Annales Academiæ Scientiarum Fennicæ
Mathematica
Volumen 45, 2020, 1171–1185

ON A_p–A_q WEIGHTED ESTIMATES FOR MAXIMAL OPERATORS

Adam Osękowski

University of Warsaw, Department of Mathematics, Informatics and Mechanics
Banacha 2, 02-097 Warsaw, Poland; ados@mimuw.edu.pl

Abstract. The paper is devoted to the study of sharp versions of mixed A_p–A_q weighted estimates for the dyadic maximal function $\mathcal{M}_d$ on $\mathbf{R}^n$. For given parameters $1 < p < \infty$ and $1 \le q \le \infty$, if a weight w satisfies Muckenhoupt's condition A_p, then we have the sharp A_p–A_q bound

$$\|\mathcal{M}_d\|_{L^p(w)\to L^p(w)} \le \frac{p^{1+1/p}}{p-1}\left(\frac{q}{q-1}\right)^{(q-1)/p}[w]_{A_p}^{1/p}[w^{1/(1-p)}]_{A_q}^{1/p}$$

(for $q \in \{1,\infty\}$, the constant is understood as an appropriate limit). Actually, a wider class of related sharp two-weight estimates for $\mathcal{M}_d$ is established. The results hold true in a more general context of maximal operators on probability spaces associated with a tree-like structure.

1. Introduction

The principal goal of this paper is to study L^p-boundedness of the dyadic maximal operator and to measure the size of the norm in terms of various mixed characteristics of the underlying weight. We start with recalling the necessary background and notation. The dyadic maximal operator $\mathcal{M}$ on $\mathbf{R}^n$ is an operator acting on locally integrable functions $\varphi\colon \mathbf{R}^n \to \mathbf{R}$ by the formula

$$\mathcal{M}\varphi(x) = \sup\left\{\frac{1}{|Q|}\int_Q |\varphi(y)|\,\mathrm{d}y\colon x \in Q,\ Q \subset \mathbf{R}^n \text{ is a dyadic cube}\right\}.$$

Here the dyadic cubes are those formed by the grids $2^{-N}\mathbf{Z}^n$, $N = 0,1,2,\ldots$, and $|Q|$ denotes the Lebesgue measure of Q. This maximal operator is of fundamental importance to analysis and PDEs, and in many applications it is of interest to control it efficiently, i.e., to have optimal or at least good bounds for its norms. For instance, $\mathcal{M}$ satisfies the weak-type $(1,1)$ inequality

$$\lambda\,|\{x \in \mathbf{R}^n \quad \mathcal{M}\varphi(x) \ge \lambda\}| \le \int_{\{\mathcal{M}\varphi\ge\lambda\}} |\phi(u)|\,\mathrm{d}u, \quad \phi \in L^1(\mathbf{R}^n), \tag{1.1}$$

which, after integration, yields the corresponding L^p estimate

$$\|\mathcal{M}\varphi\|_{L^p(\mathbf{R}^n)} \le \frac{p}{p-1}\|\varphi\|_{L^p(\mathbf{R}^n)}, \quad 1 < p \le \infty. \tag{1.2}$$

Both estimates are sharp: the constant 1 in (1.1) and the constant $p/(p-1)$ in (1.2) cannot be decreased. These two results have been successfully extended in numerous directions and applied in various contexts of harmonic analysis. See e.g. [5, 6, 7, 8, 9, 14, 15] and the monograph [3], consult also references therein.

As we have already mentioned above, we will be interested in certain mixed weighted versions of (1.2). In what follows, the word 'weight' will refer to a nonnegative, integrable function on the underlying measure space. The following statement

https://doi.org/10.5186/aasfm.2020.4544
2010 Mathematics Subject Classification: Primary 42B25; Secondary 46E30, 60G42.
Key words: Maximal, dyadic, Bellman function, best constants.

is a consequence of the classical work of Muckenhoupt [10]. Suppose that $1 < p < \infty$ is given and fixed, and let w be a weight on $\mathbf{R}^n$. Then $\mathcal{M}$ is bounded as an operator on the weighted space

$$L^p(w) = \left\{ f\colon \mathbf{R}^n \to \mathbf{R}\colon \|f\|_{L^p(w)} = \left(\int_{\mathbf{R}^n} |f|^p w \,\mathrm{d}x \right)^{1/p} < \infty \right\}$$

if and only if w belongs to the dyadic A_p class, i.e.,

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

where the supremum is taken over all dyadic cubes Q in $\mathbf{R}^n$. The classes A_p can be extended to the cases $p = 1$ and $p = \infty$ by a straightforward limiting procedure. A weight w satisfies Muckenhoupt's condition A_1, if

$$[w]_{A_1} := \sup \operatorname*{essup}_{x \in Q} \frac{\mathcal{M}w(x)}{w(x)} < \infty.$$

Furthermore, w is an A_∞ weight if

$$[w]_{A_\infty} := \sup \left(\frac{1}{|Q|} \int_Q w \,\mathrm{d}x \right) \exp \left(\frac{1}{|Q|} \int_Q \log w^{-1} \,\mathrm{d}x \right) < \infty.$$

Both suprema above are taken over all dyadic cubes Q in $\mathbf{R}^n$.

The above result of Muckenhoupt is a starting point for many interesting further questions. For example, one can ask about the dependence of $\|\mathcal{M}\|_{L^p(w)\to L^p(w)}$ on the size of the characteristic $[w]_{A_p}$. More precisely, for a given $1 < p < \infty$, the problem is to find the least number $\alpha = \alpha(p)$ such that

$$\|\mathcal{M}\|_{L^p(w)\to L^p(w)} \leq C_p [w]_{A_p}^{\alpha(p)}$$

for some C_p depending only on p. This problem was solved in the nineties by Buckley [1], who showed that the optimal exponent $\alpha(p)$ is equal to $1/(p-1)$. This result has been recently strengthened significantly by Osękowski in [13]. That paper contains, for a given $1 < p < \infty$ and $c \in [1, \infty)$, the identification of the smallest constant $C_{p,c}$ such that the following holds: if w is an A_p weight satisfying $[w]_{A_p} = c$, then

$$\|\mathcal{M}\|_{L^p(w)\to L^p(w)} \leq C_{p,c}.$$

Another extension of Buckley's result, which also serves as our motivation here, is the following two-weight estimate obtained by Hytönen and Pérez in [4]. For any $1 < p < \infty$ and any pair (w, v) of weights on $\mathbf{R}^n$,

$$\|\mathcal{M}\|_{L^p(v)\to L^p(w)} \leq \frac{4ep}{p-1} \left([w, v^{1/(1-p)}]_{A_p} [v^{1/(1-p)}]_{A_\infty} \right)^{1/p}, \tag{1.3}$$

where

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

the supremum being taken over all dyadic cubes Q in $\mathbf{R}^n$. To see that this statement does generalize Buckley's result, apply it to $w = v$ being an A_p weight: then $[w, v^{1/(1-p)}]_{A_p} = [w]_{A_p}$ and $[v^{1/(1-p)}]_{A_\infty} \leq [v^{1/(1-p)}]_{A_{p/(p-1)}} = [w]_{A_p}^{1/(p-1)}$.

We will be interested in the sharp version of (1.3) in a much wider context. Let us start with an appropriate definition of tree structures on probability spaces, following [5].

Definition 1.1. Suppose that (X,μ) is a nonatomic probability space. A set $\mathcal{T}$ of measurable subsets of X will be called a tree if the following conditions are satisfied:

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

(ii) For every $Q\in\mathcal{T}$ there is a finite subset $C(Q)\subset\mathcal{T}$ containing at least two elements such that

  (a) the elements of $C(Q)$ are pairwise disjoint subsets of Q,

  (b) $Q=\bigcup C(Q)$.

(iii) $\mathcal{T}=\bigcup_{m\ge 0}\mathcal{T}^m$, where $\mathcal{T}^0=\{X\}$ and $T^{m+1}=\bigcup_{Q\in\mathcal{T}^m}C(Q)$.

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

An important example, which links this definition with the preceding considerations, is the cube $X=[0,1)^n$ endowed with Lebesgue measure and the tree of its dyadic subcubes. Any probability space equipped with a tree gives rise to the corresponding maximal operator $\mathcal{M}_{\mathcal{T}}$, acting on integrable functions $f\colon X\to\mathbf{R}$ by the formula

$$\mathcal{M}_{\mathcal{T}}f(x)=\sup\left\{\frac{1}{\mu(Q)}\int_Q|f|\,\mathrm{d}\mu\colon x\in Q,\ Q\in\mathcal{T}\right\}.$$

In analogy to the dyadic setting described above, we say that a weight w on X satisfies Muckenhoupt's condition A_p (where $1<p<\infty$ is a fixed parameter), if

$$[w]_{A_p}:=\sup_{Q\in\mathcal{T}}\left(\frac{1}{\mu(Q)}\int_Q w\,\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q w^{-1/(p-1)}\mathrm{d}\mu\right)^{p-1}<\infty.$$

The characteristics $[w]_{A_1}$, $[w]_{A_\infty}$ and $[w,\sigma]_{A_p}$ are defined analogously. Furthermore, the weighted space $L^p(w)$ is given by

$$L^p(w)=\left\{f\colon X\to\mathbf{R}\colon \|f\|_{L^p(w)}=\left(\int_X|f|^p w\,\mathrm{d}\mu\right)^{1/p}<\infty\right\}.$$

Our main result is the following sharp version of (1.3).

Theorem 1.2. *Let X be a probability space equipped with a tree $\mathcal{T}$. Suppose that $1<p<\infty$, $1\le q\le\infty$. If (w,v) is a pair of weights on X satisfying $[w,v^{1/(1-p)}]_{A_p}<\infty$ and $[v^{1/(1-p)}]_{A_q}<\infty$ (with respect to $\mathcal{T}$), then we have*

$$\|\mathcal{M}_{\mathcal{T}}\|_{L^p(v)\to L^p(w)}\le\frac{p^{1+1/p}}{p-1}\left(\frac{q}{q-1}\right)^{(q-1)/p}[w,v^{1/(1-p)}]_{A_p}^{1/p}[v^{1/(1-p)}]_{A_q}^{1/p}\tag{1.4}$$

(for $q\in\{1,\infty\}$, the constant is understood as the appropriate limit). For each p and q the constant cannot be decreased, even if $w=v$.

In particular, the estimate (1.4) is valid and sharp also in the classical setting of $[0,1)^n$ equipped with Lebesgue's measure and the tree of dyadic subcubes; by straightforward dilation and scaling, this result extends to the whole $\mathbf{R}^n$.

A few words about the proof are in order. Our approach to (1.4) will exploit the theory of two-weight estimates. A classical result of Sawyer [17] asserts that if w, v are two weights on $\mathbf{R}^n$, then the (dyadic) maximal operator $\mathcal{M}$ is bounded as an operator from $L^p(v)$ to $L^p(w)$ if and only if the weights satisfy the so-called testing condition

$$\int_Q\big(\mathcal{M}(v^{-1/(p-1)}\chi_Q)\big)^p w\,\mathrm{d}x\le C\int_Q v^{-1/(p-1)}\,\mathrm{d}x$$

for all dyadic cubes Q, where C depends only on p, w and v. We will study a sharp version of this testing condition for the weights satisfying the assumptions of Theorem 1.2 in the above context of probability spaces. Then we will combine this estimate with the weighted version of Carleson embedding theorem (cf. [12], [22], [13]) and obtain the desired bound (1.4). Fortunately, the combination of these sharp estimates yield the inequality in which the constant is still optimal. We will handle the testing conditions with the use of the so-called Bellman function method. The technique reduces the problem of proving a given inequality to the search for a certain special function, enjoying appropriate size conditions and concavity. The literature on this subject is extremely large, for more information and the exemplary applications, we refer the interested reader to the works [11, 12, 16, 18, 19, 20, 21] and the references therein.

The next section contains the proof of (1.4). Section 3 is devoted to the construction of an example showing that the estimate is sharp.

2. Proof of (1.4)

Our main result will be deduced from a slightly more general estimate formulated in Theorem 2.1 below. For the precise statement, we need to introduce a technical parameter d, a key object in our further considerations. The geometric interpretation of this parameter is explained on Figure 1 below.

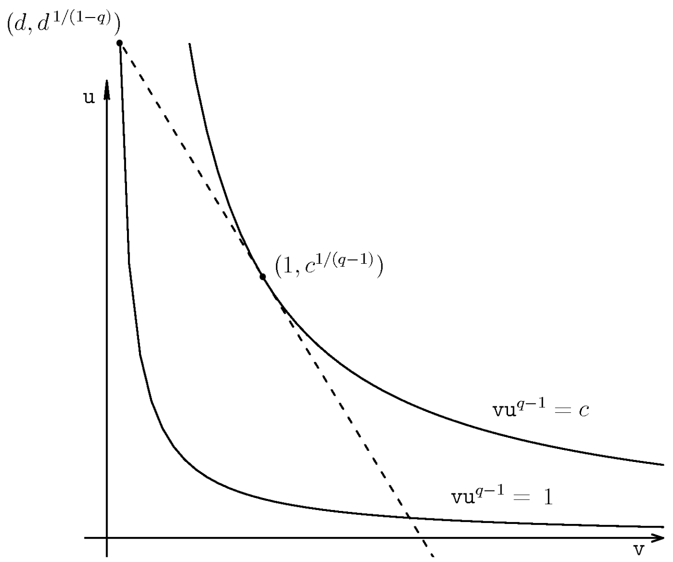

Figure 1. The geometric interpretation of the parameter $d = d(q, c)$.

Let $c \geq 1$ and $1 < q < \infty$ be fixed. Then the line, tangent to the curve $\mathtt{v}\mathtt{u}^{q-1} = c$ at the point $(1, c^{1/(q-1)})$, intersects the curve $\mathtt{v}\mathtt{u}^{q-1} = 1$ at one point (if $c = 1$) or two points (if $c > 1$). Take the intersection point with smaller $\mathtt{v}$-coordinate, and denote this coordinate by $d(q, c)$. Formally, $d = d(q, c)$ is the unique number in $(0, 1]$ satisfying the equation

$$cd\left(\frac{q-d}{q-1}\right)^{q-1} = 1. \tag{2.1}$$

We extend this definition to the boundary cases $q \in \{1, \infty\}$ by limiting procedure. Namely, we set $d(1, c) = 1/c$ and define $d(\infty, c)$ to be the unique number $d \in (0, 1]$ satisfying $cde^{1-d} = 1$.

Theorem 2.1. *Let X be a probability space equipped with a tree $\mathcal{T}$. If (w,v) is a pair of weights on X satisfying $[w, v^{1/(1-p)}]_{A_p} < \infty$ and $[v^{1/(1-p)}]_{A_q} < \infty$ (with respect to $\mathcal{T}$), then we have*

$$\|\mathcal{M}_{\mathcal{T}}\|_{L^p(v)\to L^p(w)} \leq \frac{p[w, v^{1/(1-p)}]_{A_p}^{1/p}}{p-1}\big(1-p+pd(q,[v^{1/(1-p)}]_{A_q})^{-1}\big)^{1/p}. \tag{2.2}$$

The proof of the above statement rests on two lemmas.

Lemma 2.2. *Suppose that a pair (w,σ) of weights satisfies $[w,\sigma]_{A_p} \leq c$ (with respect to $\mathcal{T}$). Then for any $R \in \mathcal{T}$,*

$$\int_R \big(\mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\big)^p w\,\mathrm{d}\mu \leq pc\int_R \mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\,\mathrm{d}\mu + (1-p)c\int_R \sigma\,\mathrm{d}\mu. \tag{2.3}$$

Both constants pc and $(1-p)c$ are the best possible.

Proof. We split the reasoning into four parts.

Step 1. An associated Bellman function. For any $c \geq 1$, introduce the domain

$$\mathcal{D}_{p,c} = \{(\mathtt{w},\mathtt{v},\mathtt{z}) \in (0,\infty)^3 \colon \mathtt{w}\mathtt{v}^{p-1} \leq c\}$$

and let $B\colon \mathcal{D}_{p,c} \to \mathbf{R}$ be given by the formula

$$B(\mathtt{w},\mathtt{v},\mathtt{z}) = \mathtt{z}^p\mathtt{w} - pc\mathtt{z}.$$

Step 2. Auxiliary notation. The set R belongs to some generation of the tree $\mathcal{T}$: say, $R \in \mathcal{T}^m$. For any n and any $x \in X$, let $Q^n(x)$ be the element of $\mathcal{T}^n$ which contains x; such a set is uniquely defined for almost all x. Next, introduce the notation

$$\mathtt{w}_n = \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} w\,\mathrm{d}\mu, \quad \mathtt{v}_n = \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} \sigma\,\mathrm{d}\mu, \quad \mathtt{z}_n = \max_{m\leq k\leq n} \mathtt{v}_k.$$

In the probabilistic language, the functional sequences $(\mathtt{w}_n)_{n\geq m}$ and $(\mathtt{v}_n)_{n\geq m}$ are martingales (on the probability space $(R, \mu(\cdot)/\mu(R))$) corresponding to the terminal variables w and σ, while $(\mathtt{z}_n)_{n\geq m}$ is the maximal function of $(\mathtt{v}_n)_{n\geq m}$. Note that for any $n \geq m$ and any $Q \in \mathcal{T}^n$, the functions $\mathtt{w}_n$, $\mathtt{v}_n$ and $\mathtt{z}_n$ are constant on Q and we have

$$\int_Q \mathtt{w}_{n+1}\,\mathrm{d}\mu = \mu(Q)\mathtt{w}_n|_Q. \tag{2.4}$$

Furthermore, the sequence $(\mathtt{z}_n)_{n\geq m}$ is nondecreasing and satisfies

$$\begin{aligned}\lim_{n\to\infty} \mathtt{z}_n(x) &= \sup_{n\geq m} \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} \sigma\,\mathrm{d}\mu \\ &= \sup_{n\geq 0} \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} \sigma\chi_R\,\mathrm{d}\mu = \mathcal{M}_{\mathcal{T}}(\sigma\chi_R)(x)\end{aligned} \tag{2.5}$$

almost everywhere.

Step 3. Monotonicity property. The main part of the proof is to show that the sequence $\big(\int_R B(\mathtt{w}_n,\mathtt{v}_n,\mathtt{z}_n)\,\mathrm{d}\mu\big)_{n\geq m}$ is nonincreasing. Observe that

$$B(\mathtt{w}_{n+1},\mathtt{v}_{n+1},\mathtt{z}_{n+1}) = \mathtt{z}_{n+1}^p\mathtt{w}_{n+1} - pc\mathtt{z}_{n+1} \leq \mathtt{z}_n^p\mathtt{w}_{n+1} - pc\mathtt{z}_n. \tag{2.6}$$

Indeed, if $\mathtt{z}_{n+1} = \mathtt{z}_n$, there is nothing to prove; therefore, assume that $\mathtt{z}_{n+1} > \mathtt{z}_n$. By the mean-value property,

$$\mathtt{z}_{n+1}^p\mathtt{w}_{n+1} - pc\mathtt{z}_{n+1} - \big(\mathtt{z}_n^p\mathtt{w}_{n+1} - pc\mathtt{z}_n\big) = p(\mathtt{z}_{n+1}-\mathtt{z}_n)\big(a^{p-1}\mathtt{w}_{n+1} - c\big),$$

for some $a \in (\mathtt{z}_n, \mathtt{z}_{n+1})$. However, since $\mathtt{z}_{n+1} = \max\{\mathtt{v}_{n+1}, \mathtt{z}_n\}$, we see that $\mathtt{v}_{n+1} = \mathtt{z}_{n+1}$ and hence $a < \mathtt{v}_{n+1}$. Therefore, the condition $[w,\sigma]_{A_p} \leq c$ implies $a^{p-1}\mathtt{w}_{n+1} - c \leq \mathtt{w}_{n+1}\mathtt{v}_{n+1}^{p-1} - c \leq 0$ and hence the bound (2.6) follows. Consequently, by (2.4), we get that for any $Q \in \mathcal{T}^n$,

$$\int_Q B(\mathtt{w}_{n+1}, \mathtt{v}_{n+1}, \mathtt{z}_{n+1})\,\mathrm{d}\mu \leq \int_Q B(\mathtt{w}_n, \mathtt{v}_n, \mathtt{z}_n)\,\mathrm{d}\mu,$$

and summing over all Q contained in R yields the desired monotonicity.

Step 4. Completion of the proof. By the previous step and the inequality $\mathtt{w}_m\mathtt{v}_m^{p-1} \leq c$, we get

$$\begin{aligned}\int_R \left(\mathtt{z}_n^p w - pc\mathtt{z}_n\right)\mathrm{d}\mu &= \int_R \left(\mathtt{z}_n^p \mathtt{w}_n - pc\mathtt{z}_n\right)\mathrm{d}\mu \leq \int_R \left(\mathtt{z}_m^p \mathtt{w}_m - pc\mathtt{z}_m\right)\mathrm{d}\mu \\ &= \int_R \left(\mathtt{v}_m^p \mathtt{w}_m - pc\mathtt{v}_m\right)\mathrm{d}\mu \\ &\leq (1-p)c\mathtt{v}_m\mu(R) = (1-p)c\int_R \sigma\,\mathrm{d}\mu.\end{aligned}$$

To deduce (2.3), it suffices to let $n \to \infty$ and combine (2.5) with Lebesgue's monotone convergence theorem. It remains to handle the sharpness of this estimate. If any of the constants pc or $(1-p)c$ could be decreased, this would lead to the improvement of the constant in (2.2) (which is impossible, as we shall see in the next section). □

The second lemma concerns the following sharp maximal inequality for A_q weights. Recall the definition of the parameter $d(q,c)$ given in (2.1) above.

Lemma 2.3. *For any A_q weight σ on X and any $R \in \mathcal{T}$ we have the inequality*

$$\int_R \mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\,d\mu \leq d(q,[\sigma]_{A_q})^{-1}\int_R \sigma\,d\mu. \tag{2.7}$$

Proof of Lemma 2.3 for $q = 1$. For this particular value of q the argument is very simple. Fix an A_1 weight σ. By the very definition of the A_1 condition, we have

$$\int_R \mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\,\mathrm{d}\mu \leq \int_R \mathcal{M}_{\mathcal{T}}\sigma\,\mathrm{d}\mu \leq [\sigma]_{A_1}\int_R \sigma\,\mathrm{d}\mu = d(1,[\sigma]_{A_1})^{-1}\int_R \sigma\,\mathrm{d}\mu,$$

as desired. □

Proof of Lemma 2.3 for $q \in (1,\infty)$. Pick an arbitrary A_q weight σ. For brevity, we will denote $[\sigma]_{A_q}$ by c and write d instead of $d(q,c)$. If $c = 1$, then $d = 1$, the weight is constant and the claim is evident; therefore, from now on, we assume that $c > 1$ (and hence $d < 1$). Let R be an arbitrary element of the tree $\mathcal{T}$; it belongs to some generation $\mathcal{T}^m$. As in the proof of Theorem 2.2, the reasoning is split into four parts.

Step 1. An associated Bellman function and its properties. Consider the domain

$$\mathcal{D}_q = \{(\mathtt{u},\mathtt{v},\mathtt{z}) \in (0,\infty)^3 \colon 1 \leq \mathtt{u}^{q-1}\mathtt{v} \leq c,\ \mathtt{z} \geq \mathtt{v}\}$$

and introduce the Bellman function $\mathcal{B}\colon \mathcal{D}_q \to \mathbf{R}$ by

$$\mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{z}) = \alpha\left[(q-1)c^{-1/(q-1)}\mathtt{z}^{q/(q-1)}\mathtt{u} - q\mathtt{z} + \mathtt{v}\right] + \mathtt{v}d^{-1},$$

where

$$\alpha = \frac{d^{1/(q-1)}}{c^{-1/(q-1)} - d^{q/(q-1)}} > 0.$$

The latter inequality is equivalent to $c < d^{-q}$, which follows immediately from the definition of d: indeed, otherwise we would have $(\frac{q-d}{qd-d})^{q-1} \leq 1$, a contradiction.

Let us prove the majorizations

$$\mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{v}) \leq \mathtt{v}d^{-1} \tag{2.8}$$

and

$$\mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{z}) \geq \mathtt{z}. \tag{2.9}$$

The first estimate follows at once from the condition $c^{-1/(q-1)}\mathtt{v}^{q/(q-1)}\mathtt{u} \leq \mathtt{v}$. To prove the second bound, note that $\mathcal{B}$ increases as $\mathtt{v}$ increases and hence it is enough to establish (2.9) for $\mathtt{v} = \mathtt{u}^{1-q}$. A straightforward calculation shows that for a fixed $\mathtt{z}$, the function

$$F(\mathtt{u}) = \alpha\left[(q-1)c^{-1/(q-1)}\mathtt{z}^{q/(q-1)}\mathtt{u} - q\mathtt{z} + \mathtt{u}^{1-q}\right] + \mathtt{u}^{1-q}d^{-1} - \mathtt{z}$$

is convex on $(0,\infty)$ and satisfies $F((\mathtt{z}d)^{-1/(q-1)}) = F'((\mathtt{z}d)^{-1/(q-1)}) = 0$. This yields (2.9).

Step 2. Monotonicity property. Define the functional sequences $(\mathtt{v}_n)_{n\geq m}$, $(\mathtt{z}_n)_{n\geq m}$ as in the proof of Theorem 2.2. Furthermore, for $n \geq m$ let

$$\mathtt{u}_n(x) = \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} \sigma^{-1/(q-1)}\,\mathrm{d}\mu.$$

By Jensen's inequality and the definition of $[\sigma]_{A_q}$, we have $1 \leq \mathtt{u}_n^{q-1}\mathtt{v}_n \leq c$. Let us show that the sequence $\big(\int_R \mathcal{B}(\mathtt{u}_n,\mathtt{v}_n,\mathtt{z}_n)\,\mathrm{d}\mu\big)_{n\geq m}$ is nonincreasing. Indeed, by the mean value property,

$$\mathcal{B}(\mathtt{u}_{n+1},\mathtt{v}_{n+1},\mathtt{z}_{n+1}) - \mathcal{B}(\mathtt{u}_{n+1},\mathtt{v}_{n+1},\mathtt{z}_n) = \alpha q(\mathtt{z}_{n+1}-\mathtt{z}_n)\big((a/c)^{1/(q-1)}\mathtt{u}_{n+1} - 1\big)$$

for some $a \in [z_n, z_{n+1}]$. This expression is nonpositive: if $\mathtt{z}_{n+1} = \mathtt{z}_n$, then this is obvious, while for $\mathtt{z}_n < \mathtt{z}_{n+1}$ we apply the bound

$$(a/c)^{1/(q-1)}\mathtt{u}_{n+1} \leq (\mathtt{z}_{n+1}/c)^{1/(q-1)}\mathtt{u}_{n+1} = (\mathtt{v}_{n+1}/c)^{1/(q-1)}\mathtt{u}_{n+1} \leq 1.$$

Consequently,

$$\int_R \mathcal{B}(\mathtt{u}_{n+1},\mathtt{v}_{n+1},\mathtt{z}_{n+1})\,\mathrm{d}\mu \leq \int_R \mathcal{B}(\mathtt{u}_{n+1},\mathtt{v}_{n+1},\mathtt{z}_n)\,\mathrm{d}\mu = \int_R \mathcal{B}(\mathtt{u}_n,\mathtt{v}_n,\mathtt{z}_n)\,\mathrm{d}\mu,$$

where the latter equality follows from the fact that the dependence of $\mathcal{B}$ on the variables $\mathtt{u}$ and $\mathtt{v}$ is linear.

Step 3. By the previous step and the estimates (2.8), (2.9), we obtain

$$\int_R \mathtt{z}_n\,\mathrm{d}\mu \leq \int_R \mathcal{B}(\mathtt{u}_n,\mathtt{v}_n,\mathtt{z}_n)\,\mathrm{d}\mu = \int_R \mathcal{B}(\mathtt{u}_m,\mathtt{v}_m,\mathtt{v}_m)\,\mathrm{d}\mu \leq d^{-1}\int_R \sigma\,\mathrm{d}\mu.$$

Since $(\mathtt{z}_n)_{n\geq m}$ increases almost everywhere to $\mathcal{M}_{\mathcal{T}}(\sigma\chi_R)$ (see (2.5)), Lebesgue's monotone convergence theorem implies

$$\frac{1}{\sigma(R)}\int_R \mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\,\mathrm{d}\mu \leq d^{-1},$$

and the claim follows, since R was taken arbitrarily. $\square$

Proof of Lemma 2.3 for $q = \infty$. The reasoning is essentially the same as previously; we will present the necessary modifications and leave the rigorous verification to the reader. We need to use a different Bellman function. Introduce the domain

$$\mathcal{D}_\infty = \{(\mathtt{u},\mathtt{v},\mathtt{z}) \in (0,\infty)^3 \colon 1 \leq \exp(-\mathtt{u})\mathtt{v} \leq c,\ \mathtt{z} \geq \mathtt{v}\}$$

and define $\mathcal{B}\colon \mathcal{D}_\infty \to \mathbf{R}$ by

$$\mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{z}) = (-\mathtt{z}\mathtt{u} - \mathtt{z} + \mathtt{z}\ln\mathtt{z} - \mathtt{z}\ln c + \mathtt{v})(1-d)^{-1} + \mathtt{v}d^{-1}.$$

It is easy to check that we have the estimates

$$\mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{v}) \leq \mathtt{v}d^{-1}, \quad \mathcal{B}(\mathtt{u},\mathtt{v},\mathtt{z}) \geq \mathtt{z}. \tag{2.10}$$

Consider the sequences $(\mathtt{v}_n)_{n\geq m}$, $(\mathtt{z}_n)_{n\geq m}$ as previously, and set

$$\mathtt{u}_n(x) = \frac{1}{\mu(Q^n(x))}\int_{Q^n(x)} \log\sigma\,\mathrm{d}\mu.$$

Arguing as above, one shows that the sequence $\big(\int_R \mathcal{B}(\mathtt{u}_n,\mathtt{v}_n,\mathtt{z}_n)\,\mathrm{d}\mu\big)_{n\geq m}$ is nonincreasing and deduces the claim by Lebesgue's monotone convergence theorem and the majorizations (2.10). □

Remark 2.4. There is a natural question how the above Bellman functions were discovered; we will give some informal argumentation about the search for $\mathcal{B}$ from the previous lemma. The desired function should satisfy (2.8) and (2.9), so it cannot be too big nor too small. The key indication is contained in Step 2 of the above proof. Since $(\mathtt{u}_n)_{n\geq m}$, $(\mathtt{v}_n)_{n\geq m}$ behave in a martingale manner and $(\mathtt{z}_n)_{n\geq 0}$ is nondecreasing, the monotonicity of $\big(\int_R \mathcal{B}(\mathtt{u}_n,\mathtt{v}_n,\mathtt{z}_n)\mathrm{d}\mu\big)_{n\geq m}$ follows if one proves that $\mathcal{B}_{\mathtt{z}}(\mathtt{w},\mathtt{z},\mathtt{z}) \leq 0$ and $\mathcal{B}(\cdot,\cdot,\mathtt{z})$ is concave. (Just inspect carefully the above proof). In our search for $\mathcal{B}$, we assumed that $\mathcal{B}$ is actually *linear* with respect to $\mathtt{u}$ and $\mathtt{v}$; in addition, we forced $B_{\mathtt{z}}(\mathtt{u},\mathtt{z},\mathtt{z})$ to vanish for the extremal choice of $\mathtt{u}$, for which $c^{-1/(q-1)}\mathtt{v}^{q/(q-1)}\mathtt{u} = \mathtt{v}$. By these two observations, it is not difficult to obtain the Bellman function, after some experimentation.

The final ingredient of the proof of (2.2) is the following sharp weighted version of Carleson embedding theorem (see [4, 13, 22]).

Theorem 2.5. *Suppose that w is an A_p weight on X. Let K be a positive constant and assume that nonnegative numbers α_Q, $Q\in\mathcal{T}$, satisfy*

$$\sum_{Q\subseteq R}\alpha_Q\left(\frac{1}{\mu(Q)}\int_Q \sigma\,d\mu\right)^p \leq K\int_R \sigma\,d\mu \tag{2.11}$$

for all $R\in\mathcal{T}$. Then for any integrable and nonnegative function f on X we have

$$\sum_{Q\in\mathcal{T}}\alpha_Q\left(\frac{1}{\mu(Q)}\int_Q f\,d\mu\right)^p \leq K\left(\frac{p}{p-1}\right)^p\int_X f^p\sigma^{1-p}\,d\mu. \tag{2.12}$$

Equipped with the above facts, we are ready for the proof of our main result.

Proof of (2.2) *and* (1.4). Put $\sigma = v^{1/(1-p)}$. The combination of Lemmas 2.2 and 2.3 shows that

$$\int_R \big(\mathcal{M}_{\mathcal{T}}(\sigma\chi_R)\big)^p w\,\mathrm{d}\mu \leq [w,\sigma]_{A_p}\Big(pd(q,[\sigma]_{A_q})^{-1} + (1-p)\Big)\int_R\sigma\,\mathrm{d}\mu. \tag{2.13}$$

It is well-known (see e.g. [13]) that this inequality implies (2.11) with the constant $K = [w,\sigma]_{A_p}\big(pd(q,[\sigma]_{A_q})^{-1} + 1 - p\big)$. Consequently the inequality (2.12) is also true, and this is precisely the desired weighted bound (2.2). To deduce (1.4), we will assume that $q\in(1,\infty)$; the proof for $q\in\{1,\infty\}$ is similar and left to the reader. It suffices to show that

$$1 - p + pd(p,[\sigma]_{A_q})^{-1} \leq \left(\frac{q}{q-1}\right)^{q-1} p[\sigma]_{A_q},$$

or equivalently, setting $c = [\sigma]_{A_q}$ and using (2.1),

$$\big((1-p)d(q,c)+p\big)\left(\frac{q-d(q,c)}{q-1}\right)^{q-1} \leq p\left(\frac{q}{q-1}\right)^{q-1}.$$

However, the left-hand side is obviously a decreasing function of $d(q,c)$ and both sides become equal if we let $d(q,c)\downarrow 0$. This gives the claim. $\square$

3. Sharpness

Now we will show that the constant in (1.4) is optimal for each choice of p and q. By continuity and the estimate $[w^{1-p}]_{A_q} \geq [w^{1-p}]_{A_\infty}$, we may restrict ourselves to the case of finite q. It is convenient to split the reasoning into a few parts.

Step 1. Auxiliary geometrical facts and parameters. Suppose that $1<q<\infty$ and $c>1$ are fixed numbers. Pick $\tilde{c}\in(1,c)$. There are two lines passing through the point $K=(1,\tilde{c}^{1/(q-1)})$ which are tangent to the curve $\mathtt{v}\mathtt{u}^{q-1}=c$; pick the line ℓ which has bigger slope (equivalently: the $\mathtt{v}$-coordinate of the tangency point is bigger than 1). This line intersects the curve $\mathtt{v}\mathtt{u}^{q-1}=1$ at two points: pick the point L with smaller $\mathtt{v}$-coordinate and denote this coordinate by $s(q,\tilde{c})$. Furthermore, the line ℓ intersects the curve $\mathtt{v}\mathtt{u}^{q-1}=\tilde{c}$ at two points: one of them is K, while the second, denoted by M, is of the form $\big(1+\delta,(\tilde{c}/(1+\delta))^{1/(q-1)}\big)$. See Figure 2 below.

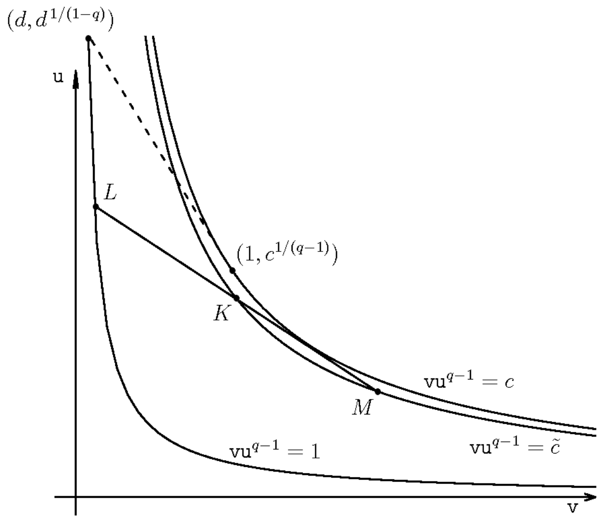

Figure 2. The crucial parameters and their geometric interpretation: $d=d(q,c)$, $K=(1,\tilde{c}^{1/(q-1)})$, $L=\big(s(q,\tilde{c}),(s(q,\tilde{c}))^{1/(1-q)}\big)$ and $M=\big(1+\delta,(\tilde{c}/(1+\delta))^{1/(q-1)}\big)$.

Let us record here two important facts. First, the points K, L, M are colinear: some simple algebra allows to transform this observation into the equality

$$\frac{1-(1+\delta)^{1/(1-q)}}{\delta}=\frac{(\tilde{c}s(q,\tilde{c}))^{1/(1-q)}-1}{1-s(q,\tilde{c})}, \tag{3.1}$$

which will be useful later. Second, it follows immediately from the geometric interpretation of $d(q,c)$ and $s(q,\tilde{c})$ that

$$d(q,c)<s(q,\tilde{c})<1, \tag{3.2}$$

and $s(q,\tilde{c})$ can be made arbitrarily close to $d(q,c)$ by picking $\tilde{c}$ sufficiently close to c.

Finally, we introduce a parameter r, which is assumed to be a positive number less than $1+d(q,c)/(p(1-d(q,c)))$. By the above discussion concerning $d(q,c)$ and

$s(q,\tilde{c})$, we see that if $\tilde{c}$ is sufficiently close to c, then we also have

$$r < 1 + \frac{s(q,\tilde{c})}{p(1-s(q,\tilde{c}))} \leq \frac{1}{1-s(q,\tilde{c})}, \tag{3.3}$$

where the latter inequality is equivalent to $p \geq 1$.

Step 2. Construction. Now, recall the following technical fact, which can be found in [5].

Lemma 3.1. *For every $Q \in \mathcal{T}$ and every $\beta \in (0,1)$ there is a subfamily $F(Q) \subset \mathcal{T}$ consisting of pairwise disjoint subsets of Q such that*

$$\mu\left(\bigcup_{R\in F(Q)} R\right) = \sum_{R\in F(Q)} \mu(R) = \beta\mu(Q).$$

We use this fact inductively, to construct an appropriate family $A_0 \supset A_1 \supset A_2 \supset \ldots$ of sets. Namely, we start with $A_0 = X$. Suppose we have successfully constructed A_n, which is a union of pairwise almost disjoint elements of $\mathcal{T}$, called the *atoms* of A_n (this condition is satisfied for $n = 0$: we have $A_0 = X \in \mathcal{T}$). Then, for each atom Q of A_n, we apply the above lemma with $\beta = (1-s(q,\tilde{c}))/(1-s(q,\tilde{c})+\delta)$ and get a subfamily $F(Q)$. Put $A_{n+1} = \bigcup_Q \bigcup_{Q'\in F(Q)} Q'$, the first union taken over all atoms Q of A_n. Directly from the definition, this set is a union of the family $\{F(Q) : Q \text{ an atom of } A_n\}$, which consists of pairwise disjoint elements of $\mathcal{T}$. We call these elements the atoms of A_{n+1} and conclude the description of the induction step. As an immediate consequence of the above construction, we see that if Q is an atom of A_m, then for any $n \geq m$ we have

$$\mu(Q \cap A_n) = \mu(Q)\left(\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\right)^{n-m}$$

and hence

$$\mu(Q \cap (A_n \setminus A_{n+1})) = \mu(Q)\left(\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\right)^{n-m} \frac{\delta}{1-s(q,\tilde{c})+\delta}. \tag{3.4}$$

Now, introduce the weights σ and w on X by the formulas

$$\sigma = s(q,\tilde{c}) \sum_{n=0}^{\infty} \chi_{A_n\setminus A_{n+1}}(1+\delta)^n$$

and $w = \sigma^{1-p}$. In addition, let $f\colon X \to \mathbf{R}$ be given by $f = \sigma^r$, where r is the number fixed at the previous step.

Step 3. Proof of the inequality $[w^{1/(1-p)}]_{A_q} = [\sigma]_{A_q} \leq c$. First observe that each $Q \in \mathcal{T}$ enjoys exactly one of the following three properties:

(i) the weight σ is constant on Q;
(ii) Q is an atom or the union of some atoms of some A_m;
(iii) there is a nonnegative integer m such that $Q \cap A_m \neq \emptyset$, $Q \setminus A_m \neq \emptyset$.

Indeed: if Q satisfies (ii), then it is divided in the inductive procedure described above and, as a result, some nontrivial part of it goes to A_{m+1}, so σ is not constant on Q. This proves that the conditions (i) and (ii) are disjoint. Now, suppose that Q does not satisfy any of these two conditions and let m be the largest integer such that $Q \subseteq A_{m-1}$. Then $Q \setminus A_m \neq \emptyset$, by the very definition of m, and $Q \cap A_m \neq \emptyset$, since otherwise (i) would hold true (see the formula for σ).

Let us now study the product $(\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu)(\frac{1}{\mu(Q)}\int_Q \sigma^{1/(1-q)}\,\mathrm{d}\mu)^{q-1}$ under each assumption (i), (ii) and (iii) separately. If σ is constant on Q, then the above product is obviously equal to 1. If Q is an atom of A_m, then, by (3.4),

$$\begin{aligned}\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu &= s(q,\tilde c)\sum_{n=m}^{\infty}\left(\frac{1-s(q,\tilde c)}{1-s(q,\tilde c)+\delta}\right)^{n-m}\frac{\delta}{1-s(q,\tilde c)+\delta}\cdot(1+\delta)^n \\ &= (1+\delta)^m\end{aligned} \tag{3.5}$$

and similarly,

$$\begin{aligned}&\frac{1}{\mu(Q)}\int_Q \sigma^{1/(1-q)}\,\mathrm{d}\mu \\ &= s(q,\tilde c)^{1/(1-q)}(1+\delta)^{m/(1-q)}\left((1-s(q,\tilde c))\cdot\frac{1-(1+\delta)^{1/(1-q)}}{\delta}+1\right)^{-1} \\ &= \tilde c^{1/(q-1)}(1+\delta)^{m/(1-q)},\end{aligned} \tag{3.6}$$

where in the last passage we have exploited (3.1). Consequently, we have the equality

$$\left(\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q \sigma^{1/(1-q)}\,\mathrm{d}\mu\right)^{q-1} = \tilde c. \tag{3.7}$$

Finally, assume that Q satisfies (iii). Pick the largest integer m such that $Q\subseteq A_{m-1}$. We have

$$\begin{aligned}\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu &= \frac{1}{\mu(Q)}\int_{Q\setminus A_m}\sigma\,\mathrm{d}\mu + \frac{1}{\mu(Q)}\int_{Q\cap A_m}\sigma\,\mathrm{d}\mu \\ &= \frac{1}{\mu(Q)}\int_{Q\setminus A_m} s(q,\tilde c)(1+\delta)^{m-1}\,\mathrm{d}\mu + \frac{1}{\mu(Q)}\int_{Q\cap A_m}\sigma\,\mathrm{d}\mu.\end{aligned}$$

By (3.5), applied to each atom R of A_m contained in Q, we get

$$\int_{Q\cap A_m}\sigma\mathrm{d}\mu = \mu(Q\cap A_m)(1+\delta)^m$$

and hence, setting $\eta := \mu(Q\cap A_m)/\mu(Q)$, we rewrite the preceding equality in the form

$$\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu = (1-\eta)s(q,\tilde c)(1+\delta)^{m-1} + \eta(1+\delta)^m. \tag{3.8}$$

A similar calculation, exploiting (3.6) instead of (3.5), shows that

$$\frac{1}{\mu(Q)}\int_Q \sigma^{1/(1-q)}\,\mathrm{d}\mu = (1-\eta)s(q,\tilde c)^{1/(1-q)}(1+\delta)^{(m-1)/(1-q)} + \eta\cdot\tilde c^{1/(q-1)}(1+\delta)^{m/(1-q)}.$$

and therefore

$$\begin{aligned}&\left(\frac{1}{\mu(Q)}\int_Q \sigma\,\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q \sigma^{1/(1-q)}\,\mathrm{d}\mu\right)^{q-1} \\ &= \Big((1-\eta)s(q,\tilde c)+\eta(1+\delta)\Big)\Big((1-\eta)s(q,\tilde c)^{1/(1-q)}+\eta\tilde c^{1/(q-1)}(1+\delta)^{1/(1-q)}\Big)^{q-1}.\end{aligned}$$

This number does not exceed c. To see this, rewrite the right-hand side in the form

$$(\eta M_v + (1-\eta)L_v)(\eta M_u + (1-\eta)L_u)^{q-1},$$

where M_v, M_u and L_v, L_u are the coordinates of the points M and L (see Figure 2). As η ranges from 0 to 1, the point $\eta M+(1-\eta)L$ runs over the line segment ML which is entirely contained in $\{(\mathtt{v},\mathtt{u})\colon \mathtt{v}\mathtt{u}^{p-1}\leq c\}$. Thus we have established the desired condition $[\sigma]_{A_q}\leq c$, and combining this with (3.7) yields the two-sided bound

$$\tilde{c}\leq[\sigma]_{A_q}\leq c.$$

Before we proceed, let us record here the information about the A_1 characteristic of σ. Pick any element $\omega\in X$ and let n be the unique integer such that $\omega\in A_n\backslash A_{n+1}$; then $\sigma(\omega)=s(q,\tilde{c})(1+\delta)^n$. Let Q be an arbitrary element of the tree $\mathcal{T}$ which contains ω; this set satisfies one of the conditions (i), (ii), (iii) listed above. If σ is constant on Q, then

$$\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu=\sigma(\omega)=s(q,\tilde{c})(1+\delta)^n.$$

If Q satisfies (ii), then $m\leq n$ and, as proved in (3.5),

$$\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu=(1+\delta)^m\leq(1+\delta)^n.$$

Finally, if Q satisfies (iii), then by (3.8),

$$\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu=(1-\eta)s(q,\tilde{c})(1+\delta)^n+\eta(1+\delta)^{n+1}\leq(1+\delta)^{n+1}.$$

Summarizing, we have $\mathcal{M}_{\mathcal{T}}\sigma(\omega)/\sigma(\omega)\leq(1+\delta)/s(q,\tilde{c})$ and hence

$$[\sigma]_{A_1}\leq\frac{1+\delta}{s(q,\tilde{c})}.$$

Furthermore, if q is sufficiently close to 1 and $\tilde{c}$ is made close enough to c, then the right hand side can be made as close to $d(1,c)^{-1}=c$ as we wish.

Step 4. On the characteristic $[w]_{A_p}=[w,\sigma]_{A_p}$. Using arguments similar to those above, we will show that if δ is sufficiently small, then $[w]_{A_p}$ can be made arbitrarily close to $\alpha=d(q,c)^{1-p}\left((1-d(q,c))\cdot(p-1)+1\right)^{-1}$. We start from the observation that any set $Q\in\mathcal{T}$ satisfies (i), (ii) or (iii) listed in the previous step. If (i) holds, then obviously

$$\left(\frac{1}{\mu(Q)}\int_Q w\,\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu\right)^{p-1}=1.$$

If Q is an atom of A_m, then arguing as in (3.6), we get

$$\frac{1}{\mu(Q)}\int_Q w\,\mathrm{d}\mu=s(q,\tilde{c})^{1-p}(1+\delta)^{m(1-p)}\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^{1-p}}{\delta}+1\right)^{-1},$$

so by (3.5),

$$\left(\frac{1}{\mu(Q)}\int_Q w\,\mathrm{d}\mu\right)\left(\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu\right)^{p-1}=s(q,\tilde{c})^{1-p}\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^{1-p}}{\delta}+1\right)^{-1}.$$

(Note that if $\delta\to 0$ - or rather $\tilde{c}\to c$ - then the expression on the right converges to α). Finally, if Q satisfies (iii), then it "mediates" between the two possibilities above: more precisely, if $Q\subseteq A_{m-1}$ and $Q\not\subseteq A_m$, then the point

$$\left(\frac{1}{\mu(Q)}\int_Q w\,\mathrm{d}\mu,\frac{1}{\mu(Q)}\int_Q\sigma\,\mathrm{d}\mu\right)$$

is contained in the line segment with endpoints

$$\left(s(q,\tilde{c})^{1-p}(1+\delta)^{(m-1)(1-p)},s(q,\tilde{c})(1+\delta)^{m-1}\right)$$

and

$$\left(s(q,\tilde{c})^{1-p}(1+\delta)^{m(1-p)}\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^{1-p}}{\delta}+1\right)^{-1},(1+\delta)^m\right).$$

Thus, it is enough to show that for any $\varepsilon>0$, this line segment is entirely contained in the hyperbolic region $\{(\mathtt{w},\mathtt{v})\in(0,\infty)^2\colon \mathtt{w}\mathtt{v}^{p-1}\le\alpha+\varepsilon\}$, provided δ is sufficiently small. We may assume that $m=0$, since the linear mapping $(\mathtt{w},\mathtt{v})\mapsto(\mathtt{w}(1+\delta)^{m(p-1)},\mathtt{v}/(1+\delta)^m)$ preserves this hyperbolic region. To show the claim, we pass to the limit $\delta\to0$ (or rather $\tilde{c}\to c$). Then the endpoints of the segment become $(d(q,c)^{1-p},d(q,c))$ and $(\alpha,1)$; as one easily verifies, this limiting line segment becomes tangent to the curve $\mathtt{w}\mathtt{v}^{p-1}=\alpha$ (and hence lies below it). Putting all the above observations together, we get the aforementioned claim concerning $[w]_{A_p}$.

Step 5. Optimality of the constant in (1.4). Suppose first that $q>1$. Repeating the calculations from (3.6), we check that if Q is an atom of A_m, then

$$\begin{aligned}\frac{1}{w(Q)}\int_Q f\,\mathrm{d}\mu &= s(q,\tilde{c})^r\sum_{n=m}^{\infty}\left(\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\right)^{n-m}\frac{\delta}{1-s(q,\tilde{c})+\delta}\cdot(1+\delta)^{rn}\\ &= s(q,\tilde{c})^r(1+\delta)^{mr}\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^r}{\delta}+1\right)^{-1}\end{aligned}$$

(the ratio of the above geometric series, equal to

$$\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\cdot(1+\delta)^r=1+\delta\left(-\frac{1}{1-s(q,\tilde{c})}+r\right)+o(\delta),$$

is less than 1, at least for sufficiently small δ: this is due to (3.3)). Consequently, we see that

$$\mathcal{M}_{\mathcal{T}}f\ge s(q,\tilde{c})^r(1+\delta)^{mr}\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^r}{\delta}+1\right)^{-1}$$

on A_m and hence, by the definition of f, we obtain

$$\mathcal{M}_{\mathcal{T}}f\ge\left((1-s(q,\tilde{c}))\cdot\frac{1-(1+\delta)^r}{\delta}+1\right)^{-1}f\ge\left((1-s(q,\tilde{c}))(-r)+1\right)^{-1}f$$

on $A_m\setminus A_{m+1}$. The latter bound does not depend on m, so we can rewrite it uniformly as

$$\mathcal{M}_{\mathcal{T}}f\ge\left((1-s(q,\tilde{c}))(-r)+1\right)^{-1}f\quad\text{ on } X. \tag{3.9}$$

Note that $f\in L^p(w)$. Indeed, we compute that

$$\begin{aligned}\|f\|^p_{L^p(w)}&=\int_X\sigma^{(r-1)p+1}\,\mathrm{d}\mu\\ &=s(q,\tilde{c})^{(r-1)p+1}\sum_{n=0}^{\infty}\left(\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\right)^n\frac{\delta}{1-s(q,\tilde{c})+\delta}\cdot(1+\delta)^{((r-1)p+1)n}\end{aligned}$$

and the ratio of this geometric series is equal to

$$\frac{1-s(q,\tilde{c})}{1-s(q,\tilde{c})+\delta}\,(1+\delta)^{(r-1)p+1}=1+\delta\left[-\frac{1}{1-s(q,\tilde{c})}+(r-1)p+1\right]+o(\delta).$$

Now recall that we take r close to (but smaller than) $1+d(q,c)/(p(1-d(q,c)))$; hence $(r-1)p+1<1/(1-d(q,c))$. If we make $\tilde{c}$ sufficiently close to c, then the expression in the square brackets above becomes negative. This proves $f \in L^p(w)$ and hence, by (3.9), we conclude that

$$\|\mathcal{M}_\mathcal{T}\|_{L^p(w)\to L^p(w)} \geq \Big((1-s(q,\tilde{c}))(-r)+1\Big)^{-1}.$$

Now if we choose r sufficiently close to $1+d(q,c)/(p(1-d(q,c)))$ and then $\tilde{c}$ sufficiently close to c, then the number $((1-s(q,\tilde{c}))(-r)+1)^{-1}$ can be made arbitrarily close to $p/((p-1)d(q,c))$. On the other hand, by the arguments presented in the previous two steps, if $\tilde{c}$ is sufficiently close to c, then $([w]_{A_p}[w^{1/(1-p)}]_{A_p})^{1/p}$ can be made arbitrarily close to $d(q,c)^{1/p-1}((1-d(q,c))\cdot(p-1)+1)^{-1/p}\cdot c^{1/p}$ and hence the ratio $\|\mathcal{M}_\mathcal{T}\|_{L^p(w)\to L^p(w)}([w]_{A_p}[w^{1/(1-p)}]_{A_q})^{-1/p}$ is as close to

$$\frac{p}{p-1}\left(\frac{(1-d(q,c))(p-1)+1}{cd(q,c)}\right)^{1/p} \tag{3.10}$$

as we wish. Now let $c\to\infty$: it follows directly from (2.1) that then $d(q,c)\to 0$ and $cd(q,c)\to(\frac{q}{q-1})^{1-q}$. Consequently, the expression (3.10) converges to the constant in (1.4), and this establishes the desired sharpness. It remains to handle the A_1 case. Let f, w be the function and the weight constructed above (they correspond to given parameters $q>1$, $c>\tilde{c}>1$ and $0<r<1+d(q,c)/(p(1-d(q,c)))$). We know that if r is sufficiently close to $1+d(\tilde{q},c)/(p(1-d(\tilde{q},c)))$ and $\tilde{c}$ is sufficiently close to c, then the ratio $\|\mathcal{M}_\mathcal{T}f\|_{L^p(w)}/\|f\|_{L^p(w)}$ can be made arbitrarily close to $p/((p-1)d(q,c))$. The latter expression tends to $pc/(p-1)$ if we let $q\to 1$. On the other hand, by the arguments in Steps 3 and 4, if we perform the above limiting procedure for r, $\tilde{c}$ and q, we have

$$[w]_{A_p}[w^{1/(1-p)}]_{A_1} = [w]_{A_p}[\sigma]_{A_1} \to \frac{c^p}{p-(p-1)/c}.$$

This implies

$$\frac{\|\mathcal{M}_\mathcal{T}\|_{L^p(w)\to L^p(w)}}{([w]_{A_p}[w^{1/(1-p)}]_{A_1})^{1/p}} \geq \frac{p}{p-1}\left(p-\frac{p-1}{c}\right)^{1/p},$$

and letting $c\to\infty$ we get the desired lower bound.

Acknowledgments. Th author would like to thank an anonymous Referee for the careful reading of the paper and several helpful comments. The research was supported by Narodowe Centrum Nauki (Poland), grant no. DEC-2014/14/E/ST1/00532.

References

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

[2] Burkholder, D. L.: Boundary value problems and sharp inequalities for martingale transforms. - Ann. Probab. 12, 1984, 647–702.

[3] Grafakos, L.: Classical and modern Fourier analysis. - Pearson Education, Inc., New Jersey, 2004.

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

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

[6] Melas, A. D.: Dyadic-like maximal operators on $L\log L$ functions. - J. Funct. Anal. 257, 2009, 1631–1654.

[7] MELAS, A. D.: Sharp general local estimates for dyadic-like maximal operators and related Bellman functions. - Adv. Math. 220, 2009, 367–426.

[8] MELAS, A. D., and E. N. NIKOLIDAKIS: On weak-type inequalities for dyadic maximal functions. - J. Math. Anal. Appl. 367, 2008, 404–410.

[9] MELAS, A. D., and E. N. NIKOLIDAKIS: Dyadic-like maximal operators on integrable functions and Bellman functions related to Kolmogorov's inequality. - Trans. Amer. Math. Soc. 362, 2010, 1571–1597.

[10] MUCKENHOUPT, B.: Weighted norm inequalities for the Hardy maximal function. - Trans. Amer. Math. Soc. 165, 1972, 207–226.

[11] NAZAROV, F. L., and S. R. TREIL: The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. - St. Petersburg Math. J. 8, 1997, 721–824.

[12] NAZAROV, F. L., S. R. TREIL, and A. VOLBERG: The Bellman functions and two-weight inequalities for Haar multipliers. - J. Amer. Math. Soc. 12, 1999, 909–928.

[13] OSĘKOWSKI, A.: Sharp martingale and semimartingale inequalities. - Monografie Matematyczne 72, Birkhäuser, 2012.

[14] OSĘKOWSKI, A.: Sharp $L^{p,\infty} \to L^q$ estimates for the dyadic-like maximal operators. - J. Fourier Anal. Appl. 20, 2014, 911–933.

[15] OSĘKOWSKI, A.: Sharp weak-type estimates for the dyadic-like maximal operators. - Taiwanese J. Math. 19, 2015, 1031–1050.

[16] PETERMICHL, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p-characteristic. - Amer. J. Math. 129:5, 2007, 1355–1375.

[17] SAWYER, E. T.: A characterization of a two-weight norm inequality for maximal operators. - Studia Math. 75, 1982, 1–11.

[18] SLAVIN, L., A. STOKOLOS, and V. VASYUNIN: Monge–Ampère equations and Bellman functions: The dyadic maximal operator. - C. R. Acad. Sci. Paris Ser. I 346, 2008, 585–588.

[19] SLAVIN, L., and V. VASYUNIN: Sharp results in the integral-form John–Nirenberg inequality. - Trans. Amer. Math. Soc. 363, 2011, 4135–4169.

[20] SLAVIN, L., and A. VOLBERG: Bellman function and the H^1-BMO duality. - In: Harmonic analysis, partial differential equations, and related topics, 113–126, Contemp. Math. 428, Amer. Math. Soc., Providence, RI, 2007.

[21] VASYUNIN, V.: The exact constant in the inverse Hölder inequality for Muckenhoupt weights. - Algebra i Analiz 15, 2003, 73–117 (in Russian); English transl. in St. Petersburg Math. J. 15, 2004, 49–79.

[22] VASYUNIN, V., and A. VOLBERG: Monge–Ampére equation and Bellman optimization of Carleson embedding theorems. - In: Linear and complex analysis, Amer. Math. Soc. Transl. Ser. 2 226, Amer. Math. Soc., Providence, RI, 2009, 195–238.

Received 21 January 2019 • Accepted 2 March 2020