.baruch.cuny.edu/ASheffer/>.
- [39] P. Sjögren and P. Sjölin, Littlewood-Paley decompositions and Fourier multipliers with singularities on certain sets, *Ann. Inst. Fourier (Grenoble)* **31** (1981), no. 1, vii, 157–175.
- [40] J.-O. Strömberg, Weak estimates on maximal functions with rectangles in certain directions, *Ark. Mat.* **15** (1977), no. 2, 229–240.
- [41] ———, Maximal functions associated to rectangles with uniformly distributed directions, *Ann. of Math. (2)* **107** (1978), no. 2, 399–402.
- [42] T. Tao, The Bochner-Riesz conjecture implies the restriction conjecture, *Duke Math. J.* **96** (1999), no. 2, 363–375.
- [43] J. Zahl, An improved bound on the number of point-surface incidences in three dimensions, *Contrib. Discrete Math.* **8** (2013), no. 1, 100–121.
- [44] ———, A Szemerédi-Trotter type theorem in $\mathbb{R}^4$, *Discrete Comput. Geom.* **54** (2015), no. 3, 513–572.