Bernoulli **17**(4), 2011, 1327–1343
DOI: 10.3150/10-BEJ314

Sharp maximal inequalities for the moments of martingales and non-negative submartingales

ADAM OSĘKOWSKI

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

In the paper we study sharp maximal inequalities for martingales and non-negative submartingales: if f, g are martingales satisfying

$$|\mathrm{d}g_n| \le |\mathrm{d}f_n|, \qquad n = 0, 1, 2, \ldots,$$

almost surely, then

$$\Big\| \sup_{n\ge 0} |g_n| \Big\|_p \le p\|f\|_p, \qquad p \ge 2,$$

and the inequality is sharp. Furthermore, if $\alpha \in [0, 1]$, f is a non-negative submartingale and g satisfies

$$|\mathrm{d}g_n| \le |\mathrm{d}f_n| \quad \text{and} \quad |\mathbb{E}(\mathrm{d}g_{n+1}|\mathcal{F}_n)| \le \alpha\mathbb{E}(\mathrm{d}f_{n+1}|\mathcal{F}_n), \qquad n = 0, 1, 2, \ldots,$$

almost surely, then

$$\Big\| \sup_{n\ge 0} |g_n| \Big\|_p \le (\alpha + 1)p\|f\|_p, \qquad p \ge 2,$$

and the inequality is sharp. As an application, we establish related estimates for stochastic integrals and Itô processes. The inequalities strengthen the earlier classical results of Burkholder and Choi.

Keywords: differential subordination; martingale; maximal function; maximal inequality; submartingale

1. Introduction

The purpose of the paper is to provide the best constants in some maximal inequalities for martingales and non-negative submartingales. Let us start with introducing the necessary notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a non-atomic probability space, equipped with a filtration $(\mathcal{F}_n)_{n\ge 0}$, that is, a non-decreasing family of sub-σ-fields of $\mathcal{F}$. Let $f = (f_n)$ and $g = (g_n)$ be adapted, real-valued integrable processes. The difference sequences $\mathrm{d}f = (\mathrm{d}f_n)$ and $\mathrm{d}g = (\mathrm{d}g_n)$ of f and g are defined by the equations

$$f_n = \sum_{k=0}^{n} \mathrm{d}f_k, \qquad g_n = \sum_{k=0}^{n} \mathrm{d}g_k, \qquad n = 0, 1, 2, \ldots.$$

1350-7265

We are particularly interested in those pairs (f, g) for which a certain domination relation is satisfied. Following Burkholder [6], we say that g is *differentially subordinate* to f if, for any $n \geq 0$, we have

$$\mathbb{P}(|\mathrm{d}g_n| \leq |\mathrm{d}f_n|) = 1.$$

As an example, let g be a transform of f by a predictable sequence $v = (v_n)$ bounded in absolute value by 1; that is, we have $\mathbb{P}(|v_n| \leq 1) = 1$ and $\mathrm{d}f_n = v_n\,\mathrm{d}g_n$, $n \geq 0$. Here, by predictability, we mean that v_0 is $\mathcal{F}_0$-measurable and v_n is $\mathcal{F}_{n-1}$-measurable for $n \geq 1$. In the particular case when each v_n is deterministic and takes values in $\{-1, 1\}$, we will say that g is a ± 1 transform of f.

Another domination we will consider is the so-called α-strong subordination, where α is a fixed non-negative number. This notion was introduced by Burkholder in [10] in the special case $\alpha = 1$ and extended to a general case by Choi [12]: The process g is α-strongly subordinate to f if it is differentially subordinate to f and, for any $n \geq 0$,

$$|\mathbb{E}(\mathrm{d}g_{n+1}|\mathcal{F}_n)| \leq \alpha |\mathbb{E}(\mathrm{d}f_{n+1}|\mathcal{F}_n)|$$

almost surely.

There is a vast literature concerning the comparison of the sizes of f and g under the assumption of one of the dominations above and the further condition that f is a martingale or non-negative submartingale; we refer the interested reader to the papers [6,9,10,12,15,16,18–21] and the references therein. In addition, these inequalities have found their applications in many areas of mathematics: Banach space theory [4,5]; harmonic analysis [8,13,14]; functional analysis [6,7,20]; analysis [1,2]; stochastic integration [6,11,17,20,21]; and more. To present our motivation, we state here only two theorems. Let us start with a fundamental result of Burkholder [6]. We use the notation $\|f\|_p = \sup_n \|f_n\|_p$, $p \in [1, \infty]$.

Theorem 1.1 (Burkholder). *Assume that f, g are martingales and g is differentially subordinate to f. Then, for any $1 < p < \infty$,*

$$\|g\|_p \leq (p^* - 1)\|f\|_p, \tag{1.1}$$

where $p^* = \max\{p, p/(p-1)\}$. The constant $p^* - 1$ is the best possible; *it is already the best possible if g is assumed to be a ± 1 transform of f.*

Here, by the optimality of the constant, we mean that for any $r < p^* - 1$ there exists a martingale f and its ± 1 transform g, for which $\|g\|_p > r\|f\|_p$.

The submartingale version of the estimate above is the following result of Choi [12].

Theorem 1.2 (Choi). *Assume that f is a non-negative submartingale and g is α-differentially subordinate to f, $\alpha \in [0, 1]$. Then for any $1 < p < \infty$,*

$$\|g\|_p \leq (p^*_\alpha - 1)\|f\|_p, \tag{1.2}$$

where $p^*_\alpha = \max\{(\alpha + 1)p, p/(p-1)\}$. The constant is the best possible.

In the paper we deal with a considerably harder problem and determine the optimal constants in the related moment estimates involving the *maximal functions* of f and g. For $n \geq 0$, let $f_n^* = \sup_{0 \leq k \leq n} |f_k|$ and $f^* = \sup_{k \geq 0} |f_k|$. Here is our first main result.

Theorem 1.3. *Let f, g be martingales with g being differentially subordinate to f. Then for any $p \geq 2$,*

$$\|g^*\|_p \leq p\|f\|_p \tag{1.3}$$

and the constant p is the best possible. It is already the best possible in the following weaker inequality: If f is a martingale and g is its ± 1 transform, then

$$\|g^*\|_p \leq p\|f^*\|_p. \tag{1.4}$$

Note that the validity of the estimates (1.3) and (1.4) is an immediate consequence of (1.1) and Doob's bound $\|f^*\|_p \leq \frac{p}{p-1}\|f\|_p$, $p > 1$. The non-trivial (and quite surprising) part is the optimality of the constant p.

Now let us state the submartingale version of the theorem above.

Theorem 1.4. *Fix $\alpha \in [0, 1]$. Let f be a non-negative submartingale and g be real valued and α-strongly subordinate to f. Then for any $p \geq 2$,*

$$\|g^*\|_p \leq (\alpha + 1)p\|f\|_p \tag{1.5}$$

and the constant $(\alpha + 1)p$ is the best possible. It is already the best possible in the weaker estimate

$$\|g^*\|_p \leq (\alpha + 1)p\|f^*\|_p. \tag{1.6}$$

There is a natural question: What is the best constant in the inequalities above in the case $1 < p < 2$? Unfortunately, we have been unable to answer it; our reasoning works only for the case $p \geq 2$.

The proof of (1.5) is based on a technique invented by Burkholder in [11]. It enables us to translate the problem of proving a maximal inequality for martingales to that of finding a certain special function, an upper solution to a corresponding nonlinear problem. The method can be easily extended to the submartingale setting (see [17]) and we construct the function in Section 3. For the sake of construction, we need a solution to a differential equation that is analyzed in Section 2. The next two sections are devoted to the proofs of the announced results: Section 4 contains the proof of the estimate (1.5) and the final part concerns the optimality of the constants appearing in (1.4) and (1.6). In the final section, we present some applications: sharp estimates for stochastic integrals and Itô processes.

2. A differential equation

For a fixed $\alpha \in (0, 1]$ and $p \geq 2$, let $C = C_{p,\alpha} = [(\alpha+1)p]^p(p-1)$. A central role in the paper is played by a certain solution to the differential equation

$$\gamma'(x) = \frac{-1 + C(1-\gamma(x))\gamma(x)x^{p-2}}{1 + C(1-\gamma(x))x^{p-1}}. \tag{2.1}$$

Lemma 2.1. *There is a solution* $\gamma : [((\alpha+1)p)^{-1}, \infty) \to \mathbb{R}$ *of* (2.1), *satisfying the initial condition*

$$\gamma\left(\frac{1}{(\alpha+1)p}\right) = 1 - [(\alpha+1)p]^{-1}. \tag{2.2}$$

The solution is non-decreasing, concave and bounded from above by 1.

Proof. Let γ be a solution to (2.1), satisfying (2.2) and extended to a maximal subinterval I of $[((\alpha+1)p)^{-1}, \infty)$. It is convenient to split the proof into a few steps.

Step 1: $I = [((\alpha+1)p)^{-1}, \infty)$. In view of the Picard–Lindelöf theorem, this will be established if we show that $\gamma < 1$ on I. To this end, suppose that the set $\{x \in I : \gamma(x) = 1\}$ is non-empty and let y denote its smallest element. Then, by (2.1), we have $\gamma'(y) = -1$, which, by minimality of y, implies $\gamma(((\alpha+1)p)^{-1}) > 1$ and contradicts (2.2).

Step 2: *Concavity of* γ. Suppose that the set $\{x \in I : \gamma''(x) > 0\}$ is non-empty and let z denote its infimum. Consider the functions $F, G : (((\alpha+1)p)^{-1}, \infty) \to \mathbb{R}$ given by

$$F(x) = \gamma(x) - x\gamma'(x),$$
$$G(x) = \big(1-\gamma(x)\big)x^{p-2}.$$

Observe that

$$G > 0 \qquad \text{on } I \quad \text{and} \quad F > 0 \qquad \text{on } \big(\big((\alpha+1)p\big)^{-1}, z+\varepsilon\big) \tag{2.3}$$

for some $\varepsilon > 0$. The statement about G is clear, while the positivity of F follows from

$$F'(x) = -x\gamma''(x) \geq 0, \qquad x \in \big(\big((\alpha+1)p\big)^{-1}, z\big]$$

and

$$F\big(\big((\alpha+1)p\big)^{-1}+\big) = \frac{1}{p} > 0.$$

Now multiply (2.1) throughout by $1 + C(1-\gamma(x))x^{p-1}$ and differentiate both sides. We obtain an equality that is equivalent to

$$\gamma''(x)\big(1 + CxG(x)\big) = CF(x)G'(x), \qquad x > \frac{1}{(\alpha+1)p}. \tag{2.4}$$

As a first consequence, we have $z > ((\alpha+1)p)^{-1}$. To see this, tend with x down to $((\alpha+1)p)^{-1}$ and observe that F and G have strictly positive limits; furthermore,

$$G'(x) = x^{p-3}\big[(p-2)\big(1-\gamma(x)\big) - x\gamma'(x)\big] =: x^{p-3}J(x) \tag{2.5}$$

with $J(((\alpha+1)p)^{-1}) = -\frac{\alpha(p-1)}{(\alpha+1)p} < 0$. Combining (2.3) and (2.4) we see that, for some $\varepsilon > 0$, $G' \leq 0$ on $(z-\varepsilon, z)$ and $G' > 0$ on $(z, z+\varepsilon)$. Consequently, by (2.5), $J \leq 0$ on $(z-\varepsilon, z)$ and $J > 0$ on $(z, z+\varepsilon)$. This implies $J'(z) > 0$ and since $J'(z) = -(p-1)\gamma'(z)$, we get $\gamma'(z) < 0$. However, this contradicts $G'(z) = 0$, in view of (2.5) and $\gamma(z) < 1$. Let us stress that here, in the last passage, we use the inequality $p \geq 2$.

Step 3: *γ is non-decreasing.* It follows from (2.4), the concavity of γ and positivity of F and G, that $G' \leq 0$, or, by (2.5),

$$(p-2)\big(1-\gamma(x)\big) - x\gamma'(x) \leq 0. \tag{2.6}$$

The claim follows. □

Let us extend γ to the whole half-line $[0, \infty)$ by

$$\gamma(x) = [(p-1)(\alpha+1)-1]x + \frac{1}{p}, \qquad x \in \left[0, \frac{1}{(\alpha+1)p}\right).$$

It can be verified readily that γ is of class C^1 on $(0, \infty)$. For the sake of reader's convenience, the graph of γ, corresponding to $p = 3$ and $\alpha = 1$, is presented on Figure 1.

Let $H : [((\alpha+1)p)^{-1}, \infty) \to [1, \infty)$ be given by $H(x) = x + \gamma(x)$ and let h be the inverse to H. Clearly, we have

$$x - 1 \leq h(x) \leq x, \qquad x \geq 1. \tag{2.7}$$

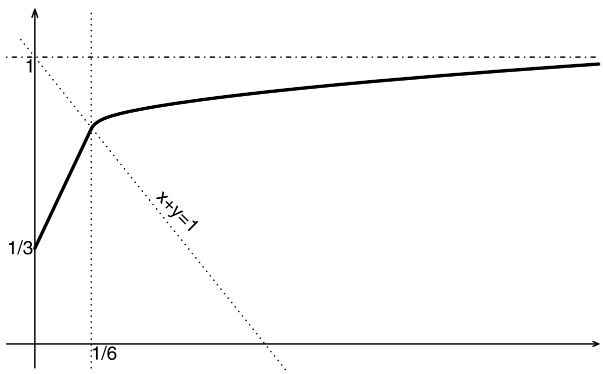

Figure 1. The graph of γ (the bold line) in the case $p = 3$, $\alpha = 1$. Note that γ is linear on $[0, 1/6]$ and solves (2.1) on $(1/6, \infty)$.

We conclude this section by providing a formula for h' to be used later. As

$$h'(x) = \frac{1}{H'(h(x))} = \frac{1}{1+\gamma'(h(x))}, \qquad x > 1, \tag{2.8}$$

it can be derived that, in view of (2.1),

$$h'(x) = \frac{1 + ((\alpha+1)p)^p (p-1)(h(x) - x + 1)h(x)^{p-1}}{((\alpha+1)p)^p (p-1)(h(x) - x + 1)h(x)^{p-2}x}. \tag{2.9}$$

3. The special function

Throughout this section, $\alpha \in (0, 1]$ and $p \geq 2$ are fixed. Let S denote the strip $[0, \infty) \times [-1, 1]$. Consider the following subsets of S.

$$\begin{aligned} D_0 &= \{(x, y) \in S : |y| \leq \gamma(x)\}, \\ D_1 &= \{(x, y) \in S : |y| > \gamma(x), x + |y| \leq 1\}, \\ D_2 &= \{(x, y) \in S : |y| > \gamma(x), x + |y| > 1\}. \end{aligned}$$

Introduce the function $u : S \to \mathbb{R}$ by

$$u(x, y) = \begin{cases} 1 - [(\alpha+1)p]^p x^p & \text{on } D_0, \\ 1 - \left(\dfrac{px + p|y| - 1}{p-1}\right)^{p-1} \big[p\big(p(\alpha+1) - 1\big)x - p|y| + 1\big] & \text{on } D_1, \\ 1 - [(\alpha+1)p]^p h(x+|y|)^{p-1}[px - (p-1)h(x+|y|)] & \text{on } D_2. \end{cases}$$

Let $U : [0, \infty) \times \mathbb{R} \times (0, \infty) \to \mathbb{R}$ be given by

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

As we will see below, the function U is the key to the inequality (1.5). Let us study the properties of this function.

Lemma 3.1. *The function U is of class C^1. Furthermore, there exists an absolute constant K such that, for all $x > 0$, $y \in \mathbb{R}$, $z > 0$, we have*

$$U(x, y, z) \leq K(x + |y| + z)^p \tag{3.1}$$

and

$$U_x(x, y, z) \leq K(x + |y| + z)^{p-1}, \qquad U_x(x, y, z) \leq K(x + |y| + z)^{p-1}. \tag{3.2}$$

Proof. The continuity of the partial derivatives can be verified readily. The inequality (3.1) is evident for those (x, y, z), for which $(\frac{x}{|y| \vee z}, \frac{y}{|y| \vee z}) \in D_0 \cup D_1$; for the remaining (x, y, z), it suffices to use (2.7). Finally, the inequality (3.2) is clear if $(\frac{x}{|y|\vee z}, \frac{y}{|y|\vee z}) \in D_0 \cup D_1$. For the remaining points one applies (2.7) and (2.8), the latter inequality implying $h' < 1$. □

Now let us deal with the following majorization property.

Lemma 3.2. *For any* $(x, y, z) \in [0, \infty) \times \mathbb{R} \times (0, \infty)$, *we have*

$$U(x, y, z) \geq (|y| \vee z)^p - [(\alpha + 1)p]^p x^p. \tag{3.3}$$

Proof. The inequality is equivalent to $u(x, y) \geq 1 - [(\alpha + 1)p]^p x^p$ and we need to establish it only on D_1 and D_2. On D_1, the substitutions $X = px$ and $Y = p|y| - 1$ (note that $Y \geq 0$) transform it into

$$(\alpha + 1)^p X^p \geq \left(\frac{X + Y}{p - 1}\right)^{p-1} \big[\big(p(\alpha + 1) - 1\big)X - Y\big].$$

This inequality is valid for all non-negative X, Y. To see this, observe that by homogeneity we may assume $X + Y = 1$, and then the estimate reads

$$F(X) := (\alpha + 1)^p X^p - (p - 1)^{-p+1}[p(\alpha + 1)X - 1] \geq 0, \qquad X \in [0, 1].$$

Now it suffices to note that F is convex on $[0, 1]$ and satisfies

$$F\left(\frac{1}{(p - 1)(\alpha + 1)}\right) = F'\left(\frac{1}{(p - 1)(\alpha + 1)}\right) = 0.$$

It remains to show the majorization on D_2. It is dealt with in a similar manner: Setting $s = x + |y| > 1$, we see that (3.3) is equivalent to

$$G(x) := x^p - h(s)^{p-1}[px - (p - 1)h(s)] \geq 0, \qquad s - 1 < x < h(s).$$

It is easily verified that G is convex and satisfies $G(h(s)) = G'(h(s)) = 0$. This completes the proof of (3.3). □

The main property of the function U is the concavity along the lines of slope belonging to $[-1, 1]$.

Lemma 3.3. *For fixed* y, z *satisfying* $z > 0$, $|y| \leq z$, *and any* $a \in [-1, 1]$, *the function* $\Phi = \Phi_{y,z,a} : [0, \infty) \to \mathbb{R}$ *given by*

$$\Phi(t) = U(t, y + at, z)$$

is concave.

Before we turn to the proof, let us first establish some useful consequences.

Corollary 3.4. (i) *The function U has the following property: For any x, y, z, k_x, k_y such that $x, x+k_x \geq 0$, $z > 0$, $|y| \leq z$ and $|k_y| \leq |k_x|$, we have*

$$U(x+k_x, y+k_y, z) \leq U(x, y, z) + U_x(x, y, z)k_x + U_y(x, y, z)k_y \tag{3.4}$$

(*for $x = 0$, we replace $U_x(0, y, z)$ by right-sided derivative $U_x(0+, y, z)$*).

(ii) *For any $x \geq 1$, we have*

$$U(x, 1, 1) \leq 0. \tag{3.5}$$

Proof. (i) This follows immediately.

(ii) We have $\Phi_{0,1,x^{-1}}(0) = U(0, 0, 1) = 1$ and $\Phi_{0,1,x^{-1}}(((\alpha+1)p)^{-1}) = U(((\alpha+1)p)^{-1}, x^{-1}((\alpha+1)p)^{-1}, 1) = 0$, since $(((\alpha+1)p)^{-1}, x^{-1}((\alpha+1)p)^{-1}, 1) \in D_0$. Since $x \geq 1 > ((\alpha+1)p)^{-1}$, the lemma above gives $U(x, 1, 1) = \Phi_{0,1,x^{-1}}(x) \leq 0$. □

Proof of Lemma 3.3. By homogeneity, we may assume $z = 1$. As Φ is of class C^1, it suffices to verify that $\Phi''(t) \leq 0$ for those t, for which $(t, y+at)$ lies in the interior of D_0, D_1, D_2 or outside the strip S. Since $U(x, y, z) = U(x, -y, z)$, we may restrict ourselves to the case $y + at \geq 0$. If $(t, y+at)$ belongs to D_0^o, the interior of D_0, then $\Phi''(t) = -[(\alpha+1)p]^p \cdot p(p-1)t^{p-2} < 0$, while for $(t, y+at) \in D_1^o$ we have

$$\Phi''(t) = -\frac{p^3(pt + p(y+at) - 1)^{p-3}(1+a)}{(p-1)^{p-2}}(I_1 + I_2),$$

where

$$\begin{aligned} I_1 &= pt\big[(p-2)(1+a)\big(p(\alpha+1) - 1\big) + 2\big(p(\alpha+1) - 1 - a\big)\big] \geq 0, \\ I_2 &= \big(p(y+at) - 1\big)(2\alpha + 1 - a) \geq 0. \end{aligned}$$

The remaining two cases are a bit more complicated. If $(t, y+at) \in D_2^o$, then

$$\frac{\Phi''(t)}{Cp(1+a)^2} = J_1 + J_2 + J_3,$$

where

$$\begin{aligned} J_1 &= h(t+y+at)^{p-2}h''(t+y+at)[h(t+y+at) - t], \\ J_2 &= h(t+y+at)^{p-3}[h'(t+y+at)]^2[(p-1)h(t+y+at) - (p-2)t], \\ J_3 &= -\frac{2}{a+1}h(t+y+at)^{p-2}h'(t+y+at). \end{aligned}$$

Now if we change y and t, keeping $s = t + y + at$ fixed, then $J_1 + J_2 + J_3$ is a linear function of $t \in [s-1, h(s)]$. Therefore, to prove it is non-positive, it suffices to verify this for $t = h(s)$ and $t = s - 1$. For $t = h(s)$, we have

$$J_1 + J_2 + J_3 = h(s)^{p-2}h'(s)\left[h'(s) - \frac{2}{a+1}\right] \leq 0,$$

since $0 \le h'(s) \le 1$ (see (2.8)). If $t = s - 1$, rewrite (2.9) in the form

$$Cs\big(h(s)+1-s\big)h(s)^{p-2}h'(s) = 1 + C\big(h(s)+1-s\big)h(s)^{p-1}$$

and differentiate both sides; as a result, we obtain

$$\begin{aligned}&Cs\left[J_1 + J_2 + J_3 + h(s)^{p-2}h'(s)\left(\frac{2}{a+1} - 1\right)\right]\\&\quad = Ch(s)^{p-2}\big[\big(h'(s)-1\big)h(s) + (p-2)\big(h(s)+1-s\big)h'(s)\big].\end{aligned}$$

As $h' \ge 0$ and $2/(a+1) \ge 1$, we will be done if we show the right-hand side is non-positive. This is equivalent to

$$h'(s)\big[h(s) + (p-2)\big(h(s)+1-s\big)\big] \le h(s).$$

Now use (2.8) and substitute $h(s) = r$, noting that $h(s)+1-s = 1-\gamma(r)$, to obtain

$$r + (p-2)\big(1-\gamma(r)\big) \le r\big(1+\gamma'(r)\big),$$

or $r\gamma'(r) \ge (p-2)(1-\gamma(r))$, which is (2.6).

Finally, suppose that $y + at > 1$. For such t we have $\Phi(t) = (y+at)^p u(t/(y+at), 1)$, hence, setting $X = t/(y+t)$, $Y = y + at$, we easily check that $\Phi''(t)$ equals

$$Y^{p-2}[p(p-1)a^2u(X,1) + 2a(p-1)(1-aX)u_x(X,1) + (1-aX)^2u_{xx}(X,1)].$$

First let us derive the expressions for the partial derivatives. Using (2.9), we have

$$\begin{aligned}u_x(X,1) &= \frac{p}{X+1}\big[1 + C\big(h(X+1)-X\big)h(X+1)^{p-1}\big] - \frac{Cph(X+1)^{p-1}}{p-1},\\ u_{xx}(X,1) &= \frac{p(p-1)}{(X+1)^2}\big[1 + C\big(h(X+1)-X\big)h(X+1)^{p-1}\big]\\ &\quad - \frac{Cph(X+1)^{p-1}}{X+1} - \frac{Cph(X+1)^{p-2}h'(X+1)}{X+1}.\end{aligned}$$

Now it can be checked that

$$\Phi''(t)Y^{2-p}/p = K_1 + K_2 + K_3,$$

where

$$\begin{aligned}K_1 &= (p-1)\left(\frac{a+1}{X+1}\right)^2\big[1 + C\big(h(X+1)-X\big)h(X+1)^{p-1}\big],\\ K_2 &= -\frac{Ch(X+1)^{p-1}}{X+1}(1+2a-a^2X),\end{aligned}$$

$$K_3 = -\left(\frac{1-aX}{X+1}\right)^2 \cdot \frac{1 + C(h(X+1)-X)h(X+1)^{p-1}}{h(X+1)-X}$$
$$\le -\left(\frac{1-aX}{X+1}\right)^2 \cdot Ch(X+1)^{p-1}.$$

We may write

$$K_2 + K_3 \le -\frac{Ch(X+1)^{p-1}}{(X+1)^2}[(1+2a-a^2X)(X+1) + (1-aX)^2]$$
$$= -\frac{Ch(X+1)^{p-1}(a+1)}{(X+1)^2}[2 + X(1-a)] \le -\left(\frac{a+1}{X+1}\right)^2 Ch(X+1)^{p-1},$$

where, in the last passage, we used $a \le 1$. On the other hand, as h is non-decreasing, we have

$$1 = \frac{Ch(1)^p}{p-1} \le \frac{Ch(X+1)^{p-1}h(1)}{p-1}.$$

Moreover, since $x \mapsto h(x+1) - x$ is non-increasing (see (2.8)), we have $h(X+1) - X \le h(1)$. Combining these two facts, we obtain

$$K_1 \le (p-1)\left(\frac{a+1}{X+1}\right)^2 [1 + Ch(1)h(X+1)^{p-1}]$$
$$\le \left(\frac{a+1}{X+1}\right)^2 Ch(X+1)^{p-1}[h(1) + (p-1)h(1)]$$
$$\le \left(\frac{a+1}{X+1}\right)^2 Ch(X+1)^{p-1},$$

as $ph(1) = (\alpha+1)^{-1} \le 1$. This implies $K_1 + K_2 + K_3 \le 0$ and completes the proof. □

The final property we will need is the following.

Lemma 3.5. *For any x, y, z such that $x \ge 0$, $z > 0$ and $|y| \le z$, we have*

$$U_x(x, y, z) \le -\alpha |U_y(x, y, z)| \tag{3.6}$$

(*if $x = 0$, then U_x is replaced by a right-sided derivative*).

Proof. It suffices to show that for fixed y, z, $|y| \le z$, and $a \in [-\alpha, \alpha]$, the function $\Phi = \Phi_{y,z,a} : [0, \infty) \to \mathbb{R}$ given by $\Phi(t) = U(t, y + at, z)$ is non-increasing. Since $\alpha \le 1$, we know from the previous lemma that Φ is concave. Hence all we need is $\Phi'(0+) \le 0$. By symmetry, we may assume $y \ge 0$. If $y \le 1/p$, then the derivative equals 0; in the remaining case, we have

$$\Phi'(0+) = -\frac{p^2(py-1)^{p-1}}{(p-1)^{p-1}}(\alpha - a) \le 0.$$

□

4. The proof of (1.5)

First let us observe that it suffices to show (1.5) for strictly positive α. This is an immediate consequence of the fact that α-strong subordination implies α'-strong subordination for $\alpha < \alpha'$.

Suppose f, g are as in Theorem 1.4. We may restrict ourselves to the case $\|f\|_p < \infty$. Hence, by Choi's inequality (1.2), we have $\|g\|_p < \infty$. It suffices to show that for any $n = 0, 1, 2, \dots$ we have

$$\mathbb{E}[(g_n^*)^p - (\alpha+1)^p p^p f_n^p] \le 0.$$

Clearly, we may assume that $\mathbb{P}(g_0 > 0) = 1$, simply replacing f, g by $f + \varepsilon$, $g + \varepsilon$ if necessary (here ε is a small positive number). In particular, this implies $f_0 > 0$ almost surely. In view of the majorization (3.3), we will be done if we show that the expectation $\mathbb{E}U(f_n, g_n, g_n^*)$ is non-positive for any n. As a matter of fact, we will show more; namely, that the process $(U(f_n, g_n, g_n^*)_{n\ge 0})$ is a supermartingale and $\mathbb{E}U(f_0, g_0, g_0^*) \le 0$.

To this end, fix $n \ge 1$ and observe that $g_n^* \le |g_0| + |g_1| + \cdots + |g_n|$, so g_n^* belongs to L^p. Thus, by Lemma 3.1 and Hölder's inequality, the variables $U(f_n, g_n, g_n^*)$, $U_x(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}f_n$ and $U_y(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}g_n$ are integrable. Moreover, by definition of U and the inequality (3.4),

$$\begin{aligned}
\mathbb{E}(U(f_n, g_n, g_n^*)|\mathcal{F}_{n-1}) &= \mathbb{E}(U_n(f_n, g_n, g_{n-1}^*)|\mathcal{F}_{n-1}) \\
&= \mathbb{E}\big(U(f_{n-1} + \mathrm{d}f_n, g_{n-1} + \mathrm{d}g_n, g_{n-1}^*)|\mathcal{F}_{n-1}\big) \\
&\le \mathbb{E}[U(f_{n-1}, g_{n-1}, g_{n-1}^*) + U_x(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}f_n \\
&\qquad + U_y(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}g_n|\mathcal{F}_{n-1}] \\
&\le U(f_{n-1}, g_{n-1}, g_{n-1}^*).
\end{aligned}$$

The latter inequality is the consequence of the following. By (3.6) and the submartingale property of f,

$$\begin{aligned}
\mathbb{E}(U_x(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}f_n|\mathcal{F}_{n-1}) &= U_x(f_{n-1}, g_{n-1}, g_{n-1}^*)\mathbb{E}(\mathrm{d}f_n|\mathcal{F}_{n-1}) \\
&\le -\alpha|U_y(f_{n-1}, g_{n-1}, g_{n-1}^*)|\mathbb{E}(\mathrm{d}f_n|\mathcal{F}_{n-1}) \\
&\le -U_y(f_{n-1}, g_{n-1}, g_{n-1}^*)\mathbb{E}(\mathrm{d}g_n|\mathcal{F}_{n-1}) \\
&= -\mathbb{E}(U_y(f_{n-1}, g_{n-1}, g_{n-1}^*)\,\mathrm{d}g_n|\mathcal{F}_{n-1}),
\end{aligned}$$

where the second inequality is due to α-domination.

To complete the proof, it suffices to show that $\mathbb{E}U(f_0, g_0, g_0^*) \le 0$. However, $U(f_0, g_0, g_0^*) = U(f_0, g_0, g_0) = g_0^p U(f_0/g_0, 1, 1)$ almost surely and the estimate follows from Corollary 3.4(ii).

5. Sharpness

We start with inequality (1.4) and restrict ourselves to the case when g is a ± 1 transform of f. Suppose the best constant in this estimate equals $\beta > 0$. This implies the existence of a function

$W : \mathbb{R} \times \mathbb{R} \times [0, \infty) \times [0, \infty) \to \mathbb{R}$, which satisfies the following properties:

$$W(1, 1, 1, 1) \leq 0, \tag{5.1}$$

$$W(x, y, z, w) = W(x, y, |x| \vee z, |y| \vee w), \qquad \text{if } x, y \in \mathbb{R}, w, z \geq 0, \tag{5.2}$$

$$(|y| \vee w)^p - \beta^p (|x| \vee z)^p \leq W(x, y, z, w), \qquad \text{if } x, y \in \mathbb{R}, w, z \geq 0 \tag{5.3}$$

and, furthermore,

$$\begin{aligned} &aW(x + t_1, y + \varepsilon t_1, z, w) + (1 - a)W(x + t_2, y + \varepsilon t_2, z, w) \leq W(x, y, z, w) \\ &\quad \text{for any } |x| \leq z, |y| \leq w, \varepsilon \in \{-1, 1\}, a \in (0, 1) \text{ and } t_1,\ t_2 \text{ with } at_1 + (1 - a)t_2 = 0. \end{aligned} \tag{5.4}$$

Indeed, one puts

$$W(x, y, z, w) = \sup\{\mathbb{E}(g_n^* \vee w)^p - \beta^p \mathbb{E}(f_n^* \vee z)^p\}, \tag{5.5}$$

where the supremum is taken over all integers n and all martingales f, g satisfying $\mathbb{P}((f_0, g_0) = (x, y)) = 1$ and $\mathrm{d}f_k = \pm \mathrm{d}g_k$, $k = 1, 2, \ldots$ (see [11] for details). This formula allows us to assume that W is homogeneous: $W(tx, ty, tz, tw) = tW(x, y, z, w)$ for all $x, y \in \mathbb{R}$, $z, w \geq 0$ and $t > 0$.

Now the idea is to exploit the above properties of W to get $\beta \geq p$. To this end, let δ be a small number belonging to $(0, 1/p)$. By (5.4) applied to $x = 0$, $y = w = 1$, $z = \delta/(1 + 2\delta)$, $\varepsilon = 1$ and $t_1 = \delta$, $t_2 = -1/p$, we obtain

$$\begin{aligned} W\left(0, 1, \frac{\delta}{1 + 2\delta}, 1\right) \geq \frac{p\delta}{1 + p\delta} W\left(-\frac{1}{p}, 1 - \frac{1}{p}, \frac{\delta}{1 + 2\delta}, 1\right) \\ + \frac{1}{1 + p\delta} W\left(\delta, 1 + \delta, \frac{\delta}{1 + 2\delta}, 1 + \delta\right). \end{aligned} \tag{5.6}$$

Now, by (5.2) and (5.3),

$$W\left(-\frac{1}{p}, 1 - \frac{1}{p}, \frac{\delta}{1 + 2\delta}, 1\right) = W\left(-\frac{1}{p}, 1 - \frac{1}{p}, \frac{1}{p}, 1\right) \geq 1 - \left(\frac{\beta}{p}\right)^p. \tag{5.7}$$

Furthermore, by (5.2),

$$W\left(\delta, 1 + \delta, \frac{\delta}{1 + 2\delta}, 1 + \delta\right) = W(\delta, 1 + \delta, \delta, 1 + \delta),$$

which, by (5.4) (with $x = z = \delta$, $y = w = 1 + \delta$, $\varepsilon = -1$ and $t_1 = -\delta$, $t_2 = \frac{1}{p} + \delta(\frac{1}{p} - 1)$), can be bounded from below by

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

Using (5.3), we get

$$W\left(\frac{1+\delta}{p}, 1-\frac{1}{p}+\delta\left(2-\frac{1}{p}\right), \delta, 1+\delta\right) \geq (1+\delta)^p\left[1-\left(\frac{\beta}{p}\right)^p\right].$$

Furthermore, by (5.2) and the homogeneity of W,

$$W(0, 1+2\delta, \delta, 1+\delta) = W(0, 1+2\delta, \delta, 1+2\delta) = (1+2\delta)^p W\left(0, 1, \frac{\delta}{1+2\delta}, 1\right).$$

Now plug all the above estimates into (5.6) to get

$$\begin{aligned} &W\left(0, 1, \frac{\delta}{1+2\delta}, 1\right)\left[1-\frac{(1+\delta-p\delta)(1+2\delta)^p}{(1+\delta)(1+p\delta)}\right] \\ &\quad \geq \frac{p\delta}{1+p\delta}\left[1-\left(\frac{\beta}{p}\right)^p\right]\bigl(1+(1+\delta)^{p-1}\bigr). \end{aligned} \tag{5.8}$$

Now it follows from the definition (5.5) of W that

$$W\left(0, 1, \frac{\delta}{1+2\delta}, 1\right) \leq W(0, 1, 0, 1).$$

Furthermore, one easily checks that the function

$$F(s) = 1-\frac{(1+s-ps)(1+2s)^p}{(1+s)(1+ps)}, \qquad s > -\frac{1}{p},$$

satisfies $F(0) = F'(0) = 0$. Hence

$$1-\left(\frac{\beta}{p}\right)^p \leq \frac{W(0,1,0,1)\cdot F(\delta)\cdot(1+p\delta)}{p\delta(1+(1+\delta)^{p-1})}$$

and letting $\delta \to 0$ yields $1-(\frac{\beta}{p})^p \leq 0$, or $\beta \geq p$.

The reasoning for the inequality (1.6) is essentially the same: suppose the best constant in the estimate equals $\gamma > 0$. Introduce the function $V : [0,\infty)\times\mathbb{R}\times[0,\infty)\times[0,\infty)\to\mathbb{R}$ by

$$V(x, y, z, w) = \sup\{\mathbb{E}(g_n^* \vee w)^p - \gamma^p\mathbb{E}(f_n^* \vee z)^p\},$$

where the supremum is taken over all integers n, all non-negative submartingales f and all integrable sequences g satisfying $\mathbb{P}((f_0, g_0) = (x, y)) = 1$ and, for $k = 1, 2, \ldots,$

$$|\mathrm{d}f_k| \geq |\mathrm{d}g_k|, \qquad \alpha\mathbb{E}(\mathrm{d}f_k|\mathcal{F}_{k-1}) \geq |\mathbb{E}(\mathrm{d}g_k|\mathcal{F}_{k-1})|$$

with probability 1. We see that V is homogeneous and satisfies the properties analogous to (5.1)–(5.4) (with obvious changes: in (5.2) and (5.3) one must assume $x \geq 0$; in (5.3) the number β

is replaced by γ; and, in (5.4), we impose $x, x+t_1, x+t_2 \geq 0$). In addition, there is an extra property of V, which corresponds to the fact that we deal with the inequality for submartingales:

$$V(x+d, y+\alpha d, z, w) \leq V(x, y, z, w), \qquad \text{if } x \geq 0, y \in \mathbb{R}, w, z \geq 0, d \geq 0. \tag{5.9}$$

Now fix $\delta \in (0, 1/p)$ and apply this property with $x = 0$, $y = w = 1$, $z = \delta/(1+(\alpha+1)p)$, $d = \delta$ and then use (5.2) to obtain

$$\begin{aligned} V\left(0, 1, \frac{\delta}{1+(\alpha+1)\delta}, 1\right) &\geq V\left(\delta, 1+\alpha\delta, \frac{\delta}{1+(\alpha+1)\delta}, 1\right) \\ &= V(\delta, 1+\alpha\delta, \delta, 1+\alpha\delta). \end{aligned} \tag{5.10}$$

Using (5.2), (5.3) and (5.4) as above, we have

$$\begin{aligned} V(\delta, 1+\alpha\delta, \delta, 1+\alpha\delta) &\geq \frac{\delta(\alpha+1)p}{1+\alpha\delta}(1+\alpha\delta)^p\left[1-\left(\frac{\gamma}{(\alpha+1)p}\right)^p\right] \\ &\quad + \frac{1+\alpha\delta-\delta(\alpha+1)p}{1+\alpha\delta}\big(1+(\alpha+1)\delta\big)^p V\left(0, 1, \frac{\delta}{1+(\alpha+1)\delta}, 1\right), \end{aligned}$$

which, combined with (5.10), gives

$$\begin{aligned} &V\left(0, 1, \frac{\delta}{1+(\alpha+1)\delta}, 1\right)\left[1-\frac{1+\alpha\delta-\delta(\alpha+1)p}{1+\alpha\delta}\big(1+(\alpha+1)\delta\big)^p\right] \\ &\geq \delta(\alpha+1)p(1+\alpha\delta)^{p-1}\left[1-\left(\frac{\gamma}{(\alpha+1)p}\right)^p\right]. \end{aligned}$$

Now it suffices to use

$$V\left(0, 1, \frac{\delta}{1+(\alpha+1)\delta}, 1\right) \leq V(0, 1, 0, 1)$$

and the fact that the function

$$G(s) = 1 - \frac{1+\alpha s - s(\alpha+1)p}{1+\alpha s}\big(1+(\alpha+1)s\big)^p, \qquad s > -1/\alpha,$$

satisfies $G(0) = G'(0) = 0$, to obtain

$$1-\left(\frac{\gamma}{(\alpha+1)p}\right)^p \leq \frac{V(0,1,0,1)G(\delta)}{\delta(\alpha+1)p(1+\alpha\delta)^{p-1}}.$$

Letting $\delta \to 0$ gives $1-(\frac{\gamma}{(\alpha+1)p})^p \leq 0$, or $\gamma \geq (\alpha+1)p$. This completes the proof.

6. Inequalities for stochastic integrals and Itô processes

In this section we present applications of the results above. Theorem 1.4 in the special case $\alpha = 1$ yields an interesting inequality for the stochastic integrals. Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, filtered by a non-decreasing right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$. In addition, let $\mathcal{F}_0$ contain all the events of probability 0. Suppose $X = (X_t)_{t\geq 0}$ is an adapted non-negative right-continuous submartingale with left limits and let Y be the Itô integral of H with respect to X,

$$Y_t = H_0 X_0 + \int_{(0,t]} H_s \,\mathrm{d}X_s, \qquad t \geq 0.$$

Here H is a predictable process with values in $[-1, 1]$. Denote $\|X\|_p = \sup_{t\geq 0} \|X_t\|_p$ and $X^* = \sup_{t\geq 0} |X_t|$. We will establish the following extension of Theorem 1.4.

Theorem 6.1. *Under the above conditions, we have, for any* $p \geq 2$,

$$\|Y^*\|_p \leq 2p\|X\|_p, \tag{6.1}$$

and the constant $2p$ *is the best possible. It is already the best possible in the weaker estimate*

$$\|Y^*\|_p \leq 2p\|X^*\|_p.$$

Proof. The constant $2p$ is optimal even in the discrete-time setting, so all we need is to show (6.1). This is a consequence of the approximation results of Bichteler [3]. We proceed as follows: Consider the family $\mathbf{Y}$ of all processes Y of the form

$$Y_t = H_0 X_0 + \sum_{k=1}^{n} h_k [X_{\tau_k \wedge t} - X_{\tau_{k-1} \wedge t}], \tag{6.2}$$

where n is a positive integer, h_k belongs to $[-1, 1]$ and the stopping times τ_k take only a finite number of finite values, with $0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_n$. Let

$$f = (X_{\tau_0}, X_{\tau_1}, \ldots, X_{\tau_n}, X_{\tau_n}, \ldots)$$

and let g be the transform of f by $(H_0, h_1, h_2, \ldots, h_n, 0, 0, \ldots)$. In virtue of Doob's optional sampling theorem, f is a submartingale. Therefore, by Theorem 1.4, if $\tau_n \leq t$ almost surely, then for Y as in (6.2),

$$\|Y_t^*\|_p = \|g_n^*\|_p \leq 2p\|f_n\|_p \leq 2p\|X_t\|_p.$$

Now we have that X and H satisfy the conditions of Proposition 4.1 of Bichteler [3]. Thus by (2) of that proposition, if Y is as in the statement of the theorem above, then there is a sequence (Y^j) of elements of $\mathbf{Y}$ such that $\lim_{j\to\infty}(Y^j - Y)^* = 0$ almost surely. Hence, by Fatou's lemma,

$$\|Y_t^*\|_p \leq 2p\|X_t\|_p.$$

Now take $t \to \infty$ to complete the proof. □

The result above can be further strengthened. Assume that X is a non-negative submartingale and $X = X_0 + M + A$ stands for its Doob–Meyer decomposition, uniquely determined by the condition that A is predictable. Let $\alpha \in [0, 1]$ be fixed and suppose ϕ, ψ are predictable processes satisfying $|\phi_s| \le 1$ and $|\psi_s| \le \alpha$ for all s. Consider the Itô process Y such that $|Y_0| \le X_0$ and

$$Y_t = Y_0 + \int_{0+}^{t} \phi_s \, \mathrm{d}M_s + \int_{0+}^{t} \psi_s \, \mathrm{d}A_s$$

for all $t \ge 0$. We have the following sharp bound.

Theorem 6.2. *For X, Y as above, we have*

$$\|Y^*\|_p \le (\alpha + 1)p\|X\|_p$$

and the inequality is sharp. So is the weaker estimate

$$\|Y^*\|_p \le (\alpha + 1)p\|X^*\|_p.$$

This result can be established using essentially the same approximation arguments as above; we omit the details. We would only like to mention here that there is an alternative way of proving Theorems 6.1 and 6.2, based on Itô's formula applied to the function u (as the function is not of class C^2, one needs some additional "smoothing" arguments to overcome this difficulty). See [19] or [20] for similar reasoning.

Acknowledgements

This work was partially supported by MEiN Grant 1 PO3A 012 29 and the Foundation for Polish Science.

References

[1] Bañuelos, R. and Bogdan, K. (2007). Lévy processes and Fourier multipliers. *J. Funct. Anal.* **250** 197–213. MR2345912

[2] Bañuelos, R. and Wang, G. (1995). Sharp inequalities for martingales with applications to the Beurling–Ahlfors and Riesz transformations. *Duke Math. J.* **80** 575–600. MR1370109

[3] Bichteler, K. (1980). Stochastic integration and L^p-theory of semimartingales. *Ann. Probab.* **9** 49–89. MR0606798

[4] Bourgain, J. (1983). Some remarks on Banach spaces in which martingale difference sequences are unconditional. *Ark. Mat.* **21** 163–168. MR0727340

[5] Burkholder, D.L. (1981). A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *Ann. Probab.* **9** 997–1011. MR0632972

[6] Burkholder, D.L. (1984). Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12** 647–702. MR0744226

[7] Burkholder, D.L. (1985). An elementary proof of an inequality of R. E. A. C. Paley. *Bull. London Math. Soc.* **17** 474–478. MR0806015

[8] Burkholder, D.L. (1989). Differential subordination of harmonic functions and martingales. In *Harmonic Analysis and Partial Differential Equations (El Escorial, 1987). Lecture Notes in Mathematics* **1384** 1–23. Berlin: Springer. MR1013814

[9] Burkholder, D.L. (1991). Explorations in martingale theory and its applications. In *Ecole d'Eté de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 1–66. Berlin: Springer. MR1108183

[10] Burkholder, D.L. (1994). Strong differential subordination and stochastic integration. *Ann. Probab.* **22** 995–1025. MR1288140

[11] Burkholder, D.L. (1997). Sharp norm comparison of martingale maximal functions and stochastic integrals. In *Proceedings of the Norbert Wiener Centenary Congress, 1994 (East Lansing, MI, 1994). Proc. Sympos. Appl. Math.* **52** 343–358. Providence, RI: Amer. Math. Soc. MR1440921

[12] Choi, C. (1996). A submartingale inequality. *Proc. Amer. Math. Soc.* **124** 2549–2553. MR1353381

[13] Choi, C. (1998). A weak-type inequality of subharmonic functions. *Proc. Amer. Math. Soc.* **126** 1149–1153. MR1425115

[14] Choi, C. (1998). A weak-type inequality for differentially subordinate harmonic functions. *Trans. Amer. Math. Soc.* **350** 2687–2696. MR1617340

[15] Hammack, W. (1995). Sharp inequalities for the distribution of a stochastic integral in which the integrator is a bounded submartingale. *Ann. Probab.* **23** 223–235. MR1330768

[16] Osękowski, A. (2007). Sharp norm inequalities for martingales and their differential subordinates. *Bull. Polish Acad. Sci. Math.* **55** 373–385. MR2369123

[17] Osękowski, A. (2008). Sharp maximal inequality for stochastic integrals. *Proc. Amer. Math. Soc.* **136** 2951–2958. MR2399063

[18] Osękowski, A. (2008). Sharp LlogL inequalities for differentially subordinated martingales. *Illinois J. Math.* **52** 745–756. MR2546005

[19] Osękowski, A. (2009). Sharp weak type inequalities for differentially subordinated martingales. *Bernoulli* **15** 871–897. MR2555203

[20] Suh, Y. (2005). A sharp weak type (p, p) inequality $(p > 2)$ for martingale transforms and other subordinate martingales. *Trans. Amer. Math. Soc.* **357** 1545–1564. MR2115376

[21] Wang, G. (1995). Differential subordination and strong differential subordination for continuous-time martingales and related sharp inequalities. *Ann. Probab.* **23** 522–551. MR1334160

Received February 2010