ON POINTWISE AND WEIGHTED ESTIMATES FOR COMMUTATORS OF CALDERÓN-ZYGMUND OPERATORS

ANDREI K. LERNER, SHELDY OMBROSI, AND ISRAEL P. RIVERA-RÍOS

Abstract. In recent years, it has been well understood that a Calderón-Zygmund operator T is pointwise controlled by a finite number of dyadic operators of a very simple structure (called the sparse operators). We obtain a similar pointwise estimate for the commutator $[b, T]$ with a locally integrable function b. This result is applied into two directions. If $b \in BMO$, we improve several weighted weak type bounds for $[b, T]$. If b belongs to the weighted BMO, we obtain a quantitative form of the two-weighted bound for $[b, T]$ due to Bloom-Holmes-Lacey-Wick.

1. Introduction

1.1. **A pointwise bound for commutators.** In the past decade, a question about sharp weighted inequalities has leaded to a much better understanding of classical Calderón-Zygmund operators. In particular, it was recently discovered by several authors (see [5, 19, 21, 24, 25], and also [1, 8] for some interesting developments) that a Calderón-Zygmund operator is dominated pointwise by a finite number of sparse operators $\mathcal{A}_{\mathcal{S}}$ defined by

$$\mathcal{A}_{\mathcal{S}} f(x) = \sum_{Q \in \mathcal{S}} f_Q \chi_Q(x),$$

where $f_Q = \frac{1}{|Q|} \int_Q f$ and $\mathcal{S}$ is a sparse family of cubes from $\mathbb{R}^n$ (the latter means that each cube $Q \in \mathcal{S}$ contains a set E_Q of comparable measure and the sets $\{E_Q\}_{Q \in \mathcal{S}}$ are pairwise disjoint).

In this paper we obtain a similar domination result for the commutator $[b, T]$ of a Calderón-Zygmund operator T with a locally integrable

2010 *Mathematics Subject Classification.* 42B20, 42B25.

Key words and phrases. Commutators, Calderón-Zygmund operators, Sparse operators, weighted inequalities.

The first author was supported by the Israel Science Foundation (grant No. 953/13). The third author was supported by Grant MTM2012-30748, Spanish Government.

function b, defined by

$$[b,T]f(x) = bTf(x) - T(bf)(x).$$

Then we apply this result in order to derive several new weighted weak and strong type inequalities for $[b,T]$.

Throughout the paper, we shall deal with ω-Calderón-Zygmund operators T on $\mathbb{R}^n$. By this we mean that T is L^2 bounded, represented as

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y)f(y)dy \quad \text{for all } x \notin \operatorname{supp} f,$$

with kernel K satisfying the size condition $|K(x,y)| \le \frac{C_K}{|x-y|^n}$, $x \neq y$, and the smoothness condition

$$|K(x,y) - K(x',y)| + |K(y,x) - K(y,x')| \le \omega\left(\frac{|x-x'|}{|x-y|}\right)\frac{1}{|x-y|^n}$$

for $|x-y| > 2|x-x'|$, where $\omega : [0,1] \to [0,\infty)$ is continuous, increasing, subadditive and $\omega(0) = 0$.

In [21], M. Lacey established a pointwise bound by sparse operators for ω-Calderón-Zygmund operators with ω satisfying the Dini condition $[\omega]_{\text{Dini}} = \int_0^1 \omega(t)\frac{dt}{t} < \infty$. For such operators we adopt the notation

$$C_T = \|T\|_{L^2 \to L^2} + C_K + [\omega]_{\text{Dini}}.$$

A quantitative version of Lacey's result due to T. Hytönen, L. Roncal and O. Tapiola [19] states that

$$|Tf(x)| \le c_n C_T \sum_{j=1}^{3^n} \mathcal{A}_{\mathcal{S}_j}|f|(x). \tag{1.1}$$

An alternative proof of this result was obtained by the first author [24].

In order to state an analogue of (1.1) for commutators, we introduce the sparse operator $\mathcal{T}_{\mathcal{S},b}$ defined by

$$\mathcal{T}_{\mathcal{S},b}f(x) = \sum_{Q\in\mathcal{S}} |b(x) - b_Q| f_Q \chi_Q(x).$$

Let $\mathcal{T}^{\star}_{\mathcal{S},b}$ denote the adjoint operator to $\mathcal{T}_{\mathcal{S},b}$:

$$\mathcal{T}^{\star}_{\mathcal{S},b}f(x) = \sum_{Q\in\mathcal{S}} \left(\frac{1}{|Q|}\int_Q |b - b_Q| f\right)\chi_Q(x).$$

Our first main result is the following. Its proof is based on ideas developed in [24].

Theorem 1.1. *Let T be an ω-Calderón-Zygmund operator with ω satisfying the Dini condition, and let $b \in L^1_{loc}$. For every compactly supported $f \in L^\infty(\mathbb{R}^n)$, there exist 3^n dyadic lattices $\mathscr{D}^{(j)}$ and $\frac{1}{2\cdot 9^n}$-sparse families $\mathcal{S}_j \subset \mathscr{D}^{(j)}$ such that for a.e. $x \in \mathbb{R}^n$,*

$$|[b,T]f(x)| \le c_n C_T \sum_{j=1}^{3^n} \big(\mathcal{T}_{\mathcal{S}_j,b}|f|(x) + \mathcal{T}^{\star}_{\mathcal{S}_j,b}|f|(x)\big). \tag{1.2}$$

Some comments about this result are in order. A classical theorem of R. Coifman, R. Rochberg and G. Weiss [4] says that the condition $b \in BMO$ is sufficient (and for some T is also necessary) for the L^p boundedness of $[b,T]$ for all $1 < p < \infty$. It is easy to see that the definition of $\mathcal{T}_{\mathcal{S},b}$ is adapted to this condition. In Lemma 4.2 below we show that if $b \in BMO$, then $\mathcal{T}_{\mathcal{S},b}$ is of weak type $(1,1)$. On the other hand, C. Pérez [29] showed that $[b,T]$ is not of weak type $(1,1)$. Therefore, the second term $\mathcal{T}^{\star}_{\mathcal{S}_j,b}$ cannot be removed from (1.2).

Notice that the first term $\mathcal{T}_{\mathcal{S},b}$ cannot be removed from (1.2), too. Indeed, a standard argument (see the proof of (2.4) in Section 2.2) based on the John-Nirenberg inequality shows that if $b \in BMO$, then

$$\mathcal{T}^{\star}_{\mathcal{S},b}f(x) \le c_n \|b\|_{BMO} \sum_{Q\in\mathcal{S}} \|f\|_{L\log L,Q}\chi_Q(x).$$

But it was recently observed [32] that $[b,T]$ cannot be pointwise bounded by an $L\log L$-sparse operator appearing here.

In the following subsections we will show applications of Theorem 1.1 to weighted weak and strong type inequalities for $[b,T]$.

1.2. **Improved weighted weak type bounds.** Given a weight w (that is, a non-negative locally integrable function) and a measurable set $E \subset \mathbb{R}^n$, denote $w(E) = \int_E w dx$ and

$$w_f(\lambda) = w\{x \in \mathbb{R}^n : |f(x)| > \lambda\}.$$

In the classical work [10], C. Fefferman and E.M. Stein obtained the following weighted weak type $(1,1)$ property of the Hardy-Littlewood maximal operator M: for an arbitrary weight w,

$$w_{Mf}(\lambda) \le \frac{c_n}{\lambda} \int_{\mathbb{R}^n} |f(x)|Mw(x)dx \quad (\lambda > 0). \tag{1.3}$$

Only forty years after that, M.C. Reguera and C. Thiele [34] gave an example showing that a similar estimate is not true for the Hilbert transform instead of M on the left-hand side of (1.3) (they disproved by this the so-called Muckenhoupt-Wheeden conjecture). On the other hand, it was shown earlier by C. Pérez [28] that an analogue of (1.3) holds for a general class of Calderón-Zygmund operators but with a slightly bigger Orlicz maximal operator $M_{L(\log L)^\varepsilon}$ instead of M on the right-hand side. This result was reproved with a better dependence on ε in [18]: if T is a Calderón-Zygmund operator and $0<\varepsilon\le 1$, then

$$w_{Tf}(\lambda)\le \frac{c(n,T)}{\varepsilon}\frac{1}{\lambda}\int_{\mathbb{R}^n}|f(x)|M_{L(\log L)^\varepsilon}w(x)dx \quad (\lambda>0). \tag{1.4}$$

A general Orlicz maximal operator $M_{\varphi(L)}$ is defined for a Young function φ by

$$M_{\varphi(L)}f(x)=\sup_{Q\ni x}\|f\|_{\varphi,Q},$$

where the supremum is taken over all cubes $Q\subset\mathbb{R}^n$ containing x, and $\|f\|_{\varphi,Q}$ is the normalized Luxemburg norm defined by

$$\|f\|_{\varphi,Q}=\inf\Big\{\lambda>0:\frac{1}{|Q|}\int_Q\varphi(|f(y)|/\lambda)dy\le 1\Big\}.$$

If $\varphi(t)=t\log^\alpha(\mathrm{e}+t),\alpha>0$, denote $M_{\varphi(L)}=M_{L(\log L)^\alpha}$.

Recently, C. Domingo-Salazar, M. Lacey and G. Rey [9] obtained the following improvement of (1.4): if $C_\varphi=\int_1^\infty\frac{\varphi^{-1}(t)}{t^2\log(\mathrm{e}+t)}dt<\infty$, then

$$w_{Tf}(\lambda)\le\frac{c(n,T)C_\varphi}{\lambda}\int_{\mathbb{R}^n}|f(x)|M_{\varphi(L)}w(x)dx. \tag{1.5}$$

It is easy to see that if $\varphi(t)=t\log^\varepsilon(\mathrm{e}+t)$, then $C_\varphi\sim\frac{1}{\varepsilon}$, and thus (1.5) contains (1.4) as a particular case. On the other hand, (1.5) holds for smaller functions than $t\log^\varepsilon(\mathrm{e}+t)$, for instance, for $\varphi(t)=t\log\log^\alpha(\mathrm{e}^\mathrm{e}+t),\alpha>1$. The key ingredient in the proof of (1.5) was a pointwise control of T by sparse operators expressed in (1.1).

Consider now the commutator $[b,T]$ of T with a BMO function b. The following analogue of (1.4) was recently obtained by the third author and C. Pérez [31]: for all $0<\varepsilon\le 1$,

$$w_{[b,T]f}(\lambda)\le\frac{c(n,T)}{\varepsilon^2}\int_{\mathbb{R}^n}\Phi\left(\|b\|_{BMO}\frac{|f(x)|}{\lambda}\right)M_{L(\log L)^{1+\varepsilon}}w(x)dx, \tag{1.6}$$

where $\Phi(t)=t\log(\mathrm{e}+t)$. Observe that Φ here reflects an unweighted $L\log L$ weak type estimate for $[b,T]$ obtained by C. Pérez [29]. Notice also that (1.6) with worst dependence on ε was proved earlier in [30].

Similarly to the above mentioned improved weak type bound for Calderón-Zygmund operators (1.5), we apply Theorem 1.1 to improve (1.6). Our next result shows that (1.6) holds with $1/\varepsilon$ instead of $1/\varepsilon^2$ and that $M_{L(\log L)^{1+\varepsilon}}$ in (1.6) can be replaced by smaller Orlicz maximal operators.

Theorem 1.2. *Let T be an ω-Calderón-Zygmund operator with ω satisfying the Dini condition, and let $b \in BMO$. Let φ be an arbitrary Young function such that $C_\varphi = \int_1^\infty \frac{\varphi^{-1}(t)}{t^2 \log(\mathrm{e}+t)} dt < \infty$. Then for every weight w and for every compactly supported $f \in L^\infty$,*

$$w_{[b,T]f}(\lambda) \leq c_n C_T C_\varphi \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO} \frac{|f(x)|}{\lambda}\right) M_{(\Phi\circ\varphi)(L)} w(x) dx, \tag{1.7}$$

where $\Phi(t) = t\log(\mathrm{e}+t)$.

By Theorem 1.1, the proof of (1.7) is based on weak type estimates for $\mathcal{T}_{\mathcal{S},b}$ and $\mathcal{T}^\star_{\mathcal{S},b}$. The maximal operator $M_{(\Phi\circ\varphi)(L)}$ appears in the weighted weak type $(1,1)$ estimate for $\mathcal{T}_{\mathcal{S},b}$. It is interesting that $\mathcal{T}^\star_{\mathcal{S},b}$, being not of weak type $(1,1)$, satisfies a better estimate than (1.7) with a smaller maximal operator than $M_{(\Phi\circ\varphi)(L)}$ (which one can deduce from Lemma 4.5 below).

We mention several particular cases of interest in Theorem 1.2. Notice that if $\varphi(t) \leq t^2$ for $t \geq t_0$, then

$$\Phi \circ \varphi(t) \leq c\varphi(t)\log(\mathrm{e}+t) \quad (t > 0).$$

Hence, if $\varphi(t) = t\log^\varepsilon(\mathrm{e}+t)$, $0 < \varepsilon \leq 1$, then simple estimates along with (1.7) imply

$$w_{[b,T]f}(\lambda) \leq \frac{c(n,T)}{\varepsilon} \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO} \frac{|f(x)|}{\lambda}\right) M_{L(\log L)^{1+\varepsilon}} w(x) dx. \tag{1.8}$$

Similarly, if $\varphi(t) = t\log\log^{1+\varepsilon}(\mathrm{e}^\mathrm{e}+t), 0 < \varepsilon \leq 1$, then

$$w_{[b,T]f}(\lambda) \leq \frac{c(n,T)}{\varepsilon} \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO} \frac{|f(x)|}{\lambda}\right) M_{L(\log L)(\log\log L)^{1+\varepsilon}} w(x) dx.$$

As an application of Theorem 1.2, we obtain an improved weighted weak type estimate for $[b,T]$ assuming that the weight $w \in A_1$. Recall that the latter condition means that

$$[w]_{A_1} = \sup_{x\in\mathbb{R}^n} \frac{Mw(x)}{w(x)} < \infty.$$

Also we define the A_∞ constant of w by

$$[w]_{A_\infty} = \sup_Q \frac{1}{w(Q)} \int_Q M(w\chi_Q) dx,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^n$. It was shown in [18] that the dependence on ε in (1.4) implies the corresponding mixed A_1-A_∞ estimate. In a similar way we have the following.

Corollary 1.3. *For every* $w \in A_1$,

$$w_{[b,T]f}(\lambda) \leq c_n C_T [w]_{A_1} \Phi([w]_{A_\infty}) \int_{\mathbb{R}^n} \Phi\left(\|b\|_{BMO}\frac{|f(x)|}{\lambda}\right) w(x)dx,$$

where $\Phi(t) = t\log(\mathrm{e}+t)$.

This provides a logarithmic improvement of the corresponding bounds in [27, 31].

1.3. **Two-weighted strong type bounds.** Let w be a weight, and let $1 < p < \infty$. Denote $\sigma_w(x) = w^{-\frac{1}{p-1}}(x)$. Given a cube $Q \subset \mathbb{R}^n$, set

$$[w]_{A_p,Q} = \frac{w(Q)}{|Q|}\left(\frac{\sigma_w(Q)}{|Q|}\right)^{p-1}.$$

We say that $w \in A_p$ if

$$[w]_{A_p} = \sup_Q [w]_{A_p,Q} < \infty.$$

As we have mentioned previously, pointwise bounds by sparse operators were motivated by sharp weighted norm inequalities. For example, (1.1) provides a simple proof of the sharp $L^p(w)$ bound for T (see [19, 24]):

$$\|T\|_{L^p(w)} \leq c_{n,p} C_T [w]_{A_p}^{\max\left(1,\frac{1}{p-1}\right)}. \quad (1 < p < \infty) \tag{1.9}$$

In the case of ω-Calderón-Zygmund operators with $\omega(t) = ct^\delta$, (1.9) was proved by T. Hytönen [15] (see also [16, 23] for the history of this result and a different proof).

An analogue of (1.9) for the commutator $[b, T]$ with a BMO function b is the following sharp $L^p(w)$ bound due to D. Chung, C. Pereyra and C. Pérez [3]:

$$\|[b,T]\|_{L^p(w)} \leq c(n,p,T)\|b\|_{BMO}[w]_{A_p}^{2\max\left(1,\frac{1}{p-1}\right)}. \quad (1 < p < \infty) \tag{1.10}$$

Much earlier, S. Bloom [2] obtained an interesting two-weighted result for the commutator of the Hilbert transform H: if $\mu, \lambda \in A_p, 1 < p < \infty, \nu = (\mu/\lambda)^{1/p}$ and $b \in BMO_\nu$, then

$$\|[b,H]f\|_{L^p(\lambda)} \leq c(p,\mu,\lambda)\|b\|_{BMO_\nu}\|f\|_{L^p(\mu)}. \tag{1.11}$$

Here BMO_ν is the weighted BMO space of locally integrable functions b such that

$$\|b\|_{BMO_\nu} = \sup_Q \frac{1}{\nu(Q)}\int_Q |b - b_Q|dx < \infty.$$

Recently, I. Holmes, M. Lacey and B. Wick [13] extended (1.11) to ω-Calderón-Zygmund operators with $\omega(t) = ct^\delta$; the key role in their proof was played by Hytönen's representation theorem [15] for such operators. In the particular case when $\mu = \lambda = w \in A_2$ the approach in [13] recovers (1.10) (this was checked in [14]; and also, (1.11) was extended in this work to higher-order commutators).

Using Theorem 1.1, we obtain the following quantitative version of the Bloom-Holmes-Lacey-Wick result. It extends (1.11) to any ω-Calderón-Zygmund operator with the Dini condition, and the explicit dependence on $[\mu]_{A_p}$ and $[\lambda]_{A_p}$ is found. Also, it can be viewed as a natural extension of (1.10) to the two-weighted setting.

Theorem 1.4. *Let T be an ω-Calderón-Zygmund operator with ω satisfying the Dini condition. Let $\mu, \lambda \in A_p, 1 < p < \infty$, and $\nu = (\mu/\lambda)^{1/p}$. If $b \in BMO_\nu$, then*

$$\|[b,T]f\|_{L^p(\lambda)} \le c_{n,p} C_T \big([\mu]_{A_p}[\lambda]_{A_p}\big)^{\max\left(1,\frac{1}{p-1}\right)} \|b\|_{BMO_\nu} \|f\|_{L^p(\mu)}.$$

The paper is organized as follows. In Section 2, we collect some preliminary information about dyadic lattices, sparse families and Young functions. Section 3 is devoted to the proof of Theorem 1.1. In Section 4, we prove Theorem 1.2 and Corollary 1.3, and Section 5 contains the proof of Theorem 1.4.

2. Preliminaries

2.1. **Dyadic lattices and sparse families.** By a cube in $\mathbb{R}^n$ we mean a half-open cube $Q = \prod_{i=1}^n [a_i, a_i + h), h > 0$. Denote by ℓ_Q the sidelength of Q. Given a cube $Q_0 \subset \mathbb{R}^n$, let $\mathcal{D}(Q_0)$ denote the set of all dyadic cubes with respect to Q_0, that is, the cubes obtained by repeated subdivision of Q_0 and each of its descendants into 2^n congruent subcubes.

A dyadic lattice $\mathscr{D}$ in $\mathbb{R}^n$ is any collection of cubes such that

(i) if $Q \in \mathscr{D}$, then each child of Q is in $\mathscr{D}$ as well;
(ii) every 2 cubes $Q', Q'' \in \mathscr{D}$ have a common ancestor, i.e., there exists $Q \in \mathscr{D}$ such that $Q', Q'' \in \mathcal{D}(Q)$;
(iii) for every compact set $K \subset \mathbb{R}^n$, there exists a cube $Q \in \mathscr{D}$ containing K.

For this definition, as well as for the next Theorem, we refer to [25].

Theorem 2.1. (The Three Lattice Theorem) *For every dyadic lattice $\mathscr{D}$, there exist 3^n dyadic lattices $\mathscr{D}^{(1)}, \dots, \mathscr{D}^{(3^n)}$ such that*

$$\{3Q : Q \in \mathscr{D}\} = \cup_{j=1}^{3^n} \mathscr{D}^{(j)}$$

and for every cube $Q \in \mathscr{D}$ and $j = 1, \dots, 3^n$, there exists a unique cube $R \in \mathscr{D}^{(j)}$ of sidelength $\ell_R = 3\ell_Q$ containing Q.

Remark 2.2. Fix a dyadic lattice $\mathscr{D}$. For an arbitrary cube $Q \subset \mathbb{R}^n$, there is a cube $Q' \in \mathscr{D}$ such that $\ell_Q/2 < \ell_{Q'} \leq \ell_Q$ and $Q \subset 3Q'$. By Theorem 2.1, there is $j = 1, \dots, 3^n$ such that $3Q' = P \in \mathscr{D}^{(j)}$. Therefore, for every cube $Q \subset \mathbb{R}^n$, there exists $P \in \mathscr{D}^{(j)}$, $j = 1, \dots, 3^n$, such that $Q \subset P$ and $\ell_P \leq 3\ell_Q$. A similar statement can be found in [17, Lemma 2.5].

We say that a family $\mathcal{S}$ of cubes from $\mathscr{D}$ is η-sparse, $0 < \eta < 1$, if for every $Q \in \mathcal{S}$, there exists a measurable set $E_Q \subset Q$ such that $|E_Q| \geq \eta|Q|$, and the sets $\{E_Q\}_{Q\in\mathcal{S}}$ are pairwise disjoint.

A family $\mathcal{S} \subset \mathscr{D}$ is called Λ-Carleson, $\Lambda > 1$, if for every cube $Q \in \mathscr{D}$,

$$\sum_{P\in\mathcal{S}, P\subset Q} |P| \leq \Lambda|Q|.$$

It is easy to see that every η-sparse family is $(1/\eta)$-Carleson. In [25, Lemma 6.3], it is shown that the converse statement is also true, namely, every Λ-Carleson family is $(1/\Lambda)$-sparse. Also, [25, Lemma 6.6] says that if $\mathcal{S}$ is Λ-Carleson and $m \in \mathbb{N}$ such that $m \geq 2$, then $\mathcal{S}$ can be written as a union of m families $\mathcal{S}_j$, each of which is $(1 + \frac{\Lambda-1}{m})$-Carleson. Using the above mentioned relation between sparse and Carleson families, one can rewrite the latter fact as follows.

Lemma 2.3. *If $\mathcal{S} \subset \mathscr{D}$ is η-sparse and $m \geq 2$, then one can represent $\mathcal{S}$ as a disjoint union $\mathcal{S} = \cup_{j=1}^m \mathcal{S}_j$, where each family $\mathcal{S}_j$ is $\frac{m}{m+(1/\eta)-1}$-sparse.*

2.2. **Young functions and normalized Luxemburg norms.** By a Young function we mean a continuous, convex, strictly increasing function $\varphi : [0,\infty) \to [0,\infty)$ with $\varphi(0) = 0$ and $\varphi(t)/t \to \infty$ as $t \to \infty$. Notice that such functions are also called in the literature the N-functions. We refer to [20, 33] for their basic properties. We will use, in particular, that $\varphi(t)/t$ is also a strictly increasing function (see, e.g., [20, p. 8]).

We will also use the fact that

$$\|f\|_{\varphi,Q} \leq 1 \Leftrightarrow \frac{1}{|Q|}\int_Q \varphi(|f(x)|)dx \leq 1. \tag{2.1}$$

Given a Young function φ, its complementary function is defined by

$$\bar{\varphi}(t) = \sup_{x\geq 0}\big(xt - \varphi(x)\big).$$

Then $\bar{\varphi}$ is also a Young function satisfying $t \le \bar{\varphi}^{-1}(t)\varphi^{-1}(t) \le 2t$. Also the following Hölder type estimate holds:

$$\frac{1}{|Q|}\int_Q |f(x)g(x)|dx \le 2\|f\|_{\varphi,Q}\|g\|_{\bar{\varphi},Q}. \tag{2.2}$$

Recall that the John-Nirenberg inequality (see, e.g., [12, p. 124]) says that for every $b \in BMO$ and for any cube $Q \subset \mathbb{R}^n$,

$$|\{x \in Q : |b(x) - b_Q| > \lambda\}| \le e|Q|e^{-\frac{\lambda}{2^n e\|b\|_{BMO}}} \quad (\lambda > 0). \tag{2.3}$$

In particular, this inequality implies (see [12, p. 128])

$$\frac{1}{|Q|}\int_Q e^{\frac{|b(x)-b_Q|}{c_n\|b\|_{BMO}}-1}dx \le 1.$$

From this and from (2.1), taking $\varphi(t) = e^t - 1$, we obtain

$$\|b - b_Q\|_{\varphi,Q} \le c_n\|b\|_{BMO}.$$

A simple computation shows that in this case $\bar{\varphi}(t) \approx t\log(\mathrm{e} + t)$, and therefore, by (2.2),

$$\frac{1}{|Q|}\int_Q |(b - b_Q)g|dx \le c_n\|b\|_{BMO}\|g\|_{L\log L,Q}. \tag{2.4}$$

Notice that many important properties of the Luxemburg normalized norms $\|f\|_{\varphi,Q}$ hold without assuming the convexity of φ. In particular, we will use the following generalized Hölder inequality.

Lemma 2.4. *Let A, B and C be non-negative, continuous, strictly increasing functions on $[0,\infty)$ satisfying $A^{-1}(t)B^{-1}(t) \le C^{-1}(t)$ for all $t \ge 0$. Assume also that C is convex. Then*

$$\|fg\|_{C,Q} \le 2\|f\|_{A,Q}\|g\|_{B,Q}. \tag{2.5}$$

This lemma was proved by R. O'Neil [26] under the assumption that A, B and C are Young functions but the same proof works under the above conditions. Indeed, by homogeneity, it suffices to assume that $\|f\|_{A,Q} = \|g\|_{B,Q} = 1$. Next, notice that the assumptions on A, B and C easily imply that $C(xy) \le A(x) + B(y)$ for all $x, y \ge 0$. Therefore, using the convexity of C and (2.1), we obtain

$$\frac{1}{|Q|}\int_Q C(|fg|/2)dx \le \frac{1}{2}\Big(\frac{1}{|Q|}\int_Q A(|f|)dx + \frac{1}{|Q|}\int_Q B(|g|)dx\Big) \le 1,$$

which, by (2.1) again, implies (2.5).

Given a dyadic lattice $\mathscr{D}$, denote

$$M_\Phi^{\mathscr{D}}f(x) = \sup_{Q\ni x, Q\in\mathscr{D}}\|f\|_{\Phi,Q}.$$

The following lemma is a generalization of the Fefferman-Stein inequality (1.3) to general Orlicz maximal functions, and it is apparently well-known. We give its proof for the sake of completeness.

Lemma 2.5. *Let Φ be a Young function. For an arbitrary weight w,*

$$w\{x\in\mathbb{R}^n : M_\Phi f(x)>\lambda\}\le 3^n\int_{\mathbb{R}^n}\Phi\left(\frac{9^n|f(x)|}{\lambda}\right)Mw(x)dx.$$

Proof. By the Calderón-Zygmund decomposition adapted to $M_\Phi^{\mathscr{D}}$ (see [6, p. 237]), there exists a family of disjoint cubes $\{Q_i\}$ such that

$$\left\{x\in\mathbb{R}^n : M_\Phi^{\mathscr{D}}f(x)>\lambda\right\}=\cup_i Q_i$$

and $\lambda<\|f\|_{\Phi,Q_i}\le 2^n\lambda$. By (2.1), we see that $\|f\|_{\Phi,Q_i}>\lambda$ implies $\int_{Q_i}\Phi(|f|/\lambda)>|Q_i|$. Therefore,

$$\begin{aligned} w\{x\in\mathbb{R}^n : M_\Phi^{\mathscr{D}}f(x)>\lambda\} &= \sum_i w(Q_i)\\ &< \sum_i w_{Q_i}\int_{Q_i}\Phi(|f(x)|/\lambda)dx\le\int_{\mathbb{R}^n}\Phi(|f(x)|/\lambda)Mw(x)dx.\end{aligned}$$

Now we observe that by the convexity of Φ and Remark 2.2, there exist 3^n dyadic lattices $\mathscr{D}^{(j)}$ such that

$$M_\Phi f(x)\le 3^n\sum_{j=1}^{3^n}M_\Phi^{\mathscr{D}^{(j)}}f(x).$$

Combining this estimate with the previous one completes the proof. □

Remark 2.6. Suppose that $\Phi(t)=t\log(\mathrm{e}+t)$. It is easy to see that for all $a,b\ge 0$,

$$\Phi(ab)\le 2\Phi(a)\Phi(b). \tag{2.6}$$

From this and from Lemma 2.5,

$$w\{x\in\mathbb{R}^n : M_{L\log L}f(x)>\lambda\}\le c_n\int_{\mathbb{R}^n}\Phi\left(\frac{|f(x)|}{\lambda}\right)Mw(x)dx.$$

3. Proof of Theorem 1.1

The proof of Theorem 1.1 is a slight modification of the argument in [24]. Although some parts of the proofs here and in [24] are almost identical, certain details are different, and hence we give a complete proof. We start by defining several important objects.

Let T be an ω-Calderón-Zygmund operator with ω satisfying the Dini condition. Recall that the maximal truncated operator $T^\star$ is defined by

$$T^\star f(x) = \sup_{\varepsilon>0} \Big| \int_{|y-x|>\varepsilon} K(x,y)f(y)dy \Big|.$$

Define the grand maximal truncated operator $\mathcal{M}_T$ by

$$\mathcal{M}_T f(x) = \sup_{Q\ni x} \operatorname*{ess\,sup}_{\xi\in Q} |T(f\chi_{\mathbb{R}^n\setminus 3Q})(\xi)|,$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^n$ containing x.

Given a cube Q_0, for $x \in Q_0$ define a local version of $\mathcal{M}_T$ by

$$\mathcal{M}_{T,Q_0} f(x) = \sup_{Q\ni x, Q\subset Q_0} \operatorname*{ess\,sup}_{\xi\in Q} |T(f\chi_{3Q_0\setminus 3Q})(\xi)|.$$

The next lemma was proved in [24].

Lemma 3.1. *The following pointwise estimates hold:*

(i) *for a.e.* $x \in Q_0$,

$$|T(f\chi_{3Q_0})(x)| \le c_n \|T\|_{L^1\to L^{1,\infty}} |f(x)| + \mathcal{M}_{T,Q_0} f(x);$$

(ii) *for all* $x \in \mathbb{R}^n$,

$$\mathcal{M}_T f(x) \le c_n(\|\omega\|_{\mathrm{Dini}} + C_K) M f(x) + T^\star f(x).$$

An examination of standard proofs (see, e.g., [12, Ch. 8.2]) shows that

$$\max(\|T\|_{L^1\to L^{1,\infty}}, \|T^\star\|_{L^1\to L^{1,\infty}}) \le c_n C_T. \tag{3.1}$$

By part (ii) of Lemma 3.1 and by (3.1),

$$\|\mathcal{M}_T\|_{L^1\to L^{1,\infty}} \le c_n C_T. \tag{3.2}$$

Proof of Theorem 1.1. By Remark 2.2, there exist 3^n dyadic lattices $\mathscr{D}^{(j)}$ such that for every $Q \subset \mathbb{R}^n$, there is a cube $R = R_Q \in \mathscr{D}^{(j)}$ for some j, for which $3Q \subset R_Q$ and $|R_Q| \le 9^n |Q|$.

Fix a cube $Q_0 \subset \mathbb{R}^n$. Let us show that there exists a $\frac{1}{2}$-sparse family $\mathcal{F} \subset \mathcal{D}(Q_0)$ such that for a.e. $x \in Q_0$,

$$\begin{aligned} &|[b,T](f\chi_{3Q_0})(x)| \\ &\le c_n C_T \sum_{Q\in\mathcal{F}} \big(|b(x) - b_{R_Q}||f|_{3Q} + |(b-b_{R_Q})f|_{3Q}\big)\chi_Q(x). \end{aligned} \tag{3.3}$$

It suffices to prove the following recursive claim: there exist pairwise disjoint cubes $P_j \in \mathcal{D}(Q_0)$ such that $\sum_j |P_j| \le \frac{1}{2}|Q_0|$ and

$$\begin{aligned}|[b,T](f\chi_{3Q_0})(x)|\chi_{Q_0} &\le c_nC_T\big(|b(x)-b_{R_{Q_0}}||f|_{3Q_0}+|(b-b_{R_{Q_0}})f|_{3Q_0}\big)\\ &+ \sum_j |[b,T](f\chi_{3P_j})(x)|\chi_{P_j}.\end{aligned}$$

a.e. on Q_0. Indeed, iterating this estimate, we immediately get (3.3) with $\mathcal{F} = \{P_j^k\}, k \in \mathbb{Z}_+$, where $\{P_j^0\} = \{Q_0\}$, $\{P_j^1\} = \{P_j\}$ and $\{P_j^k\}$ are the cubes obtained at the k-th stage of the iterative process.

Next, observe that for arbitrary pairwise disjoint cubes $P_j \in \mathcal{D}(Q_0)$,

$$\begin{aligned}&|[b,T](f\chi_{3Q_0})|\chi_{Q_0} = |[b,T](f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |[b,T](f\chi_{3Q_0})|\chi_{P_j}\\ &\le |[b,T](f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |[b,T](f\chi_{3Q_0\setminus3P_j})|\chi_{P_j}\\ &+\sum_j |[b,T](f\chi_{3P_j})|\chi_{P_j}.\end{aligned}$$

Hence, in order to prove the recursive claim, it suffices to show that one can select pairwise disjoint cubes $P_j \in \mathcal{D}(Q_0)$ with $\sum_j |P_j| \le \frac{1}{2}|Q_0|$ and such that for a.e. $x \in Q_0$,

$$\begin{aligned}&|[b,T](f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |[b,T](f\chi_{3Q_0\setminus3P_j})|\chi_{P_j}\\ &\le c_nC_T\big(|b(x)-b_{R_{Q_0}}||f|_{3Q_0}+|(b-b_{R_{Q_0}})f|_{3Q_0}\big).\end{aligned}\tag{3.4}$$

Using that $[b,T]f = [b-c,T]f$ for any $c \in \mathbb{R}$, we obtain

$$\begin{aligned}&|[b,T](f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |[b,T](f\chi_{3Q_0\setminus3P_j})|\chi_{P_j}\\ &\le |b-b_{R_{Q_0}}|\Big(|T(f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |T(f\chi_{3Q_0\setminus3P_j})|\chi_{P_j}\Big)\\ &+|T((b-b_{R_{Q_0}})f\chi_{3Q_0})|\chi_{Q_0\setminus\cup_jP_j} + \sum_j |T((b-b_{R_{Q_0}})f\chi_{3Q_0\setminus3P_j})|\chi_{P_j}.\end{aligned}\tag{3.5}$$

By (3.2), one can choose α_n such that the set $E = E_1 \cup E_2$, where

$$E_1 = \{x \in Q_0 : |f| > \alpha_n|f|_{3Q_0}\} \cup \{x \in Q_0 : \mathcal{M}_{T,Q_0}f > \alpha_nC_T|f|_{3Q_0}\}$$

and

$$\begin{aligned}E_2 &= \{x \in Q_0 : |(b-b_{R_{Q_0}})f| > \alpha_n|(b-b_{R_{Q_0}})f|_{3Q_0}\}\\ &\cup \{x \in Q_0 : \mathcal{M}_{T,Q_0}(b-b_{R_{Q_0}})f > \alpha_nC_T|(b-b_{R_{Q_0}})f|_{3Q_0}\},\end{aligned}$$

will satisfy $|E| \le \frac{1}{2^{n+2}}|Q_0|$.

The Calderón-Zygmund decomposition applied to the function χ_E on Q_0 at height $\lambda = \frac{1}{2^{n+1}}$ produces pairwise disjoint cubes $P_j \in \mathcal{D}(Q_0)$ such that

$$\frac{1}{2^{n+1}}|P_j| \leq |P_j \cap E| \leq \frac{1}{2}|P_j|$$

and $|E \setminus \cup_j P_j| = 0$. It follows that $\sum_j |P_j| \leq \frac{1}{2}|Q_0|$ and $P_j \cap E^c \neq \emptyset$. Therefore,

$$\operatorname*{ess\,sup}_{\xi \in P_j} |T(f\chi_{3Q_0 \setminus 3P_j})(\xi)| \leq c_n C_T |f|_{3Q_0}$$

and

$$\operatorname*{ess\,sup}_{\xi \in P_j} |T((b - b_{R_{Q_0}})f\chi_{3Q_0 \setminus 3P_j})(\xi)| \leq c_n C_T |(b - b_{R_{Q_0}})f|_{3Q_0}.$$

Also, by part (i) of Lemma 3.1 and by (3.1), for a.e. $x \in Q_0 \setminus \cup_j P_j$,

$$|T(f\chi_{3Q_0})(x)| \leq c_n C_T |f|_{3Q_0}$$

and

$$|T((b - b_{R_{Q_0}})f\chi_{3Q_0})(x)| \leq c_n C_T |(b - b_{R_{Q_0}})f|_{3Q_0}.$$

Combining the obtained estimates with (3.5) proves (3.4), and therefore, (3.3) is proved.

Take now a partition of $\mathbb{R}^n$ by cubes Q_j such that $\operatorname{supp}(f) \subset 3Q_j$ for each j. For example, take a cube Q_0 such that $\operatorname{supp}(f) \subset Q_0$ and cover $3Q_0 \setminus Q_0$ by $3^n - 1$ congruent cubes Q_j. Each of them satisfies $Q_0 \subset 3Q_j$. Next, in the same way cover $9Q_0 \setminus 3Q_0$, and so on. The union of resulting cubes, including Q_0, will satisfy the desired property.

Having such a partition, apply (3.3) to each Q_j. We obtain a $\frac{1}{2}$-sparse family $\mathcal{F}_j \subset \mathcal{D}(Q_j)$ such that (3.3) holds for a.e. $x \in Q_j$ with $|Tf|$ on the left-hand side. Therefore, setting $\mathcal{F} = \cup_j \mathcal{F}_j$, we obtain that $\mathcal{F}$ is a $\frac{1}{2}$-sparse family, and for a.e. $x \in \mathbb{R}^n$,

$$|[b,T]f(x)| \leq c_n C_T \sum_{Q \in \mathcal{F}} \big(|b(x) - b_{R_Q}||f|_{3Q} + |(b - b_{R_Q})f|_{3Q}\big)\chi_Q(x). \tag{3.6}$$

Since $3Q \subset R_Q$ and $|R_Q| \leq 3^n|3Q|$, we obtain $|f|_{3Q} \leq c_n|f|_{R_Q}$. Further, setting $\mathcal{S}_j = \{R_Q \in \mathscr{D}^{(j)} : Q \in \mathcal{F}\}$, and using that $\mathcal{F}$ is $\frac{1}{2}$-sparse, we obtain that each family $\mathcal{S}_j$ is $\frac{1}{2\cdot 9^n}$-sparse. It follows from (3.6) that

$$|[b,T]f(x)| \leq c_n C_T \sum_{j=1}^{3^n} \sum_{R \in \mathcal{S}_j} \big(|b(x) - b_R||f|_R + |(b - b_R)f|_R\big)\chi_R(x),$$

and therefore, the proof is complete. □

4. Proof of Theorem 1.2 and Corollary 1.3

Fix a dyadic lattice $\mathscr{D}$. Let $\mathcal{S} \subset \mathscr{D}$ be a sparse family. Define the $L \log L$ sparse operator by

$$\mathcal{A}_{\mathcal{S},L\log L}f(x) = \sum_{Q\in\mathcal{S}} \|f\|_{L\log L,Q}\chi_Q(x).$$

It follows from (2.4) that

$$|\mathcal{T}^{\star}_{b,\mathcal{S}}f(x)| \le c_n\|b\|_{BMO}\mathcal{A}_{\mathcal{S},L\log L}f(x). \tag{4.1}$$

Let $\Phi(t) = t\log(\mathrm{e}+t)$. Given a Young function φ, denote

$$C_\varphi = \int_1^\infty \frac{\varphi^{-1}(t)}{t^2\log(\mathrm{e}+t)}dt.$$

By Theorem 1.1 combined with (4.1), Lemma 2.3 and a submultiplicative property of Φ expressed in (2.6), Theorem 1.2 is an immediate consequence of the following two lemmas.

Lemma 4.1. *Suppose that $\mathcal{S}$ is $\frac{31}{32}$-sparse. Let φ be a Young function such that $C_\varphi < \infty$. Then for an arbitrary weight w,*

$$w_{\mathcal{A}_{\mathcal{S},L\log L}f}(\lambda) \le cC_\varphi\int_{\mathbb{R}^n}\Phi\left(\frac{|f(x)|}{\lambda}\right)M_{(\Phi\circ\varphi)(L)}w(x)dx \quad (\lambda>0),$$

where $c>0$ is an absolute constant.

Lemma 4.2. *Let $b \in BMO$. Suppose that $\mathcal{S}$ is $\frac{7}{8}$-sparse. Let φ be a Young function such that $C_\varphi < \infty$. Then for an arbitrary weight w,*

$$w_{\mathcal{T}_{b,\mathcal{S}}f}(\lambda) \le \frac{c_nC_\varphi\|b\|_{BMO}}{\lambda}\int_{\mathbb{R}^n}|f(x)|M_{(\Phi\circ\varphi)(L)}w(x)dx \quad (\lambda>0).$$

In the following subsection we separate a common ingredient used in the proofs of both Lemmas 4.1 and 4.2.

4.1. The key lemma.

Assume that Ψ is a Young function satisfying

$$\Psi(4t) \le \Lambda_\Psi\Psi(t) \quad (t>0, \Lambda_\Psi \ge 1). \tag{4.2}$$

Given a dyadic lattice $\mathscr{D}$ and $k \in \mathbb{N}$, denote

$$\mathcal{F}_k = \{Q \in \mathscr{D} : 4^{k-1} < \|f\|_{\Psi,Q} \le 4^k\}.$$

The following lemma in the case $\Psi(t) = t$ was proved in [9]. Our extension to any Young function satisfying (4.2) is based on similar ideas. Notice that the main cases of interest for us are $\Psi(t) = t$ and $\Psi(t) = \Phi(t)$.

Lemma 4.3. *Suppose that the family $\mathcal{F}_k$ is $\left(1-\frac{1}{2\Lambda_\Psi}\right)$-sparse. Let w be a weight and let E be an arbitrary measurable set with $w(E)<\infty$. Then, for every Young function φ,*

$$\int_E\Big(\sum_{Q\in\mathcal{F}_k}\chi_Q\Big)w\,dx\le 2^k w(E)+\frac{4\Lambda_\Psi}{\bar{\varphi}^{-1}((2\Lambda_\Psi)^{2^k})}\int_{\mathbb{R}^n}\Psi(4^k|f|)M_{\varphi(L)}w\,dx.$$

Proof. By Fatou's lemma, one can assume that the family $\mathcal{F}_k$ is finite. Split $\mathcal{F}_k$ into the layers $\mathcal{F}_{k,\nu}$, $\nu=0,1,\dots,$ where $\mathcal{F}_{k,0}$ is the family of the maximal cubes in $\mathcal{F}_k$ and $\mathcal{F}_{k,\nu+1}$ is the family of the maximal cubes in $\mathcal{F}_k\setminus\bigcup_{l=0}^{\nu}\mathcal{F}_{k,l}$.

Denote $E_Q=Q\setminus\bigcup_{Q'\in\mathcal{F}_{k,\nu+1}}Q'$ for each $Q\in\mathcal{F}_{k,\nu}$. Then the sets E_Q are pairwise disjoint for $Q\in\mathcal{F}_k$.

For $\nu\ge 0$ and $Q\in\mathcal{F}_{k,\nu}$ denote

$$A_k(Q)=\bigcup_{Q'\in\mathcal{F}_{k,\nu+2^k},\,Q'\subset Q}Q'.$$

Observe that

$$Q\setminus A_k(Q)=\bigcup_{l=0}^{2^k-1}\ \bigcup_{Q'\in\mathcal{F}_{k,\nu+l},Q'\subseteq Q}E_{Q'}.$$

Using the disjointness of the sets E_Q, we obtain

$$\begin{aligned}\sum_{Q\in\mathcal{F}_k}w\big(E\cap(Q\setminus A_k(Q))\big)&\le\sum_{\nu=0}^{\infty}\sum_{Q\in\mathcal{F}_{k,\nu}}\sum_{l=0}^{2^k-1}\sum_{\substack{Q'\in\mathcal{F}_{k,\nu+l}\\ Q'\subseteq Q}}w(E\cap E_{Q'})\\ &\le 2^k\sum_{Q\in\mathcal{F}_k}w(E\cap E_Q)\le 2^k w(E).\end{aligned}\tag{4.3}$$

Now, let us show that

$$1\le\frac{2\Lambda_\Psi}{|Q|}\int_{E_Q}\Psi(4^k|f(x)|)dx\quad(Q\in\mathcal{S}_k).\tag{4.4}$$

Fix a cube $Q\in\mathcal{F}_{k,\nu}$. Since $4^{-k-1}<\|f\|_{\Psi,Q}$, by (2.1) and by (4.2),

$$1<\frac{1}{|Q|}\int_Q\Psi(4^{k+1}|f|)\le\frac{\Lambda_\Psi}{|Q|}\int_Q\Psi(4^k|f|).\tag{4.5}$$

On the other hand, for any $P\in\mathcal{F}_k$ we have $\|f\|_{\Psi,P}\le 4^{-k}$, and hence, by (2.1),

$$\frac{1}{|P|}\int_P\Psi(4^k|f|)\le 1.$$

Using also that, by the sparseness condition, $|Q \setminus E_Q| \le \frac{1}{2\Lambda_\Psi}|Q|$, we obtain

$$\frac{1}{|Q|}\int_Q \Psi(4^k|f|) = \frac{1}{|Q|}\int_{E_Q} \Psi(4^k|f|) + \frac{1}{|Q|}\sum_{Q'\in\mathcal{S}_{k,\nu+1}}\int_{Q'} \Psi(4^k|f|)$$
$$\le \frac{1}{|Q|}\int_{E_Q} \Psi(4^k|f|) + \frac{|Q\setminus E_Q|}{|Q|} \le \frac{1}{|Q|}\int_{E_Q} \Psi(4^k|f|) + \frac{1}{2\Lambda_\Psi},$$

which, along with (4.5), proves (4.4).

Applying the sparseness assumption again, we obtain $|A_k(Q)| \le (1/2\Lambda_\Psi)^{2^k}|Q|$. From this and from Hölder's inequality (2.2),

$$\begin{aligned}\frac{w(A_k(Q))}{|Q|} &\le 2\|\chi_{A_k(Q)}\|_{\bar\varphi,Q}\|w\|_{\varphi,Q} = \frac{2}{\bar\varphi^{-1}(|Q|/|A_k(Q)|)}\|w\|_{\varphi,Q}\\ &\le \frac{2}{\bar\varphi^{-1}((2\Lambda_\Psi)^{2^k})}\|w\|_{\varphi,Q}.\end{aligned}$$

Combining this with (4.4) yields

$$w(A_k(Q)) \le \frac{4\Lambda_\Psi}{\bar\varphi^{-1}((2\Lambda_\Psi)^{2^k})}\int_{E_Q}\Psi(4^k|f|)M_{\varphi(L)}w\,dx.$$

Hence, by the disjointness of the sets E_Q,

$$\sum_{Q\in\mathcal{F}_k} w(A_k(Q)) \le \frac{4\Lambda_\Psi}{\bar\varphi^{-1}((2\Lambda_\Psi)^{2^k})}\int_{\mathbb{R}^n}\Psi(4^k|f|)M_{\varphi(L)}w\,dx,$$

which, along with (4.3), completes the proof. □

4.2. **Proof of Lemmas 4.1 and 4.2.** We first mention another common ingredient used in both proofs.

Proposition 4.4. *Let Ψ be a Young function. Assume that G is an operator such that for every Young function φ,*

$$w_{Gf}(\lambda) \le \left(\int_1^\infty \frac{\varphi^{-1}(t)}{t^2}dt\right)\int_{\mathbb{R}^n}\Psi\left(\frac{|f(x)|}{\lambda}\right)M_{\varphi(L)}w(x)dx. \tag{4.6}$$

Then

$$w_{Gf}(\lambda) \le cC_\varphi\int_{\mathbb{R}^n}\Psi\left(\frac{|f(x)|}{\lambda}\right)M_{(\Phi\circ\varphi)(L)}w(x)dx,$$

where $c>0$ is an absolute constant, and $C_\varphi = \int_1^\infty \frac{\varphi^{-1}(t)}{t^2\log(e+t)}dt$.

Indeed, this follows immediately by setting $\Phi\circ\varphi$ instead of φ in (4.6) and observing that $(\Phi\circ\varphi)^{-1} = \varphi^{-1}\circ\Phi^{-1}$ and

$$\int_1^\infty \frac{\varphi^{-1}\circ\Phi^{-1}(t)}{t^2}dt = \int_{\Phi^{-1}(1)}^\infty \frac{\varphi^{-1}(t)}{\Phi(t)^2}\Phi'(t)dt \le cC_\varphi. \tag{4.7}$$

Turn to Lemma 4.1. We actually obtain a stronger statement, namely, we will prove the following.

Lemma 4.5. *Suppose that $\mathcal{S}$ is $\frac{31}{32}$-sparse. Let φ be a Young function such that*

$$K_\varphi = \int_1^\infty \frac{\varphi^{-1}(t)\log\log(\mathrm{e}^2+t)}{t^2\log(\mathrm{e}+t)}dt < \infty.$$

Then for an arbitrary weight w,

$$w_{\mathcal{A}_{\mathcal{S},L\log L}f}(\lambda) \le cK_\varphi \int_{\mathbb{R}^n} \Phi\left(\frac{|f(x)|}{\lambda}\right) M_{\varphi(L)}w(x)dx \quad (\lambda > 0),$$

where $c > 0$ is an absolute constant.

Since $K_\varphi \le \int_1^\infty \frac{\varphi^{-1}(t)}{t^2}dt$, Proposition 4.4 shows that Lemma 4.1 follows from Lemma 4.5.

Proof of Lemma 4.5. Consider the set

$$E = \{x \in \mathbb{R}^n : \mathcal{A}_{\mathcal{S},L\log L}f(x) > 4, M_{L\log L}f(x) \le 1/4\}.$$

By homogeneity combined with Remark 2.6, it suffices to prove that

$$w(E) \le cK_\varphi \int_{\mathbb{R}^n} \Phi(|f(x)|)M_{\varphi(L)}w(x)dx. \tag{4.8}$$

One can assume that $w(E) < \infty$ (otherwise, one could first obtain (4.8) for $E \cap K$ instead of E, for any compact set K).

Denote

$$\mathcal{S}_k = \{Q \in \mathcal{S} : 4^{-k-1} < \|f\|_{L\log L,Q} \le 4^{-k}\}$$

and set

$$T_kf(x) = \sum_{Q\in\mathcal{S}_k} \|f\|_{L\log L,Q}\chi_Q(x).$$

If $E \cap Q \ne \emptyset$ for some $Q \in \mathcal{S}$, then $\|f\|_{L\log L,Q} \le 1/4$. Therefore, for $x \in E$,

$$\mathcal{A}_{\mathcal{S},L\log L}f(x) = \sum_{k=1}^\infty T_kf(x). \tag{4.9}$$

Now we apply Lemma 4.3 with $\Psi = \Phi$ and $\mathcal{F}_k = \mathcal{S}_k$. Notice that, by (2.6), one can take $\Lambda_\Psi = 16$ in (4.2) and $\Phi(4^k|f|) \le ck4^k\Phi(|f|)$. Combining this with $T_kf(x) \le 4^{-k}\sum_{Q\in\mathcal{S}_k}\chi_Q$, by Lemma 4.3 we obtain

$$\int_E (T_kf)wdx \le 2^{-k}w(E) + \frac{ck}{\bar{\varphi}^{-1}(2^{2^k})}\int_{\mathbb{R}^n} \Phi(|f(x)|)M_{\varphi(L)}w(x)dx.$$

Combining (4.9) with the latter estimate implies,

$$\begin{aligned} w(E) &\le \frac{1}{4}\int_E (\mathcal{A}_{\mathcal{S},L\log L}f)w dx \le \frac{1}{4}\sum_{k=1}^{\infty}\int_E (T_k f)w dx \\ &\le \frac{1}{4}w(E) + c\left(\sum_{k=1}^{\infty}\frac{k}{\bar\varphi^{-1}(2^{2^k})}\right)\int_{\mathbb{R}^n}\Phi(|f(x)|)M_{\varphi(L)}w(x)dx. \end{aligned}$$

From this,

$$w(E) \le c\left(\sum_{k=1}^{\infty}\frac{k}{\bar\varphi^{-1}(2^{2^k})}\right)\int_{\mathbb{R}^n}\Phi(|f(x)|)M_{\varphi(L)}w(x)dx.$$

Next, using that $\bar\varphi^{-1}(t)\varphi^{-1}(t) \approx t$, we obtain

$$\sum_{k=1}^{\infty}\frac{k}{\bar\varphi^{-1}(2^{2^k})} \le c\sum_{k=1}^{\infty}\int_{2^{2^{k-1}}}^{2^{2^k}}\frac{\log\log(\mathrm{e}^2+t)}{\bar\varphi^{-1}(t)t\log(\mathrm{e}+t)}dt \le cK_\varphi,$$

which, along with the previous estimate, yields (4.8), and therefore, the proof is complete. □

Proof of Lemma 4.2. Denote

$$E = \{x : |\mathcal{T}_{b,\mathcal{S}}f(x)| > 8, Mf(x) \le 1/4\}.$$

By the Fefferman-Stein estimate (1.3) and by homogeneity, it suffices to assume that $\|b\|_{BMO} = 1$ and to show that in this case,

$$w(E) \le cC_\varphi \int_{\mathbb{R}^n}|f|M_{(\Phi\circ\varphi)(L)}w dx.$$

Let

$$\mathcal{S}_k = \{Q \in \mathcal{S} : 4^{-k-1} < |f|_Q \le 4^{-k}\}$$

and for $Q \in \mathcal{S}_k$, set

$$F_k(Q) = \{x \in Q : |b(x) - b_Q| > (3/2)^k\}.$$

If $E \cap Q \ne \emptyset$ for some $Q \in \mathcal{S}$, then $\|f\|_Q \le 1/4$. Therefore, for $x \in E$,

$$\begin{aligned} &|\mathcal{T}_{b,\mathcal{S}}f(x)| \le \sum_{k=1}^{\infty}\sum_{Q\in\mathcal{S}_k}|b(x)-b_Q||f|_Q\chi_Q(x) \\ &\le \sum_{k=1}^{\infty}(3/2)^k\sum_{Q\in\mathcal{S}_k}|f|_Q\chi_Q(x) + \sum_{k=1}^{\infty}\sum_{Q\in\mathcal{S}_k}|b(x)-b_Q||f|_Q\chi_{F_k(Q)}(x) \\ &\equiv \mathcal{T}_1 f(x) + \mathcal{T}_2 f(x). \end{aligned}$$

Let $E_i = \{x \in E : \mathcal{T}_i f(x) > 4\}, i = 1, 2$. Then

$$w(E) \le w(E_1) + w(E_2). \tag{4.10}$$

Lemma 4.3 with $\Psi(t) = t$ yields (with any Young function φ)

$$\int_{E_1}(\mathcal{T}_1 f)w\,dx \le \Big(\sum_{k=1}^{\infty}(3/4)^k\Big)w(E_1)+16\Big(\sum_{k=1}^{\infty}\frac{(3/2)^k}{\bar\varphi^{-1}(2^{2^k})}\Big)\int_{\mathbb{R}^n}|f|M_{\varphi(L)}w\,dx.$$

This estimate, combined with $w(E_1) \le \frac{1}{4}\int_{E_1}(\mathcal{T}_1 f)w\,dx$, implies

$$w(E_1) \le 16\Big(\sum_{k=1}^{\infty}\frac{(3/2)^k}{\bar\varphi^{-1}(2^{2^k})}\Big)\int_{\mathbb{R}^n}|f|M_{\varphi(L)}w\,dx.$$

Since $\bar\varphi^{-1}(t)\varphi^{-1}(t) \approx t$, we obtain

$$\sum_{k=1}^{\infty}\frac{(3/2)^k}{\bar\varphi^{-1}(2^{2^k})} \le c\sum_{k=1}^{\infty}\int_{2^{2^{k-1}}}^{2^{2^k}}\frac{1}{\bar\varphi^{-1}(t)}\frac{dt}{t} \le c\int_1^{\infty}\frac{\varphi^{-1}(t)}{t^2}dt.$$

Hence,

$$w(E_1) \le c\Big(\int_1^{\infty}\frac{\varphi^{-1}(t)}{t^2}dt\Big)\int_{\mathbb{R}^n}|f|M_{\varphi(L)}w\,dx,$$

which by Proposition 4.4 yields

$$w(E_1) \le cC_\varphi\int_{\mathbb{R}^n}|f|M_{(\Phi\circ\varphi)(L)}w\,dx. \tag{4.11}$$

Turn to the estimate of $w(E_2)$. Exactly as in the proof of Lemma 4.3, for $Q \in \mathcal{S}_k$ define disjoint subsets E_Q. Then, by (4.4),

$$|f|_Q \le \frac{8}{|Q|}\int_{E_Q}|f|\,dx.$$

Hence,

$$\begin{aligned} w(E_2) &\le \frac{1}{4}\|\mathcal{T}_2 f\|_{L^1} \\ &\le 2\sum_{k=1}^{\infty}\sum_{Q\in\mathcal{S}_k}\Big(\frac{1}{|Q|}\int_{F_k(Q)}|b-b_Q|w\,dx\Big)\int_{E_Q}|f|. \end{aligned} \tag{4.12}$$

Now we apply twice the generalized Hölder inequality. First, by (2.4),

$$\frac{1}{|Q|}\int_{F_k(Q)}|b-b_Q|w\,dx \le c_n\|w\chi_{F_k(Q)}\|_{L\log L,Q}. \tag{4.13}$$

Second, we use (2.5) with $C(t) = \Phi(t), B(t) = \Phi\circ\varphi(t)$ and A defined by

$$A^{-1}(t) = \frac{C^{-1}(t)}{B^{-1}(t)} = \frac{\Phi^{-1}(t)}{\varphi^{-1}\circ\Phi^{-1}(t)}.$$

Since $\varphi(t)/t$ and Φ are strictly increasing functions, A is strictly increasing, too. Hence, by (2.5), we obtain

$$\begin{aligned}(4.14)\qquad \|w\chi_{F_k(Q)}\|_{L\log L,Q} &\leq 2\|\chi_{F_k(Q)}\|_{A,Q}\|w\|_{(\Phi\circ\varphi),Q}\\ &= \frac{2}{A^{-1}(|Q|/|F_k(Q)|)}\|w\|_{(\Phi\circ\varphi),Q}.\end{aligned}$$

By the John-Nirenberg inequality (2.3), $|F_k(Q)|\leq \alpha_k|Q|$, where $\alpha_k = \min(1, e^{-\frac{(3/2)^k}{2^n e}+1})$. Combining this with (4.13) and (4.14) yields

$$\frac{1}{|Q|}\int_{F_k(Q)}|b-b_Q|w\,dx \leq \frac{c_n}{A^{-1}(1/\alpha_k)}\|w\|_{(\Phi\circ\varphi),Q}.$$

From this and from (4.12) we obtain

$$\begin{aligned}w(E_2) &\leq c_n\sum_{k=1}^{\infty}\frac{1}{A^{-1}(1/\alpha_k)}\sum_{Q\in\mathcal{S}_k}\|w\|_{(\Phi\circ\varphi),Q}\int_{E_Q}|f|\\ &\leq c_n\Big(\sum_{k=1}^{\infty}\frac{1}{A^{-1}(1/\alpha_k)}\Big)\int_{\mathbb{R}^n}|f|M_{(\Phi\circ\varphi)(L)}w(x)dx.\end{aligned}$$

Further, the sum on the right-hand side can be estimated as follows:

$$\begin{aligned}&\sum_{k=1}^{\infty}\frac{1}{A^{-1}(1/\alpha_k)} \leq c\sum_{k=1}^{\infty}\int_{1/\alpha_{k-1}}^{1/\alpha_k}\frac{1}{A^{-1}(t)}\frac{1}{t\log(\mathrm{e}+t)}dt\\ &\leq c\int_1^{\infty}\frac{\varphi^{-1}\circ\Phi^{-1}(t)}{\Phi^{-1}(t)}\frac{1}{t\log(\mathrm{e}+t)}dt \leq c\int_1^{\infty}\frac{\varphi^{-1}\circ\Phi^{-1}(t)}{t^2}dt.\end{aligned}$$

Therefore, by (4.7),

$$w(E_2)\leq c_nC_\varphi\int_{\mathbb{R}^n}|f|M_{(\Phi\circ\varphi)(L)}w(x)dx,$$

which, along with (4.10) and (4.11), completes the proof. □

4.3. **Proof of Corollary 1.3.** The proof follows the same scheme as in the proof of [18, Corollary 1.4], and hence we outline it briefly.

Using that $\log t\leq \frac{t^\alpha}{\alpha}$ for $t\geq 1$ and $\alpha>0$, we obtain

$$M_{L(\log L)^{1+\varepsilon}}w(x)\leq \frac{c}{\alpha^{1+\varepsilon}}M_{L^{1+(1+\varepsilon)\alpha}}w(x).$$

Next we use that for $r_n = 1+\frac{1}{c_n[w]_{A_\infty}}$, $M_{L^{r_n}}w(x)\leq 2Mw(x)$. Hence, if α is such that $(1+\varepsilon)\alpha = \frac{1}{c_n[w]_{A_\infty}}$, then

$$\frac{1}{\varepsilon}M_{L(\log L)^{1+\varepsilon}}w(x)\leq \frac{c_n}{\varepsilon}[w]_{A_\infty}^{1+\varepsilon}Mw(x)\leq \frac{c_n}{\varepsilon}[w]_{A_\infty}^{1+\varepsilon}[w]_{A_1}w(x).$$

This estimate with $\varepsilon = 1/\log(\mathrm{e} + [w]_{A_\infty})$, along with (1.8), completes the proof of Corollary 1.3.

5. Proof of Theorem 1.4

The main role in the proof is played by the following lemma. Denote by $\Omega(b;Q)$ the standard mean oscillation,

$$\Omega(b;Q) = \frac{1}{|Q|}\int_Q |b - b_Q|dx.$$

Lemma 5.1. *Let $\mathscr{D}$ be a dyadic lattice and let $\mathcal{S} \subset \mathscr{D}$ be a γ-sparse family. Assume that $b \in L^1_{loc}$. Then there is a $\frac{\gamma}{2}$-sparse family $\widetilde{\mathcal{S}} \subset \mathscr{D}$ such that $\mathcal{S} \subset \widetilde{\mathcal{S}}$ and for every cube $Q \in \widetilde{\mathcal{S}}$,*

$$|b(x) - b_Q| \le \frac{2^{n+2}}{\gamma} \sum_{R\in\widetilde{\mathcal{S}}, R\subseteq Q} \Omega(b;R)\chi_R(x) \tag{5.1}$$

for a.e. $x \in Q$.

This lemma is based on several known ideas. The first idea is an estimate by oscillations over a sparse family (see [11, 16, 22]) and the second idea is an augmentation process (see [25, Section 5.2]).

Proof of Lemma 5.1. Fix a cube $Q \in \mathcal{S}$. Consider the set

$$E = \Big\{x \in Q : M^d_Q(b - b_Q)(x) > \frac{2^{n+2}}{\gamma}\Omega(b;Q)\Big\},$$

where M^d_Q is the standard dyadic local maximal operator restricted to a cube Q. Then $|E| \le \frac{\gamma}{2^{n+2}}|Q|$.

The Calderón-Zygmund decomposition applied to the function χ_E on Q at height $\lambda = \frac{1}{2^{n+1}}$ produces pairwise disjoint cubes $P_j \in \mathcal{D}(Q)$ such that

$$\frac{1}{2^{n+1}}|P_j| \le |P_j \cap E| \le \frac{1}{2}|P_j|$$

and $|E \setminus \cup_j P_j| = 0$. It follows that $\sum_j |P_j| \le \frac{\gamma}{2}|Q|$ and $P_j \cap E^c \ne \emptyset$. Therefore,

$$|b_{P_j} - b_Q| \le \frac{1}{|P_j|}\int_{P_j} |b - b_Q|dx \le \frac{2^{n+2}}{\gamma}\Omega(b;Q) \tag{5.2}$$

and for a.e. $x \in Q$,

$$|b(x) - b_Q|\chi_{Q\setminus\cup_j P_j} \le \frac{2^{n+2}}{\gamma}\Omega(b;Q). \tag{5.3}$$

Now, we denote by $\mathcal{M}(Q)$ the family of the maximal cubes from $\mathcal{S}$, strictly contained in Q. Let $\mathcal{R}$ be the family of the maximal cubes from

$\mathcal{M}(Q)$ and $\{P_j\}$. Denote $\mathcal{R}=\{R_i\}$. Then the cubes R_i are pairwise disjoint. We have here two possibilities. Either R_i is one of the cubes P_j or $R_i=Q'$, where $Q'\in\mathcal{S}$ and the cube Q' contains some P_j. In the latter case, $Q'\cap E^c\neq\emptyset$, and hence (5.2) holds with P_j replaced by Q'. Therefore, for every $R_i\in\mathcal{R}$,

$$|b_{R_i}-b_Q|\le\frac{1}{|R_i|}\int_{R_i}|b-b_Q|dx\le\frac{2^{n+2}}{\gamma}\Omega(b;Q).$$

Also, since $\cup_jP_j\subset\cup_iR_i$, by (5.3),

$$|b(x)-b_Q|\chi_{Q\setminus\cup_iR_i}\le\frac{2^{n+2}}{\gamma}\Omega(b;Q).$$

From this, we obtain

$$\begin{aligned}(5.4)\quad |b(x)-b_Q|\chi_Q &\le |b(x)-b_Q|\chi_{Q\setminus\cup_iR_i}(x)+\sum_i|b_Q-b_{R_i}|\chi_{R_i}\\ &+\sum_j|b(x)-b_{R_i}|\chi_{R_i}\\ &\le \frac{2^{n+2}}{\gamma}\Omega(b;Q)+\sum_i|b(x)-b_{R_i}|\chi_{R_i}.\end{aligned}$$

Moreover, since $\mathcal{S}$ is γ-sparse,

$$\begin{aligned}\sum_i|R_i| &\le \sum_{Q'\in\mathcal{M}(Q)}|Q'|+\sum_j|P_j|\le(1-\gamma)|Q|+\frac{\gamma}{2}|Q|\\ &= (1-\gamma/2)|Q|.\end{aligned}$$

We now iterate (5.4) and add all new cubes (that is, different from the cubes in $\mathcal{S}$), appearing during the process, to $\mathcal{S}$. Denote the resulting extended family by $\widetilde{\mathcal{S}}$.

Since $\sum_i|R_i|\le(1-\gamma/2)|Q|$, after the iteration of (5.4), we will arrive at (5.1). Also, the same estimate shows that for every cube $Q\in\widetilde{\mathcal{S}}$, the set $E_Q=Q\setminus\cup_{Q'\in\widetilde{\mathcal{S}},Q'\subsetneq Q}Q'$ satisfies $|E_Q|\ge\frac{\gamma}{2}|Q|$, and therefore, $\widetilde{\mathcal{S}}$ is $\frac{\gamma}{2}$-sparse. □

Recall the well-known (see [7] or [25] for a different proof) bound for the sparse operator $\mathcal{A}_{\mathcal{S}}$, where $\mathcal{S}$ is γ-sparse:

$$\|\mathcal{A}_{\mathcal{S}}\|_{L^p(w)}\le c_{\gamma,n,p}[w]_{A_p}^{\max\left(1,\frac{1}{p-1}\right)}\quad(1<p<\infty).\tag{5.5}$$

Proof of Theorem 1.4. By Theorem 1.1 and by duality,

$$
\begin{aligned}
(5.6)\qquad & \|[b,T]\|_{L^p(\mu)\to L^p(\lambda)} \\
& \leq c_n C_T \sum_{j=1}^{3^n} \big(\|\mathcal{T}_{\mathcal{S}_j,b}\|_{L^p(\mu)\to L^p(\lambda)} + \|\mathcal{T}^{\star}_{\mathcal{S}_j,b}\|_{L^p(\mu)\to L^p(\lambda)}\big) \\
& = c_n C_T \sum_{j=1}^{3^n} \big(\|\mathcal{T}^{\star}_{\mathcal{S}_j,b}\|_{L^{p'}(\sigma_\lambda)\to L^{p'}(\sigma_\mu)} + \|\mathcal{T}^{\star}_{\mathcal{S}_j,b}\|_{L^p(\mu)\to L^p(\lambda)}\big),
\end{aligned}
$$

where $\mathcal{S}_j \subset \mathscr{D}^{(j)}$ is $\frac{1}{2\cdot 9^n}$-sparse.

By Lemma 5.1, there are $\frac{1}{4\cdot 9^n}$-sparse families $\widetilde{\mathcal{S}}_j$ containing $\mathcal{S}_j$, and also, for every cube $Q \in \widetilde{\mathcal{S}}_j$,

$$
\int_Q |b(x)-b_Q||f| \leq c_n \sum_{R\in\widetilde{\mathcal{S}}_j, R\subseteq Q} \Omega(b;R)\int_R |f|
$$

$$
\leq c_n\|b\|_{BMO_\nu} \sum_{R\in\widetilde{\mathcal{S}}_j, R\subseteq Q} |f|_R \nu(R) \leq c_n\|b\|_{BMO_\nu}\int_Q \big(\mathcal{A}_{\widetilde{\mathcal{S}}_j}|f|\big)\nu dx.
$$

Therefore,

$$
\mathcal{T}^{\star}_{\widetilde{\mathcal{S}}_j,b}|f|(x) \leq c_n\|b\|_{BMO_\nu}\mathcal{A}_{\widetilde{\mathcal{S}}_j}\big((\mathcal{A}_{\widetilde{\mathcal{S}}_j}|f|)\nu\big)(x).
$$

Hence, applying (5.5) twice yields

$$
\begin{aligned}
(5.7)\quad \|\mathcal{T}^{\star}_{\widetilde{\mathcal{S}}_j,b}\|_{L^p(\mu)\to L^p(\lambda)} &\leq c_{n,p}\|b\|_{BMO_\nu}\|\mathcal{A}_{\widetilde{\mathcal{S}}_j}\|_{L^p(\lambda)}\|\mathcal{A}_{\widetilde{\mathcal{S}}_j}\|_{L^p(\mu)} \\
&\leq c_{n,p}\big([\lambda]_{A_p}[\mu]_{A_p}\big)^{\max\left(1,\frac{1}{p-1}\right)}\|b\|_{BMO_\nu}.
\end{aligned}
$$

From this and from the facts that $\nu = (\mu/\lambda)^{1/p} = (\sigma_\lambda/\sigma_\mu)^{1/p'}$ and $[\sigma_w]_{A_{p'}} = [w]_{A_p}^{\frac{1}{p-1}}$, we obtain

$$
\begin{aligned}
\|\mathcal{T}^{\star}_{\widetilde{\mathcal{S}}_j,b}\|_{L^{p'}(\sigma_\lambda)\to L^{p'}(\sigma_\mu)} &\leq c_{n,p'}\big([\sigma_\mu]_{A_{p'}}[\lambda_\mu]_{A_{p'}}\big)^{\max\left(1,\frac{1}{p'-1}\right)}\|b\|_{BMO_\nu} \\
&= c_{n,p'}\big([\mu]_{A_p}[\lambda]_{A_p}\big)^{\max\left(1,\frac{1}{p-1}\right)}\|b\|_{BMO_\nu},
\end{aligned}
$$

It remains to combine this estimate with (5.6) and (5.7), and to observe that $\mathcal{T}^{\star}_{\mathcal{S}_j,b}|f(x)| \leq \mathcal{T}^{\star}_{\widetilde{\mathcal{S}}_j,b}|f(x)|$. □

Acknowledgement. The third author thanks the Departamento de Matemática of the Universidad Nacional del Sur, for the warm hospitality shown during his visit between February and May 2016.

References

[1] F. Bernicot, D. Frey and S. Petermichl, *Sharp weighted norm estimates beyond Calderón-Zygmund theory*, preprint. Available at http://arxiv.org/abs/1510.00973

[2] S. Bloom, *A commutator theorem and weighted BMO*, Trans. Amer. Math. Soc., **292** (1985), no. 1, 103–122.

[3] D. Chung, C. Pereyra and C. Pérez, *Sharp bounds for general commutators on weighted Lebesgue spaces*, Trans. Amer. Math. Soc., **364** (2012), 1163–1177.

[4] R.R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math., **103** (1976), no. 3, 611–635.

[5] J.M. Conde-Alonso and G. Rey, *A pointwise estimate for positive dyadic shifts and some applications*, Math. Ann., doi:10.1007/s00208-015-1320-y

[6] D. Cruz-Uribe, J.M. Martell and C. Pérez, Weights, extrapolation and the theory of Rubio de Francia. Operator Theory: Advances and Applications, 215. Birkhäuser/Springer Basel AG, Basel, 2011.

[7] D. Cruz-Uribe, J.M. Martell and C. Pérez, *Sharp weighted estimates for classical operators*, Adv. Math., **229** (2012), no. 1, 408–441.

[8] A. Culiuc, F. Di Plinio and Y. Ou, *Domination of multilinear singular integrals by positive sparse forms*, preprint. Available at http://arxiv.org/abs/1603.05317

[9] C. Domingo-Salazar, M.T. Lacey and G. Rey, *Borderline weak type estimates for singular integrals and square functions*, Bull. Lond. Math. Soc., **48** (2016), no. 1, 63–73.

[10] C. Fefferman and E.M. Stein, *Some maximal inequalities*, Amer. J. Math., **93** (1971), 107–115.

[11] N. Fujii, *A proof of the Fefferman-Stein-Strömberg inequality for the sharp maximal functions*, Proc. Amer. Math. Soc. **106** (1989), no. 2, 371–377.

[12] L. Grafakos, Modern Fourier analysis. Second edition. Graduate Texts in Mathematics, 250. Springer, New York, 2009.

[13] I. Holmes, M.T. Lacey and B.D. Wick, *Commutators in the Two-Weight Setting*, preprint. Available at http://arxiv.org/abs/1506.05747

[14] I. Holmes and B.D. Wick, *Two weight inequalities for iterated commutators with Caldern-Zygmund operators*, preprint. Available at http://arxiv.org/abs/1509.03769

[15] T.P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Annals of Math. **175** (2012), no. 3, 1473–1506.

[16] T. Hytönen, *The A_2 theorem: remarks and complements*, Contemp. Math., **612** (2014), 91–106.

[17] T.P. Hytönen, M.T. Lacey and C. Pérez, *Sharp weighted bounds for the q-variation of singular integrals*, Bull. Lond. Math. Soc. **45** (2013), no. 3, 529–540.

[18] T.P. Hytönen and C. Pérez, *The $L(\log L)^\varepsilon$ endpoint estimate for maximal singular integral operators*, J. Math. Anal. Appl. **428** (2015), no. 1, 605–626.

[19] T.P. Hytönen, L. Roncal and O. Tapiola, *Quantitative weighted estimates for rough homogeneous singular integrals*, to appear in Israel J. Math. Available at http://arxiv.org/abs/1510.05789

[20] M.A. Krasnoselskiĭ and Ja.B. Rutickiĭ, Convex functions and Orlicz spaces. P. Noordhoff Ltd., Groningen, 1961.

[21] M.T. Lacey, *An elementary proof of the A_2 bound*, to appear in Israel J. Math. Available at http://arxiv.org/abs/1501.05818

[22] A.K. Lerner, *A pointwise estimate for local sharp maximal function with applications to singular integrals*, Bull. London Math. Soc., **42** (2010), no. 5, 843–856.

[23] A.K. Lerner, *A simple proof of the A_2 conjecture*, Int. Math. Res. Not. 2013, no. 14, 3159-3170.

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

[25] A.K. Lerner and F. Nazarov, *Intuitive dyadic calculus: the basics*, preprint. Available at http://arxiv.org/abs/1508.05639

[26] R. O'Neil, *Fractional integration in Orlicz spaces. I.*, Trans. Amer. Math. Soc., **115** (1965), 300–328.

[27] C. Ortiz-Caraballo, *Quadratic A_1 bounds for commutators of singular integrals with BMO functions*, Indiana Univ. Math. J., **60** (2011), no. 6, 2107–2129.

[28] C. Pérez, *Weighted norm inequalities for singular integral operators*, J. London Math. Soc., **49** (1994), 296–308.

[29] C. Pérez, *Endpoint estimates for commutators of singular integral operators*, J. Funct. Anal. **128** (1995), no. 1, 163–185.

[30] C. Pérez and G. Pradolini, *Sharp weighted endpoint estimates for commutators of singular integrals*, Michigan Math. J., **49** (2001), no. 1, 23-37.

[31] C. Pérez and I.P. Rivera-Ríos, *Borderline weighted estimates for commutators of singular integrals*, to appear in Israel J. Math. Available at http://arxiv.org/abs/1507.08568

[32] C. Pérez and I.P. Rivera-Ríos, *Three observations on commutators of Singular Integral Operators with BMO functions*, AWM-Springer Series, Harmonic Analysis, Partial Differentail Equations, Complex Analysis, Banach Spaces, and Operator Theory. Celebrating Cora Sadosky's Life. Vol. 2, to appear. Available at http://arxiv.org/abs/1601.03193

[33] M.M. Rao and Z.D. Ren, Theory of Orlicz Spaces. Monographs and Textbooks in Pure and Applied Mathematics, 146. Marcel Dekker, 1991.

[34] M.C. Reguera and C. Thiele, *The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$*, Math. Res. Lett. **19** (2012), no. 1, 1–7.

Department of Mathematics, Bar-Ilan University, 5290002 Ramat Gan, Israel
E-mail address: lernera@math.biu.ac.il

Departamento de Matemática, Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina
E-mail address: sombrosi@uns.edu.ar

IMUS & Departamento de Análisis Matemático, Universidad de Sevilla, Sevilla, Spain
E-mail address: petnapet@gmail.com