SPARSE BOUNDS FOR PSEUDODIFFERENTIAL OPERATORS

DAVID BELTRAN AND LAURA CLADEK

ABSTRACT. We prove sparse bounds for pseudodifferential operators associated to Hörmander symbol classes. Our sparse bounds are sharp up to the endpoint and rely on a single scale analysis. As a consequence, we deduce a range of weighted estimates for pseudodifferential operators. The results naturally apply to the context of oscillatory Fourier multipliers, with applications to dispersive equations and oscillatory convolution kernels.

1. INTRODUCTION AND STATEMENT OF RESULTS

Given a smooth function $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$, define the associated pseudodifferential operator T_a by

$$T_a f(x) = \int_{\mathbb{R}^n} e^{ix \cdot \xi} a(x, \xi) \widehat{f}(\xi) d\xi,$$

where $f \in \mathcal{S}$ and $\widehat{f}$ denotes the Fourier transform of f. The smooth function a is assumed to belong to the symbol classes $S_{\rho,\delta}^m$, introduced by Hörmander in [24] and further studied by many authors, see for instance [7, 21, 46, 9]. Recall that $S_{\rho,\delta}^m$ consists of all $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ satisfying the differential inequalities

$$(1) \qquad |\partial_x^\nu \partial_\xi^\sigma a(x, \xi)| \lesssim (1 + |\xi|)^{m - \rho|\sigma| + \delta|\nu|}$$

for all multi-indices $\nu, \sigma \in \mathbb{N}^n$, where $m \in \mathbb{R}$ and $0 \leqslant \delta, \rho \leqslant 1$. [1]

The main goal of this article is to prove sparse domination for such pseudodifferential operators. Obtaining sparse bounds for classical operators in harmonic analysis has been an active area of research in recent years, starting with fundamental work of Lerner [35, 36] in the context of Calderón–Zygmund operators. The original Banach space norm domination of Lerner was shortly further refined to a pointwise sparse control, see [14, 38, 31, 37]. More recently, a new perspective on the subject has been taken, which consists in a domination in the context of bilinear forms. This was introduced by Bernicot, Frey and Petermichl [5] and was further developed in the work of Culiuc, Di Plinio and Ou [16, 17]. This approach has proved to be highly successful, as it applies to operators that fall well beyond the classical Calderón–Zygmund theory. Among many examples, we may find Bochner–Riesz multipliers [4, 33], rough singular integrals [13], the bilinear Hilbert transform [16], the variational Carleson operator [19], oscillatory singular integrals [34, 28], spherical maximal functions [30], a specific singular Radon transform [44] or the recent work by Ou and the second author [12] for Hilbert transforms along curves;

---

[1]We use the notation $A \lesssim B$ to denote that there is a constant C such that $A \leqslant CB$ and the notation $A \lesssim_\epsilon B$ to specify the dependence of the implicit constant in a certain parameter ϵ. We omit the constant factors of π coming from our normalisation of the Fourier transform.

we also refer to [29, 18] for sparse domination in a discrete setting.

While the technique of sparse domination has been successfully applied to study weighted bounds for a wide variety of operators in harmonic analysis, as the ones listed above, it has not previously been used to yield any weighted bounds for pseudodifferential operators. However, weighted A_p estimates for these objects have been studied before. Chanillo and Torchinksy proved in [11] that if $a \in S_{\rho,\delta}^{-n(1-\rho)/2}$, with $0 < \delta < \rho < 1$, then T_a is bounded on $L^p(w)$ for $w \in A_{p/2}$ and $2 \leqslant p < \infty$. Later, Michalowski, Rule and Staubach [40] extended this result to $\delta = \rho$, and also proved that for the smaller classes $S_{\rho,\delta}^{-n(1-\rho)}$, with $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$, there is $L^p(w)$ boundedness for $w \in A_p$ and $1 < p < \infty$, see [41]. More recently, the first author [1] established weighted boundedness for the symbol classes $S_{\rho,\delta}^m$ with $-n(1-\rho)/2 < m < 0$ and $0 \leqslant \delta \leqslant \rho < 1$; in this case T_a is bounded on $L^p(w)$ if $w \in A_{p/2} \cap RH_{(2t'/p)'}$ and $2 \leqslant p < 2t'$, where $t = (\rho - 1)n/2m$. [2]

The weighted bounds obtained as a corollary of our sparse domination theorem yield new weighted estimates for pseudodifferential operators and also recover all the bounds mentioned in the above paragraph, except for some endpoint cases. A more detailed discussion is postponed to Section 4. The simplicity of the methods through which we are able to obtain weighted corollaries that improve existing results in the literature is a further testament to the power and versatility of the method of sparse domination.

We first give the definition of sparse collections of cubes and sparse bilinear forms, and then follow with the statement of our main theorem. It is instructive to interpret our results recalling the embeddings of the symbol classes, that is $S_{\rho_1,\delta_1}^{m_1} \subseteq S_{\rho_2,\delta_2}^{m_2}$ if $m_1 \leqslant m_2, \rho_1 \geqslant \rho_2$ and $\delta_1 \leqslant \delta_2$.

Definition 1.1. A collection of dyadic cubes $\mathcal{S}$ in $\mathbb{R}^n$ is η-sparse for $0 < \eta < 1$ if there exist sets $\{E_Q : Q \in \mathcal{S}\}$ that are pairwise disjoint, $E_Q \subset Q$, and satisfy $|E_Q| > \eta|Q|$ for all $Q \in \mathcal{S}$. For any dyadic cube Q, for $1 \leqslant p < \infty$ define $\langle f \rangle_{p,Q}^p :=$ $|Q|^{-1} \int_Q |f|^p \, dx$. Given $1 \leqslant r, s < \infty$, the (r, s) sparse form $\Lambda_{\mathcal{S},r,s}$ is defined as

$$\Lambda_{\mathcal{S},r,s}(f,g) := \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s,Q},$$

and the r-sparse operator $\mathcal{A}_{r,\mathcal{S}}$ is defined as

$$\mathcal{A}_{r,\mathcal{S}} f(x) := \sum_{Q \in \mathcal{S}} \langle f \rangle_{r,Q} \chi_Q(x).$$

Theorem 1.2. *Let $a \in S_{\rho,\delta}^m$ for some $m < 0$ and $0 < \delta \leqslant \rho < 1$. Then for any compactly supported bounded functions f, g on $\mathbb{R}^n$, there exist sparse collections $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ of dyadic cubes such that*

$$|\langle T_a f, g \rangle| \leqslant C(m, \rho, r, s)\Lambda_{\mathcal{S},r,s'}(f,g)$$

and

$$|\langle T_a f, g \rangle| \leqslant C(m, \rho, r, s)\Lambda_{\widetilde{\mathcal{S}},s',r}(f,g)$$

for all pairs (r, s') and (s', r) such that

$$m < -n(1-\rho)(1/r - 1/2), \qquad 1 \leqslant r \leqslant s \leqslant 2$$

---

[2] Given $1 \leqslant r \leqslant \infty$, r' denotes its conjugate Hölder exponent, that is $1/r + 1 + r' = 1$.

or

$$m < -n(1-\rho)(1/r - 1/s), \qquad 1 \leqslant r \leqslant 2 \leqslant s \leqslant r'.$$

The sparse bounds in Theorem 1.2 are best understood in the plane $(1/r, 1/s')$. Given a symbol $a \in S^m_{\rho,\delta}$, where $m < 0$ and $0 \leqslant \delta \leqslant \rho < 1$, the pairs (r, s') for which a sparse bound holds satisfy that $(1/r, 1/s')$ lie inside an *open* trapezoid $\mathcal{T}_{m,\rho,n}$ with vertexes depending on the parameters m, ρ and the dimension n (see Figure 1). Such vertexes are given by

$$v_1 = (1,0), \quad v_2 = \Big(1, \frac{-m}{n(1-\rho)}\Big), \quad v_3 = \Big(\frac{-m}{n(1-\rho)}, 1\Big), \quad v_4 = (0,1),$$

if $-n(1-\rho) \leqslant m \leqslant -n(1-\rho)/2$, and

$$v_1 = \Big(\frac{1}{2} + \frac{-m}{n(1-\rho)}, \frac{1}{2} + \frac{m}{n(1-\rho)}\Big), \quad v_2 = \Big(\frac{1}{2} + \frac{-m}{n(1-\rho)}, 1/2\Big),$$

$$v_3 = \Big(\frac{1}{2}, \frac{1}{2} + \frac{-m}{n(1-\rho)}\Big), \qquad\qquad v_4 = \Big(\frac{1}{2} + \frac{m}{n(1-\rho)}, \frac{1}{2} + \frac{-m}{n(1-\rho)}\Big),$$
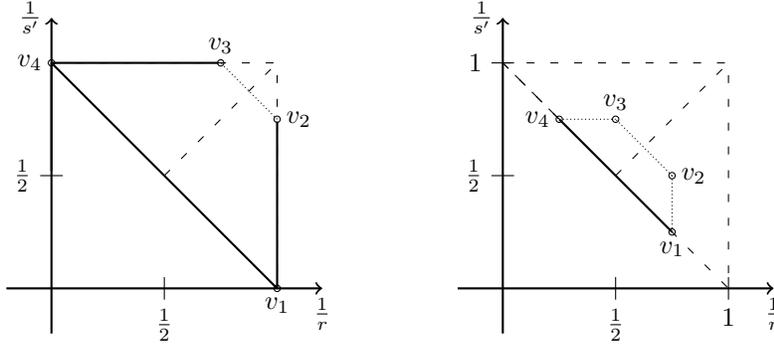
if $-n(1-\rho)/2 < m < 0$.

FIGURE 1. The region for sparse bounds. These bounds are sharp up to the boundary of the region, and also include the thick boundary segments. The trapezoid $\mathcal{T}_{m,\rho,n}$ on the left corresponds to the case $-n(1-\rho) < m \leqslant -n(1-\rho)/2$, and the one on the right to the case $-n(1-\rho)/2 < m < 0$.

The methods to prove the above domination theorem are inspired by the recent works of Lacey and Spencer [34] and Lacey, Mena and Reguera [33]. It consists in obtaining geometrically decaying sparse bounds for a single scale version of the operators under study. Its short proof will be given in Section 2.

The above sparse bounds are sharp up to the endpoints, as it will be shown in Section 3. Concerning endpoint results, we remark that the symbols in the class $S^0_{1,\delta}$ with $\delta < 1$ [3] are Calderón–Zygmund operators, and thus a pointwise sparse bound by L^1 averages (in the context of the upcoming Theorem 1.3) follows from

---

[3]The case $\delta = 1$ is naturally excluded from all statements as it is well know that there are symbols in that class that fail to be bounded on L^2; see, for instance [46].

[14, 38]. Also, for the specific classes $S_{\rho,\delta}^{-n(1-\rho)}$, with $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$, it is possible to obtain the following endpoint sparse bound in the stronger pointwise control context.

Theorem 1.3. *Let $a \in S_{\rho,\delta}^{-n(1-\rho)}$, $0 < \rho \leqslant 1$, $0 \leqslant \delta < 1$. Then for every $f \in C_0^\infty(\mathbb{R}^n)$ and any $r > 1$, there exists a sparse family $\mathcal{S}$ such that for a.e. $x \in \mathbb{R}^n$,*

$$|T_a f(x)| \lesssim \mathcal{A}_{r,\mathcal{S}} f(x).$$

Of course this implies domination for the symbol classes $S_{\rho,\delta}^{-n(1-\rho)}$ by a bilinear $(r, 1)$ sparse form. The proof of this result is a straightforward consequence of the pointwise domination principle of Lerner [37] and a sharp maximal function estimate in [41]. A sketch of the proof is given in an appendix.

Finally, we note that our techniques naturally extend to the setting of oscillatory Fourier multipliers associated to differential inequalities of the type (1). In particular, we are able to obtain sparse bounds for the model Fourier multipliers $a_{\rho,m}(\xi) := e^{i|\xi|^{1-\rho}}(1 + |\xi|)^m \chi_{\{|\xi|^{1-\rho} \geqslant 1\}}$ for any value of $\rho \in \mathbb{R}$. Interestingly, when $\rho = -1$, the results may be interpreted in the context of the free linear Schrödinger equation. We refer to Section 5 for further discussion.

2. Sparse bilinear form domination for Pseudodifferential operators

This section is devoted to prove Theorem 1.2, our sparse domination theorem for pseudodifferential operators. To this end, we first prove a domination by a geometrically decaying sequence of sparse bilinear forms, which will be later refined to a domination by a true sparse bilinear form. A key ingredient in the proof is the use of optimal L^p improving estimates for frequency and spatially localised pieces of the operator.

Such Lebesgue space estimates are obtained by interpolating first the $L^2 \to L^2$ and $L^1 \to L^1$ estimates to yield optimal $L^s \to L^s$ bounds for $1 \leqslant s \leqslant 2$. Then, the use of Bernstein's inequality lead to the optimal $L^r \to L^s$ bounds for $1 \leqslant r \leqslant s \leqslant 2$, and a further interpolation with the $L^1 \to L^\infty$ estimates leads to the optimal $L^r \to L^s$ bounds for $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$.

This will allow us to obtain domination by a (r, s') bilinear sparse form for such values of r and s. The corresponding (r, s') sparse bounds associated to the remaining possible cases for the values of r and s will follow by duality, as the Hörmander symbol classes of pseudodifferential operators are closed under adjoints.

Decomposition of the operators T_a. Let $a \in S_{\rho,\delta}^m$, where $m < 0$ and $0 \leqslant \delta \leqslant \rho < 1$, and consider its associated pseudodifferential operator T_a. Let ψ be a smooth cutoff that is identically 1 in the unit ball and supported in its double, and let $\widetilde{\psi}(\cdot) := \psi(\cdot) - \psi(2(\cdot))$. Decompose T_a into frequency localised operators T_a^j for integers $j \geqslant 0$, where T_a^j is defined by

$$\begin{cases} T_a^0 f(x) = \int a(x,\xi)\psi(\xi)\widehat{f}(\xi)e^{ix\cdot\xi}d\xi, \\ T_a^j f(x) = \int a(x,\xi)\widetilde{\psi}(2^{-j}\xi)\widehat{f}(\xi)e^{ix\cdot\xi}d\xi, \quad j > 0. \end{cases}$$

It is well known that the operator T_a^j is naturally associated to the spatial scale $2^{-j\rho}$; this will become clear in the forthcoming Lemma 2.1 via an integration by parts argument. Thus we further decompose each T_a^j into spatially localised operators $T_a^{j,l}$ defined by

$$T_a^{j,l} f(x) = \int a(x,\xi)\widetilde{\psi}(2^{-j}\xi) \int f(y)\widetilde{\psi}(2^{-l+j\rho}(x-y))e^{i(x-y)\cdot\xi} \, dy \, d\xi$$

for $l \in \mathbb{Z}$ and $j > 0$, and in an analogous manner for $T_a^{0,l}$, so that

$$T_a = \sum_{j \geqslant 0} T_a^j = \sum_{l \in \mathbb{Z}} \sum_{j \geqslant 0} T_a^{j,l}.$$

We further group the pieces $T_a^{j,l}$ according to their spatial scale. Fix some $\epsilon > 0$ and write

$$T_a^j = \sum_{l \leqslant j\epsilon} T_a^{j,l} + \sum_{l > j\epsilon} T_a^{j,l}.$$

Our L^p improving estimates for the pieces $\sum_{l \leqslant j\epsilon} T_a^{j,l}$ and $T_a^{j,l}$ for $l > j\epsilon$ in what follows, and consequently our sparse bounds for T_a, will depend on this fixed value of ϵ. This is the reason why we achieve sparse bounds only up to an endpoint (see Figure 1).

$L^2 \to L^2$ bounds. It is a classical result of Calderón and Vaillancourt [7] that the symbol classes $S_{\rho,\rho}^0$ are bounded on L^2 for $0 \leqslant \rho < 1$. This and an integration by parts argument yield the following L^2 bounds.

Lemma 2.1. *We have*

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^2 \to L^2} \lesssim_\epsilon 2^{jm}, \\ \|T_a^{j,l}\|_{L^2 \to L^2} \lesssim_\epsilon 2^{10n(m-n)(j+l)}, \quad l > j\epsilon. \end{cases}$$

Proof. For $l \geqslant j\epsilon$, we integrate by parts in the ξ variable $N(\epsilon)$ many times to obtain

$$|T_a^{j,l} f(x)| \lesssim \int\int |\nabla_\xi^{N(\epsilon)}[a(x,\xi)\widetilde{\psi}(2^{-j}\xi)]||x-y|^{-N(\epsilon)}|f(y)|\widetilde{\psi}(2^{-l+j\rho}(x-y)) \, dy \, d\xi$$

$$\lesssim_\epsilon \|2^{-lN(\epsilon)/2}(1+|\cdot|)^m\widetilde{\psi}(2^{-j}\cdot)\|_{L^1(d\xi)} \int |f(y)|2^{-lN(\epsilon)/2}\widetilde{\psi}(2^{-l+j\rho}(x-y)) \, dy$$

$$\lesssim \int |f(y)|2^{-lN(\epsilon)/2}\widetilde{\psi}(2^{-l+j\rho}(x-y)) \, dy,$$

where $N(\epsilon)$ is taken sufficiently large to guarantee that $2^{-lN(\epsilon)/2}2^{jm}\widetilde{\psi}(2^{-j}(\cdot)) \in L^1$ uniformly in j and $l > j\epsilon$. Imposing also that $N(\epsilon)$ is such that $2^{-lN(\epsilon)/4}\widetilde{\psi}(2^{-l+j\rho}(\cdot)) \in$

L^1, uniformly in j and $l > j\epsilon$, and $2^{-lN(\epsilon)/4} \leqslant 2^{10n(m-n)(j+l)}$, an application of Young's inequality yields

$$\|T_a^{j,l}f\|_2 \lesssim_\epsilon 2^{10n(m-n)(j+l)}\|f\|_2, \qquad l > j\epsilon,$$

as desired.

To prove the first statement, note that for any fixed j, we have that $T_a^j = \sum_l T_a^{j,l}$ is a pseudodifferential operator associated to a symbol in the class $S_{\rho,\delta}^0$ with constant 2^{jm}, so $\|\sum_l T_a^{j,l}\|_{L^2 \to L^2} \lesssim 2^{jm}$ by the L^2 boundedness of the symbol classes. The result then follows from noting that $\sum_{l \leqslant j\epsilon} T_a^{j,l} = \sum_l T_a^{j,l} - \sum_{l > j\epsilon} T_a^{j,l}$, as we just saw that the L^2 norm of $T_a^{j,l}$ for $l > j\epsilon$ decays geometrically in l. □

$L^1 \to L^\infty$ **bounds.** As in the proof of the $L^2 \to L^2$ bounds, integrating by parts in the ξ variable $N(\epsilon)$ many times for a sufficiently large $N(\epsilon)$ in the regime $l > j\epsilon$ leads to the estimate

$$\|T_a^{j,l}f\|_\infty \lesssim_\epsilon \sup_x \int |f(y)|2^{-lN(\epsilon)/2}\widetilde{\psi}(2^{-l+j\rho}(x-y))\,dy \lesssim_\epsilon 2^{10n(m-n)(j+l)}\|f\|_1.$$

For $l \leqslant j\epsilon$, we may trivially bound

$$\Big\|\sum_{l \leqslant j\epsilon} T_a^{j,l}f\Big\|_\infty \lesssim \int (1+|\xi|)^m\widetilde{\psi}(2^{-j}\xi)\int |f(y)|\,dy\,d\xi \lesssim 2^{jm+jn}\|f\|_1.$$

We summarize this as the following lemma.

Lemma 2.2. *We have*

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^1 \to L^\infty} \lesssim_\epsilon 2^{jm+jn}, \\ \|T_a^{j,l}\|_{L^1 \to L^\infty} \lesssim_\epsilon 2^{10n(m-n)(j+l)}, \quad l > j\epsilon. \end{cases}$$

$L^1 \to L^1$ **bounds.** As above, integrating by parts in the ξ variable $N(\epsilon)$ many times for a sufficiently large $N(\epsilon)$ in the regime $l > j\epsilon$ leads to the estimate

$$\|T_a^{j,l}f\|_1 \lesssim_\epsilon \int\int |f(y)|2^{-lN(\epsilon)/2}\widetilde{\psi}(2^{-l+j\rho}(x-y))\,dydx \lesssim_\epsilon 2^{10n(m-n)(j+l)}\|f\|_1.$$

Reasoning as in Lemma 2.1, the L^1 boundedness for $\sum_{l \leqslant j\epsilon} T_a^{j,l}$ will follow then from that of T_a^j.

Lemma 2.3. *For any $\varepsilon > 0$, we have*

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^1 \to L^1} \lesssim_\epsilon 2^{j(m+n(1-\rho)/2+\varepsilon)}, \\ \|T_a^{j,l}\|_{L^1 \to L^1} \lesssim_\epsilon 2^{10n(m-n)(j+l)}, \qquad l > j\epsilon. \end{cases}$$

Remark 2.4. It is possible to prove the above Lemma with $\varepsilon = 0$ using the result of Fefferman [21] that asserts that $T_a : H^1 \to L^1$ for $a \in S_{\rho,\delta}^{-n(1-\rho)/2}$. However, for completeness of the paper, we prefer to give the proof for the above easier lemma than appealing to the result in [21]. The presence of ε will play no role in determining the (open) trapezoid $\mathcal{T}_{m,\rho,n}$ for which (r, s') sparse bounds hold, provided $\varepsilon > 0$ is taken sufficiently small.

Proof. It suffices to show that

(2) $$\|T_a^j\|_{L^1 \to L^1} \lesssim 2^{j(m+n(1-\rho)/2+\varepsilon)}$$

for any $\varepsilon > 0$. This in turn follows from

(3) $$\|T_a^j\|_{L^\infty \to L^\infty} \lesssim 2^{j(m+n(1-\rho)/2)},$$

as then one can sum geometrically in j and obtain L^∞ boundedness for the symbol classes $S^m_{\rho,\delta}$ with $m < -n(1-\rho)/2$. By duality, such symbol classes are also bounded on L^1 and one readily deduces (2).

The inequality (3) may be found in [21]; we include the simple argument for interest of the reader. Write T^j_a as

$$T^j_a f(x) = \int_{\mathbb{R}^n} K_j(x, x-y) f(y) dy,$$

where

$$K_j(x, z) := \int_{\mathbb{R}^n} e^{iz \cdot \xi} a(x, \xi) \widetilde{\psi}(2^{-j}\xi) d\xi.$$

T^j_a may be interpreted as the convolution of the function $K(x, \cdot)$ with f evaluated at the point x, so it suffices to compute $\|K(x, \cdot)\|_1$ uniformly in x.

To this end, for some $t > 0$ to be determined later, split the integral in two parts,

$$\int_{\mathbb{R}^n} |K_j(x, z)| dz = \int_{|z|<t} |K_j(x, z)| dz + \int_{|z|>t} |K_j(x, z)| dz.$$

For the first term, the Cauchy–Schwarz inequality, Plancherel's theorem, and the differential inequalities (1) with $\gamma = 0$ yield

$$\int_{|z|<t} |K_j(x, z)| dz \leqslant t^{n/2} \Big(\int_{\mathbb{R}^n} |K_j(x, z)|^2 dz \Big)^{1/2}$$

$$= t^{n/2} \int_{\mathbb{R}^n} |a(x, \xi)|^2 |\widetilde{\psi}(2^{-j}\xi)|^2 d\xi \Big)^{1/2}$$

$$\lesssim t^{n/2} 2^{jn/2} 2^{jm}.$$

For the second term we proceed similarly, but exploiting the differential inequalities (1) for $\gamma \neq 0$. Thus,

$$\int_{|z|>t} |K_j(x, z)| dz \leqslant \Big(\int_{|z|>t} \frac{1}{|z|^{2N}} \Big)^{1/2} \Big(\int_{\mathbb{R}^n} |z|^{2N} |K_j(x, z)|^2 dz \Big)^{1/2}$$

$$\lesssim t^{n/2-N} \Big(\int_{\mathbb{R}^n} \sum_{\gamma \in \mathbb{N}^d : |\gamma|=2N} |D^\gamma[a(x, \xi) \widetilde{\psi}(2^{-j}\xi)]|^2 d\xi \Big)^{1/2}$$

$$\lesssim t^{n/2-N} 2^{jn/2} 2^{jm-j\rho N},$$

provided $N > n/2$. Choosing $t = 2^{-j\rho}$, one concludes that

$$\int_{\mathbb{R}^n} |K_j(x, z)| dz \lesssim 2^{jm+jn(1-\rho)/2}.$$

$\square$

Obtaining $L^r \to L^s$ bounds for $1 \leqslant r \leqslant s \leqslant 2$. Interpolating Lemma 2.1 and Lemma 2.3 leads to the bounds

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T^{j,l}_a\|_{L^s \to L^s} \lesssim_\epsilon 2^{jm+j(n(1-\rho)/2+\varepsilon)(\frac{2}{s}-1)} \\ \|T^{j,l}_a\|_{L^s \to L^s} \lesssim_\epsilon 2^{10n(m-n)(j+l)}, & l > j\epsilon \end{cases}$$

for $1 \leqslant s \leqslant 2$.

Using Bernstein's inequality, we have

$$\|T_a^{j,l}f\|_{L^s} \lesssim_\epsilon \|T_a^{j,l}\|_{L^s \to L^s}\|f * \mathcal{F}[(\psi(2^{-j-100}\cdot) - \psi(2^{-j+100}\cdot))]\|_{L^s}$$

$$\lesssim \|T_a^{j,l}\|_{L^s \to L^s} 2^{-jn(\frac{1}{s}-\frac{1}{r})}\|f\|_{L^r}.$$

Similarly,

$$\Big\|\sum_{l \leqslant j\epsilon} T_a^{j,l}f\Big\|_{L^s} \lesssim_\epsilon \Big\|\sum_{l \leqslant j\epsilon} T_a^{j,l}\Big\|_{L^s \to L^s}\|f * \mathcal{F}[(\psi(2^{-j-100}\cdot) - \psi(2^{-j+100}\cdot))]\|_{L^s}$$

$$\lesssim \Big\|\sum_{l \leqslant j\epsilon} T_a^{j,l}\Big\|_{L^s \to L^s} 2^{-jn(\frac{1}{s}-\frac{1}{r})}\|f\|_{L^r}.$$

This leads to the following lemma.

Lemma 2.5. *For* $1 \leqslant r \leqslant s \leqslant 2$, *we have*

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^r \to L^s} \lesssim_\epsilon 2^{jm}2^{-jn(\frac{1}{s}-\frac{1}{r})}2^{j(n(1-\rho)/2+\varepsilon)(\frac{2}{s}-1)} \\ \|T_a^{j,l}\|_{L^r \to L^s} \lesssim_\epsilon 2^{-jn(\frac{1}{s}-\frac{1}{r})}2^{10n(m-n)(j+l)}, & l > j\epsilon. \end{cases}$$

Obtaining $L^r \to L^s$ **bounds for** $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$. Interpolating Lemma 2.1 and Lemma 2.2 leads to the bounds

$$(4) \qquad \begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^r \to L^{r'}} \lesssim_\epsilon 2^{jm+jn(\frac{2}{r}-1)} \\ \|T_a^{j,l}\|_{L^r \to L^{r'}} \lesssim_\epsilon 2^{10n(m-n)(j+l)}, & l > j\epsilon \end{cases}$$

for $1 \leqslant r \leqslant 2$. Reasoning as above, Lemma 2.1 and Bernstein's inequality leads to the bounds

$$(5) \qquad \begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^r \to L^2} \lesssim_\epsilon 2^{jm-jn(\frac{1}{2}-\frac{1}{r})} \\ \|T_a^{j,l}\|_{L^r \to L^2} \lesssim_\epsilon 2^{10n(m-n)(j+l)}2^{-jn(\frac{1}{2}-\frac{1}{r})}, & l > j\epsilon. \end{cases}$$

Interpolating (4) and (5) leads to the following estimate.

Lemma 2.6. *For* $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$, *we have*

$$\begin{cases} \|\sum_{l \leqslant j\epsilon} T_a^{j,l}\|_{L^r \to L^s} \lesssim_\epsilon 2^{jm-jn(\frac{1}{s}-\frac{1}{r})} \\ \|T_a^{j,l}\|_{L^r \to L^s} \lesssim_\epsilon 2^{-jn(\frac{1}{s}-\frac{1}{r})}2^{10n(m-n)(j+l)}, & l > j\epsilon. \end{cases}$$

Obtaining geometrically decaying sparse bounds. The sparse domination for T_a follows from establishing the corresponding sparse domination results for $\sum_{l \leqslant j\epsilon} T_a^{j,l}$ and $\sum_{l > j\epsilon} T_a^{j,l}$. Note that we have much better $L^r \to L^s$ estimates for the terms $T_a^{j,l}$ with $l > j\epsilon$, so these terms will not determine the values of r and s for which we achieve a (r, s') sparse domination result, and we will deal with those tails later.

We first invoke the standard "3^n grid trick". Define the standard universal shifted dyadic grids $\mathcal{D}_{\vec{v}}$ as

$$\left\{ 2^k\left[m_1 + \frac{v_1}{3}, m_1 + 1 + \frac{v_1}{3}\right] \times \cdots \times 2^k\left[m_n + \frac{v_n}{3}, m_n + 1 + \frac{v_n}{3}\right], k \in \mathbb{Z}, \vec{m} \in \mathbb{Z}^n \right\}$$

where $\vec{v} \in \{0,1,2\}^n$. We can then decompose each operator $T_a^{j,l}$ as

$$T_a^{j,l} = \sum_{\vec{v}} T_{a,\vec{v}}^{j,l}$$

where we define $T_{a,\vec{v}}^{j,l}$ by

$$T_{a,\vec{v}}^{j,l}f := \sum_{\substack{Q \in \mathcal{D}_{\vec{v}} \\ l(Q)=2^{\lfloor -j\rho+j\epsilon+10 \rfloor}}} T_a^{j,l}(f\chi_{\frac{1}{3}Q}) \qquad \text{if } l \leqslant j\epsilon,$$

and

$$T_{a,\vec{v}}^{j,l}f := \sum_{\substack{Q \in \mathcal{D}_{\vec{v}} \\ l(Q)=2^{\lfloor -j\rho+l+10 \rfloor}}} T_a^{j,l}(f\chi_{\frac{1}{3}Q}) \qquad \text{if } l > j\epsilon.$$

We choose to use this decomposition because now the support of each term $T_a^{j,l}(f\chi_{\frac{1}{3}Q})$ lies entirely in Q. Clearly, it suffices to prove the desired results with the standard dyadic grid replaced with a specific shifted dyadic grid $\mathcal{D}_{\vec{v}}$ for some fixed choice of $\vec{v} \in \{0,1,2\}^n$, and with the operator T_a replaced with the operator $T_{a,\vec{v}}$ defined by

$$T_{a,\vec{v}} := \sum_{j,l} T_{a,\vec{v}}^{j,l}.$$

Thus, in what follows, we will suppress the additional subscript $\vec{v}$ in our notation and write T_a and $T_a^{j,l}$ in place of $T_a^{\vec{v}}$ and $T_{a,\vec{v}}^{j,l}$ and also implicitly assume that our dyadic cubes lie in the shifted dyadic grid $\mathcal{D}_{\vec{v}}$, for some specific choice of $\vec{v}$.

With this in mind, observe first that the localisation and boundedness of $\sum_{l \leqslant j\epsilon} T_a^{j,l}$ gives the domination

$$\left\langle \sum_{l \leqslant j\epsilon} T_a^{j,l}f, g \right\rangle \lesssim \sum_{j \geqslant 0} \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q)=2^{\lfloor -j\rho+j\epsilon+10 \rfloor}}} \left\| \sum_{l \leqslant j\epsilon} T_a^{j,l}f \right\|_{L^s(Q)} \|g\|_{L^{s'}(Q)}$$

$$\lesssim_\epsilon \sum_{j \geqslant 0} \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q)=2^{\lfloor -j\rho+j\epsilon+10 \rfloor}}} |Q|^{1/r+1/s'-1} \left\| \sum_{l \leqslant j\epsilon} T_a^{j,l} \right\|_{L^r \to L^s} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}.$$

To obtain (r,s') sparse bounds for $1 \leqslant r \leqslant s \leqslant 2$, we use Lemma 2.5,

$$\left\langle \sum_{l \leqslant j\epsilon} T_a^{j,l}f, g \right\rangle \lesssim_\epsilon \sum_{j \geqslant 0} 2^{(-j\rho+j\epsilon)n(1/r-1/s)} 2^{jm} 2^{-jn(\frac{1}{s}-\frac{1}{r})} 2^{j(n(1-\rho)/2+\varepsilon)(2/s-1)}$$

$$\times \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q)=2^{\lfloor -j\rho+j\epsilon+10 \rfloor}}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}.$$

Note that this gives geometrically decaying (r,s') sparse bounds for sufficiently small choice of $\epsilon > 0$ and $\varepsilon > 0$ whenever we have

$$m < -n(1-\rho)(1/r-1/2), \qquad 1 \leqslant r \leqslant s \leqslant 2.$$

To obtain (r,s') sparse bounds for $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$, we use Lemma 2.6,

$$\left\langle \sum_{l \leqslant j\epsilon} T_a^{j,l}f, g \right\rangle \lesssim_\epsilon \sum_{j \geqslant 0} 2^{(-j\rho+j\epsilon)n(1/r-1/s)} 2^{jm-jn(\frac{1}{s}-\frac{1}{r})}$$

$$\times \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q)=2^{\lfloor -j\rho+j\epsilon+10 \rfloor}}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}.$$

Note that this gives geometrically decaying (r,s') sparse bounds for sufficiently small choice of $\epsilon > 0$ whenever we have

$$m < -n(1-\rho)(1/r-1/s), \qquad 1 \leqslant r \leqslant 2 \leqslant s \leqslant r'.$$

By duality and the fact that symbol classes are closed under adjoints, (r, s') sparse bounds also imply (s, r') sparse bounds, completing the proof of Theorem 1.2 at the expense of establishing the corresponding sparse bounds for the term $\sum_{l>j\epsilon} T_a^{j,l}$ (see Figure 2).
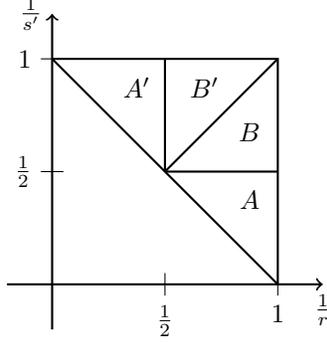

FIGURE 2. The sparse bounds in the region A corresponds to those under the constraint $1 \leqslant r \leqslant s \leqslant 2$. The ones in the region B are those for which $1 \leqslant r \leqslant 2 \leqslant s \leqslant r' \leqslant \infty$. Those in A' and B' are obtained by duality.

Dealing with the tails. We only concern ourselves with the case $1 \leqslant r \leqslant s \leqslant 2$; the case $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$ follows in an analogous manner.

For the terms $T_a^{j,l}$, we have

$$
\left\langle \sum_{j \geqslant 0} \sum_{l > j\epsilon} T_a^{j,l} f, g \right\rangle \lesssim \sum_{j \geqslant 0} \sum_{l > j\epsilon} \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q) = 2^{\lfloor -j\rho + l + 10 \rfloor}}} \| T_a^{j,l} f \|_{L^s(Q)} \| g \|_{L^{s'}(Q)}
$$

$$
\lesssim_\epsilon \sum_{j \geqslant 0} \sum_{l > j\epsilon} \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q) = 2^{\lfloor -j\rho + l + 10 \rfloor}}} |Q|^{1/r + 1/s' - 1} \| T_a^{j,l} \|_{L^r \to L^s} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}
$$

$$
\lesssim_\epsilon \sum_{j \geqslant 0} 2^{(-j\rho)n(1/r - 1/s) - jn(1/s - 1/r) + j10n(m-n)} \sum_{l > j\epsilon} 2^{ln(1/r - 1/s) + 10n(m-n)l}
$$

$$
\times \sum_{\substack{Q \text{ dyadic:} \\ \ell(Q) = 2^{\lfloor -j\rho + l + 10 \rfloor}}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}
$$

The sum in j is finite if

$$
m < n - (1 - \rho)(1/r - 1/s)/10,
$$

which trivially holds as $\rho \in [0, 1]$, $0 \leqslant 1/r - 1/s \leqslant 1$ and $m < 0$. Similarly, the sum in l is finite if

$$
10(m - n) < -(1/r - 1/s),
$$

which holds from the same considerations as above.

Obtaining true sparse bounds. Observe that for each j and l, the partition of $\mathbb{R}^n$ into dyadic cubes Q such that $\ell(Q) = 2^{\lfloor -j\rho + l + 10 \rfloor}$ is a sparse family of dyadic cubes, by disjointness of the cubes. The geometric decay in both j and l allows then to recover a true sparse bound, from the following fact about sparse bounds, which tells us that there exists a universal sparse form that controls them all. This type of result has appeared several times in the literature, and we refer the reader to, among others, the work of Lacey and Mena in [32], Lemma 4.7 and the work of Conde-Alonso and Rey in [14], Proposition 2.1.

Lemma 2.7. ([32, Lemma 4.7]). Given f, g, there is a sparse family of dyadic cubes $\mathcal{S}^0$ and a constant $C > 1$ so that for any other bilinear sparse form $\Lambda_{r,s,\mathcal{S}}$, we have
$$\Lambda_{r,s,\mathcal{S}}(f,g) \leqslant C\Lambda_{r,s,\mathcal{S}^0}(f,g).$$

For completeness, we find instructive to provide the reader with all the details, as in the case one has a geometrically decaying series of sparse bounds the obtention of the genuine sparse bounds may be done through the following simple selection process. Let $C > 1$ denote the geometrically decaying constant in our sparse bounds, so that we have

$$(6) \qquad \langle T_a f, g \rangle \lesssim \sum_{j \geqslant 0} C^{-j} \sum_{Q \text{ dyadic: } \ell(Q) = 2^{\lfloor -j\rho + j\epsilon + 10 \rfloor}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q} .$$

Since a finite union of sparse collections is sparse, it suffices to prove sparse domination for

$$(7) \qquad \sum_{j \geqslant 0 : \, j \equiv a (\text{mod } N)} C^{-j} \sum_{Q \text{ dyadic: } \ell(Q) = 2^{\lfloor -j\rho + j\epsilon + 10 \rfloor}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}$$

for some N, and for all $0 \leqslant a \leqslant N - 1$. Choose N sufficiently large so that $C^N > 4^n$ and $N > 100/\rho$ if $\rho \neq 0$, so that different j in (7) correspond to different dyadic levels of cubes; this is in contrast with the situation in (6). Fix some such a. After a re-indexing, this large choice of N allows us to rewrite (7) as

$$\sum_{j \geqslant 0} 3^{-nj} \sum_{Q \text{ dyadic: } \ell(Q) = 2^{-m(j) + j_0}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q} ,$$

where $m(j)$ is a strictly increasing sequence of integers such that $m(0) = 0$. Define our sparse collection $\mathcal{S}$ as follows. For each j, choose $\mathcal{S}_j$ to be a certain *checkerboard* collection of dyadic cubes of sidelength $2^{-m(j)+j_0}$: for each dyadic cube Q of sidelength $2^{-m(j)+j_0+j}$, choose a single dyadic cube Q' of sidelength $2^{-m(j)+j_0}$ contained in Q such that $\langle f \rangle_{r,Q'} \langle g \rangle_{s',Q'}$ is maximized. Then if we set $\mathcal{S} = \bigsqcup_j \mathcal{S}_j$, clearly

$$\sum_{j \geqslant 0} 3^{-nj} \sum_{Q \text{ dyadic: } \ell(Q) = 2^{-m(j)+j_0}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q}$$
$$\lesssim \sum_{j \geqslant 0} \sum_{Q \in \mathcal{S}_j} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q} \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{r,Q} \langle g \rangle_{s',Q} .$$

Note that $\mathcal{S}$ is clearly a Carleson collection of cubes, i.e. it is a collection of dyadic cubes such that the total measure of cubes of a given sidelength decays geometrically with the sidelength,

$$\sum_{Q \in \mathcal{S}} |Q| = \sum_{j \geqslant 0} \sum_{Q \in \mathcal{S}_j} |Q| = \sum_{j \geqslant 0} |Q_0| 2^{n[m(j) - j_0 - j]} 2^{-n[m(j) - j_0]} = \sum_{j \geqslant 0} 2^{-jn} |Q_0| \leqslant |Q_0|,$$

where Q_0 is the cube containing the supports of f and g. Therefore $\mathcal{S}$ is sparse, as it is well known that the Carleson and sparse conditions are equivalent; see [38].

3. Sharpness of the sparse bounds

Suppose towards a contradiction that a sparse bound holds outside the closure of the range given by Theorem 1.2. That is, suppose that for any symbol $a \in S_{\rho,\delta}^m$ there are some r and s such that

$$(8) \qquad\qquad m > -n(1-\rho)(1/r - 1/2), \qquad 1 \leqslant r \leqslant s \leqslant 2$$

or

$$m > -n(1-\rho)(1/r - 1/s), \qquad 1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$$

and for which a sparse (r, s') bound holds; the case for which a sparse (s', r) bound holds follows from duality. In particular, we may take this symbol $a \in S_{\rho,\delta}^m$ to be $a(x,\xi) = e^{i|\xi|^{1-\rho}}(1+|\xi|)^m$ if $1 \leqslant r \leqslant s \leqslant 2$ and $a(x,\xi) = (1+|\xi|^2)^{m/2}$ if $r \leqslant 2 \leqslant s$, since these symbols satisfy the same L^p improving boundedness properties as a general symbol for these ranges of r and s.

Consider first the case of $1 \leqslant r \leqslant s \leqslant 2$ satisfying (8). Set $m' = -j\rho n(1/r - 1/s)$ and define the symbol $b(x,\xi) = a(x,\xi)(1+|\xi|^2)^{m'/2}$. Then T_b is unbounded from L^p to L^q for all (p, q) in a sufficiently small neighbourhood of (r, s), so for sufficiently small $\epsilon' > 0$, there are infinitely many j such that

$$\|T_a^j\|_{L^r \to L^s} \geqslant 2^{j[\rho n(\frac{1}{r} - \frac{1}{s}) + \epsilon']}.$$

Let $\{j_k\}_{k=1}^\infty$ be a sequence of such j's. But by Lemma 2.5,

$$\Big\| \sum_{l > j\epsilon} T_a^{j,l} \Big\|_{L^r \to L^s} \lesssim 2^{j\rho n(\frac{1}{r} - \frac{1}{s})},$$

so reindexing the $j_k's$ to start at a sufficiently large value of j, we have for all k,

$$\Big\| \sum_{l \leqslant j_k \epsilon} T_a^{j_k,l} \Big\|_{L^r \to L^s} \gtrsim 2^{j_k[\rho n(\frac{1}{r} - \frac{1}{s}) + \epsilon']}.$$

Recall that for $1 \leqslant r \leqslant s \leqslant 2$, we considered $a(x,\xi) = e^{i|\xi|^{1-\rho}}(1+|\xi|)^m$. Integrating by parts in the ξ variable $N(\epsilon)$ times (in the "reverse" manner to Lemma 2.1, that is applying ξ derivatives to the exponential term $e^{i(x-y)\cdot\xi}$ rather than the symbol), we have

$$|T_a^{j,l}f(x)| = \Big| \int e^{i|\xi|^{1-\rho}}(1+|\xi|)^m \widetilde{\psi}(2^{-j}\xi) \int f(y)[\widetilde{\psi}(2^{-l+j\rho}(x-y))]e^{i(x-y)\cdot\xi}\, dy\, d\xi \Big|$$

$$\lesssim_\epsilon 2^{jn}(2^{j\rho}2^{l-j\rho})^{N(\epsilon)} \int |f(y)|\widetilde{\psi}(2^{-l+j\rho}(x-y))|\, dy.$$

For $N(\epsilon)$ sufficiently large,

$$\sum_{l < -j_k\epsilon} \big\| 2^{jn}(2^{j\rho}2^{l-j\rho})^{N(\epsilon)}\widetilde{\psi}(2^{-l+j\rho}\cdot) \big\|_{L^{(1+1/s-1/r)^{-1}}} \lesssim 1,$$

so by Young's convolution inequality and the triangle inequality we have

$$\Big\| \sum_{-j_k\epsilon \leqslant l \leqslant j_k\epsilon} T_a^{j_k,l} \Big\|_{L^r \to L^s} \gtrsim 2^{j_k[\rho n(\frac{1}{r} - \frac{1}{s}) + \epsilon']}.$$

By pigeonholing, it follows that there is some $l = l(k)$ with $-j_k\epsilon \leqslant l \leqslant j_k\epsilon$ such that

$$\|T_a^{j_k,l(k)}\|_{L^r \to L^s} \gtrsim 2^{j_k[\rho n(\frac{1}{r}-\frac{1}{s})+\epsilon'/2]}.$$

Moreover, by disjointness and translation invariance there is a f_k, with $\|f_k\|_{L^r} = 1$, supported in the ball centered at the origin of radius $2^{-j_k\rho+l(k)-10n}$ and such that

$$\|T_a^{j_k,l(k)}f_k\|_{L^s} \gtrsim 2^{j_k[\rho n(\frac{1}{r}-\frac{1}{s})+\epsilon'/2]}.$$

By the spatial localisation of the operator $T_a^{j_k,l(k)}$ and the support of f_k, the function $T_a^{j_k,l(k)}f_k$ is supported in a dyadic annulus at scale $2^{-j_k\rho+l(k)}$. Then, there exists a g_k supported in that annulus, with $\|g_k\|_{L^{s'}} = 1$ and such that

(9)
$$\left|\left\langle T_a^{j_k,l(k)}f_k, g_k \right\rangle\right| \gtrsim 2^{j_k[\rho n(\frac{1}{r}-\frac{1}{s})+\epsilon'/2]}.$$

Observe that the supports of f_k and g_k are disjoint, and in particular

(10)
$$\mathrm{dist}(\mathrm{supp}(f_k), \mathrm{supp}(g_k)) \approx 2^{-j_k\rho+l(k)}.$$

Now, for each k, $\sum_l T_a^{j_k,l}$ is a symbol in $S_{\rho,\delta}^m$, so by hypothesis there is a positive sparse form $\Lambda_{\mathcal{S},r,s',k}(f,g) = \sum_{Q\in\mathcal{S}} |Q| \langle f\rangle_r \langle g\rangle_{s'}$ such that

$$\Lambda_{\mathcal{S},r,s',k}(f_k, g_k) \gtrsim \left|\left\langle \sum_l T_a^{j_k,l}f_k, g_k \right\rangle\right|.$$

Then, by (9) and support considerations

$$\Lambda_{\mathcal{S},r,s',k}(f_k, g_k) \gtrsim \left|\left\langle \sum_l T_a^{j_k,l}f_k, g_k \right\rangle\right|$$
$$= \left|\left\langle T_a^{j_k,l(k)}f_k, g_k \right\rangle\right| \gtrsim 2^{j_k[\rho n(\frac{1}{r}-\frac{1}{s})+\epsilon'/2]}.$$

On the other hand, a cube contributing effectively to the sparse form $\Lambda_{\mathcal{S},r,s',k}$ must contain the supports of both f_k and g_k. By (10), the largest possible value for the sparse form will be given by a sparse family consisting of a single cube Q with smallest possible sidelength satisfying such a property; note that for $r \leqslant s$ the factor $|Q|^{1-1/r-1/s'}$ decreases with the sidelength of Q. Then

$$\Lambda_{\mathcal{S},r,s',k}(f_k, g_k) \lesssim 2^{-j_k(\rho+\epsilon)n(1/s-1/r)},$$

and so for an appropriate choice of ϵ, ϵ', we have

$$-n\rho(1/s - 1/r) < -n\rho(1/s - 1/r),$$

a contradiction.

To handle the case when $1 \leqslant r \leqslant 2 \leqslant s \leqslant r'$, the argument is similar. The only difference is that since the L^p-improving estimates are ρ-independent, we treat the symbol $(1+|\xi|^2)^{m/2}$ as belonging to the best possible symbol class that is a member of, namely $S_{1,\delta}^m$, and physically localised to scale 2^{-j}. We leave the verification of the details to the interested reader.

Sharpness along the line $v_1 - v_4$**.** We remark that a quick inspection allows one to identify certain pairs (r, s') for which sparse bounds cannot hold, as otherwise they would imply consequences which are known to be false. In particular, if $-n(1-\rho)/2 < m < 0$, the pairs (r, s') for which $(1/r, 1/s')$ lie outside the equilateral triangle formed by v_1, v_4 and v_5 in the sparse diagram (see Figure 1) are not admissible; here $v_5 := (\frac{1}{2} + \frac{-m}{n(1-\rho)}, \frac{1}{2} + \frac{-m}{n(1-\rho)})$. Indeed, by duality, it is enough to show that no sparse bounds may hold if $(1/r, 1/s')$ lie inside the triangle joining $(1/2, 1/2), (1, 1)$ and $(1, 0)$ but outside the triangle joining $(1/2, 1/2)$, v_1 and v_5.

Suppose towards a contradiction that given any $a \in S^m_{\rho,\delta}$ there exists a (r, s') sparse bound for T_a with

$$(11) \qquad \frac{1}{r} > \frac{1}{2} + \frac{-m}{n(1-\rho)}$$

and some $s \geqslant r$. By Lemma 4.1, this implies that T_a is bounded on L^p for $r < p < s$. Rearranging (11) one has

$$r < \frac{2n(1-\rho)}{n(1-\rho) - 2m},$$

so there exists $\varepsilon > 0$ such that T_a is bounded on $L^{p(\varepsilon)}$, with $p(\varepsilon) = \frac{2n(1-\rho)}{n(1-\rho)-2m} - \varepsilon$. But this is a contradiction, as it is shown in [21] that there are symbols in $S^m_{\rho,\delta}$ which fail to be bounded on L^p for $p < \frac{2n(1-\rho)}{n(1-\rho)-2m}$.

4. Consequences of sparse domination

One of the main advantages of sparse domination is that it yields weighted inequalities as a corollary. To obtain weighted estimates for sparse forms has proven to be much simpler than for many other operators, thanks to the fact that they are positive, localised and well behaved objects. We discuss some of them in the upcoming subsections.

4.1. **One-weight estimates.** Let w denote a weight, that is, a nonnegative locally integrable function. Recall that given $1 < p < \infty$, the Muckenhoupt classes of weights A_p and the reverse Hölder classes of weights RH_p consist of all w satisfying

$$[w]_{A_p} := \sup_B \langle w \rangle_B \langle w^{1-p'} \rangle_B^{p-1} < \infty, \qquad [w]_{RH_p} := \sup_B \langle w \rangle_B^{-1} \langle w \rangle_{B,p} < \infty$$

respectively, where the supremum ranges over all balls B in $\mathbb{R}^n$. The class A_1 consists of all w satisfying $Mw \leqslant Cw$ for some finite constant $C > 0$; here M denotes the Hardy–Littlewood maximal function. The infimum of such admissible constants C is denoted by $[w]_{A_1}$.

Quantitative weighted estimates in the context of the above weights may be deduced for operators dominated by bilinear sparse forms as a consequence of the following lemma.

Lemma 4.1 ([5]). *Suppose that we are given* $r, s \in [1, \infty]$ *and a sparse collection* $\mathcal{S}$ *of dyadic cubes. Then for any* $r < p < s$ *there exists a constant* $C = C(r, s, p)$ *so that for every weight* $w \in A_{p/r} \cap RH_{(s/p)'}$,

$$\Lambda_{\mathcal{S}, r, s'}(f, g) \leqslant C \big([w]_{A_{p/r}} [w]_{RH_{(s/p)'}} \big)^\alpha \|f\|_{L^p(w)} \|g\|_{L^{p'}(w^{1-p'})},$$

where $\alpha := \max \left(\frac{1}{p-r}, \frac{s-1}{s-p} \right)$.

This lemma is sharp at the level of sparse bounds and provides the best possible dependence on the A_p characteristic of the weight. The intersection of the A_p and the reverse Hölder classes may be understood as

$$(12) \qquad w \in A_{p/r} \cap RH_{(s/p)'} \iff w^{(s/p)'} \in A_{(s/p)'(p/r-1)+1}.$$

Of course Theorem 1.2 and Lemma 4.1 allows one to deduce weighted estimates for pseudodifferential operators associated to the symbol classes $S_{\rho,\delta}^m$. In views of the reverse Hölder properties of weights, one may effectively use a sparse bound at the closure of the trapezoid $\mathcal{T}_{m,\rho,n}$ in order to deduce weighted estimates. In particular, using the endpoint v_3, one has the following.

Corollary 4.2. *Let $a \in S_{\rho,\delta}^m$ for some $m < 0$ and $0 < \delta \leqslant \rho < 1$.*

(i) *Let $m = -n(1 - \rho)$. Then T_a is bounded on $L^p(w)$ for $w \in A_p$ and all $1 < p < \infty$.*

(ii) *Let $-n(1-\rho) < m \leqslant -n(1-\rho)/2$. Then T_a is bounded on $L^p(w)$ for $w \in A_{p/r_e}$ and all $r_e < p < \infty$, where $r_e := -n(1 - \rho)/m$.*

(iii) *Let $-n(1 - \rho)/2 < m < 0$. Then T_a is bounded on $L^p(w)$ for $w \in A_{p/2} \cap RH_{(s_e/p)'}$ and all $2 < p < s_e$, where $s_e := \frac{2n(1-\rho)}{n(1-\rho)+2m}$.*

This corollary recovers all the previous known weighted estimates for pseudodifferential operators in the context of Muckenhoupt weights mentioned in the Introduction, except for the cases $p = r_e$ in (ii) when $m = -n(1 - \rho)/2$ and $p = 2$ in (iii); see [41] for (i), [11, 40] for (ii) and [1] for (iii). We believe that the estimates (ii) for $-n(1 - \rho) < m < -n(1 - \rho)/2$ are new. One should note that the estimates in Corollary 4.2 only contain some of the possible weighted estimates; in particular other versions for (ii) and (iii) may be obtained by using other sparse bounds in the admissible region $\mathcal{T}_{m,\rho,n}$.

We also note that one could state a version of Corollary 4.2 with an explicit quantitative dependence of the operator norm in terms of the $[w]_{A_p}$ and $[w]_{RH_p}$ constants. Because of the use of the reverse Hölder inequality in the proof of the corollary, the behaviour of the weighted constant would not just be that of Lemma 4.1, but would also depend on how the implicit constant in the sparse domination in Theorem 1.2 depends on r and s. The resulting quantitative weighted estimate would probably be far from being sharp and we do not concern with such finer points in this paper.

Proof of Corollary 4.2. We shall only prove (ii) and (iii), as (i) is a particular case of (ii) when $m = -n(1 - \rho)$. Let $a \in S_{\rho,\delta}^m$ with $-n(1 - \rho) \leqslant m \leqslant -n(1 - \rho)/2$, $w \in A_{p/r_e}$ and $r_e < p < \infty$. By the reverse Hölder property of weights, there exists $\varepsilon := \varepsilon(n, p, [w]_{A_\infty}) > 0$ such that $w \in A_{p/r_e - \delta}$ for all $0 \leqslant \delta \leqslant \varepsilon$; see for instance [26] for its sharp version. In particular, choose $\delta < \min\{\varepsilon, (p - r_e)/2r_e\}$ and set $r := \frac{p}{p/r_e - \delta}$. Then $w \in A_{p/r}$ and $r_e < r < p$. As $(1/r, 1) \in \mathcal{T}_{m,\rho,n}$, T_a is dominated by a $(r, 1)$ sparse form by Theorem 1.2, and therefore bounded on $L^p(w)$ by Lemma 4.1. This establishes (ii).

Similarly for a given $a \in S_{\rho,\delta}^m$, with $-n(1-\rho)/2 < m < 0$, let $w \in A_{p/2} \cap RH_{(s_e/p)'}$ and $r < p < s_e$. By the characterisation (12), $w^{(s_e/p)'} \in A_{(s/p)'(p/2-1)+1}$. Using the reverse Hölder property of weights [25], there exists $\varepsilon := \varepsilon(n, [w]_{A_\infty}) > 0$ such that

$$\langle w^{(s_e/p)'} \rangle_{1+\delta,Q} \leqslant 2 \langle w^{(s_e/p)'} \rangle_{1,Q}$$

for any cube Q and $0 \leqslant \delta \leqslant \varepsilon$. Thus

$$\langle w \rangle_{(s_e/p)'(1+\delta),Q} \leqslant 2 \langle w \rangle_{(s_e/p)',Q} \lesssim \langle w \rangle_{1,Q},$$

where the last inequality follows as $w \in RH_{(s_e/p)'}$. In particular, we have that $w \in RH_{(s_e/p)'(1+\delta)}$ for $0 \leqslant \delta \leqslant \varepsilon$. It is then a computation to find $p < s < s_e$ such that $w \in A_{p/2} \cap RH_{(s/p)'}$, with $(1/2, 1/s')$ belonging to $\mathcal{T}_{m,\rho,n}$. It follows then from Theorem 1.2 that T_a is dominated by a $(2, s')$ sparse form, so by Lemma 4.1 T_a is bounded on $L^p(w)$, which concludes the proof of (iii). $\qquad \square$

4.2. Fefferman–Stein inequalities.
It is also possible to deduce Fefferman–Stein weighted inequalities as a consequence of the sparse domination. That is, given $1 \leqslant p < \infty$, to identify an operator $\mathcal{M}$ such that

$$(13) \qquad \int_{\mathbb{R}^n} |Tf|^p w \leqslant C \int_{\mathbb{R}^n} |f|^p \mathcal{M}w$$

holds for any weight w, with constant independent of the weight.

Lemma 4.1 immediately yields this type of estimates for operators which are sparse dominated. This follows from the trivial observation that $(M_\gamma w)^{(s/p)'} \in A_1$ if $\gamma > (s/p)'$, with constant independent of w; here $M_\gamma w := (Mw^\gamma)^{1/\gamma}$. As $w \leqslant M_\gamma w$, an operator T dominated by a (r, s') sparse form satisfies then the inequality (13) with $\mathcal{M} = M_\gamma$ for $\gamma > (s/p)'$ and $r < p < s$.

In the case of $(r, 1)$-sparse forms, where the condition on γ becomes $\gamma > 1$, the first author observed in [2] that M_γ may be improved to the controlling maximal function $\mathcal{M} = M^{\lfloor p \rfloor + 1}$; here $M^{\lfloor p \rfloor + 1}$ denotes the $(\lfloor p \rfloor + 1)$-fold composition of M with itself. In this case, the argument is more intrinsic and does not follow from Lemma 4.1. The question whether M_γ may be improved for general (r, s) sparse forms remains open.

Of course the above discussion yields Fefferman–Stein inequalities for pseudo-differential operators through the sparse bounds in Theorem 1.2. We remark that the first author proved in [1] sharp versions of these inequalities for $p = 2$ and for any given symbol class $S_{\rho,\delta}^m$, with the controlling operator $\mathcal{M}$ being a fractional maximal function associated to non-tangential approach regions. The proof of that result differs very much from the techniques presented in this paper. It would be interesting to explore a further connection between the two; in particular whether if sparse domination techniques could lead to Fefferman–Stein inequalities with operators $\mathcal{M}$ falling beyond the classical A_p theory, in a way that allows one to recover the results in [1].

4.3. Coifman–Fefferman estimates.
It has recently been established in [39] that if an operator is dominated by a $(1, s)$ sparse form for any $s > 1$, then a Coifman–Fefferman type estimate

$$\|Tf\|_{L^p(w)} \lesssim \|Mf\|_{L^p(w)}.$$

holds for any $w \in A_\infty$ and $1 \leqslant p < \infty$, with implicit constant depending on $[w]_{A_\infty}$. This applies then to pseudodifferential operators associated to the symbol classes $S_{\rho,\delta}^{-n(1-\rho)}$. We believe this is the first time that these sorts of estimates are obtained for pseudodifferential operators.

4.4. **Weak-type estimates.** Theorem 1.2 also allows to recover that the symbol classes $S_{\rho,\delta}^m$ with $m < -n(1-\rho)/2$ and $0 \leqslant \delta \leqslant \rho < 1$ are of weak-type $(1,1)$. This is a consequence of the following result for sparse forms.

Lemma 4.3 ([13])**.** *Suppose that a sublinear operator T has the following property: there exists $C > 0$ and $1 \leqslant r < \infty$ such that for every f, g bounded with compact support there exists a sparse collection $\mathcal{S}$ such that*

$$|\langle Tf, g \rangle| \leqslant C\Lambda_{\mathcal{S},1,r}(f,g).$$

Then $T : L^1(\mathbb{R}^n) \to L^{1,\infty}(\mathbb{R}^n)$ boundedly.

5. Oscillatory Fourier multipliers

Theorem 1.2 and the techniques presented in Section 2 naturally extend for certain classes of oscillatory multipliers. In this section, given $\alpha, \beta \in \mathbb{R}$, let m denote a Fourier multiplier on $\mathbb{R}^n$, with support in the set $\{\xi \in \mathbb{R}^d : |\xi|^\alpha \geqslant 1\}$, and satisfying the Miyachi condition

$$(14) \qquad\qquad |D^\gamma m(\xi)| \lesssim |\xi|^{-\beta + |\gamma|(\alpha - 1)}$$

for every multi-index $\gamma \in \mathbb{N}^n$.

Of course, when $0 < \alpha \leqslant 1$, the differential inequalities (14) are essentially equivalent to those defining the symbol classes $S_{1-\alpha,0}^{-\beta}$. The reason to have the support condition $\{|\xi|^\alpha \geqslant 1\}$ and the factor $|\xi|$ in (14) instead of $(1 + |\xi|)$ is for our treatment to hold also when $\alpha < 0$. In that case, the relevant behaviour of the oscillatory multipliers is determined by low frequencies. Interestingly, when $\alpha \notin [0,1]$, differentiating the multiplier implies a growth in its order; this is in contrast with the case $0 \leqslant \alpha < 1$, where differentiating implies decay. Model examples for the above Fourier multipliers are given by $m_{\alpha,\beta}(\xi) := |\xi|^{-\beta} e^{i|\xi|^\alpha} \chi_{\{|\xi|^\alpha \geqslant 1\}}(\xi)$, first studied by Hirschman [23], and later by Wainger [47], Fefferman [20], Fefferman and Stein [22], Miyachi [42, 43] and others. They satisfy the following sparse bounds.

Theorem 5.1. *Let $\alpha, \beta \in \mathbb{R}$, $\alpha \neq 0$. Let m be a multiplier satisfying* (14). *Then for any compactly supported bounded functions f, g on $\mathbb{R}^n$, there exists sparse collections $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ of dyadic cubes such that*

$$|\langle Tf, g \rangle| \leqslant C(\alpha, \beta, r, s)\Lambda_{\mathcal{S},r,s'}(f,g)$$

and

$$|\langle Tf, g \rangle| \leqslant C(\alpha, \beta, r, s)\Lambda_{\widetilde{\mathcal{S}},s',r}(f,g)$$

for all pairs (r, s') and (s', r) such that $\alpha \cdot \beta > 0$ and

$$|\beta| > n|\alpha|(1/r - 1/2), \qquad 1 \leqslant r \leqslant s \leqslant 2$$

or

$$|\beta| > n|\alpha|(1/r - 1/s), \qquad 1 \leqslant r \leqslant 2 \leqslant s \leqslant r'.$$

The missing case $\alpha = 0$, $\beta \neq 0$ in the above theorem corresponds to the classical case of fractional integrals, for which *fractional* sparse bounds were obtained by Cruz-Uribe and Moen in [15], and the case $\alpha = \beta = 0$ corresponds to the classical Mikhlin multipliers; see for instance [6] for sparse bounds. To approach the $\alpha = 0$ case with our methods, one needs to consider both low and high frequencies. It is clear that the resulting conditions on α and β coming from each case would be excluding one another.

For the case $\alpha > 0$, the proof is identical to that provided in Section 2. Indeed, as now we are in a translation-invariant setting, the L^2 bound is just the L^∞ norm of the multiplier in contrast with the more involved L^2 theory for pseudodifferential operators.

For the case $\alpha < 0$, one needs to perform a decomposition of the region $\{|\xi| \leqslant 1\}$ into dyadic annuli of width 2^j for $j < 0$. As the sum runs over $j < 0$, we require the reverse condition on the exponent of 2^j in order to sum the geometric series. Also, note that for $\alpha < 0$ we require $\beta < 0$, so for the summability on l one requires

$$\|T_a^{j,l}\|_{L^r \to L^s} \lesssim 2^{-10n(\beta-n)(j-l)}$$

for all $l > |j|\epsilon$.

As it is well known, results for the above classes of multipliers imply as corollaries bounds on oscillatory kernels and the solution operator to certain dispersive equations. Of course, Theorem 5.1 also yields weighted corollaries for oscillatory Fourier multipliers in the very same fashion as is presented in Section 4; the details are left to the interested reader.

5.1. Oscillatory kernels.
For $a > 0$, $a \neq 1$ and $b \geqslant n(1 - \frac{a}{2})$, let $K_{a,b} : \mathbb{R}^n \to \mathbb{C}$ be given by

$$K_{a,b}(x) = \frac{e^{i|x|^a}}{(1+|x|)^b}.$$

These convolution kernels satisfy the sparse bounds in Theorem 5.1 with $\alpha = \frac{a}{a-1}$ and $\beta = \frac{na/2-n+b}{a-1}$. To see this, write $K_{a,b} = K_{a,b}^0 + K_{a,b}^\infty$, where $K_{a,b}^0 = K_{a,b}\eta$ for some $\eta \in C_c^\infty$. As $|K_{a,b}^0| \lesssim \eta$, it satisfies all possible sparse bounds. For $K_{a,b}^\infty$, we use the well known-fact (see, for instance, [45, 46]) that its Fourier transform satisfies the condition (14) for $\alpha = \frac{a}{a-1}$ and $\beta = \frac{na/2-n+b}{a-1}$. We further break the multiplier in two parts, $\widehat{K_{a,b}^\infty} = m_0 + m_\infty$, so that m_∞ guarantees the support condition $\{|\xi|^\alpha > 1\}$. As m_0 is a rapidly decreasing function, so it is its associated kernel, and therefore also satisfies all possible sparse bounds. Finally, for m_∞, apply Theorem 5.1, which determines the sparse bounds for the original convolution kernel $K_{a,b}$.

We note that our results exclude the case $a = 1$, as the associated Fourier multipliers are not *oscillatory*. Indeed when $a = 1$, $b = (n+1)/2+\delta$, the kernels $K_{a,b}$ are the convolution kernels associated to the Bochner–Riesz multipliers m_δ. Sparse bounds for Bochner–Riesz multipliers have been recently obtained by Lacey, Mena and Reguera in [33]; see also the endpoint result of Kesler and Lacey [27] or the previous work of Benea, Bernicot and Luque [4]. Besides the weighted A_p estimates obtained as a consequence of the sparse domination, one should note that there are weighted Fefferman–Stein inequalities for Bochner–Riesz multipliers involving Kakeya-type maximal functions; see for instance the earlier work of Carbery [8] or Carbery and Seeger [10].

5.2. Dispersive equations.
Given $\alpha \in \mathbb{N}$, the solution of the dispersive equation

$$\begin{cases} i\partial_t u + (-\Delta)^{\alpha/2}u = 0 \\ u(\cdot,0) = f \end{cases}$$

is given by

$$u(x,t) = e^{it(-\Delta)^{\alpha/2}}f(x) := \int_{\mathbb{R}^n} e^{ix\cdot\xi}e^{it|\xi|^\alpha}\widehat{f}(\xi)d\xi.$$

As the Fourier multiplier associated to the x-variable is $m(\xi) = e^{it|\xi|^\alpha}$, a rescaling argument provides sparse bounds for $u(x, t)$ for a fixed time t. For $\alpha \in \mathbb{N}$, and under the hypothesis of Theorem 5.1 one has that there exists a sparse family $\mathcal{S}$ such that

$$|\langle u(\cdot, t), g \rangle| \leqslant C(m, \rho, r, s) \sum_{Q \in \mathcal{S}} |t^{1/\alpha}Q| \left\langle (I - t^{2/\alpha}\Delta)^{\beta/2}f \right\rangle_{r, t^{1/\alpha}Q} \langle g \rangle_{s, t^{1/\alpha}Q},$$

where $t^{1/\alpha}Q$ denotes the concentric cube to Q of volume $|t^{1/\alpha}Q| = t^{n/\alpha}|Q|$ and

$$(I - t^{2/\alpha}\Delta)^{\beta/2}f(x) := \int_{\mathbb{R}^n} e^{ix\cdot\xi}(1 + t^{2/\alpha}|\xi|^2)^{\beta/2}\widehat{f}(\xi)d\xi$$

defines a $t-$inhomogeneous Sobolev norm.

5.3. **A subdyadic Hörmander condition.** Theorem 5.1 also holds for broader classes of multipliers than those given by the Mikhlin-type condition (14). In particular, it is enough to ask that the multiplier m, with support in $\{\xi \in \mathbb{R}^n : |\xi|^\alpha \geqslant 1\}$, satisfies the Hörmander-type condition [4]

$$(15) \qquad \sup_B \, \mathrm{dist}(B, 0)^{\beta + (1-\alpha)|\gamma|} \left(\frac{1}{|B|} \int_B |D^\gamma m(\xi)|^2 d\xi \right)^{1/2} < \infty$$

for all $\gamma \in \mathbb{N}^n$. Here the supremum is taken over all euclidean balls B in $\mathbb{R}^n$ with $\mathrm{dist}(B, 0)^\alpha \geqslant 1$ such that

$$r(B) \sim \mathrm{dist}(B, 0)^{1-\alpha},$$

where $r(B)$ denotes the radius of B. The proof of Theorem 5.1 under these weaker assumptions is conceptually not harder, and only requires a reelaboration of the $L^r - L^s$ bounds for T_m^j and the decay bounds for $T_m^{j,l}$. This may be done using similar ideas to those in the proof of Lemma 2.3. The details are left to the interested reader.

The above *subdyadic* Hörmander condition was introduced by Bennett and the first author in [3], to which we refer for further discussion. Unlike in there, or in the classical Mikhlin–Hörmander theorem, our methods seem to require the conditions (14) and (15) to hold for all $\gamma \in \mathbb{N}^n$ and not only for those $|\gamma| \leqslant \lfloor \frac{n}{2} \rfloor + 1$. It would be interesting to determine if Theorem 5.1 holds under these weaker assumptions in the number of derivatives.

APPENDIX A. A POINTWISE SPARSE BOUND FOR THE CLASSES $S_{\rho,\delta}^{-n(1-\rho)}$

Given a sublinear operator T, let $\mathcal{M}_T$ be its *grand maximal function*, defined by

$$\mathcal{M}_T f(x) := \sup_{Q \ni x} \underset{z \in Q}{\mathrm{ess\,sup}} |T(f\chi_{\mathbb{R}^n \setminus 3Q})(z)|;$$

here the supremum is taken over all cubes $Q \subset \mathbb{R}^n$ containing x. In [37], Lerner proved the following abstract theorem for pointwise sparse domination, which relies on the boundedness of the operator $\mathcal{M}_T$.

Theorem A.1 ([37]). *Assume that T is a sublinear operator of weak type (q, q) and $\mathcal{M}_T$ is of weak type (r, r), where $1 \leqslant q \leqslant r < \infty$. Then, for every compactly supported $f \in L^r(\mathbb{R}^n)$, there exists a sparse family $\mathcal{S}$ such that for a.e. $x \in \mathbb{R}^n$,*

$$|Tf(x)| \leqslant C\mathcal{A}_{r,\mathcal{S}}f(x),$$

---

[4]Our results may also be seen to hold under a weaker Hörmander–Sobolev formulation, although we refrain from doing that for simplicity.

where $C = C(n, q, r)(\|T\|_{L^q \to L^{q,\infty}} + \|\mathcal{M}_T\|_{L^r \to L^{r,\infty}})$.

As is well known that T_a is of weak type $(1,1)$ for $a \in S_{\rho,\delta}^{-n(1-\rho)}$, the above Theorem of Lerner reduces the proof of Theorem 1.3 to verifying that $\mathcal{M}_{T_a}$ is of weak type (r, r) for any $r > 1$. The proof of this fact is quite standard, and is based on work of Chanillo and Torchinksy [11], and Michalowski, Rule and Staubach [41].

Proof of Theorem 1.3. Given a point x and a cube $Q \ni x$, we distinguish two cases, $|Q| \leqslant 1$ and $|Q| > 1$. The latter case is easy to deal with, and using decay estimates of the associated kernel to T_a, it is elementary to see that

$$(16) \qquad |T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(z)| \lesssim Mf(x)$$

for any $z \in Q$; see, for instance [41].

If $|Q| \leqslant 1$, a slightly more refined analysis is required. For any $z, x' \in Q$,

$$|T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(z)| \leqslant |T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(z) - T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(x')|$$
$$(17) \qquad\qquad\qquad + |T_a(f)(x')| + |T_a(f\chi_{3Q})(x')|.$$

In [41] (see also [11]), it is proven that

$$|T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(z) - T_a(f\chi_{\mathbb{R}^n \setminus 3Q})(x')| \lesssim M_p f(x)$$

for any $p := p(\rho) > 1$ sufficiently close to 1 and $Q \ni x$ with $|Q| \leqslant 1$. Raising the estimate (17) to a power $1 < s < r$, integrating with respect to $x' \in Q$, and raising it again to the power $1/s$,

$$|T_a(f\chi_{\mathbb{R} \setminus 3Q})(z)| \lesssim M_p f(x) + \left(\frac{1}{|Q|}\int_Q |T_a f(x')|^s dx'\right)^{1/s} + \left(\frac{1}{|Q|}\int_Q |T_a(f\chi_{3Q})(x')|^s dx'\right)^{1/s}$$

$$\lesssim M_p f(x) + M_s(T_a f)(x) + \|T_a\|_s \left(\frac{1}{|Q|}\int_{3Q} |f(x')|^s dx'\right)^{1/s}$$

$$(18) \qquad \lesssim M_p f(x) + M_s(T_a f)(x) + \|T_a\|_s M_s f(x),$$

after taking supremum over all $Q \ni x$ and using the boundedness of T_a on L^s. Combining (16) and (18)

$$\mathcal{M}_{T_a} f(x) \lesssim Mf(x) + M_p f(x) + M_s(T_a f)(x) + \|T_a\|_s M_s f(x).$$

Taking L^r-norms, with $r > \max(p, s)$,

$$\|\mathcal{M}_{T_a} f\|_r \lesssim \left(\|M\|_r + \|M\|_{r/p} + \|M\|_{r/s}\|T_a\|_r + \|T_a\|_s \|M\|_{r/s}\right)\|f\|_r.$$

As p and s may be chosen arbitrarily close to 1, $\mathcal{M}_{T_a}$ is bounded on L^r for any $r > 1$. Thus, an application of Theorem A.1 yields

$$|T_a f(x)| \lesssim A_{r,\mathcal{S}} f(x),$$

as required. □

REFERENCES

[1] D. Beltran. Control of pseudodifferential operators by maximal functions via weighted inequalities. To appear in *Trans. Amer. Math. Soc.* Preprint available at arXiv:1608.06571.

[2] D. Beltran. A Fefferman–Stein inequality for the Carleson operator. *Rev. Mat. Iberoam.*, 34(1):221–244, 2018.

[3] D. Beltran and J. Bennett. Subdyadic square functions and applications to weighted harmonic analysis. *Adv. Math.*, 307:72–99, 2017.

[4] C. Benea, F. Bernicot, and T. Luque. Sparse bilinear forms for Bochner Riesz multipliers and applications. *Trans. London Math. Soc.*, 4(1):110–128, 2017.

[5] F. Bernicot, D. Frey, and S. Petermichl. Sharp weighted norm estimates beyond Calderón-Zygmund theory. *Anal. PDE*, 9(5):1079–1113, 2016.

[6] T. A. Bui, J. M. Conde-Alonso, X. T. Duong, and M. Hormozi. A note on weighted bounds for singular operators with nonsmooth kernels. *Studia Math.*, 236(3):245–269, 2017.

[7] A.-P. Calderón and R. Vaillancourt. A class of bounded pseudo-differential operators. *Proc. Nat. Acad. Sci. U.S.A.*, 69:1185–1187, 1972.

[8] A. Carbery. A weighted inequality for the maximal Bochner-Riesz operator on $\mathbf{R}^2$. *Trans. Amer. Math. Soc.*, 287(2):673–680, 1985.

[9] A. Carbery and A. Seeger. Conditionally convergent series of linear operators on L^p-spaces and L^p-estimates for pseudodifferential operators. *Proc. London Math. Soc. (3)*, 57(3):481–510, 1988.

[10] A. Carbery and A. Seeger. Weighted inequalities for Bochner-Riesz means in the plane. *Q. J. Math.*, 51(2):155–167, 2000.

[11] S. Chanillo and A. Torchinsky. Sharp function and weighted L^p estimates for a class of pseudodifferential operators. *Ark. Mat.*, 24(1):1–25, 1986.

[12] L. Cladek and Y. Ou. Sparse domination of Hilbert transforms along curves. To appear in *Math. Res. Lett.* Preprint available at arXiv:1704.07810.

[13] J. M. Conde-Alonso, A. Culiuc, F. Di Plinio, and Y. Ou. A sparse domination principle for rough singular integrals. *Anal. PDE*, 10(5):1255–1284, 2017.

[14] J. M. Conde-Alonso and G. Rey. A pointwise estimate for positive dyadic shifts and some applications. *Math. Ann.*, 365(3-4):1111–1135, 2016.

[15] D. Cruz-Uribe and K. Moen. One and two weight norm inequalities for Riesz potentials. *Illinois J. Math.*, 57(1):295–323, 2013.

[16] A. Culiuc, F. Di Plinio, and Y. Ou. Domination of multilinear singular integrals by positive sparse forms. Preprint available at arXiv:1603.05317.

[17] A. Culiuc, F. Di Plinio, and Y. Ou. Uniform sparse domination of singular integrals via dyadic shifts. To appear in *Math. Res. Lett.* Preprint available at arXiv:1610.01958.

[18] A. Culiuc, R. Kesler, and M. Lacey. Sparse bounds for the discrete cubic Hilbert transform. Preprint available at arXiv:1612.08881.

[19] F. Di Plinio, Y. Do, and G. Uraltsev. Positive sparse domination of variational Carleson operators. To appear in *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*. Preprint available at arXiv:1612.03028.

[20] C. Fefferman. Inequalities for strongly singular convolution operators. *Acta Math.*, 124:9–36, 1970.

[21] C. Fefferman. L^p bounds for pseudo-differential operators. *Israel J. Math.*, 14:413–417, 1973.

[22] C. Fefferman and E. M. Stein. H^p spaces of several variables. *Acta Math.*, 129(3-4):137–193, 1972.

[23] I. I. Hirschman, Jr. On multiplier transformations. *Duke Math. J*, 26:221–242, 1959.

[24] L. Hörmander. Pseudo-differential operators and hypoelliptic equations. In *Singular integrals (Proc. Sympos. Pure Math., Vol. X, Chicago, Ill., 1966)*, pages 138–183. Amer. Math. Soc., Providence, R.I., 1967.

[25] T. Hytönen and C. Pérez. Sharp weighted bounds involving A_∞. *Anal. PDE*, 6(4):777–818, 2013.

[26] T. Hytönen, C. Pérez, and E. Rela. Sharp reverse Hölder property for A_∞ weights on spaces of homogeneous type. *J. Funct. Anal.*, 263(12):3883–3899, 2012.

[27] R. Kesler and M. Lacey. Sparse endpoint estimates for Bochner-Riesz multipliers on the plane. Preprint available at arXiv:1707.05844.

[28] B. Krause and M. Lacey. Sparse bounds for maximally truncated oscillatory singular integrals. Preprint available at arXiv:1701.05249.

[29] B. Krause and M. Lacey. Sparse bounds for random discrete Carleson theorems. Preprint available at arXiv:1609.08701.

[30] M. Lacey. Sparse bounds for spherical maximal functions. To appear in J. Anal. Math. Preprint available at arXiv:1702.08594.

[31] M. T. Lacey. An elementary proof of the A_2 bound. *Israel J. Math.*, 217(1):181–195, 2017.

[32] M. T. Lacey and D. Mena. The sparse T1 theorem. *Houston J. Math.*, 43(1):111–127, 2017.

[33] M. T. Lacey, D. Mena, and M. C. Reguera. Sparse bounds for Bochner-Riesz multipliers. Preprint available at arXiv:1705.09375.

[34] M. T. Lacey and S. Spencer. Sparse bounds for oscillatory and random singular integrals. *New York J. Math.*, 23:119–131, 2017.

[35] A. K. Lerner. On an estimate of Calderón-Zygmund operators by dyadic positive operators. *J. Anal. Math.*, 121:141–161, 2013.

[36] A. K. Lerner. A simple proof of the A_2 conjecture. *Int. Math. Res. Not. IMRN*, (14):3159–3170, 2013.

[37] A. K. Lerner. On pointwise estimates involving sparse operators. *New York J. Math.*, 22:341–349, 2016.

[38] A. K. Lerner and F. Nazarov. Intuitive dyadic calculus: the basics. Preprint available at arXiv:1508.05639.

[39] K. Li, C. Pérez, I. P. Rivera-Ríos, and L. Roncal. Weighted norm inequalities for rough singular integral operators. Preprint available at arXiv:1701.05170.

[40] N. Michalowski, D. J. Rule, and W. Staubach. Weighted norm inequalities for pseudo-pseudodifferential operators defined by amplitudes. *J. Funct. Anal.*, 258(12):4183–4209, 2010.

[41] N. Michalowski, D. J. Rule, and W. Staubach. Weighted L^p boundedness of pseudodifferential operators and applications. *Canad. Math. Bull.*, 55(3):555–570, 2012.

[42] A. Miyachi. On some Fourier multipliers for $H^p(\mathbf{R}^n)$. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 27(1):157–179, 1980.

[43] A. Miyachi. On some singular Fourier multipliers. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 28(2):267–315, 1981.

[44] R. Oberlin. Sparse bounds for a prototypical singular Radon transform. Preprint available at arXiv:1704.04297.

[45] P. Sjölin. Convolution with oscillating kernels. *Indiana Univ. Math. J.*, 30(1):47–55, 1981.

[46] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[47] S. Wainger. Special trigonometric series in k-dimensions. *Mem. Amer. Math. Soc. No.*, 59:102, 1965.

DAVID BELTRAN: BASQUE CENTER FOR APPLIED MATHEMATICS (BCAM), ALAMEDA DE MAZARREDO 14, 48009 BILBAO, SPAIN
    E-mail address: `dbeltran@bcamath.org`

LAURA CLADEK: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095
    E-mail address: `cladekl@math.ucla.edu`