Best constants in weighted estimates for dyadic shifts

Adam Osękowski

Abstract. We identify the weighted L^p-norms of shift operators in the context of nonatomic probability spaces equipped with tree-like structures.

Mathematics Subject Classification (2010). Primary 42B35; Secondary 46E30.

Keywords. Dyadic, shift, weight, Bellman function, best constant.

1. Introduction

The purpose of this paper is to determine the exact formulas for weighted L^p norms of a certain important class of operators in a wide context of measure spaces. To present the motivation, let us start with the classical, dyadic setting. Suppose that n is a fixed dimension and let $\mathcal{D}$ denote the standard grid of dyadic cubes in $\mathbb{R}^n$. A sequence $\alpha = (\alpha_Q)_{Q\in\mathcal{D}}$ of nonnegative numbers has Carleson property, if

$$\sup_{R\in\mathcal{D}} \frac{1}{|R|} \sum_{Q\in\mathcal{D}, Q\subseteq R} \alpha_Q |Q| \leq 1,$$

where $|\cdot|$ is the Lebesgue measure. It is well-known that the Carleson property is equivalent to the following fact (see e.g. [8]): there is a family $\{E(Q)\}_{Q\in\mathcal{D}}$ of pairwise disjoint sets such that $E(Q)\subseteq Q$ and $\alpha_Q = |E(Q)|/|Q|$ for each Q. For any Carleson sequence α and any positive number r, we introduce the associated shift operator $\mathcal{A}_\alpha^r$, acting on locally integrable functions $f:\mathbb{R}^n\to\mathbb{R}$ by

$$\mathcal{A}_\alpha^r f = \left(\sum_{Q\in\mathcal{D}} \alpha_Q \langle f\rangle_Q^r \chi_Q\right)^{1/r}. \tag{1.1}$$

This work was supported by NCN grant DEC-2014/14/E/ST1/00532.

Here $\langle f\rangle_Q=\frac{1}{|Q|}\int_Q f$ is the average of f over the cube Q. Such operators are closely related to the class of the so-called sparse operators. Recall that a collection $\mathcal{S}\subset\mathcal{D}$ is called sparse, if there is a family $\{E(Q)\}_{Q\in\mathcal{S}}$ of pairwise disjoint sets such that $E(Q)\subseteq Q$ and $|E(Q)|\geq |Q|/2$ for each $Q\in\mathcal{S}$. Given any such class $\mathcal{S}$ and any $r>0$, one can introduce the associated sparse operator by

$$\mathcal{T}^{\mathcal{S},r}f=\left(\sum_{Q\in\mathcal{S}}\langle f\rangle_Q^r\chi_Q\right)^{1/r}.$$

Such an operator is of the form (1.1), if one sets $\alpha_Q=1$ for $Q\in\mathcal{S}$ and $\alpha_Q=0$ otherwise; in addition, it is easy to see that the sequence $(\alpha_Q/2)_{Q\in\mathcal{D}}$ has the Carleson property. The importance of the sparse operators (and hence also the class of shift operators) lies in the fact that such objects dominate, in an appropriate sense, large families of Calderón-Zygmund singular integrals and Littlewood-Paley square functions [5, 6, 7], and thus many interesting estimates can be reduced to the context of shift operators. This observation has been exploited very intensively in the recent literature, especially in the context of weighted inequalities, as we will briefly discuss now. Here and below, the word 'weight' will refer to a positive, locally integrable function on $\mathbb{R}^n$, and the associated weighted L^p space (for $0<p<\infty$) is defined as the collection of all (equivalence classes of) functions $f:\mathbb{R}^n\to\mathbb{R}$ for which the quantity

$$\|f\|_{L^p(w)}=\left(\int_{\mathbb{R}^n}|f|^pw\mathrm{d}x\right)^{1/p}$$

is finite. Recall that given $1<p<\infty$, a weight w on $\mathbb{R}^n$ belongs to the class A_p (or satisfies the Muckenhoupt condition A_p), if

$$[w]_{A_p}=\sup\left(\frac{1}{|Q|}\int_Q w\mathrm{d}x\right)\left(\frac{1}{|Q|}\int_Q w^{1/(1-p)}\mathrm{d}x\right)^{p-1}<\infty, \tag{1.2}$$

where the supremum is taken over all cubes $Q\subset\mathbb{R}^n$ with sides parallel to the axes. There are also versions of this condition for $p\in\{1,\infty\}$, which can be obtained by passing to the appropriate limit; however, we will not present the precise definition here, as we will not work with these endpoint cases. The A_p condition appeared for the first time in the seminal work of Muckenhoupt [10] and was shown to characterize the weighted L^p boundedness of the Hardy-Littlewood maximal function. It was soon realized that the A_p condition also characterizes the boundedness of other classical operators, including singular integrals, fractional operators, square functions and area integrals, and many, many more. The literature is extremely vast here and we recall the reader to some basic works [2, 11, 3, 12] only.

Throughout this paper we will work with a slightly more general class of *dyadic* A_p weights, which are again defined by the condition (1.2), but this time the supremum is taken over the class of all dyadic cubes Q. Such classes form a natural environment for various operators whose definition refers to the dyadic lattice $\mathcal{D}$: for instance, this is the case if we consider the dyadic maximal operator, dyadic square function or the dyadic shift operator introduced above.

The motivation for the results in this paper comes from the question about the optimal dependence of the weighted L^p norms of dyadic shift operators on the A_p characteristic of the weight. For maximal operators, Buckley [1] proved that if M is a maximal operator, then for each $1 < p < \infty$ there is a constant c_p depending only on p such that

$$\|M\|_{L^p(w)\to L^p(w)} \leq c_p[w]_{A_p}^{1/(p-1)},$$

and the exponent $1/(p-1)$ is optimal (see [4, 13] for the improvement of this result). Concerning the dyadic shifts, Lerner [7] proved that the optimal dependence is

$$\|\mathcal{A}^1_\alpha\|_{L^p(w)\to L^p(w)} \lesssim_p [w]_{A_p}^{\max\{1,1/(p-1)\}},$$

and this led to the analogous weighted L^p-estimate for a large class of Calderón-Zygmund singular integral operators. In a different paper [6], Lerner showed that

$$\|\mathcal{A}^2_\alpha\|_{L^p(w)\to L^p(w)} \lesssim_p [w]_{A_p}^{\max\{1/2,(p-1)^{-1}\}},$$

which yielded the similar statement for dyadic square functions; again, the exponent $\max\{1/2, 1/(p-1)\}$ cannot be improved. We will continue the above line of research and identify the *explicit values* of the weighted L^p norms for dyadic shift operators $\mathcal{A}^r_\alpha$ for a wide range of parameter r. The precise statement of our main result requires the introduction of a certain special parameter $d(p, c)$, whose geometric interpretation is explained on Figure 1 below. Let $c \geq 1$ and $1 < p < \infty$ be fixed. Then the line, tangent to the curve $wv^{p-1} = c$ at the point $(1, c^{1/(p-1)})$, intersects the curve $wv^{p-1} = 1$ at one point (if $c = 1$) or two points (if $c > 1$). Take the intersection point with larger w-coordinate, and denote this coordinate by $1 + d(p, c)$. Formally, $d = d(p, c)$ is the unique number in $[0, p-1)$ satisfying

$$c(1+d)(p-1-d)^{p-1} = (p-1)^{p-1}. \tag{1.3}$$

The second intersection point is given by $\big(1 + d^*(p, c), (1 + d^*(p, c))^{1/(1-p)}\big)$, where $d = d^*(p, c)$ is the unique number belonging to $(-1, 0]$ satisfying (1.3).

The following theorem is one of our main results. Here and in what follows, for any $p \in (1,\infty)$, the symbol p' is the harmonic conjugate to p, i.e., $p' = p/(p-1)$.

Theorem 1.1. *Let $1 < p < \infty$ and $0 < r < p$. Then for any $w \in A_p$ and any Carleson sequence, the associated dyadic shift operator satisfies*

$$\|\mathcal{A}^r_\alpha\|_{L^p(w)\to L^p(w)} \leq C_{p,r,[w]_{A_p}}, \tag{1.4}$$

where

$$C_{p,r,c} = \begin{cases} p'\left(\dfrac{p}{r}\right)^{\frac{1}{r}} \cdot c^{\frac{1}{r}}\left(1 + d\left(p', c^{\frac{1}{p-1}}\right)\right)^{\frac{p-1}{r}-1} & \text{if } 0 < r < p-1, \\ p'\left(\dfrac{p}{r}\right)^{\frac{1}{r}} \cdot c^{\frac{1}{p-1}}\left(1 + d(p, c)\right)^{\frac{1}{p-1}-\frac{1}{r}} & \text{if } p-1 \leq r < p. \end{cases} \tag{1.5}$$

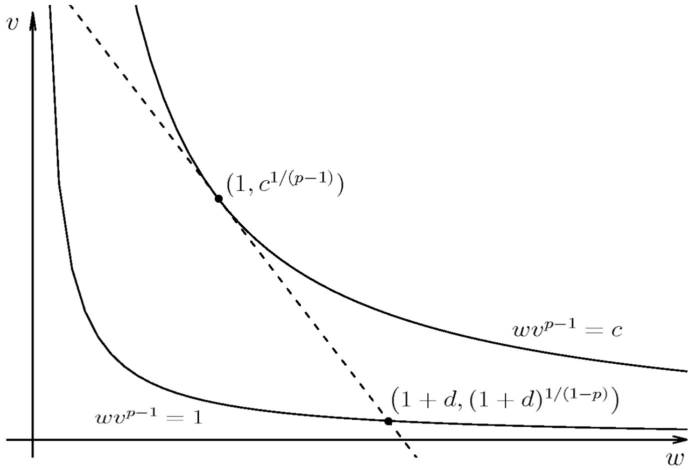

FIGURE 1. The geometric interpretation of the number $d = d(p, c)$.

The constant $C_{p,r,c}$ is the best possible in the following sense. For any p, r, c and $\varepsilon > 0$ there is a weight w satisfying $[w]_{A_p} \leq c$ and a Carleson sequence α such that

$$\|\mathcal{A}_\alpha^r\|_{L^p(w)\to L^p(w)} > C_{p,r,c} - \varepsilon. \tag{1.6}$$

Let us discuss briefly on the size of the constants. Since $0 \leq d(p,c) \leq p-1$, we easily extract the optimal dependence of $C_{p,r,c}$ on c: we have $C_{p,r,c} \asymp_p c^{\max\{1/r, 1/(p-1)\}}$ (here $A \asymp_p B$ means that the ratio A/B is bounded from below and from above by a quantity depending only on p). For $r = 1$, this is precisely the optimal dependence on weight characteristic for a wide class of singular integrals, while for $r = 2$ we recover the optimal dependence for a wide class of square functions (see above). Another special case which is worth mentioning is the L^p constant in the unweghted setting: this context corresponds to the condition $[w]_{A_p} = 1$ under which we have $C_{p,r,1} = p' \left(\frac{p}{r}\right)^{1/r}$.

As we will see below, our approach actually does not exploit any property of the dyadic lattice and hence it allows the study of shift operators in a much wider context. Suppose that (Ω, μ) is a nonatomic probability space, and call two measurable subsets A, B of Ω almost disjoint if $\mu(A \cap B) = 0$.

Definition 1.2. A set $\mathcal{T}$ of measurable subsets of Ω will be called a tree if the following conditions are satisfied:

(i) $\Omega \in \mathcal{T}$ and for every $I \in \mathcal{T}$ we have $\mu(I) > 0$.

(ii) For every $I \in \mathcal{T}$ there is a finite subset $C(I) \subset \mathcal{T}$ such that
   (a) the elements of $C(I)$ are pairwise almost disjoint subsets of I,
   (b) $I = \bigcup C(I)$.

(iii) $\mathcal{T} = \bigcup_{m\geq 0} \mathcal{T}^m$, where $\mathcal{T}^0 = \{\Omega\}$ and $T^{m+1} = \bigcup_{I\in\mathcal{T}^m} C(I)$.

(iv) We have $\lim_{m\to\infty} \sup_{I\in\mathcal{T}^m} \mu(I) = 0$.

We can redefine the notions of Carleson sequences, Muckenhoupt weights and dyadic shift operators, replacing $\mathbb{R}^n$ with Ω, the Lebesgue measure with μ and the dyadic lattice $\mathcal{D}$ with the tree $\mathcal{T}$. Then, as we shall see, the assertion of Theorem 1.1 remains valid with the same value of the constant. We should emphasize that the constants $C_{p,r,c}$ will be the best possible for each individual tree $\mathcal{T}$. It is easy to see that the normalization requirement on the measure μ can be significantly relaxed, by the use of dilation and limiting arguments; however, we will not go further into this direction, leaving it to the interested reader, and we will restrict ourselves to probabilistic setup.

Our proof can be regarded as an extension of an argument due to Moen [9]. It rests on duality and a change-of-measure, which allows to relate weighted estimates for dyadic shifts to the unweighted L^p bounds for the dyadic maximal function (with respect to an arbitrary Borel measure). However, a successful treatment of (1.4) requires the identification of a Carleson constant of a certain auxiliary sequence, which is accomplished by the so-called Bellman function method, a powerful tool used widely in probability and harmonic analysis. This intermediate result is established in the next section. The final part of the paper is devoted to the proof of Theorem 1.1 in the general probabilistic context.

2. An auxiliary bound for A_p weights

Throughout this section, $1 < p < \infty$ and $c \geq 1$ are fixed parameters and $d = d(p,c)$ is the nonnegative constant given by (1.3). Let $\beta > 0$ be another fixed number. Consider the special function $B = B_{p,c,\beta} : (0,\infty)^3 \to \mathbb{R}$ given by the formula

$$B(x,y,z) = -\frac{\beta}{d(p,c)} z^\beta + \frac{(p-1)\beta}{pd(p,c)} z^{\beta-1} y + \frac{\beta}{pcd(p,c)} x z^{p-1+\beta}.$$

We will need the following properties of this object.

Lemma 2.1. *(i) We have $B(x,y,y) \leq 0$ if $xy^{p-1} \leq c$.*

(ii) If $1 \leq xy^{p-1} \leq c$, then we have the majorization

$$B(x,y,z) \geq z^\beta - (c(1+d(p,c)))^{\beta/(p-1)} y^\beta. \tag{2.1}$$

(iii) For any $(x,y,z) \in (0,\infty)^3$ satisfying $xy^{p-1} \leq c$ and $y \leq z$, and any $r > -w$, $s > -v$ such that $(w+r)(v+s)^{p-1} \leq c$ we have

$$B(x+r, y+s, \max\{y+s, z\}) \leq B(x,y,z) + B_x(x,y,z)r + B_y(x,y,z)s. \tag{2.2}$$

Proof. (i) The function B increases as x increases, so $B(x,y,y) \leq B(cy^{1-p}, y, y)$, and the latter quantity is equal to zero.

(ii) By the above monotonicity of B with respect to x, it is enough to show the majorization for the smallest x, i.e., for $x = y^{1-p}$. Then the estimate becomes

$$\frac{(p-1)\beta}{pd(p,c)}z^{\beta-1}y+\frac{\beta}{pcd(p,c)}y^{1-p}z^{p-1+\beta}+(c(1+d(p,c)))^{\beta/(p-1)}y^{\beta}\\ -\left(1+\frac{\beta}{d(p,c)}\right)z^{\beta}\geq 0.$$

Fix z and denote the left-hand side by $F(y)$. After some lengthy but straightforward computations involving (1.3), we check that

$$F\left(z(1+d)^{1/(1-p)}\right)=F'\left(z(1+d)^{1/(1-p)}\right)=0. \tag{2.3}$$

In addition, we have

$$F''(y)=\beta y^{-1-p}\left(\frac{(p-1)z^{p-1+\beta}}{cd(p,c)}+(\beta-1)(c(1+d(p,c)))^{\beta/(p-1)}y^{p-1+\beta}\right).$$

Therefore, if $\beta\geq 1$, then F is convex on $[0,\infty)$, which combined with (2.3) gives $F\geq 0$. On the other hand, if $0<\beta<1$, then there exists $y_0>0$ such that F is convex on $(0,y_0)$ and concave on (y_0,∞). Combining this with (2.3) and the obvious observation that $F(y)\geq 0$ for sufficiently large y, we conclude that $F\geq 0$ on the whole halfline $[0,\infty)$.

(iii) For a fixed z, the function B is a linear function of x and y, so both sides of (2.2) are equal if $y+s\leq z$. If $y+s>z$, then the desired inequality is equivalent to $B(x+r,y+s,y+s)\leq B(x+r,y+s,z)$, or

$$-\frac{\beta}{d(p,c)}z^{\beta}+\frac{(p-1)\beta}{pd(p,c)}z^{\beta-1}(y+s)+\frac{\beta}{pcd(p,c)}(x+r)z^{p-1+\beta}\\ \geq -\frac{\beta}{pd(p,c)}(y+s)^{\beta}+\frac{\beta}{pcd(p,c)}(x+r)(y+s)^{p-1+\beta}.$$

Since $(y+s)^{p-1+\beta}\geq z^{p-1+\beta}$, it is enough to prove this bound when $x+r$ is as large as possible, i.e., for $x+r=c(y+s)^{1-p}$. For this choice the right-hand side vanishes and the claim reduces to

$$\frac{\beta}{d(p,c)}z^{\beta}\left[-1+\frac{p-1}{p}\cdot\frac{y+s}{z}+\frac{1}{p}\cdot\left(\frac{z}{y+s}\right)^{p-1}\right]\geq 0,$$

which follows at once from Young's inequality. □

We are ready to establish the main result of this section. We assume that (Ω,μ) is an arbitrary probability space equipped with a tree structure $\mathcal{T}$. Let M be the associated maximal operator, acting on integrable random variables φ by

$$M\varphi(\omega)=\sup\langle|\varphi|\rangle_Q,$$

where the supremum is taken over all $Q\in\mathcal{T}$ containing ω. Here and below, the averages are computed with respect to the probability measure μ.

Theorem 2.2. *Let $1 < p < \infty$, $c \geq 1$ and $\beta \geq 0$. Let $w : \Omega \to (0,\infty)$ be an A_p weight satisfying $[w]_{A_p} \leq c$ and let $v = w^{1/(1-p)}$ stand for the dual weight. Then for any $R \in \mathcal{T}$, we have*

$$\langle \big(M(v\chi_R)\big)^{\beta} \rangle_R \leq \big(c(1+d(p,c))\big)^{\frac{\beta}{p-1}} \langle v^{\beta} \rangle_R. \tag{2.4}$$

For any p, c and β, the constant $\big(c(1+d(p,c))\big)^{\frac{\beta}{p-1}}$ cannot be improved.

Proof of (2.4). We may assume that β is strictly positive, since the claim for $\beta = 0$ is obvious. We split the reasoning into a few parts.

Step 1. Some notation. Fix $R \in \mathcal{T}$: then $R \in \mathcal{T}^m$ for some nonnegative integer m. For any integer $k \geq m$ and any $\omega \in R$, let $Q^k(\omega)$ be the unique element of $\mathcal{T}^k$ which contains ω. Furthermore, define

$$x_k(\omega) = \langle w \rangle_{Q^k(\omega)}, \qquad y_k(\omega) = \langle v \rangle_{Q^k(\omega)}, \qquad z_k(\omega) = \max_{0 \leq \ell \leq k} y_\ell(\omega).$$

Of course, for any k and any $Q \in \mathcal{T}^k$, the functions x_k, y_k and z_k are constant on Q and we have

$$\int_Q x_{k+1} \mathrm{d}\mu = \mu(Q) x_k|_Q, \qquad \int_Q y_{k+1} \mathrm{d}\mu = \mu(Q) y_k|_Q. \tag{2.5}$$

Furthermore, the sequence $(z_k)_{k \geq 0}$ is non-decreasing and satisfies

$$\lim_{k\to\infty} z_k(\omega) = \sup_{k \geq 0} \langle v \rangle_{Q^k(\omega)} = \sup_{k \geq 0} \langle v\chi_R \rangle_{Q^k(\omega)} = M(v\chi_R)(\omega) \tag{2.6}$$

μ-almost everywhere. Finally, we record the double inequality

$$1 \leq x_k y_k^{p-1} \leq c, \tag{2.7}$$

noting that the left bound follows from Jensen's inequality, while the right is due to $[w]_{A_p} \leq c$.

Step 2. Monotonicity property. Let $B = B_{p,c,\beta}$ be the function constructed above. We will prove that the sequence

$$\left(\int_R B(x_k, y_k, z_k) \mathrm{d}\mu \right)_{k \geq m}$$

is non-increasing. To this end, we fix k and apply the concavity property (2.2) with $x := x_k$, $y := y_k$, $z := z_k$ and $r = x_{k+1} - x_k$, $s = y_{k+1} - y_k$. As the result, we obtain the pointwise estimate

$$\begin{aligned} &B(x_{k+1}, y_{k+1}, z_{k+1}) \\ &\leq B(x_k, y_k, z_k) + B_x(x_k, y_k, z_k)(x_{k+1} - x_k) + B_y(x_k, y_k, z_k)(y_{k+1} - y_k). \end{aligned}$$

As we have already mentioned above, if Q is an element of $\mathcal{T}^k$, then x_k, y_k and z_k are constant on Q. Furthermore, by (2.5), we have $\int_Q \big(x_{k+1} - x_k\big) d\mu = \int_Q \big(y_{k+1} - y_k\big) d\mu = 0$. Consequently, integrating the above inequality over Q, we obtain

$$\int_Q B(x_{k+1}, y_{k+1}, z_{k+1}) \mathrm{d}\mu \leq \int_Q B(x_k, y_k, z_k) \mathrm{d}\mu.$$

Summing over all $Q \in \mathcal{T}^k$ contained in R, we get the aforementioned monotonicity property of the sequence $\left(\int_R B(x_k, y_k, z_k)\mathrm{d}\mu\right)_{k\geq m}$.

Step 3. Completion of the proof. For a given k, we apply (2.1) to get

$$\begin{aligned} \int_R \left(z_k^\beta - (c(1+d(p,c)))^{\frac{\beta}{p-1}} y_k^\beta\right) \mathrm{d}\mu &\leq \int_R B(x_k, y_k, z_k)\mathrm{d}\mu \\ &\leq \int_R B(x_m, y_m, z_m)\mathrm{d}\mu. \end{aligned} \tag{2.8}$$

But the functions x_m and y_m are constant on R and $z_m = y_m$. Therefore, by Lemma 2.1 (i), $\int_R B(x_m, y_m, z_m)\mathrm{d}\mu \leq 0$. To analyze the left-hand side of (2.8), observe that by Lebesgue's monotone convergence theorem,

$$\int_R z_k^\beta \mathrm{d}\mu \to \int_R M(v\chi_R)^\beta \mathrm{d}\mu.$$

Furthermore, recall that y_k is the conditional expectation of v with respect to $\mathcal{T}^k$. Therefore, if $\beta \geq 1$, Jensen's inequality yields

$$\int_R y_k^\beta \mathrm{d}\mu \leq \int_R v^\beta \mathrm{d}\mu,$$

while for $\beta < 1$, the sequence $\left(\int_R y_k^\beta \mathrm{d}\mu\right)_{k\geq m}$ decreases to $\int_R v^\beta \mathrm{d}\mu$ (this is a simple application of Lebesgue's dominated convergence theorem). Putting all the above facts together, we get the estimate (2.4). The sharpness of this estimate will follow from our considerations below. If the constant in (2.4) could be improved, it would imply that (1.4) can also be improved which, as we shall see, is impossible. $\square$

The above estimate has the following important consequence. Introduce an auxiliary constant

$$K_{p,r,c} = \begin{cases} c\left(1 + d\left(p', c^{\frac{1}{p-1}}\right)\right)^{p-r-1} & \text{if } 0 < r < p-1, \\ c^{\frac{r}{p-1}}\left(1 + d(p,c)\right)^{\frac{1-p+r}{p-1}} & \text{if } r \geq p-1. \end{cases}$$

Theorem 2.3. *Let $1 < p < \infty$, $r > 0$ be given parameters and let w be an A_p weight satisfying $[w]_{A_p} \leq c$. Fix $R \in \mathcal{T}$ and an arbitrary Carleson sequence $(\alpha_Q)_{Q\in\mathcal{T}}$ with a finite number of nonzero terms. Then there is a family $(E(Q))_{Q\in\mathcal{T}}$ of pairwise disjoint sets such that for each Q we have $E(Q) \subseteq Q$ and*

$$\alpha_Q \mu(Q)\langle w\rangle_Q \langle v\rangle_Q^r = K_{p,r,c} \int_{E(Q)} v^{1-p+r}\, d\mu. \tag{2.9}$$

Proof. This follows by a straightforward backward induction argument. Assume first that $r \geq p-1$. Since $(\alpha_Q)_{Q\in\mathcal{T}}$ contains a finite number of nonzero terms, there is an integer N such that α_Q vanishes for all $Q \in \mathcal{T}^n$, $n \geq N$. For such Q we set $E(Q) = \emptyset$. Now, let $k < N$ and assume that we have constructed pairwise disjoint sets $E(Q) \subseteq Q$ satisfying (2.9) for all $Q \in \bigcup_{n>k} \mathcal{T}^n$.

Fix $Q \in \mathcal{T}^k$. By the A_p condition, Carleson property and (2.4), we have

$$\begin{aligned}\frac{1}{\mu(Q)} \sum_{Q' \in \mathcal{T},\, Q' \subseteq Q} \alpha_{Q'} \mu(Q') \langle w \rangle_{Q'} \langle v \rangle_{Q'}^r &\leq c \frac{1}{\mu(Q)} \sum_{Q' \in \mathcal{T},\, Q' \subseteq Q} \alpha_{Q'} \mu(Q') \langle v \rangle_{Q'}^{1-p+r} \\ &\leq c \frac{1}{\mu(Q)} \sum_{Q' \in \mathcal{T},\, Q' \subseteq Q} \mu(e(Q')) \langle v \rangle_{Q'}^{1-p+r} \\ &\leq c \frac{1}{\mu(Q)} \int_Q (M(v\chi_Q))^{1-p+r} \mathrm{d}\mu \\ &\leq c(c(1+d(p,c)))^{\frac{1-p+r}{p-1}} \frac{1}{\mu(Q)} \int_Q v^{1-p+r} \mathrm{d}\mu.\end{aligned}$$

Here $\{e(Q)\}_{Q \in \mathcal{T}}$ is the family of subsets guaranteed by the Carleson property of $(\alpha_Q)_{Q \in \mathcal{T}}$. Hence, by the induction hypothesis,

$$\begin{aligned}\alpha_Q \mu(Q) \langle w \rangle_Q \langle v \rangle_Q + \sum_{\substack{Q' \in \mathcal{T}, \\ Q' \neq Q}} c(c(1+d(p,c)))^{\frac{1-p+r}{p-1}} \int_{E(Q')} v^{1-p+r} \mathrm{d}\mu \\ \leq c(c(1+d(p,c)))^{\frac{1-p+r}{p-1}} \int_Q v^{1-p+r} \mathrm{d}\mu,\end{aligned}$$

or, equivalently,

$$\alpha_Q \mu(Q) \langle w \rangle_Q \langle v \rangle_Q \leq c(c(1+d(p,c)))^{\frac{1-p+r}{p-1}} \int_{Q \setminus \bigcup E(Q')} v^{1-p+r} \mathrm{d}\mu,$$

where the union in the lower limit of the integral is taken over all $Q' \in \mathcal{T}$, $Q' \neq Q$. Therefore, we may pick a subset $E(Q)$ with the desired properties, which completes the induction step.

The proof in the case $0 < r < p-1$ follows the same pattern, but applied to the dual weight. Namely, the weight $v = w^{1/(1-p)}$ belongs to $A_{p'}$ and satisfies $[v]_{A_{p'}} = [w]_{A_p}^{1/(p-1)} \leq c^{1/(p-1)}$, so

$$\begin{aligned}&\frac{1}{\mu(Q)} \sum_{Q' \in \mathcal{T},\, Q' \subseteq Q} \alpha_{Q'} \mu(Q') \langle w \rangle_{Q'} \langle v \rangle_{Q'}^r \\ &\qquad \leq c^{\frac{r}{p-1}} \cdot \frac{1}{\mu(Q)} \sum_{Q' \in \mathcal{T},\, Q' \subseteq Q} \alpha_{Q'} \mu(Q') \langle w \rangle_{Q'}^{1-\frac{r}{p-1}} \\ &\qquad \leq c^{\frac{r}{p-1}} \cdot \frac{1}{\mu(Q)} \int_Q (M(w\chi_Q))^{1-\frac{r}{p-1}} \mathrm{d}\mu \\ &\qquad \leq c^{\frac{r}{p-1}} \left(c^{\frac{1}{p-1}} \left(1 + d\left(p', c^{\frac{1}{p-1}}\right)\right)\right)^{p-r-1} \cdot \frac{1}{\mu(Q)} \int_Q w^{1-\frac{r}{p-1}} \mathrm{d}\mu \\ &\qquad = c^{\frac{r}{p-1}} \left(c^{\frac{1}{p-1}} \left(1 + d\left(p', c^{\frac{1}{p-1}}\right)\right)\right)^{p-r-1} \cdot \frac{1}{\mu(Q)} \int_Q v^{1-p+r} \mathrm{d}\mu.\end{aligned}$$

Here in the last inequality we exploited (2.4) applied to the weight v and the parameter p'. The remaining part of the proof is the same as above. $\square$

3. Proof of Theorem 1.1

The contents of this section splits naturally into a few parts.

3.1. Proof of (1.4)

Fix $1 < p < \infty$, $r < p$ and an A_p weight w satisfying $[w]_{A_p} = c$. By standard approximation, we may restrict ourselves to shift operators associated with Carleson sequences possessing only a finite number of nonzero terms. To prove (1.4), it is enough to show that for any $f \in L^p(w)$ and $g \in L^{p/(p-r)}(w^{-r/(p-r)})$,

$$\sum_{Q\in\mathcal{T}} \alpha_Q \mu(Q)\langle f\rangle_Q^r \langle g\rangle_Q \leq C_{p,r,c}^r \|f\|_{L^p(w)}^r \|g\|_{L^{\frac{p}{p-r}}\left(w^{-\frac{r}{p-r}}\right)}.$$

Substituting fv and gw in the places of f and g, respectively, we rewrite the above bound in the equivalent form

$$\sum_{Q\in\mathcal{T}} \alpha_Q \mu(Q)\langle w\rangle_Q \langle v\rangle_Q^r \cdot \langle f\rangle_{v,Q}^r \langle g\rangle_{w,Q} \leq C_{p,r,c}^r \|f\|_{L^p(v)}^r \|g\|_{L^{\frac{p}{p-r}}(w)}.$$

Here $\langle f\rangle_{v,Q}$ and $\langle g\rangle_{w,Q}$ denote the averages of f and g over Q with respect to the measures $vd\mu$ and $wd\mu$, respectively. Let $(E(Q))_{Q\in\mathcal{T}}$ be the collection of subsets of Ω guaranteed by Theorem 2.3. Since

$$\int_{E(Q)} v^{1-p+r}\mathrm{d}\mu = \int_{E(Q)} v^{\frac{r}{p}} w^{\frac{p-r}{p}}\mathrm{d}\mu \leq v(E(Q))^{\frac{r}{p}} w(E(Q))^{\frac{p-r}{p}},$$

we get, by Hölder's inequality and the unweighted L^p bound for maximal functions,

$$\begin{aligned}
&\sum_{Q\in\mathcal{T}} \alpha_Q \mu(Q)\langle w\rangle_Q \langle v\rangle_Q^r \cdot \langle f\rangle_{v,Q}^r \langle g\rangle_{w,Q} \\
&\qquad \leq K_{p,r,c} \sum_{Q\in\mathcal{T}} v(E(Q))^{\frac{r}{p}} w(E(Q))^{\frac{p-r}{p}} \cdot \langle f\rangle_{v,Q}^r \langle g\rangle_{w,Q} \\
&\qquad \leq K_{p,r,c} \left(\sum_{Q\in\mathcal{T}} v(E(Q))\langle f\rangle_{v,Q}^p\right)^{\frac{r}{p}} \left(\sum_{Q\in\mathcal{T}} w(E(Q))\langle g\rangle_{w,Q}^{\frac{p}{p-r}}\right)^{\frac{p-r}{p}} \\
&\qquad \leq K_{p,r,c} \left(\int_X (M_v f)^p v\mathrm{d}\mu\right)^{\frac{r}{p}} \left(\int_X (M_w g)^{\frac{p}{p-r}} w\mathrm{d}\mu\right)^{\frac{p-r}{p}} \\
&\qquad \leq K_{p,r,c} \left(\frac{p}{p-1}\right)^r \frac{p}{r} \cdot \|f\|_{L^p(v)}^r \|g\|_{L^{\frac{p}{p-r}}(w)}.
\end{aligned}$$

Here M_w, M_v stand for the maximal functions on the measure spaces $(\Omega, wd\mu)$ and $(\Omega, vd\mu)$ equipped with the tree $\mathcal{T}$, respectively. The above bound is precisely the desired estimate: the latter multiplicative constant is equal to $C_{p,r,c}^r$.

3.2. Sharpness: a special function and its properties

Now we will address the optimality of the constant in (1.4). The examples which yield the sharpness have a quite complicated, fractal-type structure and hence we will use a different method. Namely, we will show that the validity of the estimate implies the existence of a certain special function of five variables, enjoying appropriate size and concavity conditions. Then, exploiting these properties in the right order, we will get the desired lower bound for the constant.

Let $\mathcal{T}$ be a given tree on some probability space (X,μ) and let $1<p<\infty$, $r\in(0,p)$, $c\in[1,\infty)$, $C>0$ be fixed parameters. Consider the domain

$$\mathcal{D}=\mathcal{D}_{p,r,c,C}=\Big\{(x,y,u,v,t)\in[0,\infty)^4\times[0,1]\,:\,1\le uv^{p-1}\le c\Big\}$$

and the function $\mathcal{B}^{X,\mathcal{T}}:\mathcal{D}\to\mathbb{R}$ defined by the formula

$$\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t)=\sup\left\{\int_X\left(\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}}-C^pf^p\right)w\mathrm{d}\mu\right\}.$$

Here the supremum is taken over all nonnegative functions f on X satisfying $\int_X f\mathrm{d}\mu=x$ and taking a finite number of values, all Carleson sequences $\alpha=(\alpha_R)_{R\in\mathcal{T}}$ satisfying $\sum_{R\in\mathcal{T}}\alpha_R\mu(R)\le t$ and all A_p weights w on X satisfying $[w]_{A_p}\le c$, $\int_X w\mathrm{d}\mu=u$, $\int_X w^{1/(1-p)}\mathrm{d}\mu=v$. Finally, define $\mathbb{B}:\mathcal{D}\to\mathbb{R}$ by

$$\mathbb{B}(x,y,u,v,t)=\inf_{X,\mathcal{T}}\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t).$$

This functions has the following properties.

Lemma 3.1. *Fix $C'\in(0,C)$. Assume that there is a probability space (X,μ) and a tree $\mathcal{T}$ such that $\|\mathcal{A}_\alpha^{r,X}\|_{L^p(w)\to L^p(w)}\le C'$ for all Carleson sequences α and all weights w satisfying $[w]_{A_p}\le c$. Then $\mathbb{B}$ is finite on $\mathcal{D}$.*

Proof. We have

$$\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}}\le\left(\frac{C}{C'}\right)^p(\mathcal{A}_\alpha^{r,X}f)^p+\kappa y^p$$

for some constant κ depending only on C and C'. Consequently, we get

$$\begin{aligned}\mathbb{B}^{X,\mathcal{T}}(x,y,u,v,t)&\le\kappa y^pu+\left(\frac{C}{C'}\right)^p\sup\left\{\int_X\left(\left(\mathcal{A}_\alpha^{r,X}f\right)^p-(C')^pf^p\right)w\mathrm{d}\mu\right\}\\&\le\kappa y^pu,\end{aligned}$$

by the assumption on the norm $\|\mathcal{A}_\alpha^{r,X}\|_{L^p(w)\to L^p(w)}$. Therefore $\mathcal{B}^{X,\mathcal{T}}$, and hence also $\mathbb{B}$, are finite on $\mathcal{D}$. □

Lemma 3.2. *We have $\mathbb{B}(x,y,u,v,t)\ge(y^p-C^px^p)u$.*

Proof. For any X and $\mathcal{T}$, consider a constant function $f\equiv x$, a zero Carleson sequence α and any A_p weight satisfying the required conditions on the averages of w and $w^{1/(1-p)}$. Then by the very definition of $\mathcal{B}^{X,\mathcal{T}}$ we get

$$\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t)\ge\int_X\left(\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}}-C^pf^p\right)w\mathrm{d}\mu=(y^p-C^px^p)u$$

and the claim is proved, since the right-hand side does not depend on X or $\mathcal{T}$. $\square$

Next we will handle the main, concavity-type property of $\mathbb{B}$. To this end, we need two simple observations studied in two lemmas below.

Lemma 3.3. *Let (X,μ) be an arbitrary probability space equipped with a tree $\mathcal{T}$. Then for any $\lambda \in (0,1)$ there is a measurable set E such that $\mu(E)=\lambda$ and*

- *E is the union of pairwise almost disjoint elements of $\mathcal{T}$;*
- *$X\setminus E$ is the union of pairwise almost disjoint elements of $\mathcal{T}$.*

Proof. This follows from a simple inductive argument. Set $E_0=F_0=\emptyset$ and define E_n, $F_n\in\mathcal{T}^n$ as maximal sets (i.e., possessing maximal measure) satisfying $E_n\supset E_{n-1}$, $F_n\supset F_{n-1}$ and $\mu(E_n)\leq\lambda$, $\mu(F_n)\leq 1-\lambda$. It is easy to see that $E=\bigcup_n E_n$ and $X\setminus E=\bigcup_n F_n$ have the desired decomposition property; furthermore, the equality $\mu(E)=\lambda$ follows from the condition (iv) in the definition of a tree. $\square$

Lemma 3.4. *Let (X,μ) be an arbitrary probability space equipped with a tree structure $\mathcal{T}$. For any u, $v>0$ satisfying $1\leq uv^{p-1}\leq c$, there is a weight w on X satisfying $\int_X w\,d\mu=u$, $\int_X w^{1/(1-p)}\,d\mu=v$ and $[w]_{A_p}\leq c$.*

Proof. If $uv^{p-1}=1$, then the constant weight $w\equiv u$ satisfies all the requirements; so, suppose that $uv^{p-1}>1$. Let I be a line segment passing through (u,v), lying entirely in the set $\{(x,y): xy^{p-1}\leq c\}$, with endpoints (u_-,v_-) and (u_+,v_+) belonging to the curve $xy^{p-1}=1$. It is easy to see that such an interval exists (if $uv^{p-1}=c$, then it is unique - it must be then tangent to the curve $xy^{p-1}=c$). Let $\lambda\in(0,1)$ be determined by $(u,v)=\lambda(u_-,v_-)+(1-\lambda)(u_+,v_+)$, and let E be the subset of X guaranteed by the previous lemma: in particular, we have

$$E=\bigcup_j Q_j^-,\qquad X\setminus E=\bigcup_j Q_j^+$$

for some pairwise almost disjoint sets $Q_j^\pm$ of $\mathcal{T}$. Set $w=u_-\chi_E+u_+\chi_{X\setminus E}$. Then

$$\int_X w\mathrm{d}\mu=\lambda u_-+(1-\lambda)u_+=u$$

and

$$\int_X w^{1/(1-p)}\mathrm{d}\mu=\lambda u_-^{1/(1-p)}+(1-\lambda)u_+^{1/(1-p)}=\lambda v_-+(1-\lambda)v_+=v.$$

It remains to check that $[w]_{A_p}\leq c$. Pick an arbitrary $Q\in\mathcal{T}$. If $Q=X$, then

$$\left(\frac{1}{\mu(Q)}\int_Q w\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q w^{1/(1-p)}\mathrm{d}\mu\right)^{p-1}=uv^{p-1}\leq c, \tag{3.1}$$

by the assumption of the lemma. If Q is almost contained in E or in $X\setminus E$ (in the sense that $\mu(Q\setminus E)=0$ or $\mu(Q\cap E{=}0)$, then w is constant on Q and the above inequality is also true. Finally, if $Q\neq X$ and Q is not almost contained in E or $X \setminus E$, then we set $Q^- = Q \cap E$, $Q^+ = Q \setminus E$ and note that

$$\frac{1}{\mu(Q)}\int_Q w\mathrm{d}\mu = \frac{1}{\mu(Q)}\int_{Q^-} w\mathrm{d}\mu + \frac{1}{\mu(Q)}\int_{Q^+} w\mathrm{d}\mu = \frac{\mu(Q_-)}{\mu(Q)}u_- + \frac{\mu(Q_+)}{\mu(Q)}u_+,$$

and, similarly,

$$\frac{1}{\mu(Q)}\int_Q w^{1/(1-p)}\mathrm{d}\mu = \frac{\mu(Q_-)}{\mu(Q)}v_- + \frac{\mu(Q_+)}{\mu(Q)}v_+.$$

Thus we have proved that the point $\left(\frac{1}{\mu(Q)}\int_Q w\mathrm{d}\mu, \frac{1}{\mu(Q)}\int_Q w^{1/(1-p)}\mathrm{d}\mu\right)$ is a convex combination of (u_-, v_-) and (u_+, v_+), i.e., it lies on the segment I. Since I lies below the curve $xy^{p-1} \leq c$, this yields (3.1) and completes the proof. ☐

We are ready for the concavity condition of $\mathbb{B}$.

Lemma 3.5. *Suppose that two points (u_-, v_-) and (u_+, v_+) in $(0, \infty)^2$ have the property that the entire line segment joining them lies below the curve $uv^{p-1} = c$. Assume further that $x_\pm \geq 0$, $y \geq 0$, $t_\pm \in [0, 1]$ and $\lambda_\pm \in (0, 1)$ satisfy $\lambda_- + \lambda_+ = 1$ and set $x = \lambda_- x_- + \lambda_+ x_+$, $u = \lambda_- u_- + \lambda_+ u_+$, $v = \lambda_- v_- + \lambda_+ v_+$. If $\Delta t = t - (\lambda_- t_- + \lambda_+ t_+) \geq 0$ and $y' = [y^r + x^r \Delta t]^{1/r}$, then*

$$\mathbb{B}(x, y, u, v, t) \geq \lambda_- \mathbb{B}(x_-, y', u_-, v_-, t_-) + \lambda_+ \mathbb{B}(x_+, y', u_+, v_+, t_+), \tag{3.2}$$

Proof. It is convenient to split the reasoning into a few steps.

Step 1. Splicing. Pick an arbitrary probability space (X, μ) with a tree structure $\mathcal{T}$. Let E be set guaranteed by Lemma 3.3, corresponding to the measure λ_-. Setting $E_- = E$ and $E_+ = X \setminus E$, let $E_\pm = \bigcup Q^j_\pm$ be the decompositions of $E_\pm$ into pairwise almost disjoint elements of $\mathcal{T}$. For each j, treat Q^+_j as a probability space, with the normalized measure $\mu/|Q^+_j|$ and the tree $\mathcal{T}^+_j$ formed by those $Q \in \mathcal{T}$, which are contained in Q^+_j. By the definition of $\mathcal{B}^{Q^+_j, \mathcal{T}^+_j}(x_+, y', u_+, v_+, t_+)$, there exist a function f^+_j, a weight w^+_j and a Carleson sequence $\alpha^{+,j}$ satisfying the appropriate requirements, for which the integral

$$\frac{1}{\mu(Q^+_j)}\int_{Q^+_j}\left(\left[y^+_j + \mathcal{A}^r_{\alpha^{+,j}} f\right]^p - C^p f^p\right) w\mathrm{d}\mu$$

is as close to $\mathcal{B}^{Q^+_j, \mathcal{T}^+_j}(x_+, y', u_+, v_+, t_+)$ as we wish. A similar statement is true for any Q^-_j. Now we splice all the functions, weights and Carleson sequences into one function, one weight and one Carleson sequences as follows. We set $f = \sum_j f^-_j \chi_{Q^-_j} + \sum_j f^+_j \chi_{Q^+_j}$, $w = \sum_j w^-_j \chi_{Q^-_j} + \sum_j w^+_j \chi_{Q^+_j}$ and define a sequence $(\alpha_R)_{R \in \mathcal{T}}$ by taking the union of all sequences $(\alpha_{Q^\pm_j})_j$, putting $\alpha_X = \Delta t$ and letting $\alpha_R = 0$ for all remaining R (i.e., those $R \neq X$ which are not contained in any $Q^\pm_j$).

Step 2. Checking the properties. Let us verify that as the result of the above splicing procedure, we get objects which enjoy the properties needed at the definition of $\mathbb{B}(x,y,u,v,t)$. First, we see that f is a nonnegative function and

$$\int_X f\mathrm{d}\mu = \sum_j \int_{Q_j^-} f\mathrm{d}\mu + \sum_j \int_{Q_j^+} f\mathrm{d}\mu = \mu(E_-)x_- + \mu(E_+)x_+ = x,$$

since the average of f over each $Q_j^\pm$ is equal to $x_\pm$. The verification that w satisfies the appropriate conditions is the repetition of the arguments appearing in the previous lemma (see also the below analysis of the Carleson property of the sequence α). It remains to handle the sequence α. First note that

$$\begin{aligned}\frac{1}{\mu(X)}\sum_{Q\in\mathcal{T}}\alpha_Q\mu(Q) &= \alpha_X + \sum_j \sum_{Q\subseteq Q_j^-}\alpha_Q\mu(Q) + \sum_j\sum_{Q\subseteq Q_j^+}\alpha_Q\mu(Q)\\ &\leq \Delta t + \sum_j t_-\mu(Q_j^-) + \sum_j t_+\mu(Q_j^+) = t.\end{aligned}$$

To check the Carleson property, fix $R\in\mathcal{T}$. If $R=X$, then $\sum_{Q\in\mathcal{T}}\alpha_Q\mu(Q)\leq t\leq\mu(X)$, as we have just proved. If R is almost contained in some $Q_j^\pm$, then

$$\sum_{Q\in\mathcal{T},\,Q\subseteq R}\alpha_Q\mu(Q)\leq\mu(R)$$

follows from the Carleson property of the sequence $\alpha^{\pm,j}$. Finally, if $R\neq X$ is not almost contained in any $Q_j^\pm$, then we write

$$\begin{aligned}\sum_{Q\in\mathcal{T},\,Q\subseteq R}\alpha_Q\mu(Q) &= \sum_j\sum_{Q\in\mathcal{T},\,Q\subseteq R\cap Q_j^-}\alpha_Q\mu(Q) + \sum_j\sum_{Q\in\mathcal{T},\,Q\subseteq R\cap Q_j^+}\alpha_Q\mu(Q)\\ &\leq \sum_j \mu(R\cap Q_j^-) + \sum_j\mu(R\cap Q_j^+)\leq\mu(R),\end{aligned}$$

where in the first inequality we have again used the Carleson property of $\alpha_j^\pm$.

Step 3. The concavity condition. By the very definition of $\mathcal{B}^{X,\mathcal{T}}$, we obtain

$$\begin{aligned}\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t) &\geq \int_X\left(\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}} - C^pf^p\right)w\mathrm{d}\mu\\ &= \sum_j\int_{Q_j^-}\left(\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}} - C^pf^p\right)w\mathrm{d}\mu\\ &\quad + \sum_j\int_{Q_j^+}\left(\left[y^r+(\mathcal{A}_\alpha^{r,X}f)^r\right]^{\frac{p}{r}} - C^pf^p\right)w\mathrm{d}\mu.\end{aligned}$$

However, for each j we have, on $Q_j^\pm$,

$$y^r + (\mathcal{A}_\alpha^{r,X}f)^r = y^r + x^r\Delta t + \left(\mathcal{A}_{\alpha^{\pm,j}}^{r,Q_j^\pm}f\right)^r = (y')^r + \left(\mathcal{A}_{\alpha^{\pm,j}}^{r,Q_j^\pm}f\right)^r$$

and hence the right-hand side above can be made arbitrarily close to

$$\sum_j \mu(Q_j^-)\mathcal{B}^{Q_j^-,\mathcal{T}_j^-}(x_-,y',u_-,v_-,t_-) + \sum_j \mu(Q_j^+)\mathcal{B}^{Q_j^+,\mathcal{T}_j^+}(x_+,y',u_+,v_+,t_+).$$

Therefore, by the definition of $\mathbb{B}$, we see that for any $\varepsilon > 0$ we have

$$\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t) \geq \lambda_-\mathbb{B}(x_-,y_-,u_-,v_-,t_-) + \lambda_+\mathbb{B}(x_+,y_+,u_+,v_+,t_+) - \varepsilon.$$

Since the probability space and the tree were arbitrary, the claim is proved. □

We conclude the analysis by providing a certain homogeneity property.

Lemma 3.6. *For any $(x,y,u,v,t) \in \mathcal{D}$ and any λ, $\eta > 0$ we have*

$$\mathbb{B}(\lambda x, \lambda y, \eta u, \eta^{1/(1-p)}v, t) = \lambda^p \eta \mathbb{B}(x,y,u,v,t). \tag{3.3}$$

Proof. Pick an arbitrary probability space (X,μ) with some tree $\mathcal{T}$. Let f, w, α be objects as in the definition of $\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t)$. Then λf, ηw and α satisfy all the requirements of the definition of $\mathcal{B}^{X,\mathcal{T}}(\lambda x, \lambda y, \eta u, \eta^{1/(1-p)}v, t)$, so

$$\begin{aligned}
&\mathcal{B}^{X,\mathcal{T}}(\lambda x, \lambda y, \eta u, \eta^{1/(1-p)}v, t)\\
&\qquad\qquad \geq \int_X \left(\left[\lambda^r y^r + (\mathcal{A}_\alpha^{r,X}(\lambda f))^r\right]^{\frac{p}{r}} - C^p(\lambda f)^p \right)(\eta w)\mathrm{d}\mu\\
&\qquad\qquad = \lambda^p\eta \int_X \left(\left[y^r + (\mathcal{A}_\alpha^{r,X} f)^r\right]^{\frac{p}{r}} - C^p f^p \right) w\mathrm{d}\mu.
\end{aligned}$$

Since f, w and α were arbitrary, we get

$$\mathcal{B}^{X,\mathcal{T}}(\lambda x, \lambda y, \eta u, \eta^{1/(1-p)}v, t) \geq \lambda^p\eta\mathcal{B}^{X,\mathcal{T}}(x,y,u,v,t) \geq \lambda^p\eta\mathbb{B}(x,y,u,v,t).$$

Taking the infimum over all X and $\mathcal{T}$ gives

$$\mathbb{B}(\lambda x, \lambda y, \eta u, \eta^{1/(1-p)}v, t) \geq \lambda^p\eta\mathbb{B}(x,y,u,v,t).$$

To get the reverse, replace first x, y, u and v with $\lambda^{-1}x$, $\lambda^{-1}y$, $\eta^{-1}u$ and $\eta^{-1/(1-p)}v$, and then put λ^{-1} and η^{-1} in the place of λ and μ. □

3.3. Sharpness, the unweighted case

First we study the case in which the weights are constant, i.e., satisfy $[w]_{A_p} = 1$. Let $1 < p < \infty$ and $r \in (0,p)$ be fixed. Suppose that for some probability space (X,μ) with a tree structure $\mathcal{T}$ we have $\|\mathcal{A}_\alpha^{r,X}\|_{L^p\to L^p} \leq C'$ for any Carleson sequence α. We apply the machinery developed in the previous subsection, setting $c = 1$, picking an arbitrary $C > C'$ and considering the associated function $\mathbb{B}$. Pick a small positive number δ and apply (3.2), with $u = v = u_\pm = v_\pm = 1$, $t = t_+ = 1$, $t_- = 0$,

$$x_- = 1 - \frac{\delta}{(1+\delta)^p - 1}, \quad x_+ = 1+\delta, \quad \lambda_- = 1-(1+\delta)^{-p}, \quad \lambda_+ = (1+\delta)^{-p}$$

and $y = \left[\frac{1-(1+\delta)^{-p}}{(1+\delta)^r - 1}\right]^{1/r}$. Then $x = \lambda_- x_- + \lambda_+ x_+ = 1$, $\Delta t = 1-(1+\delta)^{-p}$ and $(y')^r = y^r + x^r\Delta t = y^r + 1 - (1+\delta)^{-p} = y^r(1+\delta)^r$, so (3.2) yields

$$\begin{aligned}&\mathbb{B}(1, y, 1, 1, 1)\\ &\geq \lambda_-\mathbb{B}\left(1-\frac{\delta}{(1+\delta)^p-1}, y(1+\delta), 1, 1, 0\right) + \lambda_+\mathbb{B}(1+\delta, y(1+\delta), 1, 1, 1).\end{aligned}$$

By (3.3), we have $\lambda_+\mathbb{B}(1+\delta, y(1+\delta), 1, 1, 1) = \mathbb{B}(1, y, 1, 1, 1)$. By Lemma 3.1, the latter quantity is finite, so combining this with the above estimate yields

$$\mathbb{B}\left(1-\frac{\delta}{(1+\delta)^p-1}, y(1+\delta), 1, 1, 0\right) \leq 0.$$

By Lemma 3.2, this gives $y^p(1+\delta)^p - C^p\left(1-\frac{\delta}{(1+\delta)^p-1}\right)^p \leq 0$. Letting $\delta \to 0$, we get $y \to \left(\frac{p}{r}\right)^{1/r}$ and $1-\frac{\delta}{(1+\delta)^p-1} \to \frac{p-1}{p}$, and hence $C \geq \frac{p}{p-1}\left(\frac{p}{r}\right)^{1/r}$. Since $C > C'$ was arbitrary, we get that the latter bound is true for C' as well. This is precisely the desired sharpness.

3.4. Sharpness, the weighted case, $r \geq p-1$

Suppose that $1 < p < \infty$ and $r \in [p-1, p)$ are fixed. Suppose that there is a probability space (X, μ) with a tree $\mathcal{T}$ such that $\|\mathcal{A}_\alpha^{r,X}\|_{L^p(w)\to L^p(w)} \leq C'$ for any Carleson sequence α and any A_p weight w satisfying $[w]_{A_p} \leq c$. Take an arbitrary constant $C > C'$ and let $\mathbb{B}$ be the associated special function constructed in Subsection 3.2. Let $\tilde{c} \in (1, c)$ be a constant close to c and let δ be a small positive number. Consider the line segment starting at the point $(1-\delta, \tilde{c}^{1/(p-1)}(1-\delta)^{-1/(p-1)})$, passing through $(1, \tilde{c}^{1/(p-1)})$ and ending at the curve $xy^{p-1} = 1$. If δ is sufficiently small, then the line segment lies below the curve $xy^{p-1} = c$ (this is due to the fact that the first two of the aforementioned points lie on the curve $xy^{p-1} = \tilde{c}$). A glance at Figure 1 reveals that the second endpoint of the line segment is of the form $(1+\tilde{d}, (1+\tilde{d})^{1/(p-1)})$, with $\tilde{d} \to d(p, c)$ as $\delta \to 0$ and $\tilde{c} \to c$.

Now, introduce the parameter $s = \delta^{-1}\left[\left(1+\frac{\delta}{\tilde{d}}\right)^{1/p}\left(1+\frac{\delta}{1-\delta}\right)^{1/p} - 1\right]$, so that

$$(1+\delta s)^p = \frac{\tilde{d}+\delta}{\tilde{d}(1-\delta)}. \tag{3.4}$$

Apply (3.2) with $\lambda_- = \frac{\delta}{\tilde{d}+\delta}$, $\lambda_+ = \frac{\tilde{d}}{\tilde{d}+\delta}$, $\tilde{x}_- = 1-\tilde{d}s$, $x_+ = 1+\delta s$, $y = \left(\frac{\delta}{(\tilde{d}+\delta)((1+\delta s)^r-1)}\right)^{1/r}$, $u_- = 1+\tilde{d}$, $u_+ = 1-\delta$, $v_- = (1+\tilde{d})^{1/(1-p)}$, $v_+ = \tilde{c}^{1/(p-1)}(1-\delta)^{1/(1-p)}$, $t_- = 0$ and $t_+ = 1$. Then $x = 1$, $u = 1$, $v = \tilde{c}^{1/(p-1)}$,

$\Delta t = \frac{\delta}{\tilde{d}+\delta}$ and $y' = y(1+\delta s)$, so we obtain

$$\begin{aligned}
\mathbb{B}(1, y, 1, &\tilde{c}^{1/(p-1)}, 1) \\
&\geq \frac{\delta}{\tilde{d}+\delta}\mathbb{B}(1-\tilde{d}s, y(1+\delta s), 1+\tilde{d}, (1+\tilde{d})^{1/(1-p)}, 0) \\
&\quad + \frac{\tilde{d}}{\tilde{d}+\delta}\mathbb{B}(1+\delta s, y(1+\delta s), 1-\delta, \tilde{c}^{1/(p-1)}(1-\delta)^{1/(1-p)}, 1).
\end{aligned}$$

Using the homogeneity property (3.3) and (3.4), we check that the second term on the right is equal to $\mathbb{B}(1, y, 1, \tilde{c}^{1/(p-1)}, 1)$. Since this quantity is finite (by virtue of Lemma 3.1), the above estimate implies

$$\mathbb{B}(1-\tilde{d}s, y(1+\delta s), 1+\tilde{d}, (1+\tilde{d})^{1/(1-p)}, 0) \leq 0,$$

which combined with Lemma 3.2 gives $y^p(1+\delta s)^p - C^p(1-\tilde{d}s)^p \leq 0$. Now we let $\delta \to 0$ and then $\tilde{c} \to c$. Then it is easy to see that $\tilde{d} \to d(p,c)$, $s \to \frac{d(p,c)+1}{pd(p,c)}$, $y \to \left(\frac{p}{r(d(p,c)+1)}\right)^{1/r}$ and the latter estimate implies

$$C \geq \frac{p}{p-d(p,c)-1}\left(\frac{p}{r(d(p,c)+1)}\right)^{1/r} = C_{p,r,c},$$

where the last equality follows from (1.3). As in the unweighted context, since $C > C'$ was chosen arbitrarily, we see that $C' \geq C_{p,r,c}$ and the sharpness is established.

3.5. Sharpness, the weighted case, $r < p-1$

The argument is word-by-word the same as previously, the only difference is that we pick a *negative* parameter δ (close to 0). As the result, the constant $\tilde{d}$ (which determines the second endpoint of the appropriate line segment passing through $(1, \tilde{c}^{1/(p-1)})$) is also negative. Furthermore, if we let $\delta \to 0$ and then $\tilde{c} \to c$, then, as one easily sees by looking at Figure 1, $(1+\tilde{d}, (1+\tilde{d})^{1/(1-p)})$ converges to the second intersection point of the dotted line with the curve $xy^{p-1} = 1$. Equivalently, $\tilde{d}$ converges to the unique solution $d^*(p,c) \in (-1, 0)$ of the equation (1.3). All the remaining arguments are the same and, as the result, we obtain

$$C \geq \frac{p}{p-d^*(p,c)-1}\left(\frac{p}{r(d^*(p,c)+1)}\right)^{1/r}.$$

It remains to note that the right-hand side is equal to $C_{p,r,c}$. To see this, note that the equation

$$c(1+d^*(p,c))(p-1-d^*(p,c))^{p-1} = (p-1)^{p-1}$$

is equivalent to

$$c^{\frac{1}{p-1}}\left(1-\frac{d^*(p,c)}{p-1}\right)\left(p'-1+\frac{d^*(p,c)}{p-1}\right)^{p'-1} = (p'-1)^{p'-1}.$$

In other words, we have $-\frac{d^*(p,c)}{p-1} = d\left(p', c^{\frac{1}{p-1}}\right)$. Plugging this above and using (1.3), we get the estimate $C \geq C_{p,r,c}$ and hence also $C' \geq C_{p,r,c}$, since the constant $C \in (C', \infty)$ was arbitrary. This completes the proof.

Acknowledgments

The author would like to thank an anonymous Referee for the careful reading of the paper and several helpful suggestions.

References

[1] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253–272.

[2] R. Coifman and C. Fefferman, *Weighted norm inequalities for maximal functions and singular integrals*, Studia Math. **51** (1974), 241–250.

[3] R. Gundy, R. Wheeden, *Weighted integral inequalities for the nontangential maximal function, Lusin area integral, and Walsh-Paley series*, Studia Math. **49** (1974), 107–124.

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

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

[6] A. Lerner, *Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals*, Adv. Math. **226** (2011), 3912–3926.

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

[8] A. D. Melas, *The Bellman functions of dyadic-like maximal operators and related inequalities*, Adv. Math. 192 (2005), 310–340.

[9] K. Moen, *Sharp weighted bounds without testing or extrapolation*, Arch. Math. **99**, 457–466. | Cite as

[10] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc **165** (1972), 207–226.

[11] B. Muckenhoupt and R. Wheeden, *Weighted norm inequalities for fractional integrals*, Trans. Amer. Math. Soc. **192** (1974), 261–274.

[12] F. L. Nazarov, S. R. Treil and A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc., **12** (1999), pp. 909–928.

[13] A. Osękowski, *Best constants in Muckenhoupt's inequality*, Ann. Acad. Sci. Fenn. Math. **42** (2017), 889–904.

[14] G. Rey and A. Reznikov, *Extremizers and sharp weak-type estimates for positive dyadic shifts*, Adv. Math. **254** (2014), 664–681.

Adam Osękowski
Department of Mathematics, Informatics and Mechanics
University of Warsaw
Banacha 2, 02-097 Warsaw
Poland
e-mail: ados@mimuw.edu.pl