WEIGHTED INEQUALITIES FOR MARTINGALE TRANSFORMS AND STOCHASTIC INTEGRALS

ADAM OSĘKOWSKI

Abstract. The paper is devoted to the study of Fefferman-Stein inequalities for stochastic integrals. If X is a martingale, Y is the stochastic integral, with respect to X, of some predictable process taking values in $[-1,1]$, then for any weight W belonging to the class A_1 we have the estimates $||Y_\infty||_{L^p(W)} \leq 8pp'[W]_{A_1}||X_\infty||_{L^p(W)}$, $1<p<\infty$, and $||Y_\infty||_{L^{1,\infty}(W)} \leq c[W]_{A_1}(1+\log[W]_{A_1})||X_\infty||_{L^1(W)}$. The proofs rest on the Bellman function method: the inequalities are deduced from the existence of certain special functions, enjoying appropriate majorization and concavity. As an application, related statements for Haar multipliers are indicated. The above estimates can be regarded as probabilistic counterparts of the recent results of Lerner, Ombrosi and Pérez concerning singular integral operators.

1. Introduction

Suppose that w is a weight, i.e., a nonnegative, locally integrable function on $\mathbb{R}^d$, and let M stand for the Hardy-Littlewood maximal operator. In 1971, Fefferman and Stein [4] proved the existence of a finite constant c_d, depending only on the dimension, such that

$$w\left(\left\{x\in\mathbb{R}^d : Mf(x)\geq 1\right\}\right)\leq c_d||f||_{L^1(Mw)}.$$

Here and below, we use the standard notation $w(E)=\int_E w(x)\mathrm{d}x$ and $||f||_{L^p(w)}=\left(\int_{\mathbb{R}^d}|f(x)|^p w(x)\mathrm{d}x\right)^{1/p}$, $1\leq p<\infty$. This result led to the following natural conjecture, formulated by Muckenhoupt and Wheeden in the seventies. Namely, for any Calderón-Zygmund singular integral operator T, there is a constant $c_{T,d}$, depending only on T and d, such that

$$w\left(\left\{x\in\mathbb{R}^d : |Tf(x)|\geq 1\right\}\right)\leq c_{T,d}||f||_{L^1(Mw)}. \tag{1.1}$$

This problem admits a weaker version. Recall that w satisfies Muckenhoupt's condition A_1 if there is a constant $c>0$ such that $Mw\leq cw$ almost everywhere; the smallest c with this property is denoted by $[w]_{A_1}$ and called the A_1 characteristics of w. The weaker conjecture is the following: for any Calderón-Zygmund operator T there is a constant $c'_{T,d}$ such that for any A_1 weight w,

$$w\left(\left\{x\in\mathbb{R}^d : |Tf(x)|\geq 1\right\}\right)\leq c'_{T,d}[w]_{A_1}||f||_{L^1(w)}. \tag{1.2}$$

This statement is indeed weaker than the preceding: the validity of (1.1) would imply (1.2) with $c'_{T,d}=c_{T,d}$.

2010 *Mathematics Subject Classification.* Primary: 42B25. Secondary: 46E30, 60G44.

Key words and phrases. square function, maximal operator, dyadic, weight, Bellman function.

Research supported by the National Science Centre, Poland, grant DEC-2014/14/E/ST1/00532.

Both conjectures (1.1) and (1.2) remained open for a long time, and finally, a few years ago, they were both proved to be false: see the counterexamples for the Hilbert transform provided by Reguera, Thiele, Nazarov, Reznikov, Vasyunin and Volberg in [9, 15, 16]. In the mean-time, many partial or related results in this direction were obtained. In particular, Buckley [1] showed that the strong conjecture is true for the weights $w_\delta(x) = |x|^{-d(1-\delta)}$, $0 < \delta < 1$. Lerner, Ombrosi and Pérez established the following weaker versions of (1.2). In [7] they proved the existence of an absolute c such that

$$||Tf||_{L^{1,\infty}(w)} \leq c\varphi([w]_{A_1})||f||_{L^1(w)},$$

where $\varphi(t) = t(1+\log t)(1+\log^+ \log t)$; then, in [8], they improved the estimate to the form

$$||Tf||_{L^{1,\infty}(w)} \leq c[w]_{A_1}(1+\log[w]_{A_1})||f||_{L^1(w)}, \tag{1.3}$$

where, as above, c is a certain finite universal constant. To obtain these results, they first established the L^p estimates

$$||Tf||_{L^p(w)} \leq cpp'[w]_{A_1}||f||_{L^p(w)}, \qquad 1 < p < \infty,$$

in which the linear dependence on the A_1 characteristics is optimal. Then, using some extrapolation arguments, they managed to deduce the above weak-type bounds. The improvement in the appropriate LlogL term came from the fact that the L^p-constant c_p obtained in [8] is better than that in [7].

The purpose of this paper is to study analogs of (1.3) arising in the dyadic context of Haar multipliers. Actually, it will be more convenient for us to inspect the inequalities in an even wider, probabilistic setting. Let us introduce the necessary background and notation. We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$; in what follows, we will use the well-known property of this filtration that all adapted martingales have continuous trajectories. This rather restrictive assumption on the regularity of processes is typical in the context of weighted estimates (cf. Kazamaki [6]). Let $X = (X_t)_{t\geq 0}$ be an adapted uniformly integrable martingale taking values in $\mathbb{R}$. Denote by $[X] = ([X]_t)_{t\geq 0}$ the square bracket (quadratic variation) associated to X; see Dellacherie and Meyer [3] for the appropriate definition and basic properties of this object. Given a predictable process H, we will write $Y = H \cdot X$ for the stochastic integral of H with respect to X:

$$Y_t = H_0X_0 + \int_{0+}^t H_s \mathrm{d}X_s, \qquad t \geq 0.$$

Next, let W be a weight, i.e., a nonnegative, integrable random variable; this object gives rise to a continuous-path and uniformly integrable martingale $(W_t)_{t\geq 0}$ given by $W_t = \mathbb{E}(W|\mathcal{F}_t)$; sometimes we will identify W with W_∞. Let $W^* = \sup_{s\geq 0} W_s$, $W_t^* = \sup_{0\leq s\leq t} W_s$ be the maximal and truncated maximal functions of W. More generally, for any $r > 1$, we define the r-maximal function by $M_rW = \left[(W^r)^*\right]^{1/r}$. In analogy to the analytic context, a weight W is said to satisfy Muckenhoupt's A_1 condition, if there is an absolute constant c such that $W_t^* \leq cW_t$ almost surely for all t. The least c with this property is denoted by $[W]_{A_1}$ and called the A_1 characteristics of W. See Ikeda and Kazamaki [5] and Kazamaki [6] for more information on the subject.

We are ready to formulate the main results of this paper. As in the analytic setting, we define the weighted p-th norm of X by $||X||_{L^p(W)} = (\mathbb{E}|X_\infty|^p W_\infty)^{1/p}$ for all $1 \leq p < \infty$. Furthermore, given such a p, we let $p' = p/(p-1)$ denote its harmonic conjugate, with the usual convention $1' = +\infty$.

Theorem 1.1. *Suppose that W is a weight, X is a martingale and $Y = H \cdot X$, where H is a predictable process taking values in $[-1, 1]$.*

(i) For any $1 < r \leq 2$ we have

$$||Y||_{L^p(W)} \leq \frac{c_p(r)}{(r-1)^{1/p'}} ||X||_{L^p(M_r W)}, \qquad 1 < p < \infty, \tag{1.4}$$

where

$$c_p(r) = \begin{cases} 4^{1-2/p'}(p')^{1+1/p'} r^{1/p'} & \text{if } 1 < p < 2, \\ (2(p'+r-1)/(p'-1))^{1/p'} & \text{if } 2 \leq p < \infty \end{cases} \tag{1.5}$$
$$\leq 8pp'.$$

(ii) If W belongs to the class A_1, then

$$||Y||_{L^p(W)} \leq 16pp'[W]_{A_1} ||X||_{L^p(W)}. \tag{1.6}$$

The above estimates will allow us to obtain the following weak-type bound for probabilistic A_1 weights. We define the weak first norm of Y by the formula $||Y||_{L^{1,\infty}(W)} = \sup_{\lambda > 0} \lambda \mathbb{E}\big(1_{\{|Y_\infty| > \lambda\}} W_\infty\big)$.

Theorem 1.2. *There is an absolute constant c such that the following holds. If W is an A_1 weight, X is a martingale and $Y = H \cdot X$, where H is a predictable process with values in $[-1, 1]$, then*

$$||Y||_{L^{1,\infty}(W)} \leq c[W]_{A_1} \log(1 + [W]_{A_1}) ||X||_{L^1(W)}. \tag{1.7}$$

A few comments on the proof are in order. Lerner, Ombrosi and Pérez exploit extrapolation methods and the Rubio de Francia algorithm combined with boundedness properties of the maximal operator. Our passage from (1.4) and (1.6) to (1.7) also involves extrapolation-type argument, but the approach leading to (1.4) is completely different. We will exploit the so-called Bellman function method: the inequality (1.4) (or rather its dual version) will be deduced from the existence of a certain special function, enjoying appropriate majorization and concavity. This type of approach has turned out to be extremely efficient in the study of various problems in probability and harmonic analysis (cf. [2], [10], [11], [13], [14], [17], [18] to cite just a few works out of a huge literature on the subject). One of the novel aspects of this paper is that we will manage to apply the method directly to the estimate (1.4) (or rather its dual version); in many situations, Bellman-function proofs of weighted inequalities involve also some additional reduction or decomposition arguments. We believe that the special function we construct, or at least a part of it, will be useful in other related results arising in the area.

We will require a probabilistic version of the method, which is described in the next section. Section 3 contains the proof of (1.4) and then in Section 4 we show how the L^p bound leads to (1.6) and the weak-type estimate (1.7). The final part of the paper comes back to the analytic context: we show how the martingale results obtained in the preceding sections imply the corresponding statements for Haar multipliers (or rather more general objects arising in the setting of probability spaces equipped with a tree-like structure).

2. On the method of proof

Consider the domain

$$\mathfrak{D} = \{(x, y, u, v) \in \mathbb{R}^2 \times (0, \infty)^2 : u \leq v\}.$$

Suppose $V : \mathfrak{D} \to \mathbb{R}$ is a given Borel function and assume that we want to establish the inequality

$$\mathbb{E}V(X_t, Y_t, W_t, W_t^*) \leq 0, \qquad t \geq 0, \tag{2.1}$$

for all triples (X, Y, W), where X, Y are bounded martingales such that $Y = H \cdot X$ for some predictable process H taking values in $[-1, 1]$, and W is a bounded positive weight (recall that the underlying filtration is Brownian, and hence X, Y and W have continuous paths). To study this problem, consider the family $\mathcal{U}(V)$ which consists of all continuous functions $U : \mathfrak{D} \to \mathbb{R}$ satisfying the following properties:

1° For any x, $y \in \mathbb{R}$ such that $|y| \leq |x|$ and any $u > 0$ we have

$$U(x, y, u, u) \leq 0. \tag{2.2}$$

2° For any $(x, y, u, v) \in \mathfrak{D}$ we have

$$U(x, y, u, v) \geq V(x, y, u, v). \tag{2.3}$$

3° The function U is nonincreasing with respect to its fourth variable. That is, for any x, $y \in \mathbb{R}$ and any $u \leq v_1 \leq v_2$, we have

$$U(x, y, u, v_1) \geq U(x, y, u, v_2). \tag{2.4}$$

4° For any $(x, y, u, v) \in \mathfrak{D}$ and any h, k, $\ell \in \mathbb{R}$ satisfying $|k| \leq |h|$, the function $G = G_{x,y,u,v,h,k,\ell}$, given by

$$G(t) = U(x + th, y + tk, u + t\ell, v), \tag{2.5}$$

is concave on the interval $\{t : 0 < u + t\ell \leq v\}$.

Theorem 2.1. *If the class $\mathcal{U}(V)$ is nonempty, then the inequality (2.1) holds true.*

Proof. Essentially, the argument rests on Itô's formula applied to the function U and the process (X, Y, W, W^*). However, the function U need not have the necessary regularity, which forces us to use an additional mollification. Let $g : \mathbb{R}^4 \to [0, \infty)$ be a C^∞ function, supported on the unit ball of $\mathbb{R}^4$ and satisfying $\int_{\mathbb{R}^4} g \mathrm{d}x = 1$. Given $\varepsilon > 0$ and a number $\delta \in (0, \varepsilon)$, define a function $U^\delta : \{(x, y, u, v) \in \mathfrak{D} : \delta \leq u \leq v - 2\delta\} \to \mathbb{R}$ by the convolution

$$U^\delta(x, y, u, v) = \int_{[-1,1]^4} U(x+\delta s_1, y+\delta s_2, u+\delta s_3, v+\delta s_4) g(s_1, s_2, s_3, s_4) \mathrm{d}s_1 \mathrm{d}s_2 \mathrm{d}s_3 \mathrm{d}s_4.$$

This function is of class C^∞ and inherits the properties 3° and 4° (with some obvious modifications: the inequality (2.4) holds for $u + 2\delta \leq v_1 \leq v_2$; the function G appearing in 4° is concave on $\{t : \delta \leq u + t\ell \leq v - 2\delta\}$). Introduce the process $Z_t^\varepsilon = (X_t, Y_t, W_t + \varepsilon, W_t^* + 3\varepsilon)$, $t \geq 0$. This process takes values in the domain of U^δ, so the composition $U^\delta(Z_t^\varepsilon)$ makes sense. So, if we fix $t \geq 0$ and apply Itô's formula, we get

$$U^\delta(Z_t^\varepsilon) = U^\delta(Z_0^\varepsilon) + I_1 + I_2 + I_3/2, \tag{2.6}$$

where

$$
\begin{aligned}
I_1 &= \int_0^t U_x^\delta(Z_s^\varepsilon)\mathrm{d}X_s + \int_0^t U_y^\delta(Z_s^\varepsilon)\mathrm{d}Y_s + \int_0^t U_u^\delta(Z_s^\varepsilon)\mathrm{d}W_s,\\
I_2 &= \int_0^t U_v^\delta(Z_s^\varepsilon)\mathrm{d}W_s^*,\\
I_3 &= \int_0^t D^2_{xyu}U^\delta(Z_s^\varepsilon)\mathrm{d}[X,Y,W]_s.
\end{aligned}
$$

Here the integral in I_3 is the shortened notation for the sum of all the second-order terms, i.e.,

$$
I_3 = \int_0^t U_{xx}^\delta(Z_s^\varepsilon)\mathrm{d}[X,X]_s + 2\int_0^t U_{xy}^\delta(Z_s^\varepsilon)\mathrm{d}[X,Y]_s + \dots.
$$

Note that the process W^* and partial derivatives with respect to v do not appear in I_3: this is due to the fact that W^* is a process of bounded variation and hence any square bracket involving it must vanish.

Let us look at the terms I_1, I_2 and I_3 separately. By properties of stochastic integrals, the first term is a mean-zero martingale, since X, Y and W are assumed to be bounded. The second term is nonpositive, by 3° and the fact that the process W^* is nondecreasing. Finally, I_3 is also nonpositive, which follows directly from 4°. Indeed, pick an arbitrary partition $0 = s_0 < s_1 < s_2 < \ldots < s_N = t$ and an integer $i \in \{0, 1, \ldots, N-1\}$. Next, set $\Delta X = X_{s_{i+1}} - X_{s_i}$, $\Delta Y = H_{s_i}\Delta X$, $\Delta W = W_{s_{i+1}} - W_{s_i}$ and note that

$$
\begin{aligned}
0 &\geq \frac{d^2}{dt^2}U(X_{s_i} + t\Delta X, Y_{s_i} + t\Delta Y, W_{s_i} + t\Delta W + \varepsilon, W^*_{s_i} + 2\varepsilon)\Big|_{t=0}\\
&= \Big\langle D^2_{xyu}U(X_{s_i}, Y_{s_i}, W_{s_i} + \varepsilon, W^*_{s_i} + 2\varepsilon)(\Delta X, \Delta Y, \Delta W), (\Delta X, \Delta Y, \Delta W)\Big\rangle.
\end{aligned}
$$

It remains to sum these inequalities over i and let the diameter of the partition (s_i) go to zero; as the result, one obtains $I_3 \leq 0$.

Putting all the above facts together and taking the expectation in (2.6), we obtain

$$
\mathbb{E}U^\delta(Z_t^\varepsilon) \leq \mathbb{E}U^\delta(Z_0^\varepsilon).
$$

Now we carry out an appropriate two-step limiting procedure. First, recall that U is continuous, and hence $U^\delta \to U$ pointwise as $\delta \to 0$. Therefore, if we combine this with the boundedness of the process Z^ε, we obtain

$$
\mathbb{E}U(Z_t^\varepsilon) \leq \mathbb{E}U(Z_0^\varepsilon),
$$

in the light of Lebesgue's dominated convergence theorem. Next, we let $\varepsilon \to 0$ and use the (uniform) boundedness of Z^ε and continuity of U to obtain

$$
\mathbb{E}U(X_t, Y_t, W_t, W_t^*) \leq \mathbb{E}U(X_0, Y_0, W_0, W_0^*).
$$

However, we have $|Y_0| = |H_0X_0| \leq |X_0|$ and $W_0^* = W_0$; thus, by (2.2), the right-hand side above is nonpositive. It remains to apply the majorization (2.3) to get the desired estimate (2.1). □

Remark 2.2. There is a question which arises naturally in all contexts exploiting Bellman functions. Namely, suppose that we know a priori that the inequality (2.1) holds true. Can it be proved with the use of the above method, i.e., is there a function U possessing all the required properties? The positive answer would enable, in particular, the search for the optimal constants involved in the estimates. It is not difficult to prove that the answer is indeed positive, at least if we impose some additional structural properties on V. Suppose that V is continuous and nonincreasing with respect to its fourth variable. Define $U : \mathfrak{D} \to \mathbb{R}$ by the abstract formula

$$U(x, y, u, v) = \sup \mathbb{E} V(X_t, Y_t, W_t, \max\{W_t^*, v\}),$$

where the supremum is taken over all bounded martingales X starting from x, all bounded martingales Y satisfying

$$Y_t = y + \int_{0+}^t H_s \mathrm{d} X_s,$$

for some predictable process H bounded in absolute value by 1, and all bounded weights W with $W_0 = u$. One shows that U satisfies all the required conditions with the standard use of Markov property.

Unfortunately, we have been unable to identify the "optimal" Bellman functions U mentioned in the remark above (i.e., leading to sharp version of (1.4)). Still, we have managed to find functions which yield constants of optimal order: as we shall see, these special objects have quite a complicated structure and their analysis will require a certain amount of work.

3. L^p ESTIMATES

Now we will show how the above method can be used to yield the weighted L^p estimates for martingales, announced in the introduction.

3.1. **Dual bounds.**

Actually, it will be more convenient for us to establish a dual estimate to (1.4). Let us formulate the result in a separate statement.

Theorem 3.1. *Suppose that W is a weight, X is a martingale and $Y = H \cdot X$, where H is a predictable process with values in $[-1, 1]$. Then*

$$\left|\left| \frac{Y_\infty}{M_r W} \right|\right|_{L^p(M_r W)} \leq \frac{c_{p'}(r)}{(r-1)^{1/p}} \left|\left| \frac{X_\infty}{W_\infty} \right|\right|_{L^p(W)}, \qquad 1 < p < \infty, \tag{3.1}$$

where $c_p(r)$ is given in (1.5).

To see that the above statement is indeed dual to (1.4), let ζ be an arbitrary random variable belonging to $L^{p'}$ and let $V = (\mathbb{E}(W_\infty^{1/p} \zeta | \mathcal{F}_t))_{t \geq 0}$ be the martingale generated by the variable $W_\infty^{1/p} \zeta$ (which is integrable, by Hölder's inequality). We have, by basic properties of stochastic integrals,

$$\begin{aligned}\mathbb{E}Y_\infty W_\infty^{1/p}\zeta &= \mathbb{E}(H\cdot X)_\infty V_\infty\\ &= \mathbb{E}X_\infty (H\cdot V)_\infty\\ &= \mathbb{E}X_\infty (M_rW)^{1/p}\cdot (H\cdot V)_\infty (M_rW)^{-1/p}\\ &\leq ||X||_{L^p(M_rW)}||(H\cdot V)_\infty (M_rW)^{-1/p}||_{L^{p'}}\\ &= ||X||_{L^p(M_rW)}\left|\left|\frac{(H\cdot V)_\infty}{M_rW}\right|\right|_{L^{p'}(M_rW)}\\ &\leq ||X||_{L^p(M_rW)}\cdot \frac{c_p(r)}{(r-1)^{1/p'}}\left|\left|\frac{V_\infty}{W}\right|\right|_{L^{p'}(W)}\\ &= \frac{c_p(r)}{(r-1)^{1/p'}}||X||_{L^p(M_rW)}||\zeta||_{L^{p'}}.\end{aligned}$$

Since $\zeta\in L^{p'}$ was arbitrary, (1.4) follows.

In the next two subsections, we prove the validity of (3.1) for $p<2$ and $p\geq 2$. Before we do this, let us quickly show that $c_p(r)\leq 8pp'$, which we have declared in (1.5) above. If $1<p<2$, we have

$$\frac{c_p(r)}{4pp'}=\frac{(p'r/16)^{1/p'}}{p}\leq\frac{(p'/8)^{1/p'}}{p}\leq 2,$$

since $x^{1/x}\leq 2$ for any $x>0$. On the other hand, if $p\geq 2$, then

$$c_p(r)\leq\frac{2(p'+1)}{p'-1}=\frac{2(p'+1)}{(p')^2}\cdot pp'\leq 4pp'.$$

3.2. **The case $1<p\leq 2$.** Since $M_rW=[(W^r)^*]^{1/r}$, the inequality (3.1) can be rewritten in the form

$$\mathbb{E}|Y_\infty|^p[(W^r)^*]^{(1-p)/r}\leq\frac{2(p+r-1)}{(r-1)(p-1)}\mathbb{E}|X_\infty|^pW_\infty^{1-p},$$

or, if we introduce an auxiliary weight $\tilde{W}=W^r$ and set $\alpha=1/r$,

$$\mathbb{E}|Y_\infty|^p[\tilde{W}^*]^{-\alpha(p-1)}\leq\frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}\mathbb{E}|X_\infty|^p\tilde{W}_\infty^{-\alpha(p-1)}. \tag{3.2}$$

For convenience, let us write W instead of $\tilde{W}$; then we see that the above inequality is of the form (2.1), with

$$V_p(x,y,u,v)=|y|^pv^{-\alpha(p-1)}-\frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}|x|^pu^{-\alpha(p-1)}.$$

The special function $U_p:\mathfrak{D}\to\mathbb{R}$ is given by the formula

$$U_p(x,y,u,v)=(x^2+y^2)^{p/2}v^{-\alpha(p-1)}-\frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}|x|^pu^{-\alpha(p-1)}.$$

Let us now verify that this special object enjoys all the required properties.

Lemma 3.2. *The function U_p belongs to the class $\mathcal{U}(V_p)$.*

Proof. The property 1° is equivalent to the estimate

$$2^{p/2}-\frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}\leq 0,$$

which is evident: we have

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

The majorization 2° and the property 3° are trivial. The main technical difficulty lies in checking the concavity condition 4°. Fix x, y, u, v, h, k, ℓ as in the statement of this condition; since U is of class C^1, it is enough to check that $G''(t) \leq 0$ for all t such that $x + th \neq 0$. Actually, since

$$G_{x,y,u,v,h,k,\ell}(t+s) = G_{x+th,y+tk,u+t\ell,v,h,k}(s),$$

it suffices to verify the estimate $G''(0) \leq 0$. The function G is a difference of two terms; we study the contribution of each term separately. We compute that

$$\frac{d^2}{dt^2}\big((x+th)^2 + (y+tk)^2\big)^{p/2}\bigg|_{t=0} = K_1 + K_2 + K_3,$$

where

$$\begin{aligned} K_1 &= p(x^2+y^2)^{p/2-2}((p-1)x^2+y^2)h^2, \\ K_2 &= 2p(p-2)xy(x^2+y^2)^{p/2-2}hk \leq p(2-p)(x^2+y^2)^{p/2-1}h^2, \\ K_3 &= p(x^2+y^2)^{p/2-2}(x^2+(p-1)y^2)k^2 \leq p(x^2+y^2)^{p/2-2}(x^2+(p-1)y^2)h^2. \end{aligned}$$

This implies

$$\frac{d^2}{dt^2}\big((x+th)^2+(y+tk)^2\big)^{p/2} v^{-\alpha(p-1)}\bigg|_{t=0} \leq 2p(x^2+y^2)^{p/2-1}u^{-\alpha(p-1)}h^2. \tag{3.3}$$

Next, if $x \neq 0$, we derive that

$$\frac{d^2}{dt^2}\big(|x+th|^p(u+t\ell)^{-\alpha(p-1)}\big)\bigg|_{t=0} = \big\langle A(h,\ell),(h,\ell)\big\rangle,$$

where the matrix A is given by

$$\begin{bmatrix} p(p-1)|x|^{p-2}u^{-\alpha(p-1)} & -\alpha p(p-1)|x|^{p-1}u^{-\alpha(p-1)-1}\operatorname{sgn} x \\ -\alpha p(p-1)|x|^{p-1}u^{-\alpha(p-1)-1}\operatorname{sgn} x & \alpha(p-1)(\alpha(p-1)+1)|x|^p u^{-\alpha(p-1)-2} \end{bmatrix}$$

We easily check that A is nonnegative-definite; actually, this will still be true if we multiply the number in the upper-left corner by

$$\gamma = \frac{\alpha p}{\alpha(p-1)+1}$$

(after this multiplication, the determinant of the resulting matrix vanishes). Therefore, we have

$$\begin{aligned} \frac{d^2}{dt^2}\big(|x+th|^p(u+t\ell)^{-\alpha(p-1)}\big)\bigg|_{t=0} &\geq (1-\gamma)p(p-1)|x|^{p-2}u^{-\alpha(p-1)}h^2 \\ &= \frac{p(p-1)(1-\alpha)}{\alpha(p-1)+1}|x|^{p-2}u^{-\alpha(p-1)}h^2. \end{aligned} \tag{3.4}$$

Combining this inequality with (3.3), we obtain the desired bound $G''(0) \leq 0$. This proves the claim. □

Proof of (3.1). As we have already observed above, it is enough to establish (3.2). The method described in Section 2 allows to deduce the inequality for *bounded* processes, so we need some straightforward auxiliary limiting arguments to get the claim in full generality. Take an arbitrary continuous-path, uniformly integrable martingale X and let $Y = H\cdot X$ for some predictable process H with values in $[-1,1]$. Furthermore, let W be an arbitrary nonnegative weight. For a given integer N, consider the stopping time

$$\tau_N = \{t\geq 0 : |X_t| + |Y_t| + W_t \geq N\}.$$

Apply the method of Section 2 to the bounded martingales $X_t^{(N)} = X_{\tau_N\wedge t}1_{\{\tau_N>0\}}$, $Y_t^{(N)} = Y_{\tau_N\wedge t}1_{\{\tau_N>0\}}$ and $W_t^{(N)} = N^{-1}+W_{\tau_N\wedge t}1_{\{\tau_N>0\}}$; note that $Y^{(N)} = H\cdot X^{(N)}$ with unchanged process H. We get that for each $t\geq 0$,

$$\mathbb{E}|Y^{(N)}|^p(W^{(N)*})^{-\alpha(p-1)} \leq \frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}\mathbb{E}|X^{(N)}|^p(W^{(N)})^{-\alpha(p-1)}. \tag{3.5}$$

Clearly, we have

$$\mathbb{E}|Y^{(N)}|^p(N^{-1}+W^*)^{-\alpha(p-1)} \leq \mathbb{E}|Y^{(N)}|^p(W^{(N)*})^{-\alpha(p-1)}.$$

Furthermore, the function $(x,u)\mapsto |x|^p u^{-\alpha(p-1)}$ is convex on $\mathbb{R}\times(0,\infty)$: this follows at once from the fact that the matrix A, considered above in the proof of the property 4°, is nonnegative-definite. Moreover, we have $(X_{\tau_N\wedge t}, W_{\tau_N\wedge t}) = \mathbb{E}\big[(X_\infty, W_\infty)|\mathcal{F}_{\tau_N\wedge t}\big]$; hence, by conditional Jensen's inequality, we get

$$\mathbb{E}|X^{(N)}|^p(W^{(N)})^{-\alpha(p-1)} \leq \mathbb{E}|X_\infty|^p(N^{-1}+W_\infty)^{-\alpha(p-1)}1_{\{\tau_N>0\}}.$$

Combining the above two observations with (3.5) gives

$$\mathbb{E}|Y^{(N)}|^p(N^{-1}+W^*)^{-\alpha(p-1)} \leq \frac{2(\alpha(p-1)+1)}{(1-\alpha)(p-1)}\mathbb{E}|X_\infty|^p(N^{-1}+W_\infty)^{-\alpha(p-1)}1_{\{\tau_N>0\}}$$

and it remains to let $N\to\infty$ to obtain the assertion, by virtue of Lebesgue's monotone convergence theorem and Fatou's lemma. □

3.3. **The case $p>2$.** Here the calculations will be more involved. Arguing as in the preceding case, we see that (3.1) is equivalent to the following version of (3.2):

$$\mathbb{E}|Y_\infty|^p[W^*]^{-\alpha(p-1)} \leq \frac{4^{p-2}p^{p+1}}{1-\alpha}\mathbb{E}|X_\infty|^p W_\infty^{-\alpha(p-1)}.$$

This estimate is of the form (2.1), with $V = V_p : \mathfrak{D}\to\mathbb{R}$ given by

$$V_p(x,y,u,v) = |y|^p v^{-\alpha(p-1)} - \frac{4^{p-2}p^{p+1}}{1-\alpha}|x|^p u^{-\alpha(p-1)}.$$

The special function $U_p : \mathfrak{D}\to\mathbb{R}$ is defined by

$$U_p(x,y,u,v) = 2|y|^p v^{-\alpha(p-1)} - 2\cdot\frac{4^{p-2}p^{p+1}}{1-\alpha}|x|^p u^{-\alpha(p-1)}$$

if $|y|\leq 4p|x|$, and

$$U_p(x,y,u,v) = 2\big(|y|^p - p^2x^2|y|^{p-2} + 4^{p-2}p^p|x|^p\big)v^{-\alpha(p-1)} - 2\cdot\frac{4^{p-2}p^{p+1}}{1-\alpha}|x|^p u^{-\alpha(p-1)}$$

otherwise. Actually, one can write the definition with the use of a single formula:
(3.6)

$$U_p(x,y,u,v) = 2\min\left\{|y|^p, |y|^p - p^2x^2|y|^{p-2} + 4^{p-2}p^p|x|^p\right\}v^{-\alpha(p-1)} - 2\cdot\frac{4^{p-2}p^{p+1}}{1-\alpha}|x|^pu^{-\alpha(p-1)}.$$

Now we will prove that U_p, V_p enjoy the required properties.

Lemma 3.3. *The function U_p belongs to the class $\mathcal{U}(V_p)$.*

Proof. It is clear that U_p is continuous. The inequality (2.2) is equivalent to

$$1 - 4^{p-2}p^{p+1}/(1-\alpha) \leq 0,$$

which is evident. The proof of the majorization 2° is also easy: it suffices to add the trivial inequalities

$$2\min\left\{|y|^p, |y|^p - p^2x^2|y|^{p-2} + 4^{p-2}p^p|x|^p\right\}v^{-\alpha(p-1)} \geq |y|^pv^{-\alpha(p-1)}$$

and

$$\frac{4^pp^{p+1}}{1-\alpha}|x|^pu^{-\alpha(p-1)} \geq 2\cdot\frac{4^{p-2}p^{p+1}}{1-\alpha}|x|^pu^{-\alpha(p-1)}$$

to obtain (2.3). The property 3° holds trivially. So, it remains to establish the concavity condition 4°, which will be done in a similar fashion as in the case $1 < p \leq 2$. First, note that we may use (3.4): this inequality holds true in the full range $1 < p < \infty$. Therefore, if $|y| < 4p|x|$, then

$$\begin{aligned} &G''_{x,y,u,v,h,k,\ell}(0)/2 \\ &\leq p(p-1)|y|^{p-2}k^2v^{-\alpha(p-1)} - \frac{4^{p-2}p^{p+1}}{1-\alpha}\cdot\frac{p(p-1)(1-\alpha)}{\alpha(p-1)+1}|x|^{p-2}u^{-\alpha(p-1)}h^2 \\ &\leq p(p-1)(4p|x|)^{p-2}u^{-\alpha(p-1)}h^2\left(1 - \frac{p^3}{\alpha(p-1)+1}\right) \leq 0. \end{aligned}$$

So, suppose that $|y| > 4p|x| > 0$ and set $N(x,y) = |y|^p - p^2x^2|y|^{p-2} + 4^{p-2}p^p|x|^p$. We compute that

$$N_{xx}(x,y)h^2 = \big(-2p^2|y|^{p-2} + 4^{p-2}p^p\cdot p(p-1)|x|^{p-2}\big)h^2,$$
$$N_{xy}(x,y)hk = -2p^2(p-2)x|y|^{p-4}yhk \leq 2p^2(p-2)|x||y|^{p-3}h^2$$

and

$$\begin{aligned} N_{yy}(x,y)k^2 &= \big[p(p-1)|y|^{p-2} - p^2(p-2)(p-3)|y|^{p-4}x^2\big]k^2 \\ &\leq \big[p(p-1)|y|^{p-2} - p^2(p-2)(p-3)|y|^{p-4}x^2\big]h^2 \end{aligned}$$

(clearly, the expression in the square brackets is nonnegative). Combining these estimates with (3.4), we get

$$G''_{x,y,u,v,h,k,\ell}(0)/2 \leq J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &= 4^{p-2}p^p\cdot p(p-1)|x|^{p-2}\left[v^{-\alpha(p-1)} - \frac{p}{\alpha(p-1)+1}u^{-\alpha(p-1)}\right]h^2 \leq 0, \\ J_2 &= p(p-2)|y|^{p-3}(-|y| + 4p|x|)v^{-\alpha(p-1)}h^2 \leq 0, \\ J_3 &= p|y|^{p-4}\big(-3y^2 - p(p-2)(p-3)x^2\big)v^{-\alpha(p-1)}h^2 \leq 0. \end{aligned}$$

So, we have checked 4° ,,locally", i.e., we have proved the concavity of G on the set $\{t : x + th \neq 0, |y + tk| \neq 4p|x + th|\}$. To obtain the concavity on the whole domain, it suffices to use the fact that U_p is of class C^1 on $\mathfrak{D} \setminus \{|y| = 4p|x|\}$ (which handles the troublesome points t for which $x+th = 0$) and use (3.6) (which guarantees that for $|y| = 4p|x|$, the one-sided derivatives of G behave appropriately: $G'(0+) \leq G'(0-)$).

This completes the proof of the lemma. □

Proof of (3.1), $p > 2$. The argument is word-by-word the same as for $p \leq 2$. □

4. Inequalities for A_1 weights

Now we will deduce the validity of (1.6) and (1.7). We start with an auxiliary result, which is of independent interest.

Lemma 4.1. *Suppose that W is an A_1 weight. Then for any $1 < r < (1-[W]_{A_1}^{-1})^{-1}$, we have*

$$M_r W \leq (1 + r([W]_{A_1}^{-1} - 1))^{-1/r} W$$

almost surely. In particular, if $r = 1 + (2[W]_{A_1})^{-1}$, then

$$M_r W \leq 2[W]_{A_1} W$$

with probability 1.

Proof. We will again use the method of special functions; this time the argument will be more direct. Set $c = [W]_{A_1}$ and consider the function $U : [0,\infty)^2 \to \mathbb{R}$ given by the expression

$$U(x, y) = \frac{(rx - (r-1)y)y^{r-1}}{(r - (r-1)c)c^{r-1}}.$$

Clearly, this function is of class C^∞ and satisfies $U_{xx}(x, y) = 0$, $U_y(y, y) = 0$ for all x, $y > 0$. Therefore, if $t \geq s$, then, by Itô's formula,

$$U(W_t, W_t^*) = U(W_s, W_s^*) + \int_s^t U_x(W_u, W_u^*)\mathrm{d}W_u$$

(all the second-order terms vanish, due to the above conditions on U_{xx} and U_y). Now we would like to apply the conditional expectation with respect to $\mathcal{F}_s$ (which would remove the stochastic integral on the right). However, to ensure that both sides have appropriate integrability, we need an additional localizing argument. Fix a large integer N and consider the stopping time $\tau_N = \inf\{r : W_r \geq N\}$. Multiply both sides above by $1_{\{\tau_N > s\}}$ and replace t by $\tau_N \wedge t$. Then the integrand and the integrator are bounded, and the application of the conditional expectation with respect to $\mathcal{F}_s$ gives

$$\mathbb{E}\big[U(W_{\tau_N\wedge t}, W_{\tau_N\wedge t}^*)|\mathcal{F}_s\big] \leq U(W_s, W_s^*)$$

on the set $\{\tau_N > s\}$. However, one easily checks that if $x \leq y \leq cx$, then we have the pointwise bound

$$x^r \leq U(x, y) \leq (r - (r-1)c)^{-1} c^{1-r} y^r.$$

Combining this with the previous estimate and the definition of the A_1 weight yields

$$\begin{aligned}\mathbb{E}\big[W_{\tau_N\wedge t}^r|\mathcal{F}_s\big] &\leq (r - (r-1)c)^{-1} c^{1-r} (W_s^*)^r \\ &\leq (r - (r-1)c)^{-1} c^{1-r} (W^*)^r \leq (r - (r-1)c)^{-1} c W_\infty^r\end{aligned}$$

(still on the set $\{\tau_N > s\}$). If we let $N \to \infty$ and then $t \to \infty$, Fatou's lemma gives

$$\mathbb{E}\big[W_\infty^r|\mathcal{F}_s\big] \leq (1 + r(c^{-1} - 1))^{-1} W_\infty^r$$

almost surely. Now, the martingale $\big(\mathbb{E}\big[W_\infty^r|\mathcal{F}_s\big]\big)_{s\geq 0}$ has continuous trajectories (by the assumption on the filtration). Consequently,

$$(W^r)^* = \sup_{s\geq 0} \mathbb{E}\big[W_\infty^r|\mathcal{F}_s\big] = \sup_{s\geq 0,\, s\in\mathbb{Q}} \mathbb{E}\big[W_\infty^r|\mathcal{F}_s\big] \leq (1 + r(c^{-1} - 1))^{-1} W_\infty^r$$

with probability 1. This is exactly the claim. □

Proof of (1.6). This follows at once from (1.4) and the second part of the above lemma (we get the bound with the constant $2^{1/p'} \cdot 8pp' < 16pp'$). □

Proof of (1.7). By homogeneity, it is enough to show the estimate

$$\mathbb{E}1_{\{|Y_\infty|>1\}} W_\infty \leq c[W]_{A_1} \log(1 + [W]_{A_1}) ||X_\infty||_{L^1(W)}.$$

We will actually establish the somewhat stronger bound

$$\mathbb{E}1_{\{Y^*>1\}} W_\infty \leq c[W]_{A_1} \log(1 + [W]_{A_1}) ||X_\infty||_{L^1(W)}.$$

To this end, we will exploit an extrapolation-type argument. Let W be an A_1 weight. Fix X and Y as in the statement and write

$$\mathbb{E}W_\infty 1_{\{Y^*\geq 1\}} \leq \mathbb{E}W_\infty 1_{\{X^*>1\}} + \mathbb{E}W_\infty 1_{\{Y^*>1,\, X^*\leq 1\}} = I + II.$$

We analyze I and II separately.

The term I. Let $\eta = \inf\{t \geq 0 : |X_t| > 1\}$, with the usual convention $\inf \emptyset = \infty$. We have $\eta < \infty$ on the set $\{X^* > 1\}$, so

$$I = \mathbb{E}W_\infty 1_{\{X^*>1\}} \leq \lim_{t\to\infty} \mathbb{E}W_\infty 1_{\{|X_{\eta\wedge t}|>1\}}.$$

But for any $t \geq 0$,

$$\mathbb{E}W_\infty 1_{\{|X_{\eta\wedge t}|>1\}} = \mathbb{E}W_{\eta\wedge t} 1_{\{|X_{\eta\wedge t}|>1\}} \leq \mathbb{E}W_{\eta\wedge t}|X_{\eta\wedge t}| \leq \mathbb{E}W_{\eta\wedge t}|X_\infty| \leq \mathbb{E}|X_\infty|W^*.$$

This yields $I \leq \mathbb{E}|X_\infty|W^* \leq [W]_{A_1} ||X_\infty||_{L^1(W)}$.

The term II. Look at the stopped martingales $X^\eta = (X_{\eta\wedge t})_{t\geq 0}$, $Y^\eta = (Y_{\eta\wedge t})_{t\geq 0}$ and $W^\eta = (W_{\eta\wedge t})_{t\geq 0}$. Note that

$$II = \mathbb{E}W_\infty 1_{\{Y^*\geq 1,\, \eta=\infty\}} \leq \mathbb{E}W_\infty 1_{\{(Y^\eta)^*\geq 1\}} = \mathbb{E}W_\infty^\eta 1_{\{(Y^\eta)^*\geq 1\}},$$

where the latter follows from the fact that $(Y^\eta)^*$ is $\mathcal{F}_\eta$-measurable and $W_\infty^\eta = \mathbb{E}(W_\infty|\mathcal{F}_\eta)$. Now, introduce the stopping time $\sigma = \inf\{t : |Y_t^\eta| > 1\}$. By (1.4), for any $1 < p < \infty$ we have

$$\begin{aligned}
\mathbb{E}W_\infty^\eta 1_{\{|Y_\sigma^\eta|>1\}} \leq \mathbb{E}W_\infty^\eta |Y_\sigma^\eta|^p &\leq \frac{(8pp')^p}{(r-1)^{p-1}} \mathbb{E}|X_\sigma^\eta|^p M_r W_\infty^\eta \\
&\leq \frac{(8pp')^p}{(r-1)^{p-1}} \mathbb{E}|X_\sigma^\eta| M_r W_\infty^\eta \\
&\leq \frac{(8pp')^p}{(r-1)^{p-1}} \mathbb{E}|X_\infty| M_r W_\infty^\eta \\
&\leq \frac{(8pp')^p}{(r-1)^{p-1}} \mathbb{E}|X_\infty| M_r W_\infty.
\end{aligned}$$

Here in the third passage we have used the estimate $|X_\sigma^\eta| \leq 1$ (which follows directly from the definition of η), in the fourth bound we used Jensen's inequality, while the last one follows from the pointwise bound $M_r W_\infty^\eta \leq M_r W_\infty$. To see that the latter statement is true, write, for any $t \geq 0$,

$$\begin{aligned}\mathbb{E}\big((W_\infty^\eta)^r|\mathcal{F}_t\big) &= \mathbb{E}\big((\mathbb{E}(W_\infty|\mathcal{F}_\eta))^r|\mathcal{F}_t\big)\\ &\leq \mathbb{E}\big(\mathbb{E}(W_\infty^r|\mathcal{F}_\eta)|\mathcal{F}_t\big)\\ &= \mathbb{E}(W_\infty^r|\mathcal{F}_\eta)1_{\{\eta\leq t\}} + \mathbb{E}(W_\infty^r|\mathcal{F}_t)1_{\{\eta>t\}}\\ &\leq (W_\infty^r)^* 1_{\{\eta\leq t\}} + (W_\infty^r)^* 1_{\{\eta>t\}} = (W_\infty^r)^*.\end{aligned}$$

Taking the supremum over all $t \geq 0$ gives the bound for the r-maximal functions.

We return to the preceding chain of inequalities. Take $r = 1 + (2[W]_{A_1})^{-1}$ (then $M_r W_\infty \leq 2[W]_{A_1} W_\infty$, by the previous lemma) to get

$$\mathbb{E}W_\infty^\eta 1_{\{|Y_\sigma^\eta|>1\}} \leq 2(8pp')^p [W]_{A_1}^p \mathbb{E}|X_\infty| W_\infty.$$

Putting $p = 1 + (\log(1 + [W]_{A_1}))^{-1}$ gives

$$\mathbb{E}W_\infty^\eta 1_{\{|Y_\sigma^\eta|>1\}} \leq c[W]_{A_1}(1 + \log[W]_{A_1})\mathbb{E}|X_\infty| W_\infty,$$

for some universal c. It remains to observe that $\mathbb{E}W_\infty^\eta 1_{\{(Y^\eta)^*>1\}} = \mathbb{E}W_\infty^\eta 1_{\{|Y_\sigma^\eta|>1\}}$; thus we have shown that

$$II \leq c[W]_{A_1}(1 + \log[W]_{A_1})||X_\infty||_{L^1(W)}.$$

This completes the proof. □

5. Applications

Now we will show how to extend the results of the previous sections to the discrete setting, which is perhaps more natural for applications in harmonic analysis. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a fixed, nonatomic probability space.

Definition 5.1. A set $\mathcal{T}$ of measurable subsets of Ω will be called a tree if the following conditions are satisfied:

- (i) $\Omega \in \mathcal{T}$ and for every $J \in \mathcal{T}$, we have $\mathbb{P}(J) > 0$.
- (ii) For every $J \in \mathcal{T}$ there is a finite subset $C(J) \subset \mathcal{T}$ containing at least two elements such that
  - (a) the elements of $C(J)$ are pairwise disjoint subsets of J and
  - (b) $J = \bigcup C(J)$.
- (iii) $\mathcal{T} = \bigcup_{m\geq 0} \mathcal{T}^m$, where $\mathcal{T}^0 = \{\Omega\}$ and $T^{m+1} = \bigcup_{J\in\mathcal{T}^m} C(J)$.

In what follows, we will need to work with trees satisfying certain regularity-type property.

Definition 5.2. Let $\beta \geq 1$ be a given number. A tree $\mathcal{T}$ is called β-regular, if for any nonnegative integer n and any $J_1 \in \mathcal{T}^n$, $J_2 \in C(J_1)$ we have $\mathbb{P}(J_2)/\mathbb{P}(J_1) \in [2^{-\beta}, 1 - 2^{-\beta}]$.

A classical example which illustrates the above notions, is the cube $[0, 1)^d$ with its Borel subsets and Lebesgue's measure, equipped with the tree of dyadic subcubes. Clearly, in such a case one may take $\beta = d$ and hence, in general, the parameter β (or rather the smallest value allowed in the definition) can be regarded as a dimension of the tree $\mathcal{T}$.

Any tree-like structure gives rise to the corresponding filtration $(\mathcal{F}_n)_{n\geq 0}$, given by $\mathcal{F}_n = \sigma(J : J \in \mathcal{T}^n)$. Given an integrable random variable f, one can consider the associated martingale given by $(\mathbb{E}(f|\mathcal{F}_n))_{n\geq 0}$; this sequence will be denoted by

$(f_n)_{n\geq 0}$ or, with a slight abuse of notation, again by the letter f. Note that such martingales are simple, i.e., for any nonnegative integer n, the random variable f_n takes only a finite number of values. This follows at once from the fact that $\mathcal{F}_n$ consists of finite number of sets. For a given martingale $f=(f_n)_{n\geq 0}$, we introduce its maximal function and truncated maximal function by $f^*=\sup_{k\geq 0}|f_k|$ and $f_n^*=\sup_{0\leq k\leq n}|f_k|$, $n=0,\,1,\,2,\,\ldots$. Furthermore, the difference sequence $df=(df_n)_{n\geq 0}$ associated with a martingale $(f_n)_{n\geq 0}$ is defined by $df_0=f_0$ and $df_n=f_n-f_{n-1}$ for $n\geq 1$. Given any predictable sequence $v=(v_n)_{n\geq 0}$, we say that g is the transform of f by v, if for any $n\geq 0$ we have $dg_n=v_n df_n$; equivalently, we have the identities

$$g_n=\sum_{k=0}^n v_k df_k,\qquad n=0,\,1,\,2,\,\ldots.$$

In the context when $(\Omega,\mathcal{F},\mathbb{P})$ is the interval $[0,1)$ with its Borel subsets and Lebesgue measure, equipped with the dyadic tree, the martingale $(f_n)_{n\geq 0}$ corresponds to the expansion of f into the standard Haar system and $g=(g_n)_{n\geq 0}$ reduces to the Haar multiplier of f with respect to the sequence v.

Finally, a nonnegative, integrable random variable will be called a weight; we will say that a weight w satisfies Muckenhoupt's condition A_1 if there is an absolute constant c such that $\mathbb{P}(w_n^*\leq cw_n)=1$ for all n.

We are ready to establish our main results in this direction.

Theorem 5.3. *Suppose that $(\Omega,\mathcal{F},\mathbb{P})$ is a probability space equipped with a β-regular tree $\mathcal{T}$ and let w be an A_1 weight. If g is a transform of f by a predictable sequence v satisfying $||v||_\infty=\sup_{n\geq 0}||v_n||_\infty\leq 1$, then*

$$||g||_{L^p(w)}\leq 2^\beta\cdot 16pp'[w]_{A_1}||f||_{L^p(w)},\qquad 1<p<\infty,$$

and, for some universal constant c,

$$||g||_{L^{1,\infty}(w)}\leq c2^\beta[w]_{A_1}(1+\log(2^\beta[w]_{A_1}))||f||_{L^1(w)}.$$

Proof. Let us assume first that the martingales f and w terminate after N steps. Furthermore, by a simple approximation argument, we may assume that f takes different values on different elements of $\mathcal{T}^m$, $m=0,\,1,\,\ldots,\,N$; then in particular the variable v_n is a function of f_{n-1}: say, $v_n=F_n(f_{n-1})$ for some function F_n with values in $[-1,1]$, $n=1,\,2,\,\ldots,N$.

Using standard Skorokhod embedding theorems (cf. [12]), there is a probability space $(\Omega',\mathcal{F}',\mathbb{P}')$ equipped with a filtration generated by a standard Brownian motion B and a sequence $\tau_0,\,\tau_1,\,\ldots,\,\tau_N$ of stopping times and predictable processes K^w, K^f such that

$$(w_n)_{n=0}^N\sim\left(\mathbb{E}w_N+\int_{0+}^{\tau_n}K_s^w\mathrm{d}B_s\right)_{n=0}^N,\qquad (f_n)_{n=0}^N\sim\left(\mathbb{E}f_N+\int_{0+}^{\tau_n}K_s^f\mathrm{d}B_s\right)_{n=0}^N.$$

Here $\sim$ means the equality in law. Introduce the martingales

$$X_t=\mathbb{E}f_N+\int_{0+}^t K_s^f\mathrm{d}B_s,\qquad W_t=\mathbb{E}w_N+\int_{0+}^t K_s^w\mathrm{d}B_s.$$

Since $v_n=F_n(f_{n-1})$ for each n (see the beginning of the proof), we see that there is a predictable process H with values in $[-1,1]$ such that $(g_n)_{n=0}^N\sim((H\cdot X)_{\tau_n})_{n=0}^N$. Indeed, it is enough to take $H_t=F_n(X_{\tau_{n-1}})$ when $t\in[\tau_{n-1},\tau_n)$.

Let us compare the characteristics $[w]_{A_1}$ and $[W]_{A_1}$. Let A be an element of $\mathcal{T}^n$ and let $A_1, A_2, \ldots, A_k$ be its children from $\mathcal{T}^{n+1}$. By β-regularity of $\mathcal{T}$, we see that $\mathbb{P}(A_i) \geq 2^{-\beta}\mathbb{P}(A)$ and hence

$$w_{n+1}|_{A_i} = \frac{1}{\mathbb{P}(A_i)}\int_{A_i} w\mathrm{d}\mathbb{P} \leq \frac{1}{\mathbb{P}(A_i)}\int_A w\mathrm{d}\mathbb{P} \leq 2^\beta \frac{1}{\mathbb{P}(A)}\int_A w\mathrm{d}\mathbb{P} = 2^\beta w_n|_A.$$

This yields the estimate $w_{n+1} \leq 2^\beta w_n$ on A, and since A was arbitrary, this pointwise bound is true on the whole Ω. Consequently, for any $n = 0, 1, \ldots, N-1$ we have $W_{\tau_{n+1}} \leq 2^\beta W_{\tau_n}$. Now, fix an arbitrary $t \geq 0$. On the set $\{t \in [\tau_n, \tau_{n+1}]\}$ we have

$$W_t = \mathbb{E}(W_{\tau_{n+1}}|\mathcal{F}_t) \leq 2^\beta \mathbb{E}(W_{\tau_n}|\mathcal{F}_t) = 2^\beta W_{\tau_n}$$

and hence $W_t^* \leq 2^\beta \max_{0\leq k\leq n} W_{\tau_k}$. On the other hand, the condition $w \in A_1$ implies $[w]_{A_1} w_{n+1} \geq w_{n+1}^* \geq w_n^*$ almost surely. Therefore, if $t \in [\tau_n, \tau_{n+1})$, we have

$$W_t = \mathbb{E}(W_{\tau_{n+1}}|\mathcal{F}_t) \geq [w]_{A_1}^{-1}\mathbb{E}(\max_{0\leq k\leq n} W_{\tau_k}|\mathcal{F}_t) = [w]_{A_1}^{-1}\max_{0\leq k\leq n} W_{\tau_k}.$$

Combining the above two observations, we get $W_t^* \leq 2^\beta [w]_{A_1} W_t$ on the set $t \leq \tau_N$. But W is constant on $[\tau_N, \infty)$, so this inequality is also true for $t > \tau_N$. This proves that $[W]_{A_1} \leq 2^\beta [w]_{A_1}$. It remains to apply Theorems 1.1 and 1.2 to get the claim. □

Acknowledgments

The author would like to thank anonymous referees for the careful reading of the first version of the paper and many helpful suggestions.

References

[1] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), no. 1, 253–272.

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

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

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

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

[6] N. Kazamaki, Continuous exponential martingales and BMO, Lecture Notes in Math. 1579, Springer-Verlag, Berlin, Heidelberg, 1994.

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

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

[9] F. L. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg, *Weak norm estimates of weighted singular operators and Bellman functions.* Manuscript (2010).

[10] F. L. Nazarov and S. R. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis,* St. Petersburg Math. J. **8** (1997), 721–824.

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

[12] J. Obłój, *The Skorokhod embedding problem and its offspring*, Probab. Surveys **1** (2004), pp. 321–392.

[13] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72** (2012), Birkhäuser, 462 pp.
[14] S. Petermichl and J. Wittwer, *A Sharp Estimate for the Weighted Hilbert Transform via Bellman Functions,* Michigan Math. **50** (2002), 71–87.
[15] M. C. Reguera, *On Muckenhoupt-Wheeden conjecture*, Adv. Math. 227 (2011), no. 4, pp. 1436–1450.
[16] M. C. Reguera and C. Thiele, *The Hilbert transform does not map* $L^1(Mw)$ *to* $L^{1,\infty}(w)$, Math. Res. Lett. 19 (2012), no. 1, pp. 1–7.
[17] L. Slavin, V. Vasyunin, *Sharp results in the integral-form John-Nirenberg inequality*, Trans. Amer. Math. Soc. **363** (2011), 4135–4169.
[18] V. Vasyunin, A. Volberg, *Monge-Ampére equation and Bellman optimization of Carleson Embedding Theorems*, Amer. Math. Soc. Transl. (2), vol. 226, "Linear and Complex Analysis", 2009, pp. 195-238.

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