# MAXIMAL WEAK-TYPE INEQUALITY FOR STOCHASTIC INTEGRALS

ADAM OSĘKOWSKI

Abstract. Assume that $X$ is a real-valued martingale starting from 0, $H$ is a predictable process with values in $[-1,1]$ and $Y$ is the stochastic integral of $H$ with respect to $X$. The paper contains the proofs of the following sharp weak-type estimates.

(i) If $X$ has continuous paths, then

$$\mathbb{P}\left(\sup_{t\geq 0}|Y_t|\geq 1\right)\leq 2\mathbb{E}\sup_{t\geq 0}X_t.$$

(ii) If $X$ is arbitrary, then

$$\mathbb{P}\left(\sup_{t\geq 0}|Y_t|\geq 1\right)\leq 3.477977\ldots\mathbb{E}\sup_{t\geq 0}X_t.$$

The proofs rest on Burkholder's method and exploits the existence of certain special functions possessing appropriate concavity and majorization properties.

## 1. Introduction

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

$$Y_t=H_0X_0+\int_{(0,t]}H_s\mathrm{d}X_s,\quad t\geq 0,$$

where $H$ is a predictable process with values in $[-1,1]$. Let $X^*=\sup_{t\geq 0}X_t$, $|X|^*=\sup_{t\geq 0}|X_t|$ be the one-sided and two-sided maximal functions of $X$, respectively. We will also use the notation $X_t^*=\sup_{0\leq s\leq t}X_s$ and $|X_t|^*=\sup_{0\leq s\leq t}|X_s|$, $t\geq 0$.

One of the objectives of this paper is to study maximal weak-type inequalities between $X$ and $Y$. Let us take a look at some related results from the literature, which may serve as a motivation. In the seminal paper [4], Burkholder invented a method of proving general inequalities for stochastic integrals: he showed that the validity of a given estimate can be deduced from the existence of an appropriate special function which enjoys certain majorization and concavity properties. Burkholder applied this method to establish the celebrated sharp $L^p$ bound

$$||Y||_p\leq\max\left\{p-1,(p-1)^{-1}\right\}||X||_p,\qquad 1<p<\infty$$

2000 *Mathematics Subject Classification.* Primary: 60G42. Secondary: 60G44.
*Key words and phrases.* Martingale, maximal, weak type inequality, best constant.
Partially supported by NCN grant DEC-2012/05/B/ST1/00412.

(here and below, $||X||_p = \sup_{t\geq 0} ||X_t||_p$ is the $p$-th moment of $X$). If $p = 1$, the bound does not hold with any finite constant, but we have the substitute (cf. [4])

$$\mathbb{P}(|Y|^* \geq 1) \leq 2||X||_1, \tag{1.1}$$

in which the constant 2 is optimal. These results, as well as the methodology, have been extended in many directions and applied in many areas of mathematics: see e.g. [1], [2], [5], [6], [7], [10], [11], [15], [16], [17] and references therein.

One can also ask about maximal versions of the above statements. In [8], Burkholder enhanced his method from [4], and applied it to this new class of estimates. As an illustration, he showed the bound

$$||Y||_1 \leq \gamma|| \, |X|^*||_1, \tag{1.2}$$

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

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

This constant is the best possible in (1.2). For further results in this direction and much more, see [12], [13] and Chapter 7 in [14].

The purpose of this paper is to study the maximal version of (1.1), i.e., the modification which involves the maximal function of $X$ on the right. In view of (1.1) or (1.2), the inequality holds with some constant if $|X|^*$ is used. However, what may be a little unexpected, the estimate is also valid if only the one-sided maximal function is involved. Let us formulate the precise statement; for clarity, we have decided to consider the setting of arbitrary and continuous-path martingales separately.

**Theorem 1.1.** *Suppose that $X$ is a real-valued continuous-path martingale starting from 0 and $Y$ is the integral, with respect to $X$, of a certain process $H$ with values in $[-1, 1]$. Then we have the inequality*

$$\mathbb{P}(|Y|^* \geq 1) \leq 2\mathbb{E}X^* \tag{1.3}$$

*and the constant 2 is the best possible.*

This result is perhaps not that surprising, especially in the light of the following reasoning: any continuous-path martingale can be represented as a time-changed Brownian motion $B$; furthermore, by Levy's theorem, the distributions of $|B|$ and $B^*$ coincide. Consequently, this suggests that the constants in the above maximal weak-type bound and in (1.1) should coincide. Actually, we will not try to formalize the above reasoning and present an alternative proof.

For martingales with possibly discontinuous trajectories, the situation gets much more interesting and the description of the corresponding optimal constant is far more complicated (see Section 3 below). Here is the precise statement.

**Theorem 1.2.** *Suppose that $X$ is a real-valued martingale starting from 0 and $Y$ is the integral, with respect to $X$, of a certain process $H$ with values in $[-1, 1]$. Then we have the sharp inequality*

$$\mathbb{P}(|Y|^* \geq 1) \leq 3.477977\ldots\mathbb{E}X^* \tag{1.4}$$

*(the precise definition of the constant is given in (3.8) below).*

Our approach will depend heavily on the technique of Burkholder from [8], or rather its appropriate modification. We establish Theorem 1.1 in the next section, and devote the final part of the paper to the proof of Theorem 1.2.

## 2. Weak-type inequality for continuous-path martingales

We have decided to divide this section into two subsections.

2.1. **Proof of** (1.3)**.** As announced in the introduction, the key ingredient of the reasoning is a certain special function. Let $U : \mathbb{R}^3 \to \mathbb{R}$ be given by the formula

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

It is easy to check that $U$ is continuous. Furthermore, we have the following fact.

**Lemma 2.1.** *Suppose that $x \leq z$. Then*

$$1_{\{|y| \geq 1\}} - 2z \leq U(x, y, z) \leq 1 - 2z. \tag{2.1}$$

*Proof.* Let us start with the estimate on the right. If $|y| \geq x - z + 1$, then both sides are equal. On the other hand, if $|y| < x - z + 1$, then

$$U(x, y, z) - 1 + 2z = y^2 - (x - z + 1)^2 < 0.$$

To deal with the bound on the left, assume first that $|y| \geq 1$. Since $x \leq z$, we see that $|y| \geq x - z + 1$ and hence both sides are equal. If $|y| < 1$ and $|y| \geq x - z + 1$, then the estimate takes the trivial form $-2z \leq 1 - 2z$. The final case we need to consider is $|y| < x - z + 1$. Then $x - z \in (-1, 0]$ and hence

$$U(x, y, z) + 2z = y^2 - (x - z)^2 - 2(x - z) \geq 0.$$

This completes the proof. □

We are ready to establish the estimate of Theorem 1.1.

*Proof.* It is convenient to split the reasoning into several parts.

*Step 1.* Let us start with a reduction. Namely, it is enough to prove the inequality

$$\mathbb{P}(|Y_t| \geq 1) \leq 2\mathbb{E}X^* \qquad \text{for all } t \geq 0. \tag{2.2}$$

Indeed, having done this, we fix $\varepsilon \in (0, 1)$ and apply the bound to the stopped martingales $X^\sigma = (X_{\sigma\wedge t}/\varepsilon)_{t\geq 0}$ and $Y^\sigma = (Y_{\sigma\wedge t}/\varepsilon)_{t\geq 0}$, where $\sigma = \inf\{t \geq 0 : |Y_t| \geq \varepsilon\}$ (here and below we use the convention $\inf \emptyset = \infty$). Note that $Y^\sigma$ is a stochastic integral of $X^\sigma$, so the application is permitted. Since $\{|Y|^* \geq 1\} \subseteq \{|Y_t| \geq \varepsilon \text{ for some } t \geq 0\}$, we see that

$$\mathbb{P}(|Y|^* \geq 1) \leq \lim_{t\to\infty} \mathbb{P}(|Y_t^\sigma| \geq 1) \leq 2\mathbb{E}(X^\sigma)^* \leq 2\mathbb{E}X^*/\varepsilon,$$

and letting $\varepsilon \to 1$ gives the claim.

*Step 2.* We proceed to the proof of (2.2). With no loss of generality we may assume that $\mathbb{E}X^* < \infty$; otherwise, there is nothing to prove. Introduce the process $Z = (Z_t)_{t\geq 0} = ((X_t, Y_t, X_t^*))_{t\geq 0}$, fix $T \geq 0$ and consider the stopping time

$$\tau = \inf\{t \geq 0 : |Y_t| \geq X_t - X_t^* + 1\}.$$

We will show that

$$\mathbb{E}U(Z_T) \leq \mathbb{E}U(Z_{\tau\wedge T}) \tag{2.3}$$

(the integrability of the variables on the left and on the right follows from (2.1) and the assumption $\mathbb{E}X^* < \infty$). To do this, write the trivial equality

$$\mathbb{E}U(Z_T)1_{\{\tau > T\}} = \mathbb{E}U(Z_{\tau\wedge T})1_{\{\tau > T\}}.$$

Next, observe that

$$\begin{aligned} U(Z_{\tau\wedge T})1_{\{\tau\leq T\}} = U(Z_\tau)1_{\{\tau\leq T\}} &= (1-2X_\tau^*)1_{\{\tau\leq T\}} \\ &\geq (1-2X_T^*)1_{\{\tau\leq T\}} \geq U(X_T,Y_T,X_T^*)1_{\{\tau\leq T\}}, \end{aligned}$$

where in the last passage we have used (2.1). Now take expectation throughout and add the obtained bound to the preceding equality; we get (2.3).

*Step 3.* Now we will handle the term $\mathbb{E}U(X_{\tau\wedge T},Y_{\tau\wedge T},X^*_{\tau\wedge T})$. By the very definition of $\tau$, if $t\in[0,\tau\wedge T)$, then $Z_t$ takes values in the set $\{(x,y,z):|y|<x-z+1\}$, on which $U$ is of class $C^\infty$. Therefore, we are allowed to apply Itô's formula to get

$$U(Z_{\tau\wedge T}) = U(Z_0) + I_1 + I_2 + I_3/2, \tag{2.4}$$

where

$$\begin{aligned} I_1 &= \int_{0+}^{\tau\wedge T} U_x(Z_s)\mathrm{d}X_s + \int_{0+}^{\tau\wedge T} U_y(Z_s)\mathrm{d}Y_s, \\ I_2 &= \int_{0+}^{\tau\wedge T} U_z(Z_s)\mathrm{d}X_s^*, \\ I_3 &= \int_{0+}^{\tau\wedge T} U_{xx}(Z_s)\mathrm{d}[X,X]_s + 2\int_{0+}^{\tau\wedge T} U_{xy}(Z_s)\mathrm{d}[X,Y]_s + \int_{0+}^{\tau\wedge T} U_{yy}(Z_s)\mathrm{d}[Y,Y]_s, \end{aligned}$$

where $[X,X]$, $[X,Y]$, $[Y,Y]$ denote the square brackets of the processes indicated (see Dellacherie and Meyer [9] for details). Let us take a look at the terms $I_1-I_3$. The first of them has expectation zero, by the properties of stochastic integrals. To handle the second one, note that the process $X^*$ increases on the set $\{t:X_t=X_t^*\}$. However, we easily derive that for such $t$, $U_z(X_t,Y_t,X_t^*)=0$; this proves that $I_2=0$. Finally, note that $U_{xx}(x,y,z)=-2$, $U_{xy}(x,y,z)=0$ and $U_{yy}(x,y,z)=2$ on $\{(x,y,z):|y|<x-z+1\}$. Consequently,

$$I_3 = 2[Y,Y]_{\tau\wedge T} - 2[X,X]_{\tau\wedge T} \leq 0,$$

since $d[Y,Y]_t=|H_t|^2d[X,X]_t\leq \mathrm{d}[X,X]_t$ for all $t$ (and thus $[Y,Y]_{\tau\wedge T}\leq[X,X]_{\tau\wedge T}$). Putting all the above facts together and integrating both sides of (2.4), we obtain

$$\mathbb{E}U(Z_{\tau\wedge T}) \leq \mathbb{E}U(Z_0) = U(0,0,0) = 0.$$

*Step 4.* Now we combine the latter estimate with (2.3) and get

$$\mathbb{E}U(X_T,Y_T,X_T^*) \leq 0.$$

By the left majorization in (2.1), this implies

$$\mathbb{P}(|Y_T|\geq 1) - 2\mathbb{E}X_T^* \leq 0.$$

Since $T$ was arbitrary and $\mathbb{E}X_T^*\leq\mathbb{E}X^*$, the inequality (2.2) follows. This completes the proof of (1.3). □

2.2. **Sharpness of** (1.3)**.** We will exhibit an appropriate example. Suppose that $B=(B_t)_{t\geq 0}$ is a standard, one-dimensional Brownian motion starting from 0. Fix a small $\delta\in(0,1/2)$ and introduce the following stopping times: $\tau_0\equiv 0$ and, inductively,

$$\begin{aligned} \tau_{2n+1} &= \inf\{t\geq\tau_{2n}: B_t\leq -1/2 \text{ or } B_t\geq\delta\}, \\ \tau_{2n+2} &= \inf\{t\geq\tau_{2n+1}: B_t\leq 0 \text{ or } B_t\geq 1/2\}, \qquad n=0,\,1,\,2,\,\ldots. \end{aligned}$$

Note that $(\tau_n)_{n\geq 0}$ is a nondecreasing sequence of stopping times, with

$$\tau = \lim_{n\to\infty} \tau_n = \inf\{t \geq 0 : B_t \in \{-1/2, 1/2\}\},$$

so, in particular, $\tau$ is integrable. Next, consider $\sigma = \inf\{t : |B_t| \geq 1\}$ (satisfying $\sigma > \tau$ almost surely) and the predictable process

$$H_t = \sum_{n=0}^{\infty} (-1)^n 1_{[\tau_n, \tau_{n+1})}(t) + (\operatorname{sgn} B_\tau) \cdot 1_{[\tau,\sigma)}(t).$$

Finally, for any $t \geq 0$, put

$$X_t = \int_0^t H_s \mathrm{d}B_s \quad \text{ and } \quad Y_t = B_{\sigma\wedge t}.$$

Clearly, $Y$ is a stochastic integral of $H$ with respect to $X$. Let us describe briefly the behavior of the pair $(X, Y)$. It evolves according to the following, two-stage procedure. It starts from $(0, 0)$ and moves along the line of slope 1 until $Y$ gets to $-1/2$ or to $\delta$. If the first possibility occurs, the first stage $\tau$ is over; if the latter case happens, the pair starts moving along the line of slope $-1$, and does so until $Y$ reaches $1/2$ or 0. If $Y$ gets to $1/2$, the first stage $\tau$ is over; otherwise, $(X, Y)$ begins evolving along the line of slope 1, until $Y$ visits $-1/2$ or $\delta$. The pattern of movement is then repeated. Thus, at the end of the first stage $\tau$, we see that $|Y| = 1/2$; furthermore, we also easily check that $X^*_\tau - X_\tau$ belongs to the interval $[1/2 - \delta, 1/2 + \delta]$. To describe the second stage, assume that $Y_\tau = 1/2$ (when $Y_\tau = -1/2$, the evolution is symmetric). Then $(X, Y)$ evolves along the line of slope 1 until $Y$ reaches 1 or $-1$; by the previous considerations, we see that $X$ cannot exceed $X^* + \delta$.

Now the remainder of the proof is straightforward. By the above analysis,

$$\mathbb{E}X^* = \mathbb{E}X^*_\sigma \leq \mathbb{E}X^*_\tau + \delta \leq \frac{1}{2} + \mathbb{E}X_\tau + 2\delta = \frac{1}{2} + 2\delta$$

(in the last passage we used the integrability of $\tau$) and $\mathbb{P}(Y^* \geq 1) \geq \mathbb{P}(|Y_\sigma| = 1) = 1$. So, letting $\delta \to 0$, we see that the constant 2 is indeed the best possible in (1.3).

## 3. Proof of Theorem 1.2

In the non-continuous setting, the reasoning is much more involved. As previously, we have divided the section into a few parts.

3.1. **A differential equation.** Consider the second-order differential equation

$$2(1-y)(2-y)h''(y) + (3-2y)h'(y) + h(y) = 0. \tag{3.1}$$

Standard arguments show that there is a unique smooth solution $h : (-1, 1) \to \mathbb{R}$ to (3.1), satisfying the initial condition

$$h(0) = -1 \quad \text{ and } \quad h'(0) = 1.$$

Let us study some properties of the function $h$.

**Lemma 3.1.** *For any $y \in [0, 1)$ we have*

$$h'''(y) < 0, \tag{3.2}$$

$$h(y) + h'(y) \geq 0 \tag{3.3}$$

*and*

$$-h''(y) + (1-y)h'''(y) \geq 0. \tag{3.4}$$

*Proof.* Let us first establish (3.2). By the direct differentiation, (3.1) gives

$$2(1-y)(2-y)h'''(y) = (3-2y)h''(y) + h'(y). \tag{3.5}$$

Furthermore, again by (3.1), we get $h''(0) = -1/2$ and hence, plugging this above, we obtain $h'''(0) < 0$. Now suppose that there is $y \in (0,1)$ such that $h'''(y) \geq 0$ and let $y_0 = \inf\{y \in (0,1) : h'''(y) = 0\}$. Then $h'''(y) < 0$ for $y \in (0, y_0)$, so $h''(y) < h''(0) = -1/2$ and

$$h'(y) = h'(0) + \int_0^y h''(s)\mathrm{d}s < h'(0) - y/2$$

for these $y$'s. Consequently,

$$\begin{aligned} 0 = 2(1-y_0)(2-y_0)h'''(y_0) &= (3-2y_0)h''(y_0) + h'(y_0) \\ &< -\frac{1}{2}(3-2y_0) + 1 - \frac{1}{2}y_0 = \frac{y_0-1}{2} < 0, \end{aligned}$$

a contradiction. This yields the validity of (3.2). Note that as a by-product, we get that $h$ is a concave function on $(0,1)$. We turn to (3.3). By (3.1), it is equivalent to saying that $2(1-y)(2-y)h''(y) + 2(1-y)h'(y) \leq 0$, or

$$F(y) := (2-y)h''(y) + h'(y) \leq 0 \qquad \text{for } y \in [0,1).$$

However, we have $F(0) = 2h''(0) + h'(0) = 0$ and $F'(y) = (2-y)h'''(y) < 0$; this proves (3.3). Finally, let us deal with (3.4). By (3.5), it can be rewritten as

$$(3-2y)h''(y) + h'(y) \geq 2(2-y)h''(y),$$

or $h'(y) \geq h''(y)$. However, a stronger bound $G(y) := h'(y) - yh''(y) > 0$ holds true: it follows from the inequalities $G(0) = h'(0) = 1 > 0$ and $G'(y) = -yh'''(y) > 0$. ☐

**Lemma 3.2.** *The function $h$ extends to a continuous function on $(-1, 1]$. Furthermore, the limit $h'(1-) = \lim_{y\uparrow 1} h'(y)$ exists and is finite.*

*Proof.* Let us expand the function $h$ into power series:

$$h(y) = \sum_{n=0}^{\infty} a_n y^n.$$

Directly from the initial condition, we get that $a_0 = -1$ and $a_1 = 1$. Furthermore, the differential equation (3.1) gives the recurrence

$$a_{k+2} = \frac{3(2k-1)}{4(k+2)}a_{k+1} - \frac{2k^2-4k+1}{4(k+1)(k+2)}a_k, \qquad k = 0,\, 1,\, 2,\, \ldots.$$

Substitute $b_k = k(k-1)a_k$ for $k = 0, 1, 2, \ldots$. Then $b_0 = b_1 = 0$ and, as we will show in a moment,

$$b_2 < b_3 < b_4 < \ldots < 0. \tag{3.6}$$

To do this, note that the above recurrence for $(a_k)_{k\geq 0}$ gives

$$b_{k+2} = \frac{3(2k-1)}{4k}b_{k+1} - \frac{2k^2-4k+1}{4k(k-1)}b_k, \qquad k = 2,\, 3,\, \ldots.$$

We have $b_2 = -0.5$, $b_3 = -0.125$ and, as we will now show by induction,

$$b_{k+1} \leq b_k/2 \leq 0 \qquad \text{for } k \geq 3. \tag{3.7}$$

This is true for $k = 3$, since $b_4 = -0.078125$. Assuming that this property holds for some $k$, we use the above recurrence to get

$$b_{k+2} \leq \frac{3(2k-1)}{4k} b_{k+1} - \frac{2k^2 - 4k + 1}{2k(k-1)} b_{k+1} = \frac{b_{k+1}}{2} + \frac{k+1}{4k(k-1)} b_{k+1} \leq \frac{b_{k+1}}{2},$$

and (3.7) follows. To show that the sequence $(b_k)_{k\geq 2}$ is nondecreasing, note that we have $b_2 < b_3$ (recall the above values of $b_2$ and $b_3$). Now, assuming that $b_{k+1} - b_k \geq 0$ for some $k \geq 2$, we apply the above recurrence and the negativity of $b_k$ to get that

$$b_{k+2} - b_{k+1} = \frac{2k-3}{4k}(b_{k+1} - b_k) + \frac{2-k}{4k(k-1)} b_k \geq \frac{2k-3}{4k}(b_{k+1} - b_k) \geq 0.$$

Thus, (3.6) is proved. This property shows that the sequence $(b_k)_{k\geq 0}$ is bounded, and hence $a_k$ is a sequence of negative numbers with $a_k = O(k^{-2})$. Consequently, the series $\sum_{k=0}^{\infty} a_k$ converges and by Abel's theorem, we have

$$h(y) = \sum_{k=0}^{\infty} a_k y^k \to \sum_{k=0}^{\infty} a_k < 0,$$

as $y \uparrow 1$. To deal with $h'(1-)$, note that this limit exists since $h$ is concave on $[0, 1)$ (see the proof of the previous lemma). The finiteness of this limit follows from (3.3) and the finiteness of $h(1)$ we have just proved above. □

We are ready to introduce the constant of inequality (1.4). Namely, put

$$C = -1/h(1) = 3.477979\ldots, \tag{3.8}$$

where the approximation comes from computer simulations.

3.2. **A special function.** In the proof of the inequality (1.4) and its sharpness, we will exploit the following auxiliary function $u : (-\infty, 1] \times \mathbb{R} \to \mathbb{R}$:

$$u(x,y) = \begin{cases} 1 - C & \text{if } x \leq |y|, \\ 1 - Cx + C|y|h'(x - |y|) + Ch(x - |y|) & \text{if } x > |y| \end{cases} \tag{3.9}$$

(here $h'(1)$ is understood as the limit $h'(1-)$, and $C$ is as in (3.8)). We easily see that $u$ is continuous. Some further properties are studied in the lemmas below.

**Lemma 3.3.** *The function $u$ is concave along any line segment of slope $a \in [-1, 1]$ contained in $(-\infty, 1] \times \mathbb{R}$.*

*Proof.* The function $u$ is differentiable at any point of the form $(x, 0)$, $x < 1$, and $u(x, y) = u(x, -y)$ for all $x$, $y$. Therefore, it is enough to prove the concavity along line segments contained in $(-\infty, 1] \times (0, \infty)$. Fix any point $(x, y)$ from this set, a number $a \in [-1, 1]$ and consider the function $G(t) = G_{x,y,a}(t) = u(x + t, y + at)$ (defined for those $t$, for which $x + t \leq 1$ and $y + at \geq 0$). The claim will follow if we manage to show that $G$ is concave. To do this, we will prove that $G''(t) \leq 0$ for those $t$ for which the second derivative exists, and $G'(t-) \geq G'(t+)$ for the remaining $t$. Actually, it is enough to show this for $t = 0$, due to the translation property $G_{x,y,a}(r + s) = G_{x+r,y+ar,a}(s)$. If $x < y$, then $G''(0) = 0$; if $x > y$, then

$$G''(0) = u_{xx}(x,y) + 2au_{xy}(x,y) + a^2 u_{yy}(x,y). \tag{3.10}$$

However, we have $u_{xy}(x,y) = -Cyh'''(x-y) > 0$ and

$$\begin{aligned} u_{yy}(x,y) &= -Ch''(x-y) + Cyh'''(x-y) \\ &> C\big(-h''(x-y) + (1-x+y)h'''(x-y)\big) \geq 0, \end{aligned}$$

in view of (3.4). Consequently, if we fix $(x,y)$, then $G''(0)$ in (3.10) is the largest when $a=1$. However, then $G''(0)=0$, since $u$ is linear over line segments of slope 1 contained in $(-\infty,1]\times[0,\infty)$. The final case we need to consider is that of $x=y$. Then the second derivative of $G$ at $t=0$ does not exist (unless $a=1$: but then $G''(0)=0$, as we have just pointed out), so we need to compare the left- and the right-hand derivative. We compute that $G'(0-)=0$ and

$$G'(0+) = -C + Cyh''(0) + Ch'(0) - a\cdot Cyh''(0) = C(a-1)y/2 \leq 0.$$

This completes the proof. □

**Lemma 3.4.** *For any $d \geq 0$ and any $a \in [-1,1]$ we have*

$$u(1, y+ad) - Cd \leq u(1,y) + u_x(1,y)d + au_y(1,y)d. \tag{3.11}$$

Some comments concerning the partial derivatives above are in order. Namely, the symbol $u_x(1,y)$ means the left-sided derivative $\lim_{h\to 0}(u(1,y)-u(1-h,y))/h$; moreover, the derivative $u_y(1,\pm1)$ does not exist, and we take this symbol to be 0.

*Proof.* Let us first gather some information on the function $y \mapsto u(1,y)$. This function is even and continuous on $\mathbb{R}$. Moreover, it is equal to $1-C$ when $|y|\geq 1$, and is convex on $[-1,1]$ (see the above formula for $u_{yy}$). Consequently, if $|y|\geq 1$, the estimate (3.11) is clear: the left-hand side is at most $1-C$, while the right is equal to $1-C$. Suppose then that $|y|<1$; by symmetry, it is enough to study (3.11) for $y\in(0,1)$. Rewrite the bound as

$$u(1,y+ad) \leq u(1,y) + (C + u_x(1,y))d + u_y(1,y)ad. \tag{3.12}$$

By continuity, we may assume that $a \neq 0$. As we have already written above, the function $y\mapsto u(1,y)$ is continuous on $\mathbb{R}$, convex on $[-1,1]$ and equal to $1-C$ outside $(-1,1)$. On the other hand, the right-hand side of (3.12) is linear in $d$. Since both sides are equal for $d=0$, we will be done if we show the estimate for $d$ such that $y+ad\in\{-1,1\}$. Assume that $a$ and $d$ satisfy this condition and look back at (3.12). Note that $C+u_x(1,y) = C\big(yh''(1-y)+h'(1-y)\big)\geq 0$. Indeed, by (3.1) and some manipulations, we rewrite this bound as

$$F(y) = h'(1-y) - h(1-y) \geq 0,$$

and we have $F(0)=h'(1)-h(1) = -2h(1) = 2/C > 0$, $F'(y) = -h''(1-y)+h'(1-y)\geq 0$ (see the end of the proof of Lemma 3.1). Consequently, if $ad$ and $y$ are kept fixed, the right-hand side of (3.12) is minimal when $d$ is minimal. Thus it is enough to prove this estimate when $a\in\{-1,1\}$. We consider two cases separately.

*Case I:* $a=-1$. Then $d=1+y$ and the desired bound reads

$$\begin{aligned} 1-C \leq\ & 1-C+Cyh'(1-y)+Ch(1-y) \\ &+ (Cyh''(1-y)+Ch'(1-y))(1+y) + Cy(1+y)h''(1-y), \end{aligned}$$

or, after some manipulations,

$$2y(1+y)h''(1-y) + (1+2y)h'(1-y) + h(1-y) \geq 0.$$

By (3.1), we actually have equality here: we will need this fact later on.

*Case II:* $a = 1$. Then $d = 1 - y$ and the bound becomes

$$\begin{aligned}1 - C \leq 1 - C + Cyh'(1-y) + Ch(1-y) \\ + (Cyh''(1-y) + Ch'(1-y))(1-y) - Cy(1-y)h''(1-y),\end{aligned}$$

or $h(1-y) + h'(1-y) \geq 0$. However, we have already proved this in (3.3). □

3.3. **Proof of** (1.4)**.** Using approximation arguments of Bichteler [3], it suffices to focus on the following discrete-time version of the maximal bound. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with a filtration $(\mathcal{F}_n)_{n\geq 0}$ and let $f = (f_n)_{n\geq 0}$ be an adapted real-valued martingale satisfying $f_0 = 0$ almost surely. Define the corresponding difference sequence $(df_n)_{n\geq 0}$ by $df_0 = f_0$ and $df_n = f_n - f_{n-1}$, $n \geq 1$. Let $v = (v_n)_{n\geq 0}$ be a predictable sequence, i.e., the sequence of random variables such that for each $n$, the term $v_n$ is $\mathcal{F}_{(n-1)\vee 0}$ measurable. A sequence $g = (g_n)_{n\geq 0}$ is a transform of $f$ by $v$, if for any $n = 0, 1, 2, \ldots$ we have the equality $dg_n = v_n df_n$. Note that by the predictability of $v$, the sequence $g$ is again a martingale. By the aforementioned results of Bichteler, (1.4) will follow if we establish the bound

$$\mathbb{P}(|g|^* \geq 1) \leq C\mathbb{E}f^*,$$

for all $f$, $g$ as above, where $v$ is assumed to take values in $[-1, 1]$. Here, in analogy with the continuous-time setting, $f^* = \sup_{n\geq 0} f_n$ and $|g|^* = \sup_{n\geq 0} |g_n|$. Actually, using some straightforward approximation and a stopping-time argument as in the continuous-time case, it is enough to show

$$\mathbb{P}(|g_n| \geq 1) \leq C\mathbb{E}f_n^*, \qquad n = 0, 1, 2, \ldots, \tag{3.13}$$

for simple $f$ and $g$ with $f_0 = 0$. Here by simplicity of $f$ we mean that for any $n$, the random variable $f_n$ takes only a finite number of values and there exists a deterministic number $N$ such that $f_N = f_{N+1} = \ldots$. Clearly, if $f$ and $g$ are simple, then the almost sure limits $f_\infty$ and $g_\infty$ exist and are finite.

Now we will describe the appropriate modification of Burkholder's method from [8], which will be useful in the proof of (3.13). Let $D = \mathbb{R} \times \mathbb{R} \times [0, \infty)$ and $V : D \to \mathbb{R}$ be any function satisfying $V(x, y, z) = V(x, y, x \vee z)$. Suppose we want to prove the inequality

$$\mathbb{E}V(f_n, g_n, f_n^*) \leq 0, \qquad n = 0, 1, 2, \ldots, \tag{3.14}$$

for all pairs $(f, g)$, where $f$ is a simple martingale starting from 0 and $g$ is its transform by a predictable sequence with values in $[-1, 1]$. The key idea is to study a class $\mathcal{U}(V)$ which consists of all functions $U : D \to \mathbb{R}$ satisfying the four properties below:

$$U(0, 0, 0) \leq 0, \tag{3.15}$$

$$U(x, y, z) = U(x, y, x \vee z) \quad \text{if } (x, y, z) \in D, \tag{3.16}$$

$$V(x, y, z) \leq U(x, y, z) \quad \text{if } (x, y, z) \in D \tag{3.17}$$

and the following further condition: for all $(x, y, z) \in D$ with $x \leq z$, any $a \in [-1, 1]$, $\alpha \in (0, 1)$ and any $t_1,\ t_2 \in \mathbb{R}$ satisfying $\alpha t_1 + (1 - \alpha)t_2 = 0$, we have

$$\alpha U(x + t_1, y + at_1, z) + (1 - \alpha)U(x + t_2, y + at_2, z) \leq U(x, y, z). \tag{3.18}$$

The interplay between the maximal inequality (3.14) and the class $\mathcal{U}(V)$ is described in the theorem below. It is a simple modification of Theorems 2.2 and 2.3

in [8] (see also Section 11 in [4] and Chapter 7 in [14]). We omit the proof, as it requires only some straightforward minor changes.

**Theorem 3.5.** *The inequality (3.14) holds for all $n$ and all pairs $(f,g)$ as above if and only if the class $\mathcal{U}(V)$ is nonempty.*

To handle (3.13), we will apply the above technique to the function $V(x,y,z) = 1_{\{|y|\geq 1\}} - C(x\vee z)$. Introduce $U : D\to\mathbb{R}$ by

$$U(x,y,z) = u(x+1-(x\vee z),y) - C((x\vee z)-1).$$

We will prove the following fact.

**Lemma 3.6.** *The function $U$ is an element of $\mathcal{U}(V)$.*

*Proof.* The conditions (3.15) and (3.16) are immediate. To prove the majorization (3.17), note that it is equivalent to $u(x,y)\geq 1_{\{|y|\geq 1\}} - C$ for all $x\leq 1$ and $y\in\mathbb{R}$. If $|y|\geq 1$, then both sides are equal; if $|y|<1$, then, as follows from Lemma 3.3, the function $x\mapsto u(x,y)$ is concave on $(-\infty,1]$, equal to $1-C$ for sufficiently small $x$ and at least $-C$ for $x=1$ (the latter statement follows from $u(1,y)\geq u(1,0)=-C$). This gives (3.17), and we turn our attention to the final condition (3.18). Fix the parameters as in its statement and consider the function $\Phi(t) = U(x+t,y+at,z)$. Clearly, we will be done if we manage to show the existence of a linear $\Psi:\mathbb{R}\to\mathbb{R}$ satisfying $\Psi\geq\Phi$ and $\Psi(0)=\Phi(0)$. To do this, let us look at the function $\Phi$. If $t\leq z-x$, then $\Phi(t) = u(x+t+1-z,y+at) - C(z-1)$; hence Lemma 3.3 implies that $\Phi$ is concave on $(-\infty,z-x]$. On the other hand, if $t>z-x$, then

$$\begin{aligned}\Phi(t) &= u(1,y+at) - C(x+t-1)\\ &= u\big(1,y+a(z-x)+a(t+x-z)\big) - C(t+x-z) - C(z-1),\end{aligned}$$

which, by Lemma 3.4, is not larger than

$$\begin{aligned}u(1,y+a(z-x)) &+ u_x(1,y+a(z-x))(t+x-z)\\ &\quad + au_y(1,y+a(z-x))(t+x-z) - C(z-1)\\ &= \Phi(z-x) + \Phi'\big((z-x)-\big)(t+x-z).\end{aligned}$$

By the concavity of $\Phi$ on $(-\infty,z-x]$ (which has just been established above), the latter expression does not exceed $\Phi(0)+\Phi'(0-)t$. Thus we see that the choice $\Psi(t) = \Phi(0)+\Phi'(0-)t$ works. □

3.4. **Sharpness of** (1.4)**.** Now we will prove that the constant 3.477977... is the best in (1.4), by providing an appropriate example. It is enough to show that this constant is optimal in the discrete-time case, i.e., in (3.13). Let $y_0\in(0,1)$ be a fixed number, let $N$ be a large positive integer and put $\delta = (1-y_0)/(2N)$. Consider an auxiliary discrete-time Markov martingale $(F_n,G_n)_{n\geq 0}$ with values in $\mathbb{R}^2$, whose distribution is uniquely determined by the following conditions:

(i) We have $(F_0,G_0) = (1,y_0)$.
(ii) For any $y\in[y_0,1)$, a state of the form $(1,y)$ leads to $(1-\delta,y+\delta)$ or to $(2+y,-1)$. For any $y\in(-1,-y_0]$, a state of the form $(1,-y)$ leads to $(1-\delta,y-\delta)$ or to $(2+y,1)$.
(iii) For any $y\in[y_0,1)$, a state of the form $(1-\delta,y+\delta)$ leads to $(1,y+2\delta)$ or to $(1-y-2\delta,0)$. For any $y\in(-1,-y_0]$, a state of the form $(1-\delta,y-\delta)$ leads to $(1,-y-2\delta)$ or to $(1-y-2\delta,0)$.

(iv) For any $y \in (0, 1-y_0]$, a state of the form $(y, 0)$ leads to $(y-\delta, -\delta)$ or to $(1, 1-y)$.
(v) For any $y \in (0, 1-y_0]$, a state of the form $(y-\delta, -\delta)$ leads to $(y-2\delta, 0)$ or to $(1, y-2\delta-1)$.
(vi) The state $(0,0)$ leads to $(1,1)$ or to $(-1,-1)$.
(vii) All the other states are absorbing.

We did not specify the transition probabilities, they are uniquely determined by the requirement that the pair $(F_n, G_n)_{n\geq 0}$ is a martingale. To gain some intuition about the process, it is best to look at a few first steps. The pair starts from $(1, y_0)$. In the first step, it moves, along the line of slope $-1$, either to $(2+y_0, -1)$ (and then stops), or to $(1-\delta, y_0+\delta)$. If the second possibility occurs, the next step begins, in which the pair moves along the line of slope 1. It moves either to $(1, y_0+2\delta)$ (and then it evolves as previously - along the line of slope $-1$, to one of the points $(2+y_0+2\delta, -1)$, $(1-\delta, y+3\delta)$, and so on), or to $(1-y_0-2\delta, 0)$. In the latter case, it moves, again along the line of slope 1, to $(1, y_0+2\delta)$ (and then evolves according to the above scheme), or to $(1-y_0-3\delta, -\delta)$. If the latter occurs, the pair moves along the line of slope $-1$, to $(1, -y_0-4\delta)$ (and then its movement is symmetric to that corresponding to the point $(1, y_0+4\delta)$), or to $(1-y_0-4\delta, 0)$ (in which the evolution is analogous to that corresponding to the starting point $(1-y_0-2\delta, 0)$). The pattern of the movement is then repeated.

It is clear from the above description that for any $n \geq 1$ we have $dG_n = v_n dF_n$, for some predictable sequence $v$ with values in $\{-1, 1\}$. Next, look at the process $H = (|F_n - 1| + |G_n|)_{n\geq 0}$. A careful inspection of the dynamics described in (i)-(viii) shows that this process is nondecreasing; this in particular gives that the pair $(F_n, G_n)_{n\geq 0}$ is bounded away from $(1, 0)$. Now, note that the martingale $(F_n, G_n)_{n\geq 0}$ is simple: it is clear that for each $n$ the random variable $(F_n, G_n)$ takes a finite number of values, and, as we will prove now, the number of nontrivial steps is bounded by $3N$. Suppose on contrary that there is a trajectory of length exceeding $3N$; then the first $3N$ moves of $(F_n, G_n)_{n\geq 0}$ must be of type (ii)-(vii). However, if we look back at the process $H = (|F_n - 1| + |G_n|)_{n\geq 0}$ and the conditions (ii)-(vii), we see that during each three consecutive moves of this type, the process $H$ increases in at least $2\delta$. Since $H_0 = y_0$, we get $H_{3N} \geq y_0 + 2N\delta = 1$, and this means that the evolution of $(F_n, G_n)_{n\geq 0}$ is over after $3N$ steps, a contradiction. Our final observation is that $G_\infty \in \{-1, 1\}$ with probability 1.

The key fact is that the sequence $(\mathbb{E}U(F_n, G_n, F_n^*))_{n\geq 1}$ is "almost" nondecreasing. Here is the precise statement.

**Lemma 3.7.** *There is a constant $K$ depending only on $y_0$ such that for each $n$,*

$$\mathbb{E}U(F_{n+1}, G_{n+1}, F_{n+1}^*) \geq \mathbb{E}U(F_n, G_n, F_n^*) - K\delta^2.$$

*Proof.* The idea is that $((F_n, G_n, F_n^*))_{n\geq 0}$ evolves along the directions along which $U$ is linear, or almost linear. We will present a detailed analysis for the moves described in (ii) and (iii), the remaining ones can be handled similarly. As we have proved in Lemma 3.4, for any $y \in (0, 1)$ we have

$$u(1, -1) - C(1+y) = u(1, y) + u_x(1, y)(1+y) - u_y(1, y)(1+y)$$

(see the proof of (3.11) with $a = -1$ and $d = 1+y$). Now, suppose that $\mathbb{P}((F_n, G_n) = (1, y)) > 0$. Then

$$\begin{aligned}
&\mathbb{E}\big[U(F_{n+1}, G_{n+1}, F^*_{n+1})\big|(F_n, G_n) = (1, y)\big] - U(1, y, 1)\\
&= \frac{1+y}{1+y+\delta}U(1-\delta, y+\delta, 1) + \frac{\delta}{1+y+\delta}U(2+y, -1, 1) - U(1, y, 1)\\
&= \frac{1+y}{1+y+\delta}u(1-\delta, y+\delta) + \frac{\delta}{1+y+\delta}\big(u(1,-1) - C(1+y)\big) - u(1, y)\\
&= \frac{1+y}{1+y+\delta}\big[u(1-\delta, y+\delta) - u(1, y) + u_x(1, y)\delta - u_y(1, y)\delta\big].
\end{aligned}$$

By the mean value property, this is bounded from below by $-K\delta^2$, for some $K$ depending on the second-order partial derivatives of $u$. But these partial derivatives are uniformly bounded, since the process $(F, G)$ is bounded away from $(1, 0)$; thus $K$ can be chosen to depend only on $y_0$. The boundedness of the derivatives is due to an appropriate extension argument. Consider the function $v(x, y) = 1 - Cx + C|y|h'(x-|y|) + Ch(x-|y|)$, $-1 < x-|y| < 1$. This function coincides with $u$ on the set $|y| \leq x \leq 1$ (see (3.9)) and is of class $C^2$ on $\{(x, y) : -1 < x - |y| < 1,\ |y| \neq 0\}$, so in particular it has uniformly bounded second-order derivatives on each set of the form $\{1\} \times ([-1, -y_0] \cup [y_0, 1])$ (and hence the same is true for $u$).

For the moves in (iii), the argumentation is much simpler, as then $(F, G)$ moves along line segment of slope 1 contained in $(-\infty, 1] \times \mathbb{R}$, along which $u$ is linear. Consequently, if $\mathbb{P}((F_n, G_n) = (1-\delta, y+\delta)) > 0$ for some $y \in (0, 1)$, then

$$\mathbb{E}\big[U(F_{n+1}, G_{n+1}, F^*_{n+1})\big|(F_n, G_n) = (1-\delta, y+\delta)\big] = U(1-\delta, y+\delta, 1),$$

i.e., the desired bound holds with $K = 0$. We leave the analogous analysis of the moves from (iv)-(vi) to the reader. □

Equipped with the above statements, we can now prove the sharpness of (1.4). Consider a "shifted" modification of the process $(F, G)$: assume that $(f, g)$ is a Markov martingale starting from $(0, 0)$, which, at its first move jumps to $(-1, -1)$ or to $(y_0, y_0)$; next, if it went to $(-1, 1)$, it stops; if it went to $(y_0, y_0)$, then it evolves as $(F - 1 + y_0, G)$. We easily see that $G$ is a transform of $F$ by a predictable sequence with values in $\{-1, 1\}$ and $\mathbb{P}(|G|^* \geq 1) = 1$. To handle $\mathbb{E}F^*$, we use Lemma 3.7. Namely,

$$\begin{aligned}
\mathbb{E}f^* &= \frac{1}{1+y_0}\mathbb{E}(F - 1 + y_0)^*\\
&= \frac{1}{1+y_0}\left(\mathbb{E}F^* - 1 + y_0\right)\\
&= \frac{1}{1+y_0}\left(\frac{1 - U(F_{3N}, G_{3N}, F^*_{3N})}{C} - 1 + y_0\right)\\
&\leq \frac{1}{1+y_0}\left(\frac{1 - U(F_0, G_0, F^*_0)}{C} - 1 + y_0\right) + O(\delta)\\
&= \frac{1}{1+y_0}\left(-y_0h'(1-y_0) - h(1-y_0) + y_0\right) + O(\delta).
\end{aligned}$$

Therefore, letting $N \to \infty$ (or $\delta \to 0$), we see that the best constant in (1.4) must be at least

$$\left( \frac{1}{1+y_0} \left( -y_0 h'(1-y_0) - h(1-y_0) + y_0 \right) \right)^{-1}.$$

However, $y_0$ was an arbitrary number; letting $y_0 \to 0$, we see that the above expression converges to $h(1)^{-1} = C$. This establishes the desired sharpness.

## Acknowledgments

The author would like to express his gratitude to an anonymous Referee for the careful reading of the first version of the paper and many helpful comments. The results were obtained when the author was visiting Purdue University, USA.

## References

[1] R. Bañuelos, A. Bielaszewski and K. Bogdan, *Fourier multipliers for non-symmetric Lévy processes*, Marcinkiewicz centenary volume, Banach Center Publ. **95** (2011), pp. 9–25, Polish Acad. Sci. Inst. Math., Warsaw.

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

[3] K. Bichteler, *Stochastic integration and $L^p$-theory of semimartingales*, Ann. Probab. 9 (1980), pp. 49–89.

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

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

[6] D. L. Burkholder, *Differential subordination of harmonic functions and martingales*, Harmonic Analysis and Partial Differential Equations (El Escorial, 1987), Lecture Notes in Mathematics 1384 (1989), 1–23.

[7] D. L. Burkholder, *A proof of Pełczyński's conjecture for the Haar system*, Studia Math. **91** (1988), 268–273.

[8] 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.

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

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

[11] F. Nazarov and A. Volberg, *Heating of the Ahlfors-Beurling operator and esti- mates of its norm*, St. Petersburg Math. J. **15** (2004), No. 4, pp. 563–573.

[12] A. Osękowski, *Sharp inequality for martingale maximal functions and stochastic integrals*, Illinois J. Math., volume in honour of Donald L. Burkholder **54** (2010), pp. 1133–1156.

[13] A. Osękowski, *Maximal inequalities for continuous martingales and their differential subordinates*, Proc. Amer. Math. Soc. **139** (2011), 721–734.

[14] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72** (2012), Birkhäuser Basel.

[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.

[16] V. Vasyunin, A. Volberg, *Burkholders function via MongeAmpére equation*, Illinois J. Math. **54** (2010), pp. 1393–1428.

[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

*E-mail address*: `ados@mimuw.edu.pl`