6-26-2024

Some Problems in Harmonic Analysis

Anastasios Fragkos
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Mathematics

Dissertation Examination Committee:

Brett Wick, Chair

Alan Chang

Loukas Grafakos

Gregory Knese

Henri Martikainen

Some Problems in Harmonic Analysis

by

Fragkos Anastasios

A dissertation presented to

Washington University in St. Louis

in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

August 2024

St. Louis, Missouri

Table of Contents

Acknowledgments

First, I would like to thank my advisor Brett Wick for the patience he showed with me, his constant support, trust, and our numerous discussions throughout the years that have helped me grow as a mathematician. I am also thankful to Francesco Di Plinio for his mentorship, guidance, and collaboration during the second and third year of my PhD which resulted in the first chapter of this thesis. Last but not least, I am also grateful to Walton Green for the stimulating mathematical conversations we have had over the last four years as well as his eagerness to answer any question of mine. This thesis would have been impossible to complete without the help of the aforementioned people.

I want to thank the members of my committee Alan Chang, Loukas Grafakos, Gregory Knese, and Henri Martikainen for being willing to take time out of their busy schedules to serve in this thesis defense.

In addition, to that, I would like to thank all the graduate students at WashU for creating a positive environment to do math in. Special thanks to George for our mathematical discussions during our qual years.

I would also like to thank Grigoris for the math, and non-math-related conversations and the encouragement.

Finally, this thesis is wholeheartedly dedicated to my parents Vaggelis and Stavroula for selflessly supporting me throughout my whole life.

Dedicated to my parents.

ABSTRACT OF THE DISSERTATION

Some Problems in Harmonic Analysis

by

Fragkos, Anastasios

Doctor of Philosophy in Mathematics,

Washington University in St. Louis, 2024.

Professor Brett Wick, Chair

One the most central questions in harmonic analysis of whether the Fourier series of a square integrable function on the torus $\mathbb{T}$ converges Lebesgue a.e. $x \in \mathbb{T}$ was answered positively by L. Carleson in 1966 [7], by means of a weak-L^2 inequality for the maximal operator

$$\mathcal{C}f(x) = \sup_{N\in\mathbb{Z}} \left| \sum_{|\xi|\leq N} \widehat{f}(\xi) \exp(ix\xi) \right|, \qquad x \in \mathbb{T}. \tag{0.0.1}$$

The argument of [7] estimates $\mathcal{C}$ pointwise as a maximal modulated Hilbert transform, outside appropriately constructed exceptional sets whose mass is controlled by almost-orthogonality. The implicit distributional estimate in [7] was later exploited by Hunt [39] to deduce the family of restricted weak-type L^p bounds

$$\|\mathcal{C}f\|_{L^{p,\infty}(\mathbb{T})} \leq \frac{Cp^2}{p-1}|F|^{\frac{1}{p}}, \qquad F \subset \mathbb{T},\ |f| \leq \mathbf{1}_F, \qquad 1 < p < \infty. \tag{0.0.2}$$

The estimate (1.0.2) and interpolation yield that $\mathcal{C}$ is a bounded operator on each $L^p(\mathbb{T})$, for $1 < p < \infty$. Consequently, pointwise a.e. convergence of the Fourier series holds for $f \in L^p(\mathbb{T})$ in the same range. Since [7, 39], several substantially different proofs of Carleson's theorem have appeared: in particular, the celebrated ones by Fefferman [28] and Lacey-Thiele [48], one implicit in the return times theorem of Demeter, Lacey, Tao and Thiele [19], and more recently an improvement of Fefferman's proof [28] due to Lie [56].

The primary focus of the first part of this thesis is the behavior of the Carleson operator as $p \to 1^+$. Besides its intrinsic interest, this question is deeply connected to the pointwise a.e. behavior of Fourier series in function spaces between $L^1(\mathbb{T})$ and $L^p(\mathbb{T})$. To exemplify the connection, Antonov [1] coupled the precise information on the growth rate of the restricted weak norm from (1.0.2) with an approximation argument to deduce a mixed type estimate, which is the case $w = 1$ of (2.0.3) below.

The second part of this thesis is concerned with the compactness of multilinear Calderón-Zygmund operators. Calderón-Zygmund operators are omnipresent in the field of analysis. For example, they are connected to PDEs, Complex Analysis and Geometric Measure Theory. The study of the compactness of singular integral operators stems from applications to other fields such as the study of elliptic PDEs and the characterization of Semmes-Kenig-Torro domains but also from a functional analysis point of view. In particular, we prove a wavelet $T(1)$ theorem for compactness of multilinear Calderón-Zygmund (CZ) operators. Our approach characterizes compactness in terms of testing conditions and yields a representation theorem for compact CZ forms in terms of wavelet and paraproduct forms that reflect the compact nature of the operator.

The third and final part of this thesis deals with L^p estimates on the commutator of cancellative singular integral operators. In further detail, the aim of this chapter is to provide a proof of Bloom's original inequality using the wavelet representation theorem from [25]. A particular feature of our proof is that we precisely quantify the heuristic that "the commutator of a singular integral operator is a linear combination of the compositions of the paraproduct with the singular integral operator". In addition to that, we completely avoid working with shifts of arbitrary complexity. The main technical tool we use is wavelet averaging. Furthermore, our approach is robust enough to allow us to obtain off-diagonal estimates as well. Besides that, our proof is noticably shorter than the ones existing in the literature.

1. Introduction

One the most central questions in harmonic analysis is whether the Fourier series of a square integrable function on the torus $\mathbb{T}$ converges Lebesgue a.e. $x \in \mathbb{T}$ was answered positively by L. Carleson in 1966 [7], by means of a weak-L^2 inequality for the maximal operator

$$\mathcal{C}f(x) = \sup_{N\in\mathbb{Z}} \left| \sum_{|\xi|\leq N} \widehat{f}(\xi) \exp(ix\xi) \right|, \qquad x \in \mathbb{T}. \tag{1.0.1}$$

The argument of [7] estimates $\mathcal{C}$ pointwise as a maximal modulated Hilbert transform, outside appropriately constructed exceptional sets whose mass is controlled by almost-orthogonality. The implicit distributional estimate in [7] was later exploited by Hunt [39] to deduce the family of restricted weak-type L^p bounds

$$\|\mathcal{C}f\|_{L^{p,\infty}(\mathbb{T})} \leq \frac{Cp^2}{p-1}|F|^{\frac{1}{p}}, \qquad F \subset \mathbb{T},\ |f| \leq \mathbf{1}_F, \qquad 1 < p < \infty. \tag{1.0.2}$$

The estimate (1.0.2) and interpolation yield that $\mathcal{C}$ is a bounded operator on each $L^p(\mathbb{T})$, for $1 < p < \infty$. Consequently, pointwise a.e. convergence of the Fourier series holds for $f \in L^p(\mathbb{T})$ in the same range. Since [7,39], several substantially different proofs of Carleson's theorem have appeared: in particular, the celebrated ones by Fefferman [28] and Lacey-Thiele [48], one implicit in the return times theorem of Demeter, Lacey, Tao and Thiele [19], and more recently an improvement of Fefferman's proof [28] due to Lie [56].

The primary focus of the first part of this thesis, which is based on joint work with Francesco Di Plinio [69], is the behavior of the Carleson operator as $p \to 1^+$. Besides its intrinsic interest, this question is deeply connected to the pointwise a.e. behavior of Fourier series in function spaces between $L^1(\mathbb{T})$ and $L^p(\mathbb{T})$. To exemplify the connection, Antonov [1] coupled the precise information on the growth rate of the restricted weak norm from (1.0.2) with an approximation argument to deduce a mixed type estimate, which is the case $w = 1$ of (2.0.3) below. This may be leveraged to extend the pointwise convergence result to functions in the Orlicz space $L\log_1 L\log_3 L(\mathbb{T})$. Antonov's result has been, to date, the strongest known within the Orlicz-Lorentz scale.

The result of this chapter goes beyond the Carleson-Hunt bound (1.0.2), upgrading the estimate to the weak L^p-type.

Theorem A. *The maximal operator* (1.0.1) *obeys the family of estimates*

$$\|\mathcal{C}f\|_{L^{p,\infty}(\mathbb{T})} \leq \frac{C}{p-1} \|f\|_{L^p(\mathbb{T})}, \qquad 1 < p \leq 2.$$

The same bounds hold for the maximal multiplier (1.0.5) *and for the real line analogue* (1.0.3).

In fact, we obtain Theorem A as an immediate corollary of a stronger quantitative estimate for the *sparse norms* of the operator $\mathcal{C}$

Theorem B. *Let $m \in L^\infty(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R} \setminus \{0\})$ be a smooth Hörmander-Mihlin multiplier, see* (2.1.5)*. The associated maximally modulated multiplier*

$$\mathcal{C}f(x) := \sup_{N\in\mathbb{R}} \left| \int_{\mathbb{R}} m(\xi - N)\widehat{f}(\xi)\mathrm{e}^{ix\xi}\,\mathrm{d}\xi \right| \qquad x \in \mathbb{R}, \tag{1.0.3}$$

satisfies the family of sparse bounds

$$\|\mathcal{C}\|_{(p,1)} \leq \frac{C}{p-1}, \qquad 1 < p \leq 2 \tag{1.0.4}$$

with a uniform constant C. The same estimates hold for the periodic version of (1.0.3)

$$\mathcal{C}f(x) := \sup_{N\in\mathbb{Z}} \left| \sum_{\xi\in\mathbb{Z}} m(\xi - N)\widehat{f}(\xi)\mathrm{e}^{ix\xi} \right| \qquad x \in \mathbb{T} \tag{1.0.5}$$

under the additional transference assumption that $\lim_{\varepsilon\to 0^+} \fint_{|t|<\varepsilon} m$ *exists.*

As a direct corollary of the aforementioned sparse bound we are able to obtain the following weighted estimate.

Corollary B.1. *The maximal operator* (1.0.3) *obeys the weighted norm inequality*

$$\|\mathcal{C}f\|_{L^q(w)} \leq C_q[w]_{A_q}^{\frac{\max\{q,2\}}{q-1}} \|f\|_{L^q(w)}, \qquad 1 < q < \infty.$$

The same estimates holds for the periodic version of (1.0.5).

The backbone of our treatment is a new, sharply quantified near-L^1 Carleson embedding theorem for the modulation-invariant wave packet transform. The proof of the Carleson embedding relies on a newly developed smooth multi-frequency decomposition which, near the endpoint $p = 1$, outperforms the abstract Hilbert space approach of past works, including the seminal one by Nazarov, Oberlin and Thiele. As a further example of application, we obtain a quantified version of the family of sparse bounds for the bilinear Hilbert transforms due to Culiuc, Di Plinio and Ou.

The second part of this thesis is based on joint work with Walton Green and Brett Wick [29] and is concerned with the compactness of multilinear Calderón-Zygmund operators. Calderón-Zygmund operators are omnipresent in the field of analysis. For example, they are connected to PDEs, Complex Analysis and Geometric Measure Theory. The study of the compactness of singular integral operators stems from applications to other fields such as the study of elliptic PDEs and the characterization of Semmes-Kenig-Torro domains but also from a functional analysis point of view. In the linear case, Villaroya in [76] gave a complete characterization of compact Calderón-Zygmund operators on $L^2(\mathbb{R}^d)$, which was further developed in [67, 68, 73, 77]. Recently, Mitkovski and Stockdale in [57] gave a simplified formulation of the $T(1)$ theorem for compactness of Villaroya. More precisely, they showed that a CZO T is compact if and only if $T(1)$ and $T^*(1)$ both belong to $\mathrm{CMO}(\mathbb{R}^d)$ and a vanishing version of the weak boundedness property, called the weak compactness property, is satisfied. In contrast, a multilinear version of these compactness testing theorems remains unexplored. These $T(1)$ theorems for compactness are complemented by a compact Rubio De Francia theory of extrapolation developed by Hytönen and Lappas in [40, 41]. Subsequently, in [6], Cao, Olivo, and Yabuta extended the bilinear results of [40] to the multilinear setting and to the quasi-Banach range, in which case the target space can be L^r with $r > \frac{1}{m}$. However, due to the difficulties of extrapolating to the upper endpoint in the multilinear setting [53, 63], the results of [6, 40] do not consider the case where one (or more) input spaces is $L^\infty(\mathbb{R}^d)$. We point out that our results below do yield compactness when an input space is $L^\infty(\mathbb{R}^d)$. In [29] we prove a wavelet $T(1)$ theorem for compactness of multilinear Calderón-Zygmund (CZ) operators. Our approach characterizes compactness in terms of testing conditions and yields a representation theorem for compact CZ forms in terms of wavelet and paraproduct forms that reflect the compact nature of the operator. The main result of this chapter is stated as follows

Theorem C. *Suppose Λ is an $(m+1)$-linear CZ form, with associated symbols $\mathfrak{b}_j \in \mathrm{BMO}(\mathbb{R}^d)$, $j = 0, 1, \ldots, m$ satisfying* (3.1.2). *The following are equivalent.*

A. Λ is a compact CZ form, i.e. $\mathsf{W}_\Lambda^M(z) \to 0$ as $z \to \infty$ and $\mathfrak{b}_j \in \mathrm{CMO}(\mathbb{R}^d)$.

B. There exist compact wavelet forms $\{U_k\}_{k=1}^{K_m}$ and compact paraproduct forms $\left\{\Pi_{\mathfrak{b}_j}\right\}_{j=1}^{m}$ such that for all $f_j \in \mathcal{S}(\mathbb{R}^d)$,

$$\Lambda(\mathbf{f}) = \sum_{k=1}^{K_m} U_k(\mathbf{f}) + \sum_{j=0}^{m} \Pi_{\mathfrak{b}_j}^{*,j}(\mathbf{f}), \quad \mathbf{f} = (f_0, \ldots, f_m).$$

C. Each element of T_Λ is a compact CZO.

We note that Theorem C applies to the linear case as well and a few simplifications can be made due to the greater symmetry enjoyed in this setting. For additional clarity, we restate Theorem C when $m = 1$.

Theorem D. *Let T be a linear CZO. The following are equivalent.*

A. $T(1), T^*(1) \in \mathrm{CMO}(\mathbb{R}^d)$ and

$$\lim_{\zeta \to \infty} \sup_{z \in \mathbb{B}(\zeta)} t^d \left| \langle T\phi_z, \phi_\zeta \rangle \right| = 0.$$

B. There exists a compact wavelet form U such that for all $f, g \in \mathcal{S}(\mathbb{R}^d)$,

$$\langle Tf, g \rangle = U(f, g) + \Pi_{T(1)}(f, g) + \Pi_{T^*(1)}(g, f).$$

C. T and T^* are compact CZOs.

D. $T(1), T^*(1) \in \mathrm{CMO}(\mathbb{R}^d)$ and

$$\lim_{z \to \infty} t^{\frac{d}{2}} \left\| T\phi_z \right\|_{L^2(\mathbb{R}^d)} = 0.$$

Our main result is another example of a new result in the Lebesgue setting obtained through wavelet representations. This approach differs from the ones in [57] and [76] as we do not use the machinery regarding localization, and neither do we require additional decay of the kernel, but rather we use a wavelet averaging procedure to obtain the representation formula in §3.3 and then use the Riesz-Kolmogorov criterion to obtain the precompactness of the image of the unit ball under the adjoint operators to the wavelet and paraproduct forms in §3.4 and §3.5.

The third and final part of this thesis is based on joint work with Brett Wick [30] and deals with L^p estimates on the commutator of cancellative singular integral operators. In further detail, the aim of this chapter is to provide a proof of Bloom's original inequality using the wavelet representation theorem from [25]. A particular feature of our proof is that we precisely quantify the heuristic that "the commutator of a singular integral operator is a linear combination of the compositions of the paraproduct with the singular integral operator". In addition to that, we completely avoid working with shifts of arbitrary complexity. The main technical tool we use is wavelet averaging. Furthermore, our approach is robust enough to allow us to obtain off-diagonal estimates as well. Besides that, our proof is noticably shorter than the ones existing in the literature. In particular, we can recover Bloom's inequality as a corollary of the following two weight theorem.

Theorem E. *For a* 1*-cancellative Calderón-Zygmund operator,* $w, \sigma \in A_p,\ b \in \mathrm{BMO}_\nu(\mathbb{R})$, *where* $\nu = \left(\frac{w}{\sigma}\right)^{\frac{1}{p}}$ *we have that*

$$\|[b, T]\|_{L^p(w,\mathbb{R})} \lesssim \|b\|_{\mathrm{BMO}_\nu(\mathbb{R})} \|f\|_{L^p(\sigma,\mathbb{R})} .$$

Finally, our method revisits the following one weight result.

Theorem F. *Let T be a cancellative Calderón-Zygmund operator and w weight with $w^p \in A_p$ and $w^q \in A_q$ then*

$$\|[b,T]\|_{L^p(w^p,\mathbb{R}^d)\to L^q(w^q,\mathbb{R}^d)} \lesssim \begin{cases} \|b\|_{\mathrm{BMO}}\,, & q=p,\\ \|b\|_{\dot{C}^{0,\alpha}(\mathbb{R}^d)}\,,\ \frac{\alpha}{d}=\frac{1}{p}-\frac{1}{q}, & q>p,\\ \|b\|_{\dot{L}^r(\mathbb{R}^d)}\,,\ \frac{1}{q}=\frac{1}{r}+\frac{1}{p}, & q<p. \end{cases}$$

Where the homogeneous Hölder norm of exponent a is defined as

$$\|b\|_{\dot{C}^{0,a}(\mathbb{R}^d)} \coloneqq \sup_{x\neq y} \frac{|b(x)-b(y)|}{|x-y|^a}$$

and the $\dot{L}^r(\mathbb{R}^d)$ norm is defined as $\|b\|_{\dot{L}^r(\mathbb{R}^d)} \coloneqq \inf_{c\in\mathbb{R}} \|b-c\|_{L^r(\mathbb{R}^d)}$.

The notational organization of this thesis consists of introducing the notation for the first chapter separately from the one that will be used in the second and third chapter. In particular, the notation and general definitions in the second and third chapter will be common.

2. The weak type Carleson's theorem via wave packet estimates

Recurring notation

The treatment in this part of the thesis focuses on the case of functions defined on the real line; however, the generalization to higher dimensional Euclidean spaces is merely notational and all arguments are easily transcribed to that setting. The Fourier transform on $\mathbb{R}$ obeys the normalization

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \mathrm{e}^{-ix\xi} \, \mathrm{d}x, \qquad \xi \in \mathbb{R}.$$

Throughout, the transformations

$$\mathrm{Tr}_a f := f(\cdot - a), \qquad \mathrm{Mod}_a f := \exp(ia\cdot) f(\cdot), \qquad \mathrm{Dil}_b^p f := b^{-\frac{1}{p}} f(b^{-1}\cdot)$$

for $a \in \mathbb{R}, b > 0$, $0 < p \leq \infty$, are used to describe the invariance properties of our singular operators. The symbol

$$\langle x \rangle = \sqrt{1 + |x|^2}, \qquad x \in \mathbb{R}$$

indicates the usual Japanese bracket. The center and length of an interval $I \subset \mathbb{R}$ are respectively denoted by c_I and ℓ_I. Accordingly, define the L^∞-normalized polynomial decay factor adapted to I by

$$\chi_I^M := \mathrm{Tr}_{c_I} \mathrm{Dil}_{\ell_I}^\infty \langle \cdot \rangle^{-M}, \qquad M \in 2\mathbb{N} \setminus \{0\}.$$

When we drop M and simply write χ_I instead, the parameter M is large and unimportant. As customary, for $0 < p \leq \infty$, local L^p-(quasi)norms on I, their tailed analogues and the p-th Hardy-Littlewood maximal operator follow the notation

$$\langle f \rangle_{p,I} \coloneqq |I|^{-\frac{1}{p}} \|f \mathbf{1}_I\|_p, \qquad \langle\!\langle f \rangle\!\rangle_{p,I} \coloneqq |I|^{-\frac{1}{p}} \left\| f \chi_I^{\frac{2^9}{p}} \right\|_p, \qquad \mathrm{M}_p f \coloneqq \sup_{I \subset \mathbb{R}} \langle f \rangle_{p,I} \mathbf{1}_I.$$

with most times $\mathrm{M}_1 = \mathrm{M}$ for simplicity. We clarify our notation for the weighted Lorentz and Orlicz spaces appearing in the results of Corollary F.1. A *weight* stands for a positive integrable function w on $\mathbb{T} = (-\pi, \pi]$. There is no loss in generality with assuming that $\int_{\mathbb{T}} w = 1$. As customary, we overload the notation for the weight w and the corresponding measure $\mathrm{d}w = w \mathrm{d}x$. The weak and strong weighted Lebesgue quasinorms are then defined for $p \in (0, \infty)$ by

$$\|f\|_{L^{p,\infty}(\mathbb{T};w)} \coloneqq \sup_{t>0} t \left[w\left(\{x \in \mathbb{T} : |f(x)| > t\}\right)\right]^{\frac{1}{p}}, \qquad \|f\|_{L^p(\mathbb{T};w)} \coloneqq \left(\int_{\mathbb{T}} |f|^p \, \mathrm{d}w \right)^{\frac{1}{p}}.$$

If $\Phi : [0, \infty] \to [0, \infty)$ is a fundamental function, the weighted Orlicz norm $\Phi(L)(\mathbb{T}; w)$ is

$$\|f\|_{\Phi(L)(\mathbb{T};w)} \coloneqq \inf \left\{ t > 0 : \int_{\mathbb{T}} \Phi\left(\frac{|f|}{t} \right) \mathrm{d}w \leq 1 \right\}.$$

The fundamental functions occurring are $\Phi(t) = t \log_1 t$ and $\Phi(t) = t \log_1 t \log_3 t$, with iterated logarithm notation

$$\log_1 t = \max\{1, \log t\}, \qquad \log_k t = \max\{1, \log(\log_{k-1} t)\}, \quad k \geq 2.$$

The quasinorm $\mathrm{QA}_q(w)$ appearing in (2.0.4) is defined by

$$\|f\|_{\mathrm{QA}_q(w)} = \inf \left\{ \sum_{j=1}^{\infty} \log_1 j \, \|f_j\|_{L^1(\mathbb{T};w)} \log_1 \left(\frac{\|f_j\|_{L^q(\mathbb{T};w)}}{\|f_j\|_{L^1(\mathbb{T};w)}} \right) : f = \sum_{j=1}^{\infty} f_j, \sum_{j=1}^{\infty} |f_j| < \infty \text{ a.e.} \right\}. \tag{2.0.1}$$

Finally, the symbol C and the constant implied by the almost inequality sign $\lesssim$ are meant to be absolute, unless otherwise specified via the notation $C_{a_1,\ldots,a_n}$, $\lesssim_{a_1,\ldots,a_n}$. The latter notation highlights dependence on the parameters $a_1, \ldots, a_n$.

2.0.1 Introduction

For $n \geq 2$, $\vec{p} = (p_1, \ldots, p_n) \in (0, \infty)^n$, the n-linear $\vec{p}$-maximal function of a tuple $\{f_j \in L^{p_j}_{\mathrm{loc}}(\mathbb{R}^d) : 1 \leq j \leq n\}$ is defined as

$$\mathrm{M}_{\vec{p}}(f_1, \ldots, f_n) \coloneqq \sup_Q \mathbf{1}_Q \prod_{j=1}^n \langle f_j \rangle_{p_j, Q}$$

the supremum being taken over all cubes Q of $\mathbb{R}^d$. See the final paragraph of this introduction for a summary of standard notations. An n-sublinear form Λ acting e.g. on n-tuples of functions $f_j \in L^\infty_0(\mathbb{R}^d)$ is *$\vec{p}$-sparse bounded* if there exists a constant $C > 0$ such that

$$|\Lambda(f_1, \ldots, f_n)| \leq C \left\| \mathrm{M}_{\vec{p}}(f_1, \ldots, f_n) \right\|_1$$

uniformly over all such tuples, and the *$\vec{p}$-sparse bound* $\|\Lambda\|_{\vec{p}}$ is the infimum of the set of all such constants. If T is an $(n-1)$-sublinear operator, the quantity $\|T\|_{\vec{p}}$ indicates the sparse bound $\|\Lambda\|_{\vec{p}}$ of the n-sublinear form

$$\Lambda(f_1, \ldots, f_n) = \langle T(f_1, \ldots, f_{n-1}), f_n \rangle.$$

Note that T is a specific formal adjoint of Λ, and the index n plays a distinguished role. The equivalence of this formulation with more standard notions of sparse bounds [51] is thoroughly discussed in [15, 63] and references therein. Note that the Carleson maximal operator, on the real line and on the torus respectively, correspond up to symmetries and linear combination with the identity operator to the choice $m = \mathbf{1}_{(0,\infty)}$ in (1.0.3), (1.0.5).

The $\vec{p}$-sparse bounds of T subsume a full range of quantitative weighted norm inequalities of weak and strong type. We send to the references [54,63] for a complete list of consequences and for the related extrapolation theory, and content ourselves with recalling those implications most crucial for our exposition, in the form of corollaries to this main result. Then, the estimates of Theorem A are derived from the sparse bound of Theorem B as in e.g. [12, Theorem E]. Two more corollaries are of weighted nature. The following weak type $L^p(w)$ bound for A_1 weights with controlled constant holds.

Corollary F.1. *For weights $w \in A_1$ and $1 \leq p \leq 2$, define*

$$\mathsf{K}(w,p) := [w]_{A_1}^{\frac{1}{p}} [w]_{A_\infty}^{1-\frac{1}{p}} [\log_1 [w]_{A_\infty}]^{\frac{2}{p}}.$$

For both (1.0.3), (1.0.5), *there holds*

$$\|\mathcal{C}f\|_{L^{p,\infty}(w)} \leq \frac{C\mathsf{K}(w,p)}{p-1} \|f\|_{L^p(w)}, \qquad 1 < p \leq 2. \tag{2.0.2}$$

As a further corollary of (2.0.2), (1.0.5) *satisfies the following endpoint estimates:*

$$\|\mathcal{C}f\|_{L^{1,\infty}(\mathbb{T};w)} \leq \tfrac{Cq}{q-1}\mathsf{K}(w,1) \|f\|_{L^1(\mathbb{T};w)} \log_1\left(\frac{\|f\|_{L^q(\mathbb{T};w)}}{\|f\|_{L^1(\mathbb{T};w)}}\right), \qquad 1 < q \leq \infty, \tag{2.0.3}$$

$$\|\mathcal{C}f\|_{L^{1,\infty}(\mathbb{T};w)} \leq \tfrac{Cq}{q-1}\mathsf{K}(w,1) \|f\|_{\mathrm{QA}_q(w)}, \qquad 1 < q \leq \infty, \tag{2.0.4}$$

$$\|\mathcal{C}f\|_{L^{1,\infty}(\mathbb{T};w)} \leq C\mathsf{K}(w,1) \|f\|_{L\log_1 L\log_3 L(\mathbb{T};w)}. \tag{2.0.5}$$

See (2.0.1) *for the definition of the* $\mathrm{QA}_q(w)$*-quasinorms. Additionally, as a consequence of* (2.0.4), *the Fourier series of $f \in \mathrm{QA}_q(w)$ converges pointwise a.e. whenever $w \in A_1$.*

Proof. Estimate (2.0.2) is obtained by using the $(p,1)$-sparse bound of Theorem B as the input of [32, Theorem 1.4]. For (2.0.3), a consequence of (2.0.2) is that

$$\|\mathcal{C}f\|_{L^{1,\infty}(\mathbb{T};w)} \leq \|\mathcal{C}f\|_{L^{p,\infty}(\mathbb{T};w)} \leq \frac{C\mathsf{K}(w,p)}{p-1}\|f\|_{L^p(\mathbb{T};w)} \leq \frac{C\mathsf{K}(w,1)}{p-1}\|f\|_{L^p(\mathbb{T};w)} \tag{2.0.6}$$

holds whenever $1 < p \leq 2$. For $p < q$, $\|f\|_{L^p(\mathbb{T};w)} \leq \|f\|_{L^1(\mathbb{T};w)}^{1-q'/p'}\|f\|_{L^q(\mathbb{T};w)}^{q'/p'}$, and (2.0.3) follows by using (2.0.6) for p given by the equation $p' = \max\left\{2, q'\log\left(\|f\|_{L^q(\mathbb{T};w)}/\|f\|_{L^1(\mathbb{T};w)}\right)\right\}$. Now, (2.0.4) is deduced from the definition of $\mathrm{QA}_q(w)$ and Kalton's log-convexity of $L^{1,\infty}(\mathbb{T};w)$ [45]. Estimate (2.0.4) immediately implies (2.0.5) once the (strict) continuous inclusion

$$L\log_1 L\log_3 L(\mathbb{T};w) \subsetneq \mathrm{QA}_\infty(w). \tag{2.0.7}$$

is established. This is done repeating with obvious changes the argument of [10, Sect. 3.3]. Note that the inclusion (2.0.7) is tight in the Orlicz class, under modest assumptions on the fundamental function [8]. □

Another aspect naturally arising in the pursuit of endpoint estimates and pointwise convergence of Fourier series for spaces near L^1 is the sharp quantification of the dependence on the weight constants in the weighted bounds for the Carleson operator. For instance, the next result yields that $\mathcal{C} : L(\log_1 L)^2(\mathbb{T}) \to L^1(\mathbb{T})$, via the extrapolation theory of [9]. Note that $L(\log_1 L)^2(\mathbb{T})$ is the largest Orlicz space currently known to have this property, a result originally due to Sjölin [71].

Remark 2.0.2 (Comparison with previous results)**.** This remark will place our new results in the context of past literature. First of all, the Carleson-Hunt estimate (1.0.2) is quantitatively equivalent to the generalized restricted weak type bound of Lacey and Thiele [48], and strictly stronger than the estimate proved by [28], which, when phrased as a restricted type estimate, is of type $\mathcal{C} : L^{p,1} \to L^q$ for $1 < q < p$. An alternative formulation of (1.0.2) is

$$\|\mathcal{C}f\|_{L^{1,\infty}(\mathbb{T})} \le C|F|\log_1\left(\frac{1}{|F|}\right), \qquad F \subset \mathbb{T},\ |f| \le \mathbf{1}_F. \tag{2.0.8}$$

Relying on the smoothness of the Dirichlet kernel via an approximation argument, Antonov [1] upgraded (2.0.8) to a mixed type bound which is exactly estimate (2.0.3) with $w = 1$ and $q = \infty$, and deduced the $w = 1$ case of (2.0.5). Further work of Sjölin and Soria [72] extended Antonov's approach to more general sublinear operators satisfying Carleson-Hunt type bounds as in (2.0.8); see also [35] for applications of this principle to weighted bounds. Arias de Reyna [2] introduced the quasi-Banach spaces $\mathrm{QA}_\infty := \mathrm{QA}_\infty(\mathrm{d}x)$ and noticed that Antonov's result may be phrased in terms of (2.0.4) for $w = 1$. The observation of [2] is relevant because of the strict inclusion (2.0.7). The work [56] by Lie gave a proof of the Lebesgue case of (2.0.3), with unspecified dependence on $1 < q < \infty$, without any appeal to approximation arguments of the type used in [1, 72]. In a nutshell, [56] refines the construction of the forests from Fefferman's proof of Carleson's theorem in a BMO sense. The main result of [56] thus implies the unweighted case of (2.0.4) via the same log-convexity argument. The work [56] also contains the observation[1] that QA_∞ and QA_q are equivalent quasinorms for each $1 < q < \infty$, so that the results of [56] and [2] are formally equivalent.

As far as prior weighted bounds at the endpoint $p = 1$, the work of Carro and Domingo-Salazar [9] deduces from the Carleson-Hunt bound (1.0.2) and extrapolation that the Carleson operator maps $L\log_1 L\log_3 L(\mathbb{T}; w)$ into the space $R_1(w)$, which is a logarithmic correc-

[1] In [56], the observation that QA_∞ and QA_q are the same quasi-Banach space is attributed to L. Rodriguez-Piazza.

tion of $L^{1,\infty}(\mathbb{T};w)$, and the operator norm dependence on $[w]_{A_1}$ is unspecified. In view of the strict continuous embeddings $L^{1,\infty}(\mathbb{T};w) \hookrightarrow R_1(w)$ and (2.0.7), and of the dependence of $\mathsf{K}(w,1)$ on $[w]_{A_1}$, our estimate (2.0.4) improves on [9, Theorem 4.5].

Corollary B.1 is an improvement on the previously best known quantitative estimate for the $L^q(w)$ norms of maximally modulated multipliers, due to Lerner and Di Plinio [22]. In particular, the extra $\log_2[w]_{A_q}$ term appearing in [22, Corollary 1.2 (ii)] is shown to be unnecessary.

In summary, the weak-L^p bound of Theorem A, and *a fortiori* the $(p,1)$-sparse estimate of Theorem B, sharpen the Carleson-Hunt bound (1.0.2). Theorem B also yields upgraded versions of previous results at $p=1$, which are all consequences of (1.0.2). In particular, Corollary F.1 ensures that the Fourier series of any function in the class

$$X \coloneqq \left\{ f \in L^1(\mathbb{T}) : f \in \mathrm{QA}_\infty(w) \text{ for some } w \in A_1 \right\}$$

converges almost everywhere. We stress that the class X is not just formally larger than QA_∞. For instance, for

$$w(x) = \frac{1}{|\log(|x|)|^{\frac{1}{2}}} \mathbf{1}_{\mathbb{R}\setminus\{0\}}(x), \qquad f(x) = \frac{1}{x(\log x)^2 \log\log|\log x|} \mathbf{1}_{(0,\mathrm{e}^{-\mathrm{e}^{\mathrm{e}}})}(x),$$

we have $w \in A_1$, $f \in L\log_1 L\log_3 L(\mathbb{T};w) \subset \mathrm{QA}_\infty(w)$, $f \notin \mathrm{QA}_\infty$.

2.0.3 Methods, organization and further results

The proof of Theorem B is, in essence, a version of the Lacey-Thiele argument from [48] for functions outside local L^2 that avoids interpolation and the consequent loss of constants. In Section 2.1, matters are reduced to estimating bilinear forms involving wave packet coefficients (2.1.2) associated to a *tile* P, namely a Heisenberg uncertainty box in the space-frequency plane, and their modified version (2.1.3). The wave packet coefficient (2.1.2) is roughly the L^∞ norm of the projection of f to a $O(1)$-dimensional subspace of functions space-frequency adapted to P. Using the outer L^p framework of Do-Thiele [26], described in Section 2.2, the main steps of the proof become two quantified and localized outer Carleson embedding theorems for the wave packet maps (2.1.2) and (2.1.3). The latter is essentially a localized reformulation of the mass parameter bounds of [48] and occupies Section 2.3. The former, Theorem H, is substantially new, and is stated and proved in Section 2.4. Section 2.5 then contains the short and completely elementary stopping forms argument leading to Theorem B.

The main novel technical tool behind the proof of Theorem H is a smooth space-frequency decomposition of a function f locally in L^p, $1 < p < 2$ induced by a *forest*, namely a collection of tiles organized into space-frequency *trees*. The decomposition is constructed by expanding f in Gabor series spatially localized on Calderón-Zygmund intervals associated to the forest, and selecting a principal part (2.4.11) which is locally in L^2, albeit with local norms depending on the counting function of the forest. Multi-frequency decomposition lemmas of different flavor have been used extensively in the past literature on modulation invariant singular integrals [23, 24, 66]. The construction used in all these references generates a *good part* via projection on the linear span of N pure frequencies on a spatial interval, initially due to Nazarov, Oberlin and Thiele [62], and based on a sleek Hilbert space lemma of Borwein-

Erdelyi [5]. The corresponding remainder term does have vanishing moments with respect to the relevant frequencies, but its local norms are of the same order of those of the good part, and thus also depend on the counting function. This loss may only be offset by paying an additional price on the good part. On the contrary, the smooth remainder (2.4.12) from our decomposition inherits the much smaller local norms of f and its contribution to (2.1.2) may be estimated as a pure error term, by careful exploitation of frequency decay and separation in frequency localization. We expect that our smooth decomposition will find extensive use in further problems involving modulation invariant estimates outside local L^2, such as, for instance, uniform estimates for the bilinear Hilbert transform, see [61, 66, 74] for context.

The wave packet coefficients (2.1.2) also appear in the model sums of the multiplier operators with singularity along subspaces of rank one, whose archetypal example is the bilinear Hilbert transform. The first L^p-bounds for the latter operator are due to Lacey and Thiele [46, 47], while Muscalu, Tao and Thiele address more general multipliers and higher ranks [60]. A systematic qualitative weighted theory for rank one multiplier operators was first obtained by Culiuc, Di Plinio and Ou in [16], as a corollary of a family of $\vec{p}$-sparse bounds. Subsequently, several works have deduced from the sparse bounds of [16] further qualitative weighted and vector-valued norm inequalities by developing suitable multilinear extrapolation theorems, see e.g. [14, 53, 54, 63]. On the other hand, it has proved difficult to deduce *quantitative* weighted estimates, i.e. with specified, possibly sharp dependence, from the main result of [16], mainly because the constant in the $\vec{p}$-sparse bounds blows up in an unspecified way when the vector $\vec{p}$ approaches the extremal points of the range. The wave packet embedding of Theorem H may be used to quantify the blow up rate much more precisely, leading to the following improvement of [16, Theorem 1.3].

Theorem G. *Let* $\Gamma = \{\xi = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 : \xi_1 + \xi_2 + \xi_3 = 0\}$ *and* $\Gamma' = \operatorname{span}\gamma$ *be a non-degenerate rank 1 subspace of* Γ*, in the sense that* $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ *is a unit vector with* $\gamma_j \neq 0$ *for all* $j = 1, 2, 3$*. Let* $m \in L^\infty(\mathbb{R}^3) \cap \mathcal{C}^\infty(\mathbb{R}^3 \setminus \Gamma')$ *be a symbol satisfying the estimates*

$$\sup_{\xi \in \mathbb{R}^3 \setminus \Gamma'} [\operatorname{dist}(\xi, \Gamma')]^{|\alpha|} |\nabla^\alpha m(\xi)| \leq 1$$

for all multi-indices α *up to some finite order. Then the form*[2]

$$\Lambda_m(f_1, f_2, f_3) = \int_\Gamma m(\xi) \prod_{j=1}^{3} \widehat{f}_j(\xi_j) \, \mathrm{d}\xi \tag{2.0.9}$$

satisfies the family of $\vec{p}$*-sparse bounds*

$$\|\Lambda_m\|_{\vec{p}} \lesssim \frac{1}{\varepsilon(\vec{p})}, \qquad 1 \leq p_1, p_2, p_3 < \infty, \qquad \varepsilon(\vec{p}) \coloneqq 2 - \sum_{j=1}^{2} \frac{1}{\min\{p_j, 2\}} > 0. \tag{2.0.10}$$

The proof of Theorem G is given in Section 2.6. Note that the adjoint forms to the (non-degenerate) bilinear Hilbert transforms with parameter $\beta = \gamma \times (1,1,1)$ correspond to the choices $m(\xi) = \mathbf{1}_{(0,\infty)}(\beta \cdot \xi)\psi((1,1,1)\cdot\xi)$, where ψ is any Schwartz function on $\mathbb{R}$ with $\psi(0) = 1$. We do not detail the consequences in terms of weighted bounds for Λ_m, which may be reconstructed by the interested reader via the extrapolation theorems of [53, 54, 63]. Tracking the constants in those works will lead to quantitative weighted estimates. This point is transversal to the present chapter of this thesis and will be expounded elsewhere.

[2] The action of Λ_m on tuples of functions f_j which are merely assumed to belong to $L^\infty_0(\mathbb{R})$ may be defined by smooth truncation of the integral (2.0.9) near Γ and at infinity. The obtained bounds are uniform with respect to the truncation parameter, thus allowing for the limiting argument. This is classical, and we omit details.

In a different direction, Theorem G yields precise information on the behavior of Λ_m near the extremal [49, Subsect. 2.2] pair $L^1(\mathbb{R}) \times L^2(\mathbb{R})$, fully recovering all results obtained in [24] and leading to several improvements. One of these is detailed in the following corollary, improving in particular [24, Theorem 3].

Corollary G.1. *Let T_m be an adjoint to Λ_m from* (2.0.9). *Then*

$$\left\| T_m : L^{\frac{1}{1-\varepsilon}}(\mathbb{R}) \times L^2(\mathbb{R}) \to L^{\frac{2}{3-2\varepsilon},\infty}(\mathbb{R}) \right\| \lesssim \frac{1}{\varepsilon}, \qquad 0 < \varepsilon < 2^{-6},$$
$$\left\| T_m : L^{1,\frac{2}{3}} \log L^{\frac{2}{3}}(\mathbb{R}) \times L^2(\mathbb{R}) \to \frac{L^{\frac{2}{3},\infty}}{\log L}(\mathbb{R}) \right\| \lesssim 1.$$

For the definition of the spaces appearing in the second estimate and its easy deduction from the first, see [18, Theorem 4].

Remark 2.0.4 (On the relationship between sparse and weak type)**.** Corollaries F.1 and B.1 demonstrate how sparse bounds are both formally stronger and convey additional information than Lebesgue estimates. The article [22], based on mean oscillation techniques, contains a partial converse of the sparse to weak type implication for maximal modulation singular integrals. The weighted estimates of [22] have in fact been deduced relying on weak-L^p type bounds which are strictly weaker than both Theorem B and the Carleson-Hunt bound (1.0.2), and which are in fact consequences of (1.0.2) and extrapolation; see e.g. [22, Estimate (1.7)]. On the other hand, our embedding Theorem H yields Theorem B directly, and also applies in the context of Theorem G, which is out of reach for current mean oscillation techniques: see [52] for a more recent, unified approach to sparse domination via weak type bounds.

2.1 Space-frequency analysis of modulation invariant operators

After a few preliminaries, this section introduces the wave packet transform (2.1.2) on the space-frequency tiles, its modified version (2.1.3), and their role in the discretization of maximally modulated singular integrals.

2.1.1 Dyadic grids and tiles

We say that a collection $\mathcal{D}$ of intervals of $\mathbb{R}$ is a *dyadic grid* if

a. $\{\ell_I : I \in \mathcal{D}\} \subset \rho 2^{\mathbb{Z}}$ for some $\frac{1}{2} < \rho = \rho_{\mathcal{D}} < 2$.

b. for all $k \in \mathbb{Z}$ there holds $\mathbb{R} = \bigcup \{I \in \mathcal{D}, \ell_I = \rho 2^k\}$ up to possibly a set of zero measure (covering property);

c. $I, J \in \mathcal{D} \implies I \cap J \in \{\varnothing, I, J\}$ (grid property).

The elements of a dyadic grid are referred to as *dyadic intervals*. A typical example that we will use at times are the three shifted dyadic grids

$$\mathcal{D}_g := \left\{ 2^k \left(\ell + \tfrac{g(-1)^k}{3} + [0,1) \right) : k, \ell \in \mathbb{Z} \right\}, \qquad g = 0, 1, 2.$$

Remark 2.1.2 (Parent, sibling, and children intervals)**.** Let $I \in \mathcal{D}$ be a dyadic interval and $\kappa \geq 1$. Properties a. to c. yield the existence of a unique interval $I^{\mathsf{p}(\kappa)} \in \mathcal{D}$ with $\ell_{I^{\mathsf{p}(k)}} = 2^{\kappa} \ell_I$ and $I \subset I^{\mathsf{p}(\kappa)}$. We call $I^{\mathsf{p}(\kappa)}$ the κ-th *parent* of I. Conversely, if $I \in \mathcal{D}$, we enumerate by

$$I^{\mathsf{ch}(\kappa,j)}, \qquad j = 1, \ldots, 2^{\kappa}$$

the collection of the κ-*grandchildren* of I. These are those $J \in \mathcal{D}$ with $J^{\mathsf{p}(\kappa)} = I$, with the obvious convention that $c(I^{\mathsf{ch}(\kappa,j_1)}) < c(I^{\mathsf{ch}(\kappa,j_2)})$ if $1 \leq j_1 < j_2 \leq 2^\kappa$. Finally, we denote by I^{b}, the *sibling* of I, the unique $J \in \mathcal{D}$ with $\ell_J = \ell_I$ and $I^{\mathsf{p}(1)} = I \cup J$.

Remark 2.1.3 (Shifted grids)**.** Let M be a large integer, standard shifted dyadic grid techniques, see e.g. [51], yield the existence of dyadic grids $\mathcal{G}_j$, $j = 1, \ldots, 2^{M+10}$ with the following property: for every (not necessarily dyadic) interval $Q \subset \mathbb{R}$ there exists j and $I(Q) \in \mathcal{G}_j$ with $Q \subset I(Q)$ and $\ell_{I(Q)} \leq (1 + 2^{-M})\ell_Q$. This property will be used a couple of times in what follows.

We say that the grids $\mathcal{D}, \mathcal{D}'$ are *dual* if $\rho_{\mathcal{D}}\rho_{\mathcal{D}'} = 1$. Let now $\mathcal{D} \times \mathcal{D}'$ be a fixed pair of dual dyadic grids on $\mathbb{R}$. A *tile* $P = I_P \times \omega_P \in \mathcal{D} \times \mathcal{D}'$ is the Cartesian product of dyadic intervals with reciprocal lengths, that is $\ell(I_P)\ell(\omega_P) = 1$. The intervals I_P, ω_P are referred to respectively as the *spatial support* and *frequency support* of the tile P. The set of *all tiles* in $\mathcal{D} \times \mathcal{D}'$ is denoted by $\mathbb{S}_{\mathcal{D},\mathcal{D}'}$ or simply $\mathbb{S}$ if the dyadic grids are fixed and clear from context, and referred to as *tiling* associated to $\mathcal{D} \times \mathcal{D}'$ or simply *tiling*. It is convenient to adopt the notation $\mathsf{scl}(P) = \ell_{I_P}$ for the (spatial) scale of P.

2.1.4 Wave packets and wave packet transforms.

The rationale for defining tiles as above is that they describe the space-frequency localization of the functions, referred to as *wave packets*, involved in the analysis of modulation invariant operators. Denote by Θ^M the unit ball of the Banach space

$$\left\{\vartheta \in \mathcal{C}^M(\mathbb{R}) : \|\vartheta\|_{\star,M} < \infty, \right\}, \qquad \|\vartheta\|_M \coloneqq \sup_{0\le\alpha\le M}\sup_{x\in\mathbb{R}} \left|\langle x\rangle^M D^\alpha \vartheta(x)\right|.$$

For a tile $P = I_P \times \omega_P$, define the corresponding L^1-adapted, localized classes of order M by

$$\Phi^M(P) \coloneqq \left\{\varphi = \mathrm{Mod}_{c(\omega_P)}\mathrm{Tr}_{c(I_P)}\mathrm{Dil}^1_{\mathsf{scl}(P)}\phi \text{ for some } \phi\in\Theta^M,\ \mathrm{supp}\,\widehat{\varphi}\subset\omega_P\right\}. \tag{2.1.1}$$

We stress that $\varphi \in \Phi^M(P)$ has compact frequency support in ω_P. From now on, we omit the M from the superscript and our forthcoming definitions depend on M implicitly. The *order M wave packet transform* of $f \in L^\infty_0(\mathbb{R})$ is the map

$$W[f] : \mathbb{S} \to [0,\infty), \qquad W[f](P) \coloneqq \sup_{\varphi\in\Phi^M(P)} |\langle f,\varphi\rangle|. \tag{2.1.2}$$

The dependence on M is kept implicit in the notation. We can think of $W[f](P)$ as the magnitude of the space-frequency localization of f to the tile P.

When dealing with maximally modulated singular integrals, a modified wave packet transform models the contribution of the dualizing function. Namely, define

$$A[f](P) \coloneqq \sup_{\psi\in\Psi^M(P)} \left|\left\langle f, \psi(\cdot, N(\cdot))\mathbf{1}_{\omega_P^\flat}(N(\cdot))\right\rangle\right| \tag{2.1.3}$$

where $N : \mathbb{R} \to \mathbb{R}$ stands for a fixed measurable function, and $\Psi^M(P)$ is the modified class

$$\Psi^M(P) := \left\{ \phi = \phi(x,\nu) \in \mathcal{C}^M(\mathbb{R} \times \mathbb{R}) : \left[\frac{\partial_\nu}{\mathsf{scl}(P)}\right]^a \phi(\cdot,\nu) \in \Phi^M(P),\, \forall \nu \in \mathbb{R},\, a = 0,1 \right\}. \tag{2.1.4}$$

The dependence on the function N and on the smoothness-decay parameter M is kept implicit in the notation for (2.1.2)-(2.1.3), as these will be clear from context. For this reason, unless strictly necessary, M is dropped from the notations, writing for example $\Phi(P), \Psi(P)$.

2.1.5 Analysis of maximally modulated singular multipliers

The wave packet transforms (2.1.2) and (2.1.3) enters directly the discrete models of both the Carleson operator and of rank 1 multilinear multipliers such as the bilinear Hilbert transform. For the sake of motivation, here follows the reduction of the former family to the wave packet form (2.1.9) below.

Let $m \in L^\infty(\mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R} \setminus \{0\})$ be a smooth Hörmander-Mihlin multiplier, that is

$$\sup_{0 \le \alpha \le M} \sup_{\xi \ne 0} |\xi|^\alpha \left| m^{(\alpha)}(\xi) \right| \le 1 \tag{2.1.5}$$

for some large and unimportant M. In the next paragraph, we prove the pointwise estimate

$$\mathcal{C}f(x) \le \sum_{u=1}^{95} \sum_{\star \in \{+,-\}} \sum_{g=\{0,1,2\}} \sup_{N \in \mathbb{R}} \left| \sum_{P \in \mathbb{S}_g} |I_P| \langle f, \phi_P \rangle \psi^\star_{P,u}(x,N) \right|, \qquad x \in \mathbb{R} \tag{2.1.6}$$

where

$$\mathcal{C}f(x) = \sup_{N \in \mathbb{R}} |H_N f|, \qquad H_N f(x) := \int_{\mathbb{R}} m(\xi - N)\widehat{f}(\xi) \mathrm{e}^{ix\xi}\, \mathrm{d}\xi, \quad x \in \mathbb{R}$$

is the maximally modulated multiplier operator already introduced in (1.0.3), $\mathbb{S}_g := \mathbb{S}_{\mathcal{D}_0,\mathcal{D}_g}$ is the set of all tiles associated to the grids $\mathcal{D}_0, \mathcal{D}_g$, the functions $\phi_P, \psi^{\star}_{P,u}$ are uniform multiples of adapted wave packets from respectively $\Phi(P), \Psi(P)$, cf. (2.1.1)-(2.1.4), and

$$\mathrm{supp}_2\psi^{\pm}_{P,u} := \overline{\left\{N \in \mathbb{R} : \psi^{\pm}_{P,u}(\cdot, N) \neq 0\right\}} \subset Q^{\pm,u}_P := c(\omega_P) \mp \ell_{\omega_P}\left[7 + \tfrac{u}{4}, 9 + \tfrac{u}{4}\right]. \qquad (2.1.7)$$

Fix the parameters g, u and $\star = + \in \{+,-\}$. We claim that there exist dyadic grids $\mathcal{G}_j,\ j = 1, \ldots, 2^{18}$ with the property that for all $P \in \mathbb{S}_g$ there exists $j_P \in \{1, \ldots, 2^{18}\}$ and $J_P \in \mathcal{G}_{j_P}$ with

$$\mathrm{supp}_2\psi^{+}_{P,u} \subset J_P^{\mathsf{ch}(1,1)}, \qquad \omega_P \subset J_P^{\mathsf{ch}(1,2)}.$$

This is easily obtained by applying Remark 2.1.3 with $M = 8$ to the convex hull of $Q^{+,u}_P$ and ω_P, whose leftmost fourth contains $Q^{+,u}_P$ and is contained within the left half of the smoothing interval, and whose rightmost fourth contains ω_P, and is contained within the right half of the smoothing interval. We then define the tile $\tilde{P} = \tilde{P}(P) = I_{\tilde{P}} \times \omega_{\tilde{P}} \in \mathbb{S}_{\mathcal{H}_j\times\mathcal{G}_j}$ by

$$I_{\tilde{P}} := \text{ the unique } J \in \mathcal{H}_j \text{ with } c(I_P) \in J, \ell_J \ell_{J_P^{\mathsf{ch}(1,2)}} = 1, \qquad \omega_{\tilde{P}} := J_P^{\mathsf{ch}(1,2)}$$

where $\mathcal{H}_j$ is a fixed dual grid to $\mathcal{G}_j$. With this definition,

$$\mathrm{supp}_2\psi^{+}_{P,u} \in \omega^{\mathsf{b}}_{\tilde{P}(P)}. \qquad (2.1.8)$$

For $\tilde{P} \in \mathbb{S}_{\mathcal{H}_j\times\mathcal{G}_j}$, let $\mathbb{S}_g(\tilde{P}) := \{P \in \mathbb{S}_g : \tilde{P}(P) = \tilde{P}\}$. For each $N \in \mathbb{R}$, we then have

$$\sum_{P\in\mathbb{S}_g(\tilde{P})} |I_P|\langle f, \phi_P\rangle\psi^{+}_{P,u}(x, N) = \sum_{P\in\mathbb{S}_g(\tilde{P})} |I_P|\langle f, \phi_P\rangle\psi^{+}_{P,u}(x, N)\mathbf{1}_{\omega^{\mathsf{b}}_{\tilde{P}}}(N)$$

having used (2.1.8) in the first equality. By construction, it is then easily verified that

$$\#\mathbb{S}_g(\tilde{P}) \lesssim 1, \qquad \phi_P \in C\Phi(\tilde{P}),\ \psi^{+}_{P,u} \in C\Psi(\tilde{P}) \quad \forall P \in \mathbb{S}_g(\tilde{P})$$

with uniform constants over $\tilde{P} \in \mathbb{S}_{\mathcal{H}_j \times \mathcal{G}_j}$. Linearization of the suprema in (2.1.6), a passage to the adjoint followed by using the definitions of (2.1.2), (2.1.3), and a limiting argument thus allow us to reduce estimation of the operator (1.0.3) to proving uniform bounds for the forms

$$\mathsf{C}_{\mathbb{P}}(f_1, f_2) := \sum_{P \in \mathbb{P}} |I_P| W[f_1](P) A[f_2](P) \tag{2.1.9}$$

where $\mathbb{P}$ is a finite subset of $\mathbb{S} = \mathbb{S}_{\mathcal{D},\mathcal{D}'}$ for a fixed pair of dual grids $\mathcal{D}, \mathcal{D}'$, and the function $N(\cdot)$ in the definition (2.1.3) of $A[f_2](\cdot)$ is a fixed but arbitrary measurable function.

Proof of estimate (2.1.6). By splitting and symmetry, we may assume that m is supported on the positive half-line, and obtain the $\star = +$ term, whose superscript is omitted throughout. Let $\psi_u \in \mathcal{S}(\mathbb{R})$, $u = 1, \ldots, 95$ with

$$\operatorname{supp} \widehat{\psi_u} \subseteq \left[\tfrac{1}{2} + \tfrac{u-1}{64}, \tfrac{1}{2} + \tfrac{u+1}{64}\right], \qquad \sum_{u=1}^{95} \sum_{k \in \mathbb{Z}} \widehat{\psi_u}(2^k \xi) = \mathbf{1}_{(0,\infty)}(\xi)$$

and perform the corresponding Littlewood-Paley decomposition of the multiplier $H_0 f$ as

$$H_0 f = \sum_{u=1}^{95} \sum_{k \in \mathbb{Z}} f * \Psi_{k,u}, \qquad \Psi_{k,u}(x) := \int_{\mathbb{R}} m(\xi) \widehat{\psi_u}(2^k \xi) \mathrm{e}^{ix\xi} \, \mathrm{d}\xi, \quad x \in \mathbb{R}.$$

Further, let $\mathcal{D}_g$, $g = 0, 1, 2$ be the three 1/3-shifted dyadic grids on $\mathbb{R}$, and $\mathbb{S}_g(k) = \{P \in \mathbb{S}_g : \mathsf{scl}(P) = 2^k\}$ be the corresponding scale k tiles. Performing the standard Gabor decomposition, we pick $\phi \in \mathcal{S}(\mathbb{R})$ with $\operatorname{supp} \widehat{\phi} \subseteq \left[0, \tfrac{2}{3}\right]$ such that

$$\sum_{\lambda \in \mathbb{Z}} \left| \widehat{\phi}\left(\xi - \tfrac{\lambda}{3}\right) \right|^2 = 1, \qquad \xi \in \mathbb{R}$$

so that for each $k \in \mathbb{Z}$

$$f = \sum_{g=0}^{2} \sum_{P \in \mathbb{S}_g(k)} |I_P| \langle f, \phi_P \rangle \phi_P, \qquad \phi_P := \operatorname{Mod}_{c_{\omega_P}} \operatorname{Tr}_{c_{I_P}} \operatorname{Dil}^1_{\mathsf{scl}(P)} \phi$$

holds. Note that $\phi_P \in C_a\Phi^a(P)$ for all a. Combining and using the frequency support property of ψ_k to restrict the summation,

$$\begin{aligned} H_N f &= \sum_{u=1}^{95}\sum_{k\in\mathbb{Z}} f * \mathrm{Mod}_N \Psi_{k,u} = \sum_{u=1}^{95}\sum_{g=0}^{2}\sum_{k\in\mathbb{Z}}\sum_{P\in\mathbb{S}_g(k+4)} |I_P|\langle f,\phi_P\rangle \phi_P * \mathrm{Mod}_N\Psi_{k,u} \\ &= \sum_{u=1}^{95}\sum_{g=0}^{2}\sum_{P\in\mathbb{S}_g} |I_P|\langle f,\phi_P\rangle \psi_{P,u}(\cdot,N)\mathbf{1}_{[7+\frac{u}{4},9+\frac{u}{4}]}\left(\tfrac{c(\omega_P)-N}{\ell_{\omega_P}}\right) \end{aligned} \tag{2.1.10}$$

having defined the functions

$$\psi_{P,u}(x,N) \coloneqq \phi_P * \mathrm{Mod}_N \Psi_{k,u}(x), \qquad 2^{k+4} = \mathsf{scl}(P).$$

To obtain (2.1.10), we have used that $\operatorname{supp}\widehat{\mathrm{Mod}_N\Psi}_{k,u} \subset N + 2^{-k}[\frac12+\frac{u-1}{64}, \frac12+\frac{u+1}{64}]$ and that when $P\in\mathbb{S}_g(k+4)$ the frequency support of ϕ_P is an interval ω_P of length 2^{-k-4}. Thus, in order for $\psi_{P,u}(\cdot,N)$ to be nonzero, N must belong to the interval $Q_P^{+,u}$ as claimed in (2.1.7).

We are left with proving that $\psi_{P,u} \in C_a\Psi^a(P)$ for all $0\le a\le M-1$. To this aim, fix $P\in\mathbb{S}_g(k+4)$. We treat both cases $\alpha=0,1$ at the same time. First of all, using the Hörmander-Mihlin condition (2.1.5)

$$\operatorname{supp}\widehat{\Psi_{k,u}} \subset [2^{-k-1}, 2^{-k+1}], \qquad \left| D^a\widehat{\Psi_{k,u}}(\xi)\right| \lesssim_M 2^{ka} \sim_M |\xi|^{-a}$$

for all $0\le a\le M$, $k\in\mathbb{Z}$. Let also β be an auxiliary Schwartz function with the property that $1_{[-\frac12,\frac12]} \le \beta \le 1_{[-1,1]}$ and define

$$\Phi_{k,u,P}(x,N) \coloneqq \int_{\mathbb{R}} \left(-\tfrac{D}{\mathsf{scl}(P)}\right)^{\alpha} \widehat{\Psi_{k,u}}(\xi)\beta\left(\frac{\xi+N-c_{\omega_P}}{\ell_{\omega_P}}\right) \mathrm{e}^{ix\xi}\,\frac{\mathrm{d}\xi}{\sqrt{2\pi}}, \qquad x\in\mathbb{R}.$$

Using the Fourier transform and the definition, we check that

$$\left[\frac{\partial_N}{\mathsf{scl}(P)}\right]^{\alpha} \psi_{P,u}(\cdot,N) = \phi_P * \mathrm{Mod}_N \Phi_{k,u,P}$$

so our claim follows easily from the scale $\mathsf{scl}(P) \sim 2^k$ bump function estimates for the function $\Xi = \mathrm{Mod}_{N-c_{\omega_P}} \Phi_{k,u,P}$, whose Fourier transform is supported on $|\xi| \le \ell_{\omega_P} \sim 2^{-k}$ and satisfies

$$\widehat{\Xi}(\xi) = \left(-\tfrac{D}{\mathsf{scl}(P)}\right)^{\alpha} \widehat{\Psi}_{k,u}\big(\xi - (N - c_{\omega_P})\big)\beta\left(\frac{\xi}{\ell_{\omega_P}}\right),$$

$$\left|\widehat{\Xi}^{(a)}(\xi)\right| \lesssim_a \sum_{b+c=a} 2^{-k\alpha} \left|\widehat{\Psi}^{(b)}_{k,u}\big(\xi - (N - c_{\omega_P})\big)\right| 2^{kc} \lesssim_a 2^{-k\alpha} \left|\xi - (N - c_{\omega_P})\right|^{-b} 2^{kc} \lesssim_a 1$$

for $0 \le a \le M-1$, having used that $c_{\omega_P} - N \ge 3\ell_{\omega_P}$, while $|\xi| \le \ell_{\omega_P}$ on the support of $\beta(\cdot/\ell_{\omega_P})$. This completes the proof of (2.1.6). $\square$

2.2 Outer L^p estimates for the wave packet transforms

Outer L^p spaces, introduced in this context by Do and Thiele [26], provide the functional setting for our estimates on the wave packet transforms. In this section, after particularizing the main definitions, we introduce two new outer L^p norms enjoying a weaker, but more precisely quantified form of the outer Hölder inequality. In what follows, we refer to a fixed tiling $\mathbb{S} = \mathbb{S}_{\mathcal{D},\mathcal{D}'}$.

2.2.1 Trees

Let κ be a nonnegative integer. We say that $T \subset \mathbb{S}$ is a κ-tree if there exists an interval $I_T \in \mathcal{D}$ and a frequency $\xi_T \in \mathbb{R}$ such that

$$I_P \subset I_T, \quad \xi_T \in \omega_P^{\mathsf{p}(\kappa)} \qquad \forall P \in T.$$

The pair (I_T, ξ_T) is referred to as *top data* of T. The notation

$$\mathcal{I}(T) := \{I \in D : I = I_P \text{ for some } P \in T\}, \qquad \Omega(T) := \{\omega \in \mathcal{D}' : \omega = \omega_P \text{ for some } P \in T\}$$

is used for the spatial and frequency components of a κ-tree T.

Let $1 \le j \le 2^\kappa$. We say that a κ-tree T is of type j if $\omega_P = [\omega_P^{\mathsf{p}(\kappa)}]^{\mathsf{ch}(\kappa,j)}$, that is, equals the j-th κ-grandchild of its κ-parent, for all $P \in T$. Clearly any κ-tree T splits as the disjoint union $T = \bigsqcup_{j=1}^{2^\kappa} T_{|j}$, with each $T_{|j}$ being a κ-tree of type j with the same top data.

Remark 2.2.2. The structure of $\mathbb{S}$ and the above definition entails that the intervals $\{\omega^{\mathsf{p}(\kappa)} : \omega \in \Omega(T)\}$ are nested. Therefore, $\#\{\omega \in \Omega(T) : \ell_\omega = \rho\} \le 2^\kappa$ for all $\rho > 0$. As a first consequence,

$$\#\{P \in T : I_P = I\} \le 2^\kappa \qquad \forall I \in \mathcal{I}(T). \tag{2.2.1}$$

In general, each tree T contains both a Littlewood-Paley type and a maximal function type component. The next definition isolates the Littlewood-Paley part. Say that a κ-tree T is *lacunary* if

$$\omega, \omega' \in \Omega(T), \omega \ne \omega' \implies \omega \cap \omega' = \varnothing.$$

and for every tree T, split

$$T = T^{\mathsf{ov}} \cup T^{\mathsf{lac}}, \qquad T^{\mathsf{ov}} := \{P \in T : \xi_T \in \omega_P\}, \qquad T^{\mathsf{lac}} := T \setminus T^{\mathsf{ov}}. \tag{2.2.2}$$

The next lemma tells us in particular that T^{lac} is a union of at most $\kappa 2^\kappa$ lacunary trees, and that the residual part T^{ov} has additional structure.

Lemma 2.2.3 (Structure of trees)**.** *Let T be a κ-tree with top data (I_T, ξ_T). Then $T = \bigsqcup_{u=1}^{\kappa} T^u$, with each T^u also a κ-tree with the same top data and such that, for all $j = 1, \dots, 2^\kappa$*

(i) $\left[T^u_{|j}\right]^{\mathsf{lac}}$ *is a lacunary tree;*

(ii) *whenever* $j' \neq j$, *the intervals* $\left\{\left[\omega_P^{\mathsf{p}(\kappa)}\right]^{\mathsf{ch}(\kappa,j')} : P \in \left[T^u_{|j}\right]^{\mathsf{ov}}\right\}$ *are pairwise disjoint.*

Proof. We set $T^u := \left\{P \in T : \mathsf{scl}(P) \in 2^{\kappa\mathbb{Z}+u}\right\}$ for $1 \leq u \leq \kappa$. Since every $P \in T^u$ belongs in T we have that $\xi_T \in \omega_P^{\mathsf{p}(\kappa)}$ and $I_P \subset I_T$ therefore T^u is a tree with top data (I_T, ξ_T). We prove the first claim by taking $\omega, \omega' \in \Omega\left(\left[T^u_{|j}\right]^{\mathsf{lac}}\right)$ for which we can assume without loss of generality that $\ell_\omega \leq \ell_{\omega'}$ and $\omega \neq \omega'$. If $\ell_\omega = \ell_{\omega'}$ it is clear that they cannot intersect. In the case $\ell_\omega < \ell_{\omega'}$ then $2^\kappa \ell_\omega < \ell_{\omega'}$ henceforth if $\omega \cap \omega' \neq \varnothing$ we would have that $\omega^{\mathsf{p}(\kappa)} \subset \omega'$ which implies that $\xi_T \in \omega'$ which is absurd by the defintion of $\left[T^u_{|j}\right]^{\mathsf{lac}}$. For the second claim we proceed similarly, noting that the case $\ell_\omega = \ell_{\omega'}$ we would have that $\omega = \omega'$, and therefore we can assume $2^\kappa \ell_\omega < \ell_{\omega'}$ so that if ω and ω' intersect we would have $\omega^{\mathsf{p}(k)} \subset \omega' \Rightarrow \xi_T \in \omega'$ but by the definition of $T_{|j}$ we have that $\xi_T \in \left[\omega^{\mathsf{p}(\kappa)}\right]^{\mathsf{ch}(\kappa,j)}$ which does not intersect ω' so we arrive at a contradiction. □

2.2.4 Outer L^p on the space of local tiles

For $J \in \mathcal{D}$, let $\mathbb{S}^J$ be the collection of all tiles $P \in \mathbb{S}$ with $I_P \subset J$. Below, the notation $\ell^p(\mathbb{S}^J)$ stands for the ℓ^p spaces on $\mathbb{S}^J$endowed with the weighted counting measure

$$A \mapsto \sum_{P \in A} |I_P|, \qquad A \subset \mathbb{S}^J.$$

The collection $\mathcal{T}^{J,\kappa}$ of all κ-trees $T \subset \mathbb{S}^J$ concurs to the definition of the outer measure space $(\mathbb{S}^J, \mathcal{T}^{J,\kappa}, \mu^{J,\kappa})$, with outer measure $\mu^{J,\kappa}$ defined by

$$\mu^{J,\kappa} : \mathcal{P}(\mathbb{S}^J) \to [0,\infty], \qquad \mu^{J,\kappa}(A) := \inf\left\{\frac{1}{|J|}\sum_{T \in \mathcal{T}} |I_T| : \mathcal{T} \subset \mathcal{T}^{J,\kappa}, A \subset \bigcup_{T \in \mathcal{T}} T\right\}; \quad (2.2.3)$$

to wit, the infimum above is taken over all collections $\mathcal{T} \subset \mathcal{T}^{J,\kappa}$ of κ-trees whose union covers A. Below, for a quasi-subadditive size map s as defined in [26, Def. 2.3],

$$\left[F : \mathbb{S}^J \to \mathbb{C}\right] \mapsto \left\{\mathsf{s}(F,T) : T \in \mathcal{T}^{J,\kappa}\right\}$$

we consider the outer $L^{p,r}$ space on $(\mathbb{S}^J, \mathcal{T}^{J,\kappa}, \mu^{J,\kappa})$,

$$L^{p,r}(J,\kappa,\mathsf{s}) = L^{p,r}(\mathbb{S}^J, \mu^{J,\kappa}, \mathsf{s})$$

as defined in [26, Def. 3.2], for exponents $1 \leq p, q \leq \infty$. The definition therein may be summarized as follows. First of all, define the outer essential supremum

$$\operatorname*{outsup}_{\mathsf{s}} F := \sup_{T \in \mathcal{T}^{J,\kappa}} \mathsf{s}(F,T) =: \|F\|_{L^\infty(\mathsf{s})} = \|F\|_{L^{\infty,\infty}(\mathsf{s})}.$$

Secondly, define the super level measure $\mu_{\mathsf{s}}[F] : [0,\infty) \to [0,\infty]$ and the corresponding nondecreasing rearrangement $F^{*,\mathsf{s}} : [0,\infty) \to [0,\infty]$ respectively by

$$\mu_{\mathsf{s}}[F](\tau) := \inf \left\{ \mu^{J,\kappa}(A) : \operatorname*{outsup}_{\mathsf{s}}(F \mathbf{1}_{\mathbb{S}^J \setminus A}) \leq \tau \right\},$$

$$F^{*,\mathsf{s}}(t) := \inf \left\{ \tau \in [0,\infty) : \mu_{\mathsf{s}}[F](\tau) \leq t \right\}.$$

We then set

$$\|F\|_{L^{p,r}(J,\kappa,\mathsf{s})} := \|F^{*,\mathsf{s}}\|_{p,q} = \left\| t^{\frac{1}{p}} F^{*,\mathsf{s}}(t) \right\|_{L^q\left([0,\infty), \frac{\mathrm{d}t}{t}\right)}; \tag{2.2.4}$$

recall that the right hand side is the standard Lorentz $L^{p,r}$ quasinorm on $[0,\infty]$, see e.g. [34, Sect. 1.4]. As customary, when $q = p$ we omit q from the subscripts and superscripts.

The main examples of sizes and associated outer $L^{p,r}$ spaces that arise in our applications are the following. For $1 \leq p \leq \infty$, set

$$\mathsf{size}_p(F,T) := \frac{\|F \mathbf{1}_T\|_{\ell^p(\mathbb{S}^J)}}{|I_T|^{\frac{1}{p}}}, \qquad T \in \mathcal{T}^{J,\kappa}.$$

For $p=2$, we define the variant

$$\mathsf{size}_{2,\star}(F,T) \coloneqq \sup\left\{\mathsf{size}_2(F,U) : U \in \mathcal{T}^{J,\kappa} \text{ lacunary}, U \subset T\right\}, \qquad T \in \mathcal{T}^{J,\kappa}, \tag{2.2.5}$$

which is also a size. The definition of $\mathsf{size}_p(F,\cdot)$ and $\mathsf{size}_{2,\star}(F,\cdot)$ depends on κ via the domain $\mathcal{T}^{J,\kappa}$, though we do not keep this dependence explicit in the notation.

The modified wave packet transform acting on the dual side of the Carleson operator, in accordance with the definition (2.1.3) involving ω_P^{b}, will be estimated in outer $L^{p,r}$-spaces (2.2.4) where the parameter κ is naturally chosen to be 1. On the outer measure space (2.2.3) we thus define, with reference to(2.2.2)

$$\mathsf{size}_{\mathsf{C}}(F,T) \coloneqq \mathsf{size}_2(F,T^{\mathsf{lac}}) + \mathsf{size}_1(F,T^{\mathsf{ov}}), \qquad T \in \mathcal{T}^{J,1}.$$

The next proposition is a generalization to the Lorentz scale of the outer Hölder inequality, which plays a pivotal rôle in the applications of outer spaces to modulation invariant singular integrals.

Proposition 2.2.5. *Let $m \in \mathbb{N}_{\geq 2}$ and $\mathsf{s}, \mathsf{s}_1, \mathsf{s}_2, \cdots, \mathsf{s}_m$ be sizes on $(\mathbb{S}^J, \mu^{J,\kappa}, \mathcal{T}^{J,\kappa})$ with the property that for all function m-tuples $F_1, \cdots, F_m : \mathbb{S}^J \to \mathbb{C}$,*

$$\mathsf{s}\left(\prod_{j=1}^m F_j, T\right) \leq \prod_{j=1}^m \mathsf{s}_j(F_j,T) \qquad \forall T \in \mathcal{T}^{J,\kappa}. \tag{2.2.6}$$

Then for all tuples $0 < p, p_1, \ldots, p_m,\ q, q_1, \ldots, q_m \leq \infty$, $\frac{1}{p} = \sum_{j=1}^m \frac{1}{p_j}$, $\frac{1}{q} = \sum_{j=1}^m \frac{1}{q_j}$ there holds

$$\left\|\prod_{j=1}^m F_j\right\|_{L^{p,q}(J,\kappa,\mathsf{s}_\ell)} \leq m^{\frac{1}{p}} \prod_{j=1}^m \|F_j\|_{L^{p_j,q_j}(J,\mathsf{s}_j)}$$

Proof. Chasing definitions, it is immediate to see that

$$\left[\prod_{j=1}^{m} F_j\right]^{*,\mathsf{s}} \left(\frac{t}{m}\right) \leq \prod_{j=1}^{m} F_j^{*,\mathsf{s}_k}(t), \qquad 0<t<\infty$$

and the claim follows from the usual Hölder inequality on the spaces $L^{q_j}\left([0,\infty), \frac{\mathrm{d}t}{t}\right)$. $\square$

Remark 2.2.6. Let the assumptions of Proposition 2.2.5 stand, and particularize to the case $\mathsf{s} = \mathsf{size}_1$ and $p = q = 1$. Then,

$$\frac{1}{|J|}\left\|\prod_{j=1}^{m} F_j\right\|_{\ell^1(\mathbb{S}^J)} \lesssim \left\|\prod_{j=1}^{m} F_j\right\|_{L^1(J,\kappa,\mathsf{size}_1)} \leq m\prod_{j=1}^{m}\|F_j\|_{L^{p_j}(J,\kappa,\mathsf{s}_j)} \tag{2.2.7}$$

where [26, Prop. 3.6] has been used to get the first bound.

Remark 2.2.7. Let $A \subset \mathbb{S}^J$ be a set of finite outer measure $\mu^{J,\kappa}$. It may be checked directly that $[\mathbf{1}_A]^{*,\mathsf{size}_\infty} = \mathbf{1}_{[0,\mu^{J,\kappa}(A))}$, so that in particular $\|\mathbf{1}_A\|_{L^{p,\infty}(J,\kappa,\mathsf{size}_\infty)} = \mu^{J,\kappa}(A)^{\frac{1}{p}}$. Using monotonicity of the size s, a particular case of Proposition 2.2.5 is

$$\|F\mathbf{1}_A\|_{L^{p,q}(J,\kappa,\mathsf{s})} \leq 2^{\frac{1}{p}}\|F\mathbf{1}_A\|_{L^{p_1,q}(J,\kappa,\mathsf{s})}\,\mu^{J,\kappa}(A)^{\frac{1}{p}-\frac{1}{p_1}}, \qquad 0<p\leq p_1\leq\infty,\ 0<q\leq\infty. \tag{2.2.8}$$

2.2.8 Reverse Hölder outer L^p norms

The next definition is inspired by Remark 2.2.7. Let s be any size on $(\mathbb{S}^J, \mu^{J,\kappa}, \mathcal{T}^{J,\kappa})$, cf. [26, Def. 2.3]. Define, for $F : \mathbb{S}^J \to \mathbb{C}$, $1 \leq a \leq p < \infty$, $1 \leq q \leq \infty$ and $\varepsilon > 0$, the quasi-norms

$$\|F\|_{X_a^{p,q}(J,\kappa,\mathsf{s})} := \sup_{A\subset\mathbb{S}^J}\frac{\|F\mathbf{1}_A\|_{L^{a,q}(J,\kappa,\mathsf{s})}}{\mu^{J,\kappa}(A)^{\frac{1}{a}-\frac{1}{p}}}, \qquad \|F\|_{Y^{p,q}(J,\kappa,\mathsf{s})} := \max\left\{\|F\|_{L^{p,q}(J,\kappa,\mathsf{s})}, \|F\|_{L^\infty(J,\kappa,\mathsf{s})}\right\}.$$

Remark 2.2.7 tells us immediately that $\|F\|_{X_a^{p,q}(J,\kappa,\mathsf{s})} \leq 2^{\frac{1}{p}}\|F\|_{L^{p,q}(J,\kappa,\mathsf{s})}$ in the range of the definition. The next proposition should be interpreted as a partial converse of this control and as a substitute for Proposition 2.2.5 with a smaller right hand side. The $Y^{q,\infty}(J,\kappa,\mathsf{s})$-norm is the quantity appearing in our applications.

Proposition 2.2.9. *Let $m \in \mathbb{N}_{\geq 2}$ and $\mathsf{s}_1, \mathsf{s}_2, \cdots, \mathsf{s}_m$ be m sizes on $(\mathbb{S}^J, \mu^{J,\kappa}, \mathcal{T}^{J,\kappa})$ with the property that* (2.2.6) *holds with $\mathsf{s} = \mathsf{size}_1$. Suppose that*

$$1 < a \leq p_1 < \infty, \qquad 1 \leq p_2, \ldots, p_m < \infty, \qquad \varepsilon := \left(\sum_{\ell=1}^{m} \tfrac{1}{p_\ell}\right) - 1 > 0.$$

Then, with implicit constant possibly depending on m only, there holds

$$\frac{1}{|J|}\left\|\prod_{j=1}^{m} F_j\right\|_{\ell^1(\mathbb{S}^J)} \lesssim \frac{a}{\varepsilon\,(a-1)}\,\|F_1\|_{X_a^{p_1,\infty}(J,\kappa,\mathsf{s}_1)} \prod_{\ell=2}^{m} \|F_\ell\|_{Y^{p_\ell,\infty}(J,\kappa,\mathsf{s}_\ell)}$$

Proof of Proposition 2.2.9. Throughout the proof, the constant implied by $\lesssim$ is allowed to depend on m only and vary at each occurrence By scaling we can assume

$$\|F_1\|_{X_a^{p_1,\infty}(J,\kappa,\mathsf{s}_1)} = \|F_2\|_{Y^{p_2,\infty}(J,\kappa,\mathsf{s}_2)} = \cdots = \|F_m\|_{Y^{p_m,\infty}(J,\kappa,\mathsf{s}_m)} = 1.$$

Under this assumption, we must prove

$$\frac{1}{|J|}\sum_{P \in \mathbb{S}^J} |I_P||F_1F_2F_3\cdots F_m(P)| \lesssim \frac{a}{\varepsilon\,(a-1)}. \tag{2.2.9}$$

Relying on the controls $\|F_\ell\|_{L^{p_\ell,\infty}(J,\kappa,\mathsf{s}_\ell)}, \|F_\ell\|_{L^\infty(J,\kappa,\mathsf{s}_\ell)} \leq 1$ for all $2 \leq \ell \leq m$, we iteratively decompose the support of $F_2F_3\cdots F_m$ into pairwise disjoint sets A_j, $j \in \mathbb{N}$ such that

$$\mu^{J,\kappa}(A_j) \leq 2^j, \qquad \max_{2\leq \ell \leq m} 2^{\frac{j}{p_\ell}} \operatorname*{outsup}_{\mathsf{s}_\ell}(F_\ell \mathbf{1}_{A_j}) \lesssim 1. \tag{2.2.10}$$

For $j \in \mathbb{N}$, let $k(j)$ be the largest integer k with $k \leq \frac{aj}{p_1}$. From the first estimate in (2.2.10) and the definition of $X_a^{p_1,\infty}(J,\kappa,\mathsf{s}_1)$-norm, we learn that

$$\|F_1 \mathbf{1}_{A_j}\|_{L^{a,\infty}(J,\kappa,\mathsf{s}_1)} \leq 2^{j\left(\frac{1}{a}-\frac{1}{p_1}\right)}.$$

Thus, we may further decompose A_j into pairwise disjoint sets $\{B_{j,k} : -N \leq k \leq k(j)\}$, where N is an unimportant parameter related to the outer essential supremum of F_1, with

$$\operatorname*{outsup}_{\mathsf{s}_1}(F_1 \mathbf{1}_{B_{j,k}}) \leq 2^{-\frac{k}{a}}, \qquad \mu^{J,\kappa}(B_{j,k}) \leq 2^{k+j\left(1-\frac{a}{p_1}\right)}, \tag{2.2.11}$$

which means that we may find $\mathcal{T}_{j,k} \subset \mathcal{T}^{J,\kappa}$ with the property

$$B_{j,k} \subset \bigcup_{T \in \mathcal{T}_{j,k}} T, \qquad \sum_{T \in \mathcal{T}_{j,k}} \frac{|I_T|}{|J|} \leq 2\mu^{J,\kappa}(B_{j,k}) \lesssim 2^{k+j\left(1-\frac{a}{p_1}\right)}. \tag{2.2.12}$$

We then estimate, using (2.2.10), (2.2.11) and (2.2.12) and subsequently summing in k,

$$\begin{aligned}
&\frac{1}{|J|}\sum_{P \in A_j} |I_P||F_1F_2F_3\cdots F_m(P)| \leq \sum_{-N\leq k \leq k(j)} \sum_{T \in \mathcal{T}_{j,k}} |I_T| \mathsf{size}_1(F_1F_2F_3\cdots F_m \mathbf{1}_{B_{j,k}}, T) \\
&\lesssim \sum_{-N\leq k \leq k(j)} \sum_{T \in \mathcal{T}_{j,k}} |I_T| \mathsf{s}_1(F_1 \mathbf{1}_{B_{j,k}}, T) \prod_{\ell=2}^{m} \mathsf{s}_\ell(F_\ell \mathbf{1}_{A_j}, T) \leq 2^{j\left(1-\frac{a}{p_1}-\sum_{\ell=2}^m \frac{1}{p_\ell}\right)} \sum_{-N\leq k\leq k(j)} 2^{\frac{a-1}{a}k} \\
&\lesssim \frac{a}{a-1} 2^{j\left(1-\frac{a}{p_1}-\sum_{\ell=2}^m \frac{1}{p_\ell}\right)} 2^{\frac{a-1}{a}k(j)} \lesssim \frac{a}{a-1} 2^{j\left(1-\sum_{\ell=1}^m \frac{1}{p_\ell}\right)} = \frac{a}{a-1} 2^{-\varepsilon j}.
\end{aligned}$$

The claimed bound (2.2.9) follows by summing the estimate of the last display over $j \in \mathbb{N}$. □

2.2.10 Lacunary tree estimates

This paragraph contains some $\text{size}_{2,\star,\kappa}$ estimates for $W[f]$ restricted to lacunary trees, which we use to explain the role played by this type of trees, and that will also be of use later.

Throughout our first discussion, let T be a lacunary tree with top data (I_T, ξ_T). For simplicity, we assume $\xi_T = 0$, as the general case of our observations can be recovered by suitably pre- and post-composing with $\mathrm{Mod}_{\pm\xi_T}$. Disjointness of frequency support and rapid decay tell us that whenever $P, P' \in T$ and $\phi_P \in \Phi(P), \phi_{P'} \in \Phi(P')$,

$$\ell_{I_P} = \ell_{I_{P'}} \implies |\langle \phi_P, \phi_{P'} \rangle| \lesssim |I_P|^{-1} \left\langle \frac{c_{I_P} - c_{I_{P'}}}{\ell_{I_P}} \right\rangle^{-M}, \qquad \ell_{I_P} \neq \ell_{I_{P'}} \implies \langle \phi_P, \phi_{P'} \rangle = 0.$$

This observation and standard kernel estimates tell us that the operator

$$f \mapsto H_T f := \sum_{P \in T} |I_P| \langle f, \phi_P \rangle \varphi_P, \qquad \phi_P, \varphi_P \in \Phi(P) \quad \forall P \in T$$

and its adjoint are standard L^2-bounded Calderón-Zygmund operators. Thus, Calderón-Zygmund theory and the localization trick yield in particular that

$$\frac{1}{|I_T|} \|H_T f\|_{1,\infty} \lesssim \langle\!\langle f \rangle\!\rangle_{1,I_T}, \qquad \frac{1}{|I_T|^{\frac{1}{p}}} \|H_T f\|_p \lesssim_p \langle\!\langle f \rangle\!\rangle_{p,I_T},$$

the latter inequality being true for all $1 < p < \infty$. In particular

$$\mathsf{size}_2(W[f], T) \sim |I_T|^{-\frac{1}{2}} \|H_T f\|_2 \lesssim \langle\!\langle f \rangle\!\rangle_{2,I_T} \lesssim \|f\|_\infty$$

with ϕ_P, φ_P suitably chosen so that the first absolute equivalence holds. We have just proved the outer estimate

$$\|W[f]\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \lesssim_p \|f\|_\infty. \tag{2.2.13}$$

The more precise localized estimate of the next proposition may be proved using a semi-discrete analogue of H_T and the John-Strömberg inequality. The argument is a variation on [44, Prop. 9.3]. Associate to a collection of tiles $\mathbb{P} \subset \mathbb{S}$ and $f \in L_0^\infty(\mathbb{R})$ the quasinorms

$$[f]_{p,\mathbb{P}} := \sup_{P \in \mathbb{P}} \inf_{I_P} \mathrm{M}_p f, \qquad 0 < p < \infty. \tag{2.2.14}$$

Proposition 2.2.11. $\|W[f]\mathbf{1}_{\mathbb{P}}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \lesssim \langle \operatorname{dist}(J, \operatorname{supp} f)\rangle^{-2^8} [f]_{1,\mathbb{P}}$.

Proof. There is no loss in generality with assuming $\mathbb{P} \subset \mathbb{S}^J$. For $\xi \in \mathbb{R}$, denote by $T_\xi = \{P \in \mathbb{P} : \xi \in \omega_P^{\mathsf{p}(\kappa)}\}$. Note that T_ξ is a tree with top data (J, ξ). Then

$$\|W[f]\mathbf{1}_{\mathbb{P}}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \leq 2 \sup_{\xi\in\mathbb{R}} \sup_{T\subset T_\xi} |I_T|^{-\frac{1}{2}} \|\langle f, \phi_P\rangle \mathbf{1}_T(P)\|_{\ell^2_P(\mathbb{S}^J)} \tag{2.2.15}$$

for suitably chosen $\phi_P \in \Phi(P)$. So we fix ξ and estimate $\sup_{T\subset T_\xi} \|\langle f, \phi_P\rangle \mathbf{1}_T(P)\|_{\ell^2_P(\mathbb{S}^J)}$. By composing with modulations, we may reduce to $\xi = 0$, and by (2.2.1) and finite splitting, we may also reduce to having $\#\{P \in T_\xi : I_P = I\} = 1$ for all $I \in \mathcal{I}(T_\xi)$. Then

$$\sup_{T\subset T_\xi} \frac{\|\langle f, \phi_P\rangle \mathbf{1}_T(P)\|_{\ell^2_P(\mathbb{S}^J)}}{|I_T|^{\frac{1}{2}}} \leq \sup_{\substack{K\in\mathcal{D}\\K\subset J}} \frac{\left\|\langle f, \phi_P\rangle \mathbf{1}_{\{P\in T_\xi : I_P\subset K\}}\right\|_{\ell^2_P(\mathbb{S}^J)}}{|K|^{\frac{1}{2}}} = \left\| \sum_{I\in\mathcal{I}(T_\xi)} \langle f, \varphi_I\rangle h_I \right\|_{\mathrm{BMO}} \tag{2.2.16}$$

where we have set $\varphi_I = \sqrt{|I_P|}\phi_P$ for the unique $P \in T_\xi$ with $I_P = I$, h_I stands for the L^2-normalized Haar wavelet on I, and we mean the dyadic BMO. For $K \in \mathcal{D}, K \subset J$, let $\mathcal{I}^*(K)$ be the collection of maximal intervals in $I \in \mathcal{I}(T_\xi)$ with $I \subset K$. The John-Strömberg inequality, followed by disjointness of $I \in \mathcal{I}^*(K)$ tells us that

$$\left\| \sum_{I\in\mathcal{I}(T_\xi)} \langle f, \varphi_I\rangle h_I \right\|_{\mathrm{BMO}} \lesssim \sup_{\substack{K\in\mathcal{D}\\K\subset J}} \frac{1}{|K|} \left\| \sum_{\substack{I\in\mathcal{I}(T_\xi)\\I\subset K}} \langle f, \varphi_I\rangle h_I \right\|_{1,\infty} = \sup_{\substack{K\in\mathcal{D}\\K\subset J}} \sum_{I\in\mathcal{I}^*(K)} \frac{\|H_{I,\mathrm{semi}} f\|_{1,\infty}}{|K|} \tag{2.2.17}$$

having set

$$H_{I,\mathrm{semi}} f := \sum_{\substack{J\in\mathcal{I}(T_\xi)\\J\subset I}} \langle f, \phi_J\rangle h_J.$$

Standard kernel computations tell us that $H_{I,\mathrm{semi}}$ is also an L^2-bounded Calderón-Zygmund operator and in particular is uniformly of type weak-$(1,1)$. Combining with the localization trick on $I \in \mathcal{I}^*(K)$,

$$\|H_{I,\mathrm{semi}} f\|_{1,\infty} \lesssim |I| \langle\!\langle f \rangle\!\rangle_{1,I} \lesssim |I| \inf_I \mathrm{M}_1 f \leq |I| [f]_{1,\mathbb{P}}. \tag{2.2.18}$$

Inserting the estimate (2.2.18) into (2.2.17), summing over the disjoint $I \in \mathcal{I}^*(K)$, and perusing (2.2.15)-(2.2.16) yields the partial bound $\|W[f]\mathbf{1}_{\mathbb{P}}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \lesssim [f]_{1,\mathbb{P}}$. The additional decay factor may be easily obtained by a localization trick followed by the partial result applied to $f\chi_J$ in place of f. $\square$

The following technical lemma will allow us to estimate the $L^\infty(J,\kappa,\mathsf{size}_{2,\star})$ norm of the wave packet transform restricted to a collection $\mathbb{P}$ which is covered by a certain set of top data. It will not be used until Section 2.4, but this is the most appropriate location for its proof. Notice that $T(I,\xi)$ appearing in the statement that follows is a κ-tree with top data (I,ξ).

Lemma 2.2.12. *Let $\mathbb{P} \subset \mathbb{S}$ and $\mathcal{F} \subset \mathcal{D} \times \mathbb{R}$ be a collection of top data covering $\mathbb{P}$, in the sense that*

$$\mathbb{P} = \bigcup_{(I,\xi)\in\mathcal{F}} T(I,\xi), \qquad T(I,\xi) \coloneqq \left\{ P \in \mathbb{P} : I_P \subset I, \xi \in \omega_P^{\mathsf{p}(\kappa)} \right\}.$$

Then

$$\|W[f]\mathbf{1}_{\mathbb{P}}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \leq 2^{\frac{\kappa}{2}} \sup_{(I,\xi)\in\mathcal{F}} \mathsf{size}_{2,\star,\kappa}(W[f], T(I,\xi)).$$

Proof. There is no loss in generality with assuming $\mathbb{P} \subset \mathbb{S}^J$, and we do so. Fix a lacunary κ-tree $T \subseteq \mathbb{P}$ and let (I_T, ξ_T) be its top data. Note that (I_T, ξ_T) does not necessarily belong to $\mathcal{F}$. Say that $P \in \mathbb{P}^{T,\star}$ if $P \in T$ and I_P is a maximal element of $\mathcal{I}(T)$ with respect to inclusion. By assumption, for each $P \in \mathbb{P}^{T,\star}$ we may find $(I(P), \xi(P)) \in \mathcal{F}$ with $I_P \subset I(P)$ and $\xi(P) \in \omega_P^{\mathsf{p}(\kappa)}$. Clearly

$$T = \bigcup_{P \in \mathbb{P}^{T,\star}} T(P), \qquad T(P) \coloneqq \{Q = I_Q \times \omega_Q \in T : I_Q \subset I_P\}.$$

The fact that T is a tree guarantees if $Q \in T(P)$ then $\xi_T \in \omega_P^{\mathsf{p}(\kappa)} \cap \omega_Q^{\mathsf{p}(\kappa)}$, and comparing scales $\xi(P) \in \omega_P^{\mathsf{p}(\kappa)} \subset \omega_Q^{\mathsf{p}(\kappa)}$. Therefore $T(P)$ is a κ-lacunary tree with top data $(I_P, \xi(P))$, whence the inclusion $T(P) \subset T(I(P), \xi(P))$ for all $P \in \mathbb{P}^{T,\star}$, and

$$\mathsf{size}_2(W[f], T(P)) \leq \mathsf{size}_{2,\star,\kappa}(W[f], T(I(P), \xi(P))) \leq \sup_{(I,\xi) \in \mathcal{F}} \mathsf{size}_{2,\star,\kappa}(W[f], T(I, \xi)).$$

Using (2.2.1) and disjointness of the maximal elements of $\mathcal{I}(T)$, which are all contained in I_T,

$$\mathsf{size}_2(W[f], T) \leq \left(\frac{1}{|I_T|} \sum_{P \in \mathbb{P}^{T,\star}} |I_P| \big[\mathsf{size}_2(W[f], T(P))\big]^2 \right)^{\frac{1}{2}} \leq 2^{\frac{\kappa}{2}} \sup_{(I,\xi) \in \mathcal{F}} \mathsf{size}_{2,\star,\kappa}(W[f], T(I, \xi))$$

which completes the proof of our main claim. □

2.2.13 Local L^2-bound for maximal modulations via wave packet estimates

In this paragraph, as a motivating example, two more outer L^p estimates for the wave packet transforms (2.1.2)-(2.1.2) are stated and combined into a proof of L^p-boundedness for the maximal modulated singular multiplier of (1.0.3) in the local L^2-range. The first concerns the wave packet transform (2.1.2)

Proposition 2.2.14. *Let $J \in \mathcal{D}$ and $f \in L_0^\infty(\mathbb{R})$. Then*

$$\|W[f]\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} \lesssim \langle\!\langle f \rangle\!\rangle_{2,3J} \tag{2.2.19}$$

$$\|W[f]\|_{L^{p}(J,\kappa,\mathsf{size}_{2,\star})} \lesssim_p \langle\!\langle f \rangle\!\rangle_{p,3J}, \qquad 2 < p \leq \infty. \tag{2.2.20}$$

The bound (2.2.19) is a restatement of [26, Theorem 5.1], see also [16,23]. Once (2.2.19) is at disposal, (2.2.20) follows immediately from its outer L^p interpolation with e.g. (2.2.13); an appropriate interpolation theorem is [26, Prop. 3.5]. A similar, but broader set of estimates is available for the $L^p(\mathsf{size}_\mathsf{C}, J)$ norms of (2.1.3). As anticipated, the outer norms below refer to the case $\kappa = 1$.

Proposition 2.2.15. *Let $J \in \mathcal{D}$ and $f \in L_0^\infty(\mathbb{R})$. Then*

$$\|A[f\mathbf{1}_{3J}]\|_{L^{1,\infty}(J,1,\mathsf{size}_\mathsf{C})} \lesssim \langle f \rangle_{1,3J},$$

$$\|A[f\mathbf{1}_{3J}]\|_{L^{p}(J,1,\mathsf{size}_\mathsf{C})} \lesssim_p \langle f \rangle_{p,3J}, \qquad 1 < p \leq \infty.$$

Proposition 2.2.15 is obtained as a consequence of the localized estimate (2.3.1) of Proposition 2.3.2. We send to Section 2.3 for statements and proofs. Propositions 2.2.14 and 2.2.15 may be combined to prove the estimate

$$\mathsf{C}_\mathbb{P}(f_1, f_2) \lesssim_p \|f_1\|_p \|f_2\|_{p'}, \qquad 2 < p < \infty \tag{2.2.21}$$

uniformly over all $f_1, f_2 \in L_0^\infty(\mathbb{R})$ and finite $\mathbb{P} \subset \mathbb{S}$. In turn, via (2.1.6), (2.2.21) entails the $L^p(\mathbb{R})$-boundedness of (1.0.3) in the same range.

Proof of (2.2.21). Fix $f_1, f_2 \in L_0^\infty(\mathbb{R})$ and a finite $\mathbb{P}$. Using grid property (ii), find $J \in \mathcal{D}$ such that, denoting $J_j = J + j|J|$, and setting $\mathbb{P}_j := \mathbb{P} \cap \mathbb{S}^{J_j}$, there holds

$$\mathbb{P} = \mathbb{P}_{-1} \cup \mathbb{P}_0 \cup \mathbb{P}_1, \qquad \operatorname{supp} f_1, \operatorname{supp} f_2 \subset 3J_j \quad \forall j = 0, \pm 1.$$

The easy consideration $\mathsf{size}_\infty(F,T) \le \mathsf{size}_{2,\star,1}(F,T)$ and the definitions tell us that

$$\begin{aligned}\mathsf{size}_1(F_1F_2,T) &\le \mathsf{size}_2(F_1,T^{\mathsf{lac}})\mathsf{size}_2(F_2,T^{\mathsf{lac}}) + \mathsf{size}_\infty(F_1,T^{\mathsf{ov}})\mathsf{size}_1(F_2,T^{\mathsf{ov}}) \\ &\le 2\mathsf{size}_{2,\star,1}(F_1,T)\mathsf{size}_{\mathsf{C}}(F_2,T)\end{aligned} \tag{2.2.22}$$

so that a form of (2.2.6) is verified. Applying the outer Hölder inequality to $F_1 = W[f_1], F_2 = A[f_2]$ in the form of (2.2.7) followed by Propositions 2.2.14 and 2.2.15 thus leads to

$$\begin{aligned}&\mathsf{C}_{\mathbb{P}_i}(f_1,f_2) \le \|F_1F_2\|_{\ell^1(\mathbb{S}^{J_i})} \lesssim |J_i| \, \|W[f_1]\|_{L^p(J_i,1,\mathsf{size}_{2,\star})} \, \|A[f_2]\|_{L^{p'}(J_i,1,\mathsf{size}_{\mathsf{C}})} \\ &\lesssim |J_i| \langle f_1 \rangle_{p,3J_i} \langle f_2 \rangle_{p',3J_i} \lesssim \|f_1\|_p \, \|f_2\|_{p'}\end{aligned}$$

and the proof is completed by the observations that $\mathsf{C}_{\mathbb{P}} = \mathsf{C}_{\mathbb{P}_{-1}} + \mathsf{C}_{\mathbb{P}_0} + \mathsf{C}_{\mathbb{P}_1}$. □

2.3 Localized embeddings for the modified wave packet transforms

This section contains the statement and proof of the embedding theorems for the modified wave packet transform (2.1.3), see Proposition 2.3.2. The analysis behind this proposition is essentially based on a combination of the *tree* and *mass* lemmata from [48]. We claim no particular originality, but choose to present a full argument given the additional complications brought by the explicit dependence on $N(\cdot)$ of the wavelets in the map (2.1.3), cf. also the definition of the wavelet classes $\Psi(P)$ from (2.1.4). To handle this dependence, we borrow a continuity estimate idea from the paper [50] on Stein's conjecture for the Hilbert transform along vector fields.

Remark 2.3.1. Before we begin, we make the standing assumption that the function f playing the role of the argument in (2.1.3) belongs to $L_0^\infty(\mathbb{R})$ and that $\mathbb{P}$ is a finite subset of the collection of all tiles $\mathbb{S} = \mathbb{S}_{\mathcal{D},\mathcal{D}'}$. The finiteness assumption in the estimates does not change the scope of our applications, and may in fact be removed via a limiting argument when additional regularity assumptions on f are posed; for instance $f \in \mathcal{C}_0^2(\mathbb{R})$ will suffice.

Proposition 2.3.2. *We have*

$$\|A[f]\mathbf{1}_{\mathbb{P}}\|_{L^{p,\infty}(J,1,\mathsf{size}_{\mathsf{C}})} \lesssim [f]_{1,\mathbb{P}}, \qquad 1 \le p \le \infty \tag{2.3.1}$$

with uniform constant. In particular the above estimate yields the control

$$\|A[f]\mathbf{1}_{\mathbb{P}}\|_{Y^{p,\infty}(J,1,\mathsf{size}_{\mathsf{C}})} \lesssim [f]_{1,\mathbb{P}}, \qquad 1 \le p \le \infty.$$

In Proposition 2.3.2, as anticipated in Section 2.2, the tree parameter κ equals 1 and all trees referred to below are 1-trees, without further explicit mention. The proposition is proved by combining the next two lemmata, involving the auxiliary quantity

$$\mathsf{dense}(f, \mathbb{P}) := \sup_{P \in \mathbb{P}} \sup_{\substack{P \lesssim P' \\ P' \in \mathbb{S}}} \langle\!\langle f \mathbf{1}_{N^{-1}(\omega_{P'}^{\mathsf{p}(1)})} \rangle\!\rangle_{1, I_{P'}} \tag{2.3.2}$$

defined e.g. for $f \in L_0^\infty(\mathbb{R})$ and $\mathbb{P} \subset \mathbb{S}$. The order relation in (2.3.2) is a modification of the Fefferman ordering defined by

$$P \lesssim_\kappa P' \iff I_P \subset I_{P'},\ \omega_{P'}^{\mathsf{p}(\kappa)} \subset \omega_P^{\mathsf{p}(\kappa)}. \tag{2.3.3}$$

As we use (2.3.3) with $\kappa = 1$ throughout this section, we write $\lesssim$ instead of $\lesssim_1$.

Remark 2.3.3. A moment's thought yields $\mathsf{dense}(f, \mathbb{P}) \lesssim [f]_{1,\mathbb{P}}$ uniformly over $\mathbb{P} \subset \mathbb{S}$.

Lemma 2.3.4. $\|A[f]\mathbf{1}_{\mathbb{P}}\|_{L^\infty(J,1,\mathsf{size}_{\mathsf{C}})} \lesssim \mathsf{dense}(f, \mathbb{P})$.

Lemma 2.3.5. *Let $\mathbb{P} \subseteq \mathbb{S}$ and $\delta > 0$. There exists a decomposition $\mathbb{P} = \mathbb{P}_- \cup \bigcup_{T\in\mathcal{F}} T$, where $\mathsf{dense}(f, \mathbb{P}_-) \leq \delta$, each T is a tree with top interval I_T, and the forest $\mathcal{F} = \mathcal{F}(\delta, f)$ satisfies*

$$\frac{\delta}{|J|} \sum_{\substack{T\in\mathcal{F} \\ I_T\subset J}} |I_T| \lesssim \inf_J \mathrm{M}_1 f, \qquad \forall J \in \mathcal{D}. \tag{2.3.4}$$

The proofs of Lemmata 2.3.4 and 2.3.5 are respectively postponed to Subsections 2.3.6 and 2.3.7. We now show how a combination of these yields Proposition 2.3.2. Fix $J \in \mathcal{D}$, $\mathbb{P} \subset \mathbb{S}$. The bound (2.3.1) is an immediate consequence of

$$\sup_{t>0} \max\{1, t\} \left(A[f]\mathbf{1}_{\mathbb{P}}\right)^{*,\mathsf{size}_{\mathsf{C}}}(t) \leq C[f]_{1,\mathbb{P}} \tag{2.3.5}$$

where C is an absolute constant explicitly computed below and $\mathbb{P} \subset \mathbb{S}^J$. The range $t \leq 1$ of estimate (2.3.5) is readily obtained by combining Remark 2.3.3 with the conclusion of Lemma 2.3.4 and choosing C to be larger than the product of the respective absolute implicit constants. Now, notice that the right hand side of (2.3.4) is also controlled by $[f]_{1,\mathbb{P}}$. Applying Lemma 2.3.5 to $\mathbb{P}$ with the choice $\delta = C\frac{[f]_{1,\mathbb{P}}}{t}$, provided C is larger than twice the implicit constant in (2.3.4) yields (2.3.5) in the range $t \geq 1$.

2.3.6 Proof of Lemma 2.3.4

The proof of the Lemma consists in showing that

$$\mathsf{size}_{\mathsf{C}}(A[f], T) = \frac{1}{|I_T|} \sum_{P\in T^{\mathsf{ov}}} |I_P| A[f](P) + \left(\frac{1}{|I_T|} \sum_{P\in T^{\mathsf{lac}}} |I_P| A[f](P)^2\right)^{\frac{1}{2}} \lesssim \mathsf{dense}(f, \mathbb{P}) \tag{2.3.6}$$

whenever $T \in \mathcal{T}^{J,\star}$ is a tree with $T \subset \mathbb{P}$. For any such tree, we introduce the support intervals

$$\Omega^{\mathsf{b}}(T) := \{\omega_P^{\mathsf{b}} : P \in T\}.$$

We may assume, by splitting, that T is a type 2 tree, which means that ω_P is the right child of its dyadic parent $\omega_P^{\mathsf{p}(1)}$ for all $P \in T$. Lemma 2.2.3 thus tells us that the collection $\Omega^{\mathsf{b}}(T^{\mathsf{ov}})$ consists of pairwise disjoint intervals, while T^{lac} is a lacunary tree, so that in particular $\Omega^{\mathsf{b}}(T^{\mathsf{lac}})$ is a nested collection of intervals containing ξ_T. This follows immediately by combining

$$\xi_T \in \omega_P^{\mathsf{p}(1)}\ \forall P \in T, \qquad \omega_P \neq \omega_{P'} \implies \omega_P \cap \omega_{P'} = \varnothing \quad \forall P, P' \in T^{\mathsf{lac}}.$$

Accordingly, the quantity $\delta(x) \coloneqq \inf \left\{ \ell_{\omega_P} : N(x) \in \omega_P^{\mathsf{b}}, P \in T^{\mathsf{lac}} \right\}$ records the minimal active frequency scale of T^{lac} at each $N(x) \in \mathbb{R}$ and satisfies

$$|N(x) - \xi_T| \leq \delta(x), \qquad x \in \mathbb{R}. \tag{2.3.7}$$

In estimating both contributions, a key role is played by the collection $\mathcal{G}$ of maximal elements in $\{G \in \mathcal{D} : 3G \not\supseteq I_P\ \forall P \in T\}$. Accordingly, for $G \in \mathcal{G}$, $\mathsf{j} \in \{\mathsf{ov}, \mathsf{lac}\}$ decompose

$$T^{\mathsf{j}} = T_G^{\mathsf{j},+} \cup T_G^{\mathsf{j},-}, \qquad T_G^{\mathsf{j},+} = \{P \in T^{\mathsf{j}} : \mathsf{scl}(P) > \ell_G\}, \qquad T_G^{\mathsf{j},-} = \{P \in T^{\mathsf{j}} : \mathsf{scl}(P) \leq \ell_G\}.$$

We begin to estimate the T^{ov} term in (2.3.6). Using the definition and the fact that $\mathcal{G}$ is a partition of $\mathbb{R}$ leads to

$$\sum_{P \in T^{\mathsf{ov}}} |I_P| A[f](P) \leq 2 \sum_{* \in \{+,-\}} \sum_{G \in \mathcal{G}} \sum_{P \in T_G^{\mathsf{ov},*}} |I_P| \langle f, \varphi_P \mathbf{1}_{G \cap N^{-1}(\omega_P^{\mathsf{b}})} \rangle, \qquad \varphi_P \coloneqq \psi_P(\cdot, N(\cdot)) \tag{2.3.8}$$

for suitable $\psi_P \in \Psi(P)$. Note that φ_P are not standard wavelets as they carry the dependence on the measurable function N from the second argument of ψ_P. The basic estimate

$$|\langle f, \varphi_P \mathbf{1}_{G \cap N^{-1}(\omega_P^{\mathsf{b}})} \rangle| \lesssim \chi_{I_P}^{10}(c_G) \mathsf{dense}(f, \mathbb{P}) \tag{2.3.9}$$

reveals that the $*=-$ sum in (2.3.8) is a tail term. Indeed, also relying on the defining property of $\mathcal{G}$ for the first estimate, and later on (2.2.1),

$$\begin{aligned}&\sum_{G\in\mathcal{G}}\sum_{P\in T^{\mathsf{ov},-}}|I_P|\langle f,\varphi_P\mathbf{1}_{G\cap N^{-1}(\omega_P^{\mathsf{b}})}\rangle \lesssim \mathsf{dense}(f,\mathbb{P})\sum_{G\in\mathcal{G}}\sum_{k\geq 0}\sum_{\substack{P\in T^{\mathsf{ov}}\\ |I_P|=2^{-k}|G|, I_P\cap G=\varnothing}}|I_P|\chi_{I_P}^{10}(c_G)\\ &\lesssim \mathsf{dense}(f,\mathbb{P})\sum_{G\in\mathcal{G}}|G|\chi_{I_T}^{9}(c_G)\lesssim \mathsf{dense}(f,\mathbb{P})\int\chi_{I_T}^{9}\lesssim \mathsf{dense}(f,\mathbb{P})|I_T|\end{aligned} \tag{2.3.10}$$

which is compliant with (2.3.6). The $*=+$ term is estimated as follows. First, note that $T^{\mathsf{ov},+}(G)=\varnothing$ unless $G\subset 9I_T$ and there exists $P(G)\in T$ with

$$\mathsf{scl}(P(G))=2\ell_G,\qquad \operatorname{dist}(G,I_{P(G)})\leq \mathsf{scl}(P(G)).$$

Let $P'(G)\in\mathbb{S}_{\mathcal{D},\mathcal{D}'}$ be the unique tile with $I_{P'(G)}=I_{P(G)}$ and $\xi_T\in\omega_{P'(G)}$. As the intervals $\Omega^{\mathsf{b}}(T^{\mathsf{ov},+}(G))$ are pairwise disjoint and contained in $\omega_{P'(G)}^{\mathsf{p}(1)}$,

$$\begin{aligned}&\sum_{G\in\mathcal{G}}\sum_{P\in T_G^{\mathsf{ov},+}}|I_P|\langle f,\varphi_P\mathbf{1}_{G\cap N^{-1}(\omega_P^{\mathsf{b}})}\rangle \lesssim \sum_{\substack{G\in\mathcal{G}\\ G\subset 9I_T}}|G|\langle\!\langle f\mathbf{1}_{N^{-1}(\omega_{P'(G)}^{\mathsf{p}(1)})}\rangle\!\rangle_{1,I_{P'(G)}}\\ &\lesssim \sum_{\substack{G\in\mathcal{G}\\ G\subset 9I_T}}|G|\mathsf{dense}(f,\{P(G)\})\lesssim \mathsf{dense}(f,\mathbb{P})|I_T|,\end{aligned}$$

which also complies with (2.3.6). The ov term in (2.3.8) is thus fully handled.

We move onto the lac term in (2.3.8). With the same notation of the T^{ov} the term, we estimate

$$\sum_{P\in T^{\mathsf{lac}}}A[f](P)^2|I_P|\leq 2\sum_{*\in\{+,-\}}\sum_{G\in\mathcal{G}}\sum_{P\in T_G^{\mathsf{ov},*}}A[f](P)\langle f,\varphi_P\mathbf{1}_{G\cap N^{-1}(\omega_P^{\mathsf{b}})}\rangle|I_P|. \tag{2.3.11}$$

The $*=-$ sum in (2.3.11) is handled along the lines of (2.3.10), with an additional application of (2.3.9): we omit the details. The rest of the analysis deals with the $*=+$

summand in (2.3.11). The explicit dependence of $\varphi_P = \psi_P(\cdot, N(\cdot))$ on $N(\cdot)$ prohibits us to use orthogonality methods directly. This is obviated by replacing φ_P with the standard wavelets

$$\phi_P \coloneqq \psi_P(\cdot, \xi_T) \in \Phi(P), \qquad P \in T^{\mathsf{lac}}.$$

Setting $\zeta_P \coloneqq |I_P|[\varphi_P - \phi_P] = |I_P|[\psi_P(\cdot, N(\cdot)) - \phi_P], P \in T^{\mathsf{lac}}$, th error term created by the replacement is

$$\begin{aligned}
&\sum_{G\in\mathcal{G}, G\subset 9I_T} \sum_{P\in T_G^{\mathsf{lac},+}} A[f](P) \left|\langle f, \zeta_P \mathbf{1}_{G\cap N^{-1}(\omega_P^{\mathsf{b}})}\rangle\right| \\
&\lesssim \mathsf{dense}(f,\mathbb{P}) \sum_{G\in\mathcal{G}, G\subset 9I_T} \left\langle |f|\mathbf{1}_{G\cap N^{-1}(\omega_{P'(G)}^{\mathsf{p}(1)})}, \sum_{P\in T_G^{\mathsf{lac},+}} |\zeta_P| \mathbf{1}_{\omega_P^{\mathsf{b}}}(N(\cdot)) \right\rangle \\
&\lesssim \sum_{G\in\mathcal{G}, G\subset 9I_T} |G| \mathsf{dense}(f,\mathbb{P})^2 \lesssim |I_T| \mathsf{dense}(f,\mathbb{P})^2.
\end{aligned}$$

For the passage to the second line, note that the intervals $\Omega^{\mathsf{b}}(T^{\mathsf{ov},+}(G))$ are all contained in $\omega_{P'(G)}^{\mathsf{p}(1)}$. The subsequent step was obtained via a Lipschitz estimate in the second argument of $\psi_P \in \Psi(P)$ and subsequently taking advantage of (2.3.7), so that

$$\sum_{P\in T^{\mathsf{lac}}} |\zeta_P(x)| \mathbf{1}_{\omega_P^{\mathsf{b}}}(N(x)) \lesssim \sum_{P\in T^{\mathsf{lac}}: \ell_{\omega_P} \geq \delta(x)} \chi_{I_P}(x) \frac{\delta(x)}{\ell_{\omega_P}} \lesssim 1.$$

We are left with estimating the $* = +$ summand in (2.3.11), where the φ_P have been replaced by the almost orthogonal wavelets ϕ_P. A principal role is played by the tree operator

$$H_T f \coloneqq \sum_{P\in T^{\mathsf{lac}}} |I_P| A[f](P) \phi_P.$$

As $\xi_T \in \omega^{\mathsf{b}}$ for all $\omega \in \Omega(T^{\mathsf{lac}})$, the intervals $\Omega(T^{\mathsf{lac}})$ form a lacunary sequence, that is

$$\omega \subset \left\{\xi \in \mathbb{R} : \tfrac{\ell_\omega}{2} \leq \mathrm{dist}(\xi, \xi_T) \leq 2\ell_\omega\right\} \quad \forall \omega \in \Omega(T^{\mathsf{lac}}),\ \ell_\omega > \ell_\star \coloneqq \min\left\{\ell_{\omega'} : \omega' \in \Omega(T^{\mathsf{lac}})\right\}.$$

For $\alpha \in \{\ell_\omega : \omega \in \omega(T^{\mathsf{lac}})\}$, let Ψ_α be even, real valued Schwartz functions with

$$1_{[\frac{1}{2}\alpha, 2\alpha]} \le \widehat{\Psi}_\alpha \le 1_{[\frac{2}{5}\alpha, \frac{11}{5}\alpha]} \quad \alpha > \ell_\star, \qquad 1_{[0,2\ell_\star]} \le \widehat{\Psi}_{\ell_\star} \le 1_{[0,\frac{11}{5}\ell_\star]}.$$

Assuming that $\{\ell_\omega : \omega \in T^{\mathsf{lac}}\}$ are separated by a factor of 4, and arguing by finite splitting otherwise, we obtain for all $\ell_\star \le \alpha \le \beta$,

$$\left| H_{T,\alpha,\beta} f := \sum_{\substack{P \in T^{\mathsf{lac}} \\ \alpha \le \ell_{\omega_P} < \beta}} |I_P| A[f](P) \phi_P = [\Psi_\beta - \Psi_\alpha] * H_T f \right| \lesssim \mathrm{M}[H_T f],$$

due to the frequency support conditions $\phi_P \subset \omega_P$. Relying on the definition of $\delta(\cdot)$, cf. (2.3.7), the modified $* = +$ summand in (2.3.11) is then estimated by

$$\sum_{\substack{G \in \mathcal{G} \\ G \subset 9I_T}} \int_G |f| \mathbf{1}_{N^{-1}(\omega^{\mathsf{p}(1)}_{P'(G)})} \left| H_{T,\delta(\cdot),\frac{1}{\ell_G}} f \right| \lesssim \mathsf{dense}(f,\mathbb{P})^2 \int_{9I_T} \mathrm{M}[H_T f]$$
$$\lesssim |I_T| \mathsf{dense}(f,\mathbb{P})^2 \|H_T f\|_2 \lesssim |I_T| \mathsf{dense}(f,\mathbb{P})^2 \left(\sum_{P \in T^{\mathsf{lac}}} |I_P| A[f](P)^2 \right)^{\frac{1}{2}}.$$

Balancing out the obtained bounds completes the estimation of lac term in (2.3.8), and in turn, the proof of Lemma 2.3.6.

2.3.7 Proof of Lemma 2.3.5

The selection of the trees $T \in \mathcal{F}$ and consequent estimation of $\mathsf{dense}(f, \mathbb{P}_-)$ is identical to [48, Proposition 3.1] and is thus omitted. To prove (2.3.4), it suffices to show that whenever $\mathbb{P}' \subset \mathbb{S}^J$ is a set of pairwise incomparable tiles with respect to (2.3.3)

$$\inf_{P \in \mathbb{P}'} \left\langle\!\!\left\langle f \mathbf{1}_{N^{-1}(\omega_P^{\mathsf{p}(1)})} \right\rangle\!\!\right\rangle > \delta \implies \sum_{P \in \mathbb{P}'} |I_P| \lesssim \delta^{-1} |J| \inf_J \mathrm{M} f. \tag{2.3.12}$$

Due to the premise of (2.3.12), for each $P \in \mathbb{P}'$ there exists $k = k_P \geq 0$ with the property that

$$\int_{2^k I_P \cap N^{-1}(\omega_P^{\mathsf{p}(1)})} |f| \geq 2^{6k}\delta |I_P|$$

and k_P is minimal with this property. Let $\mathbb{P}'_k$ be the collection of all $P \in \mathbb{P}'$ with $k_P = k$. Perform the following iterative selection. Initialize $\mathbb{A} := \mathbb{P}'_k, \mathbb{B} = \varnothing$. Among those $P^\star \in \mathbb{A}$ with

$$2^k I_{P^\star} \times \omega_{P^\star}^{\mathsf{p}(1)} \cap 2^k I_P \times \omega_P^{\mathsf{p}(1)} = \varnothing \qquad \forall P \in \mathbb{B}$$

select one with $\mathsf{scl}(P^\star)$ maximal, and set $\mathbb{A} := \mathbb{A} \setminus \{P^\star\}$, $\mathbb{B} := \mathbb{B} \cup \{P^\star\}$. Repeat until no such $P^\star \in \mathbb{A}$ is available. At this point, we may partition $\mathbb{P}'_k = \bigcup \{\mathbb{P}'_k(P^\star) : P^\star \in \mathbb{B}\}$ where $P \in \mathbb{P}'_k(P^\star)$ if

$$2^k I_P \times \omega_P^{\mathsf{p}(1)} \cap 2^k I_{P^\star} \times \omega_{P^\star}^{\mathsf{p}(1)} \neq \varnothing, \qquad \mathsf{scl}(P) \leq \mathsf{scl}(P^\star).$$

Notice that if $P, P' \in \mathbb{P}'_k(P^\star)$ then $\omega_P^{\mathsf{p}(1)} \cap \omega_{P'}^{\mathsf{p}(1)} \supset \omega_{P^\star}^{\mathsf{p}(1)}$, and P, P' are incomparable, so that the intervals $\{I_P : P \in \mathbb{P}'_k(P^\star)\}$ are pairwise disjoint and contained in $2^{k+2} I_{P^\star}$. We then have

$$\begin{aligned}\sum_{P \in \mathbb{P}'_k} |I_P| &\leq \sum_{P^\star \in \mathbb{B}} \sum_{P \in \mathbb{P}'_k(P^\star)} |I_P| \lesssim 2^k \sum_{P^\star \in \mathbb{B}} |I_{P^\star}| \lesssim 2^{-5k}\delta^{-1} \sum_{P^\star \in \mathbb{B}} \int_{2^k I_{P^\star} \cap N^{-1}(\omega_{P^\star}^{\mathsf{p}(1)})} |f| \\ &\lesssim 2^{-4k}\delta^{-1}|J| \langle f \rangle_{1,2^{k+2}J} \lesssim 2^{-4k}\delta^{-1}|J| \inf_J \mathrm{M}f.\end{aligned}$$

To pass to the second line, note that the sets $2^k I_{P^\star} \cap N^{-1}(\omega_{P^\star}^{\mathsf{p}(1)})$, $P^\star \in \mathbb{B}$ are pairwise disjoint and contained in $2^{k+3}J$. Then (2.3.12) follows by summing over k.

2.4 Localized wave packet estimates near L^1

If the local L^2-averages of f are under control, we may combine the bound of Proposition 2.2.11 with (2.2.19) in the single localized estimate

$$\|W[f]\mathbf{1}_{\mathbb{P}}\|_{Y^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} \leq C_\kappa [f]_{2,\mathbb{P}}. \tag{2.4.1}$$

The quantities $[f]_{p,\mathbb{P}}$ have been introduced in (2.2.14). This section contains the statement and main line of proof of a localized estimate for the wave packet transform in terms of local L^p norms in the range $1 < p \leq 2$, with good control on the estimate as $p \to 1^+$. Throughout the remainder of this section, we enforce the formal assumptions of Remark 2.3.1 without further explicit mention.

The main result of [23], a first substitute for (2.4.1) outside local L^2, is recalled in the next proposition.

Proposition 2.4.1. *Let $1 < p \leq 2$. For all $t > 1$ there exists $C_{t,p,\kappa} > 1$ such that the following holds. Let J be any interval, $f \in L_0^\infty(\mathbb{R})$ and $\mathbb{P} \subset \mathbb{S}$. Then*

$$\|W[f\mathbf{1}_{3J}]\mathbf{1}_{\mathbb{P}}\|_{L^{tp'}(J,\kappa,\mathsf{size}_{2,\star})} \leq C_{t,p,\kappa}[f]_{p,\mathbb{P}}.$$

Proposition 2.4.1 has been used to prove sparse and localized estimates for the Carleson operator [20] and the bilinear Hilbert transform [16]. However, an inspection of the proof shows that having fixed $t > 1$, the constant $C_{t,p}$ blows up polynomially in $(p-1)^{-1}$ as $p \to 1^+$.

The next theorem, which is the main technical novelty of this work, provides us with a substitute embedding that does not blow up near $p = 1$. Remark 2.2.7 tells us that the norms $X_2^{tp',\infty}$ are weaker than the ones appearing on the left hand side of Proposition 2.4.1. Nonetheless, the generalized Hölder inequality of Proposition 2.2.9 makes Theorem H applicable for our purposes.

Theorem H. *For all $t > 1$ there exists $C_{t,\kappa} > 1$ such that the following holds. Let $1 < p \leq 2$, J be any interval, $f \in L_0^\infty(\mathbb{R})$ and $\mathbb{P} \subset \mathbb{S}$. Then*

$$\|W[f]\mathbf{1}_{\mathbb{P}}\|_{X_2^{tp',\infty}(J,\kappa,\mathsf{size}_{2,\star})} \leq C_{t,\kappa}[f]_{p,\mathbb{P}}.$$

Remark 2.4.2. We clarify a delicate point in the statement of Theorem H. Fixing t, the smoothness level of the wave packet transform, as defined in (2.1.2), required for Theorem H must be greater or equal to, say, $M = 10 \cdot \lceil 2^8 t' \rceil$. Theorem H will be applied below with the fixed choice $t = 2$, so that a fixed level of smoothness, say $M = 10 \cdot 2^9$, is sufficient.

The proof of Theorem H occupies the remainder of this section and is structured as follows. Subsection 2.4.3 introduces a generalization of the wavelet classes $\Phi^M(P)$ of (2.1.1) where the compact frequency support assumption is relaxed to requiring instead vanishing moments with respect to a fixed frequency.

2.4.3 Relaxed wavelet classes

For an interval $I \subset \mathbb{R}$ and $\xi \in \mathbb{R}$, define the normalized classes

$$\Theta_1^M(I,\xi) := \left\{\mathrm{Mod}_\xi \mathrm{Tr}_{c(I)} \mathrm{Dil}^1_{\mathsf{scl}(I)}\vartheta : \vartheta \in \Phi^M\right\}$$

$$\Theta_0^M(I,\xi) := \left\{\mathrm{Mod}_\xi \mathrm{Tr}_{c(I)} \mathrm{Dil}^1_{\mathsf{scl}(I)}\vartheta : \upsilon \in \Phi^M, \widehat{\upsilon}(0) = 0\right\}.$$

As usual we drop the M when irrelevant or clear from context. If $P \in \mathbb{S}$ is a tile and $\phi_P \in \Phi(P)$ we have the inclusions

$$\xi \in \omega_P^{\mathsf{p}(\kappa)} \implies \phi_P \in C_\kappa \Theta_1(I_P, \xi), \qquad \xi \in \omega_P^{\mathsf{p}(\kappa)} \setminus \omega_P \implies \phi_P \in C_\kappa \Theta_0(I_P, \xi).$$

The next lemma is a restatement of [23, Lemma 5.2].

Lemma 2.4.4. *Suppose* $\imath \in \{0, 1\}$, $\phi \in \Theta_\imath^{3M}(I, \xi)$, *and* $K \geq 1$. *Then*

$$\phi = \psi + K^{-M} \upsilon, \qquad \psi, \upsilon \in C_M \Theta_\imath^M(I, \xi), \qquad \operatorname{supp} \psi \subset KI.$$

Remark 2.4.5. Let P be a tile, and suppose either $\imath = 1$, $\xi \in \kappa\omega_P$ or $\imath = 0$, $\xi \in \kappa\omega_P \setminus \omega_P$. Then Lemma 2.4.4 may be iterated to deduce the expansion of $\varphi_P \in \Phi^{3M}(P)$

$$\varphi_P = \sum_{k \geq 0} 2^{-Mk} \varphi_{P,k,\xi}, \qquad \varphi_{P,k,\xi} \in C_M \Theta_\imath^M(I_P, \xi), \qquad \operatorname{supp} \varphi_{P,k,\xi} \subset 2^k I_P. \tag{2.4.2}$$

The expansion (2.4.2) is the form of Lemma 2.4.4 we will use in the sequel.

2.4.6 Space-frequency decomposition on minimal tiles

Our aim in this paragraph is to provide a space-frequency decomposition induced by a finite collection of spatial intervals $\mathcal{I} \subset \mathcal{D}$, where $\mathcal{D}$ is a fixed dyadic grid. The definition also involves a dilation factor $K \geq 1$. The spatial components of the forthcoming decomposition will come from the collection

$$\mathrm{CZ}_K(\mathcal{J}) := \text{maximal elements of } \left\{ G \in \mathcal{D} : 9K^2 G \not\supset J \text{ for all } J \in \mathcal{J} \right\}. \tag{2.4.3}$$

When clear from context, the subscript K is dropped from the notation. The following properties, which will be of use to us below, are deduced from (2.4.3).

(i) $\mathrm{CZ}_K(\mathcal{J})$ partitions $\mathbb{R}$ up to a set of zero measure.

(ii) The collection $\{3G : G \in \mathrm{CZ}_K(\mathcal{J})\}$ has finite overlap.

(iii) If $J \in \mathcal{J}, G \in \mathrm{CZ}_K(\mathcal{J})$ and $G \not\subset 9KJ$ then $G \subset \mathbb{R} \setminus 3KJ$.

(iv) If $J \in \mathcal{J}, G \in \mathrm{CZ}_K(\mathcal{J})$ and $G \subset 3KJ$ then $K\ell_G \leq \ell_J$.

(v) whenever $h \in L^\infty_0(\mathbb{R})$ say, there holds

$$\sup_{J\in\mathcal{J}} \inf_J \mathrm{M}h \lesssim \sup_{G\in\mathrm{CZ}_K(\mathcal{J})} \inf_G \mathrm{M}h \lesssim K^2 \sup_{J\in\mathcal{J}} \inf_J \mathrm{M}h.$$

We briefly comment on the proof of (ii). Indeed up to finite splitting it suffices to check (ii) for the collections $\mathrm{CZ}^r_K(\mathcal{J}) \coloneqq \left\{G : G \in \mathrm{CZ}_K(\mathcal{J}), \ell_G \in 2^{8\mathbb{Z}+r}\right\},\ r = 0, \ldots, 7$. Indeed let $x \in 3G_0$ for some $G_0 \in \mathrm{CZ}^r_K$ and suppose that for some $G_1 \in \mathrm{CZ}^r_K(\mathcal{J})$ with $\ell_{G_1} \neq \ell_{G_0}$ we have that $x \in 3G_1$ then $9K^2G_1 \supset 27K^2G_0 \supset 9K^2G_0^{\mathfrak{p}(1)}$ or $9K^2G_0 \supset 27K^2G_1 \supset 9K^2G_1^{\mathfrak{p}(1)}$ which is absurd by the definition of $\mathrm{CZ}_K(\mathcal{J})$ therefore

$$\sum_{G\in\mathrm{CZ}^r_K(\mathcal{J})} \mathbf{1}_{3G}(x) = \sum_{G\in\mathrm{CZ}^r_K(\mathcal{J}),\, \ell_G=\ell_{G_0}} \mathbf{1}_{3G}(x) \leq \#\left\{G : G \in \mathrm{CZ}^r_K(\mathcal{J}),\ \ell_G = \ell_{G_0}, G \subset 9G_0\right\} \lesssim 1.$$

The corresponding collection of *minimal space-frequency tiles* is then defined by

$$\mathbb{M} = \mathbb{M}(\mathcal{J}) \coloneqq \left\{G \times \left[\xi, \xi + \tfrac{1}{\ell_G}\right) : G \in \mathrm{CZ}(\mathcal{J}), \xi \in \tfrac{\mathbb{Z}}{\ell_G}\right\} \subset \mathbb{S}_{\mathcal{D}\times\mathcal{D}_0}. \tag{2.4.4}$$

It is clear that $\mathbb{M}$ depends on $\mathcal{J}$, but we choose to keep the latter implicit in the notation when clear from context. Pick $\eta \in \mathcal{S}(\mathbb{R})$ with $\operatorname{supp}\eta \subset (-1, 1)$ and $\eta(0) = \frac{1}{2\pi}$. For $P \in \mathbb{M}$ define the approximate projection operator Π_P, acting on $f \in L^2(\mathbb{R})$ by

$$\Pi_P f \coloneqq [f\mathbf{1}_{I_P}] * \eta_P, \qquad \eta_P \coloneqq \mathrm{Mod}_{\inf\omega_P}\mathrm{Dil}^1_{\mathsf{scl}(P)}\eta.$$

The Poisson summation formula tells us that

$$f = \sum_{P\in\mathbb{M}} \Pi_P f \tag{2.4.5}$$

with convergence in $L^2(\mathbb{R})$ and almost everywhere. The decomposition (2.4.5) is approximately space-frequency localized in the sense that

$$\operatorname{supp}\Pi_P f \subset 3I_P, \qquad \Pi_P f \in C[f]_{1,P}\Theta_1(I_P, c_{\omega_P})$$

for some absolute constant C. The approximate projection onto the space-frequency region associated to some $\mathbb{W}\subset\mathbb{M}$ is then defined, for say $f\in L^2(\mathbb{R})$, by

$$\Pi_{\mathbb{W}} f \coloneqq \sum_{P\in\mathbb{W}} \Pi_P f.$$

Below, whenever $\mathbb{W}\subset\mathbb{M}$, by

$$\mathbb{W}[G] = \{P\in\mathbb{W} : I_P = G\} \tag{2.4.6}$$

we indicate the tiles of $\mathbb{W}$ having a fixed spatial interval $G\in \mathrm{CZ}_K(\mathcal{J})$.

2.4.7 Main line of proof of Theorem H

This paragraph reduces Theorem H to a Calderón-Zygmund type decomposition of f with respect to an arbitrary family of top data. Details are as follows. To prove the estimate of Theorem H, having fixed

$$\mathbb{P}\subset\mathbb{S}^J \text{ finite}, \qquad \varnothing \subsetneq A\subset \mathbb{S}^J,\ \mu^{J,\kappa}(A) =: N < \infty,$$

we need to prove the control

$$\|W[f]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} \lesssim N^{\frac12-\frac{1}{tp'}}[f]_{p,\mathbb{P}} \tag{2.4.7}$$

for $f \in L_0^\infty(\mathbb{R})$. If $N \leq 1$ then (2.4.7) follows immediately from an application of Proposition 2.2.11 and (2.2.8). We deal with the difficult case $N > 1$. To do so, we select an almost optimal collection of trees $\mathcal{T} \subset \mathcal{T}^{J,\star}$ covering A, that is

$$A \subset \bigcup_{T \in \mathcal{T}} T, \qquad \sum_{T \in \mathcal{T}} |I_T| \leq 2N|J|,$$

and denote by $\mathcal{F} = \{(I_T, \xi_T) : T \in \mathcal{T}\}$ the corresponding collection of top data. Relying on the collection $\mathcal{F}$, for a given $f \in L_0^\infty(\mathbb{R})$, we produce the decomposition

$$f = g + b, \tag{2.4.8}$$

$$\|g\|_2 \leq C_t |J|^{\frac{1}{2}} N^{\frac{1}{2} - \frac{1}{tp'}} [f]_{p,\mathbb{P}} \tag{2.4.9}$$

$$\|W[b]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} \leq C_t N^{-\frac{1}{p'}} [f]_{p,\mathbb{P}} \tag{2.4.10}$$

where the constant C_t depends only on the fixed parameter $t > 1$ and is allowed to vary at each occurrence. With (2.4.9)-(2.4.10) in hand, we use quasi-subadditivity of the $L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})$-quasinorm to obtain

$$\begin{aligned}
\|W[f]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} &\leq 2\,\|W[g]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} + 2\,\|W[b]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} \\
&\leq 2\,\|W[g]\|_{L^{2,\infty}(J,\kappa,\mathsf{size}_{2,\star})} + 4\,\|W[b]\mathbf{1}_{\mathbb{P}\cap A}\|_{L^\infty(J,\kappa,\mathsf{size}_{2,\star})} N^{\frac{1}{2}} \\
&\leq C|J|^{-\frac{1}{2}}\,\|g\|_2 + C_t N^{\frac{1}{2}-\frac{1}{p'}} [f]_{p,\mathbb{P}} \leq C_t N^{\frac{1}{2}-\frac{1}{tp'}} [f]_{p,\mathbb{P}}.
\end{aligned}$$

To pass to the second line, we have employed monotonicity on both terms and (2.2.8). The subsequent bound follows from an application of Proposition 2.2.14, in particular (2.2.19) and by taking advantage of (2.4.10), while the final estimate is a consequence of (2.4.9). This completes the proof of (2.4.7), and in turn of Theorem H, up to actual construction of the splitting $f = g + b$ with properties (2.4.9)-(2.4.10). This task is conducted in the upcoming paragraphs. The first step towards (2.4.9)-(2.4.10) is to construct a suitable collection of minimal space-frequency tiles adapted to the collection $\mathbb{P}$. To do so, take

$$\mathcal{J} = \{J \in \mathcal{D} : J = I_P \text{ for some } P \in \mathbb{P}\}$$

in (2.4.3). The choice of the constant $K \geq 1$ depends on N, p and t and will be made explicit in (2.4.22) below. From now on, $\mathbb{M} = \mathbb{M}(\mathcal{J})$ refers to the collection obtained from (2.4.4) for this choice of $\mathcal{J}, K$. Note that the spatial components of the tiles in $\mathbb{M}$ come from the collection $\mathrm{CZ}_K(\mathcal{J})$. This fact will be employed in the proof quite a few times.

Below, the notation T is used, with meaning clear from context, for both the top data pair itself $T = (I_T, \xi_T) \in \mathcal{F}$ and to the set $T = T(I_T, \xi_T) = \{P \in \mathbb{P} : I_P \subset I_T,\ \xi_T \in \omega_P^{\mathsf{p}(\kappa)}\}$. The collection of top data $\mathcal{F}$ induces a certain decomposition of the minimal tiles $\mathbb{M}$, as follows. First, the *principal region* $\mathbb{Q}$ is defined by

$$\mathbb{Q} = \bigcup_{T \in \mathcal{F}} \{Q \in \mathbb{M} : \mathsf{scl}(Q) |\inf \omega_Q - \xi_T| \leq K, I_Q \subset 3K I_T\}. \tag{2.4.11}$$

Each $T = (I_T, \xi_T) \in \mathcal{F}$ then partitions the *tail region* $\mathbb{M} \setminus \mathbb{Q}$ into the two components

$$\begin{aligned} \mathbb{Q}'(T) &\coloneqq \{Q \in \mathbb{M} \setminus \mathbb{Q} : \mathsf{scl}(Q) |\inf \omega_Q - \xi_T| > K\}, \\ \mathbb{Q}''(T) &\coloneqq \{Q \in \mathbb{M} \setminus \mathbb{Q} : \mathsf{scl}(Q) |\inf \omega_Q - \xi_T| \leq K\} \end{aligned} \tag{2.4.12}$$

roughly corresponding to the frequency tails and spatial tails with respect to T. The definitions guarantee that $\mathbb{M} = \mathbb{Q} \sqcup \mathbb{Q}'(T) \sqcup \mathbb{Q}''(T)$ for each $T \in \mathcal{F}$.

2.4.8 Space-frequency tail estimates

The following technical lemma, via a suitable decomposition, shows how the action of the (adjoint) frequency tails projection $\Pi_{\mathbb{Q}'(T)}$ on wave packets localized to T is exponentially small in the separation parameter K.

Lemma 2.4.9. *Let M be a large integer. There exists a positive constant $C = C(M)$ and a decomposition*

$$\Pi^*_{\mathbb{Q}'(T)} = \Pi^{*,\mathsf{avg}}_{\mathbb{Q}'(T)} + \Pi^{*,\mathsf{osc}}_{\mathbb{Q}'(T)}$$

with the following properties.

(i) *For each pair $f, g \in L^2(\mathbb{R})$, there exists $h \in L^2(\mathbb{R})$ such that $|h| \leq C|f|$ and*

$$\langle f, \Pi^{*,\mathsf{avg}}_{\mathbb{Q}'(T)} g\rangle = K^{-M}\langle h, g\rangle. \tag{2.4.13}$$

(ii) *If $I \in \mathcal{D}$, the pointwise inequality*

$$\sum_{\substack{P\in T\\ I_P\subset I}} \left|\Pi^{*,\mathsf{osc}}_{\mathbb{Q}'(T)}\phi_P\right| \leq CK^{-\frac{M}{10}}\chi_I^{\frac{M}{10}} \tag{2.4.14}$$

holds for each L^∞-normalized collection $\{\phi_P : P \in T, |I_P|^{-1}\phi_P \in \Theta_1^M(I_P, \xi_T)\}$.

Proof. With the notation of (2.4.6),

$$\Pi^*_{\mathbb{Q}'(T)} = \sum_{G\in \mathrm{CZ}_K(\mathcal{J})} \left[f * \left(\overline{Z_G}(-\cdot)\right)\right]\mathbf{1}_G, \qquad Z_G \coloneqq \sum_{Q\in\mathbb{Q}'(T)[G]} \eta_Q.$$

The claimed decomposition is

$$\Pi^{*,\mathsf{avg}}_{\mathbb{Q}'(T)} f := \sum_{G\in \mathrm{CZ}_K(\mathcal{J})} \overline{\widehat{Z_G}(\xi_T)}\,(\mathbf{1}_G f)\,, \qquad \Pi^{*,\mathsf{osc}}_{\mathbb{Q}'(T)} := \Pi^{*,\mathsf{poly}_n}_{\mathbb{Q}'(T)} + \Pi^{*,\mathsf{cancel}_n}_{\mathbb{Q}'(T)}$$

$$\Pi^{*,\mathsf{poly}_n}_{\mathbb{Q}'(T)} f := \sum_{G\in \mathrm{CZ}_K(\mathcal{J})} \mathbf{1}_G \int \mathrm{Mod}_x \widehat{f}\,\overline{\left[\mathsf{P}^{\xi_T}_n(\widehat{Z_G}) - \widehat{Z_G}(\xi_T)\mathbf{1}_{\mathbb{R}}\right]}$$

$$\Pi^{*,\mathsf{cancel}_n}_{\mathbb{Q}'(T)} f := \sum_{G\in \mathrm{CZ}_K(\mathcal{J})} \mathbf{1}_G \int \mathrm{Mod}_x \widehat{f}\,\overline{\left[\widehat{Z_G} - \mathsf{P}^{\xi_T}_n(\widehat{Z_G})\right]}$$

where $n := \frac{M}{10}$ and $\mathsf{P}^{\xi_T}_n(\widehat{Z_G})$ is the order n Taylor polynomial of $\widehat{Z_G}$ centered at ξ_T. First, observe that (2.4.13) follows by taking advantage of the trivial estimate

$$\left|\partial^\nu(\widehat{Z_G})(\xi_T)\right| \lesssim_{M,\nu} K^{-M}(\ell_G)^\nu, \tag{2.4.15}$$

and subsequently setting

$$h := \frac{1}{K^{-M}} \sum_{G\in \mathrm{CZ}_K(\mathcal{J})} \widehat{Z_G}(\xi_T)(\mathbf{1}_G f) \Rightarrow |h| \lesssim_M |f|. \tag{2.4.16}$$

As an intermediate step towards (2.4.14), we first prove a preliminary result under a temporary spatial compact support assumption. Namely, for $\mathsf{i} \in \{\mathsf{poly_n}, \mathsf{cancel_n}\}$, $\Pi^{*,\mathsf{i}}_{\mathbb{Q}'(T)}$ has the property that if $\varphi_P \in \Theta^{\frac{M}{3}}_\iota(I_P,\xi_T)$ with $\operatorname{supp}\varphi_P \subset \Delta I_P$

$$\left|\Pi^{*,\mathsf{poly}_n}_{\mathbb{Q}'(T)}\varphi_P\right| \lesssim \ell_{I_P}^{-1}\Delta K^{-M} \sum_{m=1}^{n} \sum_{G\in \mathrm{CZ}_K(\mathcal{J}),3G\cap\Delta I_P\neq\varnothing} \mathbf{1}_G \left(\frac{\ell_G}{\ell_{I_P}}\right)^m \tag{2.4.17}$$

$$\left|\Pi^{*,\mathsf{cancel}_n}_{\mathbb{Q}'(T)}\varphi_P\right| \lesssim \ell_{I_P}^{-1}\Delta \sum_{G\in \mathrm{CZ}_K(\mathcal{J}),3G\cap\Delta I_P\neq\varnothing} \mathbf{1}_G \left(\frac{\ell_G}{\ell_{I_P}}\right)^{n+1}. \tag{2.4.18}$$

We check (2.4.17). The adaptation $|\widehat{\varphi_P}| \lesssim \Delta\chi^{\frac{M}{3}}_{\omega_{\xi_T,P}}(\xi)$ allows us to estimate

$$\left\|(\xi-\xi_T)^j\,\widehat{\varphi_P}\right\|_1 \lesssim \left(\int_{|\xi-\xi_T|\le \ell_{\omega_P}} |\xi-\xi_T|^j + \int_{|\xi-\xi_T|>\ell_{\omega_P}} |\omega_P|^{\frac{M}{3}}|\xi-\xi_T|^{j-\frac{M}{3}}\right) \lesssim \Delta(\ell_{I_P})^{-(j+1)} \tag{2.4.19}$$

for $1 \leq j \leq n+1$. Combining (2.4.19) with (2.4.15) yields (2.4.17). Finally, an application of (2.4.19) for $j = n+1$ yields (2.4.18). In order to prove (2.4.14), apply Remark 2.4.2 to write ϕ_P as rapidly decaying superposition of wave packets with compact support and use the intermediate estimate (2.4.17) as in

$$\begin{aligned}
&\sum_{k\geq 0} 2^{-\frac{M}{3}k} \sum_{P\in T} \left| \Pi^{*,\mathsf{osc}}_{\mathbb{Q}'(T)} \ \phi_{P,k,\xi_T} \right| \\
&\lesssim \sum_{k\geq 0} 2^{-\frac{M}{3}k} 2^{k(n+1)} \sum_{P\in T} \sum_{\substack{G\in \mathrm{CZ}_K(\mathcal{J}) \\ 3G\subset 3\cdot 2^k I_P}} \mathbf{1}_G \max\left\{ K^{-M}, \left(\frac{\ell_G}{2^k \ell_{I_P}}\right)^n \right\} \left(\frac{\ell_G}{2^k \ell_{I_P}}\right) \\
&\lesssim K^{-\frac{M}{10}} \sum_{k\geq 0} 2^{-\frac{M}{5}k} \sum_{\substack{G\in \mathrm{CZ}_K(\mathcal{J}) \\ 3G\subset 6\cdot 2^k I}} \mathbf{1}_G \sum_{\substack{P\in T \\ 3G\subset 3\cdot 2^k I_P}} \left(\frac{\ell_G}{2^k \ell_{I_P}}\right).
\end{aligned} \tag{2.4.20}$$

The proof of (2.4.14) is then completed by summing up, and taking advantage of the next two observations. First, when $G \in \mathrm{CZ}_K(\mathcal{J})$ and $j \in \mathbb{Z}$ the cardinality estimate

$$\#\left\{P \in T : 3G \subset 3\cdot 2^k I_P, \mathsf{scl}(P) = 2^j\right\} \lesssim (k+2)2^k$$

holds uniformly in j, G. Next, when $G \in \mathrm{CZ}_K(\mathcal{J})$, $J \in \mathcal{J}$ and $3G \cap 2^k J \neq \varnothing$, then $G \subset 3\cdot 2^k J$ necessarily, and in particular $3K^2 \ell_G \leq \ell_{2^k J}$. Therefore, the counting estimate in the last display allows us to perform a single scale analysis in the innermost sum of (2.4.20), and using the disjointness of $G \in \mathrm{CZ}(\mathcal{J})$, we can estimate (2.4.20) by

$$K^{-\frac{M}{10}} \sum_{k\geq 0} 2^{-\frac{M}{6}k} \mathbf{1}_{2^{k+3}I} \lesssim K^{-\frac{M}{10}} \chi_I^{\frac{M}{10}}. \tag{2.4.21}$$

Finally, the proof of the lemma is finished by taking $C(M)$ to be the larger of the two implied constants in (2.4.16), (2.4.21). $\square$

Lemma 2.4.10. *If $f \in L^\infty_0(\mathbb{R})$ and $T \in \mathcal{F}$ there holds*

$$\mathsf{size}_{2,\star,\kappa}(W[\Pi_{\mathbb{M}\setminus\mathbb{Q}}(f)], T) \lesssim K^{-\frac{M}{20}} [f]_{1,\mathbb{P}}.$$

Proof. The proof is carried by splitting $\Pi_{\mathbb{M}\setminus\mathbb{Q}}$ into frequency and spatial tails. Namely,

$$\Pi_{\mathbb{M}\setminus\mathbb{Q}} f = \Pi_{\mathbb{Q}'(T)} f + \Pi_{\mathbb{Q}''(T)} f,$$

and it suffices to check that

$$\max_{\mathsf{i}\in\{',''\}} \mathsf{size}_{2,\star,\kappa}(W[\Pi_{\mathbb{Q}^{\mathsf{i}}(T)} f], T) \lesssim K^{-\frac{M}{20}} [f]_{1,\mathbb{P}}.$$

The case $\mathsf{i} = '$ is dealt with first. By identical considerations to those from the proof of Proposition 2.2.11, it suffices to bound

$$\left\| \sum_{\substack{P\in T'\\ I_P\subset I}} |I_P| \langle f, \Pi^*_{\mathbb{Q}'(T)} \phi_P\rangle h_{I_P} \right\|_{1,\infty} \lesssim K^{-\frac{M}{20}} [f]_{1,\mathbb{P}} |I|$$

for an arbitrary interval $I \in \mathcal{D}$, lacunary tree $T' \subset T$, and collection $\{\phi_P \in \Phi(P) : P \in T'\}$. To this purpose, Lemma 2.4.9 entails

$$\left\| \sum_{\substack{P\in T'\\ I_P\subset I}} |I_P| \langle f, \Pi^{*,\mathsf{osc}}_{Q'(T)} \phi_P\rangle h_{I_P} \right\|_{1} \leq \left\langle |f|, \sum_{\substack{P\in T'\\ I_P\subset I}} \left| \Pi^{*,\mathsf{osc}}_{Q'(T)} [|I_P|\phi_P] \right| \right\rangle \lesssim K^{-\frac{M}{20}} \left\| f \chi_I^{\frac{M}{10}} \right\|_1 \lesssim K^{-\frac{M}{20}} |I| [f]_{1,\mathbb{P}}.$$

Furthermore, (2.4.13) of Lemma 2.4.9 may be used to find $|h| \leq C|f|$ such that the estimate

$$\left\| \sum_{\substack{P\in T'\\ I_P\subset I}} |I_P| \langle f, \Pi^{*,\mathsf{avg}}_{\mathbb{Q}'(T)} \phi_P\rangle h_{I_P} \right\|_{1,\infty} \lesssim K^{-M} |I| [h]_{1,\mathbb{P}} \lesssim K^{-M} |I| [f]_{1,\mathbb{P}}$$

holds. This completes the handling of the term $\mathsf{i} = '$. The term $\mathsf{i} = ''$ is much easier, The definition of $\mathbb{Q}''(T)$ guarantees that $\operatorname{supp} \Pi_{\mathbb{Q}''(T)} f \cap K I_T = \varnothing$. Therefore, an application of Lemma 2.2.11 in the first step yields

$$\mathsf{size}_{2,\star,\kappa}(W[\Pi_{\mathbb{Q}''(T)} f], T) \lesssim K^{-\frac{M}{2}} \left\| \Pi_{\mathbb{Q}''(T)} f \right\|_\infty .$$

while properties (ii) and (v) of the spatial intervals $\mathrm{CZ}_K(\mathcal{J})$ guarantee the bound

$$\left\|\Pi_{\mathbb{Q}''(T)}f\right\|_\infty \lesssim K \sup_{G\in\mathrm{CZ}_K(\mathcal{J})} \inf_G M(f) \lesssim K^3[f]_{1,\mathbb{P}}.$$

The claim of the lemma for the $\mathbb{Q}''(T)$ component is then an immediate consequence of the last two displays. $\square$

2.4.11 Conclusion of the proof

The choices

$$K := N^{\frac{1}{p'}\left(\frac{10}{M}+\frac{1}{6t'}\right)}, \qquad M := M(t) = 30\lceil 2^8 t'\rceil \tag{2.4.22}$$

and the decomposition (2.4.8) are now made explicit. The choice of M, anticipated in Remark 2.4.2, ensures (2.4.9), (2.4.10) both hold. In view of the minimal tiles expansion (2.4.5), set in (2.4.8)

$$g := \Pi_{\mathbb{Q}}f, \qquad b := \Pi_{\mathbb{M}\setminus\mathbb{Q}}f.$$

Turn to the verification of (2.4.9)-(2.4.10). For the first, write

$$\Pi_{\mathbb{Q}}f = \sum_{G\in\mathrm{CZ}_K(\mathcal{J})} \Pi_{\mathbb{Q}[G]}f = \sum_{G\in\mathrm{CZ}_K(\mathcal{J})} (f\mathbf{1}_G) * U_G, \quad U_G := \sum_{P\in\mathbb{Q}[G]} \eta_P.$$

A straightforward use of Plancherel's theorem entails that $\left\|\Pi_{\mathbb{Q}[G]}\right\|_{2\to 2} \lesssim 1$. Furthermore, observe that $\widehat{U_G}$ is a sum of $\#\mathbb{Q}\,[G]$ Schwartz functions uniformly adapted to disjoint intervals of length ℓ_G^{-1}, leading to the estimates

$$\left\|\widehat{U_G}\right\|_1 \lesssim \frac{\#\mathbb{Q}[G]}{\ell_G}, \quad \left\|\widehat{U_G}\right\|_\infty \lesssim 1 \implies \left\|\Pi_{\mathbb{Q}[G]}\right\|_{p\to 2} \lesssim \left(\frac{\#\mathbb{Q}\,[G]}{\ell_G}\right)^{\frac{1}{p}-\frac{1}{2}}$$

where the implication is obtained by log-convexity, Young's inequality, and finally Riesz-Thorin interpolation of the $(2,1)$ and $(2,2)$ estimates. Preliminarily, also note

$$\#\mathbb{Q}\,[G] \leq 3K \inf_{G} \sum_{T\in\mathcal{T}} \mathbf{1}_{3KI_T} \qquad \forall G \in \mathrm{CZ}_K(\mathcal{J}). \tag{2.4.23}$$

Estimate (2.4.9) then follows by combining (2.4.23), the fact that $\operatorname{supp}\Pi_{\mathbb{Q}[G]}f \subset 3G$ and the finite overlap (ii) of $\{3G : G \in \mathrm{CZ}_K(\mathcal{J})\}$ in the string of inequalities

$$\begin{aligned} \|\Pi_{\mathbb{Q}}f\|_2 &\lesssim \left(\sum_{G\in \mathrm{CZ}_K(\mathcal{J}), G\subset 9KJ} \left\|\Pi_{\mathbb{Q}[G]}f\right\|_2^2 \right)^{\frac{1}{2}} \lesssim |J|^{\frac{1}{p'}} K^3 \left(\sum_{T\in\mathcal{T}} |I_T|\right)^{\frac{1}{2}-\frac{1}{p'}} [f]_{p,\mathbb{P}} \\ &\lesssim |J|^{\frac{1}{2}} N^{\frac{1}{2}-\frac{1}{tp'}} [f]_{p,\mathbb{P}} \end{aligned}$$

as claimed. For the property (2.4.10), by Lemma 2.2.12 it suffices to check that for each $T \in \mathcal{T}$ there holds

$$\mathsf{size}_{2,\star}(\Pi_{\mathbb{M}\setminus\mathbb{Q}}f, T) \lesssim N^{-\frac{1}{p'}} [f]_{p,\mathbb{P}}.$$

Taking notice of the relation between M and t in (2.4.22), this was proved in Lemma 2.4.10.

2.5 Proof of Theorem B

Fix a tiling $\mathbb{S} = \mathbb{S}_{\mathcal{D},\mathcal{D}'}$, $f_j \in L^\infty_0(\mathbb{R})$, $j = 1, 2$. The crux of the matter is to establish the estimate

$$\sup_{\mathbb{P}\subset\mathbb{S}\text{ finite}} \mathsf{C}_{\mathbb{P}}(f_1, f_2) \lesssim \frac{1}{\varepsilon} \left\| \mathrm{M}_{\left(\frac{1}{1-\varepsilon},1\right)}(f_1, f_2) \right\|_1 \tag{2.5.1}$$

with implied constant independent of $\varepsilon > 0$, referring to the model sums (2.1.9). In fact, if $\mathcal{C}$ stands for (1.0.3), in view of (2.1.6), the form $\langle \mathcal{C}f_1, f_2\rangle$ is controlled by the sum of $\lesssim 1$ terms of the type appearing in the left hand side of (2.5.1). The same sparse bound for the periodic operator (1.0.5) then follows from (1.0.4) via a standard transference type argument based on the Stein-Weiss lemma, see e.g [65, Appendix A].

We turn to the proof of (2.5.1), fixing $0 < \varepsilon \leq \frac{1}{2}$, and a finite collection $\mathbb{P} \subset \mathbb{S}$. To unify notation below, it is convenient to write $q_1 = \frac{1}{1-\varepsilon}, q_2 = 1$ and $\vec{q} = (q_1, q_2)$ below. Let $\mathcal{Q} \subset \mathcal{D}$ be a partition of $\mathbb{R}$ with the property that $\operatorname{supp} f_j \subset 3Q$ for $j = 1, 2$, $Q \in \mathcal{Q}$. For each $Q \in \mathcal{Q}$, define $\S_0(Q) = \{Q\}$ and inductively for $m \geq 0$

$$\mathcal{B}(S) \coloneqq \text{maximal } B \in \mathcal{D} \text{ with } B \subset S \cap \bigcup_{j=1}^{2} \left\{ \mathrm{M}_{q_j} \left[f_j \mathbf{1}_{3S} \right] > \Theta \langle f_j \rangle_{q,3S} \right\}, \quad S \in \S_m(Q),$$

$$\mathcal{S}_{m+1}(Q) \coloneqq \bigcup_{S \in \S_m(Q)} \mathcal{B}(S).$$

Finish by setting $\mathcal{S} \coloneqq \bigcup_{Q \in \mathcal{Q}} \bigcup_{m \geq 0} \mathcal{S}_m(Q)$. Note that the sets $\{E_S : S \in \mathcal{S}\}$ defined by $E_S \coloneqq S \setminus [\bigcup \{B : B \in \mathcal{B}(S)\}]$ are pairwise disjoint, and the packing condition

$$\sum_{B \in \mathcal{B}(S)} |B| \leq \frac{1}{4}|S|, \qquad S \in \mathcal{S}$$

which holds provided the absolute constant Θ is picked suitably large, guarantees that $|S| \leq 2|E_S|$ for all $S \in \mathcal{S}$. Also, the iterated stopping interval nature of the collection $\mathcal{S}$ yields

$$\sup_{\substack{I \subset S \\ I \not\subset \bigcup_{B \in \mathcal{B}(S)} B}} \inf_{I} \mathrm{M}_{q_j} f_j \lesssim \langle f_j \rangle_{q_j,3S}, \qquad S \in \mathcal{S},\ j = 1, 2.$$

Therefore, the partition

$$\mathbb{P} = \bigsqcup_{S \in \mathcal{S}} \mathbb{P}(S), \qquad \mathbb{P}(S) \coloneqq \left\{ P \in \mathbb{P} : I_P \subset S, I_P \not\subset \bigcup_{B \in \mathcal{B}(S)} B \right\}$$

inherits the property

$$[f_j]_{q_j,\mathbb{P}(S)} \lesssim \langle f_j \rangle_{q_j,3S}, \qquad S \in \mathcal{S},\ j = 1, 2. \tag{2.5.2}$$

By virtue of (2.2.22), we may apply Proposition 2.2.9 with $p_1 = 2q_1' = \frac{2}{\varepsilon}$, $a = 2$ and $p_2 = 1$ in the second step, and pass to the second line through an appeal to Theorem H with $t = 2$ and $p = q_1$ for W, and Proposition 2.3.1 for A, obtaining

$$\begin{aligned}\mathsf{C}_{\mathbb{P}}(f_1, f_2) = \sum_{S\in\mathcal{S}} \mathsf{C}_{\mathbb{P}(S)}(f_1, f_2) &\lesssim \frac{1}{\varepsilon} \sum_{S\in\mathcal{S}} |S| \left\|W[f_1]\mathbf{1}_{\mathbb{P}(S)}\right\|_{X_2^{2q_1',\infty}(S,\mathsf{size}_{2,1,\star})} \left\|A[f_2]\mathbf{1}_{\mathbb{P}(S)}\right\|_{Y^{1,\infty}(S,\mathsf{size}_{\mathsf{C}})} \\ &\lesssim \frac{1}{\varepsilon} \sum_{S\in\mathcal{S}} |S| \prod_{j=1,2} [f_j]_{q_j,\mathbb{P}(S)} \lesssim \frac{1}{\varepsilon} \sum_{S\in\mathcal{S}} |E_S| \inf_{E_S} \mathrm{M}_{\vec{q}}(f_1, f_2) \le \frac{1}{\varepsilon} \|\mathrm{M}_{\vec{q}}(f_1, f_2)\|_1.\end{aligned}$$

The middle almost-inequality in the last line relies on (2.5.2) as well as $|E_S| \gtrsim |S|$, while the final step is due to the pairwise disjointness of E_S, $S \in \mathcal{S}$. The proof is thus complete.

2.6 Proof of Theorem G

2.6.1 Rank 1 forms

This paragraph devises a reformulation, within our framework, of the trilinear forms discretizing multipliers with singularity along a rank one subspace, such as the bilinear Hilbert transforms. Although these date back in essence to the works of Lacey-Thiele [46,47], they appear in a form closer to ours in [60]. The main change in our definition with respect to the usual one is that our does not involve multi-tiles, at least explicitly, to avoid reformulating outer L^p-spaces and use our embedding theorems in the most direct way possible.

Fix $\kappa \ge 1$, two dual dyadic grids $\mathcal{D}, \mathcal{D}'$ and the tiling $\mathbb{S} = \mathbb{S}_{\mathcal{D},\mathcal{D}'}$. Our construction, similarly to [60], relies on the two order relations on $\mathbb{S}$

$$P \lesssim_\kappa P' \iff I_P \subset I_{P'},\ \omega_{P'}^{\mathsf{p}(\kappa)} \subset \omega_P^{\mathsf{p}(\kappa)}, \qquad \kappa \ge 1,$$

$$P \lesssim_\kappa' P' \iff P \lesssim_\kappa P' \text{ and } P \not\lesssim_1 P', \qquad \kappa \ge 2.$$

Note that $\lesssim_\kappa$ has been already defined in (2.3.3) and is recalled here for the reader's convenience. Let $\kappa \geq 10$, $\mathbb{P}$ be a finite subset of $\mathbb{S}$ with scales separated by a factor of $2^{5\kappa}$, and $\eta = (\eta_1, \eta_2, \eta_3) : \mathbb{P} \to \mathbb{S} \times \mathbb{S} \times \mathbb{S}$ have the properties

r1. the components $\eta_j : \mathbb{P} \to \mathbb{S}$ are injective maps for $j = 1, 2, 3$;

r2. $I_{\eta_j(P)} = I_P$ for $j = 1, 2, 3$;

r3. if $P, P' \in \mathbb{P}$ are such that $\eta_j(P) \lesssim_1 \eta_j(P')$ for some $j \in \{1, 2, 3\}$ then $\eta_k(P) \lesssim_\kappa \eta_j(P')$ for all $k \in \{1, 2, 3\}$;

r4. if $P, P' \in \mathbb{P}$ are such that $\eta_j(P) \lesssim_1 \eta_j(P')$ for some $j \in \{1, 2, 3\}$ then $\eta_k(P) \lesssim_\kappa \eta_k(P')$ for all $k \in \{1, 2, 3\}$, and in fact $\eta_k(P) \lesssim'_\kappa \eta_k(P')$ for at least two indices $k \in \{1, 2, 3\}$.

It is convenient to denote by $\mathbb{P}_j$, $j \in \{1, 2, 3\}$ the ranges of η_j. The *rank 1 form* of parameter κ associated to η and $\mathbb{Q} \subset P$ acts on a triple $f_j \in L_0^\infty(\mathbb{R})$ by

$$\Lambda_{\eta,\mathbb{Q}}(f_1, f_2, f_3) = \sum_{P \in \mathbb{Q}} |I_P| \prod_{j \in \{1,2,3\}} W[f_j](\eta_j(P))$$

where W stands for the wave packet transform (2.1.2).

A typical example of map η satisfying r1. to r4. and thus giving rise to rank 1 forms is the following. Let Γ' be a 1-dimensional subspace of $\Gamma = \{\xi \in \mathbb{R}^3 : \xi_1 + \xi_2 + \xi_3 = 0\}$ as in the statement of Theorem G and $\mathcal{Q} \subset \mathcal{D}' \times \mathcal{D}' \times \mathcal{D}'$ be a finite collection satisfying

g1. $\ell_{Q_1} = \ell_{Q_2} = \ell_{Q_3} \in 2^{H\mathbb{Z}+h}$ for all $Q = Q_1 \times Q_2 \times Q_3 \in \mathcal{Q}$,

g2. $Q \cap \Gamma \neq \varnothing$ for all $Q \in \mathcal{Q}$, g3. $K\ell_{Q_1} \leq \operatorname{dist}(Q,\Gamma') \leq K^2\ell_{Q_1}$ for all $Q \in \mathcal{Q}$

for parameters $H, K \in \mathbb{N}$ and $h \in \{0,\dots,H-1\}$. If H, K are sufficiently large parameters depending on Γ', conditions g1. to g3. tell us that the collection $\{Q \in \mathcal{Q} : Q_1 = \omega\}$ has at most one element for each $\omega \in \mathcal{D}'$, see e.g. [60, Lemma 6.2]. If such collection is nonempty, we may then write $Q = \omega \times Q_2(\omega) \times Q_3(\omega)$ for its unique element. Of course, the index $j = 1$ can be replaced by any other index in a symmetric statement. In this setting, if $\mathbb{P} = \mathbb{P}_1$ is a finite subset of $\{P \in \mathbb{S} : \omega_P = Q_1 \text{ for some } Q \in \mathcal{Q}\}$ the map

$$\eta : \mathbb{P} \to \mathbb{S} \times \mathbb{S} \times \mathbb{S}, \qquad \eta(P) = (P, I_P \times Q_2(\omega_P), I_P \times Q_3(\omega_P))$$

satisfies r1. to r4. The usual model sum reduction of [60] may be then summarized in the statement that the singular multipliers (2.0.9) lie in the closed convex hull of rank 1 forms as defined above, with parameter κ chosen sufficiently large depending on the parameter K in g3. Therefore Theorem G will follow from the estimate

$$\Lambda_{\eta,\mathbb{P}}(f_1, f_2, f_3) \leq \frac{C}{\varepsilon(\vec{p})} \left\| \mathrm{M}_{\vec{p}}(f_1, f_2, f_3) \right\|_1 \tag{2.6.1}$$

uniformly over all rank 1 forms, for all tuples $\vec{p} = (p_1, p_2, p_3)$ satisfying the conditions in (2.0.10). Symmetry in the indices p_1, p_2, p_3 and a complex interpolation argument allow us to restrict ourselves to tackling (2.6.1) in the extremal case $p_1 = \frac{1}{1-\varepsilon}, p_2 = 2 = p_3$, for which $\varepsilon(\vec{p}) = \varepsilon$. We do so in the next paragraph.

2.6.2 Using the wave packet embeddding

We now prove (2.6.1) in the above mentioned extremal case. By eventually composing η with the inverse of η_1, we reduce to the case where $\mathbb{P} = \mathbb{P}_1$ and η_1 is the identity map. Properties r1. to r4. of the map η associated to a rank 1 form of fixed constant κ come into play via the following observation. If $\varnothing \subsetneq S \subset \mathbb{S}$, $Q \in S$, then the set

$$S(Q) \coloneqq \{P \in S : P \lesssim_1 Q\}$$

is a 1-tree with top (I_Q, c_{ω_Q}). Property r3. tells us that the sets $\eta_j(S(Q)) \coloneqq \{\eta_j(P) : P \in \mathbb{S}(Q)\}$ are κ-trees with top $(I_{\eta_j(Q)}, c_{\omega_{\eta_j(Q)}}) = (I_Q, c_{\omega_{\eta_j(Q)}})$, $j = 1, 2, 3$. Furthermore, property r4. and scale separation as in the proof of Lemma 2.2.3 allows us to decompose

$$S(Q) = \bigcup_{j=1}^{3} S(Q, j)$$

with each $S(Q, j)$ having the property that $\eta_k(S(Q, j))$, obviously contained in $\eta_k(S(Q))$, is a lacunary κ-tree with top $(I_Q, c_{\omega_{\eta_k(Q)}})$ for $k \in \{1, 2, 3\} \setminus \{j\}$. In accordance with this property, we define three new variants of (2.2.5) on the outer measure space $(\mathbb{S}^J, \mathcal{T}^{J,1}, \mu^{J,1})$. Setting for $k = 1, 2, 3$

$$\mathsf{size}_{2,\star,k}(F, T) \coloneqq \sup \left\{ \mathsf{size}_2(F \circ \eta_k^{-1}, U) : U \subset \eta_k(T), U \text{ lacunary } \kappa\text{-tree} \right\}, \qquad T \in \mathcal{T}^{J,1}$$

we have the estimate

$$\mathsf{size}_1(F_1 F_2 F_3, T) \lesssim \prod_{k=1,2,3} \mathsf{size}_{2,\star,k}(F_k, T), \qquad \forall T \in \mathcal{T}^{J,1}. \tag{2.6.2}$$

This inequality, proved at the end of this section, is our analogue of the usual *tree estimate*, see e.g. [60, Lemma 7.3], and it is essentially the only additional piece of machinery we were left to set up. Indeed, if $\mathbb{Q} \subset P \cap \mathbb{S}^J$ is arbitrary, (2.6.2) allows us to appeal to Proposition 2.2.9 with the obvious choice of exponents, and obtain the chain of inequalities

$$
\begin{aligned}
&\Lambda_{\eta,\mathbb{Q}}(f_1, f_2, f_3) \leq \left\| \mathbf{1}_{\mathbb{Q}} \prod_{k=1}^{3} (W[f_k] \circ \eta_k) \right\|_{\ell^1(\mathbb{S}^J)} \\
&\lesssim \frac{|J|}{\varepsilon} \left\| \mathbf{1}_{\mathbb{Q}} (W[f_1] \circ \eta_k) \right\|_{X_2^{\frac{2}{\varepsilon},\infty}(\mathbb{S}^J,1,\mathsf{size}_{2,\star,1})} \prod_{k=2,3} \left\| \mathbf{1}_{\mathbb{Q}} (W[f_k] \circ \eta_k) \right\|_{Y^{2,\infty}(\mathbb{S}^J,1,\mathsf{size}_{2,\star,k})} \\
&\leq \frac{|J|}{\varepsilon} \left\| \mathbf{1}_{\eta_1(\mathbb{Q})} (W[f_1]) \right\|_{X_2^{\frac{2}{\varepsilon},\infty}(\mathbb{S}^J,\kappa,\mathsf{size}_{2,\star})} \prod_{k=2,3} \left\| \mathbf{1}_{\eta_k(\mathbb{Q})} (W[f_k]) \right\|_{Y^{2,\infty}(\mathbb{S}^J,\kappa,\mathsf{size}_{2,\star})} \\
&\lesssim \frac{|J|}{\varepsilon} [f_1]_{\frac{1}{1-\varepsilon},\mathbb{Q}} \prod_{k=2,3} [f_k]_{2,\mathbb{Q}}.
\end{aligned}
\tag{2.6.3}
$$

The passage to the third line follows by transport of structure, while for the subsequent step we have applied Theorem H and estimate (2.4.1), and used that the spatial components, and thus the corresponding local tile norms on $\mathbb{Q}$, are invariant under η. With (2.6.3) in hand, a stopping procedure akin to that devised in Section 2.5 easily leads to (2.6.1). Details are left to the interested reader.

Proof of (2.6.2). Let $T \in \mathcal{T}^{J,1}$ be a 1-tree with top (I_T, ξ_T), and $\mathsf{m}(T)$ be the set of those $Q \in T$ which are maximal with respect to $\lesssim_1$. As $\xi_T \in \omega_P^{\mathsf{p}(1)}$ for all $P \in T$, it must hold that $I_Q \cap I_{Q'} = \varnothing$ whenever $Q, Q' \in \mathsf{m}(T)$ with $Q \neq Q'$. Clearly, T is the disjoint union of the 1-trees $\{T(Q) : Q \in \mathsf{m}(T)\}$. Simply from the definitions and the disjointness we just stressed

$$\mathsf{size}_1(F,T) \leq \sum_{Q \in \mathsf{m}(T)} \frac{|I_Q| \mathsf{size}_1(F, T(Q))}{|I_T|} \leq \sup_{Q \in \mathsf{m}(T)} \mathsf{size}_1(F, T(Q)),$$

$$\sup_{Q \in \mathsf{m}(T)} \mathsf{size}_{2,\star,k}(F, T(Q)) \leq \mathsf{size}_{2,\star,k}(F, T)$$

where the second bound holds for $k \in \{1, 2, 3\}$ and follows by obvious inclusion considerations. The last two inequalities tell us that it suffices to prove (2.6.2) for $T = T(Q)$. In that case,

$$
\begin{aligned}
&\mathsf{size}_1(F_1F_2F_3, T(Q)) \leq \sum_{j=1}^{3} \mathsf{size}_1(F_1F_2F_3, T(Q, j)) \\
&\leq \sum_{j=1}^{3} \mathsf{size}_\infty(F_j, T(Q, j)) \prod_{k \neq j} \mathsf{size}_2(F_k, T(Q, j)) \\
&= \sum_{j=1}^{3} \sup_{P \in T(Q,j)} \mathsf{size}_2 \left(F_j \circ \eta_j^{-1}, \eta_j(\{P\})\right) \prod_{k \neq j} \mathsf{size}_2 \left(F_k \circ \eta_k^{-1}, \eta_k(T(Q, j))\right) \\
&\leq 3 \prod_{k=1}^{3} \mathsf{size}_{2,\star,k} \left(F_k, T(Q)\right)
\end{aligned}
$$

as desired. We have used in the last step the lacunarity of $\eta_k(T(Q, j))$ and of the single tile trees $\{\{P\} : P \in T(Q, j)\}$. $\square$

3. Multilinear wavelet T(1) theorem

3.1 Preliminaries

An $(m+1)$-linear form Λ defined on the $(m+1)$-fold product of the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is a singular integral form if its off-diagonal kernel satisfies the standard size and smoothness estimates (see Definition 3.3.1 below). A singular integral form Λ is bounded on $L^{p_0}(\mathbb{R}^d) \times \cdots \times L^{p_m}(\mathbb{R}^d)$ for all

$$1 < p_j \leq \infty, \quad \sum_{j=0}^{m} \frac{1}{p_j} = 1, \tag{3.1.1}$$

if and only if Λ is *Calderón-Zygmund*, which we take to mean that it satisfies the weak boundedness property (see Definition 3.3.2 below) and the following multilinear $T(1)$ condition: There exists $\mathfrak{b}_j \in \mathrm{BMO}(\mathbb{R}^d)$ such that for every ϕ in $\mathcal{S}(\mathbb{R}^d)$ with mean zero,

$$\Lambda^{*,j}(\phi, 1, \ldots, 1) = \langle \phi, \mathfrak{b}_j \rangle, \tag{3.1.2}$$

where $\Lambda^{*,j}$ permutes the 0th and jth argument (see (3.1.3) below). When $m = 1$, this is the well known $T(1)$ theorem of David and Journé [17] which was extended to $m \geq 2$ by Grafakos and Torres [36].

Our goal in this chapter is to prove a $T(1)$ theorem for compactness of multilinear singular integral forms. The first difference between the boundedness problem and compactness problem is that compactness is a property of operators, while boundedness in the reflexive range

(3.1.1) can be equivalently stated in terms of forms. Accordingly for each $j = 0, 1, \ldots, m$, we associate to Λ the m-linear adjoint operators $T^{*,j}$ and transposed forms $\Lambda^{*,j}$ by

$$\langle f_0, T^{*,j}(f_1, \ldots, f_m)\rangle = \Lambda^{*,j}(f_0, f_1, \ldots, f_m) = \Lambda(f_j, f_1, \ldots, f_{j-1}, f_0, f_{j+1}, \ldots, f_m). \quad (3.1.3)$$

If Λ is Calderón-Zygmund, then we say each $T^{*,j}$ is an m-linear Calderón-Zygmund operator (CZO). Furthermore, define $\mathsf{T}_\Lambda = \{T^{*,j}\}_{j=0}^m$.

In addition to that, for each $\sigma \in \mathrm{S}_{m+1}$, the permutation group on $\{0, 1, \ldots, m\}$, and Λ a $m+1$- linear form we define

$$\Lambda^\sigma(f_0, \ldots, f_m) = \Lambda(f_{\sigma(0)}, \ldots, f_{\sigma(m)}).$$

3.2 Wavelets

In this section we will review some preliminaries regarding wavelets and introduce one of the building blocks of our representation, wavelet forms. The notation which will be expanded below will also be used in Chapter 4.

3.2.1 Analysis of the parameter space

Introduce the parameter space

$$Z^d = \{z = (w, t) : w \in \mathbb{R}^d,\ t > 0\},$$

whose elements $z = (w, t)$ act on functions $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ functions by the formula

$$\mathsf{Sy}_z f := f_z = \frac{1}{t^d} f\left(\frac{\cdot - w}{t}\right).$$

Let μ be the measure on Z^d given by

$$\int_{Z^d} F(z)\,\mathrm{d}\mu(z) = \int_0^\infty \int_{\mathbb{R}^d} F(w,t)\,\frac{\mathrm{d}w\,\mathrm{d}t}{t}, \quad F \in \mathcal{C}_0(Z^d).$$

Notice that μ is invariant under Sy_z. To analyze multilinear operators, we will use wavelets adapted to two parameters, one in Z^{md} and the other in Z^d. First, given $\mathbf{w} \in \mathbb{R}^{md}$ and $w_0 \in \mathbb{R}^d$, define

$$|\mathbf{w} - w_0|_2 = \sqrt{\sum_{i=1}^m |w_i - w_0|^2}, \quad \mathbf{w} = (w_1, \dots, w_m), \quad w_i \in \mathbb{R}^d.$$

For $\mathbf{z} = (\mathbf{w}, s) \in \mathbb{Z}^{md}$ and $\zeta = (w_0, t) \in \mathbb{Z}^d$, define

$$[\mathbf{z}, \zeta]_\delta = \frac{\min\{s,t\}^\delta}{\max\{t, s, |\mathbf{w} - w_0|_2\}^{md+\delta}}.$$

We will also say $\mathbf{z} \geq \zeta$ if $s \geq t$. Notice that if $\delta \geq \delta'$ then $[\mathbf{z}, \zeta]_\delta \leq [\mathbf{z}, \zeta]_{\delta'}$. For $M \geq 1$ and $\zeta \in Z^d$, introduce $\mathbb{B}_M^m(\zeta)$ which are the following approximate balls in the hyperbolic metric,

$$\mathbb{B}_M^m(\zeta) = \left\{\mathbf{z} \in Z^{md} : t2^{-M} \leq s \leq t2^M,\ |\mathbf{w} - w_0|_2 \leq t2^M\right\}$$

when $\zeta = (0,1)$, $M = 1$, or $m = 1$, those parameters are omitted from the notation. Given a function $F : Z^d \to \mathbb{C}$ we say $\lim\limits_{z\to\infty} F(z) = L$ if

$$\lim_{M\to\infty} \sup_{z \notin \mathbb{B}_M} |F(z) - L| = 0.$$

3.2.2 Wavelet classes and forms

The building blocks of our representation theorem are wavelets for which we are going to introduce notation and relevant classes as well as the averaging lemmata from [21, 25]. We denote the space of Schwartz functions by $\mathcal{S}(\mathbb{R}^d)$ and the mean-zero subspace

$$\mathcal{S}^0(\mathbb{R}^d) = \left\{ \varphi \in \mathcal{S}(\mathbb{R}^d) : \int \varphi(x)\,\mathrm{d}x = 0 \right\}.$$

We fix a radial function $\phi \in \mathcal{S}^0(\mathbb{R}^d)$, supported in a ball and appropriately normalized which we will call the mother wavelet, in which case the Calderón reproducing formula holds, namely

$$f = \int_{Z^d} \langle f, \phi_z \rangle \phi_z \,\mathrm{d}\mu(z) \qquad \forall f \in \mathcal{S}(\mathbb{R}^d). \tag{3.2.1}$$

For the convenience of the reader, we restate the setup from [21] on which we will base our analysis. For $0 < \delta \leq 1$ we introduce the norm on functions $\varphi \in \mathcal{C}^\delta(\mathbb{R}^{md})$,

$$\|\varphi\|_{\star,\delta} = \sup_{\substack{x,h \in \mathbb{R}^{md} \\ 0<|h|\leq 1}} \langle x \rangle^{md+\delta} \left(|\varphi(x)| + \frac{|\varphi(x+h) - \varphi(x)|}{|h|^\delta} \right), \quad \langle x \rangle = \sqrt{1+|x|^2}.$$

Definition 3.2.3. For $z = (w,t) \in \mathbb{Z}^d$, the wavelet class $\Psi_z^{m,\delta}$ is defined by

$$\Psi_z^{m,\delta} = \left\{ \varphi \in \mathcal{C}^\delta(\mathbb{R}^{md}) : \|(\mathsf{Sy}_{\mathbf{z}})^{-1}\varphi\|_{\star,\delta} \leq 1 \right\}, \quad \mathbf{z} = (w, \dots, w, t) \in Z^{md},$$

and its cancellative subclass, for $j = 1, \dots, m$ is denoted by $\Psi_z^{m,\delta;j}$ and consists of $\varphi \in \Psi_z^{m,\delta}$ such that

$$\int_{\mathbb{R}^d} \varphi(x_1, x_2, \dots, x_m)\,\mathrm{d}x_j = 0.$$

Let χ_z denote the L^∞-normalized decay factor adapted to the parameter $z = (w,t)$,

$$\chi_z(x) = \left\langle \frac{x-w}{t} \right\rangle^{-1}.$$

With this notation, we can recast $\varphi \in \Psi_z^{m,\delta}$ as

$$|\varphi| \leq \frac{1}{t^{md}} \chi_{\mathbf{z}}^{md+\delta}, \quad |\varphi - \varphi(\cdot + h)| \leq \frac{|h|^\delta}{t^{md+\delta}} \chi_{\mathbf{z}}^{md+\delta}.$$

With the goal of making this thesis as self contained as possible, as well as to set the stage for the representation in Proposition 3.3.8, we state the averaging lemmata from [21] and a refinement of the averaging procedure in [25].

Lemma 3.2.4. *Let ϕ be the mother wavelet and $k \geq 0$. There exist functions ψ^i, $i = 1, 2, 3, 4$, satisfying*

(i) $\operatorname{supp} \psi^i \subset B(0,1)$*;*

(ii) $\psi^1, \psi^3 \in \mathcal{C}^k(\mathbb{R}^d)$*;*

(iii) $\psi^2, \psi^4 \in \mathcal{S}^0(\mathbb{R}^d)$*;*

(iv) For any $s > 0$ and $f \in \mathcal{S}(\mathbb{R}^d)$,

$$\int_{r \geq s} \int_{u \in \mathbb{R}^d} \langle f, \phi_{u,r} \rangle \phi_{u,r} \frac{\mathrm{d}u \ \mathrm{d}r}{r} = \int_{\mathbb{R}^d} \langle f, \psi^1_{u,s} \rangle \psi^2_{u,s} + \langle f, \psi^3_{u,s} \rangle \psi^4_{u,s} \, \mathrm{d}u.$$

Lemma 3.2.5. *Let $\varphi_j \in \mathcal{S}(\mathbb{R}^d)$ for $j = 1, \dots, m$ and $0 < \eta < \delta \leq 1$. There exists $C > 0$ such that for any $H : Z^{md} \times Z^d \to \mathbb{C}$ satisfying*

$$|H(\mathbf{z}, \zeta)| \leq [\mathbf{z}, \zeta]_\delta \,,$$

there holds

$$\nu_\zeta = \int_{\substack{\mathbf{z} \in Z^{md} \\ \mathbf{z} \geq \zeta}} H(\mathbf{z}, \zeta) \left(\varphi_1 \otimes \cdots \otimes \varphi_m\right)_{\mathbf{z}} \mathrm{d}\mu(\mathbf{z}) \in C \Psi_\zeta^{m,\eta}.$$

Furthermore, if $\phi_j \in \mathcal{S}^0$, then $\nu_\zeta \in C \Psi_\zeta^{m,\eta;j}$.

Finally, we end this section by introducing the wavelet forms, which will be used to systematically study m-linear CZOs.

Definition 3.2.6. Given a collection $\{\varepsilon_z \in \mathbb{C}, \nu_z \in \Psi_z^{m,\delta;1} : z \in Z^d\}$ define the associated $(m+1)$-linear *canonical wavelet form* by

$$\int_{Z^d} \varepsilon_z \langle f_0, \phi_z \rangle \langle f_1 \otimes \ldots \otimes f_m, \nu_z \rangle \, \mathrm{d}\mu(z).$$

More generally, we say U is a wavelet form if U^σ is a canonical wavelet form for some $\sigma \in \mathrm{S}_{m+1}$. We say U is a *bounded wavelet form* if $\sup_{z \in Z^d} |\varepsilon_z| < \infty$. Additionally, we say U is a *compact wavelet form* if it is a bounded wavelet form for which

$$\lim_{z \to \infty} \varepsilon_z = 0.$$

Wavelet forms may be viewed as a generalization of the Calderón-Toeplitz operators considered in [64, 70]. More generally though, all cancellative CZ forms can be realized as wavelet forms [21, 25]. Bounded wavelet forms are bounded in the following sense. In fact, they themselves are cancellative CZ forms so Proposition 3.2.8 below follows from any number of results [36, 51, 53, 55, 63]. See also [21, Proposition 5.1] for a direct proof of the sparse $(1, \ldots, 1)$ bound for bounded wavelet forms. Let us introduce some bookkeeping to concisely describe the full range of Lebesgue space estimates for CZOs.

Definition 3.2.7. Let

$$P = \{(p_1, \ldots, p_m) : 1 < p_j \leq \infty\}.$$

We introduce the shorthand for $\vec{p} \in P$,

$$L^{\vec{p}}(\mathbb{R}^d) = \bigtimes_{j=1}^{m} L^{p_j}(\mathbb{R}^d), \quad \mathcal{B}^{\vec{p}} = \left\{(f_1, \ldots, f_m) \in L^{\vec{p}} : \|f_j\|_{L^{p_j}(\mathbb{R}^d)} \leq 1\right\}, \quad r(\vec{p}) = \left(\sum_{j=1}^{m} \frac{1}{p_j}\right)^{-1}.$$

Then the admissible classes of Hölder tuples we consider are

$$Q = \{(\vec{p}, r(\vec{p})) : \vec{p} \in P,\ r(\vec{p}) < \infty\}\,.$$

Given $(\vec{p}, r) \in Q$, and an m-linear operator T, denote by $\|T\|_{\vec{p},r}$ the operator norm from $L^{\vec{p}}(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$, i.e.

$$\|T\|_{\vec{p},r} = \sup_{(f_1,\ldots,f_m)\in\mathcal{B}^{\vec{p}}} \|T(f_1,\ldots,f_m)\|_{L^r(\mathbb{R}^d)}\,.$$

Furthermore, we define the following modification at the endpoint,

$$\|T\|_{\infty,\mathrm{BMO}} = \sup_{(f_1,\ldots,f_m)\in\mathcal{B}^{\vec{p}}} \|T(f_1,\ldots,f_m)\|_{\mathrm{BMO}(\mathbb{R}^d)}\,, \quad \vec{p} = (\infty,\ldots,\infty).$$

Proposition 3.2.8. *For each $(\vec{p}, r) \in Q$, there exists $C_{\vec{p},r} > 0$ such that for any bounded wavelet form U and any $T \in \mathsf{T}_U$,*

$$\|T\|_{\vec{p},r} \leq C_{\vec{p},r} \sup_{z\in Z^d} |\varepsilon_z|\,. \tag{3.2.2}$$

Furthermore, there exists $C_\infty > 0$ such that

$$\|T\|_{\infty,\mathrm{BMO}} \leq C_\infty \sup_{z\in Z^d} |\varepsilon_z|\,. \tag{3.2.3}$$

3.3 Wavelet representation of compact Calderón-Zygmund forms

Let us make rigorous the informal definitions given in the introduction.

Definition 3.3.1. A function $K \in L^1_{\mathrm{loc}}(\mathbb{R}^{(m+1)d} \setminus \{x \in (\mathbb{R}^d)^{m+1} : x_0 = \cdots = x_m\})$ is a *δ-singular integral kernel* if there exist $C_K, \delta > 0$ such that

$$|K(x_0,\ldots,x_m)| \leq \frac{C_K}{(\sum_{j=1}^m |x_0 - x_j|)^{md}},$$

$$\max_{j=0,\ldots,m} |\Delta_h^j K(x_0,\ldots,x_m)| \le \frac{C_K|h|^\delta}{(\sum_{j=1}^m |x_0-x_j|)^{md+\delta}},$$

were Δ_h^j denotes the difference operator in the j-th position. We say Λ is an m-linear *singular integral form* if there exists a singular integral kernel K such that for any $f_0,\ldots,f_m \in \mathcal{S}$ with $\cap_{j=0}^m \operatorname{supp} f_j = \varnothing$ one has

$$\Lambda(f_0,\ldots,f_m) = \int_{(\mathbb{R}^d)^{m+1}} K(x) \prod_{j=0}^m f_j(x_j)\,\mathrm{d}x.$$

When m is understood, we simply say Λ is a singular integral form.

We need the following function spaces in order to define Calderón-Zygmund forms. For $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, define the $\mathrm{BMO}(\mathbb{R}^d)$ norm

$$\|f\|_{\mathrm{BMO}(\mathbb{R}^d)} = \sup_{Q\,\mathrm{cube}} \frac{1}{|Q|}\int_Q |f(x)-f_Q|\,\mathrm{d}x, \quad f_Q = \frac{1}{|Q|}\int_Q f(y)\,\mathrm{d}y.$$

Then, the functions of bounded mean oscillation (BMO) are those with finite $\mathrm{BMO}(\mathbb{R}^d)$ norm. Let $\mathcal{C}_{\mathrm{v}}(\mathbb{R}^d)$ be the space of all continuous functions f on $\mathbb{R}^d$ for which $\lim_{|x|\to\infty} f(x) = 0$. Then, the Banach space of functions with continuous mean oscillation, $\mathrm{CMO}(\mathbb{R}^d)$, is defined to be the closure of $\mathcal{C}_{\mathrm{v}}(\mathbb{R}^d)$ in the norm $\|\cdot\|_{\mathrm{BMO}(\mathbb{R}^d)}$. Furthermore, $\mathrm{BMO}(\mathbb{R}^d)$ becomes a a Banach space upon identifying functions which differ by a constant.

Definition 3.3.2. We say a singular integral form Λ is a *Calderón-Zygmund (CZ) form* if there exists C_{W} such that

$$t^{md}\,|\Lambda(\mathsf{Sy}_z\varphi_0,\ldots,\mathsf{Sy}_z\varphi_m)| \le C_{\mathsf{W}}, \quad \forall \varphi_j \in \mathcal{C}_0^\infty(B(0,1)) \cap \Psi^{1,1}_{(0,1)},$$

and furthermore, there exist $\mathfrak{b}_j \in \mathrm{BMO}(\mathbb{R}^d)$ such that (3.1.2) holds. The rigorous definition of (3.1.2) is as follows. Let $\theta \in \mathcal{C}_0^\infty(\mathbb{R}^d)$ with $\theta = 1$ near the origin. Then, for $t > 0$, set

$\theta_t = \theta(\cdot t)$. We say (3.1.2) holds, and sometimes use the language $T^{*,j}(1,\ldots,1) = \mathfrak{b}_j$, if for all $\varphi \in \mathcal{S}^0(\mathbb{R}^d)$,

$$\lim_{t\to 0} \Lambda^{*,j}(\varphi, \theta_t, \theta_t, \ldots, \theta_t) = \langle \varphi, \mathfrak{b}_j \rangle . \tag{3.3.1}$$

Furthermore, let ϕ be a mother wavelet and ψ^2, ψ^4 be the Schwartz functions from Lemma 3.2.4. For $j = 1, \ldots, m$ define

$$\vec{\psi}_j = \left(\bigotimes_{i=1}^{j-1} \psi^2 \right) \otimes \left(\bigotimes_{i=j}^{m-1} \psi^4 \right) \otimes \phi. \tag{3.3.2}$$

For each $M \geq 1$, define

$$\mathsf{W}_\Lambda^M(\zeta) = \sup_{\mathbf{z}\in \mathbb{B}_M^m(\zeta)} \sup_{\substack{\sigma\in \mathrm{S}_{m+1} \\ j=1,\ldots,m}} \left| \Lambda^\sigma \left((\vec{\psi}_j)_{\mathbf{z}}, \phi_\zeta \right) \right| t^{md}, \quad \zeta = (w,t). \tag{3.3.3}$$

A CZ form Λ is said to be a compact CZ form if for some (all) $M \geq 1$

$$\mathfrak{b}_j \in \mathrm{CMO}(\mathbb{R}^d), \quad \lim_{\zeta\to\infty} \mathsf{W}_\Lambda^M(\zeta) = 0.$$

Finally, we say a CZ form is cancellative if $\mathfrak{b}_j = 0$.

Remark 3.3.3. It is not too difficult to see that if $\mathsf{W}_\Lambda^M \to 0$ for some M then the same holds for each M. Furthermore, since $\mathcal{S}^0(\mathbb{R}^d)$ is dense in the Hardy space $H^1(\mathbb{R}^d)$, which is the dual space of $\mathrm{BMO}(\mathbb{R}^d)$, (3.3.1) should be interpreted as a weak limit, i.e. $T^{*,j}(\theta_t, \ldots, \theta_t)$ converges weakly to $\mathfrak{b}_j$ in $\mathrm{BMO}(\mathbb{R}^d)$.

Now we justify our description of such forms as compact.

Definition 3.3.4. Let T be an m-linear CZO. We say T is a compact CZO if for each $(\vec{p}, r) \in Q$, $T(\mathcal{B}^{\vec{p}})$ is precompact in $L^r(\mathbb{R}^d)$, and at the upper endpoint, $T(\mathcal{B}_0^\infty)$ is precompact in $\mathrm{CMO}(\mathbb{R}^d)$, where

$$\mathcal{B}_0^\infty = \left\{ (f_1, \ldots, f_m) \in \mathcal{B}^{(\infty,\ldots,\infty)} : f_j \text{ compactly supported} \right\} .$$

Remark 3.3.5. The classical definition of compactness of an abstract m-linear operator on quasi-normed spaces [3, 6] is that it maps bounded sets to precompact sets. In Definition 3.3.4, we are imposing this definition of compactness on T acting from $L^{\vec{p}}(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$ for *all* $(\vec{p}, r)$ in Q and at the upper endpoint because our testing conditions allow us to conclude compactness in this full range. It is tempting to only require $T : L^{\vec{p}}(\mathbb{R}^d) \to L^r(\mathbb{R}^d)$ be compact for a single $(\vec{p}, r)$, but the current state of multilinear extrapolation of compactness [6, 40] does not include the endpoints.

Proposition 3.3.6. *Let Λ be an $(m+1)$-linear CZ form such that each $T \in \mathsf{T}_\Lambda$ is a compact CZO. Then, each $\mathfrak{b}_j \in \mathrm{CMO}(\mathbb{R}^d)$, $j = 0, \ldots, m$, and for each $M \geq 1$,*

$$\lim_{\zeta \to \infty} \mathsf{W}_\Lambda^M(\zeta) = 0. \tag{3.3.4}$$

Proof. To prove the first conclusion, since Λ is a CZ form, (3.3.1) holds. Let us fix $j = 0, \ldots, m$. Up to an absolute constant, $(\theta_t, \ldots, \theta_t) \in \mathcal{B}_0^\infty$, so there exists $t_n \to 0$ such that $T^{*,j}(\theta_{t_n}, \ldots, \theta_{t_n})$ converges in $\mathrm{CMO}(\mathbb{R}^d)$. However, the second statement in Remark 3.3.3 requires this limit to coincide with $\mathfrak{b}_j$, and the first part is proved. To prove the second conclusion, let us suppose, towards a contradiction, that there exists $M \geq 1$ such that (3.3.4) does not hold. Therefore we can find $\sigma \in \mathrm{S}_{m+1}$, $\vec{\psi} \in \mathcal{S}(\mathbb{R}^d)^m$, $\varepsilon > 0$ and sequences $\zeta_n = (w_n, t_n) \in Z^d$ and $\mathbf{z}_n \in Z^{md}$ such that

$$\zeta_n \to \infty, \quad \mathbf{z}_n \in \mathbb{B}_M^m(\zeta_n), \quad \left|\Lambda^\sigma\left(\phi_n, \vec{\psi}_n\right)\right| \geq \varepsilon, \quad \phi_n = t_n^{\frac{d}{r}} \phi_{\zeta_n}, \quad \vec{\psi}_n = t_n^{d(m - \frac{1}{r})} \vec{\psi}_{\mathbf{z}_n}, \tag{3.3.5}$$

and $r \in (1, \infty)$. Let us suppose, for simplicity that σ is the identity. If not, then in what follows we would replace $T\vec{\psi}_n$ by $T^{*,\sigma(0)}(\vec{\psi}_n^{\sigma'})$ where $\vec{\psi}_n^{\sigma'}$ is an appropriate permutation of the

m elements of $\vec{\psi}_n$. Since $\{\vec{\psi}_n\}_{n\in\mathbb{N}}$ is a bounded sequence in $L^{\vec{p}}(\mathbb{R}^d)$, for some $\vec{p} \in P$ with $r(\vec{p}) = r$, by the compactness of T we find a subsequence so that $\{T\vec{\psi}_{n_k}\}_{k\in\mathbb{N}}$ is convergent in $L^r(\mathbb{R}^d)$. Finally, it is not hard to check that since $\zeta_n \to \infty$, ϕ_n converges weakly to zero in $L^{\frac{r}{r-1}}(\mathbb{R}^d)$ which, since $r > 1$, is the dual space of $L^r(\mathbb{R}^d)$. Therefore,

$$\Lambda^\sigma(\phi_{n_k}, \vec{\psi}_{n_k}) = \left\langle \phi_{n_k}, T\vec{\psi}_{n_k} \right\rangle \to 0, \text{ as } k \to \infty,$$

which contradicts (3.3.5). Finally, we remark that if $r \in (1, \infty)$ the weak convergence of ϕ_n to zero in $L^{\frac{r}{r-1}}(\mathbb{R}^d)$ follows from the fact that for any $K \geq 1$ and $g \in C_c^\infty(\mathbb{R}^d)$ there holds

$$\lim_{z\to\infty} \left\langle \frac{\mathbf{1}_{KB_z}}{(Kt)^{\frac{d}{r'}}}, g \right\rangle = 0$$

which follows from the inequality

$$\left|\left\langle \frac{\mathbf{1}_{KB_z}}{(Kt)^{\frac{d}{r'}}}, g \right\rangle\right| \lesssim \min\left\{ \|g\|_r, (Kt)^{-\frac{d}{r'}} \|g\|_1, (Kt)^{\frac{d}{r}} \|g\|_\infty \right\}$$

and the fact this inner product vanishes when $KB_z \cap \operatorname{supp}(g) \neq \varnothing$. □

The main step in the representation theorem is given in Proposition **??** below. There we will show that a compact cancellative CZ form enjoys additional decay in the wavelet basis. For a general cancellative CZ form, we recall the following lemma from [21, Lemma 3.3] regarding its decay when applied to a $(m+1)$-tuple of wavelets. When we want to specify the value of the smoothness parameter $\delta > 0$ from Definition 3.3.1, we say Λ is a δ-CZ form.

Lemma 3.3.7. *Let Λ be a cancellative δ-CZ form, $\eta \in (0, \delta)$, and $\psi_j \in \mathcal{C}_0^\infty(B(0,1))$, $j = 1, \ldots, m-1$. Then, there exists $C > 0$ such that for all $\sigma \in \mathrm{S}_{m+1}$, $\zeta \in Z^d$, and $\mathbf{z} \in Z^{md}$ with $\mathbf{z} \geq \zeta$,*

$$\left|\Lambda^\sigma\left(\vec{\psi}_{\mathbf{z}}, \phi_\zeta\right)\right| \leq C\,[\mathbf{z}, \zeta]_\eta, \quad \vec{\psi} = \psi_1 \otimes \cdots \otimes \psi_{m-1} \otimes \phi.$$

Proof. By symmetry, we can assume σ is the identity element and $j = 1$. Furthermore, set $\vec{\psi} = \vec{\psi}_1$ to declutter the notation. For each $n \in \mathbb{N}$, we will construct $\left\{\varepsilon_\zeta^n \in \mathbb{C} : \zeta \in Z^d\right\}$ and $\rho > 0$ with the properties that

$$\sup_{\zeta \in Z^d} |\varepsilon_\zeta^n| \lesssim 1, \quad \lim_{\zeta \to \infty} |\varepsilon_\zeta^n| = 0,$$

$$\left|\Lambda\left(\vec{\psi}_{\mathbf{z}}, \phi_\zeta\right)\right| \lesssim 2^{-\rho n} \left|\varepsilon_\zeta^n\right| [\mathbf{z}, \zeta]_\eta, \quad \mathbf{z} \in \begin{cases} \mathbb{B}_1^m(\zeta) & n = 1, \\ \mathbb{B}_n^m(\zeta) \setminus \mathbb{B}_{n-1}^m(\zeta) & n \geq 2, \end{cases} \quad \mathbf{z} \geq \zeta.$$

Assuming we have such ε_ζ^n, define

$$\varepsilon_\zeta = \sum_{n=1}^{\infty} 2^{-\rho n} \varepsilon_\zeta^n.$$

The first and third properties of ε_ζ in (**??**) are immediate and the second follows by Lebesgue's dominated convergence theorem since each ε_ζ^n approaches zero. Now, to construct ε_ζ^n, let $\delta_j > 0$ such that $\eta < \delta_1 < \delta_2 < \delta$. By Lemma 3.3.7 and the definition of W_Λ^M from (3.3.3), we have for any $\theta \in (0, 1)$,

$$\left|\Lambda\left(\vec{\psi}_{\mathbf{z}}, \phi_\zeta\right)\right| \lesssim \left(t^{-md}\right)^{1-\theta} [\mathbf{z}, \zeta]_{\delta_2}^\theta \min\{1, \mathsf{W}_\Lambda^n(\zeta)\}^{1-\theta}, \quad \mathbf{z} \in \mathbb{B}_n^m(\zeta),\ \mathbf{z} \geq \zeta.$$

We choose θ close enough to 1 that $md\theta + \delta_2\theta - md = \delta_1$. With this specific choice of θ, one can easily verify that $\left(t^{-md}\right)^{1-\theta} [\mathbf{z}, \zeta]_{\delta_2}^\theta = [\mathbf{z}, \zeta]_{\delta_1}$. Setting $\varepsilon_\zeta^1 = \min\{1, \mathsf{W}_\Lambda^1(\zeta)\}^{1-\theta}$ handles the case $n = 1$. For $n \geq 2$, we factor, with $\rho = \delta_1 - \eta$,

$$[\mathbf{z}, \zeta]_{\delta_1} = \left(\frac{t}{\max\{s, |\mathbf{w} - w_0|\}}\right)^\rho [\mathbf{z}, \zeta]_\eta.$$

Therefore, one only needs to verify that for $\mathbf{z} \in \mathbb{B}_n^m(\zeta) \setminus \mathbb{B}_{n-1}^m(\zeta)$, the first factor is comparable to $2^{-\rho n}$. The proof is concluded by setting $\varepsilon_\zeta^n = \min\{1, \mathsf{W}_\Lambda^n(\zeta)\}^{1-\theta}$. □

Proposition 3.3.8. *Every compact cancellative CZ form is a finite sum of compact wavelet forms.*

Proof. We recycle the proof of the representation theorem in [21] relying on Lemma 3.2.4 and Lemma 3.2.5. First expand each f_j using (3.2.1) to obtain

$$\Lambda(f_0, \dots, f_m) = \int_{(Z^d)^{m+1}} \Lambda(\phi_{z_0}, \dots, \phi_{z_m}) \langle f_0, \phi_{z_0} \rangle \, \mathrm{d}\mu(z_0) \dots \langle f_m, \phi_{z_m} \rangle \, \mathrm{d}\mu(z_m).$$

Split the integration region into $m(m+1)$ components defined by

$$\{(z_0, \dots, z_m) \in (Z^d)^{m+1} : z_i \geq z_j \geq z_k, \, i \neq j, k\}, \quad k = 0, \dots, m, \, j = 0, \dots, k-1, k+1, \dots, m.$$

For each j, k and each $i \neq j, k$, we apply Lemma 3.2.4 to $\langle f_i, \phi_{z_i} \rangle \phi_{z_i}$. Furthermore, setting $f^\sigma = \bigotimes_{j=0}^m f_{\sigma(j)}$ and relabelling the variables, we obtain

$$\Lambda(f_0, \dots, f_m) = \sum \int_{\zeta \in Z^d} \int_{\substack{\mathbf{z} \in Z^{md} \\ \mathbf{z} \geq \zeta}} \Lambda^\sigma((\vec{\psi}_\mathrm{e})_\mathbf{z}, \phi_\zeta) \left\langle f^\sigma, (\vec{\psi}_\mathrm{o})_\mathbf{z} \otimes \phi_\zeta \right\rangle \mathrm{d}\mu(\mathbf{z}) \, \mathrm{d}\mu(\zeta),$$

where the sum is taken over all $\sigma \in \mathrm{S}_{m+1}$ and over the combinations $\vec{\psi}_\mathrm{e}$ is of the form (3.3.2) for some $j = 1, \dots, m$ and $\vec{\psi}_\mathrm{o}$ is of the same form but with ψ^2 and ψ^4 replaced by ψ^1 and ψ^3 from Lemma 3.2.4. Each of these summands (of which there are only finitely many depending on m), will now be converted to a compact wavelet form by Proposition **??** and Lemma 3.2.5. Fix now one σ, $\vec{\psi}_\mathrm{e}$, and $\vec{\psi}_\mathrm{o}$. We define

$$\vartheta_\zeta = \int_{\substack{\mathbf{z} \in Z^{md} \\ \mathbf{z} \geq \zeta}} \Lambda^\sigma((\vec{\psi}_\mathrm{e})_\mathbf{z}, \phi_\zeta) \vec{\psi}_\mathrm{o} \, d\mu(\mathbf{z})$$

and the result will be proved if we can show $\vartheta_\zeta = \varepsilon_\zeta \nu_\zeta$ for some $\nu_\zeta \in \Psi_\zeta^{m,\delta;m}$ and ε_ζ approaching zero as $\zeta \to \infty$. Let ε_ζ be the collection provided by Proposition **??**. In particular, the third property in (**??**) guarantees

$$\Lambda^\sigma((\vec{\psi}_\mathrm{e})_\mathbf{z}, \phi_\zeta) = \varepsilon_\zeta H(\mathbf{z}, \zeta), \quad |H(\mathbf{z}, \zeta)| \leq [\mathbf{z}, \zeta]_\eta,$$

whence the proof is concluded by Lemma 3.2.5 and recalling that the m-th component of $\vec{\psi}_{\mathrm{o}}$ is the mother wavelet and thus belongs to $\mathcal{S}^0(\mathbb{R}^d)$. $\square$

3.4 Compact wavelet forms

In this section we will prove the compact analogue of Proposition 3.2.8.

Proposition 3.4.1. *Let U be a compact wavelet form. Then each $T \in \mathsf{T}_U$ is a compact CZO.*

Proof. Let us fix a compact wavelet form U, $T \in \mathsf{T}_U$, and $(\vec{p}, r) \in Q$. $T(f_1, \ldots, f_m)(x)$ is either of the form

$$\int_{Z^d} \varepsilon_z \langle f_1 \otimes \cdots \otimes f_m, \nu_z \rangle \phi_z(x)\, \mathrm{d}\mu(z) \quad \text{or} \int_{Z^d} \varepsilon_z \left\langle f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(m)}, \nu_z(x, \cdot) \otimes \phi_z \right\rangle \mathrm{d}\mu(z) \tag{3.4.1}$$

for some $\sigma \in \mathrm{S}_m$ and $\nu_z \in \Psi_z^{m,\delta;j}$; by $\nu_z(x, \cdot)$ we mean for each x it returns the function $(x_2, \ldots, x_m) \mapsto \nu_z(x, x_2, \ldots, x_m)$. We will only handle the second case, and we will reduce, by symmetry, to the case where σ is the identity. The first case in (3.4.1) is simpler, though in fact they are handled in exactly the same fashion. Let $\rho > 0$ and split the integral defining T over $\mathbb{B}_M$ and $\mathbb{Z}^d \backslash \mathbb{B}_M$ where M is chosen large enough that $|\varepsilon_z| \leq \rho$ for $z \notin \mathbb{B}_M$. Therefore, the operator norm of the second component is, by (3.2.2) controlled by ρ. The proof will be concluded if we can show R_ρ defined by

$$R_\rho \mathbf{f}(x) = \int_{\mathbb{B}_M} \langle f_1 \otimes \cdots \otimes f_m, \nu_z(x, \cdot) \otimes \phi_z \rangle \, \mathrm{d}\mu(z), \quad \mathbf{f} = (f_1, \ldots, f_m),$$

is compact. By the Riesz-Kolmogorov compactness criteria (see e.g. [58] and [75] for the case $0 < r < 1$), we need to prove that

$$\lim_{N\to\infty} \sup_{\mathbf{f}\in\mathcal{B}^{\vec{p}}} \int_{|x|>N} |R_\rho \mathbf{f}(x)|^r \mathrm{d}x = 0, \tag{3.4.2}$$

$$\lim_{h\to 0} \sup_{\mathbf{f}\in\mathcal{B}^{\vec{p}}} \int_{\mathbb{R}^d} |R_\rho \mathbf{f}(x+h) - R_\rho \mathbf{f}(x)|^r \mathrm{d}x = 0. \tag{3.4.3}$$

To this end we will give suitable pointwise bounds on the operator R_ρ and the differences induced by it. Since M is fixed, we will crucially use that $z = (w,t) \in \mathbb{B}_M$ satisfies

$$t \sim 1, \quad |w| \lesssim 1, \quad \chi_z \lesssim \chi, \tag{3.4.4}$$

where $\chi = \chi_{(0,1)}$ while $\sim$ and $\lesssim$ now denote comparability with constants depending on M. Let us now note the preliminary trivial bounds that can be obtained via applying Hölder's inequality and (3.4.4). To this end, introduce, for $j = 1, \ldots, m-1$,

$$\lambda_0 = \frac{d}{r} + \eta, \quad \lambda_j = \frac{d}{p_j'} + \eta, \quad \eta = \frac{1}{m}\left(\frac{d}{p_m'} + \delta\right) > 0.$$

It is easy to check that $md + \delta = \sum_{j=0}^{m-1} \lambda_j$, $\lambda_0 r > d$, and $\lambda_j p_j' > d$, thus for $z \in \mathbb{B}_M$,

$$\begin{aligned} &|\langle f_m, \phi_z\rangle| \le \|f_m\|_{p_m} \|\phi_z\|_{p_m'} \lesssim 1 \\ &|\langle v_z(x,\cdot), f_1 \otimes \ldots \otimes f_{m-1}\rangle| \le \frac{1}{t^d}\chi_z(x)^{\lambda_0} \prod_{j=1}^{m-1} \|f_j\|_{p_j} \left\| \frac{1}{t^d}\chi_z^{\lambda_j} \right\|_{p_j'} \lesssim \chi(x)^{\lambda_0}. \end{aligned} \tag{3.4.5}$$

With these estimates in hand we have that $|R_\rho \mathbf{f}| \lesssim \chi^{\lambda_0}$ and therefore (3.4.2) holds since $\lambda_0 r > d$. In the same way, we have that

$$|\langle v_z(x,\cdot), f_1 \otimes \ldots \otimes f_{m-1}\rangle - \langle v_z(x+h,\cdot), f_1 \otimes \ldots \otimes f_{m-1}\rangle| \lesssim |h|^\delta \chi(x)^{\lambda_0}$$

and therefore $|R_\rho \mathbf{f}(x) - R_\rho \mathbf{f}(x-h)| \lesssim |h|^\delta \chi(x)^{\lambda_0}$ from which (3.4.3) now follows. It remains to handle the endpoint case. We claim it suffices to establish that for each sequence $\{\mathbf{f}_n\}_{n\in\mathbb{N}} \subset$

$\mathcal{B}_0^\infty$, $R_\rho \mathbf{f}_n$ has a convergent subsequence in $\mathrm{CMO}(\mathbb{R}^d)$. Indeed, given such an $\{\mathbf{f}_n\}_{n\in\mathbb{N}}$, by a diagonalization argument, we may extract a subsequence $\{\mathbf{f}_{n_k}\}_{k\in\mathbb{N}}$ such that for each $\rho_n = \frac{1}{n}$, $\{R_{\rho_n}\mathbf{f}_{n_k}\}_{k\in\mathbb{N}}$ is Cauchy in $\mathrm{CMO}(\mathbb{R}^d)$. Then, for any $\epsilon > 0$ pick n large enough that

$$\|T - R_{\rho_n}\|_{\infty,\mathrm{BMO}} < \frac{\epsilon}{3}.$$

Such an n exists by the condition $\varepsilon_z \to 0$ and (3.2.3). Then, pick N large enough that for all $i, k \geq N$, $\|R_{\rho_n}\mathbf{f}_{n_i} - R_{\rho_n}\mathbf{f}_{n_k}\|_{\mathrm{BMO}(\mathbb{R}^d)} < \frac{\epsilon}{3}$. Therefore, by the triangle inequality,

$$\|T\mathbf{f}_{n_k} - T\mathbf{f}_{n_i}\|_{\mathrm{BMO}(\mathbb{R}^d)} < \epsilon$$

whence $\{T\mathbf{f}_{n_k}\}_{k\in\mathbb{N}}$ is Cauchy and has a limit in $\mathrm{CMO}(\mathbb{R}^d)$. Now it remains to show each R_ρ is compact. Applying (3.4.5) with all $p_j = \infty$ implies the same pointwise estimates as above, which implies $R_\rho : \mathcal{B}_0^\infty \to \mathcal{C}_{\mathrm{v}}(\mathbb{R}^d)$ and that the family $\{R_\rho \mathbf{f} : \mathbf{f} \in \mathcal{B}_0^\infty\}$ is equicontinuous. Therefore, by the Arzela-Ascoli theorem and a diagonalization argument, given a sequence $\{\mathbf{f}_n\}_{n\in\mathbb{N}} \in \mathcal{B}_0^\infty$, we can obtain a subsequence such that $\{R_\rho \mathbf{f}_{n_k}\}_{k\in\mathbb{N}}$ is Cauchy in $\|\cdot\|_{L^\infty([-n,n]^d)}$ for each $n \in \mathbb{N}$. However, the pointwise estimate for $R_\rho \mathbf{f}$ shows that given $\epsilon > 0$ we can find n large enough that $|R_\rho \mathbf{f}_{n_k}(x)| < \frac{\epsilon}{3}$ for x outside $[-n, n]^d$. Combining these two facts with the triangle inequality shows that $\{R_\rho \mathbf{f}_{n_k}\}_{k\in\mathbb{N}}$ is Cauchy in $\|\cdot\|_{L^\infty}$ which is a stronger norm than $\|\cdot\|_{\mathrm{BMO}(\mathbb{R}^d)}$. Finally, recalling that $R_\rho \mathbf{f}_{n_k} \in \mathcal{C}_{\mathrm{v}}(\mathbb{R}^d)$ establishes that the limit belongs to $\mathrm{CMO}(\mathbb{R}^d)$.

Remark 3.4.2. We remark that (3.4.4) follows from the fact that if $B_z \subseteq CB_{z'}$ then we have the following elementary pointwise inequality

$$\chi_{z'} \lesssim_C \chi_z.$$

Indeed if $y \in B_z \Rightarrow |y-x| \leq t \Rightarrow |y-x'| \leq |y-x| + |x-x'| \lesssim t'$, (where the last inequality was obtained because $|x-x'| \lesssim t'$) therefore

$$\chi_{z'}(y) \sim 1, \quad \chi_z(y) \sim 1$$

so the desired inequality is trivially true. If $y \notin B_z$ we have that $|y-x| \geq t$ therefore

$$|y-x'| \leq |y-x| + |x-x'| \lesssim |y-x| + t'$$

hence we can estimate

$$\frac{1+\frac{|y-x'|}{t'}}{1+\frac{|y-x|}{t}} \lesssim \frac{1+\frac{|y-x|}{t'}+\frac{|x-x'|}{t}}{\frac{|y-x|}{t}} \leq 1+1+\frac{\frac{|y-x|}{t'}}{\frac{|y-x|}{t}} = 2+\frac{t}{t'} \lesssim_C 1$$

and the claim is proved.

$\square$

3.5 Compactness of paraproducts

In this section we will deal with the compactness of the paraproducts that arise from our representation theorem. Specifically, we will prove that the membership of the symbols in $\mathrm{CMO}(\mathbb{R}^d)$ is sufficient for the compactness of the associated paraproduct by giving an essential norm estimate. Given $(\vec{p}, r) \in Q$, and an m-linear operator T, define the essential norm

$$\|T\|_{\mathrm{ess}(\vec{p},r)} = \inf_{K \text{ compact}} \|T-K\|_{\vec{p},r},$$

and the natural modification at the endpoint which we denote by $\|\cdot\|_{\mathrm{ess}(\infty,\mathrm{BMO})}$ where the operator norm of $T-K$ is measured in $\|\cdot\|_{\infty,\mathrm{BMO}}$.

Definition 3.5.1. Let $\vartheta \in C_0^\infty(B(0,1))$ with $\int \vartheta(x)\,\mathrm{d}x = 1$. For each $z \in Z^d$, set $\vartheta_z = \mathsf{Sy}_z\vartheta$. Given $\mathfrak{b} \in \mathrm{BMO}(\mathbb{R}^d)$, define the $(m+1)$-linear form $\Pi_{\mathfrak{b}}$ by

$$\Pi_{\mathfrak{b}}(f_0, f_1, \ldots, f_m) = \int_{Z^d} \langle \mathfrak{b}, \phi_z \rangle \langle f_0, \phi_z \rangle \prod_{j=1}^{m} \langle f_j, \vartheta_z \rangle \ \mathrm{d}\mu(z).$$

Π_b is called a paraproduct form with symbol $\mathfrak{b}$. Any m-linear operator in T_{Π_b} is called a paraproduct with symbol $\mathfrak{b}$, which we denote by $S_{\mathfrak{b}}$.

In this section we aim to prove the following.

Proposition 3.5.2. *If $\mathfrak{b} \in \mathrm{CMO}(\mathbb{R}^d)$, then any paraproduct $S_{\mathfrak{b}}$ is a compact CZO.*

Before proving this, let us review the standard boundedness theory of paraproducts, analogous to Proposition 3.2.8 for wavelet forms. To do so, it is convenient to view $\Pi_{\mathfrak{b}}$ as an $(m+2)$-linear form, where the extra input function is $\mathfrak{b}$ itself. In fact, in this way $\Pi_{\mathfrak{b}}(f_0, \ldots, f_m) = U(\mathfrak{b}, f_0, \ldots, f_m)$ where U is a canonical $(m+2)$-linear wavelet form with $|\varepsilon_z| \lesssim 1$. Such forms are cancellative in the first and second positions, for which a slight strengthening of (3.2.2) and (3.2.3) holds:

$$\begin{gathered} T \in \mathsf{T}_{V_{\mathfrak{b}}}, \quad V_{\mathfrak{b}}(f_0, \ldots, f_m) = U(\mathfrak{b}, f_0, \ldots, f_m), \\ \|T\|_{\vec{p},r} \le C_{\vec{p},r} \|\mathfrak{b}\|_{\mathrm{BMO}(\mathbb{R}^d)}, \quad \|T\|_{\infty,\mathrm{BMO}} \le C_\infty \|\mathfrak{b}\|_{\mathrm{BMO}(\mathbb{R}^d)}, \quad (\vec{p}, r) \in Q. \end{gathered} \tag{3.5.1}$$

In particular, (3.5.1) applies to $T = S_{\mathfrak{b}}$. A proof of (3.5.1) is omitted since it follows from standard considerations; see e.g. the proofs and comments following Propositions 2.5 and 2.7 in [25].

Now, to give a description of $\mathrm{CMO}(\mathbb{R}^d)$ which is more amenable to $S_{\mathfrak{b}}$, let us introduce an orthonormal wavelet system $\{\psi_I\}_{I\in\mathcal{D}}$. Here $\mathcal{D}$ is a dyadic grid on $\mathbb{R}^d$ and for each $I \in \mathcal{D}$, set $\zeta(I) = (c(I), \ell(I)) \in Z^d$, where $c(I)$ is center of the cube I and $\ell(I)$ the side length. Then $\psi_I = \mathsf{Sy}_{\zeta(I)}\psi$ for a specific $\psi \in C\Psi^{1,1;1}_{(0,1)}$. We have kept ψ_I to be L^1-normalized, so the reproducing formula is

$$f = \sum_{I \in \mathcal{D}} |I| \langle f, \psi_I \rangle \psi_I.$$

Now introduce the family of orthogonal projections for $M \geq 1$,

$$P_M f = \sum_{I \in \mathcal{D}_M} |I| \langle f, \psi_I \rangle \psi_I, \quad \mathcal{D}_M := \{I \in \mathcal{D} : \zeta(I) \in \mathbb{B}_M\}, \quad P_M^\perp = \mathrm{Id} - P_M.$$

From [76, Lemma 2.20], an equivalent characterization of $\mathfrak{b} \in \mathrm{CMO}(\mathbb{R}^d)$ is that

$$\lim_{M \to \infty} \left\| P_M^\perp \mathfrak{b} \right\|_{\mathrm{BMO}(\mathbb{R}^d)} = 0.$$

Therefore, Proposition 3.5.2 will be a consequence of the following essential norm estimate.

Proposition 3.5.3. *Let $S_\mathfrak{b}$ be a paraproduct with symbol $\mathfrak{b} \in \mathrm{BMO}(\mathbb{R}^d)$. Then for each $(\vec{p}, r) \in Q \cup \{(\infty, \mathrm{BMO})\}$,*

$$\|S_\mathfrak{b}\|_{\mathrm{ess}(\vec{p},r)} \lesssim \liminf_{M \to \infty} \left\| P_M^\perp \mathfrak{b} \right\|_{\mathrm{BMO}(\mathbb{R}^d)}. \tag{3.5.2}$$

Proof. Let $M \geq 1$ large, and perform the splitting $S_\mathfrak{b} = R_M + T_M$ where

$$T_M(f_1, \dots, f_m) = \int_{Z^d \setminus \mathbb{B}_{100M}} \langle \mathfrak{b}, \phi_z \rangle \prod_{j=1}^m \langle f_j, \vartheta_z \rangle \phi_z \, \mathrm{d}\mu(z),$$

and R_M is same but the integration is taken over $\mathbb{B}_{100M}$. The operator R_M is clearly compact by repeating the discussion made in §3.4 to show that R_ρ there was compact. Therefore the essential norm of $S_\mathfrak{b}$ is controlled by the operator norm of T_M. Now, for $z \notin \mathbb{B}_{100M}$, we can calculate the pairings appearing in the above equation by expanding $\mathfrak{b}$ with the aid of the wavelet basis $\{\psi_I\}_{I\in\mathcal{D}}$:

$$\langle \mathfrak{b}, \phi_z\rangle = \langle P_M\mathfrak{b}, \phi_z\rangle + \langle P_M^{\perp}\mathfrak{b}, \phi_z\rangle.$$

For the first term we will employ the linear wavelet averaging process from Lemma 3.2.5. We expand

$$\langle P_M\mathfrak{b}, \phi_z\rangle = \sum_{I\in\mathcal{D}_M} |I|\langle \mathfrak{b}, \psi_I\rangle\langle \psi_I, \phi_z\rangle = \left\langle \mathfrak{b}, \sum_{I\in\mathcal{D}_M} |I|\langle \psi_I^1, \phi_z\rangle\psi_I^1\right\rangle.$$

Since $\psi_I \in \Psi^{1,1;1}_{\zeta(I)}$ and $\phi_z \in \Psi^{1,1;1}_z$, one can compute $|\langle \psi_I, \phi_z\rangle| \lesssim [z, \zeta(I)]_{\frac{1}{2}}$ (see e.g. [25, Lemma 2.3] or [31, Appendix, Lemmata 2 and 4]). Furthermore, since $\zeta(I) \in \mathbb{B}_M$ and $z \notin \mathbb{B}_{100M}$, $[z, \zeta]_{\frac{1}{2}} \lesssim M^{-\frac{1}{4}}[z, \zeta]_{\frac{1}{4}}$. Now, to apply Lemma 3.2.5, rewrite

$$\sum_{I\in\mathcal{D}_M} |I|\langle \psi_I, \phi_z\rangle\psi_I = \int_{Z^d} H(\zeta, z)\tilde{\psi}_\zeta \, \mathrm{d}\mu(\zeta),$$

where for each $I \in \mathcal{D}_M$ and $\zeta = (w, t) \in I \times (\frac{\ell(I)}{2}, \ell(I)]$, we define

$$\tilde{\psi}_\zeta = \psi_I, \quad H(\zeta, z) = \frac{|I|}{\mu(I \times (\frac{\ell(I)}{2}, \ell(I)])}\langle \psi_I^1, \phi_z\rangle,$$

and $H(\zeta, z) = 0$ if $\zeta \notin \cup_{I\in\mathcal{D}_M} I \times (\frac{\ell(I)}{2}, \ell(I)]$. Since $|H(\zeta, z)| \lesssim M^{-\frac{1}{4}}[z, \zeta]_{\frac{1}{4}}$, Lemma 3.2.5 provides a universal constant C and $\lambda_z \in C\Psi_z^{1,\frac{1}{8};1}$ such that

$$T_M(f_1, \ldots, f_m) = \int_{Z^d\setminus\mathbb{B}_{100M}} \left(M^{-\frac{1}{4}}\left\langle \mathfrak{b}, \lambda_z\right\rangle + \left\langle P_M^{\perp}\mathfrak{b}, \phi_z\right\rangle\right) \prod_{j=1}^m \left\langle f_j, \vartheta_z\right\rangle \phi_z \, d\mu(z).$$

Therefore, (3.5.2) follows by the triangle inequality and (3.5.1). □

3.6 Proofs of Theorems C and D

Let us now put together the pieces from the previous sections to prove the following compact $T(1)$ wavelet representation theorem.

Proof of Theorem C. To show A. implies B. we isolate the cancellative part of Λ, namely

$$\Lambda_{\mathsf{c}} = \Lambda - \Pi_\Lambda, \quad \Pi_\Lambda = \sum_{j=0}^{m} \Pi_{\mathfrak{b}_j}^{*,j}.$$

Λ_{c} is definitely a CZ form, and to verify that it is cancellative, simply note that since $\langle \vartheta, 1\rangle = 1$, by the reproducing formula (3.2.1), for $\varphi \in \mathcal{S}^0(\mathbb{R}^d)$,

$$\Pi_{\mathfrak{b}}(\varphi, 1, \dots, 1) = \int_{Z^d} \langle \mathfrak{b}, \phi_z\rangle \langle \phi_z, \varphi\rangle \, \mathrm{d}\mu(z) = \langle \mathfrak{b}, \varphi\rangle ,$$

and since ϕ is cancellative, $\Pi_{\mathfrak{b}}^{*,j}(\varphi, 1, \dots, 1) = 0$ for $j = 1, \dots, m$. We want to apply Proposition 3.3.8 to Λ_{c} so we must establish that it is a compact CZ form. Since $\mathfrak{b}_j \in \mathrm{CMO}(\mathbb{R}^d)$, each $S \in \mathsf{T}_{\Pi_\Lambda}$ is compact by Proposition 3.5.2, so by Proposition 3.3.6, Π_Λ is a compact CZ form. Since we also know that Λ is a compact CZ form, Λ_{c} must indeed be compact, and B. follows by applying Proposition 3.3.8 to Λ_{c}. B. implies C. is a consequence of Propositions 3.4.1 and 3.5.2. Finally, C. implies A. is the content of Proposition 3.3.6. $\square$

4. Bloom's inequality via the wavelet representation theorem

4.1 Introduction

The commutator of an operator T with the multiplication operator, given by a symbol b, is defined as $[b,T]f \coloneqq bTf - T(bf)$. Coifman-Rochberg-Weiss in [11] characterized BMO in terms of the commutators of Riesz transforms. Subsequently, Bloom proved, in [4], that the commutator of the Hilbert transform,

$$Hf \coloneqq \text{p.v. } \frac{1}{\pi}\int \frac{f(y)}{x-y}\mathrm{d}y$$

is bounded from $L^p(w)$ to $L^p(\sigma)$, with $w, \sigma \in A_p$, if and only if $b \in \mathrm{BMO}(\nu)$ namely if

$$\|b\|_{\mathrm{BMO}(\nu)} \coloneqq \sup_Q \left(\frac{\int_Q |b - \langle b \rangle_Q| \mathrm{d}x}{\nu(Q)} \right) < \infty$$

where $\nu = \left(\frac{w}{\sigma}\right)^{\frac{1}{p}}$. Holmes, Lacey and Wick in [38] proved the upper bound for general Calderón-Zygmund operators and characterized $\mathrm{BMO}(\nu)$ in terms of the boundedness of the commutators with Riesz transforms. Except for the characterization of function spaces, commutator theorems imply the so-called div-curl lemmata and weak factorization results for H^1_ν. Finally, off-diagonal results for commutators have applications in characterizing the norm of certain function spaces such as BMO, the homogeneous Hölder space $\dot{C}^{0,\alpha}$ and $\dot{L}^r$ by dualizing against functions in the image of the Jacobian [43].

The foundational tool used in [37] was the representation theorem of Hytönen [42] to analyze a Calderón-Zygmund operator as a rapidly decaying superposition of dyadic shifts, with the most elementary example being the martingale transform, of arbitrary complexity which was used to give an affirmative answer to the A_2 conjecture. Recently, Di Plinio, Wick and Williams in [25] devised a wavelet representation leveraging the fact that a Calderón-Zygmund operator applied to a wavelet is a rougher wavepacket with smoothness and localization reflecting the kernel estimates and smoothness of the operator. One of the advantages of their approach is that the representation formula only consists of a single complexity zero cancellative operator, a single paraproduct, and a single adjoint paraproduct.

4.1.1 Wavelet coefficients , the intrinsic square function and averaging lemmata

We will need the notation for the intrinsic wavelet coefficient for $z \in Z^d$, and its cancellative counterpart, namely we define

$$\Psi_z^\delta f := \sup_{\phi \in \Psi_z^\delta} |\langle f, \phi \rangle|, \quad \Psi_z^{\delta;0} f := \sup_{\phi \in \Psi_z^{\delta;0}} |\langle f, \phi \rangle|.$$

Sometimes, given a cube Q centered at $c_Q \in \mathbb{R}^d$ with sidelength ℓ_Q we will use the notation $\Psi_Q^\delta f$ instead of the notation $\Psi_{(c_Q,\ell_Q)}^\delta f$ and likewise for the cancellative intrinsic wavelet coefficient. Furthermore, for $\delta \in (0,1)$ we introduce the intrinsic square function of smoothness δ

$$S_\delta f = \left(\int_0^\infty \left(\Psi_z^{\delta;0} f \right)^2 \frac{\mathrm{d}t}{t} \right)^{\frac{1}{2}}.$$

By [25, Proposition 2.6] we have the $L^p(w, \mathbb{R}^d),\ p > 1$ boundedness of S_δ for $w \in A_p$.

To end this section we will give a wavelet averaging lemma. This lemma and its successor is a slight generalization of [25, Lemma 3.2]. In vague terms, this lemma asserts that if one averages elements of a wavelet class with respect to the measure of the form $a(z,z')\mathrm{d}\mu(z)$, when $a(z,z')$ decays in a certain way, the output is a slightly rougher wavelet.

Lemma 4.1.2. *Let $\varphi_z \in \Psi_z^\delta$ and let $u(z,z') : Z^d \times Z^d \to \mathbb{C}$ a Borel measurable function with $|u(z,z')| \leq [z,z']_\delta$ then for all $0 < \eta < \delta$ we have the membership*

$$\psi_z := \int_Z u(z,z')\varphi_{z'}\mathrm{d}\mu(z') = \int_{\mathbb{R}^d}\int_0^\infty u((x,t),(x+at,\beta t))t^d\varphi_{(x+at,\beta t)}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a \in C_\eta\Psi_z^\eta.$$

In particular, we have the following memberships

$$\psi_z^{\mathsf{n}} = \int_{a\in\mathbb{R}^d}\int_0^1 u((x,t),(x+at,\beta t))t^d\varphi_{(x+at,\beta t)}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a \in C_\eta\Psi_z^\eta$$
$$\psi_z^{\mathsf{f}} = \int_{a\in\mathbb{R}^d}\int_1^\infty u((x,t),(x+at,\beta t))t^d\varphi_{(x+at,\beta t)}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a \in C_\eta\Psi_z^\eta.$$

Proposition 4.1.3. *Let φ_z have the property that $|\varphi_z| \leq \frac{1}{t^d}\chi_z^{-d-\delta}$ and $\lambda \in (0,d)$. Finally, let $u(z,z') : Z^d \times Z^d \to \mathbb{C}$ a Borel measurable function satisfying the inequality*

$$|u(z,z')| \leq \mathbf{1}_{\mathbb{F}}(z,z')\frac{\min\{t,t'\}^\delta}{|x-x'|^{d+\delta-\lambda}}$$

where the region $\mathbb{F} \subset Z^d \times Z^d$ is defined by

$$\mathbb{F} := \left\{(z,z') \in Z^d \times Z^d : |x-x'| \gtrsim \max\{t,t'\}\right\}.$$

Then

$$\psi_z^{\mathsf{n}} = \int_{a\in\mathbb{R}^d}\int_0^1 u((x,t),(x+at,\beta t))t^d\varphi_{(x+at,\beta t)}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a$$
$$\psi_z^{\mathsf{f}} = \int_{a\in\mathbb{R}^d}\int_1^\infty u((x,t),(x+at,\beta t))t^d\varphi_{(x+at,\beta t)}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a$$

satisfy the following estimates

$$\mathsf{i} \in \{\mathsf{n},\mathsf{f}\} \Longrightarrow \left|\psi_z^{\mathsf{i}}(y)\right| \lesssim_\eta t^\lambda\frac{1}{t^d}\chi_z^{-d-\eta-\lambda}, \quad \eta \in (0,\delta).$$

4.2 Paraproduct decomposition and two weight estimates

The treatment of commutators of singular integral operators via representation theorems has been based on the exploitation of the formula of the product of two functions as a sum of paraproducts and their adjoints. We start with $\phi \in \mathcal{S}(\mathbb{R}^d)$ with the properties

$$\operatorname{supp}(\widehat{\phi}) \subset \left\{x \in \mathbb{R}^d : |x| < 2\right\} \text{ and } \widehat{\phi} = 1 \text{ on } \left\{x \in \mathbb{R}^d : |x| < 1\right\}.$$

Hence following the approach in [33, Section 1.2.2] we may write

$$\begin{aligned} fg &= \int_0^\infty f * \psi_t \, g * \phi_t \frac{\mathrm{d}t}{t} + \int_0^\infty f * \phi_t \, g * \psi_t \frac{\mathrm{d}t}{t} = \mathcal{P}(f,g) + \mathcal{P}(g,f) \\ \mathcal{P}(f,g) &= \int_0^\infty f * \phi_t \, g * \psi_t \frac{\mathrm{d}t}{t}, \quad \psi(x) = -\sum_{|a|=1} \partial^a (x^a \phi(x)). \end{aligned} \tag{4.2.1}$$

It is clear that $\operatorname{supp}(\widehat{\psi}) \subset \left\{x \in \mathbb{R}^d : 1 < |x| < 2\right\}$.

Proceeding in the same manner as in [33] we further decompose the adjoint form to $\mathcal{P}$

$$\langle \mathcal{P}(f,g), h\rangle = \langle \mathcal{P}_1(f,g), h\rangle + \langle \mathcal{P}_2(f,g), h\rangle$$

by introducing the decomposition $\phi = \phi^{(1)} + \psi^{(1)}$ and $\phi^{(1)}, \psi^{(1)}, \phi^{(3)}, \psi^{(3)} \in \mathcal{S}(\mathbb{R}^d)$ with the properties

$$\begin{aligned} &\operatorname{supp}(\widehat{\phi^{(1)}}) \subset \left\{\xi \in \mathbb{R}^d : |\xi| < \frac{1}{2}\right\}, \quad \operatorname{supp}\left(\widehat{\psi^{(1)}}\right) \subset \left\{\xi \in \mathbb{R}^d : \frac{1}{4} < |\xi| < 2\right\} \\ &\widehat{\psi^{(3)}}(0) = 0, \quad \widehat{\psi^{(3)}} = 1 \text{ on } \left(\operatorname{supp}\widehat{\phi^{(1)}} + \operatorname{supp}\widehat{\psi}\right) \subset \left\{\xi : \frac{1}{2} < |\xi| < 4\right\} \end{aligned}$$

and so

$$
\begin{aligned}
\langle \mathcal{P}(f,g),h\rangle &= \int_{\mathbb{R}^d}\int_0^\infty \left(\int_{\mathbb{R}^d}\widehat{f}(\xi_1)\widehat{\phi}(t\xi_1)e^{2\pi i x\xi_1}\mathrm{d}\xi_1\right)\left(\int_{\mathbb{R}^d}\widehat{g}(\xi_2)\widehat{\psi}(t\xi_2)e^{2\pi i x\xi_2}\mathrm{d}\xi_2\right)\overline{h}(x)\mathrm{d}x\\
&= \int_{\mathbb{R}^d}\int_0^\infty f*\phi_t^{(1)}(x)\; g*\psi_t(x)\; \overline{h}*\psi_t^{(3)}(x)\frac{\mathrm{d}t}{t}\mathrm{d}x\\
&+ \int_{\mathbb{R}^d}\int_0^\infty f*\psi_t^{(1)}(x)\; g*\psi_t(x)\; \overline{h}*\phi_t^{(3)}(x)\frac{\mathrm{d}t}{t}\mathrm{d}x\\
&\coloneqq \langle \mathcal{P}_1(f,g),h\rangle + \langle \mathcal{P}_2(f,g),h\rangle.
\end{aligned}
$$

Where

$$
\begin{aligned}
\mathcal{P}_1(f,g)(y) &\coloneqq \int_{\mathbb{R}^d}\int_0^\infty f*\phi_t^{(1)}(x)\; g*\psi_t(x)\; \psi_t^{(3)}(x-y)\frac{\mathrm{d}t}{t}\mathrm{d}x,\\
\mathcal{P}_2(f,g)(y) &\coloneqq \int_{\mathbb{R}^d}\int_0^\infty f*\psi_t^{(1)}(x)\; g*\psi_t(x)\; \phi_t^{(3)}(x-y)\frac{\mathrm{d}t}{t}\mathrm{d}x.
\end{aligned}
$$

In the proposition below we give two weight estimates for paraproducts that will be used later on to estimate the main term of the commutator. In the propositions that follow S_2 and S_3 are going to denote the permutation groups of 2 and 3 elements respectively.

Proposition 4.2.1. *Let $p>1$ $w,\sigma\in A_p$ and $\nu=\left(\frac{w}{\sigma}\right)^{\frac{1}{p}}$ then we have that*

$$
|\langle \mathcal{P}_1(f_1,f_2),f_3\rangle| \lesssim \begin{cases} \min_{\sigma\in S_2}\|f_3\|_{BMO_\nu}\left\|f_{\sigma(1)}\right\|_{L^p(w,\mathbb{R}^d)}\left\|f_{\sigma(2)}\right\|_{L^{p'}(\sigma^{1-p'},\mathbb{R}^d)}\\ \min_{\sigma\in S_2}\|f_2\|_{BMO_\nu}\left\|f_{\sigma(1)}\right\|_{L^p(w,\mathbb{R}^d)}\left\|f_{\sigma(3)}\right\|_{L^{p'}(\sigma^{1-p'},\mathbb{R}^d)}\end{cases}
$$

$$
|\langle \mathcal{P}_2(f_1,f_2),f_3\rangle| \lesssim \begin{cases} \min_{\sigma\in S_2}\|f_1\|_{BMO_\nu}\left\|f_{\sigma(2)}\right\|_{L^p(w,\mathbb{R}^d)}\left\|f_{\sigma(3)}\right\|_{L^{p'}(\sigma^{1-p'},\mathbb{R}^d)}\\ \min_{\sigma\in S_2}\|f_2\|_{BMO_\nu}\left\|f_{\sigma(1)}\right\|_{L^p(w,\mathbb{R}^d)}\left\|f_{\sigma(3)}\right\|_{L^{p'}(\sigma^{1-p'},\mathbb{R}^d)}\end{cases}
$$

$$
|\langle \mathcal{P}(f_1,f_2),f_3\rangle| \lesssim \min_{\sigma\in S_2}\|f_2\|_{BMO_\nu}\left\|f_{\sigma(1)}\right\|_{L^p(w,\mathbb{R}^d)}\left\|f_{\sigma(3)}\right\|_{L^{p'}(\sigma^{1-p'},\mathbb{R}^d)}.
$$

Proof. The proposition is well known in the literature and we will only prove the first implication as the rest have an similar treatment, however, for completeness purposes we include a proof. Indeed by $H^1_w(\mathbb{R}^d)$-$\mathrm{BMO}_w(\mathbb{R}^d)$ duality, see for example [27, Theorem 5.2, Theorem 5.5] we have that

$$|\langle \mathcal{P}_1(f_1,f_2),f_3\rangle| \leq \|f_3\|_{\mathrm{BMO}_w}\|\mathcal{P}_1(f_1,f_2)\|_{H^1_w} \lesssim \|f_3\|_{\mathrm{BMO}_w}\|S_\zeta(\mathcal{P}_1(f_1,f_2))\|_{L^1(w)}.$$

Where S_ζ denotes the square function with respect to a wavelet system $\left\{\zeta^\varepsilon_Q\right\}_{\varepsilon=1,\dots,2^d-1,Q\in\mathcal{D}}$ of order $10d$. We have for $\varepsilon\in\left\{1,\dots,2^d-1\right\}$ and $Q\in\mathcal{D}$

$$\langle \mathcal{P}_1(f_1,f_2),\zeta^\varepsilon_Q\rangle = \int_{\mathbb{R}^d}\int_0^\infty f_1*\phi^{(1)}_t f_2*\psi_t(x)\overline{\zeta^\epsilon_Q}*\psi^{(3)}_t = \int_{Z^d}\langle f_1,\phi^{(1)}_z\rangle\langle f_2,\psi_z\rangle\langle\overline{\psi^{(3)}_z},\zeta^\epsilon_Q\rangle \mathrm{d}\mu(z)$$

$$\rho_z = \overline{\widetilde{\mathsf{Sy}_z\rho}},\quad \rho\in\left\{\phi^{(1)},\psi,\overline{\psi^{(3)}}\right\},\quad z=(x,t).$$

From [25, Lemma 2.3] we have the estimate

$$\left|\langle\zeta_Q,\psi^{(3)}_z\rangle\right| \lesssim [(c_Q,\ell_Q),z]_\eta$$

for any $\eta\in(0,10d)$. In addition to that for any $q>1$

$$|\langle f,\phi^{(1)}_z\rangle| \lesssim \inf_{B_z} M_q(f) \lesssim \left(\frac{\max\left\{t,\ell_Q,|x-x'|\right\}}{\min\left\{t,\ell_Q\right\}}\right)^{\frac{d}{q}}\inf_Q M_q(f).$$

Henceforth we have that, for η close to $10d$ that

$$[(c_Q,\ell_Q),z]_\eta\left(\frac{\max\left\{t,\ell_Q,|x-x'|\right\}}{\min\left\{t,\ell_Q\right\}}\right)^{\frac{d}{q}} = [(c_Q,\ell_Q),z]_{\eta-\frac{d}{q}}.$$

Therefore using wavelet averaging lemma 4.1.2 we have that

$$\int_{Z^d}\langle f_1,\phi^{(1)}_z\rangle\langle f_2,\psi_z\rangle\langle\zeta_Q,\psi^{(3)}_z\rangle\mathrm{d}\mu(z) = \inf_Q M_q(f_1)\langle f_2,v_Q\rangle,\quad v_Q\in C\Psi^{\delta,0}_Q$$

for some absolute constant C. Therefore

$$S_\zeta(\mathcal{P}_1(f_1,f_2)) \lesssim M_q(f_1)S(f_2)$$

where S is the intrinsic square function. From this we learn that

$$\|S_\zeta(\mathcal{P}_1(f_1,f_2))\|_{L^1(w)} \lesssim \|M_q(f_1)\|_{L^p(\mu)} \|S(f_2)\|_{L^{p'}(\lambda^{1-p'})} \lesssim \|f_1\|_{L^p(\mu)} \|f_2\|_{L^{p'}(\lambda^{1-p'})}$$

provided that q has been chosen close to 1. $\square$

In addition to the diagonal estimates we obtained in the previous proposition, standard techniques and the usage of the intrinsic square function allows us to get off-diagonal estimates when the symbol belongs in $\dot{L}^r(\mathbb{R}^d)$.

Proposition 4.2.2. *Let $1<q<p<\infty$, r with the property that $\frac{1}{q}=\frac{1}{r}+\frac{1}{p}$ and a weight with $w^p\in A_p, w^q\in A_q$ Then the following estimates hold*

$$|\langle \mathcal{P}_1(f_1,f_2),f_3\rangle| \lesssim \begin{cases} \min\limits_{\sigma\in S_2} \|f_3\|_{\dot{L}^r(\mathbb{R}^d)} \left\|f_{\sigma(1)}\right\|_{L^p(w^p,\mathbb{R}^d)} \left\|f_{\sigma(2)}\right\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \\ \min\limits_{\sigma\in S_2} \|f_2\|_{\dot{L}^r(\mathbb{R}^d)} \left\|f_{\sigma(1)}\right\|_{w^p,L^p(\mathbb{R}^d)} \left\|f_{\sigma(3)}\right\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \end{cases}$$

$$|\langle \mathcal{P}_2(f_1,f_2),f_3\rangle| \lesssim \begin{cases} \min\limits_{\sigma\in S_2} \|f_2\|_{\dot{L}^r(\mathbb{R}^d)} \left\|f_{\sigma(1)}\right\|_{L^p(w^p,\mathbb{R}^d)} \left\|f_{\sigma(3)}\right\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \\ \min\limits_{\sigma\in S_2} \|f_1\|_{\dot{L}^r(\mathbb{R}^d)} \left\|f_{\sigma(2)}\right\|_{L^p(w^p,\mathbb{R}^d)} \left\|f_{\sigma(3)}\right\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \end{cases}$$

$$|\langle \mathcal{P}(f_1,f_2),f_3\rangle| \lesssim \min_{\sigma\in S_2} \|f_2\|_{\dot{L}^r(\mathbb{R}^d)} \left\|f_{\sigma(1)}\right\|_{L^p(w^p,\mathbb{R}^d)} \left\|f_{\sigma(3)}\right\|_{L^{q'}(w^{-q'},\mathbb{R}^d)}.$$

Proof. We only prove the first estimate where $\sigma(1)=1,\sigma(2)=2$ and $f_3\in\dot{L}^r(\mathbb{R}^d)$ we choose c with the property $f-c\in\dot{L}^r(\mathbb{R}^d)$ so by the cancellation properties of $\psi_t^{(3)}$ we can rewrite

$$\langle \mathcal{P}_1(f_1,f_2),f_3\rangle = \langle \mathcal{P}_1(f_1,f_2),f_3-c\rangle$$

and therefore

$$\begin{aligned} |\langle \mathcal{P}_1(f_1,f_2),f_3-c\rangle| &\lesssim \int_{\mathbb{R}^d} \mathrm{M}(f_1)S(f_2)S(f_3-c) \\ &\le \|\mathrm{M}(f_1)\|_{L^p(w^p,\mathbb{R}^d)} \|S(f_2)\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \|S(f_3-c)\|_{L^r(\mathbb{R}^d)} \\ &\lesssim \|f_1\|_{L^p(w^p,\mathbb{R}^d)} \|f_2\|_{L^{q'}(w^{-q'},\mathbb{R}^d)} \|f_3-c\|_{L^r(\mathbb{R}^d)} \end{aligned}$$

which completes the proof. $\square$

With the aid of (4.2.1) we can write

$$\begin{aligned}
&[b,T]f = T(bf) - bTf = \mathcal{M}\mathcal{T}_b(f) + \mathcal{R}_b(f) \\
&\mathcal{M}_{T,b}(f) := T(\mathcal{P}(f,b)) + T(\mathcal{P}_2(b,f)) - \mathcal{P}_2(b,Tf) - \mathcal{P}(Tf,b) \qquad (4.2.2)\\
&\mathcal{R}_{T,b}(f) := T(\mathcal{P}_1(b,f) - \mathcal{P}_1(b,Tf)).
\end{aligned}$$

We will separately prove estimates for $\mathcal{R}_b$ and $\mathcal{M}\mathcal{T}_b(f)$. Immediately from Corollary 4.2.1, combined with the boundedness of Calderón-Zygmund operators on weighted L^p spaces associated to A_p weights, is the following.

Corollary H.1. *Let $p > 1, w, \sigma \in A_p$ and $\nu = \left(\frac{w}{\sigma}\right)^{\frac{1}{p}}$ then we have the estimate*

$$\|\mathcal{M}_{T,b}(f)\|_{L^p(w)} \lesssim \|b\|_{BMO_\nu} \|f\|_{L^p(\sigma)} .$$

In addition to that, if $w^p \in A_p$ and $w^q \in A_q$ and $\frac{1}{q} = \frac{1}{r} + \frac{1}{p}$ then

$$\|\mathcal{M}_{T,b}(f)\|_{L^p(w^p,\mathbb{R}^d)} \lesssim \|b\|_{\dot{L}^r(\mathbb{R}^d)} \|f\|_{L^q(w^q,\mathbb{R}^d)} .$$

4.3 One weight estimates

In this section we will handle the dual form to the remainder term $\mathcal{R}_b$. It is organized as follows. First, we will calculate the remainder term and rewrite it in a form that allows us to use wavelet averaging. Next, we will give certain estimates in the coefficients that appear that allow us to perform this averaging.

Proposition 4.3.1. *Let T be a δ-cancellative Calderón-Zygmund operator. Then for each $\eta \in (0,\delta)$ there exist $u_z, v_z \in C\Psi_z^{\eta;0}$, for some absolute constant C, with the property that*

$$\langle \mathcal{R}_{T,b}(f), g\rangle = \int_{Z^d\times Z^d} \langle b, \phi_{z'}^{(1)} - \phi_z^{(1)}\rangle \langle f, \psi_{z'}\rangle \langle \overline{\psi_{z'}^{(3)}}, v_z\rangle \langle u_z, g\rangle \mathrm{d}\mu(z)\mathrm{d}\mu(z')$$
$$- \int_{Z^d\times Z^d} \langle b, \phi_z^{(1)} - \phi_{z'}^{(1)}\rangle \langle f, v_{z'}\rangle \langle u_{z'}, \psi_z\rangle \langle \overline{\psi_z^{(3)}}, g\rangle \mathrm{d}\mu(z)\mathrm{d}\mu(z').$$

Where $\rho_w = \overline{\widetilde{\mathsf{Sy}_w}\rho}, \quad \rho \in \left\{\phi^{(1)}, \psi, \overline{\psi^{(3)}}\right\}, \quad w = (y,s).$

Proof. Using [25, Theorem A] for T we learn that for each $\eta \in (0,\eta)$ there exist $u_z, v_z \in C\Psi_z^{\eta;0}$

$$T(f) = \int_{Z^d} \langle f, v_z\rangle u_z \mathrm{d}\mu(z).$$

Coupling this information with the fact that

$$f = \int_0^\infty f * \psi_t \,\frac{\mathrm{d}t}{t}$$

and subtracting $\phi_z^{(1)}$ and $\phi_{z'}^{(1)}$ inside the pairings involving b in the definitions of the adjoint forms to $T(\mathcal{P}_1(b,f))$ and $\mathcal{P}_1(b,Tf)$ respectively we obtain the desired formula after performing a routine calculation.

□

The next lemma allows us to estimate the pairings that involve b that appear in Proposition 4.3.1 by $\|b\|_{\mathrm{BMO}}$ conceding only a logarithm factor of the ratio of the distance to the smallest scale and the ratio of the scales. A variant when the function belongs in the homogeneous Hölder space is obtained. Before stating the lemma we will introduce the following quantity that will help us declutter notation

$$d(z,z')_{\mathrm{BMO}} = \begin{cases} \max\left\{1, \log\left(\frac{\max\{t,t'\}}{\min\{t,t'\}}\right)\right\} & \text{if } \frac{|x-x'|}{\max\{t,t'\}} \leq 1 \\ \max\left\{1, \log\left(\frac{|x-x'|}{\min\{t,t'\}}\right)\right\} & \text{if else.} \end{cases}$$

Lemma 4.3.2. *Let $\zeta \in \mathcal{S}(\mathbb{R}^d)$ with the property that $\int_{\mathbb{R}^d} \zeta = 1$. Then the following estimate holds*

$$|\langle b, \mathsf{Sy}_z\zeta - \mathsf{Sy}_{z'}\zeta\rangle| \lesssim d(z,z')_{\mathrm{BMO}} \max\left\{\inf_{B_z} \mathrm{M}_\# b, \inf_{B_{z'}} \mathrm{M}_\# b\right\}.$$

In addition to that,

$$|\langle b, \mathsf{Sy}_z\zeta - \mathsf{Sy}_{z'}\zeta\rangle| \lesssim [b]_{\dot{C}^{0,a}(\mathbb{R}^d)} \max\{t, t', |x-x'|\}^a.$$

Proof. We start off by recalling the trivial and well known estimates

$$|\langle b, \zeta_z\rangle - b_{B_z}| \lesssim \inf_{B_z} \mathrm{M}_\# b, \quad b_{B_z} := \frac{1}{t^d}\int_{B(x,t)} b(y)\mathrm{d}y$$

$$|b_{B_z} - b_{aB_z}| \lesssim \max\{1, \log(a)\} \inf_{B_z} \mathrm{M}_\# b, \quad a > 1$$

$$\left|b_{B_z} - b_{B_{z'}}\right| \lesssim \inf_{B_z} \mathrm{M}_\# b, \text{ when } B_z \cap B_{z'} \neq \emptyset, \ t \sim t'.$$

Henceforth by virtue of the first estimate, it suffices to estimate the quantity $\left|b_{B_z} - b_{B_{z'}}\right|$. Without loss of generality we assume that $t \geq t'$. First, we consider the case $\frac{|x-x'|}{t} \leq 1$ then $B(x,t) \cap B(x',t) \neq \varnothing$ so

$$|b_{B(x,t)} - b_{B(x',t')}| \leq |b_{B(x,t)} - b_{B(x',t)}| + |b_{B(x',t)} - b_{B(x',t')}| \lesssim \max\left\{1, \log\left(\frac{t}{t'}\right)\right\} \max\left\{\inf_{B_z} \mathrm{M}_\# b, \inf_{B_{z'}} \mathrm{M}_\# b\right\}.$$

To conclude the proof of the first part of the lemma we may focus, from now in the case that $\frac{|x-'x|}{t} > 1$.

$$\begin{aligned}|b_{B(x,t)} - b_{B(x',t')}| &\leq |b_{B(x,t)} - b_{B(x,|x-x'|)}| + |b_{B(x,|x-x'|)} - b_{B(x',|x-x'|)}| + |b_{B(x',t')} - b_{B(x',|x-x'|)}| \\ &\lesssim \max\left\{1, \log\left(\frac{|x-x'|}{t'}\right)\right\} \max\left\{\inf_{B_z} \mathrm{M}_{\#}b, \inf_{B_{z'}} \mathrm{M}_{\#}b\right\}.\end{aligned}$$

The second part of the lemma is easily obtained as a combination of the inequality

$$|\langle b, \zeta_{(x,t)}\rangle - b(x)| \lesssim [b]_{\dot{C}^{0,a}(\mathbb{R}^d)}$$

and the quantitative membership of b in $\dot{C}^{0,a}(\mathbb{R}^d)$. □

Proposition 4.3.3. *Let $\delta > 0$ then we have that for each $\eta \in (0, \delta)$ there holds*

$$d(z, z')_{\mathrm{BMO}}[z, z']_\delta \lesssim_{\delta,\eta} \|b\|_{\mathrm{BMO}} [z, z']_\eta.$$

Proof. The proof of this proposition is elementary and is based on the fact that $\log(t) \lesssim_\varepsilon t^\varepsilon$ when $t > 1$.

□

In the following lemma we estimate the $L^p(w, \mathbb{R}^d)$ norm of $\mathcal{R}_{T,b}$.

Lemma 4.3.4. *If $w \in A_p$ we have that*

$$\|\mathcal{R}_{T,b}(f)\|_{L^p(w,\mathbb{R}^d)} \lesssim \|b\|_{BMO} \|f\|_{L^p(w,\mathbb{R}^d)}.$$

Proof. We claim that for each $\eta \in (0, \delta)$ there exist $\lambda_z, \kappa_z \in C\Psi_z^{\eta;0}$ for some absolute positive constant C with the property that

$$\langle \mathcal{R}_{T,b}(f), g\rangle = \|b\|_{\mathrm{BMO}} \left(\int_{Z^d} \langle f, \lambda_z\rangle\langle u_z, g\rangle \mathrm{d}\mu(z) + \int_{Z^d} \langle f, v_z\rangle\langle \kappa_z, g\rangle \mathrm{d}\mu(z)\right).$$

Indeed, combining Lemmata 4.3.2, 4.1.2, [25, Lemma 2.3], the fact that $\overline{\psi_z^{(3)}}, \psi_z, u_z, v_z \in C\Psi_z^{\delta;0}$ and Proposition 4.3.3 we obtain that

$$\lambda_z := \frac{1}{\|b\|_{\mathrm{BMO}}} \int_{Z^d} \langle b, \phi_{z'}^{(1)} - \phi_z^{(1)} \rangle \langle \overline{\psi_{z'}^{(3)}}, v_z \rangle \psi_{z'} \mathrm{d}\mu(z') \in C\Psi_z^{\eta;0}$$
$$\kappa_z := \frac{1}{\|b\|_{\mathrm{BMO}}} \int_{Z^d} \langle b, \phi_z^{(1)} - \phi_{z'}^{(1)} \rangle \langle u_z, \psi_{z'} \rangle \overline{\psi_{z'}^{(3)}} \mathrm{d}\mu(z') \in C\Psi_z^{\eta;0}.$$

Using [25, Proposition 2.6] we obtain that the $(1,1)$ sparse norm of $\mathcal{R}_{T,b}$ is controlled by $\|b\|_{\mathrm{BMO}}$ henceforth by the well known theory of sparse forms, see for example [59] we obtain that

$$|\langle \mathcal{R}_{T,b} f, g \rangle| \lesssim \|b\|_{\mathrm{BMO}} \|f\|_{L^p(w)} \|g\|_{L^{p'}(w^{1-p'})}$$

from which we the conclusion of the lemma readily follows. □

The rest of this section is going to be concerned with the obtaining off-diagonal estimates. Initially, we will prove a sparse type estimate for bi-sublinear forms that arise either from the paraproducts or the error term $\mathcal{R}_{T,b}$ related to symbols that belong in the homogeneous Hölder space.

Lemma 4.3.5. *Let $\frac{a}{d} = \frac{1}{p} - \frac{1}{q}$ with $p < q$ and $\delta > 0$ then*

$$\sum_{Q \in \mathcal{D}} \ell_Q^{a+d} \Psi_Q^\delta f \Phi_Q^\delta g \lesssim \sup_{\mathcal{S}\text{ .5-sparse}} \sum_{R \in \mathcal{S}} \ell_R^{a+d} \langle f \rangle_{1,R} \langle g \rangle_{1,R}.$$

In particular, we have that

$$\sum_{Q \in \mathcal{D}} \ell_Q^{a+d} \Psi_Q^\delta f \Phi_Q^\delta g \lesssim \|f\|_{L^p(w^p, \mathbb{R}^d)} \|g\|_{L^{q'}(w^{-q'}, \mathbb{R}^d)}.$$

Proof. As the value of δ is not important for the proof of this lemma we will omit it from the notation of the intrinsic wavelet coefficients. By standard limiting arguments it suffices to show that for a finite collection of dyadic cubes, $\mathsf{T}(Q)$, contained in Q and $f, g \in L_0^\infty(\mathbb{R}^d)$ with the additional property that there holds $\operatorname{supp}(f), \operatorname{supp}(g) \subset 3Q$

$$\sum_{R \in \mathsf{T}(Q)} \ell_R^{a+d} \Psi_R f \ \Psi_R g \lesssim \sup_{\mathcal{S}\ .5\text{-sparse}} \sum_{R \in \mathcal{S}} \ell_R^{a+d} \langle f \rangle_{1,R} \langle g \rangle_{1,R}.$$

To initiate our stopping time argument, we make the initial observation that

$$\sum_{R \in \mathsf{T}(Q)} \ell_R^{a+d} \Psi_R f \ \Psi_R g \lesssim \ell_Q^{a+d} [f]_{1,\mathsf{T}(Q)} [g]_{1,\mathsf{T}(Q)}$$

where for each $\mathcal{K} \subset \mathcal{D}$ and $r > 0$ we adopt the notation $[f]_{r,\mathcal{K}} = \sup_{R \in \mathcal{K}} \inf_R \mathrm{M}_r f$. This estimate is an immediate consequence of the calculation

$$\sum_{R \subset Q} \ell_R^{a+d} = \sum_{2^k < 1} \sum_{\ell_R = 2^k \ell_Q, R \subset Q} \ell_R^{a+d} = \sum_{2^k < 1} (2^{-k})^d (2^k \ell_Q)^{a+d} \lesssim \ell_Q^{a+d}. \tag{4.3.1}$$

We construct $\mathcal{S}(Q)$ inductively as follows. We initially set $\mathcal{S}_0(Q) = \{Q\}$ and for $m \geq 1,\ R \in \mathcal{S}_{m-1}(Q)$

$\mathcal{I}(R) \coloneqq$ maximal $I \in \mathcal{D}$ with $9I \subset R \cap \left(\left\{ \mathrm{M}\left[\mathbf{1}_{3R} f\right] > C \left\langle f \right\rangle_{1,3R} \right\} \cup \left\{ \mathrm{M}\left[\mathbf{1}_{3R} g\right] > C \langle g \rangle_{1,3R} \right\} \right)$

$$\mathcal{S}_m(Q) \coloneqq \bigcup_{R \in \mathcal{S}_{m-1}(Q)} \mathcal{I}(R), \quad \mathcal{S}(Q) \coloneqq \bigcup_{m \geq 0} \mathcal{S}_m(Q).$$

Note that the packing condition

$$\sum_{I \in \mathcal{I}(R)} |I| \leq \frac{|R|}{2}$$

is a direct consequence of the weak $(1,1)$ type of the Hardy-Littlewood maximal operator and guarantees that the collection $\mathcal{S}(Q)$ is sparse. We proceed with the proof in an iterative scheme. We make the conscious choice to present only the first step of the iteration as the subsequent ones pass in an identical fashion.

It is easy to observe that $\mathcal{I}(Q)$ induces a natural splitting of the underlying collection $\mathsf{T}(Q)$. To be more precise,

$$\mathsf{T}(Q) = \mathsf{T}_{\mathsf{stop},\mathcal{I}}(Q) \sqcup \bigcup_{I \in \mathcal{I}(Q)} \mathsf{T}(I)$$

$$\mathsf{T}_{\mathsf{stop},\mathcal{I}}(Q) \coloneqq \{R \in \mathsf{T}(Q) : R \not\subset I \ , \ \forall I \in \mathcal{I}(Q)\} , \quad \mathsf{T}(I) \coloneqq \{R \in \mathsf{T}(Q) : R \subset I\} .$$

By the definition of $\mathcal{I}(Q)$ we learn that

$$[f]_{1,\mathsf{T}_{\mathsf{stop},\mathcal{I}}(Q)} \lesssim \langle f \rangle_{1,3Q}, \quad [g]_{1,\mathsf{T}_{\mathsf{stop},\mathcal{I}}(Q)} \lesssim \langle g \rangle_{1,3Q},$$

so that

$$\sum_{R \in \mathsf{T}(Q)} \ell_R^{a+d} \Psi_R f \ \Psi_R g \leq C \ell_Q^{a+d} \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q} + \sum_{I \in \mathcal{I}(Q)} \sum_{R \in \mathsf{T}(I)} \ell_R^{a+d} \Psi_R(f\mathbf{1}_{3Q}) \Psi_R(g\mathbf{1}_{3Q}). \quad (4.3.2)$$

For the tail terms we calculate

$$\begin{aligned}
\sum_{R \in \mathsf{T}(I)} \ell_R^{a+d} \Psi_R(f\mathbf{1}_{3Q \setminus 3I}) \Psi_R(g\mathbf{1}_{3I}) &\lesssim \sum_{k \geq 0} \sum_{R \in \mathsf{T}(I), \ell_R = 2^{-k}\ell_I} \ell_R^{a+d} 2^{-\delta k} \inf_I \mathrm{M}(f\mathbf{1}_{3Q}) \Psi_R(g\mathbf{1}_{3I}) \\
&\lesssim \sum_{k \geq 0} 2^{-k(a+d)} \ell_I^a \langle f \rangle_{1,3Q} \sum_{R \in \mathsf{T}(I), \ell_R = 2^{-k}\ell_Q} \ell_I^d \Psi_R(g\mathbf{1}_{3I}) \\
&\lesssim \ell_I^{a+d} \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q}.
\end{aligned}$$

So that

$$\begin{aligned}
\sum_{I \in \mathcal{I}(Q)} \sum_{R \in \mathsf{T}(I)} \ell_R^{a+d} \Psi_R(f\mathbf{1}_{3Q}) \Psi_R(g\mathbf{1}_{3Q}) \leq C &\sum_{I \in \mathcal{I}(Q)} \ell_I^{a+d} \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q} \\
&+ \sum_{I \in \mathcal{I}(\mathcal{Q})} \sum_{R \in \mathsf{T}(I)} \ell_R^{a+d} \Psi_R(f\mathbf{1}_{3I}) \Psi_R(g\mathbf{1}_{3I})
\end{aligned}$$

where the first term is controlled as follows using the same single scale analysis as in (4.3.1)

$$\sum_{I \in \mathcal{I}} \ell_I^{a+d} \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q} \lesssim \ell_Q^a \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q} \sum_{\mathcal{I}(Q)} |I| \lesssim \ell_Q^{a+d} \langle f \rangle_{1,3Q} \langle g \rangle_{1,3Q}$$

and the proof is completed by iterating (4.3.2). Finally, the second part of our lemma is available to us from the proof of [13, Theorem 4.5] and the 3^d grid trick.

$\square$

Proposition 4.3.6. *Let T be a δ-cancellative Calderón-Zygmund operator and $b \in \dot{C}^{0,a}(\mathbb{R}^d)$. Then for each $\eta \in (0,\delta)$ there exist $\alpha_z^\iota, \beta_z^\iota, \in C\Psi_z^{\eta;0}$ and γ_z^ι such that $|\gamma_z| \le Ct^{a-d}\chi_z^{-d-\eta+a}$, $\iota \in \{1,2\}$ with the property that*

$$\begin{aligned}\langle \mathcal{R}_{T,b} f, g\rangle = &\int_{Z^d} [b]_{\dot{C}^{0,a}(\mathbb{R}^d)} \left(t^a \langle f, \psi_z\rangle\langle \alpha_z^1 . g\rangle + t^a\langle f, \beta_z^1\rangle\langle u_z, g\rangle + \langle f, \psi_z\rangle\langle \gamma_z^1, g\rangle\right) \mathrm{d}\mu(z)\\ &- \int_{Z^d} [b]_{\dot{C}^{0,a}(\mathbb{R}^d)} \left(t^a\langle f, v_z\rangle\langle \alpha_z^2 . g\rangle + t^a\langle f, \beta_z^2\rangle\langle \overline{\psi_z^{(3)}}, g\rangle + \langle f, v_z\rangle\langle \gamma_z^2, g\rangle\right) \mathrm{d}\mu(z).\end{aligned}$$

Proof. Initially, given $z \in Z^d$ we partition the parameter space in the following manner

$$Z^d = \mathcal{A}(z) \sqcup \mathcal{B}(z) \sqcup \mathcal{C}(z)$$

$$\mathcal{A}(z) := \left\{z' \in Z^d : \max\{|x-x'|, t, t'\} = t\right\}, \;\; \mathcal{B}(z) := \left\{z' \in Z^d : \max\{|x-x'|, t, t'\} = t'\right\},$$

$$\mathcal{C}(z) := \left\{z' \in Z^d : \max\{|x-x'|, t, t'\} = |x-x'|\right\}.$$

We will only calculate α_z^1, β_z^1 and γ_z^1 as the arguments for the rest are identical. We write

$$\begin{aligned}\alpha_z^1 &= \frac{1}{t^a [b]_{\dot{C}^{0,a}(\mathbb{R}^d)}} \int_{Z^d} \mathbf{1}_{\mathcal{A}(z)}(z')\langle b, \phi_{z'}^{(1)} - \phi_z^{(1)}\rangle\langle \overline{\psi_z^{(3)}}, v_{z'}\rangle u_{z'} \mathrm{d}\mu(z')\\ \beta_z^1 &= \frac{1}{t^a [b]_{\dot{C}^{0,a}(\mathbb{R}^d)}} \int_{Z^d} \mathbf{1}_{\mathcal{A}(z')}(z)\langle b, \phi_{z'}^{(1)} - \phi_z^{(1)}\rangle\langle \overline{\psi_{z'}^{(3)}}, v_z\rangle \psi_{z'} \mathrm{d}\mu(z')\\ \gamma_z^1 &= \frac{1}{[b]_{\dot{C}^{0,a}(\mathbb{R}^d)}} \int_{Z^d} \mathbf{1}_{\mathcal{A}(z)}(z')\langle b, \phi_{z'}^{(1)} - \phi_z^{(1)}\rangle\langle \overline{\psi_z^{(3)}}, v_{z'}\rangle u_{z'} \mathrm{d}\mu(z').\end{aligned}$$

We learn the memberships $\alpha_z^1, \beta_z^1 \in C\Psi_z^{\delta;0}$ and the size estimate on γ_z^1 via combining Lemma 4.3.2 with Lemmata 4.1.2 and 4.1.3 respectively. $\square$

The above lemma yields readily the following corollary that allows us to give off-diagonal estimates.

Lemma 4.3.7. *Let T be δ-cancellative Calderón-Zygmund operator and $b \in \dot{L}^r(\mathbb{R}^d)$. Then for each $\eta \in (0,\delta)$ there exist $\sigma_z^\iota, \tau_z^\iota \in C\Psi_z^{\eta;0}, \iota \in \{1,2\}$ with the property that*

$$\langle \mathcal{R}_{T,b} f, g\rangle = \int_{Z^d} \inf_{B_z} \mathrm{M}_\# b \left(\langle f, \psi_z\rangle\langle \sigma_z^1, g\rangle + \langle u_z, g\rangle\langle f, \tau_z^1\rangle - \langle f, \sigma_z^2\rangle\langle \overline{\psi_z^{(3)}}, g\rangle - \langle f, v_z\rangle\langle \tau_z^2, g\rangle \right) \mathrm{d}\mu(z).$$

Proof. Given $z \in Z^d$ we will partition the parameter space Z^d according to the sharp maximal function. In particular

$$Z^d = \mathcal{H}(z) \sqcup (\mathcal{H}(z))^c \quad \mathcal{H}(z) := \left\{ z' \in Z : \inf_{B_z} \mathrm{M}_\# b \geq \inf_{B_{z'}} \mathrm{M}_\# b \right\}.$$

We proceed similarly as in the proof of Lemma 4.3.6. □

We end this section with the proof of Theorem F.

Proof of Theorem F. Using Corollary H.1, Lemma 4.3.5 and the latter part of Lemma 4.3.2 we obtain the desired estimates for $\mathcal{M}_{T,b}$ for the cases $p = q, q < p$ and $p > q$ respectively. By the decomposition of the commutator as expanded in (4.2.2) it suffices to prove the relevant estimates for the term $\mathcal{R}_{T,b}$. The case $p = q$ is contained in Lemma 4.3.4. The case $p > q$ follows by combining Proposition 4.3.6 and Lemma 4.3.5 since

$$|\langle \mathcal{R}_{T,b}(f), g\rangle| \lesssim [b]_{\dot{C}^{0,a}(\mathbb{R}^d)} \left(\sum_{Q \in \mathcal{D}} |Q| \Psi_Q^{\frac{\delta}{2}} f \Psi_Q^{\frac{\delta}{2}} g + \langle I_a(|f|), |g|\rangle \right)$$

where the first term comes from the $a_z^\iota, \beta_z^\iota,\ \iota \in \{1,2\}$ and the second term comes from the calculation of the kernel of the relevant operator, see Lemma 4.5.1. Finally, from Lemma 4.3.7 we learn that

$$|\langle \mathcal{R}_{T,b}(f), g\rangle| \lesssim \int_{\mathbb{R}^d} \mathrm{M}_\# b S(f) S(g) \tag{4.3.3}$$

which completes the case $p < q$. □

Remark 4.3.8. We remark that we could have concluded the cases $p < q$ and $p = q$ from (4.3.3) but we want to emphasize the point that the "correct" quantitative estimate in the case that $p = q$ of the term $\mathcal{R}_{T,b}$ is that its sparse $(1,1)$ norm controlled by $\|b\|_{\mathrm{BMO}}$.

4.4 Two weight estimates

In this section we will reprove Bloom's original theorem for the Hilbert transform or for any fully cancellative Calderón-Zygmund operator with smoothness 1.

The first ingredient in the proof of the two weight inequality is an analogue of Lemma 4.3.2 in the Bloom BMO setting. For the sake of completeness we recall the following John-Nirenberg type inequality from [38].

Lemma 4.4.1. *If* $w \in A_2$ *the following inequality holds*

$$\sup_Q \left(\frac{1}{w(Q)} \int_Q |b(x) - \langle b \rangle_Q|^2 w^{-1}(x) \mathrm{d}x \right)^{\frac{1}{2}} \lesssim \|b\|_{\mathrm{BMO}_w}.$$

At this point we will introduce the analogue of $d(z,z')_{\mathrm{BMO}}$ in the Bloom BMO setting. Given $w \in A_{2-\varepsilon}$ we set

$$d(z,z')_{\mathrm{BMO}_w} = \begin{cases} \max\{\langle w \rangle_{1,B(x,t)}, \langle w \rangle_{1,B(x',t')}\} \max\left\{ \left(\frac{|x-x'|}{t'}\right)^{1-\frac{\epsilon}{2}}, \left(\frac{t}{t'}\right)^{1-\frac{\epsilon}{2}} \right\} & \text{if } t \geq t' \\ \max\{\langle w \rangle_{1,B(x,t)}, \langle w \rangle_{1,B(x',t')}\} \max\left\{ \left(\frac{|x-x'|}{t}\right)^{1-\frac{\epsilon}{2}}, \left(\frac{t'}{t}\right)^{1-\frac{\epsilon}{2}} \right\} & \text{if else.} \end{cases}$$

Lemma 4.4.2. *Let* $w \in A_{2-\varepsilon}$ *and* $\zeta \in \mathcal{S}(\mathbb{R})$ *with* $\int_{\mathbb{R}} \zeta = 1$. *then we have that*

$$|\langle b, \mathsf{Sy}_z \zeta - \mathsf{Sy}_{z'} \zeta \rangle| \lesssim d(z,z')_{\mathrm{BMO}_w} \|b\|_{\mathrm{BMO}_w}.$$

Proof. Using the doubling property of the measure w the Schwartz decay of ζ and its mean 1 property we learn that it is sufficient to show this estimate for rough averages namely, it is enough to show that

$$\left|b_{B_z} - b_{B_{z'}}\right| \lesssim d(z,z')_{\mathrm{BMO}_w} \|b\|_{\mathrm{BMO}_w}.$$

Without loss of generality we can assume that $t \geq t'$. Initially, if $|x - x'| \leq t$ we are able to write

$$|b_{B(x,t)} - b_{B(x',t')}| \leq |b_{B(x,t)} - b_{B(x,3t)}| + |b_{B(x,3t)} - b_{B(x',t')}|.$$

For the first term we have that

$$|b_{B(x,t)} - b_{B(x,3t)}| \lesssim \|b\|_{\mathrm{BMO}_w} \langle w\rangle_{1,B(x,t)}. \tag{4.4.1}$$

For the second term we have that

$$|b_{B(x,3t)} - b_{B(x',t')}| \leq \frac{1}{t'}\int_{B(x',t')} |b(y) - b_{B(x,3t)}|\mathrm{d}y \leq \langle (b - b_{B(x,3t)})w^{-\frac{1}{2}}\rangle_{2,B(x',t')}\langle w^{\frac{1}{2}}\rangle_{2,B(x',t')}.$$

Observe that $B(x,t)\cap B(x',t') \neq \varnothing$ because $x' \in B(x,t)\cap B(x',t')$ since $|x-x'| \leq t$ so that $B(x',t') \subset B(x,3t)$

$$\begin{aligned}
\langle (b - b_{B(x,3t)})w^{-\frac{1}{2}}\rangle_{2,B(x',t')} &\leq \left(\frac{1}{t'}\int_{B(x,3t)} |b(y) - b_{B(x,3t)}|^2 w^{-1}(y)\mathrm{d}y\right)^{\frac{1}{2}} \\
&\leq \left(\frac{w(B(x,3t))}{t'}\right)^{\frac{1}{2}} \left(\frac{1}{w(B(x,3t))}\int_{B(x,3t)} |b(y) - b_{B(x,3t)}|^2 w^{-1}(y)\mathrm{d}y\right)^{\frac{1}{2}} \\
&\lesssim \left(\frac{w(B(x',9t))}{t'}\right)^{\frac{1}{2}} \|b\|_{\mathrm{BMO}_w} \lesssim \left(\frac{w(B(x',t'))}{t'}\left[\frac{t}{t'}\right]^{2-\epsilon}\right)^{\frac{1}{2}} \|b\|_{\mathrm{BMO}_w} \\
&\lesssim \left(\langle w\rangle_{1,B(x',t')}\right)^{\frac{1}{2}} \left(\frac{t}{t'}\right)^{1-\frac{\epsilon}{2}} \|b\|_{\mathrm{BMO}_w}.
\end{aligned}$$

Where the third inequality was obtained since $B(x, 3t) \cap B(x', 3t) \neq \varnothing \Rightarrow B(x, 3t) \subset B(x', 9t)$ and the final one was obtained by the doubling property of w. Therefore, all in all we have that

$$|b_{B(x,3t)} - b_{B(x',t')}| \lesssim \left(\frac{t}{t'}\right)^{1-\frac{\epsilon}{2}} \langle w \rangle_{1,B(x',t')} \|b\|_{\mathrm{BMO}_w}. \tag{4.4.2}$$

Combining (4.4.1) and (4.4.2) yields the desired estimate. To finish the proof we check the case $|x - x'| > t$ by estimating

$$|b_{B(x,t)} - b_{B(x',t')}| \leq |b_{B(x,t)} - b_{B(x',|x-x'|)}| + |b_{B(x',|x-x'|)} - b_{B(x',t')}|$$

and using the previous estimate and w's doubling property. $\square$

4.5 Appendix

Lemma 4.5.1. *Let $\lambda_\zeta \in \Psi_\zeta^{\delta;0}$ and κ_ζ with the property that $|\kappa_\zeta(y)| \lesssim \sigma^{a-d} \left\langle \frac{y-\xi}{\sigma} \right\rangle^{-(d+\delta-a)}$. Then*

$$\left| \int_{Z^d} \langle f, \lambda_z \rangle \kappa_z \mathrm{d}\mu(z) \right| \lesssim I_a(|f|).$$

Here I_a is the fractional integral operator with parameter a.

Proof.

$$\int_{Z^d} \langle f, \varphi_\zeta \rangle \kappa_\zeta(x) \mathrm{d}\mu(\zeta) = \int_{Z^d} \left(\int_{\mathbb{R}^d} f(y) \overline{\varphi_\zeta}(y) \kappa_\zeta(x) \mathrm{d}x \right) \mathrm{d}\mu(\zeta) \leq \int_{\mathbb{R}^d} |f| \left(\int_{Z^d} |\varphi_\zeta(y) \kappa_\zeta(x)| \mathrm{d}\mu(\zeta) \right) \mathrm{d}y.$$

The inner integral, i.e the kernel, is estimated as follows

$$\begin{aligned}\int_{Z^d} |\varphi_\zeta(y)\kappa_\zeta(x)|\mathrm{d}\mu(\zeta) &\lesssim \int_0^\infty \int_{\mathbb{R}^d} \frac{1}{\sigma^d}\left\langle \frac{y-\xi}{\sigma}\right\rangle^{-(d+\delta)} \sigma^{a-d}\left\langle \frac{x-\xi}{\sigma}\right\rangle^{-(d+\delta-a)} \mathrm{d}\xi\frac{\mathrm{d}\sigma}{\sigma}\\ &\lesssim \int_0^\infty \sigma^{a-d}\left\langle \frac{y-x}{\sigma}\right\rangle^{-(d+\delta-a)}\frac{\mathrm{d}\sigma}{\sigma}\\ &= \int_0^{|y-x|} \sigma^{a-d}\left\langle \frac{y-x}{\sigma}\right\rangle^{-(d+\delta-a)}\frac{\mathrm{d}\sigma}{\sigma} + \int_{|y-x|}^\infty \sigma^{a-d}\left\langle \frac{y-x}{\sigma}\right\rangle^{-(d+\delta-a)}\frac{\mathrm{d}\sigma}{\sigma}.\end{aligned}$$

We estimate each integral separately

$$\int_0^{|y-x|} \sigma^{a-d}\left\langle \frac{y-x}{\sigma}\right\rangle^{-(d+\delta-a)}\frac{\mathrm{d}\sigma}{\sigma} \lesssim \int_0^{|y-x|}\sigma^{a-d-1}\frac{\sigma^{d+\delta-a}}{|y-x|^{d+\delta-a}}\mathrm{d}\sigma \lesssim \frac{|y-x|^\delta}{|y-x|^{d+\delta-a}} = \frac{1}{|y-x|^{d-a}}$$

$$\int_{|y-x|}^\infty \sigma^{a-d}\left\langle \frac{y-x}{\sigma}\right\rangle^{-(d+\delta-a)}\frac{\mathrm{d}\sigma}{\sigma} \le \int_{|y-x|}^\infty \frac{1}{\sigma^{1+d-a}}\mathrm{d}\sigma = \left[\frac{\sigma^{a-d}}{a-d}\right]_{|y-x|}^\infty \lesssim |y-x|^{a-d}$$

and the proof is complete. □

At this point we give the proof of the generalization of the wavelet averaging lemma appearing in [25] and its fractional analogue as stated in Lemma 4.1.3.

Proof of Lemma 4.1.2.

$$\begin{aligned}|\psi_z^{\mathsf{n}}(y)| &\le \int_{a\in\mathbb{R}^d}\int_0^1 \frac{\min\{t,\beta t\}^\delta}{\max\{t,\beta t,|a|t\}^{d+\delta}}t^d\frac{1}{(\beta t)^{d+}}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a\\ &\le \int_{a\in\mathbb{R}^d}\int_0^1 \frac{\beta^\delta}{\langle a\rangle^{d+\delta}}\frac{1}{(\beta t)^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a \qquad (4.5.1)\\ &\le \int_0^1\frac{\beta^{\delta-1-d}}{t^d}\int_{a\in\mathbb{R}^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}a\mathrm{d}\beta \lesssim \frac{1}{t^d}\int_0^1\beta^{\delta-1}\mathrm{d}\beta\lesssim\frac{1}{t^d}.\end{aligned}$$

To conclude the localization estimate it suffices to check it when $|y-x|\ge 2t$. Indeed,

$$\begin{aligned}|\partial^\gamma\psi_z^{\mathsf{n}}(y)| &\le \int_{a\in\mathbb{R}^d}\int_0^1 \frac{\min\{t,\beta t\}^\delta}{\max\{t,\beta t,|a|t\}^{d+\delta}}t^d\frac{1}{(\beta t)^{d}}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a\\ &\le \int_{a\in\mathbb{R}^d}\int_0^1 \frac{\beta^\delta}{\langle a\rangle^{d+\delta}}\frac{1}{(\beta t)^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a\\ &\int_{|a|\le\frac{|y-x|}{2t}}\int_0^1\frac{\beta^{+\delta-d-1}}{\langle a\rangle^{d+\delta}}\frac{1}{t^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}\beta\mathrm{d}a\\ &+\int_{|a|\ge\frac{|y-x|}{2t}}\int_0^1\frac{\beta^{\delta-d-1}}{\langle a\rangle^{d+\delta}}\frac{1}{t^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}\beta\mathrm{d}a = \mathsf{I}+\mathsf{II}.\end{aligned}$$

We will estimate each term separately. In particular we start off with the term I. Therefore,

$$
\begin{aligned}
\mathsf{I} &\lesssim \frac{1}{t^d}\int_{|a|\le\frac{|y-x|}{2t}}\frac{1}{\langle a\rangle^{d+\delta}}\left(\int_0^1\beta^{\delta-d-1+d+\delta}\left(\frac{t}{|y-x|}\right)^{d+\delta}\mathrm{d}\beta\right)\mathrm{d}a\\
&\le \frac{1}{t^d}\int_{a\in\mathbb{R}^d}\frac{1}{\langle a\rangle^{d+\delta}}\left(\frac{2t}{|y-x|}\right)^{d+\delta}\mathrm{d}a\lesssim\frac{1}{t^d}\left\langle\frac{y-x}{t}\right\rangle^{-d-\delta}
\end{aligned}
$$

We now turn to the second term II. We observe that $\langle a\rangle\gtrsim\left\langle\frac{y-x}{t}\right\rangle$ and therefore

$$
\mathsf{II}\lesssim\frac{1}{t^d}\left\langle\frac{y-x}{t}\right\rangle^{-d-\delta}\int_{|a|\ge\frac{|y-x|}{2t}}\int_0^1\beta^{\delta-d-1}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}\beta\mathrm{d}a
$$

which concludes the proof by the computation in (4.5.1). We will know prove the smoothness estimate. For $|h|\le t$ we compute the modulus of the difference $\psi_z^{\mathsf{n}}(y+h)-\psi_z^{\mathsf{n}}(y)$ as

$$
\left|\int_{a\in\mathbb{R}^d}\left(\int_{|h|<\beta t\le t}+\int_{|h|\ge\beta t>0}\right)\frac{\beta^\delta}{\langle a\rangle^{d+\delta}}\left\{\varphi_{(x+at,\beta t)}(y+h)-\varphi_{(x+at,\beta t)}(y)\right\}\mathrm{d}a\frac{\mathrm{d}\beta}{\beta}\right|:=|\mathsf{I}+\mathsf{II}|
$$

We will use the Hölder continuity estimate on regime $|h|<\beta t\le t$ and rely solely on the localization principle on the regime $|h|\ge\beta t>0$. So we compute

$$
\begin{aligned}
&\int_{|h|<\beta t\le t}\int_{a\in\mathbb{R}^d}\frac{\beta^{\delta-1}}{\langle a\rangle^{d+\delta}}\frac{|h|^\delta}{(\beta t)^{d+\delta}}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}a\mathrm{d}\beta\lesssim\frac{|h|^\delta}{t^{d+\delta}}\int_{\frac{|h|}{t}}^1\beta^{-1}\mathrm{d}\beta\\
&\lesssim\frac{|h|^\delta}{t^{d+\delta}}\log_1\left(\frac{t}{|h|}\right)\lesssim\frac{|h|^\eta}{t^{d+\eta}},\quad 0<\eta<\delta.
\end{aligned}
$$

We continue with estimating the second term by

$$
\begin{aligned}
&\int_0^{\frac{|h|}{t}}\beta^{\delta-1}\int_{a\in\mathbb{R}^d}\frac{1}{\langle a\rangle^{d+\delta}}\frac{1}{(\beta t)^d}\left\langle\frac{y+h-x-at}{\beta t}\right\rangle^{-d-\delta}+\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\mathrm{d}a\mathrm{d}\beta\\
&\lesssim\int_0^{\frac{|h|}{t}}\beta^{\delta-1-|\gamma|}\mathrm{d}\beta\lesssim\frac{1}{t^d}\left(\frac{|h|}{t}\right)^\delta.
\end{aligned}
$$

□

Proof of Lemma 4.1.3. Due to similarity in the treatment of ψ_z^{f} we choose to sketch the proof of the localization estimates of ψ_z^{n}. Initially there holds,

$$
|\psi_z^{\mathsf{n}}(y)|\le\int_{|a|\gtrsim 1}\int_0^1\frac{\beta^\delta}{\langle a\rangle^{d+\delta-\lambda}}t^{-d+\lambda}\frac{1}{(\beta)^d}\left\langle\frac{y-x-at}{\beta t}\right\rangle^{-d-\delta}\frac{\mathrm{d}\beta}{\beta}\mathrm{d}a
$$

Take $p(d+\delta-\lambda) > 1$ and $\frac{d}{p'} + \delta - d > 0$ which is possible because

$$p(d+\delta-\lambda) > 1 \Leftrightarrow p > \frac{1}{d+\delta-\lambda}$$
$$\frac{d}{p'} + \delta - d > 0 \Leftrightarrow p > \frac{d}{\delta}$$

we switch the integrals and perfrorm the Hölder in a to obtain that

$$\begin{aligned} |\psi_z^{\mathsf{n}}(y)| &\le t^{\lambda-d} \int_0^1 \beta^{\delta-d-1} \left(\int_{|a|\gtrsim 1} \frac{1}{\langle a \rangle^{p(d+\delta-\lambda)}} \mathrm{d}a \right)^{\frac{1}{p}} \left(\int_{|a|\gtrsim 1} \left\langle \frac{y-x-at}{\beta t} \right\rangle^{-p'(d+\delta)} \mathrm{d}a \right)^{\frac{1}{p'}} \mathrm{d}\beta \\ &\lesssim t^{\lambda-d} \int_0^1 \beta^{\delta-d-1} \beta^{\frac{d}{p'}} \mathrm{d}\beta = t^{\lambda-d} \left[\frac{\beta^{\delta+\frac{d}{p'}-d}}{\delta+\frac{d}{p'}-d} \right]_0^1 \lesssim t^{\lambda-d} \end{aligned}$$

where the last line is justified from the fact that $\delta + \frac{d}{p'} - d = \delta - \frac{d}{p} > 0$ For the range $|y-x| \ge 2t$ we proceed to the splitting of the integral as in the proof of the previous lemma and use the convolution inequality

$$\int_{\mathbb{R}^d} \frac{1}{t^d} \left\langle \frac{y-x}{t} \right\rangle^{-(d+\delta)} \frac{t^\lambda}{t^d} \left\langle \frac{y-x'}{t} \right\rangle^{-(d+\delta-\lambda)} \mathrm{d}y \lesssim t^{\lambda-d} \left\langle \frac{x'-x}{t} \right\rangle^{-(d+\delta-\lambda)}.$$

□

Bibliography

[1] N. Yu. Antonov. Convergence of Fourier series. In *Proceedings of the XX Workshop on Function Theory (Moscow, 1995)*, volume 2, pages 187–196, 1996. vii, 2, 13

[2] J. Arias-de Reyna. Pointwise convergence of Fourier series. *J. London Math. Soc. (2)*, 65(1):139–153, 2002. 13

[3] Árpád Bényi and Rodolfo H. Torres. Compact bilinear operators and commutators. *Proc. Amer. Math. Soc.*, 141(10):3609–3621, 2013. 76

[4] Steven Bloom. A commutator theorem and weighted BMO. *Trans. Amer. Math. Soc.*, 292(1):103–122, 1985. 88

[5] Peter Borwein and Tamás Erdélyi. Nikolskii-type inequalities for shift invariant function spaces. *Proc. Amer. Math. Soc.*, 134(11):3243–3246, 2006. 16

[6] Mingming Cao, Andrea Olivo, and Kôzô Yabuta. Extrapolation for multilinear compact operators and applications. *Trans. Amer. Math. Soc.*, 375(7):5011–5070, 2022. 4, 76

[7] Lennart Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966. vi, vii, 1

[8] M. J. Carro, M. Mastylo, and L. Rodríguez-Piazza. Almost everywhere convergent Fourier series. *J. Fourier Anal. Appl.*, 18(2):266–286, 2012. 12

[9] María J. Carro and Carlos Domingo-Salazar. Endpoint estimates for Rubio de Francia operators. *Trans. Amer. Math. Soc.*, 371(3):1621–1648, 2019. 12, 13, 14

[10] María Jesús Carro, Loukas Grafakos, José María Martell, and Fernando Soria. Multilinear extrapolation and applications to the bilinear Hilbert transform. *J. Math. Anal. Appl.*, 357(2):479–497, 2009. 12

[11] R. R. Coifman, R. Rochberg, and Guido Weiss. Factorization theorems for Hardy spaces in several variables. *Ann. of Math. (2)*, 103(3):611–635, 1976. 88

[12] José M. Conde-Alonso, Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. A sparse domination principle for rough singular integrals. *Anal. PDE*, 10(5):1255–1284, 2017. 11

[13] David Cruz-Uribe. Two weight norm inequalities for fractional integral operators and commutators, 2014. 102

[14] David Cruz-Uribe and José María Martell. Limited range multilinear extrapolation with applications to the bilinear Hilbert transform. *Math. Ann.*, 371(1-2):615–653, 2018. 16

[15] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. A sparse estimate for multisublinear forms involving vector-valued maximal functions. In *Bruno Pini Mathematical Analysis Seminar 2017*, volume 8 of *Bruno Pini Math. Anal. Semin.*, pages 168–184. Univ. Bologna, Alma Mater Stud., Bologna, 2017. 10

[16] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. Domination of multilinear singular integrals by positive sparse forms. *J. Lond. Math. Soc. (2)*, 98(2):369–392, 2018. 16, 17, 38, 47

[17] Guy David and Jean-Lin Journé. A boundedness criterion for generalized Calderón-Zygmund operators. *Ann. of Math. (2)*, 120(2):371–397, 1984. 67

[18] Ciprian Demeter and Francesco Di Plinio. Endpoint bounds for the quartile operator. *J. Fourier Anal. Appl.*, 19(4):836–856, 2013. 18

[19] Ciprian Demeter, Michael T. Lacey, Terence Tao, and Christoph Thiele. Breaking the duality in the return times theorem. *Duke Math. J.*, 143(2):281–355, 2008. vii, 1

[20] Francesco Di Plinio, Yen Q. Do, and Gennady N. Uraltsev. Positive sparse domination of variational Carleson operators. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 18(4):1443–1458, 2018. 47

[21] Francesco Di Plinio, Walton Green, and Brett D. Wick. Bilinear wavelet representation of Calderón-Zygmund forms. *Pure Appl. Anal.*, 5(1):47–83, 2023. 70, 71, 72, 77, 79

[22] Francesco Di Plinio and Andrei K. Lerner. On weighted norm inequalities for the Carleson and Walsh-Carleson operator. *J. Lond. Math. Soc. (2)*, 90(3):654–674, 2014. 14, 18

[23] Francesco Di Plinio and Yumeng Ou. A modulation invariant Carleson embedding theorem outside local L^2. *J. Anal. Math.*, 135(2):675–711, 2018. 15, 38, 47, 49

[24] Francesco Di Plinio and Christoph Thiele. Endpoint bounds for the bilinear Hilbert transform. *Trans. Amer. Math. Soc.*, 368(6):3931–3972, 2016. 15, 18

[25] Francesco Di Plinio, Brett D. Wick, and Tyler Williams. Wavelet representation of singular integral operators. *Math. Ann.*, 386(3-4):1829–1889, 2023. viii, 6, 70, 71, 72, 84, 86, 89, 90, 93, 96, 99, 107

[26] Yen Do and Christoph Thiele. L^p theory for outer measures and two themes of Lennart Carleson united. *Bull. Amer. Math. Soc. (N.S.)*, 52(2):249–296, 2015. 15, 26, 29, 31, 38

[27] Xuan Thinh Duong, Irina Holmes, Ji Li, Brett D. Wick, and Dongyong Yang. Two weight commutators in the dirichlet and neumann laplacian settings. *Journal of Functional Analysis*, 276(4):1007–1060, 2019. 93

[28] Charles Fefferman. Pointwise convergence of Fourier series. *Ann. of Math. (2)*, 98:551–571, 1973. vii, 1, 13

[29] Anastasios Fragkos, A. Walton Green, and Brett D. Wick. Multilinear wavelet compact t(1) theorem, 2023. 3, 4

[30] Anastasios Fragkos and Brett D. Wick. Commutators via the wavelet representation theorem of calderón-zygmund operators, 2023. 6

[31] Michael Frazier, Björn Jawerth, and Guido Weiss. *Littlewood-Paley theory and the study of function spaces*, volume 79 of *CBMS Regional Conference Series in Mathematics*. Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1991. 86

[32] Dorothee Frey and Zoe Nieraeth. Weak and strong type $A_1 - A_\infty$ estimates for sparsely dominated operators. *J. Geom. Anal.*, 29(1):247–282, 2019. 12

[33] Gennady Uraltsev. *Time-Frequency Analysis of the Variational Carleson Operator using outer-measure Lp spaces*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, August 2017. 91

[34] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014. 29

[35] Loukas Grafakos, José María Martell, and Fernando Soria. Weighted norm inequalities for maximally modulated singular integral operators. *Math. Ann.*, 331(2):359–394, 2005. 13

[36] Loukas Grafakos and Rodolfo H. Torres. Multilinear Calderón-Zygmund theory. *Adv. Math.*, 165(1):124–164, 2002. 67, 72

[37] Irina Holmes, Michael T. Lacey, and Brett D. Wick. Bloom's inequality: commutators in a two-weight setting. *Arch. Math. (Basel)*, 106(1):53–63, 2016. 89

[38] Irina Holmes, Michael T. Lacey, and Brett D. Wick. Commutators in the two-weight setting. *Math. Ann.*, 367(1-2):51–80, 2017. 88, 104

[39] Richard A. Hunt. On the convergence of Fourier series. In *Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)*, pages 235–255. Southern Illinois Univ. Press, Carbondale, Ill., 1968. vi, vii, 1

[40] Tuomas Hytönen and Stefanos Lappas. Extrapolation of compactness on weighted spaces: bilinear operators. *Indag. Math. (N.S.)*, 33(2):397–420, 2022. 4, 76

[41] Tuomas Hytönen and Stefanos Lappas. Extrapolation of compactness on weighted spaces. *Rev. Mat. Iberoam.*, 39(1):91–122, 2023. 4

[42] Tuomas P. Hytönen. Representation of singular integrals by dyadic operators, and the A_2 theorem. *Expo. Math.* 89

[43] Tuomas P. Hytönen. The L^p-to-L^q boundedness of commutators with applications to the Jacobian operator. *J. Math. Pures Appl. (9)*, 156:351–391, 2021. 88

[44] Tuomas P. Hytönen and Michael T. Lacey. Pointwise convergence of vector-valued Fourier series. *Math. Ann.*, 357(4):1329–1361, 2013. 34

[45] N. J. Kalton. Convexity, type and the three space problem. *Studia Math.*, 69(3):247–287, 1980/81. 12

[46] Michael Lacey and Christoph Thiele. L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. *Ann. of Math. (2)*, 146(3):693–724, 1997. 16, 61

[47] Michael Lacey and Christoph Thiele. On Calderón's conjecture. *Ann. of Math. (2)*, 149(2):475–496, 1999. 16, 61

[48] Michael Lacey and Christoph Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000. vii, 1, 12, 14, 15, 39, 45

[49] Michael T. Lacey. The bilinear maximal functions map into L^p for $2/3 < p \leq 1$. *Ann. of Math. (2)*, 151(1):35–57, 2000. 18

[50] Michael T. Lacey and Xiaochun Li. Maximal theorems for the directional Hilbert transform on the plane. *Trans. Amer. Math. Soc.*, 358(9):4099–4117, 2006. 39

[51] Andrei K. Lerner and Fedor Nazarov. Intuitive dyadic calculus: the basics. *Expo. Math.*, 37(3):225–265, 2019. 10, 20, 72

[52] Andrei K. Lerner and Sheldy Ombrosi. Some remarks on the pointwise sparse domination. *J. Geom. Anal.*, 30(1):1011–1027, 2020. 18

[53] Kangwei Li, José María Martell, Henri Martikainen, Sheldy Ombrosi, and Emil Vuorinen. End-point estimates, extrapolation for multilinear Muckenhoupt classes, and applications. *Trans. Amer. Math. Soc.*, 374(1):97–135, 2021. 4, 16, 17, 72

[54] Kangwei Li, José María Martell, and Sheldy Ombrosi. Extrapolation for multilinear Muckenhoupt classes and applications. *Adv. Math.*, 373:107286, 43, 2020. 11, 16, 17

[55] Kangwei Li, Kabe Moen, and Wenchang Sun. The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators. *J. Fourier Anal. Appl.*, 20(4):751–765, 2014. 72

[56] Victor Lie. On the boundedness of the Carleson operator near L^1. *Rev. Mat. Iberoam.*, 29(4):1239–1262, 2013. vii, 1, 13

[57] Mishko Mitkovski and Cody B. Stockdale. On the $T1$ theorem for compactness of Calderón-Zygmund operators. *preprint*, 2023. 4, 6

[58] Mishko Mitkovski, Cody B. Stockdale, Nathan A. Wagner, and Brett D. Wick. Riesz-Kolmogorov type compactness criteria in function spaces with applications. *Complex Anal. Oper. Theory*, 17(3):Paper No. 40, 31, 2023. 81

[59] Kabe Moen. Sharp weighted bounds without testing or extrapolation. *Arch. Math. (Basel)*, 99(5):457–466, 2012. 99

[60] Camil Muscalu, Terence Tao, and Christoph Thiele. Multi-linear operators given by singular multipliers. *J. Amer. Math. Soc.*, 15(2):469–496, 2002. 16, 61, 63, 64

[61] Camil Muscalu, Terence Tao, and Christoph Thiele. Uniform estimates on multi-linear operators with modulation symmetry. volume 88, pages 255–309. 2002. Dedicated to the memory of Tom Wolff. 16

[62] Fedor Nazarov, Richard Oberlin, and Christoph Thiele. A Calderón Zygmund decomposition for multiple frequencies and an application to an extension of a lemma of Bourgain. *Math. Res. Lett.*, 17(3):529–545, 2010. 15

[63] Zoe Nieraeth. Quantitative estimates and extrapolation for multilinear weight classes. *Math. Ann.*, 375(1-2):453–507, 2019. 4, 10, 11, 16, 17, 72

[64] Krzysztof Nowak. On Calderón-Toeplitz operators. *Monatsh. Math.*, 116(1):49–72, 1993. 72

[65] Richard Oberlin, Andreas Seeger, Terence Tao, Christoph Thiele, and James Wright. A variation norm Carleson theorem. *J. Eur. Math. Soc. (JEMS)*, 14(2):421–464, 2012. 60

[66] Richard Oberlin and Christoph Thiele. New uniform bounds for a Walsh model of the bilinear Hilbert transform. *Indiana Univ. Math. J.*, 60(5):1693–1712, 2011. 15, 16

[67] Jan-Fredrik Olsen and Paco Villarroya. Endpoint estimates for compact Calderón-Zygmund operators. *Rev. Mat. Iberoam.*, 33(4):1285–1308, 2017. 4

[68] Karl-Mikael Perfekt, Sandra Pott, and Paco Villarroya. Endpoint compactness of singular integrals and perturbations of the Cauchy integral. *Kyoto J. Math.*, 57(2):365–393, 2017. 4

[69] Francesco Di Plinio and Anastasios Fragkos. The weak-type carleson theorem via wave packet estimates, 2022. 1

[70] Richard Rochberg. Toeplitz and Hankel operators, wavelets, NWO sequences, and almost diagonalization of operators. In *Operator theory: operator algebras and applications, Part 1 (Durham, NH, 1988)*, volume 51, Part 1 of *Proc. Sympos. Pure Math.*, pages 425–444. Amer. Math. Soc., Providence, RI, 1990. 72

[71] Per Sjölin. An inequality of Paley and convergence a.e. of Walsh-Fourier series. *Ark. Mat.*, 7:551–570, 1969. 12

[72] Per Sjölin and Fernando Soria. Remarks on a theorem by N. Yu. Antonov. *Studia Math.*, 158(1):79–97, 2003. 13

[73] Cody B. Stockdale, Paco Villarroya, and Brett D. Wick. Sparse domination results for compactness on weighted spaces. *Collect. Math.*, 73(3):535–563, 2022. 4

[74] Christoph Thiele. A uniform estimate. *Ann. of Math. (2)*, 156(2):519–563, 2002. 16

[75] Masatsugu Tsuji. On the compactness of space $L^p(p > 0)$ and its application to integral equations. *Kodai Math. Sem. Rep.*, 3:33–36, 1951. 81

[76] Paco Villarroya. A characterization of compactness for singular integrals. *J. Math. Pures Appl. (9)*, 104(3):485–532, 2015. 4, 6, 85

[77] Paco Villarroya. A global Tb theorem for compactness and boundedness of Calderón-Zygmund operators. *J. Math. Anal. Appl.*, 480(1):123323, 41, 2019. 4