

SHARP TWO-WEIGHT INEQUALITIES FOR SINGULAR INTEGRALS, WITH APPLICATIONS TO THE HILBERT TRANSFORM AND THE SARASON CONJECTURE

D. CRUZ-URIBE, SFO, J. M. MARTELL, AND C. PÉREZ

ABSTRACT. We prove two-weight norm inequalities for Calderón-Zygmund singular integrals that are sharp for the Hilbert transform and for the Riesz transforms. In addition, we give results for the dyadic square function and for commutators of singular integrals. As an application we give new results for the Sarason conjecture on the product of unbounded Toeplitz operators on Hardy spaces

1. INTRODUCTION

1.1. Background. A long-standing problem in harmonic analysis has been to characterize the weights governing strong-type norm inequalities for classical operators. To be precise: given an operator T and p , $1 < p < \infty$, determine sufficient conditions on a pair of weights (i.e., non-negative, measurable functions) (u, v) such that for all $f \in L^p(v^p)$,

$$(1.1) \quad \int_{\mathbb{R}^n} |u(x) T f(x)|^p dx \leq C \int_{\mathbb{R}^n} |v(x) f(x)|^p dx.$$

This problem was originally posed in the early 1970's for the Hardy-Littlewood maximal operator and for the Hilbert transform on the real line, but it was soon expanded to include a variety of operators—singular integrals, fractional integrals, and square functions—on \mathbb{R}^n . While a great deal of progress has been made, many questions remain open even for the Hilbert transform.

For many of these problems, inequality (1.1) is usually stated in an equivalent form:

$$(1.2) \quad \int_{\mathbb{R}^n} |T f(x)|^p U(x) dx \leq C \int_{\mathbb{R}^n} |f(x)|^p V(x) dx,$$

where $U = u^p$ and $V = v^p$. But for our purposes (1.1) is a more suitable form as it makes the statement of our main results more elegant.

2000 *Mathematics Subject Classification.* 42B20,42B25,47B35.

Key words and phrases. Weights, Hilbert transform, singular integral operators, Sarason conjecture, Toeplitz operator, maximal functions, commutators.

The first author is partially supported by the Stewart faculty development fund of Trinity College; he would also like to thank his advisor, Donald Sarason, for posing the problem discussed in Section 1.4. The second author is partially supported by Spanish Ministerio de Educación y Ciencia “Programa Ramón y Cajal, 2005” and by Grant MTM2004-00678 from the same institution; the third author is partially supported by Grant MTM2006-05622 from the same institution.

The purpose of this paper is to give new two-weight norm inequalities for singular integrals and other operators that are sharp for the Hilbert and Riesz transforms. To put our results into context, we will sketch the outlines of some earlier work. For more information on the history of this problem, we refer the reader to Muckenhoupt [28], Dynkin and Osilenker [15], García-Cuerva and Rubio de Francia [18], and Duoandikoetxea [13].

The earliest weighted norm inequalities were for the one-weight problem (i.e., when $u = v$). Muckenhoupt [27], and Hunt, Muckenhoupt and Wheeden [20] showed that for the maximal operator and for the Hilbert transform on the real line, (1.1) held if and only if u^p satisfied the so-called A_p condition: there exists a finite constant C such that for all intervals Q ,

$$(1.3) \quad \left(\frac{1}{|Q|} \int_Q u(x)^p dx \right)^{\frac{1}{p}} \left(\frac{1}{|Q|} \int_Q u(x)^{-p'} dx \right)^{\frac{1}{p'}} \leq C.$$

The proof was simplified by Coifman and Fefferman [3] and extended to Calderón-Zygmund singular integrals on \mathbb{R}^n (with intervals replaced by cubes in (1.3)).

It was immediately conjectured that in the two-weight case, the corresponding two-weight A_p condition,

$$(1.4) \quad \left(\frac{1}{|Q|} \int_Q u(x)^p dx \right)^{\frac{1}{p}} \left(\frac{1}{|Q|} \int_Q v(x)^{-p'} dx \right)^{\frac{1}{p'}} \leq C < \infty,$$

was necessary and sufficient for these operators to be bounded from $L^p(v^p)$ to $L^p(u^p)$. However, while this condition is necessary for the maximal operator and for the Hilbert transform, it is not sufficient: see Muckenhoupt and Wheeden [30]. Sawyer [44] gave a necessary and sufficient condition for the maximal operator which involves the operator itself. Cotlar and Sadosky [4, 5] gave a necessary and sufficient condition for the Hilbert transform which is reminiscent of the Helson-Szegö theorem and is grounded in operator theory. However, their condition is difficult to check and does not readily extend to higher dimensions and general singular integrals.

Following these results, a great deal of effort was devoted to finding stronger conditions related to the more geometric two-weight A_p condition and that are sufficient for (1.1) to hold for a variety of operators, especially singular integrals. In passing, we note the work of Muckenhoupt and Wheeden [30], Fujii [17], Katz and Pereyra [22], Leckband [25], Rakotondratsimba [39, 40], Wilson [52], and Pérez [34].

An important result in this direction is due to Neugebauer [32]: he showed that if the pair of weights (u, v) is such that for some $r > 1$ the pair (u^r, v^r) satisfies (1.4), then (1.1) holds for singular integrals. He did not prove this directly; rather, by applying the ideas on factorization of weights due to Rubio de Francia, he showed that there exists $w \in A_p$ such that $c_1 u \leq w \leq c_2 v$ if and only if $(u^r, v^r) \in A_p$ for some $r > 1$. Two-weight inequalities for singular integrals and other operators then follow immediately from the one-weight case.

We can restate Neugebauer's result as follows. Given a cube Q , write

$$\|u\|_{p,Q} = \left(\frac{1}{|Q|} \int_Q |u(x)|^p dx \right)^{1/p}$$

for the normalized L^p norm on Q . The A_p condition is then equivalent to

$$\|u\|_{p,Q} \|v^{-1}\|_{p',Q} \leq C < \infty,$$

and the condition that $(u^r, v^r) \in A_p$ can be rewritten as

$$\|u\|_{rp,Q} \|v^{-1}\|_{rp',Q} \leq C < \infty.$$

In other words, if we replace the normalized L^p and $L^{p'}$ norms in the A_p condition by larger norms (in the scale of Lebesgue spaces), then we get a sufficient condition for (1.1) to hold for singular integrals and other operators. We refer to these larger norms as ‘‘power bumps.’’

P erez [35, 36] first considered the question of whether power bumps could be replaced by other function space norms larger than the L^p norm but smaller than the L^{rp} norm. He showed that for the maximal operator and fractional integrals certain norms in the scale of Orlicz spaces, the so-called ‘‘Orlicz bumps’’, are sufficient.

To state his results we need several definitions. Given a Young function $B : [0, \infty) \rightarrow [0, \infty)$, and a cube Q , define the normalized Luxemburg norm on Q by

$$\|u\|_{B,Q} = \inf \left\{ \lambda > 0 : \frac{1}{|Q|} \int_Q B \left(\frac{|u(x)|}{\lambda} \right) dx \leq 1 \right\}.$$

If $B(t) = t^p$, then $\|u\|_{B,Q} = \|u\|_{p,Q}$ and the Luxemburg norm reduces to the L^p norm. When $B(t) = t^p \log(e+t)^a$ we get the norm on the Zygmund spaces $L^p(\log L)^a$. When used to define an A_p type condition, this norm is referred to as a ‘‘log bump.’’

Given a Young function B , let \bar{B} denote its associate function: the Young function with the property that $t \leq B^{-1}(t)\bar{B}^{-1}(t) \leq 2t$, $t > 0$. If $B(t) = t^p$, then $\bar{B}(t) = t^{p'}$; if $B(t) = t^p \log(e+t)^a$, then $\bar{B}(t) \approx t^{p'} \log(e+t)^{-ap'/p}$.

The following growth condition on Young functions plays an important role in determining suitable Orlicz bumps for generalizing the A_p condition.

Definition 1.1. *Given p , $1 < p < \infty$, a Young function B satisfies the B_p condition if for some $c > 0$,*

$$(1.5) \quad \int_c^\infty \frac{B(t)}{t^p} \frac{dt}{t} < \infty.$$

If $B(t) = t^q$, $1 < q < p$, then it is immediate that $B \in B_p$. More interesting examples are given by the functions

$$B(t) = \frac{t^p}{\log(e+t)^{1+\delta}}, \quad \delta > 0,$$

$$B(t) = \frac{t^p}{\log(e+t) \log \log(e^e+t)^{1+\delta}}, \quad \delta > 0.$$

The B_p condition was introduced in [36] where it was used in to state and prove sharp two-weight norm inequalities for the Hardy-Littlewood maximal function. If B is a Young function such that $\bar{B} \in B_p$, and the pair of weights (u, v) is such that for every cube Q ,

$$(1.6) \quad \|u\|_{p,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

then (1.1) holds for the Hardy-Littlewood maximal function. Furthermore, the B_p condition is necessary: if (1.1) holds and (u, v) satisfy (1.6), then $\bar{B} \in B_p$. Note that unlike in the original result by Neugebauer, there is no bump on the weight u .

Via a discretization argument, the same techniques were applied in [34] to prove weighted norm inequalities for the fractional integral operators I_α , $0 < \alpha < n$. Let A and B be Young functions such that $\bar{A} \in B_{p'}$ and $\bar{B} \in B_p$. If (u, v) is a pair of weights such that

$$(1.7) \quad \ell(Q)^\alpha \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

then

$$\int_{\mathbb{R}^n} |u(x)I_\alpha f(x)|^p dx \leq C \int_{\mathbb{R}^n} |v(x)f(x)|^p dx.$$

The condition (1.7) can be viewed as a two-weight version of the Chang-Wilson-Wolff condition [2] for Schrödinger operators which is an improvement of the well-known Fefferman-Phong condition [16]. This result for fractional integrals immediately suggested the following conjecture:

Conjecture. *If A and B are Young functions such that $\bar{A} \in B_{p'}$ and $\bar{B} \in B_p$, and if the pair of weights (u, v) is such that for every cube Q ,*

$$(1.8) \quad \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

then (1.1) holds for Calderón-Zygmund singular integrals.

An important special case of this conjecture is when A and B are log bumps:

$$A(t) = t^p \log(e + t)^{p-1+\delta}, \quad B(t) = t^{p'} \log(e + t)^{p'-1+\delta}, \quad \delta > 0.$$

Our conjecture is closely connected to an old conjecture of Muckenhoupt and Wheeden [29]: if the pair (u, v) is such that the maximal operator M satisfies

$$(1.9) \quad M : L^p(v^p) \rightarrow L^p(u^p) \quad \text{and} \quad M : L^{p'}(u^{-p'}) \rightarrow L^{p'}(v^{-p'}),$$

then the Hilbert transform is bounded from $L^p(v^p)$ to $L^p(u^p)$. By the results in [36] described above, (1.8) is sufficient for both inequalities in (1.9) to hold, so our conjecture is a special case of theirs.

Our conjecture is known to be true in a number of special cases. When A and B are power bumps—i.e., $A(t) = t^{rp}$ and $B(t) = t^{rp'}$, $r > 1$ —then our conjecture reduces to the theorem of Neugebauer stated above. His result was improved in [11], where it was shown that it is sufficient to take A a power bump and B such that $\bar{B} \in B_p$. In [7] it was shown that if A is a large Orlicz bump, e.g., if

$$A(t) \approx t^p \exp[\log(e + t^p)^r], \quad 0 < r < 1,$$

then the conjecture is true. However, it was also shown in this paper that such functions represent the best that can be gotten using the techniques in [11]; they cannot be used to prove the full conjecture or even the case when A is a log bump.

A related but weaker version of our conjecture was proved by Treil, Volberg and Zheng [48] for the periodic Hilbert transform (i.e., the conjugate function) on the unit circle. For $z \in \mathbb{D}$, let ϕ_z be the Möbius transform in the closed unit disk,

$$\phi_z(w) = \frac{z - w}{1 - \bar{z}w}, \quad w \in \bar{\mathbb{D}}.$$

If A and B are Young functions such that $\bar{A} \in B_{p'}$ and $\bar{B} \in B_p$, and if (u, v) is a pair of weights such that

$$(1.10) \quad \sup_{z \in \mathbb{D}} \|u \circ \phi_z\|_{A, \partial \mathbb{D}} \|v^{-1} \circ \phi_z\|_{B, \partial \mathbb{D}} < \infty,$$

then the periodic Hilbert transform is bounded from $L^p(v^p, \partial \mathbb{D})$ to $L^p(u^p, \partial \mathbb{D})$.

Another result closely related to our conjecture was proved in [9]. There it was shown that if A is the log bump $A(t) = t^p \log(e + t)^{p-1+\delta}$ and if the pair of weights (u, v) is such that for every cube Q ,

$$(1.11) \quad \|u\|_{A, Q} \|v^{-1}\|_{p', Q} \leq C < \infty,$$

then Calderón-Zygmund singular integrals satisfy the weak (p, p) inequality

$$(1.12) \quad u^p(\{x \in \mathbb{R}^n : |Tf(x)| > t\}) \leq \frac{C}{t^p} \int_{\mathbb{R}^n} |v(x)f(x)|^p dx.$$

Note that condition (1.11) is a special case of (1.6), and it is natural to conjecture that (1.12) holds if A such that $\bar{A} \in B_{p'}$. This is a special case of another conjecture due to Muckenhoupt and Wheeden [29]: if the maximal operator satisfies $M : L^{p'}(u^{-p'}) \rightarrow L^p(v^{-p})$, then the Hilbert transform satisfies (1.12).

1.2. Results for singular integrals. Our main results improve all previous work by allowing us to take A to be a log bump. Our first theorem is a sharp inequality for the Hilbert transform.

Theorem 1.2. *Given p , $1 < p < \infty$, suppose the pair of weights (u, v) satisfies*

$$(1.13) \quad \|u\|_{A, Q} \|v^{-1}\|_{B, Q} \leq C < \infty,$$

where $A(t) = t^p \log(e + t)^{p-1+\delta}$, $\delta > 0$, and $\bar{B} \in B_p$. Then

$$(1.14) \quad \int_{\mathbb{R}} |u(x) Hf(x)|^p dx \leq C \int_{\mathbb{R}} |v(x) f(x)|^p dx.$$

Further, this inequality is sharp since it does not hold in general if we take $\delta = 0$ in the definition of A .

A counter-example showing that (1.14) need not hold if $\delta = 0$ when $p = 2$ is given in [9]. The example there is a pair of weights for which (1.2) does not hold: $(U, M_\Phi U)$, where $\Phi(t) = t \log(e + t)$, and M_Φ is the Orlicz maximal operator

$$(1.15) \quad M_\Phi f(x) = \sup_{Q \ni x} \|f\|_{\Phi, Q}.$$

(See Lemma 2.8 below.) By a change of variables in the definition of the Luxemburg norm it is easy to see that the pair of weights $u = U^{1/2}$, $v = (M_{\Phi}U)^{1/2}$ satisfies (1.13) with $A(t) = t^2 \log(e + t)$.

Theorem 1.2 is a special case of a more general result which holds on \mathbb{R}^n , provided $p > n$. Recall that a Calderón-Zygmund singular integral T is a singular convolution operator,

$$Tf(x) = \text{p.v.} \int_{\mathbb{R}^n} K(x-y)f(y) dy,$$

where the kernel K is continuously differentiable on $\mathbb{R}^n \setminus \{0\}$, has zero average on the unit sphere, and for all $x \neq 0$,

$$|K(x)| \leq \frac{C}{|x|^n} \quad \text{and} \quad |\nabla K(x)| \leq \frac{C}{|x|^{n+1}}.$$

More generally, we may assume that T is a Calderón-Zygmund operator. For a precise definition see Duoandikoetxea [13].

Theorem 1.3. *Let T be a Calderón-Zygmund singular integral. Fix p , $n < p < \infty$. Suppose (u, v) is a pair of weights such that for all cubes Q ,*

$$(1.16) \quad \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

where $A(t) = t^p \log(e+t)^{p-1+\delta}$, $\delta > 0$, and $\bar{B} \in B_p$. Then T satisfies the strong (p, p) inequality

$$(1.17) \quad \int_{\mathbb{R}^n} |u(x) Tf(x)|^p dx \leq C \int_{\mathbb{R}^n} |v(x) f(x)|^p dx.$$

Further, this result is sharp in the sense that there exists a family of pairs of weights (u, v) such that (1.16) holds with $\delta = 0$, but (1.17) does not hold for all of the Riesz transforms.

The sharpness of Theorem 1.3 comes from a necessary condition proved in [34]. Translated to our setting (the results there are stated in terms of inequality (1.2)), it shows that if the pairs of weights $(u, M_A u)$ (which clearly satisfy (1.16)) are such that (1.17) holds for all n of the Riesz transforms, then $\delta > 0$. By contraposition, if $\delta = 0$ then (1.17) must fail for at least one of the Riesz transforms.

The restriction that $p > n$ in Theorem 1.3 seems unnatural, but despite repeated efforts we cannot eliminate it. If $n \geq 2$, then by duality we have that (1.17) holds for $1 < p < n'$ or $p > n$ if A and B are both log bumps: $A(t) = t^p \log(e+t)^{p-1+\delta}$ and $B(t) = t^{p'} \log(e+t)^{p'-1+\delta}$, $\delta > 0$. However, this still leaves the gap $n' \leq p \leq n$.

Our next result shows that we can fill this gap if we replace A by a larger log bump.

Theorem 1.4. *Let T be a Calderón-Zygmund singular integral. Given p , $1 < p < \infty$, suppose (u, v) is a pair of weights such that for all cubes Q ,*

$$(1.18) \quad \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

where $A(t) = t^p \log(e + t)^{2p-1+\delta}$, $\delta > 0$, and $\bar{B} \in B_p$. Then T satisfies the strong (p, p) inequality

$$(1.19) \quad \int_{\mathbb{R}^n} |u(x) T f(x)|^p dx \leq C \int_{\mathbb{R}^n} |v(x) f(x)|^p dx.$$

The proofs of both Theorems 1.3 and 1.4 involve careful discretization arguments using the properties of Calderón-Zygmund cubes. They are very similar in spirit, though not in detail, to the discretization argument used to prove two-weight norm inequalities for fractional integrals in [35]. The problem with this approach is that there does not exist as good a technique for discretizing singular integrals as exists for fractional integrals. Consequently, we need to argue more obliquely using the sharp maximal operator (explicitly in the proof of Theorem 1.4 and in essence in the proof of Theorem 1.3). This leads directly to the technical obstacles which prevent us from proving the full conjecture we described above.

In particular, in both proofs we use the following property of log bumps: given $A(t) = t^p \log(e + t)^{p-1+\delta}$, $\delta > 0$, then $\bar{A} \in B_{p'}$ and there exists q , $0 < q < 1$, such that if $C(t) = A(t^{1/q})$, then $\bar{C} \in B_{(p/q)'}$. This property does not hold for arbitrary Young functions: a counter-example is given by $A(t) = t^p \log(e + t)^{p-1} \log \log(e + t)^{p-1+\delta}$. Details are left to the reader.

Key to the proof of Theorem 1.4 is the pointwise inequality [1]:

$$(1.20) \quad M_q^\#(Tf)(x) = M^\#(|Tf|^q)(x)^{1/q} \leq CMf(x),$$

for some $0 < q < 1$, where $M^\#$ is the sharp maximal operator of Fefferman-Stein. Vector-valued singular integrals satisfy essentially the same inequality [38]: if $0 < q < 1$ and $1 < r < \infty$ there exists a constant such that

$$M_q^\#(\|\{Tf_j\}_j\|_{\ell^r})(x) \leq CM(\|\{f_j\}_j\|_{\ell^r})(x).$$

Therefore, as a corollary to the proof of Theorem 1.4 we get two-weight estimates for vector-valued singular integrals. On the other hand, it is not difficult to observe that the proof of Theorem 1.3 can be carried out for vector-valued singular integrals and thus we get better conditions on (u, v) in the range $n < p < \infty$. Details are left to the reader.

Corollary 1.5. *Let T be a Calderón-Zygmund singular integral. Given p, r with $1 < p, r < \infty$, suppose (u, v) satisfy (1.18). Then*

$$\left\| \left(\sum_j |u T f_j|^r \right)^{\frac{1}{r}} \right\|_{L^p(\mathbb{R}^n)} \leq C \left\| \left(\sum_j |v f_j|^r \right)^{\frac{1}{r}} \right\|_{L^p(\mathbb{R}^n)}.$$

Moreover, the same estimate holds if $1 < r < \infty$, $p > n$ and (u, v) satisfy (1.16).

Remark 1.6. Other operators, including some pseudo-differential operators and square functions, satisfy inequality (1.20), and so similar weighted norm inequalities hold for them. For examples see [1] and [11].

1.3. Results for other operators. The proofs of Theorems 1.3 and 1.4 can be adapted to give results for other operators. Here we consider two: the dyadic square function and commutators of singular integrals.

Dyadic square functions. We first consider the dyadic square function. Let Δ denote the set of dyadic cubes in \mathbb{R}^n , and for each $m \in \mathbb{Z}$, let $\Delta_m = \{Q \in \Delta : \ell(Q) = 2^m\}$. For each $Q \in \Delta$, let \widehat{Q} denote the dyadic parent of Q : if $Q \in \Delta_m$, the unique cube $\widehat{Q} \in \Delta_{m+1}$ such that $Q \subset \widehat{Q}$. Given a function f , let $f_Q = |Q|^{-1} \int_Q f(x) dx$. For each f , the dyadic square function, $S_d f$, is defined by

$$S_d f(x) = \left(\sum_{Q \in \Delta} |f_Q - f_{\widehat{Q}}|^2 \chi_Q(x) \right)^{1/2}.$$

Theorem 1.7. *Given p , $1 < p < \infty$, suppose (u, v) is a pair of weights such that for all dyadic cubes Q ,*

$$(1.21) \quad \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

where $A(t) = t^p \log(e+t)^{p-1+\delta}$, $\delta > 0$, and $\bar{B} \in B_p$. Then the dyadic square function satisfies the strong (p, p) inequality

$$(1.22) \quad \int_{\mathbb{R}^n} (u(x) S_d f(x))^p dx \leq C \int_{\mathbb{R}^n} |v(x) f(x)|^p dx.$$

The proof of Theorem 1.7 is nearly identical to that of Theorem 1.3; the difference is that the square function is sufficiently localized that we can eliminate the restriction on p . Given the close connection between square functions and singular integrals, we take this result as evidence that the restriction on p in Theorem 1.3 is not necessary.

Theorem 1.7 is related to two-weight norm inequalities for the dyadic square function due to Uchiyama [49] and Cruz-Uribe and Pérez [10]. They showed that for any weight u ,

$$\begin{aligned} \int_{\mathbb{R}^n} S_d f(x)^p u(x) dx &\leq C \int_{\mathbb{R}^n} |f(x)|^p M u(x) dx, & 1 < p \leq 2, \\ \int_{\mathbb{R}^n} S_d f(x)^p u(x) dx &\leq C \int_{\mathbb{R}^n} |f(x)|^p M_C u(x) dx, & 2 < p < \infty, \end{aligned}$$

where $C(t) = t \log(e+t)^{p/2-1+\delta}$, $\delta > 0$, and M_C is the Orlicz maximal operator (1.15). Similar but weaker inequalities follow from Theorem 1.7: it is straightforward to see that weights of the form $(u^{1/p}, (M_D u)^{1/p})$, where $D(t) = t \log(e+t)^{p-1+\delta}$, $\delta > 0$, satisfy (1.21). On the other hand, one can also find pairs of weights which satisfy (1.21) which cannot be written in this form. It is tempting to speculate that Theorem 1.7 can be improved to include all of these results as special cases.

1.3.1. Commutators. The second class of operators we consider are commutators of singular integrals. Given a Calderón-Zygmund singular integral T and $b \in BMO$, define the first order commutator, $[b, T]$, by

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

These operators are more singular than the associated singular integrals, and so a larger log bump is required on both weights.

Theorem 1.8. *Let T be a Calderón-Zygmund singular integral and let $b \in BMO$. Given p , $1 < p < \infty$, suppose that for all cubes Q the pair of weights (u, v) satisfies*

$$(1.23) \quad \|u\|_{A,Q} \|v^{-1}\|_{B,Q} \leq C < \infty,$$

where $A(t) = t^p \log(e+t)^{3p-1+\delta}$, and $B(t) = t^{p'} \log(e+t)^{2p'-1+\delta}$, $\delta > 0$. Then

$$(1.24) \quad \int_{\mathbb{R}^n} |u(x)[b, T]f(x)|^p dx \leq C \int_{\mathbb{R}^n} |v(x)f(x)|^p dx.$$

Theorem 1.8 improves a result in [11], where the same inequality was proved assuming that A is a power bump: $A(t) = t^{rp}$, $r > 1$.

Remark 1.9. An analogous result holds for higher order commutators T_b^k , with $k \geq 2$. (For $k = 1$, $T_b^1 = [b, T]$.) These are defined inductively by $T_b^k = [b, T_b^{k-1}]$. In this case the condition imposed on the pair of weights (u, v) is (1.23) with $A(t) = t^p \log(e+t)^{(k+2)p-1+\delta}$ and $B(t) = t^{p'} \log(e+t)^{(k+1)p'-1+\delta}$, $\delta > 0$. The proof is essentially the same and some details are given in Remark 5.7 below.

Remark 1.10. We conjecture that Theorem 1.8 can be improved by taking $A(t) = t^p \log(e+t)^{2p-1+\delta}$ —the commutator should require one more log term on each weight than the associated singular integral.

1.4. Application to the Sarason conjecture. Theorem 1.2 has an application to an open problem in operator theory on the unit disk. This problem was first posed by Sarason (see Khavin and Nikol'skiĭ [23]) and is referred to as the Sarason conjecture. To state it we recall some basic facts about operator theory on the unit circle. (For complete information, see Koosis [24].)

Given a function $f \in L^1(\partial\mathbb{D})$, we define the periodic Hilbert transform of f , also known as the conjugate function of f , by

$$\tilde{f}(e^{i\theta}) = \tilde{H}f(e^{i\theta}) = \frac{1}{\pi} \int_0^\pi \frac{f(e^{i(\theta-t)}) - f(e^{i(\theta+t)})}{2 \tan(t/2)} dt.$$

The periodic Hilbert transform is a Calderón-Zygmund singular integral and so is a bounded operator on $L^2(\partial\mathbb{D})$. Define the Riesz projection operator P by

$$Pf(e^{i\theta}) = \frac{f(e^{i\theta}) + \tilde{H}f(e^{i\theta}) + \hat{f}(0)}{2}.$$

Then P is also bounded on $L^2(\partial\mathbb{D})$, and in fact is the orthogonal projection from $L^2(\partial\mathbb{D})$ to the Hardy space $H^2(\partial\mathbb{D})$, the closure of the analytic polynomials in $L^2(\partial\mathbb{D})$.

Given a function $h \in L^2(\partial\mathbb{D})$, define the Toeplitz operator with symbol h by

$$T_h f(e^{i\theta}) = P(hf)(e^{i\theta}).$$

The Toeplitz operator T_h is densely defined on $H^2(\partial\mathbb{D})$ and is a bounded operator on $H^2(\partial\mathbb{D})$ if and only if $h \in L^\infty(\partial\mathbb{D})$. Toeplitz operators have been intensively studied and appear in many problems in operator theory.

The composition of unbounded Toeplitz operators arises in the application of de Brange spaces to the study of the exposed points of $H^1(\partial\mathbb{D})$. An exposed point of H^1 is a point on the unit ball such that there exists a real linear functional which attains its maximum on the unit ball at that point and nowhere else. In [42], Sarason conjectured a deep characterization of the exposed points of H^1 in terms of de Brange spaces. In [43] he proved part of this conjecture; central to his proof was showing that certain explicit examples of unbounded Toeplitz operators had a product that was a bounded operator on $H^2(\partial\mathbb{D})$. Based on these examples, he made the following conjecture [23, p. 318]: if f and g are outer functions in $H^2(\partial\mathbb{D})$, then the product $T_f T_{\bar{g}}$ is a bounded operator on $H^2(\partial\mathbb{D})$ if and only if

$$(1.25) \quad \sup_{z \in \mathbb{D}} P_z(|f|^2) P_z(|g|^2) < \infty,$$

where, if $z = re^{i\theta}$, $P_z(\cdot)$ denotes convolution with the Poisson kernel

$$P_r(\theta) = \frac{1 - r^2}{|1 - re^{i\theta}|^2}.$$

Sarason also pointed out that (1.25) is very similar to the two-weight A_2 condition.

Initially it was widely believed that the Sarason conjecture was true. Treil (see [23]) showed that (1.25) is a necessary condition for $T_f T_{\bar{g}}$ to be a bounded operator on $H^2(\partial\mathbb{D})$. It was shown in [6] that (1.25) is necessary and sufficient for $T_f T_{\bar{g}}$ to be bounded and invertible provided that fg and $(fg)^{-1}$ are bounded functions. Zheng [55] showed that if for some $\epsilon > 0$,

$$(1.26) \quad \sup_{z \in \mathbb{D}} P_z(|f|^{2+\epsilon}) P_z(|g|^{2+\epsilon}) < \infty,$$

then $T_f T_{\bar{g}}$ is a bounded operator.

However, Nazarov [31] constructed a delicate counter-example which showed that the Sarason conjecture, as stated, is false. As was the case for two-weight norm inequalities, attention then shifted to finding sufficient conditions for $T_f T_{\bar{g}}$ to be bounded which resemble (1.25) and (1.26). This question is still referred to (loosely) as the Sarason conjecture.

There is a close connection between the Sarason conjecture and two-weight norm inequalities. This connection is best shown by the following diagram, which first appeared in [6]:

$$\begin{array}{ccc} H^2 & \xrightarrow{T_f T_{\bar{g}}} & H^2 \\ M_{\bar{g}} \downarrow & & \uparrow T_f = M_f \\ L^2(|g|^{-2}) & \xrightarrow{P} & H^2(|f|^2) \end{array}$$

Here $M_{\bar{g}}$ denotes multiplication by \bar{g} , M_f multiplication by f , and $H^2(|f|^2)$ is the closure in $L^2(\partial\mathbb{D})$ of the set of functions $p_0 f$, where p_0 is an analytic polynomial. Since f is analytic, it is clear that on $H^2(|f|^2)$, $T_f = M_f$ and it is an isometry onto $H^2(\partial\mathbb{D})$. Similarly, $M_{\bar{g}}$ is a bounded map from H^2 into $L^2(|g|^{-2})$. Hence, a sufficient condition for $T_f T_{\bar{g}}$ to be bounded is for P to be bounded, or equivalently, for the periodic Hilbert transform to be a bounded operator from $L^2(|g|^{-2})$ to $L^2(|f|^2)$.

Furthermore, the converse is true. To see this, replace H^2 by L^2 in the diagram. then $M_{\bar{g}}$ becomes an isometry on the lefthand side, so $T_f T_{\bar{g}}$ is bounded on this larger domain if and only if P is bounded. However, $L^2 = H^2 \oplus \overline{H^2}$, and since \bar{g} is co-analytic, $\bar{g}\overline{H^2}$ is in the kernel of the projection operator. Thus, $\overline{H^2}$ is in the kernel of $T_f T_{\bar{g}}$, and so $T_f T_{\bar{g}}$ is bounded on L^2 exactly when it is bounded on H^2 .

Viewed from this perspective, Zheng's result becomes an immediate consequence of the theorem of Neugebauer discussed above. By this theorem (adapted to the unit circle), the periodic Hilbert transform is bounded from $L^2(|g|^{-2})$ to $L^2(|f|^2)$ if for some $\epsilon > 0$ there is a finite constant C such that for every arc $I \subset \partial\mathbb{D}$,

$$(1.27) \quad \left(\frac{1}{|I|} \int_I |f(e^{i\theta})|^{2+\epsilon} d\theta \right) \left(\frac{1}{|I|} \int_I |g(e^{i\theta})|^{2+\epsilon} d\theta \right) \leq C.$$

If we let $I_z = (-|z|, |z|)$, then $P_z(\theta) \geq c\chi_{I_z}/|I_z|$, so we have that (1.26) implies (1.27).

Similarly, Treil, Volberg and Zheng [48] applied condition (1.10) discussed above to show that $T_f T_{\bar{g}}$ is bounded if

$$\sup_{z \in \mathbb{D}} \left\| |f| \circ \phi_z \right\|_{A, \partial\mathbb{D}} \left\| |g| \circ \phi_z \right\|_{B, \partial\mathbb{D}} < \infty,$$

where A and B are Young functions such that $\bar{A}, \bar{B} \in B_2$.

As a consequence of Theorem 1.2 (whose proof immediately extends to the unit circle) we can improve these results and give a new solution to the Sarason conjecture.

Theorem 1.11. *Let f, g be outer functions in $H^2(\partial\mathbb{D})$. If for every arc $I \subset \partial\mathbb{D}$,*

$$(1.28) \quad \|f\|_{A, I} \|g\|_{B, I} \leq C < \infty,$$

where $A(t) = B(t) = t^2 \log(e + t)^{1+\delta}$, $\delta > 0$ (or more generally, $\bar{B} \in B_2$), then $T_f T_{\bar{g}}$ is a bounded operator on $H^2(\partial\mathbb{D})$. Furthermore, this result is sharp in the sense that it does not hold in general when $\delta = 0$.

The counter-example when $\delta = 0$ is actually for the boundedness of the periodic Hilbert transform from $L^2(|g|^{-2})$ to $L^2(|f|^2)$, which, as we noted above, is equivalent to the boundedness of $T_f T_{\bar{g}}$. It is a modification of the counter-example for Theorem 1.2 from [9]. The example there has its bad behavior at infinity; it can be converted to an example on the interval $[-\pi, \pi]$ (equivalently, on the unit circle) by making the change of variables $x \mapsto 1/x$. The details are straightforward and are left to the reader.

Remark 1.12. While the original Sarason conjecture is cast in terms of complex analysis, Theorem 1.11 is strictly a real-variable result. This is not unreasonable: since f and g are outer functions, they are determined by their boundary values on the

unit circle, so complex analysis does not necessarily come into play. Nevertheless, the necessary condition (1.25) is strictly stronger than the two-weight A_2 condition (they are equivalent if u, v^{-1} are doubling), and the connection between (1.28) and (1.25) remains unclear. We suspect that it is related to the equally mysterious connection between the Muckenhoupt A_2 condition and the Helson-Szegö condition. (See [19, 18] for more information.)

Remark 1.13. Xia [53], using a combination of real and complex analytic techniques, found another sufficient condition similar to (1.25). It is not directly comparable to Theorem 1.11 but appears to include many of the same pairs f and g .

1.4.1. *A Bergman space conjecture.* In [23] Sarason also asked the analogous question for the product of Toeplitz operators $T_f T_{\bar{g}}$ on the Bergman space $L_a^2(\mathbb{D})$, the space of analytic functions on \mathbb{D} that are square integrable with respect to area measure. (The Hardy space $H^2(\partial\mathbb{D})$ is a proper subspace of the Bergman space. For more information on the Bergman spaces, see [14].) On the Bergman space, the Toeplitz operator T_h is defined exactly as on the Hardy space, but with the Riesz projection operator P replaced by P_a , the Bergman projection from $L^2(\mathbb{D})$ to $L_a^2(\mathbb{D})$.

Stroethoff and Zheng [46] showed that a necessary condition for $T_f T_{\bar{g}}$, $f, g \in L_a^2(\mathbb{D})$, to be bounded on $L_a^2(\mathbb{D})$ is

$$\sup_{z \in \mathbb{D}} B_z(|f|^2) B_z(|g|^2) < \infty,$$

where $B_z(\cdot)$ denotes the Berezin transform,

$$B_z f(\omega) = \int_{\mathbb{D}} f(z) |k_\omega(z)|^2 dA(z), \quad k_\omega(z) = \frac{1 - |\omega|^2}{(1 - \bar{\omega}z)^2}.$$

This is the natural analog of (1.25) since the Berezin transform plays a role in Bergman spaces similar to that of convolution with the Poisson kernel in Hardy spaces.

Stroethoff and Zheng further proved that the analog of (1.27),

$$\sup_{z \in \mathbb{D}} B_z(|f|^{2+\epsilon}) B_z(|g|^{2+\epsilon}) < \infty,$$

is a sufficient condition for $T_f T_{\bar{g}}$ to be bounded on $L_a^2(\mathbb{D})$.

The factorization diagram given above adapts immediately to the Bergman space case, so to prove a sufficient condition for $T_f T_{\bar{g}}$ to be bounded it suffices to prove a weighted norm inequality for the Bergman projection. This reduces to a real-variable problem since $P_a = I - TT^*$, where T is a two-dimensional Calderón-Zygmund singular integral. (See [50] and [54].) Therefore, we conjecture that the techniques used to prove Theorem 1.4, which, as we noted above, adapt to a variety of other operators, can be adapted to prove some version of Theorem 1.11 in the setting of Bergman spaces. A key tool for proving such a result—a Calderón-Zygmund decomposition adapted to the disk—has already been developed by Stroethoff and Zheng [47].

1.5. Organization. The remainder of this paper is organized as follows. In Section 2 we gather a number of preliminary results which are needed in subsequent sections. In Section 3 we prove Theorems 1.2, 1.3 and 1.7. In Section 4, we prove Theorems 1.4 and 1.8.

Throughout this paper all notation will be standard or defined as needed. All cubes are assumed to have their sides parallel to the coordinate axes. Given a cube Q and $r > 0$, rQ will denote the cube with the same center as Q and whose sides are r times as long. Given $1 < p < \infty$, $p' = p/(p - 1)$ will denote the conjugate exponent of p . C will denote a positive constant whose value may change at each appearance. By weights we will always mean non-negative, measurable functions which are positive on a set of positive measure. Given a Lebesgue measurable set E and a weight w , $|E|$ will denote the Lebesgue measure of E and $w(E) = \int_E w \, dx$.

In the theorems stated above we assumed that the weights satisfied conditions such as (1.8) with A and B being certain Young functions. Such conditions always imply that $u, v^{-1} \in L^1_{\text{loc}}(\mathbb{R}^n)$, and we will make use of this without further comment. We do not, however, assume that v is locally integrable. This is important for our results on the Sarason conjecture, since there are simple examples of outer functions g such that $|g|^{-1}$ is not in L^1 , and $|g|^{-1}$ corresponds to the weight v . In Section 2 below we will indicate how we can reduce to the special case of weights that are bounded functions.

2. SOME PRELIMINARY LEMMAS

2.1. The Calderón-Zygmund decomposition.

Definition 2.1. Given a non-negative function $f \in L^1(\mathbb{R}^n)$ (e.g., $f \in L^\infty(\mathbb{R}^n)$) and $\lambda > 0$, define the Calderón-Zygmund (CZ) cubes of f at height λ to be the maximal disjoint dyadic subcubes of the set

$$\Omega_\lambda = \{x \in \mathbb{R}^n : M_d f(x) > \lambda\}.$$

Lemma 2.2. Given $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$ non-negative, let $\{Q_j\}$ be the set of CZ cubes of f at height λ . Then for all j ,

$$(2.1) \quad \lambda < \frac{1}{|Q_j|} \int_{Q_j} f(x) \, dx \leq 2^n \lambda,$$

and for $x \in \mathbb{R}^n \setminus \Omega_\lambda$, $f(x) \leq \lambda$. Further, for all $x \in Q_j$,

$$(2.2) \quad M_d f(x) = M_d(f\chi_{Q_j})(x).$$

Inequality (2.1) follows immediately from the definition. Inequality (2.2) is an observation due to Journé [21] and follows from the maximality of Q_j : if Q is any dyadic cube containing Q_j , then

$$\frac{1}{|Q|} \int_Q f(x) \, dx \leq \lambda < M_d f(x)$$

and hence (2.2) must hold.

Lemma 2.3. *Let $f \in L^1(\mathbb{R}^n)$. Fix $a > 2^n$, and for $k \in \mathbb{Z}$, let $\{Q_j^k\}$ be the CZ cubes of f at height a^k . Then there exist sets $\{\tilde{Q}_j^k\}$, $\tilde{Q}_j^k \subset Q_j^k$, which are pairwise disjoint for all j and k , and such that $|Q_j^k| \leq \alpha |\tilde{Q}_j^k|$ with $\alpha > 1$ depending only on a and n .*

For a proof, see [35].

Lemma 2.4. *Let $f \in L^1(\mathbb{R}^n)$ and fix $\lambda > 0$. Let $\{Q_j\}$ be the set of CZ cubes of f at height $\lambda/4^n$. Then*

$$\{x \in \mathbb{R}^n : Mf(x) > \lambda\} \subset \bigcup_j 3Q_j.$$

A proof can be found in Duoandikoetxea [13, Lemma 2.12] for the centered maximal operator; the same argument works for the uncentered maximal operator.

2.2. The sharp maximal operator. Let $M^\#$ denote the sharp maximal operator of Fefferman-Stein and for $0 < q < 1$ define $M_q^\# f(x) = M^\#(|f|^q)(x)^{\frac{1}{q}}$. In the next two lemmas, T denotes a Calderón-Zygmund singular integral operator.

Lemma 2.5. *Given q , $0 < q < 1$, there exists C such that for any $f \in L_c^\infty(\mathbb{R}^n)$,*

$$M_q^\#(Tf)(x) \leq C Mf(x) \quad \text{and} \quad M_q^\#(Mf)(x) \leq C M^\# f(x).$$

The first estimate can be found in [1] and the second in [8].

Lemma 2.6. *Let $f \geq 0$ be such that its level sets $\{x : f(x) > \lambda\}$ have finite measure for all $\lambda > 0$ (e.g., $f \in L_c^\infty$). Then for all weights w ,*

$$(2.3) \quad \int_{\mathbb{R}^n} f(x) w(x) dx \leq C \int_{\mathbb{R}^n} M^\# f(x) Mw(x) dx.$$

As a consequence, for each q , $0 < q < 1$,

$$\int_{\mathbb{R}^n} Mf(x)^q w(x) dx \leq C \int_{\mathbb{R}^n} M^\# f(x)^q Mw(x) dx,$$

and

$$\int_{\mathbb{R}^n} |Tf(x)|^q w(x) dx \leq C \int_{\mathbb{R}^n} Mf(x)^q Mw(x) dx.$$

The first estimate is due to Lerner [26]. The other two follow from it combined with Lemma 2.5.

2.3. Orlicz spaces. For more information on Orlicz spaces, see Rao and Ren [41].

Lemma 2.7. *If A , B and C are Young functions such that $A^{-1}(t)B^{-1}(t) \leq C^{-1}(t)$, then for all functions f and g and any cube Q ,*

$$(2.4) \quad \|fg\|_{C,Q} \leq 2 \|f\|_{A,Q} \|g\|_{B,Q}.$$

In particular, given any Young function A ,

$$(2.5) \quad \frac{1}{|Q|} \int_Q |f(x)g(x)| dx \leq 2 \|f\|_{A,Q} \|g\|_{\bar{A},Q}.$$

Inequality (2.5) is due to Weiss [51]; (2.4) is due to O'Neil [33].

Given a Young function B , recall that we define the Orlicz maximal operator associated with B by

$$M_B f(x) = \sup_{Q \ni x} \|f\|_{B,Q}.$$

We have the following result taken from [36] that characterizes the boundedness of these maximal functions on $L^p(\mathbb{R}^n)$. This will play an important role in the proofs of our main results.

Lemma 2.8. *Given p , $1 < p < \infty$, and a Young function B , then*

$$M_B : L^p(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n) \quad \text{if and only if} \quad B \in B_p.$$

2.4. Reduction to bounded weights and bounded functions of compact support. To insure that the integrals which appear in our estimates are finite, we need to assume that the functions involved are bounded. First note that we can assume without loss of generality that both $u, v \in L^\infty(\mathbb{R}^n)$. Given Young functions A, B , assume the pair of weights (u, v) satisfies condition (1.8) with constant C_0 . For $N > 0$ set $u_N = \min\{u, N\}$, $v_N = \min\{v, N\}$. Then the pair (u_N, v_N) satisfies the same estimate with constant at most $C_0 + 1$:

$$\|u_N\|_{A,Q} \|v_N^{-1}\|_{B,Q} \leq \|u\|_{A,Q} \|v^{-1}\|_{B,Q} + \|N\|_{A,Q} \|N^{-1}\|_{B,Q} \leq C_0 + 1.$$

Therefore, we can work with the weights $u_N, v_N \in L^\infty(\mathbb{R}^n)$ and get estimates for them. We can then let $N \rightarrow \infty$ and apply the monotone convergence theorem to get the desired result for the pair of weights (u, v) .

Finally, by standard density arguments we will also be able to assume without loss of generality that $f \in L_c^\infty(\mathbb{R}^n)$.

3. PROOF OF THEOREMS 1.3 AND 1.7

The heart of this section is the proof of Theorem 1.3. Theorem 1.2 is just a special case of this result (the case $n = 1$) so there is nothing to prove. The proof of Theorem 1.7 is very similar to that of Theorem 1.3, and at the end of this section we will describe the necessary changes.

3.1. Proof of Theorem 1.3. Fix p , $1 < p < \infty$, the pair of weights (u, v) and f . As discussed above, we may assume without loss of generality that $u, v \in L^\infty(\mathbb{R}^n)$ and $f \in L_c^\infty(\mathbb{R}^n)$. Fix q , $0 < q < 1$, such that if $r = p/q$, then there exists $\epsilon > 0$ so that $p - 1 + \delta = r - 1 + \epsilon$. Then by duality,

$$\left(\int_{\mathbb{R}^n} |u(x) T f(x)|^p dx \right)^{q/p} = \sup \int_{\mathbb{R}^n} |u(x) T f(x)|^q h(x) dx,$$

where the supremum is taken over all non-negative functions $h \in L_c^\infty(\mathbb{R}^n)$ such that $\|h\|_{L^{r'}(\mathbb{R}^n)} = 1$. Fix such a function h . We will bound the integral on the righthand side with a constant independent of h .

Let $w = u^q h$; then $w \in L_c^\infty$ since $h \in L_c^\infty(\mathbb{R}^n)$ and $u \in L^\infty(\mathbb{R}^n)$. We will now form a kind of atomic decomposition of w that is due to Lerner [26] and lies at the heart of his proof of Lemma 2.6. Fix $a > 2^n$ and $m > 0$ such that $\|w\|_{L^\infty} \leq a^m$. For each $k \leq m$, let $\{Q_j^k\}$ be the Calderón-Zygmund cubes of w at height a^k (Lemma 2.2). Let $w_{Q_j^k} = |Q_j^k|^{-1} \int_{Q_j^k} w(x) dx$, and for each k define the functions

$$b_k(x) = \sum_j (w(x) - w_{Q_j^k}) \chi_{Q_j^k}(x), \quad g_k(x) = w(x) - b_k(x) = \begin{cases} w_{Q_j^k} & x \in Q_j^k \\ w(x) & x \in \mathbb{R}^n \setminus \Omega_{a^k}. \end{cases}$$

Again by Lemma 2.2, for all k we have $g_k(x) \leq 2^n a^k$ and $\|g_k\|_1 = \|w\|_1$.

Since the set Ω_{a^m} is empty, $b_m = 0$. Therefore, for every integer $l < 0$, we have the telescoping sequence

$$w(x) = \sum_{k=l}^{m-1} (b_k(x) - b_{k+1}(x)) + g_l(x).$$

By (2.1), $w_{Q_j^k} \leq 2^n a^k$. Since for each j and k ,

$$(3.1) \quad (b_k(x) - b_{k+1}(x)) \chi_{Q_j^k}(x) = (w(x) - w_{Q_j^k}) \chi_{Q_j^k}(x) - \sum_{Q_i^{k+1} \subset Q_j^k} (w(x) - w_{Q_i^{k+1}}) \chi_{Q_i^{k+1}}(x),$$

it follows immediately that for all x ,

$$(3.2) \quad |b_k(x) - b_{k+1}(x)| \leq (1+a) 2^n a^k.$$

Further, by integrating (3.1) we see that

$$(3.3) \quad \int_{Q_j^k} (b_k(x) - b_{k+1}(x)) dx = 0.$$

We can now estimate as follows: for any $l < 0$,

$$(3.4) \quad \int_{\mathbb{R}^n} |Tf(x)|^q u(x)^q h(x) dx = \int_{\mathbb{R}^n} |Tf(x)|^q w(x) dx \\ = \sum_{k=l}^{m-1} \int_{\mathbb{R}^n} |Tf(x)|^q (b_k(x) - b_{k+1}(x)) dx + \int_{\mathbb{R}^n} |Tf(x)|^q g_l(x) dx.$$

We now claim that the last term on the righthand side tends to 0 as $l \rightarrow -\infty$. This follows at once from Hölder's inequality, the fact that T is bounded on $L^2(\mathbb{R}^n)$, and that f and w are bounded functions with compact support:

$$0 \leq \int_{\mathbb{R}^n} |Tf(x)|^q g_l(x) dx \leq \left(\int_{\mathbb{R}^n} |Tf(x)|^2 dx \right)^{q/2} \left(\int_{\mathbb{R}^n} g_l(x)^{(2/q)'} dx \right)^{1/(2/q)'} \\ \leq C \|f\|_2^{q/2} (2^n a^l)^{(q/2)} \|g_l\|_1^{1/(2/q)'} = C \|f\|_2^{q/2} (2^n a^l)^{(q/2)} \|w\|_1^{1/(2/q)'}. \quad \square$$

As $l \rightarrow -\infty$ the last term tends to zero. Therefore, taking the limit in (3.4) we get

$$(3.5) \quad \int_{\mathbb{R}^n} |Tf(x)|^q u(x)^q h(x) dx = \sum_{k=-\infty}^{m-1} \int_{\mathbb{R}^n} |Tf(x)|^q (b_k(x) - b_{k+1}(x)) dx.$$

We estimate the righthand side of (3.5) as follows. For each j, k , let c_j^k be a constant whose value will be specified below. Since $q < 1$, $||a|^q - |b|^q| \leq |a - b|^q$. Therefore, by (3.3) and (3.2),

$$\begin{aligned} & \sum_{k=-\infty}^{m-1} \int_{\mathbb{R}^n} |Tf(x)|^q (b_k(x) - b_{k+1}(x)) dx = \sum_{k,j} \int_{Q_j^k} |Tf(x)|^q (b_k(x) - b_{k+1}(x)) dx \\ & = \sum_{k,j} \int_{Q_j^k} (|Tf(x)|^q - |c_j^k|^q) (b_k(x) - b_{k+1}(x)) dx \\ & \leq C \sum_{k,j} (1+a) 2^n a^k \int_{Q_j^k} ||Tf(x)|^q - |c_j^k|^q| dx \leq C \sum_{k,j} w_{Q_j^k} \int_{Q_j^k} |Tf(x) - c_j^k|^q dx \\ & \leq C \sum_{k,j} w_{Q_j^k} \int_{Q_j^k} |T(f\chi_{2Q_j^k})(x)|^q dx + C \sum_{k,j} w_{Q_j^k} \int_{Q_j^k} |T(f\chi_{\mathbb{R}^n \setminus 2Q_j^k})(x) - c_j^k|^q dx \\ & = C(I_1 + I_2). \end{aligned}$$

We consider each term separately. To estimate I_1 we use Kolmogorov's inequality (since $q < 1$) and Lemmas 2.3 and 2.7:

$$\begin{aligned} I_1 & \leq C \sum_{k,j} \frac{1}{|2Q_j^k|} \int_{2Q_j^k} w(x) dx \left(\frac{1}{|2Q_j^k|} \int_{2Q_j^k} |f(x)| dx \right)^q |Q_j^k| \\ & = C \sum_{k,j} \frac{1}{|2Q_j^k|} \int_{2Q_j^k} u(x)^q h(x) dx \left(\frac{1}{|2Q_j^k|} \int_{2Q_j^k} v(x) |f(x)| v(x)^{-1} dx \right)^q |Q_j^k| \\ & \leq C \sum_{k,j} \|u^q\|_{C,2Q_j^k} \|h\|_{\bar{C},2Q_j^k} \|v f\|_{\bar{B},2Q_j^k}^q \|v^{-1}\|_{B,2Q_j^k}^q |\tilde{Q}_j^k|, \end{aligned}$$

where $C(t) = t^r \log(e+t)^{r-1+\epsilon}$. Let $C_q(t) = C(t^q) = t^p \log(e+t^q)^{r-1+\epsilon} \approx A(t)$. Therefore, by a change of variables in the definition of the Orlicz norm,

$$\|u^q\|_{C,2Q_j^k} = \|u\|_{C_q,2Q_j^k}^q \approx \|u\|_{A,2Q_j^k}^q.$$

Hence, (1.16), the fact that the sets \tilde{Q}_j^k are disjoint, and Hölder's inequality yield

$$\begin{aligned} I_1 & \leq C \sum_{k,j} \|h\|_{\bar{C},2Q_j^k} \|v f\|_{\bar{B},2Q_j^k}^q |\tilde{Q}_j^k| \\ & \leq C \sum_{k,j} \int_{\tilde{Q}_j^k} M_{\bar{C}} h(x) M_{\bar{B}}(v f)(x)^q dx \\ & \leq C \left(\int_{\mathbb{R}^n} M_{\bar{C}} h(x)^{r'} dx \right)^{1/r'} \left(\int_{\mathbb{R}^n} M_{\bar{B}}(v f)(x)^p dx \right)^{q/p} \end{aligned}$$

$$\begin{aligned}
&\leq C \left(\int_{\mathbb{R}^n} h(x)^{r'} dx \right)^{1/r'} \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p} \\
&= C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p}.
\end{aligned}$$

The last inequality holds since $\bar{C} \in B_{r'}$ and $\bar{B} \in B_p$, and so by Lemma 2.8 $M_{\bar{C}}$ is bounded on $L^{r'}$ and $M_{\bar{B}}$ is bounded on L^p . This completes the estimate of I_1 .

To estimate I_2 we choose the value of the constant c_j^k to be

$$c_j^k = \frac{1}{|Q_j^k|} \int_{Q_j^k} T(f \chi_{\mathbb{R}^n \setminus 2Q_j^k})(y) dy.$$

Let $C(t)$ be as in the estimate of I_1 . Then, by a standard estimate for Calderón-Zygmund singular integrals (see [13, 18]), since $q < 1$ and by Lemmas 2.3 and 2.7, we obtain

$$\begin{aligned}
I_2 &\leq C \sum_{k,j} \frac{1}{|Q_j^k|} \int_{Q_j^k} u(x)^q h(x) dx \left(\sum_{i=1}^{\infty} 2^{-i} \frac{1}{|2^i Q_j^k|} \int_{2^i Q_j^k} |f(x)| dx \right)^q |Q_j^k| \\
&\leq C \sum_{k,j} \frac{1}{|Q_j^k|} \int_{Q_j^k} u(x)^q h(x) dx \sum_{i=1}^{\infty} 2^{-iq} \left(\frac{1}{|2^i Q_j^k|} \int_{2^i Q_j^k} v(x) f(x) v(x)^{-1} dx \right)^q |Q_j^k| \\
&\leq C \sum_{k,j} \|u^q\|_{C, Q_j^k} \|h\|_{\bar{C}, Q_j^k} |\tilde{Q}_j^k| \sum_{i=1}^{\infty} 2^{-iq} \|v f\|_{\bar{B}, 2^i Q_j^k}^q \|v^{-1}\|_{B, 2^i Q_j^k}^q \\
&\leq C \sum_{k,j} \|u\|_{A, Q_j^k}^q \|h\|_{\bar{C}, Q_j^k} |\tilde{Q}_j^k| \sum_{i=1}^{\infty} 2^{-iq} \|v f\|_{\bar{B}, 2^i Q_j^k}^q \|v^{-1}\|_{B, 2^i Q_j^k}^q.
\end{aligned}$$

For $0 < \beta < 1$, $A(\beta t) \leq \beta^p A(t)$, so by the definition of the Luxemburg norm, we have that $\|u\|_{A, Q_j^k} \leq C 2^{in/p} \|u\|_{A, 2^i Q_j^k}$. Thus, by (1.16) and since $p > n$ it follows

$$\begin{aligned}
I_2 &\leq C \sum_{k,j} \|h\|_{\bar{C}, Q_j^k} |\tilde{Q}_j^k| \sum_{i=1}^{\infty} 2^{-iq} 2^{inq/p} \|u\|_{A, 2^i Q_j^k}^q \|v f\|_{\bar{B}, 2^i Q_j^k}^q \|v^{-1}\|_{B, 2^i Q_j^k}^q \\
&\leq C \sum_{k,j} \|h\|_{\bar{C}, Q_j^k} |\tilde{Q}_j^k| \inf_{x \in Q_j^k} M_{\bar{B}}(v f)(x)^q \\
&\leq C \sum_{k,j} \int_{\tilde{Q}_j^k} M_{\bar{C}} h(x) M_{\bar{B}}(v f)(x)^q dx.
\end{aligned}$$

We can now argue as we did above for I_1 to obtain the desired estimate for I_2 .

3.2. Proof of Theorem 1.7. The proof is almost identical to the one just given and we only indicate the minor changes. We proceed in the same manner with S_d in place of T . We observe that S_d is bounded on $L^2(\mathbb{R}^n)$ and so it suffices to get the appropriate estimates for I_1 and I_2 , where now in I_1 we write $f \chi_{Q_j^k}$ in place of $f \chi_{2Q_j^k}$ and in I_2 we put $f \chi_{\mathbb{R}^n \setminus Q_j^k}$ in place of $f \chi_{\mathbb{R}^n \setminus 2Q_j^k}$. The estimate for I_1 adapts

immediately to the dyadic square function since S_d is of weak-type $(1, 1)$ and thus satisfies Kolmogorov's inequality.

Since the dyadic square function is more localized than a singular integral, the estimate for I_2 is much easier. Fix a cube Q_j^k and set

$$c_j^k = \left(\sum_{\substack{Q \in \Delta \\ Q \supseteq Q_j^k}} |(f\chi_{\mathbb{R}^n \setminus Q_j^k})_Q - (f\chi_{\mathbb{R}^n \setminus Q_j^k})_{\widehat{Q}}|^2 \right)^{1/2}.$$

Then for any $x \in Q_j^k$ we have that $S_d(f\chi_{\mathbb{R}^n \setminus Q_j^k})(x) \equiv c_j^k$; thus $I_2 = 0$ and we are done.

4. PROOF OF THEOREM 1.4

At the heart of the proof of Theorem 1.4 is the following lemma, whose proof we defer for the moment.

Lemma 4.1. *Given p and (u, v) as in the hypotheses of Theorem 1.4, there exists q , $0 < q < 1$, such that for all $f, h \in L_c^\infty(\mathbb{R}^n)$,*

$$(4.1) \quad \int_{\mathbb{R}^n} Mf(x)^q M(u^q h)(x) dx \leq C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p} \left(\int_{\mathbb{R}^n} |h(x)|^{(p/q)'} dx \right)^{1/(p/q)'}$$

Proof of Theorem 1.4. Fix q as in Lemma 4.1 and let $r = \frac{p}{q} > 1$. Then by duality,

$$\left(\int_{\mathbb{R}^n} |u(x) Tf(x)|^p dx \right)^{q/p} = \sup \int_{\mathbb{R}^n} |Tf(x)|^q u(x)^q h(x) dx,$$

where the supremum is taken over all non-negative functions $h \in L_c^\infty(\mathbb{R}^n)$ such that $\|h\|_{L^{r'}(\mathbb{R}^n)} = 1$. Fix such a function h . Then by Lemmas 2.6 and 4.1,

$$\begin{aligned} \int_{\mathbb{R}^n} |Tf(x)|^q u(x)^p h(x) dx &\leq C \int_{\mathbb{R}^n} Mf(x)^q M(u^q h)(x) dx \\ &\leq C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p} \left(\int_{\mathbb{R}^n} |h(x)|^{r'} dx \right)^{1/r'} = C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p}. \end{aligned}$$

This completes the proof of Theorem 1.4. \square

4.1. Proof of Lemma 4.1. Fix f ; by a standard argument we may assume without loss of generality that $f \geq 0$. Further, as we noted above, we may assume without loss of generality that $f \in L_c^\infty$ and $u, v \in L^\infty$. Fix q , $0 < q < 1$, sufficiently close to 1 that there exists $\epsilon > 0$ such that $2p - 1 + \delta = 2(p/q) - 1 + \epsilon$. Let $r = p/q$, $w = u^q h$ and $a = 4^n > 2^n$. For each j, k let

$$\Omega_{kj} = \{a^{k-j-1} < Mw(x) \leq a^{k-j+1}\} \cap \{a^j < Mf(x)^q \leq a^{j+1}\};$$

then

$$\int_{\mathbb{R}^n} Mf(x)^q Mw(x) dx = \sum_{k,j} \int_{\{a^k < (Mf)^q Mw \leq a^{k+1}\} \cap \{a^j < (Mf)^q \leq a^{j+1}\}} Mf(x)^q Mw(x) dx$$

$$\leq \sum_{k,j} \int_{\Omega_{kj}} Mf(x)^q Mw(x) dx.$$

For each integer l, m , let $\{R_l^r\}_r$ be the CZ cubes of w at height a^l , and let $\{S_m^s\}_s$ be the CZ cubes of f at height $a^{m/q}$. Then by Lemma 2.4, for each pair (k, j) ,

$$\{x : Mw(x) > a^{k-j-1}\} \subset \bigcup_r 3R_{k-j-2}^r, \quad \{x : Mf(x)^q > a^j\} \subset \bigcup_s 3S_{j-1}^s.$$

If $x \in \Omega_{kj}$, there exists at least one pair (r, s) such that $x \in 3R_{k-j-2}^r \cap 3S_{j-1}^s$. Let $E_{kj}^{rs} = \{x \in \Omega_{kj} : x \in 3R_{k-j-2}^r \cap 3S_{j-1}^s\}$. If the set E_{kj}^{rs} is not empty, then $3R_{k-j-2}^r \cap 3S_{j-1}^s \neq \emptyset$. Therefore, depending on their relative sizes, we either have $3R_{k-j-2}^r \subset 9S_{j-1}^s$, or $3S_{j-1}^s \subset 9R_{k-j-2}^r$. If the first inclusion holds we say that $(k, j, r, s) \in \Gamma_1$; if the second holds we say that $(k, j, r, s) \in \Gamma_2$. Hence,

$$\begin{aligned} \int_{\mathbb{R}^n} Mf(x)^q Mw(x) dx &\leq \sum_{k,j} \sum_{r,s} \int_{E_{kj}^{rs}} Mf(x)^q Mw(x) dx \leq \sum_{k,j} \sum_{r,s} a^{k-j+1} a^{j+1} |E_{kj}^{rs}| \\ &\leq \sum_{(k,j,r,s) \in \Gamma_1} a^{k-j+1} a^{j+1} |E_{kj}^{rs}| + \sum_{(k,j,r,s) \in \Gamma_2} a^{k-j+1} a^{j+1} |E_{kj}^{rs}| = I_1 + II_2. \end{aligned}$$

To complete the proof we will estimate each term separately. We consider first I_1 . Since $E_{kj}^{rs} \subset 3R_{k-j-2}^r$, by Lemma 2.3, $|E_{kj}^{rs}| \leq 3^n |R_{k-j-2}^r| \leq C |\tilde{R}_{k-j-2}^r|$. On the other hand $3R_{k-j-2}^r \subset 9S_{j-1}^s$. Thus by Lemma 2.2,

$$\begin{aligned} I_1 &\leq a^5 \sum_{(k,j,r,s) \in \Gamma_1} \left(\frac{1}{|R_{k-j-2}^r|} \int_{R_{k-j-2}^r} w(x) dx \right) \left(\frac{1}{|S_{j-1}^s|} \int_{S_{j-1}^s} f(x) dx \right)^q |E_{kj}^{rs}| \\ &\leq C \sum_{j,s} \left(\sum_{k,r} \frac{1}{|R_{k-j-2}^r|} \int_{R_{k-j-2}^r} w(x) dx \cdot |\tilde{R}_{k-j-2}^r| \right) \left(\frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} f(x) dx \right)^q \\ &\leq C \sum_{j,s} \left(\sum_{k,r} \int_{\tilde{R}_{k-j-2}^r} M(w\chi_{9S_{j-1}^s})(x) dx \right) \left(\frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} f(x) dx \right)^q. \end{aligned}$$

Since the sets \tilde{R}_{k-j-2}^r are disjoint and contained in $9S_{j-1}^s$, we can apply Yano's theorem (see Zygmund [56]) to get

$$\begin{aligned} I_1 &\leq C \sum_{j,s} \frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} M(w\chi_{9S_{j-1}^s})(x) dx \left(\frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} f(x) dx \right)^q |\tilde{S}_{j-1}^s| \\ &\leq C \sum_{j,s} \|w\|_{\Phi, 9S_{j-1}^s} \left(\frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} f(x) dx \right)^q |\tilde{S}_{j-1}^s|, \end{aligned}$$

where $\Phi(t) = t \log(e+t)$ and the constant depends only on n and not on the cube S_{j-1}^s . Recall that $2p-1+\delta = 2r-1+\epsilon$; hence, if we define $D(t) = t^r \log(e+t)^{2r-1+\epsilon}$,

then $D(t^q) \approx A(t)$. Now define $\tilde{D}(t) = t^{r'} \log(e+t)^{-1-\epsilon(r'-1)} \in B_{r'}$. Then we have that

$$\Phi^{-1}(t) \approx \frac{t}{\log(e+t)} = \frac{t^{\frac{1}{r}}}{\log(e+t)^{\frac{2r-1+\epsilon}{r}}} \times t^{\frac{1}{r'}} \log(e+t)^{-1+\frac{2r-1+\epsilon}{r}} \approx D^{-1}(t) \cdot \tilde{D}^{-1}(t).$$

Therefore, recalling that $w = u^q h$, we can apply Lemma 2.7 and (1.18) to get

$$\begin{aligned} I_1 &\leq C \sum_{j,s} \|u^q\|_{D,9S_{j-1}^s} \|h\|_{\tilde{D},9S_{j-1}^s} \|v f\|_{\tilde{B},9S_{j-1}^s}^q \|v^{-1}\|_{B,9S_{j-1}^s}^q |\tilde{S}_{j-1}^s| \\ &\leq C \sum_{j,s} \|u\|_{A,9S_{j-1}^s}^q \|h\|_{\tilde{D},9S_{j-1}^s} \|v f\|_{\tilde{B},9S_{j-1}^s}^q \|v^{-1}\|_{B,9S_{j-1}^s}^q |\tilde{S}_{j-1}^s| \\ &\leq C \sum_{j,s} \int_{\tilde{S}_{j-1}^s} M_{\tilde{D}} h(x) M_{\tilde{B}}(v f)(x)^q dx \\ &\leq C \int_{\mathbb{R}^n} M_{\tilde{D}} h(x) M_{\tilde{B}}(v f)(x)^q dx \\ &\leq C \left(\int_{\mathbb{R}^n} M_{\tilde{D}} h(x)^{r'} dx \right)^{1/r'} \left(\int_{\mathbb{R}^n} M_{\tilde{B}}(v f)(x)^p dx \right)^{q/p} \\ &\leq C \left(\int_{\mathbb{R}^n} |h(x)|^{r'} dx \right)^{1/r'} \left(\int_{\mathbb{R}^n} (v(x)f(x))^p dx \right)^{q/p}, \end{aligned}$$

where the third inequality holds because the sets \tilde{S}_{j-1}^s are disjoint, and the last inequality holds since by Lemma 2.8, $\tilde{B} \in B_p$ so $M_{\tilde{B}}$ is bounded on L^p , and, as we noted above, $\tilde{D} \in B_{r'}$, so $M_{\tilde{D}}$ is bounded in $L^{r'}$. Thus we get the desired bound for I_1 .

We will now estimate I_2 . The ideas are the same, except that at the key step we will use Kolmogorov's inequality instead of Yano's theorem. Since $E_{kj}^{rs} \subset 3S_{j-1}^s$, by Lemma 2.3, $|E_{kj}^{rs}| \leq C |\tilde{S}_{j-1}^s|$. Further, $M_d f(x)^q > a^{j-1}$ on S_{j-1}^s . Thus

$$\begin{aligned} I_2 &\leq a^3 \sum_{(k,j,r,s) \in \Gamma_2} a^{j+1} |E_{kj}^{rs}| \left(\frac{1}{|R_{k-j-2}^r|} \int_{R_{k-j-2}^r} w(x) dx \right) \\ &\leq C \sum_{(k,j,r,s) \in \Gamma_2} a^{j+1} |\tilde{S}_{j-1}^s| \left(\frac{1}{|9R_{k-j-2}^r|} \int_{9R_{k-j-2}^r} w(x) dx \right) \\ &\leq C \sum_{(k,j,r,s) \in \Gamma_2} \int_{\tilde{S}_{j-1}^s} M_d f(x)^q dx \left(\frac{1}{|9R_{k-j-2}^r|} \int_{9R_{k-j-2}^r} w(x) dx \right) \\ &= C \sum_{l,r} \left(\sum_{\substack{(k,j,r,s) \in \Gamma_2 \\ k-j-2=l}} \int_{\tilde{S}_{j-1}^s} M_d f(x)^q dx \right) \frac{1}{|9R_l^r|} \int_{9R_l^r} w(x) dx. \end{aligned}$$

For fixed l and r , $\tilde{S}_{j-1}^s \subset 9R_l^r$. Thus, by Lemma 2.2, for all $x \in \tilde{S}_{j-1}^s$,

$$(4.2) \quad M_d f(x) = M_d(f \chi_{S_{j-1}^s})(x) \leq M_d(f \chi_{9R_l^r})(x).$$

Since $t \leq \Phi(t)$, $\|w\|_{L^1, 9R_l^r} \leq \|w\|_{\Phi, 9R_l^r}$ (see [41]). Therefore, using this, (4.2), Lemma 2.3, and Kolmogorov's inequality, we get that

$$\begin{aligned} I_2 &\leq C \sum_{l,r} \left(\sum_{\substack{(k,j,r,s) \in \Gamma_2 \\ k-j-2=l}} \int_{\tilde{S}_{j-1}^s} M_d(f \chi_{9R_l^r})(x)^q dx \right) \|w\|_{\Phi, 9R_l^r} \\ &\leq C \sum_{l,r} \frac{1}{|9R_l^r|} \int_{9R_l^r} M_d(f \chi_{9R_l^r})(x)^q dx \cdot \|w\|_{\Phi, 9R_l^r} \cdot |\tilde{R}_l^r| \\ &\leq C \sum_{l,r} \left(\frac{1}{|9R_l^r|} \int_{9R_l^r} f(x) dx \right)^q \|w\|_{\Phi, 9R_l^r} \cdot |\tilde{R}_l^r|. \end{aligned}$$

We can now argue exactly as we did in the estimate of I_1 to get the desired bound for I_2 . This completes the proof.

Remark 4.2. The term I_2 is less singular than the term I_1 : if we did not replace $\|w\|_{L^1, 9R_l^r}$ by $\|w\|_{\Phi, 9R_l^r}$, then a slight modification of our argument would show that we get the desired bound for I_2 assuming only the weaker condition (1.16).

5. PROOF OF THEOREM 1.8

The proof of Theorem 1.8 is identical in basic idea and organization to the proof of Theorem 1.4, differing only in details. Therefore, rather than give the complete argument, we will outline the changes necessary in the proof of Theorem 1.4.

The key changes are in the statement and proof of Lemma 4.1. The new lemma is the following.

Lemma 5.1. *Given p and (u, v) as in the hypotheses of Theorem 1.8, there exists q , $0 < q < 1$, such that for all $f, h \in L_c^\infty(\mathbb{R}^n)$,*

$$\int_{\mathbb{R}^n} M^2 f(x)^q M^2(u^q h)(x) dx \leq C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p} \left(\int_{\mathbb{R}^n} |h(x)|^{(p/q)'} dx \right)^{1/(p/q)'}$$

Given this inequality, the proof of Theorem 1.8 begins with the same duality argument as the proof of Theorem 1.4. But, instead of Lemma 2.5 we use the following pointwise estimate from [37]: given $0 < q < \epsilon < 1$,

$$(5.1) \quad M_q^\#([b, T]f)(x) \leq C M_\epsilon(Tf)(x) + C M^2 f(x).$$

Thus by Lemma 2.6, we have for every weight w that

$$\begin{aligned} \int_{\mathbb{R}^n} |[b, T]f(x)|^q w(x) dx &\leq C \int_{\mathbb{R}^n} M_q^\#([b, T]f)(x)^q M w(x) dx \\ &\leq C \int_{\mathbb{R}^n} M_\epsilon(Tf)^q M w(x) dx + C \int_{\mathbb{R}^n} (M^2 f)^q M w(x) dx. \end{aligned}$$

The second integral in the last term is dominant. To see this we use the fact that $q/\epsilon < 1$, and Lemmas 2.6 and 2.5 to get

$$\int_{\mathbb{R}^n} M_\epsilon(Tf)(x)^q M w(x) dx = \int_{\mathbb{R}^n} M(|Tf|^\epsilon)(x)^{\frac{q}{\epsilon}} M w(x) dx$$

$$\begin{aligned}
&\leq C \int_{\mathbb{R}^n} M^\#(|Tf|^\epsilon)(x)^\frac{q}{\epsilon} M^2w(x) dx = C \int_{\mathbb{R}^n} M_\epsilon^\#(Tf)(x)^q M^2w(x) dx \\
&\leq C \int_{\mathbb{R}^n} Mf(x)^q M^2w(x) dx \leq C \int_{\mathbb{R}^n} M^2f(x)^q M^2w(x) dx.
\end{aligned}$$

Hence, we have shown that

$$(5.2) \quad \int_{\mathbb{R}^n} M_\epsilon(Tf)(x)^q Mw(x) dx \leq C \int_{\mathbb{R}^n} M^2f(x)^q M^2w(x) dx.$$

Fix q as in Lemma 5.1 and let $r = \frac{p}{q} > 1$. Then by duality,

$$\left(\int_{\mathbb{R}^n} |u(x) [b, T]f(x)|^p dx \right)^{q/p} = \sup \int_{\mathbb{R}^n} |[b, T]f(x)|^q u(x)^q h(x) dx,$$

where the supremum is taken over all non-negative functions $h \in L_c^\infty(\mathbb{R}^n)$ such that $\|h\|_{L^{r'}(\mathbb{R}^n)} = 1$. Fix such a function h . By (5.2) and Lemma 5.1 it follows that

$$\begin{aligned}
&\int_{\mathbb{R}^n} |[b, T]f(x)|^q u(x)^q h(x) dx \leq C \int_{\mathbb{R}^n} M^2f(x)^q M^2(u^qh)(x) dx \\
&\leq C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{\frac{q}{p}} \left(\int_{\mathbb{R}^n} |h(x)|^{r'} dx \right)^{\frac{1}{r'}} = C \left(\int_{\mathbb{R}^n} |v(x) f(x)|^p dx \right)^{q/p}.
\end{aligned}$$

This completes the proof of Theorem 1.8.

5.1. Proof of Lemma 5.1. Define, as before, $\Phi(t) = t \log(e + t)$. It is well known (see, for instance, [45, 12]) that $M^2f \approx M_\Phi f$, so in the desired estimate we can replace M^2 by the maximal function M_Φ . We follow the same steps as in the proof of Lemma 4.1, but we replace the Calderón-Zygmund decomposition by a more general decomposition based on the Orlicz maximal operator M_Φ . Its essential properties are exactly the same and are captured in the following definition and lemmas.

Definition 5.2. *Given a Young function Φ , a non-negative function $f \in L^1(\mathbb{R}^n)$ (e.g., $f \in L_c^\infty(\mathbb{R}^n)$), and $\lambda > 0$, we define the CZ cubes of f at height λ with respect to Φ to be the maximal disjoint dyadic subcubes of the set $\{x \in \mathbb{R}^n : M_{d,\Phi}f(x) > \lambda\}$.*

Lemma 5.3. *Given $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$ non-negative, let $\{Q_j\}$ be the set of CZ cubes of f with respect to Φ . Then for all j , we have $\lambda < \|f\|_{\Phi, Q_j} \leq 2^n \lambda$. Further, $M_{d,\Phi}f(x) = M_{d,\Phi}(f\chi_{Q_j})(x)$ for all $x \in Q_j$.*

Lemma 5.4. *Let $f \in L^1(\mathbb{R}^n)$. Fix $a > 2^n$, and for $k \in \mathbb{Z}$, let $\{Q_j^k\}$ be the CZ cubes of f at height a^k with respect to Φ . Then there exist sets $\{\tilde{Q}_j^k\}, \tilde{Q}_j^k \subset Q_j^k$, which are pairwise disjoint for all j and k , and such that there exists $\alpha > 1$ depending only on a and n such that $|Q_j^k| \leq \alpha |\tilde{Q}_j^k|$.*

Lemma 5.5. *Let $f \in L^1(\mathbb{R}^n)$ and fix $\lambda > 0$. Let $\{Q_j\}$ be the set of CZ cubes of f at height $\lambda/4^n$ with respect to Φ . Then $\{x \in \mathbb{R}^n : M_\Phi f(x) > \lambda\} \subset \cup_j 3Q_j$.*

The proof of each of these lemmas is given in [9] except for the identity in Lemma 5.3, whose proof is identical to the proof of (2.2).

We now obtain the desired estimate with M_Φ in place of M^2 . Proceed as in Lemma 4.1 but f and $w = u^q h$ are decomposed with respect to M_Φ (in place of M). We estimate I_1 and I_2 using the previous lemmas and repeating the computations in Lemma 4.1. We get that

$$\begin{aligned} I_1 &\leq C \sum_{j,s} \frac{1}{|9S_{j-1}^s|} \int_{9S_{j-1}^s} M_\Phi(w\chi_{9S_{j-1}^s})(x) dx \cdot \|f\|_{\Phi,9S_{j-1}^s}^q |\tilde{S}_{j-1}^s|, \\ I_2 &\leq C \sum_{l,r} \frac{1}{|9R_l^r|} \int_{9R_l^r} M_{d,\Phi}(f\chi_{9R_l^r})(x)^q dx \cdot \|w\|_{\Phi,9R_l^r} |\tilde{R}_l^r|. \end{aligned}$$

For $k \geq 0$ define $\Phi_k(t) = t \log(e+t)^k$; then $\Phi = \Phi_1$. We have the following auxiliary result: the first inequality generalizes Yano's theorem and is well known (see for instance [12]), and the proof of the second is given below.

Lemma 5.6. *Let $k \geq 0$ and $0 < q < 1$. Then there exists a constant C such that for any cube Q*

$$\frac{1}{|Q|} \int_Q M_{\Phi_k}(g\chi_Q)(x) dx \leq C \|g\|_{\Phi_{k+1},Q}, \quad \frac{1}{|Q|} \int_Q M_{\Phi_k}(g\chi_Q)(x)^q dx \leq C \|g\|_{\Phi_k,Q}^q.$$

If we apply Lemma 5.6 to the estimates for I_1 and I_2 we get

$$\begin{aligned} I_1 &\leq C \sum_{j,s} \|w\|_{\Phi_2,9S_{j-1}^s} \|f\|_{\Phi,9S_{j-1}^s}^q |\tilde{S}_{j-1}^s| \\ I_2 &\leq C \sum_{l,r} \|f\|_{\Phi,9R_l^r}^q \|w\|_{\Phi,9R_l^r} |\tilde{R}_l^r| \leq C \sum_{l,r} \|f\|_{\Phi,9R_l^r}^q \|w\|_{\Phi_2,9R_l^r} |\tilde{R}_l^r|. \end{aligned}$$

Hence, both these estimates can be handled in the same way. Let

$$\begin{aligned} D(t) &= t^{p'} \log(e+t)^{2p'-1+\delta}, & \tilde{D}(t) &= t^p \log(e+t)^{-1-\delta(p-1)} \in B_p, \\ E(t) &= t^r \log(e+t)^{3r-1+\epsilon}, & \tilde{E}(t) &= t^{r'} \log(e+t)^{-1-\epsilon(r'-1)} \in B_{r'}. \end{aligned}$$

Then $\Phi^{-1}(t) \approx D^{-1}(t) \cdot \tilde{D}^{-1}(t)$, $\Phi_2^{-1}(t) \approx E^{-1}(t) \cdot \tilde{E}^{-1}(t)$, $D(t) = B(t)$ and $E(t^q) \approx A(t)$. Therefore, by Lemma 2.7 we have for every cube Q that

$$\|f\|_{\Phi,Q} \leq C \|f v\|_{\tilde{D},Q} \|v^{-1}\|_{B,Q}, \quad \|w\|_{\Phi_2,Q} \leq C \|u^q\|_{E,Q} \|h\|_{\tilde{E},Q} \leq C \|u\|_{A,Q}^q \|h\|_{\tilde{E},Q}.$$

Substitute these values into the above estimates for I_1, I_2 ; since $\tilde{D} \in B_p, \tilde{E} \in B_{r'}$, the proof can now be completed exactly as in the proof of Lemma 4.1.

Remark 5.7. As we noted in Remark 1.9, the proof of Theorem 1.8 can be adapted to treat the higher order commutators T_b^k , $k \geq 2$. The ideas are essentially the same. Beginning with the duality argument and applying the analog of (5.1) for higher order commutators (also found in [37]), it is not difficult to see that the proof reduces to obtaining a version of Lemma 5.1 with M^{k+1} in place of M^2 . As $M^{k+1} \approx M_{\Phi_k}$, the decompositions of f and w are made with respect to this Orlicz maximal function. If we let $D(t) = t^{p'} \log(e+t)^{(k+1)p'-1+\delta}$, $E(t) = t^r \log(e+t)^{(k+2)r-1+\epsilon}$ (\tilde{D}, \tilde{E} remain the same), then by means of Lemma 5.6 we get that the bumps for u and v are, respectively, $A(t) = t^p \log(e+t)^{(k+2)p-1+\delta}$ and $B(t) = t^{p'} \log(e+t)^{(k+1)p'-1+\delta}$.

Proof of Lemma 5.6. We only need to prove show the second inequality. By homogeneity it suffices to assume that $\|g\|_{\Phi_k, Q} = 1$. By the properties of Orlicz norms (see [41]), this implies that

$$(5.3) \quad \frac{1}{|Q|} \int_Q \Phi_k(|g(x)|) dx \leq 1.$$

The maximal operator M_{Φ_k} satisfies the modular inequality

$$(5.4) \quad |\{x \in \mathbb{R}^n : M_{\Phi_k} h(x) > \lambda\}| \leq C \int_{\mathbb{R}^n} \Phi_k(|h(x)|/\lambda) dx.$$

The proof is standard; see, for instance, [36]. Finally, we note that Φ_k is submultiplicative: $\Phi_k(st) \leq \Phi_k(s)\Phi_k(t)$. Therefore, since $0 < q < 1$, if we write the L^q -norm in terms of the level sets, then by (5.3) and (5.4) we have that

$$\begin{aligned} \frac{1}{|Q|} \int_Q M_{\Phi_k}(g \chi_Q)^q dx &\leq \int_0^1 q \lambda^q \frac{d\lambda}{\lambda} + C \int_1^\infty \lambda^q \frac{1}{|Q|} \int_{\mathbb{R}^n} \Phi_k(|g(x)\chi_Q(x)|/\lambda) dx \frac{d\lambda}{\lambda} \\ &\leq C + C \int_1^\infty \lambda^q \Phi_k(1/\lambda) \frac{d\lambda}{\lambda} \frac{1}{|Q|} \int_Q \Phi_k(|g(x)|) dx \leq C. \end{aligned}$$

□

REFERENCES

- [1] J. Álvarez and C. Pérez, *Estimates with A_∞ weights for various singular integral operators*, Boll. Un. Mat. Ital. A (7) 8 (1994), 123–133.
- [2] S. Y. A. Chang, J. M. Wilson, and T. H. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helvetici 60 (1985), 217–286.
- [3] R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. 51 (1974), 241–250.
- [4] M. Cotlar and C. Sadosky, *On the Helson-Szegö theorem and a related class of modified Toeplitz kernels*, Harmonic Analysis in Euclidean Spaces, G. Weiss and S. Wainger eds., vol. 1, 383–407, Proc. Symp. Pure Math. 35, AMS, Providence, 1979.
- [5] M. Cotlar and C. Sadosky, *On some L^p versions of the Helson-Szegö theorem*, Conference on Harmonic Analysis in Honor of Antoni Zygmund, W. Beckner *et. al.* eds., vol. 1, 306–317, Wadsworth International Mathematics Series, Wadsworth, Belmont, 1983.
- [6] D. Cruz-Uribe, SFO, *The product of unbounded Toeplitz operators*, Integral Equations Operator Theory 20 (1994), 231–237.
- [7] D. Cruz-Uribe, SFO, and A. Fiorenza, *The A_∞ property for Young functions and weighted norm inequalities*, Houston J. Math., 28 (2002), 169–182.
- [8] D. Cruz-Uribe, SFO, J.M. Martell, and C. Pérez, *Extrapolation from A_∞ weights and applications*, J. Funct. Anal. 213 (2004), 412–439.
- [9] D. Cruz-Uribe, SFO, and C. Pérez, *Sharp two-weight, weak-type norm inequalities for singular integral operators*, Math. Res. Let. 6 (1999), 417–428.
- [10] D. Cruz-Uribe, SFO, and C. Pérez, *Two weight extrapolation via the maximal operator*, J. Funct. Anal. 174 (2000), 1–17.
- [11] D. Cruz-Uribe, SFO, and C. Pérez, *On the two-weight problem for singular integral operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5) Vol. I (2002), 821–849.
- [12] G. Curbera, J. García-Cuerva, J.M. Martell, and C. Pérez, *Extrapolation with Weights, Rearrangement Invariant Function Spaces, Modular inequalities and applications to Singular Integrals*, to appear in Adv. Math.

- [13] J. Duoandikoetxea, *Fourier Analysis*, Grad. Studies Math. 29, Amer. Math. Soc., Providence, 2000.
- [14] P. Duren and A. Schuster, *Bergman Spaces*, Mathematical surveys and monographs 100) Amer. Math. Soc., Providence, 2004.
- [15] E. M. Dynkin and B. P. Osilenker, *Weighted estimates for singular integrals and their applications*, J. Soviet Math. 30 (1985), 2094–2154; translated from Itogi Nauki i Tekhniki, Akad. Nauk SSSR, Moscow, 1983.
- [16] C. Fefferman, *The uncertainty principle*, Bull. Amer. Math. Soc. 9 (1983), 129-206.
- [17] N. Fujii, *A condition for a two-weight norm inequality for singular integral operators*, Studia Math. 98 (1991), 175-190.
- [18] J. García-Cuerva and J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North Holland Math. Studies 116, North Holland, Amsterdam, 1985.
- [19] H. Helson and G. Szegő, *A problem in prediction theory*, Ann. Mat. Pura Appl. 51 (1960), 107-138.
- [20] R. Hunt, B. Muckenhoupt, and R. Wheeden, *Weighted norm inequalities for the conjugate function and Hilbert transform*, Trans. Amer. Math. Soc. 176 (1973), 227-252.
- [21] J.-L. Journé, *Zygmund Operators, Pseudo-Differential Operators and the Cauchy Integral of Calderón*, Lecture Notes in Mathematics, 994, Springer Verlag, New York, 1983.
- [22] N.H. Katz and C. Pereyra, *On the two weights problem for the Hilbert transform*, Rev. Mat. Iberoamericana 13 (1997), 189-210.
- [23] V. Khavin and N. Nikol'skiĭ, eds. *Linear and complex analysis. Problem book 3. Part I*, Lecture Notes in Mathematics, 1573, Springer-Verlag, Berlin, 1994.
- [24] P. Koosis, *Introduction to H_p spaces*, Second ed., Cambridge Tracts in Mathematics, 115, Cambridge University Press, Cambridge, 1998.
- [25] M.A. Leckband, *Structure results on the maximal Hilbert transform and two-weight norm inequalities*, Indiana Math. J. 34 (1985), 259-275.
- [26] A.K. Lerner, *Weighted norm inequalities for the local sharp maximal function*, J. Fourier Anal. Appl. 10 (2004), 465-474.
- [27] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc. 165 (1972), 207-226.
- [28] B. Muckenhoupt, *Weighted norm inequalities for classical operators*, Harmonic Analysis in Euclidean Spaces, (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, pp. 69–83, Proc. Sympos. Pure Math., XXXV, Amer. Math. Soc., Providence, 1979.
- [29] B. Muckenhoupt, private communication.
- [30] B. Muckenhoupt and R. Wheeden, *Two weight function norm inequalities for the Hardy-Littlewood maximal function and the Hilbert transform*, Studia Math. 60 (1976), 279-294.
- [31] F. Nazarov, *A counterexample to Sarason's conjecture*, preprint, 1997.
Available at www.math.msu.edu/~fedja/prepr.html.
- [32] C.J. Neugebauer, *Inserting A_p -weights*, Proc. Amer. Math. Soc. 87 (1983), 644-648.
- [33] R. O'Neil, *Fractional integration in Orlicz spaces*, Trans. Amer. Math. Soc. 115 (1965), 300-328.
- [34] C. Pérez, *Weighted norm inequalities for singular integral operators*, J. London Math. Soc. 49 (1994), 296-308.
- [35] C. Pérez, *Two weighted inequalities for potential and fractional type maximal operators*, Indiana Math. J. 43 (1994), 663-683.
- [36] C. Pérez, *On sufficient conditions for the boundedness of the Hardy-Littlewood maximal operator between weighted L^p -spaces with different weights*, Proc. London Math. Soc. 71 (1995), 135-57.
- [37] C. Pérez, *Endpoint estimates for commutators of singular integral operators*, J. Func. Anal. 128 (1995), 163-185.
- [38] C. Pérez and R. Trujillo-Gonzalez, *Sharp weighted estimates for vector-valued singular integral operators and commutators*, Tohoku Math. J. (2) **55** (2003), 109–129.
- [39] Y. Rakotonratsimba, *Two weight norm inequality for Calderón-Zygmund operators*, Acta Math. Hungar. 80 (1998), 39-54.

- [40] Y. Rakotondratsimba, *Two-weight inequality for commutators of singular integral operators*, Kobe J. Math. 16 (1999), 1-20.
- [41] M.M. Rao and Z.D. Ren, *Theory of Orlicz Spaces*, Marcel Dekker, New York, 1991.
- [42] D. Sarason, *Exposed points in H^1 , I*, Operator Theory: Advances and Applications, 41 (1989), Birkhäuser Verlag, Basel, 485–496.
- [43] D. Sarason, *Exposed points in H^1 , II*, Operator Theory: Advances and Applications, 48 (1990), Birkhäuser Verlag, Basel, 333–347.
- [44] E.T. Sawyer, *A characterization of a two-weight norm inequality for maximal operators*, Studia Math. 75 (1982), 1-11.
- [45] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
- [46] K. Stroethoff and D. Zheng, *Products of Hankel and Toeplitz operators on the Bergman spaces*, J. Func. Anal. 169 (1999), 289-313.
- [47] K. Stroethoff and D. Zheng, *Invertible Toeplitz products*, J. Func. Anal. 195 (2002), 48-70.
- [48] S. Treil, A. Volberg and D. Zheng, *Hilbert Transform, Toeplitz operators and Hankel operators, and invariant A_∞ weights*, Rev. Mat. Ibero. 13 (1997), 319-360.
- [49] A. Uchiyama, *L^p weighted inequalities for the dyadic square function*, Studia Math. 115 (1995), 135–149.
- [50] N. Vasilevski, *Toeplitz operators on the Bergman spaces: Inside-the-domain effects*, Second summer school in analysis and mathematical physics, Contemporary Math. 289, Amer. Math. Soc., Providence, 2000, 79-146.
- [51] G. Weiss, *A note on Orlicz spaces*, Portugal. Math. 15 (1950), 35-47.
- [52] J.M. Wilson, *Weighted norm inequalities for the continuous square functions*, Trans. Amer. Math. Soc. 314 (1989), 661-692.
- [53] J. Xia, *Rigged non-tangential maximal function associated with Toeplitz operators and Hankel operators*, Pacific J. Math. 182 (1998), 385-396.
- [54] V.P. Zaharjuta and V.I. Judovič, *The general form of a linear functional on H'_p* , Uspekhi Mat. Nauk 19 (1964), 2, 139-142. (Russian)
- [55] D. Zheng, *The distribution function inequality and products of Toeplitz operators and Hankel operators*, J. Funct. Anal. 138 (1996), 477–501.
- [56] A. Zygmund, *Trigonometric Series*, vols. I and II, 2nd ed., Cambridge University Press, London, 1959.

DAVID CRUZ-URIBE, DEPT. OF MATHEMATICS, TRINITY COLLEGE, HARTFORD, CT 06106-3100, USA

E-mail address: david.cruzuribe@trincoll.edu

JOSÉ MARÍA MARTELL, INSTITUTO DE MATEMÁTICAS Y FÍSICA FUNDAMENTAL, CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS, C/ SERRANO 123, 28006 MADRID, SPAIN

E-mail address: chema.martell@uam.es

CARLOS PÉREZ, DEPARTAMENTO DE ANÁLISIS MATEMÁTICO, FACULTAD DE MATEMÁTICAS, UNIVERSIDAD DE SEVILLA, 41080 SEVILLA, SPAIN

E-mail address: carlosperez@us.es