Osękowski, A.
Osaka J. Math.
52 (2015), 1125–1142

SHARP MAXIMAL ESTIMATES FOR BMO MARTINGALES

ADAM OSĘKOWSKI

(Received May 9, 2014, revised September 29, 2014)

Abstract

We introduce a method which can be used to study maximal inequalities for martingales of bounded mean oscillation. As an application, we establish sharp Φ-inequalities and tail inequalities for the one-sided maximal function of a BMO martingale. The results can be regarded as BMO counterparts of the classical maximal estimates of Doob.

1. Introduction

Martingales of bounded mean oscillation form an important subclass of uniformly integrable martingales, which plays a role in the study of H_p spaces, for instance via Fefferman's duality theorem, the inequalities of John and Nirenberg or the integrability properties of the corresponding exponential local martingales. Essentially, the theory is parallel to that of the BMO functions defined on $\mathbb{R}^n$, but the passage to the probabilistic setting reveals some additional structure and enables further applications, for example, in financial mathematics (see e.g. [1], [3] or [5]).

We start the exposition from recalling the necessary analytic background. A real-valued locally integrable function f defined on $\mathbb{R}^n$ is said to be in BMO, the space of functions of bounded mean oscillation, if

$$\sup_Q \frac{1}{|Q|}\int_Q \left| f(x) - \frac{1}{|Q|}\int_Q f(y)\,\mathrm{d}y \right| \mathrm{d}x < \infty,$$

where the supremum is taken over all cubes Q in $\mathbb{R}^n$. This definition is due to John and Nirenberg [8], who also established some fundamental estimates for such functions, and the celebrated result of Fefferman [4] identified the class BMO as the dual to the Hardy space H^1. In this paper we will study the probabilistic counterpart of this notion, introduced by Getoor and Sharpe [7]. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing family of sub-σ-fields of $\mathcal{F}$, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Let $X = (X_t)_{t\geq 0}$ be an adapted, continuous-path real valued martingale, satisfying $X_0 \equiv 0$. Following [7], for $1 \leq p < \infty$, the martingale X belongs

2000 Mathematics Subject Classification. Primary 60G42; Secondary 60G44.
Research partially supported by NCN grant DEC-2012/05/B/ST1/00412.

to BMO_p if it is uniformly integrable and

$$\|X\|_{\mathrm{BMO}_p} = \sup_\sigma \|\mathbb{E}[|X_\infty - X_\sigma|^p \mid \mathcal{F}_\sigma]^{1/p}\|_\infty < \infty,$$

where the supremum is taken over all stopping times σ. It turns out that all the norms $\|\cdot\|_{\mathrm{BMO}_p}$ are comparable and hence all the classes BMO_p coincide. Thus we are allowed to skip the lower index and just write BMO; furthermore, it will be convenient for us to work with the norm $\|\cdot\|_{\mathrm{BMO}_2}$, and will use the shortened notation $\|\cdot\|_{\mathrm{BMO}}$ for it.

The BMO martingales have very strong integrability properties (for an overview, see e.g. the book by Kazamaki [9]). In particular, the inclusion $\mathrm{BMO} \subset H_p$ holds true for any $1 \le p < \infty$; in fact we have the exponential bound $\mathbb{E}\exp(c|X_\infty|) < \infty$ for some $c > 0$ depending on the BMO norm of X; see e.g. Getoor and Sharpe [7], Garsia [6] and P.A. Meyer [11]. The question about sharp versions of such estimates (in the analytic setting) has gathered recently some interest in the literature: see Korenovskiĭ [10], Slavin and Vasyunin [13], Vasyunin [14] and Vasyunin and Volberg [15]. The purpose of this paper is to study the problem of this type, but concerning $X^* = \sup_{t\ge 0} X_t$, the one-sided maximal function of X. We propose a novel method which can be used to establish general sharp estimates involving X and X^* in the BMO setting. The technique rests on finding a certain appropriate special function, having some convex-type and majorizing properties, and can be regarded as a version of a well-known Burkholder's method (for the description of the latter, see e.g. [2] or [12]). The technique will be applied to establish the following sharp Φ-estimate.

Theorem 1.1. *Suppose that Φ is a convex and increasing function on $[0,\infty)$ and X is a uniformly integrable martingale. Then*

$$\mathbb{E}\Phi(X^*) \le \int_0^\infty \Phi(t\|X\|_{\mathrm{BMO}})e^{-t}\,\mathrm{d}t. \tag{1.1}$$

The constant on the right is the best possible; *more precisely, there is a martingale X with $0 < \|X\|_{\mathrm{BMO}} < \infty$ for which both sides are equal.*

In particular, if we take $\Phi(t) = t^p$, $p \ge 1$, we obtain the sharp estimate

$$\|X^*\|_p \le (\Gamma(p+1))^{1/p}\|X\|_{\mathrm{BMO}},$$

which can be regarded as a BMO version of the Doob's maximal inequality.

Our next result concerns the following bound for the tail of X^*.

Theorem 1.2. *Suppose that X is a uniformly integrable martingale. Then for any $\lambda > 0$ we have*

$$\mathbb{P}(X^* \geq \lambda) \leq \begin{cases} 1 - \lambda/(2\|X\|_{\mathrm{BMO}}) & \text{if} \quad \lambda \leq \|X\|_{\mathrm{BMO}}, \\ \dfrac{1}{2}\exp(1 - \lambda\|X\|_{\mathrm{BMO}}^{-1}) & \text{if} \quad \lambda > \|X\|_{\mathrm{BMO}}. \end{cases} \tag{1.2}$$

The bound is the best possible: for each $\lambda > 0$ there is a martingale X such that $0 < \|X\|_{\mathrm{BMO}} < \infty$, for which both sides are equal.

The above result leads to the following sharp weak-type (p, p) estimate. For $p \geq 1$, let $\|X^*\|_{p,\infty} = \sup_{\lambda>0}[\lambda^p \mathbb{P}(X^* \geq \lambda)]^{1/p}$ denote the weak p-th norm of X^*. Multiplying both sides of (1.2) by λ^p and optimizing over λ, we get

Corollary 1.3. *For any $1 \leq p < \infty$ we have*

$$\|X^*\|_{p,\infty} \leq 2^{-1/p} p \exp(p^{-1} - 1)\|X\|_{\mathrm{BMO}} \tag{1.3}$$

and the constant $2^{-1/p} p \exp(p^{-1} - 1)$ is the best possible for each p.

The paper is organized as follows. The next section is devoted to the description of the method which will be used in the proofs of Theorems 1.1 and 1.2. These two theorems are established in Sections 3 and 4.

2. A method of proof

This section contains the detailed description of the methodology which will be used to establish the results aforementioned in the introduction. In general, all the problems studied in this paper can be stated as follows. Assume that c is a fixed real number, let $V : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ be a given Borel function and suppose we are interested in proving the maximal estimate

$$\mathbb{E}V(X_\infty, X^*) \leq c \tag{2.1}$$

for all uniformly integrable martingales X satisfying $\|X\|_{\mathrm{BMO}} \leq 1$. For example, the choice

$$V(x, z) = \Phi(z) \quad \text{and} \quad c = \int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$$

corresponds to the inequality (1.1). To handle (2.1), it is convenient to interpret a martingale X with $\|X\|_{\mathrm{BMO}} \leq 1$ as an appropriate two-dimensional martingale. To be more precise, consider the set

$$\mathcal{D} = \{(x, y) \in \mathbb{R} \times [0, \infty) : 0 \le y - x^2 \le 1\} \tag{2.2}$$

and its interior

$$\mathcal{D}^o = \{(x, y) \in \mathbb{R} \times [0, \infty) : 0 < y - x^2 < 1\}.$$

Next, introduce the martingale Y by the formula $Y_t = \mathbb{E}(X_\infty^2 \mid \mathcal{F}_t)$, $t \ge 0$. Then, by conditional Jensen's inequality, we have $Y_t \ge X_t^2$ almost surely; in addition,

$$Y_t - X_t^2 = \mathbb{E}[|X_\infty - X_t|^2 \mid \mathcal{F}_t] \le \|X\|_{\mathrm{BMO}}^2 \le 1.$$

Thus, the pair (X, Y) is a two-dimensional martingale with uniformly integrable coordinates, taking values in $\mathcal{D}$ and terminating at the lower boundary of $\mathcal{D}$: $Y_\infty = X_\infty^2$ with probability 1. In fact, this correspondence can be reversed: for any such pair (X, Y), we have $Y_t = \mathbb{E}(Y_\infty \mid \mathcal{F}_t) = \mathbb{E}(X_\infty^2 \mid \mathcal{F}_t)$ for all t and hence the martingale X satisfies $\|X\|_{\mathrm{BMO}} \le 1$.

The underlying concept of our approach is to find a special function $U : \mathcal{D} \times [0, \infty) \to \mathbb{R}$ which majorizes V at the lower boundary of $\mathcal{D}$ (that is, $V(x, z) \le U(x, x^2, z)$ for all x, z) and such that for all X,

$$\mathbb{E}U(X_\infty, Y_\infty, X^*) \le c. \tag{2.3}$$

Obviously, the existence of such a function immediately yields the desired estimate (2.1). To guarantee (2.3), we will impose some conditions on U which will imply that the process $(U(X_t, Y_t, X_t^*))_{t\ge0}$ is a supermartingale such that $U(X_0, Y_0, X_0^*) \le c$ almost surely (here $X_t^* = \sup_{0\le s\le t} X_s$ is the truncated one-sided maximal function of X). We turn to the precise formulation. Introduce the class $\mathcal{U}(V)$, which consists of all functions $U : \mathcal{D} \times [0, \infty) \to \mathbb{R}$, satisfying the following conditions:

(2.4) $\quad U(0, y, 0) \le c \quad$ for all $\quad y \in [0, 1]$,

(2.5) $\quad U(x, x^2, z) \ge V(x, z) \quad$ for all $\quad x \in \mathbb{R}, \quad z \ge 0$,

(2.6) $\quad U$ is continuous on $\mathcal{D} \times [0, \infty)$ and of class C^2 on $\mathcal{D}^o \times (0, \infty)$,

(2.7) $\quad U_z(x, y, x) \le 0 \quad$ for all $\quad x > 0 \quad$ and $\quad y \in (x^2, x^2 + 1)$,

and the further property that for all $(x, y, z) \in \mathcal{D}^o \times (0, \infty)$,

$$\text{the matrix } \begin{bmatrix} U_{xx}(x, y, z) & U_{xy}(x, y, z) \\ U_{xy}(x, y, z) & U_{yy}(x, y, z) \end{bmatrix} \text{ is nonpositive-definite.} \tag{2.8}$$

The following statement is the key to handle the supermartingale property of $(U(X_t, Y_t, X_t^*))_{t\ge0}$.

Lemma 2.1. *Suppose that a function* $U\colon \mathcal{D}\times[0,\infty)\to\mathbb{R}$ *satisfies* (2.6), (2.7) *and* (2.8). *Let* X *be a uniformly martingale with* $\|X\|_{\mathrm{BMO}}<1$ *and let* τ, σ *be two stopping times such that* $\sigma\le\tau$ *almost surely. Then there is a sequence* $(\tau_n)_{n\ge 0}$ *of stopping times which starts from* σ *and increases to* τ *almost surely, such that*

$$\mathbb{E}[U(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}) - U(X_\sigma, Y_\sigma, X^*_\sigma)] \le 0, \quad n\ge 0 \tag{2.9}$$

(*here and in what follows,* $Y_t=\mathbb{E}(X^2_\infty \mid \mathcal{F}_t)$, $t\ge 0$).

Proof. Introduce the process $Z=(X,Y,X^*)$. Observe that we have the strict inequality $Y_t-X_t^2\le\|X\|^2_{\mathrm{BMO}}<1$ and that the process (X,Y) terminates at the lower boundary of $\mathcal{D}$. Thus, by (2.6), we may apply Itô's formula to obtain

$$U(Z_{t\vee\sigma})-U(Z_\sigma)=I_1+\frac{1}{2}I_2+I_3, \tag{2.10}$$

where $I_1=I_2=I_3=0$ on $\{\sigma=\infty\}$ and, on the compliment of this set,

$$\begin{aligned}
I_1&=\int_\sigma^{t\vee\sigma} U_x(Z_s)\,\mathrm{d}X_s+\int_\sigma^{t\vee\sigma} U_y(Z_s)\,\mathrm{d}Y_s,\\
I_2&=\int_\sigma^{t\vee\sigma} U_{xx}(Z_s)\,\mathrm{d}[X,X]_s+2\int_\sigma^{t\vee\sigma} U_{xy}(Z_s)\,\mathrm{d}[X,Y]_s+\int_\sigma^{t\vee\sigma} U_{yy}(Z_s)\,\mathrm{d}[Y,Y]_s,\\
I_3&=\int_\sigma^{t\vee\sigma} U_z(Z_s)\,\mathrm{d}X^*_s.
\end{aligned}$$

First note that $I_3\le 0$: the measure $\mathrm{d}X^*$ is concentrated on $\{s\colon X_s=X_s^*\}$, and on this set we have $U_z(Z_s)\le 0$, in view of (2.7). Next, we will prove that $I_2\le 0$, by showing that the process

$$\left(\int_0^t U_{xx}(Z_s)\,\mathrm{d}[X,X]_s+2\int_0^t U_{xy}(Z_s)\,\mathrm{d}[X,Y]_s+\int_0^t U_{yy}(Z_s)\,\mathrm{d}[Y,Y]_s\right)_{t\ge 0} \tag{2.11}$$

is nondecreasing. To do this, note that (2.8) implies

$$U_{xx}(Z_s)h^2+2U_{xy}(Z_s)hk+U_{yy}(Z_s)k^2\le 0 \tag{2.12}$$

for any $h,k\in\mathbb{R}$. Fix positive numbers s, u such that $s<u$. For any j, let $(t_n^{(j)})_{n=0}^{k_j}$ be a nondecreasing sequence with $t_0^{(j)}=s$ and $t_{k_j}^{(j)}=u$, such that $\lim_{j\to\infty}\sup_{0<n\le k_j}|t_n^{(j)}-t_{n-1}^{(j)}|=0$. Apply (2.12) to $h=X_{t_n^{(j)}}-X_{t_{n-1}^{(j)}}$, $k=Y_{t_n^{(j)}}-Y_{t_{n-1}^{(j)}}$, $n=1,2,\ldots,k_j$, sum the obtained inequalities and let $j\to\infty$. As the result, we get that

$$U_{xx}(Z_s)[X,X]_s^u+2U_{xy}(Z_s)[X,Y]_s^u+U_{yy}(Z_s)[Y,Y]_s^u\le 0,$$

where we have used the notation $[X,Y]_s^u = [X,Y]_u - [X,Y]_s$. This implies the monotonicity property of the process (2.11), by a simple approximation of the integrals by discrete sums, and hence $I_2 \leq 0$. Next, by the properties of stochastic integrals, the process

$$\left(\int_0^t U_x(Z_s)\,\mathrm{d}X_s + \int_0^t U_y(Z_s)\,\mathrm{d}Y_s\right)_{t\geq 0}$$

is a local martingale. Let $(\eta_n)_{n\geq 1}$ be the corresponding localizing sequence and define

$$\tau_n = (\tau \wedge \eta_n \wedge \inf\{t : |X_t| \geq n\} \wedge n) \vee \sigma, \quad n \geq 0.$$

Then $(\tau_n)_{n\geq 0}$ is a nondecreasing sequence of finite stopping times which satisfies $\tau_0 = \sigma$ and which converges almost surely to τ. Furthermore, by the martingale property,

$$\mathbb{E}\left[\int_\sigma^{\tau_n} U_x(Z_s)\,\mathrm{d}X_s + \int_\sigma^{\tau_n} U_y(Z_s)\,\mathrm{d}Y_s \,\middle|\, \mathcal{F}_\sigma\right] 1_{\{\sigma<\infty\}} = 0.$$

Thus, plugging τ_n in the place of t in (2.10) and integrating both sides gives

$$\mathbb{E}[U(Z_{\tau_n}) - U(Z_\sigma)] \leq 0,$$

which is precisely the claim. □

The above lemma leads to the following solution of the problem formulated at the beginning of this section. Suppose that $U \in \mathcal{U}(V)$ and fix a martingale X satisfying $\|X\|_{\mathrm{BMO}} \leq 1$ and $X_0 \equiv 0$. Take a number $\kappa \in (0, 1)$ and consider a martingale κX, which has the BMO norm strictly smaller than 1. An application of Lemma 2.1 with $\tau \equiv \infty$ and $\sigma \equiv 0$ yields

$$\mathbb{E}U(\kappa X_{\tau_n}, \kappa^2 Y_{\tau_n}, \kappa X_{\tau_n}^*) \leq \mathbb{E}U(\kappa X_0, \kappa^2 Y_0, \kappa X_0^*) = U(0, \kappa^2 \mathbb{E}X_\infty^2, 0),$$

for an appropriate sequence $(\tau_n)_{n\geq 1}$ of stopping times. By (2.4), the right-hand side can be bounded from above by c. Now if we can only justify the passage with n to infinity and $\kappa \to 1$ (for example, if U is nonnegative, or the random variable $\sup_n \sup_{\kappa\in(0,1)} |U(\kappa X_{\tau_n}, \kappa^2 Y_{\tau_n}, \kappa X_{\tau_n}^*)|$ is integrable), we get

$$\mathbb{E}U(X_\infty, Y_\infty, X_\infty^*) \leq c.$$

However, $Y_\infty = X_\infty^2$ almost surely; thus, by (2.5), we obtain the desired bound (2.1) and we are done.

3. Proof of Theorem 1.1

3.1. Proof of (1.1). By homogeneity, it suffices to prove the estimate under the additional assumption $\|X\|_{\mathrm{BMO}} = 1$ (indeed, having this done, we recover (1.1) in full generality by considering the martingale $X/\|X\|_{\mathrm{BMO}}$ and the function $t \mapsto \Phi(t\|X\|_{\mathrm{BMO}})$). Furthermore, by a standard approximation, we may and do assume that Φ is of class C^2. As we have already observed above, we need to take

$$V(x, z) = \Phi(z) \quad \text{and} \quad c = \int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t.$$

The corresponding special function $U\colon \mathcal{D} \times [0, \infty) \to \mathbb{R}$ is defined by the formula

$$U(x, y, z) = \Phi(z) + \frac{y - x^2 + (z - x - 1)^2}{2} \int_z^\infty \Phi'(t)e^{z-t}\,\mathrm{d}t.$$

(Some steps which lead to the discovery of U are sketched in Subsection 3.3 below). Let us verify the conditions (2.4)–(2.8). The first property follows easily from the integration by parts and the next two are evident. To check (2.7), we derive that for $x > 0$,

$$U_z(x, y, x) = \frac{y - x^2 - 1}{2}\left[\int_z^\infty \Phi'(t)e^{z-t}\,\mathrm{d}t - \Phi'(z)\right].$$

It suffices to note that $(y-x^2-1)/2 \le 0$, by the definition of $\mathcal{D}$, and that the expression in the square brackets is nonnegative: indeed, since Φ' is nondecreasing, we have

$$\int_z^\infty \Phi'(t)e^{z-t}\,\mathrm{d}t \ge \int_z^\infty \Phi'(z)e^{z-t}\,\mathrm{d}t = \Phi'(z).$$

Finally, the condition (2.8) is trivial, since all the entries of the corresponding matrix vanish. Consequently, U belongs to the class $\mathcal{U}(V)$; in addition, U is nonnegative, so the reasoning presented at the end of the previous section yields the claim.

3.2. Sharpness. Now we exhibit an appropriate example for which both sides of (1.1) are equal. Suppose that $B = (B_t)_{t\ge 0}$ is a standard, one-dimensional Brownian motion starting from the origin and let

$$\tau = \inf\{t\colon B_t^* - B_t = 1\}$$

be the first time B experiences the drop of size 1. Define $X = (B_{\tau\wedge t})_{t\ge 0}$. We have $B^*_{\tau\wedge t} - B_{\tau\wedge t} \le 1$ and by Itô's formula, the process $((B_t^*)^2 - 2B_tB_t^*)_{t\ge 0}$ is a martingale, so

$$\mathbb{E}B^2_{\tau\wedge t} = \mathbb{E}(B^*_{\tau\wedge t} - B_{\tau\wedge t})^2 - \mathbb{E}[(B^*_{\tau\wedge t})^2 - 2B_{\tau\wedge t}B^*_{\tau\wedge t}] \le 1.$$

In consequence, X is a uniformly integrable, L^2-bounded martingale. Furthermore, for any stopping time σ,

$$\begin{aligned} Y_\sigma &= \mathbb{E}(X^2_\infty \mid \mathcal{F}_\sigma) = \mathbb{E}[(B^*_\tau - B_\tau)^2 \mid \mathcal{F}_\sigma] - \mathbb{E}[(B^*_\tau)^2 - 2B_\tau B^*_\tau \mid \mathcal{F}_\sigma] \\ &= 1 + X^2_\sigma - (X^*_\sigma - X_\sigma)^2, \end{aligned} \tag{3.1}$$

which implies that $\|X\|_{\mathrm{BMO}} \le 1$. Next, observe that for any $\lambda > 0$ the process $((B_t^* - B_t + \lambda^{-1})\exp(-\lambda B_t^*))_{t\ge 0}$ is a martingale: this follows immediately from Itô's formula. Therefore, we have

$$\mathbb{E}[(X_t^* - X_t + \lambda^{-1})\exp(-\lambda X_t^*)] = \lambda^{-1},$$

and since $0 \le X_t^* - X_t \le 1$, we may let $t \to \infty$ and use Lebesgue's dominated convergence theorem to get $\mathbb{E}\exp(-\lambda X_\infty^*) = (\lambda + 1)^{-1}$. Consequently,

$$X_\infty^* \quad \text{follows the exponential law of parameter} \quad 1 \tag{3.2}$$

and hence

$$\mathbb{E}\Phi(X^*) = \int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t,$$

so the inequality (1.1) is sharp.

3.3. On the search of the suitable majorant. Let us now describe the informal reasoning which leads to the special function used above (and the optimal constant $\int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$). This function needs to satisfy the conditions (2.4)–(2.8); four of these conditions are inequalities. Since U is supposed to yield sharp results, it seems reasonable to expect that it will actually produce equalities in (some of) these conditions. Thus, at least at the very beginning, let us try to find U for which (2.5) and (2.7) hold with equality sign, and such that

$$\det\begin{bmatrix} U_{xx}(x,y,z) & U_{xy}(x,y,z) \\ U_{xy}(x,y,z) & U_{yy}(x,y,z) \end{bmatrix} = 0$$

for all $(x,y,z) \in \mathcal{D}^o \times (0,\infty)$. The latter condition means, roughly speaking, that if we fix $z > 0$, then for any $(x,y) \in \mathcal{D}$, $x \le z$, there is a line segment contained in $\mathcal{D}$, passing through (x,y), along which $U(\cdot,\cdot,z)$ is linear. This further suggests (compare this to the analogous situation occurring in the papers [13], [14] and [15]) that the whole set $\{(x,y) \in \mathcal{D} : x \le z\}$ can be "foliated", i.e., split into the union of line segments along which $U(\cdot,\cdot,z)$ is linear. It is not difficult to guess the foliation, at least for a part of the set (here a look at the papers [13], [14] and [15] is really helpful, as a similar splitting appears there). Namely, fix an arbitrary $x \le z$ and consider the line segment I_x passing through the points $(x-1,(x-1)^2)$ and (x, x^2+1). It is easy to check that this line segment is tangent to the upper boundary $\{(x,y) : y = x^2+1\}$ of the set $\mathcal{D}$ and the collection $\{I_x : x \le z\}$ splits the set $\{(x,y) \in \mathcal{D} : y \ge 2zx + 1 - z^2\}$. So, let us *assume* that U is linear along each I_x. Then for any $\lambda \in [0,1]$, and any $x \le z$,

$$\begin{aligned} &U(\lambda(x-1) + (1-\lambda)x, \lambda(x-1)^2 + (1-\lambda)(x^2+1), z) \\ &\quad = \lambda U(x-1,(x-1)^2,z) + (1-\lambda)U(x,x^2+1,z). \end{aligned} \tag{3.3}$$

But we have assumed above that both sides of (2.5) are equal. This implies $U(x-1,(x-1)^2,z)=\Phi(z)$ and hence, if we substitute $\Psi_z(x)=U(x,x^2+1,z)$ and carry out some straightforward computations, we obtain

$$U(x,y,z)=\sqrt{x^2-y+1}\Phi(z)+(1-\sqrt{x^2-y+1})\Psi_z(x+\sqrt{x^2-y+1}) \tag{3.4}$$

for any $(x,y)\in\mathcal{D}$, $y\geq 2zx+1-z^2$. To find Ψ_z, let us go back to the equation (3.3). A nice feature of the foliation we chose is that any segment I_x, $x<z$, can be lengthened a little bit "to the right" and it is still contained in $\mathcal{D}$. Thus, looking at the property (2.6), it is natural to suspect that for any $x<z$, (3.3) can be extended to some negative values of λ (in the sense that the difference of the left- and the right-hand sides should be of order $o(\lambda)$ as $\lambda\to 0$). So, take $\lambda<0$, write this difference, divide by λ and let $\lambda\to 0$. The result must be zero; using the formula (3.4), we obtain the differential equation $\Psi_z'(x)=\Psi_z(x)-\Phi(z)$, and hence

$$\Psi_z(x)=K(z)e^x+\Phi(z),$$

for some function K to be found. Now it is high time to apply (2.7) (recall that we have assumed that equality holds here). Differentiating (3.4) with respect to the variable z at the point (z,z^2+1,z), we obtain $\Phi'(z)+K'(z)e^z=0$. Hence $K(z)=\int_z^\infty \Phi'(t)e^{-t}\,\mathrm{d}t+\alpha$ for some constant α and, coming back to (3.4), we see that

$$U(x,y,z)=\Phi(z)+(1-\sqrt{x^2-y+1})\left(\int_z^\infty \Phi'(t)e^{-t}\,\mathrm{d}t+\alpha\right)e^{x+\sqrt{x^2-y+1}}.$$

With a lack of a better idea, let us take $\alpha=0$ in the above formula. Then

$$U(x,y,z)=\Phi(z)+(1-\sqrt{x^2-y+1})\int_z^\infty \Phi'(t)e^{x+\sqrt{x^2-y+1}-t}\,\mathrm{d}t \tag{3.5}$$

for $(x,y)\in\mathcal{D}$, $y\geq 2zx+1-z^2$. In particular,

$$U(0,1,0)=\Phi(z)+\int_0^\infty \Phi'(t)e^{-t}\,\mathrm{d}t=\int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t,$$

which gives us the hint about the best constant. What about the remaining part of the domain? One can, of course, proceed as above and try to find an appropriate foliation. This can be done, but the expression we get is different from that above and the function is not of class C^2. Thus, in order to use of Lemma 2.1, one has to apply some mollification to ensure the necessary regularity, and this results in a significant complication of technicalities (we will encounter some of these below, in the proof of the tail bound (1.2)). Fortunately, there is a different solution to the above problem. The key fact is that in general, the special function needed to establish a given inequality is not unique, and hence we have some freedom with choosing one. Recall that we have *imposed* the equalities in (2.5) and (2.7), while we need only inequalities. This leads to the following natural idea: let us extend U to the whole domain with the use of the formula (3.5) and verify whether all the conditions are met; if so, we will be done. Unfortunately, the condition (2.7) does not hold true and hence we need some modification of U. How should we proceed?

Some indications can be found in Subsection 3.2 above (a similar phenomenon occurs in the analytic Bellman setting: the knowledge about the (candidates for) the extremals can be very helpful in the search for the special function). Again, we stress that the arguments presented here are informal; they only serve as an intuition in the construction of U. The triple (X, Y, X^*) considered in Subsection 3.2 evolves along the set $\{(x, y, z)\colon y = 2zx + 1 - z^2\}$: see (3.1). In addition, it follows from the above construction that $U(X, Y, X^*)$ is a martingale (roughly speaking, all the inequalities which imply the supermartingale property hold with equality sign) starting from $\int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$. Thus, we have the following important observation. Suppose that $\tilde{U}$ is another special function which leads to the Φ-estimate with the constant $\int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$; hence, in particular, $\tilde{U}(0, 1, 0) \le \int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$. Then $\tilde{U}$ should coincide with U on the set $\{(x, y, z)\colon y = 2zx + 1 - z^2\}$. Otherwise, the martingale property of $\tilde{U}(X, Y, X^*)$ would not hold (only the *supermartingale* property would be valid) and this would violate the optimality of the constant. Indeed, an application of the method from Section 2 would lead to the *strict* inequality $\mathbb{E}\Phi(X^*) < \int_0^\infty \Phi(t)e^{-t}\,\mathrm{d}t$, a contradiction. Now, if we take a look at the above U, we see that if $y = 2zx + 1 - z^2$, then

$$U(x, y, z) = \Phi(z) + (1 - z + x)\int_z^\infty \Phi'(t)e^{z-t}\,\mathrm{d}t.$$

So, a natural idea is to consider U given by the above formula for *all* (x, y, z). Unfortunately, this still does not work: this time the condition (2.5) is not valid (when $z > 1 + x$ and $y = x^2$, the above expression is smaller than $\Phi(z)$). So, let us try to replace the term $1 - z + x$ by some other expression, possibly involving y too. This unknown term must be nonnegative if $y = x^2$ (because of (2.5)), and equal to 0 for $z = x + 1$ (since it coincides with $1 - z + x$ on the set $y = 2zx + 1 - z^2$). This strongly suggests to consider the expression $A \cdot (y - x^2) + B \cdot (z - x - 1)^2$ for some positive A, B. Then we must have

$$A(y - x^2) + B(z - x - 1)^2 = 1 - z + x$$

for $y = 2zx + 1 - z^2$; one easily checks that this is satisfied if and only if $A = B = 1/2$. Then we get exactly the function studied in Subsection 3.1.

4. Proof of Theorem 1.2

4.1. Proof of (1.2). Here the reasoning will be slightly more complicated. As previously, it suffices to establish the estimate only for X which have BMO norm smaller than 1, due to homogeneity reasons. Take $V(x, z) = 1_{\{z\geq\lambda\}}$ and

$$c = \begin{cases} 1 - \lambda/2 & \text{for} \quad \lambda \leq 1, \\ \exp(1-\lambda)/2 & \text{for} \quad \lambda > 1. \end{cases}$$

To define the corresponding special function, consider the following sets:

$$\begin{aligned} D_1 &= \{(x, y) \in \mathcal{D} \colon x \geq \lambda\}, \\ D_2 &= \{(x, y) \in \mathcal{D} \colon \lambda - 1 < x < \lambda,\ y > 2\lambda - 2x - \lambda^2 + 2\lambda x\}, \\ D_3 &= \{(x, y) \in \mathcal{D} \colon y \leq 2\lambda - 2x - \lambda^2 + 2\lambda x\}, \\ D_4 &= \{(x, y) \in \mathcal{D} \colon x < \lambda - 1,\ y > 2\lambda - 2x - \lambda^2 + 2\lambda x\}, \end{aligned}$$

see Fig. 1 below.

The special function $U = U_\lambda \colon \mathcal{D} \times [0, \infty) \to \mathbb{R}$ is given by

$$U(x, y, z) = \begin{cases} 1 & \text{if} \quad z \geq \lambda \text{ or } (x, y) \in D_1, \\ 1 - (\lambda - x)/2 & \text{if} \quad z < \lambda,\ (x, y) \in D_2, \\ (y - x^2)/(y - 2\lambda x + \lambda^2) & \text{if} \quad z < \lambda,\ (x, y) \in D_3, \\ \dfrac{1 - \sqrt{1 - y + x^2}}{2} \exp(x + \sqrt{1 - y + x^2} + 1 - \lambda) & \text{if} \quad z < \lambda,\ (x, y) \in D_4. \end{cases}$$

This function is constructed with the use of a similar reasoning to that in Subsection 3.3. Consult also the papers [14], [15] and the Remark 4.2 below.

The problem with U is that it is not of class C^2, so to apply the technique from Section 2, we need to use appropriate smoothing arguments, which results in some unpleasant technicalities. To overcome this problem, we will present a slightly different approach, which rests on a direct use of Lemma 2.1 and exploits three simpler special functions. Namely, introduce $U_0, U_1, U_2 \colon \mathcal{D} \times [0, \infty) \to \mathbb{R}$ by

$$U_0(x, y, z) = 1 - \frac{\lambda - x}{2}, \quad U_1(x, y, z) = \frac{y - x^2}{y - 2\lambda x + \lambda^2}$$

and

$$U_2(x, y, z) = \frac{1 - \sqrt{1 - y + x^2}}{2} \exp(x + \sqrt{1 - y + x^2} + 1 - \lambda).$$

Observe that all these functions appear as "building blocks" of the above U. These functions satisfy (2.6); moreover, none of these functions depend on the variable z and

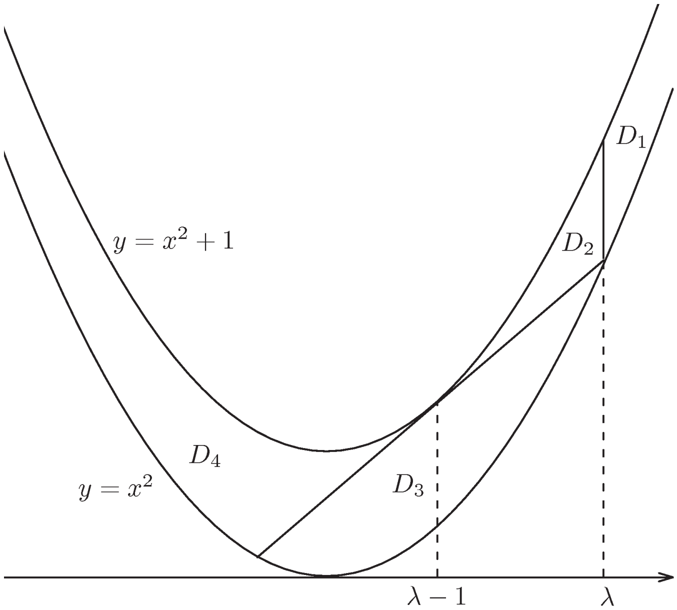

Fig. 1. The regions $D_1 - D_4$ in the case $\lambda > 1$.

thus (2.7) holds true for all of them. Finally, U_0, U_1 and U_2 satisfy (2.8). This is trivial for U_0, while for the remaining two functions, we calculate a little bit to get that

$$\begin{bmatrix} U_{1xx} & U_{1xy} \\ U_{1xy} & U_{1yy} \end{bmatrix} = \begin{bmatrix} -b(x,y)(y-\lambda^2)^2 & b(x,y)(x-\lambda)(y-\lambda^2) \\ b(x,y)(x-\lambda)(y-\lambda^2) & -b(x,y)(\lambda-x)^2 \end{bmatrix}$$

and

$$\begin{bmatrix} U_{2xx} & U_{2xy} \\ U_{2xy} & U_{2yy} \end{bmatrix} = \begin{bmatrix} -4c(x,y)(x+\sqrt{1-y+x^2})^2 & 2c(x,y)(x+\sqrt{1-y+x^2}) \\ 2c(x,y)(x+\sqrt{1-y+x^2}) & -c(x,y) \end{bmatrix},$$

where $b(x,y) = 2(y - 2\lambda x + \lambda^2)^{-3} > 0$ and

$$c(x,y) = (8\sqrt{1-y+x^2})^{-1} \exp(x + \sqrt{1-y+x^2} + 1 - \lambda) > 0.$$

Clearly, both matrices are nonpositive-definite and hence (2.8) holds true. We will also require the following properties of U_1 and U_2. First, observe that U_1 is bounded: in fact, we have

$$0 \le \frac{y - x^2}{y - 2\lambda x + \lambda^2} \le 1 \quad \text{for} \quad (x,y) \in \mathcal{D}. \tag{4.1}$$

Next, we have that

$$\text{(4.2)} \qquad \text{the functions} \quad U_1(0,\ \cdot\ ,0), U_2(0,\ \cdot\ ,0) \quad \text{are nondecreasing on} \quad [0,1],$$

which can be easily verified by differentiation.

Now we split the reasoning into two parts, corresponding to $\lambda \leq 1$ and $\lambda > 1$.

CASE $\lambda \leq 1$. Then the process (X, Y) starts from the set $D_2 \cup D_3$; suppose first that $(X_0, Y_0) \in D_2$. Introduce the stopping times

$$\tau = \inf\{t : X_t \geq \lambda\}$$

and

$$\sigma = \inf\{t : X_t \geq \text{ or } (X_t, Y_t) \in D_3\}.$$

Of course, we have $\sigma \leq \tau$ almost surely. Furthermore, U_0 and U_1 coincide at $\partial D_2 \cap \partial D_3$, the common boundary of D_2 and D_3. Therefore, applying Lemma 2.1 and using (4.1) yields

$$\mathbb{E}U_0(X_\sigma, Y_\sigma, X^*_\sigma) = \mathbb{E}U_1(X_\sigma, Y_\sigma, X^*_\sigma) \geq \mathbb{E}U_1(X_{\tau_n}, Y_{\tau_n}, X^*_{\tau_n}),$$

for an appropriate sequence $(\tau_n)_{n\geq 0}$ of stopping times. Letting $n \to \infty$ gives $\mathbb{E}U_1(X_\tau, Y_\tau, X^*_\tau) \leq \mathbb{E}U_0(X_\sigma, Y_\sigma, X^*_\sigma)$, by the use of (4.1) and Lebesgue's dominated convergence theorem. Next, applying Lemma 2.1 again, this time to the function U_0 and the stopping times 0 and σ, we obtain

$$\mathbb{E}U_0(X_{\sigma_n}, Y_{\sigma_n}, X^*_{\sigma_n}) \leq \mathbb{E}U_0(X_0, Y_0, X^*_0) = 1 - \lambda/2,$$

for some sequence $(\sigma_n)_{n\geq 0}$ of stopping times increasing to σ. However, $(X_{\sigma_n}, Y_{\sigma_n})$ belongs to the closure of D_2, so $X_{\sigma_n} \geq \lambda - 1$ and hence the random variables $U_0(X_{\sigma_n}, Y_{\sigma_n}, X^*_{\sigma_n})$ are nonnegative. Now, applying Fatou's lemma, we obtain that $\mathbb{E}U_0(X_\sigma, Y_\sigma, X^*_\sigma) \leq 1 - \lambda/2$ and combining this with the previous estimates we get

$$\text{(4.3)} \qquad \mathbb{E}U_1(X_\tau, Y_\tau, X^*_\tau) \leq 1 - \lambda/2.$$

We have obtained this bound under the assumption $(X_0, Y_0) \in D_2$; but this is also true if (X, Y) starts from D_3. Indeed, we apply Lemma 2.1 to the function U_1 and the stopping times 0 and τ, use Fatou's lemma and get

$$\text{(4.4)} \qquad \mathbb{E}U_1(X_\tau, Y_\tau, X^*_\tau) \leq \mathbb{E}U_1(X_0, Y_0, X^*_0) = U_1(0, Y_0, 0).$$

However, it is easy to check that $U_1(0, y, 0) \leq 1 - \lambda/2$ if $(0, y) \in D_3$, so the inequality (4.3) holds true.

We turn to the final step. Observe that for any fixed $\varepsilon > 0$, we have $\tau < \infty$ and $U_1(X_\tau, Y_\tau, X^*_\tau) = 1$ on the set $\{X^* \geq \lambda + \varepsilon\}$. Since U_1 is nonnegative, we get

$$\mathbb{P}(X^* \geq \lambda + \varepsilon) \leq \mathbb{E}U_1(X_\tau, Y_\tau, X^*_\tau) \leq 1 - \lambda/2.$$

Substituting $\lambda := \lambda + \varepsilon$, we see that for any $\varepsilon \in (0, \lambda)$,

$$\mathbb{P}(X^* \geq \lambda) \leq 1 - \lambda/2 + \varepsilon/2,$$

and letting $\varepsilon \to 0$ yields (1.2).

CASE $\lambda > 1$. Here the reasoning is essentially the same (and rests on properties of U_1 and U_2), so we shall be brief. The process (X, Y) starts from $D_3 \cup D_4$; suppose first that $(X_0, Y_0) \in D_4$ and introduce the stopping times

$$\tau = \inf\{t \colon X_t \geq \lambda\} \quad \text{and} \quad \sigma = \inf\{t \colon (X_t, Y_t) \in D_3\}.$$

We have $\sigma \leq \tau$ and, arguing as previously, we obtain

$$\begin{aligned}\mathbb{E}U_1(X_\tau, Y_\tau, X_\tau^*) &\leq \mathbb{E}U_1(X_\sigma, Y_\sigma, X_\sigma^*) \\ &\leq \mathbb{E}U_2(X_\sigma, Y_\sigma, X_\sigma^*) \leq \mathbb{E}U_2(X_0, Y_0, X_0^*) \leq \exp(1-\lambda)/2.\end{aligned}$$

The same bound holds true if (X, Y) starts from D_3: then (4.4) is valid and hence, using (4.2) and the fact that U_1 and U_2 coincide at $\partial D_3 \cap \partial D_4$, we get

$$\mathbb{E}U_1(X_\tau, Y_\tau, X_\tau^*) \leq \mathbb{E}U_1(X_0, Y_0, X_0^*) \leq U_2(0, 1, 0) = \exp(1-\lambda)/2.$$

It remains to repeat the above argumentation to get

$$\mathbb{P}(X^* \geq \lambda + \varepsilon) \leq \mathbb{E}U_1(X_\tau, Y_\tau, X_\tau^*) \leq \exp(1-\lambda)/2, \quad \varepsilon > 0,$$

which yields (1.2) for $\lambda > 1$.

4.2. Sharpness. Let $a \geq 0$ be a fixed number and let B be a standard Brownian motion. Introduce the stopping times $\eta = \inf\{t \colon B_t^* - B_t \geq 1 \text{ or } B_t = a\}$ and

$$\tau = \begin{cases} \eta & \text{if} \quad B_\eta < a, \\ \inf\{t > \eta \colon B_t \in \{a-1, a+1\}\} & \text{if} \quad B_\eta = a. \end{cases}$$

Of course, $\eta \leq \tau$ almost surely. We have the following fact.

Lemma 4.1. *The martingale $X = (B_{\tau \wedge t})_{t \geq 0}$ is uniformly integrable and satisfies $\|X\|_{\mathrm{BMO}} \leq 1$.*

Proof. The uniform integrability can be easily shown using the martingale $(2B_t B_t^* - (B_t^*)^2)_{t \geq 0}$; see Subsection 3.2 above. To prove the bound for the BMO norm of X, note that for any stopping time σ we have

$$\mathbb{E}(X_\infty^2 \mid \mathcal{F}_\sigma) = B_\tau^2 1_{\{\tau \leq \sigma\}} + \mathbb{E}[B_\tau^2 1_{\{\tau > \sigma \geq \eta\}} \mid \mathcal{F}_\sigma] + \mathbb{E}[B_\tau^2 1_{\{\eta > \sigma\}} \mid \mathcal{F}_\sigma]. \tag{4.5}$$

Let us analyze each term on the right separately. We have $B_\tau^2 1_{\{\tau\le\sigma\}} = X_\sigma^2 1_{\{\tau\le\sigma\}}$ and

$$\begin{aligned}\mathbb{E}[B_\tau^2 1_{\{\tau>\sigma\ge\eta\}} \mid \mathcal{F}_\sigma] &= \mathbb{E}[(B_\tau - B_\eta)^2 1_{\{\tau>\sigma\ge\eta\}} + (2B_\tau B_\eta - B_\eta^2)1_{\{\tau>\sigma\ge\eta\}} \mid \mathcal{F}_\sigma]\\ &= 1_{\{\tau>\sigma\ge\eta\}} + (2B_\sigma B_\eta - B_\eta^2)1_{\{\tau>\sigma\ge\eta\}}\\ &\le 1_{\{\tau>\sigma\ge\eta\}} + B_\sigma^2 1_{\{\tau>\sigma\ge\eta\}},\end{aligned}$$

where in the second passage we have used the equality $|B_\tau - B_\eta| = 1$ valid on $\{\tau > \eta\}$, and Doob's optional sampling theorem. To deal with the third term on the right-hand side of (4.5), we make use of the martingale $(2B_{\eta\wedge t}B^*_{\eta\wedge t} - (B^*_{\eta\wedge t})^2)_{t\ge 0}$ and write

$$\begin{aligned}\mathbb{E}[B_\tau^2 1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma] &= \mathbb{E}[(B_\tau - B_\tau^*)^2 1_{\{\eta>\sigma\}} + (2B_\tau B_\tau^* - (B_\tau^*)^2)1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma]\\ &= \mathbb{E}[(B_\tau - B_\tau^*)^2 1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma] + (2B_\sigma B_\sigma^* - (B_\sigma^*)^2)1_{\{\eta>\sigma\}}\\ &\le \mathbb{E}[(B_\tau - B_\tau^*)^2 1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma] + B_\sigma^2 1_{\{\eta>\sigma\}}.\end{aligned}\tag{4.6}$$

However, using Doob's optional sampling theorem and the equality $B_\eta = B_\eta^* = a$, valid on $\{\tau > \eta\}$, we get

$$\begin{aligned}\mathbb{E}[(B_\tau - B_\tau^*)^2 1_{\{\tau>\eta\}} \mid \mathcal{F}_\eta] &= \mathbb{E}[(B_\tau - B_\eta)^2 + 2B_\tau B_\eta - B_\eta^2 \mid \mathcal{F}_\eta]1_{\{\tau>\eta\}}\\ &\quad + (-2B_\eta B_\eta^* + (B_\eta^*)^2)1_{\{\tau>\eta\}}\\ &= 1_{\{\tau>\eta\}}.\end{aligned}$$

Plugging this into (4.6), we get

$$\begin{aligned}&\mathbb{E}[B_\tau^2 1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma]\\ &\le \mathbb{E}[(B_\tau - B_\tau^*)^2(1_{\{\tau=\eta>\sigma\}} + 1_{\{\tau>\eta>\sigma\}}) \mid \mathcal{F}_\sigma] + B_\sigma^2 1_{\{\eta>\sigma\}}\\ &= \mathbb{E}[1_{\{\tau=\eta>\sigma\}} + \mathbb{E}[(B_\tau - B_\tau^*)^2 1_{\{\tau>\eta\}} \mid \mathcal{F}_\eta]1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma] + B_\sigma^2 1_{\{\eta>\sigma\}}\\ &= \mathbb{E}[1_{\{\tau=\eta>\sigma\}} + 1_{\{\tau>\eta\}}1_{\{\eta>\sigma\}} \mid \mathcal{F}_\sigma] + B_\sigma^2 1_{\{\eta>\sigma\}}\\ &= 1_{\{\eta>\sigma\}} + B_\sigma^2 1_{\{\eta>\sigma\}}.\end{aligned}$$

Plugging all the above estimates into (4.5) yields $\mathbb{E}(X_\infty^2 \mid \mathcal{F}_\sigma) \le 1 + X_\sigma^2$, which is precisely the claim. □

Now we are ready to prove the sharpness of (1.2). First consider the case $\lambda \ge 1$. Take the martingale from the above lemma, corresponding to $a = \lambda - 1$. This martingale, and the process exploited in Subsection 3.2, coincide on the interval $[0, \eta]$, so using (3.2), we get

$$\mathbb{P}(\tau > \eta) = \mathbb{P}(X_\infty^* \ge \lambda - 1) = e^{1-\lambda}.$$

Therefore,

$$\begin{aligned}\mathbb{P}(X_\infty^* \ge \lambda) &= \mathbb{P}(X_\infty^* \ge \lambda \mid X_\infty^* \ge \lambda - 1)\mathbb{P}(X_\infty^* \ge \lambda - 1)\\ &= \mathbb{P}(B_\tau = \lambda \mid B_\eta = \lambda - 1) \cdot e^{1-\lambda} = e^{1-\lambda}/2.\end{aligned}$$

Finally, we turn to the case $\lambda \in (0, 1)$. Consider the martingale $-X$, where X comes from the above lemma applied to $a = 1 - \lambda$. Let us compute the probability $\mathbb{P}((-X)^* < \lambda)$. A closer look gives that

$$\{(-X)^* < \lambda\} = \{B \text{ reaches } 2-\lambda \text{ before getting to } \lambda\}.$$

Indeed, the inclusion "$\subseteq$" is obvious, and to get the reverse one, it suffices to observe that $(-X_\infty)^* \geq \lambda$ on the set $\{\tau = \eta\}$, since $X_\infty^* = X_\infty - 1 \leq -\lambda$ there. Consequently,

$$\mathbb{P}((-X)^* \geq \lambda) = 1 - \mathbb{P}((-X)^* < \lambda) = 1 - \lambda/2.$$

This gives the optimality of the the bound (1.2) and completes the proof of Theorem 1.2.

REMARK 4.2. The approach described in Section 2 can be also used in the case when X starts from arbitrary real number x (i.e., not necessarily from 0). Denoting $\mathbb{E}X_\infty^2$ by y, one can show that for any $\lambda > 0$,

$$\mathbb{P}(X^* \geq \lambda) \leq U_{\lambda/\|X\|_{\mathrm{BMO}}}\left(\frac{x}{\|X\|_{\mathrm{BMO}}}, \frac{y}{\|X\|_{\mathrm{BMO}}^2}, \frac{\max\{x, 0\}}{\|X\|_{\mathrm{BMO}}}\right),$$

(this is clear: see the last paragraph of Section 2) and hence also

$$\mathbb{P}(X_\infty \geq \lambda) \leq U_{\lambda/\|X\|_{\mathrm{BMO}}}\left(\frac{x}{\|X\|_{\mathrm{BMO}}}, \frac{y}{\|X\|_{\mathrm{BMO}}^2}, \frac{\max\{x, 0\}}{\|X\|_{\mathrm{BMO}}}\right). \tag{4.7}$$

Here $U_{\lambda/\|X\|_{\mathrm{BMO}}}$ is the function of Subsection 4.1, corresponding to the parameter $\lambda/\|X\|_{\mathrm{BMO}}$. Moreover, it can be proved that the bound (4.7) is sharp, with the use of similar examples as above. This should be compared to the non-maximal tail estimates for BMO functions obtained in [14] and [15]. Vasyunin and Volberg found there, for each fixed $\lambda > 0$, the least functions $B_\lambda : \mathcal{D} \to \mathbb{R}$ with the following property. If $f : [0, 1] \to \mathbb{R}$ is a function satisfying $\int_0^1 f = x$, $\int_0^1 f^2 = y$ and $\|f\|_{\mathrm{BMO}_2} \leq 1$ (that is, $\int_I \left| f(t) - \int_I f(u)\,\mathrm{d}u\right|^2 \mathrm{d}t \leq 1$ for any interval $I \subseteq [0, 1]$), then

$$|\{x \in [0, 1] : |f(x)| \geq \lambda\}| \leq B_\lambda(x, y).$$

The functions B_λ have plenty of similarities with the above U. Actually, the formulas for $U|_{D_3}$ and $U|_{D_4}$ appear also in the definitions of B_λ. Let us briefly provide an informal explanation for this phenomenon. A crucial observation is that the functions B_λ originating from "analytic BMO" can also be used in the martingale setting described in Section 2, as special functions corresponding to $V(x, y, z) = 1_{\{|x| \geq \lambda\}}$ (formally, we treat the variable z as "empty", i.e., we take $B_\lambda(x, y, z) = B_\lambda(x, y)$). Applying the approach of Section 2, one obtains the sharp bound

$$\mathbb{P}(|X_\infty| \geq \lambda) \leq B_{\lambda/\|X\|_{\mathrm{BMO}}}\left(\frac{x}{\|X\|_{\mathrm{BMO}}}, \frac{y}{\|X\|_{\mathrm{BMO}}^2}\right), \tag{4.8}$$

where, as previously, $y = \mathbb{E}X_\infty^2$. This is of course very close to (4.7); the connection becomes even closer when one notes that for *some* pairs (x, y), the extremal martingales in (4.8) satisfy $\mathbb{P}(|X_\infty| \geq \lambda) = \mathbb{P}(X_\infty \geq \lambda)$. This explains why the same expressions appear in the definitions of the functions U and B_λ (on some parts of the domains).

ACKNOWLEDGMENTS. The author would like to thank an anonymous referee for the careful reading of the paper and helpful suggestions, which greatly improved the presentation.

References

[1] J. Bion-Nadal: *Dynamic risk measures*: *time consistency and risk measures from BMO martingales*, Finance Stoch. **12** (2008), 219–244.

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

[3] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker: *Weighted norm inequalities and hedging in incomplete markets*, Finance Stoch. **1** (1997), 181–227.

[4] C. Fefferman: *Characterizations of bounded mean oscillation*, Bull. Amer. Math. Soc. **77** (1971), 587–588.

[5] S. Geiss: *Weighted BMO and discrete time hedging within the Black–Scholes model*, Probab. Theory Related Fields **132** (2005), 13–38.

[6] A.M. Garsia: Martingale Inequalities: Seminar Notes on Recent Progress, W.A. Benjamin, Inc., Reading, MA, 1973.

[7] R.K. Getoor and M.J. Sharpe: *Conformal martingales*, Invent. Math. **16** (1972), 271–308.

[8] F. John and L. Nirenberg: *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. **14** (1961), 415–426.

[9] N. Kazamaki: Continuous Exponential Martingales and BMO, Lecture Notes in Mathematics **1579**, Springer, Berlin, 1994.

[10] A.A. Korenovskiĭ, *The connection between mean oscillations and exact exponents of summability of functions*, Math. USSR-Sb. **71** (1992), 561–567.

[11] P.A. Meyer: *Un cours sur les intégrales stochastiques*; in Séminaire de Probabilités, X (Seconde partie: Théorie des intégrales stochastiques, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975), Lecture Notes in Math. **511**, Springer, Berlin, 1976, 245–400.

[12] A. Osękowski: Sharp Martingale and Semimartingale Inequalities, Monografie Matematyczne (New Series) **72**, Birkhäuser/Springer Basel AG, Basel, 2012.

[13] L. Slavin and V. Vasyunin: *Sharp results in the integral-form John–Nirenberg inequality*, Trans. Amer. Math. Soc. **363** (2011), 4135–4169.

[14] V. Vasyunin: *The sharp constant in the John–Nirenberg inequality*, preprint POMI **20** (2003).

[15] V. Vasyunin and A. Volberg: *harp constants in the classical weak form of the John–Nirenberg inequality*, Proc. Lond. Math. Soc. (3) **108** (2014), 1417–1434.

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