Electron. J. Probab. **26** (2021), article no. 21, 1–22.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP596

Sharp maximal L^p-bounds for continuous martingales and their differential subordinates

Adam Osękowski* Yahui Zuo†

Abstract

Suppose that X, Y are Hilbert-space-valued continuous-path martingales such that Y is differentially subordinate to X. The paper contains the proof of sharp estimates between p-th moments of Y and the maximal function of X for $0 < p < 1$. The proof rests on Burkholder's method and exploits a certain special function of three variables, enjoying appropriate size and concavity requirements. The analysis reveals an unexpected phase transition between the cases $0 < p < 1/2$ and $1/2 \leq p < 1$. The latter case is relatively simple: the special function is essentially quadratic and the best constant is equal to $\sqrt{2/p}$. The analysis of the former case is much more intricate and involves the study of a non-linear ordinary differential equation.

Keywords: martingale; stochastic integral; maximal inequality; differential subordination.
MSC2020 subject classifications: 60G44.
Submitted to EJP on April 18, 2020, final version accepted on February 1, 2021.

1 Introduction

Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, filtered by a nondecreasing right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$. We assume in addition that $\mathcal{F}_0$ contains all the events of probability 0. Suppose further that X, Y are two adapted martingales taking values in a given separable Hilbert space $\mathcal{H}$ with the norm $|\cdot|$ and scalar product denoted by $\langle \cdot, \cdot \rangle$. With no loss of generality, we may and do assume that the space is equal to ℓ^2. As usual, we impose standard regularity restrictions on trajectories of X and Y: we assume that the paths are right-continuous and have limits from the left. Then $[X, X]$, the quadratic covariance process of X, is given by $[X, X] = \sum_{n=1}^{\infty} [X^n, X^n]$, where X^n is the n-th coordinate of X and $[X^n, X^n]$ is the usual square bracket of the real-valued martingale X^n (see Chapters VI and VII in Dellacherie and Meyer [12] or Chapter 4 in Métivier [15] for details). We will also consider the maximal process $X^* = (X_t^*)_{t\geq 0}$, where $X_t^* = \sup_{0\leq s\leq t} |X_s|$ for each $t \geq 0$.

*Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Poland. E-mail: ados@mimuw.edu.pl
†School of Mathematics and Statistics, Central South University, China. E-mail: zuoyahui@csu.edu.cn

Throughout, we will work under the assumption that the martingale Y is *differentially subordinate* to X. This concept appeared for the first time in Burkholder's paper [8] and concerned discrete-time processes. Recall that 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 $|dg_n| \leq |df_n|$ almost surely. Here $df = (df_n)_{n\geq 0}$, $dg = (dg_n)_{n\geq 0}$ are the difference sequences of f and g, given by

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

This domination was generalized to the continuous-time context by Bañuelos and Wang [7] and Wang [24]: the martingale 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. Treating two 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 the original definition of Burkholder. The following example will be useful to us later: 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 the stochastic integral of H with respect to X, i.e., $Y_t = H_0X_0 + \int_{0+}^{t} H_s \mathrm{d}X_s$, $t \geq 0$. Then Y is differentially subordinate to X, since

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

Another example, which is very important for applications (see e.g. [5], [7], [14]), is the following. Suppose that B is a Brownian motion in $\mathbb{R}^\nu$ and H, K are predictable processes taking values in the matrices of dimensions $m \times \nu$ and $n \times \nu$, respectively. For any $t \geq 0$, define

$$X_t = \int_{0+}^{t} H_s \cdot \mathrm{d}B_s \qquad \text{and} \qquad Y_t = \int_{0+}^{t} K_s \cdot \mathrm{d}B_s.$$

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

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

The differential subordination implies many interesting inequalities comparing the sizes of X and Y. The literature on this subject is very extensive, so we refer the interested reader to the survey articles [9], [21] and the monograph [19] for the detailed description and further references. In addition, these estimates have found numerous profound applications in harmonic analysis, e.g. in the study of boundedness of wide classes of Fourier multipliers [4, 5, 6, 14], semigroup theory [11, 13] and quasiconformal mappings [1, 2, 3], to name just a few.

As the motivation for our research, we recall a celebrated result of Burkholder which gives the following information on the L^p-norms (see [8], [9] and [24]).

Theorem 1.1. *Suppose that X, Y are Hilbert-space-valued martingales such that Y is differentially subordinate to X. Then for any $1 < p < \infty$ we have*

$$\|Y_t\|_{L^p} \leq (p^* - 1)\|X_t\|_{L^p}, \qquad t \geq 0, \tag{1.1}$$

where $p^* = \max\{p, p/(p-1)\}$. The constant is the best possible, even if $\mathcal{H} = \mathbb{R}$: it cannot be replaced by any smaller number.

Besides the beauty of this result, its proof is of independent interest and connections. Burkholder showed how to deduce the validity of (1.1) from the existence of a certain special function, possessing appropriate size and concavity properties. Since the appearance of the seminal paper [8], this type of approach has turned out to be very efficient in the study of related semimartingale inequalities (for much more on the subject, see [19]).

In the boundary case $p = 1$, the above moment inequality breaks down, but, as a substitution, there are certain weak-type and logarithmic estimates; see [8], [21] and [22]. There is also a corresponding maximal L^1 bound, which will be important for our considerations below. Namely, in [10] Burkholder proved the following result.

Theorem 1.2. *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*

$$\|Y_t\|_{L^1} \leq \gamma \|X_t^*\|_{L^1}, \qquad t \geq 0, \tag{1.2}$$

where $\gamma = 2.536\ldots$ is the unique positive number satisfying $\gamma = 3 - \exp \frac{1-\gamma}{2}$.

This result was strengthened by Osękowski [17] to the case in which the first moment of Y is replaced by the first moment of its maximal function.

Theorem 1.3. *Under the assumptions of the above theorem, we have the sharp inequality*

$$||Y_t^*||_{L^1} \leq 3.4351\ldots ||X_t^*||_{L^1}, \qquad t \geq 0. \tag{1.3}$$

The precise description of the above constant involves the analysis of a complicated system of ODE's. If the dominating martingale has continuous paths, then the best constants in (1.2) and (1.3) decrease to $\sqrt{2}$ and 2, respectively (cf. [18, 20]). The following result can be found in [18].

Theorem 1.4. *Suppose that X, Y are $\mathcal{H}$-valued, continuous-path martingales such that Y is differentially subordinate to X. Then for any $1 \leq p < \infty$ we have the sharp bound*

$$\|Y_t\|_{L^p} \leq c_p ||X_t^*||_{L^p}, \qquad t \geq 0, \tag{1.4}$$

where $c_p = \sqrt{2/p}$ for $1 \leq p \leq 2$ and $c_p = p - 1$ for $p > 2$.

By Burkholder-Davis-Gundy inequalities, the estimate (1.4) holds also, with some finite constant c_p, in the range $0 < p < 1$. The purpose of this paper is to identify the corresponding optimal constant. Our main result can be stated as follows.

Theorem 1.5. *Suppose that X, Y are $\mathcal{H}$-valued continuous-path martingales such that Y is differentially subordinate to X. Then for any $0 < p < 1$ we have the estimate*

$$\|Y_t\|_{L^p} \leq C_p \|X_t^*\|_{L^p}, \qquad t \geq 0, \tag{1.5}$$

where the optimal constant is equal to

$$C_p = \begin{cases} y_0(p) + 2 & \text{if } 0 < p < 1/2, \\ \sqrt{2/p} & \text{if } 1/2 \leq p < 1. \end{cases}$$

Here $y_0(p)$ is the unique positive zero of the solution to a certain ODE: see Lemma 3.1 for the precise definition.

So, the behavior of the best constant C_p in (1.5) is different for $p < 1/2$ and $p \geq 1/2$. As we will see, the proof in the latter case ('big p') is much simpler: it can actually be extracted from the work [18]. Very interestingly (and quite unexpectedly, at least to us), there is a phase transition at $p = 1/2$. The analysis in the case $p < 1/2$ is much

more elaborate and the construction of the associated special functions (which are completely different from those for $p \geq 1/2$) will rest on solutions of the non-linear ordinary differential equation (3.4).

We would also like to point out that if we drop the assumption on the continuity of paths of the martingales, then the inequality (1.5) is no longer valid, even for $\mathcal{H} = \mathbb{R}$. It is convenient to describe appropriate examples in discrete time. Fix huge integers M and N and consider the independent random variables ξ_0, ξ_1, ξ_2, ..., ξ_N with the distribution given by $\mathbb{P}(\xi_j = 1) = 1 - \mathbb{P}(\xi_j = -M) = M/(M+1)$, $j = 0, 1, 2, \ldots, N$. These variables are centered and hence the process $f_n = \xi_0 - \xi_1 + \xi_2 - \xi_3 + \ldots + (-1)^n \xi_n$, $n = 0, 1, 2, \ldots, N$, is a martingale. On the set $A = \{\xi_j = 1 \text{ for all } j\}$, which has probability $(M/(M+1))^{N+1}$, we have $f_N^* = 1$; on the complement of this event, we obviously have $f_N^* \leq 2M(N+1)$. Consequently,

$$\|f_N^*\|_p \leq \left[\left(\frac{M}{M+1} \right)^{N+1} + \left(1 - \left(\frac{M}{M+1} \right)^{N+1} \right) \cdot (2M(N+1))^p \right]^{1/p}$$

and the right-hand side converges to 1 as $M \to \infty$ (the parameter N is fixed). Consider the martingale $g_n = \xi_0 + \xi_1 + \xi_2 + \ldots + \xi_n$, $n = 0, 1, 2, \ldots, N$. We have $|dg_n| = |df_n|$ for each n, so g is differentially subordinate to f. Furthermore, g satisfies $g_N = N + 1$ on the set A, so $\|g_N\|_p \geq (N+1)(\mathbb{P}(A))^{1/p} = (N+1)(M/(M+1))^{(N+1)/p}$. The latter expression converges to $N+1$ as $M \to \infty$, so we see that taking M sufficiently large, we may make the ratio $\|g_N\|_p / \|f_N^*\|_p$ bigger than N. Since N was arbitrary, the estimate (1.5) cannot hold with any finite constant.

We have organized this paper as follows. The next section contains the description of Burkholder's method in the abstract setting: we show how to exploit certain special appropriately smooth functions to establish maximal estimates for continuous-time processes. The contents of that section extends significantly the material in [18] (see below). Section 3 contains the proof of the inequality (1.5): we construct and study the corresponding special functions there. In Section 4 we show that the constant C_p cannot be improved, which is accomplished by exhibiting appropriate extremal examples. In the final part of the paper we describe in detail the reasoning which has led us to the discovery of the special function in the difficult case $0 < p < 1/2$.

2 Burkholder's method

The original technique developed by Burkholder in [10] concerned the discrete-time martingales (the passage to the context of stochastic integrals, as in Theorem 1.2, follows by approximation). The successful treatment of continuous-path, differentially subordinate and vector-valued martingales required the enhancement of the method, which was obtained by Osękowski [17] under strong regularity assumptions on the special function. In this section we relax this requirement and show how to deduce maximal estimates from the existence of a function satisfying a relatively mild regularity condition (class C^1). Our reasoning rests on mollification arguments, which can be tracked back to the works of Bañuelos and Wang [7] and Wang [24].

Consider the domain $\mathcal{D} = \{(x, y, z) \in \mathcal{H} \times \mathcal{H} \times (0, \infty) : |x| \leq z\}$ and fix a Borel function $V : \mathcal{D} \to \mathbb{R}$, which is bounded on bounded sets. Let us assume that we want to establish the estimate

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

for all pairs (X, Y) of continuous and bounded $\mathcal{H}$-valued martingales such that Y is differentially subordinate to X. Suppose that D_1, D_2, ..., D_n is a collection of pairwise disjoint open subsets of $\mathcal{D}$ such that $\mathcal{D} \subseteq \overline{D}_1 \cup \overline{D}_2 \cup \ldots \cup \overline{D}_n$. Assume further that

$U : \mathcal{D} \to \mathbb{R}$ is a C^1 function, which is of class C^2 on each D_i and satisfies the following four conditions.

1° For any $(x, y, z) \in \mathcal{D}$ such that $|y| \leq |x|$ we have

$$U(x, y, z) \leq 0.$$

2° We have $U \geq V$ on $\mathcal{D}$.

3° For any $(x, y, z) \in \mathcal{D}$, we have

$$U_z(x, y, z) \leq 0. \tag{2.2}$$

4° For any $j = 1, 2, \ldots, n$, there is a function $c_j : D_j \to (0, \infty)$, bounded on any set of the form $\{(x, y, z) \in D_j : |x|, |y| \leq L, 1/L \leq z \leq L\}$ for some $L > 0$, such that

$$\langle U_{xx}(x, y, z)h, h\rangle + 2\langle U_{xy}(x, y, z)h, k\rangle + \langle U_{yy}(x, y, z)k, k\rangle \leq c_j(x, y, z)(|k|^2 - |h|^2) \tag{2.3}$$

holds for any $(x, y, z) \in D_j$ and all $h, k \in \mathcal{H}$.

The interplay between the existence of such a function U and the validity of (2.1) is described in the statement below. In what follows, X_0^1 denotes the first coordinate of X_0.

Theorem 2.1. *Suppose that U satisfies the conditions 1°−4°. Assume further that X, Y are bounded, $\mathcal{H}$-valued, and continuous-path martingales such that Y is differentially subordinate to X and $\mathbb{P}(|X_0^1| \geq \eta) = 1$ for some $\eta > 0$. Then* (2.1) *holds true.*

Proof. It is convenient to split the reasoning into a few separate parts.

Step 1. Reductions. Let $t \geq 0$ be fixed. The random variable $V(X_t, Y_t, X_t^*)$ is integrable, since V is locally bounded and $X, Y \in L^\infty$. Furthermore, note that

$$C = \sup_{s \leq t} \max_j \big\|c_j(X_s, Y_s, X_s^*)\big\|_{L^\infty} < \infty, \tag{2.4}$$

by $X_s^* \geq |X_0^1| \geq \eta$ and the boundedness properties of c_1, c_2, $\ldots$, c_n. By the majorization condition 2°, the proof will be complete if we show that

$$\mathbb{E}U(X_t, Y_t, X_t^*) \leq 0. \tag{2.5}$$

Fix $\varepsilon > 0$. There is an integer $D = D(\varepsilon) \geq 1$ such that if $d \geq D$, then

$$\mathbb{E}\sum_{k>d}[X^k, X^k]_t = \mathbb{E}\sum_{k>d}|X_t^k|^2 < \varepsilon. \tag{2.6}$$

For $0 \leq s \leq t$, consider the projected processes

$$X_s^{(d)} = (X_s^1, X_s^2, \ldots, X_s^d, 0, 0, \ldots), \qquad Y_s^{(d)} = (Y_s^1, Y_s^2, \ldots, Y_s^d, 0, 0, \ldots)$$

and let

$$Z_s^{(d)} = (X_s^{(d)}, Y_s^{(d)}, X_s^{(d)*}). \tag{2.7}$$

Let $\mathcal{D}^d = \{(x, y, z) \in \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty) : |x| \leq z\}$ and let $\overline{U} : \mathcal{D}^d \to \mathbb{R}$ be the restriction of the function U, given by the formula $\overline{U}(x, y, z) = U((x, 0, 0, \ldots), (y, 0, 0, \ldots), z)$.

Step 2. A mollified version of $\overline{U}$. Pick a C^∞ function $\varphi : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to [0, \infty)$, supported on the unit ball B of $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$, satisfying $\int_B \varphi = 1$. Next, for a fixed $\delta > 0$ and any $(x, y, z) \in \mathcal{D}^d$, define $\overline{U}^\delta$ by the convolution

$$\overline{U}^\delta(x, y, z) = \int_B \overline{U}(x + \delta u, y + \delta v, z + 2\delta + \delta r)\varphi(u, v, r)\,\mathrm{d}u\mathrm{d}v\mathrm{d}r.$$

Note that the point $(x+\delta u, y+\delta v, z+2\delta+\delta r)$ lies in the domain of $\overline{U}$: indeed, we have $|x+\delta u| \le |x|+\delta|u| \le z+2\delta-\delta|r| \le z+2\delta+\delta r$, so in particular, $\overline{U}^\delta$ is well-defined. By the properties of convolutions, the function is of class C^∞; furthermore, it satisfies the appropriate versions of 3° and 4°. Indeed, the inequality

$$\overline{U}^\delta_z \le 0 \tag{2.8}$$

follows at once from the integration by parts. To check 4°, observe that $\overline{U}$ is of class C^1 (since so is U) and hence the integration by parts gives

$$\overline{U}^\delta_{xx}(x,y,z) = \int_B \overline{U}_{xx}(x+\delta u, y+\delta v, z+2\delta+\delta r)\varphi(u,v,r)\,\mathrm{d}u\mathrm{d}v\mathrm{d}r,$$

with similar formulas for other second-order partial derivatives of $\overline{U}^\delta$. Therefore, if $(x,y,z)\in\mathcal{D}^d$ and $h,\,k\in\mathbb{R}^d$, then (2.3) yields

$$\langle \overline{U}^\delta_{xx}(x,y,z)h,h\rangle + 2\langle \overline{U}^\delta_{xy}(x,y,z)h,k\rangle + \langle \overline{U}^\delta_{yy}(x,y,z)k,k\rangle \le c^\delta(x,y,z)(|k|^2-|h|^2), \tag{2.9}$$

where $c^\delta:\mathcal{D}^d\to[0,\infty)$ is given by

$$c^\delta(x,y,z) = \int_B c(x+\delta u, y+\delta v, z+2\delta+\delta r)\varphi(u,v,r)\,\mathrm{d}u\mathrm{d}v\mathrm{d}r$$

and $c=\sum_j c_j\chi_{D_j}$. In particular, note that $|c^\delta(x,y,z)|\le C$, where C was defined in (2.4).

Step 3. Proof of (2.1). We apply Itô's formula to the process $(\overline{U}^\delta(Z^{(d)}_s))_{s\in[0,t]}$ (recall that $Z^{(d)}$ is given by (2.7)), obtaining

$$\overline{U}^\delta(Z^{(d)}_t) = I_0+I_1+I_2+I_3/2, \tag{2.10}$$

where

$$\begin{aligned}
I_0 &= \overline{U}^\delta(Z^{(d)}_0),\\
I_1 &= \int_0^t \overline{U}^\delta_x(Z^{(d)}_s)\cdot dX^{(d)}_s + \int_0^t \overline{U}^\delta_y(Z^{(d)}_s)\cdot dY^{(d)}_s,\\
I_2 &= \int_0^t \overline{U}^\delta_z(Z^{(d)}_s) dX^{(d)*}_s,\\
I_3 &= \int_0^t \overline{U}^\delta_{xx}(Z^{(d)}_s) d[X^{(d)},X^{(d)}]_s + 2\int_0^t \overline{U}^\delta_{xy}(Z^{(d)}_s) d[X^{(d)},Y^{(d)}]_s\\
&\quad + \int_0^t \overline{U}^\delta_{yy}(Z^{(d)}_s) d[Y^{(d)},Y^{(d)}]_s.
\end{aligned}$$

The random variable I_1 has zero expectation, since both stochastic integrals are L^2-bounded martingales: this follows from the fact that X, Y, and hence also $X^{(d)}$ and $Y^{(d)}$, are bounded processes. The term I_2 is nonpositive by (2.8). The last term I_3 is also nonpositive, which follows from (2.9) and the approximation of the integrals by Riemann sums. Namely, let $0\le s_0<s_1\le t$. For any $j\ge 0$, let $(\eta^j_i)_{0\le i\le i_j}$ be a sequence of nondecreasing finite stopping times with $\eta^j_0=s_0$, $\eta^j_{i_j}=s_1$ such that $\lim_{j\to\infty}\max_{0\le i\le i_j-1}|\eta^j_{i+1}-\eta^j_i|=0$. Keeping j fixed, we apply, for each $i=0,1,2,\ldots,i_j$, the inequality (2.9) to $x=X^{(d)}_{s_0}$, $y=Y^{(d)}_{s_0}$, $z=X^{(d)*}_{s_0}$ and $h=h^j_i=X^{(d)}_{\eta^j_{i+1}}-X^{(d)}_{\eta^j_i}$,

$k = k_i^j = Y^{(d)}_{\eta^j_{i+1}} - Y^{(d)}_{\eta^j_i}$. Summing the obtained $i_j + 1$ inequalities and letting $j \to \infty$ yields

$$\begin{aligned}\sum_{m=1}^{d}\sum_{n=1}^{d}\Big[&\overline{U}^{\delta}_{x_m x_n}(Z^{(d)}_{s_0})[(X^{(d)})^m,(X^{(d)})^n]^{s_1}_{s_0} + 2\overline{U}^{\delta}_{x_m y_n}(Z^{(d)}_{s_0})[(X^{(d)})^m,(Y^{(d)})^n]^{s_1}_{s_0}\\ &+ \overline{U}^{\delta}_{y_m y_n}(Z^{(d)}_{s_0})[(Y^{(d)})^m,(Y^{(d)})^n]^{s_1}_{s_0}\Big]\\ &\le C\sum_{k=1}^{d}\big([(Y^{(d)})^k,(Y^{(d)})^k]^{s_1}_{s_0} - [(X^{(d)})^k,(X^{(d)})^k]^{s_1}_{s_0}\big).\end{aligned}$$

Here C is given by (2.4) and we have used the notation $[S,T]^{s_1}_{s_0} = [S,T]_{s_1} - [S,T]_{s_0}$. If we approximate I_3 by discrete sums, we see that the inequality above leads to

$$\begin{aligned}I_3 &\le C\sum_{k=1}^{d}\big([(Y^{(d)})^k,(Y^{(d)})^k]^t_0 - [(X^{(d)})^k,(X^{(d)})^k]^t_0\big)\\ &\le -C\sum_{k>d}\big([Y^k,Y^k]^t_0 - [X^k,X^k]^t_0\big),\end{aligned}$$

where the last passage is due to the differential subordination.

Now we take the expectation of the both sides of (2.10) and use (2.6) to obtain $\mathbb{E}\overline{U}^{\delta}(Z^{(d)}_t) \le \mathbb{E}\overline{U}^{\delta}(Z^{(d)}_0) + C\varepsilon$. Then let $\delta \to 0$ and finally $d \to \infty$ to get

$$\mathbb{E}U(X_t,Y_t,X^*_t) \le \mathbb{E}U(X_0,Y_0,X^*_0) + C\varepsilon \le C\varepsilon,$$

by Lebesgue's dominated convergence theorem and 1°. Since ε was chosen arbitrarily, the estimate (2.5) follows. □

3 Proof of (1.5)

3.1 Reductions

Fix $t \ge 0$. We start with some initial observations. First, note that in the proof of (1.5) we may assume that $\|X^*_t\|_{L^p} < \infty$, since otherwise the claim is evident. Second, it is enough to establish the estimate

$$\mathbb{E}|Y_t|^p \le C_p^p\mathbb{E}(X^*_t)^p \tag{3.1}$$

under the additional assumption that X and Y are bounded and $\mathbb{P}(|X^1_0| \ge \eta) = 1$ for some $\eta > 0$. Indeed, suppose that we have established the above inequality. To deduce the claim in the general case, suppose that X, Y are arbitrary $\mathcal{H}$-valued, continuous-path martingales such that Y is differentially subordinate to X. Adding one dimension to $\mathcal{H}$ if necessary, we may assume that the first coordinates of X and Y vanish: $Y^1_t = X^1_t = 0$ almost surely for all t. Given a positive integer N, consider the stopping time $\tau_N = \inf\{t : |X_t| + |Y_t| \ge N\}$ and fix an auxiliary parameter $\eta > 0$. Then the martingales $(((\eta,0,0,\ldots) + X_{\tau_N\wedge t})\cdot 1_{\{\tau_N>0\}})_{t\ge0}$, $(Y_{\tau_N\wedge t}\cdot 1_{\{\tau_N>0\}})_{t\ge0}$ are bounded and satisfy the differential subordination, so (3.1) gives

$$\mathbb{E}|Y_{\tau_N\wedge t}|^p 1_{\{\tau_N>0\}} \le C_p^p\mathbb{E}(\eta + X^*_{\tau_N\wedge t})^p 1_{\{\tau_N>0\}} \le C_p^p\mathbb{E}(\eta + X^*_t)^p$$

for any $t \ge 0$. Therefore, letting $\eta \to 0$, $N \to \infty$ and using Fatou's lemma, we get $\mathbb{E}|Y_t|^p \le C_p^p\mathbb{E}(X^*_t)^p$, which is the desired bound (1.5).

The final observation is that the inequality (3.1) is of the form (2.1), with

$$V(x,y,z) = |y|^p - C_p^p z^p. \tag{3.2}$$

Therefore, by the technique described in the previous section, all we need is the appropriate special function U.

3.2 Proof in the case $1/2 \le p < 1$

In this range of p, we use the function constructed in [18] and given by the formula

$$U(x,y,z) = \Big(\frac{p}{2}\Big)^{1-p/2}\left(|y|^2 - |x|^2 - \frac{2-p}{p}z^2\right)z^{p-2}. \tag{3.3}$$

It was proved in [18] that this object satisfies 1°, 2° and 4°, and a certain slightly weaker form of 3°. Thus, to complete the proof, we need to check (2.2), the full version of 3°. But this is easy: we have

$$U_z(x,y,z) = \Big(\frac{p}{2}\Big)^{1-p/2}(p-2)z^{p-3}\big(|y|^2 - |x|^2 + z^2\big) \le \Big(\frac{p}{2}\Big)^{1-p/2}(p-2)z^{p-3}|y|^2 \le 0.$$

Therefore (3.1), and hence also (1.5), follow. In the case $p < 1/2$ the calculations will be much more involved.

3.3 The case $0 < p < 1/2$

Throughout this subsection, p is a fixed exponent from the interval $(0, 1/2)$. To define the special function, we will need some additional objects. The central role in our further considerations is played by the non-linear ordinary differential equation

$$g'(y) = \frac{(p+y)g(y) - p}{y+1} - g^2(y), \qquad y \ge 0, \tag{3.4}$$

with the initial condition $g(0) = p/(p-1) \in (-1, 0)$. The function $(x, y) \mapsto ((p+y)x - p)/(y+1) - x^2$ is of class C^1 and so locally Lipschitz, which yields the existence and uniqueness of the solution. We extend this solution to its maximal domain $[0, M)$. It is not difficult to prove that $M < \infty$, but we will not need this; furthermore, replacing y with $-y$ in (3.4), it is easy to see that g can be extended 'backwards', i.e., to some interval of the form (m, M) with $m < 0$ (and, in particular, there is no problem with the existence of derivatives of g at zero). The following technical fact will be used later on.

Lemma 3.1. *Let $y_0 = y_0(p) = \inf\{y \ge 0 : g(y) + 1 = 0\}$. Then y_0 is well-defined (the set over which the infimum is taken, is nonempty) and satisfies $1 - 2p \le y_0 < \infty$. Furthermore, for all $y \in [0, y_0]$ we have*

$$(1 - p - y)g(y) + p \ge 0. \tag{3.5}$$

Proof. The fact that y_0 is finite is straightforward. Indeed, we infer from (3.4) that if $g(y) < 0$, then $g'(y) < 0$. But $g(0) < 0$, so g must be decreasing and negative on $[0, M)$. Consequently, again by (3.4), we obtain $g'(y) \le -g^2(y) \le -g^2(0) < 0$ for all y and hence the inequality $y_0 < \infty$ follows.

To show that $y_0 \ge 1 - 2p$, note first that $g(y) \le g(0) = p/(p-1)$, as we have just proved above. Therefore, we have $(p+y)g(y) - p \ge (y+1)g(y)$ and hence, by (3.4), we get $g'(y) \ge g(y) - g^2(y)$ or

$$\frac{g'(y)}{g(y)(1-g(y))} \le 1.$$

The latter estimate is equivalent to saying that $\Big(\ln\Big(\frac{-g(y)}{1-g(y)}\Big)\Big)' \le 1$, which, together with the identity $-g(0)/(1-g(0)) = p$, gives

$$\ln\frac{-g(y)}{1-g(y)} - \ln p \le y \qquad \text{for all } y. \tag{3.6}$$

Plugging $y = y_0$, we obtain $y_0 \ge -\ln(2p)$ and it remains to note that $-\ln u \ge 1 - u$ for all $u > 0$.

Finally, to show (3.5), it is enough to consider the case $y < 1-p$ only; indeed, for $y \geq 1-p$ both terms on the left of (3.5) are nonnegative. We rewrite (3.6) in the form $-g(y)(1-pe^y) \leq pe^y$, which, because of the elementary estimate $1-pe^y \geq 1-pe^{1-p} > 0$, is equivalent to $g(y) \geq -pe^y/(1-pe^y)$. Consequently,

$$(1-p-y)g(y)+p \geq -\frac{p(1-2p-y)e^y+p}{1-pe^y} \geq \frac{-p^2e^{1-p}+p}{1-pe^y} \geq 0. \qquad \square$$

The function g gives rise to another special function: $A : [0, y_0] \to \mathbb{R}$, defined by

$$A(y) = -\frac{p(y_0+2)^{p-1}(y_0+1)}{2p+2y_0+1} \exp\left(-\int_y^{y_0} g(u)\mathrm{d}u\right).$$

Actually, since g extends smoothly to some neighborhood of $[0, y_0]$, so does A (and hence we can speak of its derivatives at 0 and y_0).

Lemma 3.2. *The function A enjoys the following properties.*
(i) We have $A''(y)(y+1) - (p+y)A'(y) + pA(y) = 0$ for $y \in [0, y_0]$.
(ii) We have $A''(y) \geq A'''(y) \geq 0$ for $y \in [0, y_0]$.
(iii) We have $A''(y)(-y-1) + A'(y) \leq 0$ for $y \in [0, y_0]$.

Proof. (i) Obviously, we have $A < 0$ on $[0, y_0]$: dividing both sides of the desired differential equation by A and noting that $A'(y)/A(y) = g(y)$, $A''(y)/A(y) = g'(y) + g^2(y)$, we see that the claim is equivalent to (3.4).

(ii) Differentiating both sides of the equation in (i), we obtain

$$A'''(y)(y+1) = (p+y-1)A''(y) + A'(y)(1-p). \tag{3.7}$$

Again by (i), the expression on the right is nonnegative if and only if $(p+y-1)[(p+y)A'(y) - pA(y)] + (1-p)(1+y)A'(y) \geq 0$. We divide both sides by $A < 0$ and, after some straightforward manipulations, the estimate $A''' \geq 0$ becomes

$$g(y)[-(1-p)^2 - py - y^2] \geq p(1-p-y).$$

If $1-p-y \leq 0$, there is nothing to prove (the left-hand side is positive). For $1-p-y > 0$, the above estimate follows by multiplying the inequalities $g(y) \leq g(0) = p/(p-1)$ and $-(1-p)^2 - py - y^2 \leq -(1-p)^2 - py + y = (1-p)(p+y-1)$.

It remains to prove that $A''(y) \geq A'''(y)$. Multiplying both sides by $y+1$ and exploiting (3.7), we check that the estimate becomes $(2-p)A''(y) \geq (1-p)A'(y)$. Multiplying both sides by $y+1$ again, and applying part (i), we obtain the equivalent form

$$(3p-1-p^2+y)A'(y) \geq p(2-p)A(y),$$

or, if we divide both sides by $A < 0$,

$$(3p-1-p^2+y)g(y) \leq p(2-p). \tag{3.8}$$

However, by (3.5), we know that $(p+y-1)g(y) \leq p$; furthermore, we have $(2p-p^2)g(y) < 0 < p(1-p)$. Adding the latter two estimates we obtain (3.8).

(iii) By (i), the assertion amounts to saying that $(1-p-y)A'(y) + pA(y) \leq 0$ or $(1-p-y)g(y) + p \geq 0$, which we established in the previous lemma. $\square$

Let D_1, D_2, D_3 be three subsets of the strip $[-1,1] \times \mathbb{R}$, given by

$$\begin{aligned} D_1 &= \{(x,y) \in [-1,1] \times \mathbb{R} : |x| + |y| \leq 1\}, \\ D_2 &= \{(x,y) \in [-1,1] \times \mathbb{R} : 1 < |x| + |y| \leq y_0 + 1\}, \\ D_3 &= \{(x,y) \in [-1,1] \times \mathbb{R} : |x| + |y| > y_0 + 1\}. \end{aligned}$$

Consider the function $u : [-1,1] \times \mathbb{R} \to \mathbb{R}$, defined by the formula

$$u(x,y) = \begin{cases} -\frac{p}{2(1-p)}A(0)(y^2-x^2) + \frac{2-p}{2(1-p)}A(0) & \text{if } (x,y) \in D_1, \\ A(|x|+|y|-1) - |x|A'(|x|+|y|-1) & \text{if } (x,y) \in D_2, \\ A(y_0)\left\{-\frac{2p+y_0}{2(y_0+1)}[(|y|-y_0)^2 - x^2 + 1] + \frac{2p-1}{y_0+1}(|y|-y_0) + 2\right\} & \text{if } (x,y) \in D_3. \end{cases}$$

The function u is quite regular, as we will prove below.

Lemma 3.3. *The function u is continuous on the strip $[-1,1] \times \mathbb{R}$ and of class C^1 in the interior of this domain.*

Proof. Let us first handle the continuity; by symmetry, it is enough to check it on the set $[0,1] \times [0,\infty)$. If $x,\, y \geq 0$ satisfy $x + y = 1$, then

$$\begin{aligned} \lim_{D_1 \ni (x',y') \to (x,y)} u(x',y') &= -\frac{p}{2(1-p)}A(0)((1-x)^2 - x^2) + \frac{2-p}{2(1-p)}A(0) \\ &= A(0) + \frac{p}{1-p}A(0)x \\ &= A(0) - g(0)A(0)x = A(0) - A'(0)x = \lim_{D_2 \ni (x',y') \to (x,y)} u(x',y'). \end{aligned}$$

Similarly, if $x,\, y \geq 0$ satisfy $x + y = y_0 + 1$, then

$$\begin{aligned} \lim_{D_3 \ni (x',y') \to (x,y)} u(x',y') &= A(y_0)\left\{-\frac{2p+y_0}{2(y_0+1)}[(1-x)^2 - x^2 + 1] + \frac{2p-1}{y_0+1}(1-x) + 2\right\} \\ &= A(y_0)(1+x) \\ &= A(y_0) - xA'(y_0) \\ &= A(y_0)(1 - xg(y_0)) = \lim_{D_2 \ni (x',y') \to (x,y)} u(x',y'). \end{aligned}$$

This proves that u is a continuous function. To show that it is of class C^1, note that $u_x = 0$ on $\{0\} \times \mathbb{R}$ and $u_y = 0$ on $[-1,1] \times \{0\}$. Consequently, by symmetry, it is enough to check the smoothness on $(0,1) \times (0,\infty)$. If $x,\, y > 0$ satisfy $x + y = 1$, then

$$\lim_{D_1 \ni (x',y') \to (x,y)} u_x(x',y') = \frac{p}{1-p}A(0)x = -A''(0)x = \lim_{D_2 \ni (x',y') \to (x,y)} u_x(x',y'),$$

since by Lemma 3.2 (i), $A''(0) = pA'(0) - pA(0) = p \cdot \frac{p}{p-1}A(0) - pA(0) = \frac{p}{p-1}A(0)$. Analogously, if $x,\, y \geq 0$ satisfy $x + y = y_0 + 1$, then

$$\lim_{D_3 \ni (x',y') \to (x,y)} u_x(x',y') = \frac{2p+y_0}{y_0+1}A(y_0)x = -A''(y_0)x = \lim_{D_2 \ni (x',y') \to (x,y)} u_x(x',y'),$$

where we have used Lemma 3.2 (i) again: $A''(y_0)(y_0+1) = (p+y_0)A'(y_0) - pA(y_0) = \big((p+y_0)g(y_0) - p\big)A(y_0) = -(2p+y_0)A(y_0)$. The calculations for u_y are similar and are left to the reader. □

The special function $U : \mathcal{D} \to \mathbb{R}$, which will lead us to the estimate (1.5) for $p \in (0,1/2)$, is given by

$$U(x,y,z) = z^p u(|x|/z, |y|/z).$$

Let us check that this object enjoys all the necessary requirements.

Theorem 3.4. *The function U satisfies the conditions 1° – 4° listed in the previous section (with V given by* (3.2)*).*

Proof of 4°. The inequality is evident if $(|x|/z, |y|/z)$ belongs to the interior of D_1: then

$$\langle U_{xx}(x,y,z)h,h\rangle + 2\langle U_{xy}(x,y,z)h,k\rangle + \langle U_{yy}(x,y,z)k,k\rangle = -\frac{pz^{p-2}}{2(1-p)}A(0)(|k|^2-|h|^2),$$

and the coefficient in front of $|k|^2-|h|^2$ on the right is positive. The inequality (2.3) is also not difficult if $(|x|/z, |y|/z)$ lies in the interior of D_3. For such (x,y,z), we have

$$U(x,y,z) = -\frac{2p+y_0}{y_0+1}A(y_0)z^{p-2}(|y|^2-|x|^2) + A(y_0)(2p+y_0-1)z^{p-1}|y| + \eta,$$

where η does not depend on x or y. By Lemma 3.1, we have $2p+y_0-1\geq 0$ and hence $(x,y)\mapsto A(y_0)(2p+y_0-1)z^{p-1}|y|$ is a concave function in x and y. Consequently,

$$\begin{aligned}\langle U_{xx}(x,y,z)h,h\rangle + 2\langle U_{xy}(x,y,z)h,k\rangle + \langle U_{yy}(x,y,z)k,k\rangle \\ \leq -\frac{2p+y_0}{y_0+1}A(y_0)z^{p-2}(|k|^2-|h|^2),\end{aligned}$$

which is the desired bound, since $-\frac{2p+y_0}{y_0+1}A(y_0)z^{p-2}\geq 0$. The main technical difficulty lies in the verification of (2.3) for $(|x|/z, |y|/z)$ lying in the interior of D_2. A bit lengthy, but rather straightforward computations allow to obtain

$$\langle U_{xx}(x,y,z)h,h\rangle + 2\langle U_{xy}(x,y,z)h,k\rangle + \langle U_{yy}(x,y,z)k,k\rangle = I_1+I_2+I_3,$$

where

$$\begin{aligned}I_1 &= -z^{p-3}|x|A'''(u)\big(\langle x',h\rangle + \langle y',k\rangle\big)^2,\\ I_2 &= z^{p-1}\cdot\big(A''(u)(-u-1)+A'(u)\big)\cdot\frac{|k|^2-\langle y',k\rangle^2}{|y|},\\ I_3 &= z^{p-2}A''(u)(|k|^2-|h|^2)\end{aligned}$$

and $u=|x|/z+|y|/z-1$. By Lemma 3.2 (ii), the term I_1 is nonpositive. Similarly, the third part of that lemma combined with $\langle y',k\rangle^2\leq |k|^2$ yields $I_2\leq 0$. So,

$$\langle U_{xx}(x,y,z)h,h\rangle + 2\langle U_{xy}(x,y,z)h,k\rangle + \langle U_{yy}(x,y,z)k,k\rangle \leq z^{p-2}A''(u)(|k|^2-|h|^2),$$

and by Lemma 3.2 (ii) again, the factor in front of $|k|^2-|h|^2$ is positive. □

Proof of 3°. The inequality is equivalent to $pu-xu_x-yu_y\leq 0$ on $(0,1)\times(0,\infty)$. Observe that if (x,y) lies in the interior of D_1, D_2 or D_3, then

$$\frac{\partial}{\partial x}(pu-xu_x-yu_y) = (p-1)u_x(x,y) - xu_{xx}(x,y) - yu_{xy}(x,y)\geq 0. \tag{3.9}$$

Indeed, the inequality $u_{xx}(x,y)\leq 0$ follows from (2.3). The estimate $u_x(x,y)\leq 0$ is due to $u_{xx}\leq 0$ and the identity $u_x(0+,y)=0$ for all $y\geq 0$. Finally, $u_{xy}\leq 0$ follows again from (2.3): we know that u is linear along line segments of slope -1 and concave along line segments of slope $+1$, so $u_{xx}-2u_{xy}+u_{yy}=0$ and $u_{xx}+2u_{xy}+u_{yy}\leq 0$.

Therefore, (3.9) holds true, and hence, to establish $pu-xu_x-yu_y\leq 0$, it is enough to verify this estimate on $\{1\}\times[0,\infty)$. If $y\leq y_0$, then

$$pu(1,y)-u_x(1,y)-yu_y(1,y) = A''(y)(y+1)-(y+p)A'(y)+pA(y) = 0,$$

where the last equality is equivalent to (3.4). To prove the inequality for $y>y_0$, note that the function $y\mapsto pu(1,y)-u_x(1,y)-yu_y(1,y)$ is concave on $[y_0,\infty)$ (it is a quadratic function with the coefficient in front of y^2 equal to $-\frac{(2p+y_0)(p-2)}{2(y_0+1)}A(y_0)<0$) and it vanishes

for $y = y_0$, as we have just checked above. Thus, it is enough to prove that its derivative at $y = y_0$ is nonpositive. This is equivalent to $(p-1)u_y(1, y_0) - y_0 u_{yy}(1, y_0) \leq 0$, or

$$(p-1) \cdot A(y_0) \cdot \frac{2p-1}{y_0+1} + A(y_0) \cdot \frac{(2p+y_0)y_0}{y_0+1} \leq 0.$$

Since both terms on the left are nonpositive, the bound holds and hence (2.2) follows. □

Proof of 2°. We start with the observation that for any $y \in \mathcal{H}$ and any $z > 0$, the function $x \mapsto U(x, y, z)$ is concave on the set $\{x : |x| \leq z\}$: this follows from 4°. Furthermore, the analogous function $x \mapsto V(x, y, z)$ is constant. Therefore, it is enough to establish the majorization 2° under the additional assumption $|x| = z$; in the language of u, this is equivalent to saying that

$$u(1, y) \geq y^p - C_p^p \qquad \text{for all } y \geq 0. \tag{3.10}$$

The function $y \mapsto u(1, y)$ is of class C^1. Furthermore, it is convex: if $y < y_0$, its second derivative equals $A''(y) - A'''(y)$, which is nonnegative, by Lemma 3.2 (ii); if $y > y_0$, the function is quadratic, with a positive coefficient in front of y^2. On the other hand, the right-hand side is a concave function of y. Thus, the majorization follows at once from the observation that both sides match, along with the derivatives, at $y = y_0 + 2$:

$$u(1, y_0 + 2) = 0 = (y_0 + 2)^p - C_p^p$$

and

$$u_y(1, y_0 + 2) = A(y_0) \left\{ -\frac{2(2p+y_0)}{y_0+1} + \frac{2p-1}{y_0+1} \right\} = p(y_0+2)^{p-1}.$$ □

Proof of 1°. This is easy. The function $\xi : [-1, 1] \to \mathbb{R}$ given by $\xi(t) = U(tx, ty, z)$ is even and concave (by 4°). Consequently, we have

$$U(x, y, z) = \xi(1) \leq \xi(0) = \frac{2-p}{2(1-p)} A(0) < 0,$$

which is the desired estimate. □

4 Sharpness

There are two general methods which are typically used to prove that the constant involved in a given martingale inequality is the best possible. The first type of approach rests on the construction of appropriate extremal processes, for which the equality is attained, or asymptotically attained (i.e., in the limit). This natural method has been applied in many cases: see e.g. [8, 10, 22]. The second method is more abstract and indirect in nature: one assumes that an estimate under investigation holds with some constant and then proves the existence of an associated special function, enjoying appropriate size and regularity conditions. Then, using these properties in the right order, one shows that the constant in question must be bounded below by an appropriate quantity. This type of reasoning originates in Burkholder's paper [9] and has been successfully exploited in various contexts: see e.g. [9, 18, 19, 23].

We have decided to follow the first path and show the optimality of the constant C_p explicitly, with the use of examples. This approach has the additional advantage that some elements of the proof will be of key importance in the next section, where we describe some steps leading to the formula for the special function U for $0 < p < 1/2$. We consider the cases $1/2 \leq p < 1$ and $0 < p < 1/2$ separately. The first case is simpler and hence its analysis can be treated as a 'basis' for the more complicated reasoning for $p < 1/2$.

4.1 The case $1/2 \le p < 1$

Let us first describe the behavior of the extremal pair (X, Y), or rather the extremal triple (X, Y, X^*). The conditions appearing below are informal and they will be rigorously specified later. We assume that $\eta < 1$ is a fixed parameter, and consider the process (X, Y, X^*) whose evolution is governed by the following requirements.

(i) We have $(X_0, Y_0, X_0^*) = (1, 0, 1)$ almost surely.

(ii) The state of the form $(x, 0, |x|)$, $|x| \ge 1$, leads to $((1+\delta)x, 0, (1+\delta)|x|)$ (where δ is 'infinitesimally small') or to one of the points $((1-\eta)x, \eta x, |x|)$, $((1-\eta)x, -\eta x, |x|)$.

(iii) The state of the form $((1-\eta)x, \pm\eta x, |x|)$ leads to $(x, \pm 2\eta x, |x|)$ or to $(-\eta x, \pm(\eta - 1)x, |x|)$.

(iv) The state of the form $(-\eta x, \pm(\eta-1)x, |x|)$ leads to $(-x, 0, |x|)$ or to $((\eta-1)x, \mp\eta x, |x|)$.

(v) The points of the form $(x, \pm 2\eta x, |x|)$ are absorbing.

To gain some intuition about the evolution, see Figure 1 below. Actually, we have formulated the above conditions as if the triple (X, Y, X^*) was a discrete-time process. To avoid confusion, let us explain that the word "leads" means that the triple moves, in a continuous manner, along the line segments joining the appropriate states.

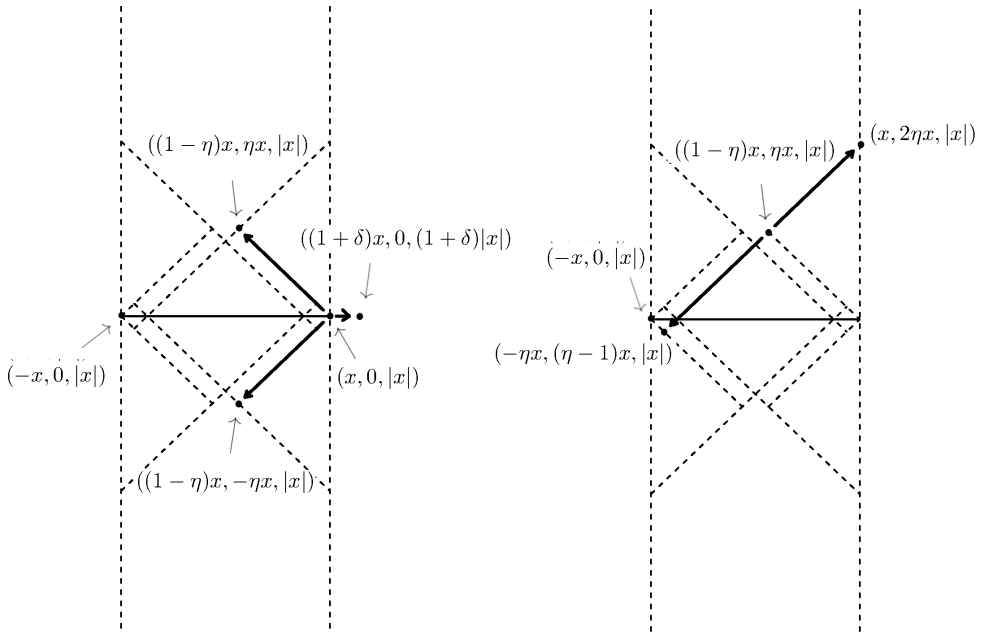

Figure 1: Bold arrows indicate the local evolution of the pair (X, Y), under the assumption that $X^* = |x|$: the rule (ii) (left) and the rule (iii) (right). Inside the strip bounded by the vertical lines, the pair evolves along the dotted lines, until it gets to a 'T-shaped' crossing: then it changes the direction to the perpendicular line segment.

The formal description of the triple (X, Y, X^*) is the following. The process Y will be the appropriately stopped one-dimensional Brownian motion started at zero and X will be a stochastic integral of Y:

$$X_t = 1 + \int_{0+}^{t} H_s \mathrm{d}Y_t, \qquad t \ge 0.$$

Thus, all we need is the specification of the process H. To this end, we will construct inductively a sequence $(\sigma_n)_{n\geq 0}$ of stopping times. Set $\sigma_0 \equiv 0$ and let

$$\sigma_{2n+1} = \inf\Big\{t \geq \sigma_{2n} : |Y_t| \geq \eta X_t^*\Big\},$$

$$\sigma_{2n+2} = \inf\Big\{t \geq \sigma_{2n+1} : |Y_t| = 2\eta X_t^* \text{ or } Y_t = (\eta-1)X^*_{\tau_{2n+1}} \cdot \operatorname{sgn} Y_{\sigma_{2n+1}}\Big\}$$

for $n = 0, 1, 2, \ldots$. Here and in what follows, we will use the (slightly unusual) convention $\operatorname{sgn} 0 = 1$. The predictable 'control process' H is given by

$$H_t = \begin{cases} \operatorname{sgn} X_{\sigma_{2n}} \operatorname{sgn} Y_t & \text{if } t \in [\sigma_{2n}, \sigma_{2n+1}), \\ \operatorname{sgn}(X_{\sigma_{2n+1}}) \operatorname{sgn}(Y_{\sigma_{2n+1}}) & \text{if } t \in [\sigma_{2n+1}, \sigma_{2n+2}). \end{cases}$$

It is easy to see that then the process (X, Y, X^*) evolves according to the rules (i)-(v) described above.

Let us gather some basic information about the triple (X, Y, X^*), which follows directly from the above construction. First, note that Y is differentially subordinate to X: indeed, we have $|Y_0| = 0 \leq 1 = |X_0|$ and the transforming process H takes values in the set $\{-1, 1\}$. Furthermore, on the set $\{s \geq 0 : X_s^* \text{ increases}\}$, we have $|X_s| = X_s^*$ and $Y_s = 0$. Finally, we have $|Y_t| \leq 2\eta X_t^*$ almost surely for all t, and the process (X, Y, X^*) terminates ultimately when the equality $|Y_t| = 2\eta X_t^*$ is experienced.

The next step of the analysis is to show that if $2\eta < \sqrt{2/p}$, then $X_\infty^* \in L^p$. The key observation is that the process $(U(X_t, Y_t, X_t^*))_{t\geq 0}$ is a local martingale. To see this, we will apply Itô's formula and check that the finite-variation part vanishes. First, for any $t \geq 0$ we have

$$\int_{0+}^{t} U_z(X_s, Y_s, X_s^*)\mathrm{d}X_s^* = 0.$$

This follows from the identity $U_z(\pm x, 0, |x|) = 0$ for all $x > 0$ and the fact that on the set where X_s^* increases, we have $|X_s| = X_s^*$ and $Y_s = 0$ (which follows directly from the above construction). Next, since H takes values in $\{-1, 1\}$, we have

$$\int_{0+}^{t} U_{xx}(\ldots)\mathrm{d}[X, X]_s + 2\int_{0+}^{t} U_{xy}(\ldots)\mathrm{d}[X, X]_s + \int_{0+}^{t} U_{yy}(\ldots)\mathrm{d}[X, X]_s = \int_{0+}^{t} (-H_s^2 + 1)\mathrm{d}s = 0.$$

This establishes the local martingale property of $(U(X_t, Y_t, X_t^*))_{t\geq 0}$. Consider the sequence $\tau_1 \leq \tau_2 \leq \ldots$ of stopping times given by $\tau_n = \inf\{t > 0 : X_t^* = n\}$. Then we obviously have $|X_{\tau_n}| = X^*_{\tau_n}$ (since this is either equal to n, or, if the triple (X, Y, X^*) terminated earlier, we have $|X_{\tau_n}| = X^*_{\tau_n} = |Y_{\tau_n}|/(2\eta)$). Furthermore, the sequence $(\tau_n)_{n\geq 1}$ is localizing for $(U(X_t, Y_t, X_t^*))_{t\geq 0}$, since for each n the process $(X^*_{\tau_n\wedge t})_{t\geq 0}$, and hence also $(X_{\tau_n\wedge t})_{t\geq 0}$ and $(Y_{\tau_n\wedge t})_{t\geq 0}$, are bounded. Consequently, we have

$$\mathbb{E}U(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}) = \lim_{t\to\infty} \mathbb{E}U(X_{\tau_n\wedge t}, Y_{\tau_n\wedge t}, X^*_{\tau_n\wedge t}) = U(1, 0, 1) = -\left(\frac{2}{p}\right)^{p/2}.$$

Plugging the formula for U and using the facts that $|Y_{\tau_n}| \leq 2\eta X^*_{\tau_n}$ and $|X_{\tau_n}| = X^*_{\tau_n}$, we obtain

$$-\frac{2}{p} = \mathbb{E}\left(Y_{\tau_n}^2 - X_{\tau_n}^2 - \frac{2-p}{p}(X^*_{\tau_n})^2\right)(X^*_{\tau_n})^{p-2} \leq \left(4\eta^2 - \frac{2}{p}\right)\mathbb{E}(X^*_{\tau_n})^p.$$

Therefore, if $2\eta < \sqrt{2/p}$, then we get

$$\mathbb{E}(X^*_{\tau_n})^p \leq \frac{2/p}{2/p - 4\eta^2}$$

and letting $n \to \infty$ we get $X^*_\infty \in L^p$. Consequently, Y converges almost surely (which follows, for example, from Burkholder-Davis-Gundy inequalities). Denoting the limit by Y_∞, we obtain $\mathbb{P}(|Y_\infty| = 2\eta X^*_\infty) = 1$ (by the construction of (X, Y, X^*)) and hence, by Fatou's lemma,

$$\liminf_{t\to\infty} \|Y_t\|_{L^p} \geq \|Y_\infty\|_{L^p} = 2\eta \|X^*_\infty\|_{L^p} = \lim_{t\to\infty} 2\eta \|X^*_t\|_{L^p}.$$

Since 2η can be made arbitrarily close to $\sqrt{2/p}$, the sharpness is established.

4.2 The case $p < 1/2$

The above reasoning cannot be carried over to exponents smaller than $1/2$. Indeed, the limiting value of the parameter η for which the above construction of the triple (X, Y, X^*) makes sense, is equal to 1 (for bigger η, the line segments 'stick out' of the picture). Recalling the key identity $2\eta = \sqrt{2/p}$, we see that the endpoint value $\eta = 1$ leads to the boundary case $p = 1/2$. However, the above discussion gives some indications about the structure of the triple (X, Y, X^*) for the small values of the parameter p. Namely, let us emphasize that for $p \geq 1/2$, we had the *pointwise* identity $|Y_\infty| = 2\eta X^*_\infty$, even stronger: $|Y_\infty| = 2\eta |X_\infty| = 2\eta X^*_\infty$, where η could be chosen arbitrarily close to C_p. It seems quite natural to search, in the case $p < 1/2$, for a triple satisfying the same condition. A little thought and experimentation lead to the following candidate. Let Y be a standard one-dimensional Brownian motion started at zero and define

$$X_t = 1 + \int_0^t H_s \mathrm{d}Y_s,$$

where

$$H_s = \begin{cases} -\operatorname{sgn}(X_s)\operatorname{sgn}(Y_s) & \text{if } |X_s| + |Y_s| < (2\eta - 1)X^*_s, \\ 1 & \text{if } |X_s| + |Y_s| \geq (2\eta - 1)X^*_s. \end{cases}$$

Furthermore, define the termination time by $\sigma = \inf\{t : |Y_t| = 2\eta |X_t| = 2\eta X^*_t\}$.

Here is some (very rough and informal) intuition about the behavior of (X, Y, X^*). The process starts at $(1, 0, 1)$. For any time t, if $|X_t| + |Y_t| < (2\eta - 1)X^*_t$, then the *pair* (X, Y) moves (locally) along a line segment of slope $-\operatorname{sgn}(X_t Y_t)$ until one of the coordinates reaches zero and the slope switches its sign. On the other hand, if we have $|X_t| + |Y_t| = (2\eta - 1)X^*_t$, then the pair (X, Y) is forced to move along the line segment of slope 1 until $|Y| = 2\eta |X| = 2\eta X^*$ (then we stop) or $|Y| = (2\eta - 2)|X| = (2\eta - 2)X^*$. If the latter happens, say, at some time τ, then the third coordinate increases instantly and the equality $|X| + |Y| = (2\eta - 1)X^*$ is no longer valid: we actually have $|X_t| + |Y_t| < (2\eta - 1)X^*_t$ almost surely for t bigger, but infinitesimally close to τ. Hence, we can apply the preceding evolution rule. See Figure 2 below.

As in the previous case, we will prove that if $2\eta < y_0 + 2$, then $X^*_\infty \in L^p$ and $\mathbb{P}(|Y_\infty| = 2\eta X^*_\infty) = 1$, which immediately yield the sharpness. To this end, we will study the local martingale property of $U(X, Y, X^*)$ with the use of Itô's formula, exactly in the same manner as in the case $p \geq 1/2$. Actually, to ensure that the function is sufficiently regular, we will need to modify it slightly. First, define $u^\eta : [-1, 1] \times \mathbb{R} \to \mathbb{R}$, by $u^\eta = u$ if $1 \leq |x| + |y| \leq 2\eta - 1$; next, assume that u^η is linear along each of the line segments $y - x = \pm(2\eta - 1)$, $x \in [-1, 1]$; finally, extend u^η to the whole $[-1, 1] \times \mathbb{R}$ (arbitrarily) so

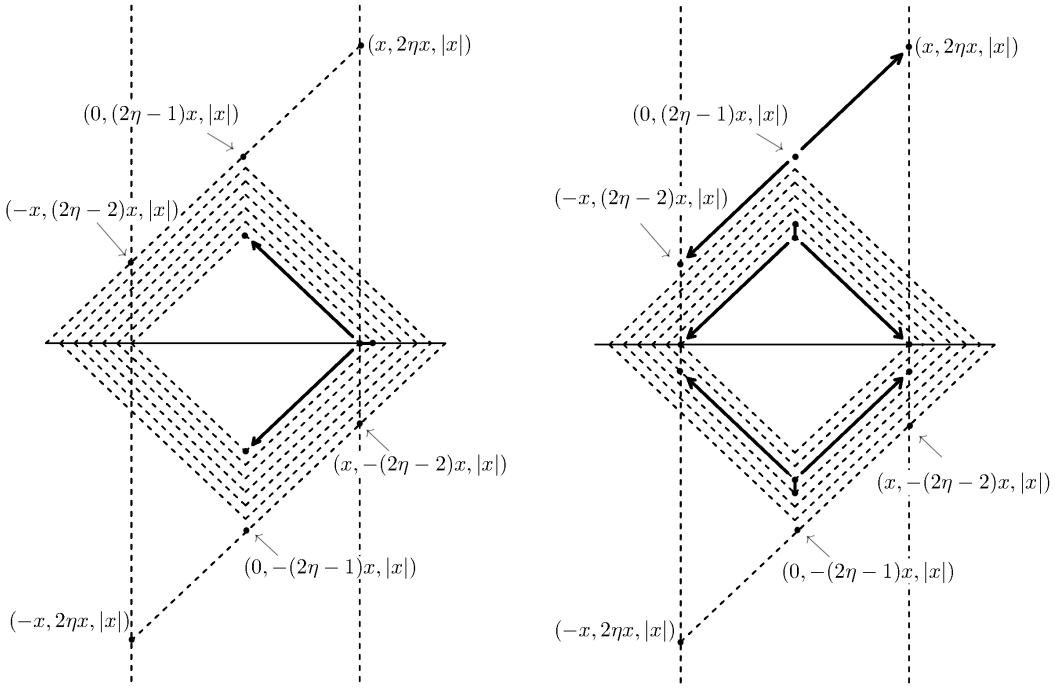

Figure 2: Bold arrows indicate the evolution of the pair (X, Y), under the assumption $X^* = |x|$. The pair (X, Y) moves locally along the dotted line segments (of slope $\operatorname{sgn}(XY)$).

that it is of class C^2. Note that in particular we have

$$\begin{aligned} u^\eta(1, 2\eta) &= 2u^\eta(0, 2\eta-1) - u(-1, 2\eta-2) \\ &= A(2\eta-2) + A'(2\eta-2) = A(2\eta-2)(1+g(2\eta-2)) < 0, \end{aligned} \tag{4.1}$$

where the last inequality follows from the fact that $2\eta - 2 < y_0$ (which gives $g(2\eta-2) > -1$, directly from the definition of y_0). Next, introduce $U^\eta : \mathcal{D} \to \mathbb{R}$ by $U^\eta(x, y, z) = z^p u^\eta(|x|/z, |y|/z)$. Observe that U coincides with U^η on the set $\{(x, y, z) : z \leq |x| + |y| \leq (2\eta-1)z\}$. Arguing as previously, we check that the process $(U^\eta(X_t, Y_t, X_t^*))_{t\geq 0}$ is a local martingale. Namely, we apply Itô's formula and verify that the finite-variation parts vanish. First, note that

$$\int_0^t U_z^\eta(X_s, Y_s, X_s^*)\mathrm{d}X_s^* = 0.$$

Indeed, we have $|X_s| = X_s^*$ and $|Y_s| \leq (2\eta-2)X_s^*$ on the set where X^* increases; furthermore, if $0 \leq |y| \leq (2\eta-2)|x|$, then $U_z^\eta(\pm x, y, |x|) = U_z(\pm x, y, |x|) = 0$. In addition, we have

$$\int_0^t U_{xx}^\eta(\ldots)\mathrm{d}[X, X]_s + 2\int_0^t U_{xy}^\eta(\ldots)\mathrm{d}[X, Y]_s + \int_0^t U_{yy}^\eta(\ldots)\mathrm{d}[Y, Y]_s = 0$$

for all $t \geq 0$. To see this, recall the directions of the evolution of the pair (X, Y), discussed above; it is easy to check that the function U^η is linear along the line segments over which (X, Y) moves.

This yields the local martingale property of $(U^\eta(X_t, Y_t, X_t^*))_{t\geq 0}$. As previously, we introduce the sequence $\tau_1 \leq \tau_2 \leq \ldots$ of stopping times, given by $\tau_n = \inf\{t > 0 :$

$X_t^* = n\}$. This sequence localizes $(U^\eta(X_t, Y_t, X_t^*))_{t\geq 0}$, since for each n, the process $(U^\eta(X_{\tau_n\wedge t}, Y_{\tau_n\wedge t}, X^*_{\tau_n\wedge t}))_{t\geq 0}$ is bounded. Consequently, we have

$$\mathbb{E}U^\eta(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}) = \lim_{t\to\infty} \mathbb{E}U^\eta(X_{\tau_n\wedge t}, Y_{\tau_n\wedge t}, X^*_{\tau_n\wedge t}) = U^\eta(1,0,1) < 0.$$

However, as in the case $p \geq 1/2$, we have $|X_{\tau_n}| = X^*_{\tau_n}$ almost surely. Actually, there are two options, either $Y_{\tau_n} = 0$ and $U^\eta(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}) = (X^*_{\tau_n})^p u^\eta(1,0) < (X^*_{\tau_n})^p u^\eta(1, 2\eta)$, or $|Y_{\tau_n}| = 2\eta X^*_{\tau_n}$ and $U^\eta(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}) = (X^*_{\tau_n})^p u^\eta(1, 2\eta)$. Plugging this above gives

$$u^\eta(1, 2\eta)\mathbb{E}(X^*_{\tau_n})^p > U^\eta(1,0,1),$$

or equivalently, by (4.1), $\mathbb{E}(X^*_{\tau_n})^p < U^\eta(1,0,1)/u^\eta(1,2\eta)$. Letting $n \to \infty$ we get $X^*_\infty \in L^p$, so Y converges almost surely to some finite random variable Y_∞ satisfying $\mathbb{P}(|Y_\infty| = 2\eta X^*_\infty) = 1$. This yields the sharpness exactly in the same manner as in the case $p \geq 1/2$.

5 On the search of the special function, $0 < p < 1/2$

Now we will describe some informal steps which have led us to the discovery of the optimal constant C_p and the special function $U : \mathcal{D} \to \mathbb{R}$ used in the case $0 < p < 1/2$. In principle, the problem of the identification of such objects can be investigated with the use of two approaches: analytically or with the use of probabilistic methods. Actually, in a given situation, it is often worth to exploit a combination of both techniques; this will also be the case here. In the considerations below, we will focus on the real-valued setting: we assume that $\mathcal{H} = \mathbb{R}$. The passage to the vector-valued case is standard: in the formula for the obtained special function, we interpret all the absolute values as norms in $\mathcal{H}$.

5.1 Analytic contribution

Let $V : \mathcal{D} \to \mathbb{R}$ be a locally bounded function. Suppose that we are interested in showing the estimate

$$\mathbb{E}V(X_t, Y_t, X_t^*) \leq 0, \qquad t \geq 0, \tag{5.1}$$

for any pair (X, Y) of differentially subordinate martingales. In particular, we want to show this estimate if Y is the stochastic integral, with respect to X, of some predictable process H with values in $\{-1, 1\}$. The fundamental property of the special function U to be found is that for any X, Y as above, the process $U(X, Y, X^*)$ must be a supermartingale majorizing $V(X, Y, X^*)$. If we also have $U(X_0, Y_0, X_0^*) \leq 0$, then we have

$$\mathbb{E}V(X_t, Y_t, X_t^*) \leq \mathbb{E}U(X_t, Y_t, X_t^*) \leq \mathbb{E}U(X_0, Y_0, X_0^*) \leq 0 \tag{5.2}$$

and the claim follows. Now *suppose* that U is of class C^2. We may freely impose this requirement, since our goal is to discover *a candidate* for the special function; even if the final function we come up with does not enjoy this regularity, we may hope that it will still lead to the desired bound anyway (and indeed, this is the case here).

If U is of class C^2 and one applies Itô's formula, the required supermartingale property gives rise to two partial differential inequalities:

$$U_z(x, y, z) \leq 0 \qquad \text{if } |x| = z \tag{5.3}$$

and

$$U_{xx}(x, y, z) \pm 2U_{xy}(x, y, z) + U_{yy}(x, y, z) \leq 0 \qquad \text{for } (x, y, z) \in \mathcal{D}. \tag{5.4}$$

Now, if U is the function leading to the sharp inequality, then there is a pair (X, Y) for which equality is attained (possibly asymptotically, in the limit). For simplicity, let us

assume that equality holds in (5.1) for some $t > 0$ and some nontrivial pair of processes. This implies that equalities must also hold in (5.2), and hence the partial differential inequalities (5.3), (5.4) also become equalities (*at least* on the range of the extremal triple (X, Y, X^*)): we have seen this, very transparently, in the previous section. As a priori we do not know how the extremal process (X, Y, X^*) will look like, a natural starting point is to assume that these equalities hold everywhere. Furthermore, the first inequality in (5.2) enforces us to impose the condition $U \geq V$ and the existence of an extremal pair implies that equality must hold on a nontrivial set.

Summarizing the discussion so far, we search for a C^2 function $U : \mathcal{D} \to \mathbb{R}$ which satisfies the following three assumptions:

(A1) We have $U_z(x, y, z) = 0$ if $|x| = z$ and $y \in \mathbb{R}$.

(A2) For any $(x, y, z) \in \mathcal{D}$ there is a sign $\mathcal{E} = \mathcal{E}(x, y, z) \in \{-1, 1\}$ such that

$$U_{xx}(x, y, z) + 2\mathcal{E}U_{xy}(x, y, z) + U_{yy}(x, y, z) = 0.$$

(A3) We have $U \geq V$ and equality holds for at least one point.

In our case, we have $V(x, y, z) = |y|^p - C_p^p z^p$ for some a priori unknown constant C_p. Note that V is homogeneous of order p, so it seems plausible that the same is true for U. Setting $U(x, y, z) = z^p u(x/z, y/z)$, we see that the problem reduces to the search of a symmetric function $u : [-1, 1] \times \mathbb{R} \to \mathbb{R}$ of class C^2 for which

(a1) We have $pu(1, y) - u_x(1, y) - yu_y(1, y) = 0$ if $y \geq 0$.

(a2) For any $(x, y) \in [-1, 1] \times \mathbb{R}$ there is a sign $\varepsilon = \varepsilon(x, y) \in \{-1, 1\}$ such that

$$u_{xx}(x, y) + 2\varepsilon u_{xy}(x, y) + u_{yy}(x, y) = 0.$$

(a3) We have $u(x, y) \geq |y|^p - C_p^p$ and equality holds for at least one point.

Let us emphasize that the above set of properties has been obtained on a base of some more or less reasonable assumptions and guesses. One should also keep in mind that there might be no function u satisfying the above requirements (for instance, equalities in (a1), (a2) may hold only on some part of the domain; the function u may not be of class C^2; etc.). So, one should treat the above requirements flexibly, allowing some small modifications if necessary. We would also like to mention here that in general, there is no uniqueness; so, there might be several (actually, infinitely many) special functions leading to the estimate (5.1). In some situations, one might be interested in the best (i.e., pointwise smallest) special function, but in practice one often searches for the simplest function, having as non-complex formula as possible. Then one can expect the analysis to avoid many tedious technical issues. This will also be the case here.

To proceed, we need to have more information about the structure of the signs in (a2) and the set of those (x, y) for which equality in (a3) is attained. It is convenient to incorporate probabilistic arguments.

5.2 Probabilistic contribution

There is an abstract formula for special function U corresponding to (5.1). For any x, $y \in \mathbb{R}$, consider the class $M(x, y)$ which consists of all pairs (X, Y) of continuous-time bounded martingales such that $X_0 = x$ and

$$Y_t = y + \int_0^t H_s \mathrm{d}X_s, \qquad t \geq 0, \tag{5.5}$$

for some predictable process H with values in $\{-1, 1\}$. Here the probability space and the filtration are also assumed to vary. Define $U^0 : \mathcal{D} \to \mathbb{R}$ by the formula

$$U^0(x, y, z) = \sup \mathbb{E}V(X_\infty, Y_\infty, X^*\infty \vee z), \tag{5.6}$$

where the supremum is taken over all $(X,Y)\in M(x,y)$. Here X_∞, Y_∞ are the almost sure limits of X and Y as $t\to\infty$, which exist due to the boundedness assumption. By a standard change-of-time argument, one may assume that X is a stopped one-dimensional Brownian motion, started at x and run at some fixed speed a (that is, $X_t=x+B_{at}$ for $t\geq 0$, where B is a standard Brownian motion). The function U^0 has all the required properties. Indeed, we have $U^0\geq V$ (simply consider the constant pair $(X,Y)\equiv(x,y)$ in the definition) and, using standard Markovian arguments, one checks that $U^0(X,Y,X^*)$ is a supermartingale if Y is a stochastic integral of X. The reason why we have decided to denote this object by U^0 (that is, with the additional superscript) is that this is actually the best (i.e., pointwise smallest) special function. Of course, there is no guarantee that U^0 is of class C^2, or even continuous; however, we do not worry about any regularity issues. Again, we stress that the main purpose of this section is to provide some ideas which led us to the function of Section 3 (which was rigorously studied there).

In our case, when $V(x,y,z)=|y|^p-C_p^pz^p$, the formula (5.6) becomes

$$U^0(x,y,z)=\sup\mathbb{E}\Big\{|Y_\infty|^p-C_p^p[(X_\infty^*)^p\vee z^p]\Big\},\tag{5.7}$$

where the supremum is taken over appropriate processes as above. Now it can be proved formally that the special function is homogeneous of order p (which was *assumed* in the analytic setting): it follows from the simple scaling property that $(X,Y)\in M(x,y)$ if and only if $(\lambda X,\lambda Y)\in M(\lambda x,\lambda y)$ for any $\lambda>0$. Next, the following ideas turn out to be helpful. By (5.7), given (x,y,z), we want to maximize the expectation $\mathbb{E}\Big\{|Y_\infty|^p-C_p^p(X_\infty^*\vee z)^p\Big\}$; naively speaking, we would like to try to maximize $\mathbb{E}|Y_\infty|^p$, keeping $\mathbb{E}(X_\infty^*\vee z)^p$ relatively small. As we mentioned above, we may assume that X is a stopped Brownian motion started at x and hence all we need is the identification of the 'control' process H in (5.5). There is a perfect analogy between this problem and the problem of finding the sign function $\mathcal{E}$ in the above analytic context. Namely, both these objects (H and $\mathcal{E}$) determine the directions of line segments along which the extremal pairs (X,Y) evolve.

First we will try to identify 'the shape' of the termination time for X. We look at the quantity $\mathbb{E}(X_\infty^*\vee z)^p$, which we want to keep relatively small. Obviously, the function $t\mapsto(X_t^*\vee z)^p$ is nondecreasing and if $|X_t|$ is strictly less than X_t^*, then $(X_t^*\vee z)$ does not increase for some time. On the other hand, let us inspect the behavior of the function $t\mapsto\mathbb{E}|Y_t|^p$, which we want to make as large as possible. If Y_t is away from zero, then this expectation decreases (which is not 'profitable' from our viewpoint): this follows from the local concavity of the function $y\mapsto|y|^p$. On the other hand, when Y visits zero, then the expectation will experience a huge increase in a short time, because of the cusp of the function $y\mapsto|y|^p$.

The above discussion suggests the following behavior of the extremal pair (X,Y): we should try to keep Y near zero, and stop when it gets far from the origin. What does 'far' mean? In the light of the homogeneity, it seems natural to conjecture that we should stop if $|Y|\geq\eta X^*$ for some unknown parameter η. Motivated by the above analysis of $t\mapsto\mathbb{E}X_t^*$, it seems plausible to guess that at the moment τ of termination, we should have $|X_\tau|=X_\tau^*$ (otherwise, it might be profitable to wait for a while...). Can we say anything about the threshold η? Take the extremal pair (X,Y) of differentially subordinate martingales (for which equality is attained). If our intuition is correct, then we have $|Y_\infty|\geq\eta|X_\infty|=\eta X_\infty^*$ almost surely and hence

$$\eta\|X_\infty^*\|_{L^p}\leq\|Y_\infty\|_{L^p}\leq C_p\|X_\infty^*\|_{L^p}.$$

This leads to the natural guess $\eta = C_p$.

Finally, we address the structure of the 'control' process H. We want to keep Y close to zero but, on the other hand, we want it to reach $\{-C_pX^*, C_pX^*\}$ after some finite time. It is not difficult to guess what H should be for some particular points (formally: how to define H if the triple (X, Y, X^*) lies in some particular set). For example, suppose that at some time t, (X_t, Y_t) got to the point lying on line segment joining $(0, (C_p - 1)z)$ and (z, C_pz), where $z = X_t^*$. Then the pair should evolve along this line segment until it gets to one of its endpoints: if it reaches (z, C_pz) then the whole process terminates ultimately, if it gets to the other endpoint, then the evolution is continued. A similar analysis can be carried over for the symmetric line segments, i.e., those joining $(0, (C_p-1)z)$ and $(-z, C_pz)$; $(0, -(C_p-1)z)$ and $(z, -C_pz)$; $(0, -(C_p-1)z)$ and $(-z, -C_pz)$. In the language of H, this means that if (X, Y) visits one of these segments, then we should take $H = \operatorname{sgn}(XY)$ (until we leave the segment); in the language of the sign function ε from the analytic setting, this means $\varepsilon(x, y) = \operatorname{sgn}(xy)$ for (x, y) belonging to the union of the segments (with $z = 1$).

Actually, it is quite tempting to take $H = \operatorname{sgn}(XY)$ everywhere, but this is not the right choice. Indeed, if we took $(X_0, Y_0, X_0^*) = (1, 1, 1)$, then this identity for H would simply give $X = Y$. This cannot be the extremal pair for (5.1), since it would imply that the best constant C_p is equal to 1 (and, in the light of the case $p \geq 1/2$, one expects it to be bigger). So, let us try to switch the sign of the process H to $-\operatorname{sgn}(XY)$: this leads to the conjecture that for x, y lying between the line segments, we have $\varepsilon(x, y) = -\operatorname{sgn}(xy)$. This turns out to be the right guess, as it leads to the appropriate special function on the crucial part of the domain (see Figure 3 below).

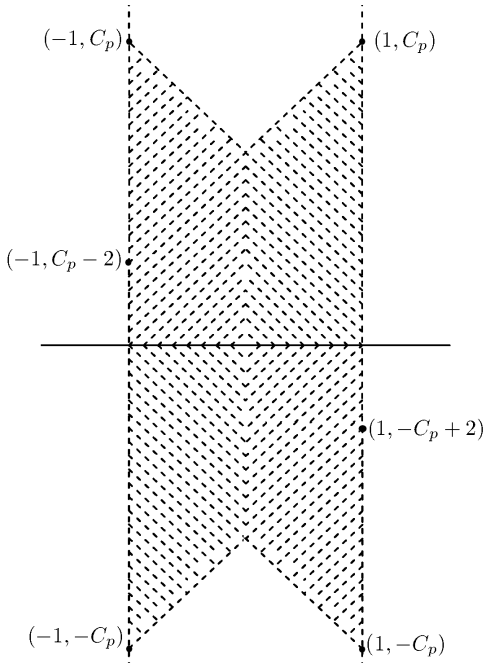

Figure 3: The slopes of the dotted line segments determine the sign function ε on the set $\{(x, y) : -|x| + |y| \leq C_p - 1\}$, and the directions along which the extremal pairs (X, Y) move along.

5.3 Explicit formula

We return to the analytic setting. By symmetry, it is enough to find the function u on the quadrant $[0,1]\times[0,\infty)$ of the strip $[-1,1]\times\mathbb{R}$. Introduce the auxiliary functions $a(y)=u(0,y)$ and $b(y)=u(1,y)$ for $y\geq 0$. The above discussion suggests

$$u(x,y)=(1-x)a(x+y)+xb(x+y-1) \qquad \text{if } 1\leq x+y\leq C_p-1, \tag{5.8}$$

$$u(x,y)=(1-x)a(C_p-1) \qquad \text{if } y-x=C_p-1. \tag{5.9}$$

Since we want U to be of class C^2, the equality (5.8) implies

$$-a(y)+a'(y)+b(y-1)=u_x(0,y)=0 \qquad \text{for } 1\leq y\leq y_0-1. \tag{5.10}$$

Furthermore, the assumption (a1) above yields

$$a(y+1)+(p-1)b(y)-b'(y)(y+1)=0 \quad \text{for } 0\leq y\leq y_0-2. \tag{5.11}$$

Combining these two differential equations, we obtain

$$a''(y+1)(y+1)-(p+y)a'(y+1)+pa(y+1)=0, \tag{5.12}$$

which is the main building block. If we substitute $g(y)=a'(y+1)/a(y+1)$, then the above equation gives (3.4). The initial condition $g(0)=p/(p-1)$ follows from (5.10), (5.11) and the equality $b'(y)=u_y(1,0)=0$ which must hold due to the symmetry and regularity of U. To see how the parameter y_0 arises, note that by (5.9),

$$a'(C_p-1)=u_x(0,C_p-1)+u_y(0,C_p-1)=-a(C_p-1),$$

or $g(C_p-2)=-1$. In other words, one conjectures that $C_p=y_0+2$.

The above discussion leads to the appropriate definition of u on the set $\{(x,y)\in[0,1]\times[0,\infty): 1\leq x+y\leq y_0-1 \text{ or } y-x=y_0-1\}$. What about the remaining part of the domain? Motivated by related constructions from the literature, we assumed that u is given by the quadratic function of the form $\alpha(y^2-x^2)+\beta$ if $x+y\leq 1$, and $\gamma(y^2-x^2)-\delta y-\eta$ for $x+y>y_0+1$. The parameters α, β, γ, δ and η were determined by the requirement $u\in C_1$ and $u_x(0,y)=0$. Fortunately, the function we come up with enjoys all the required properties and thus leads to the main result of this paper.

References

[1] Astala, K., Iwaniec, T., Prause I., and Saksman, E.: Burkholder integrals, Morrey's problem and quasiconformal mappings. *J. Amer. Math. Soc.* **25**, (2012), 507–531. MR-2869025

[2] Arcozzi, N.: Riesz transforms on compact Lie groups, spheres and Gauss space. *Ark. Mat.* **36**, (1998), 201–231. MR-1650585

[3] Baernstein II, A., and Montgomery-Smith, S. J.: Some conjectures about integral means of ∂f and $\overline{\partial} f$. *Complex analysis and differential equations (Uppsala, 1997)*, 92–109, Acta Univ. Upsaliensis Skr. Uppsala Univ. C Organ. Hist., **64**, Uppsala Univ., Uppsala, 1999. MR-1758918

[4] Bañuelos, R., and Bogdan, K.: Lévy processes and Fourier multipliers. *J. Funct. Anal.* **250**, (2007), 197–213. MR-2345912

[5] Bañuelos, R., and Méndez-Hernandez, P. J.: Space-time Brownian motion and the Beurling-Ahlfors transform. *Indiana Univ. Math. J.* **52**, (2003), 981–990. MR-2001941

[6] 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. MR-3373431

[7] 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. MR-1370109

[8] Burkholder, D. L.: Boundary value problems and sharp inequalities for martingale transforms. *Ann. Probab.* **12**, (1984), 647–702. MR-0744226

[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. MR-1108183

[10] Burkholder, D. L.: Sharp norm comparison of martingale maximal functions and stochastic integrals. *Proceedings of the Norbert Wiener Centenary Congress*, 1994 (East Lansing, MI, 1994), 343–358, Proc. Sympos. Appl. Math., **52**, Amer. Math. Soc., Providence, RI, 1997. MR-1440921

[11] Carbonaro, A., and Dragičević, O.: Functional calculus for generators of symmetric contraction semigroups. *Duke Math. J.* **166**, (2017), 937–974. MR-3626567

[12] Dellacherie, C. and Meyer, P.-A.: Probabilities and potential B: Theory of martingales. *North Holland*, Amsterdam, 1982. MR-0745449

[13] Dragičević, O., and Volberg, A.: Bellman functions and dimensionless estimates of Littlewood-Paley type. *J. Op. Theory* **56**, (2006), 167–198. MR-2261616

[14] Geiss, S., Montgomery-Smith, S. J. and Saksman, E.: On singular integral and martingale transforms. *Trans. Amer. Math. Soc.* **362** (2010), 553–575. MR-2551497

[15] Métivier, M.: Semimartingales: a course on stochastic processes, volume 2. *Walter de Gruyter*, Berlin-New York, 1982. MR-0688144

[16] Osękowski, A.: Sharp LlogL inequality for Differentially Subordinated Martingales. *Illinois J. Math.* **52**, (2008), 745–756. MR-2546005

[17] Osękowski, A.: Sharp inequality for martingale maximal functions and stochastic integrals. *Illinois J. Math.* **54**, (2010), 1133–1156. MR-2928348

[18] Osękowski, A.: Maximal inequalities for continuous martingales and their differential subordinates. *Proc. Amer. Math. Soc.* **139**, (2011), 721–734. MR-2736351

[19] Osękowski, A.: Sharp martingale and semimartingale inequalities. *Monografie Matematyczne* **72**, Birkhäuser, Basel 2012. MR-2964297

[20] Osękowski, A.: A sharp maximal inequality for continuous martingales and their differential subordinates. *Czechoslovak Math. J.* **63**, (2013), 1001–1018. MR-3165511

[21] Osękowski, A.: Survey article: Bellman function method and sharp inequalities for martingales. *Rocky Mountain J. Math.* **43**, (2013), 1759–1823. MR-3178444

[22] 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. MR-2115376

[23] Wang, G.: Sharp inequalities for the conditional square function of a martingale. *Ann. Probab.* **19**, (1991), 1679–1688. MR-1127721

[24] Wang, G.: Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities. *Ann. Probab.* **23**, (1995), 522–551. MR-1334160

Acknowledgments. The authors would like to thank an anonymous Referee for the careful reading of the first version of the paper and several helpful suggestions, which improved the presentation. A. Osękowski was supported by Narodowe Centrum Nauki (Poland), grant DEC-2014/14/E/ST1/00532, Y. Zuo was supported by China Scholarship Council.