Ars Inveniendi Analytica (2024), Paper No. 2, 53 pp.
DOI 10.15781/md28-ws10
ISSN: 2769-8505

Singular integrals along variable codimension one subspaces

Odysseas Bakas
University of Patras

Francesco Di Plinio
Università di Napoli "Federico II"

Ioannis Parissis
Universidad del País Vasco and Ikerbasque

Luz Roncal
Basque Center for Applied Mathematics, Universidad del País Vasco and Ikerbasque

Communicated by Monica Visan

Abstract. *This article deals with maximal operators on $\mathbb{R}^n$ formed by taking arbitrary rotations of tensor products of a d-dimensional Hörmander–Mihlin multiplier with the identity in $n-d$ coordinates, in the particular* codimension 1 *case $d = n-1$. These maximal operators are naturally connected to differentiation problems and maximally modulated singular integrals such as Sjölin's generalization of Carleson's maximal operator. Our main result, a weak-type $L^2(\mathbb{R}^n)$-estimate on band-limited functions, leads to several corollaries. The first is a sharp $L^2(\mathbb{R}^n)$ estimate for the maximal operator restricted to a finite set of rotations in terms of the cardinality of the finite set. The second is a version of the Carleson–Sjölin theorem. In addition, we obtain that functions in the Besov space $B^0_{p,1}(\mathbb{R}^n)$, $2 \leqslant p < \infty$, may be recovered from their averages along a measurable choice of codimension 1 subspaces, a form of Zygmund's conjecture in general dimension n.*

Keywords. Directional operators, Zygmund's conjecture, Stein's conjecture, maximally rotated singular integrals, time-frequency analysis.

1. Introduction

Let $n \geqslant 2$ be the linear dimension of the Euclidean space $\mathbb{R}^n$ endowed with the canonical basis $\{e_1, \dots, e_n\}$ and let $d \in \{1, \dots, n-1\}$. Singular integrals in d variables embed in $\mathbb{R}^n$ when made to act trivially in the $n-d$ variables perpendicular to a chosen d-dimensional subspace. More specifically, each pair $(U, \mathscr{K})$ of orientation preserving rotation $U \in \mathrm{SO}(n)$ and tempered distribution $\mathscr{K} \in \mathscr{S}'(\mathbb{R}^d)$ induces a map $T : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ by

$$\langle Tf, g\rangle := \int_{\mathbb{R}^{n-d}} \big\langle f \circ U(\cdot, y) * \mathscr{K}, g \circ U(\cdot, y)\big\rangle \, \mathrm{d}y = \big\langle f \circ U * (\mathscr{K} \otimes \delta), g \circ U\big\rangle, \tag{1.1}$$

with $\delta \in \mathscr{S}'(\mathbb{R}^{n-d})$ denoting the Dirac distribution supported at $0 \in \mathbb{R}^{n-d}$. If the Fourier transform of $\mathscr{K}$ is a bounded function $m \in L^\infty(\mathbb{R}^d)$ then Tf coincides with the continuous bounded function

$$Tf(x) = \int_{\mathbb{R}^n} m(V\Pi_\sigma \xi) \widehat{f}(\xi) e^{2\pi i \langle x, \xi\rangle} \, \mathrm{d}\xi, \qquad x \in \mathbb{R}^n, \tag{1.2}$$

where $V = U^{-1}$, σ is the preimage through V of the subspace $\mathbb{R}^d := \mathrm{span}\{e_1, \dots, e_d\}$, and Π_σ denotes orthogonal projection on σ. Our interest lies in the further particular case where m is a Hörmander–Mihlin multiplier on $\mathbb{R}^d$, namely

$$\|m\|_{\mathscr{M}_A(d)} := \sup_{0 \leqslant |\alpha| \leqslant A} \sup_{\eta \in \mathbb{R}^d} |\eta|^{|\alpha|} \left| D^\alpha m(\eta) \right| \tag{1.3}$$

is finite. If $A \geqslant d+1$ say, an application of the Hörmander–Mihlin theorem on each σ-fiber of $\mathbb{R}^n$ and Fubini's theorem tell us that (1.2) extends to a bounded operator on $L^p(\mathbb{R}^n), 1 < p < \infty$.

The broad object of interest of this article are maximal versions of (1.2) where the rotation $V \in \mathrm{SO}(n)$ depends measurably, or in some other specific fashion, on the point $x \in \mathbb{R}^n$. Note that the multiplier (1.2) is allowed to exhibit singularities on the $(n-d)$-dimensional subspace $\sigma^\perp = (V^{-1}\mathbb{R}^d)^\perp$. In light of this fact and of the invariance of the Mihlin norms (1.3), an efficient parametrization of the multipliers (1.2) must involve the corresponding oriented Grassmanian $\mathbf{Gr}(d, n)$, namely the space of oriented d-dimensional subspaces of $\mathbb{R}^n$. When equipped with the canonical distance, $\mathbf{Gr}(d, n)$ is a compact metric space which is isometrically isomorphic to the quotient group

$$\frac{\mathrm{SO}(n)}{\mathrm{SO}(d) \otimes \mathrm{SO}(n-d)}.$$

The description of the isomorphism is useful to us and can be loosely given as follows. Thinking again $\mathbb{R}^d = \mathrm{span}\{e_1, \dots, e_d\} \in \mathbf{Gr}(d, n)$, a subspace $\sigma \in \mathbf{Gr}(d, n)$ is identified with the class of matrices $\mathscr{V}(\sigma) \subset \mathrm{SO}(n)$ which map σ to $\mathbb{R}^d$. Any two elements $V, O \in \mathscr{V}(\sigma)$ are related by

$$V = (Q \otimes P)O, \qquad Q \in \mathrm{SO}(d), \quad P \in \mathrm{SO}(n-d). \tag{1.4}$$

For this reason, fixing in the remainder of our discussion a family

$$\{O_\sigma \in \mathscr{V}(\sigma) : \sigma \in \mathbf{Gr}(d, n)\} \tag{1.5}$$

of representatives, the class of rotated singular integrals (1.2) associated to a single multiplier $m \in \mathscr{M}_A(d)$ may be parametrized as

$$T_m f(x;\sigma,Q) := \int_{\mathbb{R}^n} m(QO_\sigma \Pi_\sigma \xi)\widehat{f}(\xi) e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \qquad x\in\mathbb{R}^n, \tag{1.6}$$

with parameters $\sigma \in \mathbf{Gr}(d,n)$ and $Q \in \mathrm{SO}(d)$. Note that (1.6) does not depend on the component $P \in \mathrm{SO}(n-d)$ in (1.4), reflecting the trivial action of T in the $\sigma^\perp$-variables, cf. (1.2). Also note that the family in (1.5) can be chosen arbitrarily. A specific family is explicitly constructed in Lemma 2.1 with the additional property that the map $\sigma \mapsto O_\sigma$ is C^1; see Remark 2.4. Finally, the directional operator (1.6) may be further interpreted by rewriting formally

$$T_m f(x;\sigma,Q) = \int_{\mathbb{R}^d} f(x - V^{-1}t)\,K(t)\mathrm{d}t, \qquad V = QO_\sigma, \quad x\in\mathbb{R}^n, \tag{1.7}$$

with $K(t)$ the kernel of $\mathscr{K} = \widehat{m}$.

The maps (1.6), or alternatively (1.7), are termed $\mathbf{Gr}(d,n)$*-subspace singular integral operators*. For instance, they encompass σ-subspace averaging operators, corresponding to the choice $\mathscr{K} \in L^1(\mathbb{R}^d)$ in (1.1), as well as the d-dimensional Riesz transform in the directions of σ, obtained by choosing the vector distribution

$$\mathscr{K}(y) := \mathrm{p.v.}\frac{y}{|y|^{d+1}} \in \mathscr{S}'(\mathbb{R}^d)^d.$$

Subspace singular integrals, along *variable choices* of d-dimensional subspaces σ are connected to several deep lines of investigations in harmonic analysis and partial differential equations. This article focuses on the *codimension 1* case $n = d+1$. Higher codimension cases are more singular and very few instances of operators have been treated, none beyond the L^2 threshold. A full account of the history of this circle of problems is postponed to the end of this introduction.

Below, we consider the maximal operator associated to a family of multipliers $\mathbf{m} = \{m_\sigma \in \mathscr{M}_A(d) : \sigma \in \mathbf{Gr}(d,n)\}$,

$$T^\star_{\mathbf{m}} f(x) := \sup_{\sigma\in\mathbf{Gr}(d,n)} \sup_{Q\in\mathrm{SO(d)}} \big|T_{m_\sigma} f(x;\sigma,Q)\big|, \qquad x\in\mathbb{R}^n, \tag{1.8}$$

or linearized variants thereof where the variable choices of σ, Q are expressed by measurable functions $\sigma(\cdot):\mathbb{R}^n \to \mathbf{Gr}(d,n)$ and $Q(\cdot):\mathbb{R}^n \to \mathrm{SO}(d)$, possibly under additional regularity assumptions on these functions. When $d=1$, the operators $T^\star_{\mathbf{m}}$ recover the familiar directional maximal averaging operator, and the maximal directional Hilbert transform, for suitable choices of symbols $\mathbf{m}$. Concerning (1.8), its maximal nature shows how the choice of the family O_σ in (1.5) is immaterial.

Our main result concerns the action of (1.8) on frequency band-limited functions. In order to state it we fix a smooth radial function $\zeta:\mathbb{R}^n \to \mathbb{R}_+$ such that $\mathrm{supp}(\zeta) \subset \{\xi\in\mathbb{R}^n : 1 < |\xi| < \frac{3}{2}\}$ and define for $k\in\mathbb{R}$

$$P_k f(x) := \int_{\mathbb{R}^n} \zeta\left(2^{-k}|\xi|\right)\widehat{f}(\xi) e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \qquad x\in\mathbb{R}^n. \tag{1.9}$$

Theorem 1.1. *There exists $A = A(d)$ such that the following holds. Suppose that the family $\mathbf{m} := \{m_\sigma \in L^\infty(\mathbb{R}^d) : \sigma \in \mathbf{Gr}(d,n)\}$ is such that*

$$\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} := \sup_{\sigma,\tau\in\mathbf{Gr}(d,n)} \Big[\|m_\sigma\|_{\mathscr{M}_A(d)} + \log\Big(\mathrm{e} + \tfrac{1}{\mathrm{dist}(\sigma,\tau)}\Big) \|m_\tau - m_\sigma\|_{\mathscr{M}_A(d)} \Big] \leqslant 1. \tag{1.10}$$

Referring to (1.8)*, there holds*

$$\big\| T^\star_{\mathbf{m}} \circ P_0 : L^2(\mathbb{R}^n) \to L^{2,\infty}(\mathbb{R}^n) \big\| \lesssim 1, \qquad \big\| T^\star_{\mathbf{m}} \circ P_0 : L^p(\mathbb{R}^n) \big\| \lesssim_p 1, \quad 2 < p < \infty. \tag{1.11}$$

Remark 1.2. *The norm* (1.3) *is invariant under isotropic scaling of $\mathbb{R}^d$. Condition* (1.10) *inherits such invariance, whence* (1.11) *holds with the same constants for $T^\star_{\mathbf{m}} \circ P_k$ as well, $k \in \mathbb{R}$.*

Remark 1.3. *A sufficient condition for assumption* (1.10) *to hold is the following. Let*

$$\|\mathbf{m}\|_{\mathscr{M}_{A,1}(\mathbf{Gr}(d,n))} := \sup_{\sigma\in\mathbf{Gr}(d,n)} \sup_{0\leqslant|\beta|\leqslant 1} \|\mathscr{D}^\beta_\sigma m_\sigma\|_{\mathscr{M}_A(d)}$$

where the differential operator $\mathscr{D}_\sigma$ is the natural invariant vector of σ-derivatives; an explicit description is given in Lemma 2.2. Since

$$\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \lesssim \|\mathbf{m}\|_{\mathscr{M}_{A,1}(\mathbf{Gr}(d,n))},$$

the conclusion of Theorem 1.1 holds under the assumption $\|\mathbf{m}\|_{\mathscr{M}_{A,1}(\mathbf{Gr}(d,n))} \leqslant 1$.

In order to further emphasize the interpretation of our operators as directional multipliers and draw a formal parallel to the two-dimensional case of [31], let us rewrite (1.7) in yet another form. Equip $\mathbb{R}^n$ with the standard basis $(e_1,\dots,e_d,e_n)$ and define an orthonormal basis on σ by setting

$$v^\sigma_j := V^{-1} e_j, \qquad j \in \{1,\dots,d\}, \quad \mathrm{span}(v^\sigma_1,\dots,v^\sigma_d) = \sigma, \qquad V \in \mathscr{V}(\sigma).$$

Then, calculate

$$V\Pi_\sigma \xi = V \sum_{j=1}^d \langle \xi, v^\sigma_j\rangle v^\sigma_j = \sum_{j=1}^d \langle \xi, v^\sigma_j\rangle e_j = \big(\langle \xi, v^\sigma_1\rangle,\dots,\langle \xi, v^\sigma_d\rangle\big) \in \mathbb{R}^d$$

so that (1.6) takes the form

$$T_m f(x;\sigma,Q) = \int_{\mathbb{R}^n} m\big(\langle \xi, v^\sigma_1\rangle,\dots,\langle \xi, v^\sigma_d\rangle\big)\, \widehat{f}(\xi) e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \qquad x \in \mathbb{R}^n.$$

Then Theorem 1.1 tells us that the operator

$$f \mapsto \sup_{\sigma\in\mathbf{Gr}(d,n)} \sup_{v^\sigma_1,\dots,v^\sigma_d} \Big| \int_{\mathbb{R}^n} m_\sigma\big(\langle \xi, v^\sigma_1\rangle,\dots,\langle \xi, v^\sigma_d\rangle\big)\, \widehat{f}(\xi) e^{2\pi i\langle \cdot,\xi\rangle}\,\mathrm{d}\xi \Big|$$

is of weak-type $(2,2)$ and strong type (p,p), $p > 2$, whenever $\mathbf{m} = \{m_\sigma \in L^\infty(\mathbb{R}^d) : \sigma \in \mathbf{Gr}(d,n)\}$ satisfies $\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1$; above, the inner supremum is over all orthonormal bases of σ. Specializing to a single multiplier $m \in \mathscr{M}_A(d)$ we record below an immediate corollary.

Corollary 1.4. *Let $m \in \mathscr{M}_A(d)$ and consider the maximal directional multiplier*

$$T_m^\star f(x) := \sup_{(v_1,\dots,v_d)} \left| \int_{\mathbb{R}^n} m\left(\langle \xi, v_1\rangle, \dots, \langle \xi, v_d\rangle\right) \widehat{f}(\xi) e^{2\pi i \langle x,\xi\rangle} \, \mathrm{d}\xi \right|, \qquad x \in \mathbb{R}^n,$$

where the supremum is taken over all orthonormal d-tuples $(v_1,\dots,v_d) \in (\mathbb{R}^n)^d$. Then $T_m^\star \circ P_0$ maps $L^2(\mathbb{R}^n)$ to $L^{2,\infty}(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ for all $2 < p < \infty$.

We encourage the reader to compare with the very familiar two-dimensional directional multipliers, [11, 31], given in the form

$$T f(x, v) := \int_{\mathbb{R}^2} m\left(\langle \xi, v\rangle\right) \widehat{f}(\xi) e^{2\pi i \langle x,\xi\rangle} \, \mathrm{d}\xi, \qquad (x, v) \in \mathbb{R}^2 \times \mathbb{S}^1,$$

when $m \in \mathscr{M}_A(1)$ is a one-dimensional Hörmander–Mihlin multiplier. In the two-dimensional case, SO(1) is trivial, so that such dependence may be omitted.

When the band limited restriction imposed by precomposing with P_k is lifted and no regularity assumptions are placed on the subspace choice function, such function $\sigma(\cdot)$ is allowed to be oriented along a suitable Kakeya–Nikodym set and the subspace singular integrals (1.8) are in general unbounded on $L^p(\mathbb{R}^n)$. One particularly deep line of investigation is seeking for suitable regularity assumptions on σ bypassing the above mentioned counterexample and ultimately leading to L^p-bounds. For instance, if $\sigma(\cdot)$ is 1-Lipschitz and $m_\sigma(\cdot)$ is supported on frequency scales $\gg 1$, then Kakeya-type counterexamples are avoided. Zygmund suggested that — in this context — a suitably truncated version of the averaging directional operator along a Lipschitz choice of subspace should be bounded, at least above a critical L^p-space which can be identified for each dimension n and codimension $n-d$, see [17]. The corresponding version of this conjecture for multipliers m_σ, allowed to be singular along $\sigma^\perp$ is usually attributed to Stein. These are the Zygmund and Stein conjectures alluded to above and underlying the investigations in this paper.

For the case of directional singular integrals, namely when m_σ is allowed to be singular along $\sigma^\perp$, further counterexamples of non-Kakeya type do exist even when the range of $\sigma(\cdot)$ has special (e.g. lacunary) structure as has been exhibited in [22, 28]. For this reason it is customary, at least in the case $d = 1$, to study σ with finite range and prove optimal bounds in terms of the cardinality of such range. Optimal cardinality bounds for directional singular integrals are mostly known in the two-dimensional case, see [11, 12], while in higher dimensions $n \geqslant 2$ and $d = 1$ optimal bounds for directional multipliers and σ with lacunary range are contained in [1, 15].

In the present paper we prove the first such result in higher dimensions $n \geqslant 2$ and codimension $n - d = 1$. It is worth mentioning that while most of the results in the literature deal with maximal singular integrals generated by a single multiplier, the statement below allows for a log-Hölder dependence of the family m_σ in $\sigma \in \mathbf{Gr}(d, n)$. A two-dimensional partial analogue has appeared in [23], where the authors even allow for a measurable with respect to σ choice of multipliers, albeit with sub-optimal bounds in terms of the cardinality.

Corollary 1.5. *Suppose that the family* $\mathbf{m} := \{m_\sigma \in L^\infty(\mathbb{R}^d) : \sigma \in \mathbf{Gr}(d,n)\}$ *satisfies*

$$\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1.$$

With reference to (1.6)*, the maximal operator*

$$T^\star_{\Sigma,\mathbf{m}} f(x) := \sup_{\sigma\in\Sigma} \sup_{Q\in \mathrm{SO}(d)} \left| T_{m_\sigma} f(x;\sigma,Q)\right|, \qquad x\in\mathbb{R}^n, \qquad \Sigma\subset \mathbf{Gr}(d,n), \tag{1.12}$$

satisfies the norm inequality

$$\sup_{\substack{\Sigma\subset\mathbf{Gr}(d,n)\\ \#\Sigma\leqslant N}} \left\| T^\star_{\Sigma,\mathbf{m}}\right\|_{L^2(\mathbb{R}^n)} \lesssim_n \log N$$

and this is best possible up to the implicit dimensional constant.

From the proof of Theorem 1.1, it is also possible to deduce weak $(2,2)$ and strong (p,p) bounds, with $2<p<\infty$, for maximally truncated versions of (1.8). When considering the multiplier $m_\sigma \equiv 1$, this corresponds to a maximal subspace averaging operator. The obtained mapping properties can then be used to deduce subspace Lebesgue differentiation theorems for functions in the Besov spaces $B^s_{p,1}(\mathbb{R}^n)$, $s\geqslant 0$. These recover past results due to Murcko [33], see also [2, 35], for the case $s>0$, $n=2$, and appear to be new in the cases $s=0$ or $n>2$. A detailed statement is given in Theorem 7.1, Subsection 7.1.

Further motivation, as well as influence on the proof techniques, for Theorem 1.1 comes from the connection with maximally modulated singular integrals. In the two-dimensional case, $n=2=d+1$, there are well-explored [5, 31] ties between the band-limited behavior of the maximal directional Hilbert transform and the boundedness of the Carleson maximal partial inverse Fourier transform operator [7]. In a similar fashion, Theorem 1.1 implies L^p-estimates for maximally modulated Hörmander–Mihlin multipliers in the spirit of Sjölin [38].

Theorem 1.6. *Suppose* $m\in\mathscr{M}_A(d)$*. Then, the conclusion of Theorem 1.1 implies that the Carleson–Sjölin operator*

$$\mathrm{CS}[f](x) := \sup_{N\in\mathbb{R}^d} \left| \int_{\mathbb{R}^d} \widehat{f}(\eta)\, m(\eta+N)\, e^{2\pi i\langle \eta,x\rangle}\, \mathrm{d}\eta\right|, \qquad x\in\mathbb{R}^d,$$

maps $L^2(\mathbb{R}^d)$ *to* $L^{2,\infty}(\mathbb{R}^d)$ *and* $L^p(\mathbb{R}^d)$ *to* $L^p(\mathbb{R}^d)$ *for all* $2<p<\infty$*.*

The transference type argument leading to the proof of Theorem 1.6 is a modified version of the observation, commonplace in the literature, that any horizontal frequency cutoff to a half-line of a function supported on a thin vertical frequency strip may be obtained by applying a directional Hilbert transform with a suitably chosen slope. The connection to Carleson-type theorems and, more generally, modulation invariant operators, is in fact also apparent from the proof of Theorem 1.1 itself, which, much like its predecessors in [29, 4], borrows from the Lacey–Thiele argument for Carleson's maximal operator, [30].

Due to the higher dimensional nature of the problem considered, several new conceptual and technical difficulties arise. A first element of proof, which is only relevant in ambient dimension $n\geqslant 3$, addresses the possibility of choosing several rotations mapping $\mathbb{R}^d$ to a given $\sigma\in\mathbf{Gr}(d,n)$. The underlying rotation invariance allows for arbitrary choices of

coordinates on each σ, introducing artificial discontinuities to the problem in hand and is readily appreciated in formula (1.6). The necessary reduction which allows us to dispose of the arbitrary rotation Q in (1.6) and instead use canonical smooth rotations is contained in §3; the study of the model case of maximally rotated translation invariant Calderón–Zygmund operators is instructive and is also presented in §3. Additionally, the higher dimensional time-frequency analysis of the current paper combined with the directional nature of the operator under study leads to a novel model sum in terms of wave-packets conforming to the geometry of the operator, and requiring us to work with a choice function, dictating the measurable choice of subspaces from $\mathbf{Gr}(d,n)$, which localizes in higher dimensional frequency caps.

The final paragraph of this introduction serves as a more comprehensive historical account of past progress on maximal directional singular and averaging operators. This subject has attracted considerable attention in the last fifty years, mainly in connection to the Kakeya maximal conjecture. In this context, the dimension d of the averaging manifold is 1, and the connection may be described through the L^p-bounds, $1<p<\infty$, for the single and multi-scale directional maximal functions

$$\mathrm{M}_{V,1}f(x) := \sup_{v\in V}\int_{-\frac{1}{2}}^{\frac{1}{2}} |f(x-vt)|\,\mathrm{d}t, \quad \mathrm{M}_V f(x) := \sup_{r>0}\sup_{v\in V}\frac{1}{2r}\int_{-r}^{r}|f(x-vt)|\,\mathrm{d}t, \quad x\in\mathbb{R}^n$$

associated to a generic $V\subseteq\mathbb{S}^{n-1}$. The existence of Kakeya sets implies that these bounds blow up as $\delta\to 0^+$ when V is a δ-net on $\mathbb{S}^{n-1}$. When $n=2$, tight upper bounds for δ-nets are known and due by Strömberg, [40], and Córdoba, [10], who e.g. prove sharp logarithmic blowup in terms of δ above $L^2(\mathbb{R}^2)$, which is the critical space for the Kakeya maximal function in $\mathbb{R}^2$. These results imply the maximal Kakeya conjecture in two dimensions and for example we have

$$\|M_{V,1}\|_{L^2(\mathbb{R}^2)}\eqsim\sqrt{|\log\delta|}, \qquad \|\mathrm{M}_V\|_{L^2(\mathbb{R}^2)}\eqsim|\log\delta|.$$

In parallel it was known that if V has additional structure, namely if it is a *lacunary set of directions*, then M_V is bounded on all $L^p(\mathbb{R}^2)$ spaces, $1<p\leqslant\infty$, as was shown collectively by [9, 34, 37]. Bateman [3] showed that in fact M_V is bounded on some (equivalently all) $L^p(\mathbb{R}^2)$ spaces, $1<p\leqslant\infty$, if and only if V is a lacunary set. This means that if V is an arbitrary infinite set of directions then M_V will be generically unbounded so that one can assume that $\#V<+\infty$ and seek for best possible bounds in terms of the cardinality $\#V$. This was done by Katz who showed the best possible bounds for $\mathrm{M}_{V,1}$ in [25], and for M_V in [24]. These bounds match the optimal bounds for δ-nets stated above.

The maximal directional Hilbert transform

$$H_V f(x) := \sup_{v\in V}|H_v f(x)|, \qquad H_v f(x) := \text{p.v.}\int_{\mathbb{R}} f(x-vt)\frac{\mathrm{d}t}{t}, \qquad x\in\mathbb{R}^2,$$

is always unbounded if $\#V=\infty$; this was shown for $L^2(\mathbb{R}^2)$ by Karagulyan in [22] and for all $L^p(\mathbb{R}^n)$ by Łaba, Marinelli and Pramanik in [28]. The upper bounds for H_V and V lacunary were shown by Demeter and one of the authors, [12]; see also [14], resulting to the sharp

estimate

$$\sup_{\substack{V \text{ lacunary} \\ \#V \leqslant N}} \|H_V\|_{L^2(\mathbb{R}^n)} \eqsim \sqrt{\log N}.$$

The contributions [11, 12], containing what is essentially the $n = 2$ case of Corollary 1.5, have already been recalled. Finally the $n = 2$ case of Theorem 1.1 has a partial analogue in a result of Lacey and X. Li [31]. Therein, the authors pursued this question as an intermediate step towards the two-dimensional version of the Stein conjecture. Further results are available under other specific structural assumptions on the linearizing choice function for the directions v. The articles [4] by Bateman and [5] by Bateman and Thiele deal with the case where v is almost horizontal and does not depend on the vertical variable. In particular, the latter article showed that this structure entails $L^p(\mathbb{R}^2)$-bounds for all $3/2 < p < \infty$. Variants of these assumptions have been dealt with by S. Guo [19, 21]; see also [13] for a comprehensive result of this type. In a different direction Bourgain [6] has proved that a suitably truncated version of $M_{v(\cdot)}$ is bounded on $L^2(\mathbb{R}^2)$ when v satisfies a certain curvature condition, which in particular is verified by real analytic vector fields; see [20] for a geometric proof. Analogous results for the corresponding truncated directional Hilbert transform along v real analytic fields are due to obtained by Stein and Street, [39].

In higher dimensions, $n \geqslant 3$ most of the known results concern the case of codimension $n-1$ so that $d = 1$. Parcet and Rogers [36] extended the notion of lacunarity to any dimension and showed that M_V is bounded on all $L^p(\mathbb{R}^n)$-spaces under the assumption that V is lacunary. A partial converse in the spirit of Bateman is also proved in [36] but a full characterization is still pending; for $d > 1$ there is currently no notion of lacunary subsets of $\mathbf{Gr}(d,n)$ in the literature. For arbitrary $V \subset \mathbb{S}^{n-1}$ sharp $L^2(\mathbb{R}^n)$-bounds were shown for $\mathrm{M}_{V,1}$ by two of the authors using the polynomial method, [16]. Regarding M_V, the special case where $V \subset \mathbb{S}^{n-1}$ is a δ-net is considered in [26]. Sharp $L^2(\mathbb{R}^n)$-bounds for maximal single-scale subspace averages are proved in [17] for any $1 \leqslant d \leqslant n-1$; these include the case of codimension $n-d = 1$ where $L^2(\mathbb{R}^n)$ is the critical space.

Before this work, maximal directional singular multipliers had only been studied for $d = 1$. When V is lacunary and $n \geqslant 3$, Accomazzo and two of the authors have proved sharp $L^p(\mathbb{R}^n)$-bounds in [1]. In [27], Kim and Pramanik prove the sharp $L^2(\mathbb{R}^n)$-bounds for $V \subset \mathbb{S}^{n-1}$ arbitrary with $\#V < +\infty$. It should be noted that lacunary sets yield the same bound of the order $\sqrt{\log \#V}$ for all L^p-norms of the maximal directional singular integral in all dimensions. For $V \subset \mathbb{S}^{n-1}$ arbitrary this is no longer the case and logarithmic cardinality bounds can only be expected above the critical exponent $p = n$. This problem would however entail the resolution of the Kakeya conjecture and is currently wide open. Note that Corollary 1.5 above provides the first sharp bound for maximal directional multipliers and $d > 1$; since in our case the codimension $n-d = 1$ the critical space for this problem is still $L^2(\mathbb{R}^n)$, as in the case $n = 2 = d+1$.

Structure of the paper. Section 2 contains a few geometric preliminaries on the metric and differential structure of $\mathbf{Gr}(d,n)$. In Section 3, Theorem 1.1 is reduced to a simpler

version by removing the $\mathrm{SO}(d)$ maximality, see Proposition 3.3 for a precise statement. Section 4 is devoted to the deduction of Corollary 1.5 and Theorem 1.6. Section 5 further reduces Proposition 3.3 to the estimation of a time-frequency model operator, which is finally performed in Section 6. The concluding Section 7 is devoted to complementary results and questions on maximally truncated and bi-parametric variable $\mathbf{Gr}(d,n)$-subspace singular integrals.

Acknowledgments. The authors would like to thank the anonymous referee for an expert reading and for providing several helpful and detailed comments.

O. Bakas was partially supported by the projects CEX2021-001142-S, RYC2018-025477-I, PID2021-122156NB-I00/AEI/10.13039/501100011033 funded by Agencia Estatal de Investigación and acronym "HAMIP", Juan de la Cierva Incorporación IJC2020-043082-I, grant BERC 2022-2025 of the Basque Government, and by the funding programme "MEDICUS" of the University of Patras.

F. Di Plinio has been partially supported by the National Science Foundation under the grants NSF-DMS-2000510, NSF-DMS-2054863, as well as NSF-DMS-1929284 while in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the *Harmonic Analysis and Convexity* program in Fall 2022. F. Di Plinio is also partially supported by the FRA 2022 Program of University of Napoli Federico II, project ReSinAPAS - Regularity and Singularity in Analysis, PDEs, and Applied Sciences.

I. Parissis is partially supported by the project funded by Agencia Estatal de Investigación PID2021-122156NB-I00/AEI/10.13039/501100011033 and with acronym "HAMIP", grant T1615-22 of the Basque Government and IKERBASQUE.

L. Roncal is partially supported by the projects CEX2021-001142-S, CNS2023-143893, RYC2018-025477-I and PID2020-113156GB-I00/AEI/10.13039/501100011033 and with acronym "HAPDE" funded by Agencia Estatal de Investigación, grant BERC 2022-2025 of the Basque Government and IKERBASQUE.

2. Geometric preliminaries

The remainder of this article is concerned with the codimension 1 case, $d = n-1$. In this case $\mathbf{Gr}(d,n)$ is naturally identified with $\mathbb{S}^d$ via the isometric isomorphism

$$\sigma \mapsto \sigma^{\perp}, \qquad \mathbf{Gr}(d,n) \to \mathbf{Gr}(n-d,n),$$

This identification is exploited with the notation $v_\sigma \in \mathbb{S}^d$ as the unit normal vector to σ. We equip $\mathbf{Gr}(d,n) = \mathbf{Gr}(n-1,n)$ with the metric

$$\operatorname{dist}(\sigma,\tau) := |v_\sigma - v_\tau| = \sqrt{2}\sqrt{1-\cos[\theta(v_\sigma,v_\tau)]} \eqsim |\sin[\theta(v_\sigma,v_\tau)]|$$

where $\theta(u,v)$ stands for the convex angle between $u,v \in \mathbb{S}^d$. The equivalence constant implied by $\eqsim$ is obviously independent of the dimension. We set up a unified notation for balls in metric spaces. The standing convention is that $\xi \in X$ then $B(\xi,r)$ is the ball in the metric space X, centered at ξ and of radius $r>0$. The default metric on linear subspaces of $\mathbb{R}^n$ and $\mathbb{S}^d$ is the Euclidean metric and we write $B^m(R)$ for the Euclidean ball in $\mathbb{R}^m$, centered at 0 and having radius $R>0$. We also adopt the following notation for annuli in a

generic metric space X

$$\mathrm{Ann}(x,r,R) := B(x,R) \setminus \overline{B(x,r)}, \qquad 0 < r < R, \quad x \in X.$$

When X is a linear space the shorthand $\mathrm{Ann}(r,R)$ is also used in place of $\mathrm{Ann}(0,r,R)$ and $B(R)$ is used in place of $B(0,R)$.

2.1. **Smooth families of rotations.** The next lemma constructs a typical family of Lipschitz rotations of $\mathbb{R}^n$. These are essentially two-dimensional rotations between σ and τ acting in the plane perpendicular to the common $(d-1)$-dimensional subspace $\sigma \cap \tau = (\mathrm{span}\{v_\sigma, v_\tau\})^\perp$.

Lemma 2.1. *Let $\sigma, \tau \in \mathbf{Gr}(d,n)$ with $\mathrm{dist}(\sigma,\tau) < 1/2$. There exist $O_{\tau,\sigma} \in \mathrm{SO}(n)$ such that*

$$O_{\tau,\sigma}\sigma = \tau, \qquad O_{\tau,\sigma} v_\sigma = v_\tau, \qquad \big\| O_{\tau,\sigma} - \mathrm{Id} \big\| = \mathrm{dist}(\sigma,\tau).$$

Proof. If $\sigma = \tau$ take $O_{\tau,\sigma} = \mathrm{Id}$. Otherwise define $\zeta := \sigma \cap \tau \in \mathbf{Gr}(d-1,n)$; see Figure 1. Since $\zeta^\perp = \mathrm{span}\{v_\sigma, v_\tau\}$ and $\sigma \cap \zeta^\perp$ is a one-dimensional subspace of $\zeta^\perp$, we may pick the unique unit vector $u_\sigma \in \sigma \cap \zeta^\perp$ satisfying the equation

$$v_\tau = -(\sin\theta)\, u_\sigma + (\cos\theta)\, v_\sigma, \qquad \theta := \arccos(\langle v_\sigma, v_\tau \rangle).$$

Let $O_{\tau,\sigma}$ be the rotation that acts as the identity on ζ and such that

$$O_{\tau,\sigma} u_\sigma = u_\tau := (\cos\theta)\, u_\sigma + (\sin\theta)\, v_\sigma, \qquad O_{\tau,\sigma} v_\sigma = v_\tau.$$

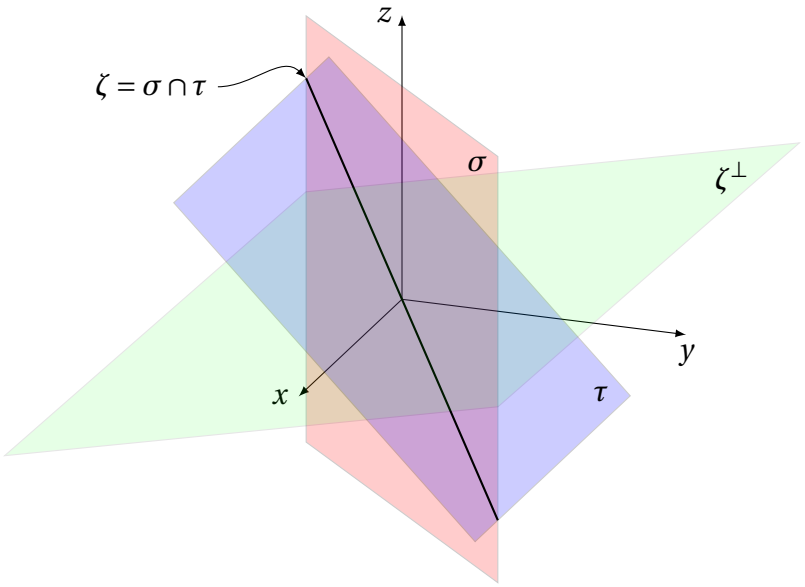

FIGURE 1. A figure for the proof of Lemma 2.1

Note that u_τ is a unit vector orthogonal to ζ and v_τ, therefore $\tau = \mathrm{span}\,\{\zeta, u_\tau\} = O_{\tau,\sigma}\sigma$. As $O_{\tau,\sigma} - \mathrm{Id} = [O_{\tau,\sigma} - \mathrm{Id}]\Pi_{\zeta^\perp}$, the equality

$$\big\| O_{\tau,\sigma} - \mathrm{Id} \big\| = \sqrt{2}\sqrt{1 - \langle v_\sigma, v_\tau \rangle} = |v_\sigma - v_\tau| = \mathrm{dist}(\sigma,\tau)$$

follows by immediate two-dimensional trigonometry. □

We give below an explicit description of the tangent bundle $\mathrm{T}\mathbf{Gr}(d,n)$. This is used in the sequel to give a concrete notion of derivatives of functions defined on $\mathbf{Gr}(d,n)$. In turn this will help us establish the smoothness of the map $\mathbf{Gr}(d,n) \ni \sigma \mapsto O_{\rho,\sigma}$ for fixed $\rho \in \mathbf{Gr}(d,n)$.

Lemma 2.2. *Let $\sigma \in \mathbf{Gr}(d,n)$ and let $\{v_1^\sigma, \cdots, v_d^\sigma\}$ be an orthonormal basis of σ and as usual $\sigma^\perp = \mathbb{R} v_\sigma$. The tangent space $\mathrm{T}_\sigma\mathbf{Gr}(d,n)$ is the linear span of*

$$X_{v_j^\sigma} \in \mathrm{so}(n), \quad j = 1,\dots,d, \qquad X_{v_j^\sigma}[v_k^\sigma] = \begin{cases} \vec{0}, & k \neq j \\ v_\sigma, & k = j \end{cases}, \quad 1 \leqslant k \leqslant d, \qquad X_{v_j^\sigma} v_\sigma = -v_j^\sigma,$$

where $\mathrm{so}(n)$ stands for $n \times n$ real skew-symmetric matrices. Furthermore, if $\xi \in \mathbb{R}^n$ is a fixed vector then the following hold,

$$\partial_{v_j^\sigma}(\Pi_\sigma \xi) = X_{v_j^\sigma}\xi, \qquad \partial_{v_j^\sigma}(\Pi_{\sigma^\perp}\xi) = -X_{v_j^\sigma}\xi, \qquad j = 1,\dots,d,$$

and the vector of derivatives $\mathscr{D}_\sigma$ can be described as

$$\mathscr{D}_\sigma = \left(\partial_{v_1^\sigma},\dots,\partial_{v_d^\sigma}\right), \qquad \sigma \in \mathbf{Gr}(d,n).$$

Proof. A basis of the tangent space $\mathrm{T}_\sigma\mathbf{Gr}(d,n)$ at $\sigma \in \mathbf{Gr}(d,n)$ can be given by $(X_{v_1^\sigma},\dots,X_{v_d^\sigma})$ where we define $X_{v_j^\sigma}$ for $1 \leqslant j \leqslant d$ to be the tangent vector at $t = 0$ to the $\mathbf{Gr}(d,n)$-valued curve

$$\sigma(t;j) \coloneqq \mathrm{span}\{v_1^\sigma,\dots,(\cos t)\, v_j^\sigma + (\sin t)\, v_\sigma,\dots,v_d^\sigma\}.$$

For the second claim we compute

$$\begin{aligned}
&\Pi_{\sigma(t;j)}\xi - \Pi_\sigma \xi \\
&= \langle \xi, (\cos t)\, v_j^\sigma + (\sin t)\, v_\sigma\rangle((\cos t)\, v_j^\sigma + (\sin t)\, v_\sigma) - \langle \xi, v_j^\sigma\rangle v_j^\sigma \\
&= (\sin t \cos t)\left[\langle \xi, v_j^\sigma\rangle v_\sigma + \langle \xi, v_\sigma\rangle v_j^\sigma\right] + [(\cos t)^2 - 1]\langle \xi, v_j^\sigma\rangle v_j^\sigma + (\sin t)^2\langle \xi, v_\sigma\rangle v_\sigma \\
&= (\sin t \cos t)\left[\langle \xi, v_j^\sigma\rangle v_\sigma + \langle \xi, v_\sigma\rangle v_j^\sigma\right] + O(t^2)
\end{aligned}$$

so that

$$\partial_{v_j^\sigma}(\Pi_\sigma\xi) = \lim_{t\to 0}\frac{\Pi_{\sigma(t;j)}\xi - \Pi_\sigma\xi}{t} = \langle \xi, v_j^\sigma\rangle v_\sigma + \langle \xi, v_\sigma\rangle v_j^\sigma = X_{v_j^\sigma}\xi$$

as claimed. Finally,

$$0 = \partial_{v_j^\sigma}\xi = \partial_{v_j^\sigma}(\Pi_\sigma\xi) + \partial_{v_j^\sigma}(\Pi_{\sigma^\perp}\xi)$$

whence the corresponding conclusion. □

The last lemma of this short section contains a computation of the σ-derivatives of the $\mathrm{SO}(n)$-valued map

$$\mathbf{Gr}(d,n) \ni \sigma \mapsto R_\sigma \coloneqq O_{\sigma,\rho} \qquad \sigma \in \mathrm{Gr}(d,n),$$

where $O_{\sigma,\rho}$ references Lemma 2.1 and $\rho \in \mathrm{Gr}(d,n)$ is kept fixed and will henceforth be omitted from the notation. In general, for $\rho \in \mathbf{Gr}(d,n)$ fixed we consider the following mappings:

- $\sigma \mapsto R_\sigma$ is the rotation on the plane $\mathrm{span}\{v_\sigma, v\}$ with $R_\sigma v = v_\sigma$; remember that $v = v_\rho$.
- $\sigma \mapsto \theta_\sigma \coloneqq \arccos(\langle v_\sigma, v\rangle)$.

· $\sigma \mapsto u_\sigma$, where u_σ is the unit vector of ρ which is perpendicular to $\sigma \cap \rho$ with $\langle u_\sigma, v_\sigma \rangle < 0$, and therefore lies in the plane $\mathrm{span}\{v_\sigma, v\}$.
· $\sigma \mapsto U_\sigma := R_\sigma u_\sigma$.

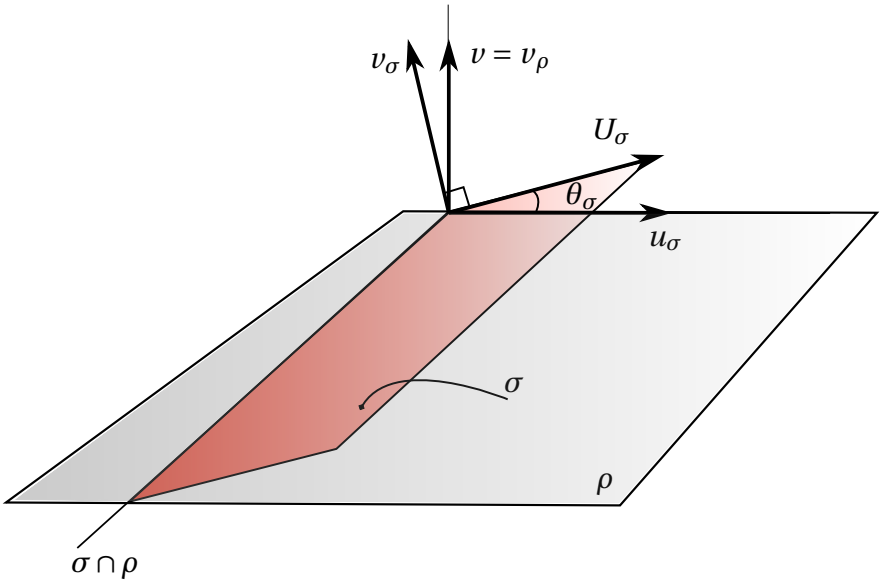

FIGURE 2. A figure for the proof of Lemma 2.3.

With these notations, depicted in Figure 2, there holds

$$U_\sigma = (\cos\theta_\sigma)\, u_\sigma + (\sin\theta_\sigma)\, v, \quad R_\sigma v = v_\sigma, \quad R_\sigma w = w \quad \forall\, w \in \sigma \cap \rho. \tag{2.1}$$

For the statement and proof of the lemma below it will be convenient to write $\mathbb{R}^n$ as the direct sum

$$\mathbb{R}^n = (\sigma \cap \rho) \oplus \mathbb{R} u_\sigma \oplus \mathbb{R} v.$$

Note then that σ can be described in the form $\sigma = \sigma \cap \rho \oplus \mathbb{R} U_\sigma$. Now Lemma 2.2 shows that in order to determine $\mathscr{D}^\beta_\sigma O_\rho(\sigma)$ for $|\beta| \leqslant 1$ it suffices to determine $\partial_{v^\sigma_j} R_\sigma$ for each vector from a suitably chosen basis of σ. This is effectively accomplished by the next lemma, as one may complete U_σ to an orthonormal basis of σ by choosing $d-1$ orthonormal vectors in $\sigma \cap \rho$.

Lemma 2.3. *Let $a, b, c \in \mathbb{R}$ and $\rho \in \mathbf{Gr}(d,n)$ be fixed. For every $\sigma \in \mathbf{Gr}(d,n)$ we have*

$$\partial_{U_\sigma} R_\sigma [w + b u_\sigma + c v] = b v_\sigma - c U_\sigma, \qquad \forall\, w \in \sigma \cap \rho.$$

Furthermore for all $w \in \sigma \cap \rho$ there holds

$$\partial_w R_\sigma [z + a w + b u_\sigma + c v] = [b\varpi(\theta_\sigma) - c]\, w + a\varpi(\theta_\sigma) u_\sigma + a v \qquad \forall\, z \in \sigma \cap \rho \cap w^\perp,$$

for the bounded function $\varpi(\beta) := \frac{\cos\beta - 1}{\sin\beta}$.

Remark 2.4. *In particular, $\sup_{0 \leqslant |\beta| \leqslant 1} \|\mathscr{D}^\beta_\sigma R_\sigma\| \lesssim 1 + |\varpi(\theta)| \lesssim_d 1$. Integrating along the geodesic connecting $\sigma, \tau \in \mathbf{Gr}(d,n)$, we obtain the operator norm estimate*

$$\|R_\tau - R_\sigma\| = \left\| R_\tau R_\sigma^{-1} - \mathrm{Id} \right\| \lesssim \mathrm{dist}(\tau, \sigma).$$

Throughout the paper we reserve the notation $O_\sigma \in \mathcal{V}(\sigma)$ for the rotations

$$O_\sigma = O_{\mathbb{R}^d,\sigma} \in \mathrm{SO}(n), \qquad O_\sigma \sigma = \mathbb{R}^d, \qquad O_\sigma v_\sigma = e_n.$$

Then, if the family $\{O_\sigma \in \mathcal{V}(\sigma) : \sigma \in \mathbf{Gr}(d,n)\}$ is constructed via Lemma 2.1, we will have

$$\sup_{0\leqslant|\beta|\leqslant 1} \|\mathscr{D}_\sigma^\beta O_\sigma\| \lesssim 1, \qquad \|O_\sigma - O_\tau\| \lesssim \mathrm{dist}(\sigma,\tau), \tag{2.2}$$

and these rotations will be used as the canonical family in (1.4).

Proof of Lemma 2.3. In order to compute $\partial_{U_\sigma} R_\sigma$, we need to compute the limit

$$\partial_{U_\sigma} R_\sigma = \lim_{t\to 0} \frac{R_{\tau(t)} - R_\sigma}{t} \tag{2.3}$$

where $\tau : [0,\pi/2] \to \mathbf{Gr}(d,n)$ is the $\mathbf{Gr}(d,n)$-valued curve

$$\tau = \tau(t) := (\sigma\cap\rho)\oplus \mathrm{span}\{U_\tau\}, \quad U_\tau := (\cos t)U_\sigma + (\sin t)v_\sigma = R_\tau u_\sigma,$$
$$v_\tau = -(\sin t)U_\sigma + (\cos t)v_\sigma,$$

and note that

$$\tau(0) = \sigma, \qquad U_{\tau(0)} = U_\sigma, \qquad v_{\tau(0)} = v_\sigma,$$

and that R_τ is a rotation by an angle $\arccos(\langle v_\tau, v\rangle) = \theta_\sigma + t$ in the plane $\mathrm{span}\{U_\sigma, v_\sigma\} = \mathrm{span}\{v_\tau, v\}$, and $v_\tau = R_\tau v$. It is then quite easy to see that

$$\partial_{U_\sigma} R_\sigma v = -U_\sigma, \qquad \partial_{U_\sigma} R_\sigma u_\sigma = v_\sigma, \qquad \partial_{U_\sigma} R_\sigma w = 0 \quad \forall\, w \in \sigma\cap\rho.$$

This proves the first identity in the statement of the lemma.

In order to prove the second claim of the statement it will suffice to compute $\partial_w R_\sigma$ for a generic vector $w \in \sigma\cap\rho$ of unit norm. Fix one such w and set $\zeta := \sigma\cap\rho\cap w^\perp$. The derivative $\partial_w R_\sigma$ will be calculated by repeating the limiting procedure (2.3) for

$$\tau = \tau(t) := \zeta \oplus \mathrm{span}\{w_{\tau(t)}, U_\sigma\}, \quad w_\tau := (\cos t)w + (\sin t)v_\sigma, \quad v_\tau := -(\sin t)w + (\cos t)v_\sigma.$$

Note again that we have

$$\tau(0) = \sigma, \qquad w_{\tau(0)} = w, \qquad v_{\tau(0)} = v_\sigma.$$

As R_σ, R_τ both act as the identity on ζ, one has

$$\partial_w R_\sigma z = 0 \qquad \forall\, z \in \zeta$$

and it suffices to test the action of $\partial_w R_\sigma$ on the basis vectors of $\zeta^\perp = \mathrm{span}\{u_\sigma, w, v\}$. The vectors $\{\widetilde{w}, \widetilde{u}\}$ given by

$$\widetilde{w} := \frac{1}{\beta(t)}\left[(\cos t)w - \frac{\sin t}{\sin\theta_\sigma}u_\sigma\right], \qquad \widetilde{u} := \frac{1}{\beta(t)}\left[\frac{\sin t}{\sin\theta_\sigma}w + (\cos t)u_\sigma\right],$$
$$\beta(t) := \sqrt{1 + (\sin t\cot\theta_\sigma)^2} \tag{2.4}$$

form an orthonormal basis of $\zeta^\perp\cap\rho = \mathrm{span}\{u_\sigma, w\}$ with, in addition, $\mathrm{span}\{\widetilde{w}\} = \zeta^\perp\cap\tau(t)\cap\rho$. Therefore R_τ acts as a rotation by θ_τ in the plane spanned by $\widetilde{u}, v$, where

$$\cos\theta_\tau = \langle v_\tau, v\rangle = \cos t\cos\theta_\sigma, \qquad \sin\theta_\tau = \sqrt{1 - (\cos t\cos\theta)^2},$$

namely,

$$R_\tau \widetilde{w} = \widetilde{w}, \qquad R_\tau \widetilde{u} = \cos[\theta_\tau]\,\widetilde{u} + \sin\theta_\tau\, v, \qquad R_\tau v = v_\tau.$$

Recalling our notation this means that $u_\tau = \widetilde{u}$ and $U_\tau = R_\tau u_\tau = R_\tau \widetilde{u}$.

Computing $\partial_w R_\sigma u_\sigma$. The system (2.4) may be easily inverted, giving

$$w = \frac{1}{\beta(t)}\left[(\cos t)\,\widetilde{w} + \frac{\sin t}{\sin\theta}\,\widetilde{u}\right], \qquad u_\sigma = \frac{1}{\beta(t)}\left[-\frac{\sin t}{\sin\theta}\,\widetilde{w} + (\cos t)\,\widetilde{u}\right].$$

Also recalling (2.1), (2.4),

$$R_\tau u_\sigma = \frac{1}{\beta(t)}\left[-\frac{\sin t}{\sin\theta_\sigma}\,\widetilde{w} + ((\cos t)^2\cos\theta_\sigma)\,\widetilde{u} + (\cos t)\sqrt{1-(\cos t\cos\theta_\sigma)^2}\,v\right],$$
$$R_\sigma u_\sigma = \frac{1}{\beta(t)}\left[-\frac{\cos\theta_\sigma\sin t}{\sin\theta_\sigma}\,\widetilde{w} + (\cos t\cos\theta_\sigma)\,\widetilde{u} + \beta(t)(\sin\theta_\sigma)\,v\right].$$

Taking the difference and using some calculus

$$[R_\tau - R_\sigma]\,u_\sigma = \frac{(\sin t)\varpi(\theta_\sigma)}{\beta(t)}\,\widetilde{w} + o(t)\,\widetilde{u} + o(t)\,v.$$

Therefore, dividing by t and taking the limit

$$\partial_w R_\sigma u_\sigma = \varpi(\theta_\sigma)\,w = \frac{\langle v_\sigma, v\rangle - 1}{\sqrt{1-\langle v_\sigma, v\rangle^2}}\,w.$$

Computing $\partial_w R_\sigma w$. Similarly

$$R_\tau w = \frac{1}{\beta(t)}\left[(\cos t)\,\widetilde{w} + \frac{\sin t\cos t\cos\theta_\sigma}{\sin\theta_\sigma}\,\widetilde{u} + \frac{\sin t}{\sin\theta_\sigma}\sqrt{1-(\cos t\cos\theta_\sigma)^2}\,v\right],$$
$$R_\sigma w = w = \frac{1}{\beta(t)}\left[(\cos t)\,\widetilde{w} + \frac{\sin t}{\sin\theta_\sigma}\,\widetilde{u}\right].$$

Taking the difference and using some more calculus

$$[R_\tau - R_\sigma]\,w = \frac{(\sin t)\varpi(\theta_\sigma)}{\beta(t)}\,\widetilde{u} + (\sin t)\,v + o(t)\,v(t).$$

Therefore, dividing by t and taking the limit

$$\partial_w R_\sigma w = \varpi(\theta_\sigma)\,u_\sigma + v = \frac{\langle v_\sigma, v\rangle - 1}{\sqrt{1-\langle v_\sigma, v\rangle^2}}\,u_\sigma + v.$$

Computing $\partial_w R_\sigma v_\rho$. This one is the easiest, as

$$\partial_w R_\sigma v = \lim_{t\to 0}\frac{[(\cos t)-1]\,v_\sigma - (\sin t)\,w}{t} = -w$$

and the proof is complete. □

3. Removing the SO(d) invariance

The maximal operator (1.8) is invariant under precompositions of the family m_σ with elements $Q_\sigma \in \mathrm{SO}(d)$. In other words, $T^\star$ does not change if each m_σ is replaced by $m_\sigma \circ Q_\sigma$ for some $Q_\sigma \in \mathrm{SO}(d)$ for each $\sigma \in \mathbf{Gr}(d,n)$. In this section, Theorem 1.1 is deduced from an

analogous estimate for the formally weaker maximal operator (3.1) where $Q_\sigma \equiv$ the identity in $\mathrm{SO}(d)$. This deduction is summarized in Proposition 3.2 below. The essential idea of the argument comes from an averaging argument showing the $L^p(\mathbb{R}^d)$-boundedness of maximally rotated d-dimensional Hörmander–Mihlin multipliers, which is presented below as a warm-up.

Proposition 3.1. *Let $m \in \mathscr{M}_A(d)$ for $A \geqslant d+1+\frac{d(d-1)}{2}$. Then*

$$T^\star f(x) := \sup_{Q\in \mathrm{SO}(d)} \left| \int_{\mathbb{R}^d} \widehat{f}(\eta) m(Q\eta) e^{2\pi i x\cdot\eta}\, \mathrm{d}\eta \right|, \qquad x \in \mathbb{R}^d,$$

maps $L^p(\mathbb{R}^d)$ into itself for all $1 < p < \infty$.

Proof. The idea of the proof is the same in all dimensions $d > 1$. In order to make the argument more concrete we first prove the case $d = 2$. A generic element $O \in \mathrm{SO}(2)$ may be written in the canonical basis of $\mathbb{R}^2$ as

$$O = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

for some $\theta \in [0, 2\pi)$. Let $O(\theta)$ be the element corresponding to $\theta \in [0,2\pi)$. We note that the coordinates of $O(\theta)$ are smooth functions of θ and we adopt the notation

$$O'(\theta) = \begin{bmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{bmatrix}$$

for the derivative. It is then useful to introduce the family of multipliers

$$m_\theta(\eta) := \langle \nabla m(O(\theta)\eta), O'(\theta)\eta \rangle, \qquad \eta \in \mathbb{R}^2.$$

It is immediate to see that there exists a constant c independent of θ such that $cm_\theta \in \mathscr{M}_{A-1}(2)$. Defining

$$Tf(x,\theta) := \int_{\mathbb{R}^2} \widehat{f}(\eta) m(O(\theta)\eta) e^{2\pi i x\cdot\eta}\, \mathrm{d}\eta, \quad Sf(x,\theta) := \partial_\theta Tf(x,\theta) = \int_{\mathbb{R}^2} \widehat{f}(\eta) m_\theta(\eta) e^{2\pi i x\cdot\eta}\, \mathrm{d}\eta,$$

one has the equality

$$Tf(x,\theta) = Tf(x,0) + \int_0^\theta Sf(x,\tau)\, \mathrm{d}\tau, \qquad 0 \leqslant \theta < 2\pi,$$

whence

$$|Tf(\cdot,\theta)| \leqslant |Tf(\cdot,0)| + \int_0^{2\pi} |Sf(\cdot,\tau)|\, \mathrm{d}\tau, \qquad 0 \leqslant \theta < 2\pi.$$

Therefore, using Minkowski's inequality

$$\begin{aligned} \left\| T^* f \right\|_p &= \left\| \sup_{\theta\in[0,2\pi)} \left| Tf(\cdot,\theta) \right| \right\|_p \leqslant \| Tf(\cdot,0) \|_p + \left\| \int_0^{2\pi} |Sf(\cdot,\tau)|\, \mathrm{d}\tau \right\|_p \\ &\leqslant \| Tf(\cdot,0) \|_p + 2\pi \sup_{\tau\in[0,2\pi)} \| Sf(\cdot,\tau) \|_p \lesssim \| f \|_p \end{aligned}$$

using that $m \in \mathscr{M}_A(2)$, $cm_\theta \in \mathscr{M}_{A-1}(2)$ uniformly, and the Hörmander–Mihlin theorem.

We then sketch the argument for $d > 2$. Recall that $\mathrm{SO}(d)$ is a compact Lie group of dimension $D = \frac{d(d-1)}{2}$. Hence, there exist $C > 0$ and $0 < \varepsilon < 1$ so that $\mathrm{SO}(d)$ may be covered by smooth charts

$$Q_j = Q_j(\theta) = Q_j(\theta_1, \dots, \theta_D) : [0,\varepsilon)^D \to \mathrm{SO}(d), \qquad 1 \leqslant j \leqslant C.$$

For example, if $d = 3$ one may use axis-angle pairs to parametrize rotations. Namely, $O_j(\theta_1, \theta_2, \theta_3)$ is the rotation by angle θ_3 in the plane perpendicular to the unit vector

$$N(\theta_1, \theta_2) = (\cos\theta_1 \sin\theta_2, \sin\theta_1 \sin\theta_2, \cos\theta_2).$$

Then $\mathrm{SO}(3)$ may be covered by $O(1)$ smooth axis angle charts with parameter $\varepsilon = 2^{-3}\pi$.

By finite splitting it thus suffices to bound the maximal operator

$$\sup_{\theta \in [0,\varepsilon)^D} |Tf(x,\theta)| \quad \text{where} \quad Tf(x,\theta) \coloneqq \int_{\mathbb{R}^d} \widehat{f}(\eta)\, m(Q_j(\theta)\eta)\, e^{2\pi i x\cdot\eta}\, \mathrm{d}\eta, \quad x \in \mathbb{R}^n$$

for each $j = 1, \dots, C$ fixed. For $S = \{k_1 < \dots < k_{\#S}\} \subset \{1, \dots, D\}$ write $\partial_S = \partial_{k_1} \cdots \partial_{k_{\#S}}$. Let also Π_S be the orthogonal projection on $\mathbb{R}^S \coloneqq \mathrm{span}\{e_k : k \in S\}$, so that $\theta = \Pi_S\theta \oplus \Pi_{S^c}\theta$. If $\tau \in \mathbb{R}^S$, it makes sense to write $\theta = \tau \oplus 0_{S^c} \in \mathbb{R}^D$ with the meaning that $\Pi_S\theta = \tau$ and $\Pi_{S^c}\theta = 0_{S^c}$. The key of our argument is again that the multiplier operators

$$T_S f(x,\theta) = \int_{\mathbb{R}^d} \widehat{f}(\eta)\, m_{\theta,S}(\eta)\, e^{2\pi i x\cdot\eta}\, \mathrm{d}\eta, \qquad m_{\theta,S} \coloneqq \partial_S m(Q_j(\theta)\cdot),$$

satisfy $c\, m_{\theta,S} \in \mathscr{M}_{A-\#S}(d) \subset \mathscr{M}_{A-D}(d)$ uniformly. Then one has the equality

$$Tf(x,\theta) = Tf(x,0) + \sum_{\varnothing \subsetneq S \subseteq \{1,\dots,D\}} \int_{\prod_{k\in S}[0,\theta_k]} T_S f(x, \tau \oplus 0_{S^c})\mathrm{d}\tau, \qquad \theta \in [0,\varepsilon)^D,$$

and similarly to what we have done before,

$$\Big\| \sup_{\theta\in[0,\varepsilon)^D} \big|Tf(\cdot,\theta)\big| \Big\|_p \leqslant \big\|Tf(\cdot,0)\big\|_p + 2^D \sup_{\varnothing \subsetneq S \subset \{1,\dots,D\}} \sup_{\tau \in [0,\varepsilon]^{\#S}} \big\| T_S f(\cdot, \tau \oplus 0_{S^c}) \big\|_p$$

so that we end up with the latter supremum being controlled by $C\|f\|_p$ via an application of the Hörmander–Mihlin theorem. The proof is complete. □

We are now ready to state and prove the anticipated reduction of Theorem 1.1.

Proposition 3.2. *Let $\Sigma \subset \mathbf{Gr}(d,n)$ be arbitrary. Let $O_\sigma = O_{\mathbb{R}^d,\sigma}$, with reference to the family of rotations constructed in Lemma 2.1; cf. Remark 2.4. Consider the maximal operator*

$$U^\star_{\Sigma,\mathbf{m}} f(x) \coloneqq \sup_{\sigma\in\Sigma} |T_{m_\sigma} f(x,\sigma)| \tag{3.1}$$

where

$$T_{m_\sigma} f(x,\sigma) \coloneqq \int_{\mathbb{R}^n} m_\sigma(O_\sigma \Pi_\sigma \xi)\, \widehat{f}(\xi)\, e^{2\pi i\langle x,\xi\rangle}\, \mathrm{d}\xi, \quad x \in \mathbb{R}^n.$$

For $1 < p < \infty$, $1 \leqslant q \leqslant \infty$, $D = \frac{d(d-1)}{2}$ and $T^\star_{\Sigma,\mathbf{m}}$ from (1.12)*, there holds*

$$\sup\{\|T^\star_{\Sigma,\mathbf{m}}\|_{L^p(\mathbb{R}^n)\to L^{p,q}(\mathbb{R}^n)} : \|\mathbf{m}\|_{\mathscr{M}_{A+D}(\mathbf{Gr}(d,n))} \leqslant 1\}$$
$$\lesssim \sup\{\|U^\star_{\Sigma,\mathbf{m}}\|_{L^p(\mathbb{R}^n)\to L^{p,q}(\mathbb{R}^n)} : \|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1\}$$

with implicit constant depending upon dimension only. The same statement holds if $U^\star_{\Sigma,\mathbf{m}}$ is replaced by $U^\star_{\Sigma,\mathbf{m}} \circ P_0$ in the right hand side and $T^\star_{\Sigma,\mathbf{m}}$ is replaced by $T^\star_{\Sigma,\mathbf{m}} \circ P_0$ in the left hand side of the estimate above.

Proof. Fix a family $\mathbf{m}$ with $\|\mathbf{m}\|_{\mathscr{M}_{A+D}(\mathbf{Gr}(d,n))} \leqslant 1$. By finite splitting, we obtain the claimed estimate for (1.12) from the same claim on the restricted maximal operator

$$T^\star_{\Sigma,\mathbf{m}} f(x) \coloneqq \sup_{\sigma\in\Sigma} \sup_{Q\in\mathscr{Q}} |T_{m_\sigma} f(x;\sigma,Q)|$$

with

$$T_{m_\sigma} f(x;\sigma,Q) \coloneqq \int_{\mathbb{R}^n} m_\sigma(QO(\sigma)\Pi_\sigma\xi)\widehat{f}(\xi) e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \quad x\in\mathbb{R}^n,$$

where $\mathscr{Q}$ is range of the chart Q_j for some $j = 1,\dots,C$; the overloading of the symbol $T^\star$ creates no confusion. At this point, for $Q\in\mathscr{Q}$, let $\theta(Q)$ be such that $Q = Q_j(\theta)$. Define the families

$$\mathbf{m}_{\theta,S} = \{m_{\sigma,\theta,S} : \sigma\in\mathbf{Gr}(d,n)\}, \quad m_{\sigma,\theta,S}(\eta) \coloneqq \partial^S m_\sigma(Q_j(\vec{\theta})\eta), \quad \theta\in[0,\varepsilon)^D,\ S\subset\{1,\dots,D\}.$$

With this definition, observe that

$$T_{m_\sigma} f(x;\sigma,Q) = T_{m_{\sigma,\theta(Q),\varnothing}} f(x;\sigma,\mathrm{Id}_{\mathbb{R}^d}) = T_{m_{\sigma,\theta(Q),\varnothing}} f(x,\sigma).$$

It is then routine to verify that

$$\sup_{\theta\in[0,\varepsilon)^D} \sup_{S\subset\{1,\dots,d\}} \left\|\mathbf{m}_{\theta,S}\right\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \lesssim \|\mathbf{m}\|_{\mathscr{M}_{A+D}(\mathbf{Gr}(d,n))} \leqslant 1. \tag{3.2}$$

Arguing as in the previous proof we then have

$$T^\star_{\mathbf{m},\Sigma} f(x) \leqslant U^\star_{\mathbf{m}_{0,\varnothing},\Sigma} f(x) + \sum_{\varnothing\subsetneq S\subseteq\{1,\dots,D\}} \int_{\prod_{k\in S}[0,\varepsilon]} U^\star_{\Sigma,\mathbf{m}_{\tau\oplus 0_{S^c},S}} f(x)\mathrm{d}\tau, \qquad \vec{\theta}\in[0,\varepsilon)^D,$$

which turns into the norm inequality

$$\|T^\star_{\mathbf{m}} f\|_{p,q} \lesssim_{p,q} \|U^\star_{\mathbf{m}_{0,\varnothing},\Sigma} f\|_{p,q} + 2^D \sup_{\varnothing\subsetneq S\subset\{1,\dots,D\}} \sup_{\tau\in[0,\varepsilon]^{\#S}} \left\|U^\star_{\Sigma,\mathbf{m}_{\tau\oplus 0_{S^c},S}} f\right\|_{p,q}.$$

The norms above are then controlled by $\|f\|_p$ by assumption in view of estimate (3.2), and the proof of the first claim is complete. For the second claim, it suffices to apply the above argument with $P_0 f$ in place of f. □

Because of the reduction devised in Proposition 3.2 above, Theorem 1.1 with $A = A(d) + \frac{d(d-1)}{2}$ is obtained from the following proposition.

Proposition 3.3. *Let $U^\star_{\Sigma,\mathbf{m}}$ be defined as in the statement of Proposition 3.2. There exists $A = A(d)$ such that the operator $U^\star_{\Sigma,\mathbf{m}} \circ P_0$ for $\Sigma = \mathbf{Gr}(d,n)$ maps $L^2(\mathbb{R}^n)$ to $L^{2,\infty}(\mathbb{R}^n)$ and $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$, $p>2$, uniformly over all families $\mathbf{m}$ satisfying $\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1$.*

The main line of proof of Proposition 3.3 occupies §5 and §6.

4. Maximal codimension-1 multipliers and the Carleson–Sjölin theorem

In this section we discuss the proofs of some of the consequences of our main result.

4.1. **Maximal codimension-1 multipliers and the proof of Corollary 1.5.** Let us recall our setup. Given $\Sigma \subset \mathbf{Gr}(d,n)$ with $\#\Sigma = N$, consider the maximal directional multiplier operator

$$f \mapsto T^{\star}_{\Sigma,\mathbf{m}} f(x) := \sup_{\sigma\in\Sigma} \sup_{Q\in \mathrm{SO}(d)} |T_{m_\sigma} f(x;\sigma,Q)|, \qquad x \in \mathbb{R}^n, \qquad \Sigma \subset \mathbf{Gr}(d,n),$$

where $T_{m_\sigma} f(x,\sigma,Q)$ is given by (1.6) and the family of multipliers $\mathbf{m} = \{m_\sigma \in \mathscr{M}_A(d) : \sigma \in \mathbf{Gr}(d,n)\}$ satisfies $\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1$. We will prove that

$$\sup_{\substack{\Sigma\subset\mathbf{Gr}(d,n)\\ \#\Sigma\leqslant N}} \left\| U^{\star}_{\Sigma,\mathbf{m}} f \right\|_{L^2(\mathbb{R}^n)} \lesssim \log N \|f\|_{L^2(\mathbb{R}^n)}$$

which implies the conclusion of Corollary 1.5 by Proposition 3.2.

A first well known consequence of Theorem 1.1 is that, in the case of N-multipliers, we can upgrade the single-annulus weak $(2,2)$ bound to a strong $(2,2)$ single-annulus bound, at a cost of $\sqrt{\log N}$.

Lemma 4.1. *We have the bound*

$$\sup_{\substack{\Sigma\subset\mathbf{Gr}(d,n)\\ \#\Sigma\leqslant N}} \sup_{k\in\mathbb{Z}} \|U^{\star}_{\Sigma,\mathbf{m}} \circ P_k f\|_{L^2(\mathbb{R}^n)} \lesssim \sqrt{\log N}\|f\|_{L^2(\mathbb{R}^n)},$$

with implicit constant depending only upon dimension.

The lemma follows easily by the weak $(2,2)$ bound

$$\sup_{k\in\mathbb{R}} \|U^{\star}_{\Sigma,\mathbf{m}}(P_k f)\|_{L^{2,\infty}(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\mathbb{R}^n)},$$

which is a special case of Theorem 1.1, together with trivial weak $(1,1)$ and strong (p,p) estimates for $U^{\star}_{\Sigma,\mathbf{m}} \circ P_k$ for $2 < p < \infty$, with bounds which are polynomial in N, and a Marcinkiewicz interpolation-type argument. The details can be found for example in [11]*Lemma 3.1, but essentially the same argument has been used in several places, for example in [24, 40].

With a strong $(2,2)$ bound in hand the proof of Corollary 1.5 follows a well known reduction, based on the Chang–Wilson–Wolff inequality [8], which allows us to essentially commute the supremum over N Fourier multipliers with a Littlewood–Paley square function. The argument leading to this reduction was introduced in [18], while in the context of directional multipliers it has been extensively used; see for example [1, 11]. The punchline proposition is stated below. Here, we use a smooth Littlewood–Paley partition of $\mathbb{R}^n$ given by means of

$$(S_t f)^{\wedge}(\xi) := \Psi\left(\frac{|\xi|}{t}\right)\widehat{f}(\xi), \qquad \xi \in \mathbb{R}^n, \quad t \in 2^{\mathbb{Z}},$$

where Ψ is a smooth radial function with compact support $\operatorname{supp}\Psi \subset \{\frac{1}{2} < |\xi| < 2\}$ with the additional requirement that

$$\sum_{t\in 2^{\mathbb{Z}}} \Psi\left(\frac{|\xi|}{t}\right) = 1, \qquad \xi \in \mathbb{R}^n \setminus \{0\}.$$

Proposition 4.2. *Let $\{T_1,\dots,T_N\}$ be Fourier multiplier operators in $\mathbb{R}^n$ with uniform bound*

$$\sup_{1\leqslant \nu\leqslant N} \|T_\nu\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} \leqslant 1.$$

If $\{S_t\}$ is a smooth Littlewood–Paley decomposition as described above then for $1<p<\infty$ there holds

$$\Big\| \sup_{1\leqslant \nu\leqslant N} |T_\nu f| \Big\|_{L^p(\mathbb{R}^n)} \lesssim_{p,n} \|f\|_{L^p(\mathbb{R}^n)} + \sqrt{\log(N+10)} \Big\| \Big(\sum_{t\in 2^{\mathbb{Z}}} \sup_{1\leqslant \nu\leqslant N} |(T_\nu \circ S_t) f|^2 \Big)^{\frac12} \Big\|_{L^p(\mathbb{R}^n)}.$$

The proof of this proposition can be found within [13]*Proof of Corollary 1.14; see also [11]. Now note that our main result, Theorem 1.1, readily implies that

$$\sup_{t\in 2^{\mathbb{Z}}} \big\| U^\star_{\Sigma,\mathbf{m}} \circ S_t f \big\|_{L^{2,\infty}(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\mathbb{R}^n)}$$

by splitting the support of Ψ into $O(1)$ pieces that match the size of the support of the auxiliary function ζ in the statement of Theorem 1.1, applying the conclusion of the theorem to each one of them, and adding them appropriately. Thus Lemma 4.1 yields the strong $L^2(\mathbb{R}^n)$-bound

$$\sup_{t\in 2^{\mathbb{Z}}} \big\| U^\star_{\Sigma,\mathbf{m}} \circ S_t f \big\|_{L^2(\mathbb{R}^n)} \lesssim (\log \#\Sigma)^{\frac12} \|f\|_{L^2(\mathbb{R}^n)}$$

for $N=\#\Sigma$ large. The conclusion of Corollary 1.5 now follows immediately by inserting the L^2-bound above in the conclusion of Proposition 4.2 for $p=2$. Note that the uniform L^2-bound, which is necessary for the application of Proposition 4.2, is an immediate consequence of the standing assumption $\|\mathbf{m}\|_{\mathscr{M}_A(\mathbf{Gr}(d,n))} \leqslant 1$ for the family of multipliers $\mathbf{m}$.

4.2. **The Carleson–Sjölin theorem.** The purpose of this subsection is to provide the connection of our main theorem, Theorem 1.1, with the Carleson–Sjölin theorem, as formulated in Theorem 1.6. We begin with a simple but useful geometric lemma.

Lemma 4.3. *Let $\varepsilon>0$ be a small parameter, $\Sigma\subset \mathbf{Gr}(d,n)$ be a ε-neighborhood of $\mathbb{R}^d = e_n^\perp$. For all $R>0$ the map*

$$N:\Sigma\to\mathbb{R}^d, \qquad N(\sigma) := R\,[\langle v_\sigma, e_n\rangle e_n - v_\sigma]$$

is onto the ball $B^d(c_o\varepsilon R)\subset\mathbb{R}^d$, for some dimensional constant $0<c_o<1$. Furthermore

$$|N(\sigma)-N(\tau)| \lesssim R\,\mathrm{dist}(\sigma,\tau), \qquad \sigma,\tau\in\Sigma$$

with implicit absolute constant.

Proof. Clearly $N(\mathbb{R}^d)=0$. If $N\neq 0$ belongs to the ball $B^d(c_o\varepsilon R)$, set

$$v_\sigma := -\frac{N}{R} + \sqrt{1-\frac{|N|^2}{R^2}}\, e_n, \qquad \sigma := v_\sigma^\perp.$$

As $\mathrm{dist}(\sigma,\mathbb{R}^d)\sim \arccos(\langle v_\sigma, e_n\rangle) \sim \frac{|N|}{R}$, and $|N|\leqslant c_o\varepsilon R$, it follows that $\sigma\in\Sigma$ if c_o is suitably chosen. Also, it is immediate to verify that $N(\sigma)=N$. The proof of the surjectivity claim is thus complete. The Lipschitz estimate is immediate, and its proof is thus omitted. □

Note that the unit vector

$$U_\sigma := \frac{[e_n - \langle v_\sigma, e_n\rangle v_\sigma]}{\sqrt{1-\langle v_\sigma, e_n\rangle^2}}, \qquad v_\sigma \neq e_n$$

spans $\sigma \cap \operatorname{span}\{v_\sigma, e_n\}$. For $\sigma \neq e_n$, it is also convenient to define the unit vector $u_\sigma := N(\sigma)/|N(\sigma)|$ spanning $\mathbb{R}^d \cap \operatorname{span}\{v_\sigma, e_n\}$. Below, let $\{O_\sigma : \sigma \in \mathbf{Gr}(d,n)\}$ be the C^1 family of rotations of Remark 2.4; we remember that $O_{\mathbb{R}^d} \equiv \mathrm{Id}$ and each O_σ acts as the identity on $\sigma \cap \mathbb{R}^d$ and, in the orthogonal complement $\operatorname{span}\{v_\sigma, e_n\}$ as the rotation mapping U_σ to u_σ, and consequently v_σ to e_n. With the notation of Lemma 4.3 we have that, for $|N| \leqslant c_o \varepsilon R$, $\cos(|N|/R) \sim \langle v_\sigma, e_n\rangle$ and $\sin(|N|/R) \sim |N|/R$. A direct calculation entails the equality

$$O_\sigma \Pi_\sigma \xi = \Pi_{\mathbb{R}^d}\xi + [\langle v_\sigma, e_n\rangle - 1]\,\langle \xi, u_\sigma\rangle u_\sigma + \langle \xi, e_n\rangle \frac{N(\sigma)}{R}. \tag{4.1}$$

The following definitions and observations readily yield the following lemma.

Lemma 4.4. *Let $m \in \mathscr{M}_A(d)$ be supported in $B^d(R_o)$ for some $R_o > 1$ and consider the ε-neighborhood Σ of $\mathbb{R}^d$ in $\mathbf{Gr}(d,n)$, constructed in Lemma 4.3 with ε sufficiently small depending upon dimension only. Let also $\{O_\sigma : \sigma \in \Sigma\}$ be the family defined in Remark 2.4. Suppose F is a Schwartz function on $\mathbb{R}^n$ with $\operatorname{supp}\widehat{F} \subset \{\xi \in \mathbb{R}^n : R/2 < |\xi_n| < 2R\}$ for some $R > R_o$. Define*

$$M(\xi,\sigma) := m\left(\Pi_{\mathbb{R}^d}\xi + [\langle v_\sigma, e_n\rangle - 1]\,\langle \xi, u_\sigma\rangle u_\sigma + \xi_n \frac{N(\sigma)}{R}\right), \qquad (\xi,\sigma) \in \mathbb{R}^n \times \Sigma,$$

and let $T_\Sigma F(x,\sigma) := (M(\cdot,\sigma)\widehat{F})^\vee(x)$. There holds

$$\left\| \sup_{\sigma\in\Sigma} |T_\Sigma F(\cdot,\sigma)| \right\|_{L^{2,\infty}(\mathbb{R}^n)} \lesssim \|F\|_{L^2(\mathbb{R}^n)}, \quad \left\| \sup_{\sigma\in\Sigma} |T_\Sigma F(\cdot,\sigma)| \right\|_{L^p(\mathbb{R}^n)} \lesssim \|F\|_{L^p(\mathbb{R}^n)}, \quad 2 < p < \infty.$$

Proof. Consider the multiplier operator

$$F \mapsto \int_{\mathbb{R}^n} \widehat{F}(\xi) m(O_\sigma \Pi_\sigma \xi) \mathrm{e}^{2\pi i \langle \xi, x\rangle}\, \mathrm{d}\xi.$$

For $\xi \in \mathbb{R}^n$ such that $m(O_\sigma\Pi_\sigma\xi)\widehat{F}(\xi) \neq 0$, the assumptions on the support on m and $\widehat{F}$ imply that

$$|\Pi_{\mathbb{R}^d}\xi| \leqslant |v_\sigma - e_n||\xi| + |\Pi_\sigma \xi| \leqslant \varepsilon|\xi| + R_o \leqslant \varepsilon|\Pi_{\mathbb{R}^d}\xi| + \varepsilon|\xi_n| + R_o \leqslant \frac{1}{2}|\Pi_{\mathbb{R}^d}\xi| + 2R$$

if $\varepsilon < 1/2$. This shows that $|\Pi_{\mathbb{R}_d}\xi| \leqslant 4R$ and thus $R/2 < |\xi| < 6R$ for ξ as above. It is then clear that $T_\Sigma F = T_\Sigma S_R F$, with S_R a smooth frequency projection onto the annulus $\{R/10 < |\xi| < 10R\}$ whose symbol is identically one on $\{R/6 \leqslant |\xi| \leqslant 6R\}$. Because of (4.1) the lemma follows by a suitable application of Theorem 1.1. □

Proof of Theorem 1.6. Let us fix a Hörmander–Mihlin multiplier $m : \mathbb{R}^d \to \mathbb{C}$ as in the statement. We will make the qualitative assumption that m is a smooth function with compact support in $\mathbb{R}^d$, which is always possible by a suitable approximation; this assumption will be removed at the end of the proof. Our goal is to show that the operator

$$\mathrm{CS}[f](x) := \sup_{N\in\mathbb{R}^d} |\mathrm{CS}_N[f](x)|, \qquad \mathrm{CS}_N[f](x) := \int_{\mathbb{R}^d} \widehat{f}(\eta) m(\eta + N) e^{2\pi i\langle \eta, x\rangle}\, \mathrm{d}\eta, \qquad x \in \mathbb{R}^d,$$

maps $L^2(\mathbb{R}^d)$ to $L^{2,\infty}(\mathbb{R}^d)$ and $L^p(\mathbb{R}^d)$ to itself for all $2 < p < \infty$. To that end let us also fix $p \in [2,\infty)$ and a function $f \in \mathscr{S}(\mathbb{R}^d)$ with compact frequency support. By our qualitative assumptions there exists $R_o > 0$ such that

$$\operatorname{supp}(\widehat{f}) \subset B^d(R_o), \qquad \operatorname{supp}(m) \subset B^d(R_o),$$

and thus only values $N \in B^d(2R_o)$ need be considered in the definition of $\mathrm{CS}[f]$. We now fix $\varepsilon > 0$ sufficiently small, as coming from the statement of Lemma 4.4, but also satisfying the smallness condition

$$\varepsilon < \Big[\max\Big(1, 100R_o^d \|m\|_{C^{20d+1}}, (100R_o)\Big)\Big]^{-1}.$$

Above we have denoted for positive integers ν

$$\|g\|_{C^\nu} := \sum_{0\leqslant|\alpha|\leqslant\nu} \|\partial^\alpha g\|_{L^\infty(\mathbb{R}^d)}.$$

We will use an auxiliary function ψ to lift f from $\mathbb{R}^d$ to $\mathbb{R}^n$. More precisely we let $\psi : \mathbb{R} \to \mathbb{R}$ be a smooth function with compact frequency support $\operatorname{supp}(\widehat{\psi}) \subset (-1,1)$ and such that $\psi \geqslant 0$, $\psi \geqslant 1$ on $(-2^{-3}, 2^{-3})$, and $\widehat{\psi} \geqslant 0$, and define

$$\psi_R(t) := \int_{\mathbb{R}} \widehat{\psi}(u-R)\, e^{2\pi i t u}\, \mathrm{d}u, \qquad t \in \mathbb{R},$$

with $R > 0$ to be chosen momentarily. We readily see that $\operatorname{supp}(\widehat{\psi}_R) \subset (R-1, R+1)$. We then define the smooth function with compact frequency support

$$F(x) := f(y)\psi_R(x_n), \qquad x = (y, x_n) \in \mathbb{R}^n = \mathbb{R}^d \times \mathbb{R},$$

where we use the conventions $x = (\Pi_{\mathbb{R}^d} x, x_n) =: (y, x_n) \in \mathbb{R}^n$ and $\xi = (\Pi_{\mathbb{R}^d}\xi, \xi_n) =: (\eta, \xi_n) \in \mathbb{R}^n$ in order to simplify our formulas.

For any $N \in B^d(2R_o)$ we apply Lemma 4.3 with parameters ε, R such that $2R_o < c_o\varepsilon R < R$. We then know that there exists $\sigma = \sigma(N) \in \Sigma$ such that $N(\sigma) = N$. We fix such a sufficiently large value of R and note that

$$\operatorname{supp}(\widehat{\psi}_R) \subset \{t : R/2 < t < 2R\}.$$

Now set

$$\mathscr{E}(\xi,\sigma) := [\langle v_\sigma, e_n\rangle - 1]\,\langle \xi, u_\sigma\rangle u_\sigma + \left(\frac{\xi_n}{R} - 1\right) N(\sigma), \qquad (\xi,\sigma) \in \mathbb{R}^n \times \Sigma,$$

and immediately observe that for $\xi \in \operatorname{supp}(\widehat{F})$ we have, taking into account the facts in Lemma 4.3,

$$\begin{aligned}
&|\mathscr{E}(\xi,\sigma)| \lesssim \operatorname{dist}(\sigma,\mathbb{R}^d)^2 R_o + \operatorname{dist}(\sigma,\mathbb{R}^d),\\
&|\partial_{\xi_j}\mathscr{E}(\xi,\sigma)| \lesssim \operatorname{dist}(\sigma,\mathbb{R}^d)^2, \qquad 1 \leqslant j \leqslant d,\\
&|\partial_{\xi_n}\mathscr{E}(\xi,\sigma)| \lesssim \operatorname{dist}(\sigma,\mathbb{R}^d),\\
&|\nabla^\alpha \mathscr{E}(\xi,\sigma)| = 0 \quad \text{if} \quad |\alpha| > 1.
\end{aligned} \tag{4.2}$$

Defining $TF(x,\sigma)$ as in the statement of Lemma 4.4, with the choice of $\sigma \in \Sigma$ as above, we calculate

$$\begin{aligned} TF(x,\sigma) &= \int_{\mathbb{R}} \int_{\mathbb{R}^d} m\big(\eta + N(\sigma) + \mathscr{E}(\xi,\sigma)\big)\, \widehat{f}(\eta)\widehat{\psi}_R(\xi_n) e^{2\pi i \langle \eta, y\rangle} e^{2\pi i x_n \xi_n}\, \mathrm{d}\eta\, \mathrm{d}\xi_n \\ &= \mathrm{CS}_N[f](y)\psi_R(x_n) + \int_{\mathbb{R}} \int_{\mathbb{R}^d} \widehat{f}(\eta)\widehat{\psi}_R(\xi_n)\Delta(\xi,\sigma) e^{2\pi i \langle \eta, y\rangle} e^{2\pi i x_n \xi_n}\, \mathrm{d}\eta\, \mathrm{d}\xi_n, \end{aligned}$$

where

$$\Delta(\xi,\sigma) := m(\eta + N(\sigma) + \mathscr{E}(\xi,\sigma)) - m(\eta + N(\sigma)), \qquad (\xi,\sigma) \in \mathbb{R}^n \times \Sigma.$$

Note that $\operatorname{supp}(\widehat{F}) \subseteq B^d(R_o) \times \{R-1 < |\xi_n| < R+1\}$ which we want to keep memory of. For that reason we let $\lambda \in \mathscr{S}(\mathbb{R}^n)$ be identically one on $\operatorname{supp}(\widehat{F})$ and vanish outside $B^d(10R_o) \times \{R-10 < |\xi_n| < R+10\}$, with the natural derivative bounds. Define the operator

$$\mathrm{Err}_\sigma F(x) := \int_{\mathbb{R}^n} \widehat{F}(\xi)\lambda(\xi)\Delta(\xi,\sigma) e^{2\pi i \langle \xi, x\rangle}\, \mathrm{d}\xi, \qquad x \in \mathbb{R}^n.$$

We then have

$$|\mathrm{CS}[f](y)\psi_R(x_n)| \leqslant \sup_{\sigma\in\Sigma} |TF(x,\sigma)| + \sup_{\sigma\in\Sigma} |\mathrm{Err}_\sigma F(x)|$$

and the proof of the theorem will be complete once we control the error term $\mathrm{Err}_\sigma F$ in $L^p(\mathbb{R}^n)$. We will do so by proving

$$\sup_{\sigma\in\Sigma} |\mathrm{Err}_\sigma F(x)| \lesssim \mathrm{M}F(x), \qquad x \in \mathbb{R}^n; \tag{4.3}$$

here M denotes the n-dimensional Hardy–Littlewood maximal function. To that end let us define

$$E(x,\sigma) := \int_{\mathbb{R}} \int_{\mathbb{R}^d} \lambda(\xi)\Delta(\xi,\sigma) e^{2\pi i \langle \eta, y\rangle} e^{2\pi i x_n \xi_n}\, \mathrm{d}\eta\, \mathrm{d}\xi_n, \qquad x \in \mathbb{R}^n,$$

and note that (4.2) remains valid for $\xi \in \operatorname{supp}(\lambda)$. Estimates (4.2) combined with a calculation involving the chain rule, the mean value theorem, and the derivatives of m, $\mathscr{E}(\xi,\sigma)$ and λ, imply

$$\begin{aligned} |\operatorname{supp}(\lambda(\cdot)\Delta(\cdot,\sigma))| &\lesssim R_o^d \quad \text{(using that } \operatorname{supp} m \subset B^d(R_o)\text{)}, \\ |\nabla^\gamma(\lambda(\xi)\Delta(\xi,\sigma))| &\lesssim \|m\|_{C^{|\gamma|+1}} \Big[\operatorname{dist}(\sigma,\mathbb{R}^d)^2 R_o + \operatorname{dist}(\sigma,\mathbb{R}^d)\Big]. \end{aligned}$$

Using these bounds and integrating by parts inside the integral defining $E(x,\sigma)$ we gather

$$|E(x,\sigma)| \lesssim \frac{\|m\|_{C^{20d+1}} R_o^d \big[\operatorname{dist}(\sigma,\mathbb{R}^d) R_o + 1\big] \operatorname{dist}(\sigma,\mathbb{R}^d)}{(1+|y|)^{10d}(1+|x_n|)^{10d}}, \qquad x = (y,x_n) \in \mathbb{R}^d \times \mathbb{R}.$$

Remembering our choice of $\varepsilon > 0$ and that $\operatorname{dist}(\sigma,\mathbb{R}^d) \lesssim \varepsilon$, this proves (4.3) with implicit constant depending only upon dimension. Thus the error term is under control

$$\|\sup_{\sigma\in\Sigma} |\mathrm{Err}_\sigma F|\|_{L^p(\mathbb{R}^n)} \lesssim \|F\|_{L^p(\mathbb{R}^n)} \eqsim \|f\|_{L^p(\mathbb{R}^d)}.$$

We have proved the estimates

$$\|\mathrm{CS}[f]\|_{L^{2,\infty}(\mathbb{R}^d)} \lesssim \|f\|_{L^2(\mathbb{R}^d)}, \qquad \|\mathrm{CS}[f]\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)}, \qquad 2 < p < \infty,$$

whenever $f \in \mathscr{S}(\mathbb{R}^d)$ has compact Fourier support and m is a $C^{20d+1}(\mathbb{R}^d)$-function with compact support. One easily removes these assumptions on m by approximating with a sequence m_k which is smooth and has compact Fourier support, and converges pointwise to m a.e. in $\mathbb{R}^d$, and satisfies $\|m_k\|_{\mathscr{M}_A(d)} \lesssim \|m\|_{\mathscr{M}_A(d)}$ uniformly in k. Then dominated convergence shows that for any measurable map $\mathbb{R}^d \ni y \mapsto N(y)$ we have

$$\int_{\mathbb{R}^d} m(\xi + N(y))\widehat{f}(\xi)e^{2\pi i\langle y,\eta\rangle}\,\mathrm{d}\eta = \lim_k \int_{\mathbb{R}^d} m_k(\xi + N(y))\widehat{f}(\xi)e^{2\pi i\langle y,\eta\rangle}\,\mathrm{d}\eta$$

whenever $f \in \mathscr{S}(\mathbb{R}^d)$ has compact Fourier support. Taking absolute values and applying Fatou's lemma inside the $L^p(\mathbb{R}^d)$-norm yields the conclusion of the theorem for $f \in \mathscr{S}(\mathbb{R}^d)$ with compact Fourier support. As such functions are dense in $L^p(\mathbb{R}^d)$ the proof is complete. □

5. Tiles, adapted functions, and model operator

This and the following sections are dedicated to the proof of Proposition 3.3 which, by the reductions of §3, implies the main theorem, Theorem 1.1. In the remaining of the paper we use the C^1 family of rotations $\{O_\sigma : \sigma \in \mathbf{Gr}(d,n)\}$ of Remark 2.4, underlying the definition of the operator $U^\star_{\Sigma,\mathbf{m}}$ in (3.1).

5.1. **Preliminaries.** The family $\mathbf{m} = \{m_\sigma \in L^\infty(\mathbb{R}^d) : \sigma \in \mathbf{Gr}(d,n)\}$ satisfying (1.10) is fixed for the rest of the paper. Recall, cf. (3.1), that the definition of $U^\star_{\Sigma,\mathbf{m}}$ involves the singular integrals

$$T_{m_\sigma} f(x,\sigma) := \int_{\mathbb{R}^n} m_\sigma(O_\sigma \Pi_\sigma \xi)\widehat{f}(\xi)e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \qquad x \in \mathbb{R}^n. \tag{5.1}$$

The subscript m_σ in (5.1) will be omitted from now on. Let $\alpha \in 3^{-\mathbb{N}}$ and denote by

$$\Sigma_\alpha := \{\sigma \in \mathbf{Gr}(d,n) : |v_\sigma - e_n| = \operatorname{dist}(\sigma, \mathbb{R}^d \otimes \{0\}) < \alpha\},$$

a small neighborhood of $e_n^\perp$. The small constant α will depend on dimension only and its choice will be explained below. With such a choice, the parameter α will be fixed along the paper, so we will drop the subindex α and we will write Σ from now on. Consider the maximal operator

$$f \mapsto \sup_{\sigma\in\Sigma} |T[P_0 f](\cdot,\sigma)| \tag{5.2}$$

with $Tf(\cdot,\sigma)$ as defined in (5.1) and P_0 given by (1.9). Note that the operator $U^\star_{\mathbf{Gr}(d,n),\mathbf{m}}$ of Proposition 3.2 is controlled by $O_\alpha(1)$ rotated copies of (5.2). In order to prove Proposition 3.3, it thus suffices to obtain the corresponding estimate for (5.2). Restricting the choice of variable subspace to Σ allows us to precompose $T \circ P_0$ in (5.2) with a smooth restriction to a small frequency cone about e_n as follows. Let $\Psi_{\mathrm{cn}} \in \mathscr{S}(\mathbb{R}^n)$ satisfying

$$\begin{aligned} &\Psi_{\mathrm{cn}} \equiv 1 \text{ on } \Gamma_0 \cap \operatorname{Ann}\left(1,\tfrac{3}{2}\right), \qquad \Gamma_0 := \{\xi \in \mathbb{R}^n \setminus \{0\} : |\xi' - e_n| < 3^4\alpha\},\\ &\operatorname{supp}\Psi_{\mathrm{cn}} \subset \Gamma_1 \cap \operatorname{Ann}\left(\tfrac{1}{2},2\right), \qquad \Gamma_1 := \{\xi \in \mathbb{R}^n \setminus \{0\} : |\xi' - e_n| < 3^5\alpha\}, \end{aligned} \tag{5.3}$$

where $\xi' := \xi/|\xi|$, and let $P_{\mathrm{cn}}f$ be the corresponding frequency cutoff $P_{\mathrm{cn}}f := (\Psi_{\mathrm{c}n}\widehat{f})^\vee$. Then

$$\sup_{\sigma\in\Sigma}|T[(\mathrm{Id}-P_{\mathrm{cn}})\circ P_0 f](\cdot,\sigma)| \lesssim \mathrm{M}f \tag{5.4}$$

where M denotes the Hardy–Littlewood maximal operator; indeed, the singularities of the symbols of the operators $f\mapsto Tf(\cdot,\sigma)$ for $\sigma\in\Sigma$ are all contained in Γ_0, namely away from the frequency support of $(\mathrm{Id}-P_{\mathrm{cn}})\circ P_0 f$. This reduces the proof of Proposition 3.3 to the proof of the corresponding bound for the maximal operator

$$U^\star f := \sup_{\sigma\in\Sigma}\big|T[P_0\circ P_{\mathrm{cn}}f](\cdot,\sigma)\big|. \tag{5.5}$$

The multiplier operators $g\mapsto T[P_0\circ P_{\mathrm{cn}}g](\cdot,\sigma)$ have Fourier support in the truncated cone

$$\Delta := \Gamma_1\cap\mathrm{Ann}\left(1,\frac{3}{2}\right). \tag{5.6}$$

In our analysis, in accordance with the uncertainty principle and the Fourier support of the restriction to Δ, the relevant spatial scales for each $\sigma\in\Sigma$ will then be

$$s\in\mathbb{S} := \big\{s\in 3^{\mathbb{Z}}, 3^6 s\cdot\alpha\geqslant 1\big\} \tag{5.7}$$

in the σ variables, while the spatial scale will always be ~ 1 in the coordinate $\sigma^\perp$. Note that by taking α sufficiently small, depending on dimension only, we can and will always assume that $s\geqslant 3^{10d}$ for all relevant scales $s\in\mathbb{S}$. The rest of the section is devoted to construction of a model sum representation for $U^\star$.

5.2. **Grids and tiles.** One of the ways in which the simultaneous localization to $\sigma\in\Sigma$ and to the homogeneous Fourier region $\Delta' = \{\xi' : \xi\in\Delta\}\subset\mathbb{S}^d$ is exploited below is to conflate either region with a copy of $\mathbb{R}^d$ as follows. Let $\widetilde{\Delta}'\subset\mathbb{S}^d$ be the $3^6\sqrt{d}$-neighborhood of Δ'. Then, provided the constant α is chosen sufficiently small depending on dimension, there is an approximate isometry

$$\xi\in\widetilde{\Delta}'\mapsto \Pi_{e_n^\perp}\xi\in e_n^\perp\equiv\mathbb{R}^d$$

in the sense that

$$\frac{\theta(\xi,\eta)}{\left|\Pi_{e_n^\perp}\xi-\Pi_{e_n^\perp}\eta\right|}, \frac{|\xi-\eta|}{\left|\Pi_{e_n^\perp}\xi-\Pi_{e_n^\perp}\eta\right|}\in\left[\tfrac{1}{3},3\right], \qquad \xi,\eta\in\widetilde{\Delta}',\ \xi\neq\eta, \tag{5.8}$$

where we remember that $\theta(\xi,\eta)$ denotes the convex angle between $\xi,\eta\in\mathbb{S}^d$. This together with our previous restriction in the remark following (5.7) fixes the choice of α to be a small dimensional constant. When $\sigma\in\Sigma$ we have that $v_\sigma\in\Delta'$ and the approximate isometry extends to Σ as well. This allows us to construct a grid structure localizing the Σ and Δ' components by pulling back the corresponding Euclidean structure. Rotated Euclidean grids will also localize the spatial component: relevant definitions follow.

A standard triadic grid $\mathscr{G}$ on $\mathbb{R}^d$ of pace (K,m), where $K\geqslant 1$ is an integer and $m\in[0,K)$ is *not necessarily an integer*, is a collection of cubes $Q\subset\mathbb{R}^d$ satisfying the properties

g1. (quantized length) $\mathscr{G}=\bigcup_{j\in\mathbb{Z}}\mathscr{G}_{Kj+m}$, where $\mathscr{G}_u=\{Q\in\mathscr{G}:\ell(Q)=3^u\}$,

g2. (partition) $\mathbb{R}^d = \bigcup_{Q\in\mathscr{G}_{Kj+m}} Q$ for all $j\in\mathbb{Z}$,

g3. (grid) $L\cap Q\in\{L,Q,\varnothing\}$ for every $L,Q\in\mathscr{G}$.

In particular, $\mathscr{G}$ is partially ordered by inclusion and we may define for each $Q\in\mathscr{G}$

$$Q^{(1)} := \text{minimal } L\in\mathscr{G} \text{ with } Q\subsetneq L,$$

and inductively $Q^{(j)} := [Q^{(j-1)}]^{(1)}$ for all $j>1$ to be the j-th parent of Q in the grid $\mathscr{G}$. With these definitions, properties g2. and g3. yield for any positive integer κ the partition

$$L = \bigsqcup_{Q\in \mathrm{ch}_\kappa(L)} Q, \qquad \mathrm{ch}_\kappa(L) := \{Q\in\mathscr{G} : Q^{(\kappa)} = L\}$$

with the collection $\mathrm{ch}_\kappa(L)$ being referred to as the κ-th generation children of $L\in\mathscr{G}$. Triadic grids have been chosen because of the following convenient property. For each $L\in\mathscr{G}$ there is a unique $Q = L^{\circ,\kappa}\in\mathrm{ch}_\kappa(L)$ which is concentric with L, which is referred to as the κ-*center* of L. The next definition highlights a different role for certain other elements of $\mathrm{ch}_\kappa(L)$. Define

$$\mathrm{ch}_{\kappa,\square}(L) := \{Q\in\mathrm{ch}_\kappa(L) : \mathrm{dist}(Q, L^{\circ,\kappa}) \geqslant 3^{3-\kappa}\ell(L)\}$$

referring to this collection as the *peripheral* κ-th children of L. In the following lemma and hereafter we fix a large value of $\kappa := 12+\log_3 d$.

Lemma 5.1. *Let $\kappa := 12+\log_3 d$. There exist standard triadic grids $\mathscr{G}_1,\dots,\mathscr{G}_C$ of pace $(1,m_j)$ on $\mathbb{R}^d\equiv e_n^\perp$, $1\leqslant j\leqslant C$, $C = C_{\kappa,d}$, with the following property: For each $\beta\in\Delta'$, $s\in\mathbb{S}$, there exists $j = j(\beta,s)\in\{1,\dots,C\}$ and a cube $L = L(\beta,s)\in\mathscr{G}_j$ such that*

1. $3^3 s^{-1}\leqslant\ell(L)\leqslant 3^4 s^{-1}$,
2. $B\left(\beta, 3^{-\kappa}s^{-1}\right)\subset\Pi_{e_n^\perp}^{-1}(L^{\circ,\kappa})$,
3. $\mathrm{Ann}\left(\beta, 3^{-1}s^{-1}, 3^2 s^{-1}\right)\subset\bigcup\left\{\Pi_{e_n^\perp}^{-1}(Q) : Q\in\mathrm{ch}_{\kappa,\square}(L)\right\}$.

Remark 5.2. *With reference to the previous lemma, say that the pair $(\beta,s)\in\Delta'\times\mathbb{S}$ is of type $j\in\{1,\dots,C\}$ if $j(\beta,s) = j$. The restriction $s\in\mathbb{S}$, see* (5.7)*, and properties 1. and 2. ensure the inclusion*

$$L(\beta,s)\subset B\left(\Pi_{e_n^\perp}\beta, 3^5\sqrt{d}s^{-1}\right)\subset\Pi_{e_n^\perp}\widetilde{\Delta}'.$$

Proof of Lemma 5.1. For this proof we denote by $Q(\xi,r)$ the cube with center $\xi\in\mathbb{R}^d$ and sidelength $2r$ and by $\|\cdot\|_\infty$ the ℓ^∞-norm on $\mathbb{R}^d$. First of all, using standard grid techniques as in [32] we may find triadic grids $\mathscr{G}_1,\dots,\mathscr{G}_C$ with the property that for each cube $P\subset\mathbb{R}^d$ there is $j\in\{1,\dots,C\}$ and $L(P)\in\mathscr{G}_j$ with $P\subset L(P)\subset(1+3^{-(\kappa+9)})P$. As a particular consequence

$$|c(P)-c(L(P))|\leqslant\sqrt{d}3^{-(\kappa+9)}\ell(P)\leqslant\sqrt{d}3^{-(\kappa+9)}\ell(L(P)),$$
$$\|c(P)-c(L(P))\|_\infty\leqslant 3^{-(\kappa+9)}\ell(P)\leqslant 3^{-(\kappa+9)}\ell(L(P)).$$

We have that $\Pi_{e_n^\perp} B(\beta, 3^2 s^{-1}) \subset B(\Pi_{e_n^\perp}\beta, 3^3 s^{-1}) \subset Q(\Pi_{e_n^\perp}\beta, 3^3 s^{-1})$. There exists $L = L(\beta, s) \in \mathscr{G}_j$ for some $j \in \{1, \dots, C\}$ such that

$$Q(\Pi_{e_n^\perp}\beta, 3^3 s^{-1}) \subseteq L \subseteq (1 + 3^{-(\kappa+9)}) Q(\Pi_{e_n^\perp}\beta, 3^3 s^{-1})$$

from which 1. follows immediately. In order to see 2. let $\eta \in B(\beta, 3^{-\kappa} s^{-1})$ so that $\Pi_{e_n^\perp}\eta \in B(\Pi_{e_n^\perp}\beta, 3^{-\kappa+1} s^{-1})$. Then we have the chain of inequalities

$$\|\Pi_{e_n^\perp}\eta - c(L)\|_\infty \leqslant |\Pi_{e_n^\perp}\beta - \Pi_{e_n^\perp}\eta| + \|\Pi_{e_n^\perp}\beta - c(L)\|_\infty \leqslant 3^{-\kappa+1} s^{-1} + 3^{-(\kappa+9)}\ell(L) \leqslant 3^{-\kappa-1}\ell(L).$$

This shows that $\Pi_{e_n^\perp}\eta \in Q(c(L), 3^{-\kappa-1}\ell(L)) = 3^{-1} L^{\circ,\kappa}$ and so $\eta \in \Pi_{e_n^\perp}^{-1}(L^{\circ,\kappa})$ as desired.

In order to complete the proof it suffices to show that if $\eta \in B(\beta, 3^2 s^{-1}) \setminus B(\beta, 3^{-1} s^{-1})$ then $\Pi_{e_n^\perp}\eta \in Q$ for some $Q \in \mathrm{ch}_{\kappa,\square}(L)$. To that end let $\eta \notin B(\beta, 3^{-1} s^{-1})$; we will have that $\Pi_{e_n^\perp}\eta \notin B(\Pi_{e_n^\perp}\beta, 3^{-2} s^{-1})$ and so

$$\begin{aligned}\|\Pi_{e_n^\perp}\eta - c(L)\|_\infty &\geqslant \frac{1}{\sqrt{d}} |\Pi_{e_n^\perp}\eta - \Pi_{e_n^\perp}\beta| - \|\Pi_{e_n^\perp}\beta - c(L)\|_\infty \\ &\geqslant \frac{3^{-2} s^{-1}\sqrt{d}}{d} - 3^{-(\kappa+9)} 3^4 s^{-1} \geqslant \sqrt{d} 3^{5-\kappa}\ell(L).\end{aligned}$$

Since $\Pi_{e_n^\perp}\eta \in B(\Pi_{e_n^\perp}\beta, 3^3 s^{-1}) \subseteq L$ there exists $Q \in \mathrm{ch}_{\kappa,\square}(L)$ such that $\Pi_{e_n^\perp}\eta \in Q$. Then for all $y_1 \in Q$ and $y_2 \in L^{\circ,\kappa}$ we calculate

$$\begin{aligned}\|y_1 - y_2\|_\infty &\geqslant \sqrt{d} 3^{5-\kappa}\ell(L) - \|y_1 - \Pi_{e_n^\perp}\eta\|_\infty - \|y_2 - c(L)\|_\infty \geqslant \sqrt{d} 3^{5-\kappa}\ell(L) - 2 3^{-\kappa}\ell(L) \\ &\geqslant \sqrt{d} 3^{4-\kappa}\ell(L)\end{aligned}$$

which in turn implies that $\mathrm{dist}(Q, L^{\circ,\kappa}) \geqslant 3^{4-\kappa}\ell(L)$. We have showed that $Q \in \mathrm{ch}_{\kappa,\square}(L)$ so the proof is complete. □

More generally, if X is any set, any collection $\mathscr{G} \subset \mathscr{P}(X)$ satisfying the grid property g3. is referred to as a *grid*. The forthcoming definitions construct suitable rotated grid systems in $\mathbb{R}^n$ whose purpose is to provide spatial localization.

5.2.1. *Spatial grids.* Fix a standard triadic grid $\mathscr{L}$ of pace $(1,0)$ in $\mathbb{R}^d$ and define the subsets of $\mathbb{R}^n$

$$\mathscr{R} := \left\{L \times [j, j+1)) : L \in \mathscr{L}, \ell(L) \geqslant 1, j \in \mathbb{Z}\right\}.$$

The collection $\mathscr{R} = \mathscr{R}_{e_n}$ of *plates parallel to* $e_n^\perp \in \mathbf{Gr}(d,n)$ obviously inherits property g3. from $\mathscr{L}$ and the unit length partition of $\mathbb{R}$. If $\mathscr{R} \ni R = L \times J$, call L the *horizontal component* and J the *vertical component* of R. The condition $\ell(L) \geqslant 1$ in the definition of $\mathscr{R}$ guarantees that the elements $L \in \mathscr{R}$ have horizontal components of larger scale than their vertical components, so they can be in general understood as plates in $\mathbb{R}^{d+1}$ of *scale* $\mathrm{scl}(R) := \ell(L) = 3^k$, $k \in \mathbb{N}$, and thickness 1 along the direction e_n. For $\beta \in \mathbb{S}^d$ the rotated grid $\mathscr{R}_\beta$ of *plates parallel to* $\beta^\perp \in \mathbf{Gr}(d,n)$ is then defined by

$$\mathscr{R}_\beta := \left\{O_{\beta^\perp, e_n^\perp} Q : Q \in \mathscr{R}\right\}.$$

It is convenient to associate to an element $R \in \mathscr{R}_\beta$ the translation and anisotropic rescaling operator

$$\mathrm{Sy}_R^p f(x) := \frac{1}{|R|^{\frac{1}{p}}} f\left(\frac{\Pi_{\beta^\perp}(x - c(R))}{|\Pi_{\beta^\perp} R|} + \frac{\Pi_\beta(x - c(R))}{|\Pi_\beta(R)|} \right), \qquad x \in \mathbb{R}^n.$$

The sense of this definition is that if f is a bump function around the origin of $\mathbb{R}^n$ then $\mathrm{Sy}_R^p f$ is an L^p-normalized bump adapted to the plate $R \in \mathscr{R}_\beta$.

5.2.2. *Tiles.* Let $\mathscr{G}$ be any standard triadic grid on $\mathbb{R}^d$. Define the *admissible tiles* $t \in \mathscr{T}_\mathscr{G}$ generated by $\mathscr{G}$ as the collection of cartesian products $t = R_t \times Q_t$ where

1. $Q_t \in \mathscr{G}$, $Q_t \subset \Pi_{e_n^\perp} \widetilde{\Delta}'$;
2. $R_t = L_t \times I_t \in \mathscr{R}_{v_t}$, where v_t is any fixed choice of vector in the set $\Pi_{e_n^\perp}^{-1} Q_t^{\circ,\kappa} \subset \mathbb{S}^d$; we will make that choice specific in (5.19) below;
3. $\mathrm{scl}(R_t)\ell(Q_t) \in [1,3)$.

The *scale* of $t \in \mathscr{T}_\mathscr{G}$ is $\mathrm{scl}(t) := \mathrm{scl}(R_t)$. The *frequency support* $\omega_t \subset \mathbb{R}^n$ of a tile t is given by

$$\omega_t := \left\{ \xi \in \mathrm{Ann}(\tfrac{1}{2}, 2) : \Pi_{e_n^\perp} \xi' \in Q_t^{\circ,\kappa} \right\} \subset \mathbb{R}^n.$$

Namely, ω_t consists of those points of $\mathrm{Ann}(\frac{1}{2}, 2)$ whose projection on the sphere is contained in the preimage of the κ-th center of Q_t, denoted now by $Q_t^{\circ,\kappa}$; see Figure 3. Note that the collection $\Omega_\mathscr{G} := \{\omega_t : t \in \mathscr{T}_\mathscr{G}\}$ is a grid on $\mathbb{R}^n$, as it inherits property g3. from $\mathscr{G}$.

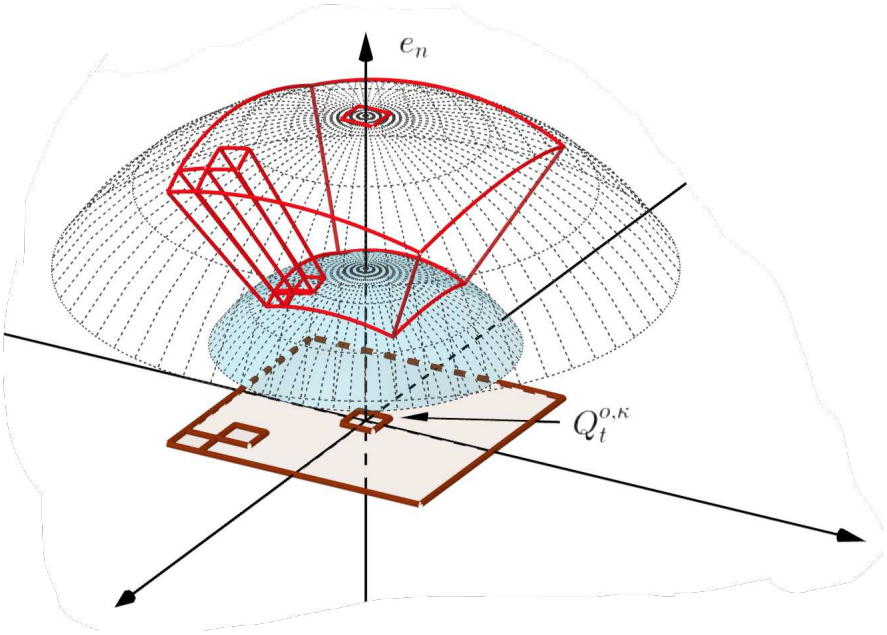

FIGURE 3. The frequency component of a tile and its κ-children

Let $Q_0 = [0,1)^d$ be considered as a triadic cube and let us fix an arbitrary enumeration of the elements of $\mathrm{ch}_{\kappa,\square}(Q_0)$

$$\mathrm{ch}_{\kappa,\square}(Q_0) = \bigsqcup_{\tau=1}^{N} Q_{0,\tau}, \qquad N = c_d 3^{\kappa d},$$

where c_d is a dimensional constant. Note that this indexing of the peripheral κ-children of Q_0 induces an indexing of the peripheral κ-children in $\mathrm{ch}_{\kappa,\square}(Q_t)$ for any tile t, in a way that is independent of the particular choice of tile t and grid $\mathscr{G}$. The *directional support* $\alpha_t \subset \mathbb{S}^d$ of a tile t is given by

$$\begin{aligned} &\alpha_t := \bigsqcup_{\tau=1,\dots,N} \alpha_{t,\tau}, \\ &\alpha_{t,\tau} := \Pi_{e_n^\perp}^{-1}(Q_{t;\tau}), \qquad Q_{t;\tau} \in \mathrm{ch}_{\kappa,\square}(Q_t), \qquad \tau \in \{1,\dots,N\}. \end{aligned} \tag{5.9}$$

The collections $A_{\mathscr{G},\tau} := \{\alpha_{t,\tau} : t \in \mathscr{T}_{\mathscr{G}}\}$ are grids on $\mathbb{S}^d$, for each $1 \leqslant \tau \leqslant N$.

5.2.3. *Adapted classes and model sums.* Hereafter, let Θ_M stand for the unit ball of the Banach space of functions

$$u \in \mathscr{C}^M(\mathbb{R}^n), \qquad \|u\|_{\star,M} := \sup_{0\leqslant|\alpha|\leqslant M} \big\| \langle x\rangle^M \partial_x^\alpha u(x) \big\|_\infty < \infty,$$

where $\langle x\rangle := (1+|x|^2)^{\frac{1}{2}}$. We will often use the special bump function $\chi_M \in \mathscr{C}^\infty$ given by

$$\chi_M(x) := \langle x\rangle^{-M}.$$

Given a tile $t = R_t \times Q_t$ and a large positive integer $M \gg 1$ we define $\mathscr{F}_t^M$ as the collection of complex-valued functions φ on $\mathbb{R}^n$ with the properties (φ are L^2-normalized bump functions adapted to R_t)

(i) $\varphi \in \Big\{ \mathrm{Mod}_{v_t} \mathrm{Sy}_{R_t}^2 \Phi : \Phi \in \Theta_M \Big\}, \qquad \mathrm{Mod}_v g(x) := e^{2\pi i \langle x, v\rangle} g(x), \qquad x \in \mathbb{R}^n, \quad v \in \mathbb{S}^d;$

(ii) $\mathrm{supp}\, \widehat{\varphi} \subseteq \omega_t$.

Further, define $\mathscr{A}_t^M$ as the collection of complex-valued functions $\vartheta = \vartheta(x,\sigma)$ on $\mathbb{R}^n \times \Sigma$ satisfying

(iii) $\vartheta(\cdot,\sigma) \in \mathscr{F}_t^M$ for all fixed $\sigma \in \Sigma$;

(iv) $\vartheta(\cdot,\sigma) = 0$ for all $\sigma \in \Sigma$ such that $v_\sigma \notin \alpha_t$;

(v) ϑ verifies the adaptedness condition

$$|\vartheta(\cdot,\sigma) - \vartheta(\cdot,\rho)| \leqslant \max\left\{ \mathrm{scl}(t)\mathrm{dist}(\sigma,\rho), \frac{1}{\log(\mathrm{e} + \big[\mathrm{dist}(\sigma,\rho)\big]^{-1})} \right\} \mathrm{Sy}_{R_t}^2 \chi_M$$

for all $\sigma, \rho \in \mathbf{Gr}(d,n)$ with $\mathrm{scl}(t)\mathrm{dist}(\sigma,\rho) \lesssim 1$.

Note that both classes are L^2-normalized.

Remark 5.3. *The point of condition (v) is that in a tree, cf. Definition 6.1, with top direction σ there are $\lesssim |\log(\mathrm{dist}(\sigma(x),\sigma))|$ frequency scales contributing at each point, and the total contribution of the difference $|\vartheta_t(x,\sigma(x)) - \vartheta_t(x,\sigma)|$ is summable.*

At this point, define the tile coefficient maps

$$F_M[f](t) := \sup_{\varphi\in\mathscr{F}_t^M} \big|\langle f,\varphi\rangle\big| \tag{5.10}$$

and, for a fixed measurable choice $\sigma:\mathbb{R}^n\to\mathbf{Gr}(d,n)$ with the property that $\sigma(x)\in\Sigma$ for all $x\in\mathbb{R}^n$, and $1\leqslant\tau\leqslant N$

$$A_{\sigma,\tau,M}[g](t) := \sup_{\vartheta\in\mathscr{A}_t^M} \big|\big\langle g,\vartheta(\cdot,\sigma(\cdot))\,\mathbf{1}_{\alpha_{t,\tau}}\big(v_{\sigma(\cdot)}\big)\big\rangle\big|. \tag{5.11}$$

The model bisublinear operator

$$\Lambda_{\mathbb{P};\sigma,\tau,M}(f,g)=\sum_{t\in\mathbb{P}} F_M[f](t)\,A_{\sigma,\tau,M}[g](t) \tag{5.12}$$

is then associated to each subset $\mathbb{P}\subset\mathscr{T}_{\mathscr{G}}$. Momentarily, it will be shown that the duality form for the maximal operator (5.5) is controlled by a suitable combination of finitely many forms (5.12).

5.3. **A homogeneous partition of unity.** Our goal here is to produce a single scale Gabor decomposition of the frequency region Δ defined in (5.6) which contains the Fourier support of the multiplier $P_0\circ P_{\mathrm{cn}}$.

Fix a (spatial) scale $s\in\mathbb{S}$, cf. (5.7), and recall that if $\beta\in\Delta'$, this choice guarantees that $B(\beta,3^6s^{-1})\subset\widetilde{\Delta}'$. We remember that $\kappa=12+\log_3 d$ is fixed. Let $\mathscr{B}_{s^{-1}}$ be a $3^{-\kappa}s^{-1}$-net on Δ', in the sense that

- the sets $\big\{B(\beta,3^{-\kappa}s^{-1}):\beta\in\mathscr{B}_{s^{-1}}\big\}$ form a finitely overlapping cover of Δ';
- the sets $\big\{B(\beta,3^{-(\kappa+1)}s^{-1}):\beta\in\mathscr{B}_{s^{-1}}\big\}$ are pairwise disjoint.

Now let $\{\theta_\beta\}_{\beta\in\mathscr{B}_{s^{-1}}}$ be a smooth, 0-homogeneous partition of unity on Γ_1, consisting of nonnegative real-valued functions, which is subordinate to $\{B(\beta,3^{-\kappa}s^{-1}):\beta\in\mathscr{B}_{s^{-1}}\}$. By this, we mean

$$\operatorname{supp}\theta_\beta\subset B(\beta,3^{-\kappa}s^{-1})\qquad\text{and}\qquad\sum_{\beta\in\mathscr{B}_{s^{-1}}}\big[\theta_\beta(\eta')\big]^2=1\qquad\forall\eta\in\Gamma_1,\quad \eta':=\frac{\eta}{|\eta|}, \tag{5.13}$$

and that the k-th order tangential derivatives of θ_β are of the order $O_{k,n,\delta}(s^k)$ uniformly in $\beta\in\mathscr{B}_{s^{-1}}$. Note also that (5.13) implies that $\theta_\beta\equiv 1$ on $B(\beta,3^{-(\kappa+1)}s^{-1})$. For $\beta\in\mathscr{B}_{s^{-1}}$, consider the function $\phi_\beta\in\mathscr{S}(\mathbb{R}^n)$ whose Fourier transform is given by

$$\widehat{\phi_\beta}(\xi):=\theta_\beta\big(\xi'\big)\zeta(|\xi|),\qquad \xi\in\mathbb{R}^n\setminus\{0\}. \tag{5.14}$$

By construction,

$$\operatorname{supp}\widehat{\phi_\beta}\subset\big\{\xi\in\operatorname{Ann}\big(\tfrac12,2\big):\xi'\in B\big(\beta,3^{-\kappa}s^{-1}\big)\big\}. \tag{5.15}$$

and (5.13) then ensures

$$(P_0\circ P_{\mathrm{cn}})f=\sum_{\beta\in\mathscr{B}_{s^{-1}}} P_0\circ P_{\mathrm{cn}}f*\phi_\beta*\phi_\beta.$$

For each $\beta \in \mathscr{B}_{s^{-1}}$, consider the lattice $Z(\beta) := O_{\beta^\perp, e_n^\perp}\left[s\mathbb{Z}^d \times \mathbb{Z}\right]$. Using Fourier series on the support of each $\widehat{\phi_{s,\beta}}$ we obtain

$$P_0 \circ P_{\mathrm{cn}} f = \lim_{T\to+\infty} \frac{1}{T^n} \int_{[0,T]^n} s^d \sum_{\beta\in\mathscr{B}_{s^{-1}}} \sum_{z\in Z(\beta)} \langle (P_0\circ P_{\mathrm{cn}})f, \mathrm{Tr}_{z+y}\phi_\beta\rangle \mathrm{Tr}_{z+y}\phi_\beta \,\mathrm{d}y, \tag{5.16}$$

for all $f \in \mathscr{S}(\mathbb{R}^n)$, where $\mathrm{Tr}_u h(x) := h(x-u)$ for $x, u \in \mathbb{R}^n$.

5.4. **Discretization of the kernel.** We turn to the discretization of the variable kernels in (5.5). Choose a non-negative, smooth radial function Ψ on $\mathbb{R}^n$ with

$$\operatorname{supp}\Psi \subset \mathrm{Ann}\left(\tfrac{8}{9}, \tfrac{28}{9}\right), \qquad \sum_{s\in 3^{\mathbb{Z}}} \Psi(s\xi) = 1 \qquad \forall \xi \in \mathbb{R}^n \setminus \{0\},$$

while Φ is a smooth radial cutoff equal to 1 on $\mathrm{Ann}(1, \frac{3}{2})$ and vanishing off $\mathrm{Ann}(\frac{1}{2}, 2)$. These definitions entail

$$\begin{aligned} T[P_0\circ P_{\mathrm{cn}} f](\cdot,\sigma) &= \sum_{s\in 3^{\mathbb{Z}}} P_0\circ P_{\mathrm{cn}} f * \psi_s(\cdot,\sigma), \\ \psi_s(x,\sigma) &:= \int_{\mathbb{R}^n} m_\sigma(O_\sigma \Pi_\sigma \xi)\Psi(s\Pi_\sigma\xi)\Phi(\xi) e^{2\pi i\langle x,\xi\rangle}\,\mathrm{d}\xi, \qquad (x,\sigma)\in\mathbb{R}^n\times\Sigma. \end{aligned} \tag{5.17}$$

First, we use condition (5.3) on the frequency support of $P_{\mathrm{cn}} f$ to deduce that $P_{\mathrm{cn}} f * \psi_s = 0$ unless $s \in \mathbb{S}$, cf. (5.7), and consequently restrict the sum in (5.17) to $s \in \mathbb{S}$. Subsequently, use (5.16) to estimate

$$|T[P_0\circ P_{\mathrm{cn}} f])(\cdot,\sigma)| \leqslant \lim_{T\to\infty} \int_{[0,T]^n} \Big| \sum_{s\in\mathbb{S}} s^d \sum_{\beta\in \mathrm{B}_{s^{-1}}} \sum_{z\in Z(\beta)} \langle \mathrm{Tr}_{-y} g, \mathrm{Tr}_z\phi_\beta\rangle \mathrm{Tr}_y(\mathrm{Tr}_z\phi_\beta * \psi_s(\cdot,\sigma))\Big| \frac{\mathrm{d}y}{T^n},$$

with $g := P_0\circ P_{\mathrm{cn}} f$. The norm estimates for $U^\star$ of (5.5) will follow from corresponding estimates for the maximal operator

$$f \mapsto \sup_{\sigma\in\Sigma} \Big| \sum_{s\in\mathbb{S}} s^d \sum_{\beta\in\mathscr{B}_{s^{-1}}} \sum_{z\in Z(\beta)} \langle f, \mathrm{Tr}_z\phi_\beta\rangle \mathrm{Tr}_z\phi_\beta * \psi_s(\cdot,\sigma)\Big|. \tag{5.18}$$

To each triplet (β, s, z) appearing in the sum (5.18), assume (β, s) is of type j in accordance to Remark 5.2 and associate a tile

$$t = R_t \times Q_t = t(\beta, s, z) \in \mathscr{T}_{\mathscr{G}_j}, \tag{5.19}$$

where $\mathscr{G}_j$ is one of the grids appearing in the conclusion of Lemma 5.1, as follows. Firstly, $Q_t = L(\beta, s)$ and $v_t := \beta \in \Pi_{e_n^\perp}^{-1} Q_t^{\circ,\kappa} \subset \mathbb{S}^d$. Secondly, R_t is the unique element of $\mathscr{R}_\beta$ with $z \in R_t$ and $\mathrm{scl}(R_t)\ell(L(\beta,s)) \in [1,3)$. Of course, $\mathrm{scl}(t) \sim s$. Recall the definitions of the collections $\mathscr{F}_t^M$ and $\mathscr{A}_t^M$ from §5.2.3. The following lemma will be used in §6 for $M = 50(d+1)$.

Lemma 5.4. *Let M be a large integer. There is $A = A(M,d)$ such that if* $\mathbf{m}$ *satisfies assumption* (1.10) *the following holds. With reference to the expressions in* (5.18)*, if $t = t(\beta, s, z)$ define the functions*

$$\varphi_t(x) := \mathrm{scl}(t)^{\frac{d}{2}} \mathrm{Tr}_z\phi_\beta(x), \quad \vartheta_t(x,\sigma) := \mathrm{scl}(t)^{\frac{d}{2}} \int_{\mathbb{R}^n} \phi_\beta(y-z)\psi_s(x-y,\sigma)\,\mathrm{d}y, \quad (x,\sigma)\in\mathbb{R}^n\times\Sigma.$$

Then

$$c\varphi_t \in \mathscr{F}_t^M, \qquad c\vartheta_t \in \mathscr{A}_t^{M-1}$$

where the constant $c>0$ may be chosen uniformly in β, s, z.

Proof. Firstly, property (ii) in the definition of the class $\mathscr{F}_t^M$ follows by (5.15) together with condition 2. of Lemma 5.1. We now check condition (i) for φ_t. By rotation and translation we can assume that $t = t(e_n, s, 0)$, namely $v_t = \beta = e_n$, and write $\mathbb{R}^d \otimes \mathbb{R} \ni x =: (y, x_n)$, where $s = \mathrm{scl}(t)$. Using property (ii) we readily see that the function

$$\widetilde{1}_B(x) := \mathrm{Mod}_{-e_n} s^{d/2} \varphi_t(sy, x_n), \qquad x = (y, x_n) \in \mathbb{R}^d \times \mathbb{R},$$

is a $\mathscr{C}^\infty$ bump function with $\|\widetilde{\mathbf{1}}_B\|_1 \eqsim 1$, which is adapted of any order to the unit ball of $\mathbb{R}^n$, centered at the origin. Thus $c\widetilde{\mathbf{1}}_B \in \Theta_M$ as defined in §5.2.3 for any order M, where $c>0$ is a normalization constant depending only upon dimension. We get that

$$c\varphi_t(x) = c\mathrm{Mod}_{e_n} s^{-d/2} \widetilde{1}_B(y/s, x_n), \qquad x = (y, x_n) \in \mathbb{R}^d \times \mathbb{R},$$

which shows (i). These two remarks show that $c\varphi_t \in \mathscr{F}_t^M$ for any desired large positive integer M.

We move to the verification of property $c\vartheta_t \in \mathscr{A}_t^M$. For this let us fix $\sigma \in \Sigma$ and note that $\vartheta_t(\cdot, \sigma) = \phi_\beta(\cdot - z) * \psi_s(\cdot, \sigma)$. In view of (5.14) and (5.17), this immediately implies the Fourier support condition

$$\mathrm{supp}\, \widehat{\vartheta_t(\cdot, \sigma)} \subseteq \mathrm{supp}\, \widehat{\phi_t} \subseteq \omega_t.$$

We now check condition (iv) for ϑ_t. Recall that the tile t has been chosen so that α_t contains $\mathrm{Ann}\big(\beta, 3^{-1}s^{-1}, 3^2 s^{-1}\big)$, and that $\mathrm{supp}\, \widehat{\phi}_\beta \subset \{\xi \in \mathrm{Ann}(1/2, 2) : \xi' \in B(\beta, 3^{-\kappa}s^{-1})\}$ which tells us that whenever $\xi \in \mathrm{supp}\, (\mathrm{Tr}_z \phi_\beta * \psi_s(\cdot, \sigma))^\wedge$ the following inequalities

$$\frac{4}{9} < s|\Pi_\sigma \xi'| < \frac{56}{9}, \qquad |\xi' - \beta| \leqslant 3^{-\kappa} s^{-1}, \qquad \xi' := \xi/|\xi| \in \mathbb{S}^d,$$

must hold. Furthermore, we have

$$|\Pi_\sigma \xi'| + 3^{-\kappa} s^{-1} \geqslant |\Pi_\sigma \xi'| + |\xi' - \beta| \geqslant |\Pi_\sigma \beta| \geqslant |\Pi_\sigma \xi'| - |\xi' - \beta| \geqslant |\Pi_\sigma \xi'| - 3^{-\kappa} s^{-1}$$

which, in view that $\kappa = 12 + \log_3 d$ yields

$$|\Pi_\sigma \beta| \in s^{-1}\big[\tfrac{1}{3}, 7\big] \quad \text{which implies} \quad v_\sigma \in \mathrm{Ann}\big(\beta, \tfrac{1}{3}s^{-1}, 9s^{-1}\big).$$

This proves that $\vartheta_t(\cdot, \sigma) = 0$ unless $v_\sigma \in \mathrm{Ann}\big(\beta, 3^{-1}s^{-1}, 3^2 s^{-1}\big)$. Condition 3. of Lemma 5.1 then entails (iv) for $\vartheta = \vartheta_t$.

We now move to our main task which is to verify that $c\vartheta_t(\cdot, \sigma)$ verifies conditions (iii) and (v) in the definition of the class $\mathscr{A}_t^M$. We can again reduce to the case $v_t = \beta = e_n$ and $z = 0$, where we recall that $t = t(\beta, s, z)$; we write $x = (y, x_n) \in \mathbb{R}^d \otimes \mathbb{R}$ and $\sigma \in \Sigma$, and consider $v_\sigma \in \mathrm{Ann}\big(e_n, 3^{-1}s^{-1}, 3^2 s^{-1}\big)$. Now, express x in terms of $\Pi_\sigma x, \Pi_{\sigma^\perp} x$ as in Lemma 2.2,

$$\begin{aligned} y - \Pi_\sigma x &= q_1(\langle x, u\rangle, \langle x, v_\sigma\rangle)\, u + q_2(\langle x, u\rangle, \langle x, v_\sigma\rangle)\, v_\sigma, \\ x_n - \Pi_{\sigma^\perp} x &= q_2(\langle -x, v_\sigma\rangle, \langle x, u\rangle)\, u + q_1(\langle x, v_\sigma\rangle, \langle -x, u\rangle)\, v_\sigma, \\ q_1(t, w) &:= \left[\frac{\sin 2\theta}{2} w - (\sin\theta)^2 t\right], \qquad q_2(t, w) := \left[\frac{\sin 2\theta}{2} t + (\sin\theta)^2 w\right]. \end{aligned} \tag{5.20}$$

Above θ is the angle between v_σ and e_n and $u \in \sigma$ is a unit vector orthogonal to $\sigma \cap e_n^\perp$. By the already verified directional support condition (iv) we know that $\theta \simeq |v_\sigma - e_n| \lesssim s^{-1}$, which implies that

$$|q_1(y,w)| + |q_2(y,w)| \lesssim s^{-1}\big(|y| + |w|\big). \tag{5.21}$$

Combining (5.20) with (5.21) shows that the decay rates

$$\big[\big(1 + s^{-1}|\Pi_\sigma x|\big)\big(1 + |\Pi_{\sigma^\perp} x|\big)\big]^{-M}, \qquad \big[\big(1 + s^{-1}|y|\big)(1 + |x_n|)\big]^{-M} \tag{5.22}$$

are equivalent. Let us define the dilation operator Dil_σ^s for $\sigma \in \mathbf{Gr}(d,n)$ as

$$\mathrm{Dil}_\sigma^s g(x) := s^{d/2} g(s\Pi_\sigma x + \Pi_{\sigma^\perp} x), \qquad x = (\Pi_\sigma x, \Pi_{\sigma^\perp} x) \in \mathbb{R}^n.$$

In view of (5.22) it will be enough to show that for some dimensional constant $c > 0$, the rescaled function $\mathbb{R}^n \ni x \mapsto c\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s F(x)$ is a bump function adapted to the unit ball of $\mathbb{R}^n$, centered at the origin, where

$$F = \vartheta_t(\cdot,\sigma) \qquad \text{or} \qquad F = \vartheta_t(\cdot,\sigma) - \vartheta_t(\cdot,\rho), \qquad \rho,\sigma \in \Sigma, \quad \rho \neq \sigma.$$

We begin with the first case where $F = \vartheta_t(\cdot,\sigma)$. Assuming again that $t = t(e_n, s, 0)$ we note that for every multiindex α

$$\begin{aligned}\partial_x^\alpha[\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s \vartheta_t(\cdot,\sigma)] &= [\partial_x^\alpha \mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s \varphi_t] * \mathrm{Mod}_{-e_n} s^{d/2}\mathrm{Dil}_\sigma^s \psi_s(\cdot,\sigma)\\ &=: [\partial_x^\alpha \mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s \varphi_t] * \mathrm{Mod}_{-e_n} G_1.\end{aligned}$$

We have already seen that the function $\mathrm{Mod}_{-e_n} s^{d/2}\varphi_t(sy, x_n)$, with $x = (y, x_n) \in \mathbb{R}^d \times \mathbb{R}$ is a smooth bump adapted at every order to the unit ball of $\mathbb{R}^n$ centered at the origin. By the remark following (5.21) this is equivalent to the fact that the function $\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s \varphi_t(x)$ has the same adaptedness property. Thus, in order to show that $\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s \vartheta_t$ is adapted to the unit ball of $\mathbb{R}^n$ centered at the origin at order M it suffices to show that the function G_1 has decay of order M at scale 1 away from the origin. Then

$$\widehat{G_1}(\xi) := (s^{d/2}\mathrm{Dil}_\sigma^s \psi_s(\cdot,\sigma))^\wedge(\xi) = m_\sigma(\tfrac{1}{s} O_\sigma \Pi_\sigma \xi)\Psi(\Pi_\sigma \xi)\Phi(\tfrac{1}{s}\Pi_\sigma \xi + \Pi_{\sigma^\perp}\xi).$$

We record the easy estimate $|\mathrm{supp}\,\widehat{G}_1| \lesssim_n 1$ which is due to the fact that on the support of the function $\widehat{G}_1$ we have

$$|\Pi_\sigma \xi| \eqsim 1 \quad \text{and} \quad |\Pi_\sigma \xi|^2/s^2 + |\Pi_{\sigma^\perp}\xi|^2 \eqsim 1 \implies |\Pi_{\sigma^\perp}\xi| \eqsim 1. \tag{5.23}$$

It is then apparent that for $M \leqslant A$ we have

$$\Big\| \partial_\xi^\gamma \widehat{G}_1(\xi) \Big\|_{L^\infty_\xi} \lesssim 1, \qquad 0 \leqslant |\gamma| \leqslant M.$$

Here, we crucially use (5.23), allowing the exploitation of the Hörmander–Mihlin condition on m_σ in the form

$$\Big|\partial_\xi^\gamma m_\sigma(s^{-1} O_\sigma \Pi_\sigma \xi)\Big| \lesssim s^{-|\gamma|}\big|\tfrac{1}{s}\Pi_\sigma \xi\big|^{-|\gamma|} \|m_\sigma\|_{\mathscr{M}_A(d)} \lesssim 1, \qquad |\gamma| \leqslant A.$$

This completes the proof of adaptation of $F = \vartheta_t(\sigma,\cdot)$.

Let now $F = \vartheta_t(\cdot,\sigma) - \vartheta_t(\cdot,\rho)$. Firstly note that we can assume that ρ, σ are such that $v_\sigma, v_\rho \in \alpha_t$ (which was defined in (5.9)) for $t = t(e_n, s, 0)$, see (5.19). Indeed, if say $\rho \notin$

α_t then $\vartheta_t(\cdot,\rho)=0$ by the directional support condition (iv), which is already verified. Then $\mathrm{dist}(\rho,\sigma)\gtrsim \mathrm{scl}(t)^{-1}$ and the desired adaptedness property follows from the case $F=\vartheta_t(\cdot,\sigma)$ proved above.

Now, assume that $\rho,\sigma\in\alpha_t$. Using (5.8) and the definition of α_t in (5.9), $\mathrm{dist}(\rho,\sigma)\leqslant 3\sqrt{d}3^{-\kappa}\mathrm{scl}(t)^{-1}\leqslant 3^{-8}s^{-1}$. We can apply the same argument as above to write

$$\begin{aligned}\partial_x^\alpha[\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s F]&=\partial_x^\alpha[\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s\varphi_t]*\mathrm{Mod}_{-e_n}s^{d/2}\mathrm{Dil}_\sigma^s[\psi_s(\cdot,\sigma)-\psi_s(\cdot,\rho)]\\&=:\partial_x^\alpha[\mathrm{Mod}_{-e_n}\mathrm{Dil}_\sigma^s\varphi_t]*\mathrm{Mod}_{-e_n}G_2.\end{aligned}$$

As before it will be enough to show that the function G_2 decays with order M at scale 1 away from the origin; more precisely we will show that $G_2\lesssim c(\sigma,\rho,s)\chi_M$, with $c(\sigma,\rho,s)$ the multiplicative factor appearing in condition (v) of §5.2.3. We will do so by estimating derivatives of sufficiently high order of the Fourier transform of G_2. To calculate the Fourier transform of G_2 it is convenient to first note that the anisotropic rescaling of $\psi_s(\rho,\cdot)$ in the directions of σ creates a small error that will be kept under control by the closeness of σ and ρ. Setting

$$B(\xi):=\Pi_\rho\Pi_\sigma\xi+s\Pi_\rho\Pi_{\sigma^\perp}\xi-\Pi_\rho\xi=(s-1)\Pi_\rho\Pi_{\sigma^\perp}\xi,\quad |B(\xi)|\lesssim s\mathrm{dist}(\sigma,\rho)|\xi|\ll|\xi|,$$

an explicit computation tells us that

$$\begin{aligned}\widehat{G_2}(\xi)=\big[m_\sigma\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)\Psi(\Pi_\sigma\xi)-m_\rho\big(\tfrac{1}{s}O_\rho(\Pi_\rho\xi+B(\xi))\big)\Psi(\Pi_\rho\xi+B(\xi))\big]\\\times\Phi\big(\tfrac{1}{s}\Pi_\sigma\xi+\Pi_{\sigma^\perp}\xi\big).\end{aligned}$$

Note that since $\sigma\in\Sigma$ we always have that $|\xi|\eqsim|\Pi_{\sigma^\perp}\xi|\eqsim 1$ for $s^{-1}\Pi_\sigma\xi+\Pi_{\sigma^\perp}\xi\in\mathrm{supp}(\Phi)$ and in particular we recover that $|\Pi_\sigma\xi|\eqsim|\Pi_\rho\xi|\eqsim|\xi|\eqsim 1$ on the support of $\widehat{G_2}$. Three mean value estimates combined with (1.10) and (2.2) tell us that

$$\begin{aligned}&\Big|\partial_\xi^\gamma\big[\Psi(\Pi_\rho\xi+B(\xi))-\Psi(\Pi_\sigma\xi)\big]\Big|\lesssim|B(\xi)|+|\Pi_\rho\xi-\Pi_\sigma\xi|\lesssim s\mathrm{dist}(\sigma,\rho),\\&\Big|\partial_\xi^\gamma\big[m_\rho\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)-m_\rho\big(\tfrac{1}{s}O_\rho\Pi_\sigma\xi\big)\big]\Big|\lesssim\|O_\rho-O_\sigma\|\lesssim\mathrm{dist}(\sigma,\rho),\\&\Big|\partial_\xi^\gamma\big[m_\rho\big(\tfrac{1}{s}O_\rho\Pi_\sigma\xi\big)-m_\rho\big(\tfrac{1}{s}O_\rho(\Pi_\rho\xi+B(\xi))\big)\big]\Big|\lesssim|B(\xi)|+|\Pi_\rho\xi-\Pi_\sigma\xi|\lesssim s\mathrm{dist}(\sigma,\rho).\end{aligned}\tag{5.24}$$

Hence, we may replace $\widehat{G_2}$ by

$$\widehat{G_3}(\xi):=\big[m_\sigma\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)-m_\rho\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)\big]\Psi(\Pi_\rho\xi+B(\xi))\Phi(s^{-1}\Pi_\sigma\xi+\Pi_{\sigma^\perp}\xi)$$

controlling the difference by means of (5.24). Finally, the main assumption (1.10) on the family $\mathbf{m}$ tells us that

$$\partial_\xi^\gamma\big[m_\sigma\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)-m_\rho\big(\tfrac{1}{s}O_\sigma\Pi_\sigma\xi\big)\big]\lesssim\frac{1}{\log\big(\mathrm{e}+[\mathrm{dist}(\sigma,\rho)]^{-1}\big)}$$

on the support of $\widehat{G_3}$; collecting the above inequalities leads to the conclusion

$$\Big\|\partial_\xi^\gamma\widehat{G}_2(\xi)\Big\|_{L^\infty_\xi}\lesssim(1+s)\mathrm{dist}(\sigma,\rho)+\frac{1}{\log\big(\mathrm{e}+[\mathrm{dist}(\sigma,\rho)]^{-1}\big)},\qquad 0\leqslant|\gamma|\leqslant A-1,$$

which completes the proof of (v). □

With the notation introduced in Lemma 5.4, the right hand side of (5.18) may be rewritten as, and subsequently bounded by, as follows:

$$\sup_{\sigma\in\Sigma}\Big|\sum_{s\in\mathbb{S}}\sum_{\beta\in \mathrm{B}_{s-1}}\sum_{z\in Z(\beta)}\langle f,\varphi_{t(\beta,s,z)}\rangle\vartheta_{t(\beta,s,z)}(\cdot,\sigma)\Big|\leqslant\sum_{j=1}^{C}\sum_{\tau=1}^{N}\mathrm{T}_{j,\tau}f,$$

$$\mathrm{T}_{j,\tau}f:=\sup_{\sigma\in\Sigma}\left|\sum_{\substack{(\beta,s,z)\\ j(\beta,s)=j}}\langle f,\varphi_{t(\beta,s,z)}\rangle\vartheta_{t(\beta,s,z)}(\cdot,\sigma)\mathbf{1}_{\alpha_{t(\beta,s,z),\tau}}(v_\sigma)\right|.$$

Property (iv) of the $\mathscr{A}_t$ class was used above. It thus suffices to control one of the $\mathrm{T}_{j_o,\tau_o}f$ summands for an arbitrary but fixed value of $(j_o,\tau_o)\in\{1,\dots,C\}\times\{1,\dots,N\}$. Choose a measurable function $\sigma=\sigma(x):\mathbb{R}^n\to\Sigma$, linearizing the corresponding supremum in the definition of $\mathrm{T}_{j_o,\tau_o}f$, and $g\in\mathscr{S}(\mathbb{R}^n)$ of unit norm in $L^{p',q'}(\mathbb{R}^n)$ so that

$$\begin{aligned}\|\mathrm{T}_{j_o,\tau_o}f\|_{L^{p,q}(\mathbb{R}^n)}&\leqslant 2\left|\int\sum_{s\in\mathbb{S}}\sum_{\beta\in\mathscr{B}_{s-1}}\sum_{z\in Z(\beta)}\langle f,\varphi_{t(\beta,s,z)}\rangle\vartheta_{t(\beta,s,z)}(x,\sigma(x))\mathbf{1}_{\alpha_{t(\beta,s,z),\tau_o}}(v_{\sigma(x)})g(x)\,\mathrm{d}x\right|\\&\lesssim\Lambda_{\mathscr{T}_{\mathscr{G}_{j_o},M};\sigma,\tau_o}(f,g)\end{aligned}$$

5.5. **Final reductions to a scale-separated model sum.** By a limiting argument and restricted weak-type interpolation, Proposition 3.3 is reduced to proving the two estimates

$$\Lambda_{\mathbb{P};\sigma,\tau_o,M}(f,g\mathbf{1}_E)\lesssim\|f\|_2|E|^{\frac12}\tag{5.25}$$

$$\Lambda_{\mathbb{P};\sigma,\tau_o,M}(\mathbf{1}_F,g\mathbf{1}_E)\lesssim|F|^{\frac1p}|E|^{1-\frac1p}\tag{5.26}$$

uniformly over all $f,g\in L^\infty_0(\mathbb{R}^n)$ with the normalization $\|g\|_\infty=1$, all sets $F,E\subset\mathbb{R}^d$ of finite measure, a fixed choice of triadic grid $\mathscr{G}=\mathscr{G}_{j_o}$ and $\mathbb{P}\subset\mathscr{T}_{\mathscr{G}}$.

Henceforth, we fix $M=50n$ and a pair (j_o,τ_o) as above and write $\Lambda_{\mathbb{P};\tau_o}(f,g)$ in place of $\Lambda_{\mathbb{P};\sigma,\tau_o,M}(f,g)$. By an additional splitting of the dyadic grid $\mathscr{G}$, we can and will assume that the scales are 3^κ separated: namely if $Q,Q'\in\mathscr{G}$ with $\ell(Q)<\ell(Q')$ then $\ell(Q)<3^{-\kappa}\ell(Q')$. Note that this implies a corresponding separation of the spatial scales of tiles in $\mathscr{T}_{\mathscr{G}}$ because of the uncertainty constraint. We will also make the qualitative assumption that $\mathbb{P}$ is a finite collection and prove estimates which are uniform in $\mathbb{P}$. The proofs of (5.25) and (5.26) will be sketched at the end of §6, relying upon the size, density and tree lemmas, also proved in §6 below.

6. Trees, size and density

We fix a finite collection of tiles $\mathbb{P}\subset\mathscr{T}_{\mathscr{G}}$, for a fixed choice of triadic grid $\mathscr{G}$. We remember that tiles are sets of the form $t=R_t\times Q_t$ where $Q_t\in\mathscr{G}$ is a triadic cube in $\mathbb{R}^d\times\{0\}$. Recall the roles of $Q_t^{\circ,\kappa}$ and $\mathrm{ch}_{\kappa,\square}(Q)$; since κ is a fixed large dimensional constant, we will omit it from the notation from now on and write instead Q_t° and $\mathrm{ch}_\square(Q)$ for the *center* and the *peripheral children* of Q, respectively.

We will be sorting these arbitrary collections of tiles into *trees*, defined below. It will be useful to remember the definition of the directional support $Q_{t,\tau}$ from (5.9).

Definition 6.1 (trees)**.** *A* tree *is a collection of tiles* $\mathbf{T} \subseteq \mathbb{P}$ *such that, there exists a pair* $(\xi_{\mathbf{T}}, R_{\mathbf{T}})$ *with* $\xi_{\mathbf{T}} \in \mathbb{R}^d$ *and* $R_{\mathbf{T}} \in \{R_t : t \in \mathscr{T}_{\mathscr{G}}\}$ *such that*

1. $\xi_{\mathbf{T}} \in Q_t$ *for all* $t \in \mathbf{T}$*;*
2. $\mathrm{scl}(R_t) \leqslant \mathrm{scl}(R_{\mathbf{T}})$ *and* $R_t \cap R_{\mathbf{T}} \neq \varnothing$ *for every* $t \in \mathbf{T}$.

We call $(\xi_{\mathbf{T}}, R_{\mathbf{T}})$ *the* top data *of* $\mathbf{T}$*. Let* $\tau \in \{1, \dots, N\}$*; a* tree $\mathbf{T}$ *will be called* τ-lacunary *if* $\xi_{\mathbf{T}} \in Q_{t,\tau} \in \mathrm{ch}_{\square}(Q_t)$ *for all* $t \in \mathbf{T}$*. The tree* $\mathbf{T}$ *will be called* lacunary *if it is* τ*-lacunary for some* $\tau \in \{1, \dots, N\}$ *while it will be called* overlapping *if* $\xi_{\mathbf{T}} \in Q_t \setminus \cup_{\tau=1}^{N} Q_{t,\tau}$.

The following relation of order will be useful: For two tiles $t, t' \in \mathscr{T}_{\mathscr{G}}$ we will define

$$t \leqslant t' \overset{\text{def}}{\iff} Q_{t'} \subseteq Q_t \quad \text{and} \quad R_t \cap R_{t'} \neq \varnothing$$

Note that this is *not* a partial order as it is not transitive. The following geometric lemma will be used in several places throughout the paper.

Lemma 6.2. *There exists a constant* K_n *depending only upon dimension, such that for every* $t, t' \in \mathbb{P}$ *with* $Q_{t'} \subseteq Q_t$*, the following hold.*

- *If* $R_t \cap R_{t'} \neq \varnothing$ *then* $R_t \subseteq K_n R_{t'}$.
- *If* $K_n R_t \cap K_n R_{t'} \neq \varnothing$ *then* $K_n R_t \subseteq K_n^2 R_{t'}$.

Remark 6.3. *The following facts about trees will be used without particular mention.*

- *If* $\mathbf{T}$ *is a tree then there exists a* top tile $\mathrm{top}(\mathbf{T}) = R_{\mathbf{T}} \times Q_{\mathbf{T}}$ *such that* $\xi_{\mathbf{T}} \in Q_t$ *and* $t \leqslant \mathrm{top}(\mathbf{T})$ *for all* $t \in \mathbf{T}$*. In particular we have that for lacunary* $\mathbf{T}$ *there holds* $t \leqslant \mathrm{top}(\mathbf{T})$ *and* $Q_{\mathbf{T}} \subseteq Q_t \setminus Q_t^{\circ}$ *for all* $t \in \mathbf{T}$*. Indeed, note that if* $\mathbf{T}$ *is lacunary then* $\ell(Q_t) > 3^{\kappa} \ell(Q_{\mathbf{T}})$ *for all* $t \in \mathbf{T}$ *because of the separation of scales assumption for the grid* $\mathscr{G}$.
- *Because of the point above, we can and will always assume that* $\xi_{\mathbf{T}}$ *has no triadic coordinates.*
- *If* $\mathbf{T}$ *is a lacunary tree then the collections* $\{Q : Q = Q_t^{\circ} \text{ for some } t \in \mathbf{T}\}$ *and* $\{\omega : \omega = \omega_t \text{ for some } t \in \mathbf{T}\}$ *are pairwise disjoint.*

For the following definition, $E \subset \mathbb{R}^n$ is a measurable set of finite measure. This set remains fixed throughout the paper. Also for any collection $\mathbf{R}$ consisting of of rectangular parallelepipeds in $\mathbb{R}^n$ we define the *shadow* of the collection

$$\mathbf{sh}(\mathbf{R}) := \bigcup_{R \in \mathbf{R}} R.$$

Recall also the definition of the intrinsic coefficients $F[f](t)$ from (5.10). The next step consists of defining two coefficient maps respectively tied to the Carleson measure properties of the coefficients (5.10), (5.11)

Definition 6.4 (Size)**.** *Let* $\mathbb{P} \subset \mathscr{T}_{\mathscr{G}}$ *be a finite collection of tiles. For fixed* $f \in L_0^{\infty}(\mathbb{R}^n)$*, define the map*

$$\mathrm{size} : \mathscr{P}(\mathbb{P}) \to [0, \infty), \qquad \mathrm{size}(\mathbb{Q}) := \sup_{\substack{\mathbf{T} \subseteq \mathbb{Q} \\ \mathbf{T} \text{ lac. tree}}} \Big(\frac{1}{|R_{\mathbf{T}}|} \sum_{t \in \mathbf{T}} F_{10n}[f](t)^2 \Big)^{\frac{1}{2}}.$$

Definition 6.5 (Density)**.** *Let $\mathbb{P}$ be a finite collection of tiles and $t \in \mathbb{P}$. For a fixed measurable set $E \subset \mathbb{R}^n$ of finite measure, we define*

$$E_t := \{x \in E : v_{\sigma(x)} \in \alpha_t\}, \qquad \operatorname{dense}(t) := \sup_{\substack{t' \geqslant t \\ t' \in \mathcal{T}_g}} \int_{E_{t'}} \operatorname{Sy}^1_{R_{t'}} \chi_{10n},$$

as well as the map

$$\operatorname{dense} : \mathscr{P}(\mathbb{P}) \to [0,\infty), \qquad \operatorname{dense}(\mathbb{Q}) := \sup_{t \in \mathbb{Q}} \operatorname{dense}(t).$$

6.1. **Orthogonality estimates for lacunary trees.** We will use different orthogonality estimates for wave packets adapted to special collections of tiles. Most of them are rather standard in the literature but we include them here for completeness.

Lemma 6.6. *Let $\mathbf{T} \subset \mathbb{P}$ be a lacunary tree. Then the following hold*

1. *For adapted families $\{\varphi_t : \varphi_t \in \mathscr{F}_t^M,\ t \in \mathbf{T}\}$, $\{\psi_t : \psi_t \in \mathscr{F}_t^M,\ t \in \mathbf{T}\}$ with $M > 2n$, we have*

$$\sup_{|\varepsilon_t|=1} \Big\| \sum_{t \in \mathbf{T}} \varepsilon_t \langle f, \varphi_t \rangle \psi_t \Big\|_{L^2(\mathbb{R}^n)} \lesssim \Big(\sum_{t \in \mathbf{T}} F_M[f](t)^2 \Big)^{\frac{1}{2}} \lesssim \|f\|_{L^2(\mathbb{R}^n)}.$$

   with implicit dimensional constants.

2. *There holds*

$$\sum_{t \in \mathbf{T}} F_{10n}[f](t)^2 \lesssim \|f\|^2_{L^\infty(\mathbb{R}^n)} |R_{\mathbf{T}}|.$$

Proof. Let $S_{\mathbf{T}} f := \sum_{t \in \mathbf{T}} \varepsilon_t \langle f, \varphi_t \rangle \psi_t$ and $Q(\mathbf{T}) := \{Q_t^\circ : t \in \mathbf{T}\}$. For $Q \in Q(\mathbf{T})$ we also write $\mathbf{T}(Q) := \{t \in \mathbf{T} : Q_t^\circ = Q\}$. Then

$$S_{\mathbf{T}} f = \sum_{Q \in Q(\mathbf{T})} \sum_{t \in \mathbf{T}(Q)} \varepsilon_t \langle f, \varphi_t \rangle \psi_t =: \sum_{Q \in Q(\mathbf{T})} S_Q(f)$$

and because of Remark 6.3 we have

$$\|S_{\mathbf{T}} f\|_2^2 \leqslant \sum_{Q \in Q(\mathbf{T})} \|S_Q f\|_2^2.$$

We next show that for each $Q \in Q(\mathbf{T})$ we have

$$\|S_Q(g)\|_2^2 \lesssim \sum_{t \in \mathbf{T}(Q)} F_M[g](t)^2.$$

Expanding the square we get for each $Q \in Q(\mathbf{T})$

$$\|S_Q(g)\|_2^2 \leqslant \sum_{t,t' \in \mathbf{T}(Q)} |\langle g, \varphi_t \rangle| |\langle g, \varphi_{t'} \rangle| |\langle \psi_t, \psi_{t'} \rangle| \lesssim \sum_{t \in \mathbf{T}(Q)} F_M[g](t)^2 \sup_{t' \in \mathbf{T}(Q)} \sum_{t \in \mathbf{T}(Q)} |\langle \psi_t, \psi_{t'} \rangle|.$$

However $\sup_{t' \in \mathbf{T}(Q)} \sum_{t \in \mathbf{T}(Q)} |\langle \psi_t, \psi_{t'} \rangle| \lesssim 1$ by the fast decay of the wave packets $\{\psi_t : t \in \mathbf{T}(Q)\}$ and the fact that the spatial components $\{R_t : t \in \mathbf{T}(Q)\}$ tile $\mathbb{R}^n$. This proves the first estimate of 1.

To see the second estimate we pick for each $t \in \mathbf{T}$ a $\varphi_t \in \mathscr{F}_t^M$ and write

$$\sum_{t \in \mathbf{T}} \big| \langle f, \varphi_t \rangle \big|^2 = \Big\langle f, \sum_{t \in \mathbf{T}} \langle f, \varphi_t \rangle \varphi_t \Big\rangle \lesssim \|f\|_{L^2(\mathbb{R}^n)} \Big(\sum_{t \in \mathbf{T}} F_M[f](t)^2 \Big)^{\frac{1}{2}}$$

where we used 1. in the last approximate inequality. This readily yields the desired estimate.

Finally for 2. let $\varphi_t \in \mathscr{F}_t^{10n}$ for each $t \in \mathbf{T}$ and note that $R_t \subseteq K_n R_{\mathbf{T}}$ by Lemma 6.2 and let $t_{\mathbf{T}} := \mathrm{Sy}^{\infty}_{K_n R_{\mathbf{T}}} \chi_{4n}$. Now for $c > 0$ we use 2. to estimate

$$\sum_{t\in\mathbf{T}} \big|\langle f, \varphi_t\rangle\big|^2 = \sum_{t\in\mathbf{T}} \big|\langle c^{-1} t_{\mathbf{T}} f, c t_{\mathbf{T}}^{-1} \varphi_t\rangle\big|^2 \lesssim \|f t_{\mathbf{T}}\|^2_{L^2(\mathbb{R}^n)}$$

by noticing that for a suitable choice of c the wave packets $c t_{\mathbf{T}}^{-1} \varphi_t \in \mathscr{F}_t^{6n}$ for each $t \in \mathbf{T}$. The estimate in the last display readily yields 3. □

6.2. **Strongly disjoint families of trees and almost orthogonality.** In what follows we will be sorting our collection of tiles $\mathbb{P}$ into trees selected by means of a greedy algorithm. The selection process will imply certain disjointness properties which we encode in the definition below.

Definition 6.7 (strongly disjoint lacunary families)**.** *Let $\mathscr{T}$ be a family of trees with $\mathbf{T} \subset \mathbb{P}$ for every $\mathbf{T} \in \mathscr{T}$. The family $\mathscr{T}$ is called (lacunary) strongly disjoint if*

1. *Every tree $\mathbf{T} \in \mathscr{T}$ is lacunary.*
2. *If $t \in \mathbf{T} \in \mathscr{T}$ and $t' \in \mathbf{T}' \in \mathscr{T}$ with $\mathbf{T} \neq \mathbf{T}'$ and $Q_t \subseteq Q_{t'}^{\circ}$ then $R_{t'} \cap K_n^2 R_{\mathbf{T}} = \varnothing$*

with K_n the constant of Lemma 6.6.

The point of the definition above is that if conclusion 2. failed then the tile t' would essentially qualify to be included in a completion of the tree $\mathbf{T}$, if that tree was chosen first via a greedy selection algorithm. This point will become apparent in the proof of the size lemma, Lemma 6.13 below. Furthermore, the definition above implies the following property.

Lemma 6.8. *Let $\mathbb{T} := \cup_{\mathbf{T}\in\mathscr{T}} \mathbf{T}$ where $\mathscr{T}$ is a lacunary strongly disjoint family. Then the collection*

$$\mathbb{T}^{\circ} := \{R_t \times Q_t^{\circ} : t \in \mathbb{T}\}$$

is pairwise disjoint.

Proof. Let $t, t' \in \mathbb{T}$. If $Q_t^{\circ} = Q_{t'}^{\circ}$ then either $t \equiv t'$ or $R_{t'} \cap R_t = \varnothing$. Assume then that $Q_t^{\circ} \subsetneq Q_{t'}^{\circ}$ so that $\ell(Q_t) < 3^{-\kappa}\ell(Q_{t'})$, because of separation of scales assumption, and consequently $Q_t \subseteq Q_{t'}^{\circ}$. Since the trees in the collection are lacunary this implies that t, t' must come from different trees, say $\mathbf{T} \neq \mathbf{T}'$, respectively. If $R_t \cap R_{t'} \neq \varnothing$ then by consecutive applications of Lemma 6.2 we would have that $R_{t'} \subseteq K_n^2 R_{\mathbf{T}}$, contradicting the definition of a strongly disjoint family. □

It is well known that strongly disjoint families of trees obey certain almost orthogonality estimates for the corresponding tree projections. The precise statement in the context of this paper is contained in Lemma 6.9 below.

Lemma 6.9. *Let $M > 4n$. For a (lacunary) strongly disjoint family $\mathscr{T}$ with $\mathbb{T} := \cup_{\mathbf{T}\in\mathscr{T}}\mathbf{T}$, there holds*

$$\left(\sum_{\mathbf{T}\in\mathscr{T}}\sum_{t\in\mathbf{T}} F_M[f](t)^2\right)^{\frac12} \lesssim \|f\|_{L^2(\mathbb{R}^n)} + \left(\sup_{t\in\mathbb{T}}\frac{F_M[f](t)}{|R_t|^{\frac12}}\left[\sum_{\mathbf{T}\in\mathscr{T}}|R_{\mathbf{T}}|\right]^{\frac12}\right)^{\frac13}\|f\|^{\frac23}_{L^2(\mathbb{R}^n)}.$$

Proof. Let $\{\varphi_t : \varphi_t \in \mathscr{F}_t^M,\ t\in\mathbb{T}\}$ be any adapted family and note that

$$S(f)^2 := \sum_{t\in\mathbb{T}}\big|\langle f,\varphi_t\rangle\big|^2 \leqslant \|f\|_2\left\|\sum_{t\in\mathbb{T}}\langle f,\varphi_t\rangle\varphi_t\right\|_2.$$

Letting $Q(\mathbb{T}) := \{Q_t : t\in\mathbb{T}\}$ and $\mathbb{T}(Q) := \{t : Q_t = Q\}$ for $Q\in Q(\mathbb{T})$, we have

$$\begin{aligned}\left\|\sum_{t\in\mathbb{T}}\langle f,\varphi_t\rangle\varphi_t\right\|_2^2 &\leqslant \sum_{Q\in Q(\mathbb{T})}\sum_{\substack{t,t'\in\mathbb{T}\\ Q_t=Q_{t'}=Q}}\big|\langle f,\varphi_t\rangle\big|\big|\langle f,\varphi_{t'}\rangle\big|\big|\langle\varphi_t,\varphi_{t'}\rangle\big|\\ &\quad + 2\sum_{t\in\mathbb{T}}\sum_{\substack{t'\in\mathbb{T}\\ Q_{t'}\supsetneq Q_t}}\big|\langle f,\varphi_t\rangle\big|\big|\langle f,\varphi_{t'}\rangle\big|\big|\langle\varphi_t,\varphi_{t'}\rangle\big|\\ &:= S_1(f)^2 + 2S_2(f)^2.\end{aligned}$$

As in the proof of Lemma 6.6 it is relatively easy to see that $|S_1(f)|\lesssim S(f)$ since for a fixed $Q\in Q(\mathbb{T})$, the spatial components $\{R_t : Q_t = Q\}$ tile $\mathbb{R}^n$. We thus focus on $S_2(f)$ which can be estimated in the form

$$S_2(f)^2 \leqslant \sup_{t'\in\mathbb{T}}\frac{|\langle f,\phi_{t'}\rangle|}{|R_{t'}|^{\frac12}}\sum_{t\in\mathbb{T}}|\langle f,\varphi_t\rangle|\sum_{\substack{t'\in\mathbb{T}\\ Q_{t'}^\circ\supseteq Q_t}}|R_{t'}|^{\frac12}|\langle\varphi_t,\varphi_{t'}\rangle|.$$

Note that for fixed $t\in\mathbb{T}$ the collection $\mathbf{R}_{\mathbb{T}}(t) := \{R_{t'} : t'\in\mathbb{T}, Q_t\subseteq Q_{t'}^\circ\}$ is pairwise disjoint and $\mathbf{sh}(\mathbf{R}_{\mathbb{T}}(t))\subseteq(K_n^2R_{\mathbf{T}})^{\mathrm{c}}$; this follows immediately from Lemma 6.8 since $\mathscr{T}$ is a (lacunary) strongly disjoint family, and from the definition of a strongly disjoint family itself. At this point we use the following estimate: For $t,t'\in\mathbb{P}$ such that $R_t\cap R_{t'}=\varnothing$ and $Q_t\subseteq Q_{t'}$ there holds

$$|\langle\varphi_t,\varphi_{t'}\rangle|\lesssim\left(\frac{|R_{t'}|}{|R_t|}\right)^{\frac12}\inf_{x\in R_{t'}}(1+\rho_{R_t}(x))^{-M+n}$$

where $\rho_R(x) := \inf\{r>0 : x\in rR\}$ for any rectangular parallelepiped R and $M>2n$, say. Combining these observations and estimates with the estimate for $S_2(f)^2$ above yields

$$\begin{aligned}S_2(f)^2 &\lesssim \sup_{t'\in\mathbb{T}}\frac{|\langle f,\phi_{t'}\rangle|}{|R_{t'}|^{\frac12}}\sum_{\mathbf{T}\in\mathscr{T}}\sum_{t\in\mathbf{T}}|\langle f,\varphi_t\rangle||R_t|^{-\frac12}\sum_{\substack{t'\in\mathbb{T}\\ Q_{t'}^\circ\supseteq Q_t}}|R_{t'}|\inf_{R_{t'}}(1+\rho_{R_t})^{-M+n}\\ &\lesssim \sup_{t'\in\mathbb{T}}\frac{|\langle f,\phi_{t'}\rangle|}{|R_{t'}|^{\frac12}}\sum_{\mathbf{T}\in\mathscr{T}}\sum_{t\in\mathbf{T}}|\langle f,\varphi_t\rangle||R_t|^{-\frac12}\int_{(K_n^2R_{\mathbf{T}(t)})^{\mathrm{c}}}(1+\rho_{R_t}(x))^{-M+n}\,\mathrm{d}x\\ &\lesssim \sup_{t'\in\mathbb{T}}\frac{|\langle f,\phi_{t'}\rangle|}{|R_{t'}|^{\frac12}}S(f)\left(\sum_{t\in\mathbb{T}}\int_{(K_n^2R_{\mathbf{T}(t)})^{\mathrm{c}}}(1+\rho_{R_t}(x))^{-M+n}\right)^{\frac12},\end{aligned}$$

where $\mathbf{T}(t)$ denotes the unique tree $\mathbf{T} \in \mathscr{T}$ such that $t \in \mathbf{T}$. In passing to the last line we used the Cauchy–Schwarz inequality together with the simple estimate

$$\int_{(K_n^2 R_{\mathbf{T}})^{\mathrm{c}}} (1+\rho_{R_t}(x))^{-M+n} \lesssim |R_t|, \qquad M > 2n.$$

The estimate for S_2 is completed by involving the estimate

$$\sum_{t\in\mathbb{T}} \int_{(K_n^2 R_{\mathbf{T}(t)})^{\mathrm{c}}} (1+\rho_{R_t}(x))^{-M+n} = \sum_{\mathbf{T}\in\mathscr{T}} \sum_{t\in\mathbf{T}} \int_{(K_n^2 R_{\mathbf{T}})^{\mathrm{c}}} (1+\rho_{R_t}(x))^{-M+n} \lesssim \sum_{\mathbf{T}\in\mathscr{T}} |R_{\mathbf{T}}|$$

for $M > 2n$. Balancing the estimates for $S(f)$ in terms of $S_1(f), S_2(f)$ according to which term dominates, and taking into account (5.10), yield the desired bound. □

6.3. **Selection by density.** The sorting of a collection of trees by density follows the standard procedure as for example in [30]; we include the proof for completeness.

Lemma 6.10 (density lemma)**.** *Let $\mathbb{T} \subset \mathbb{P}$ be a finite collection of tiles. There exists a disjoint decomposition*

$$\mathbb{T} = \mathbb{T}^{\mathrm{light}} \cup \bigcup_{\tau=1}^{N} \mathbf{T}_\tau$$

where $\{\mathbf{T}_\tau : 1 \leqslant \tau \leqslant N\}$ is a finite collection of trees and

$$\sum_{\tau=1}^{N} |R_{\mathbf{T}_\tau}| \lesssim \mathrm{dense}(\mathbb{T})^{-1}|E|, \qquad \mathrm{dense}(\mathbb{T}^{\mathrm{light}}) \leqslant \frac{1}{2}\mathrm{dense}(\mathbb{T}),$$

where E refers to the measurable $E \subset \mathbb{R}^d$ in the definition of density, with implicit constant depending only upon dimension.

Proof. We let $\mathbb{T}^{\mathrm{heavy}}$ be the collection of tiles

$$\mathbb{T}^{\mathrm{heavy}} := \{t \in \mathbb{T} : \mathrm{dense}(t) > \mathrm{dense}(\mathbb{T})/2\}.$$

By the definition of density, for each tile $t \in \mathbb{T}^{\mathrm{heavy}}$ there exists some $\mathscr{T}_{\mathscr{G}} \ni t' \geqslant t$ such that

$$\int_{E_{t'}} \mathrm{Sy}^{1}_{R_{t'}} \chi_{10n} > \mathrm{dense}(\mathbb{T})/2.$$

We let $\mathbb{T}'$ denote the maximal, with respect to $\leqslant$, elements of $\{t' : t' \in \mathbb{T}^{\mathrm{heavy}}\}$: namely, for all $t' \in \mathbb{T}'$ there does not exist $s' \in \mathbb{T}^{\mathrm{heavy}}$ with $t' \leqslant s'$. It will then suffice to prove

$$\sum_{t'\in\mathbb{T}'} |R_{t'}| \lesssim \mathrm{dense}(\mathbb{T})^{-1}|E|$$

as the tiles in $\mathbb{T}^{\mathrm{heavy}}$ can be organized into trees with top datas in $\mathbb{T}'$ and by definition

$$\mathrm{dense}(\mathbb{T}^{\mathrm{light}}) = \mathrm{dense}(\mathbb{T} \setminus \mathbb{T}^{\mathrm{heavy}}) \leqslant \frac{1}{2}\mathrm{dense}(\mathbb{T}).$$

To that end we define for $k \geqslant 0$ the collection $\mathbb{T}'_k$ to be set of all tiles $t' \in \mathbb{T}'$ such that

$$\left|E_{t'} \cap 2^k R_{t'}\right| \geqslant \frac{1}{4} 2^{5kn} \mathrm{dense}(\mathbb{T})|R_{t'}|. \tag{6.1}$$

Each element of $\mathbb{T}'$ is contained in at least one of the sets $\mathbb{T}'_k$; indeed if not we would have

$$\int_{E_{t'}} \mathrm{Sy}^1_{R_{t'}} \chi_{10n} \leqslant \frac{|E_{t'} \cap R_{t'}|}{|R_{t'}|} + \sum_{k \geqslant 0} 2^{-10kn} \frac{|E_{t'} \cap (2^{k+1} R_{t'} \setminus 2^k R_{t'})|}{|R_{t'}|} < \frac{1}{2} \mathrm{dense}(\mathbb{T}).$$

We fix k for the moment. We will decompose each collection $\mathbb{T}_k$ into subcollections of tiles with pairwise disjoint spatial components. We do this by choosing $t' \in \mathbb{T}'_k$ such that $R_{t'}$ has maximal length and let

$$\mathscr{E}(t') := \Big\{ t'' \in \mathbb{T}'_k : 2^k R_{t'} \cap 2^k R_{t''} \neq \varnothing,\, Q_{t'} \cap Q_{t''} \neq \varnothing \Big\}.$$

Note that the collection $\{R_{t''} : t'' \in \mathscr{E}(t')\}$ is pairwise disjoint. Indeed if for some $t''_1, t''_2 \in \mathscr{E}(t')$ we had $R_{t''_1} \cap R_{t''_2} \neq \varnothing$ then, since $Q_{t''_1} \cap Q_{t''_2} \neq \varnothing$ we would conclude that either $t''_1 \leqslant t''_2$ or $t''_1 \geqslant t''_2$ which is impossible because all the tiles in $\mathbb{T}_k$ were maximal to begin with. Furthermore, we have that for all $t'' \in \mathscr{E}(t')$ there holds $2^k R_{t'} \cap 2^k R_{t''} \neq \varnothing$ and $Q_{t'} \subseteq Q_{t''}$ and so a variation of Lemma 6.2 implies that $2^{2k} R_{t''} \subseteq K_n 2^{2k} R_{t'}$ for all $t'' \in \mathscr{E}(t')$. Now we replace $\mathbb{T}_k$ by $\mathbb{T}_k \setminus \mathscr{E}(t')$ and repeat so that $\mathbb{T}_k = \cup_j \mathscr{E}(t'_j)$ with $\{2^k R_{t'_j}\}_j$ disjoint and

$$\begin{aligned} \sum_{t' \in \mathbb{T}_k} |R_{t'}| &\leqslant \sum_j \sum_{t'' \in \mathscr{E}(t'_j)} |R_{t''}| \leqslant K_n^n 2^{2kn} \sum_j |R_{t'_j}| \\ &\lesssim 2^{-3kn} \mathrm{dense}(\mathbb{T})^{-1} \sum_j |2^k R_{t'_j} \cap E_{t'_j}| \lesssim 2^{-3kn} \mathrm{dense}(\mathbb{T})^{-1} |E| \end{aligned}$$

where we used (6.1) in passing to the second line, and the fact that $\{2^k R_{t'_j}\}_j$ is a pairwise disjoint collection in the last approximate inequality. The conclusion now follows by summing in k. □

6.4. **Selection by size.** We describe in the subsection the suitable version of the size (or energy) lemma, that will eventually allows us to sort a collection of tiles into trees of controlled size. The argument is standard in the literature but we introduce a small variation to account for the lack of transitivity in the relation of order that is implicit in the definition of trees. A technical device used in order to prioritize the selection of trees appears only in our higher dimensional setting and is defined below.

Definition 6.11. *(signature) Let $\mathscr{G}$ be a triadic grid of pace (κ, m) in $\mathbb{R}^d$, $m \in [0, \kappa)$, and $\xi \in \mathbb{R}^d$ having no triadic coordinates. The* signature *of ξ is the number* $\mathrm{sig}(\xi)$ *defined as follows*

$$\int\rangle\}(\xi) := \sum_{j=1}^{\infty} \frac{a_j}{3^j}, \qquad a_j := \begin{cases} 0, & \text{if} \quad \xi \in Q^o \quad \text{for some} \quad Q \in \mathscr{G}_{-\kappa j + m}, \\ 1, & \text{if} \quad \xi \notin Q^o \quad \forall Q \in \mathscr{G}_{-\kappa j + m}, \end{cases}$$

where we remember that $\mathscr{G}_u = \{Q \in \mathscr{G}, \ell(Q) = 3^u\}$; see Subsection 5.2.

Remark 6.12. *The following point motivates the definition above: if $\mathbf{T}, \mathbf{T}'$ are two different lacunary trees and $t \in \mathbf{T}, t' \in \mathbf{T}'$ with $Q_t \subseteq Q^\circ_{t'}$, then* $\mathrm{sig}(\xi_{\mathbf{T}}) < \mathrm{sig}(\xi_{\mathbf{T}'})$. *This fact will play a role in proving that the family of trees extracted in the size lemma below is strongly disjoint.*

Lemma 6.13 (size lemma)**.** *Let $\mathbb{T} \subset \mathbb{P}$ be a finite collection of tiles. There exists a disjoint decomposition*

$$\mathbb{T} = \mathbb{T}^{\text{small}} \cup \bigcup_{\tau=1}^{N} \mathbf{T}_\tau$$

where $\{\mathbf{T}_\tau : 1 \leqslant \tau \leqslant N\}$ is a finite collection of trees and

$$\sum_{\tau=1}^{N} |R_{\mathbf{T}_\tau}| \lesssim \text{size}(\mathbb{T})^{-2} \|f\|^2_{L^2(\mathbb{R}^n)}, \qquad \text{size}(\mathbb{T}^{\text{small}}) \leqslant \frac{\text{size}(\mathbb{T})}{\sqrt{2}},$$

with implicit constant depending only upon dimension.

Proof. Throughout the proof we abbreviate $\sigma := \text{size}(\mathbb{T})$. We remove trees from $\mathbb{T}$ via the following recursive procedure. First consider all the maximal (with respect to set inclusion) lacunary trees $\mathbf{T} \subseteq \mathbb{T}$ such that

$$\sum_{t \in \mathbf{T}} F[f](t)^2 > \frac{\sigma^2}{2} |R_{\mathbf{T}}| \tag{6.2}$$

and among those let $\mathbf{T}_1$ be the one with $\xi_{\mathbf{T}_1}$ having the *smallest signature*, where $(\xi_{\mathbf{T}_1}, R_{\mathbf{T}_1})$ is the top data of $\mathbf{T}_1$. We define the collection of tiles

$$\mathscr{E}(\mathbf{T}_1) := \{t \in \mathbb{T} : \xi_{\mathbf{T}_1} \in Q_t, \text{scl}(R_t) \leqslant \text{scl}(R_{\mathbf{T}_1}), R_t \cap K_n^2 R_{\mathbf{T}_1} \neq \varnothing\}.$$

Clearly $\mathbf{T}_1 \subseteq \mathscr{E}(\mathbf{T}_1)$. Now replace $\mathbb{T}$ by $\mathbb{T} \setminus \mathscr{E}(\mathbf{T}_1)$ and recurse. The selection algorithm terminates when $\text{size}(\mathbb{T}^{small}) = \text{size}(\mathbb{T} \setminus \bigcup_k \mathscr{E}(\mathbf{T}_k)) < \sigma/\sqrt{2}$.

Note that the elements of the collection $\{\mathscr{E}(\mathbf{T}_k)\}_k$ are not — strictly speaking — trees according to Definition 6.1 but we can easily remedy that. Indeed consider the rectangular parallelepipeds $R_{\mathbf{T}_{k,j}}$ which belong to the same grid as $R_{\mathbf{T}_k}$ and are such that $R_{\mathbf{T}_{k,j}} \cap K_n^2 R_{\mathbf{T}_k} \neq \varnothing$. Clearly there are at most $O_d(1)$ indices j for which this happens, uniformly in k. Now we set

$$\mathbf{T}_{k,1} := \big\{t \in \mathscr{E}(\mathbf{T}_k) : R_t \cap R_{\mathbf{T}_{k,1}} \neq \varnothing\big\}$$

and recursively let $\mathbf{T}_{k,j+1}$ be the maximal tree with top $(\xi_{\mathbf{T}_k}, R_{\mathbf{T}_{k,j+1}})$ contained in $\mathscr{E}(\mathbf{T}_k) \setminus \cup_{j' \leqslant j} \mathbf{T}_{k,j'}$. Each $\mathbf{T}_{k,j}$ is a tree and we have $\sum_k \sum_j |R_{\mathbf{T}_{j,k}}| \lesssim \sum_k |R_{\mathbf{T}_k}|$ so the proof will be complete once we estimate the latter sum. An important fact is that the collection $\mathscr{T} := \{\mathbf{T}_k\}_k$ is strongly disjoint. Indeed if $t \in \mathbf{T} \neq \mathbf{T}' \ni t'$ with $\mathbf{T}, \mathbf{T}' \in \mathscr{T}$ and $Q_t \subseteq Q_{t'}^\circ$, then the tree $\mathbf{T}$ was selected first because of Remark 6.12. Thus if we had that $R_{t'} \cap K_n^2 R_{\mathbf{T}} \neq \varnothing$ then the tile t' would have been included in the family $\mathscr{E}(\mathbf{T})$ and consequently would not be available for inclusion in the tree $\mathbf{T}'$. Now using the selection condition (6.2) together with Lemma 6.9 we get

$$\sum_{\mathbf{T} \in \mathscr{T}} |R_{\mathbf{T}}| \lesssim \sigma^{-2} \sum_{\mathbf{T} \in \mathscr{T}} \sum_{t \in \mathbf{T}} F[f](t)^2 \lesssim \sigma^{-2} \max\left(\|f\|_2^2, \sigma^{\frac{2}{3}} \left(\sum_{\mathbf{T} \in \mathscr{T}} |R_{\mathbf{T}}| \right)^{\frac{1}{3}} \|f\|_2^{\frac{4}{3}} \right)$$

which readily yields the desired estimate. □

6.5. **The single tree estimate.** The size and density lemmas combined with the discretization of the operator in (5.25) reduce matters to the estimate for a single tree which is stated and proved below.

Lemma 6.14 (tree lemma)**.** *For every tree* $\mathbf{T}$ *and every* $\tau_o \in \{1,\dots,N\}$*, there holds*

$$\Lambda_{\mathbf{T};\tau_o}(f, g\mathbf{1}_E) \lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})|R_T|$$

where size *is defined with respect to* $f \in L^\infty_0(\mathbb{R}^n)$*,* dense *is defined with respect to* $E \subset \mathbb{R}^n$ *and* $g \in L^\infty_0(\mathbb{R}^n)$ *with* $\|g\|_\infty = 1$ *is arbitrary.*

Proof. In view of the definition of the form $\Lambda_{\mathbf{T};\tau_o}$ it suffices to prove that

$$\sum_{t\in\mathbf{T}} \big|\langle f,\varphi_t\rangle\big| \Big|\Big\langle \vartheta_t(\cdot,\sigma(\cdot)), g\mathbf{1}_{E_{t,\tau_0}}\Big\rangle\Big| \lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})|R_{\mathbf{T}}|,$$

where $E_{t,\tau_0} := \{x \in E : \nu_\sigma(x) \in \alpha_{t,\tau_0}\}$, uniformly in choices of adapted families $\{\varphi_t : \varphi_t \in \mathscr{F}_t^M,\ t \in \mathbf{T}\}$ and $\{\psi_t : \varphi_t \in \mathscr{A}_t^M,\ t \in \mathbf{T}\}$. We remember that we have fixed $M = 50n$.

We begin the proof setup with a reduction. Since the single tree estimate is rotation invariant we can and will assume that $\xi_{\mathbf{T}} = \Pi_{e_n^\perp} e_n = 0 \in \mathbb{R}^d$. A simple geometric argument similar to the one needed in the proof of Lemma 6.2 implies that there exists a dimensional constant $C_n > 0$ such that for any tile $t \in \mathbf{T}$ there exists a rectangle R_t^* with sides parallel to the coordinate axes, $\Pi_{e_n^\perp}(R_t^*)$ is a cube in $\mathbb{R}^d$ and $R_t \subset R_t^* \subset C_n R_t$. Furthermore, by e.g. splitting the tree into $O_n(1)$ trees we may assume that all R_t^* come from a single grid $\mathscr{L}$ in $\mathbb{R}^n$ that consists of rectangles with sides parallel to the axes and is such that

$$\{\Pi_{e_n^\perp}(R) : R \in \mathscr{L}\} \subseteq \mathscr{D}_{(\kappa,m)},$$

where $\mathscr{D}_{(\kappa,m)}$ is a dyadic grid of pace (κ,m) in $\mathbb{R}^d$ with $m \in [0,\kappa)$, and that $\Pi_{e_n}(R^*)$ are standard dyadic intervals from $\mathscr{D}$ of length 4. We also fix $R_{\mathbf{T}}^* \in \mathscr{L}$ with $R_{\mathbf{T}} \subseteq R_{\mathbf{T}}^* \subseteq C_n R_{\mathbf{T}}$ and $R_t^* \subseteq R_{\mathbf{T}}^*$ for all $t \in \mathbf{T}$.

We note that $\mathbf{T}^* := \{R_t^* \times Q_t :\ t \in \mathbf{T}\}$ is not — strictly speaking — a collection of tiles according to our definition in §5.2.2. However the previously defined relation of order is still meaningful with the same formal definition and we have that

$$\mathrm{size}(\mathbf{T}) \eqsim \mathrm{size}(\mathbf{T}^*), \qquad \mathrm{dense}(\mathbf{T}) \eqsim \mathrm{dense}(\mathbf{T}^*)$$

with implicit dimensional constants, where in the definitions for $\mathrm{size}(\mathbf{T}^*), \mathrm{dense}(\mathbf{T}^*)$ above we use $t \in \mathbf{T}^*$ instead of the actual tiles of $\mathbf{T}$. With these reductions and remarks taken as understood we will henceforth write $\mathbf{T} = \mathbf{T}^*$ and assume that $\{R_t :\ t \in \mathbf{T}\} \subseteq \mathscr{L}$ with $\mathscr{L}$ being a grid as above. Note that the uncertainty relation will change in an inconsequential way as we will now have $\mathrm{scl}(R_t)\ell(Q_t) \in [c_n^{-1}, c_n)$ for some dimensional constant $c_n > 3$.

Now let $\mathscr{P}(e_n^\perp)$ be the maximal dyadic cubes Q in $\mathbb{R}^d$ such that $3Q \not\supseteq \Pi_{e_n^\perp} R_t$ for any $t \in \mathbf{T}$; note that $\mathscr{P}(e_n^\perp)$ partitions $\mathbb{R}^d$. Let $\mathscr{P}(e_n)$ be the dyadic intervals of length 2 on the real line and define $\mathscr{P} := \mathscr{P}(e_n^\perp) \times \mathscr{P}(e_n)$. We will write $\mathrm{scl}(P) := \ell(\Pi_{e_n^\perp} P)$. For $P \in \mathscr{P}$ we write $\mathbf{T} = \mathbf{T}_P^+ \cup \mathbf{T}_P^-$, where

$$\mathbf{T}_P^- := \{t \in \mathbf{T} : \mathrm{scl}(R_t) \leqslant \mathrm{scl}(P)\} \quad \text{and} \quad \mathbf{T}_P^+ := \{t \in \mathbf{T} : \mathrm{scl}(R_t) > \mathrm{scl}(P)\}.$$

We write the basic estimate

$$\sum_{t\in\mathbf{T}} \big|\langle f,\phi_t\rangle\big| \Big|\Big\langle \vartheta_t(\cdot,\sigma(\cdot)), g\mathbf{1}_{E_{t,\tau_0}}\Big\rangle\Big| = \sum_{P\in\mathscr{P}} \int_P L_P^-(x)\,\mathrm{d}x + \sum_{P\in\mathscr{P}} \int_P L_P^+(x)\,\mathrm{d}x,$$

where

$$L_P^\omega(x) := \sum_{t\in\mathbf{T}_P^\omega} \varepsilon_t \langle f,\varphi_t\rangle \vartheta_t(x,\sigma(x))g(x)\mathbf{1}_{E_{t,\tau_0}}(x), \quad \omega\in\{-,+\}, \quad x\in\mathbb{R}^n,$$

for an appropriate sequence $\{\varepsilon_t\}_{t\in\mathbf{T}}$ on the unit circle in $\mathbb{C}$. We shall prove that

$$\left|\sum_{P\in\mathscr{P}}\int_P L_P^\omega(x)\,\mathrm{d}x\right| \lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})|R_{\mathbf{T}}|, \quad \omega\in\{-,+\}. \tag{6.3}$$

Proof of (6.3) *for* $\omega=-$. For $P\in\mathscr{P}$ and $\|g\|_\infty\leqslant 1$

$$\begin{aligned}\left|\int_P L_P^-(x)\,\mathrm{d}x\right| &\leqslant \mathrm{size}(\mathbf{T})\sum_{t\in\mathbf{T}_P^-}\int_{P\cap E_t}|R_t|^{1/2}|\vartheta_t(x,\sigma(x))||g(x)|\,\mathrm{d}x\\ &\lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})\sum_{t\in\mathbf{T}_P^-}|R_t|\sup_{x\in P}\mathrm{Sy}_{R_t}^\infty\chi_{40n}(x).\end{aligned}$$

For $R\in\{R_t : t\in\mathbf{T}_P^-\}$, set

$$\rho_R(P) := \inf_{x\in P}\rho_R(x), \qquad \rho_R(x) := \inf\{r>0 : x\in rR\},$$

and note that $\rho_{R_t}(P)\geqslant \max(1,\rho_{R_{\mathbf{T}}}(P)) =: \rho(P,\mathbf{T})$ by the construction of $\mathscr{P}$ and the facts that $\mathrm{scl}(P)\geqslant\mathrm{scl}(R_t)$ and $R_t\subseteq R_{\mathbf{T}}$ for $t\in\mathbf{T}_P^-$. It follows that

$$\left|\sum_{P\in\mathscr{P}}\int_P L_P^-(x)\,\mathrm{d}x\right| \lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})\sum_{P\in\mathscr{P}}\sum_{t\in\mathbf{T}_P^-}|R_t|(1+\rho_{R_t}(P))^{-40n}.$$

Note that for $t\in\mathbf{T}_P^-$ with $\rho_{R_t}(P)\in[2^\ell,2^{\ell+1})$ and $\ell(R_t)=2^m$ we have

$$R_t\subseteq 2^{\ell+4}\frac{2^{m+1}}{\ell\big(\Pi_{e_n^\perp}(P)\big)}\Pi_{e_n^\perp}(P)\times 2^{\ell+4}\Pi_{e_n}(P).$$

We can thus estimate

$$\begin{aligned}\sum_{P\in\mathscr{P}}\sum_{t\in\mathbf{T}_P^-}|R_t|\rho_{R_t}(P)^{-40n} &\lesssim \sum_{P\in\mathscr{P}}\sum_{2^{\ell+1}\geqslant\rho(P,\mathbf{T})}2^{-40n\ell}\sum_{2^m=4}^{\ell(\Pi_{e_n^\perp}(P))}\left(\frac{2^{\ell+1}2^m}{\ell\big(\Pi_{e_n^\perp}(P)\big)}\right)^d 2^\ell|P|\\ &= \sum_{\substack{P\in\mathscr{P}\\ P\subseteq 100R_{\mathbf{T}}}}\sum_{\ell\geqslant -1}2^{-39n\ell}|P| + \sum_{\substack{P\in\mathscr{P}\\ P\nsubseteq 100R_{\mathbf{T}}}}\sum_{2^{\ell+1}\geqslant\rho_{R_{\mathbf{T}}}(P)}2^{-39n\ell}|P| =: \mathrm{I}+\mathrm{II}.\end{aligned}$$

Since the collection $\{P\}_{P\in\mathscr{P}}$ is pairwise disjoint we readily get that $\mathrm{I}\lesssim|R_{\mathbf{T}}|$. For II note that $P\nsubseteq 100R_{\mathbf{T}}$ implies that $P\subseteq(4R_{\mathbf{T}})^{\mathrm{c}}$ so that

$$\mathrm{II}\lesssim\sum_{\substack{P\in\mathscr{P}\\ P\subseteq(4R_{\mathbf{T}})^{\mathrm{c}}}}\int_P\rho_{R_{\mathbf{T}}}(x)^{-39n}\,\mathrm{d}x\lesssim\int_{(4R_{\mathbf{T}})^{\mathrm{c}}}\rho_{R_{\mathbf{T}}}(x)^{-39n}\,\mathrm{d}x\lesssim|R_{\mathbf{T}}|$$

and the proof of (6.3) for $\omega=-$ is complete.

Proof of (6.3) *for* $\omega = +$. We make the preliminary observation that if $\mathbf{T}_P^+ \neq \varnothing$ then $\mathrm{scl}(R_{\mathbf{T}}) \geqslant 2\mathrm{scl}(P)$ and $\Pi_{e_n^\perp} P \subseteq 5\Pi_{e_n^\perp} R_{\mathbf{T}}$. For each $P \in \mathscr{P}$ we define

$$G_P := P \cap \bigcup_{t \in \mathbf{T}_P^+} E_t.$$

Lemma 6.15. *For* $P \in \mathscr{P}$ *there holds* $|G_P| \lesssim (1+\rho_{R_{\mathbf{T}}}(P))^{10n} \mathrm{dense}(\mathbf{T})|P|$.

Proof. Let $\widehat{P} \in \mathscr{L}$ be such that $\Pi_{e_n^\perp}(\widehat{P})$ is the dyadic parent of $\Pi_{e_n^\perp}(P)$. By the maximality of P one has $3\Pi_{e_n^\perp}(\widehat{P}) \supseteq \Pi_{e_n^\perp}(R_{t_0})$ for some $t_0 \in \mathbf{T}$. Hence, there exists an $R \in \mathscr{L}$ such that $\ell(\Pi_{e_n^\perp}(R)) = \ell(\Pi_{e_n^\perp}(\widehat{P}))$ and $R \supseteq R_{t_0}$.

Let Q be the largest cube in $\mathscr{G}$ such that $\xi_{\mathbf{T}} \in Q \subseteq Q_{t_0}$ and $\mathrm{scl}(R)\mathrm{scl}(Q) \in [c_n^{-1}, c_n)$. Define $t := R \times Q$ and note that $t \geqslant t_0$ and $G_P \subseteq P \cap E_t$. As

$$|P| \eqsim |R|, \qquad \mathrm{Sy}_R^1 \chi_{10n}(x) \gtrsim (1+\rho_{R_{\mathbf{T}}}(P))^{-10n} |R|^{-1} \qquad \forall\, x \in P,$$

we have

$$\begin{aligned} |G_P| &\lesssim |P| \int_{E_t} \frac{1}{|R|} \mathbf{1}_P(x) \mathrm{d}x \\ &\lesssim (1+\rho_{R_{\mathbf{T}}}(P))^{10n} |P| \int_{E_t} \mathrm{Sy}_R^1 \chi_{10n}(x) \mathrm{d}x \lesssim (1+\rho_{R_{\mathbf{T}}}(P))^{10n} |P| \mathrm{dense}(\mathbf{T}), \end{aligned}$$

and the proof is complete. □

We can decompose $\mathbf{T} = \mathbf{T}_{\mathrm{ov}} \cup_{\tau=1}^N \mathbf{T}_{\mathrm{lac}_\tau}$, where $\mathbf{T}_{\mathrm{ov}}$ is an overlapping tree and $\mathbf{T}_{\mathrm{lac}_\tau}$ is a τ-lacunary tree for each $\tau \in \{1, \dots, N\}$. For $P \in \mathscr{P}$, we set $\mathbf{T}_{P,\mathrm{ov}}^+ := \mathbf{T}_{\mathrm{ov}} \cap \mathbf{T}_P^+$ and $\mathbf{T}_{P,\mathrm{lac}_\tau}^+ := \mathbf{T}_{\mathrm{lac}_\tau} \cap \mathbf{T}_P^+$ for $\tau \in \{1, \dots, N\}$. We define

$$L_{P,\mathrm{type}}^+(x) := \sum_{t \in \mathbf{T}_{P,\mathrm{type}}^+} \varepsilon_t \langle f, \varphi_t \rangle \vartheta_t(x, \sigma(x)) g(x) \mathbf{1}_{E_{t,\tau_0}}(x), \quad \mathrm{type} \in \{\mathrm{ov}, \cup_\tau \mathrm{lac}_\tau\}, \quad x \in \mathbb{R}^n.$$

For reasons of space, until the end of the proof we use the local notation

$$F := \sum_{t \in \mathbf{T}_{\mathrm{lac}_{\tau_o}}^+} \varepsilon_t \langle f, \varphi_t \rangle \widetilde{\vartheta}_t.$$

We shall prove the estimates

$$\|L_{P,\mathrm{type}}^+\|_{L^\infty(P)} \lesssim \mathrm{size}(\mathbf{T})(1+\rho_{R_{\mathbf{T}}}(P))^{-49n}, \quad \mathrm{type} \in \{\mathrm{ov}, \cup_{\tau \neq \tau_o} \mathrm{lac}_\tau\} \tag{6.4}$$

$$\|L_{P,\mathrm{lac}_{\tau_o}}^+\|_{L^\infty(P)} \lesssim \frac{1}{(1+\rho_{R_{\mathbf{T}}}(P))^{20n}} \inf_P \mathrm{M_S} F + \frac{\mathrm{size}(\mathbf{T})}{(1+\rho_{R_{\mathbf{T}}}(P))^{49n}}, \tag{6.5}$$

where $\{\widetilde{\vartheta}_t : t \in \mathbf{T}_{\mathrm{lac}}^+\} \subset \mathscr{A}_t^{30n}$ and $\mathrm{M_S}$ is the strong maximal function. Before doing so let us show that (6.4) and (6.5) imply (6.3) for $\omega = +$. Indeed, using (6.4) and Lemma 6.15 we have for $\mathrm{type} \in \{\mathrm{ov}, \cup_{\tau \neq \tau_o} \mathrm{lac}_\tau\}$

$$\left| \sum_{P \in \mathscr{P}} \int_P L_{P,\mathrm{type}}^+(x) \, \mathrm{d}x \right| \lesssim \mathrm{size}(\mathbf{T}) \mathrm{dense}(\mathbf{T}) \sum_{P \in \mathscr{P}} |P| (1+\rho_{R_{\mathbf{T}}}(P))^{-39n} \lesssim \mathrm{size}(\mathbf{T}) \mathrm{dense}(\mathbf{T}) |R_{\mathbf{T}}|.$$

On the other hand, using (6.5) and Lemma 6.15 and the same calculation as in the last display we get

$$\int \left| \sum_{P\in\mathscr{P}} L^+_{P,\mathrm{lac}_{\tau_o}}(x)\,\mathrm{d}x \right| \lesssim \mathrm{size}(\mathbf{T})\mathrm{dense}(\mathbf{T})|R_{\mathbf{T}}| + \mathrm{dense}(\mathbf{T}) \sum_{P\in\mathscr{P}} (1+\rho_{R_{\mathbf{T}}}(P))^{-10n} \int_P \mathrm{M_S}F.$$

Observe that

$$\sum_{\substack{P\in\mathscr{P}\\ P\subseteq 100R_{\mathbf{T}}}} \int_P \mathrm{M_S}F \lesssim |R_{\mathbf{T}}|\mathrm{size}(\mathbf{T})$$

by the Cauchy–Schwarz inequality, the boundedness of the strong maximal function on $L^2(\mathbb{R}^n)$ and Lemma 6.6. Remembering that $\mathrm{scl}(R_{\mathbf{T}}) \geqslant 2\mathrm{scl}(P)$ we get

$$\begin{aligned}\sum_{\substack{P\in\mathscr{P}\\ P\not\subseteq 100R_{\mathbf{T}}}} \tfrac{1}{(1+\rho_{R_{\mathbf{T}}}(P))^{10n}} \int_P \mathrm{M_S}F &\lesssim \sum_{\ell\geqslant 2} 2^{-10\ell n} \sum_{\substack{P\in\mathscr{P}\\ \rho_{R_{\mathbf{T}}}(P)\eqsim 2^\ell}} \int_P \mathrm{M_S}F\\ &\lesssim \sum_{\ell\geqslant 2} 2^{-10\ell n}|2^{\ell+2}R_{\mathbf{T}}|^{\frac12}|R_{\mathbf{T}}|^{\frac12}\mathrm{size}(\mathbf{T}) \lesssim |R_{\mathbf{T}}|\mathrm{size}(\mathbf{T})\end{aligned}$$

where we also used that $\rho_{R_{\mathbf{T}}}(P)\eqsim 2^\ell \implies P\subseteq 2^{\ell+2}R_{\mathbf{T}}$ in passing to the penultimate approximate inequality.

Proof of (6.4). Let $\mathrm{type}\in\{\mathrm{ov}, \cup_{\tau\neq\tau_o}\mathrm{lac}_\tau\}$ be fixed. Observe that for fixed $x\in\mathbb{R}^n$ if, for $t,t'\in \mathbf{T}^+_{P,\mathrm{type}}$, $\vartheta_t(x,\sigma(x))\mathbf{1}_{E_{t,\tau_0}}(x)\neq 0$ and $\vartheta_{t'}(x,\sigma(x))\mathbf{1}_{E_{t',\tau_0}}(x)\neq 0$ then by property (iv) one has $\mathrm{scl}(R_t)=\mathrm{scl}(R_{t'})$. Hence,

$$\|L^+_{P,\mathrm{type}}\|_{L^\infty(P)} \leqslant \mathrm{size}(\mathbf{T}) \sup_{\substack{x\in P\\ \mathrm{scl}\in 2^{\mathbb{N}}}} \sum_{\substack{t\in\mathbf{T}^+_{P,\mathrm{type}}\\ \mathrm{scl}(R_t)=\mathrm{scl}}} |R_t|^{\frac12}|\vartheta_t(x,\sigma(x))||g(x)|\mathbf{1}_{E_t}(x).$$

Define $\omega_t(y,x_n) := t^{-d}\chi_{50n}(y/t,x_n)$. For any $\mathrm{scl}\in 2^{\mathbb{N}}$ and fixed σ,by adaptedness of $\vartheta_t(\cdot,\sigma)$

$$\sum_{\substack{t\in\mathbf{T}^+_{P,\mathrm{type}}\\ \mathrm{scl}(R_t)=\mathrm{scl}}} |R_t|^{\frac12}|\vartheta_t(\cdot,\sigma)| \lesssim \sum_{\substack{t\in\mathbf{T}^+_{P,\mathrm{type}}\\ \mathrm{scl}(R_t)=\mathrm{scl}}} \mathrm{Sy}^\infty_{R_t}\chi_{50n} \lesssim \sum_{\substack{t\in\mathbf{T}^+_{P,\mathrm{type}}\\ \mathrm{scl}(R_t)=\mathrm{scl}}} \omega_{\mathrm{scl}} * \mathbf{1}_{R_t} \lesssim \omega_{\mathrm{scl}} * \mathbf{1}_{R_{\mathbf{T}}}$$

using that the collection $\{R_t\}_{\mathrm{scl}(R_t)=\mathrm{scl}}$ is pairwise disjoint and $R_t\subseteq R_{\mathbf{T}}$ for all $t\in\mathbf{T}$. Because of the trivial estimate $\omega_{\mathrm{scl}} * \mathbf{1}_{R_{\mathbf{T}}} \lesssim 1$ it will be enough to consider P such that $P\cap 3R_{\mathbf{T}}=\varnothing$; indeed if $P\cap 3R_{\mathbf{T}}\neq\varnothing$ then $\rho_{R_{\mathbf{T}}}(P)\lesssim 1$ and the desired estimate follows. So fix such a P and let $c_{R_{\mathbf{T}}}$ denote the center of $R_{\mathbf{T}}$. If $x=(y,x_n)\in P$ then

$$\begin{aligned}\omega_{\mathrm{scl}} * \mathbf{1}_{R_{\mathbf{T}}}(y,x_n) &\lesssim \frac{|R_{\mathbf{T}}|}{\mathrm{scl}^d}\left(|x_n-c_{R_{\mathbf{T}}}| + \frac{|y-c_{R_{\mathbf{T}}}|}{\mathrm{scl}}\right)^{-50n}\\ &\lesssim \frac{\mathrm{scl}(R_{\mathbf{T}})^d}{\mathrm{scl}^d}\left(|x_n-c_{R_{\mathbf{T}}}| + \frac{|y-c_{R_{\mathbf{T}}}|}{\mathrm{scl}}\right)^{-49n}\left(\frac{\mathrm{scl}(R_{\mathbf{T}})}{\mathrm{scl}}\right)^{-n} \lesssim \rho_{R_{\mathbf{T},\mathrm{scl}}}(P)^{-49n}\end{aligned}$$

where $R_{\mathbf{T},\mathrm{scl}}$ is an axis parallel rectangular parallelepiped centered at $c_{R_{\mathbf{T}}}$ with sidelengths $\mathrm{scl}^d\times 4$. Thus we have proved

$$\|\omega_{\mathrm{scl}} * \mathbf{1}_{R_{\mathbf{T}}}\|_{L^\infty(P)} \lesssim \big(1+\rho_{R_{\mathbf{T},\mathrm{scl}}}(P)\big)^{-49n} \leqslant \big(1+\rho_{R_{\mathbf{T}}}(P)\big)^{-49n}$$

since $R_{\mathbf{T},\mathrm{scl}} \subseteq R_{\mathbf{T}}$. This completes the proof of (6.4) since $\|g\|_\infty \leqslant 1$.

Proof of (6.5). Remember that $\xi_{\mathbf{T}} = \Pi_{e_n^\perp} e_n = 0 \in \mathbb{R}^d$ which means that property (v) of §5.2.3 applies for $\vartheta(\cdot,\sigma(\cdot)) - \vartheta(\cdot,\mathbb{R}^d)$ where $\mathbb{R}^d \in \mathbf{Gr}(d,n)$ is regarded as an element of the Grassmanian. Writing $\vartheta_t := \vartheta_t(\cdot,\mathbb{R}^d)$ we preliminary split

$$L^+_{P,\mathrm{lac}} = g\left(\sum_{t\in \mathbf{T}^+_{P,\mathrm{lac}_{\tau_o}}} \varepsilon_t \langle f,\varphi_t\rangle \vartheta_t \mathbf{1}_{E_{t,\tau_0}}(x)\right) + g\left(\sum_{t\in \mathbf{T}^+_{P,\mathrm{lac}_{\tau_o}}} \varepsilon_t \langle f,\varphi_t\rangle [\vartheta_t(\cdot,\sigma(\cdot)) - \vartheta_t] \mathbf{1}_{E\cap \alpha_{t,\tau_o}}(x)\right)$$
$$=: gL_{P,1} + gL_{P,2}.$$

As $\|g\|_\infty \leqslant 1$ it suffices to prove the required pointwise estimate is satisfied by $L_{P,j}$, $j = 1,2$. We first deal with the term $L_{P,1}$. Note that the sets $\{Q_{t,\tau_o}\}_{t\in \mathbf{T}_{\mathrm{lac},\tau_o}} = \{\Pi_{e_n^\perp}(\alpha_{t,\tau_o})\}_{t\in \mathbf{T}_{\mathrm{lac}_{\tau_o}}}$ are nested since they all contain $\xi_{\mathbf{T}}$. Furthermore, for a fixed $x\in\mathbb{R}^n$ a term in the sum defining $L_{P,1}(x)$ is nonzero if and only if $v_\sigma(x) \in \alpha_{t,\tau_0}$. We can then conclude that there exist positive numbers $\mathrm{scl}(P) < m(x) \leqslant M(x)$ such that the condition $\mathbf{1}_{E_{t,\tau_0}}(x) \neq 0$ is satisfied if and only if $m(x) \leqslant \mathrm{scl}(R_t) \leqslant M(x)$. Letting $t_{R_{\mathbf{T}}} := \mathrm{Sy}^\infty_{R_{\mathbf{T}}} \chi_{20n}$ we then have for any $c > 0$

$$L_{P,1}(x) = c^{-1} t_{R_{\mathbf{T}}}(x) \mathbf{1}_E(x) \sum_{\substack{t\in \mathbf{T}^+_{\mathrm{lac}_{\tau_o}} \\ m(x)\leqslant \mathrm{scl}(R_t)\leqslant M(x)}} \varepsilon_t \langle f,\varphi_t\rangle \widetilde{\vartheta}_t(x)$$

where $\widetilde{\vartheta}_t := c t^{-1}_{R_{\mathbf{T}}} \vartheta_t$. Note that for a choice of c depending only upon dimension we will have that $\{\widetilde{\vartheta}_t : t \in \mathbf{T}^+_{\mathrm{lac}_{\tau_o}}\} \subset \mathscr{A}^{30n}_t$. In order to conclude the desired estimate for $L_{P,1}$ we consider $\eta \in \mathscr{S}(\mathbb{R}^n)$ with $\mathrm{supp}\,(\widehat{\eta}) \subseteq [-3,3]^d \times [1/4,4]$ and $\widehat{\eta}(\xi) = 1$ for all $\xi \in [-2,2]^d \times [1/2,2]$. Then, as $\mathrm{supp}\,(\widetilde{\vartheta}_t)^\wedge \subseteq \omega_t$, we may write for any $B > \beta > 0$

$$\sum_{\substack{t\in \mathbf{T}^+_{\mathrm{lac}_{\tau_o}} \\ \beta\leqslant \mathrm{scl}(R_t)\leqslant B}} \varepsilon_t \langle f,\phi_t\rangle \widetilde{\vartheta}_t = \left(\eta_{\gamma_n B} - \eta_{\gamma_n \beta}\right) * \sum_{t\in \mathbf{T}^+_{\mathrm{lac}_{\tau_o}}} \varepsilon_t \langle f,\phi_t\rangle \widetilde{\vartheta}_t$$

for a constant $\gamma_n > 0$, where $\eta_r(x) := r^{-d}\sigma(r^{-1}\Pi_{e_n^\perp} x + \Pi_{e_n} x)$ for $r > 0$ and $x \in \mathbb{R}^n$. Thus

$$\sup_{x\in P} |L_{P,1}(x)| \lesssim (1 + \rho_{R_{\mathbf{T}}}(P))^{-20n} \inf_P \mathrm{M_S}\left(\sum_{t\in \mathbf{T}^+_{P,\mathrm{lac}_{\tau_o}}} \varepsilon_t \langle f,\varphi_t\rangle \widetilde{\vartheta}_t\right)$$

as desired.

We turn to the estimate for $L_{P,2}$. We use the fact that $\operatorname{scl}(t) \lesssim \operatorname{dist}(\sigma(x), \mathbb{R}^d)^{-1}$ for $t \in \mathbf{T}$ together with (v) of §5.2.3 to estimate for $x \in P$

$$\begin{aligned}
|L_{P,2}(x)| &\leqslant \operatorname{size}(\mathbf{T}) \sum_{t \in \mathbf{T}^+_{P,\operatorname{lac}\tau_o}} |R_t|^{1/2} |\vartheta_t(x, \sigma(x)) - \vartheta_t(x)| \\
&\lesssim \operatorname{size}(\mathbf{T}) \operatorname{dist}(\sigma(x), \mathbb{R}^d) \sum_{\substack{\operatorname{scl} \in 2^{\mathbb{N}} \\ \operatorname{scl} \lesssim \operatorname{dist}(\sigma(x), \mathbb{R}^d)^{-1}}} \operatorname{scl} \sum_{\substack{t \in \mathbf{T}^+_{P,\operatorname{lac}\tau_o} \\ \operatorname{scl}(R_t) = \operatorname{scl}}} \operatorname{Sy}^{\infty}_{R_t} \chi_{50n}(x) + \\
&\quad + \frac{\operatorname{size}(\mathbf{T})}{\log\big(e + \operatorname{dist}(\sigma(x), \mathbb{R}^d)^{-1}\big)} \sum_{\substack{\operatorname{scl} \in 2^{\mathbb{N}} \\ \operatorname{scl} \lesssim \operatorname{dist}(\sigma(x), \mathbb{R}^d)^{-1}}} \sum_{\substack{t \in \mathbf{T}^+_{P,\operatorname{lac}\tau_o} \\ \operatorname{scl}(R_t) = \operatorname{scl}}} \operatorname{Sy}^{\infty}_{R_t} \chi_{50n}(x) \\
&\lesssim \operatorname{size}(\mathbf{T}) (1 + \rho_{R_{\mathbf{T}}}(P))^{-49n}.
\end{aligned}$$

This concludes the proof of (6.5) and with that the proof of the tree estimate. □

6.6. **The proof of** (5.25)-(5.26)**.** Fix a finite collection $\mathbb{P}$, $g \in L^\infty_0(\mathbb{R}^n)$ with $\|g\|_\infty = 1$ throughout this discussion. Also fix a set $E \subset \mathbb{R}^n$ of finite measure. The function dense below corresponds to this choice of E.

We first prove (5.25). For this, fix $f \in L^\infty_0(\mathbb{R}^n)$ so that the size map is computed with respect to f. We note that for any collection $\mathbb{P}$ we have the *apriori* bounds

$$\operatorname{dense}(\mathbb{P}) \lesssim \min(1, |E|), \qquad \operatorname{size}(\mathbb{P}) \lesssim \|f\|_\infty =: 2^{-\frac{K(f)}{2}} \|f\|_2, \tag{6.6}$$

the first one being easily verified by checking the definition of dense and the second one being a consequence of 2. from Lemma 6.6. By consecutive applications of the size lemma, Lemma 6.13, and the density lemma, Lemma 6.10, and starting from the initial bounds for dense and size above, we can decompose the collection of tiles $\mathbb{P}$ in the form

$$\begin{aligned}
&\mathbb{P} = \bigcup_{k \geqslant K(f)} \bigcup_{\nu=1}^{N_k} \mathbf{T}_{k,\nu}, \\
&\operatorname{size}(\mathbf{T}_{k,\nu}) \leqslant 2^{-\frac{k}{2}} \|f\|_2, \qquad \operatorname{dense}(\mathbf{T}_{k,\nu}) \lesssim \min\Big\{1, 2^{-k}|E|\Big\}, \qquad \sum_{\nu=1}^{N_k} |R_{\mathbf{T}_{k,\nu}}| \lesssim 2^k.
\end{aligned} \tag{6.7}$$

Consequently, employing the tree Lemma 6.14 we have for each $\tau \in \{1, \dots, N\}$

$$\begin{aligned}
\Lambda_{\mathbb{P};\tau}(f, g\mathbf{1}_E) &\lesssim \sum_{k \geqslant K(f)} \sum_{\nu=1}^{N_k} \operatorname{size}(\mathbf{T}_{k,\tau}) \operatorname{dense}(\mathbf{T}_{k,\nu}) |R_{\mathbf{T}_{k,\nu}}| \lesssim \|f\|_2 \sum_{k \in \mathbb{Z}} 2^{\frac{-k}{2}} \min(1, 2^{-k}|E|) \sum_{\nu=1}^{N_k} |R_{\mathbf{T}_{k,\nu}}| \\
&\lesssim \|f\|_2 |E|^{\frac{1}{2}},
\end{aligned}$$

which is the sought after estimate (5.25).

Turning to (5.26), fix $F \subset \mathbb{R}^n$ of finite measure. Accordingly, in what follows, we compute the size with $\mathbf{1}_F$ in place of f. If $|F| \leqslant |E|$ the desired estimate (5.26) actually follows from (5.25), as

$$\Lambda_{\mathbb{P};\tau}(\mathbf{1}_F, g\mathbf{1}_E) \lesssim \|\mathbf{1}_F\|_2 |E|^{\frac{1}{2}} = |F|^{\frac{1}{p}} |F|^{\frac{1}{2} - \frac{1}{p}} |E|^{\frac{1}{2}} \leqslant |F|^{\frac{1}{p}} |E|^{1 - \frac{1}{p}}.$$

We can therefore assume that $|F| > |E|$. Then decomposition (6.7), estimates (6.6) with $\mathbf{1}_F$ in place of f, and the tree lemma imply that for each $\tau \in \{1,\dots,N\}$ we have

$$\begin{aligned}\Lambda_{\mathbb{P};\tau}(\mathbf{1}_F, g\mathbf{1}_E) &\lesssim \sum_{k\in\mathbb{Z}} \min(1, 2^{-\frac{k}{2}}|F|^{\frac12}) \min(1, 2^{-k}|E|) 2^k \\ &\lesssim \sum_{2^k \geqslant |F|} 2^{-\frac{k}{2}}|F|^{\frac12}|E| + \sum_{|F|>2^k\geqslant |E|} |E| + \sum_{2^k<|E|} 2^k \lesssim |E|\left(1+\log\frac{|F|}{|E|}\right) \lesssim |F|^{\frac1p}|E|^{1-\frac1p}.\end{aligned}$$

whence (5.26) follows.

7. Complements

7.1. **Maximal truncations and differentiation of Besov spaces.** Let $\gamma \in \mathscr{S}(\mathbb{R}^d)$ be a radial function satisfying

$$\operatorname{supp}\widehat{\gamma} \subseteq B^d(1), \qquad \widehat{\gamma}(0) = 1, \qquad \|\gamma\|_N := \sup_{|\alpha|,|\beta|\leqslant N} \|x^\alpha D^\beta \gamma\|_{L^\infty(\mathbb{R}^n)} \leqslant C_d. \tag{7.1}$$

A few minor tweaks in the proof yield that the maximally truncated version of (1.8)

$$T^{\star,\star}_{\mathbf{m}} f(x) := \sup_{\sigma\in\mathbf{Gr}(d,n)} \sup_{Q\in\mathrm{SO(d)}} \sup_{h>0} \left| \int_{\mathbb{R}^n} m_\sigma(QO_\sigma\Pi_\sigma\xi)\widehat{\gamma}(hO_\sigma\Pi_\sigma\xi)\widehat{f}(\xi) \mathrm{e}^{2\pi i\langle x,\xi\rangle} \mathrm{d}\xi \right|, \quad x\in\mathbb{R}^n, \tag{7.2}$$

also satisfies the conclusions of Theorem 1.1, under the same assumptions therein, provided that N in (7.1) is sufficiently large, depending upon dimension. Indeed, reductions similar to those in Section 5 may be performed, observing that the smooth cutoff at frequency scale h^{-1} does not modify adaptedness of wave packets at frequency scales $s^{-1} \lesssim h^{-1}$ while erasing the contribution of frequency scales $\gg h^{-1}$. This leads, in analogy with (5.12), to a model sum of type

$$\Lambda^\star_{\mathbb{P};\sigma,h,\tau,M}(f,g) = \sum_{t\in\mathbb{P}} F_M[f](t) A^\star_{\sigma,\tau,M}[g](t),$$

where the coefficients $A_{\sigma,h,\tau,M}$ have been replaced by the modified version

$$A^\star_{\sigma,h,\tau,M}[g](t) := \sup_{\vartheta\in\mathscr{A}_t^M} \left| \left\langle g, \vartheta(\cdot,\sigma(\cdot))\, \mathbf{1}_{\alpha_{t,\tau}}\left(v_{\sigma(\cdot)}\right) \mathbf{1}_{[\mathrm{scl}(t),\infty)}(h(\cdot)) \right\rangle \right|,$$

corresponding to measurable selector functions $h:\mathbb{R}^n \to (0,\infty)$ and $\sigma:\mathbb{R}^n \to \mathbf{Gr}(d,n)$. All arguments of Section 6 work just as well for the modified coefficients with purely notational change: details are left to the interested reader.

If $m_\sigma \equiv 1$, the maximal operator (7.2) coincides up to a constant factor with the maximal subspace averaging operator

$$T^{\star,\star} f(x) = \sup_{\sigma\in\mathbf{Gr}(d,n)} \sup_{h>0} \left| A_{\sigma,h} f(x) \right|, \quad A_{\sigma,h} f(x) := \int_{\mathbb{R}^d} f(x - O_\sigma^{-1} t) \mathrm{Dil}_h^1 \gamma(t)\, \mathrm{d}t, \quad x\in\mathbb{R}^n.$$

Thus Theorem 1.1 implies the weak $(2,2)$ and strong (p,p), $p>2$, bounds for $T^{\star,\star}\circ P_0$, uniformly over γ satisfying (7.1) with N sufficiently large; for example $N = 50(d+1)$ is sufficient. Moreover, using scale invariance and the decay of Schwartz tails we can get that the same boundedness property holds for $T^{\star,\star}$ if $A_{\sigma,h}$ is defined via a Schwartz function

$\gamma \in \mathscr{S}(\mathbb{R}^d)$ satisfying

$$\widehat{\gamma}(0) = 1, \qquad \|\gamma\|_{100d} \leqslant 1;$$

namely we can drop the compact Fourier support assumption. We can use these bounds to conclude Lebesgue differentiation theorems for functions in the nonhomogeneous Besov spaces $B^s_{p,1}(\mathbb{R}^n)$, $s \geqslant 0$, with norm

$$\|f\|_{B^s_{p,1}(\mathbb{R}^n)} \coloneqq \|Q_0 f\|_{L^p(\mathbb{R}^n)} + \sum_{k \geqslant 0} 2^{ks} \|P_k f\|_{L^p(\mathbb{R}^n)},$$

where $Q_0, \{P_k : k \geqslant 0\}$ is any fixed smooth nonhomogeneous Littlewood–Paley decomposition. In the case that $s \in (0,1)$ we can use an alternative characterization of Besov spaces in terms of finite differences, as for example in [41], that shows that $|f| \in B^s_{p,1}$ whenever $f \in B^s_{p,1}$, with a corresponding inequality of the norms, and thus we can conclude that the rough maximal averages

$$M^{\star,\star} f(x) = \sup_{\sigma \in \mathbf{Gr}(d,n)} \sup_{h>0} \langle |f| \rangle_{x,\sigma,h}, \quad \langle f \rangle_{x,\sigma,h} \coloneqq \fint_{B^n(0,h)\cap\sigma} f(x - O_\sigma^{-1} t)\, \mathrm{d}t, \quad x \in \mathbb{R}^n,$$

have the same boundedness properties as $T^{\star,\star}$.

Theorem 7.1. *Let $\sigma : \mathbb{R}^n \to \mathrm{Gr}(n-1,n)$ be any measurable function and let $2 \leqslant p < \infty$. If $f \in B^0_{p,1}(\mathbb{R}^n)$, then*

$$f(x) = \lim_{h \to 0^+} A_{\sigma(x),h} f(x) \qquad \text{a.e. } x \in \mathbb{R}^n,$$

whenever $A_{\sigma,h}$ is defined via a bump function with $\widehat{\gamma}(0) = 1$ and $\|\gamma\|_{100n} \leqslant 1$. Moreover, if $f \in B^s_{p,1}$ for some $s > 0$ then

$$f(x) = \lim_{h \to 0^+} \langle f \rangle_{x,\sigma,h} \qquad \text{a.e. } x \in \mathbb{R}^n.$$

Proof. By the comments preceding the statement of the theorem it suffices to consider the case $s = 0$. By assumption, for each $\varepsilon > 0$ fixed there exists $k = k(\varepsilon) \geqslant 0$ so that

$$\sum_{k > k(\varepsilon)} \|P_k f\|_p < \varepsilon.$$

Let $g = Q_0 f + \sum_{k=0}^{k(\varepsilon)} P_k f$. Note that $g \in C^\infty(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and that

$$\|T^{\star,\star}(f-g)\|_{p,\infty} \lesssim \sum_{k \geqslant k(\varepsilon)} \|T^{\star,\star} P_k f\|_{p,\infty} \lesssim \sum_{k \geqslant k(\varepsilon)} \|P_k f\|_p < \varepsilon$$

by an application of the uniform weak-$L^p(\mathbb{R}^n)$ estimate for $T^{\star,\star} P_k$. The remaining part of the argument is totally analogous to that of the proof of the Lebesgue differentiation theorem using the Hardy–Littlewood maximal theorem and is therefore omitted. □

The results of Theorem 7.1 recover a corresponding result from [33] for Besov spaces $B^s_{p,1}$ with $s > 0$ and $d = 1$; see also [2, 35]. The codimension $n - d = 1$ results as well as the results for zero smoothness appear to be new.

7.2. **Remarks on the bi-parameter problem.** In this paragraph, d is no longer necessarily equal to $n-1$ and we go back to considering values $1 \leqslant d < n$. The rotated codimension

$n-d$ singular integral (1.2), (1.6) may be naturally generalized to the bi-parameter version

$$Tf(x;\sigma,Q,P)=\int_{\mathbb{R}^n} m(QO_\sigma\Pi_\sigma\xi,PO_\sigma\Pi_{\sigma^\perp}\xi)\widehat{f}(\xi)\mathrm{e}^{2\pi i\langle x,\xi\rangle}\mathrm{d}\xi$$

associated to a bi-parameter Hörmander–Mihlin multiplier $m=m(\eta,\zeta)$ on $\mathbb{R}^n=\mathbb{R}^d\otimes\mathbb{R}^{n-d}$, that is, satisfying the estimates

$$\sup_{0\leqslant|\alpha|,|\beta|\leqslant A}\sup_{\xi\in\mathbb{R}^d}\sup_{\zeta\in\mathbb{R}^{n-d}}|\eta|^{|\alpha|}|\zeta|^{|\beta|}\left|D_\eta^\alpha D_\zeta^\beta m(\eta,\zeta)\right|<\infty$$

and parametrized by $Q\in\mathrm{SO}(d),P\in\mathrm{SO}(n-d)$ and $\sigma\in\mathrm{Gr}(d,n)$ or equivalently $\sigma^\perp\in\mathbf{Gr}(n-d,n)$. By arguments analogous to those in Section 3 for the removal of the $\mathrm{SO}(d),\mathrm{SO}(n-d)$ invariances and finite splitting, the study of the corresponding maximal operator may be reduced to estimates for

$$T^\star f(x)=\sup_{\sigma\in\Sigma}|Tf(x,\sigma)|,\qquad Tf(x,\sigma):=\int_{\mathbb{R}^n} m(O_\sigma\xi)\widehat{f}(\xi)\mathrm{e}^{2\pi i\langle x,\xi\rangle}\mathrm{d}\xi,\quad x\in\mathbb{R}^n$$

where $\Sigma\subset\mathrm{Gr}(d,n)$ is a small $2^{-9}\varepsilon$-neighborhood of $\mathbb{R}^d$. When acting on band-limited functions, the maximal operator $T^\star$ may be further reduced, up to L^p-bounded differences, to the sum of a codimension $n-d$ and a codimension d maximal singular integral operator. More specifically, let $\{\Gamma_j:j=-1,0,1\}$ be a partition of unity on the sphere $\mathbb{S}^{n-1}$ subordinated to the covering

$$\begin{aligned}\Gamma_{-1}&:=\{\xi\in\mathbb{S}^{n-1}:|\Pi_{\mathbb{R}^d}\xi|>1-2\varepsilon\},\\ \Gamma_0&:=\{\xi\in\mathbb{S}^{n-1}:\varepsilon<|\Pi_{\mathbb{R}^d}\xi|<1-\varepsilon\},\\ \Gamma_1&:=\{\xi\in\mathbb{S}^{n-1}:|\Pi_{\mathbb{R}^d}\xi|<2\varepsilon\}.\end{aligned}$$

and $P_{\mathrm{cn},j}$, $j=-1,0,1$ be the corresponding smooth conical Fourier cutoff to Γ_j. In analogy to what was done in (5.4), the splitting

$$T^\star\circ P_0 f\leqslant\sum_{j=-1,0,1}T^\star[P_0P_{\mathrm{cn},j}f]$$

may be performed. As the singularities along $\sigma,\sigma^\perp$ sit away from the support of $P_0P_{\mathrm{cn},0}f$ when $\sigma\in\Sigma$, $T^\star[P_0P_{\mathrm{cn},0}f]$ may be easily controlled by the strong maximal function. Furthermore, when $\sigma\in\Sigma$ $|\Pi_{\sigma^\perp}\xi|\sim|\xi|\sim1$ on the support of $P_0P_{\mathrm{cn},-1}$ while $|\Pi_\sigma\xi|\sim|\xi|\sim1$ on the support of $P_0P_{\mathrm{cn},1}$. A discretization procedure similar to that of Subsection 5.4 then shows that e.g. $T^\star[P_0P_{\mathrm{cn},-1}f]$ lies in the convex hull of the model operators

$$\mathrm{T}f(x)=\sum_{t\in\mathbb{P}}\langle f,\varphi_t\rangle\psi_t(x,\sigma(x))\mathbf{1}_{\left[\frac{a}{\mathrm{scl}(t)},\frac{b}{\mathrm{scl}(t)}\right]}(\mathrm{dist}(\sigma(x),\sigma(t))$$

for suitable constants $b>a\gg1$, $\mathbb{P}$ is a collection of tiles t whose spatial localization is a parallelepiped with d sides oriented along the subspace $\sigma(t)$ of sidelength $\mathrm{scl}(t)\gg1$ and $n-d$ short sides of sidelength 1 along $\sigma(t)^\perp$. Also, for a fixed dyadic value of $\mathrm{scl}(t)=\mathrm{scl}$, $\sigma(t)$ varies within a $\sim\mathrm{scl}^{-1}$ net in $\mathrm{Gr}(d,n)$ and for each fixed σ the spatial localizations of the tiles $t\in\mathbb{P}$ with $\sigma(t)=\sigma$ tile $\mathbb{R}^n$. The model for $T^\star[P_0P_{\mathrm{cn},1}f]$ is analogous, up to exchanging $\mathrm{Gr}(d,n)$.

When $n = 2, d = 1$, then $T^\star \circ [P_0 P_{\mathrm{cn},\pm 1}]$ are thus completely symmetric and both may be handled with the methods of Section 6. When either d or $n-d$ are greater than 1, the corresponding model sums are substantially harder. One essential reason is that their scope includes Nikodym maximal averaging operators along subspaces of codimension k larger than one, whose critical exponent, conjectured to be $p = k+1$, is above L^2: see [17] for more details on this conjecture.

REFERENCES

1. Natalia Accomazzo, Francesco Di Plinio, and Ioannis Parissis, *Singular integrals along lacunary directions in* $\mathbb{R}^n$, Adv. Math. **380** (2021), Paper No. 107580, 21. MR 4200473
2. Hugo Aimar, Liliana Forzani, and Virginia Naibo, *Rectangular differentiation of integrals of Besov functions*, Math. Res. Lett. **9** (2002), no. 2-3, 173–189. MR 1909636
3. Michael Bateman, *Kakeya sets and directional maximal operators in the plane*, Duke Math. J. **147** (2009), no. 1, 55–77. MR 2494456
4. ———, *Single annulus L^p estimates for Hilbert transforms along vector fields*, Rev. Mat. Iberoam. **29** (2013), no. 3, 1021–1069. MR 3090145
5. Michael Bateman and Christoph Thiele, *L^p estimates for the Hilbert transforms along a one-variable vector field*, Anal. PDE **6** (2013), no. 7, 1577–1600. MR 3148061
6. J. Bourgain, *A remark on the maximal function associated to an analytic vector field*, Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987), London Math. Soc. Lecture Note Ser., vol. 137, Cambridge Univ. Press, Cambridge, 1989, pp. 111–132. MR 1009171
7. Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157. MR 199631
8. S.-Y. A. Chang, J. M. Wilson, and T. H. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helv. **60** (1985), no. 2, 217–246. MR 800004
9. A. Córdoba and R. Fefferman, *On differentiation of integrals*, Proc. Nat. Acad. Sci. U.S.A. **74** (1977), no. 6, 2211–2213. MR 0476977
10. Antonio Córdoba, *The Kakeya maximal function and the spherical summation multipliers*, Amer. J. Math. **99** (1977), no. 1, 1–22. MR 0447949
11. Ciprian Demeter, *Singular integrals along N directions in* $\mathbb{R}^2$, Proc. Amer. Math. Soc. **138** (2010), no. 12, 4433–4442. MR 2680067
12. Ciprian Demeter and Francesco Di Plinio, *Logarithmic L^p bounds for maximal directional singular integrals in the plane*, J. Geom. Anal. **24** (2014), no. 1, 375–416. MR 3145928
13. Francesco Di Plinio, Shaoming Guo, Christoph Thiele, and Pavel Zorin-Kranich, *Square functions for bi-Lipschitz maps and directional operators*, J. Funct. Anal. **275** (2018), no. 8, 2015–2058. MR 3841536
14. Francesco Di Plinio and Ioannis Parissis, *A sharp estimate for the Hilbert transform along finite order lacunary sets of directions*, Israel J. Math. **227** (2018), no. 1, 189–214. MR 3846321
15. ———, *On the maximal directional Hilbert transform in three dimensions*, Int. Math. Res. Not. IMRN (2020), no. 14, 4324–4356. MR 4126303
16. ———, *Maximal directional operators along algebraic varieties*, Amer. J. Math. **143** (2021), no. 5, 1463–1503. MR 4334401
17. ———, *Maximal subspace averages*, Adv. Math. **410** (2022), Paper No. 108749. MR 4498204
18. Loukas Grafakos, Petr Honzík, and Andreas Seeger, *On maximal functions for Mikhlin-Hörmander multipliers*, Adv. Math. **204** (2006), no. 2, 363–378. MR 2249617
19. Shaoming Guo, *Hilbert transform along measurable vector fields constant on Lipschitz curves: L^2 boundedness*, Anal. PDE **8** (2015), no. 5, 1263–1288. MR 3393679
20. ———, *A geometric proof of Bourgain's L^2 estimate of maximal operators along analytic vector fields*, J. Geom. Anal. **27** (2017), no. 2, 968–985. MR 3625139

21. ———, *Hilbert transform along measurable vector fields constant on Lipschitz curves: L^p boundedness*, Trans. Amer. Math. Soc. **369** (2017), no. 4, 2493–2519. MR 3592519
22. G. A. Karagulyan, *On unboundedness of maximal operators for directional Hilbert transforms*, Proc. Amer. Math. Soc. **135** (2007), no. 10, 3133–3141. MR 2322743
23. Grigori A. Karagulyan and Michael T. Lacey, *On logarithmic bounds of maximal sparse operators*, Math. Z. **294** (2020), no. 3-4, 1271–1281. MR 4074040
24. Nets Hawk Katz, *Maximal operators over arbitrary sets of directions*, Duke Math. J. **97** (1999), no. 1, 67–79. MR 1681088
25. ———, *Remarks on maximal operators over arbitrary sets of directions*, Bull. London Math. Soc. **31** (1999), no. 6, 700–710. MR 1711029
26. Jongchon Kim, *Almost-orthogonality principles for certain directional maximal functions*, J. Geom. Anal. **31** (2021), no. 7, 7320–7332. MR 4289260
27. Jongchon Kim and Malabika Pramanik, *L^2 bounds for a maximal directional Hilbert transform*, Anal. PDE **15** (2022), no. 3, 753–794. MR 4442840
28. I. Łaba, A. Marinelli, and M. Pramanik, *On the maximal directional Hilbert transform*, Anal. Math. **45** (2019), no. 3, 535–568. MR 3995379
29. Michael Lacey and Xiaochun Li, *On a conjecture of E. M. Stein on the Hilbert transform on vector fields*, Mem. Amer. Math. Soc. **205** (2010), no. 965, viii+72. MR 2654385
30. Michael Lacey and Christoph Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370. MR 1783613
31. Michael T. Lacey and Xiaochun Li, *Maximal theorems for the directional Hilbert transform on the plane*, Trans. Amer. Math. Soc. **358** (2006), no. 9, 4099–4117. MR 2219012
32. Andrei K. Lerner and Fedor Nazarov, *Intuitive dyadic calculus: the basics*, Expo. Math. **37** (2019), no. 3, 225–265. MR 4007575
33. Jason Murcko, *Differentiation of Besov spaces and the Nikodym maximal operator*, Proc. Amer. Math. Soc. **145** (2017), no. 5, 2139–2153. MR 3611327
34. 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. MR 0466470
35. Virginia Naibo, *Bessel capacities and rectangular differentiation in Besov spaces*, J. Math. Anal. Appl. **324** (2006), no. 2, 834–840. MR 2265084
36. Javier Parcet and Keith M. Rogers, *Directional maximal operators and lacunarity in higher dimensions*, Amer. J. Math. **137** (2015), no. 6, 1535–1557. MR 3432267
37. 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. MR 613033
38. Per Sjölin, *Convergence almost everywhere of certain singular integrals and multiple Fourier series*, Ark. Mat. **9** (1971), 65–90. MR 336222
39. Elias M. Stein and Brian Street, *Multi-parameter singular Radon transforms III: Real analytic surfaces*, Adv. Math. **229** (2012), no. 4, 2210–2238. MR 2880220
40. Jan-Olov Strömberg, *Maximal functions associated to rectangles with uniformly distributed directions*, Ann. Math. (2) **107** (1978), no. 2, 399–402. MR 0481883
41. Hans Triebel, *Theory of function spaces. II*, Monographs in Mathematics, vol. 84, Birkhäuser Verlag, Basel, 1992. MR 1163193

ODYSSEAS BAKAS, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PATRAS, 26504 PATRAS, GREECE
Email address: obakas@upatras.gr

FRANCESCO DI PLINIO, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI "R. CACCIOPPOLI" UNIVERSITÀ DI NAPOLI "FEDERICO II", VIA CINTIA, MONTE S. ANGELO 80126 NAPOLI, ITALY
Email address: francesco.diplinio@unina.it

IOANNIS PARISSIS, DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DEL PAÍS VASCO, APTDO. 644, 48080 BILBAO, SPAIN AND IKERBASQUE, BASQUE FOUNDATION FOR SCIENCE, BILBAO, SPAIN
Email address: ioannis.parissis@ehu.es

LUZ RONCAL, BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS 48009 BILBAO, SPAIN, IKERBASQUE, BASQUE FOUNDATION FOR SCIENCE, BILBAO, SPAIN, AND DEPARTAMENTO DE MATEMÁTICAS, UNIVERSIDAD DEL PAÍS VASCO, APTDO. 644, 48080 BILBAO, SPAIN
Email address: lroncal@bcamath.org