PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 36, Fasc. 2 (2016), pp. 311–333

A MAXIMAL INEQUALITY FOR STOCHASTIC INTEGRALS*

BY

MATEUSZ RAPICKI (WARSZAWA)

Abstract. Assume that X is a *càdlàg*, real-valued martingale starting from zero, H is a predictable process with values in $[-1,1]$ and $Y=\int H dX$. This article contains the proofs of the following inequalities:

(i) If X has continuous paths, then

$$\mathbb{P}(\sup_{t\geqslant 0} Y_t \geqslant 1) \leqslant 2\mathbb{E}\sup_{t\geqslant 0} X_t,$$

where the constant 2 is the best possible.

(ii) If X is arbitrary, then

$$\mathbb{P}(\sup_{t\geqslant 0} Y_t \geqslant 1) \leqslant c\mathbb{E}\sup_{t\geqslant 0} X_t,$$

where $c=3.0446...$ is the unique positive number satisfying the equation $3c^4-8c^3-32=0$. This constant is the best possible.

2010 AMS Mathematics Subject Classification: Primary: 60G42; Secondary: 60G44.

Key words and phrases: Martingale, sharp inequality.

1. INTRODUCTION

Since the classical works of Kolmogorov, Hardy, Littlewood, Wiener and Doob from the first half of the twentieth century, maximal inequalities have played an important role in probability and analysis. The purpose of this paper is to establish sharp versions of some weak-type inequalities arising in the context of stochastic integrals with respect to *càdlàg* martingales.

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t\geqslant 0}$ be a filtration satisfying the usual conditions, i.e., it is right-continuous and $\mathcal{F}_t$ is complete for every $t\geqslant 0$. Let $(X_t)_{t\geqslant 0}$ be an adapted, *càdlàg*, real-valued martingale starting from zero and let $(H_t)_{t\geqslant 0}$ be a predictable process with values in $[-1,1]$. We put $Y_t=\int_0^t H_s dX_s$ and $X_t^*=\sup_{s\leqslant t} X_s$ for all $t\geqslant 0$. We will also use the notation $X^*=\sup_{t\geqslant 0} X_t$ and, analogously, $|X|^*=\sup_{t\geqslant 0}|X_t|$. Furthermore, $\langle X\rangle$ will stand for the quadratic covariance process of X; see Dellacherie and Meyer [8] for the definition and some basic properties of this object.

* Research supported in part by the NCN grant DEC-2014/14/E/ST1/00532.

Various inequalities between X and Y satisfying the above assumptions have been studied intensively in the literature. For example, Burkholder proved (see [6]) the sharp bound on the L^p-norm, i.e.,

$$\sup_{t\geqslant 0}\left(\mathbb{E}|Y_t|^p\right)^{1/p} \leqslant \max\left(p-1, \frac{1}{p-1}\right)\sup_{t\geqslant 0}\left(\mathbb{E}|X_t|^p\right)^{1/p}, \qquad 1<p<\infty, \tag{1.1}$$

and the corresponding weak-type inequality for $1 \leqslant p \leqslant 2$:

$$\mathbb{P}(|Y|^* \geqslant 1) \leqslant \frac{2}{\Gamma(p+1)}\sup_{t\geqslant 0}\mathbb{E}|X_t|^p. \tag{1.2}$$

When $p \geqslant 2$, the corresponding sharp estimate was established by Suh [15]:

$$\mathbb{P}(|Y|^* \geqslant 1) \leqslant \frac{p^{p-1}}{2}\sup_{t\geqslant 0}\mathbb{E}|X_t|^p.$$

For an overview of results in this direction, consult [16] or the monograph [12] and the references therein; see also [1]–[4], [9], [10] for applications in the study of various classes of Fourier multipliers.

There exist also maximal versions of the above estimates. In [7] Burkholder proved the sharp inequality

$$\sup_{t\geqslant 0}\mathbb{E}|Y_t| \leqslant c\mathbb{E}|X|^*,$$

which can be regarded as a substitute of (1.1) for $p = 1$. Here $c = 2.536...$ is the unique solution of the equation

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

If X is assumed to be nonnegative, then the optimal constant decreases to $2 + (3e)^{-1} = 2.1226...$, see Osękowski [11]. The paper [14] contains the proof of a one-sided maximal version of (1.2), i.e.,

$$\mathbb{P}(|Y|^* \geqslant 1) \leqslant c\mathbb{E}X^*, \tag{1.3}$$

and identifies the best absolute constant c in this inequality. It is equal to $-1/h(1) = 3.4779...$, where h is the solution of the equation

$$2(1-y)(2-y)h''(y) + (3-2y)h'(y) + h(y) = 0,$$

with the initial conditions $h(0) = -1$ and $h'(0) = 1$.

The proofs of all the above inequalities were carried out by using the same general technique, invented by Burkholder (cf. [7]), and its modifications. This method relies on coming up with a special function satisfying certain majorization and concavity conditions and then deducing the estimate from the existence of such a function. More information on this topic can be found in [12] and [13]. In the present article we will employ the same method to prove a modification of (1.3) where we bound $\mathbb{P}(Y^* \geqslant 1)$ instead of $\mathbb{P}(|Y|^* \geqslant 1)$. We will consider both the case where X has continuous paths and the case without this assumption. We will also deal with the case where the process (X, Y) starts from an arbitrary point $(x, y) \in \mathbb{R}^2$ rather than from $(0, 0)$. The precise formulation is given in the two theorems below.

THEOREM 1.1. *Assume that X is a continuous, real-valued martingale starting from zero, H is a predictable process with values in $[-1, 1]$ and $Y = \int H dX$. Then*

$$\mathbb{P}(Y^* \geqslant 1) \leqslant 2\mathbb{E}X^*, \tag{1.4}$$

and the constant 2 is the least possible. Moreover, if for any $x, y \in \mathbb{R}$ we define $X'_t = x + X_t$ and $Y'_t = y + Y_t$, then

$$\mathbb{P}\big((Y')^* \geqslant 1\big) \leqslant 2\mathbb{E}(X')^* + U(x, y, x), \tag{1.5}$$

where U is given in Section 2.1 below.

Notice that while (1.4) follows immediately from what is proven in [14], the inequality (1.5) is a new result.

THEOREM 1.2. *Assume that X is a* càdlàg, *real-valued martingale starting from zero, H is a predictable process with values in $[-1, 1]$ and $Y = \int H dX$. Then*

$$\mathbb{P}(Y^* \geqslant 1) \leqslant c\mathbb{E}X^*, \tag{1.6}$$

where $c = 3.0446...$ is the unique positive number satisfying $3c^4 - 8c^3 - 32 = 0$. The constant is the best possible. Moreover, if for any $x, y \in \mathbb{R}$ we define $X'_t = x + X_t$ and $Y'_t = y + Y_t$, then

$$\mathbb{P}\big((Y')^* \geqslant 1\big) \leqslant c\mathbb{E}(X')^* + U(x, y, x), \tag{1.7}$$

where U is given in Section 3.2 below.

Both (1.6) and (1.7) are novel results.

The reminder of the article is split into two parts. In the next section we study the continuous case and establish Theorem 1.1. The final part of the paper is devoted to the proof of Theorem 1.2.

2. CONTINUOUS CASE

2.1. A special function. The primary goal of this section is to study the inequality

$$\mathbb{P}(Y^* \geqslant 1) \leqslant 2\mathbb{E}X^*.$$

As we have already mentioned in the previous section, it follows directly from (1.3), but our analysis will provide an additional insight, which will lead to important extensions. Using a standard stopping time argument (see, e.g., [14]), we see that the above bound follows from the estimate

$$\mathbb{P}(Y_t \geqslant 1) \leqslant 2\mathbb{E}X_t^*, \qquad t \geqslant 0.$$

The latter bound can be rewritten in the equivalent form

$$\mathbb{E}V(X_t, Y_t, X_t^*) \leqslant 0, \qquad t \geqslant 0, \tag{2.1}$$

where $V : \mathbb{R}^3 \to \mathbb{R}$ is given by $V(x, y, z) = \mathbf{1}_{\{y \geqslant 1\}} - 2(x \vee z)$. As shown by Burkholder (cf. [7]), the maximal inequality of this type can be deduced from the existence of a special function, enjoying certain majorization and concavity properties. To introduce this special object, we need some extra notation.

Let $D = \{(x, y, z) \in \mathbb{R}^3 : x \leqslant z\}$. We divide D into three parts:

$$\begin{aligned}
D_1 &= \{(x, y, z) \in D : 1 - y \leqslant z - x\},\\
D_2 &= \{(x, y, z) \in D : 1 - y > z - x \text{ and } 1 - y + z - x > 1\},\\
D_3 &= \{(x, y, z) \in D : 1 - y > z - x \text{ and } 1 - y + z - x \leqslant 1\}.
\end{aligned}$$

We see that D_i are pairwise disjoint and their union is D.

Introduce a function $U : D \to \mathbb{R}$ by the formula

$$U(x, y, z) = \begin{cases} 1 - 2z & \text{for } (x, y, z) \in D_1,\\ -2z + \frac{2(z-x)}{z-x+1-y} & \text{for } (x, y, z) \in D_2,\\ 1 - 2z - (1 - y + x - z)(1 + y + x - z) & \text{for } (x, y, z) \in D_3. \end{cases}$$

Later on, we will need the following properties of U:

LEMMA 2.1. *U satisfies the following conditions*:

1. *U is continuous.*
2. *$U(x, y, z) \geqslant V(x, y, z)$ for all $(x, y, z) \in D$.*
3. *For all $(x, y, z) \in D$ and $H \in [-1, 1]$ the function $f_{x,y,z,H} : (-\infty, z-x] \to \mathbb{R}$ given by*

$$f_{x,y,z,H}(t) = U(x + t, y + Ht, z)$$

is concave.

4. *U is nonincreasing in z.*
5. *$U(0, 0, 0) = 0$.*

Proof. This is elementary. Let us handle each property separately below.

1. It is obvious that U is continuous in the interior of every D_i, so it is enough to check that the formulas for U match on the boundaries. This simple check is left to the reader.

2. In D_1 we have $U(x,y,z) = 1-2z \geqslant \mathbf{1}_{\{y\geqslant 1\}} - 2z = V(x,y,z)$. In D_2 we have $V(x,y,z) = -2z$ and $U(x,y,z) > -2z$ because both $z-x$ and $z-x+1-y$ are positive. In D_3 we have $(1-y+x-z)(1+y+x-z) \leqslant (1+x-z)^2 \leqslant 1$, so we get

$$V(x,y,z) = -2z \leqslant 1-2z-(1-y+x-z)(1+y+x-z) = U(x,y,z).$$

3. Notice that U is smooth in the interior of each D_i. First we will check that $f''(t) \leqslant 0$ if $(x+t, y+Ht, z)$ belongs to the interior of some D_i. Since $f_{x,y,z,H}(t+s) = f_{x+t,y+Ht,z,H}(s)$ for all s,t, it suffices to check the sign of $f''(0)$. To simplify calculations we will perform a change of variables. Fix z. Let $\bar{U} : \{(u,v) \in \mathbb{R}^2 : u+v \geqslant 0\} \to \mathbb{R}$ be given by the formula $\bar{U}(u,v) = 1-2z + \min\big(0, (u/v)\big(1-(v-1)^2\mathbf{1}_{\{v\leqslant 1\}}\big)\big)$ (this minimum equals zero for $u \geqslant 0$ only). If we put $u = z-x-1+y,\ v = z-x+1-y$, then we will have $\bar{U}(u,v) = U(x,y,z)$. Therefore, $f(t) = \bar{U}\big(u-(1-H)t, v-(1+H)t\big)$. This means that

$$f'' = a^2\bar{U}''_{uu} + 2ab\bar{U}''_{uv} + b^2\bar{U}''_{vv},$$

where $a = -(1-H),\ b = -(1+H)$. Simple calculation shows that this expression is nonpositive for any $a, b \leqslant 0$.

So, to complete the verification of the concavity of f, we need to check that

$$f'_{x,y,z,H}(0^-) \geqslant f'_{x,y,z,H}(0^+)$$

for all (x,y,z) on a common boundary of D_i and D_j, $i \neq j$. A little calculation shows that on the boundary between D_2 and D_3 left and right derivatives of f at zero are equal:

$$f'(0^-) = -(1-H) + u(1+H) = f'(0^+).$$

On the boundary between D_1 and $D_2 \cup D_3$ the inequality $f'(0^-) \geqslant f'(0^+)$ follows from the fact that the minimum of two concave functions is concave.

4. Since U is continuous, it suffices to check the sign of U'_z inside D_1, D_2 and D_3. In D_1 we have $U'_z = -2 \leqslant 0$. In D_2 we have

$$U'_z = \frac{2(1-y)}{(z-x+1-y)^2} - 2.$$

Notice that $1-y \leqslant 1-y+z-x < (1-y+z-x)^2$, so $U'_z < 0$. In D_3 we have $U'_z = -2(z-x) \leqslant 0$.

5. This is trivial. ∎

The proof of (1.4) will rest on the application of Itô's formula to the process

$$\big(U(X_t, Y_t, X^*_t)\big)_{t\geqslant 0}.$$

At the first glance, this operation is not permitted as U does not have the necessary smoothness; to overcome this difficulty, we use a standard mollification argument. Let $g : \mathbb{R}^3 \to [0, \infty)$ be a C^∞-function such that

$$\mathrm{supp} g \subset B(0,1) \quad \text{and} \quad \int_{B(0,1)} g = 1.$$

Next, for any $\delta > 0$, consider the function $U^\delta : \{(x,y,z) \in \mathbb{R}^3 : x \leqslant z + \delta\} \to \mathbb{R}$ given by

$$U^\delta(x,y,z) = \int_{B(0,1)} U(x - \delta u, y - \delta v, z + 3\delta - \delta w) g(u,v,w) du dv dw.$$

Analogously, we define

$$V^\delta(x,y,z) = \int_{B(0,1)} V(x - \delta u, y - \delta v, z + 3\delta - \delta w) g(u,v,w) du dv dw.$$

We immediately see that U^δ inherits most of the properties of U. We formulate the list of conditions in a separate statement.

LEMMA 2.2. *The function U^δ satisfies the following conditions*:
1. *U^δ is of class C^∞.*
2. *$U^\delta(x,y,z) \geqslant V^\delta(x,y,z)$ for $x \leqslant z + \delta$.*
3. *For all $x \leqslant z + \delta$ and $|H| \leqslant 1$ the function $f_{x,y,z,H,\delta} : (-\infty, z - x + \delta]$ given by*

$$f_{x,y,z,H,\delta}(t) = U^\delta(x + t, y + Ht, z)$$

is concave.
4. *U^δ is nonincreasing in z.*
5. *$\lim_{\delta \to 0^+} U^\delta(x,y,z) = U(x,y,z)$ for all $(x,y,z) \in D$.*

We leave the straightforward proof to the reader.

2.2. Proof of (1.4) and (1.5). Introduce the process $Z_t = (X_t, Y_t, X_t^*)$, $t \geqslant 0$. The Itô formula states that $U^\delta(Z_t) = U^\delta(0,0,0) + I_1 + I_2 + I_3 + I_4$, where

$$\begin{aligned}
I_1 &= \int_0^t U_x^\delta(Z_s) dX_s + \int_0^t U_y^\delta(Z_s) dY_s, \\
I_2 &= \int_0^t U_z^\delta(Z_s) dX_s^*, \\
I_3 &= \frac{1}{2} \int_0^t U_{xx}^\delta(Z_s) d\langle X \rangle_s + \int_0^t U_{xy}^\delta(Z_s) d\langle X, Y \rangle_s + \frac{1}{2} \int_0^t U_{yy}^\delta(Z_s) d\langle Y \rangle_s, \\
I_4 &= \int_0^t U_{xz}^\delta(Z_s) d\langle X, X^* \rangle_s + \int_0^t U_{yz}^\delta(Z_s) d\langle Y, X^* \rangle_s + \frac{1}{2} \int_0^t U_{zz}^\delta(Z_s) d\langle X^* \rangle_s.
\end{aligned}$$

Let us study the terms I_1, I_2, I_3 and I_4. We know that

$$\mathbb{E}I_1 = \mathbb{E}\int_0^t U_x^\delta(Z_s)dX_s + \mathbb{E}\int_0^t U_y^\delta(Z_s)dY_s = 0$$

because X and Y are martingales. Next,

$$I_2 = \int_0^t U_z^\delta(Z_s)dX_s^* \leqslant 0$$

because it is an integral of a nonpositive function with respect to a nondecreasing process. We have $Y = \int H dX$, so

$$\begin{aligned} I_3 &= \frac{1}{2}\int_0^t U_{xx}^\delta(Z_s)d\langle X\rangle_s + \int_0^t U_{xy}^\delta(Z_s)H_s d\langle X\rangle_s + \frac{1}{2}\int_0^t U_{yy}^\delta(Z_s)H_s^2 d\langle X\rangle_s \\ &= \frac{1}{2}\int_0^t \big(U_{xx}^\delta(Z_s) + 2U_{xy}^\delta(Z_s)H_s + U_{yy}^\delta(Z_s)H_s^2\big) d\left\langle X\right\rangle_s. \end{aligned}$$

When s is fixed, the expression under the sign of the integral is the second derivative of

$$u \mapsto U^\delta(X_s + u, Y_s + uH_s, X_s^*),$$

which is concave. Therefore,

$$U_{xx}^\delta(Z_s) + 2U_{xy}^\delta(Z_s)H_s + U_{yy}^\delta(Z_s)H_s^2 \leqslant 0$$

and the whole integral is nonpositive. Lastly, $I_4 = 0$ because X^* has finite variation. Thus, we end up with $\mathbb{E}U^\delta(X_t, Y_t, X_t^*) \leqslant U^\delta(0,0,0)$. Moreover,

$$V^\delta(x,y,z) \geqslant \mathbf{1}_{\{y \geqslant 1+\delta\}} - 2z - 8\delta,$$

so

$$\begin{aligned} \mathbb{E}U^\delta(X_t, Y_t, X_t^*) &\geqslant \mathbb{E}V^\delta(X_t, Y_t, X_t^*) \\ &\geqslant \mathbb{E}(\mathbf{1}_{\{Y_t \geqslant 1+\delta\}} - 2X_t^* - 8\delta) = \mathbb{P}(Y_t \geqslant 1+\delta) - 2\mathbb{E}X_t^* - 8\delta. \end{aligned}$$

Now let δ go to zero. We have $U^\delta(0,0,0) \to U(0,0,0) = 0$, so putting all the above facts together yields

$$\mathbb{P}(Y_t > 1) - 2\mathbb{E}X_t^* \leqslant 0.$$

To end the proof take any $\epsilon > 0$. Notice that the process $(1+\epsilon)\cdot X$ also satisfies the assumptions of Theorem 1.1 and $\int H d\big((1+\epsilon)X\big) = (1+\epsilon)\int H dX = (1+\epsilon)Y$. Therefore,

$$\mathbb{P}(Y_t \geqslant 1) \leqslant \mathbb{P}\big(Y_t(1+\epsilon) > 1\big) \leqslant 2(1+\epsilon)\mathbb{E}X_t^*.$$

Since ϵ was arbitrary, this gives $\mathbb{P}(Y_t \geqslant 1) \leqslant 2\mathbb{E}X_t^*$, and the proof of (1.4) is now complete.

In order to prove (1.5) we conduct the same reasoning as above. The Itô formula yields

$$\mathbb{E}V^\delta\big(X'_t, Y'_t, (X')^*_t\big) \leqslant \mathbb{E}U^\delta\big(X'_t, Y'_t, (X')^*_t\big) \leqslant U^\delta\big(X'_0, Y'_0, (X')^*_0\big) = U^\delta(x, y, x),$$

and letting $\delta \to 0$ gives

$$\mathbb{P}(Y'_t > 1) \leqslant 2\mathbb{E}(X')^*_t + U(x, y, x).$$

Then we take the processes $(1+\epsilon)\cdot X'$ and $(1+\epsilon)\cdot Y'$ to find out that

$$\begin{aligned}\mathbb{P}(Y'_t \geqslant 1) &\leqslant \mathbb{P}\big(Y'_t(1+\epsilon) > 1\big)\\ &\leqslant 2(1+\epsilon)\mathbb{E}(X')^*_t + U\big((1+\epsilon)x, (1+\epsilon)y, (1+\epsilon)x\big),\end{aligned}$$

which converges to $2\mathbb{E}(X')^*_t + U(x, y, x)$ as $\epsilon \to 0$. This completes the proof.

2.3. Sharpness of the inequality. Let n be an arbitrary positive integer and let W be a standard Wiener process. We define stopping times τ, σ, ρ by

$$\tau = \inf\left\{t > 0 : W_t = \frac{1}{2n}\right\}, \qquad \sigma = \inf\left\{t > 0 : W_t = -\frac{1}{2}\right\},$$
$$\rho = \inf\{t > \sigma : W_t = 0\}.$$

Let $X_t = W_{t\wedge\tau\wedge\rho}$ and let $H_t = -1$ for $0 \leqslant t \leqslant \sigma$ and $H_t = 1$ for $t > \sigma$.

The above definition implies that the pair (X, Y) evolves as follows: it starts from $(0,0)$ and moves along the line of slope -1 until X gets to $1/(2n)$ or to $-1/2$. If the first case occurs (the probability of this event is $n/(n+1)$), it stops. In the other case the pair starts moving along the line of slope 1 and stops only if it reaches $(0,1)$.

The value of $Y_t = \int_0^t H_s dX_s$ for $0 \leqslant t \leqslant \sigma$ is $Y_t = -X_t$, and for $\sigma < t$ we have $Y_t = 1 + X_t$. Hence $\mathbb{P}(Y^* \geqslant 1) = \frac{1}{n+1}$ and $\mathbb{E}X^* < \frac{1}{2n}$, so

$$\frac{\mathbb{P}(Y^* \geqslant 1)}{\mathbb{E}X^*} > \frac{2n}{n+1} \to 2.$$

This shows that the constant 2 cannot be improved.

3. GENERAL CASE

3.1. An alternative discrete-time setup. We will study the inequalities (1.6) and (1.7) in a slightly different, discrete-time context. Let $(X_n)_{n\in\mathbb{N}}$ be a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$, starting from zero and let $dX = (dX_n)_{n\geqslant 0}$ be the difference sequence of X given by $dX_0 = X_0 = 0$ and $dX_n = X_n - X_{n-1}$ for $n \geqslant 1$.

Then the one-sided maximal function of X is given by $X^* = \sup_{k\geqslant 0} X_k$, and its truncated version is defined by the equality $X_n^* = \max_{0\leqslant k\leqslant n} X_k$, $n = 0, 1, 2, \ldots$ Let $(H_n)_{n\geqslant 0}$ be a predictable sequence of variables with values in $[-1, 1]$; here by *predictability* we mean that for all n the variable H_n is $\mathcal{F}_{(n-1)\vee 0}$-measurable. A martingale $Y = (Y_n)_{n\geqslant 0}$ is the *transform* of X by H if for any $n \geqslant 0$ we have the equality $dY_n = H_n dX_n$; alternatively, Y is given by the identity $Y_n = \sum_{k=0}^n H_k dX_k$, $n = 0, 1, 2, \ldots$ We will prove that for all $n = 0, 1, 2, \ldots$

$$\mathbb{P}(Y_n \geqslant 1) \leqslant c\mathbb{E}X_n^*, \tag{3.1}$$

where c is defined in Theorem 1.2. By a usual stopping time argument, this yields

$$\mathbb{P}(Y^* \geqslant 1) \leqslant c\mathbb{E}X^*.$$

Now, there is a standard argument showing that the above bound implies (1.6): see the paper [7], in which it is shown how the results of Bichteler [5] allow the deduction of various inequalities for stochastic integrals from their counterparts in the above discrete-time setting.

3.2. Definition and properties of the special function. The reasoning is similar to that used in the preceding section, but this time the calculations will be much more elaborate. Let $V : \mathbb{R}^3 \to \mathbb{R}$ be given by $V(x, y, z) = \mathbf{1}_{\{y\geqslant 1\}} - c(x \vee z)$. We see that (3.1) is equivalent to

$$\mathbb{E}V(X_n, Y_n, X_n^*) \leqslant 0, \quad n \in \mathbb{N}.$$

Let $D = \{(x, y, z) \in \mathbb{R}^3 : x \leqslant z\}$. We divide D into five parts (see Figure 1):

$$D_1 = \{(x, y, z) \in D : y \geqslant 1 - z + x\},$$
$$D_2 = \left\{(x, y, z) \in D : y < 1 - (z - x) \text{ and } y < 1 - \frac{4}{c} + z - x\right\},$$
$$D_3 = \left\{(x, y, z) \in D : 1 - \frac{4}{c} + z - x \leqslant y < z - x \text{ and } y < 1 - \frac{8}{3c} + \frac{z - x}{3}\right\},$$
$$D_4 = \left\{(x, y, z) \in D : 3 - \frac{8}{c} - z + x \leqslant y < 1 - z + x \text{ and } y \geqslant 1 - \frac{8}{3c} + \frac{z - x}{3}\right\},$$
$$D_5 = \left\{(x, y, z) \in D : z - x \leqslant y < 3 - \frac{8}{c} - z + x\right\}.$$

We see that D_i are pairwise disjoint and their union is D.

The following special function $U : \mathbb{R}^3 \to \mathbb{R}$ will be crucial in the proof of the inequality (3.1):

$$U|_{D_1}(x, y, z) = -cz + 1,$$
$$U|_{D_2}(x, y, z) = -cz + \frac{2(z - x)}{z - x + 1 - y},$$

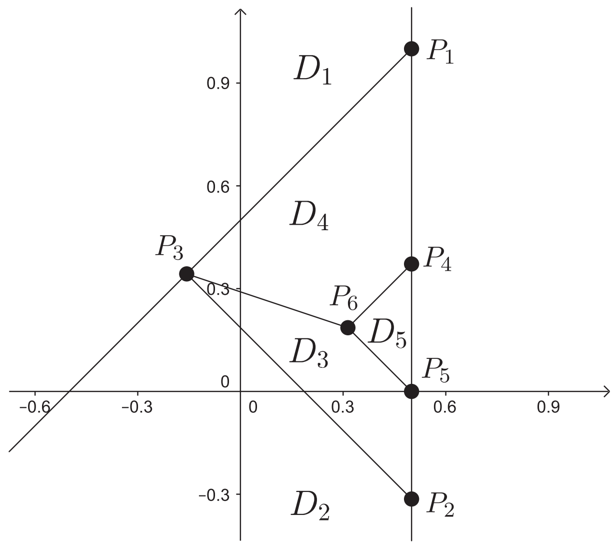

FIGURE 1. The set D intersected with the plane $z = z_0$. Coordinates of the intersection points are the following: $P_1 = (z_0, 1)$, $P_2 = (z_0, 1 - 4/c)$, $P_3 = (z_0 - 2/c, 1 - 2/c)$, $P_4 = (z_0, 3 - 8/c)$, $P_5 = (z_0, 0)$, $P_6 = (z_0 - 1.5 + 4/c, 1.5 - 4/c)$

$$\begin{aligned}
U|_{D_3}(x,y,z) &= -cz - \frac{c}{4}(z-x)\big(6c(z-x+1-y)-16\big)^{1/3} + c(z-x),\\
U|_{D_4}(x,y,z) &= -cz + \frac{1}{16}\big(8-3c(1-y-z+x)\big)^{1/3}\big(8-3c(1-y-z+x)\\
&\quad - 8c(z-x)\big) + c(z-x),\\
U|_{D_5}(x,y,z) &= -cz + \frac{c^2}{4}(6c-16)^{-2/3}\big(y^2-(z-x)^2\big)\\
&\quad - \frac{c}{4}(6c-16)^{1/3}(z-x) + c(z-x),\\
U|_{\mathbb{R}^3\setminus D}(x,y,z) &= U(x,y,x).
\end{aligned}$$

LEMMA 3.1. *The function U satisfies the following conditions*:

1. *U is continuous.*
2. *$U(x,y,z) = U(x+t, y, z+t) + ct$ for all $(x,y,z) \in \mathbb{R}^3$ and $t \in \mathbb{R}$.*
3. *For all $(x,y,z) \in D$ and $|H| \leqslant 1$ the function $f_{x,y,z,H} : (-\infty, z-x] \to \mathbb{R}$ given by*
$$f_{x,y,z,H}(t) = U(x+t, y+Ht, z)$$
is concave.
4. *$U(x,y,z) \geqslant V(x,y,z)$ for all $(x,y,z) \in D$.*

P r o o f. Essentially, the proof requires only some straightforward calculation.

1. It is clear that U is continuous in the interior of each D_i, so it suffices to check that the formulas match each other on the boundaries of D_i's. One easily verifies that this is indeed the case.

2. It is easy to see that U is a sum of $-cz$ and some function depending only on $x-z$ and y.

3. Notice that U is smooth in the interior of each D_i and

$$f' = U'_x + HU'_y, \quad f'' = U''_{xx} + 2HU''_{xy} + H^2U''_{yy}.$$

First we will check that $f''(t) \leqslant 0$ if $(x+t, y+Ht, z)$ belongs to the interior of some D_i. Since $f_{x,y,z,H}(t+s) = f_{x+t,y+Ht,z,H}(s)$ for all s, t, it suffices to check the sign of $f''(0)$. If $(x,y,z) \in \operatorname{int} D_1$, then $f''(0) = 0$ since f is locally constant. If $(x,y,z) \in \operatorname{int} D_2$, then

$$f''(0) = \frac{4(1+H)\big(H(z-x)-(1-y)\big)}{(z-x+1-y)^3}.$$

It is nonpositive because $H(z-x)-(1-y) \leqslant z-x-(1-y) < 0$. If $(x,y,z) \in \operatorname{int} D_3$, then

$$f''(0) = 2c^2(1+H)\frac{(H-2)c(z-x)-3c(1-y)+8}{\big(6c(z-x+1-y)-16\big)^{5/3}}.$$

It is not greater than zero since $1-y > \big(8-c(z-x)\big)/(3c)$. If $(x,y,z) \in \operatorname{int} D_4$, then

$$f''(0) = \frac{c^2}{4}(H-1)\frac{(H+3)\big(8-3c(1-y)\big)+(7H+5)c(z-x)}{\big(8-3c(1-y-z+x)\big)^{5/3}}.$$

In the interior of D_4 we have $1-y < \big(8-c(z-x)\big)/(3c)$, so this derivative is nonpositive. Finally, if $(x,y,z) \in \operatorname{int} D_5$, then

$$f''(0) = c^2/2(6c-16)^{-2/3}(H^2-1) \leqslant 0.$$

To complete the proof of concavity of f, we have to check that

$$f'_{x,y,z,H}(0^-) \geqslant f'_{x,y,z,H}(0^+)$$

for each (x,y,z) that lies on a common boundary of D_i and D_j, $i \neq j$. On the boundary between D_1 and D_2 we have

$$f'(0^+) = \frac{2\big(H(z-x)-(1-y)\big)}{(z-x+1-y)^2} \leqslant 0 = f'(0^-).$$

On the boundary between D_2 and D_3, given by $1-y = x-z+4/c$, we have

$$f'(0^-) = \frac{c}{8}\big((1+H)c(z-x)-4\big) = f'(0^+).$$

Between the sets D_3 and D_4 we have

$$f'(0) = \frac{c}{2}\big(4c(z-x)\big)^{-2/3}(3+H)c(z-x) - c,$$

and between D_1 and D_4 the right derivative is

$$f'(0^+) = \frac{c}{16}\big(2(1-H)c(z-x) + (1+H)8\big) - c \leqslant 0 = f'(0^-).$$

Finally, when we consider the set D_5, which borders only with D_3 and D_4, we calculate that

$$f'(0) = \frac{c}{2}(6c-16)^{-2/3}\big((1+H)c(z-x) + 3c - 8\big)$$

on the boundary with D_3 and

$$f'(0) = \frac{c}{2}(6c-16)^{-2/3}\big((1+H)(3c-8) + (1-H)c(z-x)\big)$$

on the boundary with D_4.

4. In the previous point we have proved that U is nonincreasing in x. Therefore, it is enough to check the inequality $V(x,y,z) \leqslant U(x,y,z)$ for $x = z$. So, for $y < 0$ we have $U(z,y,z) = -cz = V(z,y,z)$. For $y \in [0, 3-8/c)$ we have

$$U(z,y,z) = \frac{c^2}{4}(6c-16)^{-2/3}y^2 - cz \geqslant -cz = V(z,y,z).$$

For $3 - 8/c \leqslant y < 1$ it is true that

$$U(z,y,z) = \frac{1}{16}\big(8-3c(1-y)\big)^{4/3} - cz \geqslant \frac{1}{16}(6c-16)^{4/3} - cz \geqslant -cz = V(z,y,z).$$

Finally, for $y \geqslant 1$ we have $U(z,y,z) = 1 - cz = V(z,y,z)$. ∎

LEMMA 3.2. *Take any* $(x,y,z) \in D$ *and* $|H| \leqslant 1$. *The function* $f_{x,y,z,H}\colon \mathbb{R} \to \mathbb{R}$ *given by* $f_{x,y,z,H}(t) = U(x+t, y+Ht, z)$ *has the following property*:

$$f_{x,y,z,H}(0) + tf'_{x,y,z,H}(0^-) \geqslant f_{x,y,z,H}(t) \tag{3.2}$$

for all $t \in \mathbb{R}$.

Proof. Reduction 1. If $t \leqslant z - x$, then (3.2) follows trivially from condition 4 in Lemma 3.1. Indeed, setting $f := f_{x,y,z,H}$, we know that f' is nonincreasing (on the set where f' is well defined), so for $t > 0$ we have

$$f(t) - f(0) = \int_0^t f'(s)\,ds \leqslant \int_0^t f'(0^-)\,ds = tf'(0^-).$$

Similarly, for $t > 0$ we have

$$f(0) - f(-t) = \int_{-t}^{0} f'(s)ds \geqslant \int_{-t}^{0} f'(0^-)ds = tf'(0^-).$$

Therefore, the nontrivial part of Lemma 3.2 is for $t > z - x$.

R e d u c t i o n 2. It is enough to prove (3.2) for $z = x$. Indeed, if $z > x$ and $t > z - x$, then

$$\begin{aligned} f_{x,y,z,H}(0) &+ tf'_{x,y,z,H}(0^-) \\ &= f_{x,y,z,H}(0) + (z-x)f'_{x,y,z,H}(0^-) + \big(t - (z-x)\big)f'_{x,y,z,H}(0^-) \\ &\geqslant f_{x,y,z,H}(z-x) + \big(t-(z-x)\big)f'_{z,y+H(z-x),z,H}(0^-) \\ &= f_{z,y+H(z-x),z,H}(0) + \big(t-(z-x)\big)f'_{z,y+H(z-x),z,H}(0^-), \end{aligned}$$

so it is enough to prove $f(0) + tf'(0^-) \geqslant f(t)$ for $z = x$.

R e d u c t i o n 3. Condition 2 in Lemma 3.1 tells us that it is enough to study the case $x = z = 0$. We put $f = f_{0,y,0,H}$, $B(y) = U(0, y, 0)$ and $b_H(y) = f'(0^-)$. We have

$$\begin{aligned} f(t) &= U(t, y + Ht, 0) = U(t, y + Ht, t) = U(0, y + Ht, 0) - ct \\ &= B(y + Ht) - ct. \end{aligned}$$

With this notation, (3.2) is equivalent to

$$t\big(b_H(y) + c\big) \geqslant B(y + Ht) - B(y). \tag{3.3}$$

The formulas for B and b_H are as follows:

$$B(y) = \begin{cases} 1 & \text{for } y \geqslant 1, \\ \frac{1}{16}\big(8 - 3c(1-y)\big)^{4/3} & \text{for } 3 - \frac{8}{c} \leqslant y < 1, \\ \frac{c^2}{4}(6c-16)^{-2/3}y^2 & \text{for } 0 \leqslant y < 3 - \frac{8}{c}, \\ 0 & \text{for } y < 0 \end{cases}$$

and

$$b_H(y) = \begin{cases} 0 & \text{for } y \geqslant 1, \\ \frac{c}{4}\big(8 - 3c(1-y)\big)^{1/3}(1+H) - c & \text{for } 3 - \frac{8}{c} \leqslant y < 1, \\ \frac{c^2}{2}(6c-16)^{-2/3}Hy + \frac{c}{4}(6c-16)^{1/3} - c & \text{for } 0 \leqslant y < 3 - \frac{8}{c}, \\ \frac{c}{4}\big(6c(1-y) - 16\big)^{1/3} - c & \text{for } 1 - \frac{4}{c} \leqslant y < 0, \\ -\frac{2}{1-y} & \text{for } y < 1 - \frac{4}{c}. \end{cases}$$

R e d u c t i o n 4. It suffices to show (3.3) for $H = 1$. Notice that B is nondecreasing, therefore $B(y + Ht) - B(y) \leqslant 0$ for $H \leqslant 0$. Moreover, notice that

$b_H(y) \geqslant -c$ for all y. Hence for $H \leqslant 0$ we can write $t\big(b_H(y)+c\big) \geqslant 0 \geqslant B(y+Ht) - B(y)$. Now suppose we have successfully proved the estimate

$$t\big(b_1(y)+c\big) \geqslant B(y+t) - B(y) \tag{3.4}$$

(which is (3.3) for $H = 1$) and take any $H \in (0, 1]$. We can check that for each y the inequality $H\big(b_1(y)+c\big) \leqslant b_H(y)+c$ holds. This implies

$$\frac{t}{H}\big(b_H(y)+c\big) \geqslant t\big(b_1(y)+c\big) \geqslant B(y+t) - B(y),$$

and by replacing t with Ht we obtain $t\big(b_H(y)+c\big) \geqslant B(y+Ht) - B(y)$. The above reasoning justifies that it is enough to show (3.3) for $H = 1$.

The rest of the proof of Lemma 3.2 relies on checking the validity of (3.4) for all y. This will be done in four cases:

1. $y \geqslant 1$,
2. $3 - \frac{8}{c} \leqslant y < 1$,
3. $0 \leqslant y < 3 - \frac{8}{c}$,
4. $y < 0$.

There will also be subcases depending on which of the intervals contains the value of $y + t$.

Case 1. For $y \geqslant 1$ we have $t\big(b_1(y)+c\big) = tc > 0 = 1 - 1 = B(y+t) - B(y)$.

Case 2. For $3 - 8/c \leqslant y < 1$ we have two subcases: $y + t \geqslant 1$ or $y + t < 1$. Suppose $y + t < 1$. For $t = 0$ both sides of (3.4) are equal to zero, so it suffices to compare the derivatives with respect to t. We have

$$\Big(t\big(b_1(y)+c\big)\Big)' = b_1(y) + c = \frac{c}{2}\big(8 - 3c(1-y)\big)^{1/3} \geqslant \frac{c}{2}$$

(the last inequality is due to $y \geqslant 3 - 8/c$ and $c > 17/6$) and, on the other hand,

$$\big(B(y+t) - B(y)\big)' = \frac{c}{4}\big(8 - 3c(1-y-t)\big)^{1/3} \leqslant \frac{c}{2},$$

so the inequality is proven. Now consider the other subcase, where $y + t \geqslant 1$. We have

$$B(y+t) - B(y) = 1 - B(y) = \sup_{s \in (0, 1-y)} B(y+s) - B(y),$$

so this subcase follows trivially from the previous one.

Case 3. For $0 \leqslant y < 3 - 8/c$ we have three subcases: $y + t \geqslant 1$ or $y + t \in [3 - 8/c, 1)$ or $y + t < 3 - 8/c$. The first subcase, where $y + t \geqslant 1$, is trivial and follows from the second one in a similar fashion to Case 2. Assume that we have $y + t \in [3 - 8/c, 1)$. Then the inequality (3.4) is equivalent to

$$\frac{c^2}{2}(6c-16)^{-2/3}\left(t\left(3 - \frac{8}{c} + y\right) + \frac{y^2}{2}\right) \geqslant \frac{1}{16}\big(8 - 3c(1-y-t)\big)^{4/3}. \tag{3.5}$$

The proof of this inequality employs the following strategy of checking the values on the boundary and differentiating: for each t we will check that the inequality is true for the smallest possible y (which is either $y = 3 - 8/c - t$ for $t < 3 - 8/c$ or $y = 0$ for $t \in [3 - 8/c, 1)$) and then we will compare the derivatives of both sides with respect to y. For $t < 3 - 8/c$ and $y = 3 - 8/c - t$ this inequality is equivalent to

$$\frac{c^2}{2}(6c-16)^{-2/3}\left(\left(3-\frac{8}{c}\right)^2-\frac{y^2}{2}\right) \geqslant \frac{1}{16}(6c-16)^{4/3},$$

which is equivalent to $3 - 8/c \geqslant y$, and this is true. For $y = 0$ and $t \in [3 - 8/c, 1)$ the inequality is

$$\frac{c^2 t}{2}(6c-16)^{-2/3}\left(3-\frac{8}{c}\right) \geqslant \frac{1}{16}\big(8-3c(1-t)\big)^{4/3}.$$

Notice that the function on the left is linear and the function on the right is convex. Therefore, it is enough to check if the inequality holds at the ends of the interval (namely, in $t = 3 - 8/c$ and $t = 1$) to know that it is true on the whole interval. We have already checked $t = 3 - 8/c$, and for $t = 1$ the inequality is equivalent to

$$\frac{c}{4}(6c-16)^{1/3} \geqslant 1$$

or

$$3c^4 - 8c^3 - 32 \geqslant 0,$$

which is true from the definition of c. This is actually the only point of the whole proof when we use the exact value of c. Now we derive that the y-derivatives of both sides of (3.5) are given by

$$\frac{d}{dy}\frac{c^2}{2}(6c-16)^{-2/3}\left(t\left(3-\frac{8}{c}+y\right)+\frac{y^2}{2}\right) = \frac{c^2}{2}(6c-16)^{-2/3}(t+y),$$

$$\frac{d}{dy}\frac{1}{16}\big(8-3c(1-y-t)\big)^{4/3} = \frac{c}{4}\big(8-3c(1-y-t)\big)^{1/3}.$$

Let us put $u = y + t \in [3 - 8/c, 1)$. As we have announced above, we will be done if we show that

$$\frac{c^2 u}{2}(6c-16)^{-2/3} \geqslant \frac{c}{4}\big(8-3c(1-u)\big)^{1/3}.$$

It is easy to check that this inequality holds for $u = 3 - 8/c$, and by differentiating with respect to u we obtain

$$\frac{c^2}{2}(6c-16)^{-2/3}$$

on the left and

$$\frac{c^2}{4}\big(8-3c(1-u)\big)^{-2/3}$$

on the right. But we easily see that the inequality

$$\frac{c^2}{2}(6c-16)^{-2/3} \geqslant \frac{c^2}{4}\big(8-3c(1-u)\big)^{-2/3}$$

is true for $u \geqslant 3-8/c$, and this ends the proof of the subcase. For the third subcase assume that $y+t<3-8/c$. The inequality (3.4) reads

$$\frac{c^2}{2}(6c-16)^{-2/3}\left(t\left(3-\frac{8}{c}\right)+ty+\frac{y^2}{2}-\frac{(y+t)^2}{2}\right) \geqslant 0,$$

which is equivalent to $3-8/c \geqslant t/2$. This is clearly true since both y and t are nonnegative, and thus $y+t \geqslant t/2$.

C a s e 4. For $y<0$ and $y+t \leqslant 0$ the inequality is trivial:

$$t\big(b_1(y)+c\big) > 0 = 0-0 = B(y+t)-B(y).$$

If $y<0$ and $y+t>0$, consider the following reasoning. Suppose that

$$t_2\big(b_1(y+t_1)+c\big) \geqslant B(y+t_1+t_2)-B(y+t_1)$$

and

$$t_1\big(b_1(y)+c\big) \geqslant B(y+t_1)-B(y)$$

are true for some $t_1, t_2 > 0$. Then, since B is nondecreasing, we may write

$$\begin{aligned}
(t_1+t_2)\big(b_1(y)+c\big) &\geqslant t_2\big(b_1(y)+c\big)+B(y+t_1)-B(y)\\
&= t_2\big(b_1(y+t_1)+c\big)+B(y+t_1)-B(y)+t_2\big(b_1(y)-b_1(y+t_1)\big)\\
&\geqslant B(y+t_1+t_2)-B(y)+t_2\big(b_1(y)-b_1(y+t_1)\big)\\
&\geqslant B(y+t_1+t_2)-B(y)
\end{aligned}$$

provided that $b_1(y) \geqslant b_1(y+t_1)$. Notice that the least value of b_1 is at $y=0$:

$$\inf b_1(y) = b_1(0) = \frac{c}{4}(6c-16)^{1/3}-c;$$

so if we put $t_1=-y$ and $t_2=t+y$ in the above inequality, then we will end up with $t\big(b_1(y)+c\big) \geqslant B(t+y)-B(y)$. Since we have already dealt with all the cases where $y=0$ and $t>0$, this completes the whole proof. ■

3.3. Proof of (3.1) and its extension. We have

$$\begin{aligned}
&\mathbb{E}\big(U(X_{n+1},Y_{n+1},X^*_{n+1})|\mathcal{F}_n\big)\\
&\quad=\mathbb{E}\Big(U\big(X_n+dX_{n+1},Y_n+H_{n+1}dX_{n+1},X^*_n\vee(X_n+dX_{n+1})\big)|\mathcal{F}_n\Big)\\
&\quad=\mathbb{E}\big(f_{X_n,Y_n,X^*_n,H_{n+1}}(dX_{n+1})|\mathcal{F}_n\big)\\
&\quad\leqslant\mathbb{E}\big(f_{X_n,Y_n,X^*_n,H_{n+1}}(0)+f'_{X_n,Y_n,X^*_n,H_{n+1}}(0^-)dX_{n+1}|\mathcal{F}_n\big)\\
&\quad=f_{X_n,Y_n,X^*_n,H_{n+1}}(0)=U(X_n,Y_n,X^*_n).
\end{aligned}$$

We take the expected value of both sides and infer that the sequence

$$\mathbb{E}U(X_n,Y_n,X^*_n),\quad n=0,1,2,\ldots,$$

is nonincreasing. Therefore,

$$\mathbb{E}V(X_n,Y_n,X^*_n)\leqslant\mathbb{E}U(X_n,Y_n,X^*_n)\leqslant\mathbb{E}U(X_0,Y_0,X^*_0)=U(0,0,0)=0.$$

This completes the proof of (3.1). Observe that if we set $X'=x+X$ and $Y'=y+Y$, then an analogous reasoning shows that

$$\mathbb{E}V(x+X_n,y+Y_n,x+X^*_n)\leqslant\mathbb{E}U(x+X_n,y+Y_n,x+X^*_n)\leqslant U(x,y,x)$$

or

$$\mathbb{P}\big((Y')_n\geqslant 1\big)\leqslant c\mathbb{E}(X')^*_n+U(x,y,x).$$

3.4. Sharpness of the inequality. Finally, let us address the optimality of the constant c. Take any $K,M,N\in\mathbb{N}$. Let us put $\delta=(4-c)/(2cN)$ and $\eta=K\delta$. Introduce the function $G:[3-8/c,1]\to\mathbb{R}^2$ by the formula

$$G(y)=\left(\frac{3c(1-y)-8}{4c},\frac{c(3+y)-8}{4c}\right).$$

Observe that $G(y_0)$ is the intersection point of the line given by $y-x=y_0$ with the line $y=(3c-8-cx)/(3c)$. Notice that had we put zero on the third coordinate, then the image of G would be the boundary between D_3 and D_4 intersected with the plane $\{(x,y,0):x,y\in\mathbb{R}\}$. Moreover, $G^{-1}(x,y)=y-x$. We define a few subsets of $\mathbb{R}^2$:

$$\begin{aligned}
A_0&=\{(0,0),(-2\eta,-2\eta)\},\\
A_1&=\left\{(-2\delta,y-2\delta):3-\frac{8}{c}<y<1\right\},\\
A_2&=\left\{(0,y):3-\frac{8}{c}<y<1\right\},\\
A_3&=\left\{G(y)+(\delta,\delta):3-\frac{8}{c}<y<1\right\},
\end{aligned}$$

$$A_4 = \left\{ G(y) : 3 - \frac{8}{c} < y < 1 \right\},$$
$$A_5 = \{(M, y) : y < 0\} \cup \{(0, y) : y < 0\} \cup \{(0,1), (1,1), (-M, 1-M)\},$$
$$A_6 = \{(x, 1+x) : -1 < x < 0\}.$$

Consider a discrete-time Markov martingale $(X_n, Y_n)_{n \in \mathbb{N}}$ with values in the union of the above sets, whose distribution is uniquely determined by the following conditions:

1. $(X_0, Y_0) = 0$.
2. $(X_1, Y_1) = (-2\eta, -2\eta)$ or $(X_1, Y_1) = (1, 1)$.
3. If $(X_1, Y_1) = (-2\eta, -2\eta)$, then $(X_2, Y_2) = (0, -4\eta)$ or $(X_2, Y_2) = (\delta, \delta) + G(3 - 8/c + 8\eta + 4\delta)$.
4. If $(X_n, Y_n) \in A_1$, then $(X_{n+1}, Y_{n+1}) = (0, Y_n - X_n)$ or $(X_{n+1}, Y_{n+1}) = G(Y_n - X_n)$.
5. If $(X_n, Y_n) \in A_2$, then $(X_{n+1}, Y_{n+1}) = (-2\delta, Y_n + 2\delta)$ or $(X_{n+1}, Y_{n+1}) = (M, Y_n - M)$.
6. If $(X_n, Y_n) \in A_3$, then $(X_{n+1}, Y_{n+1}) = G\,(Y_n - X_n)$ or $(X_{n+1}, Y_{n+1}) = (0, Y_n - X_n)$.
7. If $(X_n, Y_n) \in A_4$, then (X_{n+1}, Y_{n+1}) is either $G\,(Y_n - X_n + 4\delta) + (\delta, \delta)$ or $(0, X_n + Y_n)$.
8. If $(X_n, Y_n) \in A_5$, then $(X_{n+1}, Y_{n+1}) = (X_n, Y_n)$.
9. If $(X_n, Y_n) \in A_6$, then (X_{n+1}, Y_{n+1}) is either $(0, 1)$ or $(-M, 1 - M)$.

Although we do not write the formulas for H_n explicitly, it is clear that Y has the form $Y = \sum H dX$ for some H satisfying the assumptions of (3.1). We do not specify the transition probabilities, as they are uniquely determined by (X, Y) being a martingale.

This martingale is constructed in such a way (see Figure 2) that in every step the pair (X, Y) moves along such a line of slope 1 or -1 that $U(\cdot, \cdot, 0)$ is nearly linear on this line. This will allow us to bound the difference

$$U(X_n, Y_n, X_n^*) - \mathbb{E}\big(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n\big)$$

(which we already proved to be nonnegative) by an expression proportional to δ^2.

The difference $Y_{2n} - X_{2n}$ is increasing unless we reach an absorbing state. Indeed, notice that after the first two steps, if (X_{2n}, Y_{2n}) is in A_1 or in A_3, then (X_{2n+1}, Y_{2n+1}) is in A_2 or in A_4, and $Y_{2n} - X_{2n} = Y_{2n+1} - X_{2n+1}$. Moreover, after the first two steps, if (X_{2n-1}, Y_{2n-1}) is in A_2 or A_4, then (X_{2n}, Y_{2n}) either goes directly to A_5 and reaches an absorbing state or (X_{2n}, Y_{2n}) is in A_1 or A_3 and $Y_{2n} - X_{2n} = Y_{2n-1} - X_{2n-1} + 4\delta$. The difference $Y_n - X_n$ is at most one provided that $(X_n, Y_n) \notin A_5$. Actually, $Y_{2n} - X_{2n} = 3 - 8/c + (4n + 8K)\delta$ provided that $(X_{2n}, Y_{2n}) \notin A_5$. Therefore, $(X_{2N}, Y_{2N}) \in A_5$, which means that this maringale stops after at most $2N$ steps.

Now we will find a bound on $U(X_n, Y_n, X_n^*) - \mathbb{E}\big(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n\big)$.

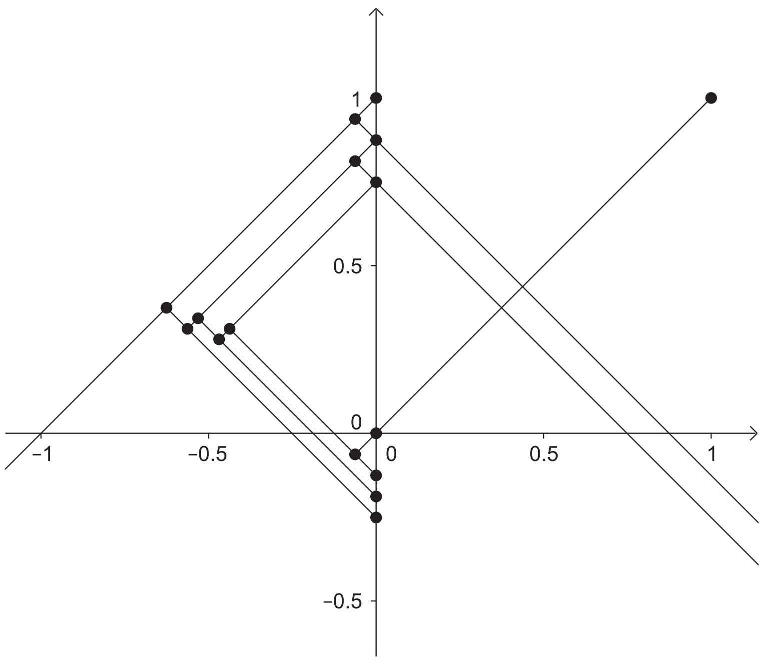

FIGURE 2. The pair (X, Y) moves along the black lines

C a s e 1. $(X_n, Y_n) \in A_1$. We have

$$U(X_n, Y_n, X_n^*) = \mathbb{E}\big(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n\big)$$

because $X_n^* = X_{n+1}^* = 0$ and $U(\cdot,\cdot,0)$ is linear on the interval from $(0, Y_n - X_n)$ to $G(Y_n - X_n)$.

C a s e 2. $(X_n, Y_n) \in A_2$. We will need a simple auxiliary inequality: if $a > 0$, then for sufficiently small $\epsilon > 0$ the following inequalities are true:

$$\begin{aligned} 6a^2\epsilon^2 &\geqslant 6a^2\epsilon^2 - 8a\epsilon^3 + 3\epsilon^4 = a^4 - (a+3\epsilon)(a-\epsilon)^3 \\ &= \big(a^{4/3} - (a+3\epsilon)^{1/3}(a-\epsilon)\big) \\ &\quad \times \big(a^{8/3} + a^{4/3}(a+3\epsilon)^{1/3}(a-\epsilon) + (a+3\epsilon)^{2/3}(a-\epsilon)^2\big) \\ &\geqslant \big(a^{4/3} - (a+3\epsilon)^{1/3}(a-\epsilon)\big) \cdot 2a^{8/3}. \end{aligned}$$

Therefore,

$$a^{4/3} - (a+3\epsilon)^{1/3}(a-\epsilon) \leqslant 3a^{-2/3}\epsilon^2. \tag{3.6}$$

If $(X_n, Y_n) = (x, y) \in A_2$, then

$$U(X_n, Y_n, X_n^*) = U(0, y, 0) = \frac{1}{16}\big(8 - 3c(1-y)\big)^{4/3}$$

and

$$\begin{aligned}
&\mathbb{E}\big(U(X_{n+1},Y_{n+1},X^*_{n+1})|X_n=x,Y_n=y\big)\\
&=\frac{2\delta}{M+2\delta}U(M,y-M,M)+\frac{M}{M+2\delta}U(-2\delta,y+2\delta,0)\\
&=\frac{M\Big(\frac{1}{16}\big(8-3c(1-y)+12c\delta\big)^{1/3}\big(8-3c(1-y)-4c\delta\big)+2c\delta\Big)-2cM\delta}{M+2\delta}.
\end{aligned}$$

To simplify things, we put $a=8-3c(1-y)$ and $\epsilon=4c\delta$. Then

$$\begin{aligned}
&\mathbb{E}\big(U(X_n,Y_n,X^*_n)-U(X_{n+1},Y_{n+1},X^*_{n+1})|X_n=x,Y_n=y\big)\\
&=\frac{1}{16}a^{4/3}+\frac{2cM\delta}{M+2\delta}+\frac{M}{M+2\delta}\cdot\frac{1}{16}(a+3\epsilon)^{1/3}(a-\epsilon)-\frac{M}{M+2\delta}\cdot 2c\delta\\
&=\frac{M}{M+2\delta}\frac{1}{16}(a+3\epsilon)^{1/3}(a-\epsilon)\leqslant\frac{1}{16}3a^{-2/3}\epsilon^2\leqslant 3(6c-16)^{-2/3}\delta^2<20\delta^2.
\end{aligned}$$

C a s e 3. $(X_n,Y_n)\in A_3$. Here the reasoning is analogous to that of Case 1.

C a s e 4. $(X_n,Y_n)\in A_4$. This case relies solely on the analysis of the derivative of U along a certain line. For $y_0\in(3-8/c,1)$ and $(x_1,y_1)=G(y_0)$ we define $f:[x_1-2\delta,0]\to\mathbb{R}$ by $f(x)=U(x,x_1+y_1-x,0)$. Explicit values of (x_1,y_1) are:

$$x_1=\frac{3(1-y_0)}{4}-\frac{2}{c},\qquad y_1=\frac{3+y_0}{4}-\frac{2}{c},\qquad f(x)=U\left(x,\frac{3-y_0}{2}-\frac{4}{c}-x\right).$$

We have $f'=U'_x-U'_y$, which implies

$$f'(x)=\frac{c}{4}\big(8-3c(1-y_0)\big)^{1/3}-c\qquad\text{for } x_1\leqslant x<0$$

and

$$f'(x_1-2\delta)=\frac{c}{4}\big(8-3c(1-y_0)+12c\delta\big)^{1/3}\cdot\left(1-\frac{4c\delta}{8-3c(1-y_0)+12c\delta}\right)-c.$$

From property 4 in Lemma 3.1 we know that f is concave, and so f' is nonincreasing, therefore on the interval $(x_1-2\delta,x_1)$ we have $f'(x)\leqslant f'(x_1-2\delta)$. This means that for $x\in(x_1-2\delta,x_1)$ we have

$$\begin{aligned}
0&\leqslant f'(x)-f'(x_1)\leqslant f'(x_1-2\delta)-f'(x_1)\\
&=\frac{c}{4}\big(8-3c(1-y_0)+12c\delta\big)^{1/3}\cdot\left(1-\frac{4c\delta}{8-3c(1-y_0)+12c\delta}\right)\\
&\quad-\frac{c}{4}\big(8-3c(1-y_0)\big)^{1/3}\\
&\leqslant\frac{c}{4}\Big(\big(8-3c(1-y_0)+12c\delta\big)^{1/3}-\big(8-3c(1-y_0)\big)^{1/3}\Big)\\
&\leqslant\frac{c}{4}\cdot\frac{1}{3}\big(8-3c(1-y_0)\big)^{-2/3}\cdot 12c\delta\leqslant c^2(6c-16)^{-2/3}\delta<7\delta.
\end{aligned}$$

We are now ready to write the inequality (assume $(X_n, Y_n) = (x_1, y_1) \in A_4$):

$$
\begin{aligned}
&\mathbb{E}\big(U(X_n, Y_n, X_n^*) - U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|X_n = x_1, Y_n = y_1\big)\\
&= U(x_1, y_1, 0) - \frac{-x_1}{2\delta - x_1} U(x_1 - 2\delta, y_1 + 2\delta, 0) - \frac{2\delta}{2\delta - x_1} U(0, x_1 + y_1, 0)\\
&= f(x_1) - \frac{-x_1}{-x_1 + 2\delta} f(x_1 - 2\delta)\\
&= f(x_1) - \frac{-x_1}{-x_1 + 2\delta}\big(f(0) - \int_{x_1-2\delta}^{0} f'(x)dx\big)\\
&= f(x_1) + \frac{-x_1}{-x_1 + 2\delta} \int_{x_1-2\delta}^{0} f'(x)dx\\
&= f(x_1) + \frac{-x_1}{-x_1 + 2\delta}\big(\int_{x_1}^{0} f'(x)dx + \int_{x_1-2\delta}^{x_1} f'(x)dx\big)\\
&= f(x_1) + \frac{-x_1}{-x_1 + 2\delta}\Big(-x_1 f'(x_1) + \int_{x_1-2\delta}^{x_1} \big(f'(x) - f'(x_1)\big)dx + 2\delta f'(x_1)\Big)\\
&= f(x_1) - x_1 f'(x_1) + \frac{-x_1}{-x_1 + 2\delta} \int_{x_1-2\delta}^{x_1} \big(f'(x) - f'(x_1)\big)dx < 2\delta \cdot 7\delta = 14\delta^2.
\end{aligned}
$$

C a s e 5. $(X_n, Y_n) \in A_5$. There is nothing to prove since (X, Y) is already in an absorbing state.

C a s e 6. $(X_n, Y_n) \in A_6$. Here the reasoning is analogous to that of Case 1.

C a s e 7. $(X_n, Y_n) \in A_0$. This only happens for $n \in \{0, 1\}$, and these are actually two separate subcases. Assume $n = 0$. Then $(X_0, Y_0) = (0, 0)$ and

$$
\begin{aligned}
U(0,0,0) - \mathbb{E}U(X_1, Y_1, X_1^*) &= -\frac{2\eta}{1+2\eta} U(1,1) - \frac{1}{1+2\eta} U(-2\eta, -2\eta)\\
&= \frac{2\eta(c-1)}{1+2\eta} + \frac{1}{1+2\eta}\left(\frac{c\eta}{2}(6c - 16 + 24c\eta)^{1/3} - 2c\eta\right)\\
&= \frac{2\eta}{1+2\eta}\left(\frac{c}{4}(6c - 16 + 24c\eta)^{1/3} - 1\right)\\
&\leqslant \frac{2\eta}{1+2\eta}\left(\frac{c}{4} \cdot \frac{1}{3}(6c-16)^{-2/3} \cdot 24c\eta + \frac{c}{4}(6c-16)^{1/3} - 1\right)\\
&= \frac{2\eta}{1+2\eta}\big(2c^2\eta(6c-16)^{-2/3}\big) < 25\eta^2.
\end{aligned}
$$

Now let us assume $n = 1$ and $(X_1, Y_1) = (-2\eta, -2\eta)$. The step from (X_1, Y_1) to (X_2, Y_2) can be divided into two substeps: firstly (X, Y) goes to the point $G(3 - 8/c + 8\eta)$ or to $(0, -4\eta)$ and secondly it either performs the routine for $(X_n, Y_n) \in A_4$ if it went to $G(3 - 8/c + 8\eta)$ or for $(X_n, Y_n) \in A_5$ if it went to

$(0, -4\eta)$. The first substep is trivial (analogously to Case 1) and the second substep is something we have already dealt with in Cases 4 and 5.

All the above cases prove that $U(X_n, Y_n, X_n^*) - \mathbb{E}\big(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n\big) < 20\delta^2$ except for the case $n = 0$, where we bounded $U(0,0,0) - \mathbb{E}U(X_1, Y_1, X_1^*)$ by $25\eta^2$. This means that

$$\begin{aligned}\mathbb{E}U(X_{2N}, Y_{2N}, X_{2N}^*) &> -2N \cdot 20\delta^2 - 25\eta^2 = -40\frac{(4-c)^2}{4c^2N} - 25\eta^2 \\ &> -2/N - 25\eta^2.\end{aligned}$$

Notice that $(X_{2N}, Y_{2N}) \in A_5$ and for any point $(x,y) \in A_5$ we have $V(x,y,0) = U(x,y,0)$ with the only exception of $(-M, 1-M)$, where $V(-M, 1-M, 0) = 0 < 1 = U(-M, 1-M, 0)$. But we have $\mathbb{P}\big((X_{2N}, Y_{2N}) = (-M, 1-M)\big) < 1/M$, so

$$\mathbb{E}V(X_{2N}, Y_{2N}, X_{2N}^*) > \mathbb{E}U(X_{2N}, Y_{2N}, X_{2N}^*) - 1/M > -2/N - 1/M - 25\eta^2$$

or

$$\mathbb{P}(Y_{2N} \geqslant 1) - c\mathbb{E}X_{2N}^* > -\frac{2}{N} - \frac{1}{M} - 25\eta^2. \tag{3.7}$$

Notice that in the first step (X, Y) jumps to $(1,1)$ with probability $2\eta/(1+2\eta)$, so we have

$$\mathbb{P}(Y_{2N} \geqslant 1) \geqslant \frac{2\eta}{1+2\eta}. \tag{3.8}$$

Both sides of (3.7) are negative and both sides of (3.8) are positive, so we can divide the inequalities to obtain

$$1 - \frac{c\mathbb{E}X_{2N}^*}{\mathbb{P}(Y_{2N} \geqslant 1)} > -\frac{1+2\eta}{2\eta}\left(\frac{2}{N} + \frac{1}{M}\right) - \frac{1+2\eta}{2} \cdot 25\eta. \tag{3.9}$$

Take $N = M = K^2$ and let K go to infinity. The number η is of order $K/N = 1/K$, and hence it is clear that the right-hand side above converges to zero. Therefore, for any $\epsilon > 0$ there exist K, M, N such that

$$1 - \frac{c\mathbb{E}X_{2N}^*}{\mathbb{P}(Y_{2N} \geqslant 1)} > -\epsilon$$

or, equivalently,

$$\mathbb{P}(Y_{2N} \geqslant 1) > \frac{c}{1+\epsilon}\mathbb{E}X_{2N}^*.$$

This shows that the constant c cannot be improved.

Acknowledgments. I am sincerely grateful to Adam Osękowski, who introduced me to Burkholder's method of proving martingale inequalities, for his guidance and patience.

REFERENCES

[1] R. Bañuelos and K. Bogdan, *Lévy processes and Fourier multipliers*, J. Funct. Anal. 250 (2007), pp. 197–212.

[2] R. Bañuelos and A. Osękowski, *Martingales and sharp bounds for Fourier multipliers*, Ann. Acad. Sci. Fenn. Math. 37 (2012), pp. 251–263.

[3] R. Bañuelos and A. Osękowski, *On Astala's theorem for martingales and Fourier multipliers*, Adv. Math. 283 (2015), pp. 275–302.

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

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

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

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

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

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

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

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

[12] A. Osękowski, *Sharp Martingale and Semimartingale Inequalities*, Monografie Matematyczne, Vol. 72, Birkhäuser, Basel 2012.

[13] A. Osękowski, *Survey article: Bellman function method and sharp inequalities for martingales*, Rocky Mountain J. Math. 43 (2013), pp. 1759–1823.

[14] A. Osękowski, *Maximal weak-type inequality for stochastic integrals*, Electron. Commun. Probab. 19 (2014), Article 25, pp. 1–13.

[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), pp. 1545–1564.

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

Mateusz Rapicki
Faculty of Mathematics, Informatics and Mechanics
University of Warsaw
ul. Banacha 2
02-097 Warsaw, Poland
E-mail: mateusz.rapicki@mimuw.edu.pl

Received on 5.12.2015;
revised version on 10.2.2016