DOI: 10.1214/ECP.v15-1582

Elect. Comm. in Probab. **15** (2010), 508–10

ELECTRONIC COMMUNICATIONS in PROBABILITY

SHARP TAIL INEQUALITIES FOR NONNEGATIVE SUBMARTINGALES AND THEIR STRONG DIFFERENTIAL SUBORDINATES

ADAM OSĘKOWSKI [1]

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

email: `ados@mimuw.edu.pl`

Submitted June 29, 2010, *accepted in final form* October 19, 2010

AMS 2000 Subject classification: 60G42, 60G44

Keywords: Submartingale, Weak-type inequality, Strong differential subordination

Abstract

Let $f = (f_n)_{n\geq 0}$ be a nonnegative submartingale starting from x and let $g = (g_n)_{n\geq 0}$ be a sequence starting from y and satisfying

$$|dg_n| \leq |df_n|, \quad |\mathbb{E}(dg_n|\mathscr{F}_{n-1})| \leq \mathbb{E}(df_n|\mathscr{F}_{n-1})$$

for $n \geq 1$. We determine the best universal constant $U(x, y)$ such that

$$\mathbb{P}(\sup_n g_n \geq 0) \leq ||f||_1 + U(x, y).$$

As an application, we deduce a sharp weak type $(1, 1)$ inequality for the one-sided maximal function of g and determine, for any $t \in [0, 1]$ and $\beta \in \mathbb{R}$, the number

$$L(x, y, t, \beta) = \inf\{||f||_1 : \mathbb{P}(\sup_n g_n \geq \beta) \geq t\}.$$

The estimates above yield analogous statements for stochastic integrals in which the integrator is a nonnegative submartingale. The results extend some earlier work of Burkholder and Choi in the martingale setting.

1 Introduction

The purpose of this paper is to study some new sharp estimates for submartingales and their differential subordinates. Let us start with introducing the necessary background and notation. In what follows, $(\Omega, \mathscr{F}, \mathbb{P})$ is a non-atomic probability space, filtered by a nondecreasing family $(\mathscr{F}_n)_{n\geq 0}$ of sub-σ-fields of $\mathscr{F}$. Let $f = (f_n)_{n\geq 0}$ be an adapted sequence of integrable variables. Then $df = (df_n)_{n\geq 0}$, the difference sequence of f, is given by

$$df_0 = f_0, \quad \text{and} \quad df_n = f_n - f_{n-1} \text{ for } n \geq 1.$$

[1] PARTIALLY SUPPORTED BY MNISW GRANT N N201 397437 AND FOUNDATION FOR POLISH SCIENCE.

Assume that $g=(g_n)_{n\geq 0}$ is another adapted integrable sequence, satisfying

$$|dg_n|\leq |df_n| \quad \text{and} \quad |\mathbb{E}(dg_n|\mathscr{F}_{n-1})|\leq |\mathbb{E}(df_n|\mathscr{F}_{n-1})|, \quad n\geq 1. \tag{1.1}$$

Following Burkholder [5], we say that g is strongly differentially subordinate to f, if $|g_0|\leq |f_0|$ and the condition (1.1) holds. For example, this is the case when g is a transform of f by a predictable sequence $v=(v_n)_{n\geq 0}$, bounded in absolute value by 1. That is, we have $dg_n=v_n df_n$ for $n\geq 0$ and by predictability we mean that for each n the variable v_n is measurable with respect to $\mathscr{F}_{(n-1)\vee 0}$. Let us also mention that if f is a martingale, then the strong differential subordination is equivalent to saying that g is a martingale satisfying $|dg_n|\leq |df_n|$ for all $n\geq 0$.
Let $|g|^*=\sup_n |g_n|$, $g^*=\sup_n g_n$ denote the maximal function of g and the one-sided maximal function of g, respectively. We will also use the notation $||f||_p=\sup_n ||f_n||_p$ for the p-th norm of the sequence f, $p\geq 1$.
The problem of a sharp comparison of the sizes of f and g under various assumptions on f has been studied extensively by many authors. The literature on the subject is very rich, we refer the interested reader to the papers [3]–[9], [11]–[16] and references therein for more information on the subject, and [1], [2], [10] for applications to Riesz systems and the Beurling-Ahlfors transform. We only mention here a few classical estimates, related to the problem investigated in the paper. In the martingale setting, Burkholder [3] proved the following weak-type inequality.

Theorem 1.1. *Suppose that f is a martingale and g is strongly differentially subordinate to f. Then for any $\lambda>0$,*

$$\lambda\mathbb{P}(|g|^*\geq\lambda)\leq 2||f||_1, \tag{1.2}$$

and the constant 2 is the best possible.

A natural question about the optimal constant above for nonnegative submartingales f was answered by Burkholder in [5].

Theorem 1.2. *Suppose that f is a nonnegative submartingale and g is strongly differentially subordinate to f. Then for any $\lambda>0$,*

$$\lambda\mathbb{P}(|g|^*\geq\lambda)\leq 3||f||_1 \tag{1.3}$$

and the constant 3 is the best possible.

The two results above have been extended and generalized in many directions, see e.g. [3], [6], [9], [11], [12], [15] and [16]. We take the line of research related to the following question, raised by Burkholder in [3]. Suppose that g, a strong differential subordinate to f, has at least probability t of exceeding β; how small can $||f||_1$ be? In the particular case when $t=1$ and f is a martingale, the answer is the following (cf. [3]).

Theorem 1.3. *Suppose that f is a martingale starting from x and g is strongly differentially subordinate to f. If g satisfies the one-sided bound*

$$\mathbb{P}(g^*\geq\beta)=1,$$

then

$$||f||_1\geq |x|\vee(\beta-x)$$

and the expression on the right is the best possible.

This result was generalized by Choi [9] to the case when $t\in[0,1]$ is arbitrary. Precisely, we have the following.

Theorem 1.4. *Suppose that f is a martingale starting from x and g is strongly differentially subordinate to f. If g satisfies the one-sided bound*

$$\mathbb{P}(g^* \geq \beta) \geq t,$$

where $t \in [0,1]$ is a fixed number, then

$$||f||_1 \geq |x| \vee \{\beta - x - [\beta^+(\beta - 2x)^+(1-t)]^{1/2}\}.$$

Again, the bound on the right is the best possible.

2 Main results

Our contribution will be, among other things, to establish a submartingale version of the theorems above. First, we study a more general problem and provide a sharp upper bound for the tail of g^*, which depends not only on $||f||_1$ and f_0, but also on the starting point of g. Throughout, the function $U : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ is given by (3.1) below.

Theorem 2.1. *Let f be a nonnegative submartingale starting from $x \geq 0$ and let g be a sequence starting from $y \in \mathbb{R}$ such that the condition* (1.1) *is satisfied. Then*

$$\mathbb{P}(g^* \geq 0) \leq ||f||_1 + U(x, y) \tag{2.1}$$

and the inequality is sharp.

This will be proved in Sections 3 and 4 below. As an application, we will obtain in Section 5 the following extension of Theorem 1.4. Throughout the paper, the function $L : [0,\infty)\times\mathbb{R}\times[0,1] \to \mathbb{R}$ is given by (5.4) below.

Theorem 2.2. *Let f be a nonnegative submartingale starting from $x \geq 0$ and let g start from $y \in \mathbb{R}$. Suppose that* (1.1) *holds. If g satisfies the one-sided estimate*

$$\mathbb{P}(g^* \geq \beta) \geq t, \tag{2.2}$$

where $t \in [0,1]$ is a fixed number, then

$$||f||_1 \geq L(x, y - \beta, t) \tag{2.3}$$

and the bound is the best possible. In particular, if g is strongly differentially subordinate to f, then we have a sharp inequality

$$||f||_1 \geq L(x, x - \beta, t). \tag{2.4}$$

Finally, Theorem 2.1 leads to another interesting variation of the inequality (1.3), to be proved in Section 5.

Theorem 2.3. *Assume that f is a nonnegative submartingale and g is strongly differentially subordinate to f. Then for any $\lambda > 0$ we have*

$$\lambda\mathbb{P}(g^* \geq \lambda) \leq \frac{8}{3}||f||_1 \tag{2.5}$$

and the constant $8/3$ *is the best possible.*

Comparing this to Theorem 1.2 we see that the constant decreases in the one-sided setting. This is not surprising: a careful study of Burkholder's example in [5] shows that the extremal pair (f,g) in (1.3) is symmetric in the sense that $\mathbb{P}(g^* \geq \lambda) = \mathbb{P}((-g)^* \geq \lambda) = 1/2$. In other words, half of the tail of $|g|^*$ comes from dropping to $-\lambda$; however, in (2.5) we do not take this part into account.

We conclude this section by the observation that the results above yield some new and interesting sharp estimates for stochastic integrals in which the integrator is a nonnegative submartingale. To be more precise, suppose that $(\Omega, \mathscr{F}, \mathbb{P})$ is complete and is equipped with a right-continuous filtration $(\mathscr{F}_t)_{t\geq 0}$. Assume that $X = (X_t)$ is an adapted nonnegative cadlag submartingale and let $H = (H_t)_{t\geq 0}$ be a predictable process taking values in $[-1,1]$. Let $Y = (Y_t)_{t\geq 0}$ denote the Itô integral of H with respect to X, that is,

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

Then standard approximation arguments (see [5]) yield the following.

Theorem 2.4. *(i) If Y satisfies the one-sided bound*

$$\mathbb{P}(Y^* \geq \beta) \geq p,$$

where $p \in [0,1]$ is a fixed number, then

$$||X||_1 \geq L(x, y-\beta, p)$$

and the bound is the best possible.
(ii) For any $\lambda > 0$ we have

$$\lambda \mathbb{P}(Y^* \geq \lambda) \leq \frac{8}{3} ||X||_1$$

and the constant $8/3$ is the best possible.

3 A special function

Consider the following subsets of $[0,\infty) \times \mathbb{R}$:

$$\begin{aligned} D_0 &= \{(x,y) : x + y \geq 0\}, \\ D_1 &= \{(x,y) : (x-8)/3 \leq y < -x\}, \\ D_2 &= \{(x,y) : x - 4 < y < (x-8)/3\}, \\ D_3 &= \big([0,\infty) \times \mathbb{R}\big) \setminus (D_0 \cup D_1 \cup D_2) \end{aligned}$$

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

$$U(x,y) = \begin{cases} 1 - x & \text{if } (x,y) \in D_0, \\ \frac{1}{16}(3x+3y+8)^{1/3}(-5x+3y+8) & \text{if } (x,y) \in D_1, \\ -\frac{1}{4}x(6x-6y-16)^{1/3} & \text{if } (x,y) \in D_2, \\ \frac{2x}{x-y} - x & \text{if } (x,y) \in D_3. \end{cases} \tag{3.1}$$

We also introduce the functions ϕ, $\psi : [0,\infty)\times\mathbb{R}\to\mathbb{R}$ by

$$\phi(x,y)=\begin{cases}-1 & \text{if } (x,y)\in D_0,\\ \frac{1}{4}(3x+3y+8)^{-2/3}(-5x-3y-8) & \text{if } (x,y)\in D_1,\\ \frac{1}{2}(6x-6y-16)^{-2/3}(-4x+3y+8) & \text{if } (x,y)\in D_2,\\ -\frac{2y}{(x-y)^2}-1 & \text{if } (x,y)\in D_3\end{cases}$$

and

$$\psi(x,y)=\begin{cases}0 & \text{if } (x,y)\in D_0,\\ \frac{1}{4}(3x+3y+8)^{-2/3}(x+3y+8) & \text{if } (x,y)\in D_1,\\ \frac{1}{2}x(6x-6y-16)^{-2/3} & \text{if } (x,y)\in D_2,\\ \frac{2x}{(x-y)^2} & \text{if } (x,y)\in D_3.\end{cases}$$

Later on, we will need the following properties of these objects.

Lemma 3.1. *(i) The function U is continuous on its domain and of class C^1 on the set $E=\{(x,y): x>0,\ x+y\neq 0\}$. Furthermore, $\phi=U_x$ and $\psi=U_y$ on E.*
(ii) There is an absolute constant A such that for all $x\geq 0$ and $y\in\mathbb{R}$,

$$|U(x,y)|\leq A+A|x|,\quad |\phi(x,y)|\leq A,\quad |\psi(x,y)|\leq A.$$

(iii) For all $x\geq 0$ and $y\in\mathbb{R}$.

$$U(x,y)\geq 1_{\{y\geq 0\}}-x. \tag{3.2}$$

(iv) For any $x\geq 0$ and $y\in\mathbb{R}$,

$$\phi(x,y)\leq -|\psi(x,y)|. \tag{3.3}$$

Proof. (i) This is straightforward and reduces to tedious verification that the function U and its partial derivatives match appropriately at the common boundaries of the sets D_0-D_3. We omit the details.
(ii) This follows immediately from the formulas for U, ϕ and ψ above (in fact $A=1$ suffices, but we will not need this).
(iii) Observe that $\psi\geq 0$, which gives that the function $y\mapsto U(x,y)$ is nondecreasing for a fixed x. It suffices to note that $\lim_{y\to-\infty}U(x,y)=-x$ and $U(x,0)=1-x$.
(iv) The desired estimate reduces to $\phi+\psi\leq 0$. However, we have

$$\phi(x,y)+\psi(x,y)=\begin{cases}-1 & \text{if } (x,y)\in D_0,\\ -x(3x+3y+8)^{-2/3} & \text{if } (x,y)\in D_1,\\ -\frac{1}{4}(6x-6y-16)^{1/3} & \text{if } (x,y)\in D_2,\\ 2(x-y)^{-1}-1 & \text{if } (x,y)\in D_3\end{cases}$$

and it is evident that all the expressions are nonpositive. □

Lemma 3.2. *(i) For any $x\geq 0$, $y\in\mathbb{R}$ and h, $k\in\mathbb{R}$ satisfying $|h|\geq|k|$ and $x+h\geq 0$ we have*

$$U(x+h,y+k)\leq U(x,y)+\phi(x,y)h+\psi(x,y)k. \tag{3.4}$$

Proof. There is a well-known procedure to establish such an estimate (see e.g. [4]): fix $x \geq 0$, $y \in \mathbb{R}$, $a \in [-1,1]$ and consider a function $G = G_{x,y,a} : [-x,\infty) \to \mathbb{R}$ given by $G(t) = U(x+t, y+at)$. Then the condition (3.4) is equivalent to saying that G is concave. Since U is of class C^1 on the set E (in virtue of part (i) of Lemma 3.1), the concavity is the consequence of the two conditions, which will be proved below:

(a) $G''(t) \leq 0$ for those t, for which $(x+t, y+at)$ lies in the interior of one of the sets D_i, $i = 0, 1, 2, 3$,
(b) $G'(t-) \geq G'(t+)$ for t satisfying $x+t > 0$ and $(x+t, y+at) \notin E$.

By the translation property $G_{x,y,a}(t+s) = G_{x+t,y+at,a}(s)$, valid for all $t \geq -x$ and $s \geq -t-x$, it suffices to establish (a) and (b) for $t = 0$. Let us verify the first condition. If (x,y) lies in D_0^o, the interior of D_0, then $G''(0) = 0$. For $(x,y) \in D_1^o$, a little calculation yields

$$G''(0) = \frac{1}{4}(3x+3y+8)^{-5/3}(a+1)\big[(a-3)(-x+3y+8)+8(a-1)x\big],$$

which is nonpositive: this follows from $|a| \leq 1$ and $-x+3y+8 \geq 0$, coming from the definition of D_1. If $(x,y) \in D_2^o$, then

$$G''(0) = 2(1-a)(6x-6y-16)^{-5/3}\big[-x(a+1)+(-x+3y+8)\big] \leq 0,$$

since $-x+3y+8 \leq 0$, by the definition of D_2. Finally, if (x,y) belongs to the interior of D_3, we have

$$G''(0) = 4(1-a)(x-y)^{-3}\big[-(a+1)x+(x+y)\big] \leq 0,$$

because $x+y \leq 0$. It remains to check (b). The condition $x > 0$, $(x,y) \notin E$ is equivalent to $x > 0$ and $x+y = 0$. If $(x,y) \in \partial D_0 \cap \partial D_1$ (so $y = -x \in [-2,0)$), then after some straightforward computations,

$$G'(0-) = \frac{1}{8}[(a+1)y+4a-4] \geq \frac{1}{4}(a+1)-1 \geq -1 = G'(0+).$$

On the other hand, if $(x,y) \in \partial D_3 \cap \partial D_0$ and $x > 2$, then

$$G'(0-) = \frac{a+1}{8x} - 1 \geq -1 = G'(0+)$$

and we are done. □

Remark 3.3. *By* (3.3) *and the proof of the above lemma, we have that for any y the function $t \mapsto U(t, y+t)$, $t \geq 0$, is nonincreasing.*

Now we turn to the proof of Theorem 2.1.

Proof of (2.1)**.** We will prove that for any nonnegative integer n we have

$$\mathbb{P}(g_n \geq 0) - ||f_n||_1 \leq U(x,y). \tag{3.5}$$

This will yield the claim: to see this, fix $\varepsilon > 0$ and introduce the stopping time $\tau = \inf\{n : g_n \geq -\varepsilon\}$. Note that

$$\{g^* \geq 0\} \subseteq \{g_n \geq -\varepsilon \text{ for some } n\} = \bigcup_{n=0}^{\infty}\{g_{\tau\wedge n} + \varepsilon \geq 0\}.$$

Obviously, the family $(\{g_{\tau\wedge n}+\varepsilon \geq 0\})_{n\geq 0}$ is nondecreasing. In addition, the modified pair $(f,g')=(f_n, g_{\tau\wedge n}+\varepsilon)_{n\geq 0}$ still satisfies the domination relation (1.1): this follows from the identity $dg_n' = 1_{\{\tau\geq n\}}dg_n$, valid for all $n\geq 1$. Hence, applying (3.5) to this pair yields

$$\mathbb{P}(g_n' \geq 0) \leq ||f_n||_1 + U(x,y+\varepsilon) \leq ||f||_1 + U(x,y+\varepsilon)$$

and, in consequence, $\mathbb{P}(g^* \geq 0) \leq \lim_{n\to\infty} \mathbb{P}(g_n' \geq 0) \leq ||f||_1 + U(x,y+\varepsilon)$. It suffices to let $\varepsilon \to 0$ to get (2.1).

Thus it remains to establish (3.5). The key observation is that the sequence $(U(f_n,g_n))_{n=0}^{\infty}$ is an $(\mathscr{F}_n)$-supermartingale. Indeed, by (3.4), applied to $x=f_n$, $y=g_n$, $h=df_{n+1}$ and $k=dg_{n+1}$, we get

$$U(f_{n+1},g_{n+1}) \leq U(f_n,g_n) + \phi(f_n,g_n)df_{n+1} + \psi(f_n,g_n)dg_{n+1}. \tag{3.6}$$

By part (ii) of Lemma 3.1, both sides above are integrable. Apply the conditional expectation with respect to $\mathscr{F}_n$ to obtain that

$$\mathbb{E}(U(f_{n+1},g_{n+1})|\mathscr{F}_n) \leq U(f_n,g_n) + \phi(f_n,g_n)\mathbb{E}(df_{n+1}|\mathscr{F}_n) + \psi(f_n,g_n)\mathbb{E}(dg_{n+1}|\mathscr{F}_n).$$

By (1.1) and (3.3), we have $\phi(f_n,g_n)\mathbb{E}(df_{n+1}|\mathscr{F}_n) \leq -|\psi(f_n,g_n)\mathbb{E}(dg_{n+1}|\mathscr{F}_n)|$, which gives the supermartingale property. Now use the majorization (3.2) to get

$$\mathbb{P}(g_n \geq 0) - \mathbb{E}|f_n| \leq \mathbb{E}U(f_n,g_n) \leq \mathbb{E}U(f_0,g_0) = U(x,y), \tag{3.7}$$

which completes the proof. □

Before we proceed, let us mention here how we have constructed the function U. Note that the assertion of Theorem 2.1 (see also (3.5)) can be rephrased as

$$U(x,y) = \sup\{\mathbb{E}V(f_n,g_n)\}. \tag{3.8}$$

Here $V(x,y) = 1_{\{y\geq 0\}} - x$ and the supremum is taken over all n and all pairs (f,g) starting from (x,y), such that f is a nonnegative submartingale and (1.1) is satisfied. Repeating the arguments from [3] and [4], we are led to the corresponding boundary value problem. Namely, U is the least function on $[0,\infty)\times\mathbb{R}$ which majorizes V on the whole domain and satisfies the following condition: for any $y\in\mathbb{R}$, the functions $t\mapsto U(t,y+t)$ and $t\mapsto U(t,y-t)$ are concave and nonincreasing on $[0,\infty)$. Some experimentation with these two assumptions leads to (3.1). For example, to get that $U(x,y)=1-x$ on D_0, we use the following argument. First, note that if U_1, U_2 are any solutions to the above boundary value problem, then so is their minimum $\min\{U_1,U_2\}$. Applying this to $U_1=U$ and U_2 given by $U_2(x,y)=1-x$, we obtain $U(x,y)\leq 1-x$ for all $x\geq 0$, $y\in\mathbb{R}$ (this bound can also be directly derived from (3.8)). Therefore, equality must hold for $y\geq 0$, since $V(x,y)=1-x$ for these y. If $0<x+y<x$, then consider the half line H of slope -1, passing through (x,y). Let $u(t)=U(t,y-x-t)$, $t\geq 0$, be the restriction of U to H. Then u is concave, $u(t)=1-t$ for small t and $u(t)\geq 1_{\{y-x-t\geq 0\}}-t$ for all t. This implies that $u(t)=1-t$ for all t and hence $U(x,y)=1-x$ if $x+y>0$. Finally, the continuity of U along the lines of slope 1 gives $U(x,y)=1-x$ on the whole D_0. Similar reasoning yields the explicit formula for U on the remaining D_i's.

4 Sharpness of (2.1)

Let $\delta>0$ be a fixed small number, to be specified later. Consider a Markov family (f_n,g_n) on $[0,\infty)\times\mathbb{R}$, with the transities described as follows.

(i) The states $\{(x,y):y\geq 0\}$ and $\{(0,y):y\leq -8/3\}$ are absorbing.
(ii) For $-x\leq y<0$, the state (x,y) leads to $(x+y,0)$ or to $(x-y,2y)$, with probabilities $1/2$.
(iii) The state $(x,y)\in D_1$, $x>0$, leads to $(0,y+x)$ or to $(\frac{3x+3y}{4}+2+\delta, \frac{x+y}{4}-2-\delta)$, with probabilities p_1 and $1-p_1$, where $p_1=(-x+3y+8+4\delta)/(3x+3y+8+4\delta)$.
(iv) The state $(x,y)\in D_2$, $x>0$, leads to $(0,y-x)$ or to $(\frac{3x-3y}{2}-4, \frac{x-y}{2}-4)$, with probabilities p_2 and $1-p_2$, where $p_2=(x-3y-8)/(3x-3y-8)$.
(v) The state $(x,y)\in D_3$, $x>0$, leads to $(0,y-x)$ or to $(\frac{x-y}{2}, \frac{y-x}{2})$, with probabilities p_3 and $1-p_3$, where $p_3=-(x+y)/(x-y)$.
(vi) The state $(0,y)$, $y\in(-8/3,0)$, leads to $(0+2\delta, y+2\delta)$.

Let $x\geq 0$ and $y\in\mathbb{R}$. It is not difficult to check that under the probability measure $\mathbb{P}_{x,y}=\mathbb{P}(\cdot|(f_0,g_0)=(x,y))$, the sequence f is a nonnegative submartingale and g satisfies $dg_n=\pm df_n$ for $n\geq 1$ (so (1.1) is satisfied). In fact, the steps described in (i)–(v) are martingale moves in the sense that

$$\mathbb{E}_{x,y}\big((f_{n+1},g_{n+1})|(f_n,g_n)=(x',y')\big)=(x',y'),$$

provided the conditioning event has nonzero probability and (x',y') belongs to one of the sets from (i)–(v). Using this Markov process we will show that the bound (2.1) is optimal. How did we obtain the transity function? A natural idea which comes into one's mind is to search for such a pair (f,g), for which both estimates in (3.7) become equalities, or "almost" equalities. This implies that equality must also hold in (3.6), so, in other words, the Markov process must move according to the following rule. Assume that $(f_0,g_0)=(x,y)$, $x\neq 0$, and let I be the line segment with endpoints given by the possible values of (f_1,g_1). Then U is linear, or "almost" linear when restricted to I. One easily checks that the steps described in (i)–(v) satisfy this condition.
Set

$$P^\delta(x,y)=\mathbb{P}_{x,y}(g^*\geq 0),\qquad M^\delta(x,y)=\lim_{n\to\infty}\mathbb{E}_{x,y}f_n. \tag{4.1}$$

Usually we will skip the upper index and write P, M instead of P^δ, M^δ, but it should be kept in mind that these functions do depend on δ. We will prove that if this parameter is sufficiently small, then

$$P(x,y)-M(x,y)\ \text{is arbitrarily close to}\ U(x,y). \tag{4.2}$$

This will clearly yield the claim. It is convenient to split the remaining part of the proof into a few steps.
1°. The case $y\geq 0$. Here (4.2) is trivial: $P(x,y)=1$, $M(x,y)=x$ for all δ, and $U(x,y)=1-x$.
2°. The case $x=0$, $y\leq -8/3$. Again, (4.2) is obvious: $P(x,y)=M(x,y)=0$ for all δ, and $U(x,y)=0$.
3°. The case $-x\leq y<0$. For any $\delta>0$ and $n\geq 1$, it is easy to see that $\mathbb{P}_{x,y}(g_n\geq 0)=1-2^{-n}$ and $\mathbb{E}_{x,y}f_n=x$. Letting $n\to\infty$, we get $P(x,y)=1$ and $M(x,y)=x$, which yields (4.2), since $U(x,y)=1-x$.
4°. The case $\{(x,y):y=x/3-8/3,\ x\in(0,2)\}\cup\{(x,y):x=0, y\in(-8/3,0)\}$. This is the most technical part. Let us first deal with

$$A(x):=P(x,x/3-8/3),\qquad \text{and}\qquad B(x):=P(0,4x/3-8/3)$$

for $0<x<2$. We will prove that

$$\lim_{\delta\to 0}A(x)=\frac{2}{3}\left(\frac{x}{2}\right)^{1/3}+\frac{1}{3}\left(\frac{x}{2}\right)^{4/3},\quad \lim_{\delta\to 0}B(x)=\frac{4}{3}\left(\frac{x}{2}\right)^{1/3}-\frac{1}{3}\left(\frac{x}{2}\right)^{4/3}. \tag{4.3}$$

This will be done by showing that

$$\lim_{\delta\to 0}\left[A(x)+B(x)\right]=2(x/2)^{1/3},\quad \lim_{\delta\to 0}[2A(x)-B(x)]=(x/2)^{4/3}. \tag{4.4}$$

To get the first statement above, note that by (iii) and Markov property, we have

$$P\left(x,\frac{x}{3}-\frac{8}{3}\right)=\frac{x}{x+\delta}P\left(x+\delta,\frac{x}{3}-\frac{8}{3}-\delta\right)+\frac{\delta}{x+\delta}P\left(0,\frac{4}{3}x-\frac{8}{3}\right),$$

which can be rewritten in the form

$$A(x)=\frac{x}{x+\delta}P\left(x+\delta,\frac{x}{3}-\frac{8}{3}-\delta\right)+\frac{\delta}{x+\delta}B(x). \tag{4.5}$$

Similarly, using (iv) and Markov property, we get

$$\begin{aligned}P\left(x+\delta,\frac{x}{3}-\frac{8}{3}-\delta\right)&=\frac{x+\delta}{x+3\delta}A(x+3\delta)+\frac{2\delta}{x+3\delta}P\left(0,-\frac{2}{3}x-\frac{8}{3}-2\delta\right)\\&=\frac{x+\delta}{x+3\delta}A(x+3\delta),\end{aligned}$$

where in the last passage we have used 2°. Plugging this into (4.5) yields

$$A(x)=\frac{x}{x+3\delta}A(x+3\delta)+\frac{\delta}{x+\delta}B(x). \tag{4.6}$$

Analogous argumentation, with the use of (iii), (vi) and Markov property, leads to the equation

$$B(x)=\frac{2\delta}{x+3\delta}A(x+3\delta)+\frac{x+\delta}{x+3\delta}B(x+3\delta). \tag{4.7}$$

Adding (4.6) to (4.7) gives

$$A(x)+\frac{x}{x+\delta}B(x)=\frac{x+2\delta}{x+3\delta}\left[A(x+3\delta)+\frac{x+3\delta}{x+4\delta}B(x+3\delta)\right]+c(x,\delta)\cdot\delta^2, \tag{4.8}$$

where

$$c(x,\delta)=-\frac{2B(x+3\delta)}{(x+3\delta)(x+4\delta)},$$

which can be bounded in absolute value by $2/x^2$. Now we use (4.8) several times: if N is the largest integer such that $x+3N\delta<2$, then

$$\begin{aligned}A(x)+\frac{x}{x+\delta}B(x)=&\frac{x+2\delta}{x+3\delta}\cdot\frac{x+5\delta}{x+6\delta}\cdot\ldots\cdot\frac{x+3N\delta-\delta}{x+3N\delta}\times\\&\times\left[A(x+3N\delta)+\frac{x+3N\delta}{x+3N\delta+\delta}B(x+3N\delta)\right]+cN\delta^2,\end{aligned} \tag{4.9}$$

where $|c|<2/x^2$. Now we will study the limit behavior of the terms on the right as $\delta\to 0$. First, note that for any $k=1,2,\ldots,N$,

$$\frac{x+3k\delta-\delta}{x+3k\delta}=\exp\left(-\frac{\delta}{x+3k\delta}\right)\cdot\exp(d(k)\delta^2),$$

where $|d(k)| \leq 1/x^2$; consequently,

$$\prod_{k=1}^{N} \frac{x+3k\delta-\delta}{x+3k\delta} = \exp\left(-\sum_{k=1}^{N} \frac{\delta}{x+3k\delta}\right) \exp(\tilde{d} N\delta^2), \tag{4.10}$$

for some $\tilde{d}$ satisfying $|\tilde{d}| \leq 1/x^2$. Since $N = O(1/\delta)$ for small δ, we conclude that the product in (4.10) converges to $\exp(-\frac{1}{3}\int_x^2 t^{-1}\mathrm{d}t) = (x/2)^{1/3}$ as $\delta \to 0$.
The next step is to show that the expression in the square brackets in (4.9) converges to 2 as δ tends to 0. First observe that

$$B(x+3N\delta) = P\left(0, 4x/3+4N\delta - \frac{8}{3}\right) = P\left(2\delta, 4x/3+4N\delta - \frac{8}{3}+2\delta\right) = 1,$$

where in the first passage we have used the definition of B, in the second we have exploited (vi), and the latter is a consequence of 1° and 3°. To show that $A(x+3N\delta)$ converges to 1, use (4.5), with x replaced by $x+3N\delta$, to get

$$A(x+3N\delta) = \frac{x+3N\delta}{x+3N\delta+\delta} P\left(x+3N\delta+\delta, \frac{x+3N\delta}{3} - \frac{8}{3} - \delta\right) + \frac{\delta}{x+3N\delta+\delta}.$$

Note that by the definition of N, the point under P lies in D_3, and arbitrarily close to the line $y = -x$, if δ is sufficiently small. Thus, by (v) and 3°,

$$P\left(x+3N\delta+\delta, \frac{x+3N\delta}{3} - \frac{8}{3} - \delta\right)$$

can be made arbitrary close to 1, provided δ is small enough. Summarizing, letting $\delta \to 0$ in (4.9) yields the first limit in (4.4). To get the second one, multiply both sides of (4.7) by 1/2, subtract it from (4.6) and proceed as previously.
Next we show that $C(x) := M(x, x/3-8/3)$, $D(x) := M(0, 4x/3-8/3)$, $x \in (0,2)$, satisfy

$$\lim_{\delta\to 0} C(x) = \frac{2}{3}\left(\frac{x}{2}\right)^{1/3} + \frac{4}{3}\left(\frac{x}{2}\right)^{4/3}, \quad \lim_{\delta\to 0} D(x) = \frac{4}{3}\left(\frac{x}{2}\right)^{1/3} - \frac{4}{3}\left(\frac{x}{2}\right)^{4/3}. \tag{4.11}$$

We proceed exactly in the similar manner: arguing as reviously, C and D satisfy the same system of equations as A and B, that is, (4.6) and (4.7). The only difference in the further considerations is that $C(x+3N\delta) \to 2$ and $D(x+3N\delta) \to 0$ as $\delta \to 0$.
The final step is to combine (4.3) and (4.11). We get

$$\lim_{\delta\to 0}\left[A(x) - C(x)\right] = U(x, x/3-8/3), \quad \lim_{\delta\to 0}\left[B(x) - D(x)\right] = U(0, 4x/3-8/3),$$

as desired.
5°. *The remaining* (x,y)*'s.* Now (4.2) is easily deduced from the previous cases using Markov property: the point (x,y) leads, in at most two steps, to the state for which we have already calculated the (limiting) values of P and M. For example, if $(x,y) \in D_3$, we have

$$P(x,y) = p_3 P(0, y-x) + (1-p_3) P\left(\frac{x-y}{2}, \frac{y-x}{2}\right) = 1 - p_3 = \frac{2x}{x-y},$$

$$M(x,y) = p_3 M(0, y-x) + (1-p_3) M\left(\frac{x-y}{2}, \frac{y-x}{2}\right) = x,$$

in view of 1° and 3°. Since $U(x,y) = P(x,y) - M(x,y)$, (4.2) follows. The other states are checked similarly. The proof of the sharpness is complete.

We conclude this section with an observation which follows immediately from the above considerations. It will be needed later in the proof of Theorem 2.2.

Remark 4.1. *The function $Q : [0,\infty)\times\mathbb{R} \to [0,1]$, given by $Q(x,y) = \lim_{\delta\to 0} P(x,y)$, is continuous.*

5 Applications

We start with the following auxiliary fact.

Lemma 5.1. *(i) Suppose that*

$$y < -x < 0 \quad \textit{and} \quad \left(\frac{4x}{x-3y}\right)^{1/3}\frac{2x-2y}{x-3y} < t < 1. \tag{5.1}$$

Then there is a unique positive number $C_0 = C_0(x,y,t)$ satisfying $C_0 \le 8/(x-3y)$ (equivalently, $(C_0x, C_0y) \in \overline{D_1}$) and

$$\frac{1}{16}C_0^2(x+y)(5x-3y) + C_0(x+y) + 4 = t(3C_0(x+y)+8)^{2/3}. \tag{5.2}$$

(ii) Suppose that

$$y < -x < 0 \quad \textit{and} \quad \frac{2x}{x-y} < t \le \left(\frac{4x}{x-3y}\right)^{1/3}\frac{2x-2y}{x-3y}. \tag{5.3}$$

Then there is a unique positive number $C_1 = C_1(x,y,t)$ such that $8/(x-3y) \le C_1 \le 4/(x-y)$ (equivalently, $(C_1x, C_1y) \in \overline{D_2}$) and

$$C_1^2x(x-y) = 2t(6C_1(x-y)-16)^{2/3}.$$

Proof. We will only prove (i), the second part can be established essentially in the same manner. Let

$$F(C) = \frac{1}{16}C^2(x+y)(5x-3y) + C(x+y) + 4 - t(3C(x+y)+8)^{2/3}.$$

It can be verified readily that F' is convex and satisfies $F'(0+) = (x+y)(1-t) < 0$: thus F is either decreasing on $(0, 8/(x-3y))$, or decreasing on $(0, x_0)$ and increasing on $(x_0, 8/(x-3y)$ for some x_0 from $(0, 8/(x-3y))$. To complete the proof, it suffices to note that $F(0) = 4-4t > 0$ and

$$F\left(\frac{8}{x-3y}\right) = 4\left(\frac{4x}{x-3y}\right)^{2/3}\left[\left(\frac{4x}{x-3y}\right)^{1/3}\frac{2x-2y}{x-3y} - t\right] < 0. \quad \square$$

Let $L : [0,\infty)\times\mathbb{R}\times[0,1] \to \mathbb{R}$ be given by

$$L(x,y,t) = \begin{cases} (x-y)/2 & \text{if } t = 1, \\ \frac{t}{C_0} - (3C_0(x+y)+8)^{1/3}(\frac{-5x+3y}{16} + \frac{1}{2C_0}) & \text{if (5.1) holds,} \\ \frac{t}{C_1} + \frac{C_1x}{4}\left(\frac{x(x-y)}{2t}\right)^{1/2} & \text{if (5.3) holds,} \\ x & \text{if } t \le \frac{2x}{x-y}. \end{cases} \tag{5.4}$$

Proof of (2.3) **and** (2.4)**.** Clearly, it suffices to prove the inequality for $\beta = 0$, replacing (f, g) by $(f, g - \beta)$, if necessary. Let $C > 0$ be an arbitrary constant. Application of (2.1) to the sequences Cf, Cg yields $\mathbb{P}(g^* \geq 0) - C||f||_1 \leq U(Cx, Cy)$, which, by (2.2), leads to the bound

$$||f||_1 \geq \frac{t - U(Cx, Cy)}{C}. \tag{5.5}$$

If one maximizes the right-hand side over C, one gets precisely $L(x, y, t)$. This follows from a straightforward but lengthy analysis of the derivative with an aid of the previous lemma. We omit the details. To get (2.4), note that for fixed x and t, the function $L(x, \cdot, t)$ is nonincreasing. This follows immediately from (5.5) and the fact that $U(x, \cdot)$ is nondecreasing, which we have already exploited. □

Sharpness of (2.3) As previously, we may restrict ourselves to $\beta = 0$. Fix $x \geq 0$, $y \in \mathbb{R}$ and $t \in [0, 1]$. If $x + y \geq 0$, then the examples studied in 1° and 3° in the preceding section give equality in (2.3). Hence we may and do assume that $x + y < 0$. Consider three cases.

(i) Suppose that $t \leq 2x/(x - y)$. Take $C > 0$ such that $(Cx, Cy) \in D_3$ and take the Markov pair (f, g), with $(f_0, g_0) = (Cx, Cy)$, from the previous section. Then the pair $(f/C, g/C)$ gives equality in (2.3) (see 5°).

(ii) Let $2x/(x - y) < t < 1$ and take $0 < \varepsilon < 1 - t$. Recall the function Q defined in Remark 4.1. First we will show that

$$Q(Cx, Cy) = t + \varepsilon \qquad \text{for some } C = C(\varepsilon, t) > 0. \tag{5.6}$$

Indeed, for large C we have $(Cx, Cy) \in D_3$, so $P(Cx, Cy) = 2x/(x - y)$ regardless of the value of δ (see 5°), and, in consequence, $Q(Cx, Cy) = 2x/(x - y)$. Similarly, $P(0, 0) = 1$ for any δ, so $Q(0, 0) = 1$. Thus (5.6) follows from Remark 4.1. Another observation, to be needed at the end of the proof, is that

$$\liminf_{\varepsilon \to 0} C(\varepsilon, t) > 0. \tag{5.7}$$

Otherwise, we would have a contradiction with (5.6), Remark 4.1 and the equality $Q(0, 0) = 1$.

Now fix $\delta > 0$ and consider a Markov pair (f, g), starting from (Cx, Cy), studied in the previous section. If δ is taken sufficiently small, then the following two conditions are satisfied: first, by (4.2), we have $\mathbb{P}(g^* \geq 0) - ||f||_1 \geq U(Cx, Cy) - \varepsilon$; second, by the definition of Q, $\mathbb{P}(g^* \geq 0) = P(Cx, Cy) \in (t, t+2\varepsilon)$. In other words, for this choice of δ, the pair $(f/C, g/C)$ starts from (x, y), satisfies (1.1), we have $\mathbb{P}((g/C)^* \geq 0) \geq t$ and

$$||f/C||_1 \leq (t - U(Cx, Cy) + 3\varepsilon)/C \leq L(x, y, t) + 3\varepsilon/C.$$

To get the claim, it suffices to note that ε was arbitrary and that (5.7) holds.

(iii) Finally, assume that $t = 1$. Then the following Markov pair (f, g), starting from (x, y), gives equality in (2.3):

- For $y \geq 0$, the state $(0, y)$ is absorbing.
- if $x \neq 0$, then (x, y) leads to $(0, y + x)$ and $(2x, y - x)$, with probabilities $1/2$.
- if $y < 0$, the state $(0, y)$ leads to $(-y/2, y/2)$.

The analysis is similar to the one presented in the case 3° in the previous section. The details are left to the reader. □

We turn to the proof of the weak type inequality from Theorem 2.3.

Proof of Theorem 2.3. Fix $\lambda > 0$ and apply Theorem 2.1 to martingales $8f/(3\lambda)$ and $8g/(3\lambda) - 8/3$, conditionally on $\mathscr{F}_0$. Taking expectation of both sides, we get

$$\mathbb{P}(g^* \geq \lambda) = \mathbb{P}\left(\left(\frac{8g}{3\lambda} - \frac{8}{3}\right)^* \geq 0\right) \leq \left\|\frac{8f}{3\lambda}\right\|_1 + \mathbb{E}U\left(\frac{8f_0}{3\lambda}, \frac{8g_0}{3\lambda} - \frac{8}{3}\right).$$

Thus it suffices to show that for all points $(x, y) \in [0, \infty) \times \mathbb{R}$ satisfying $|y| \leq |x|$ we have

$$U\left(x, y - \frac{8}{3}\right) \leq 0 \tag{5.8}$$

and that the equality holds for at least one such point. This is straightforward: as already mentioned above, the function $y \mapsto U(x, y)$ is nondecreasing, so

$$U\left(x, y - \frac{8}{3}\right) \leq U\left(x, x - \frac{8}{3}\right) \leq U\left(0, -\frac{8}{3}\right) = 0,$$

where, in the latter inequality, we have used Remark 3.3. □

Acknowledgment

The author would like to thank the anonymous referees for the careful reading of the first version of the manuscript.

References

[1] R. Bañuelos and P. Janakiraman, *L^p-bounds for the Beurling-Ahlfors transform*, Trans. Amer. Math. Soc. **360** No. 7 (2008), 3603–3612. MR2386238

[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. MR1370109

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

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

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

[6] C. Choi, *A weak-type submartingale inequality*, Kobe J. Math. **14** (1997), pp. 109–121. MR1612162

[7] C. Choi, *A norm inequality for Itô processes*, J. Math. Kyoto Univ. **37** Vol. 2 (1997), pp. 229–240. MR1463846

[8] C. Choi, *A sharp bound for Itô processes*, J. Korean Math. Soc. **35** (1998), No. 3, pp. 713–725 MR1660804

[9] K. P. Choi, *Some sharp inequalities for martingale transforms*, Trans. Amer. Math. Soc. **307** No. 1 (1988), 279–300. MR0936817

[10] S. Geiss, S. Montgomery-Smith and E. Saksman, *On singular integral and martingale transforms*, Trans. Amer. Math. Soc. **362** No. 2 (2010), pp. 553–575. MR2551497

[11] W. Hammack, *Sharp maximal inequalities for stochastic integrals in which the integrator is a submartingale*, Proc. Amer. Math. Soc. **124** Vol. 3 (1996), pp. 931–938. MR1307522

[12] A. Osękowski, *Weak type inequality for noncommutative differentially subordinated martingales*, Probability Theory Related Fields **140** (2008), 553 – 568. MR2365484

[13] A. Osękowski, *Sharp LlogL inequalities for differentially subordinated martingales*, Illinois J. Math., **52** Vol. 3 (2009), 745–756. MR2546005

[14] A. Osękowski, *Sharp weak type inequalities for differentially subordinated martingales*, Bernoulli **15** Vol. 3 (2009), 871–897. MR2555203

[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), 1545–1564 (electronic). MR2115376

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