ANALYSIS & PDE

Volume 16 No. 7 2023

NATALIA ACCOMAZZO, FRANCESCO DI PLINIO,
PAUL HAGELSTEIN, IOANNIS PARISSIS AND LUZ RONCAL

DIRECTIONAL SQUARE FUNCTIONS

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 7, pp. 1651–1699
DOI: 10.2140/apde.2023.16.1651

DIRECTIONAL SQUARE FUNCTIONS

NATALIA ACCOMAZZO, FRANCESCO DI PLINIO,
PAUL HAGELSTEIN, IOANNIS PARISSIS AND LUZ RONCAL

Quantitative formulations of Fefferman's counterexample for the ball multiplier are naturally linked to square function estimates for conical and directional multipliers. We develop a novel framework for these square function estimates, based on a directional embedding theorem for Carleson sequences and multiparameter time-frequency analysis techniques. As applications we prove sharp or quantified bounds for Rubio-de Francia-type square functions of conical multipliers and of multipliers adapted to rectangles pointing along N directions. A suitable combination of these estimates yields a new and currently best-known logarithmic bound for the Fourier restriction to an N-gon, improving on previous results of A. Córdoba. Our directional Carleson embedding extends to the weighted setting, yielding previously unknown weighted estimates for directional maximal functions and singular integrals.

1. Motivation and main results

The celebrated theorem of Charles Fefferman [1971] shows that the ball multiplier is an unbounded operator on $L^p(\mathbb{R}^n)$ for all $p \neq 2$ whenever $n \geq 2$. A well-known argument, originally due to Yves Meyer [de Guzmán 1981], exhibits the intimate relationship of the ball multiplier with vector-valued estimates for directional singular integrals along all possible directions. Fefferman [1971] proved the impossibility of such estimates by testing these vector-valued inequalities on a Kakeya set.

Besicovitch or Kakeya sets are compact sets in the Euclidean space that contain a line segment of unit length in every direction. Sets of this type with zero Lebesgue measure do exist. However, in two dimensions, Kakeya sets are necessarily of full Hausdorff dimension. The question of the Hausdorff

MSC2020: primary 42B20; secondary 42B25.
Keywords: directional operators, directional square functions, Rubio de Francia inequalities, directional Carleson embedding theorems, polygon multiplier.

dimension of Kakeya sets can be then formulated as a question of quantitative boundedness of the Kakeya maximal function, which is a maximal directional average along rectangles of fixed eccentricity and pointing along arbitrary directions.

The importance of the ball multiplier for the summation of higher dimensional Fourier series, as well as its intimate connection to Kakeya sets, have motivated a host of problems in harmonic analysis which have been driving relevant research since the 1970s. Finitary or smooth models of the ball multiplier such as the polygon multiplier and the Bochner–Riesz means quantify the failure of boundedness of the ball multiplier and formalize the close relation of these operators with directional maximal and singular averages.

This paper is dedicated to the study of a variety of operators in the plane that are all connected in one way or another with the ball multiplier. Our point of view is through the analysis of directional operators mapping into $L^p(\mathbb{R}^2; \ell^q)$-spaces where the inner ℓ^q-norm is taken with respect to the set of directions. Different values of q are relevant in our analysis but the cases $q = 2$ and $q = \infty$ are of particular interest. On one hand, the case $q = \infty$ arises when considering maximal directional averages and the corresponding differentiation theory along directions; see [Bateman 2013; Christ et al. 1986; Di Plinio and Parissis 2021; Katz 1999] for classical and recent work on the subject. On the other hand, the case $q = 2$ is especially relevant for Meyer's argument that bounds the norm of a vector-valued directional Hilbert transform by the norm of the ball multiplier. It also arises when dealing with square functions associated to conical or directional Fourier multipliers of the type

$$f \mapsto \{C_j f : j = 1, \dots, N\},$$

where each C_j is adapted to a different coordinate pair and the C_j have disjoint or well-separated Fourier support. These estimates are directional analogues of the celebrated square function estimate for Fourier restriction to families of disjoint cubes, due to Rubio de Francia [1985], and they appear naturally when seeking quantitative estimates on the N-gon Fourier multiplier.

While such square function estimates have been considered previously in the literature, and usually approached directly via weighted norm inequalities, our treatment is novel and leads to improved and in certain cases sharp estimates in terms of the cardinality of the set of directions. It rests on a new directional Carleson measure condition and corresponding embedding theorem, which is subsequently applied to intrinsic directional square functions of time-frequency nature. The link between the abstract Carleson embedding theorem and the applications is provided by directional, one- and two-parameter time-frequency analysis models. The latter allow us to reduce estimates for directional operators to those of the corresponding intrinsic square functions involving directional wave packet coefficients. We note that in the fixed coordinate system case, related square functions have appeared in [Lacey 2007], while a single-scale directional square function similar to those of Section 4 is present in [Di Plinio et al. 2018] by Guo, Thiele, Zorin-Kranich and the second author.

Having clarified the context of our investigation, we turn to the detailed description of our main results and techniques.

A new approach to directional square functions. While we address several types of square functions associated to directional multipliers, our analysis of each relies on a common first step. This is an

L^4-square function inequality for abstract Carleson measures associated with one- and two-parameter collections of rectangles in $\mathbb{R}^2$, pointing along a finite set of N directions; this setup is presented in Section 2 and the central result is Theorem C. Section 2 builds upon the proof technique first introduced in [Katz 1999] and revisited in [Bateman 2013] in the study of sharp weak L^2-bounds for maximal directional operators. Our main novel contributions are the formulation of an abstract directional Carleson condition which is flexible enough to be applied in the context of time-frequency square functions, and the realization that square functions in L^4 can be treated in a TT^*-like fashion. The advancements over [Bateman 2013; Katz 1999] also include the possibility of handling two-parameter collections of rectangles.

In Section 4, we verify that the Carleson condition, which is a necessary assumption in the directional embedding of Theorem C, is satisfied by the intrinsic directional wave packet coefficients associated with certain time-frequency tile configurations, and Theorem C may be thus applied to obtain sharp estimates for discrete time-frequency models of directional Rubio de Francia square functions (for instance). Establishing the Carleson condition requires a precise control of spatial tails of the wave packets; this control is obtained by a careful use of Journé's product theory lemma.

The estimates obtained for the time-frequency model square functions are then applied to three main families of operators described below. All of them are defined in terms of an underlying set of N directions. As in Fefferman's counterexample for the ball multiplier, the Kakeya set is the main obstruction for obtaining uniform estimates. Depending on the type of operator, the usable estimates will be restricted in the range $2 < p < 4$ for square function estimates or in the range $\frac{3}{4} < p < 4$ for the self-adjoint case of the polygon multiplier. The fact that the estimates should be logarithmic in N in the L^p-ranges above is directed by the Besicovitch construction of the Kakeya set. It is easy to see that for p outside this range the only available estimates are essentially trivial polynomial estimates. Further obstructions deter any estimates for Rubio-de-Francia-type square function in the range $p < 2$ already in the one-directional case.

Sharp Rubio de Francia square function estimates in the directional setting. Section 5 concerns quantitative estimates of Rubio de Francia type for the square function associated with N finitely overlapping cone multipliers, of both rough and smooth type. Beginning with the seminal article of Nagel, Stein and Wainger [Nagel et al. 1978], square functions of this type are crucial in the theory of maximal operators, in particular along lacunary directions; see for instance [Parcet and Rogers 2015; Sjögren and Sjölin 1981]. In the case of N uniformly spaced cones, logarithmic estimates with unspecified dependence were proved by A. Córdoba [1982] using weighted theory.

In order to make the discussion above more precise, and to give a flavor of the results of this paper, we introduce some basic notation. Let $\tau \subset [0, 2\pi)$ be an interval and consider the corresponding smooth restriction to the frequency cone subtended by τ, namely

$$C_\tau^\circ f(x) := \int_0^{2\pi}\int_0^\infty \hat{f}(\varrho e^{i\vartheta})\beta_\tau(\vartheta) e^{ix\cdot\varrho e^{i\vartheta}}\varrho \, d\varrho \, d\vartheta, \qquad x \in \mathbb{R}^2,$$

where β_τ is a smooth indicator on τ; namely it is supported in τ and is identically 1 on the middle half of τ.

One of the main results of this paper is a quantitative estimate for a square function associated with the smooth conical multipliers of a finite collection of intervals with bounded overlap. In the statement of the theorem below $\ell^2_{\boldsymbol{\tau}}$ denotes the ℓ^2-norm on the finite set of directions $\boldsymbol{\tau}$.

Theorem A. *Let $\boldsymbol{\tau} = \{\tau\}$ be a finite collection of intervals in $[0, 2\pi)$ with bounded overlap, namely*

$$\Big\| \sum_{\tau \in \boldsymbol{\tau}} \mathbf{1}_\tau \Big\|_\infty \lesssim 1.$$

We then have the square function estimate

$$\|\{C_\tau^\circ f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim_p (\log \#\boldsymbol{\tau})^{\frac{1}{2}-\frac{1}{p}} \|f\|_p$$

for $2 \le p < 4$, as well as the restricted-type analogue valid for all measurable sets E

$$\|\{C_\tau^\circ (f \mathbf{1}_E)\}\|_{L^4(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim (\log \#\boldsymbol{\tau})^{\frac{1}{4}} |E|^{\frac{1}{4}} \|f\|_\infty.$$

The dependence on $\#\boldsymbol{\tau}$ in the estimates above is best possible.

The sharp estimate of Theorem A above can be suitably bootstrapped in order to provide an estimate for rough conical frequency projections; the precise statement can be found in Theorem J of Section 5. The sharpness of the estimates in Theorem A above is discussed in Section 8.6.

A similar square function estimate associated with disjoint rectangular directional frequency projections is presented in Section 6. This is a square function that is very close in spirit to the one originally considered in [Rubio de Francia 1985], and especially to the two-parameter version from [Journé 1985] and revisited in [Lacey 2007]. The novel element is the directional aspect which comes from the fact that the frequency rectangles are allowed to point along a set of N different directions. Our method of proof can deal equally well with one-parameter rectangular projections or collections of arbitrary eccentricities. As before we prove a sharp — in terms of the number of directions — estimate for the smooth square function associated with rectangular frequency projections along N directions; this is the content of Theorem K. The main term in the upper bound of Theorem K matches the logarithmic lower bound associated with the Kakeya set.

The polygon multiplier. The square function estimates discussed above may be combined with suitable vector-valued estimates in the directional setting in order to obtain a quantitative estimate for the operator norm of the N-gon multiplier, namely the Fourier restriction to a regular N-gon $\mathcal{P}_N$,

$$T_{\mathcal{P}_N} f(x) := \int_{\mathcal{P}_N} \hat{f}(\xi) e^{i x \cdot \xi} \, d\xi, \qquad x \in \mathbb{R}^2. \tag{1.1}$$

In Section 7 we give the details and proof of the following quantitative estimate for the polygon multiplier.

Theorem B. *Let $\mathcal{P}_N$ be a regular N-gon in $\mathbb{R}^2$ and $T_{\mathcal{P}_N}$ be the corresponding Fourier restriction operator defined in (1.1). We have the estimate*

$$\|T_{\mathcal{P}_N} : L^p(\mathbb{R}^2)\| \lesssim (\log N)^{4\left|\frac{1}{2}-\frac{1}{p}\right|}, \qquad \tfrac{4}{3} < p < 4.$$

We limit ourselves to treating the regular N-gon case; however, it will be clear from the proof that this restriction may be significantly weakened by requiring instead a well-distribution-type assumption on the arcs defining the polygon, similar to the one that is implicit in Theorem A.

Precise L^p-bounds for the N-gon multiplier as a function of N quantify Fefferman's counterexample and so the failure of boundedness of the ball multiplier when $p \neq 2$. A logarithmic-type estimate for $T_{\mathcal{P}_N}$

was first obtained in [Córdoba 1977]. While the exact dependence in that work is not explicitly tracked, the upper bound on the operator norm obtained there must be necessarily larger than $O(\log N)^{5/4}$ for p close to the endpoints of the relevant interval; see Remark 7.12 and Section 8.4 for details. While the dependence obtained in Theorem B is a significant improvement over previous results, it does not match the currently best-known lower bound, which is the same as that for the Meyer lemma constant in Lemma 7.21 and Section 8.1.

Remark. Let $\delta > 0$ and T_j be a smooth frequency restriction to one of the $O(\delta^{-1})$ tangential $\delta \times \delta^2$ boxes covering the δ^2 neighborhood of $\mathbb{S}^1$. Unlike the sharp forward square function estimate we prove in this article, the *reverse square function* estimate

$$\|f\|_p \leq C_{p,\delta} \|\{T_j f : 1 \leq j \leq O(1/\delta)\}\|_{L^p(\mathbb{R}^2:\ell^2_j)} \tag{1.2}$$

holds with $C_{4,\delta} = O(1)$ at the endpoint $p = 4$. For the proof of this L^4-decoupling estimate, see [Córdoba 1977; Fefferman 1973]. An extension to the range $2 < p < 4$ is at the moment only possible via vector-valued methods, which introduce the loss $C_{p,\delta} = O(|\log \delta|^{1/2-1/p})$. In fact (1.2) with the loss $C_{p,\delta}$ claimed above follows easily from Lemma 7.18; the details are contained in Remark 7.22.

Reverse square function inequalities of the type (1.2) have been popularized by Wolff in his proof of local smoothing estimates in the large p regime; see also [Garrigós and Seeger 2010; Łaba and Pramanik 2006; Łaba and Wolff 2002; Pramanik and Seeger 2007]. We refer to [Carbery 2015] for a proof that the $p = 2n/(n-1)$ case of the $\mathbb{S}^{n-1}$ reverse square function estimate implies the corresponding $L^n(\mathbb{R}^n)$ Kakeya maximal inequality, as well as the Bochner–Riesz conjecture. In [Carbery 2015], the author also asks whether a δ-free estimate holds in the range $2 < p < 2n/(n-1)$. At the moment this is not known in any dimension.

On a different but related note, weakening (1.2) by replacing the right-hand side with the larger square function of $\|f_j\|_p$ yields a sample (weak) *decoupling* inequality: a full range of sharp decoupling inequalities for hypersurfaces with curvature have been established starting from the recent, seminal paper [Bourgain and Demeter 2015]. In the case of $\mathbb{S}^1$, the weak decoupling inequality holds in the wider range $2 \leq p \leq 6$, with $C_\varepsilon \delta^{-\varepsilon}$-type bounds outside of $[2, 4]$; our methods do not seem to provide insights on the quantitative character of weak decoupling in this wider range.

Weighted estimates for the maximal directional function. The simplest example of an application of the directional Carleson embedding theorem is the adjoint of the directional maximal function; this was already noticed by Bateman [2013], re-elaborating on the approach of [Katz 1999]. By duality, the L^2-directional Carleson embedding theorem of Section 2 yields the sharp bound for the weak-$(2,2)$-norm of the maximal Hardy–Littlewood maximal function M_N along N *arbitrary directions*

$$\|M_N : L^2(\mathbb{R}^2) \to L^{2,\infty}(\mathbb{R}^2)\| \sim \sqrt{\log N};$$

this result first appeared in the quoted article [Katz 1999].

Theorem C may be extended to the directional weighted setting. We describe this extension in Section 3, see Theorem D, and derive several novel weighted estimates for directional maximal and singular integrals as an application.

More specifically, our weighted Carleson embedding Theorem D yields a Fefferman–Stein-type inequality for the operator M_N with sharp dependence on the number of directions; this result is the content of Theorem E. Specializing to A_1-weights in the directional setting yields the first sharp weighted result for the maximal function along arbitrary directions. Furthermore, Theorem F contains an $L^{2,\infty}(w)$-estimate for the maximal directional singular integrals along N directions, for suitable directional weights w, with a quantified logarithmic dependence in N. This is a weighted counterpart of the results of [Demeter 2010; Demeter and Di Plinio 2014].

2. An L^2-inequality for directional Carleson sequences

In this section we prove an abstract L^2-inequality for certain Carleson sequences adapted to sets of directions: the main result is Theorem C below. The Carleson sequences we will consider are indexed by parallelograms with long side pointing in a given set of directions in $\mathbb{R}^2$, and possessing certain natural properties. The definitions below are motivated by the applications we have in mind, all of them lying in the realm of directional singular and averaging operators.

2.1. ***Parallelograms and sheared grids.*** Fix a coordinate system and the associated horizontal and vertical projections of $A \subset \mathbb{R}^2$:

$$\pi_1(A) := \{x \in \mathbb{R} : \{x\} \times \mathbb{R} \cap A \neq \varnothing\}, \quad \pi_2(A) := \{y \in \mathbb{R} : \mathbb{R} \times \{y\} \cap A \neq \varnothing\}.$$

Fix a finite set of slopes $S \subset [-1, 1]$. Throughout, we indicate by $N = \#S$ the number of elements of S. In general we will deal with sets of directions

$$V := \{(1, s) : s \in S\}, \quad V^\perp := \{(-s, 1) : s \in S\}.$$

We will conflate the descriptions of directions in terms of slopes in S and in terms of vectors in V with no particular mention.

For each $s \in S$ let

$$\begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix}$$

be the corresponding shearing matrix. A *parallelogram along s* is the image $P = A_s(I \times J)$ of the rectangular box $I \times J$ in the fixed coordinate system with $|I| \geq |J|$. We denote the collection of parallelograms along s by $\mathcal{P}^2_s$ and

$$\mathcal{P}^2_S := \bigcup_{s \in S} \mathcal{P}^2_s.$$

In order to describe the setup for our general result we introduce a collection of directional dyadic grids of parallelograms. In order to define these grids we consider the two-parameter product dyadic grid

$$\mathcal{D}^2_0 := \{R = I \times J : I, J \in \mathcal{D}(\mathbb{R}),\ |I| \geq |J|\}$$

obtained by taking the cartesian product of the standard dyadic grid $\mathcal{D}(\mathbb{R})$ with itself; we note that we only consider the rectangles in $\mathcal{D} \times \mathcal{D}$ whose horizontal side is longer than their vertical one. Define the

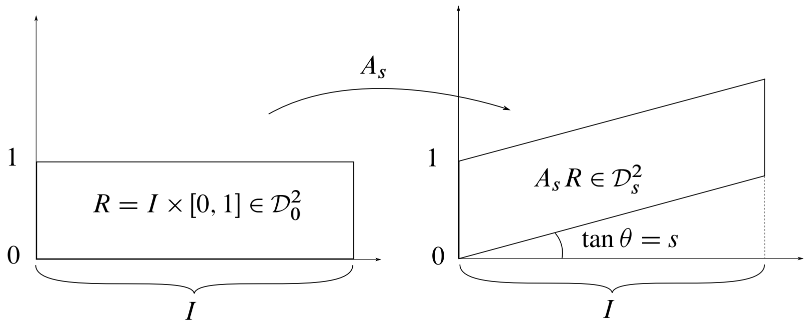

Figure 1. The axis-parallel rectangle $R \in \mathcal{D}^2_0$ is mapped to the slanted parallelogram $A_s R \in \mathcal{D}^2_s$.

sheared grids

$$\mathcal{D}^2_s := \{A_s R : R \in \mathcal{D}^2_0\}, \quad s \in S, \qquad \mathcal{D}^2_S := \bigcup_{s\in S} \mathcal{D}^2_s.$$

We will also use the notation

$$\mathcal{D}^2_{s,k_1,k_2} := \{A_s R : R = I \times J \in \mathcal{D}^2_0,\ |I| = 2^{-k_1},\ |J| = 2^{-k_2}\}, \quad s \in S,\ k_1, k_2 \in \mathbb{Z},\ k_1 \le k_2.$$

Note that $\mathcal{D}^2_s$ is a special subcollection of $\mathcal{P}^2_s$. In particular, $R \in \mathcal{D}^2_s$ is a parallelogram oriented along $v = (1, s)$ with vertical sides parallel to the y-axis and such that $\pi_1(R)$ is a standard dyadic interval. Furthermore our assumptions on S and the definition of $\mathcal{D}^2_0$ imply that the parallelograms in $\mathcal{D}^2_S$ have long side with slope $|s| \le 1$ and a vertical short side. See Figure 1. With a slight abuse of language we will continue referring to the rectangles in $\mathcal{D}^2_S$ as *dyadic*.

Several results in this paper will involve collections of parallelograms $\mathcal{R} \subset \mathcal{D}^2_S$. Writing $\mathcal{R}_s := \mathcal{R} \cap \mathcal{D}^2_s$ we have the natural decomposition of $\mathcal{R}$ into $\#S = N$ subcollections

$$\mathcal{R} = \bigcup_{s\in S} \mathcal{R}_s.$$

In general for any collection $\mathcal{R}$ of parallelograms we will use the notation

$$\mathsf{sh}(\mathcal{R}) := \bigcup_{R\in\mathcal{R}} R$$

for the *shadow* of the collection. Finally, for any collection of parallelograms $\mathcal{R}$ we define the corresponding maximal operator

$$\mathrm{M}_{\mathcal{R}} f(x) := \sup_{R\in\mathcal{R}} \langle |f| \rangle_R \mathbf{1}_R(x), \quad f \in L^1_{\mathrm{loc}}(\mathbb{R}^2),\ x \in \mathbb{R}^2. \tag{2.2}$$

We will also use the following notation for directional maximal functions:

$$\mathrm{M}_v f(x) := \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |f(x + tv)|\,\mathrm{d}t, \quad \mathrm{M}_j f(x) := \mathrm{M}_{e_j} f(x), \quad j \in \{1,2\},\ x \in \mathbb{R}^2. \tag{2.3}$$

If $V \subset \mathbb{R}^2$ is a compact set of directions with $0 \notin V$, we write

$$\mathrm{M}_V f := \sup_{v\in V} \mathrm{M}_v f. \tag{2.4}$$

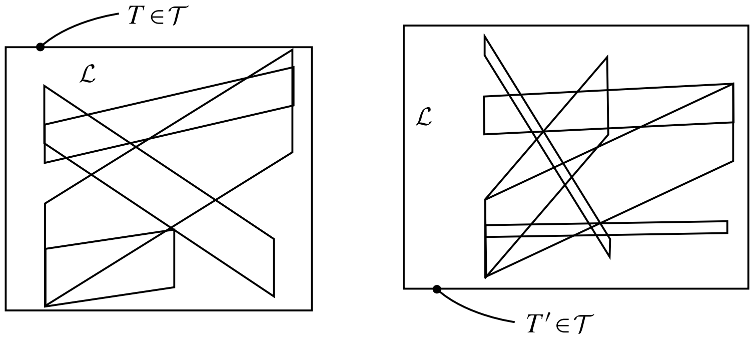

Figure 2. A collection $\mathcal{L}$ subordinate to a collection $\mathcal{T} \subset \mathcal{P}^2_0$.

In the definitions above and throughout the paper we use the notation

$$\langle g \rangle_E = \fint_E g := \frac{1}{|E|} \int_E g(x)\,\mathrm{d}x$$

whenever g is a locally integrable function in $\mathbb{R}^2$ and $E \subset \mathbb{R}^2$ has finite measure.

2.5. ***An embedding theorem for directional Carleson sequences.*** In this section we will be dealing with Carleson-type sequences $a = \{a_R\}_{R \in \mathcal{D}^2_S}$, indexed by dyadic parallelograms. In order to define them precisely we need a preliminary notion.

Definition 2.6. Let $\mathcal{L} \subset \mathcal{P}^2_S$ be a collection of parallelograms and let $s \in S$. We will say that $\mathcal{L}$ is subordinate to a collection $\mathcal{T} \subset \mathcal{P}^2_s$ if for each $L \in \mathcal{L}$ there exists $T \in \mathcal{T}$ such that $L \subset T$; see Figure 2.

It is important to stress that collections $\mathcal{L}$ are subordinate to rectangles $\mathcal{T} \subset \mathcal{P}^2_s$ *having a fixed slope* s. The Carleson sequences $a = \{a_R\}_{R \in \mathcal{R}}$ we will be considering will fall under the scope of the following definition.

Definition 2.7. Let $a = \{a_R\}_{R \in \mathcal{D}^2_S}$ be a sequence of nonnegative numbers. Then a will be called an L^∞*-normalized* Carleson sequence if for every $\mathcal{L} \subset \mathcal{D}^2_S$ which is subordinate to some collection $\mathcal{T} \subset \mathcal{P}^2_\tau$ for some fixed $\tau \in S$, we have

$$\sum_{L \in \mathcal{L}} a_L \leq |\mathrm{sh}(\mathcal{T})|$$

and the quantity

$$\mathsf{mass}_a := \sum_{R \in \mathcal{D}^2_S} a_R$$

is finite. Given a Carleson sequence $a = \{a_R : R \in \mathcal{D}^2_S\}$ and a collection $\mathcal{R} \subset \mathcal{D}^2_S$ we define the corresponding *balayage*

$$T_{\mathcal{R}}(a)(x) := \sum_{R \in \mathcal{R}} a_R \frac{\mathbf{1}_R(x)}{|R|}, \quad x \in \mathbb{R}^2. \tag{2.8}$$

We write $T(a)$ for $T_{\mathcal{R}}(a)$ when $\mathcal{R} = \mathcal{D}^2_S$. For $1 \leq p \leq 2$ we then define the balayage norms

$$\mathsf{mass}_{a,p}(\mathcal{R}) := \|T_{\mathcal{R}}(a)\|_{L^p}.$$

Note that $\mathsf{mass}_{a,1}(\mathcal{R}) = \sum_{R \in \mathcal{R}} a_R \leq \mathsf{mass}_a$.

Remark 2.9 (elementary properties of mass). Let $\mathcal{R} \subset \mathcal{D}^2_\tau$ for some fixed $\tau \in S$. Then $\mathcal{R}$ is subordinate to itself and if a is an L^∞-normalized Carleson sequence we have

$$\mathsf{mass}_{a,1}(\mathcal{R}) = \sum_{R\in\mathcal{R}} a_R \le |\mathsf{sh}(\mathcal{R})|, \quad \mathcal{R} \subset \mathcal{D}^2_\tau \text{ for some fixed } \tau \in S.$$

Also, the very definition of mass and the log-convexity of the L^p-norm imply

$$\mathsf{mass}_{a,p}(\mathcal{R}) \le \mathsf{mass}_{a,1}(\mathcal{R})^{1-\frac{2}{p'}} \, \mathsf{mass}_{a,2}(\mathcal{R})^{\frac{2}{p'}} \tag{2.10}$$

for all $1 \le p \le 2$, with p' its dual exponent.

We are now ready to state the main result of this section. The result below should be interpreted as a reverse Hölder-type bound for the balayages of directional Carleson sequences.

Theorem C. *Let $S \subset [-1, 1]$ be a finite set of N slopes and $\mathcal{R} \subset \mathcal{D}^2_S$. Suppose that the maximal operators $\{\mathrm{M}_{\mathcal{R}_s} : s \in S\}$ satisfy*

$$\sup_{s\in S} \|\mathrm{M}_{\mathcal{R}_s} : L^p \to L^{p,\infty}\| \lesssim (p')^\gamma, \quad p \to 1^+,$$

for some $\gamma \ge 0$. Then for every L^∞-normalized Carleson sequence $a = \{a_R\}_{R\in\mathcal{D}^2_S}$

$$\mathsf{mass}_{a,2}(\mathcal{R}) \lesssim (\log N)^{\frac{1}{2}} ((1+\gamma) \log\log N)^{\frac{\gamma}{2}} \, \mathsf{mass}_{a,1}(\mathcal{R})^{\frac{1}{2}}.$$

The proof of Theorem C occupies the next subsection. The argument relies on several lemmas, whose proof is postponed to Section 2.23.

Remark 2.11. There are essentially two cases in the assumption of Theorem C above. If for each $s \in S$ the family $\mathcal{R}_s$ happens to be a one-parameter family, then the corresponding maximal operator $\mathrm{M}_{\mathcal{R}_s}$ is of weak-type-$(1, 1)$, whence the assumption holds with $\gamma = 0$. In the generic case that $\mathcal{R} = \mathcal{D}^2_S$, for each s the operator $\mathrm{M}_{\mathcal{R}_s} = \mathrm{M}_{\mathcal{D}^2_S}$ is a skewed copy of the strong maximal operator and the assumption holds with $\gamma = 1$.

2.12. ***Main line of proof of Theorem C.*** Throughout the proof, we use the following partial order between parallelograms $Q, R \in \mathcal{D}^2_S$:

$$Q \le R \quad \overset{\text{def}}{\Longleftrightarrow} \quad Q \cap R \ne \varnothing, \;\; \pi_1(Q) \subseteq \pi_1(R). \tag{2.13}$$

Notice that, since $Q, R \in \mathcal{D}^2_S$, we have that $\pi_1(R), \pi_1(Q)$ belong to the standard dyadic grid $\mathcal{D}$ on $\mathbb{R}$.

It is convenient to encode the main inequality of Theorem C by means of the following dimensionless quantity associated with a collection $\mathcal{R} \subset \mathcal{D}^2_S$ and a Carleson sequence $a = \{a_R\}_{R\in\mathcal{D}^2_S}$:

$$\mathsf{U}_p(\mathcal{R}) := \sup_{\substack{\mathcal{L}\subset\mathcal{R} \\ a=\{a_R\}}} \frac{\mathsf{mass}_{a,p}(\mathcal{L})}{\mathsf{mass}_{a,1}(\mathcal{L})^{\frac{1}{p}}},$$

where the supremum is taken over all finite subcollections $\mathcal{L} \subset \mathcal{R}$ and all L^∞-normalized Carleson sequences $a = \{a_R\}_{R\in\mathcal{D}^2_S}$. There is an easy, albeit lossy, a priori estimate for $\mathsf{U}_p(\mathcal{R})$ for general $\mathcal{R} \subset \mathcal{D}^2_S$.

Lemma 2.14. *Let $S \subset [-1,1]$ be a finite set of N slopes and $a = \{a_R\}_{R\in\mathcal{R}}$ be a normalized Carleson sequence as above. For every $\mathcal{R} \subset \mathcal{D}^2_S$ we have the estimate*

$$\mathsf{U}_p(\mathcal{R}) \lesssim N^{\frac{1}{p'}} \sup_{s\in S} \|\mathrm{M}_{\mathcal{R}_s} : L^{p'} \to L^{p',\infty}\|, \quad 1 < p < \infty.$$

Theorem C is then an easy consequence of the following bootstrap-type estimate. For an arbitrary finite collection of parallelograms $\mathcal{R} \subset \mathcal{D}^2_S$ we will prove the estimate

$$\mathsf{U}_2(\mathcal{R})^2 \lesssim (\log \mathsf{U}_2(\mathcal{R}))^\gamma \log N, \tag{2.15}$$

with absolute implicit constant. Note also that the boundedness assumption on $\mathrm{M}_{\mathcal{R}_s}$ for some $p < 2$ and Lemma 2.14 yield the a priori estimate $\mathsf{U}_2(\mathcal{R}) \lesssim N^{1/2}$. Inserting this a priori estimate into (2.15) and bootstrapping will then complete the proof of Theorem C. It thus suffices to prove (2.15) to obtain Theorem C.

The remainder of the section is dedicated to the proof of (2.15). We begin by expanding the square of the L^2-norm of $T_\mathcal{R}(a)$ as follows:

$$\mathsf{mass}_{a,2}(\mathcal{R})^2 = \|T_\mathcal{R}(a)\|_2^2 \le 2 \sum_{R\in\mathcal{R}} a_R \frac{1}{|R|} \int_R \sum_{\substack{Q\in\mathcal{R}\\ Q\le R}} a_Q \frac{\mathbf{1}_Q}{|Q|} =: 2 \sum_{R\in\mathcal{R}} a_R B_R^{\mathcal{R}}. \tag{2.16}$$

For any $\mathcal{L} \subset \mathcal{R}$ and $R \in \mathcal{R}$ we have implicitly defined

$$B_R^{\mathcal{L}} := \frac{1}{|R|} \int_R \sum_{\substack{Q\in\mathcal{L}\\ Q\le R}} a_Q \frac{\mathbf{1}_Q}{|Q|}. \tag{2.17}$$

Remark 2.18. Observe that for any $\mathcal{L} \subset \mathcal{R}$ and every fixed $s \in S$ we have

$$\bigcup \{R \in \mathcal{R}_s : B_R^{\mathcal{L}} > \lambda\} \subset \left\{x \in \mathbb{R}^2 : \mathrm{M}_{\mathcal{R}_s}\Big[\sum_{Q\in\mathcal{L}} a_Q \frac{\mathbf{1}_Q}{|Q|}\Big](x) > \lambda\right\},$$

which by our assumption on the weak (p,p) norm of $\mathrm{M}_{\mathcal{R}_s}$ implies

$$\sup_{s\in S} \left| \bigcup \left\{ R \in \mathcal{R}_s : B_R^{\mathcal{L}} > \lambda \right\} \right| \lesssim (p')^\gamma \frac{\mathsf{mass}_{a,p}(\mathcal{L})^p}{\lambda^p}, \quad p \to 1^+.$$

For a numerical constant $\lambda \ge 1$, to be chosen at the end of the proof, a nonnegative integer k and $s \in S$ we consider subcollections of $\mathcal{R}_s$ as follows:

$$\mathcal{R}_{s,k} := \{R : R \in \mathcal{R}_s,\ \lambda k \le B_R^{\mathcal{R}} < \lambda(k+1)\}, \quad k \in \mathbb{N},\ s \in S. \tag{2.19}$$

Using (2.16) we have

$$\begin{aligned} \|T_\mathcal{R}(a)\|_2^2 &\lesssim \sum_{s\in S} \sum_{k=0}^{N} k\lambda \sum_{R\in\mathcal{R}_{s,k}} a_R + N \sup_{s\in S} \Big[\sum_{k>N} k\lambda \sum_{R\in\mathcal{R}_{s,k}} a_R \Big] \\ &\lesssim \lambda (\log N)\, \mathsf{mass}_{a,1}(\mathcal{R}) + \lambda N \sum_{k>N} k \sup_{s\in S} |\mathsf{sh}(\mathcal{R}_{s,k})|. \end{aligned} \tag{2.20}$$

Here $\lambda > 0$ is the constant used to define the collections $\mathcal{R}_{s,k}$ and in the last lines we used the definition of a Carleson sequence and Remark 2.9.

The following lemma encodes the exponential decay relation between mass and $B_R^{\mathcal{L}}$ and is in fact the main step of the proof of Theorem C.

Lemma 2.21. *Let $a = \{a_R : R \in \mathcal{D}_S^2\}$ be an L^∞-normalized Carleson sequence, $S \subset [-1,1]$, and $\mathcal{L}, \mathcal{R} \subset \mathcal{D}_S^2$ with $\mathcal{L} \subseteq \mathcal{R}$. We assume that for some $p \in [1,2)$*

$$A_p := \sup_{s\in S} \|\mathrm{M}_{\mathcal{R}_s} : L^p \to L^{p,\infty}\| < +\infty.$$

If $\lambda \geq C \max(1, A_p \mathsf{U}_2(\mathcal{L})^{2/p'})$ for a sufficiently large numerical constant $C > 1$ then there exists $\mathcal{L}_1 \subset \mathcal{L}$ such that

(i) $\mathsf{mass}_{a,1}(\mathcal{L}_1) \leq \frac{1}{2}\mathsf{mass}_{a,1}(\mathcal{L})$,

(ii) *fixing $s \in S$ and denoting by $\mathcal{R}'_s$ the collection of rectangles R in $\mathcal{R}_s$ with $B_R^{\mathcal{L}} > \lambda$, see (2.17), we have that*

$$B_R^{\mathcal{L}} \leq \lambda + B_R^{\mathcal{L}_1} \quad \textit{for all } R \in \mathcal{R}'_s.$$

The final lemma we make use of in the argument translates the exponential decay of the mass of each $\mathcal{R}_{s,k}$ into exponential decay of the support size, which is what we need in the estimate (2.20).

Lemma 2.22. *Let $S \subset [-1,1]$ and define the collections $\mathcal{R}_{s,k}$ by (2.19) with λ defined as in Lemma 2.21 for $\mathcal{L} = \mathcal{R}$*

$$\lambda := C \max(1, A_p \mathsf{U}_2(\mathcal{R})^{\frac{2}{p'}}).$$

We assume that the operators $\{\mathrm{M}_{\mathcal{R}_s} : s \in S\}$ map $L^p(\mathbb{R}^2)$ to $L^{p,\infty}(\mathbb{R}^2)$ uniformly with constant A_p. For $k \geq 1$ we then have the estimate

$$|\mathsf{sh}(\mathcal{R}_{s,k})| \lesssim 2^{-k}\,\mathsf{mass}_{a,1}(\mathcal{R}),$$

with absolute implicit constant.

With these lemmas in hand we now return to the proof of (2.15). Substituting the estimate of Lemma 2.22 into (2.20) yields

$$\|T_{\mathcal{R}}(a)\|_2^2 \lesssim \lambda\,\mathsf{mass}_{a,1}(\mathcal{R})\Bigg[(\log N) + N \sum_{k \geq \log N} k2^{-k}\Bigg] \lesssim \lambda\,\mathsf{mass}_{a,1}(\mathcal{R})(\log N).$$

This was proved for an arbitrary collection $\mathcal{R}$ and so also for every $\mathcal{L} \subset \mathcal{R}$. Thus the estimate above and our assumption $A_p \lesssim (p')^\gamma$ imply

$$\mathsf{U}_2(\mathcal{R})^2 \lesssim \lambda(\log N), \qquad \lambda \gtrsim \max(1, (p')^\gamma \mathsf{U}_2(\mathcal{R})^{\frac{2}{p'}}).$$

Now observe that we can assume $\mathsf{U}_2(\mathcal{R}) \gtrsim 1$; otherwise there is nothing to prove. In this case we can take

$$\lambda \simeq (p')^\gamma \mathsf{U}_2(\mathcal{R})^{\frac{2}{p'}}$$

for every $p > 1$. The choice $p' := (\log \mathsf{U}_2(\mathcal{R}))$ guarantees that $[\mathsf{U}_2(\mathcal{R})]^{1/p'} \lesssim 1$ and leads to

$$\mathsf{U}_2(\mathcal{R})^2 \lesssim (\log \mathsf{U}_2(\mathcal{R}))^\gamma \log N.$$

This is the desired estimate (2.15) and so the proof of Theorem C is complete.

2.23. *Proofs of the lemmas.*

Proof of Lemma 2.14. We follow the proof of [Lacey 2007, Lemma 3.11]. Take $\mathcal{R}$ to be some finite collection and $\|g\|_{p'} = 1$ such that

$$\Big\| \sum_{R \in \mathcal{R}} a_R \frac{\mathbf{1}_R}{|R|} \Big\|_p = \int \sum_{R \in \mathcal{R}} a_R \frac{\mathbf{1}_R}{|R|} g.$$

Define $\mathcal{R}' := \{R \in \mathcal{R} : \langle g \rangle_R > [cN/\operatorname{mass}_{a,1}(\mathcal{R})]^{1/p'}\}$ for some $c > 1$ and $\mathcal{R}'_s := \mathcal{R}' \cap \mathcal{D}^2_s$ for $s \in S$. Then,

$$\int \sum_{R \in \mathcal{R}} a_R \frac{\mathbf{1}_R}{|R|} g \le \sum_{R \in \mathcal{R} \setminus \mathcal{R}'} a_R \langle g \rangle_R + \Big\| \sum_{R \in \mathcal{R}'} a_R \frac{\mathbf{1}_R}{|R|} \Big\|_p \le (cN)^{\frac{1}{p'}} \Big(\sum_{R \in \mathcal{R}} a_R \Big)^{\frac{1}{p}} + N \sup_{s \in S} \Big\| \sum_{R \in \mathcal{R}'_s} a_R \frac{\mathbf{1}_R}{|R|} \Big\|_p.$$

This means

$$\Big\| \sum_{R \in \mathcal{R}} a_R \frac{\mathbf{1}_R}{|R|} \Big\|_p \lesssim (cN)^{\frac{1}{p'}} \Bigg(1 + \frac{N^{\frac{1}{p}}}{c^{\frac{1}{p'}}} \sup_{s \in S} \frac{\big\| \sum_{R \in \mathcal{R}'_s} a_R (\mathbf{1}_R/|R|) \big\|_p}{\big(\sum_{R \in \mathcal{R}'_s} a_R \big)^{\frac{1}{p}}} \frac{\big(\sum_{R \in \mathcal{R}'_s} a_R \big)^{\frac{1}{p}}}{\big(\sum_{R \in \mathcal{R}_s} a_R \big)^{\frac{1}{p}}} \Bigg) \Big(\sum_{R \in \mathcal{R}} a_R \Big)^{\frac{1}{p}}.$$

We have proved that for an arbitrary collection $\mathcal{R}$ we have

$$\mathsf{U}_p(\mathcal{R}) \le (cN)^{\frac{1}{p'}} \Bigg(1 + \frac{N^{\frac{1}{p}}}{c^{\frac{1}{p'}}} \sup_s \mathsf{U}_p(\mathcal{R}'_s) \frac{\operatorname{mass}_{a,1}(\mathcal{R}'_s)^{\frac{1}{p}}}{\operatorname{mass}_{a,1}(\mathcal{R})^{\frac{1}{p}}} \Bigg).$$

We claim that $\sup_{s \in S} \mathsf{U}_p(\mathcal{R}'_s) \lesssim \sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_s} : L^{p'} \to L^{p',\infty}\|$. Assuming this for a moment and using Remark 2.9 we can estimate

$$\begin{aligned} \sum_{R \in \mathcal{R}'_s} a_R \le |\operatorname{sh}(\mathcal{R}'_s)| &\le |\{\mathrm{M}_{\mathcal{R}_s}(g) > (cN/\operatorname{mass}_{a,1}(\mathcal{R}))^{1/p'}\}| \\ &\le \sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_s} : L^{p'} \to L^{p',\infty}\|^{p'} \frac{\operatorname{mass}_{a,1}(\mathcal{R})}{cN}. \end{aligned}$$

This proves the proposition upon choosing $c \gtrsim \sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_s} : L^{p'} \to L^{p',\infty}\|^{p'}$.

We have to prove the claim. Note that since $\mathcal{R}'_s$ is a collection in a fixed direction, the inequality $\mathsf{U}_{\mathcal{R}'_s} \lesssim \sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_s} : L^{p'} \to L^{p',\infty}\|$ follows by the John–Nirenberg inequality in the product setting and Remark 2.9; see [Lacey 2007, Lemma 3.11]. □

Proof of Lemma 2.21. By the invariance under shearing of our statement, we can work in the case $s = 0$. Therefore, $\mathcal{R}'_0$ will stand for the collection of rectangles in $\mathcal{R}_0$ such that $B^{\mathcal{L}}_R > \lambda$, where $\lambda \ge C$ and $C > 1$ will be specified at the end of the proof. We write $R = I_R \times L_R$ for $R \in \mathcal{R}_0$.

Inside-outside splitting. For $I \in \{\pi_1(R) : R \in \mathcal{R}'_0\}$ and any interval K we define

$$\mathcal{L}^{\text{in}}_{I,K} := \{Q \in \mathcal{L} : Q \le I \times K,\ \pi_2(Q) \subset 3K\}, \quad \mathcal{L}^{\text{out}}_{I,K} := \{Q \in \mathcal{L} : Q \le I \times K,\ \pi_2(Q) \not\subseteq 3K\},$$

where we recall that the definition of partial order $Q \le R$ was given in (2.13). Set also

$$B^{\text{in}}_{I,K} := \fint_{I \times K} \sum_{Q \in \mathcal{L}^{\text{in}}_{I,K}} \frac{a_Q}{|Q|} \mathbf{1}_Q, \quad B^{\text{out}}_{I,K} := \fint_{I \times K} \sum_{Q \in \mathcal{L}^{\text{out}}_{I,K}} \frac{a_Q}{|Q|} \mathbf{1}_Q.$$

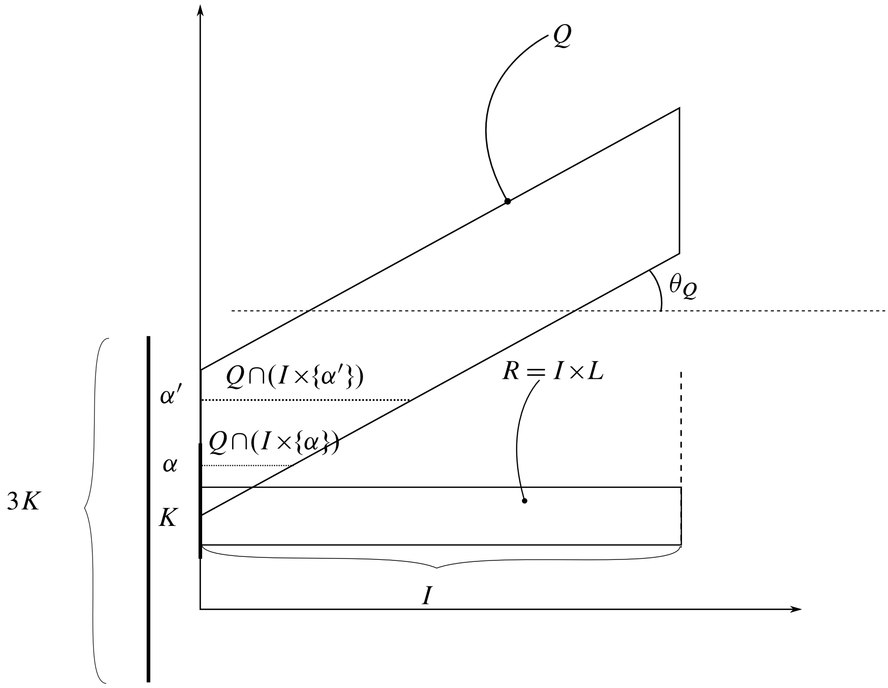

Figure 3. A rectangle Q with angle θ_Q intersecting $R = I \times L \subset I \times K$.

We claim that if $K \subset \mathbb{R}$ is any interval then for all $\alpha \in K$ we have

$$\fint_{I\times\{\alpha\}} \sum_{Q\in\mathcal{L}^{\mathrm{out}}_{I,K}} a_Q \frac{\mathbf{1}_Q}{|Q|} = \sum_{Q\in\mathcal{L}^{\mathrm{out}}_{I,K}} a_Q \frac{|Q\cap(I\times\{\alpha\})|}{|Q|} \lesssim \fint_{I\times 3K} \sum_{Q\in\mathcal{L}^{\mathrm{out}}_{I,K}} a_Q \frac{\mathbf{1}_Q}{|Q|}. \tag{2.24}$$

To see this note that in order for a Q-term appearing in the sum of the left-hand side above to be nonzero we must have

$$\pi_1(Q) \subset I, \quad \pi_2(Q)\cap K \neq \varnothing, \quad \pi_2(Q)\cap\mathbb{R}\setminus 3K \neq \varnothing.$$

Let us write $\theta_Q = \arctan\sigma$ if $Q \in \mathcal{D}^2_\sigma$ for some $\sigma \in S$. A computation then reveals that

$$|Q\cap(I\times\{\alpha\})| = \min(|J_Q|, \operatorname{dist}(\alpha, \mathbb{R}\setminus\pi_2(Q)))\cot\theta_Q.$$

We also observe that $\pi_2(Q)\cap(3K\setminus K)$ contains an interval $A = A(\alpha)$ of length $|K|/3$, whence for all $\alpha' \in A$ we have

$$\operatorname{dist}(\alpha,\mathbb{R}\setminus\pi_2(Q)) \le \operatorname{dist}(\alpha,\alpha') + \operatorname{dist}(\alpha',\mathbb{R}\setminus\pi_2(Q)) \lesssim |K| + \operatorname{dist}(\alpha',\mathbb{R}\setminus\pi_2(Q)) \lesssim \operatorname{dist}(\alpha',\mathbb{R}\setminus\pi_2(Q));$$

see Figure 3. This clearly implies that for every $\alpha \in K$ we have

$$|Q\cap(I\times\{\alpha\})| \lesssim \fint_A |Q\cap(I\times\{\alpha'\})|\,\mathrm{d}\alpha' \lesssim \fint_{3K} |Q\cap(I\times\{\alpha'\})|\,\mathrm{d}\alpha',$$

which proves the claim.

Smallness of the local average. We now use the previously obtained (2.24) to prove (ii). Let $\mathcal{R}_0^\star$ denote the family of parallelograms $R = I_R \times L_R \in \mathcal{R}_0'$ such that $B^{\mathrm{out}}_{I_R,L_R} > \lambda$. For each such R let K_R be the maximal interval $K \in \{L_R, 3L_R, \dots, 3^k L_R, \dots\}$ such that $B^{\mathrm{out}}_{I_R,K} > \lambda$; the existence of the maximal interval K_R is guaranteed for example by the a priori estimate of Lemma 2.14 and the assumption $R \in \mathcal{R}_0^\star$. Obviously $K_R \supseteq L_R$ and $B^{\mathrm{out}}_{I_R,3K_R} \leq \lambda$.

We show that for $R \in \mathcal{R}_0^\star$ we have

$$\fint_R \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \leq \kappa\lambda \tag{2.25}$$

for some numerical constant $\kappa \geq 1$. Indeed it is a consequence of (2.24) that

$$\begin{aligned}\fint_{I_R\times\{\alpha\}} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} &\lesssim \fint_{I_R\times 3K_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \\ &\leq \fint_{I_R\times 3K_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} + \fint_{I_R\times 3K_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}\setminus \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|}.\end{aligned}$$

The first summand is estimated using the maximality of K_R:

$$\fint_{I_R\times 3K_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} = B^{\mathrm{out}}_{I_R,3K_R} \leq \lambda.$$

The second summand can be further analyzed by observing that the cubes Q appearing in the sum above satisfy $\pi_1(Q) \subset I$ and $\pi_2(Q) \subset 9K_R$ since $Q \notin \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}$, that is, $\mathcal{L}^{\mathrm{out}}_{I_R,3K_R} \setminus \mathcal{L}^{\mathrm{out}}_{I_R,K_R}$ is subordinate to the singleton collection $\{I_R \times 9K_R\}$. Applying the Carleson sequence property

$$\fint_{I_R\times 3K_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}\setminus \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \leq \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}\setminus \mathcal{L}^{\mathrm{out}}_{I_R,3K_R}} a_Q \frac{|Q\cap(I_R\times 3K_R)|}{|Q||I_R\times 3K_R|} \lesssim 1 \leq \lambda \tag{2.26}$$

by our assumption on λ. Combining the estimates above shows that

$$\fint_{I_R\times\{\alpha\}} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \lesssim \lambda$$

for all $\alpha \in K_R$. Since $\pi_2(R) \subset K$ this implies (2.25).

Observe that if $R = I_R \times L_R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star$ then

$$B^{\mathrm{out}}_{I_R,L_R} = \fint_{I_R\times L_R} \sum_{Q \in \mathcal{L}^{\mathrm{out}}_{I_R,L_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \leq \lambda.$$

Defining the subcollection $\mathcal{L}_1$. We set

$$\mathcal{L}_1{}' := \bigcup_{R\in\mathcal{R}_0^\star} \mathcal{L}^{\mathrm{in}}_{I_R,K_R}, \quad \mathcal{L}_1{}'' := \bigcup_{R\in\mathcal{R}_0{}'\setminus\mathcal{R}_0^\star} \mathcal{L}^{\mathrm{in}}_{I_R,L_R}, \quad \mathcal{L}_1 := \mathcal{L}_1{}' \cup \mathcal{L}_1{}''.$$

Now note that for each $R \in \mathcal{R}_0^\star$ and $K = K_R \in \mathcal{K}_{\pi_1(R)}$ we have that

$$B_R^{\mathcal{L}} \le \fint_R \sum_{Q \in \mathcal{L}^{\rm out}_{I_R \times K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} + \fint_R \sum_{Q \in \mathcal{L}^{\rm in}_{I_R \times K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \le \kappa\lambda + B_R^{\mathcal{L}_1},$$

while for $R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star$ the same estimate holds using L_R in place of K_R. It remains to show the desired estimate for $\mathsf{mass}_{a,1}(\mathcal{L}_1)$ in (i) of the lemma.

Smallness of $\mathsf{mass}_{a,1}(\mathcal{L}_1)$. By the definition of the collections $\mathcal{L}^{\rm in}_{I,K}$ we have that

$$\mathsf{sh}(\mathcal{L}_1) \subset \bigcup_{R \in \mathcal{R}_0^\star} I_R \times 3K_R \cup \bigcup_{R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star} I_R \times 3L_R.$$

If $K = K_R$ for some $R \in \mathcal{R}_0^\star$ we have by definition that $B^{\rm out}_{I_R,K_R} > \lambda$. On the other hand for $R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star$ we have that $B_R^{\mathcal{L}} = B^{\mathcal{L}}_{I_R,L_R} > \lambda$.

Define

$$E := \left\{ (x,y) \in \mathbb{R}^2 : \mathrm{M}_v \Bigg[\sum_{Q \in \mathcal{L}} a_Q \frac{\mathbf{1}_Q}{|Q|} \Bigg](x,y) \ge \frac{\lambda}{2} \right\},$$

where $\mathrm{M}_v = \mathrm{M}_{(1,s)} = \mathrm{M}_1$ is the directional Hardy–Littlewood maximal operator acting in the direction $v = (1,s) = (1,0)$, see (2.3), since we have assumed $s = 0$. We will show that

$$\bigcup_{R \in \mathcal{R}_0^\star} I_R \times 3K_R \subset \{(x,y) \in \mathbb{R}^2 : \mathrm{M}_2(\mathbf{1}_E)(x,y) \ge C\}$$

for a sufficiently small constant $C > 0$, where M_2 is as in (2.3). To this end let us define

$$\psi(\alpha) := \frac{1}{|I_R|} \int_{I_R \times \{\alpha\}} \sum_{Q \in \mathcal{L}^{\rm out}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|}.$$

Note that

$$\lambda < B^{\rm out}_{I_R,K_R} = \fint_{K_R} \psi(\alpha)\,\mathrm{d}\alpha \le \frac{1}{|K_R|} \int_{\{K_R : \psi(\alpha) > \lambda/2\}} \psi(\alpha)\,\mathrm{d}\alpha + \frac{\lambda}{2} \le \frac{c\lambda}{|K_R|} \left| \left\{ K_R : \psi(\alpha) > \frac{\lambda}{2} \right\} \right| + \frac{\lambda}{2},$$

which readily yields the existence of $K' \subset K_R$, with

$$|K_R| \lesssim |K'|, \qquad \inf_{x \in I_R} \inf_{y \in K'} \mathrm{M}_v \Bigg[\sum_{Q \in \mathcal{L}^{\rm out}_{I_R,K_R}} a_Q \frac{\mathbf{1}_Q}{|Q|} \Bigg](x,y) > \frac{\lambda}{2}.$$

This in turn implies that $\mathrm{M}_2(\mathbf{1}_E) \gtrsim 1$ on $I_R \times 3K_R$. Now we can conclude

$$\Bigg| \bigcup_{R \in \mathcal{R}_0^\star} I_R \times 3K_R \Bigg| \le |\{\mathrm{M}_2(\mathbf{1}_E) \gtrsim 1\}| \lesssim |E| \lesssim \frac{1}{\lambda} \mathsf{mass}_{a,1}(\mathcal{L})$$

by the weak-$(1,1)$ inequality of the directional Hardy–Littlewood maximal operator $\mathrm{M}_{(1,0)}$.

On the other hand we have for the rectangles $R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star$ that

$$\bigcup_{R \in \mathcal{R}_0' \setminus \mathcal{R}_0^\star} I_R \times 3L_R \subset \left\{ \mathrm{M}_{\mathcal{R}_0} \Bigg(\sum_{Q \in \mathcal{L}} a_Q \frac{\mathbf{1}_Q}{|Q|} \Bigg) > \frac{\lambda}{3} \right\}.$$

Thus we get by the weak (p, p) assumption for $\mathrm{M}_{\mathcal{R}_0}$ that

$$\Big|\bigcup_{R\in\mathcal{R}_0'\setminus\mathcal{R}_0^\star} I_R\times 3L_R\Big| \le \Big|\Big\{\mathrm{M}_{\mathcal{R}_0}\Big(\sum_{Q\in\mathcal{L}} a_Q\frac{\mathbf{1}_Q}{|Q|} > \frac{\lambda}{3}\Big)\Big\}\Big| \\ \lesssim \frac{A_p^p}{\lambda^p}\,\mathsf{mass}_{a,p}(\mathcal{L}) \lesssim \frac{A_p^p}{\lambda^p}\,\mathsf{mass}_{a,1}(\mathcal{L})\mathsf{U}_2(\mathcal{L})^{2(p-1)}.$$

By the subordination property of $\mathcal{L}_1$ we get

$$\mathsf{mass}_{a,1}(\mathcal{L}_1) \le \Big|\bigcup_{R\in\mathcal{R}_0^\star} I_R\times 3K_R \cup \bigcup_{R\in\mathcal{R}_0'\setminus\mathcal{R}_0^\star} I_R\times 3L_R\Big| \le \frac{1}{2}\,\mathsf{mass}_{a,1}(\mathcal{L}),$$

upon choosing $\lambda \ge C\max(1, A_p\mathsf{U}_2(\mathcal{L})^{2/p'})$ with sufficiently large $C > 1$. □

Proof of Lemma 2.22. Fix $s\in S$ and choose λ in the definition of $\mathcal{R}_{s,k}$ to be the value given by Lemma 2.21 with $\mathcal{L} = \mathcal{R} = \bigcup_{s\in S}\mathcal{R}_s$. Let $j = 0$ and $\mathcal{L}_0 = \mathcal{L}_j := \mathcal{R}$. Construct $\mathcal{L}_1 = \mathcal{L}_{j+1}\subset\mathcal{R}$ such that $\mathsf{mass}_{a,1}(\mathcal{L}_1) \le \frac{1}{2}\,\mathsf{mass}_{a,1}(\mathcal{L}_0)$. Since $B_R^{\mathcal{L}_0} > k\lambda$ for all $R\in\mathcal{R}_{s,k}$, we have

$$\lambda k < B_R^{\mathcal{L}_0} \le \lambda + B_R^{\mathcal{L}_1} \quad\Longrightarrow\quad B_R^{\mathcal{L}_1} > \lambda(k-1).$$

Repeat the procedure recursively with $j+1$ in place of j. When $j = k-1$, we have reached the collection $\mathcal{L}_{k-1}$ with $\mathsf{mass}_{a,1}(\mathcal{L}_{k-1}) \lesssim 2^{-k}\,\mathsf{mass}_{a,1}(\mathcal{L}_0)$ and $B_R^{\mathcal{L}_{k-1}} > \lambda$. This last condition and Remark 2.18 imply that

$$\mathsf{sh}(\mathcal{R}_{s,k}) \subset \Big\{\mathrm{M}_{\mathcal{R}_s}\Big[\sum_{Q\in\mathcal{L}_{k-1}} a_Q\frac{\mathbf{1}_Q}{|Q|}\Big] > \lambda\Big\}$$

and so, using (2.10),

$$|\mathsf{sh}(\mathcal{R}_{s,k})| \le \frac{A_p^p}{\lambda^p}\,\mathsf{mass}_{a,p}(\mathcal{L}_{k-1})^p \le \frac{A_p^p}{\lambda^p}\,\mathsf{mass}_{a,1}(\mathcal{L}_{k-1})^{p-\frac{2p}{p'}}\,\mathsf{mass}_{a,2}(\mathcal{L}_{k-1})^{\frac{2p}{p'}} \\ \le 2^{-k}\,\mathsf{mass}_{a,1}(\mathcal{L}_0)\frac{CA_p^p}{\lambda^p}\Big(\frac{\mathsf{mass}_{a,2}(\mathcal{L}_0)^2}{\mathsf{mass}_{a,1}(\mathcal{L}_0)}\Big)^{p-1} = 2^{-k}\,\mathsf{mass}_{a,1}(\mathcal{L}_0)\frac{CA_p^p}{\lambda^p}\mathsf{U}_2(\mathcal{L}_0)^{2(p-1)}$$

and the lemma follows by the definition of λ since $\mathcal{L}_0 = \mathcal{R}$. □

3. A weighted Carleson embedding and applications to directional maximal operators

In this section, we provide a weighted version of the directional Carleson embedding theorem. We then derive, as applications, novel weighted norm inequalities for maximal and singular directional operators.

The proof of the weighted Carleson embedding follows the strategy used for Theorem C, with suitable modifications. In order to simplify the presentation, we restrict our scope to collections of parallelograms $\mathcal{R} = \{\bigcup\mathcal{R}_s : s\in S\}$ with the property that the maximal operator $\mathrm{M}_{\mathcal{R}_s}$ associated to each collection $\mathcal{R}_s$ satisfies the appropriate weighted weak-$(1,1)$ inequality. This is the case, for instance, when the collections $\mathcal{R}_s$ are of the form

$$\mathcal{R}_s \subset \mathcal{D}^2_{s,k}, \quad \mathcal{D}^2_{s,k} := \bigcup_{k_1\le k}\mathcal{D}^2_{s,k_1,k} \tag{3.1}$$

for a fixed $k \in \mathbb{Z}$. In other words, the parallelograms in direction s have fixed vertical sidelength and arbitrary eccentricity.

3.2. ***Directional weights.*** Let S be a set of slopes and $w, u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ be nonnegative functions, which we refer to as *weights* from now on. Our weight classes are related to the maximal operator

$$\mathrm{M}_{S;2} := \mathrm{M}_V \circ \mathrm{M}_{(0,1)},$$

recalling that $M_V = M_{\{(1,s):s\in S\}}$ is the directional maximal operator defined in (2.4). We introduce the two-weight directional constant

$$[w,u]_S := \sup_{x\in\mathbb{R}^2} \frac{\mathrm{M}_{S;2}w(x)}{u(x)}.$$

We pause to point out some relevant examples of pairs w, u with $[w,u]_S < \infty$. Recall that, for $p > 2$, $\|\mathrm{M}_{S;2}\|_{p\to p} \lesssim (\log \#S)^{1/p}$; this is actually a special case of Theorem C and interpolation. Therefore, if $g \geq 0$ belongs to the unit sphere of $L^p(\mathbb{R}^2)$,

$$w := \sum_{\ell=0}^{\infty} \frac{\mathrm{M}^{[\ell]}_{S;2} g}{2^\ell \|\mathrm{M}_{S;2}\|^\ell_{p\to p}}$$

satisfies $[w,w]_S \leq 2\|\mathrm{M}_{S;2}\|_{p\to p}$; here $T^{[\ell]}$ denotes ℓ-fold composition of an operator T with itself. We also highlight the relevance of $[w,u]_S$ in Theorem D below by noticing that

$$\sup_{s\in S} \|\mathrm{M}_{\mathcal{D}^2_{s,k}} : L^1(u) \to L^{1,\infty}(w)\| \lesssim [w,u]_S,$$

with absolute implicit constant. This result is obtained via the classical Fefferman–Stein inequality in direction s paired with the remark that $\mathrm{M}_{\mathcal{D}^2_{s,k}} w \lesssim \mathrm{M}_{S;2} w \leq [w,u]_S u$.

3.3. ***Weighted Carleson sequences.*** We begin with the weighted analogue of Definition 2.7, which is given with respect to a fixed weight w.

Definition 3.4. Let $a = \{a_R\}_{R\in\mathcal{D}^2_S}$ be a sequence of nonnegative numbers. Then a will be called an L^∞*-normalized* w-Carleson sequence if for every $\mathcal{L} \subset \mathcal{D}^2_S$ which is subordinate to some collection $\mathcal{T} \subset \mathcal{P}^2_\tau$ for some fixed $\tau \in S$, we have

$$\sum_{L\in\mathcal{L}} a_L \leq w(\mathsf{sh}(\mathcal{T})), \qquad \mathsf{mass}_a := \sum_{R\in\mathcal{D}^2_S} a_R < \infty.$$

As before, if $\mathcal{R} \subset \mathcal{D}^2_\tau$ for some fixed $\tau \in S$ then $\mathcal{R}$ is subordinate to itself and

$$\mathsf{mass}_{a,1}(\mathcal{R}) = \sum_{R\in\mathcal{R}} a_R \leq w(\mathsf{sh}(\mathcal{R})), \quad \mathcal{R} \subset \mathcal{D}^2_\tau \text{ for some fixed } \tau \in S.$$

Throughout this section all Carleson sequences and related quantities are taken with respect to some fixed weight w which is suppressed from the notation. We can now state our weighted Carleson embedding theorem.

Theorem D. *Let $S \subset [-1,1]$ be a finite set of N slopes and $\mathcal{R} \subset \mathcal{D}^2_S$. Let w, u be weights with $[w,u]_S < \infty$ and such that*

$$\sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_s} : L^1(u) \to L^{1,\infty}(w)\| \lesssim [w,u]_S.$$

Then for every L^∞-normalized w-Carleson sequence $a = \{a_R\}_{R \in \mathcal{D}^2_S}$ we have

$$\left(\int |T_{\mathcal{R}}(a)(x)|^2 \frac{\mathrm{d}x}{\mathrm{M}_{\mathcal{R}} u(x)} \right)^{\frac{1}{2}} \lesssim (\log N)^{\frac{1}{2}} [w,u]_S \,\mathsf{mass}_{a,1}(\mathcal{R})^{\frac{1}{2}}.$$

3.5. ***Proof of Theorem D.*** We follow the proof of Theorem C and only highlight the differences to accommodate the weighted setting. Write $\sigma := [\mathrm{M}_{\mathcal{R}} u]^{-1}$. Expanding the $L^2(\sigma)$-norm we have

$$\|T_{\mathcal{R}}(a)\|^2_{L^2(\sigma)} \le 2 \sum_{R \in \mathcal{R}} a_R \sum_{\substack{Q \in \mathcal{R} \\ Q \le R}} a_Q \frac{\sigma(Q \cap R)}{|Q||R|}.$$

From the definition of σ we have that

$$\sigma(Q \cap R) \le \frac{|Q \cap R|}{\inf_Q \mathrm{M}_{\mathcal{R}} u} \le \frac{|Q|}{u(Q)} |Q \cap R|,$$

whence

$$\|T_{\mathcal{R}}(a)\|^2_{L^2(\sigma)} \le 2 \sum_{R \in \mathcal{R}} a_R \fint_R \sum_{\substack{Q \in \mathcal{R} \\ Q \le R}} a_Q \frac{\mathbf{1}_Q}{u(Q)} := 2 \sum_{R \in \mathcal{R}} a_R B^{\mathcal{R}}_R,$$

where now for any $\mathcal{L} \subset \mathcal{R}$ we have defined

$$B^{\mathcal{L}}_R := \fint_R \sum_{\substack{Q \in \mathcal{L} \\ Q \le R}} a_Q \frac{\mathbf{1}_Q}{u(Q)}.$$

Defining the families $\mathcal{R}_{s,k}$ for $s \in S$ and $k \in \mathbb{N}$ as in (2.19) we then have the estimate

$$\|T_{\mathcal{R}}(a)\|^2_{L^2(\sigma)} \le 2\lambda \left[(\log N) \,\mathsf{mass}_{a,1}(\mathcal{R}) + N \sum_{k > \log N} k \sup_{s \in S} w(\mathsf{sh}(\mathcal{R}_{s,k})) \right].$$

Again $\lambda > 0$ is a constant that will be determined later in the proof and in the last line we used the w-Carleson assumption for the sequence $a = \{a_R\}$ for rectangles in a fixed direction.

We need the weighted version of Lemma 2.21, which is given under the standing assumptions of Theorem D.

Lemma 3.6. *Let $a = \{a_R : R \in \mathcal{D}^2_S\}$ be an L^∞-normalized w-Carleson sequence, $s \in S \subset [-1,1]$, and $\mathcal{L}, \mathcal{R} \subset \mathcal{D}^2_S$ with $\mathcal{L} \subseteq \mathcal{R}$. For every $\lambda > C[w,u]_S$, where C is a suitably chosen absolute constant, there exists $\mathcal{L}_1 \subset \mathcal{L}$ such that*

(i) $\mathsf{mass}_{a,1}(\mathcal{L}_1) \le \frac{1}{2} \mathsf{mass}_{a,1}(\mathcal{L})$,

(ii) *denoting by $\mathcal{R}'_s$ the collection of rectangles R in $\mathcal{R}_s$ with $B^{\mathcal{L}}_R > \lambda$ we have that*

$$B^{\mathcal{L}}_R \le \lambda + B^{\mathcal{L}_1}_R \quad \textit{for all } R \in \mathcal{R}'_s.$$

Proof. We can assume that $s = 0$ and let $\mathcal{R}_0'$ be the collection of rectangles in $\mathcal{R}_0$ such that $B_R^{\mathcal{L}} > \lambda$, where λ is as in the statement of the lemma and C will be specified at the end of the proof. For $I \in \{\pi_1(R) : R \in \mathcal{R}_0'\}$ and any interval $K \subset \mathbb{R}$ we define $\mathcal{L}_{I,K}^{\mathrm{in}}$ and $\mathcal{L}_{I,K}^{\mathrm{out}}$ as in the proof of Theorem C, but now we set

$$B_{I,K}^{\mathrm{in}} := \fint_{I\times K} \sum_{Q\in\mathcal{L}_{I,K}^{\mathrm{in}}} \frac{a_Q}{u(Q)}\mathbf{1}_Q, \quad B_{I,K}^{\mathrm{out}} := \fint_{I\times K} \sum_{Q\in\mathcal{L}_{I,K}^{\mathrm{out}}} \frac{a_Q}{u(Q)}\mathbf{1}_Q.$$

We define $\mathcal{R}_0''$ to be the subcollection of those $R = I \times L \in \mathcal{R}_0'$ such that $B_{I,L}^{\mathrm{out}} \le \lambda$. By linearity we get for each $R \in \mathcal{R}_0''$ that $B_R^{\mathcal{L}} \le \lambda + B_{I,L}^{\mathrm{in}} \le \lambda + B_R^{\mathcal{L}_1''}$, where

$$\mathcal{L}_1{}'' := \bigcup_{R=I\times L\in\mathcal{R}_0''} \mathcal{L}_{I,L}^{\mathrm{in}}, \quad \mathsf{sh}(\mathcal{L}_1'') \subset \bigcup_{R=I\times L\in\mathcal{R}_0''} I\times 3L.$$

Since $\mathcal{R}_0'' \subset \mathcal{R}_0'$ we conclude as before that

$$\begin{aligned} w(\mathsf{sh}(\mathcal{L}_1'')) &\le w\Big(\bigcup_{R=I\times L\in\mathcal{R}_0''} I\times 3L\Big) \le w\Big(\Big\{\mathrm{M}_{\mathcal{R}_0}\Big(\sum_{Q\in\mathcal{L}} \frac{a_Q\mathbf{1}_Q}{u(Q)}\Big) > \frac{\lambda}{3}\Big\}\Big) \\ &\lesssim \frac{[w,u]_S}{\lambda}\int_{\mathbb{R}^2} \sum_{Q\in\mathcal{R}} a_Q \frac{\mathbf{1}_Q}{u(Q)}\,\mathrm{d}u = \frac{[w,u]_S}{\lambda}\,\mathsf{mass}_{a,1}(\mathcal{L}) \end{aligned}$$

by the two-weight weak-type-$(1,1)$ inequality for $\mathrm{M}_{\mathcal{R}_s} = \mathrm{M}_{\mathcal{R}_0}$. Now $\mathcal{L}_1''$ is subordinate to the collection $\{I\times 3L : I\times L \in \mathcal{R}_0''\}$. Using the definition of a Carleson sequence we have

$$\sum_{Q\in\mathcal{L}_1''} a_Q \le w\Big(\bigcup_{R=I\times L\in\mathcal{R}_0''} I\times 3L\Big) \lesssim \frac{[w,u]_S}{\lambda}\,\mathsf{mass}_{a,1}(\mathcal{L}),$$

and so $\mathsf{mass}_{a,1}(\mathcal{L}_1'') \lesssim [w,u]_S\,\mathsf{mass}_{a,1}(\mathcal{L})/\lambda$.

It remains to deal with parallelograms

$$R = I\times L \in \mathcal{R}_0^\star := \mathcal{R}_0' \setminus \mathcal{R}_0'', \quad B_{I,L}^{\mathrm{out}} > \lambda.$$

We define the maximal K_R such that $B_{I,K_R}^{\mathrm{out}} > \lambda$ as before; the existence of this maximal interval can be guaranteed for example by assuming the collection $\mathcal{R}$ is finite. We have for each $R = I\times L \in \mathcal{R}_0^\star$ that $B_{I,L}^{\mathrm{out}} > \lambda$ so $K_R \supset L$ and $B_{I,3K_R}^{\mathrm{out}} \le \lambda$ by maximality.

Now using (2.24) we get that

$$\Delta := \sum_{Q\in\mathcal{L}_{I,3K_R}^{\mathrm{out}}} a_Q \frac{|Q\cap(I\times\{\alpha\})|}{u(Q)|I|} \lesssim \sum_{Q\in\mathcal{L}_{I,3K_R}^{\mathrm{out}}} a_Q \frac{|Q\cap(I\times 3K_R)|}{u(Q)|3K_R||I|} = \fint_{I\times 3K_R} \sum_{Q\in\mathcal{L}_{I,3K_R}^{\mathrm{out}}} a_Q \frac{\mathbf{1}_Q}{u(Q)} \lesssim \lambda$$

by the maximality of K_R. On the other hand

$$\Xi := \sum_{Q\in\mathcal{L}_{I,K}^{\mathrm{out}}\setminus\mathcal{L}_{I,3K}^{\mathrm{out}}} a_Q \fint_{I\times\{\alpha\}} \frac{\mathbf{1}_Q}{u(Q)} \lesssim \sum_{Q\subset I\times 9K} a_Q \frac{|Q\cap(I\times 3K)|}{|I\times 3K|\,u(Q)}.$$

Since $\mathrm{M}_{\mathcal{R}_s} w \leq \mathrm{M}_V \mathrm{M}_2 w \leq [w,u]_S \, u$ uniformly in s we get that for $Q \subset I \times 9K$

$$u(Q) \gtrsim [w,u]_S^{-1} \frac{w(I \times 9K)}{|I \times 9K|} |Q|$$

and by this and the w-Carleson property for all Q subordinate to $I \times 9K$ we get

$$\Xi \lesssim [w,u]_S \leq \lambda$$

provided $\lambda \geq [w,u]_S$. We now define

$$\mathcal{L}_1{}' := \bigcup_{R \in \mathcal{R}_0^\star} L^{\mathrm{in}}_{\pi_1(R), K_R}$$

so that

$$\mathsf{sh}(\mathcal{L}_1') \subset \bigcup_{R \in \mathcal{R}_0^\star} \pi_1(R) \times K_R.$$

Arguing as in the unweighted case of Theorem C we can estimate

$$w(\mathsf{sh}(\mathcal{L}_1')) \leq w\Bigg(\bigcup_{R \in \mathcal{R}_0^\star} \pi_1(R) \times K_R \Bigg) \lesssim w(\{\mathrm{M}_2(\mathbf{1}_E) \gtrsim 1\}),$$

where

$$E := \left\{ (x,y) \in \mathbb{R}^2 : \mathrm{M}_v \Bigg[\sum_{Q \in \mathcal{L}} a_Q \frac{\mathbf{1}_Q}{u(Q)} \Bigg](x,y) \geq \frac{\lambda}{2} \right\}.$$

In the definition of E above we have that $\mathrm{M}_v = \mathrm{M}_{(1,s)} = \mathrm{M}_1$ since we have reduced to the case $v = (1,s) = (1,0)$. Using the subordination property of $\mathcal{L}_1'$ and the Fefferman–Stein inequality once in the direction e_2 for M_2 and once in the direction $v = (1,s) = (1,0)$ for M_v we estimate

$$\mathsf{mass}_{a,1}(\mathcal{L}_1') \leq w\Bigg(\bigcup_{R \in \mathcal{R}_0^\star} \pi_1(R) \times K_R \Bigg) \lesssim \frac{1}{\lambda} \sum_{Q \in \mathcal{L}} a_Q \frac{\mathrm{M}_V \mathrm{M}_2 w(Q)}{u(Q)} \leq \frac{[w,u]_S}{\lambda} \mathsf{mass}_{a,1}(\mathcal{L}).$$

We have thus proved the lemma upon setting $\mathcal{L}_1 := \mathcal{L}_1'' \cup \mathcal{L}_1'$ and choosing $\lambda \geq C[w,u]_S$ for a sufficiently large numerical constant $C > 1$. □

Repeating the steps in the proof of Lemma 2.22 for λ as in the statement of Lemma 3.6 we get for the sets $\mathcal{R}_{s,k}$ defined with respect to this λ that

$$w(\mathsf{sh}(\mathcal{R}_{s,k})) \lesssim 2^{-k} \, \mathsf{mass}_{a,1}(\mathcal{R}),$$

and this completes the proof of Theorem D.

3.7. ***Applications of Theorem D.*** The first corollary of Theorem D is a two-weighted estimate for the directional maximal operator M_V from (2.4).

Theorem E. *Let $V \subset \mathbb{S}^1$ be a finite set of N slopes and w be a weight on $\mathbb{R}^2$. Then*

$$\|\mathrm{M}_V : L^2(\widetilde{\mathrm{M}}_V w) \to L^{2,\infty}(w)\| \lesssim \sqrt{\log N}, \quad \widetilde{\mathrm{M}}_V := \mathrm{M}_V \circ \mathrm{M}_V \circ \max\{\mathrm{M}_{(1,0)}, \mathrm{M}_{(0,1)}\}.$$

Remark 3.8. In the proof below, we argue for almost horizontal V, and in place of $\max\{\mathrm{M}_{(1,0)}, \mathrm{M}_{(0,1)}\}$ we use $\mathrm{M}_{(0,1)}$. The usage of $\max\{\mathrm{M}_{(1,0)}, \mathrm{M}_{(0,1)}\}$ enables the statement of the theorem to be invariant under rotation of V.

Proof of Theorem E. By standard limiting arguments, it suffices to prove that for each $k \in \mathbb{Z}$ the estimate

$$\|\mathrm{M}_{\mathcal{R}} : L^2(z) \to L^{2,\infty}(w)\| \lesssim \sqrt{\log N}, \quad z := \mathrm{M}_{\mathcal{R}} \circ \mathrm{M}_V \circ \mathrm{M}_{(0,1)} w, \tag{3.9}$$

when $\mathcal{R}$ is a one-parameter collection as in (3.1), holds uniformly in k.

For a nonnegative function $f \in \mathcal{S}(\mathbb{R}^2)$ let Uf be a linearization of $\mathrm{M}_{\mathcal{R}} f$, namely

$$\mathrm{M}_{\mathcal{R}} f(x) = Uf(x) = \frac{1}{|R(x)|} \int_{R(x)} f(y)\,\mathrm{d}y = \sum_{R \in \mathcal{R}} \langle f \rangle_R \mathbf{1}_{F_R}(x), \quad F_R := \{x \in R : R(x) = R\}.$$

By duality, (3.9) turns into

$$\|U^*(w\mathbf{1}_E)\|_{L^2(z^{-1})} \lesssim \sqrt{\log N}\sqrt{w(E)} \quad \text{for all } E \subset \mathbb{R}^2. \tag{3.10}$$

We can easily calculate

$$U^*(w\mathbf{1}_E) = \sum_{R \in \mathcal{R}} w(E \cap F_R) \frac{\mathbf{1}_R}{|R|}$$

and it is routine to check that $\{w(E \cap F_R)\}_{R \in \mathcal{R}}$ is a w-Carleson sequence according to Definition 3.4. The main point here is that the sets $\{E \cap F_R\}_{R \in \mathcal{R}}$ are by definition pairwise disjoint and $F_R \subseteq R$ for each $R \in \mathcal{R}$.

Setting $u := \mathrm{M}_V \circ \mathrm{M}_{(0,1)} w$, if S are the slopes of V, it is clear that $[w,u]_S \lesssim 1$ and that $z^{-1} = (\mathrm{M}_{\mathcal{R}} u)^{-1}$. Therefore (3.10) follows from an application of Theorem D. $\square$

We may in turn use Theorem E to establish a weighted norm inequality for maximal directional singular integrals with controlled dependence on the cardinality $\#V = N$. Similar considerations may be used to yield weighted bounds for directional singular integrals in $L^p(\mathbb{R}^2)$ for $p > 2$; we do not pursue this issue.

Theorem F. *Let K be a standard Calderón–Zygmund convolution kernel on $\mathbb{R}$ and $V \subset \mathbb{S}^1$ be a finite set of N slopes. For $v \in V$ we define*

$$T_v f(x) = \sup_{\varepsilon > 0} \left| \int_{\varepsilon < t < \frac{1}{\varepsilon}} f(x + tv) K(t)\,\mathrm{d}t \right|, \quad T_V f(x) = \sup_{v \in V} |T_v f(x)|.$$

Let w be a weight on $\mathbb{R}^2$ with $[w]_{A_1^V} := \|\mathrm{M}_V w / w\|_\infty < \infty$. Then

$$\|T_V : L^2(w) \to L^{2,\infty}(w)\| \lesssim (\log N)^{\frac{3}{2}} [w]_{A_1^V}^{\frac{5}{2}}.$$

We sketch the proof, which is a weighted modification of the arguments for [Demeter and Di Plinio 2014, Theorem 1]. Hunt's classical exponential good-λ inequality, see [Demeter and Di Plinio 2014, Proposition 2.2] for a proof, may be upgraded to

$$w(\{x \in \mathbb{R}^2 : T_v f(x) > 2\lambda, \mathrm{M}_v f(x) \le \gamma\lambda\}) \lesssim \exp\left(-\frac{c}{\gamma [w]_{A_1^V}}\right) w(\{x \in \mathbb{R}^2 : T_v f(x) > \lambda\}) \tag{3.11}$$

by using that $[w]_{A_1^V}$ dominates the A_∞ constant of the one-dimensional weight $t \mapsto w(x+tv)$ for all $x \in \mathbb{R}^2$, $v \in V$, together with Fubini's theorem. With (3.11) in hand, Theorem F follows from Theorem E via standard good-λ inequalities, selecting $(\gamma)^{-1} \sim [w]_{A_1^V} \log N$. Note that the right-hand side of the estimate in the conclusion of Theorem E becomes $[w]_{A_1}^{3/2} \sqrt{\log N}$ when the estimate is specified to A_1^V weights as the ones we consider here.

4. Tiles, adapted families, and intrinsic square functions

We define here some general notions of tiles and adapted families of wave-packets: definitions in this spirit have appeared in, among others [Barrionuevo and Lacey 2003; Demeter and Di Plinio 2014; Lacey and Li 2006; 2010; Lacey 2007]. These will be essential for the time-frequency analysis square functions we use in this paper in order to model the main operators of interest. After presenting these abstract definitions we show some general orthogonality estimates for wave packet coefficients. We then detail how these notions are specialized in three particular cases of interest.

4.1. ***Tiles and wavelet coefficients.*** Throughout this section we fix a finite set of slopes $S \subset [-1, 1]$. Remember that alternatively we will refer to the set of vectors $V := \{(1, s) : s \in S\}$. A *tile* is a set $t := R_t \times \Omega_t \subset \mathbb{R}^2 \times \mathbb{R}^2$, where $R_t \in \mathcal{D}_S^2$ and $\Omega_t \subset \mathbb{R}^2$ is a measurable set, and $|R_t||\Omega_t| \gtrsim 1$. We denote by $s(t) \in S$ the slope such that $R_t \in \mathcal{D}_{s(t)}^2$, and then

$$R_t = A_{s(t)}(I_t \times J_t), \quad \text{with } I_t \times J_t \in \mathcal{D}_0^2.$$

We also use the notation $v_t := (1, s(t))$. There are several different collections of tiles used in this paper, they will generically be denoted by $\boldsymbol{T}, \boldsymbol{T}_1, \boldsymbol{T}'$ or similar. Given any collection of tiles $\boldsymbol{T}$ we will often use the notation $\mathcal{R}_{\boldsymbol{T}} := \{R_t : t \in \boldsymbol{T}\}$ to denote the collection of spatial components of the tiles in $\boldsymbol{T}$. The exact geometry of these tiles will be clear from context; however, several estimates hold for generic collections of tiles, as we will see in Section 4.3.

Let $t = R_t \times \Omega_t$ be a tile and $M \geq 2$. We denote by $\mathcal{A}_t^M$ the collection of Schwartz functions ϕ on $\mathbb{R}^2$ such that

(i) $\operatorname{supp}(\hat{\phi}) \subset \Omega_t$,

(ii) there holds

$$\sup_{0 \leq \alpha, \beta \leq M} \sup_{x \in \mathbb{R}^2} |R_t|^{\frac{1}{2}} |I|^\alpha |J|^\beta \left(1 + \frac{|x \cdot v_t|}{|I||v_t|}\right)^M \left(1 + \frac{|x \cdot e_2|}{|J|}\right)^M |\partial_{v_t}^\alpha \partial_{e_2}^\beta \phi(x + c_{R_t})| \leq 1.$$

In the above display c_{R_t} refers to the center of R_t and

$$\partial_{v_t}(\cdot) := \frac{v_t}{|v_t|} \cdot \nabla(\cdot).$$

An immediate consequence of property (ii) is the normalization

$$\sup_{\phi \in \mathcal{A}_t^M} \|\phi\|_2 \lesssim 1.$$

We thus refer to $\mathcal{A}_t^M$ as the collection of *L^2-normalized wave packets adapted to t of order M*. For our purposes, it will suffice to work with moderate values of M, say $2^3 \leq M \leq 2^{50}$. In fact, we use

$M = M_0 = 2^{50}$ in the definition of the *intrinsic wavelet coefficient* associated with the tile t and the Schwartz function f:

$$a_t(f) := \sup_{\phi \in \mathcal{A}_t^{M_0}} |\langle f, \phi\rangle|^2, \quad M_0 = 2^{50}. \tag{4.2}$$

This section is dedicated to square functions involving wavelet coefficients associated with particular collections of tiles which formally look like

$$\Delta_{\boldsymbol{T}}(f)^2 := \sum_{t \in \boldsymbol{T}} a_t(f) \frac{\mathbf{1}_{R_t}}{|R_t|}, \quad \boldsymbol{T} \text{ is a collection of tiles.}$$

We begin by proving some general global and local orthogonality estimates for collections of tiles with finitely overlapping frequency components. These estimates will be essential in showing that the sequence $\{a_t(f)\}_{t \in \boldsymbol{T}}$ is Carleson in the sense of Section 2, when $|f| \leq \mathbf{1}_E$ for some measurable set $E \subset \mathbb{R}^2$ with $0 < |E| < \infty$. This in turn will allow us to use the directional Carleson embedding of Theorem C in order to conclude corresponding estimates for intrinsic square functions defined on collections of tiles.

4.3. ***Orthogonality estimates for collections of tiles.*** We begin with an easy orthogonality estimate for wave packet coefficients. For completeness we present a sketch of proof which has a TT^* flavor. The argument follows the lines of proof of [Lacey 2007, Proposition 3.3].

Lemma 4.4. *Let $\boldsymbol{T}$ be a set of tiles such that $\sum_{t \in \boldsymbol{T}} \mathbf{1}_{\Omega_t} \lesssim 1$, let $M \geq 2^3$ and $\{\phi_t : t \in \boldsymbol{T}\}$ be such that $\phi_t \in \mathcal{A}_t^{\boldsymbol{M}}$ for all $t \in \boldsymbol{T}$. We have the estimate*

$$\sum_{t \in \boldsymbol{T}} |\langle f, \phi_t\rangle|^2 \lesssim \|f\|_2^2, \tag{4.5}$$

and as a consequence

$$\sum_{t \in \boldsymbol{T}} a_t(f) \lesssim \|f\|_2^2.$$

Proof. Fix $M \geq 2^3$. It suffices to prove that for $\|f\|_2 = 1$ and an arbitrary adapted family of wave packets $\{\phi_t : \phi_t \in \mathcal{A}_t^{\boldsymbol{M}}, t \in \boldsymbol{T}\}$ there holds

$$B := \sum_{t \in \boldsymbol{T}} |\langle f, \phi_t\rangle|^2 \lesssim 1. \tag{4.6}$$

Let us first fix some $\Omega \in \Omega(\boldsymbol{T}) := \{\Omega_t : t \in \boldsymbol{T}\}$ and consider the family

$$\boldsymbol{T}(\{\Omega\}) := \{t \in \boldsymbol{T} : \Omega_t = \Omega\}.$$

To prove (4.6), we introduce

$$B_\Omega(g) := \sum_{t \in \boldsymbol{T}(\{\Omega\})} |\langle g, \phi_t\rangle|^2, \quad S_\Omega(g) := (\hat{g}\mathbf{1}_\Omega)^\vee.$$

We claim that $B_\Omega(g) \lesssim \|g\|_2^2$ for all g, uniformly in $\Omega \in \Omega(\boldsymbol{T})$. Assuming the claim for a moment and remembering the finite overlap assumption on the frequency components of the tiles we have

$$B = \sum_{\Omega \in \Omega(\boldsymbol{T})} B_\Omega(S_\Omega f) \lesssim \sum_{\Omega \in \Omega(\boldsymbol{T})} \|S_\Omega(f)\|_2^2 \leq \Big\| \sum_{\Omega \in \Omega(\boldsymbol{T})} \mathbf{1}_\Omega \Big\|_\infty^2 \|f\|_2^2 \lesssim 1$$

as desired. It thus suffices to prove the claim. To this end let

$$P_\Omega(g) := \sum_{t \in \boldsymbol{T}(\{\Omega\})} \langle g, \phi_t \rangle \phi_t.$$

Then for any g with $\|g\|_2 = 1$ we have that $B_\Omega(g) = \langle P_\Omega(g), g \rangle \leq \|P_\Omega(g)\|_2$ and it suffices to prove that $\|P_\Omega(g)\|_2^2 \lesssim B_\Omega(g)$. A direct computation reveals that

$$\|P_\Omega(g)\|_2^2 \leq B_\Omega(g) \sup_{t' \in \boldsymbol{T}(\{\Omega\})} \sum_{t \in \boldsymbol{T}(\{\Omega\})} |\langle \phi_t, \phi_{t'} \rangle| \lesssim B,$$

where the second inequality in the last display above follows by the polynomial decay of the wave packets $\{\phi_t : \Omega_t = \Omega\}$. This completes the proof of the lemma. □

We present below a localized orthogonality statement which is needed in order to verify that the coefficients $a_t(f)$ form a Carleson sequence in the sense of Section 2. Verifying this Carleson condition relies on a variation of Journé's lemma that can be found in [Cabrelli et al. 2006, Lemma 3.23]; we rephrase it here adjusted to our notation. In the statement of the lemma below we denote by $\mathrm{M}_{\mathcal{P}_s^2}$ the *maximal operator* corresponding to the collection $\mathcal{P}_s^2$, where $s \in S$ is a fixed slope. Note that the proof in [Cabrelli et al. 2006] corresponds to the case of slope $s = 0$ but the general case $s \in S$ follows easily by a change of variables. Remember here that we have $S \subset [-1, 1]$.

In the statement of the lemma below two parallelograms are called *incomparable* if none of them is contained in the other.

Lemma 4.7. *Let $s \in S$ be a slope and $\mathcal{T} \subset \mathcal{D}_s^2$ be a collection of pairwise incomparable parallelograms. Define*

$$\mathsf{sh}^\star(\mathcal{T}) := \{\mathrm{M}_{\mathcal{P}_s^2} \mathbf{1}_{\mathsf{sh}(\mathcal{T})} > 2^{-6}\}$$

and for each $R \in \mathcal{T}$ let u_R be the least integer u such that $2^u R \not\subset \mathsf{sh}^\star(\mathcal{T})$. Then

$$\sum_{\substack{R \in \mathcal{T} \\ u_R = u}} |R| \lesssim 2^u |\mathsf{sh}(\mathcal{T})|.$$

With the suitable analogue of Journé's lemma in hand we are ready to state and prove the localized orthogonality condition for the coefficients $a_t(f)$.

Lemma 4.8. *Let $s \in S$ be a slope, $\mathcal{T} \subset \mathcal{P}_s^2$ be a given collection of parallelograms and $\boldsymbol{T}$ be a collection of tiles such that*

$$\mathcal{R}_{\boldsymbol{T}} := \{R_t : t \in \boldsymbol{T}\}$$

is subordinate to $\mathcal{T}$. Then we have

$$\sum_{t \in \boldsymbol{T}} a_t(f) \lesssim |\mathsf{sh}(\mathcal{T})| \|f\|_\infty^2.$$

Proof. We first make a standard reduction that allows us to pass to a collection of dyadic rectangles. To do this we use that there exist at most 9^2 shifted dyadic grids $\mathcal{D}_{s,j}^2$ such that for each parallelogram $T \in \mathcal{T}$ there exists $\widetilde{T} \in \bigcup_j \mathcal{D}_{s,j}^2$ with $T \subset \widetilde{T}$ and $|T| \leq |\widetilde{T}| \lesssim |T|$; see for example [Hytönen et al. 2013]. Now note that for each $\widetilde{T} \in \widetilde{\mathcal{T}}$ we have

$$\frac{|T \cap \widetilde{T}|}{|\widetilde{T}|} \gtrsim 1, \quad \mathsf{sh}(\widetilde{\mathcal{T}}) \subset \{\mathrm{M}_{\mathcal{P}^2_S}(\mathbf{1}_{\mathsf{sh}(\mathcal{T})}) \gtrsim 1\}$$

and so $|\mathsf{sh}(\widetilde{\mathcal{T}})| \lesssim |\mathsf{sh}(\mathcal{T})|$. Now it is clear that we can replace $\mathcal{T}$ with the dyadic collection $\widetilde{\mathcal{T}}$ in the assumption. Furthermore there is no loss in generality with assuming that $\mathcal{T}$ is a pairwise incomparable collection. We do so in the rest of the proof and continue using the notation $\mathcal{T}$ assuming it is a dyadic collection.

Since $\mathcal{R}_{\boldsymbol{T}}$ is subordinate to $\mathcal{T}$ we have the decomposition

$$\boldsymbol{T} = \bigcup_{T \in \mathcal{T}} \boldsymbol{T}(T), \quad \boldsymbol{T}(T) := \{t \in \boldsymbol{T} : R_t \subset T\}.$$

Now if f is supported on $\mathsf{sh}^\star(\mathcal{T})$ and $\phi_t \in \mathcal{A}_t^{M_0}$ for each $t \in \boldsymbol{T}$ then

$$\sum_{t \in \boldsymbol{T}} |\langle f, \phi_t \rangle|^2 \lesssim \|f\|_2^2 \leq |\mathsf{sh}^\star(\mathcal{T})| \|f\|_\infty^2 \lesssim |\mathsf{sh}(\mathcal{T})| \|f\|_\infty^2$$

by Lemma 4.4. We may thus assume that f is supported outside $\mathsf{sh}^\star(\mathcal{T})$. By Lemma 4.7 it then suffices to prove that

$$\sum_{t \in \boldsymbol{T}(T)} |\langle f, \phi_t \rangle|^2 \lesssim 2^{-10u} |T|$$

whenever u is the least integer such that $2^u T \not\subset \mathsf{sh}^\star(\mathcal{T})$ and $\|f\|_\infty = 1$. As f is supported off $\mathsf{sh}^\star(\mathcal{T})$ we have for this choice of u that

$$f = \sum_{n \geq 0} f_n, \quad f_n := f \mathbf{1}_{2^{u+n}T \setminus 2^{u+n-1}T}.$$

Let z_T be the center of T and suppose that $T = A_s(I_T \times J_T)$, with $I_T \times J_T \in \mathcal{D}_0^2$; remember that we write $v_s := (1, s)$. Let

$$\chi_T(x) := \left(1 + \frac{(x - z_T) \cdot v_s}{|I_T||v_s|}\right)^{-20} (1 + |J_T|^{-1}(x - z_T) \cdot e_2)^{-20}.$$

Observe preliminarily that

$$\|f_n \chi_T\|_\infty \lesssim 2^{-20(u+n)}$$

so that for any constant $c > 0$ we have

$$\begin{aligned}\Big(\sum_{t \in \boldsymbol{T}(T)} |\langle f, \phi_t \rangle|^2\Big)^{\frac{1}{2}} &\leq \sum_{n \geq 0} \Big(\sum_{t \in \boldsymbol{T}(T)} |\langle f_n, \phi_t \rangle|^2\Big)^{\frac{1}{2}} = \sum_{n \geq 0} \Big(\sum_{t \in \boldsymbol{T}(T)} |\langle f_n c^{-1} \chi_T, c \chi_T^{-1} \phi_t \rangle|^2\Big)^{\frac{1}{2}} \\ &\lesssim \sum_{n \geq 0} \|f_n \chi_T\|_2 \lesssim \sum_{n \geq 0} \|f_n \chi_T\|_\infty |2^{u+n} T|^{\frac{1}{2}} \lesssim 2^{-5u} |T|^{\frac{1}{2}}\end{aligned}$$

as claimed. To pass to the second line we have used estimate (4.5) of Lemma 4.4 together with the easily verifiable fact that for each $t \in \boldsymbol{T}(T)$ the wave-packet $c \chi_T^{-1} \phi_t$ is adapted to t with order $M_0 - 20 \geq 2^3$ provided the absolute constant c is chosen small enough. $\square$

4.9. ***The intrinsic square function associated with rough frequency cones.*** Let $s \in S$ be our finite set of slopes. As usual we write $v_s := (1,s)$ for $s \in S$ and $V := \{v_s : s \in S\}$ and switch between the description of directions as slopes or vectors as desired with no particular mention. Now assume we are given a finitely overlapping collection of arcs $\{\omega_s\}_{s\in S}$ with each $\omega_s \subset \mathbb{S}^1$ centered at $(v_s/|v_s|)^\perp$. We will adopt the notation

$$\omega_s := \left(\left(\frac{v_{s^-}}{|v_{s^-}|}\right)^\perp, \left(\frac{v_{s^+}}{|v_{s^+}|}\right)^\perp\right)$$

assuming that the positive direction on the circle is counterclockwise and $s^- < s < s^+$.

For $s \in S$ we define the conical sectors

$$\Omega_{s,k} := \left\{\xi \in \mathbb{R}^2 : 2^{k-1} < |\xi| < 2^{k+1},\ \frac{\xi}{|\xi|} \in \omega_s\right\}, \quad k \in \mathbb{Z}; \tag{4.10}$$

these are an overlapping cover of the cone

$$C_s := \left\{\xi \in \mathbb{R}^2 \setminus \{0\} : \frac{\xi}{|\xi|} \in \omega_s\right\},$$

with $k \in \mathbb{Z}$ playing the role of the annular parameter. Each sector $\Omega_{s,k}$ is strictly contained in the cone C_s.

For each $s \in S$ let $\ell_s \in \mathbb{Z}$ be chosen such that $2^{-\ell_s} < |\omega_s| \le 2^{-\ell_s+1}$. We perform a further discretization of each conical sector $\Omega_{s,k}$ by considering Whitney-type decompositions with respect to the distance to the lines determined by the boundary rays r_{s^-} and r_{s^+}; here r_{s^+} denotes the ray emanating from the origin in the direction of $v_{s^+}^\perp$ and similarly for r_{s^-}. For each sector $\Omega_{s,k}$ a central piece which we call $\Omega_{s,k,0}$ is left uncovered by these Whitney decompositions. This is merely a technical issue and we will treat these central pieces separately in what follows.

To make this precise let s, k be fixed and define the regions

$$\begin{aligned} \Omega_{s,k,m} &:= \left\{\xi \in \Omega_{s,k} : \frac{1}{3}2^{-|m|-1} \le \frac{\operatorname{dist}(\xi, r_{s^+})}{|\omega_s|} \le \frac{1}{3}2^{-|m|+1}\right\}, \quad m > 0,\\ \Omega_{s,k,m} &:= \left\{\xi \in \Omega_{s,k} : \frac{1}{3}2^{-|m|-1} \le \frac{\operatorname{dist}(\xi, r_{s^-})}{|\omega_s|} \le \frac{1}{3}2^{-|m|+1}\right\}, \quad m < 0. \end{aligned} \tag{4.11}$$

The central part that was left uncovered corresponds to $m = 0$ and is described as

$$\Omega_{s,k,0} := \left\{\xi \in \Omega_{s,k} : \min(\operatorname{dist}(\xi, r_{s^-}), \operatorname{dist}(\xi, r_{s^+})) \ge \frac{1}{2}\frac{1}{3}|\omega_s|\right\}. \tag{4.12}$$

Notice that the collection $\{\Omega_{s,k,m}\}_{m\in\mathbb{N}}$ is a finitely overlapping cover of $\Omega_{s,k}$. Furthermore the family $\{\Omega_{s,k,m}\}_{s,k,m}$ has finite overlap as the cones $\{C_s\}_{s\in S}$ have finite overlap and for fixed s the family $\{\Omega_{s,k,m}\}_{k,m}$ is Whitney both in k and m.

These geometric considerations are depicted in Figure 4.

The collection of tiles $\boldsymbol{T}$ corresponding to this decomposition is obtained as

$$\boldsymbol{T} := \bigcup_{s\in S} \boldsymbol{T}_s^- \cup \boldsymbol{T}_s^0 \cup \boldsymbol{T}_s^+ \tag{4.13}$$

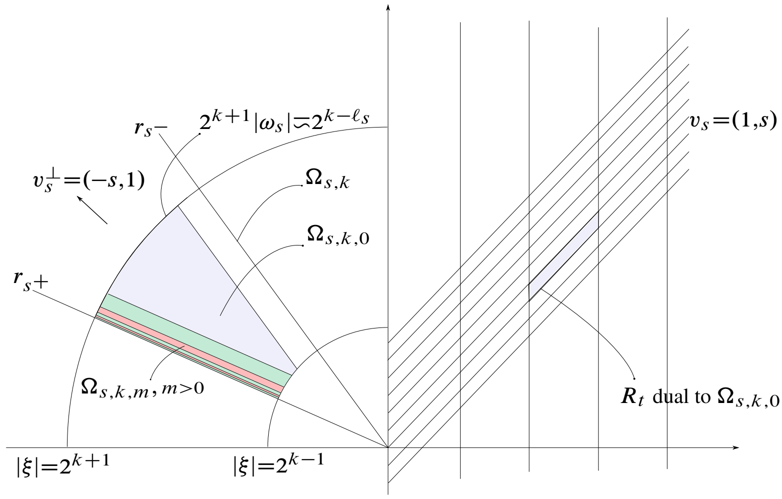

Figure 4. The decomposition of the sector $\Omega_{s,k}$ into Whitney regions, and the spatial grid corresponding to the middle region $\Omega_{s,k,0}$.

where

$$\begin{aligned} \boldsymbol{T}_s^- &:= \bigcup_{k\in\mathbb{Z},\, m<0} \boldsymbol{T}_{s^-,k,m}, \quad \boldsymbol{T}_{s^-,k,m} := \{t = R_t \times \Omega_{s,k,m} : R_t \in \mathcal{D}_{s^-,k,k-\ell_s+|m|}\}, \quad m<0, \\ \boldsymbol{T}_s^0 &:= \bigcup_{k\in\mathbb{Z}} \boldsymbol{T}_{s,k,0}, \qquad\quad \boldsymbol{T}_{s,k,0} := \{t = R_t \times \Omega_{s,k,0} : R_t \in \mathcal{D}_{s,k,k-\ell_s}\}, \\ \boldsymbol{T}_s^+ &:= \bigcup_{k\in\mathbb{Z},\, m>0} \boldsymbol{T}_{s^+,k,m}, \quad \boldsymbol{T}_{s^+,k,m} := \{t = R_t \times \Omega_{s,k,m} : R_t \in \mathcal{D}_{s^+,k,k-\ell_s+|m|}\}, \quad m>0. \end{aligned} \tag{4.14}$$

We stress here that for each cone C_s we introduce tiles in three possible directions v_{s^-}, v_s, v_{s^+}. This turns out to be a technical nuisance more than anything else as the total number of directions is still comparable to $\#S$, and our estimates will be uniform over all S with the same cardinality. However in order to avoid confusion we set

$$S^* := S \cup \{s^- : s \in S\} \cup \{s^+ : s \in S\} =: S^- \cup S \cup S^+. \tag{4.15}$$

Note also that for fixed s, k, m the choice of scales for R_t yields that the tile $t = R_t \times \Omega_{s,k,m}$ obeys the uncertainty principle in both radial and tangential directions.

We then define the associated intrinsic square function by

$$\Delta_{\boldsymbol{T}}(f) := \left(\sum_{t\in\boldsymbol{T}} a_t(f) \frac{\mathbf{1}_{R_t}}{|R_t|} \right)^{\frac{1}{2}}, \tag{4.16}$$

where the set of slopes S are kept implicit in the notation. Here we remember the notation $a_t(f)$ that was introduced in (4.2). Using the orthogonality estimates of Section 4.3 as input for Theorem C, we readily obtain the estimates of the following theorem.

Theorem G. *We have the estimates*

$$\|\Delta_{\boldsymbol{T}} : L^p(\mathbb{R}^2)\| \lesssim_p (\log \#S)^{\frac{1}{2}-\frac{1}{p}} (\log\log \#S)^{\frac{1}{2}-\frac{1}{p}}, \quad 2 \le p < 4, \tag{4.17}$$

$$\sup_{E,f} \frac{\|\Delta_{\boldsymbol{T}}(f\mathbf{1}_E)\|_4}{|E|^{\frac{1}{4}}} \lesssim (\log \#S)^{\frac{1}{4}} (\log\log \#S)^{\frac{1}{4}}, \tag{4.18}$$

where the supremum in the last display is taken over all measurable sets $E \subset \mathbb{R}^2$ of finite positive measure and all Schwartz functions f on $\mathbb{R}^2$ with $\|f\|_\infty \le 1$.

Proof. First of all, observe that the case $p = 2$ of (4.17) is exactly the conclusion of Lemma 4.4. By restricted weak-type interpolation it thus suffices to prove (4.18) to obtain the remaining cases of (4.17); we turn to the former task.

For convenience define $S^* := S \cup \{s^- : s \in S\} \cup \{s^+ : s \in S\} =: S^- \cup S \cup S^+$; note that this is the actual set of slopes of tiles in $\boldsymbol{T}$. Let

$$\mathcal{R}_{\boldsymbol{T}} := \{R_t : t \in \boldsymbol{T}\} \subset \mathcal{D}^2_{S^*}.$$

Observe that we can write

$$\Delta_{\boldsymbol{T}}(f\mathbf{1}_E)^2 = \sum_{R\in\mathcal{R}_{\boldsymbol{T}}} \Bigg(\sum_{t\in\boldsymbol{T}:R_t=R} a_t(f\mathbf{1}_E)\Bigg)\frac{\mathbf{1}_R}{|R|} =: \sum_{R\in\mathcal{R}_{\boldsymbol{T}}} a_R \frac{\mathbf{1}_R}{|R|},$$

where

$$a := \Bigg\{a_R = \sum_{t\in\boldsymbol{T}:R_t=R} a_t(f\mathbf{1}_E) : R \in \mathcal{R}_{\boldsymbol{T}}\Bigg\}.$$

We fix E and f as in the statement and we will obtain (4.18) from an application of Theorem C to the Carleson sequence $a = \{a_R\}_{R\in\mathcal{R}_{\boldsymbol{T}}}$.

First, $\mathsf{mass}_a \lesssim |E|$ as a consequence of Lemma 4.4 since

$$\sum_{R\in\mathcal{R}_{\boldsymbol{T}}} a_R = \sum_{R\in\mathcal{R}_{\boldsymbol{T}}} \sum_{t\in\boldsymbol{T}:R_t=R} a_t(f\mathbf{1}_E) = \sum_{t\in\boldsymbol{T}} a_t(f\mathbf{1}_E) \lesssim \|f\mathbf{1}_E\|_2^2 \lesssim |E|.$$

Further, the fact that a is (a constant multiple of) an L^∞-normalized Carleson sequence is a consequence of the localized estimate of Lemma 4.8. To verify this we need to check the validity of Definition 2.7 for the sequence a above. To that end let $\mathcal{L} \subset \mathcal{D}^2_{S^*}$ be a collection of parallelograms which is subordinate to $\mathcal{T} \subset \mathcal{D}^2_\sigma$ for some fixed $\sigma \in S^*$. Then

$$\sum_{R\in\mathcal{L}} a_R = \sum_{R\in\mathcal{L}} \sum_{t\in\boldsymbol{T}:R_t=R} a_t(f\mathbf{1}_E) = \sum_{t\in\boldsymbol{T}_{\mathcal{L}}} a_t(f\mathbf{1}_E),$$

where $\boldsymbol{T}_{\mathcal{L}} := \{t \in \boldsymbol{T} : R_t \in \mathcal{L}\}$. By Lemma 4.8 the right-hand side of the display above can be estimated by a constant multiple of $|\mathsf{sh}(\mathcal{T})|\|f\mathbf{1}_E\|_\infty^2 \le |\mathsf{sh}(\mathcal{T})|$. This shows the desired property in the definition of a Carleson sequence.

Finally if $\boldsymbol{T}_\sigma := \{t \in \boldsymbol{T} : s(t) = \sigma\}$ for $\sigma \in S^*$, we have that

$$\sup_{\sigma\in S^*} \|\mathrm{M}_{\mathcal{R}_{\boldsymbol{T}_\sigma}} : L^p(\mathbb{R}^2) \to L^{p,\infty}(\mathbb{R}^2)\| \lesssim p', \quad p \to 1^+.$$

Indeed note that for fixed direction $\sigma \in S^*$ each maximal operator appearing in the estimate above is bounded by the strong maximal operator in the coordinates (v, e_2) with $v = (1, \sigma)$.

Now Theorem C applies to the Carleson sequence $a = \{a_R\}_{R \in \mathcal{R}_{\boldsymbol{T}}}$ yielding

$$\|\Delta_{\boldsymbol{T}}(f \mathbf{1}_E)\|_4^4 = \|T_{\mathcal{R}_{\boldsymbol{T}}}(a)\|_2^2 \lesssim (\log \# S^*)(\log \log \# S^*)\, \mathsf{mass}_a \lesssim (\log \# S)(\log \log \# S)|E|,$$

which is the claimed estimate (4.18) as $\# S^* \simeq \# S$. The proof of Theorem G is thus complete. $\square$

4.19. ***The intrinsic square function associated with smooth frequency cones.*** The tiles in the previous subsection were used to model rough frequency projections on a collection of essentially disjoint cones. Indeed note that all decompositions were of Whitney type with respect to all the singular sets of the corresponding rough multiplier. In the case of smooth frequency projections on cones we need a simplified collection of tiles that we briefly describe below.

Assuming S is a finite set of slopes and the arcs $\{\omega_s\}_{s \in S}$ on $\mathbb{S}^1$ have finite overlap as before we now define for $s \in S$ and $k \in \mathbb{Z}$ the collections

$$\boldsymbol{T}_{s,k} := \{t = R_t \times \Omega_{s,k} : R_t \in \mathcal{D}_{s,k-\ell_s,k}\}, \quad \boldsymbol{T}_s := \bigcup_{k \in \mathbb{Z}} \boldsymbol{T}_{s,k}, \quad \boldsymbol{T} := \bigcup_{s \in S} \boldsymbol{T}_s, \tag{4.20}$$

with $\Omega_{s,k}$ given by (4.10). Here we also assume that $2^{-\ell_s} \leq |\omega_s| \leq 2^{-\ell_s+1}$. Notice that each conical sector $\Omega_{s,k}$ now generates exactly one frequency component of possible tiles in contrast with the previous subsection where we need a whole Whitney collection for every s and every k; in fact the tiles $\boldsymbol{T}_{s,k}$ are for all practical purposes the same as the tiles $\boldsymbol{T}_{s,k,0}$ considered in Section 4.9. It is of some importance to note here that for each fixed $s \in S$ the collection $\mathcal{R}_{\boldsymbol{T}} := \{R_t : t \in \boldsymbol{T}\}$ consists of parallelograms of fixed eccentricity 2^{ℓ_s} and thus the corresponding maximal operator $\mathrm{M}_{\mathcal{R}_{\boldsymbol{T}_s}}$ is of weak-type-$(1,1)$ uniformly in $s \in S$:

$$\sup_{s \in S} \|\mathrm{M}_{\mathcal{R}_{\boldsymbol{T}_s}} : L^1(\mathbb{R}^2) \to L^{1,\infty}(\mathbb{R}^2)\| \lesssim 1.$$

The intrinsic square function $\Delta_{\boldsymbol{T}}$ is formally given as in (4.16) but defined with respect to the new collection of tiles defined in (4.20). A repetition of the arguments that led to the proof of Theorem G yields the following.

Theorem H. *For $\boldsymbol{T}$ defined by (4.20) we have the estimates*

$$\|\Delta_{\boldsymbol{T}} : L^p(\mathbb{R}^2)\| \lesssim_p (\log \# S)^{\frac{1}{2}-\frac{1}{p}}, \quad 2 \leq p < 4,$$

$$\sup_{E,f} \frac{\|\Delta_{\boldsymbol{T}}(f \mathbf{1}_E)\|_4}{|E|^{\frac{1}{4}}} \lesssim (\log \# S)^{\frac{1}{4}},$$

where the supremum in the last display is taken over all measurable sets $E \subset \mathbb{R}^2$ of finite positive measure and all Schwartz functions f on $\mathbb{R}^2$ with $\|f\|_\infty \leq 1$.

4.21. ***The intrinsic square function associated with rough frequency rectangles.*** The considerations in this subsection aim at providing the appropriate time-frequency analysis in order to deal with a Rubio-de-Francia-type square function, given by frequency projections on disjoint rectangles in finitely many directions. The intrinsic setup is described by considering again a finite set of slopes S and corresponding directions V. Suppose that we are given a finitely overlapping collection of rectangles $\mathcal{F}=\bigcup_{s\in S}\mathcal{F}_s$, consisting of rectangles which are tensor products of intervals in the coordinates $v, v^{\perp},\ v=(1,s)$, for some $s\in S$. Namely a rectangle $F\in\mathcal{F}_s$ is a *rotation* by s of an axis-parallel rectangle. We stress that the rectangles in each collection $\mathcal{F}_s$ are generic two-parameter rectangles, namely their sides have independent lengths (there is no restriction on their eccentricity).

We also note that $\mathcal{F}_s$ consists of rectangles rather than parallelograms and this difference is important when one deals with rough frequency projections. Our techniques are sufficient to deal with the case of parallelograms as well but we just choose to detail the setup for the rectangular case. The interested reader will have no trouble adjusting the proof for variations of our main statement below for the case of parallelograms, or for the case that the families $\mathcal{F}_s$ are in fact one-parameter families.

Given $F\in\mathcal{F}_s$ we define a two-parameter Whitney discretization as follows. Let $F=\operatorname{rot}_s(I\times J)+y_F$ for some $y_F\in\mathbb{R}^2$, where rot_s denotes counterclockwise rotation by s about the origin and $I\times J$ is an axis parallel rectangle centered at the origin. Note that $I=(-|I|/2,|I|/2)$ and similarly for J. Then we define for $(k_1,k_2)\in\mathbb{N}^2,\ k_1,k_2\neq 0$,

$$W_{k_1,k_2}(F):=\left\{\xi\in I\times J:\frac{1}{3}2^{-k_1-1}\leq\frac{1}{2}-\frac{|\xi_1|}{|I|}\leq\frac{1}{3}2^{-k_1+1},\ \frac{1}{3}2^{-k_2-1}\leq\frac{1}{2}-\frac{|\xi_2|}{|J|}\leq\frac{1}{3}2^{-k_2+1}\right\}.$$

The definition has to be adjusted for $k_1=0$ or $k_2=0$. For example we define for $k_2\neq 0$

$$W_{0,k_2}(F):=\left\{\xi\in I\times J:\frac{1}{2}|I|-|\xi_1|\geq\frac{1}{2}\frac{1}{3}|I|,\ \frac{1}{3}2^{-k_2-1}|J|\leq\frac{1}{2}|J|-|\xi_2|\leq\frac{1}{3}2^{-k_2+1}|J|\right\}$$

and symmetrically for $k_1\neq 0$ and $k_2=0$. Finally

$$W_{0,0}(F):=\left\{\xi\in I\times J:\frac{1}{2}|I|-|\xi_1|\geq\frac{1}{2}\frac{1}{3}|I|,\ \frac{1}{2}|J|-|\xi_2|\geq\frac{1}{2}\frac{1}{3}|J|\right\}.$$

Then for $k=(k_1,k_2)\in\mathbb{N}^2$ we set $\Omega_{s,k_1,k_2}(F):=\operatorname{rot}_s(W_{k_1,k_2}(F))+y_F$.

We can define tiles for this system as follows. If $F\in\mathcal{F}_s$ for some $s\in S$ and $F=\operatorname{rot}_s(I\times J)+y_F$ with $I\times J$ as above, then we choose $\ell_I^F,\ell_J^F\in\mathbb{Z}$ such that $2^{\ell_I^F}<|I|\leq 2^{\ell_I^F+1}$ and $2^{\ell_J^F}<|J|\leq 2^{\ell_J^F+1}$. We will have

$$\boldsymbol{T}^{\mathcal{F}}:=\bigcup_{s\in S}\boldsymbol{T}_s^{\mathcal{F}},\quad \boldsymbol{T}_s^{\mathcal{F}}:=\bigcup_{F\in\mathcal{F}_s}\boldsymbol{T}_s(F),\quad \boldsymbol{T}_s(F):=\bigcup_{(k_1,k_2)\in\mathbb{N}^2}\boldsymbol{T}_{s,k_1,k_2}(F),\quad F\in\mathcal{F}_s, \tag{4.22}$$

where

$$\boldsymbol{T}_{s,k_1,k_2}(F):=\{t=R_t\times\Omega_{s,k_1,k_2}(F):R_t\in\mathcal{D}_{s,-k_2+\ell_J^F,-k_1+\ell_I^F}\},\quad F\in\mathcal{F}_s.$$

Note again that the tiles defined above obey the uncertainty principle in both $v,v^{\perp}$ for every fixed $v=(1,s)$ with $s\in S$.

The intrinsic square function associated with the collection $\mathcal{F}$ is denoted by $\Delta_{\boldsymbol{T}^{\mathcal{F}}}$ and formally has the same definition as (4.16), where now the $\boldsymbol{T}$ are given by the collection $\boldsymbol{T}^{\mathcal{F}}$ of (4.22). The corresponding theorem is the intrinsic analogue of a multiparameter directional Rubio de Francia square function estimate.

Theorem I. *Let $\mathcal{F}$ be a finitely overlapping collection of two-parameter rectangles in directions given by S*

$$\Big\|\sum_{F\in\mathcal{F}}\mathbf{1}_F\Big\|_\infty \lesssim 1.$$

Consider the collection of tiles $\boldsymbol{T}^{\mathcal{F}}$ defined in (4.22) and let $\Delta_{\boldsymbol{T}^{\mathcal{F}}}$ be the corresponding intrinsic square function. We have the estimates

$$\begin{aligned}
\|\Delta_{\boldsymbol{T}^{\mathcal{F}}} : L^p(\mathbb{R}^2)\| &\lesssim_p (\log\#S)^{\frac12-\frac1p}(\log\log\#S)^{\frac12-\frac1p}, \quad 2\le p<4,\\
\sup_{E,f}\frac{\|\Delta_{\boldsymbol{T}^{\mathcal{F}}}(f\mathbf{1}_E)\|_4}{|E|^{\frac14}} &\lesssim (\log\#S)^{\frac14}(\log\log\#S)^{\frac14},
\end{aligned}$$

where the supremum in the last display is taken over all measurable sets $E\subset\mathbb{R}^2$ of finite positive measure and all Schwartz functions f on $\mathbb{R}^2$ with $\|f\|_\infty\le 1$.

Remark 4.23. As before, there is slight improvement in the case of one-parameter spatial components in each direction. More precisely suppose that $\mathcal{F}=\bigcup_{s\in S}\mathcal{F}_s$ is a given collection of disjoint rectangles in directions given by S. If for each $s\in S$ the family $\mathcal{R}_{\mathcal{F}_s}:=\{R_t : t\in\boldsymbol{T}_{\mathcal{F}_s}\}$ yields a weak-type-$(1,1)$ maximal operator then the estimates of Theorem I hold without the $\log\log$-terms.

Remark 4.24. Suppose that $\mathcal{R}=\bigcup_{s\in S}\mathcal{R}_s\subset\mathcal{P}^2_S$ is a family of parallelograms in directions given by s; namely we have that if $R\in\mathcal{R}_s$ then $R=A_s(I\times J)+y_R$ for some rectangle $I\times J$ in $\mathbb{R}^2$ with sides parallel to the coordinate axes and centered at 0, and $y_R\in\mathbb{R}^2$. Now there is an obvious way to construct a Whitney partition of each $R\in\mathcal{R}$. Indeed we just define the frequency components

$$\Omega_{s,k_1,k_2}(R):=A_s(W_{k_1,k_2}(I\times J))+y_R,$$

with $W_{k_1,k_2}(I\times J)$ as constructed before. Then

$$\boldsymbol{T}_{s,k_1,k_2}(R):=\{R_t\times\Omega_{s,k_1,k_2}(R) : R_t\in\mathcal{D}_{s,-k_2+\ell^F_J,-k_1+\ell^F_I}\},\quad R\in\mathcal{R}_s,$$

and $\boldsymbol{T}$ are given as in (4.22). With this definition there is a corresponding intrinsic square function $\Delta_{\boldsymbol{T}_{\mathcal{R}}}$ which satisfies the bounds of Theorem I. The improvement of Remark 4.23 is also valid if $\mathcal{R}=\bigcup_{s\in S}\mathcal{R}_s$ and each $\mathcal{R}_s$ consists of rectangles of fixed eccentricity.

The proof of Theorem I relies again on the global and local orthogonality estimates of Section 4.3 and a subsequent application of the directional Carleson embedding theorem, Theorem C. We omit the details.

5. Sharp bounds for conical square functions

We begin this section by recalling the definition for the smooth conical frequency projections given in Section 1. Let $\tau\subset[0,2\pi)$ be an interval and consider the corresponding rough cone multiplier

$$C_\tau f(x):=\int_0^{2\pi}\int_0^\infty \hat f(\varrho e^{i\vartheta})\mathbf{1}_\tau(\vartheta)e^{ix\cdot\varrho e^{i\vartheta}}\varrho\,d\varrho\,d\vartheta,\quad x\in\mathbb{R}^2,$$

and its smooth analogue

$$C_\tau^\circ f(x) := \int_0^{2\pi}\int_0^\infty \hat{f}(\varrho e^{i\vartheta})\beta\left(\frac{\vartheta - c_\tau}{|\tau|/2}\right)e^{ix\cdot\varrho e^{i\vartheta}}\varrho\,d\varrho\,d\vartheta, \quad x \in \mathbb{R}^2, \tag{5.1}$$

where β is a smooth function on $\mathbb{R}$ supported on $[-1,1]$ and equal to 1 on $\left[-\frac{1}{2},\frac{1}{2}\right]$ and c_τ, $|\tau|$ stand respectively for the center and length of τ.

This section is dedicated to the proofs of two related theorems concerning conical square functions. The first is a quantitative estimate for a square function associated with the smooth conical multipliers of a finite collection of intervals with bounded overlap given in Theorem A, namely the estimates

$$\|\{C_\tau^\circ f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim_p (\log \#\boldsymbol{\tau})^{\frac{1}{2}-\frac{1}{p}}\|f\|_p$$

for $2 \le p < 4$, as well as the restricted-type analogue valid for all measurable sets E

$$\|\{C_\tau^\circ(f\mathbf{1}_E)\}\|_{L^4(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim (\log \#\boldsymbol{\tau})^{\frac{1}{4}}|E|^{\frac{1}{4}}\|f\|_\infty,$$

under the assumption of finite overlap

$$\Big\|\sum_{\tau\in\boldsymbol{\tau}} \mathbf{1}_\tau\Big\|_\infty \lesssim 1. \tag{5.2}$$

The second theorem concerns an estimate for the rough conical square function for a collection of finitely overlapping cones $\boldsymbol{\tau}$.

Theorem J. *Let $\boldsymbol{\tau}$ be a finite collection intervals in $[0,2\pi)$ with finite overlap as in (5.2). Then the square function estimate*

$$\|\{C_\tau f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim_p (\log \#\boldsymbol{\tau})^{1-\frac{2}{p}}(\log\log \#\boldsymbol{\tau})^{\frac{1}{2}-\frac{1}{p}}\|f\|_p \tag{5.3}$$

holds for each $2 \le p < 4$.

Theorem A is sharp, in terms of $\log \#\boldsymbol{\omega}$-dependence, for all $2 \le p < 4$ and for $p = 4$ up to the restricted type. Theorem J improves on [Córdoba 1982, Theorem 1], where the dependence on cardinality is unspecified. Examples providing a lower bound of $(\log \#\boldsymbol{\omega})^{1/2-1/p}\|f\|_p$ for the left-hand side of (5.3), and showing the sharpness of Theorem A, are detailed in Section 8.

The remainder of the section is articulated as follows. In the upcoming Section 5.4 we show Theorem A. The subsequent subsection is dedicated to the proof of Theorem J.

5.4. ***Proof of Theorem A.*** We are given a finite collection of intervals $\omega \in \boldsymbol{\omega}$ having bounded overlap as in (5.2). By finite splitting we may reduce to the case of $\omega \in \boldsymbol{\omega}$ being pairwise disjoint; we treat this case throughout.

The first step in the proof of Theorem A is a radial decoupling. Let ψ be a smooth radial function on $\mathbb{R}^2$ with

$$\mathbf{1}_{[1,2]}(|\xi|) \le \psi(\xi) \le \mathbf{1}_{[2^{-1},2^2]}(|\xi|)$$

and define the Littlewood–Paley projection

$$S_k f(x) := \int \psi(2^{-k}\xi)\hat{f}(\xi)\,e^{ix\cdot\xi}\,d\xi, \quad x \in \mathbb{R}^2.$$

The following weighted Littlewood–Paley inequality is contained in [Bennett and Harrison 2012, Proposition 4.1].

Proposition 5.5 [Bennett and Harrison 2012]. *Let w be a nonnegative locally integrable function. Then*

$$\int_{\mathbb{R}^2} |f|^2 w \lesssim \int_{\mathbb{R}^2} \sum_{k\in\mathbb{Z}} |S_k(f)|^2 \mathrm{M}^{[3]} w,$$

with implicit constant independent of w, f, where we recall that $\mathrm{M}^{[3]}$ denotes the three-fold iteration of the Hardy–Littlewood maximal operator M with itself.

We may easily deduce the next lemma from the proposition.

Lemma 5.6. *For any $p \geq 2$ we have*

$$\|\{C_\tau^\circ f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} \lesssim \Big\| \Big(\sum_{k\in\mathbb{Z},\, \tau\in\boldsymbol{\tau}} |C_\tau^\circ S_k(f)|^2 \Big)^{\frac{1}{2}} \Big\|_p. \tag{5.7}$$

Proof. The case $p = 2$ is trivial so we assume $p > 2$. Letting $r := \frac{p}{2} > 1$ there exists some $w \in L^{r'}(\mathbb{R}^2)$ with $\|w\|_{r'} = 1$ such that

$$\|\{C_\tau^\circ f\}\|^2_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\tau}})} = \sum_{\tau\in\boldsymbol{\tau}} \int_{\mathbb{R}^2} |C_\tau^\circ f|^2 w \lesssim \sum_{k\in\mathbb{Z},\, \tau\in\boldsymbol{\tau}} \int_{\mathbb{R}^2} |C_\tau^\circ S_k(f)|^2 M^{[3]} w$$

and the lemma follows by Hölder's inequality and the boundedness of $M^{[3]}$ on $L^{r'}(\mathbb{R}^2)$. □

The second and final step of the proof of Theorem A is the reduction of the operator appearing in the right-hand side of (5.7) to the model operator of Theorem H.

In order to match the notation of Section 4.9 we write $\{\omega_s\}_{s\in S}$ for the collection of arcs in $\mathbb{S}^1$ corresponding to the collection of intervals $\boldsymbol{\tau}$, namely for $\tau \in \boldsymbol{\tau}$ we implicitly define $s = s_\tau$ by means of $v_s^\perp/|v_s^\perp| := \mathrm{e}^{ic_\tau} = (1, s)/|(1, s)|$. We set $S := \{s_\tau : \tau \in \boldsymbol{\tau}\}$ and define the corresponding arcs in $\mathbb{S}^1$ as

$$\omega_{s_\tau} := \{\mathrm{e}^{i\theta} : \theta \in \tau\}.$$

Now the cone C_τ is the same thing as the cone C_s and $\#S = \#\boldsymbol{\tau}$. Similarly we write $C_\tau^\circ = C_{s_\tau}^\circ$ so the cones can now be indexed by $s \in S$. Define ℓ_s such that $2^{-\ell_s} \leq |\omega_s| \leq 2^{-\ell_s+1}$.

By finite splitting and rotational invariance there is no loss in generality with assuming that $S \subset [-1, 1]$. Notice that the support of the multiplier of $C_s^\circ S_k$ is contained in the frequency sector $\Omega_{s,k}$ defined in (4.10). By standard procedures of time-frequency analysis, as for example in [Demeter and Di Plinio 2014, Section 6], the operator $C_s^\circ S_k$ can be recovered by appropriate averages of operators

$$\mathsf{C}_{s,k} f := \sum_{t\in\boldsymbol{T}_{s,k}} \langle f, \phi_t \rangle \phi_t,$$

where $\phi_t \in \mathcal{A}_t^{8M_0}$ for all $t \in \boldsymbol{T}_{s_\omega,k}$ and $\boldsymbol{T}_{s,k}$ is defined in (4.20). Here $M_0 = 2^{50}$ is as chosen in (4.2). Fixing s, k for the moment we preliminarily observe that for each $\nu \geq 1$ the collection

$\mathcal{R}_{s,k} := \mathcal{R}_{\boldsymbol{T}_{s,k}} = \{R_t : t \in \boldsymbol{T}_{s,k}\}$ can be partitioned into subcollections $\{\mathcal{R}^j_{s,k,\nu} : 1 \le j \le 2^{8\nu}\}$ with the property that

$$R_1, R_2 \in \mathcal{R}^j_{s,k,\nu} \quad \Longrightarrow \quad 2^{2\nu+4} R_1 \cap 2^{2\nu+4} R_2 = \varnothing.$$

We will also use below the Schwartz decay of $\phi_t \in \mathcal{A}^{M_0}_t$ in the form

$$\sqrt{|R_t|}|\phi_t| \lesssim \mathbf{1}_{R_t} + \sum_{\nu \ge 0} 2^{-8M_0\nu} \sum_{\substack{\rho \in \mathcal{R}_{s,k} \\ \rho \not\subset 2^\nu R_t,\, \rho \subset 2^{\nu+1} R_t}} \mathbf{1}_\rho.$$

Using Schwartz decay of ϕ_t twice, in particular to bound by an absolute constant the second factor obtained by Cauchy–Schwarz after the first step, we get

$$\begin{aligned}
|\mathsf{C}_{s,k} f|^2 &\lesssim \Bigg(\sum_{t \in \boldsymbol{T}_{s,k}} |\langle f, \phi_t\rangle|^2 \frac{|\phi_t|}{\sqrt{|R_t|}}\Bigg)\Bigg(\sum_{t \in \boldsymbol{T}_{s,k}} \sqrt{|R_t|}|\phi_t|\Bigg) \\
&\lesssim \sum_{t \in \boldsymbol{T}_{s,k}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_{R_t}}{|R_t|} + \sum_{\nu \ge 0} 2^{-8M_0\nu} \sum_{t \in \boldsymbol{T}_{s,k}} \sum_{\substack{\rho \in \mathcal{R}_{s,k} \\ \rho \not\subset 2^\nu R_t, \rho \subset 2^{\nu+1} R_t}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_\rho}{|\rho|} \\
&\le \sum_{t \in \boldsymbol{T}_{s,k}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_{R_t}}{|R_t|} + \sum_{\nu \ge 0} 2^{-8M_0\nu} \sum_{R \in \mathcal{R}_{s,k}} \sum_{j=1}^{2^{8\nu}} \sum_{\substack{\rho \in \mathcal{R}_{s,k} \\ \rho \not\subset 2^\nu R_t, \rho \subset 2^{\nu+1} R_t}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_\rho}{|\rho|}.
\end{aligned}$$

Now for fixed ω, k, ν, j and $t \in \boldsymbol{T}_{s,k}$ observe that there is at most one $\rho = \rho^j_{s,k,\nu}(t) \in \mathcal{R}^j_{\omega,k,\nu}$ such that $\rho \not\subset 2^\nu R_t,\ \rho \subset 2^{\nu+1} R_t$. Thus the estimate above can be written in the form

$$|\mathsf{C}_{s,k} f|^2 \lesssim \sum_{t \in \boldsymbol{T}_{s,k}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_{R_t}}{|R_t|} + \sum_{\nu \ge 0} 2^{-8M_0\nu} \sum_{j=1}^{2^{8\nu}} \sum_{t \in \boldsymbol{T}_{s,k}} |\langle f, \phi_t\rangle|^2 \frac{\mathbf{1}_{\rho^j_{s,k,\nu}(t)}}{|\rho^j_{s,k,\nu}(t)|}.$$

Observe that if $t \in \boldsymbol{T}_{s,k}$,

$$\phi_t \in \mathcal{A}^{8M_0}_t, \ \ \rho \in \mathcal{R}_{s,k}, \ \ \rho \subset 2^{\nu+1} R_t \quad \Longrightarrow \quad 2^{-4M\nu}|\langle f, \phi_t\rangle|^2 \le a_{t_\rho}(f),$$

where $t_\rho = \rho \times \Omega_{s,k} \in \boldsymbol{T}_{s,k}$ is the unique tile with spatial localization given by ρ; this is because $2^{-4M\nu}\phi_t \in \mathcal{A}^{M_0}_{t_\rho}$. We thus conclude that

$$|\mathsf{C}_{s,k} f|^2 \lesssim \sum_{t \in \boldsymbol{T}_{s,k}} a_t(f) \frac{\mathbf{1}_{R_t}}{|R_t|}. \tag{5.8}$$

Comparing with the definition of $\Delta_{\boldsymbol{T}}$ given in (4.16) we may summarize the discussion in the lemma below.

Lemma 5.9. *Let $1 < p < \infty$. Then*

$$\sup_{\|f\|_p = 1} \Bigg\| \Bigg(\sum_{k \in \mathbb{Z},\, \tau \in \boldsymbol{\tau}} |C^\circ_\tau S_k(f)|^2 \Bigg)^{\frac{1}{2}} \Bigg\|_p \lesssim \sup_{\|f\|_p = 1} \|\Delta_{\boldsymbol{T}}(f)\|_p,$$

where

$$\boldsymbol{T} := \bigcup_{s\in S}\bigcup_{k\in\mathbb{Z}} \boldsymbol{T}_{s,k}$$

and $\boldsymbol{T}_{s,k}$ is defined in (4.20).

The proof of the upper bound in Theorem A is then completed by juxtaposing the estimates of Lemmas 5.6 and 5.9 with Theorem H. For the optimality of the estimate see Section 8.6.

5.10. *Proof of Theorem J.* The proof of Theorem J is necessarily more involved than its smooth counterpart Theorem A. In particular we need to decompose each cone not only in the radial direction as before, but also in the directions perpendicular to the singular boundary of each cone. We describe this procedure below.

Consider a collection of intervals $\boldsymbol{\tau} = \{\tau\}$ as in the statement. By the same correspondence as in the proof of Theorem A we pass to a family $\{\omega_s\}_{s\in S}$ consisting of finitely overlapping arcs on $\mathbb{S}^1$ centered at $v_s^\perp/|v_s^\perp|$ and corresponding cones C_s. Note that the sectors $\{\Omega_{s,k}\}_{s\in S,k\in\mathbb{Z}}$ defined in (4.10) form a finitely overlapping cover of $\bigcup_{s\in S} C_s$. We remember here that $v_s = (1,s)$, that the interval ω_s is given by $(v_{s-}^\perp, v_{s+}^\perp)$, and that the positive direction is counterclockwise.

Now, for each fixed $s \in S$ the cover $\{\Omega_{s,k,m}\}_{(k,m)\in\mathbb{Z}^2}$ defined in (4.11), (4.12), is a Whitney cover of $\Omega_{s,k}$ in the product sense: for each $\Omega_{s,k,m}$ the distance from the origin is comparable to 2^k and the distance to the boundary is comparable to $2^{-|m||\omega_s|}$.

The radial decomposition in k will be taken care of by the Littlewood–Paley decomposition $\{S_k\}_{k\in\mathbb{Z}}$, defined as in the proof of Theorem J. Now for fixed s,k we consider a smooth partition of unity subordinated to the cover $\{\Omega_{s,k,m}\}_{m\in\mathbb{Z}}$. Note that one can easily achieve that by choosing $\{\varphi_{s,m}\}_{m<0}$ to be a one-sided (contained in C_s) Littlewood–Paley decomposition in the negative direction $v^- = v_{s-}$, and constant in the direction $(v^-)^\perp$ when $m < 0$, and similarly one can define $\varphi_{s,m}$ when $m > 0$, with respect to the positive direction v^+. The central piece $\Omega_{s,k,0}$ corresponds to φ_{s_0} defined implicitly as

$$\varphi_{s,0} = \mathbf{1}_{C_s} - \sum_{m\in\mathbb{Z}} \varphi_{s,m}.$$

Now the desired partition of unity is

$$\pi_{s,k,m}(\xi) := \mathbf{1}_{C_s}(\xi)\varphi_{s,m}(\xi)\psi_k(\xi) = \varphi_{s,m}(\xi)\psi_k(\xi),$$

where $\psi_k := \psi(2^{-k}\,\cdot)$, with the ψ constructed in the proof of Theorem A. Remember that $S_k f := (\psi_k \hat{f})^\vee$ and let us define $\Phi_{s,m} f := (\varphi_{s,m}\hat{f})^\vee$.

An important step in the proof is the following square function estimate in $L^p(\mathbb{R}^2)$, with $2 \le p < 4$, that decouples the Whitney pieces in every cone C_s. It comes at a loss in N, which appears to be inevitable because of the directional nature of the problem.

Lemma 5.11. *Let $\{C_s\}_{s\in S}$ be a family of frequency cones, given by a family of finitely overlapping arcs $\boldsymbol{\omega} := \{\omega_s\}_{s\in S}$ as above. For $2 \le p < 4$ there holds*

$$\|\{C_s f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\omega}})} \lesssim \frac{1}{4-p}(\log \#S)^{\frac{1}{2}-\frac{1}{p}} \|\{S_k \Phi_{s,m} f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\boldsymbol{\omega}\times\mathbb{Z}\times\mathbb{Z}\}})}.$$

Proof. Observe that the desired estimate is trivial for $p = 2$ so let us fix some $p \in (2, 4)$. There exists some $g \in L^q$ with $q = (p/2)' = p/(p-2)$ such that

$$A^2 := \|\{C_s f\}\|^2_{L^p(\mathbb{R}^2;\ell^2_\omega)} = \int_{\mathbb{R}^2} \sum_{s\in S} |C_s f|^2 g$$

and so by Proposition 5.5 we get

$$A^2 \lesssim \sum_{k\in\mathbb{Z}} \sum_{s\in S} \int_{\mathbb{R}^2} |C_s S_k f|^2 \mathrm{M}^{[3]} g,$$

where we recall that $\mathrm{M}^{[3]}$ denotes three iterations of the Hardy–Littlewood maximal operator M. Fixing s for a moment we use Proposition 5.5 in the directions v_{s^-}, v_s and v_{s^+} to further estimate

$$\int_{\mathbb{R}^2} |C_s f|^2 \mathrm{M}^{[3]} g \lesssim \sum_{m\in\mathbb{Z}} \sum_{\varepsilon\in\{-,0,+\}} \int_{\mathbb{R}^2} |S_k \Phi_{s,m} f|^2 \mathrm{M}^{[3]}_{v_{s^\varepsilon}} \mathrm{M}^{[3]} g,$$

where we adopted the convention $s^0 := s$ for brevity, and M_v is given by (2.3). Remember also that $\Phi_{s,m}$ for $m > 0$ corresponds to directions s^+, while $\Phi_{s,m}$ corresponds to directions s^- for $m < 0$, and to directions $s^0 = s$ for $m = 0$. Now for any $v \in \mathbb{S}^1$ and $r > 1$ we have that

$$\mathrm{M}^{[3]}_v G \lesssim (r')^2 [\mathrm{M}_v G^r]^{\frac{1}{r}};$$

see for example [Pérez 1994]. Thus $\mathrm{M}^{[3]}_{v_{s^\varepsilon}} \mathrm{M}^{[3]} g \lesssim (r')^2 [\mathrm{M}_{V^*}[\mathrm{M}^{[3]} G]^r]^{1/r}$, where $\mathrm{M}_{V^*} f := \sup_{v\in V^*} \mathrm{M}_v f$, where here we use $V^* := \{(1, s) : s \in S^*\}$ with S^* as in (4.15), and $\mathrm{M}_{V^*} f := \sup_{w\in V^*} \mathrm{M}_w(f)$.

It is known [Katz 1999] that M_{V^*} maps $L^p(\mathbb{R}^2)$ to $L^p(\mathbb{R}^2)$ with a bound $(\log \# V^*)^{1/p}$ for $p > 2$. As $p < 4$ there exists a choice of $1 < r < p/(2(p-2))$ so that $p/(r(p-2)) > 2$ and a theorem from [Katz 1999] applies. Using this fact together with Hölder's inequality proves the lemma. □

The proof of Theorem J can now be completed as follows. For each $(s, k, m) \in S \times \mathbb{Z} \times \mathbb{Z}$ the operator $S_k \Phi_{s,m}$ is a smooth frequency projection adapted to the rectangular box $\Omega_{s,k,m}$. Following the same procedure that led to (5.8) in the proof of Theorem A we can approximate each piece $S_k \Phi_{s,m} f$ by an operator of the form

$$\mathsf{C}_{s^\varepsilon,k,m} f := \sum_{t\in \boldsymbol{T}_{s^\varepsilon,k,m}} \langle f, \phi_t\rangle \phi_t, \quad |\mathsf{C}_{s^\varepsilon,k,m} f|^2 \lesssim \sum_{t\in \boldsymbol{T}_{s^\varepsilon,k,m}} a_t(f) \frac{\mathbf{1}_{R_t}}{|R_t|},$$

where s^ε follows the sign of m and coincides with s if $m = 0$. The collections of tiles $\boldsymbol{T}_{s^\varepsilon,k,m}$ are the ones given in (4.14). Now Lemma 5.11 and Theorem G are combined to complete the proof of Theorem J.

6. Directional Rubio de Francia square functions

In his seminal paper Rubio de Francia [1985] proved a one-sided Littlewood–Paley inequality for arbitrary intervals on the line. This estimate was later extended by Journé [1985] to the case of rectangles (n-dimensional intervals) in $\mathbb{R}^n$; a proof more akin to the arguments of the present paper appears in [Lacey 2007]. The aim of this subsection is to present a generalization of the one-sided Littlewood–Paley inequality to the case of rectangles in $\mathbb{R}^2$ with sides parallel to a given set of directions. The set of directions is to be finite, necessarily, because of Kakeya counterexamples.

As in the case of cones of Section 5 we will present two versions, one associated with smooth frequency projections and one with rough. To set things up let S be a finite set of slopes and V be the corresponding directions. We consider a family of rotated rectangles $\mathcal{F}$ as in Section 4.21, where $\mathcal{F} = \bigcup_{s\in S} \mathcal{F}_s$. For each $s \in S$ a rectangle $F \in \mathcal{F}_s$ is a rotation by s of an axis parallel rectangle, so that the sides of R are parallel to $(v, v^{\perp})$ with $v = (1, s)$. We will write $F = \mathrm{rot}_s(I_F \times J_F) + y_F$ for some $y_F \in \mathbb{R}^2$ in order to identify the axes-parallel rectangle $I_F \times J_F$ producing F by an s-rotation; this writing assumes that $I_F \times J_F$ is centered at the origin.

Now for each $F \in \mathcal{F}$ we consider the rough frequency projection

$$P_F f(x) := \int_{\mathbb{R}^2} \hat{f}(\xi)\mathbf{1}_F(\xi)e^{ix\cdot\xi}\,\mathrm{d}\xi, \quad x \in \mathbb{R}^2,$$

and its smooth analogue

$$P_F^{\circ} f(x) := \int_{\mathbb{R}^2} \hat{f}(\xi)\gamma_F(\xi)e^{ix\cdot\xi}\,\mathrm{d}\xi, \quad x \in \mathbb{R}^2,$$

where γ_R is a smooth function on $\mathbb{R}^2$, supported in R, and identically 1 on $\mathrm{rot}_s\big(\frac{1}{2}I \times \frac{1}{2}J\big)$.

We first state the smooth square function estimate.

Theorem K. *Let $\mathcal{F}$ be a collection of rectangles in $\mathbb{R}^2$ with sides parallel to $(v, v^{\perp})$ for some v in a finite set of directions V. Assume that $\mathcal{F}$ has finite overlap. Then*

$$\|\{P_F^{\circ} f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \lesssim_p (\log \#V)^{\frac{1}{2}-\frac{1}{p}}(\log\log \#V)^{\frac{1}{2}-\frac{1}{p}}\|f\|_p$$

for $2 \le p < 4$, as well as the restricted-type analogue valid for all measurable sets E

$$\|\{P_F^{\circ}(f\mathbf{1}_E)\}\|_{L^4(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \lesssim (\log \#V)^{\frac{1}{4}}(\log\log \#V)^{\frac{1}{4}}|E|^{\frac{1}{4}}\|f\|_{\infty}.$$

The dependence on $\#V$ in the estimates above is best possible up the doubly logarithmic term.

Remark 6.1. We record a small improvement of the estimates above in some special cases. Suppose that for fixed $s \in S$ all the rectangles $F \in \mathcal{F}_s$ have one side-length fixed, or that they have fixed eccentricity. In both these cases the collections of spatial components of the tiles needed to discretize these operators, $\mathcal{R}_{\boldsymbol{T}_s^{\mathcal{F}}} := \{R_t : t \in \boldsymbol{T}_s^{\mathcal{F}}\}$, with $\boldsymbol{T}_s^{\mathcal{F}}$ as in (4.22), give rise to maximal operators that are of weak-type $(1, 1)$. Then Remark 4.23 shows that the estimates of Theorem K hold without the doubly logarithmic terms, and as shown in Section 8.2 this is best possible.

The rough version of this Rubio-de-Francia-type theorem is slightly worse in terms of the dependence on the number of directions. The reason for that is that, as in the case of conical projections, passing from rough to smooth in the directional setting incurs a loss of logarithmic terms, essentially originating in the corresponding maximal function bound.

Theorem L. *Let $\mathcal{F}$ be a collection of rectangles in $\mathbb{R}^2$ with sides parallel to $(v, v^{\perp})$ for some v in a finite set of directions V. Assume that $\mathcal{F}$ has finite overlap. Then the following square function estimate holds for $2 \le p < 4$:*

$$\|\{P_F f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \lesssim_p (\log \#V)^{\frac{3}{2}-\frac{3}{p}}(\log\log \#V)^{\frac{1}{2}-\frac{1}{p}}\|f\|_p.$$

The proofs of these theorems follow the by now familiar path of introducing local Littlewood–Paley decompositions on each multiplier, approximating with time-frequency analysis operators, establishing a directional Carleson condition on the wave-packet coefficients and finally applying Theorem C. We will very briefly comment on the proofs below.

Proof of Theorems L and K. We first sketch the proof of Theorem L, which is slightly more involved. The first step here is a decoupling lemma which is completely analogous to Lemma 5.11 with the difference that now we need to use two directional Littlewood–Paley decompositions, while in the case of cones only one. This explains the extra logarithmic term of the statement.

Remember that $\mathcal{F} = \bigcup_s \mathcal{F}_s$, with $s = (1, v)$ for some $v \in V$; here s gives the directions $(v, v^\perp)$ of the rectangles in $\mathcal{F}_s$. Using the finitely overlapping Whitney decomposition of Section 4.21 we have for each $F \in \mathcal{F}_s$ a collection of tiles

$$\boldsymbol{T}_s(F) := \bigcup_{(k_1,k_2)\in\mathbb{Z}^2} \boldsymbol{T}_{s,k_1,k_2}(F)$$

as in (4.22). Let us for a moment fix s and $F \in \mathcal{F}_s$. The frequency components of the tiles in $\boldsymbol{T}_s(F)$ form a two-parameter Whitney decomposition of F, so let $\{\phi_{F,k_1,k_2}\}_{(k_1,k_2)\in\mathbb{Z}^2}$ be a smooth partition of unity subordinated to this cover and denote by Φ_{F,k_1,k_2} the Fourier multiplier with symbol ϕ_{F,k_1,k_2}.

The promised analogue of Lemma 5.11 is the following estimate: for $2 \le p < 4$ there holds

$$\|\{P_F f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \lesssim \frac{1}{(4-p)^2}(\log \#V)^{1-\frac{2}{p}} \|\{\Phi_{s,k_1,k_2} f\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}\times\mathbb{Z}\times\mathbb{Z}\}})}. \tag{6.2}$$

The proof of this estimate is a two-parameter repetition of the proof of Lemma 5.11, where one applies Proposition 5.5 once in the direction of v and once in the direction of $v^\perp$. Using the familiar scheme we can approximate each $\Phi_{s,k_1,k_2} f$ by time-frequency analysis operators

$$\mathsf{P}_{F,k_1,k_2} f := \sum_{t\in\boldsymbol{T}_{s,k_1,k_2}(F)} \langle f, \phi_t\rangle \phi_t, \quad |\mathsf{P}_{F,k_1,k_2} f|^2 \lesssim \sum_{t\in\boldsymbol{T}_{s,k_1,k_2}(F)} a_t(f)\frac{\boldsymbol{1}_{R_t}}{|R_t|}$$

and by (6.2) the proof of Theorem L follows by corresponding bounds for the intrinsic square function of Theorem I, defined with respect to the tiles $\boldsymbol{T}^{\mathcal{F}}$ given by (4.22).

For Theorem K things are a bit simpler as the decoupling step of (6.2) is not needed. Apart from that one needs to consider for each F a new set of tiles which is very easy to define: If $F \in \mathcal{F}_s$ with $F = \operatorname{rot}_s(I_F \times J_F) + y_F$,

$$\boldsymbol{T}'(F) := \{t = R_t \times F : R_t \in \mathcal{D}^2_{s,\ell_J,\ell_I}\},$$

and then $\boldsymbol{T}' := \bigcup_{F\in\mathcal{F}} \boldsymbol{T}'(F)$. One can recover P°_F by operators of the form

$$\mathsf{P}^\circ_F f := \sum_{t\in\boldsymbol{T}_s(F)} \langle f, \phi_t\rangle \phi_t, \quad |\mathsf{P}^\circ_F f|^2 \lesssim \sum_{t\in\boldsymbol{T}_s(F)} a_t(f)\frac{\boldsymbol{1}_{R_t}}{|R_t|}$$

as before. Using the orthogonality estimates of Section 4.3 in Theorem C yields the upper bound in Theorem K. The optimality of the estimates in the statement of Theorem K is discussed in Section 8.2. □

7. The multiplier problem for the polygon

Let $\mathcal{P}=\mathcal{P}_N$ be a regular N-gon and $T_{\mathcal{P}_N}$ be the corresponding Fourier restriction operator on $\mathcal{P}$

$$T_{\mathcal{P}} f(x) := \int_{\mathbb{R}^2} \hat{f}(\xi)\mathbf{1}_{\mathcal{P}}(\xi)\mathrm{e}^{ix\cdot\xi}\,\mathrm{d}\xi, \quad x\in\mathbb{R}^2.$$

In this subsection we prove Theorem B, namely we will prove the estimate

$$\|T_{\mathcal{P}_N} : L^p(\mathbb{R}^2)\| \lesssim (\log N)^{4\left|\frac{1}{2}-\frac{1}{p}\right|}, \quad \tfrac{4}{3}<p<4.$$

The idea is to reduce the multiplier problem for the polygon to the directional square function estimates of Theorem K and combine those with vector-valued inequalities for directional averages and directional Hilbert transforms.

We introduce some notation. The large integer N is fixed throughout and left implicit in the notation. By scaling, it will be enough to consider a regular polygon $\mathcal{P}$ with the following geometric properties: First, $\mathcal{P}$ has vertices

$$\{v_j=\mathrm{e}^{i\vartheta_j} : 1\le j\le N+1\}, \quad v_j := \exp(2\pi j/N),$$

on the unit circle $\mathbb{S}^1$, with $\vartheta_1=\vartheta_{N+1}=0$ and oriented counterclockwise so that $\vartheta_{j+1}-\vartheta_j>0$. The associated Fourier restriction operator is then defined by

$$T_{\mathcal{P}} f := (\mathbf{1}_{\mathcal{P}}\hat{f})^{\vee}.$$

The proof of the estimate of Theorem B for $T_{\mathcal{P}}$ occupies the remainder of this section; by self-duality of the estimate it will suffice to consider the range $2\le p<4$.

7.1. ***A preliminary decomposition.*** Let N be a large positive integer and take κ such that $2^{\kappa-1}<N\le 2^{\kappa}$. For each $-2\kappa\le k\le 0$ consider a smooth radial multiplier m_k which is supported on the annulus

$$A_k := \left\{\xi\in\mathbb{R}^2 : 1-\frac{2^{-k-1}}{2^{2\kappa}}<|\xi|<1-\frac{2^{-k-5}}{2^{2\kappa}}\right\}$$

and is identically 1 on the smaller annulus

$$a_k := \left\{\xi\in\mathbb{R}^2 : 1-\frac{2^{-k-2}}{2^{2\kappa}}<|\xi|<1-\frac{2^{-k-4}}{2^{2\kappa}}\right\}.$$

Now consider the corresponding radial multiplier operators T_k

$$T_k f := (m_k\hat{f})^{\vee}, \quad m_\kappa := \sum_{k=-2\kappa}^{0} m_k.$$

We note that m_κ is supported in the annulus

$$\left\{\xi\in\mathbb{R}^2 : \frac{1}{2}<|\xi|<1-\frac{2^{-5}}{2^{2\kappa}}\right\}.$$

With this in mind let us consider radial functions $m_0, m_{\mathcal{P}}\in\mathcal{S}(\mathbb{R}^2)$, with $0\le m_0, m_{\mathcal{P}}\le 1$, such that

$$(m_0+m_\kappa+m_{\mathcal{P}})\mathbf{1}_{\mathcal{P}}=\mathbf{1}_{\mathcal{P}}, \tag{7.2}$$

with the additional requirement that

$$\operatorname{supp}(m_{\mathcal{P}}) \subset A_{\mathcal{P}} := \{\xi \in \mathbb{R}^2 : 1 - 2^{-2\kappa-3} \le |\xi| \le 1 + 2^{-2\kappa-3}\}. \tag{7.3}$$

Defining

$$\widehat{T_0 f} := \hat{f} m_0, \quad \widehat{T_\kappa f} := \hat{f} m_\kappa, \quad \widehat{O_{\mathcal{P}} f} := \hat{f} m_{\mathcal{P}} \mathbf{1}_{\mathcal{P}},$$

identity (7.2) implies that $T_{\mathcal{P}} = T_0 + T_\kappa + O_{\mathcal{P}}$. Observing that T_0 is bounded on p for all $1 < p < \infty$ with bounds $O_p(1)$ we have

$$\|T_{\mathcal{P}}\|_{L^p(\mathbb{R}^2)} \lesssim_p 1 + \|T_\kappa\|_{L^p(\mathbb{R}^2)} + \|O_{\mathcal{P}}\|_{L^p(\mathbb{R}^2)}, \quad 1 < p < \infty. \tag{7.4}$$

7.5. ***Estimating T_κ.*** We aim for the estimate

$$\|T_\kappa f\|_p \lesssim \kappa^{4(\frac{1}{2}-\frac{1}{p})} \|f\|_p, \quad 2 \le p < 4. \tag{7.6}$$

The case $p = 2$ is obvious, whence it suffices to prove the restricted-type version at the endpoint $p = 4$

$$\|T_\kappa(f \mathbf{1}_E)\|_4 \lesssim \kappa |E|^{\frac{1}{4}} \|f\|_\infty. \tag{7.7}$$

Now we have that for any g

$$|T_\kappa g| = \left| \sum_{k=-2\kappa}^{0} T_k g \right| \lesssim \left(\sum_{k=-2\kappa}^{0} |T_k g|^4 \right)^{\frac{1}{4}} \kappa^{\frac{3}{4}}$$

and thus

$$\|T_\kappa g\|_4 \lesssim \kappa^{\frac{3}{4}} \left(\sum_{k=-2\kappa}^{0} \|T_k g\|_4^4 \right)^{\frac{1}{4}}. \tag{7.8}$$

Let $\{\omega_j : j \in J\}$ be the collection of intervals on $\mathbb{S}^1$ centered at $v_j := \exp(2\pi i j/N)$ and of length $2^{-\kappa}$. Note that these intervals have finite overlap and their centers v_j form a $\sim 1/N$-net on $\mathbb{S}^1$. Now let $\{\beta_j : j \in J\}$ be a smooth partition of unity subordinated to the finitely overlapping open cover $\{\omega_j : j \in J\}$ so that each β_j is supported in ω_j. We can decompose each T_k as

$$(\widehat{T_k f})(\xi) = \sum_{j \in J} m_k(|\xi|) \beta_j \left(\frac{\xi}{|\xi|} \right) \hat{f}(\xi) =: \sum_{j \in J} m_{j,k}(\xi) \hat{f}(\xi), =: \sum_{j \in J} (\widehat{T_{j,k} f})(\xi), \quad \xi \in \mathbb{R}^2.$$

For $s_j \in S$ and $-2\kappa \le k \le 0$ we define the conical sectors

$$\Omega_{j,k} := \{\xi \in \mathbb{R}^2 : \xi \in A_k, \, \xi/|\xi| \in \omega_j\}$$

and note that each one of the multipliers $m_{j,k}$ is supported in $\Omega_{j,k}$. Each $\Omega_{j,k}$ is an annular sector around the circle of radius $1 - 2^{-k}/2^{2\kappa}$ of width $\sim 2^{-k}/2^{2\kappa}$, where $-2\kappa \le k \le 0$. It is a known observation, usually attributed to Córdoba [1977, Theorem 2] or C. Fefferman [1973], that for such parameters we have

$$\sum_{j,j' \in J} \mathbf{1}_{\Omega_{j,k} + \Omega_{j',k}} \lesssim 1. \tag{7.9}$$

This pointwise inequality and Plancherel's theorem allow us to decouple the pieces $T_{j,k}$ in L^4; for each fixed k as above we have

$$\|T_k f\|_4 \lesssim \Big\| \Big(\sum_{j\in J} |T_{j,k} f|^2 \Big)^{\frac12} \Big\|_4; \tag{7.10}$$

see also the proof of Lemma 7.18 below for a vector-valued version of this estimate. Combining the last estimate with (7.8) and dominating the ℓ^2-norm by the ℓ^1-norm yields

$$\begin{aligned}\|T_\kappa f\|_4 &\lesssim \kappa^{\frac34} \Big(\int_{\mathbb{R}^2} \sum_{k=-2\kappa}^{0} \Big(\sum_{j\in J} |T_{j,k} f|^2 \Big)^2 \Big)^{\frac14} \lesssim \kappa^{\frac34} \Big(\int_{\mathbb{R}^2} \Big[\sum_{k=-2\kappa}^{0} \Big(\sum_{j\in J} |T_{j,k} f|^2 \Big)^2 \Big]^{\frac12 2} \Big)^{\frac14} \\ &\le \kappa^{\frac34} \Big(\int_{\mathbb{R}^2} \Big[\sum_{k=-2\kappa}^{0} \sum_{j\in J} |T_{j,k} f|^2 \Big]^2 \Big)^{\frac14} =: \kappa^{\frac34} \|\Delta_{J,\kappa} f\|_4,\end{aligned}$$

with

$$\Delta_{J,\kappa} f := \Big(\sum_{k=-2\kappa}^{0} \sum_{j\in J} |T_{j,k} f|^2 \Big)^{\frac12}.$$

But now note that $\{T_{j,k}\}_{j,k}$ is a finitely overlapping family of smooth frequency projections on a family of rectangles in at most $\sim N$ directions. Furthermore all these rectangles have one side of fixed length since $|\omega_j| = 2^{-\kappa}$ for all $j \in J$. So Theorem K with the improvement of Remark 6.1 applies to yield

$$\|\Delta_{J,\kappa} f\|_4 \lesssim (\log \#N)^{\frac14} \|f\|_\infty |E|^{\frac14} \simeq \kappa^{\frac14} \|f\|_\infty |E|^{\frac14}. \tag{7.11}$$

The last two displays establish (7.7) and thus (7.6).

Remark 7.12. The term T_κ is also present in the argument of [Córdoba 1977]. Therein, an upper estimate of order $O(\kappa^{5/4})$ for p near 4 is obtained, by using the triangle inequality and the bound $\sup\{\|T_k\|_{L^4(\mathbb{R}^2)} : -2\kappa \le k \le 0\} \sim \kappa^{1/4}$ for the smooth restriction to a single annulus.

7.13. ***Estimating $O_{\mathcal{P}}$.*** In this subsection we will prove the estimate

$$\|O_{\mathcal{P}} f\|_p \lesssim \kappa^{4(\frac12 - \frac1p)} \|f\|_p. \tag{7.14}$$

Let Φ be a smooth radial function with support in the annular region $\{\xi \in \mathbb{R}^2 : 1 - c2^{-2\kappa} < |\xi| < 1 + c2^{-2\kappa}\}$, where c is a fixed small constant, and satisfying $0 \le \Phi \le 1$. Let $\{\beta_j : j \in J\}$ be a partition of unity on $\mathbb{S}^1$ relative to intervals ω_j as in Section 7.5. Define the Fourier multiplier operators on $\mathbb{R}^2$

$$\widehat{T_j f}(\xi) := \Phi(\xi) \beta_j \Big(\frac{\xi}{|\xi|} \Big) \hat{f}(\xi), \quad \xi \in \mathbb{R}^2. \tag{7.15}$$

The operators T_j satisfy a square function estimate

$$\begin{aligned} \|\{T_j f\}\|_{L^p(\mathbb{R}^2;\ell^2_J)} &\lesssim \kappa^{\frac12 - \frac1p} \|f\|_p, \quad 2 \le p < 4, \\ \|\{T_j (f \mathbf{1}_E)\}\|_{L^4(\mathbb{R}^2;\ell^2_J)} &\lesssim \kappa^{\frac14} |E|^{\frac14} \|f\|_\infty, \end{aligned} \tag{7.16}$$

which follows in the same way as (7.11), by using Theorem K with the improvement of Remark 6.1. They also obey a vector-valued estimate

$$\begin{aligned} \|\{T_j f_j\}\|_{L^p(\mathbb{R}^2;\ell^2_J)} &\lesssim \kappa^{\frac{1}{2}-\frac{1}{p}} \|\{f_j\}\|_{L^p(\mathbb{R}^2;\ell^2_J)}, \quad 2 \le p < 4, \\ \|\{T_j(f_j \mathbf{1}_F)\}\|_{L^4(\mathbb{R}^2;\ell^2_J)} &\lesssim \kappa^{\frac{1}{4}} |F|^{\frac{1}{4}} \|\{f_j\}\|_{L^\infty(\mathbb{R}^2;\ell^2_J)}. \end{aligned} \tag{7.17}$$

These estimates are easy to prove. Indeed note that it suffices to prove the endpoint-restricted estimate at $p = 4$. Using the Fefferman–Stein inequality for fixed $j \in J$ we can estimate for each function g with $\|g\|_2 = 1$

$$\begin{aligned} \int_{\mathbb{R}^2} \sum_{j\in J} |T_j(f_j \mathbf{1}_F)|^2 g &\lesssim \sum_{j\in J} \int_{\mathbb{R}^2} |f_j \mathbf{1}_F|^2 \mathrm{M}_j g \le \|\{f_j\}\|^2_{L^\infty(\mathbb{R}^2;\ell^2_J)} \int_F \sup_{j\in J} \mathrm{M}_j g \\ &\lesssim |F|^{\frac{1}{2}} \|\sup_{j\in J} \mathrm{M}_j g\|_{L^{2,\infty}(\mathbb{R}^2)}, \end{aligned}$$

where M_j is the Hardy–Littlewood maximal operator with respect to the collection of parallelograms in $\mathcal{D}^2_{s_j,-2\kappa,-\kappa}$ with s_j defined through $(-s_j, 1) := v_j$. Now $\sup_{j\in J} M_j$ is the maximal directional maximal operator and the number of directions involved in its definition is comparable to $N \sim 2^\kappa$. Then the maximal theorem from [Katz 1999] applies to give the estimate

$$\|\sup_{j\in J} \mathrm{M}_j g\|_{L^{2,\infty}(\mathbb{R}^2)} \lesssim \kappa^{\frac{1}{2}}.$$

This proves the second of the estimates (7.17) and thus both of them by interpolation.

In the estimate for $O_{\mathcal{P}}$ we will also need the following decoupling result.

Lemma 7.18. *Let $2 \le p < 4$. Then*

$$\Big\|\sum_j T_j f_j\Big\|_p \lesssim \kappa^{\frac{1}{2}-\frac{1}{p}} \|\{f_j\}\|_{L^p(\mathbb{R}^2;\ell^2_J)}.$$

Proof. Note that the case $p = 2$ of the conclusion is trivial due to the finite overlap of the supports of the multipliers of the operators T_j. Thus by vector-valued restricted-type interpolation of the operator

$$\{f_j\} \mapsto O(\{f_j\}) := \sum_{j\in J} T_j f_j$$

it suffices to prove a restricted type $L^{4,1} \to L^4$ estimate:

$$\|O(\{f_j\})\|_4 \lesssim \kappa^{\frac{1}{4}} |E|^{\frac{1}{4}} \tag{7.19}$$

for functions with $\|\{f_j\}\|_{\ell^2} \le \mathbf{1}_E$. To do so note that the finite overlap of the supports of $\widehat{T_j f_j} * \widehat{T_k f_k}$ over j, k, as in (7.9), gives

$$\|O(\{f_j\})\|_4 \lesssim \|\{T_j f_j\}\|_{L^4(\mathbb{R}^2;\ell^2_J)}$$

and the restricted-type estimate (7.19) follows from (7.17). □

We come to the main argument for $O_{\mathcal{P}}$. Let $m_{\mathcal{P}}$ be as in (7.2)–(7.3) and T_j be the multiplier operators from (7.15) corresponding to the choice $\Phi = m_{\mathcal{P}}$. Then obviously

$$m_{\mathcal{P}} \hat{f} = \sum_{j\in J} \widehat{T_j f}.$$

We may also tweak Φ and the partition of unity on $\mathbb{S}^1$ to obtain further multiplier operators $\widetilde{T}_j$ as in (7.15) and such that the Fourier transform of the symbol of $\widetilde{T}_j$ equals 1 on the support of the symbol of T_j. With these definitions in hand we estimate for $2 < p < 4$

$$\begin{aligned}\|O_{\mathcal{P}} f\|_p = \Big\|\sum_j \widetilde{T}_j(T_j T_{\mathcal{P}} f)\Big\|_p &\lesssim \kappa^{\frac{1}{2}-\frac{1}{p}} \|\{T_{\mathcal{P}}(T_j f)\}\|_{L^p(\mathbb{R}^2;\ell^2_J)} \\ &= \kappa^{\frac{1}{2}-\frac{1}{p}} \|\{H_j H_{j+1}(T_j f)\}\|_{L^p(\mathbb{R}^2;\ell^2_J)}. \end{aligned} \tag{7.20}$$

The first inequality is an application of Lemma 7.18 for $\widetilde{T}_j$. The last equality is obtained by observing that the polygon multiplier $T_{\mathcal{P}}$ on the support of each T_j may be written as a (sum of $O(1)$) directional biparameter multipliers $H_j H_{j+1}$ of iterated Hilbert transform type, where H_j is a Hilbert transform along the direction ν_j, which is the unit vector perpendicular to the j-th side of the polygon, and pointing inside the polygon; these are at most $\sim N$ such directions.

In order to complete our estimate for $O_{\mathcal{P}}$ we need the following Meyer-type lemma for directional Hilbert transforms of the form

$$H_v f(x) := \int_{\mathbb{R}^2} \hat{f}(\xi) \mathbf{1}_{\{\xi\cdot v>0\}} e^{ix\cdot\xi}\, \mathrm{d}\xi, \qquad x \in \mathbb{R}^2.$$

Lemma 7.21. *Let $V \subset \mathbb{S}^1$ be a finite set of directions and H_v be the Hilbert transform in the direction v. Then for $\frac{4}{3} < p < 4$ we have*

$$\|\{H_v f_v\}\|_{L^p(\mathbb{R}^2;\ell^2_V)} \lesssim (\log \# V)^{\left|\frac{1}{2}-\frac{1}{p}\right|} \|\{f_v\}\|_{L^p(\mathbb{R}^2;\ell^2_V)}.$$

The dependence on $\# V$ is best possible.

Proof. It suffices to prove the estimate for $2 < p < 4$. The proof is by way of duality and uses the following inequality for the Hilbert transform: for $r > 1$ and w a nonnegative locally integrable function we have

$$\int_{\mathbb{R}^2} |H_v f|^2 w \lesssim \int_{\mathbb{R}^2} |f|^2 (\mathrm{M}_v |w|^r)^{\frac{1}{r}},$$

with M_v given by (2.3). See for example [Pérez 1994]. Using this we have for a suitable $g \in L^{(p/2)'}$ of norm 1 that

$$\begin{aligned}\|\{H_v f_v\}\|^2_{L^p(\mathbb{R}^2;\ell^2_V)} = \int_{\mathbb{R}^2} \sum_{v\in V} |H_v f_v|^2 g &\lesssim \sum_{v\in V} \int_{\mathbb{R}^2} |f_v|^2 (\mathrm{M}_v |g|^r)^{\frac{1}{r}} \\ &\lesssim \|\{f_v\}\|^2_{L^p(\mathbb{R}^2;\ell^2_V)} \|(\mathrm{M}_V |g|^r)^{\frac{1}{r}}\|_{L^{(p/2)'}(\mathbb{R}^2)},\end{aligned}$$

with $\mathrm{M}_V g := \sup_{v\in V} \mathrm{M}_v g$. Now for $2 < p < 4$ there is a choice of $1 < r < p/(2(p-2))$ so that $p/(r(p-2)) > 2$. This means that the maximal theorem from [Katz 1999] applies again to give

$$\|(\mathrm{M}_V |g|^r)^{\frac{1}{r}}\|_{L^{(p/2)'}(\mathbb{R}^2)} \lesssim (\log \# V)^{1-\frac{2}{p}},$$

and so the proof of the upper bound is complete. The optimality is discussed in Section 8.1. □

Let us now go back to the estimate for $O_{\mathcal{P}}$. The left-hand side of (7.20) contains a double Hilbert transform. By an iterated application of Lemma 7.21 we thus have

$$\|\{H_j H_{j+1}(T_j f)\}\|_{L^p(\mathbb{R}^2;\ell^2_J)} \lesssim \kappa^{1-\frac{2}{p}} \|\{T_j f)\}\|_{L^p(\mathbb{R}^2;\ell^2_J)}$$

since the number of directions is $N = 2^\kappa$. The final estimate for the right-hand side of the display above is a direct application of (7.16), which together with (7.20) yields the estimate for $\|O_{\mathcal{P}} f\|_p$ claimed in (7.14).

Now the decomposition (7.4), together with the estimate of Section 7.5 for T_κ and the estimate (7.14) for $O_{\mathcal{P}}$, completes the proof of Theorem B.

Remark 7.22. Consider a function f in $\mathbb{R}^2$ such that $\operatorname{supp}(\hat{f}) \subseteq A_\delta$, where A_δ is an annulus of width δ^2 around $\mathbb{S}^1$. Decomposing A_δ into a union of $O(1/\delta)$ finitely overlapping annular boxes of radial width δ^2 and tangential width δ, we can write $f = \sum_{j \in J} T_j f$, where each T_j is a smooth frequency projection onto one of these annular boxes, indexed by j. Then if $\widetilde{T}_j$ is a multiplier operator whose symbol is identically 1 on the frequency support of $T_j f$ and supported on a slightly larger box, we can write $f = \sum_j \widetilde{T}_j T_j f$, as in (7.20) above. Then Lemma 7.18 yields

$$\|f\|_{L^p(\mathbb{R}^2)} \lesssim (\log(1/\delta))^{\frac{1}{2}-\frac{1}{p}} \|\{T_j f\}\|_{L^p(\mathbb{R}^2;\ell^2_J)}.$$

This is the inverse square function estimate claimed in the remark after Theorem B in Section 1.

8. Lower bounds and concluding remarks

8.1. ***Sharpness of Meyer's lemma.*** We briefly sketch the quantitative form of Fefferman's counterexample [1971] proving the sharpness of Lemma 7.21. Let N be a large dyadic integer. Using a standard Besicovitch-type construction we produce rectangles $\{R_j : j = 1, \dots, N\}$ with sidelengths $1 \times 1/N$, so that the long side of R_j is oriented along $v_j := \exp(2\pi i j/N)$. Now we consider the set E to be the union of these rectangles and

$$\left| E := \bigcup_{j=1}^{N} R_j \right| \lesssim \frac{1}{\log N}.$$

Denoting by $\widetilde{R}_j$ the 2-translate of R_j in the direction of v_j we gather that $\{\widetilde{R}_j : j = 1, \dots, N\}$ is a pairwise disjoint collection. Furthermore if H_j is the Hilbert transform in direction v_j, there holds

$$|H_j \mathbf{1}_{R_j}| \geq c \mathbf{1}_{\widetilde{R}_j}.$$

Therefore for all $1 < p < \infty$

$$\left\| \left(\sum_{j=1}^{N} |H_j \mathbf{1}_{R_j}|^2 \right)^{\frac{1}{2}} \right\|_p \geq c \left| \bigcup_{j=1}^{N} \widetilde{R}_j \right|^{\frac{1}{p}} \geq c,$$

while for $p \leq 2$

$$\left\| \left(\sum_{j=1}^{N} |\mathbf{1}_{R_j}|^2 \right)^{\frac{1}{2}} \right\|_p \leq \left(\sum_{j=1}^{N} |R_j| \right)^{\frac{1}{2}} |E|^{\frac{1}{p}-\frac{1}{2}} \lesssim (\log N)^{\frac{1}{2}-\frac{1}{p}}.$$

Self-duality of the square function estimate then gives the optimality of the estimate of Lemma 7.21.

8.2. ***Sharpness of the directional square function bound.*** In this subsection we prove that the bound of Theorem L is best possible, up to the doubly logarithmic terms. In particular we prove that the bound of Remark 6.1 is best possible.

We begin by showing a lower bound for the rough square function estimate

$$\|\{P_F g\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \le \|\{P_F\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})\| \|g\|_p, \quad 2 \le p < 4, \tag{8.3}$$

where the notation is as in Section 6. Now as in [Fefferman 1971] one can easily show that the estimate above implies the vector-valued inequality for directional averages, for directions corresponding to the directions of rectangles in $\mathcal{F}$. For this let $\#V = N$, where V is the set of directions of rectangles in $\mathcal{F}$. Now consider functions $\{g_F\}_{F\in\mathcal{F}}$ with compact Fourier support; by modulating these functions we can assume that $\operatorname{supp}(\hat{g}_F) \subset B(c_F, A)$ for some $A > 1$ and $\{c_F\}_{F\in\mathcal{F}}$ a $100AN$-net in $\mathbb{R}^2$. Then if F is a rectangle centered at c_F with short side 1 parallel to a direction $v_F \in V$ and long side of length N parallel to $v_F^\perp$, then we have that $|P_F g_F| = |A_{v_F} g_F|$, where A_{v_F} is the averaging operator

$$A_{v_F} f(x) := 2N \int_{|t|\le 1/2} \int_{N|s|<1} f(x - t v_F - s v_F^\perp)\, \mathrm{d}t\, \mathrm{d}s, \quad x \in \mathbb{R}^2.$$

Note that this is a single-scale average with respect to rectangles of dimensions $1 \times 1/N$ in the directions $v_F, v_F^\perp$ respectively. Since the frequency supports of these functions are well-separated we gather that for all choices of signs $\varepsilon_F \in \{-1, 1\}$ we have

$$\sum_{T\in\mathcal{F}} |P_T G|^2 := \sum_{T\in\mathcal{F}} \Big| P_T\Big(\sum_{F\in\mathcal{F}} \varepsilon_F g_F\Big)\Big|^2 = \sum_{T\in\mathcal{F}} |P_T g_T|^2.$$

Thus applying (8.3) with the function G as above and averaging over random signs we get

$$\|\{A_{v_F} g_F\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \le \|\{P_F\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})\| \|\{g_F\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})}, \quad 2 \le p < 4.$$

Now we just need to note that as in Section 8.1 we have that

$$A_{v_F} \mathbf{1}_{R_F} \gtrsim \mathbf{1}_{\tilde{R}_F},$$

where $\{R_F\}_{F\in\mathcal{F}}$ are the rectangles used in the Besicovitch construction in Section 8.1. As before we get

$$\|\{P_F\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})\| \gtrsim (\log \#V)^{\frac{1}{2}-\frac{1}{p}}.$$

For $p < 2$ the square function estimate (8.3) is known to fail even in the case of a single directions; see for example the counterexample in [Rubio de Francia 1985, §1.5].

One can use the same argument in order to show a lower bound for the norm of the smooth square function

$$\|\{P_F^\circ g\}\|_{L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})} \le \|\{P_F^\circ\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2;\ell^2_{\mathcal{F}})\| \|g\|_p, \quad 2 \le p < 4.$$

Indeed, following the exact same steps we can deduce a vector-valued inequality for smooth averages

$$A_{v_F}^\circ f(x) := \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - t v_F - s v_F^\perp) \gamma_F(t, s)\, \mathrm{d}t\, \mathrm{d}s, \quad x \in \mathbb{R}^2,$$

where γ_F is the smooth product bump function used in the definition of P_F° in Section 6. By a direct computation one easily shows the analogous lower bound $A_{v_F}^\circ \mathbf{1}_{R_F} \gtrsim \mathbf{1}_{\tilde{R}_F}$ for the rectangles of the Besicovitch construction and this completes the proof of the lower bound for smooth projections as well.

8.4. ***Sharpness of Córdoba's bound for radial multipliers.*** Firstly we remember the definition of each radial multiplier P_δ: Let $\Phi : \mathbb{R} \to \mathbb{R}$ be a smooth function which is supported in $[-1, 1]$ and define

$$P_\delta f(x) := \int_{\mathbb{R}^2} \hat{f}(\xi)\Phi(\delta^{-1}(1-|\xi|))\mathrm{e}^{ix\cdot\xi}\,\mathrm{d}\xi, \quad x \in \mathbb{R}^2.$$

These smooth radial multipliers were used extensively in Section 7. Córdoba [1979] proved the bound

$$\|P_\delta f\|_p \lesssim (\log 1/\delta)^{\left|\frac{1}{2}-\frac{1}{p}\right|}\|f\|_p, \quad \tfrac{3}{4} \le p \le 4.$$

In fact the same bound is implicitly proved in Section 7 in a more refined form, but only in the open range $p \in \left(\frac{3}{4}, 4\right)$ with weak-type analogues at the endpoints. More precisely we have discretized P_δ into a sum of pieces $\{P_{\delta,j}\}_{j\in J}$, where each $P_{\delta,j}$ is a smooth projection onto an annular box of width δ and length $\sqrt{\delta}$, pointing along one of N equispaced directions v_j. Then it follows from the considerations in Section 7 that

$$\begin{aligned} \|\{P_{\delta,j} f\}\|_{L^p(\mathbb{R}^2;\ell^2_j)} &\lesssim \log(1/\delta)^{\frac{1}{2}-\frac{1}{p}}\|f\|_p, \quad 2 < p < 4, \\ \|\{P_{\delta,j} f\mathbf{1}_F\}\|_{L^4(\mathbb{R}^2;\ell^2_j)} &\lesssim \log(1/\delta)^{\frac{1}{4}}\|f\|_\infty |F|^{\frac{1}{4}}. \end{aligned} \tag{8.5}$$

Obviously one gets the same bound by duality for $\frac{4}{3} < p < 2$, while the L^2-bound is trivial. Now these estimates imply Córdoba's estimate for P_δ in the open range $\left(\frac{3}{4}, 4\right)$ by the decoupling inequality (7.10), also due to Córdoba. On the other hand Córdoba's estimate is sharp. Indeed one uses the same rescaling and modulation arguments as in the previous subsection in order to deduce a vector-valued inequality for smooth averages starting by Córdoba's estimate. Testing this vector-valued estimate against the rectangles of the Besicovitch construction proves the familiar lower bound for P_δ and thus also shows the optimality of the estimates in (8.5). We omit the details.

8.6. ***Lower bounds for the conical square function.*** We conclude this section with a simple example that provides a lower bound for the operator norm of the conical square function $\|C_\omega(f) : \ell^2_{\boldsymbol{\omega}}\|$ of Theorem J and the smooth conical square function $\|C_\omega^\circ : \ell^2_{\boldsymbol{\omega}}\|$ of Theorem A. The considerations in this subsection also rely on the Besicovitch construction so we adopt again the notation of Section 8.1 for the rectangles $\{R_j : 1 \le j \le N\}$ and their union E. Let H_j^+ denote the frequency projection in the half-space $\{\xi \in \mathbb{R}^2 : \xi \cdot v_j > 0\}$, where $v_j := \exp(2\pi i j/N)$. We begin by observing that

$$H_j^+ f - H_{j+1}^+ f = C_j P_+ f - C_j P_- f, \tag{8.7}$$

where P_+, P_- denote the rough frequency projections in the upper and lower half-space respectively and C_{v_j} is the multiplier associated with the cone bordered by v_j, v_{j+1}. Since H_j^+ is a linear combination of the identity with the usual directional Hilbert transform H_j along v_j we conclude that

$$\left\|\left(\sum_{j=1}^N |(H_{j+1}-H_j)f|^2\right)^{\frac{1}{2}}\right\|_p \lesssim \|\{C_j\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2;\ell^2_j)\|\|f\|_p, \quad 2 \le p < 4.$$

Now note that for each fixed $1 \le k \le N$ we have

$$\mathbf{1}_{\tilde{R}_k} \sum_j (H_j - H_{j+1}) \mathbf{1}_{R_j} = \mathbf{1}_{\tilde{R}_k} H_k \mathbf{1}_{R_k} \gtrsim \mathbf{1}_{\tilde{R}_k} \tag{8.8}$$

if $\widetilde{R}_k$ is a sufficiently large translation of R_k in the positive direction v_k. Thus

$$\left| \int \mathbf{1}_{\bigcup_k \tilde{R}_k} \sum_{j=1}^{N} (H_{j+1} - H_j) \mathbf{1}_{R_j} \right| \gtrsim \left| \sum_k \int_{\tilde{R}_k} \mathbf{1}_{\tilde{R}_k} \right| \simeq 1.$$

On the other hand the left-hand side of the display above is bounded by a constant multiple of

$$\|\{C_j\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2; \ell_j^2)\| \left\| \left(\sum_j \mathbf{1}_{R_j}^2 \right)^{\frac{1}{2}} \right\|_{p'} \lesssim \|C_V : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2; \ell^2)\| (\log N)^{\frac{1}{2} - \frac{1}{p'}}$$

for all $2 \le p < 4$. We thus conclude that

$$\|\{C_j\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2; \ell^2)\| \gtrsim (\log N)^{\frac{1}{2} - \frac{1}{p}}, \quad 2 \le p < 4.$$

We explain how this counterexample can be modified to get a lower square function estimate for the smooth cone multipliers C_ω° from (5.1) matching the upper bound of Theorem A. For $t \in \mathbb{R}$ write $v_j^t := \exp(2\pi i (j+t)/N)$ and let H_j^t and $H_j^{t,+}$ be the directional Hilbert transform and analytic projection along v_j^t, respectively. Let $\delta > 0$ be a small parameter to be chosen later and for each $1 \le j \le N$ let ω_j be an interval of size δN^{-1} centered around $2\pi j/N$. Arguing as in (8.7),

$$C_{\omega_j}^\circ P_+ f - C_{\omega_j}^\circ P_- f = \fint_{N|t|<\delta} \alpha\left(\frac{Nt}{\delta}\right) (H_j^{t,+} f - H_{j+1}^{t,+} f) \, \mathrm{d}t$$

for a suitable nonnegative averaging function α which equals 1 on $\left[-\frac{1}{4}, \frac{1}{4}\right]$. Now, if $\widetilde{R}_k$ is again a sufficiently large translation of R_k in the positive direction v_k and δ is chosen sufficiently small depending only on the translation amount, the analogue of (8.8) is

$$\mathbf{1}_{\tilde{R}_k} \inf_{N|t|<\delta} \sum_{j=1}^{N} (H_j^t - H_{j+1}^t) \mathbf{1}_{R_j} = \mathbf{1}_{\tilde{R}_k} \inf_{N|t|<\delta} H_k^t [\mathbf{1}_{R_k}] \gtrsim \mathbf{1}_{\tilde{R}_k}.$$

The lower bound for $\|\{C_{\omega_j}\} : L^p(\mathbb{R}^2) \to L^p(\mathbb{R}^2; \ell_j^2)\|$ then follows exactly as in the previous case.

Acknowledgements

The authors are grateful to Ciprian Demeter and Jongchon Kim for fruitful discussions on reverse square function estimates, and for providing additional references on the subject.

Accomazzo is partially supported by the project PGC2018-094528-B-I00 (AEI/FEDER, UE) with acronym "IHAIP", project MTM-2017-82160-C2-2-P of the Ministerio de Economía y Competitividad (Spain), and grant T1247-19 of the Basque Government. Di Plinio is partially supported by the National Science Foundation under the grants DMS-1800628, DMS-2000510, NSF-DMS-2054863. Hagelstein is partially supported by a grant from the Simons Foundation (#521719 to Paul Hagelstein).

Parissis is partially supported by the project PGC2018-094528-B-I00 (AEI/FEDER, UE) with acronym "IHAIP", grant T1247-19 of the Basque Government and IKERBASQUE. Roncal is supported by the Basque Government through the BERC 2018-2021 program, by the Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2017-2018 and through project PID2020-113156GB-I00. She also acknowledges the RyC project RYC2018-025477-I and IKERBASQUE.

References

[Barrionuevo and Lacey 2003] J. Barrionuevo and M. T. Lacey, "A weak-type orthogonality principle", *Proc. Amer. Math. Soc.* **131**:6 (2003), 1763–1769. MR Zbl

[Bateman 2013] M. Bateman, "Maximal averages along a planar vector field depending on one variable", *Trans. Amer. Math. Soc.* **365**:8 (2013), 4063–4079. MR Zbl

[Bennett and Harrison 2012] J. Bennett and S. Harrison, "Weighted norm inequalities for oscillatory integrals with finite type phases on the line", *Adv. Math.* **229**:4 (2012), 2159–2183. MR Zbl

[Bourgain and Demeter 2015] J. Bourgain and C. Demeter, "The proof of the l^2 decoupling conjecture", *Ann. of Math.* (2) **182**:1 (2015), 351–389. MR Zbl

[Cabrelli et al. 2006] C. Cabrelli, M. T. Lacey, U. Molter, and J. C. Pipher, "Variations on the theme of Journé's lemma", *Houston J. Math.* **32**:3 (2006), 833–861. MR Zbl

[Carbery 2015] A. Carbery, "A remark on reverse Littlewood–Paley, restriction and Kakeya", preprint, 2015. arXiv 1507.02515

[Christ et al. 1986] M. Christ, J. Duoandikoetxea, and J. L. Rubio de Francia, "Maximal operators related to the Radon transform and the Calderón–Zygmund method of rotations", *Duke Math. J.* **53**:1 (1986), 189–209. MR Zbl

[Córdoba 1977] A. Córdoba, "The multiplier problem for the polygon", *Ann. of Math.* (2) **105**:3 (1977), 581–588. MR

[Córdoba 1979] A. Córdoba, "A note on Bochner–Riesz operators", *Duke Math. J.* **46**:3 (1979), 505–511. MR Zbl

[Córdoba 1982] A. Córdoba, "Geometric Fourier analysis", *Ann. Inst. Fourier* (*Grenoble*) **32**:3 (1982), 215–226. MR Zbl

[Demeter 2010] C. Demeter, "Singular integrals along N directions in $\mathbb{R}^2$", *Proc. Amer. Math. Soc.* **138**:12 (2010), 4433–4442. MR Zbl

[Demeter and Di Plinio 2014] C. Demeter and F. Di Plinio, "Logarithmic L^p bounds for maximal directional singular integrals in the plane", *J. Geom. Anal.* **24**:1 (2014), 375–416. MR Zbl

[Di Plinio and Parissis 2021] F. Di Plinio and I. Parissis, "Maximal directional operators along algebraic varieties", *Amer. J. Math.* **143**:5 (2021), 1463–1503. MR Zbl

[Di Plinio et al. 2018] F. Di Plinio, S. Guo, C. Thiele, and P. Zorin-Kranich, "Square functions for bi-Lipschitz maps and directional operators", *J. Funct. Anal.* **275**:8 (2018), 2015–2058. MR Zbl

[Fefferman 1971] C. Fefferman, "The multiplier problem for the ball", *Ann. of Math.* (2) **94** (1971), 330–336. MR Zbl

[Fefferman 1973] C. Fefferman, "A note on spherical summation multipliers", *Israel J. Math.* **15** (1973), 44–52. MR Zbl

[Garrigós and Seeger 2010] G. Garrigós and A. Seeger, "A mixed norm variant of Wolff's inequality for paraboloids", pp. 179–197 in *Harmonic analysis and partial differential equations* (El Escorial, Spain, 2008), edited by P. Cifuentes et al., Contemp. Math. **505**, Amer. Math. Soc., Providence, RI, 2010. MR

[de Guzmán 1981] M. de Guzmán, *Real variable methods in Fourier analysis*, Notas Mat. **75**, North-Holland, Amsterdam, 1981. MR Zbl

[Hytönen et al. 2013] T. P. Hytönen, M. T. Lacey, and C. Pérez, "Sharp weighted bounds for the q-variation of singular integrals", *Bull. Lond. Math. Soc.* **45**:3 (2013), 529–540. MR Zbl

[Journé 1985] J.-L. Journé, "Calderón–Zygmund operators on product spaces", *Rev. Mat. Iberoam.* **1**:3 (1985), 55–91. MR Zbl

[Katz 1999] N. H. Katz, "Maximal operators over arbitrary sets of directions", *Duke Math. J.* **97**:1 (1999), 67–79. MR Zbl

[Łaba and Pramanik 2006] I. Łaba and M. Pramanik, "Wolff's inequality for hypersurfaces", *Collect. Math.* **2006**:Extra (2006), 293–326. MR Zbl

[Łaba and Wolff 2002] I. Łaba and T. Wolff, "A local smoothing estimate in higher dimensions", *J. Anal. Math.* **88** (2002), 149–171. MR Zbl

[Lacey 2007] M. T. Lacey, *Issues related to Rubio de Francia's Littlewood–Paley inequality*, New York J. Math. Monogr. **2**, State Univ. New York, Albany, 2007. MR Zbl

[Lacey and Li 2006] M. T. Lacey and X. Li, "Maximal theorems for the directional Hilbert transform on the plane", *Trans. Amer. Math. Soc.* **358**:9 (2006), 4099–4117. MR Zbl

[Lacey and Li 2010] M. Lacey and X. Li, *On a conjecture of E. M. Stein on the Hilbert transform on vector fields*, Mem. Amer. Math. Soc. **965**, Amer. Math. Soc., Providence, RI, 2010. MR Zbl

[Nagel et al. 1978] A. Nagel, E. M. Stein, and S. Wainger, "Differentiation in lacunary directions", *Proc. Natl. Acad. Sci. USA* **75**:3 (1978), 1060–1062. MR Zbl

[Parcet and Rogers 2015] J. Parcet and K. M. Rogers, "Directional maximal operators and lacunarity in higher dimensions", *Amer. J. Math.* **137**:6 (2015), 1535–1557. MR Zbl

[Pérez 1994] C. Pérez, "Weighted norm inequalities for singular integral operators", *J. Lond. Math. Soc.* (2) **49**:2 (1994), 296–308. MR Zbl

[Pramanik and Seeger 2007] M. Pramanik and A. Seeger, "L^p regularity of averages over curves and bounds for associated maximal operators", *Amer. J. Math.* **129**:1 (2007), 61–103. MR Zbl

[Rubio de Francia 1985] J. L. Rubio de Francia, "A Littlewood–Paley inequality for arbitrary intervals", *Rev. Mat. Iberoam.* **1**:2 (1985), 1–14. MR Zbl

[Sjögren and Sjölin 1981] P. Sjögren and P. Sjölin, "Littlewood–Paley decompositions and Fourier multipliers with singularities on certain sets", *Ann. Inst. Fourier* (*Grenoble*) **31**:1 (1981), 157–175. MR Zbl

Received 28 Jun 2021. Revised 9 Dec 2021. Accepted 24 Jan 2022.

NATALIA ACCOMAZZO: naccomazzo@math.ubc.ca
Department of Mathematics, University of British Columbia, Vancouver, BC, Canada

FRANCESCO DI PLINIO: francesco.diplinio@unina.it
Dipartimento di Matematica e Applicazioni, Università di Napoli, Napoli, Italy

PAUL HAGELSTEIN: paul_hagelstein@baylor.edu
Department of Mathematics, Baylor University, Waco, TX, United States

IOANNIS PARISSIS: ioannis.parissis@ehu.es
Departamento de Matemáticas, Universidad del País Vasco, Bilbao, Spain

and

Ikerbasque, Basque Foundation for Science, Bilbao, Spain

LUZ RONCAL: lroncal@bcamath.org
BCAM - Basque Center for Applied Mathematics, Bilbao, Spain

and

Ikerbasque, Basque Foundation for Science, Bilbao, Spain

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 16 No. 7 2023