SHARP WEIGHTED WEAK-NORM ESTIMATES FOR MAXIMAL FUNCTIONS

MICHAŁ BRZOZOWSKI, ADAM OSĘKOWSKI, AND MATEUSZ RAPICKI

Abstract. For any $1 < p < \infty$ and any $c \geq 1$ we identify the least constant $C_{p,c}$ with the following property. If $X = (X_t)_{t\geq 0}$ is a uniformly integrable martingale and $W = (W_t)_{t\geq 0}$ is a weight satisfying Muckenhoupt's condition A_p with $[W]_{A_p} \leq c$, then we have the Lorentz-norm estimate

$$\left|\left|\sup_{t\geq 0} |X_t|\right|\right|_{L^{p,\infty}(W)} \leq C_{p,c}||X_\infty||_{L^{p,\infty}(W)}.$$

The proof exploits related sharp weak-type estimates and optimization arguments.

1. Introduction

The paper is devoted to the study of sharp weighted versions of the classical maximal estimates for real-valued martingales obtained by Doob [2]. Let us start with the necessary background, notation and the statement of related results. Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, filtered by $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing family of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all the events of probability 0. Let X be an adapted, real-valued, uniformly integrable martingale with right-continuous trajectories that have limits from the left; such a martingale converges almost surely to an integrable variable which will be denoted by X_∞. The maximal function of X is given by $X^* = \sup_{s\geq 0} |X_s|$ and the square bracket of X is denoted by $[X, X]$ (see e.g. Dellacherie and Meyer [1] for the definition). A classical result of Doob [2] asserts that the maximal function satisfies the weak-type (p, p) estimate

$$||X^*||_{L^{p,\infty}} \leq ||X_\infty||_{L^p}, \qquad 1 \leq p < \infty,$$

where $||X^*||_{L^{p,\infty}} = \sup_{\lambda>0} \left[\lambda^p \mathbb{P}(X^* \geq \lambda)\right]^{1/p}$ is the usual weak p-th norm of X^*. Furthermore, if $1 < p \leq \infty$, then we have the strong-type bound

$$||X^*||_{L^p} \leq \frac{p}{p-1}||X_\infty||_{L^p}.$$

Both estimates above are sharp: for any value of p, the constants 1 and $p/(p-1)$ cannot be improved. There is a related result proved by Osękowski in [8] which provides a sharp comparison of weak p-th norms of X and X^*: for any $1 < p < \infty$ we have

$$||X^*||_{L^{p,\infty}} \leq \frac{p}{p-1}||X_\infty||_{L^{p,\infty}}, \tag{1.1}$$

and the constant $p/(p-1)$ is again the best possible. See also [7] for a related sharp $L^{q,\infty} \to L^p$ bound.

2010 *Mathematics Subject Classification.* Primary: 42B25. Secondary: 60G44.
Key words and phrases. Maximal, martingale, Bellman function, best constants.

The primary goal of this paper is to study weighted version of the estimate (1.1). Here the word 'weight' will refer to a positive, uniformly integrable martingale $W = (W_t)_{t\geq 0}$. Such a process gives rise to a new (not necessarily probability) measure $W_\infty \mathrm{d}\mathbb{P}$. For technical reasons, we will also assume that W has continuous paths with probability 1. It is well-known that without any regularity assumptions on the trajectories of the weight almost all reasonable inequalities fail to hold (cf. the paper [4] for a related fact for BMO martingales).

When studying weighted L^p or weak-L^p estimates for maximal functions, one has to restrict oneself to the so-called A_p weights. Let us discuss this issue a little here. Assume that W is a given and fixed weight. Following Izumisawa and Kazamaki [3], we say that W satisfies Muckenhoupt's condition A_p (where $1 < p < \infty$ is a fixed parameter), if

$$[W]_{A_p} := \sup_\tau \left|\left|\mathbb{E}\Big[\big\{W_\tau/W_\infty\big\}^{1/(p-1)}\big|\mathcal{F}_\tau\Big]^{p-1}\right|\right|_{L^\infty} < \infty, \tag{1.2}$$

where the supremum is taken over the class of all adapted stopping times. It turns out that the weak-type estimate

$$||X^*||_{L^{p,\infty}(W)} \leq c_{p,[W]_{A_p}} ||X_\infty||_{L^p(W)} \tag{1.3}$$

holds for all martingales X (with some constant $c_{p,[W]_{A_p}}$ depending only on the parameters indicated) if and only if W satisfies (1.2). Here we use the notation $||X_\infty||_{L^p(W)} = \left(\mathbb{E}|X_\infty|^p W_\infty\right)^{1/p}$ and $||X^*||_{L^{p,\infty}(W)} = \sup_{\lambda>0} \lambda \left[W(X^* \geq \lambda)\right]^{1/p}$ for the weighted strong and weak weighted p-th norms of X (for $A \in \mathcal{F}$, we write $W(A) = \int_A W_\infty \mathrm{d}\mathbb{P}$). A similar phenomenon occurs in the context of strong type inequalities: the estimate

$$||X^*||_{L^p(W)} \leq C_{p,[W]_{A_p}} ||X_\infty||_{L^p(W)}$$

holds for all X with some $C_{p,[W]_{A_p}}$ independent of X if and only if W is an A_p weight. These results, proved by Izumisawa and Kazamaki [3], are in perfect correspondence with the classical theorems of Muckenhoupt concerning weighted inequalities for the Hardy-Littlewood maximal function on $\mathbb{R}^d$: cf. [5].

We will provide the proof of the weighted counterpart of (1.1). In fact, we will establish a much stronger result: we will identify the best constant involved in this weighted estimate. To describe this constant, we need to introduce some auxiliary parameters. For the geometric interpretation of these objects, we refer the reader to Figure 1 below. Let $c \geq 1$ and $1 < p < \infty$ be fixed. Then the line, tangent to the curve $x_1 x_2^{p-1} = c$ at the point $(1, c^{1/(p-1)})$, intersects the curve $x_1 x_2^{p-1} = 1$ at one point (if $c = 1$) or two points (if $c > 1$). Take the intersection point with larger x_1-coordinate, and denote this coordinate by $1 + d(p, c)$. Formally, $d = d(p, c)$ is the unique number in $[0, p-1)$ satisfying the equation

$$c(1 + d)(p - 1 - d)^{p-1} = (p - 1)^{p-1}. \tag{1.4}$$

We are ready to state the main result of the paper.

Theorem 1.1. *If $1 < p < \infty$ and W is an A_p weight, then for any martingale X we have the estimate*

$$||X^*||_{L^{p,\infty}(W)} \leq \frac{p}{p - 1 - d(p, [W]_{A_p})} ||X_\infty||_{L^{p,\infty}(W)}. \tag{1.5}$$

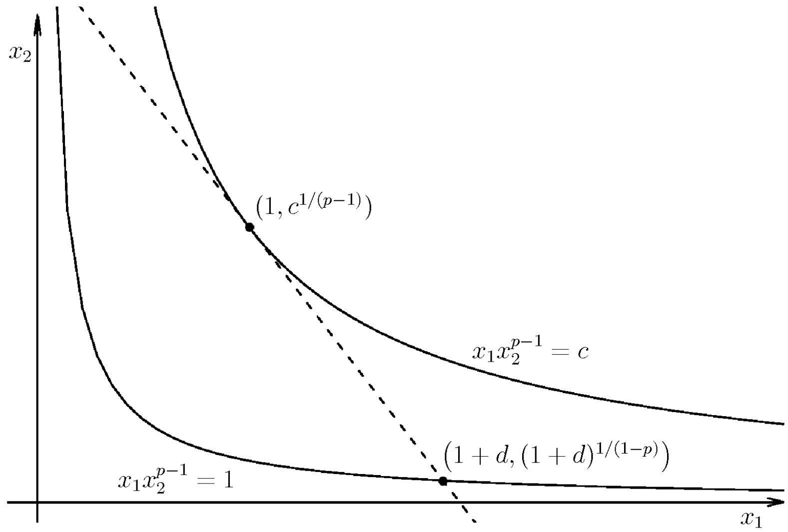

FIGURE 1. The geometric interpretation of the number $d = d(p, c)$.

The constant $p/(p - 1 - d(p, [W]_{A_p}))$ is the best possible: for any $\varepsilon > 0$, any $1 < p < \infty$ and any $c \geq 1$, there is an A_p weight W satisfying $[W]_{A_p} = c$ and a martingale X such that

$$||X^*||_{L^{p,\infty}(W)} > \left(\frac{p}{p-1-d(p,[W]_{A_p})} - \varepsilon\right) ||X_\infty||_{L^{p,\infty}(W)}.$$

The proof of the above result will rest on the sharp version of the weighted weak-type estimate (1.3). Such an estimate was proved in [6], but under more restrictive conditions on the martingale X (the continuity of paths was required). We establish the more general version in Section 2 below, using an appropriate sharp reverse Hölder inequality for A_p weights which is of independent interest. Theorem 1.1 is proved in Section 3.

2. A SHARP WEAK-TYPE ESTIMATE

2.1. **A special function.** We will work with a certain function constructed in [10]. Let us briefly recall the definition of this object. Consider the hyperbolic-type domain

$$\mathcal{D}_{p,c} = \{(x_1, x_2) \in \mathbb{R}_+^2 : 1 \leq x_1 x_2^{p-1} \leq c\},$$

foliated by the family of curves $\gamma_b = \{(x_1, x_2) \in \mathbb{R}_+^2 : x_1 x_2^{p-1} = b\}$, $1 \leq b \leq c$. Then $\partial^+\mathcal{D}_{p,c} = \gamma_c$ and $\partial^-\mathcal{D}_{p,c} = \gamma_1$ are the upper and the lower parts of the boundary of $\mathcal{D}_{p,c}$. We will also need the following geometrical object. For any $x > 0$, consider the line ℓ tangent to the upper boundary $\partial^+\mathcal{D}_{p,c}$, passing through the point $(x, (c/x)^{1/(p-1)})$. This line has the equation

$$x_2 = -\frac{c^{1/(p-1)} x^{-p/(p-1)}}{p-1} x_1 + \frac{p}{p-1}\left(\frac{c}{x}\right)^{1/(p-1)} \tag{2.1}$$

and intersects the lower boundary $\partial^-\mathcal{D}_{p,c}$ at two points. Take the point with the larger x_1-coordinate: since the point lies on $\partial^-\mathcal{D}_{p,c}$, its coordinates can be expressed in the form

$$(x(1+d),(x(1+d))^{1/(1-p)})$$

for some $d>0$. It is straightforward to check that d does not depend on x and hence (by taking $x=1$) it must be equal to $d(p,c)$ defined in (1.4). Let I_x be the line segment tangent to $\partial^+\mathcal{D}_{p,c}$ with the endpoints $(x,(c/x)^{1/(p-1)})$ and $(x(1+d),(x(1+d))^{1/(1-p)})$.

We are ready to introduce the special function from [10]. For a given $d+1<r<p$, consider $B=B_{p,r,c}:\mathcal{D}_{p,c}\to\mathbb{R}$ uniquely determined by the following three requirements:

(i) For any $(x_1,x_2)\in\partial^+\mathcal{D}_{p,c}$ we have

$$B(x_1,x_2)=\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}x_1^{1/(1-r)}.$$

(ii) For any $(x_1,x_2)\in\partial^-\mathcal{D}_{p,c}$ we have

$$B(x_1,x_2)=x_1^{1/(1-r)}.$$

(iii) The function B is linear along any line segment I_x, $x>0$.

It is not difficult to show that B is continuous and of class C^∞ in the interior of $\mathcal{D}_{p,c}$. In [10], the following further property of B was established.

Lemma 2.1. *For any p, c and $r\in(d(p,c)+1,p)$, the function $B_{p,r,c}$ is locally concave, i.e., concave along any line segment entirely contained in $\mathcal{D}_{p,c}$.*

We will also need the following majorization condition.

Lemma 2.2. *For any p, c and $r\in(d(p,c)+1,p)$, we have*

$$x_1^{1/(1-r)}\le B_{p,r,c}(x_1,x_2)\le\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}x_1^{1/(1-r)}. \tag{2.2}$$

Proof. For brevity, set $B=B_{p,r,c}$. Fix x_1 and let $y>0$, $z>0$ be chosen so that $(x_1,y)\in\partial^+\mathcal{D}_{p,c}$ and $(x_1,z)\in\partial^-\mathcal{D}_{p,c}$; then the desired inequality is equivalent to $B(x_1,z)\le B(x_1,x_2)\le B(x_1,y)$. Therefore, we will be done if we show that the function $t\mapsto B(x_1,t)$ is nondecreasing on the interval $\{t:(x_1,t)\in\mathcal{D}_{p,c}\}=[z;y]$. Pick $x_1/(1+d)<x<x_1$. Using the equation (2.1), we see that the point

$$P=\left(x_1,-\frac{c^{1/(p-1)}x^{-p/(p-1)}}{p-1}x_1+\frac{p}{p-1}\left(\frac{c}{x}\right)^{1/(p-1)}\right)$$

lies on I_x and hence, by the definition of $B_{p,r,c}$,

$$\begin{aligned}B(P)&=\frac{x_1-x}{dx}B\left(x(1+d),(x(1+d))^{1/(1-p)}\right)+\frac{x(1+d)-x_1}{dx}B\left(x,\left(\frac{c}{x}\right)^{1/(p-1)}\right)\\&=\frac{x_1-x}{dx}(x(1+d))^{1/(1-r)}+\frac{x(1+d)-x_1}{dx}\cdot\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}x^{1/(1-r)}.\end{aligned}$$

Differentiating this equality with respect to x and calculating a little bit, we get

$$B_{x_2}(P)\frac{p}{(p-1)^2}\left(\frac{c}{x}\right)^{1/(p-1)}\cdot\frac{x_1-x}{x^2}=\frac{r(x(1+d))^{1/(1-r)}}{(1-r)(d+1-r)}\cdot\frac{x_1-x}{x^2}.$$

(Here and below, B_{x_i} stands for the partial derivative of B with respect to the variable x_i, $i=1,2$. Similarly, $B_{x_ix_j}$ denotes the second-order derivative with respect to x_i and x_j, i, $j\in\{1,2\}$). This gives the claim, since taking all x ranging from $x_1/(1+d)$ to x_1, we get all points from the interior of $\mathcal{D}_{p,c}$ with the first coordinate equal to x_1. $\square$

2.2. **A reverse Hölder inequality for weights.** We start with the following useful interpretation of A_p weights in terms of appropriate two-dimensional martingales. Fix such a weight W and let $c=[W]_{A_p}$. Furthermore, let $V=(V_t)_{t\geq 0}$ be the martingale given by $V_t=\mathbb{E}(W_\infty^{1/(1-p)}|\mathcal{F}_t)$, $t\geq 0$. Note that Jensen's inequality implies $W_\tau V_\tau^{p-1}\geq 1$ almost surely for any stopping time τ; furthermore, the A_p condition is equivalent to the reverse bound

$$W_\tau V_\tau^{p-1}\leq c \qquad \text{with probability 1.}$$

In other words, an A_p weight of characteristic equal to c gives rise to a two-dimensional martingale (W,V) taking values in the domain $\mathcal{D}_{p,c}$. In addition, this martingale terminates at the lower boundary $\partial^-\mathcal{D}_{p,c}$: $W_\infty V_\infty^{p-1}=1$ almost surely. A nice feature is that this is a full characterization: given any martingale pair (W,V) (with W having continuous paths) taking values in $\mathcal{D}_{p,c}$ and terminating at $\partial^-\mathcal{D}_{p,c}$, one easily checks that its first coordinate is an A_p weight with $[W]_{A_p}\leq c$.

Equipped with the above interpretation, we are ready for the proof of the following statement.

Lemma 2.3. *Suppose that W is an A_p weight, put $c=[W]_{A_p}$ and let $d=d(p,c)$ be the positive constant given by (1.4). Then for $r\in(d+1,p)$ we have*

$$[W]_{A_r}\leq\left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1}.$$

Proof. Let $B=B_{p,r,c}$ and pick an arbitrary stopping time τ. Furthermore, for a given $\delta>0$, consider the stopping time

$$\sigma=\sigma_\delta=\inf\{t: W_t\leq\delta \text{ or } W_t+[W,W]_t+[V,V]_t\geq\delta^{-1}\}.$$

The process (W,V) takes values in the set $\mathcal{D}_{p,c}$ and terminates when reaching the lower boundary $\partial^-\mathcal{D}_{p,c}$. Thus, by the definition of σ, we see that on the set $\{\sigma>0\}$ the stopped process (W^σ,V^σ) is bounded. Since B is of class C^∞ in the interior of $\mathcal{D}_{p,c}$, we may apply Itô's formula to obtain that on $\{\sigma>0\}$,

$$B(W_\infty^\sigma,V_\infty^\sigma)=I_0+I_1+I_2/2,$$

where

$$\begin{aligned}
I_0&=B(W_\tau^\sigma,V_\tau^\sigma),\\
I_1&=\int_\tau^\infty B_{x_1}(W_s^\sigma,V_s^\sigma)\mathrm{d}W_s^\sigma+\int_\tau^\infty B_{x_2}(W_s^\sigma,V_s^\sigma)\mathrm{d}V_s^\sigma,\\
I_2&=\int_\tau^\infty B_{x_1x_1}(W_s^\sigma,V_s^\sigma)\mathrm{d}[W^\sigma,W^\sigma]_s+2\int_\tau^\infty B_{x_1x_2}(W_s^\sigma,V_s^\sigma)\mathrm{d}[W^\sigma,V^\sigma]_s\\
&\quad+\int_\tau^\infty B_{x_2x_2}(W_s^\sigma,V_s^\sigma)\mathrm{d}[V^\sigma,V^\sigma]_s.
\end{aligned}$$

By the definition of the stopping time σ and properties of stochastic integrals, the processes $\left(\int_0^u B_{x_1}(W_s^\sigma,V_s^\sigma)\mathrm{d}W_s^\sigma\right)_{u\geq 0}$ and $\left(\int_0^u B_{x_2}(W_s^\sigma,V_s^\sigma)\mathrm{d}V_s^\sigma\right)_{u\geq 0}$ are L^2-bounded martingales. Moreover, since B is locally concave, its Hessian matrix is nonpositive-definite and hence $I_2\leq 0$. Putting the above facts together, we get

$$\mathbb{E}\big[B(W_\infty^\sigma,V_\infty^\sigma)|\mathcal{F}_\tau\big]1_{\{\sigma>0\}}\leq B(W_\tau^\sigma,V_\tau^\sigma)1_{\{\sigma>0\}},$$

which combined with (2.2) yields

$$\mathbb{E}\big((W_\infty^\sigma)^{1/(1-r)}|\mathcal{F}_\tau\big)1_{\{\sigma>0\}} \leq \frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}(W_\tau^\sigma)^{1/(1-r)}1_{\{\sigma>0\}}.$$

Now recall that σ depends on δ and send this parameter to 0. Clearly, then $\sigma \to \infty$ and hence, by the path-continuity of W and Fatou's lemma, we obtain

$$\mathbb{E}\big(W^{1/(1-r)}|\mathcal{F}_\tau\big) \leq \frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}W_\tau^{1/(1-r)}. \tag{2.3}$$

This is precisely the assertion, since τ was arbitrary. □

The above lemma will enable us to establish the following fact (recall that for $A \in \mathcal{F}$ we write $W(A) = \int_A W_\infty \mathrm{d}\mathbb{P}$).

Theorem 2.4. *Let $1 < p < \infty$. Suppose that W is an A_p weight with $c = [W]_{A_p}$ and let $d = d(p,c)$. Then for any $r \in (d+1, p)$ and any cádlág martingale $(X_t)_{t\geq 0}$ we have*

$$W(X^* \geq 1) \leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} \mathbb{E}(|X_\infty|^r V_\infty^{1-r} 1_{\{X^*\geq 1\}}), \tag{2.4}$$

where $V_t = \mathbb{E}(W^{1/(1-r)}|\mathcal{F}_t)$ for $t \in [0,\infty]$.

Proof. Let $\varepsilon > 0$ and set $\tau = \inf\{t : |X_t| \geq 1-\varepsilon\}$. Our starting point is the inequality

$$1_{\{\tau<\infty\}} \leq \frac{|X_\tau|}{1-\varepsilon}1_{\{\tau<\infty\}} = \frac{1}{1-\varepsilon} \cdot V_\tau^{1-1/r} \cdot \left(|X_\tau|^r V_\tau^{1-r} 1_{\{\tau<\infty\}}\right)^{1/r}. \tag{2.5}$$

There are two factors on the right which must be handled appropriately. By (2.3), we see that

$$\begin{aligned} V_\tau^{1-1/r} = \mathbb{E}(V_\infty|\mathcal{F}_\tau)^{1-1/r} &= \mathbb{E}(W_\infty^{1/(1-r)}|\mathcal{F}_\tau)^{1-1/r} \\ &\leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{1-1/r} W_\tau^{-1/r}. \end{aligned}$$

To treat the term $|X_\tau|^r V_\tau^{1-r} 1_{\{\tau<\infty\}}$, observe that the function $(x,v) \mapsto |x|^r v^{1-r}$ is convex on $\mathbb{R}\times(0,\infty)$. Indeed, it is of class C^1, symmetric with respect to the variable x and its Hessian matrix at $(x,v) \in (0,\infty)\times(0,\infty)$, which is equal to

$$\begin{bmatrix} r(r-1)x^{r-2}v^{1-r} & r(1-r)x^{r-1}v^{-r} \\ r(1-r)x^{r-1}v^{-r} & r(r-1)x^r v^{-1-r} \end{bmatrix},$$

is nonnegative definite. Consequently, we may write

$$|X_\infty|^r V_\infty^{1-r} \geq |X_\tau|^r V_\tau^{1-r} + r|X_\tau|^{r-2}X_\tau V_\tau^{1-r}(X_\infty - X_\tau) + (1-r)|X_\tau|^r V_\tau^{-r}(V_\infty - V_\tau),$$

since the graph of a convex function lies above any tangent plane. Pick an arbitrary event $A \in \mathcal{F}_\tau$, a parameter $\delta > 0$ and put $A_\delta = A \cap \{|X_\tau| \leq \delta^{-1}, \delta \leq V_\tau \leq \delta^{-1}\}$. Then A_δ also belongs to $\mathcal{F}_\tau$ and the above inequality implies

$$\int_{A_\delta} |X_\infty|^r V_\infty^{1-r}\mathrm{d}\mathbb{P} \geq \int_{A_\delta} |X_\tau|^r V_\tau^{1-r}\mathrm{d}\mathbb{P},$$

since $\mathbb{E}(X_\infty - X_\tau|\mathcal{F}_\tau) = 0$ and $\mathbb{E}(V_\infty - V_\tau|\mathcal{F}_\tau) = 0$. The events A_δ increase as δ decreases and we have $A = \bigcup_{n=2}^\infty A_{1/n}$. Therefore, Lebesgue's monotone convergence theorem gives

$$\int_A |X_\infty|^r V_\infty^{1-r} \mathrm{d}\mathbb{P} \geq \int_A |X_\tau|^r V_\tau^{1-r} \mathrm{d}\mathbb{P}.$$

In particular, this implies that $|X_\tau|^r V_\tau^{1-r}$ is integrable (by taking $A = \Omega$) and $|X_\tau|^r V_\tau^{1-r} \leq \mathbb{E}(|X_\infty|^r V_\infty^{1-r}|\mathcal{F}_\tau)$. Plugging the above observations to (2.5), we obtain

$$1_{\{\tau<\infty\}} \leq \frac{1}{1-\varepsilon}\left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{1-1/r} \left(\mathbb{E}(|X_\infty|^r V_\infty^{1-r} 1_{\{\tau<\infty\}}|\mathcal{F}_\tau)\right)^{1/r} W_\tau^{-1/r},$$

or, equivalently,

$$W_\tau 1_{\{\tau<\infty\}} \leq \left(\frac{1}{1-\varepsilon}\right)^r \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} \mathbb{E}(|X_\infty|^r V_\infty^{1-r} 1_{\{\tau<\infty\}}|\mathcal{F}_\tau).$$

Integrating and using the equality $W_\tau = \mathbb{E}(W|\mathcal{F}_\tau)$, we get

$$W(\tau<\infty) \leq \left(\frac{1}{1-\varepsilon}\right)^r \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} \mathbb{E}(|X_\infty|^r V_\infty^{1-r} 1_{\{\tau<\infty\}}).$$

But $\{X^* \geq 1\} \subseteq \{\tau < \infty\}$ for any $\varepsilon > 0$ (where ε is the parameter appearing in the definition of τ). Furthermore, we see that $\bigcap_{\varepsilon>0}\{\tau<\infty\} \subseteq \{X^* \geq 1\}$. Therefore, letting $\varepsilon \to 0$ and applying Lebesgue's dominated convergence theorem, we obtain the assertion. □

3. Proof of Theorem 1.1

3.1. **Proof of** (1.5)**.** Let $1 < p < \infty$, take an A_p weight W, put $c = [W]_{A_p}$ and let $d = d(p,c)$ be given by (1.4). In addition, fix a martingale X as in the statement. Using (2.4), we obtain

$$W(X^* \geq 1) \leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} r\int_0^\infty \lambda^{r-1} W(X^* \geq 1, |X| \geq \lambda)\mathrm{d}\lambda.$$

Set $K = ||X||_{L^{p,\infty}(W)}$: then for any $\lambda > 0$ we have the inequality

$$W(X^* \geq 1,\, |X| \geq \lambda) \leq \min\{(K/\lambda)^p, W(X^* \geq 1)\}$$

and therefore, plugging it above, we get

$$\begin{aligned} &W(X^* \geq 1)\\ &\leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} r\left[\int_0^{\lambda_0} \lambda^{r-1} W(X^* \geq 1)\mathrm{d}\lambda + K^p \int_{\lambda_0}^\infty \lambda^{r-1-p}\mathrm{d}\lambda\right], \end{aligned}$$

where $\lambda_0 = KW(X^* \geq 1)^{-1/p}$. This, after some straightforward manipulations, is equivalent to

$$W(X^* \geq 1) \leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{r-1} \cdot \frac{p}{p-r} K^r W(X^* \geq 1)^{1-r/p},$$

or

$$\left[W(X^* \geq 1)\right]^{1/p} \leq \left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{1-1/r} \left(\frac{p}{p-r}\right)^{1/r} ||X||_{L^{p,\infty}(W)}.$$

The constant on the right, considered as a function of $r \in (d+1, p)$, attains its minimal value $p/(p-1-d)$ at the point $r = 1 + pd/(d+1)$. To show this, denote this constant by $\exp(G(r))$ and compute that

$$G(r) = -\log(1+d) + \left(1 - \frac{1}{r}\right)\log\left[\frac{(1+d)(1-r)}{d+1-r}\right] + \frac{1}{r}\log\frac{p}{p-r}$$

and

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

We easily check that $G'(1+pd/(d+1)) = 0$; furthermore, the function $H(r) = r^2G'(r)$ has the same sign as G' and

$$\begin{aligned} H'(r) &= \left(\frac{p}{(p-r)^2} - \frac{1}{p-r}\right) + \left(\frac{1}{r-1} - \frac{1}{r-d-1} + \frac{d(d+1)}{(d+1-r)^2}\right) \\ &= \frac{r}{(p-r)^2} + \frac{rd^2}{(r-1)(r-1-d)^2}. \end{aligned}$$

Clearly, H' is nonnegative, so H is increasing and hence G' is negative on $(d+1, 1+pd/(d+1))$ and positive on $(1+pd/(d+1), p)$. Therefore, the choice $r = 1+pd/(d+1)$ indeed minimizes the constant in the weak-norm estimate. It remains to check that for $r = 1 + pd/(d+1)$ we have

$$\left(\frac{(1+d)^{1/(1-r)}(1-r)}{d+1-r}\right)^{1-1/r}\left(\frac{p}{p-r}\right)^{1/r} = \frac{p}{p-1-d}.$$

This is just a matter of simple manipulations; we leave the straightforward calculation to the reader.

3.2. **Sharpness.** Now we will prove that the constant $p/(p-1-d(p,[W]_{A_p}))$ in (1.5) is indeed the best possible. The unweighted case, corresponding to $[W]_{A_p} = 1$, is simple. Fix $1 < p < \infty$, $\varepsilon > 0$ and let B be a one-dimensional Brownian motion starting from 1. If we set $\tau = \inf\{t : B_t \leq \big((p-1)/p + \varepsilon\big)B_t^*\}$, then τ is an $L^{p/2}$-integrable stopping time (cf. Peskir [9]). Consequently, the stopped process $X = B^\tau$ is a martingale converging in L^p and hence $||X_\infty||_{L^{p,\infty}} \leq ||X_\infty||_{L^p} < \infty$. On the other hand, we have $X_\infty = ((p-1)/p + \varepsilon)X^*$ almost surely, so

$$||X_\infty||_{L^{p,\infty}} = \left(\frac{p-1}{p} + \varepsilon\right)||X^*||_{L^{p,\infty}}$$

and the sharpness follows from the fact that ε is arbitrary.

If $c := [W]_{A_p} > 1$, the construction and the analysis of the appropriate counterexample is more elaborate. Some parts of the construction are taken from [6], but there are several crucial modifications, so we have decided to provide all details here. We split the reasoning into a few separate parts.

1° Special points in the first quadrant. Let, as usual, $d = d(p,c)$ be given by (1.4). For two given numbers b and c satisfying $1 < b < c$, draw three curves: $\gamma_1 = \{(x_1,x_2) \in \mathbb{R}_+^2 : x_1x_2^{p-1} = 1\}$, $\gamma_b = \{(x_1,x_2) \in \mathbb{R}_+^2 : x_1x_2^{p-1} = b\}$ and $\gamma_c = \{(x_1,x_2) \in \mathbb{R}_+^2 : x_1x_2^{p-1} = c\}$. Next, consider the line passing through $(1, c^{1/(p-1)})$, tangent to γ_c. This line intersects γ_b in two points: $P_0 = (x_{1+}, x_{2+})$ and $P_1 = (x_{1-}, x_{2-})$, where $x_{1+} > 1$ and $x_{1-} < 1$. Furthermore, it intersects γ_1 at a point $Z_1 = (1+d, (1+d)^{1/(1-p)})$ where d is given by (1.4). Next, construct inductively the sequences $(P_n)_{n\geq 2}$ and $(Z_n)_{n\geq 2}$ of points as follows. Having constructed P_{n-1} and

P_{n-2}, consider a line passing through P_{n-1}, tangent to γ_c, different from $P_{n-2}P_{n-1}$; this line intersects γ_b in P_{n-1} and yet another point, which we denote by P_n. Furthermore, let Z_n be the point of intersection of the line $P_{n-1}P_n$ with γ_1, having a bigger x-coordinate than P_n. We hope that the Figure 2 below clarifies the construction.

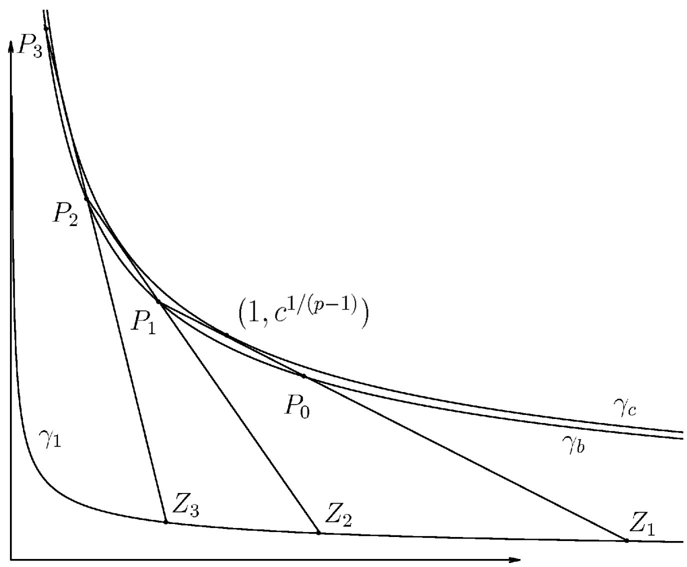

FIGURE 2. Special points $P_0 = (x_{1+}, x_{2+})$, $P_1 = (x_{1-}, x_{2-})$, P_2, $\ldots$ and $Z_1 = (1+d, (1+d)^{1/(1-p)})$, Z_2, Z_3, $\ldots$

It is clear that $x_{1\pm}$, $x_{2\pm}$ are functions of b, c and p; furthermore, if we keep c and p fixed, and let $b \uparrow c$, then $x_{1+} \downarrow 1$ and $x_{1-} \uparrow 1$, so in particular the difference $x_{1+} - x_{1-}$ converges to 0. Observe also that the picture has a self-similarity property. Clearly, for any $\lambda > 0$ we have $(x_1, x_2) \in \gamma_c$ if and only if $(\lambda x_1, \lambda^{1/(1-p)} x_2) \in \gamma_c$, and a similar equivalence holds for γ_b and γ_1. In consequence, for each $n \geq 0$ we have

$$P_n = \Big(x_{1+}(x_{1-}/x_{1+})^n, x_{2+}(x_{1+}/x_{1-})^{n/(p-1)} \Big)$$

and

$$Z_{n+1} = \Big((1+d)(x_{1-}/x_{1+})^n, (1+d)^{1/(1-p)}(x_{1+}/x_{1-})^{n/(p-1)} \Big).$$

In particular, this implies that for each $n \geq 0$ the point P_n splits the segment $P_{n+1}Z_{n+1}$ in the same ratio:

$$\frac{|Z_{n+1} - P_n|}{|Z_{n+1} - P_{n+1}|} = \frac{1 + d - x_{1+}}{1 + d - x_{1-}}. \tag{3.1}$$

2° Construction of the weight W. Introduce the two-dimensional continuous-path martingale (W, V) with distribution uniquely determined by the following requirements.

- W is a stopped Brownian motion,
- $(W_0, V_0) = P_0$ almost surely.

- The range of (W,V) is equal to the union of the segments P_nZ_n, $n = 1,2,\ldots$.

A more explicit description is in order. The process (W,V) starts from P_0 and first, it evolves along the line segment P_1Z_1, hitting eventually one of the endpoints. Denote

$$\tau_1 = \inf\{t : (W_t,V_t) \in \{P_1,Z_1\}\}.$$

If the ending point is Z_1, then the process (W,V) stops and we define its lifetime to be $\tau = \tau_1$. Otherwise, it continues its movement, but now it evolves along the line segment P_2Z_2, ending after some time in the set $\{P_2,Z_2\}$. Let

$$\tau_2 = \inf\{t : (W_t,V_t) \in \{P_2,Z_2\}\}.$$

If $(W_{\tau_2},V_{\tau_2}) = Z_2$, then the evolution is over and the lifetime τ of (W,V) equals τ_2. If $(W_{\tau_2},V_{\tau_2}) = P_2$, the process starts moving along P_3Z_3, and so on. Thus, we end up with a sequence $(\tau_n)_{n\geq 0}$ of stopping times (we set $\tau_0 \equiv 0$) and the lifetime variable $\tau = \sup_{n\geq 0}\tau_n$. Furthermore, directly from the self-similarity of the picture mentioned above (see (3.1)), we get

$$\mathbb{P}(\tau > \tau_n) = \left(\frac{1+d-x_{1+}}{1+d-x_{1-}}\right)^n, \tag{3.2}$$

so in particular τ is finite with probability 1 (since all τ_n's are). Finally, one easily checks that the pair (W,V) is uniformly integrable with values in $\{(w,v) : 1 \leq wv^{p-1} \leq c\}$, and hence W is an A_p weight satisfying $[W]_{A_p} \leq c$. Actually, the A_p characteristic is equal to c, since the trajectory of (W,V) touches the curve γ_c.

3° Construction of the martingale X. The process X will be an appropriate affine transformation of W. Let δ be a small positive number to be specified later. Define the points $\tilde{P}_n = (1+\delta)^n$, $\tilde{Z}_{n+1} = (1+\delta)^n(1-s)$ for $n = 0, 1, 2, \ldots$, where $s = \delta(1+d-x_{1+})/(x_{1+}-x_{1-})$. Then $\tilde{Z}_n < \tilde{P}_{n-1} < \tilde{P}_n$ for each $n \geq 1$ and

$$\frac{|P_{n-1}-Z_n|}{|P_n-Z_n|} = \frac{|\tilde{P}_{n-1}-\tilde{Z}_n|}{|\tilde{P}_n-\tilde{Z}_n|} = \frac{1+d-x_{1+}}{1+d-x_{1-}}. \tag{3.3}$$

We construct the martingale X separately on each interval $[\tau_n,\tau_{n+1}]$ (where τ_n's are the stopping times introduced in Step 2°). First, we let X start from $\tilde{P}_0$ and on the interval $[\tau_0,\tau_1]$, let it move along $[\tilde{Z}_1,\tilde{P}_1]$ so that $\tau_1 = \inf\{t : X_t \in \{\tilde{P}_1,\tilde{Z}_1\}\}$. Clearly, this is possible because of (3.3); actually, we may even require that

$$\{X_{\tau_1} = \tilde{P}_1\} = \{(W_{\tau_1},V_{\tau_1}) = P_1\} \quad \text{and} \quad \{X_{\tau_1} = \tilde{Z}_1\} = \{(W_{\tau_1},V_{\tau_1}) = Z_1\}.$$

Indeed, it suffices to put $X_t = \varphi(W_t)$, $t \in [\tau_0,\tau_1]$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is an affine mapping sending $1+d$ to $1-s$ and x_{1-} to $1+\delta$.

If $X_{\tau_1} = \tilde{Z}_1$, the process stops (and so does (W,V)); otherwise, on the set $\{\tau > \tau_1\}$, the movement is continued, along the segment $[\tilde{Z}_2,\tilde{P}_2]$ on the time interval $[\tau_1,\tau_2]$ so that $\tau_2 = \inf\{t > \tau_1 : X_t \in \{\tilde{P}_2,\tilde{Z}_2\}$ and

$$\{X_{\tau_2} = \tilde{P}_2\} = \{(W_{\tau_2},V_{\tau_2}) = P_2\} \quad \text{and} \quad \{X_{\tau_2} = \tilde{Z}_2\} = \{(W_{\tau_2},V_{\tau_2}) = Z_2\}.$$

This can be guaranteed due to (3.3): if φ is an affine mapping which sends $(1+d)x_{1-}/x_{1+}$ to $(1+\delta)(1-s)$ and x_{1-}^2/x_{1+} to $(1+\delta)^2$, then the formula $X_t = \varphi(W_t)$ on $\{\tau > \tau_1\}$, $t \in [\tau_1,\tau_2]$, gives us the desired process X. We continue the construction using this pattern on each $[\tau_n,\tau_{n+1}]$, requiring

$$\{X_{\tau_n} = \tilde{P}_n\} = \{(W_{\tau_n},V_{\tau_n}) = P_n\} \quad \text{and} \quad \{X_{\tau_n} = \tilde{Z}_n\} = \{(W_{\tau_n},V_{\tau_n}) = Z_n\}$$

for all $n \geq 1$. It is clear from the construction that X converges almost surely and satisfies

$$X_\infty = (1-s)X^*. \tag{3.4}$$

4° Calculation. Now we will check that for appropriate δ the variable X_∞ belongs to $L^p(W)$ (and hence also to $L^{p,\infty}(W)$). We have

$$\begin{aligned}\mathbb{E}|X_\infty|^p W_\infty &= \sum_{n=0}^\infty (1-s)^p(1+\delta)^{np}\cdot(1+d)\left(\frac{x_{1-}}{x_{1+}}\right)^n\cdot\left(\frac{1+d-x_{1+}}{1+d-x_{1-}}\right)^n\frac{x_{1+}-x_{1-}}{1+d-x_{1-}}\\ &= (1-s)^p(1+d)\cdot\frac{x_{1+}-x_{1-}}{1+d-x_{1-}}\sum_{n=0}^\infty\left[(1+\delta)^p\cdot\frac{x_{1-}}{x_{1+}}\cdot\frac{1+d-x_{1+}}{1+d-x_{1-}}\right]^n.\end{aligned}$$

If δ is chosen so that the expression in the square brackets is less than 1, then the above geometric series converges. So, take δ smaller than

$$\left(\frac{1+d-x_{1-}}{1+d-x_{1+}}\cdot\frac{x_{1+}}{x_{1-}}\right)^{1/p}-1$$

and let us look at the factor $1-s$ appearing in (3.4). If δ is sufficiently close to the above quantity, we see that $1-s$ can be made arbitrarily close to

$$1-\frac{\left(1+\frac{x_{1+}-x_{1-}}{1+d-x_{1+}}\right)^{1/p}\left(\frac{x_{1+}-x_{1-}}{x_{1-}}+1\right)^{1/p}-1}{x_{1+}-x_{1-}}(1+d-x_{1+}). \tag{3.5}$$

Recalling our discussion in 1°, if we let $b\to c$, then x_{1+}, x_{1-} converge to 1 and $x_{1+}-x_{1-}$ tends to 0. Consequently, the above expression converges to

$$1-\left(\frac{1}{pd}+\frac{1}{p}\right)d=\frac{p-d-1}{p}.$$

Therefore, if we choose first b sufficiently close to c and then δ sufficiently close to (3.5), then the factor $1-s$ can be made arbitrarily close to $(p-d-1)/p$ and, simultaneously, $X_\infty\in L^{p,\infty}(W)$. This shows that the constant $p/(p-1-d)$ is indeed the best possible.

Acknowledgments

The authors would like to thank anonymous referees for the careful reading of the first version of the paper and many helpful suggestions. The research was supported by Narodowe Centrum Nauki (Poland), grant no. DEC-2014/14/E/ST1/00532.

References

[1] C. Dellacherie and P.-A. Meyer, Probabilities and potential B: Theory of martingales, North Holland, Amsterdam, 1982.

[2] J. L. Doob, Stochastic processes, Reprint of the 1953 original, Wiley Classics Library, A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1990.

[3] M. Izumisawa and N. Kazamaki, Weighted norm inequalities for martingales, Tohoku Math. J. 29 (1977), 115–124.

[4] M. Izumisawa, T. Sekiguchi and Y. Shiota, *Remark on characterization of BMO-martingales*, Tohoku Math. J. **31** (1979), 281–284.

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

[6] A. Osękowski, *Weighted maximal inequalities for martingales*, Tohoku Math. J. **65** (2013), pp. 75–91.

[7] A. Osękowski, *Sharp $L^{p,\infty} \to L^q$ estimates for the dyadic-like maximal operators*, J. Fourier Anal. Appl. **20** (2014), pp. 911–933.

[8] A. Osękowski, *Sharp weak-type estimates for the dyadic-like maximal operators*, Taiwanese J. Math. **19** (2015), pp. 1031–1050.

[9] G. Peskir, *Optimal stopping of the maximum process: the maximality principle*, Ann. Probab. **26** (1998), 1614–1640.

[10] V. Vasyunin, *The exact constant in the inverse Hölder inequality for Muckenhoupt weights (Russian)*, Algebra i Analiz **15** (2003), 73–117; translation in St. Petersburg Math. J. **15** (2004), 49–79.

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

E-mail address: M.Brzozowski@mimuw.edu.pl

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

E-mail address: A.Osekowski@mimuw.edu.pl

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

E-mail address: M.Rapicki@mimuw.edu.pl