A WEIGHTED MAXIMAL WEAK-TYPE INEQUALITY

ADAM OSĘKOWSKI AND MATEUSZ RAPICKI

Abstract. Let w be a dyadic A_p weight ($1 \leq p < \infty$) and let $M^{\mathscr{D}}$ be the dyadic Hardy-Littlewood maximal function on $\mathbb{R}^d$. The paper contains the proof of the estimate

$$w\big(\{x \in \mathbb{R}^d : M^{\mathscr{D}} f(x) > w(x)\}\big) \leq C_p [w]_{A_p} \int_{\mathbb{R}^d} |f| \mathrm{d}x,$$

where the constant C_p does not depend on the dimension d. Furthermore, the linear dependence on $[w]_{A_p}$ is optimal, which is a novel result for $1 < p < \infty$. The estimate is shown to hold in a wider context of probability spaces equipped with an arbitrary tree-like structure. The proof rests on the Bellman function method: we construct an abstract special function satisfying certain size and concavity requirements.

1. Introduction

Let M be the Hardy-Littlewood maximal operator, acting on locally integrable functions on $\mathbb{R}^d$ by the formula

$$Mf(x) = \sup \Big\{ \langle |f| \rangle_Q \chi_Q(x) \Big\}, \tag{1.1}$$

where the supremum is taken over all cubes $Q \subset \mathbb{R}^d$ with sides parallel to the axes, and $\langle f \rangle_Q = \frac{1}{|Q|} \int_Q f \mathrm{d}x$ denotes the average of f over Q. A celebrated result of Fefferman and Stein [2] established in 1971 asserts that if w is an arbitrary weight on $\mathbb{R}^d$, i.e., a nonnegative, locally integrable function, then

$$w\big(\{x \in \mathbb{R}^d : Mf(x) \geq 1\}\big) \leq C_d \int_{\mathbb{R}^d} |f| Mw \mathrm{d}x. \tag{1.2}$$

Here we use the notation $w(E) = \int_E w \mathrm{d}x$ for the measure associated with w and the constant C_d depends only on the dimension d. This immediately yields the corresponding weak-type one-weight bound

$$w\big(\{x \in \mathbb{R}^d : Mf(x) \geq 1\}\big) \leq C_d [w]_{A_1} \int_{\mathbb{R}^d} |f| w \mathrm{d}x, \tag{1.3}$$

under the assumption that the weight w satisfies the so-called A_1 condition

$$[w]_{A_1} = \operatorname*{esssup}_{\mathbb{R}^d} Mw/w < \infty.$$

The estimates (1.2) and (1.3) play an important role in harmonic analysis, in particular, they can be used in the study of vector-valued maximal functions (cf. [2]).

2010 *Mathematics Subject Classification.* Primary: 42B25. Secondary: 46E30.

Key words and phrases. Dyadic, shift, weight, Bellman function, best constant.

Adam Osękowski is supported by National Science Centre Poland grant DEC-2014/14/E/ST1/00532.

Mateusz Rapicki is supported by National Science Centre Poland grant 2018/29/N/ST1/02840.

They have also been extended to the setting in which the maximal function M is replaced by a general Calderón-Zygmund singular integral operator (cf. [3, 4, 5]).

There are dual counterparts of the estimates (1.2) and (1.3), see e.g. the work of Lerner, Ombrosi and Pérez [4, Proposition 2.1]. For a related statement, consult Muckenhoupt and Wheeden [6, Theorem 3]. The strong version is

$$w\left(\{x \in \mathbb{R}^d : Mf(x) \geq Mw(x)\}\right) \leq C_d \int_{\mathbb{R}^d} |f| \mathrm{d}x, \tag{1.4}$$

where w is an arbitrary weight and C_d depends only on the dimension. The weaker inequality concerns A_1 weights and reads

$$w\left(\{x \in \mathbb{R}^d : Mf(x) \geq w(x)\}\right) \leq C_d [w]_{A_1} \int_{\mathbb{R}^d} |f| \mathrm{d}x. \tag{1.5}$$

The primary goal of this paper is to study the version of (1.5) in the less restrictive context of A_p weights. Recall that a weight w satisfies Muckenhoupt's condition A_p (or belongs to the class A_p), if the A_p characteristic of w, given by

$$[w]_{A_p} = \sup_Q \langle w \rangle_Q \langle w^{1/(1-p)} \rangle_Q^{p-1}, \tag{1.6}$$

is finite. Actually, we will study the estimate (1.5) in the dyadic context. Recall that the dyadic maximal function $M^{\mathscr{D}}$ is defined by the same formula (1.1) as for the usual maximal operator, but the supremum is taken over all dyadic cubes Q contained in $\mathbb{R}^d$; similarly, a weight w satisfies the dyadic A_p condition, if its A_p characteristic $[w]_{A_p^d}$, given by (1.6) (with the supremum taken over all dyadic cubes $Q \subset \mathbb{R}^d$), is finite.

We will prove the following statement.

Theorem 1.1. *Let $1 < p < \infty$. Then for any dyadic A_p weight w on $\mathbb{R}^d$ and any locally integrable function $f : \mathbb{R}^d \to \mathbb{R}$, we have*

$$w\big(\{x \in \mathbb{R}^d : M^{\mathscr{D}} f(x) > w(x)\}\big) \leq 2ep[w]_{A_p^d} \int_{\mathbb{R}^d} |f| dx. \tag{1.7}$$

The linear dependence on the A_p characteristic is optimal.

It should be emphasized that the multiplicative constant $2ep$ appearing in (1.7) does not depend on the dimension. Actually, we will prove the above statement in a much more general setting: we will study the estimate in the context of probability measures equipped with a tree-like structure. Here is the precise definition.

Definition 1.2. 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) $\mathcal{T} = \bigcup_{m \geq 0} \mathcal{T}^m$, where $\mathcal{T}^0 = \{X\}$ and $\mathcal{T}^{m+1} = \bigcup_{Q \in \mathcal{T}^m} C(Q)$.
- (iv) We have $\lim_{m \to \infty} \sup_{Q \in \mathcal{T}^m} \mu(Q) = 0$.

All the objects introduced above in the dyadic setting can be generalized to the probabilistic context, simply by replacing $\mathscr{D}$, the family of dyadic cubes in $\mathbb{R}^d$, with $\mathcal{T}$ and $(\mathbb{R}^d, |\cdot|)$ with (X, μ). The associated maximal operator $M^{X,\mathcal{T}}$ is given by

$$M^{X,\mathcal{T}} f = \sup_{Q \in \mathcal{T}} \Big(\langle |f| \rangle_{Q,\mu} \chi_Q \Big),$$

where $\langle f \rangle_{Q,\mu} = (\mu(Q))^{-1} \int_Q f \mathrm{d}\mu$ is the average of f over Q with respect to the measure μ. Furthermore, the A_p characteristic of a weight (i.e., a positive and integrable random variable) w on X is given by

$$[w]_{A_p} = \sup_{Q \in \mathcal{T}} \langle w \rangle_{Q,\mu} \langle w^{1/(1-p)} \rangle_{Q,\mu}^{p-1}.$$

Our main result can be stated as follows.

Theorem 1.3. *Let $1 < p < \infty$ and let (X, μ) be a probability space with a tree structure $\mathcal{T}$. Then for any A_p weight w on X and any integrable random variable f we have the estimate*

$$w\Big(\big\{x \in X : M^{X,\mathcal{T}} f(x) > w(x)\big\}\Big) \leq 2ep[w]_{A_p} \int_X |f| d\mu. \tag{1.8}$$

The linear dependence on the A_p characteristic is optimal for each individual triple $(X, \mathcal{T}, \mu)$.

Let us stress here that we do not impose any regularity condition on $\mathcal{T}$: for any element Q of $\mathcal{T}$ and any child Q' of Q, the ratio $\mu(Q')/\mu(Q)$ need not be bounded away from 0 or 1. It is easy to see that the above result is an extension of Theorem 1.1. Indeed, given a dyadic lattice $\mathscr{D}$, we pick an arbitrary base cube $Q \in \mathscr{D}$ and consider the probability space $(Q, |\cdot|/|Q|)$ equipped with the dyadic tree. Now, any dyadic A_p weight w on $\mathbb{R}^d$, when restricted to Q, becomes the probabilistic weight with the characteristic less or equal to $[w]_{A_p}$ and hence (1.8) holds true. Multiplying both sides by $|Q|$ and letting $|Q| \to \infty$ gives (1.7).

A few words about the proof of the inequality (1.8) are in order. Our approach will make use of a certain novel aspect of the Bellman function method, a powerful tool used widely in harmonic analysis and probability theory. This technique has its origins in the theory of optimal stochastic control, and its connections with other areas of mathematics were firstly observed by Burkholder, who used it to identify the unconditional constants of the Haar system. Soon after the appearance of [1], Burkholder's arguments were extended by a number of mathematicians to investigate numerous estimates for semimartingales: see e.g. [9, 10] for an overview. In the nineties, the seminal paper [7] by Nazarov and Treil (inspired by the preprint version of [8]) pushed the technique towards applications in harmonic analysis; since then, the method has been used in many contexts, including BMO inequalities, weighted estimates and many more. Roughly speaking, the Bellman function method relates the validity of a given estimate to the existence of a certain special function which enjoys appropriate size and concavity conditions.

The following important comment is worth emphasizing. Typically, the Bellman function is quite complicated and its discovery, as well as the verification of the required properties, is quite an elaborate issue. Our approach will enable us to overcome this difficulty: we will obtain an *abstract*, non-explicit formula for the Bellman function corresponding to (1.8). This argument was motivated by a similar phenomenon which occurs in the classical, well-understood context of Haar multipliers on the interval $[0,1]$. We strongly believe that this novel argument is applicable in a number of related results in the area.

The next section is devoted to the explanation of the above idea of obtaining abstract Bellman functions for weak-type estimates from the corresponding objects coming from L^p estimates. Section 3 contains the detailed exposition of the Bellman function method in the context of maximal operators $M^{\mathcal{X},\mathcal{T}}$. In the final part we provide the proof of Theorem 1.3. In particular, the optimality of the linear dependence on $[w]_{A_p}$ in (1.8) will be handled at the very end of the paper, by providing appropriate examples.

2. A motivating example

Let $(h_n)_{n\geq 0}$ be the standard Haar system on $[0,1)$, i.e., the collection of functions given by $h_0 = \chi_{[0,1)}$, $h_1 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$, $h_2 = \chi_{[0,1/4)} - \chi_{[1/4,1/2)}$, $h_3 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$, and so on. Suppose further that $V:\mathbb{R}^2\to\mathbb{R}$ is a fixed function and assume that we want to establish the inequality

$$\int_{[0,1)} V\left(\sum_{k=0}^n a_kh_k, \sum_{k=0}^n \epsilon_k a_k h_k\right)\mathrm{d}x \leq 0 \qquad n=0,\,1,\,2,\,\ldots, \tag{2.1}$$

for any sequence $(a_k)_{k\geq 0}$ of integers and any sequence $(\epsilon_k)_{k\geq 0}$ of signs. For example, the choice $V(x,y)=|y|^p - C_p^p|x|^p$ (where $1<p<\infty$) is related to the unconditionality of the Haar system; the choice $V(x,y)=\lambda\chi_{|y|\geq 1} - C|x|$ leads to weak-type estimates for Haar multipliers.

The key to handle (2.1) is to consider the class of all functions $\mathcal{B}:\mathbb{R}^2\to\mathbb{R}$ which enjoy the following conditions:

1° (Initial condition) $\mathcal{B}(x,\pm x)\leq 0$ for all $x\in\mathbb{R}$;
2° (Majorization) $\mathcal{B}\geq V$ on $\mathbb{R}^2$;
3° (Concavity-type property) $\mathcal{B}$ is concave along any line of slope ± 1.

The existence of a function $\mathcal{B}$ with the above properties implies the validity of (2.1). Indeed, by Jensen's inequality, the concavity 3° gives that for any $n\geq 0$ we have

$$\int_0^1 \mathcal{B}\left(\sum_{k=0}^{n+1} a_kh_k, \sum_{k=0}^{n+1}\epsilon_k a_kh_k\right)\mathrm{d}x \leq \int_0^1 \mathcal{B}\left(\sum_{k=0}^{n} a_kh_k, \sum_{k=0}^{n}\epsilon_k a_kh_k\right)\mathrm{d}x.$$

Combining this with 2° and finally 1°, we get

$$\begin{aligned}\int_0^1 V\left(\sum_{k=0}^{n} a_kh_k, \sum_{k=0}^{n}\epsilon_k a_kh_k\right)\mathrm{d}x &\leq \int_0^1 \mathcal{B}\left(\sum_{k=0}^{n} a_kh_k, \sum_{k=0}^{n}\epsilon_k a_kh_k\right)\mathrm{d}x\\ &\leq \int_0^1 \mathcal{B}(a_0,\epsilon a_0)\mathrm{d}x \leq 0.\end{aligned}$$

The important feature of the approach is that the implication can be reversed: if we know a priori that the estimate (2.1) holds, then the corresponding special function $\mathcal{B}$ exists (one can actually write an abstract formula for it).

For example, consider the L^2 bound

$$\left\|\sum_{k=0}^n \epsilon_k a_kh_k\right\|_{L^2}^2 \leq \left\|\sum_{k=0}^n a_kh_k\right\|_{L^2}^2, \qquad n=0,\,1,\,2,\,\ldots.$$

This follows at once from the orthogonality of the Haar system, but let us apply the above approach. The corresponding function V, i.e., the one which transforms the L^2 bound into (2.1), is given by $V(x,y) = y^2 - x^2$, and it turns out that $\mathcal{B} = V$ is the corresponding special function. Let us see what happens for the weak-type $(1,1)$ estimate

$$\left|\left\{x \in [0,1) : \left|\sum_{k=0}^{n} \epsilon_k a_k h_k(x)\right| \geq 1\right\}\right| \leq C \left\|\sum_{k=0}^{n} a_k h_k\right\|_{L^1},$$

for $n = 0, 1, 2, \ldots$. This inequality is of the form (2.1), with $V(x,y) = \chi_{\{|y|\geq 1\}} - C|x|$, and using the above approach, Burkholder showed the estimate with the optimal constant $C = 2$. The special function $\mathcal{B}$ is slightly more complicated:

$$\mathcal{B}(x,y) = \begin{cases} y^2 - x^2 & \text{if } |x| + |y| \leq 1, \\ 1 - 2|x| & \text{if } |x| + |y| > 1. \end{cases}$$

For some more or less formal arguments which lead to the discovery of this function, see e.g. [9, 10]. For our further considerations concerning the estimate (1.8), let us make here some important observations. We see that $\mathcal{B}$ is built of two components: if (x,y) is close to $(0,0)$, then it coincides with the special function corresponding to the L^2 estimate; for remaining (x,y), it is an affine expression (in $|x|$), which is almost equal to V. One easily checks 1° and 2°; to verify 3°, one rewrites the above formula as

$$\mathcal{B}(x,y) = \begin{cases} \min\left\{y^2 - x^2, 1 - 2|x|\right\} & \text{if } |x| \leq 1, \\ 1 - 2|x| & \text{if } |x| > 1, \end{cases} \tag{2.2}$$

from which it is clear that the concavity holds: both $(x,y) \mapsto y^2 - x^2$ and $(x,y) \mapsto 1 - 2|x|$ are concave along the lines of slope ± 1, and hence so is $\mathcal{B}$, being essentially the minimum of the two.

As we will see in Section 4, the inequality (1.8) can be efficiently studied in a similar manner: it will be handled with a certain Bellman function given as the minimum of special functions associated with L^p estimates and the appropriate affine expressions. More precisely, we will proceed as follows: first we will recall a certain weighted L^p estimate for maximal operators; this will give us the existence of the associated Bellman function $\mathfrak{B}$. Then we will take an appropriate modification of the formula (2.2), with the term $y^2 - x^2$ replaced with $\mathfrak{B}$, to obtain the function for the weak-type estimate.

3. Bellman function method for maximal operators

We return to the context of arbitrary probability space (X,μ) equipped with a tree-like structure $\mathcal{T}$. Let $c \in [1,\infty)$, $p \in (1,\infty)$ be given parameters and let $V : [0,\infty)^3 \to \mathbb{R}$ be a fixed function. Suppose further that we are interested in showing the estimate

$$\int_X V\left(f, M^{X,\mathcal{T}} f, w\right) \mathrm{d}\mu \leq 0 \tag{3.1}$$

for any integrable function $f : X \to [0,\infty)$ and any A_p weight w on X satisfying $[w]_{A_p} \leq c$. Here the probability space (X,μ) and a tree structure $\mathcal{T}$ are also allowed to vary. To handle this problem, consider the four-dimensional domain

$$D = D_{p,c} = \left\{ (x,y,u,v) \in [0,\infty)^4 \,:\, x \leq y,\, 1 \leq uv^{p-1} \leq c \right\}$$

and consider the class of special functions $\mathcal{B} : D \to \mathbb{R}$ which enjoy the following structural properties.

1° (Initial condition) We have

$$\mathcal{B}(x,x,u,v) \leq 0 \qquad \text{if } (x,x,u,v) \in D. \tag{3.2}$$

2° (Majorization) If $0 \leq x \leq y$, then

$$\mathcal{B}(x,y,u,u^{1/(1-p)}) \geq V(x,y,u). \tag{3.3}$$

3° (Concavity-type property) Let $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ be nonnegative numbers summing up to 1 and let (x,y,u,v), (x_1,y_1,u_1,v_1), $\ldots$, (x_m,y_m,u_m,v_m) be elements of D enjoying the following conditions: we have $y_j = \max\{x_j, y\}$ for all $j = 1, 2, \ldots, m$ and

$$x = \sum_{k=1}^m \lambda_k x_k, \qquad u = \sum_{k=1}^m \lambda_k u_k, \qquad v = \sum_{k=1}^m \lambda_k v_k.$$

Then we have

$$\mathcal{B}\left(x,y,u,v\right) \geq \sum_{k=1}^m \lambda_k \mathcal{B}(x_k,y_k,u_k,v_k). \tag{3.4}$$

In what follows, we say that a function f on X is $\mathcal{T}$-simple, if it is measurable with respect to the σ-algebra generated by $\mathcal{T}^N$ for some integer N.

Theorem 3.1. *Let $1 < p < \infty$ be fixed. If there is a function $\mathcal{B}$ satisfying 1°, 2° and 3°, then* (3.1) *holds true for any probability space (X,μ) with a tree $\mathcal{T}$, any $\mathcal{T}$-simple function $f : X \to [0,\infty)$ and any $\mathcal{T}$-simple weight $w \in A_p$ satisfying $[w]_{A_p} \leq c$.*

Proof. Fix (X,μ), $\mathcal{T}$ and any f, w as in the statement. We split the reasoning into three intermediate parts.

Step 1. Auxiliary notation. For any $n \geq 0$, define the functions f_n, g_n, w_n and z_n on X as follows: if $\omega \in X$ and $Q = Q_n(\omega)$ denotes the unique element of $\mathcal{T}^n$ which contains ω, then

$$f_n(\omega) = \langle f \rangle_{Q,\mu}, \quad g_n(\omega) = \max_{k \leq n} f_k(\omega), \quad w_n(\omega) = \langle w \rangle_{Q,\mu}, \quad z_n(\omega) = \langle w^{1/(1-p)} \rangle_{Q,\mu}.$$

It is easy to see that (f_n, g_n, w_n, z_n) takes values in the set D: this is the consequence of the inequality $[w]_{A_p} \leq c$.

Step 2. Monotonicity. Now we will prove that

$$\text{the sequence } \left(\int_X \mathcal{B}(f_n, g_n, w_n, z_n) \mathrm{d}\mu \right)_{n \geq 0} \text{ is nonincreasing.} \tag{3.5}$$

This is a simple combination of the inequality (3.4) and the evolution rules of (f,g,w,z). Namely, fix $n \geq 0$, an element $Q \in \mathcal{T}^n$ and denote the children of Q in $\mathcal{T}^{n+1}$ by Q_1, Q_2, $\ldots$, Q_m. The functions f_n, g_n, w_n and z_n are constant on Q: denote the corresponding values by x, y, u and v. Similarly, f_{n+1}, g_{n+1}, w_{n+1} and z_{n+1} are constant on each Q_j: denote the values by x_j, y_j, u_j and v_j, respectively. Let us check that the conditions listed below (3.4) are satisfied, with $\lambda_j = \mu(Q_j)/\mu(Q)$. The numbers λ_j sum up to 1 and

$$x = \frac{1}{\mu(Q)}\int_Q f\mathrm{d}\mu = \sum_{k=1}^m \frac{\mu(Q_k)}{\mu(Q)}\cdot\frac{1}{\mu(Q_k)}\int_{Q_k} f\mathrm{d}\mu = \sum_{k=1}^m \lambda_k x_k.$$

The identities $u = \sum_{k=1}^m \lambda_k u_k$ and $v = \sum_{k=1}^m \lambda_k v_k$ are verified analogously. Moreover, for each j we obviously have

$$y_j = \max_{k\leq n+1} f_k|_{Q_j} = \max\left\{f_{n+1}|_{Q_j}, \max_{k\leq n} f_k|_{Q_j}\right\} = \max\{x_j, y\}.$$

Consequently, we may apply (3.4), and this estimate is equivalent to

$$\int_Q \mathcal{B}(f_n, g_n, w_n, z_n)\mathrm{d}\mu \geq \int_Q \mathcal{B}(f_{n+1}, g_{n+1}, w_{n+1}, z_{n+1})\mathrm{d}\mu.$$

Summing over all $Q \in \mathcal{T}^n$, we get the desired monotonicity.

Step 3. Completion of the proof. Fix a large integer N such that f, w are $\sigma(\mathcal{T}^N)$-measurable. By the previous step, we get

$$\int_X \mathcal{B}(f_N, g_N, w_N, z_N)\mathrm{d}\mu \leq \int_X \mathcal{B}(f_0, g_0, w_0, z_0)\mathrm{d}\mu.$$

But $f_0 \equiv g_0$, so by (3.2), the right-hand side is nonpositive. Furthermore, we have $f_N = f$, $g_N = M^{X,\mathcal{T}}f$, $w_N = w$ and $z_N = w^{1/(1-p)} = w_N^{1/(1-p)}$, so applying (3.3) to the left-hand side, we get the claim. □

Now we will handle the implication in the reverse direction.

Theorem 3.2. *The reverse to Theorem 3.1 holds true.*

Proof. Introduce the abstract function $\mathcal{B}: D \to \mathbb{R}$ by the formula

$$\mathcal{B}(x,y,u,v) = \sup\left\{\int_X V\Big(f, \max\left\{M^{X,\mathcal{T}}f, y\right\}, w\Big)\mathrm{d}\mu\right\}. \tag{3.6}$$

Here the supremum is taken over all probability spaces X with a tree $\mathcal{T}$, all $\mathcal{T}$-simple functions $f: X \to [0,\infty)$ satisfying $\int_X f\mathrm{d}\mu = x$, all $\mathcal{T}$-simple A_p weights w on X satisfying $[w]_{A_p} \leq c$, $\int_X w\mathrm{d}\mu = u$ and $\int_X w^{1/(1-p)}\mathrm{d}\mu = v$.

We will now verify that $\mathcal{B}$ enjoys the properties 1°, 2° and 3°. The initial condition follows directly from (3.1): indeed, for any X, $\mathcal{T}$, f and w as in the definition of $\mathcal{B}(x,x,u,v)$ we have

$$\int_X V\Big(f, \max\left\{M^{X,\mathcal{T}}f, x\right\}, w\Big)\mathrm{d}\mu = \int_X V\Big(f, M^{X,\mathcal{T}}f, w\Big)\mathrm{d}\mu \leq 0,$$

and the inequality remains valid if we take the supremum. The majorization is also very simple: pick arbitrary X, $\mathcal{T}$ and consider the constant function $f \equiv x$ and the constant weight $w \equiv u$. Then $[w]_{A_p} = 1 \leq c$ and $\int_X w^{1/(1-p)}\mathrm{d}\mu = u^{1/(1-p)}$, so by the very definition of $\mathcal{B}$, we may write

$$\mathcal{B}(x,y,u,u^{1/(1-p)}) \geq \int_X V\Big(f, \max\left\{M^{X,\mathcal{T}}f, x\right\}, w\Big)\mathrm{d}\mu = V(x,y,u).$$

It remains to prove the concavity-type condition 3°. Fix an auxiliary number $\varepsilon > 0$ and pick parameters λ_j and points (x,y,u,v), (x_j,y_j,u_j,v_j) as in the statement of 3°. By the definition of $\mathcal{B}$, there are probability spaces (X_j,μ_j) with a tree $\mathcal{T}_j$, as well as appropriate functions f_j and w_j on X_j such that

$$\mathcal{B}(x_j,y_j,u_j,v_j)\leq\int_{X_j}V\Big(f_j,\max\big\{M^{X_j,\mathcal{T}_j}f_j,y_j\big\},w_j\Big)\mathrm{d}\mu_j+\varepsilon. \tag{3.7}$$

With no loss of generality, we may assume that X_j are pairwise disjoint. We splice them into one space $X=\bigcup_{j=1}^m X_j$ with the probability measure μ given by $\mu(A)=\sum_{j=1}^m\lambda_j\mu_j(A\cap X_j)$ and the tree structure $\mathcal{T}$ such that $\mathcal{T}^0=\{X\}$ and $\mathcal{T}^n=\bigcup_{j=1}^m\mathcal{T}_j^{n-1}$ for $n\geq 1$. Next, we "splice" the functions and weights as follows: $f=\sum_{j=1}^m f_j\chi_{X_j}$ and $w=\sum_{j=1}^m w_j\chi_{X_j}$. Let us check that f and w satisfy all the requirements in the definition of $\mathcal{B}(x,y,u,v)$. First, note that

$$\int_X f\mathrm{d}\mu=\sum_{j=1}^m\int_{X_j}f\mathrm{d}\mu=\sum_{j=1}^m\lambda_j\int_{X_j}f\mathrm{d}\mu_j=\sum_{j=1}^m\lambda_jx_j=x$$

and similarly, $\int_X w\mathrm{d}\mu=u$, $\int_X w^{1/(1-p)}\mathrm{d}\mu=v$, so the averaging conditions are satisfied. Now we will verify that $[w]_{A_p}\leq c$. By the calculations we have just carried out, we see that $\langle w\rangle_{X,\mu}\langle w^{1/(1-p)}\rangle_{X,\mu}=uv^{p-1}\leq c$, where the latter bound follows from the inclusion $(x,y,u,v)\in D$. Next, if $Q\in\mathcal{T}$ is different than X, then Q belongs to $\mathcal{T}_j$ for some j; since $[w_j]_{A_p}\leq c$,

$$\langle w\rangle_{Q,\mu}\langle w^{1/(1-p)}\rangle_{Q,\mu}=\langle w_j\rangle_{Q,\mu}\langle w_j^{1/(1-p)}\rangle_{Q,\mu}\leq c.$$

This establishes the desired Muckenhoupt condition and hence, by the very definition of $\mathcal{B}$,

$$\mathcal{B}(x,y,u,v)\geq\int_X V\Big(f,\max\big\{M^{X,\mathcal{T}}f,y\big\},w\Big)\mathrm{d}\mu.$$

Now, since $x\leq y$, we have $\max\big\{M^{X,\mathcal{T}}f,y\big\}=\max\big\{M^{X_j,\mathcal{T}_j}f_j,y\big\}$ on X_j and hence

$$\begin{aligned}\mathcal{B}(x,y,u,v)&\geq\sum_{j=1}^m\lambda_j\int_{X_j}V\Big(f_j,\max\big\{M^{X_j,\mathcal{T}_j}f_j,y\big\},w_j\Big)\mathrm{d}\mu_j\\&\geq\sum_{j=1}^m\lambda_j\mathcal{B}(x_j,y_j,u_j,v_j)-\varepsilon,\end{aligned}$$

where in the last passage we have exploited (3.7). Since ε was arbitrary, the concavity condition follows. □

4. Proof of Theorem 1.3

Our starting point is the sharp dimension-free weighted L^p estimate for maximal operators established in [11]. Namely, for any $1<p<\infty$ and any probability space (X,μ) with the tree structure $\mathcal{T}$ and any A_p weight w on X, we have

$$\|M^{X,\mathcal{T}}\|_{L^p(w)\to L^p(w)}\leq\frac{p}{p-1-d(p,[w]_{A_p})}.$$

Here, for a given $1<p<\infty$ and $c\geq 1$, the constant $d(p,c)$ is the unique number in $[0,p-1)$ satisfying the equation

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

We will need the more explicit bound

$$\begin{aligned}\|M^{X,\mathcal{T}}\|_{L^p(w)\to L^p(w)} &\leq \frac{p}{p-1-d(p,[w]_{A_p})}\\ &= \frac{p}{p-1}\big(1+d(p,[w]_{A_p})[w]_{A_p}\big)^{1/(p-1)}\\ &\leq \frac{p}{p-1}p^{1/(p-1)}[w]_{A_p}^{1/(p-1)} \leq \frac{pe}{p-1}[w]_{A_p}^{1/(p-1)}.\end{aligned} \tag{4.1}$$

Let $q = p/(p-1)$ be the harmonic conjugate to p and consider the weight w^{1-q} dual to w. It follows directly from the definition of the A_p condition that $[w^{1-q}]_{A_q}^{1/(q-1)} = [w]_{A_p}$, and hence the above theorem implies that

$$\|M^{X,\mathcal{T}}\|_{L^q(w^{1-q})\to L^q(w^{1-q})} \leq \frac{qe}{q-1}[w]_{A_p} = pe[w]_{A_p}.$$

Equivalently, for any A_p weight w with $[w]_{A_p} \leq c$ and any $f \in L^q(w^{1-q})$ we have

$$\int_X V(f, M^{X,\mathcal{T}}f, w)\mathrm{d}\mu \leq 0,$$

for $V(x,y,u) = y^q u^{1-q} - (pecx)^q u^{1-q}$. In particular, the above estimate holds for all $\mathcal{T}$-simple functions f. Therefore, by Theorem 3.2 there exists an associated function $\mathcal{B}$ possessing the properties 1°, 2° and 3°. We will need the following enhanced version of the majorization.

Lemma 4.1. *For all* $(x,y,u,v) \in D$ *we have*

$$\mathcal{B}(x,y,u,v) \geq y^q v - (pecx)^q u^{1-q}. \tag{4.2}$$

Proof. Let us go back to the definition (3.6) of $B(x,y,u,v)$ (with $V(x,y,u) = y^q u^{1-q} - (pecx)^q u^{1-q}$). Take there an arbitrary weight w with the appropriate conditions on characteristic and averages, and put $f = xw/u$. Since $\int_X f\mathrm{d}\mu = x$, we have

$$\begin{aligned}\mathcal{B}(x,y,u,v) &\geq \int_X \Big[\max\{M^{X,\mathcal{T}}f, y\}\Big]^q w^{1-q}\mathrm{d}\mu - (pec)^q \int_X f^q w^{1-q}\mathrm{d}\mu\\ &\geq \int_X y^q w^{1-q}\mathrm{d}\mu - (pecx)^q u^{-q}\int_X w\mathrm{d}\mu\\ &= y^q v - (pecx)^q u^{1-q}.\end{aligned}$$

□

Now we will modify $\mathcal{B}$ to obtain the Bellman corresponding to the weak-type estimate (1.8). Define $\overline{\mathcal{B}} : D \to \mathbb{R}$ by

$$\overline{\mathcal{B}}(x,y,u,v) = \begin{cases}\min\Big\{\mathcal{B}(x,y,u,v), u-2pecx\Big\} & \text{if } pecx < u,\\ u - 2pecx & \text{if } pecx \geq u\end{cases} \tag{4.3}$$

and $\overline{V} : [0;\infty)^3 \to \mathbb{R}$ by $\overline{V}(x,y,u) = u\chi_{\{y\geq u\}} - 2pecx$. Obviously, we have

$$\overline{\mathcal{B}}(x,y,u,v) \leq u - 2pecx \qquad \text{on } D. \tag{4.4}$$

Furthermore, by (4.2), if $pecx = u$, then

$$\mathcal{B}(x,y,u,v) \geq y^q v - pecx\cdot(pecxu^{-1})^{q-1} \geq -pecx = u - 2pecx, \tag{4.5}$$

so we also have

$$\overline{\mathcal{B}}(x,y,u,v)=\begin{cases}\min\Big\{\mathcal{B}(x,y,u,v),u-2pecx\Big\} & \text{if } pecx\leq u,\\ u-2pecx & \text{if } pecx>u\end{cases}$$

(in comparison to the formula (4.3), the inequalities $pecx<u$ and $pecx\geq u$ have become non-strict and strict, respectively). We will need the following additional property of $\overline{\mathcal{B}}$.

Lemma 4.2. *For any point $(x,y,u,v)\in D$ and any $x'>x$ we have*

$$\overline{\mathcal{B}}(x',\max\{x',y\},u,v)\geq\overline{\mathcal{B}}(x,y,u,v)-2pec(x'-x).$$

Proof. We split the reasoning into a few parts.

Step 1. An easy case. If $\overline{\mathcal{B}}(x',\max\{x',y\},u,v)=u-2pecx'$, then the claim follows immediately from (4.4):

$$\overline{\mathcal{B}}(x',\max\{x',y\},u,v)=u-2pecx-2pec(x'-x)\geq\overline{\mathcal{B}}(x,y,u,v)-2pec(x'-x).$$

Hence, from now on, we assume that $\overline{\mathcal{B}}(x',\max\{x',y\},u,v)<u-2pecx'$; this in particular implies that $\mathcal{B}(x',\max\{x',y\},u,v)=\overline{\mathcal{B}}(x',\max\{x',y\},u,v)$ and $pecx'<u$, by the definition of $\overline{\mathcal{B}}$.

Step 2. Monotonicity of $\mathcal{B}$ with respect to y. Fix $(x,y,u,v)\in D$. Observe that if $y'>y$, then

$$\mathcal{B}(x,y,u,v)\leq\mathcal{B}(x,y',u,v), \tag{4.6}$$

which follows from the definition of $\mathcal{B}$. Indeed, if (X,μ), $\mathcal{T}$ is an arbitrary probability space with a tree, and f, w are functions on X as in the definition of $\mathcal{B}(x,y,u,v)$, then

$$\begin{aligned}&\int_X\Big[\max\big\{M^{X,\mathcal{T}}f,y\big\}\Big]^q w^{1-q}\mathrm{d}\mu-(pec)^q\int_X f^q w^{1-q}\mathrm{d}\mu\\ &\leq\int_X\Big[\max\big\{M^{X,\mathcal{T}}f,y'\big\}\Big]^q w^{1-q}\mathrm{d}\mu-(pec)^q\int_X f^q w^{1-q}\mathrm{d}\mu\leq\mathcal{B}(x,y',u,v).\end{aligned}$$

Taking the supremum over all f and w yields (4.6).

Step 3. An additional concavity. We have $pecx'<u$ (see the end of Step 1 above), so x' belongs to the interval $(x,u/(pec))$ and hence there is $\lambda\in(0,1)$ such that $x'=\lambda x+(1-\lambda)u/(pec)$. Therefore, an application of the concavity property of $\mathcal{B}$ yields

$$\begin{aligned}&\overline{\mathcal{B}}(x',\max\{x',y\},u,v)\\ &=\mathcal{B}(x',\max\{x',y\},u,v)\\ &\geq\lambda\mathcal{B}(x,\max\{x',y\},u,v)+(1-\lambda)\mathcal{B}(u/(pec),\max\{u/(pec),y\},u,v).\end{aligned}\tag{4.7}$$

However, by (4.6) and the inequality $pecx<pecx'<u$ we have

$$\mathcal{B}(x,\max\{x',y\},u,v)\geq\mathcal{B}(x,y,u,v)\geq\overline{\mathcal{B}}(x,y,u,v). \tag{4.8}$$

Furthermore, by (4.5) and the definition of $\overline{\mathcal{B}}$, we see that

$$\mathcal{B}(u/(pec),\max\{u/(pec),y\},u,v)\geq\overline{\mathcal{B}}(u/(pec),\max\{u/(pec),y\},u,v),$$

so by Step 1 above,

$$\mathcal{B}(u/(pec),\max\{u/(pec),y\},u,v)\geq\overline{\mathcal{B}}(x',\max\{x',y\},u,v)-2pec\left(\frac{u}{pec}-x'\right).$$

Plugging this and (4.8) into (4.7) yields the claim. □

We are ready for the main ingredient of Theorem 1.3.

Theorem 4.3. *The function $\overline{\mathcal{B}}$ satisfies the conditions 1°, 2° and 3° (with respect to $\overline{V}$).*

Proof. The property 1° is easy to check: by the initial property of $\mathcal{B}$, if $pecx \leq u$, then $\overline{\mathcal{B}}(x,x,u,v) \leq \mathcal{B}(x,x,u,v) \leq 0$; on the other hand, if $pecx \geq u$, then $\overline{\mathcal{B}}(x,x,u,v) = u - 2pecx \leq -pecx \leq 0$.

We proceed to the majorization condition 2°. If $pecx \geq u$, then there is nothing to prove, so from now on we may assume that the reverse estimate holds. Suppose first that $y \geq u$. Then, by the definition of $\overline{\mathcal{B}}$, the majorization is equivalent to $\mathcal{B}(x,y,u,v) \geq u - 2pecx$. However, applying (4.2) (and using the estimate $u^{q-1}v \geq 1$), we get

$$\mathcal{B}(x,y,u,v) \geq y^q v - pecx \cdot \left(pecxu^{-1}\right)^{q-1} \geq u - pecx \geq u - 2pecx.$$

So, it remains to verify 2° for $y < u$; then the desired bound becomes

$$\overline{\mathcal{B}}(x,y,u,v) \geq -2pecx.$$

This is obvious if $\overline{\mathcal{B}}(x,y,u,v) = u - 2pecx$; otherwise, again by (4.2),

$$\begin{aligned}\overline{\mathcal{B}}(x,y,u,v) &= \mathcal{B}(x,y,u,v)\\ &\geq y^q v - (pecx)^q u^{1-q} \geq -pecx \cdot \left(pecxu^{-1}\right)^{q-1} \geq -2pecx. \qquad \square\end{aligned}$$

It remains to establish 3°. If $\overline{\mathcal{B}}(x,y,u,v) = u - 2pecx$, then the condition follows directly from (4.4). So, suppose that $\overline{\mathcal{B}}(x,y,u,v) = \mathcal{B}(x,y,u,v) < u - 2pecx$. In particular this implies $pecx < u$ and hence we have $pecx_j \leq u_j$ for at least one j; relabelling the points if necessary, we may and do assume that there is an integer k such that $pecx_1 \leq u_1$, $pecx_2 \leq u_2$, …, $pecx_k \leq u_k$ and $pecx_{k+1} > u_{k+1}$, $pecx_{k+2} > u_{k+2}$, …, $pecx_m > u_m$. Now we will run a backward induction with respect to k. First, if $k = m$, then the claim follows from the concavity property 3° of $\mathcal{B}$:

$$\overline{\mathcal{B}}(x,y,u,v) = \mathcal{B}(x,y,u,v) \geq \sum_{j=1}^m \lambda_j \mathcal{B}(x_j,y_j,u_j,v_j) \geq \sum_{j=1}^m \lambda_j \overline{\mathcal{B}}(x_j,y_j,u_j,v_j).$$

We proceed to the induction step. Assume that $pecx_1 \leq u_1$, $pecx_2 \leq u_2$, …, $pecx_{k-1} \leq u_{k-1}$ and $pecx_k > u_k$, $pecx_{k+1} > u_{k+1}$, …, $pecx_m > u_m$. The idea is to modify x_j, but keeping their average $\sum_{j=1}^m \lambda_j x_j$ fixed. More specifically, we may increase x_1, x_2, …, x_{k-1} a little bit (so that the estimates $pecx_j \leq u_j$ remain valid) and decrease x_k to make $pecx_k > u_k$ into equality; the points x_{k+1}, x_{k+2}, …, x_m remain unchanged. For notational convenience, denote these new values by x_1', x_2', …, x_m'. Then, by the induction assumption, we have

$$\overline{\mathcal{B}}(x,y,u,v) \geq \sum_{j=1}^m \lambda_j \overline{\mathcal{B}}(x_j', \max\{x_j', y\}, u_j, v_j). \tag{4.9}$$

Now, by the previous lemma and (4.6), for any $j \leq k-1$ we have

$$\begin{aligned}\overline{\mathcal{B}}(x_j', \max\{x_j', y\}, u_j, v_j) &\geq \overline{\mathcal{B}}(x_j, \max\{x_j', y\}, u_j, v_j) - 2pec(x_j' - x_j)\\ &\geq \overline{\mathcal{B}}(x_j, \max\{x_j, y\}, u_j, v_j) - 2pec(x_j' - x_j).\end{aligned}$$

Furthermore, by (4.5),

$$\overline{\mathcal{B}}(x'_k, \max\{x'_k, y\}, u_k, v_k) \geq u_k - 2pecx'_k = u_k - 2pecx_k - 2pec(x_k - x'_k).$$

Plugging the last two estimates into (4.9), we complete the proof of the induction step: we obtain

$$\overline{\mathcal{B}}(x, y, u, v) \geq \sum_{j=1}^m \lambda_j \overline{\mathcal{B}}(x_j, y_j, u_j, v_j).$$

Thus, $\overline{\mathcal{B}}$ has the desired concavity property.

The properties of $\overline{\mathcal{B}}$ immediately yield our main weighted estimate.

Proof of (1.8). Fix $1 < p < \infty$. Let (X, μ) be a probability space with a tree structure $\mathcal{T}$ and assume that $f : X \to [0, \infty)$ is an arbitrary integrable function, w is an A_p weight on X. Set $c := [w]_{A_p}$. Let us approximate f and w by simple functions: given a large positive integer N, we let f_N, w_N be the conditional expectations of f and w with respect to $\mathcal{T}^N$ (see the proof of Theorem 3.1, Step 1). Then f_N and w_N are $\mathcal{T}$-simple and $[w_N]_{A_p} \leq [w]_{A_p}$. To see the latter estimate, simply note that

$$\langle w_N \rangle_{Q,\mu} \langle w_N^{1/(1-p)} \rangle_{Q,\mu}^{p-1} = 1$$

for $Q \in \mathcal{T}^N \cup \mathcal{T}^{N+1} \cup \mathcal{T}^{N+2} \cup \ldots$, while for remaining $Q \in \mathcal{T}$ we have

$$\langle w_N \rangle_{Q,\mu} \langle w_N^{1/(1-p)} \rangle_{Q,\mu}^{p-1} \leq \langle w \rangle_{Q,\mu} \langle w^{1/(1-p)} \rangle_{Q,\mu}^{p-1}$$

by Jensen's inequality. Therefore, by Theorem 3.1 applied to $\overline{\mathcal{B}}$ and $\overline{V}$, we get

$$\int_X w_N (M^{X,\mathcal{T}} f_N \geq w_N) \mathrm{d}\mu \leq 2pe[w]_{A_p} \int_X f_N \mathrm{d}\mu = 2pe[w]_{A_p} \int_X f \mathrm{d}\mu.$$

However, $\int_X w_N (M^{X,\mathcal{T}} f_N \geq w_N) \mathrm{d}\mu = \int_X w (M^{X,\mathcal{T}} f_N \geq w_N) \mathrm{d}\mu$ and $M^{X,\mathcal{T}} f_N \uparrow M^{X,\mathcal{T}} f$, $w_N \to w$ μ-almost surely as $N \to \infty$. Therefore, the previous estimate yields

$$\int_X w (M^{X,\mathcal{T}} f > w) \mathrm{d}\mu \leq 2pe[w]_{A_p} \int_X f \mathrm{d}\mu.$$

To obtain the non-strict inequality on the left, consider an auxiliary parameter $\theta \in (0, 1)$ and apply the above bound to the A_p weight θw:

$$\int_X w (M^{X,\mathcal{T}} f \geq w) \mathrm{d}\mu \leq \int_X w (M^{X,\mathcal{T}} f > \theta w) \mathrm{d}\mu \leq 2pe\theta^{-1} [\theta w]_{A_p} \int_X f \mathrm{d}\mu.$$

Since $[\theta w]_{A_p} = [w]_{A_p}$, letting $\theta \to 1$ completes the proof. □

It remains to show that the linear dependence on the A_p characteristic in (1.8) is optimal. Fix $1 < p < \infty$ and pick an arbitrary probability space (X, μ) with a tree $\mathcal{T}$. Let $Q_0 = X$. By a simple induction and the property (ii) in the definition of a tree, for each $n \geq 1$ there exists $Q_n \in \mathcal{T}^n$ such that $\mu(Q_n) \leq \mu(Q_{n-1})/2$. Fix a huge positive integer N and take $f = \chi_{Q_N}/\mu(Q_N)$. Then obviously $\int_X f \mathrm{d}\mu = 1$, and the maximal function of f is given by

$$M^{X,\mathcal{T}} f = \frac{\chi_{Q_N}}{\mu(Q_N)} + \sum_{n=1}^N \frac{\chi_{Q_{n-1} \setminus Q_n}}{\mu(Q_{n-1})}. \tag{4.10}$$

Indeed, we verify the latter identity by the very definition of the maximal operator. Given $x \in X$, we have two possibilities: either $x \in Q_N$, or $x \in Q_{n-1} \setminus Q_n$ for some

$n = 1, 2, \ldots, N$. In the first case, the maximal average of f is that over Q_N (and it equals $1/\mu(Q_N)$). If $x \in Q_{n-1} \setminus Q_n$, then the maximal average corresponds to the choice Q_{n-1} and is equal to $1/\mu(Q_{n-1})$.

Let $w = Mf$. Then the above considerations yield

$$w(Mf \geq w) = w(X) = 1 + \sum_{n=1}^{N} \frac{\mu(Q_{n-1} \setminus Q_n)}{\mu(Q_{n-1})} \geq 1 + \frac{N}{2},$$

where the latter estimate follows from the estimate $\mu(Q_n) \leq \mu(Q_{n-1})/2$ we assumed at the beginning. It remains to analyze the A_p characteristic of w. Let $Q \in \mathcal{T}$. If $Q \subseteq Q_N$ or $Q \subseteq Q_{n-1} \setminus Q_n$ for some n, then w is constant on Q and hence $\langle w \rangle_{Q,\mu} \langle w^{1/(1-p)} \rangle_Q^{p-1} = 1$. If Q does not satisfy any of the two above conditions, then $Q = Q_k$ for some $k = 0, 1, 2, \ldots, N-1$. Then by (4.10) we have

$$\langle w \rangle_{Q,\mu} = \frac{1}{\mu(Q_k)} \sum_{n=k+1}^{N} \frac{\mu(Q_{n-1} \setminus Q_n)}{\mu(Q_{n-1})} \leq \frac{N-k}{\mu(Q_k)} \leq \frac{N}{\mu(Q_k)}$$

and

$$\begin{aligned} \langle w^{1/(1-p)} \rangle_{Q,\mu} &= \frac{1}{\mu(Q_k)} \sum_{n=k+1}^{N} \mu(Q_{n-1})^{1/(p-1)} \mu(Q_{n-1} \setminus Q_n) \\ &\leq \frac{1}{\mu(Q_k)} \sum_{n=k+1}^{N} \left[\mu(Q_{n-1})^{p/(p-1)} - \mu(Q_n)^{p/(p-1)} \right] \leq \mu(Q_k)^{1/(p-1)}. \end{aligned}$$

Therefore $\langle w \rangle_{Q,\mu} \langle w^{1/(1-p)} \rangle_Q^{p-1} \leq N$ and hence $[w]_{A_p} \leq N$. Putting all the above facts together, we see that the inequality

$$w(Mf \geq w) \leq C_p [w]_{A_p}^{\kappa} \int_X f \, \mathrm{d}\mu$$

cannot hold with any exponent $\kappa < 1$.

Acknowledgments

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

References

[1] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. 12 (1984), 647–702.

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

[3] A. K. Lerner, S. Ombrosi and C. Pérez, *Sharp A_1 bounds for Calderón-Zygmund operators and the relationship with a problem of Muckenhoupt and Wheeden*, Int. Math. Res. Not. IMRN 6 (2008), Art. ID rnm161, 11 p.

[4] A. K. Lerner, S. Ombrosi and C. Pérez, *Weak type estimates for singular integrals related to a dual problem of Muckenhoupt-Wheeden*, J. Fourier Anal. Appl. **15** (2009), pp. 394–403.

[5] A. K. Lerner, S. Ombrosi, C. Pérez, *A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden*, Math. Res. Lett. **16** (2009), pp. 149–156.

[6] B. Muckenhoupt and R. Wheeden, *Some Weighted Weak-Type Inequalities for the Hardy-Littlewood Maximal Function and the Hilbert Transform*, Indiana Univ. Math. J. **26**, No. 5 (1977), 801–816.

[7] F.L. Nazarov 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.

[8] 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.

[9] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72** (2012), Birkhäuser, 462 pp.

[10] A. Osękowski, *Survey article: Bellman function method and sharp inequalities for martingales*, Rocky Mountain J. Math. **43** (2013), 1759–1823.

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

Email address: `A.Osekowski@mimuw.edu.pl`

Email address: `M.Rapicki@mimuw.edu.pl`

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland