Publ. Mat. **69** (2025), 319–340
DOI: 10.5565/PUBLMAT6922503

WEIGHTED STABILITY ESTIMATES FOR MAXIMAL OPERATORS

Łukasz Kamiński and Adam Osękowski

Abstract: We study sharp stability estimates associated with the boundedness of maximal operators on weighted L^p spaces.

2020 Mathematics Subject Classification: Primary: 42B25. Secondary: 46E30, 60G42.

Key words: maximal, dyadic, Bellman function, best constants.

1. Introduction

Sharp estimates arise in many contexts in probability and analysis, and the information about the best constants involved often provides an additional insight into the structure of the problem under investigation. Given a sharp bound, it is then natural to ask about the extremizers, i.e., those quantities for which the estimate becomes an equality. Having identified these, the next step is to consider the following stability question. Assume that the bound is *almost* an equality; how far is the inserted quantity from being an extremizer? Such questions appear in many places in analysis; see e.g. [**4**], [**7**], [**8**], [**9**], [**10**], and [**11**].

The purpose of this paper is to study such a stability problem in the context of weighted L^p bounds for maximal operators. We start with the necessary background and notation. The dyadic maximal operator $\mathcal{M}$ on the unit cube $[0,1]^m$ is an operator acting on integrable functions $\varphi\colon [0,1]^m \to \mathbb{R}$ by the formula

$$\mathcal{M}\varphi(x) = \sup\left\{ \frac{1}{|Q|}\int_Q |\varphi| : x\in Q,\, Q\subset [0,1]^m \text{ is a dyadic cube}\right\},$$

where the integration is with respect to the Lebesgue measure. This operator, as well as its global version on $\mathbb{R}^m$, is of fundamental importance for analysis, probability theory, and PDEs, and the questions about its boundedness on various function spaces have been studied extensively in the literature. Our motivation stems from the classical sharp L^p inequality

$$\|\mathcal{M}\varphi\|_{L^p([0,1]^m)} \leq \frac{p}{p-1}\|\varphi\|_{L^p([0,1]^m)}, \quad 1<p\leq\infty. \tag{1.1}$$

What are the extremizers of this estimate? Only trivial: it turns out that the inequality is strict, unless $\|\varphi\|_{L^p([0,1]^m)}=0$ or $\|\varphi\|_{L^p([0,1]^m)}=\infty$. However, there is an intriguing structural property of the functions for which both sides are almost equal. Namely, for

---

The results were obtained during the realization of the tutoring program, supported by Polish Ministry of Science and Higher Education project MNiSW/2019/394/DIR/KH. The second-named author was supported by University of Warsaw, Inicjatywa Doskonałości - Uczelnia Badawcza (IDUB), Nowe Idee 3B, grant no. 501-D110-20-3004310.

any $\varepsilon>0$ there is $\varphi^{(\varepsilon)}\in L^p$ for which the pointwise identity $\mathcal{M}\varphi^{(\varepsilon)}=\big(\frac{p}{p-1}-\varepsilon\big)\varphi^{(\varepsilon)}$ holds true. Therefore, the family $(\varphi^{(\varepsilon)})_{\varepsilon>0}$ can be regarded as an approximate eigenfunction of $\mathcal{M}$ associated with the eigenvalue $p/(p-1)$. Treating this family as the extremal for (1.1), Melas ([**15**]) established the following related stability result. If $2\le p<\infty$ is a fixed exponent, $\varepsilon>0$ is a small number, and φ is *any* nonnegative function satisfying

$$\|\mathcal{M}\varphi\|_{L^p([0,1]^m)}\ge\left(\frac{p}{p-1}-\varepsilon\right)\|\varphi\|_{L^p([0,1]^m)},$$

then

$$\left\|\mathcal{M}\varphi-\frac{p}{p-1}\varphi\right\|_{L^p([0,1]^m)}\le c_p\,\varepsilon^{1/p}\|\varphi\|_{L^p([0,1]^m)}$$

for some constant c_p depending only on p, and the exponent $1/p$ is the best possible. In other words, if φ is almost extremal for the L^p-estimate, then it is close, in the L^p-sense, to being an eigenfunction of $\mathcal{M}$ corresponding to the eigenvalue $p/(p-1)$. Actually, Melas proved the above statement in the wider context: he considered nonatomic probability spaces equipped with atomic filtrations (see below).

There are several natural extensions of the above statement. For example, one can study the version of Melas' result in the range $1<p<2$, investigate the stability of other types of estimates, or consider wider classes of maximal operators (e.g., associated with a given martingale structure). In what follows, we will inspect the problem of stability of *weighted* analogues of (1.1). We need more definitions. Suppose that w is a weight on $[0,1]^m$, i.e., a nonnegative, integrable function on the unit cube, and let $1<p<\infty$ be a fixed exponent. A classical result of Muckenhoupt ([**16**]) asserts that $\mathcal{M}$ is bounded as an operator on the weighted L^p space

$$L^p(w)=\left\{\varphi\colon[0,1]^m\to\mathbb{R}:\|\varphi\|_{L^p(w)}=\left(\int_{[0,1]^m}|\varphi|^pw\,\mathrm{d}x\right)^{1/p}<\infty\right\}$$

if and only if w belongs to the dyadic A_p class. The latter means that the quantity

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

where the supremum is taken over all dyadic cubes in $[0,1]^m$, is finite. In [**23**], the second-named author established the sharp quantitative bound for the norm $\|\mathcal{M}\|_{L^p(w)\to L^p(w)}$. To describe this result, we need to introduce an auxiliary parameter. For given $c\ge1$ and $1<p<\infty$, let $d=d(p,c)\in[1,p)$ be the unique solution of the equation

$$cd(p-d)^{p-1}=(p-1)^{p-1}. \tag{1.2}$$

Then we have the estimate

$$\|\mathcal{M}\|_{L^p(w)\to L^p(w)}\le\frac{p}{p-d(p,[w]_{A_p})}, \tag{1.3}$$

which is sharp in the following sense: for any $1<p<\infty$, any $c\ge1$, and any $\varepsilon>0$ there is a weight w satisfying $[w]_{A_p}\le c$ and $\|\mathcal{M}\|_{L^p(w)\to L^p(w)}>p/(p-d(p,[w]_{A_p}))-\varepsilon$. This immediately gives rise to the question about the corresponding stability result. Here is one of the main results of this paper.

Theorem 1.1. *Let $1 < p < \infty$, $c \geq 1$, and $\varepsilon > 0$ be fixed. Assume further that a weight w and a nonnegative function φ on $[0,1]^d$ satisfy $[w]_{A_p} \leq c$, $\varphi \in L^p(w)$, and*

$$\|\mathcal{M}\varphi\|_{L^p(w)} > \left(\frac{p}{p-d(p,c)} - \varepsilon\right) \|\varphi\|_{L^p(w)}.$$

Then there is a finite constant c_p depending only on p such that

$$\left\|\mathcal{M}\varphi - \frac{p}{p-d(p,c)}\varphi\right\|_{L^p(w)} \leq c_p\, \varepsilon^{1/\max\{p,2\}} \|\varphi\|_{L^p(w)} \tag{1.4}$$

and the exponent $1/\max\{p,2\}$ is the best possible.

The interpretation of this statement is the same as in the case of Melas' result: if a nonnegative φ is almost extremal for the weighted L^p bound, then it is close, in the L^p-sense, to being the eigenfunction of $\mathcal{M}$. It should also be emphasized that the choice $c = 1$ corresponds to the unweighted setting: thus, in particular the above theorem contains Melas' result, as well as its extension to the range $1 < p < 2$.

Actually, we will study the above statement in the much more general context of arbitrary probability spaces and associated maximal operators, which also appears as the setup for **[15]**. Suppose that $(\Omega, \mathcal{F}, \mu)$ is an arbitrary probability space, equipped with the *tree* (or *dyadic-like*) structure $\mathcal{T}$. That is, we have $\mathcal{T} = \bigcup_{n\in\mathbb{Z}} \mathcal{T}^{(n)}$, where $(\mathcal{T}^{(n)})_{n\in\mathbb{Z}}$ is an increasing family of finite partitions of Ω into $\mathcal{F}$-measurable sets of positive measure, with $\mathcal{T}^{(0)} = \{\Omega\}$. We will also assume that the mesh $\max_{Q\in\mathcal{T}^{(n)}} \mu(Q)$ of the partition $\mathcal{T}^{(n)}$ converges to 0 as $n \to \infty$: this will allow us to apply Lebesgue's differentiation theorem (or rather martingale convergence theorem) in our considerations.

Clearly, the concept of a tree generalizes the notion of a dyadic lattice, which is obtained by taking $\Omega = [0,1]^m$ and setting $\mathcal{T}^{(n)}$ to be the class of all dyadic subcubes of Ω of measure 2^{-nm}. Any tree structure gives rise to the corresponding maximal operator $\mathcal{M}_{\mathcal{T}}$, acting on integrable functions $\varphi\colon \Omega \to \mathbb{R}$ by

$$\mathcal{M}_{\mathcal{T}}\varphi(x) = \sup \frac{1}{\mu(Q)} \int_Q |\varphi| \,\mathrm{d}\mu, \quad x \in \Omega,$$

where the supremum is taken over all elements Q of $\mathcal{T}$ containing x. Then the word "weight" refers to a positive and integrable random variable; the associated weighted L^p space is introduced as in the previous setting. It is easy to extend the notion of A_p weights to this new context: a positive, integrable variable w on Ω satisfies Muckenhoupt's condition A_p (where $1 < p < \infty$ is a given exponent), if

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

where the supremum is taken over all $Q \in \mathcal{T}$. In **[23]** it was shown that

$$\|\mathcal{M}_{\mathcal{T}}\|_{L^p(w)\to L^p(w)} \leq \frac{p}{p - d(p,[w]_{A_{p,\mathcal{T}}})}, \tag{1.5}$$

and the estimate is sharp, in the same sense as in the case of (1.3). We may ask about the stability of this bound. Here is the generalization of Theorem 1.1.

Theorem 1.2. *Let $1 < p < \infty$, $c \geq 1$, and $\varepsilon > 0$ be fixed. Assume that a nonnegative random variable $\varphi \in L^p$ and a weight $w \in A_{p,\mathcal{T}}$ satisfy $[w]_{A_{p,\mathcal{T}}} = c$ and*

$$\|\mathcal{M}_{\mathcal{T}}\varphi\|_{L^p(w)} > \left(\frac{p}{p-d(p,c)} - \varepsilon\right)\|\varphi\|_{L^p(w)}.$$

Then there is a finite constant c_p depending only on p such that

$$\left\|\mathcal{M}_{\mathcal{T}}\varphi - \frac{p}{p-d(p,c)}\varphi\right\|_{L^p(w)} \leq c_p\,\varepsilon^{1/\max\{p,2\}}\|\varphi\|_{L^p(w)} \tag{1.6}$$

and the exponent $1/\max\{p,2\}$ is the best possible.

A few words about our approach and the organization of the paper are in order. The proof of the estimate (1.6) will make use of the so-called Bellman function method, a powerful technique used widely in probability and analysis. This approach originates in the theory of stochastic optimal control (cf. [**3**]), and its deep connection with other areas of mathematics was first observed by Burkholder ([**5**]) in the 1980s, during his study of sharp inequalities for martingale transforms, the Haar system, and monotone bases. This direction was then further explored by Burkholder and his PhD students in the more general semimartingale context; see [**19**] for more on the subject. The decisive step towards applications of the Bellman function method in harmonic analysis was made by Nazarov, Treil, and Volberg: see the seminal paper [**17**], which was inspired by the preprint version of [**18**]. This analytic direction has turned out to be very fruitful: the approach has been successfully extended to cover numerous contexts, including BMO estimates, weighted theory, and the theory of fractional operators. The literature on the subject is extremely large; see e.g. [**1, 2, 13, 14, 21, 24, 25, 26, 27**] and references therein.

The Bellman function method, both in the probabilistic and analytic setups, relates the validity of a given estimate to the existence of a certain special function, which possesses appropriate majorization and concavity-type properties. Actually, this special object often carries much more information concerning the problem: for instance, it enables the proof of a wider class of estimates and encodes the structure of the extremizers. It should be emphasized that in general, the problem of finding the Bellman function corresponding to a given sharp inequality might be a very difficult task. The construction often boils down to the intricate analysis of certain classes of partial differential inequalities, but in many cases one simply builds the function by guessing and experimenting with different algebraic expressions, some of which are taken from the large Bellman literature. Sometimes, when one is interested in a nonsharp version of the inequality, i.e., one accepts a suboptimal constant, this analysis can be significantly simpler, but it is not always the case.

The version of the Bellman function method which allows the study of stability estimates for maximal operators is presented in Section 2. In Section 3, we apply the method: we exhibit the special function and verify that it enjoys all the relevant properties. In Section 4, we present the martingale version of Theorem 1.2. The sharpness of the exponent $\min\{1/2, 1/p\}$ will be obtained in Section 5, by providing appropriate extremal examples. We conclude the paper with the description of some steps which led us to the discovery of the Bellman function.

It should be emphasized that the proof of (1.5), presented in [**23**], rests on the change-of-measure argument and the sharp version of Sawyer's testing conditions. The reasoning presented below gives a completely new, direct proof of the weighted L^p estimate. This contribution is of independent interest and connections.

2. On the method

Throughout this section, we assume that $1 < p < \infty$ and $c \geq 1$ are fixed parameters. We start by introducing the four-dimensional domain

$$\mathcal{D} = \{(x, y, u, v) \in [0, \infty)^2 \times (0, \infty)^2 : x \leq y,\ 1 \leq uv^{p-1} \leq c\}. \tag{2.1}$$

Next, suppose that $B\colon \mathcal{D} \to \mathbb{R}$ is a function satisfying the following requirements.

(a) *Initial condition.* We have $B(x, x, u, v) \leq 0$ for all $(x, x, u, v) \in \mathcal{D}$.

(b) *Majorization.* There are positive constants β, C such that

$$B(x, y, u, v) \geq y^p u \left[1 - \left(\frac{Cx}{y}\right)^p + \beta \left|1 - \frac{Cx}{y}\right|^{\max\{p,2\}}\right]$$

for all $(x, y, u, v) \in \mathcal{D}$ with $y > 0$.

(c) *Concavity.* If (x, y, u, v) and $(x + h, \max\{x + h, y\}, u + k, v + \ell)$ belong to $\mathcal{D}$, then we have

$$\begin{aligned} B(x + h, \max\{x + h, y\}, u + k, v + \ell) \leq{}& B(x, y, u, v) + B_x(x, y, u, v)h \\ &+ B_u(x, y, u, v)k + B_v(x, y, u, v)\ell. \end{aligned} \tag{2.2}$$

Here is a key statement, which links the existence of a function B as above with the stability estimates (1.4) and (1.6).

Theorem 2.1. *Suppose that B satisfies (a), (b), and (c). Fix an arbitrary weight $w \in A_{p,\mathcal{T}}$ with $[w]_{A_{p,\mathcal{T}}} \leq c$ and any nonnegative function $\varphi \in L^p(w)$ such that*

$$\|\mathcal{M}_\mathcal{T}\varphi\|_{L^p(w)} > (C - \varepsilon)\|\varphi\|_{L^p(w)}$$

for some $\varepsilon > 0$. Then we have

$$\|\mathcal{M}_\mathcal{T}\varphi - C\varphi\|_{L^p(w)} \leq C \left(\frac{\varepsilon p}{C\beta}\right)^{1/\max\{p,2\}} \|\varphi\|_{L^p(w)} \tag{2.3}$$

(where β comes from (b)).

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

Step 1. Some notation. For any n and any $\omega \in \Omega$, let $Q^n(\omega)$ stand for the unique element of $\mathcal{T}^{(n)}$ which contains ω. Furthermore, we introduce the functional sequences

$$\varphi_n(\omega) = \frac{1}{\mu(Q^n(\omega))} \int_{Q^n(\omega)} \varphi \,\mathrm{d}\mu, \quad \psi_n(\omega) = \max_{0 \leq k \leq n} \varphi_n(\omega),$$

and

$$w_n(\omega) = \frac{1}{\mu(Q^n(\omega))} \int_{Q^n(\omega)} w \,\mathrm{d}\mu, \quad v_n(\omega) = \frac{1}{\mu(Q^n(\omega))} \int_{Q^n(\omega)} w^{-1/(p-1)} \,\mathrm{d}\mu.$$

That is, φ_n, w_n, v_n are the conditional expectations of φ, w, and $w^{-1/(p-1)}$ with respect to $\mathcal{T}^{(n)}$, and ψ_n is the maximal function of φ, truncated to the subtree $\mathcal{T}^{(0)} \cup \mathcal{T}^{(1)} \cup \mathcal{T}^{(2)} \cup \cdots \cup \mathcal{T}^{(n)}$. Therefore, for any n and any $Q \in \mathcal{T}^{(n)}$ we have

$$\begin{aligned} \frac{1}{\mu(Q)} \int_Q \varphi_{n+1} \,\mathrm{d}\mu &= \varphi_n|_Q, \\ \frac{1}{\mu(Q)} \int_Q w_{n+1} \,\mathrm{d}\mu &= w_n|_Q, \\ \frac{1}{\mu(Q)} \int_Q v_{n+1} \,\mathrm{d}\mu &= v_n|_Q, \end{aligned} \tag{2.4}$$

and we have the monotone almost sure convergence $\psi_n \uparrow \mathcal{M}_\mathcal{T}\varphi$.

Step 2. Monotonicity property. The main part of the proof is to show that the sequence $\left(\int_\Omega B(\varphi_n,\psi_n,w_n,v_n)\,\mathrm{d}\mu\right)_{n\geq 0}$ is nondecreasing. To this end, we will apply the concavity condition (c). Fix n, an element Q of $\mathcal{T}^{(n)}$, and suppose that $Q_1,Q_2,\ldots,Q_m\in\mathcal{T}^{(n+1)}$ is the collection of all children of Q. The functions φ_n, ψ_n, w_n, and v_n are constant on Q; denote the corresponding values by x, y, u, and v. Similarly, for a fixed child Q_j, the functions φ_{n+1}, ψ_{n+1}, w_{n+1}, and v_{n+1} are constant on Q_j, and their values can be denoted by $x+h$, $\max\{x+h,y\}$, $u+k$, $v+\ell$. Thus (2.2) yields

$$\begin{aligned}\int_{Q_j} B(\varphi_{n+1},\psi_{n+1},w_{n+1},v_{n+1})\,\mathrm{d}\mu &\leq B(\varphi_n,\psi_n,w_n,v_n)\mu(Q_j)\\ &\quad+B_x(\varphi_n,\psi_n,w_n,v_n)\int_{Q_j}(\varphi_{n+1}-\varphi_n)\,\mathrm{d}\mu\\ &\quad+B_u(\varphi_n,\psi_n,w_n,v_n)\int_{Q_j}(w_{n+1}-w_n)\,\mathrm{d}\mu\\ &\quad+B_v(\varphi_n,\psi_n,w_n,v_n)\int_{Q_j}(v_{n+1}-v_n)\,\mathrm{d}\mu.\end{aligned}$$

Summing over all Q_j and applying (2.4), we obtain

$$\begin{aligned}\int_Q B(\varphi_{n+1},\psi_{n+1},w_{n+1},v_{n+1})\,\mathrm{d}\mu &\leq B(\varphi_n,\psi_n,w_n,v_n)\mu(Q)\\ &=\int_Q B(\varphi_n,\psi_n,w_n,v_n)\,\mathrm{d}\mu.\end{aligned}$$

It remains to sum over all $Q\in\mathcal{T}^{(n)}$ to get the desired monotonicity. Thus in particular for any n we have

$$\int_\Omega B(\varphi_n,\psi_n,w_n,v_n)\,\mathrm{d}\mu\leq\int_\Omega B(\varphi_0,\psi_0,w_0,v_0)\,\mathrm{d}\mu\leq 0,$$

where the last bound follows from $\varphi_0=\psi_0$ (by the very definition of ψ_0) and the initial condition (a).

Step 3. Proof of (2.3). Suppose first that $p>2$. Combining (b) with the last inequality of the previous step, we see that

$$\int_\Omega \psi_n^p w_n\,\mathrm{d}\mu+\beta\int_\Omega|\psi_n-C\varphi_n|^p w_n\,\mathrm{d}\mu\leq C^p\int_\Omega\varphi_n^p w_n\,\mathrm{d}\mu,$$

or

$$\int_\Omega \psi_n^p w\,\mathrm{d}\mu+\beta\int_\Omega|\psi_n-C\varphi_n|^p w\,\mathrm{d}\mu\leq C^p\int_\Omega\varphi_n^p w\,\mathrm{d}\mu,$$

since w_n is the conditional expectation of w. By (1.5) and the assumption $\varphi\in L^p(w)$, we have $\mathcal{M}_\mathcal{T}\varphi\in L^p(w)$ and hence we may let $n\to\infty$ above, obtaining

$$\|\mathcal{M}_\mathcal{T}\varphi\|^p_{L^p(w)}+\beta\|\mathcal{M}_\mathcal{T}\varphi-C\varphi\|^p_{L^p(w)}\leq C^p\|\varphi\|^p_{L^p(w)}.$$

Here we have used the fact that $\varphi_n\to\varphi$ almost surely, which is due to Lebesgue's differentiation theorem. Since $\|\mathcal{M}_\mathcal{T}\varphi\|_{L^p(w)}\geq(C-\varepsilon)\|\varphi\|_{L^p}$, this gives

$$\beta\|\mathcal{M}_\mathcal{T}\varphi-C\varphi\|^p_{L^p(w)}\leq(C^p-(C-\varepsilon)^p)\|\varphi\|^p_{L^p(w)}\leq pC^{p-1}\varepsilon\|\varphi\|^p_{L^p(w)},$$

as desired. For $1 < p \leq 2$ the reasoning is similar, but we need an additional application of Hölder's inequality. Namely, arguing as above we get

$$\beta \int_\Omega (\mathcal{M}_\mathcal{T}\varphi)^{p-2}(\mathcal{M}_\mathcal{T}\varphi - C\varphi)^2 \, \mathrm{d}\mu \leq pC^{p-1}\varepsilon \|\varphi\|^p_{L^p(w)}.$$

Since $p \leq 2$, this implies

$$\beta \|\mathcal{M}_\mathcal{T}\varphi\|^{p-2}_{L^p(w)} \|\mathcal{M}_\mathcal{T}\varphi - C\varphi\|^2_{L^p(w)} \leq pC^{p-1}\varepsilon \|\varphi\|^p_{L^p(w)}.$$

But by (1.5) and the estimate $p \leq 2$ again, we have $\|\mathcal{M}_\mathcal{T}\varphi\|^{p-2}_{L^p(w)} \geq C^{p-2}\|\varphi\|^{p-2}_{L^p(w)}$. Plugging this above, we get the assertion. □

Thus, all we need to establish the stability estimate is the existence of the appropriate special function B. It will be constructed and analyzed in the next section.

3. A special function

Throughout this section, we assume that $1 < p < \infty$ and $c > 1$ are given and fixed, and let $d = d(p, c)$ be as in (1.2); then $d > 1$. Consider the auxiliary parameters $\alpha = \alpha(p, c)$, $K = K(p, c)$, and $L = L(p, c)$, given by

$$\alpha = \frac{d}{d-1}, \quad L = d^{-1}c^{-1/p}, \quad \text{and} \quad K = L^{-1/(p-1)} - L.$$

Introduce the Bellman function $B\colon [0, \infty)^3 \times (0, \infty) \to \mathbb{R}$ by the formula

$$B(x, y, u, v) = \alpha(y^p u - c(Kx + Ly)^p v^{1-p}).$$

Now we will verify that B enjoys the properties (a), (b), and (c) listed in the previous section. To check condition (a), we note that $B(x, x, u, v) = \alpha x^p(u - c(K + L)^p v^{1-p}) \leq \alpha c x^p v^{1-p}(1 - (K + L)^p) = \alpha c x^p v^{1-p}(1 - d^{p/(p-1)}c^{1/(p-1)}) \leq 0$, where in the last passage we have used the fact that both c and d are bigger than 1. We continue with the concavity property.

Lemma 3.1. *The function B satisfies the requirement (c).*

Proof: We consider separately three major cases.

Case I: $x + h \leq y$. Under this assumption, the assertion follows at once from the fact that for a fixed y the function $(x, u, v) \mapsto B(x, y, u, v)$ is concave. To prove this property, it is enough to note that the function $(s, t) \mapsto s^p t^{1-p}$ is convex on $(0, \infty)^2$.

Case II: $x = y < x + h$. The estimate is equivalent to

$$\begin{aligned}((x + h)^p - x^p)(u + k) + c(K + L)^p[x^p v^{1-p} - (x + h)^p(v + \ell)^{1-p}] \\ \leq -pcK(K + L)^{p-1}x^{p-1}v^{1-p}h + (p - 1)c(K + L)^p x^p v^{-p}\ell.\end{aligned}$$

Now, since $x + h > x$, the estimate becomes the strongest if we take k as large as possible, i.e., we take $u + k = c(v + \ell)^{1-p}$. Under this additional assumption (and recalling that $L(K + L)^{p-1} = 1$, so that $1 - (K + L)^p = -K(K + L)^{p-1}$), we may rewrite the inequality in the equivalent form

$$\begin{aligned}&- K(x + h)^p(v + \ell)^{1-p} + pKx^{p-1}v^{1-p}h \\ &\qquad \leq Lx^p(v + \ell)^{1-p} - (K + L)x^p v^{1-p} + (p - 1)(K + L)x^p v^{-p}\ell.\end{aligned}$$

Divide throughout by $x^p v^{1-p}$ and substitute $h := h/x$, $\ell := \ell/v$, to obtain the simpler form

$$-K(1 + h)^p(1 + \ell)^{1-p} + pKh \leq L(1 + \ell)^{1-p} - (K + L) + (p - 1)(K + L)\ell.$$

A direct differentiation shows that the left-hand side, considered as a function of h, attains its maximum for $h = \ell$. Plugging this choice of h above, we obtain the bound equivalent to $(1+\ell)^{1-p} \geq 1 + (1-p)\ell$, which holds true by the convexity of the function $s \mapsto s^{1-p}$.

Case III: $x < y < x + h$. Let t be the unique number from the interval $(0,1)$ such that $x + th = y$. By Case I, we may write

$$B(x,y,u,v) + B_x(x,y,u,v)th + B_u(x,y,u,v)tk + B_v(x,y,u,v)t\ell \\ \geq B(x+th, y, u+tk, v+t\ell).$$

Next, by the reasoning used in Case I, we know that the function $\xi(s) = B(x + sh, y, u + sk, v + s\ell)$ is concave. This implies $\xi'(0) \geq \xi'(t)$, or

$$\begin{aligned} B_x(x,y,u,v)h &+ B_u(x,y,u,v)k + B_v(x,y,u,v)\ell \\ &\geq B_x(x+th,y,u+tk,v+t\ell)h + B_u(x+th,y,u+tk,v+t\ell)k \\ &\quad + B_v(x+th,y,u+tk,v+t\ell)\ell. \end{aligned}$$

Multiplying both sides by $1 - t$ and adding to the previous estimate, we obtain

$$\begin{aligned} B(x,y,u,v) &+ B_x(x,y,u,v)h + B_u(x,y,u,v)k + B_v(x,y,u,v)\ell \\ &\geq B(x+th,y,u+tk,v+t\ell) + B_x(x+th,y,u+tk,v+t\ell)(1-t)h \\ &\quad + B_u(x+th,y,u+tk,v+t\ell)(1-t)k + B_v(x+th,y,u+tk,v+t\ell)(1-t)\ell. \end{aligned}$$

Recall that $x + th = y$; thus, by Case II, the expression on the right is not smaller than

$$B(x+th+(1-t)h, y+(1-t)h, u+tk+(1-t)k, v+t\ell+(1-t)\ell) \\ = B(x+h, x+h, u+k, v+\ell).$$

This completes the proof. □

It remains to verify the majorization condition (b). Let us first establish an auxiliary technical fact.

Lemma 3.2. *Fix $p > 2$. For $t \in [1,p]$, let*

$$F_p(t) = (p-t)^p + p^p(t-1) - (p-1)^p t + t(1 - t^{p-1}). \tag{3.1}$$

Then $F_p(t) \geq 0$ for all $t \in [1,p]$; in particular, $F_p(d) \geq 0$.

Proof: We will verify the following properties of F_p: $F_p(1) = 0$; $F_p'(1) \geq 0$; $F_p'' > 0$ on $[1, p/2)$; $F_p'' < 0$ on $(p/2, p]$; and $F_p(p) \geq 0$. This will clearly yield the claim. The identity $F_p(1) = 0$ is obvious. The estimate $F_p'(1) \geq 0$ is equivalent to

$$p\left(\frac{p}{p-1}\right)^{p-1} \geq 2p - 1 + (p-1)^{2-p},$$

which follows at once from the inequalities $(p/(p-1))^{p-1} \geq 2$ and $(p-1)^{2-p} \leq 1$. The second derivative of F_p equals

$$F_p''(t) = p(p-1)((p-t)^{p-2} - t^{p-2}),$$

and hence the two estimates postulated for F_p'' above are evident. Finally, we compute that

$$F_p(p) = p(p-1)\left[\frac{p-2}{p-1}p^{p-1} + \frac{1}{p-1} - (p-1)^{p-1}\right]$$

and note that the expression in the square brackets is nonnegative: this follows at once from the Jensen inequality applied to the convex function $t \mapsto t^{p-1}$. □

Next, we will establish the following majorization. We write C for the optimal constant in (1.3): $C = p/(p-d)$.

Lemma 3.3. *Suppose that* $1 < p \leq 2$. *Then for all* (x, y, u, v) *satisfying* $x \leq y$ *and* $uv^{p-1} \geq 1$ *we have*

$$B(x, y, u, v) \geq y^p u - C^p x^p u + \beta y^{p-2}(y - Cx)^2 u,$$

where

$$\beta = \left(\frac{p-d}{d}\right)^2 \left[\left(\frac{p}{p-d}\right)^p - \frac{d}{d-1}\left(\frac{p-1}{p-d}\right)^p + \frac{1}{d-1}\right].$$

Proof: By continuity, we may assume that $y > 0$, $u > 0$, and $p < 2$. The right-hand side does not depend on v; furthermore, the left-hand side is a nondecreasing function of this parameter. Consequently, it is enough to study the majorization for $v = u^{1/(1-p)}$. Dividing throughout by $y^p u$ and substituting $s = x/y \in [0, 1]$, we obtain the equivalent estimate

$$F(s) := 1 - \alpha - C^p s^p + \beta(Cs - 1)^2 + \alpha c(Ks + L)^p \leq 0. \tag{3.2}$$

Using the definitions of α, β, C, K, L, and the identity (1.2), one verifies the equality

$$F(C^{-1}) = F'(C^{-1}) = F(1) = 0. \tag{3.3}$$

Therefore, by Rolle's theorem, F'' has a root in $(C^{-1}, 1)$; we will show that there are no other roots of F'' in $(0, 1)$. To this end, assume that $F''(s_0) = 0$ for some $s_0 \in (0, 1)$ and compute that

$$F''(s) = -p(p-1)s^{p-2}[C^p - \alpha c K^2 (K + Ls^{-1})^{p-2}] + 2C^2\beta.$$

Let us study the behavior of the expression $s^{p-2}[C^p - \alpha c K^2(K + Ls^{-1})^{p-2}]$ for $s \in (0, s_0)$. Both functions $s \mapsto s^{p-2}$ and $s \mapsto C^p - \alpha c K^2 (K + Ls^{-1})^{p-2}$ are decreasing on $(0, s_0)$. Furthermore, they are both positive at s_0: this is trivial for the first function, and for the other it follows from the equality $F''(s_0) = 0$. Thus we have $s^{p-2}[C^p - \alpha c K^2(K + Ls^{-1})^{p-2}] > s_0^{p-2}[C^p - \alpha c K^2(K + Ls_0^{-1})^{p-2}]$ for $s \in (0, s_0)$, which is equivalent to saying that $F'' < 0$ on $(0, s_0)$. This implies the existence of a unique $s_0 \in (C^{-1}, 1)$ such that $F''(s) < 0$ for $s < s_0$ and $F''(s) > 0$ for $s > s_0$. Combining this with (3.3), we easily deduce the estimate (3.2). □

Lemma 3.4. *Suppose that* $p > 2$. *Then for all* (x, y, u, v) *satisfying* $x \leq y$ *and* $uv^{p-1} \geq 1$ *we have*

$$B(x, y, u, v) \geq y^p u - C^p x^p u + \beta |y - Cx|^p u, \tag{3.4}$$

where

$$\beta = \frac{1 - d^{1-p}}{d-1} \geq \frac{1 - p^{1-p}}{p-1}. \tag{3.5}$$

Proof: Let us start with the estimate in (3.5): it follows directly from the fact that the function $s \mapsto (1 - s^{1-p})/(s-1)$ is decreasing on $(1, p)$. The latter property can be easily checked by differentiation: the monotonicity is equivalent to $(s^p - s)/(s-1) \geq p - 1$. This follows from the stronger bound $(s^{p-1} - 1)/(s-1) \geq p - 1$, which, in turn, is due to the convexity of $s \mapsto s^{p-1}$.

We proceed to the majorization (3.4). As before, we may assume that $y > 0$ and $u > 0$. The same argument as previously shows that it is enough to prove the majorization for $v = u^{1/(1-p)}$; then, by homogeneity, the claim can be rewritten in the form

$$F(s) = 1 - \alpha - C^p s^p + \beta |Cs - 1|^p + \alpha c(Ks + L)^p \leq 0, \quad s = x/y \in [0, 1]. \tag{3.6}$$

Now we consider two major cases.

Suppose first that $s > C^{-1}$. We compute that

$$F'(s) = pC^p s^{p-1}\left[\beta(1-u)^{p-1} - 1 + \frac{\alpha c K}{C}\left(\frac{K}{C} + Lu\right)^{p-1}\right],$$

where $u = (Cs)^{-1}$. Note that $u \in (C^{-1}, 1]$ if and only if $s \in [C^{-1}, 1)$. Denote the expression in the square brackets by $G(u)$; one easily checks directly that $G(1) = 0$,

$$G'(u) = (p-1)\left[-\beta(1-u)^{p-2} + \frac{\alpha c K L}{C}\left(\frac{K}{C} + Lu\right)^{p-2}\right]$$

and $G'' > 0$. Therefore, we have two possibilities. If $G'(C^{-1}) \geq 0$, then $G' > 0$ and $G < 0$ on $(C^{-1}, 1)$, and hence F is decreasing on this interval. This gives $F(s) \leq F(C^{-1}) = 0$, the desired bound. The second possibility is that $G'(C^{-1}) < 0$; since $G'(1) > 0$, there exists a unique parameter $u_0 \in (C^{-1}, 1)$ such that $G' < 0$ on (C^{-1}, u_0) and $G' > 0$ on $(u_0, 1)$. Combining this with the equality $G(1) = 0$, we see that either $G \leq 0$ on $(C^{-1}, 1)$, or there is $u_1 \in (C^{-1}, 1)$ such that $G > 0$ on (C^{-1}, u_1) and $G < 0$ on $(u_1, 1)$. Expressing these scenarios in the language of F' and the variable s, we either have $F' \leq 0$ on $(C^{-1}, 1)$ (and then $F(s) \leq F(C^{-1}) = 0$, as above), or $F' < 0$ on $(C^{-1}, (Cu_1)^{-1})$ and $F' > 0$ on $((Cu_1)^{-1}, 1)$. In the second case, since $F(C^{-1}) = 0$, it suffices to check the desired estimate (3.6) at the endpoint $s = 1$. After some straightforward manipulations, we transform the majorization into

$$\beta \leq \frac{(p-d)^p}{d^p(d-1)} + \left(\frac{p}{d}\right)^d - \frac{(p-1)^p}{d^{p-1}(d-1)}.$$

Plugging the above choice of β, we obtain an estimate equivalent to (3.1).

The second major case is $s < C^{-1}$. We proceed as above and compute that

$$F'(s) = pC^p s^{p-1}\left[-\beta(u-1)^{p-1} - 1 + \frac{\alpha c K}{C}\left(\frac{K}{C} + Lu\right)^{p-1}\right],$$

where, as above, $u = (Cs)^{-1} > 1$. Denoting the expression in the square brackets by $G(u)$, we check that $G(1) = 0$ and

$$G'(u) = (p-1)\left[-\beta(u-1)^{p-2} + \frac{\alpha c K L}{C}\left(\frac{K}{C} + Lu\right)^{p-2}\right].$$

Considering the two possibilities as above (the reasoning is repeated, word by word), we come to the conclusion that it is enough to check the estimate (3.6) in the endpoint case $s = 0$. But for this choice of s, we have $F(0) = 0$: this is actually where the formula for β comes from. □

4. A probabilistic counterpart

Theorem 1.2 has a natural extension to the context of martingales. Let us briefly discuss the necessary background on the subject. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$, i.e., a nondecreasing family of sub-σ-algebras of $\mathcal{F}$. Assume further that $X = (X_t)_{t\geq 0}$ is an adapted, real-valued martingale. We impose standard regularity assumption on the regularity of trajectories of X—we consider those X whose paths are right-continuous and have limits from the left. We will assume that X is L^1-convergent, as $t \to \infty$, to some integrable random variable, denoted by X_∞. Let $\mathcal{M}X = \sup_{t\geq 0} |X_t|$ be the maximal function of X; we will also use the notation $\mathcal{M}X_t = \sup_{0\leq s\leq t} |X_s|$ for the truncated version of $\mathcal{M}X$.

Let W be a probabilistic weight, i.e., a positive and integrable random variable. Given $1 < p < \infty$, we say that W belongs to the Muckenhoupt class A_p, if

$$[W]_{A_p} = \sup_{t \geq 0} \|\mathbb{E}(W|\mathcal{F}_t)(\mathbb{E}(W^{-1/(1-p)}|\mathcal{F}_t))^{p-1}\|_\infty < \infty.$$

As we will show below, if $X_\infty \in L^p(W)$ (i.e., $\|X_\infty\|_{L^p(W)} = (\mathbb{E}|X_\infty|^p W)^{1/p} < \infty$), then we have the following analogue of (1.5):

$$\|\mathcal{M}X\|_{L^p(W)} \leq \frac{p}{p - d(p, [W]_{A_p})} \|X\|_{L^p(W)}. \tag{4.1}$$

Note that we may restrict ourselves to nonnegative martingales. Indeed, suppose that we have established (4.1) in this case, and let X be an arbitrary martingale such that $X_\infty \in L^p(W)$. Then the process $(\tilde{X}_t)_{t\geq 0} = (\mathbb{E}(|X_\infty||\mathcal{F}_t))_{t\geq 0}$ is a positive uniformly integrable martingale satisfying $\tilde{X}_t \geq |\mathbb{E}(X_\infty|\mathcal{F}_t)|$ almost surely, by the Jensen inequality. Consequently, we have $\mathcal{M}X \leq \mathcal{M}\tilde{X}$ with probability 1, so by (4.1) applied to $\tilde{X}$,

$$\|\mathcal{M}X\|_{L^p(W)} \leq \frac{p}{p - d(p, [W]_{A_p})} \|\tilde{X}\|_{L^p(W)} = \frac{p}{p - d(p, [W]_{A_p})} \|X\|_{L^p(W)}.$$

Here in the last passage we have used the almost sure identity $\tilde{X}_\infty = |X_\infty|$.

The inequality (4.1) is sharp, since the martingale context generalizes the dyadic setup discussed in the introductory section. Thus, we may ask about the corresponding stability result. We will prove the following version of Theorem 1.2.

Theorem 4.1. *Let $1 < p < \infty$, $c \geq 1$, and $\varepsilon > 0$ be fixed. Assume further that a nonnegative martingale $X \in L^p(W)$ and a weight $W \in A_p$ satisfy $[W]_{A_p} = c$ and*

$$\|\mathcal{M}X\|_{L^p(W)} > \left(\frac{p}{p - d(p, c)} - \varepsilon\right) \|X\|_{L^p(W)}.$$

Then there is a finite constant c_p depending only on p such that

$$\left\|\mathcal{M}X - \frac{p}{p - d(p, c)} X\right\|_{L^p(W)} \leq c_p\, \varepsilon^{\min\{1/2, 1/p\}} \|X\|_{L^p(W)} \tag{4.2}$$

and the exponent $\min\{1/2, 1/p\}$ *is the best possible.*

Note that this statement yields (4.1). Indeed, suppose conversely that for a given $1 < p < \infty$ there is a martingale X for which

$$\|\mathcal{M}X\|_{L^p(W)} > \frac{p}{p - d(p, c)} \|X\|_{L^p(W)}.$$

Then the above theorem would imply

$$\|\mathcal{M}X\|_{L^p(W)} \leq \left\|\mathcal{M}X - \frac{p}{p - d(p, c)} X\right\|_{L^p(W)} + \frac{p}{p - d(p, c)} \|X\|_{L^p(W)} \xrightarrow{\varepsilon \to 0} \frac{p}{p - d(p, c)} \|X\|_{L^p(W)},$$

a contradiction.

Proof of (4.2)*:* We will make use of Itô's formula and the Bellman function B introduced in the previous section. It is convenient to split the reasoning into a few steps.

Step 1. An application of Itô's formula. Consider the auxiliary weight $V = W^{1/(1-p)}$ and the associated martingales $W_t = \mathbb{E}(W|\mathcal{F}_t)$, $V_t = \mathbb{E}(V|\mathcal{F}_t)$ for $t \geq 0$. Note that by the assumed A_p condition, the four-dimensional process $\xi = ((X_t, \mathcal{M}X_t, W_t, V_t))_{t\geq 0}$ takes values in the domain (2.1). By Itô's formula,

$$B(\xi_t) = I_0 + I_1 + I_2 + I_3/2 + I_4,$$

where

$$\begin{aligned} I_0 &= B(\xi_0),\\ I_1 &= \int_0^t B_x(\xi_{s-})\,\mathrm{d}X_s + \int_0^t B_u(\xi_{s-})\,\mathrm{d}W_s + \int_0^t B_v(\xi_{s-})\,\mathrm{d}V_s,\\ I_2 &= \int_0^t B_y(\xi_{s-})\,\mathrm{d}(\mathcal{M}X_s),\\ I_3 &= \int_0^t D^2_{x,u,v}B(\xi_s)\,\mathrm{d}[X,W,V]^c_s,\\ I_4 &= \sum_{0<s\leq t}(B(\xi_s) - B(\xi_{s-}) - B_x(\xi_{s-})\Delta X_s - B_u(\xi_{s-})\Delta W_s - B_v(\xi_{s-})\Delta V_s). \end{aligned}$$

Here in the definition of I_3 the symbol $D^2_{x,u,v}B$ stands for the Hessian matrix of B, treated as a function of x, u, and v, and we have used the shortened notation for the sum of all second-order terms. That is, we have

$$I_3 = \int_0^t B_{xx}(\xi_{s-})\,\mathrm{d}[X]^c_s + 2\int_0^t B_{xu}(\xi_{s-})\,\mathrm{d}[X,W]^c_s + 2\int_0^t B_{xv}(\xi_{s-})\,\mathrm{d}[X,V]^c_s + \cdots .$$

Step 2. Analysis of the terms I_0 through I_4. By the initial condition (a), we see that $I_0 \leq 0$. Next, the stochastic integrals appearing in I_1 are local martingales, so, applying the appropriate localization if necessary (i.e., replacing t with $\tau_n \wedge t$, where τ_n is the common localizing sequence, and letting $n \to \infty$ at the end), we may assume that $\mathbb{E}I_1 = 0$. The terms I_2 and I_3 are nonpositive, since $B_y \leq 0$ and the Hessian matrix $D^2_{x,u,v}$ is seminegative-definite, as one easily checks. Finally, each summand appearing in I_4 is nonpositive: this follows directly from (2.2), applied to $(x, y, u, v) = \xi_{s-}$ and $(h, k, \ell) = (\Delta X_s, \Delta W_s, \Delta V_s)$. Putting all these observations together, we obtain that $\mathbb{E}B(\xi_t) \leq 0$ for any $t \geq 0$.

Step 3. Completion of the proof. It remains to repeat, almost word by word, the arguments from Step 3 of the proof of Theorem 2.1. One just needs to replace the tree-maximal operator $\mathcal{M}_\mathcal{T}$ with the martingale maximal function $\mathcal{M}$ and the function φ with the martingale X. The straightforward modification is left to the reader. □

5. Sharpness

Fix $1 < p < \infty$, $c \geq 1$, and let $d = d(p,c) \in [1,p)$ be given by (1.2). For the sake of convenience, we have decided to split the contents of this section into three separate parts.

5.1. Sharpness for the martingale case. It is most convenient to start with the probabilistic estimate (4.2): the example has a "continuous" character and the calculations are easier (as we shall see later, the discretization complicates the analysis significantly).

Consider the probability space $((0,1],\mathcal{B}(0,1),|\cdot|)$, where $\mathcal{B}(0,1)$ stands for the σ-algebra of Borel subsets of $(0,1]$. We equip this space with the filtration $(\mathcal{F}_t)_{t\in[0,1]}$, where $\mathcal{F}_t$ is generated by the interval $(0,1-t]$ and all Borel subsets of $(1-t,1]$. Introduce the probabilistic weight W on $(0,1]$, given by $W(s)=s^{d-1}$. Then for any $t\in(0,1)$ and $A=(0,1-t]$ we have

$$\left(\frac{1}{|A|}\int_A W\right)\left(\frac{1}{|A|}\int_A W^{1/(1-p)}\right)^{p-1}=\frac{1}{d}\left(\frac{p-1}{p-d}\right)^{p-1}=c$$

(the last equality is due to (1.2)), which implies $W\in A_p$ and $[W]_{A_p}=c$. An important comment is in order: if we slightly modified the exponent and consider the weight $W(s)=s^{\tilde{d}-1}$ for some $\tilde{d}$ close to but smaller than d, then we would get $[W]_{A_p}<c$. This observation will be helpful later.

Now suppose that $p\geq 2$ and fix another auxiliary parameter $\eta\in(0,1)$ (which will be sent to zero). Then the random variable

$$X_\infty(s)=\frac{\eta^{-d/p}}{1-d/p}\chi_{(0,\eta]}(s)+s^{-d/p}\chi_{(\eta,1]}(s)$$

satisfies

$$\|X_\infty\|^p_{L^p(W)}=\left(\frac{p}{p-d}\right)^p\frac{1}{d}-\ln\eta. \tag{5.1}$$

The maximal function is analyzed in the separate statement below.

Lemma 5.1. *The maximal function of X admits the explicit formula*

$$\mathcal{M}X(s)=\frac{p}{p-d}\max\{s,\eta\}^{-d/p},\quad s\in(0,1].$$

Proof: We have $\mathcal{M}X=\sup_{t\in[0,1]}X_t=\sup_{t\in[0,1]}X_{1-t}=\sup_{t\in[0,1]}\mathbb{E}(X_\infty|\mathcal{F}_{1-t})$. Now, for any $s\in(0,1]$ and $t\in[0,1]$ we compute that

$$\mathbb{E}(X_\infty|\mathcal{F}_{1-t})(s)=\begin{cases}\frac{1}{t}\int_0^t X_\infty & \text{if } s\leq t,\\ X_\infty(s) & \text{if } s\geq t.\end{cases}$$

Since $s\mapsto X_\infty(s)$ is nonincreasing, we have

$$\mathcal{M}X(s)=\frac{1}{s}\int_0^s X_\infty=\begin{cases}X_\infty(s) & \text{for } s\in(0,\eta],\\ \frac{X_\infty(s)}{1-d/p} & \text{for } s\in(\eta,1]\end{cases}$$

$$=\frac{1}{1-d/p}\max\{s,\eta\}^{-d/p}.$$

□

We proceed with the sharpness. By the lemma above, we may write

$$\|\mathcal{M}X\|_{L^p(W)}^p \geq \int_\eta^1 \frac{u^{-d}}{(1-d/p)^p} W(u)\,\mathrm{d}u = \left(\frac{p}{p-d}\right)^p \cdot (-\ln\eta),$$

which combined with (5.1) yields

$$\|\mathcal{M}X\|_{L^p(W)} > \left(\frac{p}{p-d} - \varepsilon\right) \|X_\infty\|_{L^p(W)},$$

with $\varepsilon = O((-\ln\eta)^{-1})$ as $\eta \to 0$. It remains to note that

$$\begin{aligned}
\left\|\mathcal{M}X - \frac{p}{p-d}X_\infty\right\|_{L^p(W)}^p &= \int_0^\eta \left|\mathcal{M}X - \frac{p}{p-d}X_\infty\right|^p W \\
&= \left(\frac{d}{p-d}\right)^p \int_0^\eta X_\infty^p W \\
&= \left(\frac{d}{p-d}\right)^p \left(\frac{p}{p-d}\right)^p \eta^{-d} \int_0^\eta s^{d-1}\,\mathrm{d}s \\
&= \frac{1}{d}\left(\frac{d}{p-d}\right)^p \left(\frac{p}{p-d}\right)^p \\
&= O((-\ln\eta)^{-1}) \int_0^1 X_\infty^p W,
\end{aligned}$$

where the last passage is due to (5.1). This ends the proof of the case $p \geq 2$.

For $1 < p < 2$ the calculations are a bit more involved. Fix small $\eta, \varepsilon > 0$ and negative α, β determined by the conditions

$$(1+\alpha)^{-1} = \frac{p}{p-d} - \varepsilon^{1/2}, \quad (1+\beta)^{-1} = \frac{p}{p-d} + \varepsilon^{1/2}.$$

We take the same weight W as above, but now the extremal nonnegative martingale X is generated from

$$X_\infty(s) = \frac{(\alpha+1)s^\alpha}{\eta^{\alpha+1}}\chi_{(0,\eta]}(s) + \frac{(\beta+1)s^\beta}{\eta^{\beta+1}}\chi_{(\eta,1]}(s).$$

As previously, we check that

$$\frac{1}{t}\int_0^t X_\infty = \frac{t^\alpha}{\eta^{\alpha+1}}\chi_{(0,\eta]}(t) + \frac{t^\beta}{\eta^{\beta+1}}\chi_{(\eta,1]}(t) = \frac{X_\infty(t)}{\alpha+1}\chi_{(0,\eta]}(t) + \frac{X_\infty(t)}{\beta+1}\chi_{(\eta,1]}(t),$$

which implies $\mathcal{M}X - \frac{p}{p-d}X_\infty = \pm\varepsilon^{1/2}X_\infty$ and thus

$$\left\|\mathcal{M}X - \frac{p}{p-d}X_\infty\right\|_{L^p(W)} = \varepsilon^{1/2}\|X_\infty\|_{L^p(W)}.$$

Thus we will be done if we prove that $\|\mathcal{M}X\|_{L^p(W)} > \big(\frac{p}{p-d} - O(\varepsilon)\big)\|X_\infty\|_{L^p(W)}$. To this end, we compute that

$$\frac{\|\mathcal{M}X\|_{L^p(W)}}{\|X_\infty\|_{L^p(W)}} = \left(\frac{(\alpha+1)^{-p}\int_0^\eta X_\infty^p W + (\beta+1)^{-p}\int_\eta^1 X_\infty^p W}{\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W}\right)^{1/p}$$

$$= \left(\frac{p}{p-d} + \varepsilon^{1/2}\left(\frac{p}{p-d}\right)^{p-1} \cdot \frac{-\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W}{\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W} + O(\varepsilon)\right)^{1/p}$$

and hence it is enough to pick η so that the ratio

$$\left(-\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W\right) \Big/ \left(\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W\right)$$

is of order $O(\varepsilon^{1/2})$. However,

$$\int_0^\eta X_\infty^p W = \frac{(\alpha+1)^p \eta^{d-p}}{\alpha p + d}, \quad \int_\eta^1 X_\infty^p W = \frac{(\beta+1)^p \eta^{d-p}(1-\eta^{-\beta p - d})}{-\beta p - d}$$

and

$$\alpha p + d = \frac{\varepsilon^{1/2}(p-d)^2}{p - \varepsilon^{1/2}(p-d)}, \quad \beta p + d = -\frac{\varepsilon^{1/2}(p-d)^2}{p + \varepsilon^{1/2}(p-d)}.$$

This implies that the ratio (as previously, we use the notation $C = p/(p-d)$, for brevity)

$$\frac{-\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W}{\int_0^\eta X_\infty^p W + \int_\eta^1 X_\infty^p W} = \frac{-(C-\varepsilon^{1/2})(\alpha+1)^p + (C+\varepsilon^{1/2})(\beta+1)^p(1-\eta^{-\beta p - d})}{(C-\varepsilon^{1/2})(\alpha+1)^p + (C+\varepsilon^{1/2})(\beta+1)^p(1-\eta^{-\beta p - d})}$$

converges, as $\eta \to 0$, to

$$\frac{-(C-\varepsilon^{1/2})(\alpha+1)^p + (C+\varepsilon^{1/2})(\beta+1)^p}{(C-\varepsilon^{1/2})(\alpha+1)^p + (C+\varepsilon^{1/2})(\beta+1)^p}.$$

But by the very definition of α and β, the numerator, and hence the whole fraction, is of order $O(\varepsilon^{1/2})$. This gives the claim.

5.2. Sharpness for a special tree. Now we will discretize the example from the previous subsection. We will restrict ourselves to the case $c > 1$. In the unweighted context $c = 1$, the argument is much simpler and goes along the same lines.

The idea is straightforward: we take the probability space $((0,1], \mathcal{B}(0,1), |\cdot|)$ as before and consider the following natural approximation of the filtration $(\mathcal{F}_t)_{t\in[0,1]}$. Namely, for a small $\delta > 0$ and any integer n, let $\mathcal{T}^{(n)}$ consist of the interval $(0, (1-\delta)^n]$ and a collection of sufficiently small subintervals of $((1-\delta)^n, 1]$ (their size will be specified in a moment). We define w and φ by averaging of W and X_∞: for any $n \geq 0$, we assume that w and φ are constant on $J_n = ((1-\delta)^{n+1}, (1-\delta)^n]$, with

$$w = \frac{1}{|J_n|}\int_{J_n} W, \quad \varphi = \frac{1}{|J_n|}\int_{J_n} X_\infty.$$

Now it is easy to see that if δ and the intervals in $\mathcal{T}^{(n)}$ are chosen sufficiently small, then the distributions of w, φ, and $\mathcal{M}_\mathcal{T}\varphi$ are arbitrarily close to W, X_∞, and $\mathcal{M}X$, so in particular the sharpness follows. The only difficulty we must handle is the bound for the A_p characteristic of w. The above averaging process in general increases the characteristic, but not much: if $\tilde{c} \in (c, \infty)$ is an arbitrary level, then the choice of sufficiently small δ leads to the weight w satisfying $[w]_{A_p} \leq \tilde{c}$. So, to overcome the difficulty, we make use of the comment formulated in the previous subsection: the modified weight $W(s) = s^{\tilde{d}-1}$ (for some $\tilde{d} < d$) satisfies $[W]_{A_p} < c$. Furthermore, by continuity, the estimates of the previous subsection remain valid if $\tilde{d}$ is sufficiently close to d. For this modified W, we can now apply the above discretization argument and obtain the weight w with the characteristic not bigger than c.

5.3. Sharpness for a general tree. Finally, we will sketch the argument for an arbitrary tree structure $\mathcal{T}$; since a similar construction appears in many papers (e.g., see Section 4 in [**20**], Section 4 in [**22**], or Section 2 in [**23**]), so we will be brief. The idea is to embed appropriately the examples from Subsection 5.2 into the tree context. Namely, arguing as in [**23**], we show that for any $\delta > 0$ there is a decreasing sequence $X = A_0 \supset A_1 \supset A_2 \supset \cdots$ of sets which enjoy the following properties:

- each A_n is a union of pairwise disjoint elements (atoms) of $\mathcal{T}$;
- we have $\mu(Q \cap A_m) = (1-\delta)^{n-m}\mu(Q)$ for any $n \geq m$ and any atom Q of A_n.

The sets $A_0, A_1, A_2, \ldots$ are precisely the analogues of $(0,1], (0,1-\delta], (0,(1-\delta)^2], \ldots$ We repeat the construction from the previous subsection, replacing $J_0, J_1, J_2, \ldots$ with $A_0 \setminus A_1, A_1 \setminus A_2, A_2 \setminus A_3, \ldots$ Define the weight w and the function φ by

$$w = \sum_{n=0}^{\infty} \left(\frac{1}{|A_n \setminus A_{n+1}|} \int_{A_n \setminus A_{n+1}} W \right) \chi_{A_n \setminus A_{n+1}}$$

and

$$\varphi = \left(\sum_{n=0}^{\infty} \frac{1}{|A_n \setminus A_{n+1}|} \int_{A_n \setminus A_{n+1}} X_\infty \right) \chi_{A_n \setminus A_{n+1}}.$$

It is now straightforward to check that, with this choice, the estimates and conditions from the previous subsection carry over: see [**23**]. This gives the desired claim.

6. On the search for the Bellman function

The Bellman function B we have used in the proof of the estimate (1.4) might look a little mysterious and there is a natural question about how it was invented. The purpose of this section is to sketch some more or less formal argumentation which leads to the discovery of this object.

Step 1. An easier estimate. Let us first describe how the Bellman function approach can be used to yield directly the sharp weighted L^p estimates for maximal functions; that is, let us skip the question of stability for a moment. This can done with a slight modification of the method described in Section 2. Suppose that for a given $1 < p < \infty$ and $c \geq 1$ we are interested in the best constant $C = C_{p,c}$ depending only on the parameters indicated such that

$$\|\mathcal{M}_{\mathcal{T}}\varphi\|_{L^p(w)} \leq C_{p,c}\|\varphi\|_{L^p(w)}, \tag{6.1}$$

for all $w \in A_p$ with $[w]_{A_p} \leq c$ and all $\varphi \in L^p(w)$. Here the probability space, as well as the tree structure $\mathcal{T}$, is allowed to vary. To study this problem, we distinguish the domain

$$\mathcal{D} = \{(x,y,u,v) \in [0,\infty)^2 \times (0,\infty)^2 : x \leq y,\ 1 \leq uv^{p-1} \leq c\}$$

and search for a function $B\colon \mathcal{D} \to \mathbb{R}$ satisfying the following requirements.

(a)' *Initial condition.* We have $B(x,x,u,v) \leq 0$ for all $(x,x,u,v) \in \mathcal{D}$.

(b)' *Majorization.* There is a positive constant C such that

$$B(x,y,u,v) \geq y^p u \left[1 - \left(\frac{C_{p,c}x}{y}\right)^p\right]$$

for all $(x,y,u,v) \in \mathcal{D}$ with $y > 0$.

(c)' *Concavity.* If (x,y,u,v) and $(x+h,\max\{x+h,y\},u+k,v+\ell)$ belong to $\mathcal{D}$, then we have

$$B(x+h,\max\{x+h,y\},u+k,v+\ell)\le B(x,y,u,v)+B_x(x,y,u,v)h + B_u(x,y,u,v)k+B_v(x,y,u,v)\ell.$$

Comparing these conditions with those listed in Section 2, we see that the only difference can be found in the majorization property. However, let us keep the prime notation to indicate that we are dealing with the estimate (6.1), not its stability.

Repeating the proof of Theorem 2.1, if B satisfies these conditions, then the estimate (6.1) holds true. The beauty and efficiency of the method lie in the fact that the implication can be reversed: if we know a priori that the estimate (6.1) holds, then there exists a function which enjoys (a)', (b)', and (c)'. In general, there may be many such functions, and the (pointwise) smallest of them is given by

$$\mathcal{B}(x,y,u,v)=\sup\int_\Omega(\max\{\mathcal{M}\varphi,y\}^pw-C^w_{p,c}\varphi^pw)\,\mathrm{d}\mu.$$

Here the supremum is taken over all probability spaces $(\Omega,\mathcal{F},\mu)$, all tree structures $\mathcal{T}$, all simple $A_{p,\mathcal{T}}$-weights w satisfying

$$[w]_{A_{p,\mathcal{T}}}\le c,\qquad \int_\Omega w\,\mathrm{d}\mu=u,\quad\text{and}\quad \int_\Omega w^{1/(1-p)}\,\mathrm{d}\mu=v,$$

and all simple positive functions $\varphi\in L^p(w)$. Here by simplicity of a function we mean that it is measurable with respect to $\sigma(\mathcal{T}^{(n)})$ for some integer n. The above formula is abstract and nonexplicit, but it gives some structural properties of the special function. For example, one easily checks the homogeneity properties

$$\mathcal{B}(\lambda x,\lambda y,u,v)=\lambda^p\mathcal{B}(x,y,u,v)\quad\text{for }(x,y,u,v)\in\mathcal{D},\ \lambda>0, \tag{6.2}$$

and

$$\mathcal{B}(x,y,\lambda u,\lambda^{1/(1-p)}v)=\lambda\mathcal{B}(x,y,u,v)\qquad\text{for }(x,y,u,v)\in\mathcal{D},\ \lambda>0. \tag{6.3}$$

Thus, during the search for Bellman functions, we may impose additionally the conditions (6.2) and (6.3): there is at least one special function satisfying these extra properties.

It is not clear, at least to us, whether such reverse implication holds for stability result as well. However, the analysis of the easier problem (6.1) can be helpful, as we describe now. As we already observed above, if we compare the properties (a)', (b)', and (c)' with their versions (a), (b), and (c) formulated in Section 2, only the majorization condition is different. More specifically, (b) from the context of stability is slightly stronger: it involves an additional term with the factor β. Thus, a natural idea is the following: try to find the Bellman function for the weighted maximal L^p estimate, and then hope that it enjoys the stronger majorization leading to the stability result. As we will see in a moment, this idea works perfectly, though it might require some additional work with the Bellman function.

Step 2. The unweighted case. Thus, from now on, we search for a function leading to (6.1). The very natural next step is to inspect carefully the unweighted maximal L^p estimate

$$\|\mathcal{M}_\mathcal{T}\varphi\|_{L^p}\le C_p\|\varphi\|_{L^p},\quad 1<p<\infty, \tag{6.4}$$

with the approach described above. The unweighted case corresponds to the choice $c=1$, for which the domain has a much simpler, three-dimensional form. Actually, the setup can be further simplified: it is enough to construct the function B depending on x and y only; specifically, one reduces the domain to $\tilde{\mathcal{D}} = \{(x,y) \in [0,\infty)^2 : x \leq y\}$ and searches for a function $b\colon \tilde{\mathcal{D}} \to \mathbb{R}$ satisfying

(a)" *Initial condition.* We have $b(x,x) \leq 0$ for all $x > 0$.

(b)" *Majorization.* We have

$$b(x,y) \geq y^p - C_p^p x^p \quad \text{for all } (x,y) \in \tilde{\mathcal{D}}.$$

(c)" *Concavity.* If (x,y) and $(x+h, \max\{x+h,y\})$ belong to $\tilde{\mathcal{D}}$, then we have

$$b(x+h, \max\{x+h,y\}) \leq b(x,y) + b_x(x,y,u,v)h.$$

It is well known that such a function exists if and only if $C_p \geq p/(p-1)$: thus the best constant in (6.4) is $p/(p-1)$. For the optimal choice $C_p = p/(p-1)$, the *smallest* b with the above properties is given by

$$b(x,y) = \begin{cases} y^p - \left(\dfrac{p}{p-1}\right)^p x^p & \text{if } y \geq \dfrac{p}{p-1}x, \\ \dfrac{p}{p-1} y^{p-1}\left(y - \dfrac{p}{p-1}x\right) & \text{if } y \leq \dfrac{p}{p-1}x. \end{cases}$$

See [**12**], for instance. Unfortunately, this function does not satisfy the stronger inequality

$$b(x,y) \geq y^p \left[1 - \left(\frac{px}{(p-1)y}\right)^p + \beta \left|1 - \frac{px}{(p-1)y}\right|^{\max\{p,2\}}\right]$$

with any $\beta > 0$ (the reason is the "bad" behavior of the inequality for $y \geq px/(p-1)$). In other words, the above function b does not seem to lead to any stability result. Fortunately, this is not the only Bellman function leading to (6.4). As observed by Burkholder [**6**] (see also [**17**]), the function

$$b(x,y) = py^{p-1}\left(y - \frac{p}{p-1}x\right) = \frac{p}{p-1}\left(y^p - \frac{p}{p-1}y^{p-1}x\right) \tag{6.5}$$

also satisfies the conditions (a)", (b)", and (c)". One can check that this function does satisfy the stronger majorization and does produce a stability result. Actually, this is a function which we use above in the case $c = 1$ (see Remark 6.1 below).

Step 3. The weighted case. Now we proceed to the context of general A_p weights; fix $1 < p < \infty$ and $c > 1$. We will find the Bellman function, and our argumentation will also show where the optimal constant $C_{p,c} = p/(p - d(p,c))$ comes from. The first step is to try to understand condition (c) better. Clearly, this property implies that for any y the function $(x,u,v) \mapsto B(x,y,u,v)$ must be locally concave on the three-dimensional domain $\{(x,u,v) : x \leq y,\ 1 \leq uv^{p-1} \leq c\}$ (that is, concave along any line segment entirely contained in the domain). Next, if $(x,x,u,v) \in \mathcal{D}$ and $x > 0$, then for $h \in (0,x)$ condition (c) yields

$$B(x+h, x+h, u, v) \leq B(x,x,u,v) + B_x(x,x,u,v)h$$

and

$$B(x-h, x, u, v) \leq B(x,y,u,v) - B_x(x,y,u,v)h.$$

This implies $B(x+h,x+h,u,v)+B(x-h,x,u,v) \leq 2B(x,y,u,v)$ and hence, assuming that $B \in C^1$, we must have $B_y(x,x,u,v) \leq 0$. These two consequences, the local concavity of $(x,u,v) \mapsto B(x,y,u,v)$ and the requirement $B_y(x,x,u,v) \leq 0$, will be all we need during the search.

How to find the Bellman function? A natural idea is to play with the expression (6.5), adding weight components to the formula. Taking the homogeneity properties (6.2) and (6.3) into account, one could start with the function

$$B(x, y, u, v) = \alpha(y^p u - \kappa y^{p-1} xu),$$

where α, κ are some positive parameters. But this is clearly wrong: the local concavity of $(x, u, v) \mapsto B(x, y, u, v)$ does not hold. To correct the problematic term $-\kappa y^{p-1} xu$, let us think about some different expression which is homogeneous of order p with respect to x, y, and satisfies (6.3). Such a term is well known and appears in many papers (see e.g. [**25**]): it is equal to $-\kappa x^p v^{1-p}$. It also enjoys the following property: for any x, v, there is a line passing through (x, v) along which the term is linear. In other words, the term is concave and there is a direction along which the concavity degenerates. This is in perfect correspondence with the second term in (6.5), which is also linear (in x). Thus, we have arrived at the function

$$B(x, y, u, v) = \alpha(y^p u - \kappa x^p v^{1-p}).$$

Unfortunately, this does not work either, since the inequality $B_y(x, x, u, v) \leq 0$, the second part of the concavity condition, is not true. To guarantee this inequality, we need to insert somehow the variable y into the term $-\kappa x^p v^{1-p}$, so that the local concavity is not ruined. A little thought and experimentation leads to the corrected term $-(Kx + Ly)^p v^{1-p}$, for some constants K and L, and the corresponding function

$$B(x, y, u, v) = \alpha(y^p u - (Kx + Ly)^p v^{1-p}).$$

Actually, it is more convenient to work with a slightly different expression

$$B(x, y, u, v) = \alpha(y^p u - c(Kx + Ly)^p v^{1-p}),$$

involving the factor c inside. This is precisely the formula we introduced in Section 3. To identify the constants α, K, and L, we inspect the conditions (a), the weaker majorization (b)', and the concavity (c) the function B must satisfy. First, the inequality $B_y(x, x, u, v) \leq 0$ (which follows from the concavity) is equivalent to $u - cL(K+L)^{p-1} v^{1-p} \leq 0$. If we fix v and allow u to vary, then this bound becomes strongest for $u = cv^{1-p}$, and this strongest version is equivalent to $1 \leq L(K + L)^{p-1}$. We *assume* equality here: this gives us the equation

$$K = L^{-1/(p-1)} - L$$

appearing in Section 3 in the definition of B. Next, we look at the (weaker) majorization $B(x, y, u, v) \geq y^p u - C_{p,c} x^p u$. If we fix u and let v vary, this inequality is the strongest for $v^{1-p} = u$, and this strongest version is equivalent to

$$C_{p,c}^p \geq (1 - \alpha)s^p + \alpha c(L^{-1/(p-1)} - L + Ls)^p,$$

after the substitution $s = y/x \geq 1$. This gives us the following natural conjecture about the best constant in (6.1):

$$C_{p,c} = \inf_{\alpha, L} \sup_s \{(1 - \alpha)s^p + \alpha c(L^{-1/(p-1)} - L + Ls)^p\}.$$

To compute the right-hand side, fix α, L, and maximize the expression in the parentheses overs. We may assume that

$$\alpha c L^p < \alpha - 1, \tag{6.6}$$

since otherwise this maximum is infinite. In particular, we must have $cL^p < 1$. A straightforward analysis of the derivative shows that the maximum is attained for

$$s = \frac{L^{-p/(p-1)} - 1}{\left(\frac{L^p \alpha c}{\alpha - 1}\right)^{-1/(p-1)} - 1},$$

which gives

$$C_{p,c} = \inf \frac{(\alpha - 1)(L^{-p/(p-1)} - 1)^p}{\left(\left(\frac{L^p \alpha c}{\alpha - 1}\right)^{-1/(p-1)} - 1\right)^{p-1}}.$$

Here the infimum is taken over all $\alpha > 1$ and $L \in (0, c^{-1/p})$ satisfying (6.6). Fixing L and differentiating over α, we check that the expression on the right is minimal for $\alpha = \delta/(\delta - 1)$, where $\delta = L^{-1}c^{-1/p} > 1$. Plugging this above and optimizing over L, we verify that the minimum is equal to $p/(p - d(p,c))$, and it is attained for $L = d(p,c)^{-1}c^{-1/p}$ (where $d(p,c)$ was defined in (1.2)). Thus we have obtained the value of the best constant and the formula for the special function of Section 3.

We conclude with a comment about the limit of the above Bellman functions as $c \downarrow 1$.

Remark 6.1. Fix $1 < p < \infty$ and let $B^{(c)}$ be the Bellman function constructed above (note that we indicate the dependence of B on c). The calculations carried out in Step 3 above work for c strictly bigger than 1, and hence there is a natural question about the pointwise limit of these functions as $c \downarrow 1$. As c decreases to 1, the domain of $B^{(c)}$ shrinks to $\{(x, y, u, v) \in [0, \infty)^2 \times (0, \infty)^2 : x \leq y,\ uv^{p-1} = 1\}$. Let (x, y, u, v) be a fixed point from this set. To find the limit

$$B^{(1)}(x, y, u, v) := \lim_{c \downarrow 1} B^{(c)}(x, y, u, v),$$

note that the equation (1.2) is equivalent to $c = (p-1)^p d^{-1}(p-d)^{1-p}$, and hence we may carry out all computations explicitly in terms of the variable d. Recall that

$$\alpha = \frac{d}{d-1}, \quad L = d^{-1}c^{-1/p} = \frac{d^{1/p-1}(p-d)^{(p-1)/p}}{(p-1)^{(p-1)/p}}, \quad \text{and} \quad K = L^{-1/(p-1)} - L,$$

and observe that d goes to 1 as c tends to 1. Consequently,

$$\lim_{c \to 1} \frac{L-1}{d-1} = \lim_{d \to 1} \frac{L-1}{d-1} = -1, \quad \lim_{c \to 1} \frac{K}{d-1} = \frac{p}{p-1},$$

and therefore

$$\begin{aligned} \lim_{c \to 1} B^{(c)}(x, y, u, v) &= \lim_{c \to 1} \alpha(y^p - (Kx + Ly)^p)u \\ &= \lim_{d \to 1} \frac{y^p - (Kx + Ly)^p}{y - (Kx + Ly)} \cdot \frac{d(y(1-L) - Kx)}{d-1} \cdot u \\ &= py^{p-1}\left(y - \frac{p}{p-1}x\right)u. \end{aligned}$$

This is precisely the Bellman function defined in (6.5), multiplied by the factor u. Thus $(B^{(c)})_{c \geq 1}$ can be regarded as a continuous-scale extension of b.

Acknowledgments

The authors would like to thank the anonymous referees for their very thorough reading of the paper and for their numerous comments and suggestions.

References

[1] R. Bañuelos and A. Osękowski, Weighted norm inequalities for fractional maximal operators: a Bellman function approach, *Indiana Univ. Math. J.* **64(3)** (2015), 957–972. DOI: 10.1512/iumj.2015.64.5534.

[2] R. Bañuelos and A. Osękowski, Sharp weak type inequalities for fractional integral operators, *Potential Anal.* **47(1)** (2017), 103–121. DOI: 10.1007/s11118-016-9610-x.

[3] R. Bellman, *Dynamic Programming*, Reprint of the 1957 edition, With a new introduction by Stuart Dreyfus, Princeton Landmarks Math. **33**, Princeton University Press, Princeton, NJ, 2010. DOI: 10.2307/j.ctv1nxcw0f.

[4] G. Bianchi and H. Egnell, A note on the Sobolev inequality, *J. Funct. Anal.* **100(1)** (1991), 18–24. DOI: 10.1016/0022-1236(91)90099-Q.

[5] D. L. Burkholder, Boundary value problems and sharp inequalities for martingale transforms, *Ann. Probab.* **12(3)** (1984), 647–702. DOI: 10.1214/AOP/1176993220.

[6] D. L. Burkholder, Explorations in martingale theory and its applications, in: *École d'Été de Probabilités de Saint-Flour XIX—1989*, Lecture Notes in Math. **1464**, Springer-Verlag, Berlin, 1991, pp. 1–66. DOI: 10.1007/BFb0085167.

[7] E. A. Carlen, Duality and stability for functional inequalities, *Ann. Fac. Sci. Toulouse Math. (6)* **26(2)** (2017), 319–350. DOI: 10.5802/afst.1535.

[8] S. Chen, R. L. Frank, and T. Weth, Remainder terms in the fractional Sobolev inequality, *Indiana Univ. Math. J.* **62(4)** (2013), 1381–1397. DOI: 10.1512/iumj.2013.62.5065.

[9] M. Christ, A sharpened Hausdorff–Young inequality, Preprint (2014). arXiv:1406.1210.

[10] J. Dolbeault and G. Toscani, Stability results for logarithmic Sobolev and Gagliardo–Nirenberg inequalities, *Int. Math. Res. Not. IMRN* **2016(2)** (2016), 473–498. DOI: 10.1093/imrn/rnv131.

[11] M. Fathi, E. Indrei, and M. Ledoux, Quantitative logarithmic Sobolev inequalities and stability estimates, *Discrete Contin. Dyn. Syst.* **36(12)** (2016), 6835–6853. DOI: 10.3934/dcds.2016097.

[12] S. E. Graversen and G. Peškir, On Doob's maximal inequality for Brownian motion, *Stochastic Process. Appl.* **69(1)** (1997), 111–125. DOI: 10.1016/S0304-4149(97)00032-X.

[13] P. Ivanisvili, N. N. Osipov, D. M. Stolyarov, V. I. Vasyunin, and P. B. Zatitskiy, Bellman function for extremal problems in BMO, *Trans. Amer. Math. Soc.* **368(5)** (2016), 3415–3468. DOI: 10.1090/tran/6460.

[14] P. Ivanisvili, D. M. Stolyarov, V. I. Vasyunin, and P. B. Zatitskiy, Bellman function for extremal problems in BMO II: Evolution, *Mem. Amer. Math. Soc.* **255(1220)** (2018), 133 pp. DOI: 10.1090/memo/1220.

[15] A. D. Melas, An eigenfunction stability estimate for approximate extremals of the Bellman function for the dyadic maximal operator on L^p, *Proc. Amer. Math. Soc.* **147(8)** (2019), 3367–3375. DOI: 10.1090/proc12740.

[16] B. Muckenhoupt, Weighted norm inequalities for the Hardy maximal function, *Trans. Amer. Math. Soc.* **165** (1972), 207–226. DOI: 10.2307/1995882.

[17] F. L. Nazarov and S. R. Treĭl', The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis (Russian), *Algebra i Analiz* **8(5)** (1996), 32–162; translation in *St. Petersburg Math. J.* **8(5)** (1997), 721–824.

[18] F. Nazarov, S. Treil, and A. Volberg, The Bellman functions and two-weight inequalities for Haar multipliers, *J. Amer. Math. Soc.* **12(4)** (1999), 909–928. DOI: 10.1090/S0894-0347-99-00310-0.

[19] A. Osękowski, *Sharp Martingale and Semimartingale Inequalities*, IMPAN Monogr. Mat. (N. S.) **72**, Birkhäuser/Springer Basel AG, Basel, 2012. DOI: 10.1007/978-3-0348-0370-0.

[20] A. Osękowski, Sharp $L^{p,\infty} \to L^q$ estimates for the dyadic-like maximal operators, *J. Fourier Anal. Appl.* **20(5)** (2014), 911–933. DOI: 10.1007/s00041-014-9338-1.

[21] A. Osękowski, Sharp weak type inequality for fractional integral operators associated with d-dimensional Walsh–Fourier series, *Integral Equations Operator Theory* **78(4)** (2014), 589–600. DOI: 10.1007/s00020-013-2116-7.

[22] A. Osękowski, Sharp weak-type estimates for the dyadic-like maximal operators, *Taiwanese J. Math.* **19(4)** (2015), 1031–1050. DOI: 10.11650/tjm.19.2015.3550.

[23] A. Osękowski, Best constants in Muckenhoupt's inequality, *Ann. Acad. Sci. Fenn. Math.* **42(2)** (2017), 889–904. DOI: 10.5186/aasfm.2017.4256.

[24] S. PETERMICHL AND A. VOLBERG, Heating of the Ahlfors–Beurling operator: weakly quasiregular maps on the plane are quasiregular, *Duke Math. J.* **112(2)** (2002), 281–305. DOI: 10.1215/S0012-9074-02-11223-X.

[25] S. PETERMICHL AND J. WITTWER, A sharp estimate for the weighted Hilbert transform via Bellman functions, *Michigan Math. J.* **50(1)** (2002), 71–87. DOI: 10.1307/mmj/1022636751.

[26] V. I. VASYUNIN, The sharp constant in the inverse Hölder inequality for Muckenhoupt weights (Russian), *Algebra i Analiz* **15(1)** (2003), 73–117; translation in: *St. Petersburg Math. J.* **15(1)** (2004), 49–79. DOI: 10.1090/S1061-0022-03-00802-1.

[27] V. VASYUNIN AND A. VOLBERG, Monge–Ampère equation and Bellman optimization of Carleson embedding theorems, in: *Linear and Complex Analysis*, Amer. Math. Soc. Transl. Ser. 2 **226**, Adv. Math. Sci. **63**, American Mathematical Society, Providence, RI, 2009, pp. 195–238. DOI: 10.1090/trans2/226/16.

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

E-mail address: l.kaminski5@student.uw.edu.pl

E-mail address: ados@mimuw.edu.pl

Received on September 26, 2023.

Accepted on June 14, 2024.