

Maximal Inequalities for Martingales and Their Differential Subordinates

Adam Osękowski

Received: 2 September 2011 / Revised: 5 July 2012 / Published online: 8 November 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract We introduce a method of proving maximal inequalities for Hilbert-space-valued differentially subordinate local martingales. As an application, we prove that if $X = (X_t)_{t \geq 0}$, $Y = (Y_t)_{t \geq 0}$ are local martingales such that Y is differentially subordinate to X , then

$$\|Y\|_1 \leq \beta \left\| \sup_{t \geq 0} |X_t| \right\|_1,$$

where $\beta = 2.585 \dots$ is the best possible.

Keywords Martingale · Maximal inequality · Differential subordination

Mathematics Subject Classification (2000) Primary 60G44 · Secondary 60G42

1 Introduction

Since the works of Kolmogorov, Hardy and Littlewood, Wiener, Doob and many other mathematicians, maximal inequalities have played an important role in analysis and probability. One of the main goals of this paper is to present a method of proving such estimates for continuous-time Hilbert-space-valued local martingales satisfying differential subordination.

We start with introducing the necessary background and notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, filtered by a nondecreasing right-continuous family

A. Osękowski (✉)

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

$(\mathcal{F}_t)_{t \geq 0}$ of sub- σ -fields of \mathcal{F} . In addition, we assume that \mathcal{F}_0 contains all the events of probability 0. Let X, Y be two adapted local martingales, taking values in a certain separable Hilbert space \mathcal{H} with norm $|\cdot|$ and scalar product $\langle \cdot, \cdot \rangle$. With no loss of generality, we may take $\mathcal{H} = \ell^2$. As usual, we assume that the trajectories of the processes are right-continuous and have limits from the left. The symbol $[X, X]$ will stand for the quadratic covariance process of X : this object is given by $[X, X] = \sum_{n=1}^{\infty} [X^n, X^n]$, where X^n denotes the n th coordinate of X and $[X^n, X^n]$ is the usual square bracket of the real-valued martingale X^n (see e.g. Dellacherie and Meyer [15] for details). In what follows, $X^* = \sup_{t \geq 0} |X_t|$ will denote the maximal function of X , we also use the notation $X_t^* = \sup_{0 \leq s \leq t} |X_s|$. Furthermore, for $1 \leq p \leq \infty$, we shall write $\|X\|_p = \sup_{t \geq 0} \|X_t\|_p$ and $\|X\|_p = \sup_{\tau} \|X_{\tau}\|_p$, where the second supremum is taken over all adapted bounded stopping times τ .

Throughout the paper we assume that the process Y is *differentially subordinate* to X . This concept was originally introduced by Burkholder [8] in the discrete-time case: a martingale $g = (g_n)_{n \geq 0}$ is differentially subordinate to $f = (f_n)_{n \geq 0}$, if for any $n \geq 0$ we have $|dg_n| \leq |df_n|$. Here $df = (df_n)_{n \geq 0}$, $dg = (dg_n)_{n \geq 0}$ are the difference sequences of f and g , respectively, given by the equations

$$f_n = \sum_{k=0}^n df_k \quad \text{and} \quad g_n = \sum_{k=0}^n dg_k, \quad n = 0, 1, 2, \dots$$

The extension of the domination to the continuous-time setting is due to Bañuelos and Wang [3] and Wang [23]. We say that Y is *differentially subordinate* to X , if the process $([X, X]_t - [Y, Y]_t)_{t \geq 0}$ is nondecreasing and nonnegative as a function of t . If we treat given discrete-time martingales f, g as continuous-time processes (via $X_t = f_{\lfloor t \rfloor}$ and $Y_t = g_{\lfloor t \rfloor}$, $t \geq 0$), we see this domination is consistent with Burkholder's original definition of differential subordination.

To illustrate this notion, consider the following example. Suppose that X is an \mathcal{H} -valued martingale, H is a predictable process taking values in the interval $[-1, 1]$ and let Y be given as the stochastic integral $Y_t = H_0 X_0 + \int_{0+}^t H_s dX_s$, $t \geq 0$. Then Y is differentially subordinate to X : we have

$$[X, X]_t - [Y, Y]_t = (1 - H_0^2)|X_0|^2 + \int_{0+}^t (1 - H_s^2) d[X, X]_s.$$

Another example for stochastic integrals, which plays an important role in applications (see e.g., [2, 3, 16]), is the following. Suppose that B is a Brownian motion in \mathbb{R}^d and H, K are predictable processes taking values in the matrices of dimensions $m \times d$ and $n \times d$, respectively. For any $t \geq 0$, define

$$X_t = \int_{0+}^t H_s \cdot dB_s \quad \text{and} \quad Y_t = \int_{0+}^t K_s \cdot dB_s.$$

If the Hilbert–Schmidt norms of H and K satisfy $\|K_t\|_{\text{HS}} \leq \|H_t\|_{\text{HS}}$ for all $t > 0$, then Y is differentially subordinate to X : this follows from the identity

$$[X, X]_t - [Y, Y]_t = \int_{0+}^t \left(\|H_s\|_{\text{HS}}^2 - \|K_s\|_{\text{HS}}^2 \right) ds.$$

The differential subordination implies many interesting inequalities comparing the sizes of X and Y . A celebrated result of Burkholder gives the following information on the L^p -norms (see [8, 10, 12, 13, 23]).

Theorem 1.1 *Suppose that X, Y are Hilbert-space-valued local martingales such that Y is differentially subordinate to X . Then*

$$|||Y|||_p \leq (p^* - 1) |||X|||_p, \quad 1 < p < \infty, \quad (1.1)$$

where $p^* = \max\{p, p/(p-1)\}$. The constant is the best possible, even if $\mathcal{H} = \mathbb{R}$.

For $p = 1$, the above moment inequality does not hold with any finite constant, but we have the corresponding weak-type $(1, 1)$ estimate. In fact, we have the following result for a wider range of parameters p , proved by Burkholder [8] for $1 \leq p \leq 2$ and Suh [22] for $p > 2$. See also Wang [23].

Theorem 1.2 *Suppose that X, Y are Hilbert-space-valued local martingales such that Y is differentially subordinate to X . Then*

$$\mathbb{P}(Y^* \geq 1) \leq \frac{2}{\Gamma(p+1)} |||X|||_p^p, \quad 1 \leq p \leq 2,$$

and

$$\mathbb{P}(Y^* \geq 1) \leq \frac{p^{p-1}}{2} |||X|||_p^p, \quad 2 < p < \infty.$$

Both inequalities are sharp, even if $\mathcal{H} = \mathbb{R}$.

There are many other related results, see e.g., the papers [3] and [4] by Bañuelos and Wang, [11] and [13] by Burkholder and consult the references therein. For more recent works, we refer the interested reader to the papers [18–20] by the author, and [6, 7] by Borichev et al. The estimates have found numerous applications in many areas of mathematics, in particular, in the study of the boundedness of various classes of Fourier multipliers (consult, for instance, [1–3, 12, 16, 17]).

There is a general method, invented by Burkholder, which enables one not only to establish various estimates for differentially subordinated martingales, but is also very efficient in determining the optimal constants in such inequalities. The idea is to construct an appropriate special function, an upper solution to a nonlinear problem corresponding to the inequality under investigation, and then to exploit its properties. See the survey [13] for the detailed description of the technique in the discrete-time

setting and consult Wang [23] for the necessary changes which have to be implemented so that the method worked in the continuous-time setting.

The above results can be extended in another, very interesting direction. Namely, in the present paper we will be interested in inequalities involving the maximal functions of X and/or Y . Burkholder [14] modified his technique so that it could be used to study such inequalities for stochastic integrals, and applied it to obtain the following result, which can be regarded as another version of (1.1) for $p = 1$.

Theorem 1.3 *Suppose that X is a real-valued martingale and Y is the stochastic integral, with respect to X , of some predictable real-valued process H taking values in $[-1, 1]$. Then we have the sharp estimate*

$$|||Y|||_1 \leq \gamma ||X^*||_1, \quad (1.2)$$

where $\gamma = 2.536 \dots$ is the unique positive number satisfying

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

As we have already observed above, if X and Y satisfy the assumptions of this theorem, then Y is differentially subordinate to X . An appropriate modification of the proof in [14] shows that the assertion is still valid if we impose this less restrictive condition on the processes. However, the assertion does not hold any more if we pass from the real to the vector valued case. Here is one of the main results of this paper.

Theorem 1.4 *Suppose that X, Y are Hilbert-space-valued local martingales such that Y is differentially subordinate to X . Then*

$$|||Y|||_1 \leq \beta ||X^*||_1, \quad (1.3)$$

where $\beta = 2.585 \dots$ is the unique positive number satisfying

$$\beta = 2 + \log \frac{1 + \beta}{2}. \quad (1.4)$$

The constant β is the best possible, even for discrete-time martingales taking values in a two-dimensional subspace of \mathcal{H} .

This is a very surprising result. In most cases, the inequalities for stochastic integrals of real valued martingales carry over, with unchanged constants, to the corresponding bounds for vector-valued local martingales satisfying differential subordination. In other words, given a sharp inequality for \mathcal{H} -valued differentially subordinated martingales, the extremal processes, i.e. those for which the equality is (almost) attained, can be usually realized as stochastic integrals in which the integrator takes values in one-dimensional subspace of \mathcal{H} . See e.g., the statements of Theorems 1.1 and 1.2. Here the situation is different: the optimal constant does depend on the dimension of the range of X and Y .

Finally, let us mention here another related result. In general, the best constants in non-maximal inequalities for differentially subordinated local martingales do not change when we restrict ourselves to continuous-path processes; see e.g., Section 15 in [8] for the justification of this phenomenon. However, if we study the maximal estimates, the best constants may be different: for example, the passage to continuous-path local martingales reduces the constant γ in (1.2) to $\sqrt{2}$. Specifically, we have the following theorem, which is one of the principal results of [21].

Theorem 1.5 *Assume that X, Y are Hilbert-space-valued, continuous-path local martingales such that Y is differentially subordinate to X . Then*

$$\|Y\|_p \leq \sqrt{\frac{2}{p}} \|X^*\|_p, \quad 1 \leq p \leq 2,$$

and

$$\|Y\|_p \leq (p-1) \|X^*\|_p, \quad 2 < p < \infty.$$

Both inequalities are sharp, even if $\mathcal{H} = \mathbb{R}$.

We have organized the paper as follows. The next section is devoted to an extension of Burkholder's method. In Sect. 3 we apply the technique to establish (1.3). In Sect. 4 we prove that the constant β cannot be replaced in (1.3) by a smaller one. The final part of the paper contains the proofs of technical facts needed in the earlier considerations.

2 On the Method of Proof

Burkholder's method from [14] is a powerful tool for proving maximal inequalities for transforms of discrete-time real-valued martingales. Then the results for the wider setting of stochastic integrals are obtained by the use of approximation theorems of Bichteler [5]. This approach has the advantage that it avoids practically all the technicalities which arise naturally in the study of continuous-time processes. On the other hand, it does not allow to study estimates for (local) martingales under differential subordination; the purpose of this section is to present a refinement of the method which can be used to handle such problems.

The general statement is the following. Let $V : \mathcal{H} \times \mathcal{H} \times [0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$ be a given Borel function and suppose that we want to show the estimate

$$\mathbb{E}V(X_t, Y_t, X_t^*, Y_t^*) \leq 0 \tag{2.1}$$

for any $t \geq 0$ and any \mathcal{H} -valued local martingales X, Y such that Y is differentially subordinate to X . Due to some technical reasons, we shall deal with a slightly different, localized version of (2.1) (see Theorem 2.2 for the precise statement). Let $D = \mathcal{H} \times \mathcal{H} \times (0, \infty) \times (0, \infty)$. Introduce the class $\mathcal{U}(V)$, which consists of all C^2 functions $U : D \rightarrow \mathbb{R}$ satisfying (2.2)–(2.5) below: for any $(x, y, z, w) \in D$,

$$U(x, y, z, w) \leq 0 \quad \text{if } |x| \leq z, |y| \leq \min\{|x|, w\}, \quad (2.2)$$

$$U(x, y, z, w) \geq V(x, y, z, w) \quad \text{if } |x| \leq z, |y| \leq w. \quad (2.3)$$

Furthermore, there is a locally bounded measurable function $c : D \rightarrow [0, \infty)$ such that for all $(x, y, z, w) \in D$ with $|x| \leq z, |y| \leq w$ and all $h, k \in \mathcal{H}$,

$$\begin{aligned} & \langle U_{xx}(x, y, z, w)h, h \rangle + 2 \langle U_{xy}(x, y, z, w)h, k \rangle + \langle U_{yy}(x, y, z, w)k, k \rangle \\ & \leq c(x, y, z, w) (|k|^2 - |h|^2). \end{aligned} \quad (2.4)$$

Finally, for all $(x, y, z, w) \in D$ with $|x| \leq z, |y| \leq w$ and all $h, k \in \mathcal{H}$ with $|k| \leq |h|$,

$$\begin{aligned} & U(x + h, y + k, |x + h| \vee z, |y + k| \vee w) \\ & \leq U(x, y, z, w) + \langle U_x(x, y, z, w), h \rangle + \langle U_y(x, y, z, w), k \rangle. \end{aligned} \quad (2.5)$$

The latter condition implies that

$$\begin{aligned} U_z(x, y, z, w) & \leq 0 \quad \text{if } |x| = z, \\ U_w(x, y, z, w) & \leq 0 \quad \text{if } |y| = w. \end{aligned} \quad (2.6)$$

For example, let us establish the bound for U_w . Pick x, y, z, w with $|y| = w$; by the continuity of U_w , we may and do assume that $|x| < z$. Apply (2.5) to x, y, z, w and $h = k = sy$ for some $s > 0$. Then, take all the terms on one side, divide throughout by s and let $s \rightarrow 0$. Since $|x + sy| \vee z = z$ for sufficiently small s and $|y + sy| \vee w = (1 + s)w$ for all $s > 0$, we obtain the second estimate in (2.6); the bound for U_z is established similarly.

Before we turn to the main result of this section, let us mention here a technical fact, which will be needed later. Recall that for any semimartingale X there exists a unique continuous local martingale part X^c of X satisfying $X_0^c = 0$ and

$$[X, X]_t = |X_0|^2 + [X^c, X^c]_t + \sum_{0 < s \leq t} |\Delta X_s|^2 \quad \text{for } t \geq 0.$$

Here $\Delta X_s = X_s - X_{s-}$ is the jump of X at time s . Furthermore, $[X^c, X^c] = [X, X]^c$, the pathwise continuous part of $[X, X]$. Here is Lemma 1 of Wang [23].

Lemma 2.1 *If X and Y are semimartingales, then Y is differentially subordinate to X if and only if Y^c is differentially subordinate to X^c , $|\Delta Y_t| \leq |\Delta X_t|$ for all $t > 0$ and $|Y_0| \leq |X_0|$.*

We are ready to study the interplay between the class $\mathcal{U}(V)$ and the bound (2.1).

Theorem 2.2 *Assume that $\mathcal{U}(V)$ is nonempty and X, Y are Hilbert-space-valued local martingales such that Y is differentially subordinate to X . Then there is a non-decreasing sequence $(\tau_N)_{N \geq 1}$ of stopping times such that $\lim_{N \rightarrow \infty} \tau_N = \infty$ and*

$$\mathbb{E}V(X_{\tau_N \wedge t}, Y_{\tau_N \wedge t}, X_{\tau_N \wedge t}^* \vee \varepsilon, Y_{\tau_N \wedge t}^* \vee \varepsilon) \leq 0 \quad (2.7)$$

for all $N \geq 1$, $t > 0$, and $\varepsilon > 0$.

Proof Let $(\sigma_n)_{n \geq 1}$ be the localizing sequence for X and Y . Fix $t > 0$, $\varepsilon > 0$, $N \in \{1, 2, \dots\}$ and let

$$\tau_N = \sigma_N \wedge \inf\{s > 0 : |X_s| + |Y_s| + |X_s^c| + |Y_s^c| \geq N\}.$$

Since $X_{\tau_N \wedge t}^c$ is bounded, for any $\delta > 0$ there is $\mathcal{D} = \mathcal{D}(\delta, N, t) \geq 1$ such that

$$\mathbb{E} \sum_{k > \mathcal{D}} [X^{kc}, X^{kc}]_{\tau_N \wedge t} = \mathbb{E} \sum_{k > \mathcal{D}} |X_{\tau_N \wedge t}^{kc}|^2 < \delta. \quad (2.8)$$

For $0 \leq s \leq t$ and $d \geq \mathcal{D}$, put

$$\begin{aligned} X_s^{(d)} &= (X_s^1, X_s^2, \dots, X_s^d, 0, 0, \dots), \\ Y_s^{(d)} &= (Y_s^1, Y_s^2, \dots, Y_s^d, 0, 0, \dots) \end{aligned}$$

and

$$Z_s^{(d)} = (X_s^{(d)}, Y_s^{(d)}, X_s^{(d)*} \vee \varepsilon, Y_s^{(d)*} \vee \varepsilon).$$

There is a sequence $(T_{N,j})_{j \geq 1}$ of stopping times with $T_{N,j} \uparrow \tau_N$, localizing the stochastic integrals $\int U_x(Z_{s-}^{(d)}) \cdot dX_s^{(d)}$, $\int U_y(Z_{s-}^{(d)}) \cdot dY_s^{(d)}$. Since $X^{(d)}$, $Y^{(d)}$ take values in finite-dimensional subspace, we may apply Ito's formula to get

$$U(Z_{T_{N,j} \wedge t}^{(d)}) - U(Z_0^{(d)}) = I_1 + I_2 + I_3/2 + I_4, \quad (2.9)$$

where

$$\begin{aligned} I_1 &= \int_{0+}^{T_{N,j} \wedge t} U_x(Z_{s-}^{(d)}) \cdot dX_s^{(d)} + \int_{0+}^{T_{N,j} \wedge t} U_y(Z_{s-}^{(d)}) \cdot dY_s^{(d)}, \\ I_2 &= \int_{0+}^{T_{N,j} \wedge t} U_z(Z_{s-}^{(d)}) d(X^{(d)*} \vee \varepsilon)_s^c + \int_{0+}^{T_{N,j} \wedge t} U_w(Z_{s-}^{(d)}) d(Y^{(d)*} \vee \varepsilon)_s^c, \\ I_3 &= \int_{0+}^{T_{N,j} \wedge t} U_{xx}(Z_{s-}^{(d)}) d[X^{(d)}, X^{(d)}]_s^c + 2 \int_{0+}^{T_{N,j} \wedge t} U_{xy}(Z_{s-}^{(d)}) d[X^{(d)}, Y^{(d)}]_s^c \\ &\quad + \int_{0+}^{T_{N,j} \wedge t} U_{yy}(Z_{s-}^{(d)}) d[Y^{(d)}, Y^{(d)}]_s^c, \end{aligned}$$

$$I_4 = \sum_{0 < s \leq T_{N,j} \wedge t} [U(Z_{s-}^{(d)}) - U(Z_{s-}^{(d)}) - \langle U_x(Z_{s-}^{(d)}), \Delta X_s^{(d)} \rangle - \langle U_y(Z_{s-}^{(d)}), \Delta Y_s^{(d)} \rangle].$$

Note that the integrals in I_2 are with respect to the continuous parts of the processes $X^{(d)*} \vee \varepsilon$ and $Y^{(d)*} \vee \varepsilon$; this is due to the lack of the terms $U_z(Z_{s-}^{(d)})\Delta(X_s^{(d)*} \vee \varepsilon)$ and $U_w(Z_{s-}^{(d)})\Delta(Y_s^{(d)*} \vee \varepsilon)$ in I_4 .

Let us analyze the terms I_1 – I_4 . We have $\mathbb{E}I_1 = 0$, since both the stochastic integrals are martingales. Next, $I_2 \leq 0$: by (2.6), we have $U_z(Z_{s-}^{(d)}) \leq 0$ on the set $\{s : |X_{s-}^{(d)}| = X_{s-}^{(d)*} \vee \varepsilon\}$ in which the support of $d(X^{(d)*} \vee \varepsilon)_s^c$ is contained. This gives that the first integral in I_2 is nonpositive, the second one is handled analogously. To deal with I_3 , fix $0 \leq s_0 < s_1 \leq t$. For any $\ell \geq 0$, let $(\eta_i^\ell)_{1 \leq i \leq \ell}$ be a nondecreasing sequence of stopping times with $\eta_0^\ell = s_0$, $\eta_{i_\ell}^\ell = s_1$ such that $\lim_{\ell \rightarrow \infty} \max_{1 \leq i \leq \ell-1} |\eta_{i+1}^\ell - \eta_i^\ell| = 0$. Keeping ℓ fixed, we apply, for each $i = 0, 1, 2, \dots, i_\ell$, the property (2.4) to $x = X_{s_0-}^{(d)}$, $y = Y_{s_0-}^{(d)}$, $z = X_{s_0-}^{(d)*} \vee \varepsilon$, $w = Y_{s_0-}^{(d)*} \vee \varepsilon$ and $h = h_i^\ell = X_{T_{N,j} \wedge \eta_{i+1}^\ell}^{(d)c} - X_{T_{N,j} \wedge \eta_i^\ell}^{(d)c}$, $k = k_i^\ell = Y_{T_{N,j} \wedge \eta_{i+1}^\ell}^{(d)c} - Y_{T_{N,j} \wedge \eta_i^\ell}^{(d)c}$. We sum the obtained $i_\ell + 1$ inequalities and let $\ell \rightarrow \infty$. Using the notation $[S, T]_s^u = [S, T]_u - [S, T]_s$, we may write the result in the form

$$\begin{aligned} & \sum_{m=1}^d \sum_{n=1}^d \left[U_{x_m x_n}(Z_{s_0-}^{(d)}) [X^{mc}, X^{nc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} + 2U_{x_m y_n}(Z_{s_0-}^{(d)}) [X^{mc}, Y^{nc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} \right. \\ & \quad \left. + U_{y_m y_n}(Z_{s_0-}^{(d)}) [Y^{mc}, Y^{nc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} \right] \\ & \leq c \left(Z_{s_0-}^{(d)} \right) \sum_{k=1}^d \left([Y^{kc}, Y^{kc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} - [X^{kc}, X^{kc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} \right) \\ & \leq c \left(Z_{s_0-}^{(d)} \right) \left([Y^c, Y^c]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} - [X^c, X^c]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} + \sum_{k>d} [X^{kc}, X^{kc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} \right) \\ & \leq c \left(Z_{s_0-}^{(d)} \right) \sum_{k>d} [X^{kc}, X^{kc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1} \\ & = c \left(Z_{T_{N,j} \wedge s_0-}^{(d)} \right) \sum_{k>d} [X^{kc}, X^{kc}]_{T_{N,j} \wedge s_0}^{T_{N,j} \wedge s_1}, \end{aligned}$$

where in the third passage we have exploited the differential subordination of Y^c to X^c . From the local boundedness of c and the definition of τ_N , we infer that on the set $\{T_{N,j} > 0\}$, $c(Z_{T_{N,j} \wedge s_0-}^{(d)})$ is bounded by a constant C depending only on N and ε . Thus

$$I_3 \leq C \sum_{k>d} [X^{kc}, X^{kc}]_{T_{N,j} \wedge t},$$

using a standard approximation of integrals by discrete sums. Finally, we see that each summand in I_4 is nonpositive, directly from (2.5) and the fact that $|\Delta Y_s| \leq |\Delta X_s|$, see Lemma 2.1. Consequently,

$$I_4 \leq U\left(Z_{T_{N,j} \wedge t}^{(d)}\right) - U\left(Z_{T_{N,j} \wedge t-}^{(d)}\right) - \left\langle U_x(Z_{T_{N,j} \wedge t-}^{(d)}), \Delta X_{T_{N,j} \wedge t}^{(d)} \right\rangle - \left\langle U_y(Z_{T_{N,j} \wedge t-}^{(d)}), \Delta Y_{T_{N,j} \wedge t}^{(d)} \right\rangle$$

on the set $\{T_{N,j} > 0\}$. Plug all the above estimates into (2.9) and take expectation of both sides. By (2.8), the bound we obtain can be rewritten in the form

$$\begin{aligned} \mathbb{E} \left[U(Z_{T_{N,j} \wedge t-}^{(d)}) - U(Z_0^{(d)}) + \left\langle U_x(Z_{T_{N,j} \wedge t-}^{(d)}), \Delta X_{T_{N,j} \wedge t}^{(d)} \right\rangle \right. \\ \left. + \left\langle U_y(Z_{T_{N,j} \wedge t-}^{(d)}), \Delta Y_{T_{N,j} \wedge t}^{(d)} \right\rangle \right] 1_{\{T_{N,j} > 0\}} \leq C\delta. \end{aligned} \quad (2.10)$$

For fixed N , the random variables $Z_{T_{N,j} \wedge t-}^{(d)}$, $j \geq 1$, $d \geq \mathcal{D}$, are uniformly bounded on $\{\tau_N > 0\}$, in view of the definition of τ_N . Moreover, we have

$$\begin{aligned} |\Delta X_{T_{N,j} \wedge t}^{(d)}| &= |\Delta X_{T_{N,j} \wedge t}^{(d)}| 1_{\{T_{N,j} = \tau_N\}} + |\Delta X_{T_{N,j} \wedge t}^{(d)}| 1_{\{T_{N,j} < \tau_N\}} \\ &\leq |\Delta X_{\tau_N \wedge t}^{(d)}| 1_{\{T_{N,j} = \tau_N\}} + \left(|X_{T_{N,j} \wedge t}^{(d)}| + |X_{T_{N,j} \wedge t-}^{(d)}| \right) 1_{\{T_{N,j} < \tau_N\}} \\ &\leq |\Delta X_{\tau_N \wedge t}^{(d)}| + 2N \end{aligned}$$

and, similarly, $|\Delta Y_{T_{N,j} \wedge t}^{(d)}| \leq |\Delta Y_{\tau_N \wedge t}^{(d)}| + 2N$. The random variables $|\Delta X_{\tau_N \wedge t}^{(d)}|$ and $|\Delta Y_{\tau_N \wedge t}^{(d)}|$ are integrable on $1_{\{\tau_N > 0\}}$, since $(\tau_N)_{N \geq 1}$ localizes X and Y . Thus, if we let $j \rightarrow \infty$ and then $d \rightarrow \infty$ in (2.10), we obtain

$$\begin{aligned} \mathbb{E} \left[U(Z_{\tau_N \wedge t-}) - U(Z_0) + \left\langle U_x(Z_{\tau_N \wedge t-}), \Delta X_{\tau_N \wedge t} \right\rangle \right. \\ \left. + \left\langle U_y(Z_{\tau_N \wedge t-}), \Delta Y_{\tau_N \wedge t} \right\rangle \right] 1_{\{\tau_N > 0\}} \leq C\delta, \end{aligned}$$

by Lebesgue's dominated convergence theorem. Here $Z_s = (X_s, Y_s, X_s^* \vee \varepsilon, Y_s^* \vee \varepsilon)$, $s \geq 0$. Let $\delta \rightarrow 0$ and apply (2.5) to get

$$\mathbb{E} [U(Z_{\tau_N \wedge t}) - U(Z_0)] = \mathbb{E} [U(Z_{\tau_N \wedge t}) - U(Z_0)] 1_{\{\tau_N > 0\}} \leq 0.$$

It remains to use (2.2) and (2.3) to complete the proof. \square

Remark 2.3 A careful inspection of the proof of the above theorem shows that the function U need not be given on the whole $D = \mathcal{H} \times \mathcal{H} \times (0, \infty) \times (0, \infty)$. Indeed, it suffices to define it on a certain neighborhood of the set $\{(x, y, z, w) \in D : |x| \leq z, |y| \leq w\}$ in which the process Z takes its values. This can be further relaxed: if we are allowed to work with those X, Y which are bounded away from 0, then all we need is a C^2 function U given on some neighborhood of $\{(x, y, z, w) \in D : 0 < |x| \leq z, 0 < |y| \leq w\}$, satisfying (2.2)–(2.5) on this set.

3 The Special Function Corresponding to (1.3)

Now we apply the approach described in the previous section to establish (1.3). Let $V : D \rightarrow \mathbb{R}$ be given by $V(x, y, z, w) = |y| - \beta(|x| \vee z)$. Furthermore, put

$$U(x, y, z, w) = z\Phi\left(\frac{|y|^2 - |x|^2}{z^2} + 1\right) \quad (3.1)$$

for $\{(x, y, z, w) \in D : 0 < |x| < \sqrt{|y|^2 + z^2}, |y| > 0\}$, where $\Phi : [0, \infty) \rightarrow \mathbb{R}$ is defined by

$$\Phi(t) = \left(1 + \frac{1}{\beta}\right) \left[\sqrt{t} - \log(1 + \sqrt{t}) - (2 - \log 2) \right].$$

We start with four technical lemmas, which will be proved in Sect. 5.

Lemma 3.1 (i) *We have $\Phi(t) \leq \Phi(1) \leq 0$ for $t \leq 1$.*

(ii) *We have $\Phi(t) \geq \sqrt{t} - \beta$ for $t \geq 0$.*

(iii) *For any $c \geq 0$ the function*

$$f(s) = -\sqrt{s} \log\left(1 + \frac{c}{\sqrt{s}}\right) - (2 - \log 2)\sqrt{s}$$

is convex and nonincreasing.

(iv) *For any $c > 0$, the function*

$$f(s) = \sqrt{s} - c \log\left(1 + \frac{\sqrt{s}}{c}\right)$$

is concave.

Lemma 3.2 *The function $y \mapsto \Phi(|y|^2)$ is convex on \mathcal{H} .*

Lemma 3.3 (i) *For any $y, k \in \mathcal{H}$, we have*

$$(2 - \log 2)(1 - \sqrt{1 + |k|^2}) + (1 - \sqrt{1 + |k|^2}) \log(\sqrt{1 + |k|^2} + |y + k|) \\ + \sqrt{1 + |k|^2} \log(\sqrt{1 + |k|^2}) \leq 0. \quad (3.2)$$

(ii) *For any $y, k \in \mathcal{H}$ with $|y| + 1 \leq \sqrt{1 + |k|^2} + |k| - |y|$ we have*

$$(2 - \log 2)(1 - \sqrt{1 + |k|^2}) + \sqrt{1 + |k|^2} \left[2 \frac{|k| - |y|}{\sqrt{1 + |k|^2}} - \log\left(1 + \frac{|k| - |y|}{\sqrt{1 + |k|^2}}\right) \right] \\ \leq \frac{|k|}{1 + |y|} - \log(1 + |y|). \quad (3.3)$$

Lemma 3.4 Assume that $x, y, h, k \in \mathcal{H}$ and $z > 0$ satisfy $|x| = z$, $\langle x, h \rangle \geq 0$ and $|k| \leq |h|$. Then

$$U(x + h, y + k, |x + h| \vee z, |y + k| \vee w) \leq U(x, y, z, w) + \left(1 + \frac{1}{\beta}\right) \frac{\langle y, k \rangle - \langle x, h \rangle}{|x| + |y|}. \quad (3.4)$$

Equipped with these four lemmas, we turn to the following statement.

Theorem 3.5 The function U belongs to the class $\mathcal{U}(V)$.

Proof We check each of the conditions (2.2)–(2.5) separately.

The estimate (2.2): this follows immediately from the first part of Lemma 3.1.

The property (2.4): we derive that the left-hand side of the estimate equals

$$\frac{|k|^2 - |h|^2}{z + \sqrt{S}} - \frac{(\langle y, k \rangle - \langle x, h \rangle)^2}{2(z + \sqrt{S})^2 \sqrt{S}} \leq \frac{|k|^2 - |h|^2}{z + \sqrt{S}},$$

with $S = |y|^2 - |x|^2 + z^2$. The property follows.

The majorization (2.3): in particular, (2.4) implies that for any h the function $t \mapsto U(x + th, y, z, w)$ is concave on $[t_-, t_+]$, where $t_- = \inf\{t : |x + th| \leq z\}$ and $t_+ = \sup\{t : |x + th| \leq z\}$. Consequently, it suffices to verify (2.3) only for (x, y, z, w) satisfying $|x| = z$. But this reduces to the second part of Lemma 3.1.

The condition (2.5): by homogeneity and continuity of both sides, we may assume that $z = 1$ and $|x| < 1$. Define

$$H(t) = U(x + th, y + tk, |x + th| \vee 1, |y + tk| \vee w)$$

for $t \in \mathbb{R}$ and let t_-, t_+ be as above; note that $t_- < 0$ and $t_+ > 0$. By (2.4), H is concave on $[t_-, t_+]$ and hence (2.5) holds if $|x + h| \leq 1$. Suppose then that $|x + h| > 1$ or, in other words, that $t_+ < 1$. The vector $x' = x + t_+h$ satisfies $\langle x', h \rangle \geq 0$: this is equivalent to $\frac{d}{dt}|x + th|^2|_{t=t_+} \geq 0$. Hence, by (3.4), if we put $y' = y + t_+k$, then

$$\begin{aligned} &U(x + h, y + k, |x + h| \vee 1, |y + k| \vee w) \\ &= U(x' + (1 - t_+)h, y' + (1 - t_+)k, |x + h| \vee 1, |y + k| \vee w) \\ &\leq U(x', y', 1, w) + \left(1 + \frac{1}{\beta}\right) \frac{\langle y', (1 - t_+)k \rangle - \langle x', (1 - t_+)h \rangle}{1 + |y'|} \\ &= H(t_+) + H'_-(t_+)(1 - t_+) \\ &\leq H(0) + H'(0)t_+ + H'(0)(1 - t_+) = H(0) + H'(0). \end{aligned}$$

This is precisely the claim. \square

Proof of (1.3) It suffices to establish the estimate for $X^* \in L^1$, because otherwise there is nothing to prove. Furthermore, we may assume that Y is bounded away from 0. To

see this, consider a new Hilbert space $\mathbb{R} \times \mathcal{H}$ and the martingales (δ, X) and (δ, Y) , with $\delta > 0$. These martingales are bounded away from 0 and (δ, Y) is differentially subordinate to (δ, X) . Having proved (1.3) for these processes, we let $\delta \rightarrow 0$ and get the bound for X and Y , by Lebesgue's dominated convergence theorem.

We must show that for any bounded stopping time τ we have

$$\mathbb{E}|Y_\tau| \leq \beta \mathbb{E}X^*.$$

Now we make use of the methodology described in the previous section (in particular, we exploit Remark 2.3). Since $U \in \mathcal{U}(V)$, the above estimate follows immediately from (2.7), applied to the local martingales $(X_{\tau \wedge t})_{t \geq 0}$, $(Y_{\tau \wedge t})_{t \geq 0}$, and letting $N \rightarrow \infty$, $t \rightarrow \infty$ and $\varepsilon \rightarrow 0$. \square

4 Sharpness

The constant β can be shown to be optimal in (1.3) by the use of appropriate examples, but then the calculations are quite involved. To simplify the proof, we use a different approach. Assume that the probability space is the interval $[0, 1]$ equipped with its Borel subsets and Lebesgue's measure. Suppose that there is $\beta_0 \in (0, \beta)$ with the following property: for any discrete filtration $(\mathcal{F}_n)_{n \geq 0}$ and any adapted martingales f, g taking values in \mathbb{R}^2 such that g is differentially subordinate to f , we have

$$\|g\|_1 \leq \beta_0 \|f^*\|_1. \quad (4.1)$$

We shall show that the validity of this estimate implies the existence of a certain special function, with properties similar to those in the definition of the class $\mathcal{U}(V)$. Then, by proper exploitation of these conditions, we shall deduce that $\beta_0 \geq \beta$.

Recall that a sequence $(f_n)_{n \geq 0}$ is called simple if for any n the term f_n takes only a finite number of values and there is a deterministic N such that $f_N = f_{N+1} = f_{N+2} = \dots = f_\infty$. For any $(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2$, introduce the class $\mathcal{M}(x, y)$ which consists of those simple martingale pairs (f, g) with values in $\mathbb{R}^2 \times \mathbb{R}^2$, which satisfy the following two conditions.

- (i) $(f_0, g_0) \equiv (x, y)$,
- (ii) for any $n \geq 1$ we have $|dg_n| \leq |df_n|$.

Here we also allow the filtration $(\mathcal{F}_n)_{n \geq 0}$ to vary. Let $W : \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty) \rightarrow \mathbb{R} \cup \{\infty\}$ be given by the formula

$$W(x, y, z) = \sup \{ \mathbb{E}|g_\infty| - \beta_0 \mathbb{E}(f^* \vee z) \},$$

where the supremum is taken over all $(f, g) \in \mathcal{M}(x, y)$.

Lemma 4.1 *The function W enjoys the following properties.*

- (i) W is finite.
- (ii) W is homogeneous of order 1: for any $(x, y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty)$ and $\lambda \neq 0$,

$$W(\lambda x, \pm \lambda y, |\lambda|z) = |\lambda|W(x, y, z).$$

- (iii) We have $W(x, y, z) = W(x, y, |x| \vee z)$ for all $(x, y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty)$.
- (iv) We have $W(x, y, z) \geq |y| - \beta_0(|x| \vee z)$ for $(x, y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty)$.
- (v) For fixed $x \in \mathbb{R}^2$ and $z > 0$, the function $y \mapsto W(x, y, z)$ is convex on \mathbb{R}^2 .
- (vi) For any $(x, y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty)$ with $|x| \leq z$, any $h, k \in \mathbb{R}^2$ with $|k| \leq |h|$ and any $s, t > 0$,

$$\frac{s}{s+t}W(x+th, y+tk, z) + \frac{t}{s+t}W(x-sh, y-sk, z) \leq W(x, y, z). \quad (4.2)$$

Proof (i) This follows from (4.1): for any $(f, g) \in \mathcal{M}(x, y)$ the martingale $g - y = (g_n - y)_{n \geq 0}$ is differentially subordinate to f , so for any $z > 0$,

$$\mathbb{E}|g_\infty| - \beta_0\mathbb{E}(f^* \vee z) \leq |y| + \mathbb{E}|g_\infty - y| - \beta_0\mathbb{E}f^* \leq |y|.$$

Taking the supremum over $(f, g) \in \mathcal{M}(x, y)$ yields $W(x, y, z) \leq |y| < \infty$.

- (ii) Use the fact that $(f, g) \in \mathcal{M}(x, y)$ if and only if $(\lambda f, \pm \lambda g) \in \mathcal{M}(\lambda x, \pm \lambda y)$.
- (iii) This follows immediately from the very definition of W .
- (iv) The constant pair (x, y) belongs to $\mathcal{M}(x, y)$.
- (v) Take any $x, y_1, y_2 \in \mathbb{R}^2$, $\alpha \in (0, 1)$ and let $y = \alpha y_1 + (1 - \alpha)y_2$. Pick $(f, g) \in \mathcal{M}(x, y)$ and observe that $(f, g + y_i - y) \in \mathcal{M}(x, y_i)$, $i = 1, 2$. Thus,

$$\begin{aligned} \mathbb{E}|g_\infty| - \beta_0\mathbb{E}(f^* \vee z) &\leq \alpha [\mathbb{E}|g_\infty + y_1 - y| - \beta_0\mathbb{E}(f^* \vee z)] \\ &\quad + (1 - \alpha) [\mathbb{E}|g_\infty + y_2 - y| - \beta_0\mathbb{E}(f^* \vee z)] \\ &\leq \alpha W(x, y_1, z) + (1 - \alpha)W(x, y_2, z). \end{aligned}$$

Taking the supremum over $(f, g) \in \mathcal{M}(x, y)$ gives the desired convexity.

- (vi) This is a consequence of the so-called “splicing argument” of Burkholder (see e.g., in [9, p. 77]). For the convenience of the reader, let us provide the easy proof. Pick $(f^+, g^+) \in \mathcal{M}(x+th, y+tk)$, $(f^-, g^-) \in \mathcal{M}(x-sh, y-sk)$. These two pairs are spliced together into one pair (f, g) as follows: set $(f_0, g_0) \equiv (x, y)$ and (recall that $\Omega = [0, 1]$)

$$\begin{aligned} (f_n, g_n)(\omega) &= (f_{n-1}^+, g_{n-1}^+) \left(\frac{\omega(s+t)}{s} \right) \quad \text{if } \omega \leq \frac{s}{s+t}, \\ (f_{n-1}^-, g_{n-1}^-) &\left(\left(\omega - \frac{s}{s+t} \right) \frac{t+s}{t} \right) \quad \text{if } \omega > \frac{s}{s+t} \end{aligned}$$

for $n = 1, 2, \dots$. It is not difficult to see that (f, g) is a martingale pair with respect to its natural filtration. Furthermore, it is clear that this pair belongs to $\mathcal{M}(x, y)$. Finally, since $|x| \leq z$, we have $f_n^* \vee z = \sup_{1 \leq k \leq n} |f_k| \vee z$ for $n = 1, 2, \dots$ and therefore

$$\begin{aligned} W(x, y, z) &\geq \mathbb{E}|g_\infty| - \beta_0 \mathbb{E}(f^* \vee z) \\ &= \frac{s}{t+s} [\mathbb{E}|g_\infty^+| - \beta_0 \mathbb{E}(f^{+*} \vee z)] + \frac{t}{t+s} [\mathbb{E}|g_\infty^-| - \beta_0 \mathbb{E}(f^{-*} \vee z)]. \end{aligned}$$

It remains to take the supremum over all (f^-, g^-) and (f^+, g^+) to get (4.2).

It will be convenient to work with another special function: for any $r \geq 0$, define

$$\Psi(r) = \inf\{W(x, y, 1) : |x| = 1, |y| = r\}.$$

We shall establish the following property of this object.

Lemma 4.2 *For any $r > 0$ and $\varepsilon > 0$ we have*

$$\Psi(r) \geq \Psi\left(\sqrt{r^2 + \varepsilon}\right) + \frac{\varepsilon \Psi(1)}{2(r+1)}. \quad (4.3)$$

Proof Fix $\delta > 0$. Pick $(x, y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 \times (0, \infty)$ satisfying $|x| = z = 1$, $|y| = r$ and apply (4.2) with $h = x$, $k = -y/r$, $s = \delta$ and $t > 0$. We obtain

$$\begin{aligned} W(x, y, 1) &\geq \frac{\delta}{\delta+t} W(x+tx, y-ty/r, 1) + \frac{t}{\delta+t} W(x-\delta x, y+\delta y/r, 1) \\ &= \frac{\delta}{\delta+t} (1+t) W\left(x, \frac{y-ty/r}{1+t}, 1\right) + \frac{t}{\delta+t} W(x-\delta x, y+\delta y/r, 1), \end{aligned}$$

where we have used parts (ii) and (iii) of Lemma 4.1. By part (v) of that lemma, the function $s \mapsto W(x, sy, 1)$, $s \in \mathbb{R}$, is continuous. Thus, if we let $t \rightarrow \infty$, we get

$$\begin{aligned} W(x, y, 1) &\geq \delta W(x, -y/r, 1) + W(x-\delta x, y+\delta y/r, 1) \\ &\geq \delta \Psi(1) + W(x-\delta x, y+\delta y/r, 1). \end{aligned} \quad (4.4)$$

Now we have come to the point where we use the fact that we are in the vector-valued setting. Namely, we pick a vector $d \in \mathbb{R}^2 \setminus \{0\}$, orthogonal to $y+\delta y/r - (x-\delta x)$. Let $s, t > 0$ be uniquely determined by the equalities $|x-\delta x-sd| = |x-\delta x+td| = 1$. Then

$$\begin{aligned} |y+\delta y/r-sd|^2 - |x-\delta x-sd|^2 &= |y+\delta y/r|^2 - |x-\delta x|^2 \\ &= |y|^2 + 2\delta|y| - 1 + 2\delta, \end{aligned}$$

since, as we have assumed at the beginning, $|y| = r$ and $|x| = 1$. In other words, we have $|y+\delta y/r-sd| = \sqrt{|y|^2 + 2\delta|y| + 2\delta}$ and, similarly, $|y+\delta y/r+td| = \sqrt{|y|^2 + 2\delta|y| + 2\delta}$. Therefore, if we apply (4.2) with $x' := x-\delta x$, $y' := y+\delta y/r$, $z = 1$, $h = k = d$ and s, t as above, and combine it with the definition of Ψ , we get

$$W(x - \delta x, y + \delta y/r, 1) \geq \Psi \left(\sqrt{|y|^2 + 2\delta|y| + 2\delta} \right).$$

Plugging this into (4.4) and taking the infimum over x, y (satisfying $|x| = 1, |y| = r$), we arrive at the estimate

$$\Psi(r) \geq \Psi \left(\sqrt{r^2 + 2\delta(r+1)} \right) + \delta\Psi(1).$$

It suffices to put $\delta = \varepsilon/(2r+2)$ to get the claim. \square

Now we are ready to prove that $\beta_0 \geq \beta$; suppose on contrary that this inequality does not hold. By induction, (4.3) yields

$$\Psi(1) \geq \Psi(\sqrt{1+n\varepsilon}) + \frac{\varepsilon\Psi(1)}{2} \sum_{k=0}^{n-1} \frac{1}{1+\sqrt{1+k\varepsilon}}.$$

Fix $t > 1$, put $\varepsilon = (t^2 - 1)/n$ and let $n \rightarrow \infty$ to obtain

$$\Psi(1) \geq \Psi(t) + \frac{\Psi(1)}{2} \int_1^{t^2} \frac{1}{1+\sqrt{s}} ds = \Psi(t) + \Psi(1) \left(t - \log \left(\frac{1+t}{2} \right) - 1 \right).$$

We have $\Psi(t) \geq t - \beta_0$ by Lemma 4.1 (iv), so the above estimate yields

$$\beta_0 \geq t + \Psi(1) \left(t - \log \left(\frac{1+t}{2} \right) - 2 \right) \quad \text{for all } t > 1. \quad (4.5)$$

Now we shall choose an appropriate t . We have $\Psi(1) < -1$; otherwise, we would let $t \rightarrow \infty$ and obtain the contradiction with the assumption $\beta_0 < \beta$. Furthermore, $\Psi(1) \geq 1 - \beta_0 > -2$. Thus, the number t , determined by the equation $\Psi(1) = -\frac{1+t}{t}$, satisfies $t > 1$. Application of (4.5) with this choice of t gives

$$\beta_0 \geq t - \frac{1+t}{t} \left(t - \log \left(\frac{1+t}{2} \right) - 2 \right).$$

It remains to note that for any $t > 1$ the right-hand side is not smaller than β . This follows from a standard analysis of the derivative. The proof is complete.

5 Proofs of Technical Lemmas

Proof of Lemma 3.1 (i) We have $\Phi'(t) = (1 + \beta^{-1})/(2(1 + \sqrt{t})) > 0$ and $\Phi(1) = -(1 + \beta^{-1}) < 0$.

(ii) The claim is equivalent to $\Psi(t) := \Phi(t^2) - t + \beta \geq 0$ for all $t \geq 0$. We easily check that Ψ is convex on $[0, \infty)$ and, by virtue of (1.4), satisfies $\Psi(\beta) = \Psi'(\beta) = 0$.

(iii) Since $\lim_{s \rightarrow \infty} f'(s) = 0$, it suffices to prove the convexity of f . We have

$$f''(s) = \frac{1}{4s^{3/2}} \left[\log \left(1 + \frac{c}{\sqrt{s}} \right) - \frac{\sqrt{s}}{c + \sqrt{s}} + \frac{s}{(c + \sqrt{s})^2} \right] + \frac{2 - \log 2}{4s^{3/2}}$$

and the expression in the square brackets is nonnegative: indeed, the function

$$x \mapsto \log(1+x) - (1+x)^{-1} + (1+x)^{-2}, \quad x \geq 0,$$

vanishes at 0 and is nondecreasing.

(iv) We compute that $f''(s) = -[4(c + \sqrt{s})^2 \sqrt{s}]^{-1} \leq 0$. \square

Proof of Lemma 3.2 Pick $y_1, y_2 \in \mathcal{H}$ and $\alpha \in (0, 1)$. By the concavity of the logarithm, we have

$$\begin{aligned} & \alpha [|y_1| - \log(1 + |y_1|)] + (1 - \alpha) [|y_2| - \log(1 + |y_2|)] \\ & \geq |\alpha y_1| + |(1 - \alpha)y_2| - \log(1 + |\alpha y_1| + |(1 - \alpha)y_2|). \end{aligned}$$

This can be further bounded from below by

$$|\alpha y_1 + (1 - \alpha)y_2| - \log(1 + |\alpha y_1 + (1 - \alpha)y_2|),$$

since the function $t \mapsto t - \log(1 + t)$ is nondecreasing on $[0, \infty)$. We are done. \square

Proof of Lemma 3.3 (i) This follows easily from the obvious estimates

$$\log(\sqrt{1 + |k|^2} + |y + k|) \geq \log(\sqrt{1 + |k|^2})$$

and

$$1 - \sqrt{1 + |k|^2} \leq -\log(\sqrt{1 + |k|^2}).$$

(ii) For simplicity, we shall write k, y instead of $|k|, |y|$, respectively. We consider two major cases.

Case I: Suppose that

$$\sqrt{1 + k^2} \geq (2 - \log 2)(1 + y). \quad (5.1)$$

Then $\sqrt{1 + k^2} \geq 2 - \log 2$, or $k \geq k_0 := \sqrt{(2 - \log 2)^2 - 1}$. In addition, $\frac{k-y}{\sqrt{1+k^2}} \leq 1$, so using the convexity of the function $\xi(s) = 2s - \log(1 + s)$, $s \geq 0$, we have

$$\xi\left(\frac{k-y}{\sqrt{1+k^2}}\right) \leq \frac{k-y}{\sqrt{1+k^2}} \cdot \xi(1) + \left(1 - \frac{k-y}{\sqrt{1+k^2}}\right) \cdot \xi(0),$$

or

$$2 \frac{k-y}{\sqrt{1+k^2}} - \log\left(1 + \frac{k-y}{\sqrt{1+k^2}}\right) \leq (2 - \log 2) \frac{k-y}{\sqrt{1+k^2}}.$$

Hence it suffices to prove that

$$(2 - \log 2)(1 + k - \sqrt{1 + k^2}) \leq \frac{k}{1 + y} - \log(1 + y) + (2 - \log 2)y. \quad (5.2)$$

We consider three possibilities $y \leq 1$, $1 < y < 2$, and $y \geq 2$ separately.

(1) If $y \leq 1$, then the function

$$k \mapsto (2 - \log 2)(1 + k - \sqrt{1 + k^2}) - \frac{k}{1 + y}, \quad k \geq k_0,$$

is nonincreasing: its derivative at k equals

$$\begin{aligned} (2 - \log 2) \left(1 - \frac{k}{\sqrt{1 + k^2}} \right) - \frac{1}{1 + y} &\leq (2 - \log 2) \left(1 - \frac{k_0}{\sqrt{1 + k_0^2}} \right) - \frac{1}{2} \\ &= -0.03 \dots < 0. \end{aligned}$$

Thus, for $y \leq 1$ all we need is to check (5.2) for k satisfying the equation $\sqrt{1 + k^2} = (2 - \log 2)(1 + y)$. But then the estimate is equivalent to

$$\left(2 - \log 2 - \frac{1}{1 + y} \right) (k - \sqrt{1 + k^2}) \leq -\log(1 + y) + (2 - \log 2)y,$$

and the left-hand side is negative, the right-hand side is nonnegative.

(2) If $1 < y < 2$, then by (5.1) we have $k \geq \sqrt{4(2 - \log 2)^2 - 1} > 2.4$. Consequently, the left-hand side of (5.2) is smaller than $2 - \log 2$, while the right-hand side exceeds

$$\frac{k}{3} + 2 - 2 \log 2 > 0.8 + 2 - 2 \log 2 > 2 - \log 2.$$

(3) Suppose finally, that $y \geq 2$. As previously, the left-hand side of (5.2) is bounded from above by $2 - \log 2$. On the other hand, the right-hand side is larger than $-\log 3 + 2(2 - \log 2) > 2 - \log 2$.

Case II: Now we assume that

$$\sqrt{1 + k^2} < (2 - \log 2)(1 + y). \quad (5.3)$$

The inequality (3.3) is equivalent to $F(k) \leq 2y - \log(1 + y)$, where

$$\begin{aligned} F(k) &= (2 - \log 2)(1 - \sqrt{1 + k^2}) + 2(k - y) \\ &\quad - \sqrt{1 + k^2} \log \left(1 + \frac{k - y}{\sqrt{1 + k^2}} \right) - \frac{k}{1 + y}. \end{aligned}$$

We derive that $F'(k) = J_1 + J_2$, where

$$J_1 = -\frac{k}{\sqrt{1+k^2}} \left[\log \left(1 + \frac{k-y}{\sqrt{1+k^2}} \right) - \frac{\frac{k-y}{\sqrt{1+k^2}}}{1 + \frac{k-y}{\sqrt{1+k^2}}} \right],$$

$$J_2 = \frac{y}{1+y} + \frac{k-y}{\sqrt{1+k^2} + k - y} - \frac{(2 - \log 2)k}{\sqrt{1+k^2}}.$$

Since $\log(1+x) \geq x/(x+1)$ for $x > -1$, we have $J_1 \leq 0$. Furthermore, using the assumption $\sqrt{k^2+1} + k - y \geq 1+y$, we get

$$J_2 \leq \frac{y}{1+y} + \frac{k-y}{1+y} - \frac{(2 - \log 2)k}{\sqrt{1+k^2}} = \frac{k}{\sqrt{1+k^2}} \left(\frac{\sqrt{1+k^2}}{1+y} - (2 - \log 2) \right) < 0,$$

due to (5.3). Hence F is nonincreasing; thus $F(k) \leq F(k_1)$, where k_1 satisfies $\sqrt{1+k_1^2} = (2 - \log 2)(1+y)$; however, by the Case I, $F(k_1) \leq 2y - \log(1+y)$. This completes the proof. \square

Proof of Lemma 3.4 Of course, we may assume that $h \neq 0$. Furthermore, by homogeneity, it suffices to verify the estimate for $z = 1$. It is convenient to split the reasoning into three parts.

Step 1 First we shall show (3.4) in the case when x and h are linearly dependent. Introduce the function $G : [0, \infty) \rightarrow \mathbb{R}$ given by

$$G(t) = |x + th| \Phi \left(\frac{|y + tk|^2}{|x + th|^2} \right).$$

We shall prove that this function is convex. To do this, fix $t_1, t_2 \geq 0$, $\alpha_1, \alpha_2 \in (0, 1)$ with $\alpha_1 + \alpha_2 = 1$, and let $t = \alpha_1 t_1 + \alpha_2 t_2$. Using Lemma 3.2, we get

$$\begin{aligned} \alpha_1 G(t_1) + \alpha_2 G(t_2) &= \alpha_1 |x + t_1 h| \Phi \left(\frac{|y + t_1 k|^2}{|x + t_1 h|^2} \right) + \alpha_2 |x + t_2 h| \Phi \left(\frac{|y + t_2 k|^2}{|x + t_2 h|^2} \right) \\ &\geq (\alpha_1 |x + t_1 h| + \alpha_2 |x + t_2 h|) \Phi \left(\frac{|y + tk|^2}{(\alpha_1 |x + t_1 h| + \alpha_2 |x + t_2 h|)^2} \right) \\ &= |x + th| \Phi \left(\frac{|y + tk|^2}{|x + th|^2} \right) = G(t), \end{aligned}$$

where in the third passage we have exploited the linear dependence of x , h and the inequality $\langle x, h \rangle \geq 0$. Therefore, using the bound $|k| \leq |h|$ and Lemma 3.1 (i),

$$\begin{aligned} \lim_{t \rightarrow \infty} G'(t) &= \lim_{t \rightarrow \infty} \frac{G(t) - G(1)}{t - 1} \\ &= |h| \Phi \left(\frac{|k|^2}{|h|^2} \right) \\ &\leq |h| \Phi(1) \\ &= - \left(1 + \frac{1}{\beta} \right) |h| \\ &= - \left(1 + \frac{1}{\beta} \right) \frac{|y||h| + |h|}{1 + |y|} \\ &\leq \left(1 + \frac{1}{\beta} \right) \frac{\langle y, k \rangle - \langle x, h \rangle}{1 + |y|}. \end{aligned}$$

Consequently,

$$\begin{aligned} &U(x + h, y + k, |x + h| \vee 1, |y + k| \vee w) - U(x, y, 1, w) \\ &= G(1) - G(0) \leq \left(1 + \frac{1}{\beta} \right) \frac{\langle y, k \rangle - \langle x, h \rangle}{1 + |y|}. \end{aligned}$$

Step 2 Next we check (3.4) in the case when x and h are orthogonal. The inequality becomes

$$\begin{aligned} &|y + k| - \sqrt{1 + |h|^2} \log \left(1 + \frac{|y + k|}{\sqrt{1 + |h|^2}} \right) - (2 - \log 2) \sqrt{1 + |h|^2} - \frac{\langle y, k \rangle}{1 + |y|} \\ &\leq |y| - \log(1 + |y|) - (2 - \log 2). \end{aligned} \quad (5.4)$$

As a function of $|h|$, the left-hand side of the inequality is nonincreasing (see Lemma 3.1 (iii)), so it suffices to prove the bound for $|h| = |k|$. Fix $|y|$, $|k|$ and consider the left-hand side as a function F of $\langle y, k \rangle$. This function is concave (Lemma 3.1 (iv)), and

$$F'(\langle y, k \rangle) = \frac{1}{\sqrt{1 + |k|^2} + |y + k|} - \frac{1}{1 + |y|}. \quad (5.5)$$

Now, if $|y| + 1 > \sqrt{1 + |k|^2} + |k| - |y|$, then F' vanishes at $\langle y, k \rangle = (1 + |y|)(1 - \sqrt{1 + |k|^2})$ and hence it suffices to establish (5.4) for y and k satisfying this equation. A little calculation transforms the estimate into (3.2). On the other hand, if $|y| + 1 \leq \sqrt{1 + |k|^2} + |k| - |y|$, then F' is nonpositive on $[-|y||k|, |y||k|]$ and we need to verify (5.4) for y, k satisfying $\langle y, k \rangle = -|y||k|$. Then the bound reduces to (3.3).

Step 3 Finally, we treat (3.4) for general vectors. The bound is equivalent to

$$\begin{aligned} |y+k| - |x+h| \log \left(1 + \frac{|y+k|}{|x+h|} \right) - (2 - \log 2)|x+h| + \frac{\langle x, h \rangle - \langle y, k \rangle}{1 + |y|} \\ \leq |y| - \log(1 + |y|) - (2 - \log 2). \end{aligned} \quad (5.6)$$

For fixed $|x|$, y , h , and k , the left-hand side, as a function of $\langle x, h \rangle$, is convex (see Lemma 3.1 (iii)) and hence it suffices to verify the estimate in the case when $\langle x, h \rangle = \{|x||h|, 0\}$. These cases have been considered in Steps 1 and 2. \square

Acknowledgments This work was partially supported by Polish Ministry of Science and Higher Education (MNiSW) Grant N N201 397437. The author would like to thank the anonymous Referee for the careful reading of the first version of the paper, the comments and remarks, which greatly improved the presentation.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Bañuelos, R., Bogdan, K.: Lévy processes and Fourier multipliers. *J. Funct. Anal.* **250**(1), 197–213 (2007)
2. Bañuelos, R., Méndez-Hernandez, P.J.: Space–time Brownian motion and the Beurling-Ahlfors transform. *Indiana Univ. Math. J.* **52**(4), 981–990 (2003)
3. Bañuelos, R., Wang, G.: Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transforms. *Duke Math. J.* **80**(3), 575–600 (1995)
4. Bañuelos, R., Wang, G.: Davis’s inequality for orthogonal martingales under differential subordination. *Michigan Math. J.* **47**, 109–124 (2000)
5. Bichteler, K.: Stochastic integration and L^p -theory of semimartingales. *Ann. Probab.* **9**, 49–89 (1980)
6. Borichev, A., Janakiraman, P., Volberg, A.: Subordination by orthogonal martingales in L^p and zeros of Laguerre polynomials. *arXiv:1012.0943* (2010)
7. Borichev, A., Janakiraman, P., Volberg, A.: On Burkholder function for orthogonal martingales and zeros of Legendre polynomials. *Am. J. Math.* (to appear)
8. Burkholder, D.L.: Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12**, 647–702 (1984)
9. Burkholder, D.L.: Martingales and Fourier analysis in Banach spaces. In: *Probability and analysis* (Varrenna, 1985). *Lecture Notes in Mathematics*, vol. 1206, pp. 61–108. Springer, Berlin (1986)
10. Burkholder, D.L.: A sharp and strict L^p -inequality for stochastic integrals. *Ann. Probab.* **15**, 268–273 (1987)
11. Burkholder, D.L.: Sharp inequalities for martingales and stochastic integrals. *Colloque Paul Lévy* (Palaiseau, 1987). *Astérisque* **157–158**, 75–94 (1988)
12. Burkholder, D.L.: A proof of Pelczyński’s conjecture for the Haar system. *Studia Math.* **91**, 79–83 (1988)
13. Burkholder, D.L.: Explorations in martingale theory and its applications. In: *École d’Ete de Probabilités de Saint-Flour XIX–1989. Lecture Notes in Mathematics*, vol. 1464, pp. 1–66. Springer, Berlin (1991)
14. Burkholder, D.L.: Sharp norm comparison of martingale maximal functions and stochastic integrals. In: *Proceedings of the Norbert Wiener Centenary Congress, 1994* (East Lansing, MI, 1994), *Proceedings of Symposia in Applied Mathematics*, vol. 52, pp. 343–358. American Mathematical Society, Providence, RI (1997)
15. Dellacherie, C., Meyer, P.A.: *Probabilities and Potential B*. North-Holland, Amsterdam (1982)
16. Geiss, S., Montgomery-Smith, S., Saksman, E.: On singular integral and martingale transforms. *Trans. Am. Math. Soc.* **362**(2), 553–575 (2010)

17. Nazarov, F.L., Volberg, A.: Heat extension of the Beurling operator and estimates for its norm. *Rossiiskaya Akademiya Nauk. Algebra i Analiz* **15**(4), 142–158 (2003)
18. Osękowski, A.: Sharp LlogL inequality for differentially subordinated martingales. *Illinois J. Math.* **52**(3), 745–756 (2008)
19. Osękowski, A.: Sharp weak type inequalities for differentially subordinated martingales. *Bernoulli* **15**(3), 871–897 (2009)
20. Osękowski, A.: Sharp moment inequalities for differentially subordinated martingales. *Studia Math.* **201**, 103–131 (2010)
21. Osękowski, A.: Maximal inequalities for continuous martingales and their differential subordinates. *Proc. Am. Math. Soc.* **139**, 721–734 (2011)
22. Suh, Y.: A sharp weak type (p, p) inequality $(p > 2)$ for martingale transforms and other subordinate martingales. *Trans. Am. Math. Soc.* **357**(4), 1545–1564 (2005)
23. Wang, G.: Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities. *Ann. Probab.* **23**, 522–551 (1995)