PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 139, Number 2, February 2011, Pages 721–734
S 0002-9939(2010)10539-7
Article electronically published on August 13, 2010

MAXIMAL INEQUALITIES FOR CONTINUOUS MARTINGALES AND THEIR DIFFERENTIAL SUBORDINATES

ADAM OSĘKOWSKI

(Communicated by Richard C. Bradley)

Abstract. Let $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ be continuous-path martingales such that Y is differentially subordinate to X. The paper contains the proofs of the sharp inequalities

$$\sup_{t\geq 0} ||Y_t||_p \leq \sqrt{\frac{2}{p}} || \sup_{t\geq 0} |X_t| \, ||_p, \quad 1 \leq p < 2$$

and

$$\sup_{t\geq 0} ||Y_t||_p \leq (p-1) || \sup_{t\geq 0} |X_t| \, ||_p, \quad 2 \leq p < \infty.$$

1. Introduction

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with a filtration $(\mathcal{F}_t)$, a nondecreasing right-continuous family of sub-σ-fields of $\mathcal{F}$. Assume that $\mathcal{F}_0$ contains all the events of probability 0. Let $X = (X_t)_{t\geq 0}$ be an adapted real-valued right-continuous semimartingale with left limits. Let Y be the Itô integral of H with respect to X,

$$Y_t = H_0 X_0 + \int_{(0,t]} H_s dX_s, \quad t \geq 0,$$

where H is a predictable process with values in $[-1, 1]$. Let $||Y||_p = \sup_{t\geq 0} ||Y_t||_p$ and $X^* = \sup_{t\geq 0} |X_t|$.

In the present paper we will be interested in comparing the moments of Y and X^*. In [4], Burkholder described a method of proving maximal inequalities for martingales and used it to obtain the following sharp estimate.

Theorem 1.1. *If X and Y are as above, then we have*

$$||Y||_1 \leq \gamma ||X^*||_1, \tag{1.1}$$

where $\gamma = 2.536\ldots$ is the unique solution of the equation

$$\gamma - 3 = -\exp\big(\frac{1-\gamma}{2}\big).$$

The constant is the best possible.

Received by the editors June 29, 2009 and, in revised form, April 14, 2010.

2010 *Mathematics Subject Classification.* Primary 60G44; Secondary 60H05.

Key words and phrases. Martingale, stochastic integral, maximal inequality, differential subordination.

The author was partially supported by the Foundation for Polish Science and MNiSW Grant N N201 397437.

Then it was shown by the author in [6] that if X is nonnegative, then the constant decreases to $2 + (3e)^{-1} = 2.1226\ldots$; furthermore, the estimate remains valid if we assume that X runs over the class of nonnegative supermartingales.

In the paper we will study the case when the integrator X is continuous. Our contribution is to determine the best constants in the inequalities between the p-th moments of Y and X^*, $1 \leq p < \infty$. Our approach works for a wider class of processes: we allow the martingales to take values in a separable Hilbert space $\mathcal{H}$ (which can and will be taken to be ℓ^2) and, furthermore, we show the maximal estimates under the assumption of *differential subordination*. Suppose that $(X_t) = ((X_t^1, X_t^2, \ldots))$, $Y_t = ((Y_t^1, Y_t^2, \ldots))$ are two ℓ^2-valued martingales and let $[X, X] = \sum_{i=1}^{\infty}[X^i, X^i]$, $[Y, Y] = \sum_{i=1}^{\infty}[Y^i, Y^i]$, where $[X^i, X^i]$ and $[Y^i, Y^i]$ denote the quadratic variance processes of the real-valued martingales X^i and Y^i, respectively (see e.g. [5] for details). Following [8], we say that Y is differentially subordinate to X if the process $[X, X] - [Y, Y]$ is nondecreasing and nonnegative. For example, as in the beginning of this section, if Y is the Itô integral of H with respect to X, where H is predictable and taking values in $[-1, 1]$, then Y is differentially subordinate to X. This is a consequence of

$$[X, X]_t - [Y, Y]_t = \int_0^t (1 - |H_s|^2) d[X, X]_s.$$

Our main result can be stated as follows.

Theorem 1.2. *Suppose X, Y are $\mathcal{H}$-valued continuous martingales such that Y is differentially subordinate to X. Then the inequalities*

$$||Y||_p \leq \sqrt{\frac{2}{p}}||X^*||_p, \quad 1 \leq p < 2, \tag{1.2}$$

and

$$||Y||_p \leq (p-1)||X^*||_p, \quad 2 \leq p < \infty \tag{1.3}$$

hold true and the constants are the best possible. They are the best possible even when X is assumed to take values in $\mathbb{R}$ and Y is the Itô integral, with respect to X, of some predictable H taking values in $\{-1, 1\}$.

Comparing (1.1) and (1.2) for $p = 1$, we see that for continuous-path integrators, the constant γ decreases to $\sqrt{2}$.

It is worth mentioning here the following inequality, which was originally proved in the discrete-time setting by Burkholder [1] and extended to the continuous-time setting by Wang [8].

Theorem 1.3. *Suppose X, Y are $\mathcal{H}$-valued martingales such that Y is differentially subordinate to X. Then*

$$||Y||_p \leq (p^* - 1)||X||_p, \quad 1 < p < \infty, \tag{1.4}$$

where $p^* = \max\{p, p/(p-1)\}$. The inequality is sharp, even if X, Y are assumed to be real-valued.

We see that if $p \geq 2$, then $p^* - 1 = p - 1$, so (1.4) implies (1.3); it is quite surprising that the latter estimate is still sharp. On the other hand, we see that the constants in (1.2) and (1.4) have completely different behavior for $1 \leq p < 2$.

A few words about the organization of the paper. In the next section we modify Burkholder's technique so that it provides sharp constants in the continuous-path

setting. In particular, the method transforms the problem of proving the estimate (1.2) to the existence of a certain special function. In Section 3 we present such a function, and in Section 5 we describe some steps which led to its discovery. Before that, in Section 4, we show that the constants appearing in (1.2) and (1.3) are optimal.

2. Burkholder's method

The technique developed by Burkholder in [4] enables us to obtain sharp maximal estimates for general martingales. In particular, these inequalities are valid when we restrict ourselves to the class of continuous processes; however, the constants may, but need not, remain optimal. For example, as already mentioned, the constant γ from (1.1) decreases to $\sqrt{2}$ in the continuous-path case (as will be proved below). Therefore, to keep track of the best constants, we need to refine the method so that it exploits the continuity of the paths. The purpose of this section is to present this modification.

Fix a Borel function $V : \mathcal{H} \times \mathcal{H} \times [0, \infty) \to \mathbb{R}$, which is bounded on bounded sets. Suppose we are interested in the inequality

$$\mathbb{E}V(X_t, Y_t, X_t^*) \le 0 \tag{2.1}$$

being valid for all $t \ge 0$ and in the class of all pairs (X, Y) of continuous and bounded $\mathcal{H}$-valued martingales such that Y is differentially subordinate to X. The key tool to study this problem is the class $\mathcal{U} = \mathcal{U}(V)$, which consists of those functions $U : \mathcal{H} \times \mathcal{H} \times (0, \infty) \to \mathbb{R}$ of class C^2, which satisfy the following conditions:

(i) For any $(x, y, z) \in \mathcal{H} \times \mathcal{H} \times (0, \infty)$ such that $|y| \le |x| \le z$ we have

$$U(x, y, z) \le 0.$$

(ii) For any $(x, y, z) \in \mathcal{H} \times \mathcal{H} \times (0, \infty)$ such that $|x| \le z$ we have

$$U(x, y, z) \ge V(x, y, z).$$

(iii) For any $x \ne 0$,

$$U_z(x, y, |x|) \le 0.$$

(iv) There is a function $M : \mathcal{H} \times \mathcal{H} \times (0, \infty) \to (0, \infty)$, bounded on any set of the form $\{(x, y, z) : |x|, |y| \le L, 1/L \le z \le L\}$ for some $L > 0$, such that the following holds. For any $(x, y, z) \in \mathcal{H} \times \mathcal{H} \times (0, \infty)$ and $h, k \in \mathcal{H}$, we have

$$(U_{xx}(x, y, z)h, h) + 2(U_{xy}(x, y, z)h, k) + (U_{yy}(x, y, z)k, k) \le M(x, y, z)(|k|^2 - |h|^2).$$

Now we turn to the main fact in this section, relating the class $\mathcal{U}(V)$ to the inequality (2.1). Recall that X_0^1 denotes the first coordinate of X_0.

Theorem 2.1. *Suppose $U \in \mathcal{U}(V)$ and X, $Y \in L^\infty$ are $\mathcal{H}$-valued and continuous-path martingales such that Y is differentially subordinate to X and $\mathbb{P}(|X_0^1| \ge \delta) = 1$ for some $\delta > 0$. Then for any $t \ge 0$,*

$$\mathbb{E}V(X_t, Y_t, X_t^*) \le 0.$$

Proof. Let $t \ge 0$ be fixed. Observe first that $V(X_t, Y_t, X_t^*)$ is integrable: this is a consequence of local boundedness of V and the condition X, $Y \in L^\infty$. Furthermore, we have that $\overline{M} = \sup_{s \le t} ||M(X_s, Y_s, X_s^*)||_\infty$ is finite, which follows from $X_s^* \ge |X_0^1| \ge \delta$ and the boundedness property of M. By (ii), it suffices to show that

$$\mathbb{E}U(X_t, Y_t, X_t^*) \le 0. \tag{2.2}$$

For a fixed $\varepsilon > 0$, there is $D = D(\varepsilon) \geq 1$ such that if $d \geq D$, then

$$\mathbb{E} \sum_{k>d} [X^k, X^k]_t = \mathbb{E} \sum_{k>d} |X_t^k|^2 < \varepsilon. \tag{2.3}$$

For $0 \leq s \leq t$, let

$$\begin{aligned} X_s^{(d)} &= (X_s^1, X_s^2, \ldots, X_s^d, 0, 0, \ldots), \\ Y_s^{(d)} &= (Y_s^1, Y_s^2, \ldots, Y_s^d, 0, 0, \ldots) \end{aligned}$$

and

$$Z_s^{(d)} = (X_s^{(d)}, Y_s^{(d)}, X_s^{(d)*}).$$

Since $X^{(d)}$, $Y^{(d)}$ take values in a finite-dimensional subspace of $\mathcal{H}$, Ito's formula is applicable, which, together with (i), gives

$$U(Z_t^{(d)}) = U(Z_0^{(d)}) + I_1 + I_2 + I_3 \leq I_1 + I_2 + I_3, \tag{2.4}$$

where

$$\begin{aligned} I_1 &= \int_0^t U_x(Z_s^{(d)}) dX_s^{(d)} + \int_0^t U_y(Z_s^{(d)}) dY_s^{(d)}, \\ I_2 &= \int_0^t U_z(Z_s^{(d)}) dX_s^{(d)*}, \\ I_3 &= \int_0^t U_{xx}(Z_s^{(d)}) d[X^{(d)}, X^{(d)}]_s + 2 \int_0^t U_{xy}(Z_s^{(d)}) d[X^{(d)}, Y^{(d)}]_s \\ &\quad + \int_0^t U_{yy}(Z_s^{(d)}) d[Y^{(d)}, Y^{(d)}]_s. \end{aligned}$$

The random variable I_1 has zero expectation. The term I_2 is nonpositive: by (iii), we have $U_z(Z_s^{(d)*}) \leq 0$ on $\{s : X_s^{(d)} = X_s^{(d)*}\}$. However, this set is precisely the support of $dX_s^{(d)*}$. It suffices to deal with I_3. Let $0 \leq s_0 < s_1 \leq t$. For any $j \geq 0$, let $(\eta_i^j)_{1 \leq i \leq i_j}$ be a sequence of nondecreasing finite stopping times with $\eta_0^j = s_0$, $\eta_{i_j}^j = s_1$ such that $\lim_{j \to \infty} \max_{1 \leq i \leq i_j - 1} |\eta_{i+1}^j - \eta_i^j| = 0$. Keeping j fixed, we apply, for each $i = 0, 1, 2, \ldots, i_j$, the property (iv) to $x = X_{s_0}^{(d)}$, $y = Y_{s_0}^{(d)}$, $z = X_{s_0}^{(d)*}$ and $h = h_i^j = X_{\eta_{i+1}^j}^{(d)} - X_{\eta_i^j}^{(d)}$, $k = k_i^j = Y_{\eta_{i+1}^j}^{(d)} - Y_{\eta_i^j}^{(d)}$. Summing the obtained $i_j + 1$ inequalities and letting $j \to \infty$ yield

$$\begin{aligned} \sum_{m=1}^d \sum_{n=1}^d \Big[& U_{x_m x_n}(Z_{s_0}^{(d)}) [(X^{(d)})^m, (X^{(d)})^n]_{s_0}^{s_1} + 2U_{x_m y_n}(Z_{s_0}^{(d)}) [(X^{(d)})^m, (Y^{(d)})^n]_{s_0}^{s_1} \\ & + U_{y_m y_n}(Z_{s_0}^{(d)}) [(Y^{(d)})^m, (Y^{(d)})^n]_{s_0}^{s_1} \Big] \\ & \leq \overline{M} \sum_{k=1}^d \big([(Y^{(d)})^k, (Y^{(d)})^k]_{s_0}^{s_1} - [(X^{(d)})^k, (X^{(d)})^k]_{s_0}^{s_1} \big), \end{aligned}$$

where we have used the notation $[S,T]_{s_0}^{s_1} = [S,T]_{s_1} - [S,T]_{s_0}$. If we approximate I_3 by discrete sums, we see that the inequality above leads to

$$\begin{aligned} I_3 &\le \overline{M}\sum_{k=1}^{d}\big([(Y^{(d)})^k,(Y^{(d)})^k]_0^t - [(X^{(d)})^k,(X^{(d)})^k]_0^t\big) \\ &\le -\overline{M}\sum_{k>d}\big([Y^k,Y^k]_0^t - [X^k,X^k]_0^t\big), \end{aligned}$$

where the last passage is due to the differential subordination.

Now take the expectation of both sides of (2.4) and use (2.3) to obtain $\mathbb{E}U(Z_t^{(d)}) \le \overline{M}\varepsilon$. Then let $d\to\infty$ to get $\mathbb{E}U(X_t,Y_t,X_t^*)\le \overline{M}\varepsilon$, by Lebesgue's dominated convergence theorem. Since ε was chosen arbitrarily, (2.2) follows. □

3. The special function corresponding to (1.2)

Now we will see how the method described in the previous section can be used to obtain (1.2). For $p\in[1,2)$, we will write $\beta=\beta_p=\sqrt{2/p}$. Let U, $V:\mathcal{H}\times\mathcal{H}\times(0,\infty)\to\mathbb{R}$ be given by

$$U(x,y,z) = \frac{p}{2}\beta^{p-2}(|y|^2-|x|^2-(\beta^2-1)z^2)z^{p-2} \tag{3.1}$$

and $V(x,y,z)=|y|^p-\beta^p z^p$.

Lemma 3.1. *The function U belongs to the class $\mathcal{U}(V)$.*

Proof. The property (i) is evident. To prove the majorization, observe that we may assume that $z=1$, due to the homogeneity. Now, by the mean value property of the concave function $t\mapsto t^{p/2}$,

$$\begin{aligned} U(x,y,1) &= \frac{p}{2}\beta^{p-2}(|y|^2-|x|^2-(\beta^2-1)) \ge \frac{p}{2}\beta^{p-2}(|y|^2-\beta^2) \\ &= \frac{p}{2}(\beta^2)^{p/2-1}(|y|^2-\beta^2)\ge |y|^p-\beta^p = V(x,y,1). \end{aligned}$$

To check (iii), note that

$$U_z(x,y,|x|) = \frac{p(p-2)}{2}\beta^{p-2}|y|^2|x|^{p-3}\le 0.$$

Finally, observe that

$$(U_{xx}(x,y,z)h,h)+2(U_{xy}(x,y,z)h,k)+(U_{yy}(x,y,z)k,k) = p\beta^{p-2}z^{p-2}(|k|^2-|h|^2)$$

and the function $M(x,y,z)=p\beta^{p-2}z^{p-2}$ has the required boundedness property. This completes the proof. □

Proof of (1.2). We start with some standard reductions. We may assume that $||X^*||_p<\infty$. For a fixed n, consider the stopping time $T_n=\inf\{t:|X_t|\ge n \text{ or } |Y_t|\ge n\}$. For $t\ge 0$, $\delta>0$, consider the martingales $(X_t^{(n,\delta)})$ and $(Y_t^{(n,\delta)})$, taking values in the Hilbert space $\mathbb{R}\times\mathcal{H}$, given by

$$X_t^{(n,\delta)} = \begin{cases} (\delta, X_{T_n\wedge t}) & \text{if } T_n>0, \\ (\delta,0) & \text{if } T_n=0, \end{cases}$$

with a similar definition for $Y_t^{(n,\delta)}$. Since $U\in\mathcal{U}(V)$ and $X^{(n,\delta)}$, $Y^{(n,\delta)}$ satisfy the conditions of Theorem 2.1, we have, for any $t\ge 0$,

$$\mathbb{E}|Y_t^{(n,\delta)}|^p \le \beta^p\mathbb{E}(X_t^{(n,\delta)*})^p \le \beta^p\mathbb{E}(\delta^2+(X^*)^2)^{p/2}.$$

Now let $n \to \infty$ and then $\delta \to 0$ to obtain $||Y_t||_p \leq \beta ||X^*||_p$, by Fatou's lemma and Lebesgue's dominated convergence theorem. It suffices to take the supremum over t to get the claim. $\square$

Remark. The function U leading to the inequality (1.2) is not unique. For example, one can try to work with the following one. First introduce the auxiliary $\Phi : [\beta - 1, \infty) \to \mathbb{R}$ by

$$\Phi(t) = e^{-t} \left[-p \int_{\beta-1}^{t} e^s (s+1)^{p-1} ds + p\beta^{p-2}(1-\beta) e^{\beta - 1} \right], \tag{3.2}$$

and let $u : [0,1] \times [0,\infty) \to \mathbb{R}$ be given by

$$u(x,y) = \begin{cases} (y-x+1)^p + (1-x)\Phi(y-x) - \beta^p, & \text{if } y - x \geq \beta - 1, \\ \frac{1}{2} p \beta^{p-2}(y^2 - x^2 - \beta^2 + 1), & \text{if } y - x < \beta - 1. \end{cases} \tag{3.3}$$

The special function $U : \mathbb{R}^d \times \mathbb{R}^d \times (0,\infty)$ is defined by

$$U(x,y,z) = (|x| \vee z)^p u\left(\frac{|x|}{|x| \vee z}, \frac{|y|}{|y| \vee z} \right). \tag{3.4}$$

This function has almost all the properties (i) – (iv) listed in the previous section. However, it is not of class C^2 on the set $\{(x,y,z) : |y| - |x| = (\beta - 1)(|x| \vee z)\}$ and one needs to use smoothing arguments to overcome this difficulty. We omit the details.

4. On optimality of the constants

One could try to show the sharpness of the estimates (1.2) and (1.3) by providing appropriate examples. However, we will proceed in a different manner and prove, in a certain sense, the converse to Theorem 2.1: the validity of a maximal inequality for martingales implies the existence of a certain special function. Such an approach is not new and has been successful in a number of papers; see e.g. [3], [6] and [7].

For any $(x,y) \in \mathbb{R} \times \mathbb{R}$, let $\mathcal{M}(x,y)$ denote the class of all pairs (X,Y) of real-valued, bounded and continuous-path martingales such that $(X_0, Y_0) = (x,y)$ almost surely and, for some predictable process H taking values in $\{-1,1\}$, we have $Y_t = y + \int_{(0,t]} H_s dX_s$, $t \geq 0$. Here the probability space and filtration may be different for different pairs. Observe that if $(X,Y) \in \mathcal{M}(x,y)$, then the limits X_∞^*, Y_∞ exist and are finite with probability 1, due to the boundedness and continuity assumptions.

Let $1 \leq p < \infty$ be fixed and suppose that κ is the best constant in the inequality

$$||Y||_p \leq \kappa ||X^*||_p, \tag{4.1}$$

to be valid for all $(X,Y) \in \mathcal{M}(x,y)$ such that $y = \pm x$. Let $\kappa' > \kappa$ and consider a function $W : \mathbb{R} \times \mathbb{R} \times [0,\infty) \to \mathbb{R}$, given by

$$W(x,y,z) = \sup_{\mathcal{M}(x,y)} \{ \mathbb{E}|Y_\infty|^p - (\kappa')^p \mathbb{E}(X_\infty^* \vee z)^p \}.$$

Lemma 4.1. *The function W has the following properties:*

(i) $W(x,y,z) < \infty$ for any $(x,y,z) \in \mathbb{R} \times \mathbb{R} \times [0,\infty)$,

(ii) For any $(x,y,z) \in \mathbb{R} \times \mathbb{R} \times [0,\infty)$,

$$W(\lambda x, \pm \lambda y, |\lambda| z) = |\lambda|^p W(x,y,z) \text{ for } \lambda \neq 0. \tag{4.2}$$

(iii) For any $(x, y, z) \in \mathbb{R} \times \mathbb{R} \times [0, \infty)$,

$$W(x, y, z) \geq |y|^p - (\kappa')^p (|x| \vee z)^p. \tag{4.3}$$

(iv) For any $(x, y, z) \in \mathbb{R} \times \mathbb{R} \times [0, \infty)$, $|x| \leq z$, $\varepsilon \in \{-1, 1\}$, *if* $\alpha_1, \alpha_2 \in (0, 1)$ *and* $t_1, t_2 \in (-x - z, -x + z)$ *satisfy* $\alpha_1 + \alpha_2 = 1$ *and* $\alpha_1 t_1 + \alpha_2 t_2 = 0$, *then*

$$W(x, y, z) \geq \alpha_1 W(x + t_1, y + \varepsilon t_1, z) + \alpha_2 W(x + t_2, y + \varepsilon t_2, z).$$

(v) If $y \in \mathbb{R}$, $\delta > 0$ *and* $a \in (0, 2)$, *then*

$$\begin{aligned} W(1, y, 1) &\geq \frac{\delta}{a + \delta} W(1 - a, y + a, 1 + \delta) + \frac{a}{a + \delta} W(1 + \delta, y - \delta, 1 + \delta) \\ &\geq \frac{\delta}{a + \delta} [W(1 - a, y + a, 1) + (\kappa')^p (1 - (1 + \delta)^p)] \\ &\quad + \frac{a}{a + \delta} (1 + \delta)^p W(1, \frac{y - \delta}{1 + \delta}, 1). \end{aligned} \tag{4.4}$$

Proof. (i) Let $(x, y, z) \in \mathbb{R} \times \mathbb{R} \times [0, \infty)$ and $(X, Y) \in \mathcal{M}(x, y)$. Since $(Y - y + x, X) \in \mathcal{M}(x, x)$, we have, by the triangle inequality and (4.1),

$$\begin{aligned} ||Y_\infty||_p &\leq ||(Y - y + x)_\infty||_p + |y - x| \leq \kappa ||X^*_\infty||_p + |y - x| \\ &\leq ((\kappa')^p ||X^*_\infty \vee z||_p^p + C \cdot |y - x|^p)^{1/p}, \end{aligned}$$

where C depends only on κ, κ' and p. Thus $W(x, y, z) \leq C \cdot |y - x|^p$.

(ii) This is immediate: $(X, Y) \in \mathcal{M}(x, y)$ if and only if $(\lambda X, \pm \lambda Y) \in \mathcal{M}(\lambda x, \pm \lambda y)$.

(iii) This follows from the fact that the pair (x, y) of constant martingales belongs to $\mathcal{M}(x, y)$.

(iv) We will use the continuous analogue of the so-called "splicing" argument; see e.g. [1] for details. Let $(X^i, Y^i) \in \mathcal{M}(x + t_i, y + \varepsilon t_i)$ and let H^i be the corresponding predictable processes, $i = 1, 2$. Intuitively speaking, we will "glue" the two pairs using Brownian motion and obtain a pair belonging to $\mathcal{M}(x, y)$. To this end, we may and do assume that these processes are given on the same probability space equipped with the same filtration. Suppose B is a Brownian motion starting from x, independent of these two pairs. Let $\tau = \inf\{t : B_t \in \{x + t_1, x + t_2\}\}$ and set

$$X_t = \begin{cases} B_t & \text{if } t \leq \tau, \\ X^i_{t - \tau} & \text{if } t > \tau \text{ and } B_\tau = x + t_i, \end{cases}$$

$$H_t = \begin{cases} \varepsilon & \text{if } t \leq \tau, \\ H^i_{t - \tau} & \text{if } t > \tau \text{ and } B_\tau = x + t_i \end{cases}$$

and $Y_t = y + \int_{(0, t]} H_s dX_s$. It is easy to see that X is a martingale with respect to the natural filtration and $(X, Y) \in \mathcal{M}(x, y)$. Now we have, with probability 1,

$$Y_\infty = Y^1_\infty 1_{\{B_\tau = x + t_1\}} + Y^2_\infty 1_{\{B_\tau = x + t_2\}} \tag{4.5}$$

and, since $|x + t_1|$, $|x + t_2| \leq z$,

$$X^*_\infty \vee z = (X^{1*}_\infty \vee z) 1_{\{B_\tau = x + t_1\}} + (X^{2*}_\infty \vee z) 1_{\{B_\tau = x + t_2\}}. \tag{4.6}$$

Therefore

$$\begin{aligned} W(x,y,z) &\geq \mathbb{E}|Y_\infty|^p - (\kappa')^p \mathbb{E}(X_\infty^* \vee z)^p \\ &= \sum_{i=1}^{2} (\mathbb{E}|Y_\infty^i|^p - (\kappa')^p \mathbb{E}(X_\infty^{i*} \vee z)^p) \mathbb{P}(B_\tau = x + t_i) \\ &= \sum_{i=1}^{2} \alpha_i (\mathbb{E}|Y_\infty^i|^p - (\kappa')^p \mathbb{E}(X_\infty^{i*} \vee z)^p). \end{aligned} \tag{4.7}$$

Now take the supremum on the right-hand side over the classes $\mathcal{M}(x+t_1, y+\varepsilon t_1)$ and $\mathcal{M}(x+t_2, y+\varepsilon t_2)$ to obtain the desired estimate.

(v) We repeat the argument from the previous part, with $x = z = 1$, $\varepsilon = -1$, $t_1 = -a$ and $t_2 = \delta$. The equation (4.5) remains valid; however, (4.6) is no longer true: we still have $(X_\infty^* \vee 1)1_{\{B_\tau = 1+\delta\}} = (X_\infty^{2*} \vee 1)1_{\{B_\tau = 1+\delta\}}$, but the equality $(X_\infty^* \vee 1)1_{\{B_\tau = 1-a\}} = (X_\infty^{1*} \vee 1)1_{\{B_\tau = 1-a\}}$ does not hold in general. Nonetheless, we have the inequality

$$X_\infty^* \vee 1 \leq (X_\infty^{1*} \vee (1+\delta))1_{\{B_\tau = 1-a\}} + (X_\infty^{2*} \vee 1)1_{\{B_\tau = 1+\delta\}}. \tag{4.8}$$

But, by the very definition of W,

$$W(x,y,z_1) \leq W(x,y,z_2) \quad \text{for } z_1 \geq z_2, \tag{4.9}$$

so arguing as in (4.7), we get the first inequality in (4.4). To deal with the second one, we need to compare $W(1-a, y+a, 1+\delta)$ and $W(1-a, y+a, 1)$. Note that for x, $\delta \geq 0$ we have $(x \vee (1+\delta))^p - (x \vee 1)^p \leq (1+\delta)^p - 1$. Thus, for any $(X,Y) \in \mathcal{M}(1-a, y+a)$,

$$\begin{aligned} W(1-a, y+a, 1+\delta) &\geq \mathbb{E}|Y_\infty|^p - (\kappa')^p \mathbb{E}\big(X_\infty^* \vee (1+\delta)\big)^p \\ &\geq \mathbb{E}|Y_\infty|^p - (\kappa')^p \mathbb{E}\big(X_\infty^* \vee 1\big)^p + (\kappa')^p (1 - (1+\delta)^p). \end{aligned}$$

Taking the supremum over all such (X,Y) yields

$$W(1-a, y+a, 1+\delta) \geq W(1-a, y+a, 1) + (\kappa')^p (1 - (1+\delta)^p).$$

Thus the second inequality in (4.4) follows, since, by the homogeneity of W,

$$W(1+\delta, y-\delta, 1+\delta) = (1+\delta)^p W(1, \frac{y-\delta}{1+\delta}, 1). \qquad \square$$

Now we are ready to prove that the inequalities (1.2) and (1.3) are sharp.

Sharpness of (1.2). We keep the notation $\beta = \beta_p = \sqrt{2/p}$. Apply (iv) with $x = \beta/2$, $y = 1 - \beta/2$, $z = 1$, $\varepsilon = -1$, $t_1 = 1 - \beta$ and $t_2 = 1 - \beta/2$ (α_1 and α_2 are uniquely determined by t_1 and t_2) to get

$$W\left(\frac{\beta}{2}, 1 - \frac{\beta}{2}, 1\right) \geq \frac{2-\beta}{\beta} W\left(1 - \frac{\beta}{2}, \frac{\beta}{2}, 1\right) + \frac{2\beta - 2}{\beta} W(1,0,1). \tag{4.10}$$

The condition (iv) with $x = 1 - \beta/2$, $y = \beta/2$, $z = 1$, $\varepsilon = 1$, $t_1 = \beta/2$ and $t_2 = -1$ implies that

$$W\left(1 - \frac{\beta}{2}, \frac{\beta}{2}, 1\right) \geq \frac{2}{\beta+2} W(1, \beta, 1) + \frac{\beta}{\beta+2} W\left(-\frac{\beta}{2}, -1 + \frac{\beta}{2}, 1\right). \tag{4.11}$$

By homogeneity, $W\left(-\frac{\beta}{2},-1+\frac{\beta}{2},1\right)=W\left(\frac{\beta}{2},1-\frac{\beta}{2},1\right)$. Furthermore, (iii) implies that $W(1,\beta,1)\geq\beta^p-(\kappa')^p$; combining this with (4.10) and (4.11) gives

$$\frac{\beta^2}{2}W\left(\frac{\beta}{2},1-\frac{\beta}{2},1\right)\geq\left(1-\frac{\beta}{2}\right)(\beta^p-(\kappa')^p)+\left(\frac{\beta^2}{2}+\frac{\beta}{2}-1\right)W(1,0,1). \tag{4.12}$$

Now exploit (v), with $y=0$, $a=1-\beta/2$ and $\delta>0$, to obtain

$$\begin{aligned}W(1,0,1)\geq{}&\frac{2\delta}{2-\beta+2\delta}\left[W\left(\frac{\beta}{2},1-\frac{\beta}{2},1\right)+(\kappa')^p\big(1-(1+\delta)^p\big)\right]\\&+\frac{2-\beta}{2-\beta+2\delta}(1+\delta)^pW\left(1,-\delta/(1+\delta),1\right).\end{aligned} \tag{4.13}$$

By (ii), $W\left(1,-\delta/(1+\delta),1\right)=W(1,\delta/(1+\delta),1)$; moreover, if we use the first inequality in (4.4), with y, a, δ replaced by the numbers $\delta/(1+\delta)$, $(1-\frac{\beta}{2})\frac{1+2\delta}{1+\delta}-\frac{\delta}{1+\delta}$ and $\delta/(1+\delta)$, respectively, we obtain

$$\begin{aligned}W\left(1,\frac{\delta}{1+\delta},1\right)\geq{}&\frac{2\delta}{(2-\beta)(1+2\delta)}\left(\frac{1+2\delta}{1+\delta}\right)^pW\left(\frac{\beta}{2},\left(1-\frac{\beta}{2}\right),1\right)\\&+\frac{(2-\beta)(1+2\delta)-2\delta}{(2-\beta)(1+2\delta)}\left(\frac{1+2\delta}{1+\delta}\right)^pW\left(1,0,1\right).\end{aligned}$$

Plug this into (4.13), subtract $W(1,0,1)$ from both sides, divide throughout by 2δ and let $\delta\to0$ to obtain

$$0\geq\frac{2W(\beta/2,1-\beta/2,1)}{2-\beta}+\left(p-\frac{2}{2-\beta}\right)W(1,0,1).$$

Combining this with (4.12) and using $\beta=\sqrt{2/p}$ give

$$\frac{1}{p}W\left(\frac{\beta}{2},1-\frac{\beta}{2},1\right)\geq\left(1-\frac{1}{\sqrt{2p}}\right)\left(\left(\frac{2}{p}\right)^{p/2}-(\kappa')^p\right)+\frac{1}{p}W\left(\frac{\beta}{2},1-\frac{\beta}{2},1\right)$$

or $\kappa'\geq\sqrt{2/p}$. Since $\kappa'>\kappa$ was arbitrary, we conclude that the constant $\sqrt{2/p}$ cannot be replaced in (1.2) by a smaller one. □

Sharpness of (1.3). Apply (iv) with $x=0$, $y=p$, $z=1$, $\varepsilon=-1$, $t_1=1$ and $t_2=-\delta$ to get

$$\begin{aligned}W(0,p,1)&\geq\frac{\delta}{1+\delta}W(1,p-1,1)+\frac{1}{1+\delta}W(-\delta,p+\delta,1)\\&\geq\frac{\delta}{1+\delta}((p-1)^p-(\kappa')^p)+\frac{1}{1+\delta}W(\delta,p+\delta,1),\end{aligned} \tag{4.14}$$

where we have used the majorization (iii) and the homogeneity. By (iv) again, this time with $x=\delta$, $y=p+\delta$, $z=1$, $\varepsilon=-1$, $t_1=1-\delta$ and $t_2=-\delta$, we get

$$\begin{aligned}W(\delta,p+\delta,1)&\geq\delta W(1,p+2\delta-1,1)+(1-\delta)W(0,p+2\delta,1)\\&\geq\delta W(1,p+2\delta-1,1)+(1-\delta)W(0,p+2\delta,1+2\delta/p),\end{aligned} \tag{4.15}$$

where in the last passage we have exploited (4.9). Now by (v), with y, a, δ replaced by $p+2\delta-1$, 1 and $2\delta/p$, respectively, together with the majorization and

homogeneity, we may write

$$\begin{aligned} W(1,p+2\delta-1,1) &\geq \frac{2\delta}{p+2\delta}W(0,p+2\delta,1+2\delta/p) \\ &\quad + \frac{p}{p+2\delta}W(1+2\delta/p,(1+2\delta/p)(p-1),1+2\delta/p) \\ &\geq \frac{2\delta}{p+2\delta}(1+2\delta/p)^p W(0,p,1) \\ &\quad + \frac{p}{p+2\delta}(1+2\delta/p)^p((p-1)^p-(\kappa')^p). \end{aligned} \tag{4.16}$$

Now combine (4.15) with (4.16) and insert the obtained lower bound for $W(\delta,p+\delta,1)$ into (4.14). We obtain the estimate, which, after subtracting $W(0,p,1)$ from both sides, dividing throughout by δ and letting $\delta\to 0$, becomes

$$0 = W(0,p,1)\cdot\lim_{\delta\to 0}\left[1-\frac{1-\delta}{1+\delta}\left(1+\frac{2\delta}{p}\right)^p\right] \geq 2((p-1)^p-(\kappa')^p).$$

This implies that $\kappa'\geq p-1$ and, consequently, that $\kappa\geq p-1$. □

5. Concluding remarks

5.1. **On the search for the suitable majorant.** We will sketch some steps which led to the right choice of the optimal constant β_p, $1<p<2$, and the right guess of the special function U used in the proof of (1.2).

First we will exhibit arguments which yield the function described in the Remark at the end of Section 3. Let $p\in[1,2)$ be fixed and write down the desired inequality

$$||Y||_p \leq \beta||X^*||_p,$$

with the optimal $\beta\geq 1$ to be determined; some experimentation shows that β should be smaller than 2. Let us restrict ourselves to the real-valued martingales: $\mathcal{H}=\mathbb{R}$. Since the function $V(x,y,z)=|y|^p-\beta^p z^p$ is homogeneous of order p, we assume that U also has this property. Let $u(x,y)=U(x,y,1)$ for $x\in[-1,1]$ and $y\in\mathbb{R}$. In addition, assume that U is of class C^1 on $\mathbb{R}\times\mathbb{R}\times(0,\infty)$.

It is natural to expect that there should be some similarities between U and the special function used by Burkholder in [4]: both functions concern essentially the same maximal inequality (strictly speaking, this is the case if $p=1$; but for $p>1$, the conditions imposed below also lead to the right function). Burkholder's majorant suggests that we should search for u in the class of the functions satisfying the assumptions (A1)–(A3):

(A1) For all $x\in[-1,1]$ and $y\in\mathbb{R}$,

$$u(x,y)=u(-x,y)=u(x,-y). \tag{5.1}$$

(A2) For $0<x<1$ and $x\leq y$, then

$$u(x,y)=(1-x)A(-x+y)+xB(1-x+y), \tag{5.2}$$

and if $0<y<x<1$, then

$$u(x,y)=\frac{1-x}{1-x+y}C(x-y)+\frac{y}{1-x+y}B(1-x+y), \tag{5.3}$$

where $A=u(0,\cdot)$, $B=u(1,\cdot)$ and $C=u(\cdot,0)$.

(A3) For all $y\geq\beta$ we have $u(1,y)=U(1,y,1)=V(1,y,1)$.

Lemma 5.1. *If u satisfies the above conditions, then for $x \in [0,1]$ and $y \geq x+\beta-1$, we have*

$$u(x,y) = (y-x+1)^p - \beta^p + (1-x)e^{x-y}\left[A(\beta-1)e^{\beta-1} - p\int_{\beta-1}^{-x+y} e^s(s+1)^{p-1}ds\right].$$

Proof. By (5.2), for any $y \geq \beta - 1$ and $\delta \in (0,1)$, we have

$$u(\delta, y+\delta) = (1-\delta)A(y) + \delta B(1+y).$$

Subtracting $A(y)$ from both sides, dividing by δ and letting $\delta \to 0$ yields $u_x(0,y) + u_y(0,y) = -A(y) + B(1+y)$. But $u_y(0,y) = A'(y)$ and, by (5.1), $u_x(0,y) = 0$, so we obtain

$$A'(y) = -A(y) + (y+1)^p - \beta^p.$$

Solving this differential equation gives

$$A(y) = -pe^{-y}\int_{\beta-1}^{y} e^s(s+1)^{p-1}\mathrm{d}s + (y+1)^p - \beta^p + A(\beta-1)e^{-y+\beta-1},$$

and plugging this into (5.2) yields the claim. □

Now we will find the function u on the remaining part of the domain. It is easy to see that the property (iv) from the definition of $\mathcal{U}(V)$ implies that the function $w : s \mapsto u(s, 1-s)$, $s \in [0,1]$, is concave. From the lemma above, we know the explicit form of w on the interval $[0, 1-\beta/2]$. Some experimentation suggests the following assumption, which is *not* satisfied by Burkholder's majorant, but in our case leads to the right function. This is the key condition.

(A4) The function w is linear on $[1-\beta/2, 1]$.

Lemma 5.2. *Under the assumptions (A1)–(A4), for all x, $y \in [0,1]$ such that $x+y \leq 1$ and $-x+y \leq \beta-1$, we have*

$$u(x,y) = \frac{p\beta^{p-2}}{2}(y^2 - x^2 - \beta^2 + 1). \tag{5.4}$$

Proof. Let

$$a_\beta = (\beta-1)A(\beta-1) + p\beta^{p-1}(\beta-2). \tag{5.5}$$

Since u is of class C^1, we obtain that for $s \in [1-\beta/2, 1]$,

$$\begin{aligned} w(s) &= u(1-\beta/2, \beta/2) + (u_x(1-\beta/2,\beta/2) - u_y(1-\beta/2,\beta/2))(s-1+\beta/2) \\ &= \frac{\beta}{2}A(\beta-1) + a_\beta(s-1+\beta/2). \end{aligned}$$

Suppose that $y \in [0, \beta-1]$. By (5.2), the function $s \mapsto u(s, y+s)$ is linear, so, for $0 < \delta < (1-y)/2$,

$$u(\delta, y+\delta) = \frac{2\delta}{1-y}w\left(\frac{1-y}{2}\right) + \frac{1-y-2\delta}{1-y}A(y).$$

Subtract $A(y)$ from both sides, divide throughout by δ and let $\delta \to 0$ to obtain

$$A'(y) = u_x(0,y) + u_y(0,y) = -\frac{2A(y)}{1-y} + \frac{2}{1-y}w\left(\frac{1-y}{2}\right).$$

Solving the differential equation, we get

$$A(y) = \frac{\beta}{2}A(\beta-1) + a_\beta\left(\frac{\beta}{2} - y\right) + \frac{A(\beta-1) - a_\beta}{2(2-\beta)}(1-y)^2.$$

By (5.1), $A'(0)=0$; this gives $A(\beta-1)=a_\beta(\beta-1)$, so, by (5.5),

$$a_\beta=-p\beta^{p-2}\qquad\text{and}\qquad A(y)=\frac{p\beta^{p-2}}{2}(y^2+1-\beta^2).$$

This enables us to obtain (5.4) for $y\geq x$: it suffices to use

$$u(x,y)=\frac{2x}{1+x-y}w\left(\frac{1-y}{2}\right)+\frac{1-x-y}{1+x-y}A(y-x),$$

which follows directly from (5.2).

If $y<x$, we proceed similarly: by (5.3), we have, for $x\in[0,1)$ and $0<\delta<(1-x)/2$,

$$u(x+\delta,\delta)=\frac{1-x-2\delta}{1-x}C(x)+\frac{2\delta}{1-x}w\left(\frac{1+x}{2}\right).$$

Subtract $C(x)$ from both sides, divide by δ and let $\delta\to 0$ to obtain a differential equation for C. Solve it and use $C'(0)=0$ to get $C(x)=p\beta^{p-2}(-x^2-\beta^2+1)/2$. To obtain (5.4), apply the following consequence of (5.3):

$$u(x,y)=\frac{2y}{1-x+y}w\left(\frac{1+x-y}{2}\right)+\frac{1-x-y}{1-x+y}C(x-y).\qquad\square$$

Lemma 5.3. *If u satisfies (A1)–(A4), then $\beta\geq\sqrt{2/p}$.*

Proof. Since U is homogeneous of order p, we have $x^pu(1,0)=U(x,0,x)$ for $x>0$. Differentiating at 1 and using the property (iii) from the definition of class $\mathcal{U}(V)$, we get

$$pC(1)=pu(1,0)=U_x(1,0,1)+U_z(1,0,1)\leq U_x(1,0,1)=C'(1),$$

which is the claim. $\square$

We impose the following condition.

(A5) We have $\beta=\sqrt{2/p}$.

To complete the description of u, it remains to guess its values on the set $E=\{(x,y):x\in[0,1],\,1-x<y<x+\beta-1\}$. Here is our final assumption.

(A6) For $(x,y)\in E$, the formula (5.4) is valid.

As one easily checks, we have thus obtained u given by (3.3). The description of U is completed by (3.4). As already mentioned in the Remark at the end of Section 3, this function is not sufficiently smooth, so Itô's formula is not directly applicable. However, in general, the majorant corresponding to a given martingale inequality is not uniquely determined. Sometimes, when U is described by different expressions on pairwise disjoint subsets, it is worth looking at $\overline{U}$, given by one of these expressions on the whole domain. Such an approach has been successful in a number of nonmaximal estimates. For example, the inequality (1.4) in the case $1<p<2$ can be shown using a function

$$u_p(x,y)=\begin{cases}p^{2-p}((p-1)|y|-|x|)(|x|+|y|)^{p-1} & \text{if } |x|\geq(p-1)|y|,\\ |y|^p-(p-1)^{-p}|x|^p & \text{if } |x|<(p-1)|y|,\end{cases}$$

which is optimal in the sense that it is the least in the class of all majorants leading to (1.4). However, the function $\overline{u}_p$, given by

$$\overline{u}_p(x,y)=p^{2-p}((p-1)|y|-|x|)(|x|+|y|)^{p-1}$$

for all x, $y \in \mathcal{H}$, also has all the required properties and is much easier to handle (for details, see e.g. [2], [3] and [8]).

Fortunately, a similar phenomenon occurs also in our case. The function U defined by (3.1), much simpler and more regular than the one just obtained above, is sufficient for our purposes.

5.2. **On the case $p > 2$.** There is a natural question about more explicit justification of the fact that the constant $p-1$ from (1.4) remains optimal in the stronger inequality (1.3). To answer it, we follow Burkholder [1] and recall the examples which yield the sharpness of (1.4) in the discrete-time setting. Fix $x > 0$ and denote by w the unique number larger than p, which satisfies $x^p + pw^{p-1} - w^p = 0$. Set $\theta = 1 - 1/w$ and

$$\pi_n = \left[\frac{x}{x + (n-1)\delta}\right]^w,$$

where $\delta > 0$. Define $g = (g_1, g_2, \ldots)$ on $[0, 1)$ by

$$g_n(s) = \begin{cases} x + (n-1)\delta & \text{if } 0 \le s < \pi_n, \\ \theta x s^{\theta - 1} & \text{if } \pi_n \le s < 1 \end{cases}$$

and let f be the transform of g by a sequence $v = (1, -1, 1, -1, \ldots)$; that is, we have $f_n = g_1 + \sum_{k=2}^{n} (-1)^{k-1}(g_k - g_{k-1})$ for $n = 1, 2, \ldots$. Then g is the transform of f by v, the sequences f, g are martingales with respect to the natural filtration and

$$\lim_{x \to 0} \lim_{\delta \to 0} ||f||_p = 1, \qquad \lim_{x \to 0} \lim_{\delta \to 0} ||g||_p = p - 1.$$

In fact, a careful study of the proof yields $\lim_{x \to 0} \lim_{\delta \to 0} ||f^*||_p = 1$. See pages 669–670 in [1] for details. Now it suffices to embed f in Brownian motion: denoting the obtained process by X, we see that g embeds as a stochastic integral, with respect to X, of some predictable process taking values in $\{-1, 1\}$. This yields the desired sharpness of (1.3).

It should also be stressed that the above argumentation does not carry over to the case $p < 2$. Analogous examples studied in [1] satisfy

$$\lim_{x \to 0} \lim_{\delta \to 0} ||f||_p = 1, \qquad \lim_{x \to 0} \lim_{\delta \to 0} ||g||_p = (p-1)^{-1},$$

but $\lim_{x \to 0} \lim_{\delta \to 0} ||f^*||_p$ is strictly larger than 1.

References

1. D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. **12** (1984), 647–702. MR744226 (86b:60080)
2. D. L. Burkholder, *A sharp and strict L^p-inequality for stochastic integrals*, Ann. Probab. **15**, no. 1 (1987), 268–273. MR877602 (88d:60156)
3. D. L. Burkholder, *Explorations in martingale theory and its applications*, École d'Été de Probabilités de Saint-Flour XIX—1989, 1–66, Lecture Notes in Math., 1464, Springer, Berlin, 1991. MR1108183 (92m:60037)
4. D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals,* Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994), pp. 343–358, Proc. Sympos. Appl. Math., 52, Amer. Math. Soc., Providence, RI, 1997. MR1440921 (98f:60103)
5. C. Dellacherie and P. A. Meyer, *Probabilities and potential. B.* Theory of martingales, North-Holland, Amsterdam, 1982. MR745449 (85e:60001)

6. A. Osękowski, *Sharp maximal inequality for stochastic integrals*, Proc. Amer. Math. Soc. **136** (2008), 2951–2958. MR2399063 (2009c:60137)
7. A. Osękowski, *Sharp maximal inequality for martingales and stochastic integrals*, Elect. Comm. in Probab. **14** (2009), 17–30. MR2472172 (2010d:60110)
8. G. Wang, *Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities*, Ann. Probab. **23** (1995), no. 2, 522–551. MR1334160 (96b:60120)

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