A RATIO INEQUALITY FOR NONNEGATIVE MARTINGALES AND THEIR DIFFERENTIAL SUBORDINATES

ADAM OSĘKOWSKI

ABSTRACT. We prove a sharp inequality for the ratio of the maximal functions of nonnegative martingales and their differential subordinates. An application in the theory of weighted inequalities is given.

1. INTRODUCTION

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, filtered with a nondecreasing right-continuous family $(\mathcal{F}_t)_{t \geq 0}$ of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all the events of probability 0. Let X, Y be two adapted martingales, taking values in a certain separable Hilbert space $\mathcal{H}$ (which can be taken to be ℓ^2) with norm $|\cdot|$ and scalar product denoted by the dot $\cdot$. We impose the usual regularity properties on the trajectories of X and Y: the paths are assumed to be right-continuous and have limits from the left. Then $[X, X]$ will denote the quadratic covariance process of X: that is, we set $[X, X] = \sum_{n=1}^{\infty}[X^n, X^n]$, where X^n is the n-th coordinate of X and $[X^n, X^n]$ is the usual square bracket of the real-valued martingale X^n (see e.g. Dellacherie and Meyer [4] for details). In our considerations, $X^* = \sup_{t \geq 0}|X_t|$ will denote the maximal function of X, we will also use the notation $X_t^* = \sup_{0 \leq s \leq t}|X_s|$ and $\|X\|_p = \sup_{t \geq 0}\|X_t\|_{L^p}$, $1 \leq p \leq \infty$.

Throughout the paper we will assume that the process Y is dominated by X in the sense of the so-called differential subordination: following Bañuelos and Wang [1] and Wang [17], we say that Y is *differentially subordinate* to X, if the process $([X, X]_t - [Y, Y]_t)_{t \geq 0}$ is nondecreasing and nonnegative as a function of t. The origins of this domination principle go back to the works of Burkholder [3] in the discrete-time case: a martingale $g = (g_n)_{n \geq 0}$ is differentially subordinate to $f = (f_n)_{n \geq 0}$, if for any $n \geq 0$ we have $|dg_n| \leq |df_n|$ almost surely. Here $df = (df_n)_{n \geq 0}$, $dg = (dg_n)_{n \geq 0}$ are the difference sequences of f and g, respectively, uniquely determined by the requirement

$$f_n = \sum_{k=0}^{n} df_k \quad \text{and} \quad g_n = \sum_{k=0}^{n} dg_k, \qquad n = 0, 1, 2, \ldots.$$

This is a special case of the continuous-time differential subordination: simply treat given two martingales f, g as continuous-time processes (via $X_t = f_{\lfloor t \rfloor}$ and $Y_t = g_{\lfloor t \rfloor}$, $t \geq 0$). Then the inequality $|df| \geq |dg|$ implies that $[X, X] - [Y, Y]$ enjoys the required properties.

Inequalities for martingales and their differential subordinates (in the discrete and continuous time) have been studied very intensively and have found numerous and deep applications in harmonic analysis. The literature on the subject is extremely vast and it is impossible to give even a brief review here. See e.g. [1, 2, 3, 7, 13, 16, 17] and the

references therein. We will just present a few results which serve as a motivation for our research. A celebrated result of Burkholder [3, 17] gives the following sharp L^p estimate.

Theorem 1.1. *Suppose that X, Y are Hilbert-space-valued martingales such that Y is differentially subordinate to X. Then*

$$(1.1) \qquad\qquad \|Y\|_p \leq (p^* - 1)\|X\|_p, \qquad 1 < p < \infty,$$

where $p^* = \max\{p, p/(p-1)\}$. The constant is the best possible, even if $\mathcal{H} = \mathbb{R}$.

For $p = 1$, the above inequality fails to hold, but we have the corresponding weak-type $(1, 1)$ estimate. In fact, there is a result for a wider range of parameters p, proved by Burkholder [3] for $1 \leq p \leq 2$ and Suh [16] for $p > 2$. See also Wang [17].

Theorem 1.2. *Suppose that X, Y are Hilbert-space-valued martingales such that Y is differentially subordinate to X. Then*

$$\mathbb{P}(Y^* \geq 1) \leq \frac{2}{\Gamma(p+1)}\|X\|_p^p, \qquad 1 \leq p \leq 2,$$

and

$$\mathbb{P}(Y^* \geq 1) \leq \frac{p^{p-1}}{2}\|X\|_p, \qquad 2 < p < \infty.$$

Both inequalities are sharp, even if $\mathcal{H} = \mathbb{R}$.

There are also other, logarithmic, exponential, mixed-norm and other results in this direction. Moreover, there is a powerful method which enables the identification of the best constants involved: roughly speaking, it reduces the problem to the construction of a certain special function, enjoying appropriate size requirements and concavity: see [13].

We will continue the above line of research and study a novel class of estimates between X and Y. Observe that the inequalities of Theorems 1.1 and 1.2 can be regarded as comparisons of sizes of X and Y measured separately in terms of strong or weak L^p norms and, as such, they do not say anything about the *joint* behavior of X and Y. From this point of view, a very natural functional to study is the ratio $|Y|/|X|$. For example, for a given time $t > 0$ and a level $\lambda > 0$, is there any *nontrivial* upper bound for the probability $\mathbb{P}(|Y_t| > \lambda|X_t|)$? No. It is easy to construct, for any t and λ, a pair (X, Y) of real-valued martingales satisfying the differential subordination such that this probability is as close to 1 as we wish. Indeed, for $\lambda < 1$ the constant pair $(1, 1)$ does the job, while for $\lambda \geq 1$ we take a two-dimensional Brownian motion (X, Y) started at $(1, 1)$ and stopped upon reaching the set $\{(x, y) : |y| = 2\lambda|x|\}$. Then the differential subordination is satisfied and, rescaling time if necessary, we can ensure that the probability $\mathbb{P}(|Y_t| = 2\lambda|X_t|)$ is arbitrarily close to 1. The above problem becomes more interesting if the maximal functions of X and Y are involved. Our first result is negative, even in the discrete time.

Theorem 1.3. *For any $\lambda > 0$ and any $\eta < 1$ there is a pair (f, g) of discrete-time, real valued martingales starting from 0 such that g is differentially subordinate to f and $\mathbb{P}(g^* > \lambda f^*) \geq \eta$.*

However, if we assume that the dominating martingale is nonnegative, there is a non-trivial bound. Here is the precise statement.

Theorem 1.4. *Suppose that X is a nonnegative martingale and Y is an ℓ^2-valued martingale which is differentially subordinate to X. Then for any $\lambda > 0$ we have*

$$(1.2) \qquad \mathbb{P}(Y^* > \lambda X^*) \leq \begin{cases} e^{1-\lambda/2} & \text{if } \lambda > 2, \\ 1 & \text{if } 0 < \lambda \leq 2. \end{cases}$$

For any λ, the constant on the right cannot be improved, even in the discrete-time case.

Our approach will exploit Burkholder's method: the inequality (1.2) will be deduced from the properties of a certain special function constructed in the next section. Section 3 contains the proofs of Theorems 1.3 and 1.4. The final part of the paper is devoted to some applications of the estimate (1.2) in the theory of weighted inequalities.

2. A special function and its properties

Throughout this section, we assume that $\lambda > 2$ is a fixed parameter and d is a fixed dimension. Distinguish the following four sets:

$$\begin{aligned} D_1 &= \big\{(x,y) \in [0,1] \times \mathbb{R}^d : x \geq |y|\big\}, \\ D_2 &= \big\{(x,y) \in [0,1] \times \mathbb{R}^d : x < |y| \leq x + \lambda - 2\big\}, \\ D_3 &= \big\{(x,y) \in [0,1] \times \mathbb{R}^d : x + \lambda - 2 < |y| < -x + \lambda\big\}, \\ D_4 &= \big\{(x,y) \in [0,1] \times \mathbb{R}^d : |y| \geq -x + \lambda\big\}. \end{aligned}$$

Consider $u : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ given by

$$u(x,y) = \begin{cases} \left(\dfrac{|y|^2 - x^2}{4} + x\right) e^{1-\lambda/2} & \text{if } (x,y) \in D_1, \\ xe^{1+(-x+|y|-\lambda)/2} & \text{if } (x,y) \in D_2, \\ \dfrac{2x}{x - |y| + \lambda} & \text{if } (x,y) \in D_3, \\ 1 & \text{if } (x,y) \in D_4. \end{cases}$$

It is easy to check that the function u is continuous on $([0,1] \times \mathbb{R}^d) \setminus \{(x,y) : |y| = \lambda\}$. Indeed, if $(x,y) \in \partial D_1 \cap \partial D_2$, then $x = |y|$ and

$$\left(\frac{|y|^2 - x^2}{4} + x\right) e^{1-\lambda/2} = xe^{1-\lambda/2} = xe^{1+(-x+|y|-\lambda)/2},$$

so the formulas for u on D_1 and D_2 match appropriately on the common boundary. One checks analogously that u is continuous at each point from $\partial D_2 \cap \partial D_3$ and $(\partial D_3 \cap \partial D_4) \setminus \{(x,y) : |y| = \lambda\}$. This gives the aforementioned continuity, since $\partial D_i \cap \partial D_j = \emptyset$ for other choices of i and j. A similar calculation shows that u is of class C^1 on $((0,1) \times \mathbb{R}^d) \setminus \{(x,y) : x + |y| = \lambda\}$: the partial derivatives of u match at the common boundaries of the sets D_1, D_2 and D_3. In our further considerations we will need to extend these partial derivatives

to the whole strip $[0,1] \times \mathbb{R}^d$: define $\varphi : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ and $\psi : [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ by

$$(\varphi(x,y), \psi(x,y)) = \begin{cases} \left(\left(1 - \dfrac{x}{2}\right)e^{1-\lambda/2}, \dfrac{1}{2}ye^{1-\lambda/2}\right) & \text{if } (x,y) \in D_1, \\[3mm] \left(\left(1 - \dfrac{x}{2}\right)e^{1+(-x+|y|-\lambda)/2}, \dfrac{xy'}{2}e^{1+(-x+|y|-\lambda)/2}\right) & \text{if } (x,y) \in D_2, \\[3mm] \left(\dfrac{2(\lambda - |y|)}{(x - |y| + \lambda)^2}, \dfrac{2xy'}{(x - |y| + \lambda)^2}\right) & \text{if } (x,y) \in D_3, \\[3mm] (0,0) & \text{if } (x,y) \in D_4. \end{cases}$$

Here and in what follows, we use the notation $y' = y/|y|$ if $y \in \mathbb{R}^d \setminus \{0\}$ and $0' = 0$. Further important properties are studied in a sequence of lemmas below.

Lemma 2.1. *We have*

$$(2.1) \qquad\qquad u(x,y) \le e^{1-\lambda/2} \qquad \text{if } |y| \le x$$

and

$$(2.2) \qquad\qquad 1 \ge u(x,y) \ge 1_{\{x+|y| \ge \lambda\}} \qquad \text{if } (x,y) \in [0,1] \times \mathbb{R}^d.$$

Proof. The first inequality is trivial: we have $(x,y) \in D_1$, so

$$u(x,y) = \left(\frac{|y|^2 - x^2}{4} + x\right)e^{1-\lambda/2} \le xe^{1-\lambda/2} \le e^{1-\lambda/2}.$$

To show the majorization (2.2), fix $x \in [0,1]$ and observe first that the number $u(x,0) = e^{1-\lambda/2}(-x^2/4 + x)$ lies between 0 and 1. Furthermore, it follows directly from the formula for u that the function $y \mapsto u(x,y)$ increases (precisely: does not decrease) as $|y|$ increases, and it is equal to 1 for $|y| \ge \lambda - x$. This yields the claim. $\qquad\square$

Let us now establish an important bound for the Hessian matrix of u.

Lemma 2.2. *If (x,y) lies in the interior of D_1, D_2, D_3 or D_4, then for any $(h,k) \in \mathbb{R} \times \mathbb{R}^d$,*

$$(2.3) \qquad\qquad D^2u(x,y)(h,k) \cdot (h,k) \le c(x,y)(|k|^2 - h^2),$$

where

$$c(x,y) = \begin{cases} \frac{1}{2}e^{1-\lambda/2} & \text{if } (x,y) \in \text{int } D_1, \\[1mm] \frac{1}{4}e^{1+(-x+|y|-\lambda)/2} & \text{if } (x,y) \in \text{int } D_2, \\[1mm] 2(x - |y| + \lambda)^{-2} & \text{if } (x,y) \in \text{int } D_3, \\[1mm] 0 & \text{if } (x,y) \in D_4. \end{cases}$$

Proof. For (x,y) lying in the interior of D_1 or D_4, both sides of (2.3) are equal. If (x,y) lies inside D_2, then $D^2u(x,y)(h,k) \cdot (h,k) = c(x,y)(|k|^2 - h^2 + I + II)$, where both terms

$$I = -\left(1 - \frac{x}{2}\right)(h - y' \cdot k)^2, \qquad II = -\left(1 - \frac{x}{|y|}\right)(|k|^2 - (y' \cdot k)^2)$$

are nonpositive (since $x \le 1$ and $|y| \ge x$). Similarly, if (x,y) is in the interior of D_3, then $D^2u(x,y)(h,k) \cdot (h,k) = c(x,y)(|k|^2 - h^2) + I + II$, with

$$I = -\frac{2(\lambda - x - |y|)(h - y' \cdot k)^2}{(x - |y| + \lambda)^3}, \qquad II = -\frac{2}{(x - |y| + \lambda)^2}\left(1 - \frac{x}{|y|}\right)(|k|^2 - (y' \cdot k)^2).$$

Both these expressions are nonpositive, so the claim follows. $\qquad\square$

We will also need the following auxiliary estimate.

Lemma 2.3. *For any* $h \in \mathbb{R}$, y, $k \in \mathbb{R}^d$ *satisfying* $|k| \leq h$ *we have*

$$(2.4) \qquad u\left(1, \frac{y+k}{1+h}\right) \leq u(1,y) + \varphi(1,y)h + \psi(1,y) \cdot k.$$

Proof. We consider separately three cases.

Case I. $|y| > \lambda - 1$. Then the estimate is obvious: the right-hand side is equal to 1, while the left-hand side is at most one (see the left inequality in (2.2)).

Case II: $1 < |y| \leq \lambda - 1$. Since

$$(2.5) \qquad \frac{|y+k|}{1+h} \leq \frac{|y|+|k|}{1+h} \leq \frac{|y|+h}{1+h} \leq \frac{|y|+|y|h}{1+h} = |y| \leq \lambda - 1,$$

the point $(1, (y+k)/(x+h))$ belongs to D_1 or to D_2, depending on whether $|y+k|/(1+h)$ is less than 1 or not. If this point lies in D_1, the inequality (2.4) can be rewritten as

$$\frac{1}{4}\left(\frac{|y+k|^2}{(1+h)^2} - 1\right) + 1 \leq \exp\left(\frac{|y|-1}{2}\right)\left(1 + \frac{h}{2} + \frac{y' \cdot k}{2}\right)$$

and the left-hand side is not bigger than 1, while the right-hand side is at least 1 (since $y' \cdot k \geq -|k| \geq -h$). If $(1, (y+k)/(x+h)) \in D_2$, then (2.4) reads

$$\exp\left(\frac{1}{2}\left(\frac{|y+k|}{1+h} - |y|\right)\right) \leq 1 + \frac{h}{2} + \frac{y' \cdot k}{2},$$

and again: the left-hand side does not exceed 1 (see (2.5) above), while the expression on the right is 1 or more.

Case III: $|y| \leq 1$. We observe that

$$\frac{|y+k|}{1+h} \leq \frac{|y|+|k|}{1+h} \leq \frac{1+h}{1+h} = 1,$$

so the point $(1, (y+k)/(1+h))$ lies in D_1 and the inequality (2.4) becomes, after some straightforward manipulations,

$$|y+k|^2\left(\frac{1}{(1+h)^2} - 1\right) \leq -|k|^2 + 2h.$$

Clearly, if we increase h, then the left-hand side decreases and the right-hand side increases; therefore, it suffices to show the estimate for the smallest value of h, i.e., $h = |k|$. Dividing by $|k|$ and rearranging terms, we see that it is enough to show that

$$-|y+k|^2(|k|+2) \leq (1+|k|)^2(2-|k|).$$

If $|k| \leq 2$, then this bound holds true: the left-hand side is nonpositive, while the right-hand side is nonnegative. If $|k| \geq 2$, then we note that $|y+k| \geq |k| - |y| \geq |k| - 1$ and hence the desired estimate follows from $-(|k|-1)^2(|k|+2) \leq (1+|k|)^2(2-|k|)$, which can be easily verified by the direct opening of the brackets. $\qquad \square$

Combining the two lemmas above, we get the following bound.

Corollary 2.4. *For any* $(x,y) \in [0,1] \times \mathbb{R}^d$, *any* $h \in \mathbb{R}$ *and any* $k \in \mathbb{R}^d$ *such that* $x+h \geq 0$ *and* $|k| \leq |h|$, *we have*

$$(2.6) \qquad u\left((x+h) \wedge 1, \frac{y+k}{(x+h) \vee 1}\right) \leq u(x,y) + \varphi(x,y)h + \psi(x,y) \cdot k.$$

Proof. We may assume that $h \neq 0$, since otherwise the claim is obvious. Consider the function G given by the formula

$$G(t) = u\left((x+th) \wedge 1, \frac{y+tk}{(x+th) \vee 1}\right),$$

defined for those $t \in \mathbb{R}$, for which $x + th \geq 0$. Then the inequality (2.6) is equivalent to $G(1) \leq G(0) + G'(0)$ (for the points (x, y) at which u is not differentiable, $G'(0)$ is a one-sided derivative). On the set of those t, for which $x + th \in [0, 1]$ (i.e., for t lying between $-x/h$ and $(1-x)/h$), the function G is concave. To see this, note that for these particular t we have the equality $G(t) = u(x + th, y + tk)$. Now, if $(x + th, y + tk)$ lies in the interior of some D_i, then G is twice differentiable and by (2.3),

$$G''(t) = D^2 u(x + th, y + tk)(h, k) \cdot (h, k) \leq c(x + th, y + tk)(|k|^2 - h^2) \leq 0.$$

On the other hand, if $(x+th, y+tk)$ lies at the common boundary of two sets D_i and D_{i+1}, then $G'(t-) \geq G'(t+)$: indeed, we have equality for $i = 1, 2$ (since u is of class C^1 in the interior of $D_1 \cup D_2 \cup D_3$), while for $i = 3$ this follows from the fact that $u(x, y) = 1$ on D_4 and $u(x, y) \leq 1$ on D_3. Putting these observations together, we get the aforementioned concavity of G and hence if $x + h \leq 1$, the claim follows. If $x + h > 1$, then in particular $h > 0$ and the concavity of G on $[-x/h, (1-x)/h]$ (which we just established) yields

$$(2.7) \qquad G(0) + G'(0) \geq G\left(\frac{1-x}{h}\right) + G'\left(\frac{1-x}{h}\right)\left(1 - \frac{1-x}{h}\right).$$

However, the expression on the right is not smaller than the left-hand side of (2.6): this is guaranteed by the inequality (2.4). Indeed, set

$$\tilde{h} = x + h - 1 = h(1 - (1-x)/h), \quad \tilde{k} = k(1 - (1-x)/h), \quad \tilde{y} = y + k - \tilde{k}.$$

Then $|\tilde{k}| \leq \tilde{h}$ and hence

$$G(1) = u\left(1, \frac{y+k}{x+h}\right) = u\left(1, \frac{\tilde{y} + \tilde{k}}{1 + \tilde{h}}\right) \leq u(1, \tilde{y}) + \varphi(1, \tilde{y})h + \psi(1, \tilde{y}) \cdot \tilde{k}$$

$$= G\left(\frac{1-x}{h}\right) + G'\left(\frac{1-x}{h}\right)\left(1 - \frac{1-x}{h}\right).$$

By (2.7), this is bounded from above by $G(0) + G'(0)$ and the claim is proved.   □

The final property of u is the following.

Lemma 2.5. *For any $x \in (0, 1]$ and $y \in \mathbb{R}^d$, the function $F(t) = u(tx, ty)$ is nondecreasing on $[0, 1/x]$.*

Proof. By the direct differentiation, it is enough to show that $x\varphi(x, y) + \psi(x, y) \cdot y \geq 0$. However, it is clear from the formulas for φ and ψ that $\varphi(x, y) \geq 0$ and $\psi(x, y) \cdot y \geq 0$.   □

3. Proof of (1.2)

We start with recalling a simple fact from stochastic analysis [4]. Namely, for any martingale X there exists a unique continuous local martingale part X^c of X satisfying

$$[X, X]_t = |X_0|^2 + [X^c, X^c]_t + \sum_{0 < s \leq t} |\Delta X_s|^2$$

for all $t \geq 0$. Furthermore, the bracket $[X^c, X^c]$ coincides with $[X, X]^c$, the pathwise continuous part of $[X, X]$. We will also require the following statement [17, Lemma 1].

Lemma 3.1. *If X and Y are martingales, then Y is differentially subordinate to X if and only if $|Y_0| \leq |X_0|$, Y^c is differentially subordinate to X^c and we have $|\Delta Y_t| \leq |\Delta X_t|$ for all $t > 0$.*

In the proof of (1.2), we may assume that $\lambda > 2$, since otherwise the claim is obvious. Suppose that X, Y are two martingales such that X is nonnegative, Y takes values in a Hilbert space $\mathcal{H} = \ell^2$ and is differentially subordinate to X. We may restrict ourselves to X which is bounded away from 0, by adding a deterministic number $\varepsilon > 0$ to this process and letting $\varepsilon \to 0$ at the end (such a modification does not affect the differential subordination). For a fixed small positive number η, take the stopping time

$$\tau = \inf\{t \geq 0 : X_t + |Y_t| \geq \lambda X_t^*\},$$

with the usual convention $\inf \emptyset = +\infty$. Consider the domain $\mathcal{D} = \{(x, y, z) \in [0, \infty) \times \mathbb{R}^d \times (0, \infty) : x \leq z\}$ and let $U : \mathcal{D} \to \mathbb{R}$ be given by

$$U(x, y, z) = u\left(\frac{x}{z}, \frac{y}{z}\right).$$

Before we proceed to the formal proof (which involves some additional technical arguments), let us briefly *sketch* the idea. First we translate the properties of u into the language of U. By (2.3), if $(x/z, y/z)$ lies in the interior of some D_i, then

$$(3.1) \qquad D^2_{x,y} U(x, y, z)(h, k) \cdot (h, k) \leq c\left(\frac{x}{z}, \frac{y}{z}\right)(k^2 - |h|^2).$$

Furthermore, by (2.6), if $x \leq z$ and $|k| \leq |h|$, then we have

$$(3.2) \qquad U(x + h, y + k, (x + h) \vee z) \leq U(x, y, z) + U_x(x, y, z)h + U_y(x, y, z) \cdot k.$$

Finally, Lemma 2.5 implies that

$$(3.3) \qquad \text{for each } (x, y), \text{ the function } z \mapsto U(x, y, z) \text{ is nonincreasing on } [x, \infty).$$

The plan is to apply Itô's formula to the composition of U with the stopped process $(X^\tau, Y^\tau, (X^\tau)^*)$. Then the three inequalities above imply that the resulting semimartingale is actually a supermartingale: the first inequality handles the second-order terms, (3.2) deals with the jump part and (3.3) enables the control of the stochastic integral with respect to X^*. Therefore, we obtain

$$\mathbb{E}U(X_\tau, Y_\tau, X_\tau^*) \leq \mathbb{E}U(X_0, Y_0, X_0^*) = \mathbb{E}u(1, Y_0/X_0) \leq e^{1-\lambda/2},$$

where the last bound is due to (2.1) and the differential subordination. Using the right inequality in (2.2), we obtain $\mathbb{P}(|Y_\tau| \geq \lambda X_\tau^*) \leq e^{1-\lambda/2}$ and it remains to note that $\{Y^* > \lambda X^*\} \subseteq \{|Y_\tau| \geq \lambda X_\tau^*\}$.

There is a small gap in the above reasoning, but the main idea goes along the path described above. The problem is that the application of Itô's formula is not permitted, since the function U lacks the necessary regularity. To overcome this difficulty, we will make use of some additional mollifying arguments. Let $g : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \to [0, \infty)$ be a C^∞ function supported on the unit ball of $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ and satisfying $\int_{\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}} g = 1$. For a fixed parameter $\delta > 0$, we define $U^\delta : \mathcal{D} \to \mathbb{R}$ by the convolution

$$U^\delta(x, y, z) = \int_{[-1,1] \times \mathbb{R}^d \times [-1,1]} U(x + \delta + \delta r, y + \delta s, z + 4\delta + \delta t) g(r, s, t) \, dr ds dt.$$

The function u is of class C^1 in the interior of the set $D_1 \cup D_2 \cup D_3$, so U is of class C^1 inside the set $\mathcal{D}' = \{(x, y, z) : x + |y| < (\lambda - 1)z, \ x < z\}$. Therefore,

$$D_{x,y}^2 U^\delta(x, y, z) = \int_{[-1,1] \times \mathbb{R}^d \times [-1,1]} D_{x,y}^2 U(x + \delta + \delta r, y + \delta s, z + 4\delta + \delta t) g(r, s, t) \, \mathrm{d}r\mathrm{d}s\mathrm{d}t$$

on $\mathcal{D}'$, by the integration by parts. Consequently, (3.1) yields

$$(3.4) \qquad D_{x,y}^2 U^\delta(x, y, z)(h, k) \cdot (h, k) \leq c^\delta(x, y, z)(|k|^2 - h^2)$$

for $(x, y, z) \in \mathcal{D}'$, where

$$c^\delta(x, y, z) = \int_{[-1,1] \times \mathbb{R}^d \times [-1,1]} c\left(\frac{x + \delta + \delta r}{z + 4\delta + \delta t}, \frac{x + \delta + \delta r}{z + 4\delta + \delta t}\right) g(r, s, t)\mathrm{d}r\mathrm{d}s\mathrm{d}t \geq 0.$$

Moreover, the inequality (3.2) extends to U^δ without any modification (for any $(x, y, z) \in \mathcal{D}'$); the same happens to the property (3.3). Now, the function U^δ is of class C^∞, so composing it with the process $S^\tau = (X^\tau, Y^\tau, (X^\tau)^*)$ we get, by Itô's formula,

$$U^\delta(S_{\tau \wedge T}) = U^\delta(S_0) + I_1 + I_2 + I_3/2 + I_4,$$

where

$$I_1 = \int_0^{\tau \wedge T} U_x^\delta(S_{s-})\mathrm{d}X_s + \int_0^{\tau \wedge T} U_y^\delta(S_{s-})\mathrm{d}Y_s,$$

$$I_2 = \int_0^{\tau \wedge T} U_z^\delta(S_{s-})\mathrm{d}(X_s^*)^c$$

$$I_3 = \int_0^{\tau \wedge T} D_{x,y}^2 U^\delta(S_{s-}) d[X, Y]_s^c,$$

$$I_4 = \sum_{0 \leq s \leq \tau \wedge T} \left[U^\delta(S_s) - U^\delta(S_{s-}) - U_x^\delta(S_{s-})\Delta X_s - U_y^\delta(S_{s-}) \cdot \Delta Y_s \right].$$

Here in I_3 we have used a shortened form for the sum of all second-order terms. Observe that the integrals in I_2 are with respect to the continuous part of the processes X^*; this implies the lack of the term $U_z^\delta(S_{s-})\Delta X_s^*$ in I_4. By the properties of stochastic integrals, the term I_1 has zero expectation. By (3.3) we have $U_z^\delta \leq 0$, so $I_2 \leq 0$. To handle I_3, note that $(S_{t-})_{t\geq 0}$ takes values in the set $\mathcal{D}'$ (which follows directly from the form of the stopping time τ). Hence, if we approximate I_3 by Riemann sums and apply (3.4), we get

$$I_3 \leq \int_0^{\tau \wedge T} c^\delta(S_{s-}) d([Y, Y]_s^c - [X, X]_s^c) \leq 0,$$

where in the last passage we have exploited the differential subordination of Y^c to X^c (see Lemma 3.1) and the inequality $c^\delta \geq 0$. The term I_4 is also nonpositive, since by (3.2) and the inequality $|\Delta Y_s| \leq |\Delta X_s|$ (see Lemma 3.1 again), so is its each summand. Putting all the above facts together, we obtain the estimate $\mathbb{E}U^\delta(X_{\tau \wedge T}, Y_{\tau \wedge T}, X_{\tau \wedge T}^*) \leq \mathbb{E}U^\delta(X_0, Y_0, X_0^*)$. Now the function U^δ is continuous and bounded by 1 (since U has this property). Consequently, letting $T \to \infty$ we get $\mathbb{E}U^\delta(X_\tau, Y_\tau, X_\tau^*) \leq \mathbb{E}U^\delta(X_0, Y_0, X_0^*)$, by Lebesgue's dominated convergence theorem. Now, the function U is continuous on $\{(x, y, z) \in (0, \infty) \times \mathbb{R}^d \times (0, \infty) : x \leq z\}$, so letting $\delta \to 0$, using the boundedness of U and the fact that X is bounded away from 0, we arrive at $\mathbb{E}U(X_\tau, Y_\tau, X_\tau^*) \leq \mathbb{E}U(X_0, Y_0, X_0^*)$. As we have seen above, this implies the desired tail estimate $\mathbb{P}(Y^* > \lambda X^*) \leq e^{1-\lambda/2}$.

Now we turn our attention to the construction of appropriate examples.

Proof of Theorem 1.3. Pick a huge positive number M and consider the sequence ε_1, ε_2, ... of i.i.d. centered random variables with the distribution given by $\mathbb{P}(\varepsilon_i = -1) = 1 - \mathbb{P}(\varepsilon_i = M) = \frac{M}{M+1}$. For a given integer $N > \lambda$, and for any $n = 0, 1, 2, \ldots, N$, set $f_n = \varepsilon_1 - \varepsilon_2 + \varepsilon_3 - \ldots + (-1)^{n+1}\varepsilon_n$ and $g_n = -\varepsilon_1 - \varepsilon_2 - \varepsilon_3 - \ldots - \varepsilon_n$. Furthermore, for $n > N$, put $f_n = f_N$ and $g_n = g_N$. Note that f and g are mean-zero martingales (which follows from the fact that ε_i are centered). Furthermore, we have $dg_n = (-1)^n df_n$, so the differential subordination is satisfied. Consider the event $A = \{\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_N = -1\}$, on which we have $f_0 = 0$, $f_1 = -1$, $f_2 = 0$, $f_3 = -1$, ... and $g_N = N$, so $f^* = 1$ and $g^* \geq N$. Consequently, $\mathbb{P}(g^* > \lambda f^*) \geq \mathbb{P}(A) = \left(\frac{M}{M+1}\right)^N$, and the latter fraction can be made arbitrarily close to 1, by taking M sufficiently large. $\qquad\square$

Sharpness of (1.2). If $\lambda < 2$, we use the martingales $\tilde{f} = f + 1$ and $\tilde{g} = g + 1$, where f, g come from the above example with $N = 1$. Then $\tilde{f}$ is nonnegative, $\tilde{g}$ is differentially subordinate to $\tilde{f}$ and we have $\mathbb{P}(\tilde{g}^* > \lambda \tilde{f}^*) \geq \mathbb{P}((\tilde{f}^*, \tilde{g}^*) = (1, 2)) = \frac{M}{M+1}$. It remains to note that the latter fraction tends to 1 as we let $M \to \infty$.

Now we will handle the case $\lambda \geq 2$. Pick an arbitrary $\lambda' > \lambda$, a huge positive number M, a huge positive integer N and set $\delta = (\lambda' - 2)/(2N) > 0$. Consider the independent random variables ε_1, ε_2, ..., ε_{2N}, whose two-point distributions are given as follows. If n is odd, then $\mathbb{P}(\varepsilon_n = -\delta) = 1 - \mathbb{P}(\varepsilon_n = M) = \frac{M}{M+\delta}$; if n is even, then $\mathbb{P}(\varepsilon_n = \delta) = 1 - \mathbb{P}(\varepsilon_n = 1 - \delta) = 1 - \delta$. Finally, consider a variable ε_{2N+1}, independent of ε_1, ε_2, ..., ε_{2N}, satisfying $\mathbb{P}(\varepsilon_{2N+1} = -1) = 1 - \mathbb{P}(\varepsilon_{2N+1} = M) = \frac{M}{M+1}$. Introduce the stopping time $\tau = \inf\{n : |\varepsilon_n| \neq \delta\}$. Since the variables ε_1, ε_2, ..., ε_{2N+1} are centered, the sequences $f_n = 1 + \varepsilon_{\tau \wedge 1} + \varepsilon_{\tau \wedge 2} + \ldots + \varepsilon_{\tau \wedge N}$, $g_n = 1 - \varepsilon_{\tau \wedge 1} + \varepsilon_{\tau \wedge 2} - \ldots + (-1)^n \varepsilon_{\tau \wedge n}$ ($n \geq 0$) are martingales; furthermore, we see that the differential subordination is satisfied. Let us gain some intuition about the behavior of the pair (f, g). It starts from the point $(1, 1)$; then it moves to $(1 + M, 1 - M)$ or to $(1 - \delta, 1 + \delta)$. If the first scenario occurs, the pair stops ultimately; otherwise the movement is continued and in the second step the pair jumps to $(0, 2\delta)$ or to $(1, 1 + 2\delta)$. If the first possibility takes place, the pair terminates; otherwise it continues its evolution, going to $(1 + M, 1 + 2\delta - M)$ or to $(1 - \delta, 1 + 3\delta)$; in the first case, the process stops, while in the second it evolves further. The pattern is then repeated. Note that f is nonnegative. We see that after the $2N$ steps we have two possibilities: either f visited the set $\{0, 1 + M\}$ (and stopped there), or (f, g) has come to the point $(1, 1 + 2N\delta) = (1, \lambda' - 1)$. If the latter happens, the pair goes to $(1 + M, \lambda' - 1 - M)$ or to $(0, \lambda')$. We easily compute that

$$\mathbb{P}((f_{2N+1}, g_{2N+1}) = (0, \lambda')) = \left(\frac{M}{M+\delta}\right)^N (1 - \delta)^N \cdot \frac{M}{M+1}.$$

It is clear from the above that on the set $\{(f_{2N+1}, g_{2N+1}) = (0, \lambda')\}$ we have $f^* = 1$. So,

$$\mathbb{P}(g^* > \lambda f^*) \geq \mathbb{P}(g^* \geq \lambda' f^*) \geq \left(\frac{M}{M+\delta}\right)^N (1 - \delta)^N \cdot \frac{M}{M+1}.$$

But the expression on the right can be made as close to $e^{1-\lambda/2}$ as we wish. To accomplish this, take λ' close to λ, recall that $\delta = (\lambda' - 2)/(2N)$, take sufficiently large N and then sufficiently large M. $\qquad\square$

4. An application

We will show how the estimate can be used to study a certain weighted estimate. For the basic facts concerning the weighted theory, the reader is referred to any textbook on harmonic analysis (e.g. [8]); for the probabilistic aspects of the subject, consult [9]. Throughout this section, we will work with the probability space $([0,1), \mathcal{B}([0,1)), |\cdot|)$. Let $h = (h_n)_{n \geq 0}$ stand for the usual, unnormalized Haar system on $[0,1)$:

$$h_0 = [0,1), \quad h_1 = [0,1/2) - [1/2,1),$$
$$h_2 = [0,1/4) - [1/4,1/2), \quad h_3 = [1/2,3/4) - [3/4,1),$$
$$h_4 = [0,1/8) - [1/8,1/4), \quad h_5 = [1/4,3/8) - [3/8,1/2),$$
$$h_6 = [1/2,5/8) - [5/8,3/4), \quad h_7 = [3/4,7/8) - [7/8,1),$$

and so on, where we have identified a set with its indicator function. The Haar system is a martingale difference sequence with respect to its natural filtration $(\mathcal{F}_n)_{n \geq 0}$ and hence, for any given integrable function $f = \sum_{k=0}^{\infty} a_k h_k$, the associated sequence $f_n = \sum_{k=0}^{n} a_k h_k$, $n = 0, 1, 2, \ldots$, is a martingale. For a given collection $\varepsilon = (\varepsilon_n)_{n \geq 0}$ of vectors from a Hilbert space $\mathcal{H}$, we define the corresponding Haar multiplier $T = T_\varepsilon$ by $T\left(\sum_{n=0}^{\infty} a_n h_n\right) = \sum_{n=0}^{\infty} \varepsilon_n a_n h_n$, which gives rise to another martingale, given by $g_n = \sum_{k=0}^{n} \varepsilon_k a_k h_k$, $n = 0, 1, 2, \ldots$. Observe that if ε is bounded in absolute value by 1, i.e., the terms ε_0, ε_1, ε_2, $\ldots$ take values in the unit ball of $\mathcal{H}$, then g is differentially subordinate to f.

Let w be a weight, i.e., a nonnegative, integrable function on $[0,1)$. For any $A \subseteq [0,1)$, we will use the notation $w(A) = \int_A w \mathrm{d}x$. In 1971, Fefferman and Stein [6] established the following weighted version of the maximal weak-type (1,1) estimate:

$$w\left(f^* \geq 1\right) \leq c \int_0^1 |f| w^* \mathrm{d}x,$$

for some absolute constant c (here w^* is the maximal function of the martingale $(w_n)_{n \geq 0} = (\mathbb{E}(w|\mathcal{F}_n))_{n \geq 0}$ generated by w). This result gives rise to a very interesting question: does the estimate hold true if we replace the maximal function of f on the left by $|T_\varepsilon f|$, where ε is a sequence bounded in absolute value by 1? This question, known as a dyadic Muckenhoupt-Wheeden conjecture, was open for almost forty years and was finally answered in the negative by Reguera [15]. There is a dual problem concerning the estimate

$$w\left(|T_\varepsilon f| \geq w^*\right) \leq c \int_0^1 |f| \mathrm{d}x,$$

see [10, 11, 12]. The author showed in [14] that this inequality does not hold either with any finite constant c. Actually, the counterexample is given by $f = w/\lambda$, where w is an appropriately constructed weight and $\lambda > 0$ is a constant. In other words, the inequality

$$\lambda w\left(|T_\varepsilon w| \geq \lambda w^*\right) \leq c \int_0^1 w \mathrm{d}x$$

is not valid for general positive and integrable functions. Consider the related inequality

$$(4.1) \qquad\qquad \lambda v\left(|T_\varepsilon w| \geq \lambda w^*\right) \leq c \int_0^1 v \mathrm{d}x,$$

in which v is another weight on $[0,1)$. It follows from the tail estimate (1.2) that if $v \equiv 1$, then the above bound does hold, with the sharp constant $c = \sup_{\lambda \geq 2} \lambda e^{1-\lambda/2} = 2$.

We will strengthen this result to arbitrary A_p weights. Recall that a weight v satisfies Muckenhoupt's dyadic condition A_p $(1 < p < \infty)$, if

$$[v]_{A_p} := \sup \left(\frac{1}{|I|} \int_I v \right) \left(\frac{1}{|I|} \int_I v^{1/(1-p)} \right)^{p-1} < \infty,$$

where the supremum is taken over the class of all dyadic intervals I. There are versions of this condition for $p = 1$ and $p = \infty$: for $p = 1$, the above definition becomes

$$[v]_{A_1} := \sup \left(\frac{1}{|I|} \int_I v \right) \left(\operatorname*{esssup}_I v^{-1} \right) < \infty,$$

and for $p = \infty$, let $[v]_{A_\infty} := \sup \frac{1}{v(I)} \int_I (v\chi_I)^* \mathrm{d}x < \infty$. See [9] for more on the subject.

Theorem 4.1. *Let $1 \le p \le \infty$. If v is an A_p weight, w is an arbitrary weight and $\lambda > 0$, then (4.1) holds with $c = 4\kappa_p [v]_{A_p}$, for some constant $\kappa_p \ge 1$ depending only on p.*

Proof. If $\lambda < 4$, then the estimate is trivial, since $v\left(|T_\varepsilon w| \ge \lambda w^*\right) \le \int_0^1 v$, so we assume that $\lambda \ge 4$. We will use the following property of A_p weights established by Fefferman and Pipher [5] (modify the proof of Lemma 3.6 according to the remark on page 359 of that paper): there is a constant $\kappa_p \ge 1$ depending only on p such that for any dyadic interval $I \subseteq [0,1)$ and any subset A of I,

$$\frac{v(A)}{v(I)} \le \kappa_p [v]_{A_p} \left(\log \frac{|I|}{|A|} \right)^{-1}.$$

Letting $I = [0,1)$ and $A = \{|T_\varepsilon w| > \lambda w^*\}$, we get, by (1.2),

$$\frac{v(\{|T_\varepsilon w| > \lambda w^*\})}{\int_0^1 v \mathrm{d}x} \le \kappa_p [v]_{A_p} \left(\frac{\lambda}{2} - 1 \right)^{-1} \le 4\kappa_p [v]_{A_p} \lambda^{-1}.$$

This completes the proof. $\qquad\square$

References

[1] R. Bañuelos and G. Wang, *Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms*, Duke Math. J. **80** (1995), no. 3, 575–600.

[2] A. Borichev, P. Janakiraman, and A. Volberg, *On Burkholder function for orthogonal martingales and zeros of Legendre polynomials*, Amer. J. Math. **135** (2013), 207–236.

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

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

[5] R. Fefferman and J. Pipher, *Multiparameter operators and sharp weighted inequalities*, Amer. J. Math. 119 (1997), no. 2, 337–369.

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

[7] S. Geiss, S. Montgomery-Smith and E. Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. 362 No. 2 (2010), 553–575.

[8] L. Grafakos, *Classical and Modern Fourier Analysis*, Pearson Education, Inc. New Jersey, 2004.

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

[10] 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.

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

[12] 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), 149–156.

[13] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72**, Birkhäuser, 2012.

[14] A. Osękowski, *A weighted weak-type bound for Haar multipliers*, Royal Soc. Edinburgh Proc. A **148A** (2018), 643–658.

[15] M. C. Reguera, *On Muckenhoupt-Wheeden conjecture*, Adv. Math. **227** (2011), no. 4, 1436–1450.

[16] J. 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

[17] G. Wang, *Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities*, Ann. Probab. **23** (1995), 522–551.

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

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