SHARP INEQUALITY FOR MARTINGALE MAXIMAL FUNCTIONS AND STOCHASTIC INTEGRALS

ADAM OSĘKOWSKI

Abstract. The paper contains the description of the optimal constant $\beta = 3.4351\ldots$ for which the following inequality holds. Let X be a real-valued martingale, H be a predictable process taking values in $[-1,1]$ and let Y be an Itô integral of H with respect to X. Then

$$||\sup_{t\geq 0}|Y_t|||_1 \leq \beta||\sup_{t\geq 0}|X_t|||_1.$$

A version of this inequality in the discrete-time case is also established. The proof is based on Burkholder's technique, which relates the above estimate to the construction of an upper solution to a corresponding nonlinear three-dimensional problem.

1. Introduction

The paper aims at answering a natural and basic question about the stochastic integrals, stated by Burkholder in [5]. Let us start with introducing the necessary notation. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with a nondecreasing right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathcal{F}$. Assume in addition, that $\mathcal{F}_0$ contains all the events of probability 0. Let $X = (X_t)_{t\geq 0}$ be an adapted real-valued martingale, which has right-continuous paths with limits from the left and suppose $H = (H_t)_{t\geq 0}$ is a predictable process taking values in the interval $[-1, 1]$. Let $Y = (Y_t)_{t\geq 0}$ be the Itô integral of H with respect to X, that is, for $t \geq 0$,

$$Y_t = H_0X_0 + \int_{(0,t]} H_s dX_s.$$

Let $X^* = \sup_{t\geq 0}|X_t|$ stand for the maximal function of X and let $||X||_p = \sup_{t\geq 0}||X_t||_p$ denote the p-th moment of X, $1 \leq p \leq \infty$.

In the literature, there has been an interest in the sharp comparison of the sizes of X and Y defined as above. An excellent source of information on this subject is the survey [4] by Burkholder (see also references therein), which contains moment, weak-type, exponential and escape inequalities, and much more. For more recent results, see the papers [10], [11] by the author and [15] by Suh.

In what follows, we will be particularly interested in maximal estimates. In [5], Burkholder introduced a method to determine the optimal constants in the problems of this type (see Section 2 below) and used it to establish the following sharp inequality.

2000 *Mathematics Subject Classification.* Primary: 60G42. Secondary: 60G44.

Key words and phrases. martingale, maximal function, stochastic integral, maximal inequality, moment inequality.

Partially supported by Foundation for Polish Science and MNiSW Grant N N201 397437.

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

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

where $\kappa = 2.536\ldots$ is the unique positive solution to the equation

$$\kappa = 3 - \exp \frac{1-\kappa}{2}.$$

Using Burkholder's technique, the author established two further results in this direction. In [8] it was shown that if the martingale X is nonnegative, then the optimal constant in (1.1) decreases to $2+(3e)^{-1} = 2.1226\ldots$. The paper [9] contains the proof of the related estimate, where the first moment of Y is replaced by the first moment of the one-sided maximal function of Y. To be precise, it was proved that if X, Y are as above, then

$$||\sup_{t\geq 0} Y_t||_1 \leq \eta ||X^*||_1, \tag{1.2}$$

where $\eta = 2.0856\ldots$ is the unique positive solution to the equation $2\log\left(\frac{8}{3}-\eta\right) = 1-\eta$. Furthermore, if X is assumed to be nonnegative, the best constant in (1.2) equals $14/9 = 1.555\ldots$.

In the present paper we focus on the bound

$$||Y^*||_1 \leq \beta ||X^*||_1, \tag{1.3}$$

a stronger estimate than the ones considered above. By (1.2) or Davis' square function inequality (cf. [7]), we see that (1.3) holds with some finite universal β. We will determine the optimal value of this constant.

Theorem 1.2. *The constant $\beta = 3.4351\ldots$ given by* (3.5) *below, is optimal in* (1.3)*. It is already the best possible even when the process H takes values in $\{-1,1\}$.*

It suffices to establish the discrete-time version of this result; then one obtains the continuous-time extension using approximation theorems due to Bichteler [1] (see [5] for analogous argumentation). Let us reformulate our problem in this new setting. Suppose that $f=(f_n)_{n\geq 0}$ is a discrete-time real-valued martingale, with a difference sequence $(df_n)_{n\geq 0}$ defined by $df_0 = f_0$ and $df_n = f_n - f_{n-1}$ for $n\geq 1$. Let $v=(v_n)_{n\geq 0}$ be a predictable sequence taking values in $[-1,1]$ and let $g=(g_n)_{n\geq 0}$ be a transform of f by v: that is, assume that $dg_n = v_n df_n$ for $n \geq 0$. In the particular case when v_n is deterministic and takes values in $\{-1,1\}$, we will write $dg_n = \pm df_n$. If f and g satisfy $dg_n = \pm df_n$ for all n, we will say that g is a ± 1-transform of f. We will use the notation $f^* = \sup_{n\geq 0}|f_n|$ and $||f||_p = \sup_n ||f_n||_p$, analogous to the one used in the continuous-time setting.

The discrete-time version of Theorem 1.2 can be stated as follows.

Theorem 1.3. *Assume that f, g are martingales such that g is a transform of f by a predictable sequence bounded in absolute value by* 1*. Then*

$$||g^*||_1 \leq \beta ||f^*||_1, \tag{1.4}$$

where β is given by (3.5)*. The constant β is the best possible. It is already the best possible even if g is assumed to be a ± 1-transform of f.*

A few words about the structure of the proof and the organization of the paper. We will use Burkholder's technique, which described in the next section. The method relates a given maximal inequality for martingales to an upper solution to a certain nonlinear problem. In contrast with the papers [5], [8] and [9], where the corresponding problems were two-dimensional, we will have to construct a function of three variables; this results in a considerable growth of difficulty of the calculations. The solution is presented in Section 4, using an auxiliary differential equation, studied in Section 3. Section 5 is the most elaborate and is devoted to the optimality of the constant β. In Section 6 we establish some technical facts needed in the earlier sections, and in the final part of the paper we present some concluding remarks as well as some open problems, which await further research.

2. Burkholder's method

Throughout this section we deal with the discrete-time setting. In order to apply Burkholder's technique, we need some reductions in (1.4). First, using standard approximation, we restrict ourselves to *simple* martingales f: that is, we assume that for any n the random variable f_n takes only a finite number of values and there is a deterministic N such that $f_N = f_{N+1} = f_{N+2} = \dots$ almost surely. Moreover, we may assume that $|f_0| > 0$ with probability 1. The next observation is that it suffices to show (1.4) for ± 1 transforms. It is an immediate consequence of the following fact.

Lemma 2.1. *Let g be the transform of a real valued martingale f by a real-valued predictable sequence v uniformly bounded in absolute value by 1. Then there exist real valued martingales $F^j = (F^j_n)_{n\geq 0}$ and Borel measurable functions $\varphi_j : [-1,1] \to \{-1,1\}$ such that, for $j \geq 1$ and $n \geq 0$,*

$$f_n = F^j_{2n+1}, \qquad f_n^* = (F^j_{2n+1})^*,$$
$$g_n = \sum_{j=1}^{\infty} 2^{-j}\varphi_j(v_0) G^j_{2n+1},$$

where G^j is the transform of F^j by $\varepsilon = (\varepsilon_k)_{k\geq 0}$ with $\varepsilon_k = (-1)^k$.

Proof. This is essentially Lemma A.1 in [4]. The only difference, which is the equality $f_n^* = (F^j_{2n+1})^*$, follows from the construction of the sequence (F^j) there. □

The final observation is that it suffices to prove

$$||g_n^*||_1 \leq \beta ||f_n^*||_1 \tag{2.1}$$

for any n. This is a consequence of the fact that f, and hence also g, are simple.

Now we are ready to describe the method. Let us explain it in the general setting. Suppose that $V : D \to \mathbb{R}$ is a fixed function, where $D = \mathbb{R} \times \mathbb{R} \times (0,\infty) \times (0,\infty)$. Assume that we are interested in the inequality

$$\mathbb{E}V(f_n, g_n, f_n^*, g_n^*) \leq 0, \tag{2.2}$$

for all n and all pairs (f,g) of simple martingales such that $\mathbb{P}(|f_0| > 0) = 1$ and g is a ± 1-transform of f. The key idea to study this problem is to introduce the class $\mathcal{U}(V)$ which consists of all functions $U : D \to \mathbb{R}$ satisfying the following properties: if $(x,y,z,w) \in D$, then

$$U(x,y,z,w) = U(x,y,|x|\vee z, |y| \vee w), \tag{2.3}$$

$$U(x,y,|x|,|y|) \leq 0 \quad \text{ if } |x| = |y| > 0, \tag{2.4}$$

$$V(x,y,z,w) \leq U(x,y,z,w), \tag{2.5}$$

and, furthermore, for any $|x| \leq z$, $|y| \leq w$, $\varepsilon \in \{-1, 1\}$, $\alpha \in (0,1)$ and t_1, $t_2 \in \mathbb{R}$ such that $\alpha t_1 + (1-\alpha)t_2 = 0$,

$$\alpha U(x+t_1, y+\varepsilon t_1, z, w) + (1-\alpha) U(x+t_2, y+\varepsilon t_2, z, w) \leq U(x,y,z,w). \tag{2.6}$$

The connection between the class $\mathcal{U}(V)$ and the inequality (2.2) is described in the following result.

Theorem 2.2. *If $\mathcal{U}(V)$ is nonempty, then* (2.2) *is valid.*

Proof. The argumentation is similar to the one used in the proof of Theorem 2.1 in [5]. Using (2.3) and (2.6), one proves that $(U(f_k, g_k, f_k^*, g_k^*))_{k \geq 0}$ is a supermartingale. Therefore, by (2.5) and then by (2.4),

$$\mathbb{E}V(f_n, g_n, f_n^*, g_n^*) \leq \mathbb{E}U(f_n, g_n, f_n^*, g_n^*) \leq \mathbb{E}U(f_0, g_0, f_0^*, g_0^*) \leq 0. \quad \square$$

We have the following result in the reverse direction, a slight modification of Theorem 2.2 from [5]. It will be the key tool to provide the lower bound for the constant β.

Theorem 2.3. *Suppose V satisfies* (2.3) *and assume that the inequality* (2.2) *holds for all n and all pairs (f, g) of simple martingales such that g is a ± 1-transform of f. Then the class $\mathcal{U}(V)$ is nonempty. Furthermore, let $U_0 : D \to \mathbb{R}$ be given by*

$$U_0(x,y,z,w) = \sup\{\mathbb{E}V(f_n, g_n, f_n^* \vee z, g_n^* \vee w)\}, \tag{2.7}$$

where the supremum is taken over all pairs (f, g) of simple martingales such that f starts from x, g starts from y and for all $n \geq 1$, $dg_n = \pm df_n$ almost surely. Then U_0 is the least element in $\mathcal{U}(V)$.

From now on, we will consider $V = V_\gamma$ given by

$$V_\gamma(x,y,z,w) = |y| \vee w - \gamma |x| \vee z$$

and our aim is to find the smallest γ for which the class $\mathcal{U}(V_\gamma)$ is nonempty. Due to the additional homogeneity of V_γ, we see that it suffices to search for a suitable U in the class of functions satisfying $U(\lambda x, \lambda y, \lambda z, \lambda w) = \lambda U(x,y,z,w)$ for all $\lambda > 0$ and $(x,y,z,w) \in D$ (see (2.7)). Thus the problem is three-dimensional. Its solution is studied in the next sections.

3. A differential equation and the optimal value of β

In order to introduce the special function and give the description of β, we need to consider the following auxiliary differential equation. Let $w > 1$ be a fixed number. Standard argumentation yields the existence and uniqueness of $Y = Y^w : [1, w] \to \mathbb{R}$ satisfying

$$Y'(t) = \frac{1}{2}\left(1 + \frac{1}{w}\right)(1+t)^{-2}\left[t^2 + 2(1-t)\left(\exp\left(\frac{t - Y(t)}{2}\right) - 1\right)\right], \tag{3.1}$$

$t \in (1, w)$, with the terminal condition $Y(w) = w$.

Lemma 3.1. *Let $w > 1$. Then Y^w is nondecreasing and*

$$Y^w(t) \geq t \qquad \text{for all } t, \text{ with equality only for } t = w. \tag{3.2}$$

Proof. Note that $2(1-t)\exp(\frac{t-Y^w(t)}{2}) \leq 0$ for $t \in (1,w)$, so

$$(Y^w)'(t) \leq \frac{1+w^{-1}}{2(1+t)^2}(t^2+2t-2) = \frac{1+w^{-1}}{2}\left(t+\frac{3}{1+t}\right)'.$$

Hence, since $t \leq w$ and $w > 1$,

$$\begin{aligned} Y^w(t) &\geq Y^w(w) - \frac{1+w^{-1}}{2}\left[w+\frac{3}{1+w}-t-\frac{3}{1+t}\right] \\ &\geq w - \frac{1+w^{-1}}{2}(w-t) \\ &= t + \frac{1-w^{-1}}{2}(w-t), \end{aligned} \tag{3.3}$$

which gives (3.2). This also implies $2(1-t)(\exp(\frac{t-Y^w(t)}{2})-1) \geq 0$ for $t \in (1,w)$, so

$$(Y^w)'(t) \geq \frac{1+w^{-1}}{2}\left(\frac{t}{1+t}\right)^2 \geq 0 \tag{3.4}$$

and the proof is complete. □

Let β be the positive number given by

$$\beta = \min_w \left\{Y^w(1) + \frac{5}{4}\left(1+\frac{1}{w}\right)\right\}. \tag{3.5}$$

It will be shown to be equal to the best constant in (1.4). Let us provide some approximation of β.

Lemma 3.2. *We have*

$$3.4142\ldots < \beta < 3.4358\ldots. \tag{3.6}$$

Proof. The number on the left is $2+\sqrt{2}$. To prove the bound, take $w > 1$ and use the first line of (3.3) with $t = 1$ to obtain

$$\begin{aligned} Y^w(1) + \frac{5}{4}\left(1+w^{-1}\right) &\geq w - \frac{1+w^{-1}}{2}\left(w+\frac{3}{1+w}-\frac{5}{2}\right) + \frac{5}{4}(1+w^{-1}) \\ &= 2 + w^{-1} + \frac{w}{2}. \end{aligned} \tag{3.7}$$

The expression on the right, as a function of $w \in (1,\infty)$, attains its minimum $2+\sqrt{2}$ for $w = \sqrt{2}$. This gives the left inequality in (3.6). To prove the right one, we proceed as previously, using a lower bound for $(Y^w)'$ coming from (3.4). After integration, we get

$$Y^w(1) + \frac{5}{4}\left(1+w^{-1}\right) \leq \left(1+\frac{1}{w}\right)[2-\log 2+\log(1+w)] + \frac{w}{2} - 1,$$

and the upper bound in (3.6) is the minimum of the expression on the right above. □

It is clear that the function $w \mapsto Y^w(1)+5/4(1+1/w)$ is continuous. In addition, it tends to $7/2 > \beta$ as $w \downarrow 1$ and, by (3.7), tends to infinity as $w \to \infty$. Hence the minimum defining β is attained for some w_0. To avoid the question about the uniqueness of w_0, let us take the smallest number with this property. Combining (3.7) with the right inequality in (3.6), we conclude that $1/w_0 + w_0/2 < 1.436$, so $1.18 < w_0 < 1.69$. To complete the discussion about the explicit values of β

and w_0, let us record here that numerical approximation gives $\beta = 3.4351\ldots$ and $w_0 = 1.302\ldots$.

A few words about some auxiliary notation. Throughout the paper, we will set $w_1 = Y^{w_0}(1)$ and write $\gamma = -1-1/w_0$. We will also use the function $y_0 : [w_1, w_0] \to [1, w_0]$, the inverse of Y^{w_0}; we will often skip the argument and write y_0 instead of $y_0(w)$. It can be verified readily that the function y_0 satisfies the differential equation

$$y_0' = -\frac{1}{\gamma} \cdot \frac{2(1+y_0)^2}{y_0^2 + 2y_0 - 2 + 2(1-y_0)\exp(\frac{y_0-w}{2})} \tag{3.8}$$

for $w \in (w_1, w_0)$. Moreover, in view of (3.2), we have

$$y_0(w) \leq w \text{ for } w \in [w_1, w_0], \text{ with equality only for } w = w_0. \tag{3.9}$$

We conclude this section with a technical fact, which will be needed later.

Lemma 3.3. *We have*

$$y_0' \geq 1 \tag{3.10}$$

and

$$y_0'(y_0 - 1) \leq y_0 + 1. \tag{3.11}$$

Proof. The first estimate follows immediately from the fact that

$$y_0' \geq -\frac{2}{\gamma} \cdot \frac{(1+y_0)^2}{y_0^2 + 2y_0 - 2}$$

and that both the factors are bigger than 1. To prove the second inequality, observe that by (3.9), we have $2(1-y_0)\exp(\frac{y_0-w}{2}) \geq 2(1-y_0)$. Plugging this into (3.8) gives

$$y_0'(y_0 - 1) \leq -\frac{1}{\gamma} \cdot \frac{2(1+y_0)^2}{y_0^2} \cdot (y_0 - 1),$$

so we will be done if we show that $2(y_0^2 - 1) \leq -\gamma y_0^2$. But

$$2(y_0^2 - 1) + \gamma y_0^2 = y_0^2(2+\gamma) - 2 \leq w_0^2(2+\gamma) - 2 = w_0(w_0 - 1) - 2 \leq 0,$$

the latter estimate coming from the bound $w_0 < 2$. $\square$

4. Proof of (1.4)

In this section we will construct an element U of $\mathcal{U}(V_\beta)$. Consider the following subsets of $[0,1] \times [0,\infty) \times (0,\infty)$:

$$\begin{aligned}
D_1 &= \{(x,y,w) : w \leq w_1,\ y \leq x\},\\
D_2 &= \{(x,y,w) : w \leq w_1,\ x < y \leq x + w_1 - 1\},\\
D_3 &= \{(x,y,w) : w \leq w_1,\ x + w_1 - 1 < y\},\\
D_4 &= \{(x,y,w) : w_1 < w \leq w_0,\ y \leq x + y_0(w) - 1\},\\
D_5 &= \{(x,y,w) : w_1 < w \leq w_0,\ x + y_0(w) - 1 < y \leq x + w - 1\},\\
D_6 &= \{(x,y,w) : w_1 < w \leq w_0,\ x + w - 1 < y\},\\
D_7 &= \{(x,y,w) : w > w_0,\ x + y \leq 1\},\\
D_8 &= \{(x,y,w) : w > w_0,\ 1 < x + y \leq 1 + w - w_0\},\\
D_9 &= \{(x,y,w) : w > w_0,\ 1 + w - w_0 < x + y\}.
\end{aligned}$$

Now we introduce an auxiliary function $u : [-1,1] \times \mathbb{R} \times (0,\infty) \to \mathbb{R}$. First we define it on the sets D_1–D_9.

$$u(x,y,w) = \begin{cases} -\frac{\gamma}{4}(y^2 - x^2 + 1) + \frac{5}{4}\gamma & \text{on } D_1, \\ 3\gamma - \gamma y + (x-2)\gamma \exp\left(\frac{x-y}{2}\right) & \text{on } D_2, \\ \gamma(3-y) + \gamma \exp\left(\frac{1-w_1}{2}\right)\left(-1 + y - w_1 - \frac{(y-w_1+1)^2 - x^2}{4}\right) & \text{on } D_3, \\ -\frac{\gamma}{2(1+y_0)}(y^2 - x^2 + 1) + w - \beta & \text{on } D_4, \\ \frac{2\gamma}{1+y_0} \exp\left(\frac{y_0 - y - 1 + x}{2}\right)(x-2) + \alpha(y,w) & \text{on } D_5, \\ \frac{2\gamma}{1+y_0} \exp\left(\frac{y_0 - w}{2}\right)\left(-1 + y - w - \frac{(y-w+1)^2 - x^2}{4}\right) + \alpha(y,w) & \text{on } D_6, \\ \frac{\exp(w_0 - w)}{2w_0}(y^2 - x^2 + 1) + w - \beta & \text{on } D_7, \\ \frac{(1-x)}{w_0}\exp(x + y + w_0 - w - 1) + w - \beta & \text{on } D_8, \\ \frac{(y - w + w_0)^2 - x^2 + 1}{2w_0} + w - \beta & \text{on } D_9, \end{cases}$$

where

$$\alpha(y,w) = \gamma(1-y) + \frac{\gamma y_0^2 + 2\gamma}{2(1+y_0)} + w - \beta.$$

We extend u to its whole domain $[-1,1] \times \mathbb{R} \times (0,\infty)$, setting

$$u(x,y,w) = u(-x,y,w) = u(x,-y,w) = u(-x,-y,w) \tag{4.1}$$

for all $x \in [0,1]$, $y \geq 0$ and $w > 0$.

In the lemma below we describe the main properties of the function u. For the sake of convenience and clarity of the arguments, the proofs of these technical facts are postponed to the last section.

Lemma 4.1. *(i) The function u is continuous. In addition, it is of class C^1 on each of the sets $\{(x,y,w) : w < w_1\}$, $\{(x,y,w) : w \in (w_1, w_0)\}$, $\{(x,y,w) : w > w_0\}$.*

(ii) For all $w > 0$ and $|y| \leq w$,

$$\lim_{\delta \downarrow 0} \frac{u(1,y,w) - u(1-\delta, y \pm \delta, w)}{\delta} \geq \gamma. \tag{4.2}$$

Furthermore, for all $w > 0$ and $x \in (-1,1]$,

$$\lim_{\delta \downarrow 0} \frac{u(x,w,w) - u(x-\delta, w-\delta, w)}{\delta} \geq \gamma. \tag{4.3}$$

(iii) For $x \in [-1,1]$ and $w \in (0,\infty) \setminus \{w_0, w_1\}$, we have $u_w(x,w,w) \leq 0$.

(iv) For any x, the function $H_x : (-1-x, 1-x) \to \mathbb{R}$, given by $H_x(t) = u_x(x+t,t,t) + u_y(x+t,t,t)$, is nonincreasing.

(v) The function $J : (0,\infty) \to \mathbb{R}$ given by $J(y) = u(1,y,y)$ is convex.

(vi) For any fixed $w > 0$, the function $u(\cdot,\cdot,w)$, restricted to the rectangle $[-1,1] \times [-w,w]$, is diagonally concave, i.e., concave along any line of slope ± 1.

(vii) For any fixed $w > 0$, the function $y \mapsto u(1,y,w)$ is nondecreasing on $[0,w]$.

(viii) For any $w > 0$ and $|y| \leq w$,

$$u(1,y,w) - (y-1)u_y(1,y,w) - wu_w(1,y,w) \leq \gamma. \tag{4.4}$$

(ix) For any $w > 0$ and $|x| \leq 1$, $|y| \leq w$ we have

$$u(x,y,w) \geq w - \beta.$$

(x) If $w \in (0,1]$ and $|x| = |y| \leq w$, then $u(x,y,w) \leq 0$.

Now we introduce the special function $U:\mathbb{R}\times\mathbb{R}\times(0,\infty)\times(0,\infty)\to\mathbb{R}$ by

$$U(x,y,z,w)=(|x|\vee z)u\left(\frac{x}{|x|\vee z},\frac{y}{|x|\vee z},\frac{|y|\vee w}{|x|\vee z}\right). \tag{4.5}$$

In the following theorem we establish the inequality (1.4).

Theorem 4.2. *The function U belongs to the class $\mathcal{U}(V_\beta)$.*

Proof. We check the conditions (2.3) – (2.6). The first property is evident from the definition of U. The inequality (2.4) follows from Lemma 4.1 (x). The majorization (2.5) is a consequence of part (ix), due to the homogeneity of U and V_β. The main technical difficulty lies in proving the property (2.6). Fix all the variables as in the statement and first note some reductions. Since U is homogeneous, we may and do assume that $z=1$. By the continuity of U, we are allowed to take $|x|<z$ and $|y|<w$. Moreover, since U satisfies $U(x,y,z,w)=U(x,-y,z,w)$, we may assume that $\varepsilon=1$. Now let $\Phi(t)=U(x+t,y+t,z,w)$ for $t\in\mathbb{R}$. Lemma 4.1 (i) guarantees the following regularity of this function: Φ is continuous and differentiable at 0. In fact, the derivative exists except for a finite number of points and the one sided derivatives of Φ exist everywhere.

To show the condition (2.6), it suffices to prove that

$$\Phi(t)\leq\Phi(0)+\Phi'(0)t \tag{4.6}$$

for positive t. Indeed, applying this to the function $\overline{\Phi}(t)=U(-x+t,-y+t,z,w)$ we get, for $t<0$,

$$\Phi(t)=\overline{\Phi}(-t)\leq\overline{\Phi}(0)+\overline{\Phi}'(0)(-t)=\Phi(0)+\Phi'(0)t.$$

Thus there is a linear function Ψ such that $\Phi\leq\Psi$ on $\mathbb{R}$ and $\Phi(0)=\Psi(0)$; this implies (2.6).

To show (4.6), it will be convenient to consider two cases.

The case $y\geq x+w-1$. It will be proved below that

$$\Phi \text{ is concave on the set } [0,w-y], \tag{4.7}$$

$$\Phi'(t+)\leq\Phi'((w-y)-) \text{ for } t\in(w-y,1-x), \tag{4.8}$$

$$\Phi \text{ is convex on } [1-x,\infty), \tag{4.9}$$

$$\lim_{t\to\infty}\Phi'(t)\leq\Phi'((w-y)-). \tag{4.10}$$

These properties clearly yield (4.6). The first condition is a consequence of Lemma 4.1 (vi). The property (4.8) follows from parts (iii) and (iv) of this lemma; indeed,

$$\begin{aligned}\Phi'(t+)&=\lim_{s\downarrow t}(u_x+u_y+u_w)(x+s,y+s,y+s)\\&\leq\limsup_{s\downarrow t}(u_x+u_y)(x+s,y+s,y+s)\\&\leq\limsup_{s\uparrow(w-y)}(u_x+u_y)(x+s,y+s,y+s)=\Phi'((w-y)-).\end{aligned}$$

We move to (4.9). Suppose that $t_1,\ t_2 \geq 1-x$ and $\alpha_1,\ \alpha_2 \in (0,1)$ satisfy $\alpha_1+\alpha_2=1$. Using Lemma 4.1 (v), we may write, for $\alpha_i'=\alpha_i(x+t_i)/(\alpha_1(x+t_1)+\alpha_2(x+t_2))$,

$$\begin{aligned}\alpha_1\Phi(t_1)+\alpha_2\Phi(t_2)&=(\alpha_1(x+t_1)+\alpha_2(x+t_2))\left[\alpha_1'J\left(\frac{y+t_1}{x+t_1}\right)+\alpha_2'J\left(\frac{y+t_2}{x+t_2}\right)\right]\\&\geq(\alpha_1(x+t_1)+\alpha_2(x+t_2))J\left(\alpha_1'\frac{y+t_1}{x+t_1}+\alpha_2'\frac{y+t_2}{x+t_2}\right)\\&=(\alpha_1(x+t_1)+\alpha_2(x+t_2))J\left(\frac{y+\alpha_1t_1+\alpha_2t_2}{\alpha_1(x+t_1)+\alpha_2(x+t_2)}\right)\\&=\Phi(\alpha_1t_1+\alpha_2t_2).\end{aligned}$$

Finally, we turn to (4.10). We have, for sufficiently large t,

$$\begin{aligned}\Phi(t)&=U(x+t,y+t,x+t,y+t)\\&=(x+t)u\left(1,\frac{y+t}{x+t},\frac{y+t}{x+t}\right)\\&=\begin{cases}\gamma(x+t)+\frac{1}{2}\gamma(x-y)-\frac{\gamma(y-x)^2}{4(x+t)} & \text{if } y<x,\\ 3\gamma(x+t)-\gamma(y+t)-\gamma(x+t)\exp\left(\frac{x-y}{2(x+t)}\right) & \text{if } y\geq x,\end{cases}\end{aligned}$$

from which we infer that $\lim_{t\to\infty}\Phi'(t)=\gamma$. It suffices to use (4.3).

The case $y<x+w-1$. We have

$$\Phi \text{ is concave on the set } [0,1-x], \tag{4.11}$$

$$\Phi'(t)\leq\gamma \text{ for } t\in(1-x,w-y), \tag{4.12}$$

$$\Phi \text{ is convex on } [w-y,\infty), \tag{4.13}$$

$$\lim_{t\to\infty}\Phi'(t)\leq\Phi'(1-x-). \tag{4.14}$$

The properties (4.11), (4.13) and (4.14) can be established in the same manner as above. To show (4.12), note that for $t\in(1-x,w-y)$,

$$\Phi'(t)=u(1,y',w')-(y'-1)u_y(1,y',w')-wu_w(1,y',w'),$$

where $y'=(y+1)/(x+t)$ and $w'=w/(x+t)$. It suffices to apply Lemma 4.1 (viii) to complete the proof. □

5. Sharpness

The purpose of this section is to show that the constant β defined in (3.5) is optimal in (1.4). One could try to construct an example, but this leads to very complicated calculations. We take a different approach and exploit Theorem 2.3. The proof is much simpler, however, it is still quite involved: for the convenience of the reader we have split it into nine lemmas.

Suppose that the inequality (1.4) holds with some constant γ. Clearly, the set of those γ's forms an interval of the form $[\beta',\infty)$. Let U^γ be the function guaranteed by Theorem 2.3: we have

$$U^\gamma(x,y,z,w)=\sup\mathbb{E}\big[g_n^*\vee w-\gamma f_n^*\vee z\big],$$

the supremum being taken over all n and all simple martingales f, g starting from x, y, respectively and such that $dg_n=\pm df_n$ for $n\geq 1$.

Lemma 5.1. *The function $F:[\beta',\infty)\to\mathbb{R}$ given by $F(\gamma)=U^\gamma(1,1,1,1)$ is convex.*

Proof. This is straightforward. Fix $\lambda\in(0,1)$ and γ_1, $\gamma_2\geq\beta'$. Let f, g be martingales starting from 1 such that $dg_n=\pm df_n$ for $n\geq 1$. Then for any n,

$$\begin{aligned}\mathbb{E}\big[g_n^*\vee w-(\lambda\gamma_1+(1-\lambda)\gamma_2)f_n^*\vee z\big] =&\lambda\mathbb{E}\big[g_n^*\vee w-\gamma_1 f_n^*\vee z\big]\\ &+(1-\lambda)\mathbb{E}\big[g_n^*\vee w-\gamma_2 f_n^*\vee z\big]\\ \leq&\lambda F(\gamma_1)+(1-\lambda)F(\gamma_2).\end{aligned}$$

It suffices to take supremum over f, g and n to complete the proof. □

Now suppose that the inequality (1.4) holds with some constant $\beta_0<\beta$. By the previous lemma, enlarging β_0 if necessary, we may assume that $U^{\beta_0}(1,1,1,1)\leq U^\beta(1,1,1,1)+1/100$. Since U^β is the least element of $\mathcal{U}(V_\beta)$ and U belongs to this class, we have $U^\beta(1,1,1,1)\leq U(1,1,1,1)=\gamma\leq -1-(1.7)^{-1}$. The latter estimate follows from the bound $w_0<1.69$, see Section 3. In consequence,

$$U^{\beta_0}(1,1,1,1)<-3/2. \tag{5.1}$$

From now on, we will work with the function U^{β_0}. It satisfies (2.3)–(2.6) (with $V=V_{\beta_0}$). There are two extra properties which follow directly from the definition. First, U^{β_0} is homogeneous: $U^{\beta_0}(\lambda x,\lambda y,\lambda z,\lambda w)=\lambda U^{\beta_0}(x,y,z,w)$ for all x, $y\in\mathbb{R}$, z, $w>0$ and $\lambda>0$. Second, U^{β_0} is symmetric in a sense that we have $U^{\beta_0}(x,y,z,w)=U^{\beta_0}(-x,y,z,w)=U^{\beta_0}(x,-y,z,w)$ for all x, $y\in\mathbb{R}$, z, $w>0$. We will use the following notation: $u_0(x,y,w)=U^{\beta_0}(x,y,1,w)$, $A^w(y)=u_0(0,y,w)$, $B^w(y)=u_0(1,y,w)$ and $\gamma_0=u_0(1,1,1)$.

Lemma 5.2. *For any $x\in[-1,1]$, y_1, $y_2\in\mathbb{R}$ and w_1, $w_2>0$ we have*

$$|u_0(x,y_1,w_1)-u_0(x,y_2,w_2)|\leq\max\{|y_1-y_2|,|w_1-w_2|\}.$$

Proof. By the triangle inequality, for any numbers a_1, a_2, …, a_n, b_1, b_2, …, b_n,

$$\begin{aligned}|y_1+a_1|\vee|y_1+a_2|\vee\ldots\vee|y_1+a_n|\vee w_1-|y_1+a_1|\vee|y_2+a_2|\vee\ldots\vee|y_2+a_n|\vee w_2\\ \leq\max\{|y_1-y_2|,|w_1-w_2|\}.\end{aligned}$$

In consequence, if f, g are simple martingales such that f starts from x, g starts from 0 and $dg_n=\pm df_n$ for $n\geq 1$, then, by the definition of u_0,

$$\begin{aligned}&\mathbb{E}((y_1+g)_n^*\vee w_1-\beta_0 f_n^*\vee 1)-u_0(x,y_2,w_2)\\ &\leq\mathbb{E}\big[((y_1+g)_n^*\vee w_1-\beta_0 f_n^*\vee 1)-((y_2+g)_n^*\vee w_1-\beta_0 f_n^*\vee 1)\big]\\ &\leq\max\{|y_1-y_2|,|w_1-w_2|\}.\end{aligned}$$

It suffices to take supremum over f, g and n to obtain

$$u_0(x,y_1,w_1)-u_0(x,y_2,w_2)\leq\max\{|y_1-y_2|,|w_1-w_2|\},$$

and the claim follows by symmetry. □

Lemma 5.3. *For any $w>0$, $|y|\leq w$ and $\delta\in(0,1)$,*

$$B^w(y)\geq u_0(1-\delta,y+\delta,w)+\delta\gamma_0. \tag{5.2}$$

Proof. Apply (2.6) to $(x,y,z,w):=(1,y,1,w)$, $\varepsilon=-1$ and $t_1=-\delta$, $t_2>0$ (the number α is uniquely determined by t_1 and t_2: $\alpha=t_2/(t_2+\delta)$). We obtain

$$\frac{t_2}{t_2+\delta}U^{\beta_0}(1-\delta,y+\delta,1,w)+\frac{\delta}{t_2+\delta}U^{\beta_0}(1+t_2,y-t_2,1,w)\leq U^{\beta_0}(1,y,1,w).$$

Using (2.3) and the homogeneity of U^{β_0}, we have

$$\begin{aligned}U^{\beta_0}(1+t_2,y-t_2,1,w)&=U^{\beta_0}(1+t_2,y-t_2,1+t_2,w)\\&=(1+t_2)U^{\beta_0}\left(1,\frac{y-t_2}{1+t_2},1,\frac{w}{1+t_2}\right),\end{aligned}$$

so we can rewrite the above estimate in the form

$$\frac{t_2}{t_2+\delta}u_0(1-\delta,y+\delta,w)+\frac{\delta(1+t_2)}{t_2+\delta}u_0\left(1,\frac{y-t_2}{1+t_2},\frac{w}{1+t_2}\right)\leq B^w(y).$$

By (2.3), we have $u_0\left(1,\frac{y-t_2}{1+t_2},\frac{w}{1+t_2}\right)=u_0\left(1,\frac{y-t_2}{1+t_2},\left|\frac{y-t_2}{1+t_2}\right|\right)$ for sufficiently large t_2. Therefore, letting $t_2\to\infty$ and using the previous lemma together with $u_0(1,-1,1)=u_0(1,1,1)=\gamma_0$, we get (5.2). □

Lemma 5.4. *For any $w>y\geq 1$ and $\delta\in(0,1)$ satisfying $\delta\leq(w-y)/2$,*

$$B^w(y)\geq\delta A^w(y+2\delta-1)+(1-\delta)B^w(y+2\delta)+\delta\gamma_0, \tag{5.3}$$

$$\begin{aligned}A^w(y+2\delta-1)&\geq\frac{\delta}{1+\delta}B^w(y+2\delta)+\frac{1}{1+\delta}u_0(-\delta,y+\delta-1,w)\\&\geq\frac{\delta}{1+\delta}B^w(y+2\delta)+\frac{\delta}{1+\delta}B^w(y)+\frac{1-\delta}{1+\delta}A^w(y-1)\end{aligned} \tag{5.4}$$

and

$$(1-\delta)(B^w(y)-A^w(y-1)+\gamma_0)\geq B^w(y+2\delta)-A^w(y-1+2\delta)+\gamma_0. \tag{5.5}$$

Proof. Applying (2.6) to $(x,y,z,w):=(1-\delta,y+\delta,1,w)$, $\varepsilon=1$ and $t_1=\delta-1$, $t_2=\delta$, we obtain

$$\delta A^w(y+2\delta-1)+(1-\delta)B^w(y+2\delta)\leq u_0(1-\delta,y+\delta,w).$$

Combining this with (5.2) yields (5.3). Similarly, if we apply (2.6) twice, first to $(x,y,z,w):=(0,y+2\delta-1,1,w)$, $\varepsilon=1$, $t_1=-1$, $t_2=\delta$, then to $(x,y,z,w):=(-\delta,y+\delta-1,1,w)$, $\varepsilon=-1$, $t_1=\delta-1$ and $t_2=\delta$, we get (5.4). To obtain (5.5), multiply both sides of (5.4) by $1+\delta$ and add this to (5.3). □

Lemma 5.5. *For any $w>1$,*

$$B^w(w)\leq\gamma_0\left[3-w-\exp\left(\frac{1-w}{2}\right)\right]. \tag{5.6}$$

Proof. We start with the observation that $A^w(0)\geq B^w(1)$, which follows from (2.6) applied to $(x,y,z,w):=(0,0,1,w)$, $\varepsilon=1$ and $t_1=-1$, $t_2=1$. Thus, using (5.5) and induction,

$$(1-\delta)^N\gamma_0\geq(1-\delta)^N(B^w(1)-A^w(0)+\gamma_0)\geq B^w(1+2N\delta)-A^w(2N\delta)+\gamma_0.$$

Hence, if we put $\delta=(w-1)/(2N)$ and let $N\to\infty$, we arrive at

$$A^w(w-1)\geq B^w(w)+\gamma_0(1-\exp((1-w)/2)). \tag{5.7}$$

Now repeat the arguments leading to (5.3), with y replaced by w, to get

$$B^w(w)\geq(1-\delta)B^{w+2\delta}(w+2\delta)+\delta A^{w+\delta}(w+2\delta-1)+\delta\gamma_0,$$

so, by Lemma 5.2,

$$B^w(w) \geq (1-\delta)B^{w+2\delta}(w+2\delta) + \delta A^{w+2\delta}(w+2\delta-1) + \delta\gamma_0 - \delta^2. \tag{5.8}$$

Applying (5.7) yields

$$B^w(w) \geq B^{w+2\delta}(w+2\delta) + \delta\gamma_0(2-\exp((1-w-2\delta)/2) - \delta^2.$$

Using induction as in the proof of (5.7), this leads to

$$\gamma_0 = B^1(1) \geq B^w(w) + \gamma_0[w-2+\exp((1-w)/2)],$$

which is the claim. □

Lemma 5.6. *Suppose that $w \in (1,2)$ and let $1 \leq y \leq w$. Then*

$$A^w(y-1)y^2 \geq B^w(y)(y^2-2y+2) + 2(y-1)(w-\beta_0). \tag{5.9}$$

Proof. We apply (2.6) three times:

$$A^w(y-1) \geq \frac{y}{y+2}B^w(y) + \frac{2}{y+2}u_0(-y/2, y/2-1, w),$$
$$u_0(-y/2, y/2-1, w) \geq \frac{2-y}{y}u_0(-1+y/2, -y/2, w) + \frac{2y-2}{y}u_0(-1,0,w),$$
$$u_0(-1+y/2, -y/2, w) = u_0(1-y/2, y/2, w) \geq \frac{y}{2}A^w(y-1) + \frac{2-y}{2}B^w(y)$$

and combine these estimates with $u_0(-1,0,w) \geq w-\beta_0$, a consequence of (2.5), thus obtaining (5.9). □

Lemma 5.7. *Suppose that $w \in (1,2)$ and let $1 \leq y \leq w$. Then*

$$\begin{aligned}(A^w(w-1) - w + \beta_0)&\left[2\exp\left(\frac{w-y}{2}\right) - 1\right] \\ &\geq (B^w(y) - w + \beta_0)(1 - 2y^{-1} + 2y^{-2}) + (y-w)\gamma_0 \\ &\quad + 2(B^w(w) - w + \beta_0 + \gamma_0)\left[\exp\left(\frac{w-y}{2}\right) - 1\right]\end{aligned} \tag{5.10}$$

and

$$\begin{aligned}(B^w(y) - w + \beta_0)&\left[1 + \frac{2(y-1)}{y^2}\left(1 - \exp\left(\frac{y-w}{2}\right)\right)\right] \\ &\geq B^w(w) - w + \beta_0 + \gamma_0\left[w - y - 1 + \exp\left(\frac{y-w}{2}\right)\right].\end{aligned} \tag{5.11}$$

In addition,

(5.12)

$$\begin{aligned}B^{w-2\delta}(w-2\delta) \geq &B^w(w) + 2\delta\gamma_0 - \frac{\delta\exp(\frac{y-w}{2})}{1+2y^{-2}(y-1)(1-\exp(\frac{y-w}{2}))} \times \\ &\times \left\{2y^{-2}(y-1)(B^w(w) - w + \beta_0 + \gamma(w-y)) + \gamma_0\right\} - \delta^2.\end{aligned}$$

Proof. Using the inequality (5.5) inductively, as in the proof of (5.7), yields the following estimate: for $1 \leq y'' \leq y' \leq w$,

$$\exp\left(\frac{y''-y'}{2}\right)(B^w(y'') - A^w(y''-1) + \gamma_0) \geq B^w(y') - A^w(y'-1) + \gamma_0. \tag{5.13}$$

Take $y' = w$ and note that $B^w(y''+2\delta) \geq B^w(y'') - 2\delta$, a consequence of Lemma 5.2. Plug these two estimates into (5.4) to get

$$\begin{aligned}A^w(y''+2\delta-1) \geq & A^w(y''-1)\\ &+ \frac{2\delta}{1+\delta}\left[\exp\left(\frac{w-y''}{2}\right)(B^w(w)-A^w(w-1)+\gamma_0)-\gamma_0-\delta\right].\end{aligned}$$

Now set $\delta = (w-y)/(2N)$, write the above estimates for $y''=y$, $y''=y+2\delta$, $\ldots$, $y''=y+(2N-2)\delta$ and sum them up. We obtain

$$\begin{aligned}A^w(w-1) = A^w(y+2N\delta-1) &\geq A^w(y-1) - \frac{2\delta}{1+\delta}N(\gamma_0+\delta)\\ &+ \frac{2\delta}{1+\delta}(B^w(w)-A^w(w-1)+\gamma_0)\exp\left(\frac{w-y}{2}\right)\frac{1-\exp(-N\delta)}{1-\exp(-\delta)}.\end{aligned}$$

Letting $N\to\infty$ gives

$$A^w(w-1) \geq A^w(y-1) + 2(B^w(w)-A^w(w-1)+\gamma_0)\left[\exp\left(\frac{w-y}{2}\right)-1\right] + \gamma_0(y-w)$$

and combining this with (5.9) yields the first estimate. We skip the proof of (5.11), it can be established using similar argumentation. To get (5.12), plug (5.11) into (5.10) to obtain

$$\begin{aligned}A^w(w-1) \geq & B^w(w)+\gamma_0 - \frac{\exp(\frac{y-w}{2})}{1+2y^{-2}(y-1)(1-\exp(\frac{y-w}{2}))}\times\\ &\times\left\{2y^{-2}(y-1)(B^w(w)-w+\beta_0+\gamma(w-y))+\gamma_0\right\}.\end{aligned}$$

It suffices to make use of (5.8) (with w replaced by $w-2\delta$) to complete the proof. $\square$

The final estimate we will need is the following. It can be established essentially in the same manner as above; we omit the tedious and lengthy calculations.

Lemma 5.8. *For any $w \geq 1$,*

$$B^w(w) \geq \frac{w^2(1+\gamma_0)}{2} + 2w - \beta_0. \tag{5.14}$$

Now we are ready to complete the proof.

Sharpness of (1.4). The first observation is that $\gamma_0 \in (-2, -3/2)$. The inequality $\gamma_0 < -3/2$ is precisely (5.1). To get the lower bound, apply (5.14) to $w=1$ to obtain $\gamma_0 \geq 5-2\beta_0 > -2$ (we have $\beta_0 < \beta < 3.5$). Now let $v_0 = -(1+\gamma_0)^{-1} \in (1,2)$, define Y^{v_0} as in Section 3 and let $\overline{y}_0$ be the inverse to Y^{v_0}. Finally, let $v_1 = Y^{v_0}(1)$ and

$$C(w) = -\frac{2\gamma_0}{1+\overline{y}_0}\exp\left(\frac{\overline{y}_0-w}{2}\right)+\gamma_0(1-w)+\frac{\gamma_0(\overline{y}_0^2+2)}{2(1+\overline{y}_0)}+w-\beta_0$$

for $w\in[v_1,v_0]$. Observe that since $\overline{y}_0(v_0)=v_0$, we have, after some manipulations,

$$C(v_0) = \frac{3}{2}v_0 - \beta_0 = \frac{v_0^2(1+\gamma_0)}{2} + 2v_0 - \beta_0 \leq B^{v_0}(v_0), \tag{5.15}$$

by virtue of the previous lemma. Furthermore, it can be verified that C satisfies the differential equation

$$\begin{aligned}C'(w) = &-\gamma_0 + \frac{\exp(\frac{\overline{y}_0-w}{2})}{2\big[1+2\overline{y}_0^{-2}(\overline{y}_0-1)(1-\exp(\frac{\overline{y}_0-w}{2}))\big]}\times\\ &\times\big\{2\overline{y}_0^{-2}(\overline{y}_0-1)(C(w)-w+\beta_0+\gamma(w-\overline{y}_0))+\gamma_0\big\},\end{aligned}$$

for $w\in(v_1,v_0)$. Note that C'' is bounded on (v_1,v_0). To see this, observe that the solution Y^{v_0} to (3.1) can be extended to an increasing C^∞ function on a certain open interval I containing $[1,v_0]$. Consequently, y_0, y_0', y_0'' are bounded on (v_1,v_0) and hence C'' also has this property. Therefore for some absolute constant r,

$$\begin{aligned}C(w-2\delta)\leq C(w)+2\delta\gamma_0 &- \frac{\delta\exp(\frac{\overline{y}_0-w}{2})}{1+2\overline{y}_0^{-2}(\overline{y}_0-1)(1-\exp(\frac{\overline{y}_0-w}{2}))}\times\\ &\times\big\{2\overline{y}_0^{-2}(\overline{y}_0-1)(C(w)-w+\beta_0+\gamma(w-\overline{y}_0))+\gamma_0\big\}+r\delta^2.\end{aligned}\tag{5.16}$$

Combining this with (5.12), applied to $y=\overline{y}_0$ (which is allowed, since $\overline{y}_0\in(1,2)$), yields

$$B^{w-2\delta}(w-2\delta)-C(w-2\delta)\geq(B^w(w)-C^w(w))\cdot R(\delta,w)-(r+1)\delta^2,$$

where $R(\delta,w)$ is a certain constant lying in $[0,1]$. By induction and (5.15), we obtain $B(v_1)-C(v_1)\geq 0$, which implies, by (5.6), that

$$\gamma_0\left[3-v_1-\exp\left(\frac{1-v_1}{2}\right)\right]\geq C(v_1).$$

This is equivalent to

$$\beta_0\geq v_1+\frac{5}{4}\left(1+\frac{1}{v_0}\right)=Y^{v_0}(1)+\frac{5}{4}\left(1+\frac{1}{v_0}\right)$$

and gives $\beta_0\geq\beta$, by virtue of (3.5). This contradicts the assumption $\beta_0<\beta$ and completes the proof. □

6. The proof of Lemma 4.1

Proof of (i). This is straightforward. It is clear that u is of class C^1 in the interiors of D_1–D_9. To show the smoothness, one only needs to verify if the partial derivatives match at the common boundaries of D_1, D_2, D_3 (this corresponds to the set $\{(x,y,w):w<w_1\}$), D_4, D_5, D_6 (the set $\{(x,y,w):w\in(w_1,w_0)\}$) and, finally, D_7, D_8, D_9 (the set $\{(x,y,w):w>w_1\}$). We omit the tedious calculations. □

Proof of (ii). We will only show (4.2); the arguments leading to the second inequality are similar. By (4.1), it suffices to establish the estimate for $y\geq 0$. Let

$$L_-=\lim_{\delta\downarrow 0}\frac{u(1,y,w)-u(1-\delta,y-\delta,w)}{\delta},\quad L_+=\lim_{\delta\downarrow 0}\frac{u(1,y,w)-u(1-\delta,y+\delta,w)}{\delta}.$$

We have that

$$L_- = \begin{cases} -\frac{\gamma(y-1)}{2} & \text{if } y \leq 1,\ w \leq w_1, \\ -\gamma + \gamma\exp(\frac{1-y}{2}) & \text{if } 1 < y \leq w \leq w_1, \\ -\frac{\gamma(y-1)}{1+y_0(w)} & \text{if } y \leq y_0,\ w_1 < w < w_0, \\ \frac{2\gamma}{1+y_0(w)}\exp(\frac{y_0(w)-y}{2}) - \gamma & \text{if } w \geq y > y_0,\ w_1 < w < w_0, \\ -\frac{1}{w_0}\exp(y+w_0-w) & \text{if } y \leq w - w_0,\ w \geq w_0, \\ \frac{y-w+w_0-1}{w_0} & \text{if } y > w - w_0,\ w \geq w_0 \end{cases}$$

and it is easy to verify that none of these expressions is smaller than γ (to check the last two, recall that $\gamma = -1 - 1/w_0$). Similarly,

$$L_+ = \begin{cases} \frac{\gamma(y+1)}{2} & \text{for } y \leq 1,\ w \leq w_1, \\ \gamma & \text{for } 1 < y \leq w \leq w_1, \\ -\frac{\gamma(-y-1)}{1+y_0(w)} & \text{if } y \leq y_0,\ w_1 < w < w_0, \\ \gamma & \text{if } w \geq y > y_0,\ w_1 < w < w_0, \\ -\frac{1}{w_0}\exp(y+w_0-w) & \text{if } y \leq w - w_0,\ w \geq w_0, \\ \frac{-y+w-w_0-1}{w_0} & \text{if } w \geq y > w - w_0,\ w \geq w_0, \end{cases}$$

and it is not difficult to check that $L_+ \geq \gamma$. □

Proof of (iii). If $w < w_1$, we have $u_w(x, w, w) = 0$. Suppose that $w \in (w_1, w_0)$. First note that $u_w(1, w, w) = 0$; this is equivalent to (3.8). For $|x| < 1$, a little computation shows that

$$u_w(x, w, w) = u_w(1, w, w) + \frac{\gamma(x^2-1)}{4(1+y_0(w))^2}\exp\left(\frac{y_0(w)-w}{2}\right)(y_0'(y_0-1) - y_0 - 1),$$

and the latter term is nonpositive due to (3.11). Finally, if $w \geq w_0$, we have the equality again: $u_w(x, w, w) = 0$. □

Proof of (iv). There are no points of the form (x, t, t) in the sets D_4, D_5, D_7 and D_8. A little computation shows that

$$H_x(t) = \begin{cases} \gamma x/2 & \text{if } (|x+t|, t, t) \in D_1, \\ -\gamma + \gamma\exp(x/2) & \text{if } (x+t, t, t) \in D_2, \\ -\gamma + \gamma(x+t+1)\exp(-x/2-t) & \text{if } (-x-t, t, t) \in D_2, \\ -\gamma + \gamma\exp(\frac{1-w_1}{2})(\frac{1+w_1+x}{2}) & \text{if } (|x+t|, t, t) \in D_3, \\ -\gamma + \frac{2\gamma}{1+y_0(t)}\exp(\frac{y_0(t)-t}{2})(\frac{1+x+t}{2}) & \text{if } (|x+t|, t, t) \in D_6, \\ 1 - (x+t)/w_0 & \text{if } (|x+t|, t, t) \in D_9. \end{cases}$$

Now it can be easily verified that H_x is continuous. Furthermore, all the expressions above define nonincreasing functions of t. The only nontrivial case is for $(|x+t|, t, t) \in D_6$; then we have, by (3.11),

$$\begin{aligned} H_x'(t) &= a(t)\left[(1+x+t)(y_0'(1-y_0) + 1 + y_0) - 2(1+y_0)\right] \\ &\leq a(t)\left[2(y_0'(1-y_0) + 1 + y_0) - 2(1+y_0)\right] = 2a(t)y_0'(1-y_0) \leq 0, \end{aligned}$$

where $a(t) = -\frac{\gamma}{2(1+y_0)^2}\exp(\frac{y_0-t}{2}) > 0$. □

Proof of (v). A direct calculation shows that

$$J(y) = \begin{cases} -\frac{\gamma}{4}y^2 + \frac{5}{4} & \text{for } y \le 1, \\ 3\gamma - \gamma y - \gamma \exp(\frac{1-y}{2}) & \text{for } 1 < y \le w_1, \\ -\frac{2\gamma}{1+y_0}\exp(\frac{y_0-y}{2}) + \gamma(1-y) + \frac{\gamma(y_0^2+2)}{2(1+y_0)} + y - \beta & \text{for } w_1 < y < w_0, \\ w_0/2 + y - \beta & \text{for } y \ge w_0 \end{cases}$$

and

$$J'(y) = \begin{cases} -\frac{\gamma y}{2} & \text{for } y < 1, \\ -\gamma + \frac{\gamma}{2}\exp(\frac{1-y}{2}) & \text{for } 1 < y < w_1, \\ \frac{\gamma}{1+y_0}\exp(\frac{y_0-y}{2}) - \gamma & \text{for } w_1 < y < w_0, \\ 1 & \text{for } y > w_0. \end{cases}$$

From the formulas above we infer that J is of class C^1, so we will be done if we show that J'' is nonnegative on $(0,1)\cup(1,w_1)\cup(w_1,w_0)\cup(w_0,\infty)$. This is trivial on $(0,1)$, $(1,w_1)$ and (w_0,∞). For $y\in(w_1,w_0)$ we have

$$J''(y) = -\frac{\gamma}{2(1+y_0)^2}\exp\left(\frac{y_0-y}{2}\right)(y_0'(1-y_0)+1+y_0) \ge 0,$$

by virtue of (3.11). □

Proof of (vi). By part (i) and the symmetry condition (4.1), it suffices to check the concavity in $[0,1]\times[0,w]$. This clearly will follow when we show that $u_{xx} + 2|u_{xy}| + u_{yy} \le 0$ in the interiors of D_1–D_9. One can check that for $\varepsilon\in\{-1,1\}$, $u_{xx}(x,y,w) + 2\varepsilon u_{xy}(x,y,w) + u_{yy}(x,y,w)$ equals

$$\begin{cases} 0 & \text{on } D_1^o\cup D_3^o\cup D_4^o\cup D_6^o\cup D_7^o\cup D_9^o, \\ -\gamma x\exp\left(\frac{x-y}{2}\right)\left(\frac{\varepsilon-1}{2}\right) & \text{on } D_2^o, \\ -\frac{2\gamma x}{1+y_0}\exp\left(\frac{y_0-y-1+x}{2}\right)\left(\frac{\varepsilon-1}{2}\right) & \text{on } D_5^o \end{cases}$$

and we are done, since all the expressions above are nonpositive. □

Proof of (vii). We will show that $u_y(1,y,w)\ge 0$ for $y\in(0,w)$. The inequality is evident for $(1,y,w)\in D_1\cup D_2\cup D_4\cup D_8\cup D_9$; in addition, no points of the form $(1,y,w)$ lie in $D_3\cup D_6\cup D_7$. Thus it remains to check the estimate for $(1,y,w)\in D_5$. We have

$$u_y(1,y,w) = \frac{\gamma}{1+y_0}\exp\left(\frac{y_0-y}{2}\right) - \gamma \ge \frac{\gamma}{1+y_0} - \gamma,$$

which is positive. □

Proof of (viii). Suppose that $y<0$. Then, by (4.1), we may write

$$\begin{aligned} &u(1,y,w) - (y-1)u_y(1,y,w) - wu_w(1,y,w) \\ &\qquad = u(1,-y,w) + (y-1)u_y(1,-y,w) - wu_w(1,y,w) \\ &\qquad = u(1,-y,w) - (-y-1)u_y(1,-y,w) - wu_w(1,-y,w) - 2u_y(1,-y,w), \end{aligned}$$

which, by the previous part, does not exceed

$$u(1,-y,w) - (-y-1)u_y(1,-y,w) - wu_w(1,-y,w).$$

The latter is the left hand side of (4.4) for $-y$; thus it suffices to show the estimate for nonnegative y. Suppose first that $w \le w_1$. If $y \le 1$, we have

$$u(1, y, w) - (y - 1)u_y(1, y, w) - wu_w(1, y, w) = \gamma + \frac{\gamma}{4}(y - 1)^2 \le \gamma.$$

If $y \in (1, w_1]$, then

$$u(1, y, w) - (y - 1)u_y(1, y, w) - wu_w(1, y, w) = \gamma + \gamma\left[1 - \frac{y+1}{2}\exp\left(\frac{1-y}{2}\right)\right]$$

and the expression in the square brackets is nonnegative (due to the estimate $e^z \ge 1 + z$ applied to $z = (y - 1)/2$). Now let $w \in (w_1, w_0)$. The function $\psi : y \mapsto u(1, y, w) - (y - 1)u_y(1, y, w) - wu_w(1, y, w)$ is continuous; moreover, for $y \ne y_0$, we have $\psi'(y) = -(y - 1)u_{yy}(1, y, w) - u_{wy}(1, y, w)$, or, in an explicit form,

$$\psi'(y) = \begin{cases} -\gamma(1 + y_0)^{-2}[-(y - 1)(y_0 + 1) + wyy_0'] & \text{for } y < y_0, \\ \frac{\gamma}{2(1+y_0)^2}\exp(\frac{y_0-y}{2}[y_0'(y_0 - 1) - (y_0 + 1)] & \text{for } y > y_0. \end{cases}$$

Note that $\psi' \ge 0$; indeed, this is an immediate consequence of (3.10) and (3.11). Thus it suffices to show (4.4) for $y = w$. As $u_w(1, w, w) = 0$, the inequality takes the form

$$u(1, w, w) - (w - 1)u_y(1, w, w) \le \gamma.$$

Differentiating the left-hand side, one gets

$$-\frac{\gamma(w-1)}{(1+y_0)^2}\exp\left(\frac{y_0(w) - w}{2}\right)(1 + y_0 - y_0'(y_0 - 1)),$$

which is negative due to (3.11). Thus

$$u(1, w, w) - (w - 1)u_y(1, w, w) \le u(1, w_1, w_1) - (w_1 - 1)u_y(1, w_1, w_1) \le \gamma,$$

as we have already checked above. Finally, assume that $w \ge w_0$. Then, for $y \le w - w_0$, the inequality (4.4) takes the form $-\beta \le \gamma$, which is obviously true. If $y \ge w - w_0$, then

$$\begin{aligned} u(1, y, w) - (y - 1)u_y(1, y, w) - wu_w(1, y, w) &= \frac{y - w}{w_0} - \frac{(y - w)^2}{2w_0} + \frac{w_0}{2} + 1 - \beta \\ &\le \frac{w_0}{2} + 1 - \beta \end{aligned}$$

and the latter does not exceed γ: substituting $\gamma = -1 - w_0^{-1}$, we see that this is equivalent to (3.7) applied to $w = w_0$. □

Proof of (ix). By (4.1) and part (vi), it suffices to check the majorization on the sets $\{(x, y, w) : x = 1, y \ge 0\}$ and $\{(x, y, w) : x \ge 0, y = w\}$. Now, by part (vii), $u(1, y, w) \ge u(1, 0, w)$ and, as one easily verifies, the right-hand side is not smaller than $w - \beta$. The proof is completed by the observation that for any fixed $w > 0$, the function $x \mapsto u(x, w, w)$ is nonincreasing on $[0, 1]$. □

Proof of (x). This is trivial: for x, y, w as in the statement, we have $(x, y, w) \in D_1$, so $u(x, y, w) = \gamma < 0$. □

7. Concluding remarks and related open problems

The final section of the paper is devoted to various modifications and extensions of Burkholder's method. We will also formulate here a number of open problems which may be investigated using this technique. Throughout this section, unless stated otherwise, we assume that f is a martingale and g is its transform by a certain predictable sequence bounded in absolute value by 1.

For the sake of clarity, we have split the problems into several groups.

7.1. **Doob's bounds.** As observed by Burkholder in [4], the classical Doob's inequality

$$||f^*||_p \leq \frac{p}{p-1}||f||_p, \qquad 1 < p < \infty, \tag{7.1}$$

can also be established using the above approach. Due to the fact that the transform g does not appear in this estimate, the corresponding functions V and U will not depend on y and w. Namely, for a fixed $\gamma > 0$, let $V_p^\gamma : \mathbb{R} \times (0, \infty) \to \mathbb{R}$ be given by the formula

$$V_p^\gamma(x, z) = (|x| \vee z)^p - \gamma^p |x|^p$$

and let $\mathcal{U}(V_p^\gamma)$ consist of those U satisfying (2.3)–(2.6), which depend only on x and z. It can be shown that this class is nonempty if and only if $\gamma \geq p/(p-1)$, and that $U_p : \mathbb{R} \times (0, \infty) \to \mathbb{R}$, given by

$$U_p(x, z) = p(|x| \vee z)^{p-1} \left((|x| \vee z) - \frac{p}{p-1}|x| \right),$$

belongs to $\mathcal{U}(V_p^{p/(p-1)})$. For details, see pages 14–15 in [4].

For $0 < p \leq 1$ Doob's inequality does not hold with any finite constant, but we have the following fact, due to Shao [14].

Theorem 7.1. *Suppose that f is a nonnegative martingale. Then for $0 < p < 1$,*

$$||f^*||_p \leq (1 - p)^{-1/p} ||f||_p. \tag{7.2}$$

The inequality is sharp.

Proof. The original argument of Shao did not exploit Burkholder's technique; however, the method can be successfully implemented and we take the opportunity to present it here. Introduce the functions

$$V_p(x, z) = (x \vee z)^p, \quad U_p(x, z) = (x \vee z)^{p-1} \left((x \vee z) + \frac{p}{1-p} x \right),$$

given on $[0, \infty) \times (0, \infty)$. It is easy to verify that V_p and U_p satisfy the appropriate modifications of (2.3), (2.5) and (2.6): one has to restrict oneself to nonnegative x there, and to t_1, $t_2 \geq -x$ in (2.6). On the other hand, (2.4) is *not* satisfied. Nonetheless, arguing as in the proof of Theorem 2.2, we get that $(U_p(f_n, f_n^*))_{n \geq 0}$ is a supermartingale and hence

$$||f_n^*||_p^p = \mathbb{E}V_p(f_n, f_n^*) \leq \mathbb{E}U_p(f_n, f_n^*) \leq \mathbb{E}U_p(f_0, f_0^*) = (1-p)^{-1}\mathbb{E}f_0^p = (1-p)^{-1}||f||_p^p.$$

It suffices to let $n \to \infty$ to get (7.2). To see that the constant $(1-p)^{-1/p}$ is the best possible, we apply Theorem 2.3. Let $U_p^0 : [0, \infty) \times (0, \infty) \to \mathbb{R}$ be given by

$$U_p^0(x, z) = \sup \left\{ \mathbb{E}V_p(f_n, f_n^* \vee z) \right\},$$

where the supremum is taken over all n and simple nonnegative martingales f starting from x. Then V_p, U_p^0 satisfy (2.3), (2.5) and (2.6). Applying the latter condition, with $x = z = 1$, $t_1 = -1$, $t_2 = \delta$ (and any ε, y, w), we obtain

$$U_p^0(1,1) \geq \frac{\delta}{1+\delta} U_p^0(0,1) + \frac{1}{1+\delta} U_p^0(1+\delta, 1).$$

Now use (2.3), (2.5) and the fact that U_p^0 is homogeneous of order p to get

$$U_p^0(1,1) \geq \frac{\delta}{1+\delta} + (1+\delta)^{p-1} U_p^0(1,1),$$

or

$$U_p^0(1,1) \geq \frac{\delta}{(1+\delta)(1-(1+\delta)^{p-1})}.$$

Letting $\delta \to 0$ yields $U_p^0(1,1) \geq (1-p)^{-1}$. This is the claim, by the very definition of U_p^0. □

7.2. **Maximal L^p-estimates for martingale transforms.** A natural extension of the problem studied in the present paper is the following. Let $1 < p < \infty$. What are the best values of the constants κ_p, κ_p^* in the estimates

$$||g^*||_p \leq \kappa_p ||f||_p, \tag{7.3}$$

$$||g^*||_p \leq \kappa_p^* ||f^*||_p\,? \tag{7.4}$$

We can throw some light on this problem, using Doob's bound (7.1) and Burkholder's famous sharp inequality (cf. [2]):

$$||g||_p \leq (p^* - 1)||f||_p,$$

where $p^* = \max\{p, p/(p-1)\}$. As the result, we get

$$||g^*||_p \leq \frac{p}{(p-1)^2}||f||_p, \qquad ||g^*||_p \leq \frac{p}{(p-1)^2}||f^*||_p \tag{7.5}$$

for $1 < p < 2$, and

$$||g^*||_p \leq p\,||f||_p, \qquad ||g^*||_p \leq p\,||f^*||_p \tag{7.6}$$

for $p \geq 2$. Quite surprisingly, both inequalities in (7.6) are sharp: this can be seen by a careful study of the examples invented by Burkholder in [2] (see page 669–670 there). On the other hand, neither of the inequalities in (7.5) is sharp. The constant $p/(p-1)^2$ is not even of optimal order as $p \to 1$ (which is known to be $O((p-1)^{-1})$ in both the estimates). The question about the optimal values of κ_p and κ_p^* in this case is open, to the best of author's knowledge.

7.3. **Logarithmic estimates for martingale transforms.** A related and very interesting problem is to study the inequality

$$||g^*||_1 \leq K \sup_n \mathbb{E}|f_n| \log^+ |f_n| + L(K). \tag{7.7}$$

There are two questions to be answered:

(i) For which $K > 0$ there is a universal $L(K) < \infty$ such that the above holds?

(ii) What is the optimal value of $L(K)$?

Both these questions seem to be open so far. Let us mention here that a similar problem, with g^* replaced by the one-sided maximal function $\sup_{n\geq 0} g_n$, was solved by the author in [12]. Since the full answer is quite complicated, we will not present it here and refer the interested reader to that paper. See also [10] for a sharp non-maximal version of (7.7).

7.4. **Maximal inequalities for the martingale square function.** Given a martingale f, we define its square function $S(f)$ by

$$S(f) = \left(\sum_{k=0}^{\infty} |df_k|^2 \right)^{1/2}.$$

The classical problem of comparing the sizes of f, f^* and $S(f)$ plays an important role in many areas of mathematics and has been studied by many authors. As shown by Burkholder in [3], for $1 < p < \infty$ we have

$$(p^* - 1)^{-1}||S(f)||_p \leq ||f||_p \leq (p^* - 1)||S(f)||_p. \tag{7.8}$$

Furthermore, the left inequality is sharp for $1 < p \leq 2$, and the right is sharp for $p \geq 2$; in the remaining cases the optimal constants are not known. A related problem is to study the maximal estimates

$$c_p||S(f)||_p \leq ||f^*||_p \leq C_p||S(f)||_p, \qquad 1 \leq p < \infty. \tag{7.9}$$

Combining (7.8) with Doob's inequality (7.1) we get that the above estimate holds with $c_p = (p^*-1)^{-1}$ and $C_p = p(p^*-1)/(p-1)$. Surprisingly, for $p \geq 2$ the constant $C_p = p$ is the best possible (cf. [13] and page 19 in [4]); this should be compared to similar phenomenon in (7.6). An important and interesting case in (7.9) corresponds to the choice $p = 1$. Using a clever decomposition of a martingale, Davis [7] proved that the estimate holds with some finite universal c_1 and C_1. Later, Burkholder [6] invented a method, which is a modification of the one presented in Section 2 and allows to obtain sharp estimates involving f, f^* and $S(f)$. Burkholder used it to show that $c_1 = 1/\sqrt{3}$ is the best, and obtained some tight bounds for c_p when $1 < p \leq 2$. For details, we refer the reader to that paper. However, the optimal values of c_p for $1 < p < \infty$, $p \neq 2$ and C_p for $1 \leq p < 2$ seem to be unknown so far.

Acknowledgement

The author would like to thank the Referee for the careful reading of the first version of the paper and many remarks and suggestions which greatly improved the exposition.

References

[1] K. Bichteler, *Stochastic integration and L^p-theory of semimartingales*, Ann. Probab. **9** (1981), 49–89.

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

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

[4] D. L. Burkholder, *Explorations in martingale theory and its applications*, Ecole d'Eté de Probabilités de Saint Flour XIX-1989, Lecture Notes in Mathematics **1464** (1991), 1–66.

[5] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, Proceedings of the Norbert Wiener Centenary Congress, 1994 (edited by V. Mandrekar and P. R. Masani), Proceedings of Symposia in Applied Mathematics **52** (1997), 343–358.

[6] D. L. Burkholder, *The best constant in the Davis inequality for the expectation of the martingale square function*, Trans. Amer. Math. Soc. **354** (2002), pp. 91–105.

[7] B. J. Davis, *On the integrability of the martingale square function*, Israel J. Math. **8** (1970), 187–190.
[8] A. Osękowski, *Sharp maximal inequality for stochastic integrals*, Proc. Amer. Math. Soc. **136** (2008), 2951–2958.
[9] A. Osękowski, *Sharp maximal inequality for martingales and stochastic integrals*, Electr. Comm. in Probab. **14** (2009), 17–30.
[10] A. Osękowski, *Sharp LlogL inequalities for differentially subordinated martingales*, Illinois J. Math., **52** Vol. 3 (2008), 745–756.
[11] A. Osękowski, *Sharp weak type inequalities for differentially subordinated martingales*, Bernoulli **15** Vol. 3 (2009), 871–897.
[12] A. Osękowski, *Logarithmic estimates for submartingales and their differential subordinates*, J. Theoret. Probab., to appear.
[13] A. O. Pittenger, *Note on a square function inequality*, Ann. Probab. **7** (1979), 907–908.
[14] Q. Shao, *A comparison theorem on maximal inequalities between negatively associated and independent random variables*, J. Theoret. Probab. **13** (2000), 343–356.
[15] Y. Suh, *A sharp weak type* (p, p) *inequality* $(p > 2)$ *for martingale transforms and other subordinate martingales*, Trans. Amer. Math. Soc. **357** (2005), no. 4, 1545–1564 (electronic).

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