PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 41, Fasc. 1 (2021), pp. 89–114
Published online 1.4.2021
doi:10.37190/0208-4147.41.1.7

WEIGHTED MAXIMAL INEQUALITIES FOR MARTINGALE TRANSFORMS*

BY

MICHAŁ BRZOZOWSKI (WARSZAWA) AND ADAM OSĘKOWSKI (WARSZAWA)

Abstract. We study the weighted maximal L^1-inequality for martingale transforms, under the assumption that the underlying weight satisfies Muckenhoupt's condition A_∞ and that the filtration is regular. The resulting linear dependence of the constant on the A_∞ characteristic of the weight is optimal. The proof exploits certain special functions enjoying appropriate size conditions and concavity.

2020 Mathematics Subject Classification: Primary 60G44; Secondary 60G42.

Key words and phrases: martingale, weight, Bellman function, maximal function.

1. INTRODUCTION

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space filtered by $(\mathcal{F}_n)_{n\geqslant 0}$, a nondecreasing sequence of sub-σ-algebras of $\mathcal{F}$. We additionally assume that this filtration is θ-regular for some $\theta \in (0, 1/2]$, that is, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and every atom A of each $\mathcal{F}_n$ splits into a finite number $A_1, \ldots, A_k$ of atoms of $\mathcal{F}_{n+1}$ satisfying $\mathbb{P}(A_j) \geqslant \theta\mathbb{P}(A)$, $j = 1, \ldots, k$. Regular filtrations are natural extensions of dyadic filtrations widely used in harmonic analysis: for a fixed dimension d, the dyadic filtration of the space $([0, 1]^d, \mathcal{B}([0, 1]^d), |\cdot|)$ is 2^{-d}-regular in the above sense.

Next, suppose that $f = (f_n)_{n\geqslant 0}$ and $g = (g_n)_{n\geqslant 0}$ are adapted, uniformly integrable martingales. We will identify the martingales f and g with the pointwise limits f_∞, g_∞, which exist due to the uniform integrability. Define the associated difference sequences $df = (df_n)_{n\geqslant 0}$ and $dg = (dg_n)_{n\geqslant 0}$ by

$$df_0 = f_0, \quad df_n = f_n - f_{n-1}, \quad n = 1, 2, \ldots,$$

and similarly for dg. The maximal function of f is given by $|f|^* = \sup_{k\geqslant 0} |f_k|$, and the truncated maximal function is $|f|^*_n = \sup_{0\leqslant k\leqslant n} |f_k|$, $n = 0, 1, \ldots$. The

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

martingale g is a *transform* of f if there is a predictable sequence $\varepsilon = (\varepsilon_n)_{n\geqslant 0}$ such that $dg_n = \varepsilon_n df_n$ for every n; here by *predictability* we mean that for each n, the random variable ε_n is measurable with respect to $\mathcal{F}_{(n-1)\vee 0}$. Moreover if the sequence ε is deterministic and its terms take values in $\{-1,1\}$, then g is said to be a ± 1-*transform* of f.

Inequalities for martingale transforms play an important role in probability theory and have deep applications in harmonic analysis. There is a huge literature on the subject: we mention here Burkholder's papers [5], [6], [7], the monograph [20] and the papers [26], [27] for an overview of probabilistic results; for analytic applications, consult e.g. [1], [2], [10], [25]. In this paper, we will be particularly interested in maximal inequalities. In [7], Burkholder introduced a general method of proving such estimates in the context of martingale transforms and exploited it to establish the following result.

THEOREM 1.1. *If f, g are martingales satisfying $dg_n = \varepsilon_n df_n$, $n = 0, 1, \ldots,$ for some predictable sequence $\varepsilon = (\varepsilon_n)_{n\geqslant 0}$ with values in $[-1,1]$, then*

$$\|g\|_{L^1} \leqslant \eta \big\| |f|^* \big\|_{L^1}, \tag{1.1}$$

where $\eta = 2.536\ldots$ is the unique solution of the equation $\eta - 3 = -\exp\left(\frac{1-\eta}{2}\right)$. The constant is the best possible.

See also [17], [19] and [18] for related results and generalizations. In this paper we will be interested in the weighted versions of the above statement. In what follows, the word "weight" will refer to a positive, integrable random variable usually denoted by w. Given $1 < p < \infty$, we say that w satisfies *Muckenhoupt's condition* A_p (or belongs to the A_p class) if

$$[w]_{A_p} := \sup \left(\frac{1}{\mathbb{P}(A)} \int_A w \, \mathrm{d}\mathbb{P} \right) \left(\frac{1}{\mathbb{P}(A)} \int_A w^{-1/(p-1)} \, \mathrm{d}\mathbb{P} \right)^{p-1} < \infty,$$

where the supremum is taken over all n and all atoms A of $\mathcal{F}_n$. There are versions of this definition for $p \in \{1, \infty\}$; we will recall the case $p = \infty$ only, as we will not work with A_1 here. A weight w belongs to the class A_∞ if

$$[w]_{A_\infty} := \sup \left(\frac{1}{\mathbb{P}(A)} \int_A w \, \mathrm{d}\mathbb{P} \right) \exp\left(-\frac{1}{\mathbb{P}(A)} \int_A \log(w) \, \mathrm{d}\mathbb{P} \right) < \infty,$$

the supremum taken over the same class of A as above. Two comments are in order. First, note that in the dyadic context (i.e., when the probability space equals $([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|)$ and the filtration is dyadic), the above definitions lead to the classical dyadic A_p weights. The second observation is that the above definitions can be easily rephrased in the language of conditional expectations: $[w]_{A_p}$ is the least number c such that for all $n \geqslant 0$,

$$\mathbb{E}(w \,|\, \mathcal{F}_n)(\mathbb{E}(w^{1/(1-p)} \,|\, \mathcal{F}_n))^{p-1} \leqslant c$$

almost surely, while $[w]_{A_\infty}$ is the smallest c for which

$$\mathbb{E}(w\,|\,\mathcal{F}_n)\exp(\mathbb{E}(-\log(w)\,|\,\mathcal{F}_n))\leqslant c$$

almost surely, $n=0,1,\ldots$. It follows directly from Hölder's inequality that $[w]_{A_p}\geqslant 1$ and that the A_p classes grow as p increases. Furthermore, it is well-known (cf. [11]) that $A_\infty=\bigcup_{1<p<\infty}A_p$.

The main theme of this paper is to study the following weighted extension of (1.1):

$$\big\||g|^*\big\|_{L^1(w)}\leqslant C_{\theta,w}\big\||f|^*\big\|_{L^1(w)}. \tag{1.2}$$

Note that the maximal function appears on both sides of the estimate. We will show that if w belongs to the class A_∞, then (1.2) holds for all martingales f and their transforms. In addition, we will study the following aspect of the weighted bound. There is an interesting question on the sharp dependence of the constant C on the characteristic $[w]_{A_\infty}$. More precisely: what is the least exponent κ for which there exists a constant $\tilde{C}_\theta$ depending only on the regularity of the filtration such that

$$\big\||g|^*\big\|_{L^1(w)}\leqslant \tilde{C}_\theta[w]_{A_\infty}^\kappa\big\||f|^*\big\|_{L^1(w)}$$

for all f, g, w as above? Such "extraction" problems have gained a lot of interest in the literature and have been studied for various classes of operators and estimates: see e.g. [4], [13], [14], [15], [28].

The main result of this paper gives the full answer to the above question.

THEOREM 1.2. *Fix $\theta\in(0,1/2]$. Let f, g be martingales adapted to a θ-regular filtration such that g is a transform of f by means of a predictable sequence with values in $[-1,1]$. Then for any A_∞ weight w we have*

$$\big\||g|^*\big\|_{L^1(w)}\leqslant 769\theta^{-2}[w]_{A_\infty}\big\||f|^*\big\|_{L^1(w)}. \tag{1.3}$$

The dependence on the A_∞ characteristic of the weight is optimal in the sense that for any $\kappa<1$ and any $K>0$, there is a weight w, a real-valued martingale f and a predictable sequence ε with values in $\{-1,1\}$ such that

$$\big\||g|^*\big\|_{L^1(w)}> K[w]_{A_\infty}^\kappa\big\||f|^*\big\|_{L^1(w)}.$$

A weaker result for Haar multipliers and A_p weights was obtained in [23]. It was shown there that

$$\Big\|\max_{0\leqslant n\leqslant N}\Big|\sum_{k=0}^n\varepsilon_k a_k h_k\Big|\Big\|_{L^1(w)}\leqslant C_p[w]_{A_p}\Big\|\max_{0\leqslant n\leqslant N}\Big|\sum_{k=0}^n a_k h_k\Big|\Big\|_{L^1(w)}, \tag{1.4}$$

where $1<p<\infty$, w is a dyadic A_p weight, N is a nonnegative integer, $a_0,a_1,\ldots,a_N$ are real numbers, $\varepsilon_0,\varepsilon_1,\ldots,\varepsilon_N$ is a sequence of signs and $(h_k)_{k=0}^\infty$ is the

Haar system on $[0, 1)$. Moreover, it was proved in [23] that the linear dependence on the characteristic is optimal. Observe that (1.3) generalizes this result in two directions. Firstly, we consider the more general case of θ-regular filtrations. Secondly, since $[w]_{A_\infty} \leqslant [w]_{A_p}$, the estimate (1.3) is stronger; hence the optimality of the linear dependence in (1.3) follows at once from the analogous sharpness in (1.4) and all we need is to prove of (1.3).

Let us now make an important comment on the θ-regularity of the underlying filtration. The multiplicative constant in (1.3) depends on θ and goes to infinity as θ tends to 0. We will prove that this dependence is necessary, even if we consider the weaker estimate for A_p weights for any given $p > 1$. Here is the precise formulation.

THEOREM 1.3. *Let $p > 1$ and let K be an arbitrary positive constant. Then there is a positive integer d, a martingale f on a d-dimensional dyadic probability space, an A_p weight w satisfying $[w]_{A_p} \leqslant 2$ and a predictable sequence v with values in $\{-1, 1\}$ such that the associated martingale transform g satisfies*

$$\|g\|_{L^1(w)} > K \big\| |f|^* \big\|_{L^1(w)}.$$

There is a well-known method of proving maximal inequalities for martingales and their martingale transforms. This method, invented by Burkholder in [7] and modified by the second author in [19], [20], allows one to deduce an estimate from the existence of a certain special function, enjoying appropriate majorization and concavity. This method is extended in Section 2 to cover the setting of A_p weights, and successfully applied in Section 3 in the proof of (1.3). Section 5 is devoted to Theorem 1.3, which is proved again with the use of the Bellman function method.

2. ON THE METHOD OF PROOF

We will now describe a general technique which can be used to study weighted estimates for martingales.

We start with the following helpful interpretation of A_∞ weights. Suppose that w is such a weight; we will often identify it with the associated martingale $(w_n)_{n\geqslant 0} = (\mathbb{E}(w \mid \mathcal{F}_n))_{n\geqslant 0}$. Let $\sigma = (\sigma_n)_{n\geqslant 0}$ be the dual martingale given by $\sigma_n = \mathbb{E}(\log(w) \mid \mathcal{F}_n)_{n\geqslant 0}$ (the integrability of $\log w$ follows at once from the condition $w \in A_\infty$). By Jensen's inequality we have $w_n \exp(-\sigma_n) \geqslant 1$ almost surely for all $n \geqslant 0$, and the condition A_∞ implies $w_n \exp(-\sigma_n) \leqslant [w]_{A_\infty}$ with probability 1. In other words, an A_∞ weight of characteristic less than or equal to c gives rise to a two-dimensional uniformly integrable martingale (w, σ) taking values in the hyperbolic domain $\{(u, v) \in (0, \infty) \times \mathbb{R} : 1 \leqslant ue^{-v} \leqslant c\}$. Actually, the implication can be reversed: any uniformly integrable martingale pair (w, σ) taking values in the above set and terminating at its lower boundary (i.e., satisfying $w_\infty e^{-\sigma_\infty} = 1$) induces an A_∞ weight: just take the first coordinate w. A similar statement is true for A_p weights, $1 < p < \infty$: the only change is that now the dual martingale σ is generated by $w^{1/(1-p)}$ and the domain should be modified to $\{(u,v) \in (0,\infty)^2 : 1 \leqslant uv^{p-1} \leqslant c\}$.

Now, suppose that $M : \mathbb{R}^2 \times [0,\infty) \times (0,\infty) \to \mathbb{R}$ is a given continuous function and we want to show that

$$\mathbb{E}M(f_n, g_n, |f|_n^*, w_n) \leqslant 0, \qquad n \geqslant 0, \tag{2.1}$$

for all f, g, w, where $f,\ g$ are martingales such that g is the transform of f via a certain sequence with values in $[-1,1]$, and w is an A_∞ weight satisfying $[w]_{A_\infty} \leqslant c$. We additionally assume that all these processes are adapted to a θ-regular filtration on some probability space. The key to handle this problem is to consider the class $B(M)$ of all functions B defined on the five-dimensional set

$$\mathcal{D} = \{(x,y,z,u,v) \in \mathbb{R}^2 \times (0,\infty)^2 \times \mathbb{R} : |x| \leqslant z, 1 \leqslant ue^{-v} \leqslant c\}$$

and enjoying the following three properties:

0° (Initial condition) We have $B(x,y,|x|,u,v) \leqslant 0$ if $|y| \leqslant |x|$, $|x| > 0$ and $1 \leqslant ue^{-v} \leqslant c$.

1° (Majorization property) We have

$$B(x,y,z,u,v) \geqslant M(x,y,z,u) \quad \text{ for } (x,y,z,u,v) \in \mathcal{D}.$$

2° (Concavity-type property) For any $(x,y,z,u,v) \in \mathcal{D}$, any $\varepsilon \in [-1,1]$, any positive integer $k \leqslant 1/\theta$ and any sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying

$$\begin{aligned} &\alpha_j \in [\theta,1), \qquad \sum_{j=1}^k \alpha_j = 1, \\ &\sum_{j=1}^k \alpha_j h_j = \sum_{j=1}^k \alpha_j r_j = \sum_{j=1}^k \alpha_j s_j = 0, \\ &(x+h_j, y+\varepsilon h_j, |x+h_j| \vee z, u+r_j, v+s_j) \in \mathcal{D}, \end{aligned}$$

we have

$$B(x,y,z,u,v) \geqslant \sum \alpha_j B(x+h_j, y+\varepsilon h_j, |x+h_j| \vee z, u+r_j, v+s_j).$$

The relation between functions satisfying the above special properties and the validity of (2.1) is described in the statement below.

THEOREM 2.1. *If the class* $\mathcal{B}(M)$ *is nonempty, then* (2.1) *holds.*

Proof. By a standard limiting argument (using continuity of M and the fact that the variables f_n, g_n, ... take only a finite number of values), we may and do assume that $|f_0| > 0$ almost surely; then the process $z_n = (f_n, g_n, |f|_n^*, w_n, \sigma_n)$ takes values in $\mathcal{D}$. The key fact is that the process $(B(z_n))_{n\geqslant 0}$ is a supermartingale, which is an immediate consequence of the concavity-type condition 2°:

$$\begin{aligned}\mathbb{E}[B(f_n, g_n, |f|_n^*, w_n, \sigma_n) \,|\, \mathcal{F}_{n-1}] \\ = \mathbb{E}[B(f_{n-1} + df_n, g_{n-1} + dg_n, |f|_{n-1}^* \vee |f_n + df_n|, \\ w_{n-1} + dw_n, \sigma_{n-1} + d\sigma_n) \,|\, \mathcal{F}_{n-1}] \\ \leqslant B(f_{n-1}, g_{n-1}, |f|_{n-1}^*, w_{n-1}, \sigma_{n-1}).\end{aligned}$$

Therefore, if we apply the majorization 1° and then the initial condition 0°, we get

$$\begin{aligned}\mathbb{E}M(f_n, g_n, |f|_n^*, w_n) \leqslant \mathbb{E}B(f_n, g_n, |f|_n^*, w_n, \sigma_n) \leqslant \mathbb{E}B(f_0, g_0, |f|_0^*, w_0, \sigma_0) \\ \leqslant 0,\end{aligned}$$

which is the desired inequality (2.1). ∎

It is a beautiful fact that the above implication can be reversed.

THEOREM 2.2. *If* (2.1) *holds true* (*for all f, g and all weights w with $[w]_{A_\infty} \leqslant c$*), *then the class $\mathcal{B}(M)$ is nonempty.*

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

$$B(x, y, z, u, v) = \sup \mathbb{E}M(f_n, g_n, |f|_n^* \vee z, w_n).$$

Here the supremum is taken over all n, all A_∞ weights w satisfying $[w]_\infty \leqslant c$, $w_0 = u$, $\mathbb{E}\log w = v$ and all martingale pairs (f, g) satisfying $(f_0, g_0) = (x, y)$ and $dg_k = \varepsilon_k df_k$, $k \geqslant 1$, for some predictable sequence $(\varepsilon_k)_{k\geqslant 1}$ with values in $[-1, 1]$. Here the probability space as well as the θ-regular filtration are also assumed to vary. We will show that the function B satisfies conditions 0°–2°. The initial condition 0° follows immediately from (2.1). The majorization condition is also easy: it suffices to compute the expression in the definition of B for $n = 0$. The most difficult issue is the concavity-type condition 2°. We will use the so-called "splicing" argument. Fix the parameters x, y, z, u, v, k, ... as in 2° and, for each $j = 1, \ldots, k$, pick arbitrary martingales (f^j, g^j, w^j) as in the definition of $B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j)$. We may assume that these martingales are given on k pairwise disjoint probability spaces $(\Omega^j, \mathcal{F}^j, \mathbb{P}^j)$. Now we "glue" these spaces and the martingale triples into one space and one triple using the parameters $(\alpha_j)_{j=1}^k$. Namely, let $\Omega = \Omega^1 \cup \cdots \cup \Omega^k$, $\mathcal{F} = \sigma(\mathcal{F}^1, \ldots, \mathcal{F}^k)$ and define a probability measure $\mathbb{P}$ on $\mathcal{F}$ by requiring that $\mathbb{P}(\bigcup A_j) = \sum \alpha_j \mathbb{P}(A_j)$

for any $A_j \in \mathcal{F}^j$, $j = 1, \ldots, k$. Next, define (f, g, w) by $(f_0, g_0, w_0) = (x, y, w)$ and

$$(f_n(\omega), g_n(\omega), w_n(\omega)) = (f^j_{n-1}(\omega), g^j_{n-1}(\omega), w^j_{n-1}(\omega))$$

if $\omega \in \Omega^j$. Finally, let $(\mathcal{F}_n)_{n\geqslant 0}$ be the natural filtration of (f, g, w).

Directly from the above definition, we see that

$$\mathbb{E}((f_1, g_1, w_1) \,|\, \mathcal{F}_0) = \mathbb{E}(f_1, g_1, w_1) = \sum \alpha_j(x+h_j, y+\varepsilon h_j, u+r_j) = (x, y, u).$$

Furthermore, since (f^j, g^j, w^j) are martingales, so is (f, g, w). In addition,

$$\mathbb{E}\log(w) = \sum \alpha_j \mathbb{E}^j \log(w^j) = \sum \alpha_j(v_j + s_j) = v,$$

where $\mathbb{E}^j$ is the expectation with respect to $\mathbb{P}^j$. Our next observation is that w is an A_∞ weight with $[w]_{A_\infty} \leqslant c$. Indeed, $w_0 e^{-\sigma_0} = ue^{-v} \leqslant c$, and for $n \geqslant 1$ the pointwise estimate $w_n e^{-\sigma_n} \leqslant c$ follows from $[w^j]_{A_\infty} \leqslant c$. Consequently, by the very definition of B,

$$\begin{aligned} B(x, y, z, u, v) &\geqslant \mathbb{E}M(f_n, g_n, |f_n| \vee z, w_n) \\ &= \sum \alpha_j \mathbb{E}^j M(f^j_{n-1}, g^j_{n-1}, |f^j_{n-1}| \vee z, w^j_{n-1}), \end{aligned}$$

so taking the supremum over all n and all triples (f^j, g^j, w^j) as above, we obtain

$$B(x, y, z, u, v) \geqslant \sum \alpha_j B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j).$$

This is precisely the desired condition 2°. ■

Three comments are in order.

REMARK 2.1. The above method works for A_p weights as well: the only change concerns the definition of the domain $\mathcal{D}$, in which the double estimate $1 \leqslant ue^{-v} \leqslant c$ should be changed to $1 \leqslant uv^{p-1} \leqslant c$.

REMARK 2.2. Suppose that we are interested in the estimate (2.1) in the d-dimensional dyadic context. Then the above approach can be modified easily: we consider the function B given by the abstract formula as above,

$$B(x, y, z, u, v) = \sup \mathbb{E}M(f_n, g_n, |f|^*_n \vee z, w_n).$$

Here the supremum is taken over all martingales as in the above proof, the essential difference is that the probability space is fixed to be $([0, 1]^d, \mathcal{B}([0, 1]^d), |\cdot|)$ and the filtration is assumed to be dyadic. Thanks to the fractal, self-similar structure of the dyadic filtration, the above splicing argument is valid, and the function B satisfies $0°$, $1°$ and a weaker version of $2°$, with all α_j's equal to 2^{-d}. A similar modification can be applied for A_p weights (see the previous remark). This observation will be crucial in the last subsection where we show that (1.3) cannot hold universally, i.e., with a constant independent of θ.

REMARK 2.3. The technique is quite flexible and general. For instance, it can be used to study weighted nonmaximal estimates, simply by working with the functions M and B depending only on x, y, w and v. Another possible modification is that if we want to show (2.1) for processes as previously, but satisfying the additional property $\|f\|_\infty \leqslant 1$, the domain of M and B has to be changed: it is enough to consider M and B defined on $\{(x, y, z, u, v) \in [-1, 1]\times\mathbb{R}\times(0, 1]\times(0, \infty)\times\mathbb{R} : |x| \leqslant z, 1 \leqslant ue^{-v} \leqslant c\}$.

The remainder of this section contains some informal reasoning which leads to the special function corresponding to (1.3); the reader might skip it and proceed to Section 3. We have decided to insert this material, since we believe that the steps leading to the discovery of the function may become useful in the study of other related estimates.

As we will see later, the main difficulty lies in proving the estimate

$$\|g_n\|_{L^1(w)} \leqslant C[w]_{A_\infty} \big\||f|_n^*\big\|_{L^1(w)}, \qquad n \geqslant 0,$$

which is slightly weaker than (1.3), since it does not involve the maximal function of g on the left. This inequality is of the form (2.1) with $M(x, y, z, u, v) = |y|u - Cczu$, where $c = [w]_{A_\infty}$, and hence all we need is an appropriate special function B. At first glance, it is not clear at all how to search for this object. To gain some intuition and indication, let us review several results from the well-understood unweighted case.

We start with the nonmaximal $L^\infty \to L^2$ inequality (as we will see in a moment, it will be of key importance): if f, g are martingales such that $\|f\|_\infty \leqslant 1$ and $dg_n = v_n df_n$, $n = 0, 1, \ldots$, for some predictable sequence $(v_n)_{n\geqslant 0}$ taking values in $[-1, 1]$, then $\|g\|_2 \leqslant 1$. This trivial result can be proved with the use of Burkholder's method (see [5]), and the corresponding function $u : [-1, 1]\times\mathbb{R} \to \mathbb{R}$ is

$$u(x, y) = y^2 - x^2.$$

Next, we turn to maximal estimates in the unweighted setting. As shown in [19], the special function $\mathcal{U} : \{(x, y, z) : |x| \leqslant z\} \to \mathbb{R}$ corresponding to the continuous analogue of (1.1) is given by

$$\mathcal{U}(x, y, z) = \frac{y^2 - x^2 - z}{z} = z\bigg(u\Big(\frac{x}{z}, \frac{y}{z}\Big) - 1\bigg). \tag{2.2}$$

As we see, this special function uses two components: the multiplicative constant z which controls the maximal function of f, and the special function on the strip which handles the $L^\infty \to L^2$ estimate.

A natural idea is to try to follow this path in the weighted setting. Suppose that w is an A_∞ weight. The main problem is to find an appropriate weighted analogue of the function u above; indeed, having found such an object (say $\bar{u}$, a function of

x, y, u and v), it seems plausible to put

$$B(x,y,z,u,v) = z\bigg(\bar{u}\bigg(\frac{x}{z},\frac{y}{z},u,v\bigg) - Lu\bigg),$$

for some constant L to be found. The function $\bar{u}$ should encode the $L^\infty(W) \to L^2(W)$ inequality, or rather an $L^\infty(W) \to L^q(W)$ estimate for some q, for martingale transforms. Fortunately, some indications towards its discovery can be extracted from [22]. In that paper, similar inequalities in the presence of A_p weights were studied. Roughly speaking, to obtain $L^\infty(W) \to L^q(W)$ estimates in this context, the procedure is as follows. Take a special function U_r associated with a nonmaximal and unweighted $L^r \to L^r$ bound (this problem is well-understood, see Burkholder [5]) and then put

$$\bar{u}(x,y,u,v) = (U_r(x,y)+\kappa)^\beta(uv^{p-1}-a)^\alpha v^{1-p}$$

for some parameters α, β, κ and a. In the present paper we want to take $p=\infty$, so some change is needed. It turns out that the right choice for $\bar{u}$ is

$$\bar{u}(x,y,u,v) = (U_r(x,y)+\kappa)^\beta(ue^{-v}-a)^\alpha e^v.$$

To see the reason for our modification "$v^{p-1} \to e^{-v}$", compare the geometric interpretations of A_p and A_∞ weights presented at the beginning of this section.

3. BURKHOLDER'S FUNCTION OF FIVE VARIABLES

In order to prove the inequality (1.3), we will first prove the weaker estimate

$$\|g_n\|_{L^1(w)} \leqslant C[w]_{A_\infty}\big\||f|_n^*\big\|_{L^1(w)}, \quad n=0,1,\dots. \tag{3.1}$$

From the previous section it is sufficient to find a function $B:\mathcal{D}\to\mathbb{R}$ which satisfies conditions 0°–2° with $M(x,y,z,u,v)=|y|u-Cczu$. As we will see, this special object will be built from several simpler "blocks". For brevity, set

$$\beta=\theta(8c(1-\theta))^{-1}, \quad \alpha=1-(2c)^{-1}, \quad a=3/4, \quad p=1/\beta, \quad A=4/\theta-1.$$

Observe that $p=8c(1/\theta-1)\geqslant 8$. Let

$$D_c=\{(u,v)\in(0,\infty)\times\mathbb{R}: 1\leqslant ue^{-v}\leqslant c\}.$$

For $(r,u,v)\in(0,\infty)\times D_c$, set

$$F(r,u,v)=r^\beta(ue^{-v}-a)^\alpha e^v.$$

Furthermore, for any $x,y\in\mathbb{R}$, define

$$U(x,y)=\begin{cases} p(1-1/p)^{p-1}(|y|-(p-1)|x|)(|x|+|y|)^{p-1} & \text{if } |y|\geqslant(p-1)|x|,\\ |y|^p-(p-1)^p|x|^p & \text{if } |y|<(p-1)|x|.\end{cases}$$

This is the celebrated special function invented by Burkholder [5] to establish sharp L^p bounds for martingale transforms. Burkholder proved that U enjoys the following.

LEMMA 3.1. *The function U has the following properties:*

(i) (Initial condition) $U(x,y) \leqslant 0$ *if* $|y| \leqslant |x|$.

(ii) (Majorization property) $U(x,y) \geqslant |y|^p - (p-1)^p|x|^p$.

(iii) (Concavity-type property) *For any* $(x,y) \in \mathbb{R}^2$, *any* $\varepsilon \in [-1,1]$, *any positive integer* k *and any sequences* $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k$ *satisfying*

$$\alpha_j \in [0,1), \quad \sum_{j=1}^k \alpha_j = 1, \quad \sum_{j=1}^k \alpha_j h_j = 0,$$

we have

$$U(x,y) \geqslant \sum \alpha_j U(x+h_j, y+\varepsilon h_j).$$

We are ready to construct Burkholder's function B described in the previous section. Let $B : \mathcal{D} \to \mathbb{R}$ be given by

$$\begin{aligned} B(x,y,z,u,v) &= \left[F\left(U\left(\frac{x}{z},\frac{y}{z}\right) + 2(p-1)^pA^p, u, v\right) - 3Apu\right]z \\ &= F(U(x,y) + 2(p-1)^pA^pz^p, u, v) - 3Apuz. \end{aligned}$$

Here the second equality follows from the homogeneity $U(\lambda x, \lambda y) = |\lambda|^p U(x,y)$ and the relation $\beta = 1/p$.

3.1. The analysis of U and F. In this subsection we will prove some properties of the auxiliary functions U and F.

In what follows, we will also need the fact stated below.

LEMMA 3.2. *For any* $\varepsilon \in [-1,1]$, $t \geqslant 0$ *and* $\eta \in \mathbb{R}$,

$$(U(1,\eta)+2(p-1)^pA^p)^{\beta-1}(U(1,\eta)+2(p-1)^pA^p+\beta U_y(1,\eta)(\varepsilon-\eta)) \leqslant 3Ap.$$

Proof. Recall that $\beta = 1/p$. If $|\eta| < p-1$, we use the second formula in the definition of U and calculate that

$$U_y(1,\eta) = p \operatorname{sgn}(\eta)|\eta|^{p-1}.$$

Hence

$$\begin{aligned} U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta) &= \varepsilon|\eta|^{p-1}\operatorname{sgn}(\eta) - (p-1)^p \\ &\leqslant (p-1)^{p-1} - (p-1)^p \leqslant 0 \leqslant p(1+|\eta|)^{p-1}. \end{aligned} \tag{3.2}$$

Next, consider the case $|\eta| \geqslant p-1$. We use the first formula in the definition of U and calculate that

$$U_y(1,\eta) = p(1-1/p)^{p-1}\operatorname{sgn}(\eta)(1+|\eta|)^{p-2}(p|\eta| + p(2-p)|x|).$$

Hence

$$U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta) = p(1-1/p)^{p-1}(1+|\eta|)^{p-2} \\ \times [(|\eta|-(p-1))(1+|\eta|)+\operatorname{sgn}(\eta)\beta p(|\eta|+2-p)(\varepsilon-\eta)].$$

The expression in square brackets is equal to

$$\varepsilon\eta - 1 + (\varepsilon\operatorname{sgn}(\eta)+1)(2-p).$$

Thus if $|\eta| \geqslant p-1$ then

$$\begin{aligned} U(1,\eta)&+\beta U_y(1,\eta)(\varepsilon-\eta) \\ &= p(1-1/p)^{p-1}(1+|\eta|)^{p-2}(\varepsilon\eta - 1 + (\varepsilon\operatorname{sgn}(\eta)+1)(2-p)) \\ &\leqslant p(1-1/p)^{p-1}(1+|\eta|)^{p-1} \leqslant p(1+|\eta|)^{p-1}. \end{aligned} \tag{3.3}$$

In the first inequality above we have used the rough estimates $\varepsilon\eta - 1 \leqslant 1+|\eta|$ and $(\varepsilon\operatorname{sgn}(\eta)+1)(2-p) \leqslant 0$. We have shown in (3.2) and (3.3) that for every $\eta\in\mathbb{R}$,

$$U(1,\eta)+\beta U_y(1,\eta)(\varepsilon-\eta) \leqslant p(1+|\eta|)^{p-1}. \tag{3.4}$$

Hence, from Lemma 3.1(ii) (recall that the exponent $\beta-1=1/p-1$ is negative), it is sufficient to establish that

$$(\eta^p-(p-1)^p+2(p-1)^pA^p)^{\beta-1}(2(p-1)^pA^p+p(1+\eta)^{p-1}) \leqslant 3Ap \tag{3.5}$$

for every $\eta \geqslant 0$. The derivative of the expression on the left with respect to η is equal to

$$p(\eta^p+(2A^p-1)(p-1)^p)^{\beta-2}[\eta^{p-1}(\beta-1)2(p-1)^pA^p+ \\ \eta^{p-1}p(\beta-1)(1+\eta)^{p-1}+(p-1)(1+\eta)^{p-2}(\eta^p-(p-1)^p+2(p-1)^pA^p)].$$

From the identity $p(\beta-1)=1-p$, we find that the expression in square brackets is equal to

$$2(\beta-1)(p-1)^pA^p\eta^{p-1}+(1+\eta)^{p-2}((1-p)\eta^{p-1}+(2A^p-1)(p-1)^p(p-1)).$$

Now we can omit the negative summand $(1+\eta)^{p-2}((1-p)\eta^{p-1}-(p-1)^{p+1})$ and estimate this expression from above by

$$2A^p(p-1)^p((\beta-1)\eta^{p-1}+(1+\eta)^{p-2}(p-1)).$$

This is nonpositive for $\eta \geqslant 4(p-2)$. Indeed,

$$(1+\eta)^{p-2}(p-1) = \eta^{p-2}p(1-\beta)(1+1/\eta)^{p-2} \leqslant \eta^{p-2}p(1-\beta)e^{1/4} \leqslant \eta^{p-2}(1-\beta)p\frac{4(p-2)}{p} \leqslant (1-\beta)\eta^{p-1}.$$

Here we have used the assumption $\eta \geqslant 4(p-2)$ and the bound $p \geqslant 8$. We have proved that the expression on the left-hand side of (3.5) is decreasing for $\eta \geqslant 4(p-2)$. Hence to establish (3.5) it is sufficient to prove this for $\eta \in [0, 4(p-2))$. We estimate the left-hand side of this inequality from above by

$$\big((2A^p-1)(p-1)^p\big)^{\beta-1}\big(2(p-1)^pA^p + p(4p-7)^{p-1}\big) = A(2-A^{-p})^{\beta-1}\left(2(p-1) + p\left(\frac{4p-7}{Ap-A}\right)^{p-1}A^{-1}\right) \leqslant A(2-4^{-p})^{\beta-1}(2(p-1)+pA^{-1}) \leqslant 3Ap.$$

Here we have used the identity $p(\beta-1) = 1-p$, the bound $A \geqslant 4$ and the estimate $2-4^{-p} \geqslant 1$. ∎

Concerning F, we start with the following fact.

LEMMA 3.3. *For any* $(r,u,v) \in (0,\infty)^3$ *with* $1 \leqslant ue^{-v} \leqslant c$, *we have*

$$\tfrac{1}{4}ur^\beta \leqslant F(r,u,v) \leqslant ur^\beta. \tag{3.6}$$

Proof. Recall that $\alpha = 1-(2c)^{-1}$ and $a = 3/4$. We have to show the estimate

$$\frac{1}{4} \leqslant \frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leqslant 1.$$

Let $t = ue^{-v}$. Observe that the function $[1,c] \ni t \mapsto (t-a)^\alpha/t$ is increasing:

$$\left(\frac{(t-a)^\alpha}{t}\right)' = \frac{(t-a)^{\alpha-1}((\alpha-1)t+a)}{t^2} \geqslant 0.$$

Thus the assertion follows from $1/4 \leqslant (1-a)^\alpha$ and $(c-a)^\alpha/c \leqslant 1$. ∎

LEMMA 3.4. *The function F is θ-concave: for any* $x, x_1, \ldots, x_n \in (0,\infty)\times D_c$ *and any sequence* $(\alpha_j)_{j=1}^n$ *satisfying*

$$\alpha_j \in [\theta, 1), \quad \sum_{j=1}^n \alpha_j = 1, \quad \sum \alpha_j x_j = x,$$

we have

$$F(x) \geqslant \sum \alpha_j F(x_j).$$

Proof. By the homogeneity of F we may assume that $v = 0$. In other words, it is sufficient to prove

$$r^{\beta}(u-a)^{\alpha} \geqslant \sum \alpha_j r_j^{\beta}(u_j e^{-v_j} - a)^{\alpha} e^{v_j}, \tag{3.7}$$

where $\sum \alpha_j (r_j, u_j, v_j) = (r, u, 0)$ and $(r, u, 0), (r_1, u_1, v_1), \dots (r_n, u_n, v_n) \in (0,\infty) \times D_c$. Because $\alpha + \beta + \beta < 1$, the function $(0,\infty) \times (1,\infty) \times (0,\infty) \ni (k, s, t) \mapsto k^{\beta}(s-a)^{\alpha} t^{\beta}$ is concave. Hence

$$\sum (r_j e^{-v_j})^{\beta}(u_j e^{-v_j} - a)^{\alpha}(e^{v_j})^{\beta} \frac{e^{v_j}\alpha_j}{P} \leqslant \left(\frac{r}{P}\right)^{\beta}\left(\frac{u}{P} - a\right)^{\alpha}\left(\frac{Q}{P}\right)^{\beta},$$

where $P = \sum \alpha_j e^{v_j}$ and $Q = \sum \alpha_j e^{2v_j}$. Thus to prove (3.7) it is sufficient to establish the inequality

$$P^{1-\beta}\left(\frac{u}{P} - a\right)^{\alpha}\left(\frac{Q}{P}\right)^{\beta} \leqslant (u-a)^{\alpha}. \tag{3.8}$$

We will need the following estimate:

$$Q \leqslant \frac{1}{\theta}P^2 - \frac{1-\theta}{\theta}.$$

This follows from the assumption $\alpha_j \in [\theta, 1)$ by applying the convexity of e^x twice:

$$\begin{aligned} P^2 - \theta Q &= \sum_j \alpha_j e^{v_j}\Big((\alpha_j - \theta)e^{v_j} + \sum_{k \neq j} \alpha_k e^{v_k}\Big) \geqslant \sum_j \alpha_j e^{v_j}(1-\theta)e^{-v_j\theta/(1-\theta)} \\ &\geqslant 1 - \theta. \end{aligned}$$

Hence to prove (3.8) it is sufficient to establish that

$$P\left(\frac{u}{P} - a\right)^{\alpha}\left(\frac{1}{\theta} - \frac{1-\theta}{\theta P^2}\right)^{\beta} \leqslant (u-a)^{\alpha}.$$

We know that $u \in [1, c]$ and $P \in [1, u]$ (here the lower bound is just convexity of e^x and the upper bound follows from $u_j e^{-v_j} \geqslant 1$ for each j). Let $s = 1/P$. It is enough to show that

$$s^{-1}(us - a)^{\alpha}(1 - (1-\theta)s^2)^{\beta} \leqslant \theta^{\beta}(u-a)^{\alpha}$$

for any $u \in [1, c]$ and $s \in [1/u, 1]$. Observe that for $s = 1$ both sides are equal. Hence it is sufficient to show that $s \mapsto s^{-1}(us-a)^{\alpha}(1-(1-\theta)s^2)^{\beta}$ is nondecreasing. By differentiating we obtain the condition

$$((\alpha - 1)us + a)(1 - (1-\theta)s^2) - 2\beta(1-\theta)s^2(us - a) \geqslant 0.$$

Since $s \leqslant 1$, the expression on the left is greater than

$$((\alpha-1)us+a)\theta-2\beta(1-\theta)(us-a)$$
$$=\left(\alpha-1-2\beta\frac{1-\theta}{\theta}\right)\theta us+\left(1+2\beta\frac{1-\theta}{\theta}\right)\theta a \geqslant \left(\alpha-1-2\beta\frac{1-\theta}{\theta}\right)\theta c+\theta a=0.$$

Here in the last step we have just plugged in the values $a=3/4, \alpha=1-1/2c$ and $\beta=\theta(8c(1-\theta)^{-1}$. ∎

REMARK 3.1. It can be shown that the regularity assumption $\alpha_j \geqslant \theta$ is necessary here. In other words, the function F does not satisfy the concavity condition if we do not assume any lower bound on α_j.

3.2. Burkholder's function B of five variables. We are ready for the main step: we will check that the function B satisfies conditions 0°, 1° and 2°.

LEMMA 3.5. *The function B satisfies the initial condition* 0°.

Proof. Recall condition 0°: for every $(x,y,|x|,u,v) \in \mathcal{D}$ such that $|y| \leqslant |x|$ and $1 \leqslant ue^{-v} \leqslant c$ we have

$$B(x,y,|x|,u,v) \leqslant 0.$$

From the definition of B this is equivalent to

$$\left[F\left(U\left(\frac{x}{|x|},\frac{y}{|x|}\right)+2(p-1)^pA^p,u,v\right)-3Apu\right]|x| \leqslant 0. \tag{3.9}$$

From Lemma 3.3 we have

$$\frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leqslant 1.$$

Recall that $p\beta=1$. From Lemma 3.1(i), if $|y| \leqslant |x|$, then $U(x/|x|,y/|x|) \leqslant 0$ and hence

$$\left(U\left(\frac{x}{|x|},\frac{y}{|x|}\right)+2(p-1)^pA^p\right)^\beta \frac{(ue^{-v}-a)^\alpha}{ue^{-v}} \leqslant 2A(p-1) \leqslant 3Ap,$$

which is precisely the required estimate (3.9). ∎

LEMMA 3.6. *The function B satisfies the majorization condition*

$$B(x,y,z,u,v) \geqslant \tfrac{1}{4}(|y|u-12Apzu).$$

Proof. From Lemma 3.1(ii), the estimate $|x|/z \leqslant 1 \leqslant A$ and the identity $p\beta = 1$ we have

$$\left(U\left(\frac{x}{z}, \frac{y}{z}\right) + 2(p-1)^p A^p\right)^\beta \geqslant \left(\left(\frac{|y|}{z}\right)^p - (p-1)^p \left(\frac{|x|}{z}\right)^p + 2(p-1)^p A^p\right)^\beta \geqslant \frac{|y|}{z}$$

and, as we have proved in Lemma 3.3, also

$$\frac{(ue^{-v} - a)^\alpha}{ue^{-v}} \geqslant \frac{1}{4}. \tag{3.10}$$

Consequently,

$$B(x, y, z, u, v) \geqslant \tfrac{1}{4}|y|u - 3Apuz = \tfrac{1}{4}(|y|u - 12Apzu). \quad \blacksquare$$

It remains to check the most difficult condition 2°. Recall that we need to show that the function B satisfies the following concavity-type condition: for any $(x, y, z, u, v) \in \mathcal{D}$, any $\varepsilon \in [-1, 1]$, any positive integer $k \leqslant 1/\theta$ and any sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying

$$\begin{aligned}
&\alpha_j \in [\theta, 1), \quad \sum_{j=1}^k \alpha_j = 1, \\
&\sum_{j=1}^k \alpha_j h_j = \sum_{j=1}^k \alpha_j r_j = \sum_{j=1}^k \alpha_j s_j = 0, \\
&(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j) \in \mathcal{D},
\end{aligned}$$

we have

$$B(x, y, z, u, v) \geqslant \sum \alpha_j B(x + h_j, y + \varepsilon h_j, |x + h_j| \vee z, u + r_j, v + s_j). \tag{3.11}$$

We have already established that the auxiliary functions U and F have appropriate concavity properties (Lemmas 3.1 and 3.4). From this it is almost immediate to deduce that

$$B(x, y, z, u, v) \geqslant \sum \alpha_j B(x + h_j, y + k_j, z, u + r_j, v + s_j)$$

for points satisfying the additional condition $|x + h_j| \leqslant z$. The main difficulty is to prove (3.11) when $|x + h_j| > z$ for some j. To solve this problem we will consider the extension of B onto the domain

$$\bar{\mathcal{D}} = \{(x, y, z, u, v) \in \mathbb{R}^2 \times (0, \infty) \times (0, \infty) \times \mathbb{R} : |x| \leqslant Az, 1 \leqslant ue^{-v} \leqslant c\}.$$

The function $\bar{B}:\bar{\mathcal{D}}\to\mathbb{R}$ will be given by the same formula:

$$\bar{B}(x,y,z,u,v)=\left[F\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2A^p(p-1)^p,u,v\right)-3Apu\right]z.$$

In the next theorem we will prove the concavity and monotonicity properties of $\bar{B}$.

THEOREM 3.1. *The function $\bar{B}$ has the following properties:*

(1) (Concavity-type property) *For any $(x,y,z,u,v)\in\bar{\mathcal{D}}$, any $\varepsilon\in[-1,1]$, any positive integer $k\leqslant 1/\theta$ and sequences $(\alpha_j)_{j=1}^k,(h_j)_{j=1}^k,(r_j)_{j=1}^k,(s_j)_{j=1}^k$ satisfying*

$$\alpha_j\in[\theta,1),\qquad \sum_{j=1}^k\alpha_j=1,$$
$$\sum_{j=1}^k\alpha_jh_j=\sum_{j=1}^k\alpha_jr_j=\sum_{j=1}^k\alpha_js_j=0,$$
$$(x+h_j,y+\varepsilon h_j,z,u+r_j,v+s_j)\in\bar{\mathcal{D}},$$

we have

$$\bar{B}(x,y,z,u,v)\geqslant\sum\alpha_j\bar{B}(x+h_j,y+\varepsilon h_j,z,u+r_j,v+s_j). \tag{3.12}$$

(2) (Vertical monotonicity) $\bar{B}_z(x,y,z,u,v)\leqslant 0$ *for every* $(x,y,z,u,v)\in\bar{\mathcal{D}}$.

(3) (Diagonal monotonicity) *Let $(x_1,y_1,|x_1|,u,v),(x_2,y_2,|x_2|,u,v)\in\mathcal{D}$. If $|x_2|<|x_1|$ and $|y_2-y_1|\leqslant|x_2-x_1|$, then*

$$B(x_1,y_1,|x_1|,u,v)\leqslant B(x_2,y_2,|x_2|,u,v).$$

Proof. (1) follows immediately from Lemmas 3.1 and 3.4 and the monotonicity of F with respect to the variable r:

$$\begin{aligned}&F\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2A^p(p-1)^p,u,v\right)\\&\quad\geqslant F\left(\sum\alpha_jU\left(\frac{x+h_j}{z},\frac{y+\varepsilon h_j}{z}\right)+2A^p(p-1)^p,u+\sum\alpha_jr_j,v+\sum\alpha_js_j\right)\\&\quad\geqslant\sum\alpha_jF\left(U\left(\frac{x+h_j}{z},\frac{y+\varepsilon h_j}{z}\right)+2A^p(p-1)^p,u+r_j,v+s_j\right),\end{aligned}$$

which is equivalent to the desired inequality.

It remains to show the monotonicity properties. We start with (2). By symmetry, we may assume that $x\geqslant 0$. Because $\beta=1/p$, the condition $\bar{B}_z(x,y,z,u,v)\leqslant 0$ is equivalent to

$$\left(U\left(\frac{x}{z},\frac{y}{z}\right)+2A^p(p-1)^p\right)^{\beta-1}2A^p(p-1)^p\frac{(ue^{-v}-a)^\alpha}{ue^{-v}}\leqslant 3Ap.$$

From Lemmas 3.1(ii) and 3.3 the left hand side is smaller than

$$\left(\left(\frac{|y|}{|z|}\right)^p - (p-1)^p\left(\frac{|x|}{|z|}\right)^p + 2A^p(p-1)^p\right)^{\beta-1} 2A^p(p-1)^p \leqslant 2Ap.$$

This gives the assertion.

To handle (3), we first apply symmetry and homogeneity to assume that $x_1 > 0$ and $x_2 = 1$. Consider the function $\phi : [0,\infty) \to \mathbb{R}$ given by

$$\phi(t) = B(1+t, y+\varepsilon t, 1+t, u, v),$$

where $\varepsilon \in [-1,1]$ and $(u,v) \in D_c$ are fixed. It is sufficient to show that $\phi'(t) \leqslant 0$. This is equivalent to proving that the expression

$$(U(1,\eta) + 2(p-1)^p A^p)^{\beta-1} \\ \times (U(1,\eta) + 2(p-1)^p A^p + \beta U_y(1,\eta)(\varepsilon-\eta))\frac{(ue^{-v}-a)^\alpha}{ue^{-v}}u - 3Apu,$$

where $\eta = (y+\varepsilon t)/(1+t)$, is nonpositive. This follows from Lemmas 3.2 and 3.3. ■

We are ready to prove that the function B satisfies the concavity-type condition.

THEOREM 3.2. *The function B satisfies the following concavity-type condition: for any $(x,y,z,w,v) \in \mathcal{D}$, any $\varepsilon \in [-1,1]$, any positive integer $k \leqslant 1/\theta$ and sequences $(\alpha_j)_{j=1}^k, (h_j)_{j=1}^k, (r_j)_{j=1}^k, (s_j)_{j=1}^k$ satisfying*

$$\alpha_j \in [\theta,1), \qquad \sum_{j=1}^k \alpha_j = 1,$$
$$\sum_{j=1}^k \alpha_j h_j = \sum_{j=1}^k \alpha_j r_j = \sum_{j=1}^k \alpha_j s_j = 0,$$
$$(x+h_j, y+\varepsilon h_j, |x+h_j| \vee z, u+r_j, v+s_j) \in \mathcal{D},$$

we have

$$B(x,y,z,u,v) \geqslant \sum \alpha_j B(x+h_j, y+\varepsilon h_j, |x+h_j| \vee z, u+r_j, v+s_j). \tag{3.13}$$

Proof. The function B satisfies the following homogeneity condition:

$$B(\lambda x, \lambda y, \lambda z, u, v) = \lambda B(x,y,z,u,v)$$

for every $\lambda > 0$. Hence, we can divide both sides of (3.13) by z to obtain the equivalent condition:

$$B(x/z, y/z, 1, u, v) \\ \geqslant \sum \alpha_j B(x/z + h_j/z, y/z + \varepsilon h_j/z, |x/z + h_j/z| \vee 1, u+r_j, v+s_j).$$

Thus, to prove the general case, it is sufficient to prove the statement for $z = 1$.

For convenience let $x_j = x + h_j$, $y_j = y + \varepsilon h_j$, $u_j = u + r_j$ and $v_j = v + s_j$ and sort the points in increasing order: $x_1 \leqslant \cdots \leqslant x_k$. We will divide the proof into two steps.

STEP 1. Let us consider two special cases: when $x_1 \geqslant -1$ and when $x_k \leqslant 1$. By symmetry it is enough to handle the first one; the second is analogous. From $x_1 \geqslant -1$ and the lower bound on probabilities we can deduce that x_n cannot be large. More precisely,

$$\begin{aligned} x_k &= \Big(x - \sum_{j=1}^{k-1} \alpha_j x_j\Big)/\alpha_k \leqslant (x + 1 - \alpha_k)/\alpha_k = (x+1)/\alpha_k - 1 \\ &\leqslant 2/\theta - 1 \leqslant A. \end{aligned}$$

Hence $(x_j, y_j, 1, u_j, v_j) \in \bar{\mathcal{D}}$ and from Theorem 3.1(2) we obtain

$$\bar{B}(x_j, y_j, |x_j| \vee 1, u_j, v_j) \leqslant \bar{B}(x_j, y_j, 1, u_j, v_j).$$

Combining this with Theorem 3.1(1), we get

$$\begin{aligned} \sum \alpha_j B(x_j, y_j, |x_j| \vee 1, u_j, v_j) &\leqslant \sum \alpha_j \bar{B}(x_j, y_j, 1, u_j, v_j) \\ &\leqslant \bar{B}(x, y, 1, u, v) = B(x, y, 1, u, v). \end{aligned}$$

STEP 2. We now reduce the general case to the one considered before. Assume that $x_1 < -1$ and $x_k > 1$. The idea is to replace x_1, y_1, x_k, y_k by $\hat{x}_1, \hat{y}_1, \hat{x}_k, \hat{y}_k$ in such a way that:

(a) We "pull" the points closer to the center: $\hat{x}_1 \in (x_1, -1]$ and $\hat{x}_k \in [1, x_k)$.

(b) We have $\hat{y}_1 - y_1 = \varepsilon(\hat{x}_1 - x_1)$ and $\hat{y}_k - y_k = \varepsilon(\hat{x}_k - x_k)$.

(c) The average is preserved: $\alpha_1 x_1 + \alpha_k x_k = \alpha_1 \hat{x}_1 + \alpha_k \hat{x}_k$.

Then in the light of Theorem 3.1(3),

$$\begin{aligned} B(x_1, y_1, |x_1|, u_1, v_1) &\leqslant B(\hat{x}_1, \hat{y}_1, |\hat{x}_1|, u_1, v_1), \\ B(x_k, y_k, |x_k|, u_k, v_k) &\leqslant B(\hat{x}_k, \hat{y}_k, |\hat{x}_k|, u_k, v_k). \end{aligned}$$

Hence the replacement does not change the left hand side of (3.13) and does not decrease the right hand side, making the inequality stronger. Moreover we will also ensure that

(d) We "pull" the points as close as possible: $\hat{x}_1 = -1$ or $\hat{x}_k = 1$.

Now we repeat the replacement procedure until all the first coordinates $x_1, \dots, x_k$ are contained either in $[-1, \infty)$ or in $(-\infty, 1]$, which is the case solved in Step 1. Condition (d) ensures that this algorithm will stop after at most $n-1$ replacements. It remains to find the points $\hat{x}_1, \hat{y}_1, \hat{x}_k, \hat{y}_k$ satisfying (a)–(d). This will be done explicitly. Let us consider two cases. If

$$\alpha_1 x_1 + \alpha_k x_k \geqslant \alpha_k - \alpha_1,$$

then we put $\hat{x}_1 = -1$, $\hat{x}_k = (\alpha_1 x_1 + \alpha_k x_k + \alpha_1)/\alpha_k$, $\hat{y}_1 = \varepsilon(\hat{x}_1 - x_1) + y_1$ and $\hat{y}_k = \varepsilon(\hat{x}_k - x_k) + y_k$. Conditions (a)–(d) are easy to check. The case

$$\alpha_1 x_1 + \alpha_k x_k < \alpha_k - \alpha_1$$

is analogous: we put $\hat{x}_k = 1$, $\hat{x}_1 = (\alpha_1 x_1 + \alpha_k x_k - \alpha_k)/\alpha_1$, $\hat{y}_1 = \varepsilon(\hat{x}_1 - x_1) + y_1$ and $\hat{y}_k = \varepsilon(\hat{x}_k - x_k) + y_k$. Again it is easy to check the required conditions. This completes the proof. ■

We have shown that B satisfies conditions $0°-2°$. By the method of Section 2, this yields the estimate (3.1) with $C = 12Apc^{-1} \leqslant 384\theta^{-2}$.

4. PROOF OF THE MAIN INEQUALITY

To prove the main inequality (1.3) we will construct a function of *six* variables. Let

$$\mathfrak{D} = \{(x, y, z, r, u, v) \in \mathbb{R}^2 \times (0, \infty) \times \mathbb{R} \times (0, \infty) \times \mathbb{R} : \\ |x| \leqslant z, y \leqslant r,\ 1 \leqslant ue^{-v} \leqslant c\}.$$

The additional variable r is associated with the one-sided maximal function defined as $g_n^* = \sup_{n \geqslant 0} g_n$. We define Burkholder's function $\mathfrak{B} : \mathfrak{D} \to \mathbb{R}$ by

$$\mathfrak{B}(x, y, z, r, u, v) = \left[F\left(U\left(\frac{x}{z}, \frac{r-y}{z}\right) + 2A^p(p-1)^p, u, v\right) - 12cu\right] z \\ = B(x, r-y, z, u, v).$$

This new function satisfies the following properties:

$0°$ (Initial condition) $\mathfrak{B}(x, y, |x|, y, u, v) \leqslant 0$ if $1 \leqslant ue^{-v} \leqslant c$.

$1°$ (Majorization property) For any $(x, y, z, r, u, v) \in \mathfrak{D}$ we have

$$\mathfrak{B}(x, y, z, r, u, v) \geqslant \tfrac{1}{4}((r-y)u - 12Apzu).$$

2° (Concavity-type property) For any $(x,y,z,r,u,v)\in\mathfrak{D}$, any $\varepsilon\in[-1,1]$, any positive integer $k\leqslant 1/\theta$ and sequences $(\alpha_j)_{j=1}^k$, $(h_j)_{j=1}^k$, $(r_j)_{j=1}^k$, $(s_j)_{j=1}^k$ satisfying

$$\alpha_j\in[\theta,1),\qquad \sum_{j=1}^k\alpha_j=1,$$
$$\sum_{j=1}^k\alpha_jh_j=\sum_{j=1}^k\alpha_jr_j=\sum_{j=1}^k\alpha_js_j=0,$$
$$(x+h_j,y+\varepsilon h_j,|x+h_j|\vee z,(y+\varepsilon h_j)\vee r,u+r_j,v+s_j)\in\mathfrak{D},$$

we have

$$\mathfrak{B}(x,y,z,r,u,v)\geqslant\sum\alpha_j\mathfrak{B}(x+h_j,y+\varepsilon h_j,|x+h_j|\vee z,(y+\varepsilon h_j)\vee r,u+r_j,v+s_j).$$

Conditions 0° and 1° are immediate consequences of the analogous properties of B. Now consider the concavity-type condition. It is easy to check that Burkholder's function U satisfies $U(x,y)\geqslant U(x,0)$ for every $(x,y)\in\mathbb{R}^2$. Hence

$$B(x,y,z,u,v)\geqslant B(x,0,z,u,v)$$

for every $(x,y,z,u,v)\in\mathcal{D}$. From the above estimate and (3.13) we have

$$\begin{aligned}\sum\alpha_j\mathfrak{B}\big(&x+h_j,y+\varepsilon h_j,|x+h_j|\vee z,(y+\varepsilon h_j)\vee r,u+r_j,v+s_j\big)\\&=\sum\alpha_jB\big(x+h_j,(r-y-\varepsilon h_j)\vee 0,|x+h_j|\vee z,u+r_j,v+s_j\big)\\&\leqslant\sum\alpha_jB(x+h_j,r-y-\varepsilon h_j,|x+h_j|\vee z,u+r_j,v+s_j)\\&\leqslant B(x,r-y,z,u,v)=\mathfrak{B}(x,y,z,r,u,v).\end{aligned}$$

Now we repeat, word-for-word, the reasoning of Section 2: the only change is that the process $(z_n)_{n\geqslant 0}$ is six-dimensional and involves the one-sided maximal function of g: $z_n=(f_n,g_n,|f|_n^*,g_n^*,w_n,\sigma_n)$. Hence, we obtain

$$\mathbb{E}(g_n^*-g_n)w_n\leqslant 12Ap[w]_{A_\infty}\mathbb{E}|f|_n^*w_n\leqslant 384\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^*w_n$$

and, by symmetry, $\mathbb{E}((-g)_n^*+g_n)w_n\leqslant 384\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^*w_n$. Add these two bounds to get

$$\mathbb{E}(g_n^*+(-g)_n^*)w_n\leqslant 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^*w_n. \tag{4.1}$$

Now observe that if g started from 0, we would have the pointwise inequality $|g|_n^*\leqslant g_n^*+(-g)_n^*$ and (4.1) would give

$$\mathbb{E}|g|_n^*w_n\leqslant 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^*w_n,$$

as desired (see the limiting argument below). To prove (1.3) in full generality, note that if $(g_n)_{n\geqslant 0}$ is a ± 1-transform of f, then the martingale $\tilde{g} = (g_n - g_0)_{n\geqslant 0}$ also has this property and additionally starts from 0. Hence by the above estimate,

$$\begin{aligned}\mathbb{E}|g|_n^* w_n &\leqslant \mathbb{E}(|\tilde{g}|_n^* + |g_0|)w_n \leqslant 768\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n + \mathbb{E}|f_0|w_n \\ &\leqslant 769\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w_n.\end{aligned}$$

Since $w_n = \mathbb{E}(w \,|\, \mathcal{F}_n)$, this gives $\mathbb{E}|g|_n^* w \leqslant 769\theta^{-2}[w]_{A_\infty}\mathbb{E}|f|_n^* w$ and the claim follows by letting $n \to \infty$ and applying Lebesgue's monotone convergence theorem.

5. NECESSITY OF THE θ-REGULARITY CONDITION

The purpose of this section is to establish Theorem 1.3, and from now on we work with dyadic filtrations only. We could prove the theorem by constructing appropriate examples, but these seem to have quite involved, fractal-type structure and their analysis is a little complicated. Our approach will rest on Remark 2.2, which enables us to avoid most of these technical issues. Roughly speaking, the argument is as follows. First we assume that, on the contrary, the inequality does hold universally, i.e., with a constant independent of the dimension. Then the Bellman method yields the existence of an abstract function satisfying the appropriate size and concavity requirements. Finally, we exploit these properties in the right order to obtain a contradiction (to the assumption that the constant involved is dimension-free).

So, suppose that there is $1 < p < \infty$ and a constant K depending only on p such that for any dimension d, any martingales f and g adapted to the d-dimensional dyadic filtration on $[0,1)^d$ such that $dg_n = v_n dg_n$ for a predictable sequence of signs v_n, and any A_p weight w on $[0,1)^d$ with $[w]_{A_p} \leqslant 2$, we have

$$\|g\|_{L^1(w)} \leqslant K\big\||f|^*\big\|_{L^1(w)}. \tag{5.1}$$

Fix d and let B be the associated Bellman function, given by

$$B(x,y,z,u,v) = \sup \mathbb{E}\{|g_n|w - K(|f_n|^* \vee z)w\}.$$

Here the probability space is $([0,1]^d, \mathcal{B}([0,1]^d), |\cdot|)$, the filtration is dyadic and the above supremum is taken over:

- all adapted martingale pairs (f,g) satisfying $(f_0,g_0) = (x,y)$ and $dg_k = v_k df_k$ for all $k \geqslant 1$, for some deterministic sequence $v_1, v_2, \ldots$ of signs;
- all dyadic A_p weights w satisfying $[w]_{A_p} \leqslant 2$, $\mathbb{E}w = u$ and $\mathbb{E}w^{1/(1-p)} = v$.

This Bellman function enjoys the appropriate initial, majorization and concavity conditions, proved in Section 2. We will also need the following additional properties which follow from the special form of the function M.

THEOREM 5.1.

(i) *We have*

$$B(x,y,z,u,v)=B(|x|,|y|,|x|\vee z,u,v). \tag{5.2}$$

(ii) *For any* $\lambda\neq 0$ *and any* $\mu>0$ *we have*

$$B(\lambda x,\lambda y,|\lambda|z,\mu u,\mu^{-1/(p-1)}v)=|\lambda|\mu B(x,y,z,u,v). \tag{5.3}$$

(iii) *We have*

$$B(x,y,z,u,v)\geqslant B(x,0,z,u,v). \tag{5.4}$$

Proof. The symmetry $B(x,y,z,u,v) = B(|x|,|y|,z,u,v)$ follows directly from the definition. Indeed, if f, g, w are arbitrary martingales as in the definition of $B(x,y,z,u,v)$, then $-f,g,w$ satisfy all the requirements needed in the definition of $B(-x,y,z,u,v)$, so

$$B(-x,y,z,u,v)\geqslant \mathbb{E}\{|-g_n|w-K(|f_n|^*\vee z)w\}=\mathbb{E}\{|-g_n|w-K(|f_n|^*\vee z)w\}.$$

Taking the supremum over all f,g,w we get $B(-x,y,z,u,v)\geqslant B(x,y,z,u,v)$, and the passage from x to $-x$ shows that we actually have equality here. The identity $B(x,y,z,u,v)=B(x,-y,z,u,v)$ is shown in the same manner, and the equality $B(|x|,|y|,z,u,v)=B(|x|,|y|,|x|\vee z,u,v)$ follows from the fact that

$$\mathbb{E}\{|g_n|w-K(|f_n|^*\vee z)w\}=\mathbb{E}\{|g_n|w-K(|f_n|^*\vee |f_0|\vee z)w\}.$$

The proof of the homogeneity property (ii) is analogous: pick arbitrary martingales f, g, w as in the definition of $B(x,y,z,u,v)$. Then λf has average λx, λg has average λy, while μw is an A_p weight with characteristic bounded by 2 satisfying $\mathbb{E}\mu w=\mu u$ and $\mathbb{E}(\mu w)^{-1/(p-1)}=\mu^{-1/(p-1)}v$. Consequently,

$$\begin{aligned}B(\lambda x,\lambda y,|\lambda|z,\mu u,\mu^{-1/(p-1)}v)&\geqslant \mathbb{E}\{|\lambda g_n|(\mu w)-K(|\lambda f_n|^*\vee|\lambda z|)(\mu w)\}\\&=|\lambda|\mu\mathbb{E}\{|g_n|w-K(|f_n|^*\vee z)w\}.\end{aligned}$$

Hence, taking the supremum over all f,g,w as above, we get

$$B(\lambda x,\lambda y,|\lambda|z,\mu u,\mu^{-1/(p-1)}v)\geqslant |\lambda|\mu B(x,y,z,u,v). \tag{5.5}$$

To get the reverse bound, apply the above estimate to $(\lambda x,\lambda y,|\lambda|z,\mu u,\mu^{-1/(p-1)}v)$ in place of (x,y,z,u,v) and the numbers λ^{-1}, μ^{-1} in place of λ, μ.

Finally, to check (iii), we will prove that the function $y\mapsto B(x,y,z,u,v)$ is convex; together with its symmetry (which is guaranteed by (i)), we will get the claim. Pick $\alpha \in (0,1)$, $y_1, y_2 \in \mathbb{R}$ and set $y = \alpha_1 y_1 + (1-\alpha_1)y_2$. If f, g, w are martingales as in the definition of $B(x,y,z,u,v)$, then the convexity of the function $t \mapsto |t|$ yields

$$\begin{aligned}\mathbb{E}\{|g_n|w - K(|f_n|^* \vee z)w\} &\leqslant \alpha_1 \mathbb{E}\{|y_1 - y + g_n|w - K(|f_n|^* \vee z)w\} \\ &\quad + \alpha_2 \mathbb{E}\{|y_2 - y + g_n|w - K(|f_n|^* \vee z)w\} \\ &\leqslant \alpha_1 B(x, y_1, z, u, v) + \alpha_2 B(x, y_2, z, u, v).\end{aligned}$$

Taking the supremum over all f, g, w, n gives the desired convexity. ∎

We will exploit the concavity of B in appropriate directions; to this end, we need the following auxiliary geometrical fact, taken from [24]. We provide an easy proof for the sake of completeness.

LEMMA 5.1. *Suppose that N is a huge positive integer, $u = 1$ and $v = 2^{1/(p-1)}$. Then there are $R, T \in \mathbb{R}^2$ such that $R = (R_x, R_y)$ lies on the curve $xy^{p-1} = 2$, $T = (T_x, T_y)$ lies on the curve $xy^{p-1} = 1$, $R_x \leqslant T_x$ and*

$$(1 - (1-2^{-d})^N)R + (1-2^{-d})^N T = (u, v). \tag{5.6}$$

Furthermore,

$$(1 - (1-2^{-d})^N)2^d R_x < 1/2 \tag{5.7}$$

provided d is sufficiently large.

Proof. The existence of R, T follows from a very simple continuity argument. Pick any point $R = (R_x, R_y)$ on the curve $xy^{p-1} = 2$ such that $R_x \leqslant u$ and let T be defined by (5.6) (then of course $R_x \leqslant u \leqslant T_x$). Note that T is a continuous function of R. Furthermore, if R_y is huge, then T_y is negative, so T lies below the curve $xy^{p-1} = 1$. On the other hand, when $R_y = v$, then $R = T = (u, v)$, so T lies above the curve $xy^{p-1} = 1$. Thus, by the Darboux property, there must be a point R for which the desired configuration is satisfied.

To show (5.7), we exploit (5.6). Recall that $u = 1$. We have

$$1 = (1 - (1-2^{-d})^N)R_x + (1-2^{-d})^N T_x,$$

and since $R_x < 1 < T_x$,

$$\begin{aligned}2^{1/(p-1)} &= (1 - (1-2^{-d})^N)\left(\frac{2}{R_x}\right)^{1/(p-1)} + (1-2^{-d})^N T_x^{-1/(p-1)} \\ &< (1 - (1-2^{-d})^N)\left(\frac{2}{R_x}\right)^{1/(p-1)} + (1-2^{-d})^N,\end{aligned}$$

which implies

$$R_x < \left(\frac{1-(1-2^{-d})^N}{1-(1-2^{-d})^N/2^{1/(p-1)}}\right)^{p-1}.$$

Thus if $d \to \infty$, then $R_x \to 0$; on the other hand $(1-(1-2^{-d})^N)2^d \leqslant N$ for each d. This proves the assertion. ∎

Let u, v, R and T be as in (ii) above. In what follows, we will also exploit the points $T_0, T_1, \ldots, T_N$ given by $T_0 = (u, v)$ and the recursive equation

$$T_k = 2^{-d}R + (1-2^{-d})T_{k+1}. \tag{5.8}$$

By straightforward induction, we see that $(u,v) = (1-2^{-d})^k T_k + (1-(1-2^{-d})^k)R$ for each k and hence in particular $T_N = T$.

Proof of Theorem 1.3. If $x \in \mathbb{R}$, $y \in \mathbb{R}$, $z \geqslant 0$ and $P = (u, v) \in \mathcal{D}$, we will sometimes write $B(x, y, z; P) = B(x, y, z, u, v)$. Let $\bar{x} = 1/(2^{d+1}-1)$. As shown in Section 2 (see Remark 2.2), the function B satisfies the initial condition 0°: for every $|y| \leqslant |x|$ we have $B(x, y, |x|, u, v) \leqslant 0$. This condition combined with Theorem 5.1(iii) gives

$$0 \geqslant B(1,1,1,1,2^{1/(p-1)}) \geqslant B(1,0,1,1,2^{1/(p-1)}). \tag{5.9}$$

Next, the concavity property combined with (5.8) yields, for each k,

$$\begin{aligned} &B(\bar{x}, 2k\bar{x}, \bar{x}; T_k) \\ &\quad \geqslant 2^{-d}B(1, (2k+1)\bar{x}-1, \bar{x}; R) + (1-2^{-d})B(-\bar{x}, 2(k+1)\bar{x}, \bar{x}; T_{k+1}). \end{aligned}$$

By Theorem 5.1(i), this expression is equal to

$$2^{-d}B(1, (2k+1)\bar{x}-1, 1; R) + (1-2^{-d})B(\bar{x}, 2\bar{x}(k+1), \bar{x}; T_{k+1}),$$

which by (ii) and (iii) is no smaller than

$$2^{-d}R_x B(1,0,1,1,2^{1/(p-1)}) + (1-2^{-d})B(\bar{x}, 2\bar{x}(k+1), \bar{x}; T_{k+1}).$$

Hence, by induction,

$$\begin{aligned} \bar{x}B(1,0,1,1,2^{1/(p-1)}) &= B(\bar{x}, 0, \bar{x}; T_0) \\ &\geqslant (1-2^{-d})^N B(\bar{x}, 2\bar{x}N, \bar{x}; T_N) \\ &\quad + \sum_{k=0}^{N-1}(1-2^{-d})^k 2^{-d} R_x B(1,0,1,1,2^{1/(p-1)}) \\ &= (1-2^{-d})^N T_x \bar{x} B(1, 2N, 1, 1, 1) \\ &\quad + (1-(1-2^{-d})^N) R_x B(1,0,1,1,2^{1/(p-1)}). \end{aligned}$$

Now we assume that d is large; if we apply (5.9) and (5.7), we obtain

$$(5.10)\quad B(1,2N,1,1,1) \leqslant \frac{\bar{x}-(1-(1-2^{-d})^N)R_x}{(1-2^{-d})^N T_x \bar{x}} B(1,0,1,1,2^{1/(p-1)}) \leqslant 0.$$

As shown in Section 2 (see Remark 2.2), the function B satisfies the majorization condition 1°: $B(x,y,z,u,v) \geqslant |y|u - Kzu$, where K is a finite constant in our key assumption (5.1). Hence, the left-hand side of (5.10) is greater than $2N-K$. This implies $2N-K \leqslant 0$, a contradiction, since N was arbitrary. ■

Acknowledgments. The authors would like to thank the referee for the careful reading of the first version of the paper and many helpful comments.

REFERENCES

[1] R. Bañuelos and K. Bogdan, *Lévy processes and Fourier multipliers*, J. Funct. Anal. 250 (2007), 197–213.

[2] R. Bañuelos and P. J. Méndez-Hernandez, *Space-time Brownian motion and the Beurling–Ahlfors transform*, Indiana Univ. Math. J. 52 (2003), 981–990.

[3] R. Bañuelos and G. Wang, *Sharp inequalities for martingales with applications to the Beurling–Ahlfors and Riesz transforms*, Duke Math. J. 80 (1995), 575–600.

[4] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. 340 (1993), 253–272.

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

[6] D. L. Burkholder, *Sharp inequalities for martingales and stochastic integrals*, in: Colloque Paul Lévy (Palaiseau, 1987), Astérisque 157–158 (1988), 75–94.

[7] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, in: Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994), Proc. Sympos. Appl. Math. 52, Amer. Math. Soc., Providence, RI, 1997, 343–358.

[8] C. Dellacherie and P. A. Meyer, *Probabilities and Potential B*, North-Holland, Amsterdam, 1982.

[9] K. Domelevo and S. Petermichl, *Sharp L_p estimates for discrete second order Riesz transforms*, Adv. Math. 262 (2014), 932–952.

[10] S. Geiss, S. Montgomery-Smith and E. Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. 362 (2010), 553–575.

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

[12] N. Kazamaki, *Continuous Exponential Martingales and BMO*, Lecture Notes in Math. 1579, Springer, Berlin, 1994.

[13] M. T. Lacey, K. Moen, C. Pérez and R. H. Torres, *Sharp weighted bounds for fractional integral operators*, J. Funct. Anal. 259 (2010), 1073–1097.

[14] A. Lerner, *Sharp weighted norm inequalities for Littlewood–Paley operators and singular integrals*, Adv. Math. 226 (2011), 3912–3926.

[15] A. Lerner, *On an estimate of Calderón–Zygmund operators by dyadic positive operators*, J. Anal. Math. 121 (2013), 141–161.

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

[17] A. Osękowski, *Sharp maximal inequality for stochastic integrals*, Proc. Amer. Math. Soc. 136 (2008), 2951–2958.

[18] A. Osękowski, *Sharp maximal inequality for martingales and stochastic integrals*, Electron. Comm. Probab. 14 (2009), 17–30.
[19] A. Osękowski, *Maximal inequalities for continuous martingales and their differential subordinates*, Proc. Amer. Math. Soc. 139 (2011), 721–734.
[20] A. Osękowski, *Sharp Martingale and Semimartingale Inequalities*, Monografie Matematyczne 72, Birkhäuser, 2012.
[21] A. Osękowski, *Weighted maximal inequality for differentially subordinate martingales*, Electron. Comm. Probab. 21 (2016), art. 23, 10 pp.
[22] A. Osękowski, *Weighted weak-type inequality for martingales*, Bull. Polish Acad. Sci. Math. 65 (2017), 165–175.
[23] A. Osękowski, *Weighted maximal inequalities for the Haar system*, Monatsh. Math. 186 (2018), 321–336.
[24] A. Osękowski, *Weighted square function inequalities*, Publ. Mat. 62 (2018), 321–336.
[25] S. Petermichl and A. Volberg, *Heating of the Ahlfors–Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. 112 (2002), 281–305.
[26] Y. Suh, *A sharp weak type* (p, p) *inequality* $(p > 2)$ *for martingale transforms and other subordinate martingales*, Trans. Amer. Math. Soc. 357 (2005), 1545–1564
[27] G. Wang, *Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities*, Ann. Probab. 23 (1995), 522–551.
[28] J. Wittwer, *A sharp bound for the martingale transform*, Math. Res. Lett. 7 (2000), 1–12.

Michał Brzozowski, Adam Osękowski
University of Warsaw
Banacha 2
02-097 Warszawa, Poland
E-mail: M.Brzozowski@mimuw.edu.pl
ados@mimuw.edu.pl

Received 13.11.2019;
revised version 20.2.2020