Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 66, 1–20.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP953

Sharp estimates for martingale transforms with unbounded transforming sequences*

Tomasz Gałązka† Adam Osękowski‡

Abstract

Suppose that $p, q, r \geq 1$ satisfy the condition $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. The paper contains the identification of the best constants $c_{p,q,r}$ and $C_{p,q,r}$ in the estimates

$$\|g\|_{p,\infty} \leq c_{p,q,r}\|f\|_q\|v^*\|_r, \quad \|g\|_p \leq C_{p,q,r}\|f\|_q\|v^*\|_r$$

where f is an arbitrary Hilbert-space valued martingale, g is its transform by a predictable sequence v, and v^* is the maximal function of v. This is extended to the more general context of differential subordination for continuous-time processes.

Keywords: martingale; transform; differential subordination; Burkholder's method; best constants.
MSC2020 subject classifications: 60G42; 60G44.
Submitted to EJP on February 16, 2022, final version accepted on May 1, 2023.

1 Introduction

The motivation for the results studied in this paper comes from a very natural question about the boundedness of martingale transforms. Let us start with the necessary background and notation. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, equipped with a discrete-time filtration $(\mathcal{F}_n)_{n\geq 0}$, i.e., a nondecreasing family of sub-σ-fields of $\mathcal{F}$. Let $f = (f_n)_{n\geq 0}$ be a martingale taking values in a given separable Hilbert space $\mathcal{H}$, with norm and scalar product denoted by $|\cdot|$ and $\langle \cdot, \cdot \rangle$, respectively: with no loss of generality, we may and do assume that $\mathcal{H} = \ell^2$. The difference sequence $(df_n)_{n\geq 0}$ associated with f is defined by $df_0 = f_0$ and $df_n = f_n - f_{n-1}$ for $n \geq 1$. Let $g = v \cdot f$ be a transform of f by a predictable sequence $v = (v_n)_{n\geq 0}$ with values in $\{-1, 1\}$: that is, we have $dg_n = v_n df_n$ for all $n \geq 0$ and the predictability of v means that for each n, the term v_n is measurable

*T. Gałązka and A. Osękowski are supported by Narodowe Centrum Nauki (Poland), grant no. 2018/30/Q/ST1/00072.

†Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland E-mail: T.Galazka@mimuw.edu.pl

‡Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland E-mail: A.Osekowski@mimuw.edu.pl

with respect to $\mathcal{F}_{(n-1)\vee 0}$. Then, as Burkholder proved in [6], for any $1<p<\infty$ there is an absolute constant C_p for which

$$\|g\|_p \leq C_p \|f\|_p. \tag{1.1}$$

Here we have used the notation $\|f\|_p = \sup_{n\geq 0}\|f_n\|_p$. The best constant C_p above was identified in the later work [8] of Burkholder: we have $C_p = p^*-1$, where $p^* = \max\{p,p'\}$ (here and in what follows, $p' = p/(p-1)$ is the harmonic conjugate to p). This result is a starting point for numerous extensions. For example, one can study a version of (1.1) under the assumption of the so-called differential subordination. Recall that a martingale g is said to be differentially subordinate to f, if for any $n\geq 0$ we have the inequality $|dg_n|\leq |df_n|$ almost surely. Of course, if $g = v\cdot f$ is the transform of f by a predictable sequence v with values in $[-1,1]$, then g is differentially subordinate to f. Burkholder [8] proved that under this less restrictive domination, the L_p estimate holds with the same constant p^*-1. Another extension concerns the weak–type analogue of (1.1): one can ask, for any $1\leq p<\infty$, about the best constant c_p in the inequality

$$\|g\|_{p,\infty} \leq c_p \|f\|_p. \tag{1.2}$$

Here $\|\xi\|_{p,\infty} = \sup_{\lambda>0}(\lambda^p \mathbb{P}(|\xi|\geq\lambda))^{1/p}$ and $\|g\|_{p,\infty} = \sup_n \|g_n\|_{p,\infty}$. This problem was answered by Burkholder [8] for $1\leq p\leq 2$ and Suh [19] for $p>2$.

Theorem 1.1. *Suppose that g is differentially subordinate to f. Then the estimate (1.2) holds with the constant*

$$c_p^p = \begin{cases} 2/\Gamma(p+1) & \text{if } 1\leq p\leq 2,\\ p^{p-1}/2 & \text{if } p>2. \end{cases}$$

The constant is already the best possible if $\mathcal{H}=\mathbb{R}$ and g is assumed to be a transform of f by the deterministic sequence $v_n = (-1)^n$, $n = 0, 1, 2, \ldots$.

From the viewpoint of applications, it is sometimes convenient to consider a different norming of weak L_p spaces. Define

$$|||\xi|||_{p,\infty} = \sup\left\{\mathbb{P}(A)^{1/p-1}\int_A |\xi|\mathrm{d}\mathbb{P}\right\},$$

where the supremum is taken over all events $A\in\mathcal{F}$ of positive probability. It is easy to prove that for $1<p<\infty$ we have $\|\cdot\|_{p,\infty}\leq |||\cdot|||_{p,\infty}\lesssim_p \|\cdot\|_{p,\infty}$ (here and below, the symbol $A\lesssim_p B$ means that there is a constant κ, depending only on p, such that $A\leq\kappa B$). The question about the optimal weak–type constant under this different norming was answered in [15].

Theorem 1.2. *Suppose that g is differentially subordinate to f. Then for any $1<p<\infty$ we have $|||g|||_{p,\infty}\leq c_p'\|f\|_p$, where*

$$(c_p')^p = \begin{cases} \left(\frac{1}{2}\Gamma\left(\frac{2p-1}{p-1}\right)\right)^{p-1} & \text{if } 1<p\leq 2,\\ p^{p-1}/2 & \text{if } p>2. \end{cases}$$

The constant is already the best possible if $\mathcal{H}=\mathbb{R}$ and g is assumed to be a transform of f by the deterministic sequence $v_n = (-1)^n$, $n = 0, 1, 2, \ldots$.

So, we see that the constants c_p and c_p' are the same for $p\geq 2$, and we have the strict inequality $c_p<c_p'$ in the range $1<p<2$; actually, we see that c_p' explodes as $p\downarrow 1$.

We should point out that the above strong- and weak–type estimates have numerous interesting applications in harmonic analysis. For example, they can be used in the study of unconditional constants for bases in L_p or the boundedness properties of wide classes of Fourier multipliers. The literature on the subject is very extensive, we mention here the works [2], the interested reader is referred also to the bibliographic details contained in these papers.

All the results for martingale transforms formulated above concerned the case in which v was bounded by 1. There is a very interesting question about strong- and weak–type estimates under the assumption that the transforming sequence belongs to L_r for some given $r < \infty$. More specifically, the above question is to study, for given parameters p, q and r satisfying $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$, the optimal constants $C_{p,q,r}$ and $c_{p,q,r}$ in the inequalities

$$\|g\|_p \leq C_{p,q,r} \|f\|_q \|v^*\|_r \tag{1.3}$$

and

$$\||g|\|_{p,\infty} \leq c_{p,q,r} \|f\|_q \|v^*\|_r. \tag{1.4}$$

Here v^* is the maximal function of v, defined by $v^* = \sup_n |v_n|$. To the best of our knowledge, this question is completely open. One of the main results of this paper is to give the answer in the full range of parameters. In the formulation below, ϕ is a certain special concave function on $[0, \infty)$, precisely described in Theorem 5.1.

Theorem 1.3. *Suppose that g is the transform of f by a predictable sequence v and let $1 < p,\, q,\, r < \infty$ be parameters satisfying $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$. Then the estimates* (1.3) *and* (1.4) *hold with*

$$C_{p,q,r} = \begin{cases} p-1 & \text{if } p > 2, \\ (q-1)^{-1} & \text{if } p < q < 2, \\ 1 & \text{otherwise} \end{cases}$$

and

$$c_{p,q,r} = \begin{cases} (p^{p-1}/2)^{1/p} & \text{if } p > 2, \\ \left(\frac{q}{r'}\right)^{1/r'} \left(\frac{r'(2-r')}{2(q-r')}\phi(0)^{r'}\right)^{1/p'} & \text{if } p < q < 2, \\ 1 & \text{otherwise.} \end{cases}$$

The constants are the best possible even if $\mathcal{H} = \mathbb{R}$.

For the more explicit expressions controlling $c_{p,q,r}$ in the range $p < q < 2$, consult Lemma 5.2 below. There is a natural question whether the above theorem can be extended to the context of differential subordination. In particular, one needs to provide an appropriate domination principle which would generalize transforming by L_r-valued sequences. We will prove the following statement.

Theorem 1.4. *Suppose that f, g are martingales and v is the predictable sequence such that g is differentially subordinate to $v \cdot f$. Then for any parameters $1 < p, q, r < \infty$ satisfying $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$, the estimates* (1.3) *and* (1.4) *hold true.*

Note that by a simple limiting argument, the above theorem yields the aforementioned sharp weak– and strong–type inequalities of Burkholder and Suh.

In fact, one can establish a version of Theorem 1.4 in the continuous-time setting. To set up the appropriate context, suppose that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete and equip it with a right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ such that $\mathcal{F}_0$ contains all sets of probability zero. Let X, Y be two adapted càdlàg martingales taking values in a separable Hilbert space $\mathcal{H}$. Define the square bracket of X by $[X] = \sum_{j=0}^{\infty} [X^j]$, where X^j stands for the j-th coordinate of X (in $\mathcal{H} = \ell_2$) and $[X^j]$ is the usual quadratic variation process of the real-valued martingale X^j (see Chapters VI and VII in Dellacherie and Meyer [10] or Chapter 4 in Métivier [11] for details). Given a real predictable process H, the symbol $H \cdot X$ will denote the stochastic integral of H with respect to X, i.e.,

$$(H \cdot X)_t = H_0 X_0 + \int_{0+}^t H_s \cdot \mathrm{d}X_s, \qquad t \geq 0.$$

Clearly, this notion is the continuous-time extension of the concept of martingale transforms considered above. Following Bañuelos and Wang [4] and Wang [20], we say that Y is differentially subordinate to X, if the difference $[X]_t - [Y]_t$ is nonnegative and nondecreasing as a function of t. Note that this definition is consistent with the discrete-time differential subordination discussed above. Indeed, if we treat discrete-time martingales f and g as continuous-time processes (with the use of the embedding $X_t = f_{\lfloor t \rfloor}$ and $Y_t = g_{\lfloor t \rfloor}$, $t \geq 0$), then g is differentially subordinate to f if and only if X and Y satisfy the above condition. Consequently, the following result generalizes Theorems 1.3 and 1.4. In analogy to the discrete-time context, we use the notation $\|X\|_p = \sup_{t\geq 0} \|X_t\|_p$, $|||X|||_{p,\infty} = \sup_{t \geq 0} |||X_t|||_{p,\infty}$ and $H^* = \sup_{t\geq 0} |H_t|$.

Theorem 1.5. *Suppose that X, Y are $\mathcal{H}$-valued martingales and let H be a left-continuous process such that Y is differentially subordinate to $H \cdot X$. Then for any parameters $1 < p, q, r < \infty$ satisfying $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$ we have the sharp estimates*

$$\|Y\|_p \leq C_{p,q,r} \|X\|_q \|H^*\|_r \tag{1.5}$$

and

$$|||Y|||_{p,\infty} \leq c_{p,q,r} \|X\|_q \|H^*\|_r. \tag{1.6}$$

We see that in the above statement we consider integrands H with left-continuous trajectories. This is slightly more restrictive than the predictability condition which is typically imposed in the context of stochastic integrals. The reason is that the analysis of the behavior of H^* will be based on Itô's formula, for which this enhanced regularity seems to be necessary.

Let us say a few words about our approach. Typically, estimates for differentially subordinate martingales are studied with the use of the so-called Burkholder's method (sometimes referred to as the Bellman function method, especially in the analytic context). Roughly speaking, the technique allows to deduce a given estimate from the existence of a certain special function on $\mathcal{H} \times \mathcal{H}$, enjoying appropriate size and concavity constraints (see [14] for the detailed exposition). Let us briefly describe the idea in the context of transforms. Suppose that $V : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a given Borel function and we are interested in the estimate

$$\mathbb{E}V(f_n, g_n) \leq 0, \qquad n = 0, 1, 2, \ldots, \tag{1.7}$$

where f is an arbitrary real-valued martingale and g is its transform by the deterministic sequence $((-1)^n)_{n\geq 0}$ (typically, one imposes some additional boundedness properties on f which guarantee that the above expectations exist). To prove the estimate, one searches for a special function $U : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ which enjoys the following three properties:

1° We have $U(x, x) \leq 0$ for all x;

2° We have $U \geq V$ on $\mathbb{R} \times \mathbb{R}$;

3° The function U is concave along lines of slope ± 1.

Then the condition 3° implies that the process $(U(f_n, g_n))_{n\geq 0}$ is a supermartingale, which combined with 2° and 1° yields the desired bound: for any integer n we have

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}U(f_n, g_n) \leq \mathbb{E}U(f_0, g_0) = \mathbb{E}U(f_0, f_0) \leq 0.$$

A similar argument works for Hilbert-space valued martingales satisfying the differential subordination. One can also extend the approach to the continuous-time setting, but then instead of the above chain of inequalities, one needs to apply Itô's formula (see [20]). A beautiful feature of the approach is that the implication can be reversed: the validity of the estimate (1.7) implies the existence of a function U satisfying the appropriate set of requirements. Thus one can try to implement the method in the search of sharp versions of various inequalities.

Let us emphasize here that the set of slopes allowed in 3° is precisely the set in which the transforming sequence takes values. There seems to be no extension of Burkholder's method which would allow the control the L_r-norm of the sequence v. Thus, our first problem is to modify the approach so that it becomes applicable to the above strong- and weak–type estimates. As we shall see, this can be done by proving slightly stronger estimates involving the 'product processes' H_+^*X, Y/H_+^*, and then splitting them appropriately. These stronger inequalities will be established with the classical form of Burkholder's method.

The paper is organized as follows. The next section is devoted to two special estimates which serve as 'building blocks' in the proofs of (1.5) and (1.6). The strong–type estimate (1.5) is established in Section 3, we also prove the sharpness of (1.3) there. The last two sections contain the proof of the weak–type estimate (1.6) and address the sharpness of (1.4).

2 Two auxiliary inequalities

Introduce the domain $D = \{(x, y) \in \mathcal{H} \times \mathcal{H} : |x| + |y| < 1\}$ and let u_1, $u_\infty : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be two special functions, given by

$$u_1(x,y) = \begin{cases} |y|^2 - |x|^2 & \text{if } (x,y) \in D, \\ 1 - 2|x| & \text{if } (x,y) \notin D \end{cases}$$

and

$$u_\infty(x,y) = \begin{cases} 0 & \text{if } (x,y) \in D, \\ (|y|-1)^2 - |x|^2 & \text{if } (x,y) \notin D. \end{cases}$$

The function u_1 was invented by Burkholder in [8], it played the key role in the proof of the weak–type (1,1) estimate for martingale transforms. To the best of our knowledge, the function u_∞ first appeared in [1] and it can be regarded as an appropriate dual to u_1. See the monograph [14] for the detailed discussion and much more on the subject.

Later on, we will need the following property of these functions. Namely, if $(x, y) \in D$ and h, $k \in \mathcal{H}$ satisfy $|k| \le |h|$, then

$$u_1(x+h, y+k) \le u_1(x,y) + \langle u_{1x}(x,y), h\rangle + \langle u_{1y}(x,y), k\rangle \tag{2.1}$$

and similarly,

$$u_\infty(x+h, y+k) \le u_\infty(x,y) + \langle u_{\infty x}(x,y), h\rangle + \langle u_{\infty y}(x,y), k\rangle. \tag{2.2}$$

Here u_{1x}, u_{1y}, $u_{\infty x}$ and $u_{\infty y}$ stand for the appropriate partial derivatives of u_1 and u_∞. Note that (2.2) is equivalent to saying that $u_\infty(x + h, y + k) \le 0$.

Before we proceed, let us record a useful fact, proved by Wang in [20] (see Lemma 1 there). In what follows, the symbol $\Delta\xi_t = \xi_t - \xi_{t-}$ stands for the jump of a process ξ at a time t. Furthermore, for an arbitrary martingale ξ, we will write ξ^c for its unique continuous part. See [10, 11] for detailed exposition.

Lemma 2.1. *If Y is differentially subordinate to X, then Y^c is differentially subordinate to X^c and with probability 1 we have $|Y_0| \le |X_0|$ and $|\Delta Y_t| \le |\Delta X_t|$ for all $t \ge 0$.*

We are ready for the main result of this section. In what follows, H_+^* is the càdlàg maximal function of H, defined by $H_{t+}^* = \inf_{s>t} H_s^*$.

Theorem 2.2. *Let $t \geq 0$. Suppose that X, Y are martingales and H is a left-continuous process such that Y is differentially subordinate to $H \cdot X$.*

(i) If H_0 is bounded away from zero, then

$$\mathbb{E}u_1(X_t, Y_t/H^*_{t+}) \leq 0. \tag{2.3}$$

(ii) If $H^*_{t+}X_t$ and Y_t are square integrable, then

$$\mathbb{E}u_\infty(H^*_{t+}X_t, Y_t) \leq 0. \tag{2.4}$$

Proof of (2.3). Introduce the stopping time $\tau = \inf\{s \geq 0 : (X_s, Y_s/H^*_{s+}) \notin D\}$, with the usual convention $\inf \emptyset = +\infty$. Let us start with the obvious identity

$$\mathbb{E}u_1(X_t, Y_t/H^*_{t+}) = \mathbb{E}u_1(X_t, Y_t/H^*_{t+})\chi_{\{\tau \leq t\}} + \mathbb{E}u_1(X_t, Y_t/H^*_{t+})\chi_{\{\tau > t\}}.$$

Note that $u_1(x, y) \leq 1 - 2|x|$ for all x, $y \in \mathcal{H}$. Therefore, using the supermartingale property of $(1 - 2|X_s|)_{s \geq 0}$, we may write

$$\begin{aligned} \mathbb{E}u_1(X_t, Y_t/H^*_{t+})\chi_{\{\tau \leq t\}} &\leq \mathbb{E}(1 - 2|X_t|)\chi_{\{\tau \leq t\}} \\ &\leq \mathbb{E}(1 - 2|X_\tau|)\chi_{\{\tau \leq t\}} = \mathbb{E}u_1(X_\tau, Y_\tau/H^*_{\tau+})\chi_{\{\tau \leq t\}}, \end{aligned}$$

which combined with the preceding identity gives

$$\mathbb{E}u_1(X_t, Y_t/H^*_{t+}) \leq \mathbb{E}u_1(X_{\tau\wedge t}, Y_{\tau\wedge t}/H^*_{\tau\wedge t+}).$$

Hence it is enough to prove that the right-hand side is nonpositive. To this end, denote $Z_s = (X_s, Y_s/H^*_{s+})$ and apply Itô's formula to obtain

$$u_1(X_{\tau\wedge t}, Y_{\tau\wedge t}/H^*_{\tau\wedge t+}) = I_0 + I_1 + I_2 + I_3/2 + I_4, \tag{2.5}$$

where

$$\begin{aligned} I_0 &= u_1(X_0, Y_0/H^*_{0+}), \\ I_1 &= \int_{0+}^{\tau\wedge t} u_{1x}(Z_{s-}) \cdot \mathrm{d}X_s + \int_{0+}^{\tau\wedge t} \frac{u_{1y}(Z_{s-})}{H^*_s} \cdot \mathrm{d}Y_s, \\ I_2 &= -\int_{0+}^{\tau\wedge t} u_{1y}(Z_{s-}) \cdot \frac{Y_{s-}}{(H^*_s)^2} \mathrm{d}H^*_{s+} + \sum_{0<s\leq\tau\wedge t} u_{1y}(Z_{s-}) \cdot \frac{Y_{s-}}{(H^*_s)^2} \Delta H^*_{s+}, \\ I_3 &= \int_{0+}^{\tau\wedge t} u_{1xx}(Z_{s-}) \cdot \mathrm{d}[X]^c_s + \int_{0+}^{\tau\wedge t} u_{1yy}(Z_{s-})(H^*_s)^{-2} \cdot \mathrm{d}[Y]^c_s, \\ I_4 &= \sum_{0<s\leq\tau\wedge t} \Big[u_1(Z_s) - u_1(Z_{s-}) - \langle u_{1x}(Z_{s-}), \Delta X_s \rangle - \langle u_{1y}(Z_{s-}), (\Delta Y_s)/H^*_s \rangle \Big] \end{aligned}$$

Let us make several helpful observations here. The quantity I_1 and the integral in I_2 is just the sum of all first-order terms, while the expression I_3 is the sum of all second-order terms (note that for $(x, y) \in D$ we have $u_{1xy}(x, y) = 0$, so the mixed integral does not appear in I_3). The second half of I_2 and the whole I_4 correspond to the jump part.

Let us study the behavior of the terms I_0 through I_4. By the differential subordination of Y to $H \cdot X$, we have $|Y_0| \leq |H_0||X_0| \leq H^*_{0+}|X_0|$ and hence $I_0 \leq 0$ (indeed, we have $u_1(x, y) \leq 0$ if $|y| \leq |x|$). The stochastic integrals in I_1 are martingales (as processes indexed by t), since by the definition of τ, Z_- is bounded on $(0, \tau]$. The term I_2 is nonpositive: indeed, we have $u_{1y}(Z_{s-}) \cdot Y_{s-} = 2|Y_{s-}|^2/H^*_{s-} \geq 0$, the process $\left(H^*_{s+} - \sum_{0<u\leq s} \Delta H^*_{u+}\right)_{s\geq 0}$ is nondecreasing, and

$$I_2 = -\int_{0+}^{\tau\wedge t} u_{1y}(Z_{s-}) \cdot \frac{Y_{s-}}{(H^*_s)^2} \mathrm{d}\left(H^*_{s+} - \sum_{0<u\leq s} \Delta H^*_{u+}\right).$$

Next, we compute that

$$I_3 = -2\int_{0+}^{\tau\wedge t} \mathrm{d}[X]_s^c + 2\int_{0+}^{\tau\wedge t} (H_s^*)^{-2}\mathrm{d}[Y]_s^c \leq 0,$$

by the differential subordination and Lemma 2.1 above. Finally, each summand appearing in I_4 is nonpositive, by virtue of (2.1) applied to $(x,y) = Z_{s-}$ and $(h,k) = (\Delta X_s, (\Delta Y_s)/H_s^*)$ (the estimate $|k| \leq |h|$ follows by the differential subordination of Y to $H \cdot X$: see Lemma 2.1).

Putting all the above facts together, we get the desired assertion. □

Proof of (2.4). Let $\tau = \inf\{s \geq 0 : (H_{s+}^* X_s, Y_s) \notin D\}$. The first step is to show that

$$\mathbb{E}u_\infty(H_{\tau\wedge t+}^* X_{\tau\wedge t}, Y_{\tau\wedge t}) \leq 0. \tag{2.6}$$

To this end, we write the trivial identity

$$\begin{aligned}&\mathbb{E}u_\infty(H_{\tau\wedge t+}^* X_{\tau\wedge t}, Y_{\tau\wedge t})\\ &\qquad = \mathbb{E}u_\infty(H_{\tau\wedge t+}^* X_{\tau\wedge t}, Y_{\tau\wedge t})\chi_{\{\tau>t\}} + \mathbb{E}u_\infty(H_{\tau\wedge t+}^* X_{\tau\wedge t}, Y_{\tau\wedge t})\chi_{\{\tau\leq t\}}.\end{aligned}$$

The first summand on the right is equal to zero: by the definitions of u_∞ and the stopping time τ, the random variable under the expectation vanishes. To handle the second summand, we apply, on the set $\{\tau \leq t\}$, the inequality (2.2) with $x = H_\tau^* X_{\tau-}$, $y = Y_{\tau-}$, $h = H_\tau^* \Delta X_\tau$ and $k = \Delta Y_\tau$. Note that $|k| \leq |h|$, by Lemma 2.1 and hence we get $u_\infty(H_\tau^* X_\tau, Y_\tau) \leq 0$. Thus, we also have $u_\infty(H_{\tau+}^* X_\tau, Y_\tau) \leq 0$, since $u_\infty(x,y)$ decreases as $|x|$ increases. Integrating, we get $\mathbb{E}u_\infty(H_{\tau+}^* X_\tau, Y_\tau)\chi_{\{\tau\leq t\}} \leq 0$, which proves (2.6).

The next step is to establish the inequality

$$\mathbb{E}u_\infty(H_{t+}^* X_t, Y_t) \leq \mathbb{E}u_\infty(H_{\tau\wedge t+}^* X_{\tau\wedge t}, Y_{\tau\wedge t}), \tag{2.7}$$

or equivalently,

$$\mathbb{E}u_\infty(H_{t+}^* X_t, Y_t)\chi_{\{\tau\leq t\}} \leq \mathbb{E}u_\infty(H_{\tau+}^* X_\tau, Y_\tau)\chi_{\{\tau\leq t\}}. \tag{2.8}$$

To show this bound, note that $u_\infty(x,y) \leq (|y|-1)^2 - |x|^2$ for all $(x,y) \in \mathcal{H}$, and hence

$$\mathbb{E}u_\infty(H_{t+}^* X_t, Y_t)\chi_{\{\tau\leq t\}} \leq \mathbb{E}\big((|Y_t|-1)^2 - |H_{t+}^* X_t|^2\big)\chi_{\{\tau\leq t\}}.$$

Arguing as above, by Doob's optional sampling theorem and the supermartingale property of the process $(1-2|Y_s|)_{s\geq 0}$, the estimate (2.8) will follow if we manage to prove that

$$\mathbb{E}\big(|Y_t|^2 - |H_{t+}^* X_t|^2\big)\chi_{\{\tau\leq t\}} \leq \mathbb{E}\big(|Y_\tau|^2 - |H_{\tau+}^* X_\tau|^2\big)\chi_{\{\tau\leq t\}}. \tag{2.9}$$

This is done by Itô's formula. The calculations are essentially the same as in the proof of (2.3) and hence we omit the details. □

3 Strong–type estimates

3.1 Proof of (1.5)

With no loss of generality, we may assume that X is bounded in L_q and $H^* \in L_r$, since otherwise there is nothing to prove. Furthermore, we may assume that H_0 is bounded away from zero, replacing it with $|H_0| + \varepsilon$ and letting $\varepsilon \downarrow 0$ at the very end of the proof. These assumptions imply that for each t the random variables $H_{t+}^* X_t$ and Y_t belong to

L_p. Indeed, we have $\|H^*_{t+}X_t\|_p \leq \|X\|_q\|H^*\|_r$ by Young's inequality, while Y is handled with the use of Burkholder-Gundy inequality and the differential subordination:

$$\begin{aligned}\|Y_t\|_p \lesssim_p \|[Y]_t^{1/2}\|_p &\leq \left\| \left(\int_0^t (H_s^*)^2 \mathrm{d}[X]_s \right)^{1/2} \right\|_p \\ &\leq \|H^*_{t+}[X]_t^{1/2}\|_p \leq \|[X]_t^{1/2}\|_q \|H^*_{t+}\|_r \lesssim_q \|X\|_q\|H^*\|_r.\end{aligned} \tag{3.1}$$

Now we consider separately three cases.

The case $p \geq 2$. If $p = 2$, the claim follows from (3.1): all the intermediate inequalities hold with the constant 1. Hence we may restrict ourselves to p strictly bigger than 2. Consider the functions U, $V : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ given by

$$V(x,y) = |y|^p - (p-1)^p|x|^p$$

and

$$U(x,y) = p^{2-p}(p-1)^{p-1}(|y| - (p-1)|x|)(|x| + |y|)^{p-1}.$$

Burkholder [9] showed that we have the majorization

$$U \geq V \qquad \text{on } \mathcal{H} \times \mathcal{H}. \tag{3.2}$$

The function U has the following remarkable representation in the language of u_∞:

$$U(x,y) = \alpha_p \int_0^\infty \lambda^{p-1} u_\infty(x/\lambda, y/\lambda)\mathrm{d}\lambda,$$

where $\alpha_p = p^{3-p}(p-1)^p(p-2)/2$ (see [1]). Therefore, by (2.4) and Fubini's theorem,

$$\mathbb{E}U(H^*_{t+}X_t, Y_t) \leq 0, \qquad t \geq 0. \tag{3.3}$$

To see that Fubini's theorem is applicable, note that

$$|u_\infty(x,y)| \leq \begin{cases} 0 & \text{if } |x| + |y| \leq 1, \\ |x|^2 + |y|^2 & \text{if } |x| + |y| > 1, \end{cases}$$

which implies

$$\int_0^\infty \lambda^{p-1}|u_\infty(x/\lambda, y/\lambda)|\mathrm{d}\lambda \lesssim_p (|x|^2 + |y|^2)(|x| + |y|)^{p-2} \lesssim_p |x|^p + |y|^p. \tag{3.4}$$

Since $H^*_{t+}X_t$ and Y_t belong to L_p, we have the necessary integrability and (3.3) follows. Thus, by (3.2), we conclude that $\mathbb{E}V(H^*_{t+}X_t, Y_t) \leq 0$, or

$$\|Y_t\|_p \leq (p-1)\|H^*_{t+}X_t\|_p \leq (p-1)\|X\|_q\|H^*\|_r. \tag{3.5}$$

Since t was arbitrary, the inequality is established.

Remark 3.1. The latter estimate shows that under the assumptions of Theorem 1.5, we have

$$\|Y\|_p \leq (p-1)\|H^*X\|_p, \qquad p \geq 2.$$

This bound is sharp, since so is its consequence $\|Y\|_p \leq (p-1)\|X\|_q\|H^*\|_r$ (as will be proved below).

The case $p < q \leq 2$. We may actually assume that $q < 2$, the case $q = 2$ follows from a limiting argument. The reasoning goes along similar to those above, but we need some additional effort. Let U_q, $V_q : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ be defined by

$$V_q(x,y) = |y|^q - (q-1)^{-q}|x|^q, \qquad U_q(x,y) = \frac{q^{2-q}}{q-1}((q-1)|y| - |x|)(|x| + |y|)^{q-1}.$$

As shown by Burkholder in [9], we have

$$U_q \geq V_q \qquad \text{on } \mathcal{H} \times \mathcal{H}. \tag{3.6}$$

Furthermore, the function U_q admits the representation (cf. [1])

$$U_q(x,y) = \alpha_q \int_0^\infty \lambda^{q-1} u_1(x/\lambda, y/\lambda) \mathrm{d}\lambda,$$

where $\alpha_q = q^{3-q}(2-q)/2$. Now it is natural to try to use (2.3) and Fubini's theorem to obtain $\mathbb{E}U_q(X_t, Y_t/H^*_{t+}) \leq 0$ for all $t \geq 0$. The function U_q enjoys an appropriate boundedness: we have

$$|u_1(x,y)| \leq \begin{cases} |x|^2 + |y|^2 & \text{if } |x| + |y| \leq 1, \\ |x| + |y| & \text{if } |x| + |y| > 1 \end{cases}$$

and hence

$$\int_0^\infty \lambda^{q-1} |u_1(x/\lambda, y/\lambda)| \mathrm{d}\lambda \lesssim_q |x|^q + |y|^q. \tag{3.7}$$

So, to use Fubini's theorem, we need to establish the L_q-boundedness of the process Y/H^*_+. This, in contrast to the previous situation, does not seem to follow from Burkholder-Gundy inequality. To overcome this difficulty, we apply localization. Given an arbitrary positive integer M, consider the stopping time

$$\sigma_M = \inf\{s \geq 0 : |X_s| + |Y_s/H^*_s| \geq M\}.$$

By the differential subordination of Y to $H \cdot X$, we have

$$|\Delta(Y_{\sigma_M}/H^*_{\sigma_M})| = |\Delta Y_{\sigma_M}|/H^*_{\sigma_M} \leq |\Delta X_{\sigma_M}|,$$

which implies that $|Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t}| \leq M + |\Delta X_{\sigma_M \wedge t}|$, in particular, $Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t}$, and hence also $Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t+}$, belong to L_q. The stopped martingale Y^{σ_M} is differentially subordinate to $H^{\sigma_M} \cdot X^{\sigma_M}$, so (2.3) and Fubini's theorem give

$$\mathbb{E}U_q(X_{\sigma_M \wedge t}, Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t+}) \leq 0, \qquad t \geq 0. \tag{3.8}$$

Combining this estimate with (3.6), we get $\mathbb{E}V_q(X_{\sigma_M \wedge t}, Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t+}) \leq 0$ and hence

$$\begin{aligned} \|Y_{\sigma_M \wedge t}\|_p &\leq \|Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t+}\|_q \|H^*_{\sigma_M \wedge t+}\|_r \\ &\leq (q-1)^{-1}\|X_{\sigma_M \wedge t}\|_q \|H^*_{\sigma_M \wedge t+}\|_r \leq (q-1)^{-1}\|X\|_q \|H^*\|_r. \end{aligned}$$

Letting $M \to \infty$ and $t \to \infty$, we get the claim, by Fatou's lemma.

We conclude with the analogue of Remark 3.1.

Remark 3.2. Under the assumptions of Theorem 1.5, we have the sharp bound

$$\|Y/H^*\|_q \leq (q-1)^{-1}\|X\|_q, \qquad 1 < q \leq 2.$$

This follows directly from the estimate $\mathbb{E}V_q(X_{\sigma_M \wedge t}, Y_{\sigma_M \wedge t}/H^*_{\sigma_M \wedge t+}) \leq 0$ and some simple limiting arguments.

The case $p < 2 < q$. For this choice of p and q, the assertion will follow by applying (1.5) twice, in the range already covered by the above considerations. Specifically, take $s = 2p/(2-p)$, $\alpha = r(2-p)/(2p) < 1$ and write the stochastic integral $H \cdot X$ in the alternative form

$$\int H_t \mathrm{d}X_t = \int |H_t|^\alpha \frac{H_t}{|H_t|^\alpha} \mathrm{d}X_t,$$

i.e., as the stochastic integral of the process $|H|^\alpha$ with respect to the martingale $H|H|^{-\alpha} \cdot X$. So, Y is differentially subordinate to $|H|^\alpha \cdot (H|H|^{-\alpha} \cdot X)$, and hence (1.5), applied with $1/p = 1/2 + 1/s$ (then $C_{p,2,s} = 1$, as we have shown above), gives

$$\|Y\|_p \leq \left\|H|H|^{-\alpha} \cdot X\right\|_2 \|(H^*)^\alpha\|_s = \left\|H|H|^{-\alpha} \cdot X\right\|_2 \|H^*\|_r^{r/s}.$$

The term $\left\|H|H|^{-\alpha} \cdot X\right\|_2$ is again handled by (1.5). Namely, we have $1/2 = 1/q + (q-2)/(2q)$ and $C_{2,q,(2q)/(q-2)} = 1$, so

$$\left\|H|H|^{-\alpha} \cdot X\right\|_2 \leq \|X\|_q \|(H^*)^{1-\alpha}\|_{2q/(q-2)} = \|X\|_q \|H^*\|_r^{r(q-2)/(2q)}.$$

Putting all the above facts together, we get the desired estimate.

3.2 Sharpness for martingale transforms

Observe that the best constant in (1.3) is at least one, for all p, q, r satisfying $1/p = 1/q + 1/r$: this is easily seen by considering the constant sequences $f = g = v \equiv 1$. Therefore, the estimate (1.3) is sharp for $p \leq 2 \leq q$ and from now on we may assume that $p > 2$ or $p < q < 2$. Actually, by the lemma below, we may restrict ourselves to the first possibility.

Lemma 3.3. *Let $C^{tr}_{p,q,r}$ denote the optimal constant in* (1.3)*, restricted to real-valued martingales. Then we have $C^{tr}_{p,q,r} = C^{tr}_{q',p',r}$ for all p, q and r satisfying $1/p = 1/q + 1/r$.*

Proof. Recall that $p' = p/(p-1)$ is the Hölder conjugate to p. Assume that $\varphi = (\varphi_n)_{n\geq 0}$ is an arbitrary $L_{p'}$-bounded, real-valued martingale with $\|\varphi\|_{p'} \leq 1$ and let $\psi = (\psi_n)_{n\geq 0}$ be the transform of φ by v. Since the martingale differences are orthogonal, we may write

$$\mathbb{E} g_n \varphi_n = \mathbb{E} \sum_{k=0}^n dg_k d\varphi_k = \mathbb{E} \sum_{k=0}^n df_k d\psi_k = \mathbb{E} f_n \psi_n.$$

However, we have $1/q' = 1/p' + 1/r$, so

$$\mathbb{E} f_n \psi_n \leq \|f_n\|_q \|\psi_n\|_{q'} \leq C^{tr}_{q',p',r} \|f_n\|_q \|\varphi\|_{p'} \|v^*\|_r \leq C^{tr}_{q',p',r} \|f\|_q \|v^*\|_r.$$

Combining this with the previous identity and using the fact that φ was chosen arbitrarily, we conclude that $\|g_n\|_p \leq C^{tr}_{q',p',r} \|f\|_q \|v^*\|_r$ and hence, taking the supremum over n, we obtain that $C^{tr}_{p,q,r} \leq C^{tr}_{q',p',r}$. Switching from (p,q) to (q',p'), we get the reverse bound. The proof is complete. □

Thus, from now on, we assume that $p > 2$ and proceed to the construction of the appropriate extremal examples. The analysis splits naturally into several steps.

Step 1. The filtered probability space. Assume that the probability space is the interval $(0,1]$ with its Borel subsets and the Lebesgue measure. Let $a > q$ and $0 < \delta < a^{-1}$ be fixed parameters, and set $Q = 1 - a\delta$. We start with defining a certain decreasing sequence $(p_n)_{n\geq 0}$ with values in $(0,1]$. Namely, for any $n \geq 0$ we put

$$p_{2n} = Q^n \quad \text{and} \quad p_{2n+1} = \frac{Q^n + Q^{n+1}}{2} = \frac{p_{2n} + p_{2n+2}}{2}.$$

This sequence gives rise to the filtration $(\mathcal{F}_n)_{n\geq 0}$ such that for a fixed n, the σ-field $\mathcal{F}_n$ is generated by the intervals $(0,p_n]$, $(0,p_{n-1}]$, $(0,p_{n-2}]$, $\ldots$, $(0,p_0]$. That is, the atoms of $\mathcal{F}_n$ are precisely $(0,p_n]$, $(p_n,p_{n-1}]$, $(p_{n-1},p_{n-2}]$, $\ldots$, $(p_1,p_0]$.

Step 2. The variable f. Introduce the function (random variable) $f:(0,1]\to\mathbb{R}$ by

$$f=\sum_{n=0}^{\infty}(1+\delta)^n\Big(\chi_{(p_{2n+1},p_{2n}]}-\chi_{(p_{2n+2},p_{2n+1}]}\Big).$$

Note that f is measurable with respect to $\sigma(\mathcal{F}_n:n\geq 0)$. It is easy to check that f is integrable, it actually belongs to L_q, at least for sufficiently small δ. Indeed, we compute directly that

$$\mathbb{E}|f|^q=\sum_{n=0}^{\infty}(1+\delta)^{nq}(p_{2n}-p_{2n+2})=a\delta\sum_{n=0}^{\infty}\Big[(1+\delta)^q(1-a\delta)\Big]^n<\infty,$$

where the last inequality follows from the estimate $a>q$ (which guarantees that the ratio of the geometric series is less than 1). Furthermore, note that if a is chosen close to q, then

$$\lim_{\delta\to 0}\mathbb{E}|f|^q=\lim_{\delta\to 0}\frac{a\delta}{1-(1+\delta)^q(1-a\delta)}=\frac{a}{a-q}$$

and hence

$$\lim_{a\downarrow q}\lim_{\delta\downarrow 0}\|f\|_q=\infty.$$

Step 3. On the martingale $(f_n)_{n\geq 0}$ generated by f. For any nonnegative integer n, we let $f_n=\mathbb{E}(f|\mathcal{F}_n)$. By the very definition of f and $(\mathcal{F}_n)_{n\geq 0}$, we check that

$$f_{2n}=\begin{cases}0 & \text{on } (0,p_{2n}],\\ f & \text{on } (p_{2n},1].\end{cases}$$

Indeed, on $(0,p_{2n}]$ we have

$$f_{2n}=\frac{1}{|(0,p_{2n}]|}\int_0^{p_{2n}}f\,\mathrm{d}x=0,$$

by symmetry: for each k, the point p_{2k+1} is the middle of (p_{2k+2},p_{2k}). Similarly, we get

$$f_{2n+1}=\begin{cases}-\dfrac{1-Q}{1+Q}(1+\delta)^n & \text{on } (0,p_{2n+1}],\\ f & \text{on } (p_{2n+1},1].\end{cases}$$

To check the first formula, note that $\int_0^{p_{2n+2}}f=0$ (as we have seen above), so on $(0,p_{2n+1}]$,

$$\begin{aligned}f_{2n+1}=\frac{1}{|(0,p_{2n+1}]|}\int_0^{p_{2n+1}}f\,\mathrm{d}x&=\frac{2}{Q^n+Q^{n+1}}\int_{p_{2n+2}}^{p_{2n+1}}f\,\mathrm{d}x\\&=-\frac{2}{Q^n+Q^{n+1}}(1+\delta)^n(p_{2n+1}-p_{2n+2})\\&=-\frac{1-Q}{1+Q}(1+\delta)^n.\end{aligned}$$

Passing to the difference sequence df, we obtain that $df_0 = f_0 = 0$ and

$$df_{2n+1} = \begin{cases} -\dfrac{1-Q}{1+Q}(1+\delta)^n & \text{on } (0, p_{2n+1}], \\ (1+\delta)^n & \text{on } (p_{2n+1}, p_{2n}], \\ 0 & \text{on } (p_{2n}, 1], \end{cases}$$

$$df_{2n+2} = \begin{cases} \dfrac{1-Q}{1+Q}(1+\delta)^n & \text{on } (0, p_{2n+2}], \\ -(1+\delta)^n \cdot \dfrac{2Q}{1+Q} & \text{on } (p_{2n+2}, p_{2n+1}], \\ 0 & \text{on } (p_{2n+1}, 1] \end{cases}$$

for $n = 0, 1, 2, \ldots$.

Step 4. The predictable sequence v and its properties. We introduce $v = (v_n)_{n\geq 0}$ by $v_0 \equiv 1$ and, for $n \geq 0$,

$$v_{2n+1} = -(1+\delta)^{nq/r}\chi_{(0,p_{2n}]}, \qquad v_{2n+2} = (1+\delta)^{nq/r}\chi_{(0,p_{2n+1}]}.$$

Obviously, v is predictable: we have $(0, p_n] \in \mathcal{F}_n$ for each n. Furthermore, on the set $(p_{n+1}, p_n]$ we have $|v_0| \leq |v_1| \leq |v_2| \leq \ldots \leq |v_{n+1}|$ and $v_{n+2} = v_{n+3} = \ldots = 0$. Consequently,

$$v^* = \sum_{n=0}^{\infty} (1+\delta)^{nq/r}\chi_{(p_{2n+2}, p_{2n}]} = |f|^{q/r},$$

so $\|v^*\|_r = \|f\|_q^{q/r}$ and hence in particular $\|f\|_q \|v^*\|_r = \|f\|_q^{q/p}$.

Step 5. On the transform. Let g be the transform of f by v. We will compute the explicit formula for g on each interval of the form $(p_{n+1}, p_n]$. We start with an even n. Directly from the above construction, we see that $df_{2n+2} = 0$ on $(p_{2n+1}, p_{2n}]$ and hence

$$\begin{aligned} g_{2n+2} = g_{2n+1} &= v_0 df_0 + v_1 df_1 + \ldots + v_{2n} df_{2n} + v_{2n+1} df_{2n+1} \\ &= \frac{2(1-Q)}{1+Q}\left[1 + (1+\delta)^{1+q/r} + \ldots + (1+\delta)^{(n-1)(1+q/r)}\right] - (1+\delta)^{n(1+q/r)} \\ &= \frac{2(1-Q)}{1+Q} \cdot \frac{(1+\delta)^{n(1+q/r)} - 1}{(1+\delta)^{1+q/r} - 1} - (1+\delta)^{n(1+q/r)} \\ &= (1+\delta)^{n(1+q/r)}\left[\frac{2}{1+Q} \cdot \frac{1-Q}{(1+\delta)^{1+q/r} - 1} - 1\right] - \frac{2}{1+Q} \cdot \frac{1-Q}{(1+\delta)^{1+q/r} - 1}. \end{aligned}$$

On $(p_{2n+2}, p_{2n+1}]$ the calculations are similar, but slightly more complicated: we get

$$\begin{aligned} g_{2n+2} &= v_0 df_0 + v_1 df_1 + \ldots + v_{2n+1} df_{2n+1} + v_{2n+2} df_{2n+2} \\ &= \frac{2(1-Q)}{1+Q}\left[1 + (1+\delta)^{1+q/r} + \ldots + (1+\delta)^{(n-1)(1+q/r)}\right] \\ &\quad + \frac{1-Q}{1+Q}(1+\delta)^{n(1+q/r)} - (1+\delta)^{n(1+q/r)} \cdot \frac{2Q}{1+Q} \\ &= (1+\delta)^{n(1+q/r)}\left[\frac{2}{1+Q} \cdot \frac{1-Q}{(1+\delta)^{1+q/r} - 1} - 1\right] - \frac{2}{1+Q} \cdot \frac{1-Q}{(1+\delta)^{1+q/r} - 1} \\ &\quad + \frac{2(1-Q)}{1+Q}(1+\delta)^{n(1+q/r)}. \end{aligned}$$

Finally, note that $dg_{n+1} = dg_{n+2} = \ldots = 0$ on $(p_n, p_{n-1}]$. Therefore, we have that g, the pointwise limit of $(g_n)_{n\geq 0}$, can be rewritten in the form $g = g^{(1)} + g^{(2)} + g^{(3)}$, where

$$g^{(1)} = \sum_{n=0}^{\infty}(1+\delta)^{n(1+q/r)}\left[\frac{2}{1+Q}\cdot\frac{1-Q}{(1+\delta)^{1+q/r}-1}-1\right]\chi_{(p_{2n+2},p_{2n}]},$$
$$g^{(2)} = -\sum_{n=0}^{\infty}\frac{2}{1+Q}\cdot\frac{1-Q}{(1+\delta)^{1+q/r}-1}\chi_{(p_{2n+2},p_{2n}]},$$
$$g^{(3)} = \sum_{n=0}^{\infty}\frac{2(1-Q)}{1+Q}(1+\delta)^{n(1+q/r)}\chi_{(p_{2n+2},p_{2n+1}]}.$$

Step 6. The analysis of $g^{(j)}$. Observe that $|g^{(1)}| = |f|^{q/p}\cdot\left[\frac{2}{1+Q}\cdot\frac{1-Q}{(1+\delta)^{1+q/r}-1}-1\right]$ and the expression in the square brackets enjoys the following behavior:

$$\lim_{\delta\to 0}\left[\frac{2}{1+Q}\cdot\frac{1-Q}{(1+\delta)^{1+q/r}-1}-1\right] = \lim_{\delta\to 0}\left(\frac{2a\delta}{(2-a\delta)((1+\delta)^{1+q/r}-1)}-1\right) = \frac{ra}{r+q}-1.$$

The limit $ra/(r+q)-1$ can be made arbitrarily close to $rq/(r+q)-1 = p-1$, if a is chosen sufficiently close to q. Consequently,

$$\lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(1)}\|_p}{\|f\|_q^{q/p}} = \lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(1)}\|_p}{\|f\|_q\|v^*\|_r} = p-1.$$

Now we will show that the contribution of the variables $g^{(2)}$ and $g^{(3)}$ is negligible. Note that

$$|g^{(2)}| = \frac{2(1-Q)}{(1+Q)((1+\delta)^{1+q/r}-1)}$$

is deterministic and converges to $ra/(r+q)$ as $\delta\to 0$. Combining this with the analysis at the end of Step 2, we see that

$$\lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(2)}\|_p}{\|f\|_q^{q/p}} = \lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(2)}\|_p}{\|f\|_q\|v^*\|_r} = 0.$$

Finally, note that $|g^{(3)}| \leq \frac{2(1-Q)}{1+Q}|f|^{q/p}$, and hence

$$\lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(3)}\|_p}{\|f\|_q^{q/p}} = \lim_{a\downarrow q}\lim_{\delta\downarrow 0}\frac{\|g^{(3)}\|_p}{\|f\|_q\|v^*\|_r} = 0.$$

Step 7. Completion of the proof. Let us put the above facts together. We fix $\varepsilon > 0$ and take $a > q$ such that

$$\left|\frac{ra}{r+q}-1-(p-1)\right| < \varepsilon.$$

Then for sufficiently small δ we have

$$\frac{\|g^{(1)}\|_p}{\|f\|_q\|v^*\|_r} > p-1-2\varepsilon, \qquad \frac{\|g^{(2)}\|_p}{\|f\|_q\|v^*\|_r} < \varepsilon, \qquad \frac{\|g^{(3)}\|_p}{\|f\|_q\|v^*\|_r} < \varepsilon$$

and hence

$$\frac{\|g\|_p}{\|f\|_q\|v^*\|_r} > p-1-4\varepsilon.$$

Since ε was arbitrary, the sharpness follows.

4 Weak-type estimates, $p \geq 2$

4.1 Proof of (1.6)

As in the case of strong–type estimates, we may and do assume that $\|X\|_q < \infty$ and $\|H^*\|_r < \infty$; we may also assume that the norms are strictly positive since otherwise, the claim is obvious. Then $H^*_+ X$ and Y belong to L_p, as we checked in the preceding section. Consider the functions U, V on $\mathcal{H} \times \mathcal{H}$, given by

$$U(x,y) = \beta_p \int_0^{1-p^{-1}} \lambda^{p-1} u_\infty(x/\lambda, y/\lambda)\mathrm{d}\lambda$$

and

$$V(x,y) = p\,(|y| - 1 + 1/p)_+ - \frac{p^{p-1}}{2}|x|^p,$$

where $\beta_p = p^p(p-1)^{2-p}(p-2)/4$. It was proved in [3] that

$$U \geq V \qquad \text{on } \mathcal{H} \times \mathcal{H}. \tag{4.1}$$

Applying (2.4) and Fubini's theorem, we get $\mathbb{E}V(H^*_{t+}X_t, Y_t) \leq \mathbb{E}U(H^*_{t+}X_t, Y_t) \leq 0$ for $t \geq 0$. Fubini's theorem is applicable, since

$$\int_0^{1-p^{-1}} \lambda^{p-1}|u_\infty(x/\lambda, y/\lambda)|\mathrm{d}\lambda \leq \int_0^\infty \lambda^{p-1}|u_\infty(x/\lambda, y/\lambda)|\mathrm{d}\lambda \lesssim_p |x|^p + |y|^p,$$

as we already verified in (3.4). Therefore, we obtain

$$\mathbb{E}(p|Y_t| - p + 1)_+ \leq \frac{p^{p-1}}{2}\mathbb{E}|H^*_{t+}X_t|^p \leq \frac{p^{p-1}}{2}\|X\|_q^p\|H^*\|_r^p.$$

Fix an arbitrary event A of positive probability. Then

$$\mathbb{E}(p|Y_t| - p + 1)1_A \leq \mathbb{E}(p|Y_t| - p + 1)_+ \leq \frac{p^{p-1}}{2}\|X\|_q^p\|H^*\|_r^p,$$

or equivalently,

$$\int_A |Y_t|\mathrm{d}\mathbb{P} \leq \frac{p^{p-2}}{2}\|X\|_q^p\|H^*\|_r^p + \frac{p-1}{p}\mathbb{P}(A).$$

The differential subordination of Y to $H \cdot X$ is preserved if we multiply X and Y by a fixed positive constant λ. Applying the above estimate to the modified triple λX, λY and H, we obtain

$$\lambda \int_A |Y_t|\mathrm{d}\mathbb{P} \leq \lambda^p \frac{p^{p-2}}{2}\|X\|_q^p\|H^*\|_r^p + \frac{p-1}{p}\mathbb{P}(A).$$

Dividing both sides by λ and optimizing over λ (specifically, the best choice is $\lambda = (2\mathbb{P}(A)/p^{p-1})^{1/p}\|X\|_q^{-1}\|H^*\|_r^{-1}$), we get

$$\int_A |Y_t|\mathrm{d}\mathbb{P} \leq \left(\frac{p^{p-1}}{2}\right)^{1/p}\|X\|_q\|H^*\|_r \cdot \mathbb{P}(A)^{1-1/p}.$$

This yields $|||Y|||_{p,\infty} \leq (p^{p-1}/2)^{1/p}\|X\|_q\|H^*\|_r$, since A and t were arbitrary.

Here is the analogue of Remarks 3.1 and 3.2.

Remark 4.1. The above argumentation gives the sharp estimate

$$\|Y\|_{p,\infty} \leq \left(\frac{p^{p-1}}{2}\right)^{1/p}\|H^*X\|_p, \qquad p \geq 2,$$

under the assumptions of Theorem 1.5.

4.2 Sharpness for martingale transforms

The calculations are quite similar to those appearing in the previous section. We take $\delta > 0$, fix a positive integer N and set $Q = 1 - (p-1)q\delta/p$. Then we define the sequence $(p_n)_{n\geq 0}$ as before and consider the probability space $((0,1], \mathcal{B}(0,1), |\cdot|)$. We consider the σ-algebras $\mathcal{F}_0$, $\mathcal{F}_1$, $\ldots$, $\mathcal{F}_{2N}$ as previously, and $\mathcal{F}_{2N+1} = \mathcal{F}_{2N+2} = \ldots$ is the σ-field with atoms $(0, p_{2N}/2]$, $(p_{2N}/2, p_{2N}]$, $(p_{2N}, p_{2N-1}]$, $(p_{2N-1}, p_{2N-2}]$, $\ldots$, $(p_1, p_0]$, that is, $\mathcal{F}_{2N+1} = \mathcal{F}_{2N+2} = \ldots = \sigma(\mathcal{F}_{2N}, (0, p_{2N}/2])$. Consider the function f given by the finite sum

$$f = \sum_{n=0}^{N-1} (1+\delta)^n \Big(\chi_{(p_{2n+1}, p_{2n}]} - \chi_{(p_{2n+2}, p_{2n+1}]} \Big) + (1+\delta)^N \Big(\chi_{(0, p_{2N}/2]} - \chi_{(p_{2N}/2, p_{2N}]} \Big).$$

This function is measurable with respect to $\mathcal{F}_{2N+1}$ and satisfies

$$\begin{aligned} \mathbb{E}|f|^q &= \sum_{n=0}^{N-1} (1+\delta)^{qn} (p_{2n} - p_{2n+2}) + (1+\delta)^{qN} p_{2N} \\ &\leq (Q(1+\delta)^q)^N \cdot \frac{(1+\delta)^q Q - Q}{(1+\delta)^q Q - 1}. \end{aligned} \tag{4.2}$$

It is easy to see that the formulas for df_n, $n = 0, 1, 2, \ldots, 2N$, are the same as in the previous section. This follows from the fact that f has not been changed on $(p_{2N}, 1]$ and it still has a vanishing integral on $(0, p_{2N}]$. To complete the description of the difference sequence, note that $df_{2N+1} = f\chi_{(0,p_{2N}]}$ and $df_{2N+2} = df_{2N+3} = \ldots = 0$.

The transforming sequence $v = (v_n)_{n\geq 0}$ is given by $v_0 \equiv 1$; for $n = 0, 1, 2, \ldots, N-1$ we put $v_{2n+1} = -(1+\delta)^{nq/r}\chi_{(0,p_{2n}]}$ and $v_{2n+2} = (1+\delta)^{nq/r}\chi_{(0,p_{2n+1}]}$; finally, for $n > 2N$ we set $v_n = (1+\delta)^{Nq/r}\chi_{(0,p_{2N}]}$. So, in comparison to the formulas from the previous section, we see that v_0, v_1, $\ldots$, v_{2N} are the same. Consequently, we may repeat the analysis and obtain that $v^* = |f|^{q/r}$; furthermore, on $(0, p_{2N}/2]$ we have

$$\begin{aligned} g_{2N+1} &= v_0 df_0 + v_1 df_1 + \ldots + v_{2N+1} df_{2N+1} \\ &= \frac{2(1-Q)}{1+Q} \Big[1 + (1+\delta)^{1+q/r} + \ldots + (1+\delta)^{(N-1)(1+q/r)} \Big] + (1+\delta)^{N(1+q/r)} \\ &= \frac{2(1-Q)}{1+Q} \cdot \frac{(1+\delta)^{N(1+q/r)} - 1}{(1+\delta)^{1+q/r} - 1} + (1+\delta)^{N(1+q/r)}. \end{aligned}$$

Denoting the latter expression by λ, we see that

$$\frac{|||g|||_{p,\infty}}{\|f\|_q \|v^*\|_r} = \frac{|||g|||_{p,\infty}}{\|f\|_q^{1+q/r}} \geq \frac{\lambda |(0, p_{2N}/2]|^{1/p}}{\|f\|_q^{q/p}} \geq \lambda \left(\frac{(1+\delta)^q Q - 1}{2Q(1+\delta)^{qN}((1+\delta)^q - 1)} \right)^{1/p},$$

where the last inequality is due to (4.2). Now we need to perform an appropriate limiting procedure. Letting $N \to \infty$, the latter expression converges to

$$\left(\frac{2(1-Q)}{(1+Q)((1+\delta)^{1+q/r} - 1)} + 1 \right) \left(\frac{(1+\delta)^q Q - 1}{2Q((1+\delta)^q - 1)} \right)^{1/p}.$$

Now if we let $\delta \to 0$, the above quantity tends to $p \cdot (2p)^{-1/p} = (p^{p-1}/2)^{1/p}$. This yields the desired lower bound for the weak–type constant.

5 Weak-type estimates, $p < 2$

5.1 Proof of (1.4)

If $q \geq 2$, then the estimate follows at once from the strong–type bound: we have

$$|||Y|||_{p,\infty} \leq \|Y\|_p \leq \|X\|_q \|H^*\|_r.$$

The main difficulty lies in proving the weak–type inequality for $1<p<q<2$; one easily checks that $1<r'<q$ in such a case. Fix X, Y and H as in the statement; we may assume that $\|X\|_q<\infty$, $\|H^*\|_r<\infty$ and $|H_0|$ is bounded away from zero. Then $\|Y\|_p<\infty$, by the strong–type estimate which we have established in Section 3.

We will make use of Burkholder's method: this time the definitions of the appropriate special functions are much more involved. To avoid notational confusion, in our considerations below we will use the letter α instead of r'. Consider the differential equation

$$\alpha(2-\alpha)\phi'(x)+\alpha=q(q-1)x^{q-2}\phi(x)^{2-\alpha}. \tag{5.1}$$

We have the following fact, which appears as Theorem 2.1 in [16].

Theorem 5.1. *There exists a unique nondecreasing, concave solution $\phi:[0,\infty)\to[0,\infty)$ of* (5.1) *satisfying $\phi(0)>0$ and $\phi'(t)\to 0$, $\phi(t)\to\infty$, as $t\to\infty$.*

From now on, ϕ stands for the solution described in the above theorem. We have the following explicit bounds for $\phi(0)$.

Lemma 5.2. *We have*

$$\left(\frac{\alpha}{2}\right)^{1/(q-\alpha)}\leq\phi(0)\leq\frac{q-\alpha}{2-\alpha}\left(\frac{\alpha(3-q)}{q(q-1)}\right)^{1/(q-\alpha)}.$$

Proof. As shown in [17, Lemma 2.2], for any $t\geq 0$ we have

$$(\phi(t)+t)^q-t^q-qt^{q-1}\phi(t)-(\alpha-1)\phi(t)^\alpha-\frac{2-\alpha}{2}\phi(0)^\alpha\geq 0.$$

Plugging $t=0$ gives the lower bound for $\phi(0)$. To obtain the upper bound, set

$$w=\left(\frac{\alpha(3-q)}{q(q-1)}\right)^{1/(q-\alpha)}.$$

Let ψ be the solution to (5.1), satisfying $\psi(w)=w$ and extended to the maximal domain contained in $[0,\infty)$. A direct application of (5.1) gives $\psi''(w)=0$. Thus, as proved in [16, Theorem 2.1], ψ is given on the whole $[0,\infty)$ and satisfies $\psi>\phi$ there; furthermore, we have $\psi''<0$ on $(0,w)$. Consequently,

$$\phi(0)<\psi(0)<\psi(w)-w\psi'(w)=\frac{q-\alpha}{2-\alpha}\left(\frac{\alpha(3-q)}{q(q-1)}\right)^{1/(q-\alpha)},$$

where the last passage is due to $\psi(w)=w$ and the fact that ψ enjoys (5.1). □

Let $\Phi:[\phi(0),\infty)\to[0,\infty)$ be the inverse to $t\mapsto t+\phi(t)$. We have $\phi(\Phi(t))+\Phi(t)=t$, which in particular yields

$$\phi(\Phi(t))\leq t\quad\text{and}\quad\phi'(\Phi(t))\Phi'(t)\leq 1 \tag{5.2}$$

for $t>0$. For the notational convenience, let us distinguish the constant

$$L_{\alpha,q}=\frac{(2-\alpha)\phi(0)^\alpha}{2}$$

and consider the auxiliary kernel

$$w(\lambda)=\frac{\alpha(2-\alpha)}{2}\phi(\Phi(\lambda))^{\alpha-3}\phi'(\Phi(\lambda))\Phi'(\lambda)\lambda^2,\qquad\lambda>0.$$

We are ready for the definitions of the functions V, $U:\mathcal{H}\times\mathcal{H}\to\mathbb{R}$ which will lead us to the weak–type estimate. Set

$$V(x,y)=\left(|y|^\alpha-L_{\alpha,q}\right)_+-|x|^q$$

and

$$U(x,y)=\int_{\phi(0)}^{\infty} w(\lambda)u_1(x/\lambda,y/\lambda)\mathrm{d}\lambda. \tag{5.3}$$

One can derive the explicit formula for U, but it will not be needed in our considerations. The only property which matters to us is the majorization of V by U (see Lemma 3.5 in [17]). Furthermore, by (5.2) we have $w(\lambda)\lesssim_{\alpha,q}\lambda^{q-1}$ and hence, computing as in (3.7),

$$\int_{\phi(0)}^{\infty} w(\lambda)|u_1(x/\lambda,y/\lambda)|\mathrm{d}\lambda \lesssim_{\alpha,q} |x|^q+|y|^q.$$

Thus by (2.3), Fubini's theorem and the majorization $U\geq V$,

$$\mathbb{E}V(X_t,Y_t/H^*_{t+})\leq \mathbb{E}U(X_t,Y_t/H^*_{t+})\leq 0, \qquad t\geq 0. \tag{5.4}$$

Now we argue as in the case $p>2$. For an arbitrary event A of positive probability, we may write

$$\mathbb{E}\big(|Y_t/H^*_{t+}|^\alpha-L_{\alpha,q}\big)1_A\leq \mathbb{E}(|Y_t/H^*_{t+}|^\alpha-L_{\alpha,q})_+\leq \mathbb{E}|X_t|^q,$$

where the last passage is equivalent to (5.4). Therefore, we get

$$\int_A |Y_t/H^*_{t+}|^\alpha \mathrm{d}\mathbb{P}\leq \|X\|_q^q+L_{\alpha,q}\mathbb{P}(A).$$

The differential subordination of Y to $H\cdot X$ is not affected if we multiply X and Y by a fixed positive constant λ. Therefore, the above inequality gives

$$\int_A |Y_t/H^*_{t+}|^\alpha \mathrm{d}\mathbb{P}\leq \lambda^{q-\alpha}\|X\|_q^q+\lambda^{-\alpha}L_{\alpha,q}\mathbb{P}(A),$$

and the optimization over λ yields

$$\int_A |Y_t/H^*_{t+}|^\alpha \mathrm{d}\mathbb{P}\leq \frac{q}{\alpha}\left(\frac{\alpha}{q-\alpha}L_{\alpha,q}\right)^{1-\alpha/q}\|X\|_q^\alpha\mathbb{P}(A)^{1-\alpha/q}. \tag{5.5}$$

Consequently, recalling that α is the Hölder conjugate to r, we may write

$$\begin{aligned}\int_A |Y_t|\mathrm{d}\mathbb{P} &\leq \left(\int_A |Y_t/H^*_{t+}|^\alpha \mathrm{d}\mathbb{P}\right)^{1/\alpha}\|H^*\|_r\\ &\leq \left(\frac{q}{r'}\right)^{1/r'}\left(\frac{r'}{q-r'}L_{r',q}\right)^{1-1/p}\|X\|_q\|H^*\|_r\mathbb{P}(A)^{1-1/p}.\end{aligned}$$

This is precisely the desired weak–type bound, since A and t were chosen arbitrarily.

Let us conclude by stating the analogue of Remarks 3.1, 3.2 and 4.1. Namely, the estimate (5.5) immediately gives the following.

Remark 5.3. Under the assumptions of Theorem 1.5, we have the sharp bound

$$|||Y/H^*|||_{q,\infty}\leq \left(\frac{q}{r'}\right)^{1/r'}\left(\frac{r'}{q-r'}L_{r',q}\right)^{1-1/p}\|X\|_q.$$

Here $|||\cdot|||$ is the equivalent weak L^q norm

$$|||\xi|||_{q,\infty}=\sup\left\{\mathbb{P}(A)^{1/q-1/\alpha}\left(\int_A|\xi|^\alpha\mathrm{d}\mathbb{P}\right)^{1/\alpha}:\ \mathbb{P}(A)>0\right\}.$$

5.2 Sharpness for martingale transforms

As previously, we may restrict ourselves to the case $1 < p < q < 2$: for $q \geq 2$, the constant is 1, which is achieved for $f = g = v \equiv 1$.

Fix $\varepsilon > 0$. Our starting point is the strong–type estimate

$$\|\varphi\|_{r'} \leq K_{r',q}\|f\|_q,$$

where f is an arbitrary L_q-bounded martingale and φ is its transform by the deterministic sequence $w_n = (-1)^n$, $n = 0, 1, 2, \ldots$. The optimal value of the constant $K_{r',q}$ was identified in [16]: it is equal to $c_{p,q,r}$ and the almost-extremal examples have the following structure: see Figure 1 below to gain some intuition. Fix a small parameter $\delta > 0$. The pair (f, φ) starts from $(\phi(0)/2, \phi(0)/2)$ and at the first move it goes to $(0, \phi(0))$ or to $(\phi(0), 0)$. Then the evolution is governed by the following rules:

· if (f, φ) lies on one of the curves $y = \phi(x)$ or $y = -\phi(x)$, it stops ultimately;

· if we have $(f, \varphi) = (x, 0)$ for some $x > 0$, then the pair jumps, along the line of slope 1, to $(x + \delta, \delta)$ or onto the curve $y = -\phi(x)$;

· if we have $(f, \varphi) = (x + \delta, \delta)$ for some $x > 0$, then the pair jumps, along the line of slope -1, to $(x + 2\delta, 0)$ or onto the curve $y = \phi(x)$.

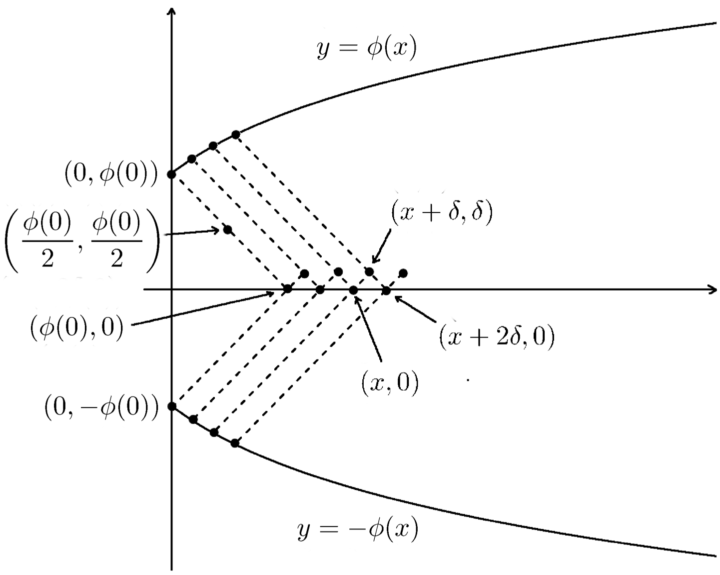

Figure 1: The structure of the extremal examples. The dots • indicate the possible locations of the pair (f, φ).

Let us gather some basic information about f and φ, which will be needed later. First, the martingales are unbounded, but they are both bounded in L_q. Furthermore, it can be extracted from [16] that

$$\lim_{\delta \downarrow 0} \frac{\|\varphi\|_{r'}}{\|f\|_q} = c_{p,q,r}.$$

Next, we make some observations concerning the behavior of the differences df_0, df_1, ... We easily see that with probability 1, first several differences are positive; then there is a negative term; and then the remaining differences are zero. Let us be more specific. We have $df_0 = \phi(0)/2 > 0$ and then there are two possible scenarios:

(a) $df_1 = -\phi(0)/2$ and $df_2 = df_3 = \ldots = 0$; then $(df)^* = \phi(0)/2$ and $\varphi^* = \varphi = \phi(0)$;

(b) $df_1 = \phi(0)/2$. Then there is an integer $m \geq 2$ such that $df_2 = df_3 = \ldots = df_{m-1} = \delta > 0$, $df_m < 0$ and $df_{m+1} = df_{m+2} = \ldots = 0$. In this case, we have $(df)^* = |df_m|$ and $\varphi^* = |\varphi| \geq (df)^*$.

We define the transforming sequence v by $v_0 = \phi(0)^{r'-1}$, $v_1 = -\phi(0)^{r'-1}$ and $v_n = (-1)^n |\varphi_{n-1}|^{r'-1}$ for $n \geq 2$. Obviously, this sequence is predictable and we have $v^* = (\varphi^*)^{r'-1} = |\varphi|^{r'-1}$. To understand the behavior of g, note that in the scenario (a),

$$g = \phi(0)^{r'-1} \cdot \phi(0)/2 - \phi(0)^{r'-1} \cdot (-\phi(0)/2) = \phi(0)^{r'} = |\varphi|^{r'}.$$

On the other hand, in the scenario (b) we have $v_0 df_0 + v_1 df_1 = 0$ and

$$g = v_2 df_2 + v_3 df_3 + \ldots + v_m df_m.$$

But the sequence $(v_n)_{n\geq 0}$ is alternating and $(|v_n|)_{n\geq 0}$ is nondecreasing, while $df_2 = df_3 = \ldots = df_{m-1} = \delta$ and $df_m < 0$. Consequently, $|g| \geq |v_m||df_m| = v^*(df)^* > (1-\varepsilon)v^*\varphi^* = (1-\varepsilon)|\varphi|^{r'}$, if δ is sufficiently small. Putting all these facts together, we obtain the inequality

$$|||g|||_{p,\infty} \geq \int_\Omega |g| \mathrm{d}\mathbb{P} \geq (1-\varepsilon)\mathbb{E}|\varphi|^{r'} = (1-\varepsilon)\|\varphi\|_{r'}\|v^*\|_r \geq (1-\varepsilon)(c_{p,q,r} - \varepsilon)\|f\|_q \|v^*\|_r,$$

provided δ is sufficiently small. This is precisely the desired claim, since ε can be chosen arbitrarily small.

References

[1] Baernstein II, A. and Montgomery-Smith, S. J.: Some conjectures about integral means of ∂f and $\bar{\partial} f$. Complex analysis and differential equations (Uppsala, 1997), Acta Univ. Upsaliensis Skr. Uppsala Univ. C Organ. Hist., vol. 64, *Uppsala Univ.*, Uppsala, 1999, 92–109. MR1758918

[2] Bañuelos, R., Bielaszewski, A., and Bogdan, K.: Fourier multipliers for non-symmetric Lévy processes. Marcinkiewicz centenary volume, *Banach Center Publ.* **95**, (2011), 9–25. MR2918086

[3] Bañuelos, R. and Osękowski, A.: Sharp martingale inequalities and applications to Riesz transforms on manifolds, Lie groups and Gauss space. *J. Funct. Anal.*, **269**, (2015), 1652–1713. MR3373431

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

[5] Bellman, R.: Dynamic programming. Reprint of the 1957 edition, Princeton Landmarks in Mathematics. *Princeton University Press*, Princeton, NJ, 2010. MR2641641

[6] Burkholder, D. L.: Martingale transforms. *Ann. Math. Statist.* **37**, (1966), 1494–1504. MR0208647

[7] Burkholder, D. L.: A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *Ann. Probab.* **9**, (1981), 997–1011. MR0208647

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

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

[10] Dellacherie, C. and Meyer, P. A.: Probabilities and potential B. North-Holland, Amsterdam, 1982. MR0745449

[11] Métivier, M.: Semimartingales: a course on stochastic processes, volume 2. Walter de Gruyter, 1982. MR0688144

[12] Nazarov, F. L. and Treil, S. R.: The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. *St. Petersburg Math. J.* **8**, (1997), 721–824. MR1428988

[13] Nazarov, F. L., Treil, S. R. and Volberg, A.: The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.* **12**, (1999), 909–928. MR1685781

[14] Osękowski, A.: Sharp martingale and semimartingale inequalities. Monografie Matematyczne **72**, Birkhäuser, Basel, 2012. MR2964297

[15] Osękowski, A.: Weak-type inequalities for Fourier multipliers with applications to the Beurling-Ahlfors transform. *J. Math. Soc. Japan* **66**, (2014), 745–764. MR3238316

[16] Osękowski, A.: Sharp moment inequalities for differentially subordinated martingales. *Studia Math.* **201**, (2010), 103–131. MR2738156

[17] Osękowski, A.: Sharp localized inequalities for Fourier multipliers. *Canadian J. Math.* **66**, (2014), 1358–1381. MR3270787

[18] Revuz, D. and Yor, M.: Continuous martingales and Brownian motion. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 239. Springer-Verlag, Berlin, third edition, 1999. MR1725357

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

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

Acknowledgments. The authors would like to thank Referee and the Associate Editor for the helpful suggestions, which led to the significant improvement of the presentation.