

Pathwise versions of the Burkholder–Davis–Gundy inequality

MATHIAS BEIGLBOCK* and PIETRO SIORPAES**

Faculty of Mathematics, University of Vienna, Nordbergstrasse 15, 1090 Wien, Austria.
E-mail: *mathias.beiglboeck@univie.ac.at; **pietro.siorpaes@maths.ox.ac.uk

We present a new proof of the Burkholder–Davis–Gundy inequalities for $1 \leq p < \infty$. The novelty of our method is that these martingale inequalities are obtained as consequences of elementary *deterministic* counterparts. The latter have a natural interpretation in terms of robust hedging.

Keywords: Burkholder–Davis–Gundy; martingale inequalities; pathwise hedging

1. Introduction

In this paper, we derive estimates which compare the running maximum of a martingale with its quadratic variation. Given real numbers $x_n, h_n, n \in \mathbb{N}$ we write

$$x_n^* := \max_{k \leq n} |x_k|, \quad [x]_n := x_0^2 + \sum_{k=0}^{n-1} (x_{k+1} - x_k)^2, \quad (h \cdot x)_n := \sum_{k=0}^{n-1} h_k (x_{k+1} - x_k).$$

We will derive pathwise versions of the famous Burkholder–Davis–Gundy inequalities.

Theorem 1.1. *For $1 \leq p < \infty$, there exist constants $a_p, b_p < \infty$ such that the following holds: for every $N \in \mathbb{N}$ and every martingale $(X_k)_{k=0}^N$*

$$\mathbb{E}[X_N]^{p/2} \leq a_p \mathbb{E}[(X_N^*)^p], \quad \mathbb{E}[(X_N^*)^p] \leq b_p \mathbb{E}[X_N]^{p/2}. \quad (\text{BDG})$$

For $p \in (1, \infty)$ this was established by Burkholder [7]. Under additional assumptions, Burkholder and Gundy [9] obtain a version for $p \in (0, 1]$, while the case $p = 1$ of (BDG) without restrictions is due to Davis [16].

For a modern account see, for instance, [11].

Trajectorial inequalities. The novelty of this note is that the above martingale inequalities are established as consequences of *deterministic* counterparts. We postpone the general statements and first state the *trajectorial version* of Davis' inequality.

Theorem 1.2. Let x_0, \dots, x_N be real numbers and set¹ $h_n := \frac{x_n}{\sqrt{[x]_n + (x_n^*)^2}}$, $n \leq N$. Then

$$\sqrt{[x]_N} \leq 3x_N^* - (h \cdot x)_N, \quad x_N^* \leq 6\sqrt{[x]_N} + 2(h \cdot x)_N. \quad (1.1)$$

While the proof of Theorem 1.2 is not trivial, we emphasize that the inequalities in (1.1) are completely elementary in nature. The significance of the result lies in the fact that it implies Davis' inequalities: indeed, if $(X_n)_{n=0}^N$ is a martingale, we may apply (1.1) to each trajectory of X and obtain a bounded and adapted process H . The decisive observation is that, by the martingale property,

$$\mathbb{E}[(H \cdot X)_N] = 0, \quad (1.2)$$

so Davis' inequalities (with $a_1 = 3, b_1 = 6$) follow from (1.1) by taking expectations.

We recall that the BDG inequalities also apply if $X = (X_t)_t$ is a cadlag local martingale, and that this follows from a straightforward limiting procedure. Moreover, the inequalities are considerably simpler to prove for *continuous* local martingales (see, for example, [32]); in this case, they also hold for $p \in (0, 1)$, as proved by Burkholder and Gundy [9].

The problem of finding the optimal values of the constants a_p, b_p is delicate, and has been open for 47 years and counting; we refer to Osekowski [29] for a discussion of the current state of research.

The rest of this paper is organized as follows. In Section 2, we discuss the history of the pathwise approach to martingale inequalities. In Section 3, we explain the intuition behind the hedging strategy $h = (h_k)_k$ used in the pathwise version of Davis' inequality. In Section 4, we give a short proof of one Davis' inequality for continuous martingales; notably, this argument leads to a better constant compared to the previous literature (to the best of our knowledge). In Section 5, we establish Theorem 1.2. In Section 6, we use Theorem 1.2 to derive trajectorial versions of the BDG-inequalities in the $p > 1$ case; these also lead to their corresponding classical probabilistic counterpart, thus concluding a fully analytic derivation of Theorem 1.1.

2. History of the trajectorial approach

The inspiration of the pathwise approach to martingale inequalities used in this paper comes from mathematical finance, more specifically, the theory of model-independent pricing. The starting point of the field is the paper [21] of Hobson, which introduces the idea to study option-prices by means of *semi-static hedging*; we explain the concepts using the inequality

$$\sqrt{[x]_N} \leq 3x_N^* - (h \cdot x)_N \quad (2.1)$$

appearing in Theorem 1.2. If the process $x = (x_n)_{n=0}^N$ describes the price evolution of a financial asset, the functions $\Phi(x) = \sqrt{[x]_N}$ and $\Psi(x) = 3x_N^*$ have the natural financial interpretation of being exotic options; specifically, here Φ is an option on realized variance, while Ψ is a look-back option. The seller of the option Φ pays the buyer the amount $\Phi(x_0, \dots, x_N)$ after the option's

¹Throughout this paper we use the convention $0/0 = 0$.

expiration at time N , and $(h \cdot x)_N$ corresponds to the gains or losses accumulated while trading in x according to the portfolio $h = (h_k)_k$.

The decisive observation of Hobson is that inequalities of the type (2.1) can be used to derive *robust bounds* on the relation of the prices of Φ and Ψ : independently of the market model, one should never trade the option Φ at a price higher than the price of Ψ , since the payoff Φ can be *super-hedged* using the option Ψ plus self-financing trading. Here the *hedge* $3x_N^* - (h \cdot x)_N$ is designated *semi-static*: it is made up of a static part – the option $3x_N^*$ which is purchased at time 0 and kept during the entire time range – plus a dynamic part which corresponds to the trading in the underlying asset according to the strategy h .

Since the publication of [21] a considerable amount of literature on the topic has evolved (e.g., [6,12–15,17,23–25,31]); we refer in particular to the survey by Hobson [22] for a very readable introduction to this area. The most important tool in model-independent finance is the Skorokhod-embedding approach; an extensive overview is given by Obłój in [27]. Starting with the papers [3,20] the field has also been linked to the theory of optimal transport, leading to a formal development of the connection between martingale theory and robust hedging ([1,18,19]). A benefit for the theory of martingale inequalities is the following guiding principle:

Every martingale inequality which compares expectations of two functionals has a deterministic counterpart.

In fact, recently Bouchard and Nutz [5] have coined this into a rigorous theorem in the discrete time setup, see also [4].

This idea served as a motivation to derive the Doob-maximal inequalities from deterministic, discrete-time inequalities in [2].² In the present article, we aim to extend the approach to the case of the Burkholder–Davis–Gundy inequalities.

3. Heuristics for the pathwise hedging approach

The aim of this section is to explain the basic intuition which lies behind the choice of the integrand in the pathwise Davis inequalities. Arguments are simpler in the case of Brownian motion, which we will now consider.

We focus on one of the two inequalities; according to the pathwise hedging approach, we should be looking for a strategy H and a constant a such that $\sqrt{t} \leq aB_t^* + (H \cdot B)_t$. Indeed, a reasonable ansatz to find a super hedging strategy is to search for a function $f(b, b^*, t)$ such that

$$\sqrt{t} \leq aB_t^* + (f(B, B^*, t) \cdot B)_t, \quad t \geq 0. \tag{3.1}$$

To make an educated guess for the function f , we argue on a *purely heuristic* level and consider paths which evolve in a very particular way. Assume first that the path $(B_t(\omega))_{t \geq 0}$ stays infinitesimally close to the value b for all $t \geq t_0$: we picture BM as a random walk on a time grid with size dt , making alternating up and down steps of height \sqrt{dt} . Thus, we assume that B evolves in the form

$$B_{t_0+2ndt} = b, \quad B_{t_0+(2n+1)dt} = b + \sqrt{dt}, \quad n \geq 0, \tag{3.2}$$

²Notably, much of the approach of [2] was already developed earlier by Obłój and Yor [28].

where necessarily b lies between $-B_{t_0}^*$ and $B_{t_0}^*$. The left-hand side of (3.1) is of course increasing, so we have to ensure the same behavior on the right side. A little calculation reveals that this means that f should have the form

$$f(B, B^*, t) \approx -\frac{B}{\sqrt{t}} \quad \text{as } t \rightarrow \infty; \tag{3.3}$$

to see this, set $H_t := f(B_t, B_t^*, t)$ and compare the value $\sqrt{t+2dt} - \sqrt{t} \approx dt/\sqrt{t}$ with

$$\begin{aligned} (H \cdot B)_{t+2dt} - (H \cdot B)_t &\approx f(t, b, b^*) dB_t + f(t+dt, b + \sqrt{dt}, b^*) dB_{t+dt} \\ &\approx f(t, b, b^*)\sqrt{dt} + f(t+dt, b + \sqrt{dt}, b^*)(-\sqrt{dt}) \\ &\approx -[f(t, b + \sqrt{dt}, b^*) - f(t, b, b^*)]\sqrt{dt} + O(dt^{3/2}) \\ &\approx -f_b dt. \end{aligned}$$

To assure that both sides of (3.1) grow at the same speed, we thus need to require $dt/\sqrt{t} \approx -f_b dt$ which leads to (3.3).

Next, we consider a path which exhibits a different kind of extreme evolution: assume that $B_t(\omega) \approx Mt$ for some number $M > 0$. Simply setting $f(B, B^*, t) \approx -B/\sqrt{t}$ would lead to $(f(B, B^*, t) \cdot B)_t \approx -2M^2t^{3/2}/3$. Taking t sufficiently large, this quantity would eventually supersede $aB_t^* \approx aMt$ independent of the choice of a , and thus (3.1) would fail. So, this argument suggest to choose a function which is bounded (at least for fixed (t, B^*)). Moreover, dealing with a bounded integrand would conveniently allow to follow the explanation after Theorem 1.1 and obtain Davis’ inequalities from the pathwise Davis’ inequalities. Thus, we could consider the function

$$f(B, B^*, t) = -\frac{B_t}{\sqrt{t} \vee B_t^*}. \tag{3.4}$$

Thanks to the additional term aB_t^* in (3.1), it is not a problem if $f(B, B^*, t) \approx -2B/\sqrt{t}$ is violated for “small” values of t ; and, if \sqrt{t} is large compared to B^* , $f(B, B^*, t) \approx -2B_t/\sqrt{t}$ holds, thus satisfying (3.3). Another similar possibility would be to use the function

$$f(B, B^*, t) = -\frac{B_t}{\sqrt{t + (B_t^*)^2}}, \tag{3.5}$$

as in Theorem 1.2; the latter turns out to lead to easier computations in the discrete time case. We choose however f given by (3.4) when dealing with continuous martingales, since this allows us to obtain Davis’ inequality with a better constant than the values we could find in the literature.

4. Davis inequality for continuous local martingales

We now derive one pathwise Davis’ inequality for continuous local martingales; integrating it yields the corresponding standard Davis’ inequality. We notice that Theorem 4.1 provides the constant $3/2$, which is smaller than the optimal constant for general cadlag martingales (which is known to be $\sqrt{3}$, see [8]). We do not address here the opposite pathwise Davis’ inequality for

continuous local martingales since we are only interested in Theorem 4.1 for illustrative purposes (since, as mentioned before, Davis inequality for cadlag local martingales follows from the case of martingales in discrete time).

Theorem 4.1. *If M is a continuous local martingale such that $M_0 = 0$ then*

$$\sqrt{[M]_t} \leq \frac{3}{2} M_t^* - \left(\frac{M_t}{\sqrt{[M]_t} \vee M_t^*} \cdot M_t \right)_t \quad \text{for all } t \geq 0. \tag{4.1}$$

Proof. By the Dambis–Dubins–Schwarz time change result, it is enough to consider the case where M is a Brownian Motion, which we will denote by B . From Ito’s formula applied to the semi-martingales B_t^2 and $\sqrt{t} \vee B_t^*$ we find

$$d \frac{B_t^2}{\sqrt{t} \vee B_t^*} = - \frac{B_t^2}{t \vee B_t^{*2}} d(\sqrt{t} \vee B_t^*) + \frac{1}{\sqrt{t} \vee B_t^*} (2B_t dB_t + dt).$$

We may thus replace the integral in (4.1) and arrive at the equivalent formulation

$$\frac{B_t^2}{\sqrt{t} \vee B_t^*} + \int_0^t \frac{B_s^2}{s \vee B_s^{*2}} d(\sqrt{s} \vee B_s^*) - \int_0^t \frac{1}{\sqrt{s} \vee B_s^*} ds \leq 3B_t^* - 2\sqrt{t}. \tag{4.2}$$

Inequality (4.2) gets stronger if we replace each occurrence of B by B^* ; thus, setting $f(t) = \sqrt{t}$, $g(t) = B_t^*$, it is enough to prove the following claim:

Let $f, g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be continuous increasing functions such that $f(0) = g(0) = 0$ and $(f \vee g)(a) > 0$ if $a > 0$. Then, for all $a > 0$

$$\left(\frac{g^2}{f \vee g} \right)(a) + \int_0^a \frac{g^2}{f^2 \vee g^2} d(f \vee g) - \int_0^a \frac{1}{f \vee g} df^2 \leq (3g - 2f)(a). \tag{4.3}$$

To show this, observe that, by a change of variables $\int \frac{g^2}{f^2 \vee g^2} d(f \vee g) = - \int g^2 d \frac{f \vee g}{f^2 \vee g^2}$. Hence, integrating by parts on the interval (ε, a) and taking the limit $\varepsilon \rightarrow 0$, we see that the left hand side of (4.3) equals

$$\int_0^a \frac{dg^2 - df^2}{f \vee g}.$$

By a change of variables and applying trivial inequalities we obtain

$$\int_0^a \frac{dg^2}{f \vee g} = \int_0^a 1_{\{g>0\}} \frac{dg^2}{f \vee g} \leq \int_0^a \frac{1_{\{g>0\}} dg^2}{g} = 2g(a), \quad \int_0^a \frac{df^2}{f \vee g} \geq \int_0^a \frac{df^2}{f(\cdot) \vee g(a)}.$$

If $f(a) \leq g(a)$, the last integral equals $f^2(a)/g(a)$; otherwise there exists some $b \in [0, a)$ such that $f(b) = g(a)$, and then evaluating separately the integral on $(0, b)$ and on $[b, a)$ we obtain that

$$\int_0^a \frac{df^2}{f(\cdot) \vee g(a)} = \frac{f^2(b)}{g(a)} + 2(f(a) - f(b)) = 2f(a) - g(a).$$

Since $2y - x^2/y \leq 3y - 2x$ holds for $y > 0$, either way (4.3) follows. □

5. Davis inequality

In this section, we prove Theorem 1.2; in fact, we will establish that³

$$\sqrt{[x]_n} \leq (\sqrt{2} + 1)x_n^* + (-h \cdot x)_n, \tag{5.1}$$

$$x_n^* \leq 6\sqrt{[x]_n} + (2h \cdot x)_n, \tag{5.2}$$

where the dynamic hedging strategy is defined by $h_n = \frac{x_n}{\sqrt{[x]_n + (x_n^*)^2}}$ as in Theorem 1.2.

To prove (5.1), (5.2) we introduce the convention, used throughout the paper, that any sequence $(y_i)_{i \geq 0}$ is defined to be 0 at time $i = -1$, and we define the auxiliary functions f, g for $m > 0, q \geq 0, |x| \leq m$ by

$$f(x, m, q) := -2\sqrt{q} + \sqrt{m^2 + q} - \frac{m^2 - x^2}{2\sqrt{m^2 + q}}, \tag{5.3}$$

$$g(x, m, q) := -2m + \sqrt{m^2 + q} + \frac{m^2 - x^2}{2\sqrt{m^2 + q}} \tag{5.4}$$

and continuously extend them to $(x, m, q) = (0, 0, 0)$ by setting $f(0, 0, 0) = g(0, 0, 0) = 0$. We will need the following lemma, whose proof is a somewhat tedious exercise in calculus.

Lemma 5.1. *For $d \in \mathbb{R}, |x| \leq m, q \geq 0, m \geq 0$ we have, with $c = \sqrt{2} - 1$,*

$$f(x + d, m \vee |x + d|, q + d^2) - f(x, m, q) \leq \frac{xd}{\sqrt{m^2 + q}} + (\sqrt{q + d^2} - \sqrt{q}), \tag{5.5}$$

$$g(x + d, m \vee |x + d|, q + d^2) - g(x, m, q) \leq -\frac{xd}{\sqrt{m^2 + q}} + c((m \vee |x + d|) - m). \tag{5.6}$$

Before proving Lemma 5.1 we explain why it implies (5.1) and (5.2).

Proof of Theorem 1.2. Since $f(x_0, |x_0|, x_0^2) \leq 0$, (5.5) gives

$$\begin{aligned} -2\sqrt{[x]_n} + x_n^*/2 &\leq f(x_n, x_n^*, [x]_n) \leq \sum_{k=0}^{n-1} f(x_{k+1}, x_{k+1}^*, [x]_{k+1}) - f(x_k, x_k^*, [x]_k) \\ &\leq (h \cdot x)_n + \sqrt{[x]_n}, \end{aligned}$$

which implies (5.1); and since $g(x_0, |x_0|, x_0^2) \leq 0$, we get (5.2) from (5.6) as follows

$$\begin{aligned} -2x_n^* + \sqrt{[x]_n} &\leq g(x_n, x_n^*, [x]_n) \leq \sum_{k=0}^{n-1} g(x_{k+1}, x_{k+1}^*, [x]_{k+1}) - g(x_k, x_k^*, [x]_k) \\ &\leq -(h \cdot x)_n + cx_n^*. \end{aligned} \quad \square$$

³Inequality (5.1) slightly improves on Inequality (1.1) by replacing the constant 3 with the smaller $1 + \sqrt{2}$.

Now we prove Lemma 5.1.

Proof of Inequality (5.6). It is enough to consider the case $m > 0$, as the one where $m = 0$ then follows by continuity. Then, we can assume that $m = 1$ through normalization. Define $h(x, q, d)$ to be the LHS minus the RHS of (5.6); since $h(x, q, d) = h(-x, q, -d)$, it is sufficient to deal with the case $d \geq 0$.

Case I [$1 \geq |x + d|$]: Here we have to show that

$$h = \sqrt{1 + q + d^2} + \frac{1 - (x + d)^2}{2\sqrt{1 + q + d^2}} - \sqrt{1 + q} - \frac{1 - x^2}{2\sqrt{1 + q}} + \frac{xd}{\sqrt{1 + q}} \leq 0. \tag{5.7}$$

Since $h_{xx} \geq 0$, h is convex, so it is sufficient to treat the boundary cases $x = -1$ and $x = 1 - d$. To simplify notation, we set $r = \sqrt{1 + q}$; notice that $r \geq 1$ and $0 \leq d \leq 2$.

Sub-case I.A [$1 \geq |x + d|, x = -1$]: Then (5.7) follows from

$$\begin{aligned} & \sqrt{r^2 + d^2} + \frac{1 - (d - 1)^2}{2\sqrt{r^2 + d^2}} - r - \frac{d}{r} \leq 0 \\ \Leftrightarrow & r^2 + d^2 + d - d^2/2 \leq (r + d/r)\sqrt{r^2 + d^2} \\ \Leftrightarrow & r^4 + d^4/4 + d^2 + r^2d^2 + d^3 + 2dr^2 \leq r^4 + 2dr^2 + d^2 + r^2d^2 + 2d^3 + d^4/r^2 \\ \Leftrightarrow & d^4/4 \leq d^3 + d^4/r^2, \end{aligned}$$

which is true since $0 \leq d \leq 2$.

Sub-case I.B [$1 \geq |x + d|, x = 1 - d$]: Here (5.7) amounts to

$$\begin{aligned} & \sqrt{r^2 + d^2} - r - \frac{1 - (1 - d)^2}{2r} + \frac{(1 - d)d}{r} \leq 0 \\ \Leftrightarrow & \sqrt{r^2 + d^2} \leq r + d^2/2r \\ \Leftrightarrow & r^2 + d^2 \leq r^2 + d^2 + d^4/4r^2. \end{aligned}$$

Case II [$1 \leq |x + d|$]: Since $|x| \leq 1$ and $d \geq 0$, we find that $|x + d| \geq 1$ implies $x + d = |x + d| \geq 1$. In this case h equals

$$-(2 + c)(x + d - 1) + \sqrt{(x + d)^2 + q + d^2} - \sqrt{1 + q} - \frac{1 - x^2}{2\sqrt{1 + q}} + \frac{xd}{\sqrt{1 + q}}. \tag{5.8}$$

Since $s \mapsto \sqrt{s^2 + 1}$ is convex, $h \leq 0$ holds iff it holds for all x on the boundary. Moreover if $-1 \leq 1 - d = x \leq 1$, then we already know that $h \leq 0$ from the corresponding sub-case 1 $\geq |x + d|$; so we only need to show that $h \leq 0$ for $x = 1, q, d \geq 0$ and for $x = -1, q \geq 0, d \geq 2$, respectively.

Sub-case II.A [$1 \leq |x + d|, x = 1$]: We have to show that, for all $q, d \geq 0$,

$$h(1, q, d) = -(2 + c)d + \sqrt{(1 + d)^2 + q + d^2} - \sqrt{1 + q} + \frac{d}{\sqrt{1 + q}} \leq 0.$$

Since $(1 + d)^2 + d^2 = 2(d + 1/2)^2 + 1/2$ and $s \mapsto \sqrt{1 + s^2}$ is convex, it follows that $h(1, q, d)$ is convex in d ; hence, the inequality has to be checked only for $d = 0$ and for $d \rightarrow \infty$. The first case is trivial, and in the latter, after dividing both sides by d , we arrive at $-(2 + c) + \sqrt{2} + 1/\sqrt{1 + q} \leq 0$, which holds by our choice of c and the fact that $q \geq 0$.

Sub-case II.B [$1 \leq x + d, x = -1$]: We have to show that, for all $q \geq 0, d \geq 2$,

$$h(-1, q, d) = -(2 + c)(d - 2) + \sqrt{(-1 + d)^2 + q + d^2} - \sqrt{1 + q} - \frac{d}{\sqrt{1 + q}} \leq 0.$$

As above, by convexity in d it suffices to consider the cases $d = 2$ and $d \rightarrow \infty$. The first one amounts to $\sqrt{5 + q} \leq \sqrt{1 + q} + 2/\sqrt{1 + q}$, which is easily proved taking the squares. The second one, after dividing by d , amounts to $-(2 + c) + \sqrt{2} - 1/\sqrt{1 + q} \leq 0$, which holds since $-(2 + c) + \sqrt{2} \leq 0$ by our choice of c . □

Proof of Inequality (5.5). As before, we can assume w.l.o.g. that $m = 1$ and $d \geq 0$. Define $k(x, q, d)$ to be the LHS minus the RHS of (5.5).

Case I [$1 \geq |x + d|$]: In this case, k equals

$$\sqrt{1 + q + d^2} - \frac{1 - (x + d)^2}{2\sqrt{1 + q + d^2}} - \sqrt{1 + q} + \frac{1 - x^2}{2\sqrt{1 + q}} - \frac{xd}{\sqrt{1 + q}} - 3(\sqrt{q + d^2} - \sqrt{q}).$$

Let us first isolate the terms that depend on x . Define $k_0 := (1 + q + d^2)^{-1/2} - (1 + q)^{-1/2}$, and $k_2 := k - k_0(x + d)^2/2$, so that

$$k_2 = \sqrt{1 + q + d^2} - \sqrt{1 + q} - \frac{1}{2\sqrt{1 + q + d^2}} + \frac{1 + d^2}{2\sqrt{1 + q}} - 3(\sqrt{q + d^2} - \sqrt{q}).$$

Notice that we can write

$$k_0 = \int_0^{d^2} k_1(s) \, ds \quad \text{for } k_1(s) := \frac{d}{ds} (1 + q + s)^{-1/2},$$

and similarly $k_2 = \int_0^{d^2} k_3(s, d^2) \, ds$ for

$$k_3(s, d^2) := \frac{d}{ds} \left(\sqrt{1 + q + s} - \frac{1 - s + d^2}{2\sqrt{1 + q + s}} - 3\sqrt{q + s} \right) \tag{5.9}$$

$$= \frac{1}{2\sqrt{1 + q + s}} + \frac{2(1 + q + s) + 1 - s + d^2}{4(1 + q + s)^{3/2}} - \frac{3}{2\sqrt{q + s}}. \tag{5.10}$$

Since the $(k_i)_i$ do not depend on x and $k_0 \leq 0$, $\max_x k = k_2 + k_0 \min_x (x + d)^2/2$. Since $\min_{-1 \leq x \leq 1} (x + d)^2$ equals 0 if $0 \leq d \leq 1$ and equals $(-1 + d)^2$ if $1 \leq d$, to show $k \leq 0$ we are lead to study the following two sub-cases.

Sub-case I.A [$1 \geq |x + d|, d \leq 1$]: In this case, $k = k_2$; to show that $k_2 \leq 0$ it is enough to show $k_3 \leq 0$. Since $0 \leq s \leq d^2 \leq 1$ we get $-s + d^2 \leq 1$, and so trivially

$$k_3 \leq \frac{2(1 + q + s) + 1 + 1}{4(1 + q + s)^{3/2}} - \frac{2}{2\sqrt{q + s}}. \tag{5.11}$$

So, calling $y := q + s$, it is enough to prove that for all $y \geq 0$

$$\frac{2y + 4}{4(1 + y)^{3/2}} - \frac{2}{2\sqrt{y}} \leq 0, \quad \text{i.e.} \quad \sqrt{y}(y + 2) \leq (1 + y)^{3/2}, \tag{5.12}$$

which is seen to be true by taking squares and bringing everything on the RHS to obtain a polynomial whose coefficients are all positive.

Sub-case I.B [$1 \geq |x + d|, d \geq 1$]: In this case $k = k_2 + k_0(1 - d)^2/2$, so it is enough to show that $k_3 + k_1(1 - d)^2/2 \leq 0$. Since from $1 \geq |x + d|, |x| \leq 1$ it follows that $d \leq 2$, computations entirely similar⁴ to the other sub-case establish the desired result.

Case II [$1 \leq |x + d|$]: In this case, $x + d = |x + d| \geq 1$ and k equals

$$\sqrt{(x + d)^2 + q + d^2} - \sqrt{1 + q} + \frac{1 - x^2}{2\sqrt{1 + q}} - \frac{xd}{\sqrt{1 + q}} - 3(\sqrt{q + d^2} - \sqrt{q}).$$

Since trivially $dk/dx \leq 0$, to show $k \leq 0$ we can assume that $x = 1 - d$, in which case we can write k as $k = \int_0^{d^2} \tilde{k}(s) ds$ for

$$\tilde{k}(s) := \frac{d}{ds} \left(\sqrt{1 + q + s} + \frac{s}{2\sqrt{1 + q}} - 3\sqrt{q + s} \right) \tag{5.13}$$

$$= \frac{1}{2\sqrt{1 + q + s}} + \frac{1}{2\sqrt{1 + q}} - \frac{3}{2\sqrt{q + s}}. \tag{5.14}$$

Since $1 - d = x \in [-1, 1]$ we have $d^2 \leq 4$, and so to get $k \leq 0$ it suffices to show that $\tilde{k} \leq 0$ for $s \leq 4$. This holds since

$$\tilde{k} \leq \frac{1}{2\sqrt{1 + q}} - \frac{2}{2\sqrt{q + s}} \leq 0 \quad \text{for } s \leq 4. \quad \square$$

6. Pathwise Burkholder–Gundy inequality

Garsia has given a simple proof of the fact that the BDG inequalities for general $p \geq 1$ are a consequence of Davis inequality ($p = 1$) and of the famous lemma by Garsia and Neveu; in this section we revisit his proof and turn it into pathwise discrete-time arguments.

Garsia’s proof (for which we refer to [26], Chapter 3, Theorems 30 and 32 or to [10]) works similarly to how the Doob L^p -inequalities for $p > 1$ follow by writing x^p as an integral, applying the (weak) Doob L^1 -inequality, using Fubini’s theorem, and finally applying Hölder’s inequality

⁴Use that in this case $0 \leq s \leq d^2 \leq 4$ implies $-s + d^2 - (d - 1)^2 \leq 3$.

(see for example [30]). The difference is that for the BDG inequalities one needs to use a different integral expression for x^p , and so one has to consider Davis' inequalities not on the time interval $[0, T]$ but on $[\tau, T]$, where τ is a stopping time.

In the pathwise setting, by the guiding principle stated in Section 2, if L is a functional of a martingale X and τ is a stopping time, a statement of the type $\mathbb{E}[L|F_\tau] \leq 0$ will have to be turned into one of the type $L + (H \cdot X)_T - (H \cdot X)_\tau \leq 0$; moreover, since there will be no expectations involved, Hölder's inequality will have to be replaced by Young's inequality.

We will need to consider discrete time stochastic integrals for which the initial time is different from 0; given $i < n$ and real numbers $(h_j)_{i \leq j \leq n-1}$ and $(x_j)_{i \leq j \leq n}$, we define

$$(h \cdot x)_i^n := \sum_{j=i}^{n-1} h_j(x_{j+1} - x_j). \tag{6.1}$$

Moreover if, for $i \leq j \leq n - 1$, h_j is a function from \mathbb{R}^{j+1} to \mathbb{R} , given real numbers $(x_j)_{0 \leq j \leq n}$ we define $(h \cdot x)_i^n$ as

$$\sum_{j=i}^{n-1} h_j(x_0, \dots, x_j)(x_{j+1} - x_j).$$

Either way, we set $(h \cdot x)_i^n := 0$ if $n = i$.

We now deduce pathwise Davis' inequalities on $\{i, i + 1, \dots, n\}$ from the ones on $\{0, 1, \dots, n\}$ by a simple time shift.

Lemma 6.1. *Assume that $\alpha, \beta > 0$ and $h_n, k_n : \mathbb{R}^{n+1} \rightarrow \mathbb{R}, n \geq 0$ satisfy*

$$\sqrt{[x]_n} \leq \alpha x_n^* + (h \cdot x)_n, \quad x_n^* \leq \beta \sqrt{[x]_n} + (k \cdot x)_n \tag{6.2}$$

for every sequence $(x_n)_{n \geq 0}$. Define, for $i \geq 0, n \geq i$, the functions $f_n^{(i)}, g_n^{(i)} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ by

$$f_n^{(i)}((x_j)_{0 \leq j \leq n}) := h_{n-i}((x_l - x_{l-1})_{i \leq l \leq n}), \quad g_n^{(i)}((x_j)_{j \leq n}) := k_{n-i}((x_l - x_{l-1})_{i \leq l \leq n}).$$

Then we have, for $n \geq i \geq 0$,

$$\sqrt{[x]_n} - \sqrt{[x]_{i-1}} \leq 2\alpha x_n^* + (f^{(i)} \cdot x)_i^n, \quad x_n^* - x_{i-1}^* \leq \beta \sqrt{[x]_n} + (g^{(i)} \cdot x)_i^n.$$

Proof. Fix $n \geq i \geq 0, (x_n)_{n \geq 0}$ and let $y_j^{(i)} := x_{j+i} - x_{i-1}$. Applying (6.2) to $(y_j^{(i)})_{j \geq 0}$ we find

$$\begin{aligned} \sqrt{[x]_n} - \sqrt{[x]_{i-1}} &\leq \sqrt{[x]_n - [x]_{i-1}} = \sqrt{[y^{(i)}]_{n-i}} \leq \alpha (y^{(i)})_{n-i}^* + (h \cdot y^{(i)})_{n-i} \\ &\leq \alpha 2x_n^* + (f^{(i)} \cdot x)_i^n, \end{aligned}$$

and (respectively)

$$x_n^* - x_{i-1}^* \leq (y^{(i)})_{n-i}^* \leq \beta \sqrt{[y^{(i)}]_{n-i}} + (k \cdot y^{(i)})_{n-i} \leq \beta \sqrt{[x]_n} + (g^{(i)} \cdot x)_i^n. \quad \square$$

Here follows the pathwise version of Garsia–Neveu’s lemma.

Lemma 6.2. *Let $p > 1$, $c_n \in \mathbb{R}$, $(x_j)_{j \leq n}$, $(h_n^{(i)})_{i \leq n} \in \mathbb{R}^{n+1}$, and assume that $0 = a_{-1} \leq a_0 \leq \dots \leq a_n < \infty$ and*

$$a_n - a_{i-1} \leq c_n + (h^{(i)} \cdot x)_i^n \quad \text{for } n \geq i \geq 0.$$

Then, if we set

$$w_j := \sum_{i=0}^j p(a_i^{p-1} - a_{i-1}^{p-1})h_j^{(i)}, \quad j \leq n,$$

we have that

$$a_n^p \leq pc_n a_n^{p-1} + (w \cdot x)_n, \tag{6.3}$$

$$a_n^p \leq (p-1)^{p-1} c_n^p + (pw \cdot x)_n. \tag{6.4}$$

Proof. From $a_n^p = p(p-1) \int_0^{a_n} s^{p-2}(a_n - s) ds = p \sum_{i=0}^n \int_{a_{i-1}}^{a_i} (p-1)s^{p-2}(a_n - s) ds$ and $a_n - s \leq a_n - a_{i-1}$ on $s \in [a_{i-1}, a_i]$, we find (6.3) by writing

$$\begin{aligned} a_n^p &\leq p \sum_{i=0}^n (a_i^{p-1} - a_{i-1}^{p-1})(a_n - a_{i-1}) \\ &\leq p \sum_{i=0}^n (a_i^{p-1} - a_{i-1}^{p-1})[c_n + (h^{(i)} \cdot x)_i^n] \\ &= pc_n a_n^{p-1} + p \sum_{i=0}^n \sum_{j=i}^{n-1} (a_i^{p-1} - a_{i-1}^{p-1})h_j^{(i)}(x_{j+1} - x_j) \\ &= pc_n a_n^{p-1} + \sum_{j=0}^{n-1} \left(\sum_{i=0}^j p(a_i^{p-1} - a_{i-1}^{p-1})h_j^{(i)} \right) (x_{j+1} - x_j) = pc_n a_n^{p-1} + (w \cdot x)_n. \end{aligned}$$

We then obtain (6.4) from (6.3) by applying Young’s inequality $ab \leq C_\epsilon a^p/p + \epsilon b^q/q$ (where $C_\epsilon^{-1} = p(\epsilon q)^{p-1}$ and $1/p + 1/q = 1$) with $\epsilon = 1/p$, $a = c_n$, $b = a_n^{p-1}$. \square

Finally, from Theorem 1.2, Lemma 6.1 and Lemma 6.2, we obtain the following discrete-time pathwise BDG inequalities for $p > 1$. We recall that, by convention, $x_{-1} = x_{-1}^* = [x]_{-1} = 0$ and $0/0 = 0$, and in particular the integrand $f_n^{(i)}$ is well defined.

Theorem 6.3. *Let x_0, \dots, x_N be real numbers, $c_p := 6^p(p-1)^{p-1}$ for $p > 1$, and define*

$$h_n := \sum_{i=0}^n p^2 \left(\sqrt{[x]_i^{p-1}} - \sqrt{[x]_{i-1}^{p-1}} \right) f_n^{(i)}, \quad g_n := \sum_{i=0}^n p^2 \left((x_i^*)^{p-1} - (x_{i-1}^*)^{p-1} \right) f_n^{(i)},$$

where

$$f_n^{(i)} := \frac{x_n - x_{i-1}}{\sqrt{[x]_n - [x]_{i-1} + \max_{i \leq k \leq n} (x_k - x_{i-1})^2}}.$$

Then

$$\sqrt{[x]_N^p} \leq c_p (x_N^*)^p - (h \cdot x)_N, \quad (x_N^*)^p \leq c_p \sqrt{[x]_N^p} + 2(g \cdot x)_N. \quad (6.5)$$

We notice that Theorem 6.3 yields (BDG); indeed, given a finite constant N and a martingale $(X_n)_{n=0}^N$, trivially $\sqrt{[X]_N}$ and X_N^* are in $L^p(\mathbb{P})$ iff X_n is in $L^p(\mathbb{P})$ for every $n \leq N$, and in this case the adapted integrands $(H_n)_{n=0}^{N-1}$ and $(G_n)_{n=0}^{N-1}$ which we obtain applying Theorem 6.3 to the paths of X are in $L^q(\mathbb{P})$ for every n (for $q = p/(p-1)$), thus $H \cdot X$ and $G \cdot X$ are martingales and so

$$\mathbb{E}[(H \cdot X)_N] = 0 = \mathbb{E}[(G \cdot X)_N],$$

and the Burkholder–Davis–Gundy inequalities for $p > 1$ (with $a_p = b_p = 6^p(p-1)^{p-1}$) follow from (6.5) by taking expectations, completing the proof of Theorem 1.1.

Acknowledgements

The authors thank Harald Oberhauser for comments on an earlier version of this paper. The first author thanks the Austrian Science Fund for support through project p21209.

References

- [1] Acciaio, B., Beiglböck, M., Penkner, F. and Schachermayer, W. (2013) A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *Math. Finance*. DOI:10.1111/mafi.12060.
- [2] Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W. and Temme, J. (2013). A trajectorial interpretation of Doob’s martingale inequalities. *Ann. Appl. Probab.* **23** 1494–1505. MR3098440
- [3] Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985
- [4] Beiglböck, M. and Nutz, M. (2014). Martingale inequalities and deterministic counterparts. arXiv preprint, arXiv:1401.4698.
- [5] Bouchard, B. and Nutz, M. (2013). Arbitrage and duality in nondominated discrete-time models. arXiv preprint, arXiv:1305.6008.
- [6] Brown, H., Hobson, D. and Rogers, L.C.G. (2001). Robust hedging of barrier options. *Math. Finance* **11** 285–314. MR1839367
- [7] Burkholder, D.L. (1966). Martingale transforms. *Ann. Math. Statist.* **37** 1494–1504. MR0208647
- [8] Burkholder, D.L. (2002). The best constant in the Davis inequality for the expectation of the martingale square function. *Trans. Amer. Math. Soc.* **354** 91–105 (electronic). MR1859027
- [9] Burkholder, D.L. and Gundy, R.F. (1970). Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.* **124** 249–304. MR0440695

- [10] Chou, C.S. (1975). Les méthodes d'A. Garsia en théorie des martingales. Extensions au cas continu. In *Séminaire de Probabilités IX Université de Strasbourg. Lecture Notes in Math.* **465** 213–225. Berlin: Springer. [MR0428427](#)
- [11] Chow, Y.S. and Teicher, H. (2012). *Probability Theory: Independence, Interchangeability, Martingales*. New York: Springer.
- [12] Cox, A.M.G., Hobson, D. and Oblój, J. (2008). Pathwise inequalities for local time: Applications to Skorokhod embeddings and optimal stopping. *Ann. Appl. Probab.* **18** 1870–1896. [MR2462552](#)
- [13] Cox, A.M.G. and Oblój, J. (2011). Robust hedging of double touch barrier options. *SIAM J. Financial Math.* **2** 141–182. [MR2772387](#)
- [14] Cox, A.M.G. and Oblój, J. (2011). Robust pricing and hedging of double no-touch options. *Finance Stoch.* **15** 573–605. [MR2833100](#)
- [15] Cox, A.M.G. and Wang, J. (2013). Root's barrier: Construction, optimality and applications to variance options. *Ann. Appl. Probab.* **23** 859–894. [MR3076672](#)
- [16] Davis, B. (1970). On the integrability of the martingale square function. *Israel J. Math.* **8** 187–190. [MR0268966](#)
- [17] Davis, M., Oblój, J. and Raval, V. (2010) Arbitrage bounds for prices of options on realized variance. ArXiv e-prints.
- [18] Dolinsky, Y. and Soner, H.M. (2013). Robust hedging with proportional transaction costs. arXiv e-prints.
- [19] Dolinsky, Y. and Soner, M.H. (2012). Robust hedging and martingale optimal transport in continuous time. arXiv e-prints.
- [20] Galichon, A., Henry-Labordère, P. and Touzi, N. (2011). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options.
- [21] Hobson, D. (1998). Robust hedging of the lookback option. *Finance Stoch.* **2** 329–347.
- [22] Hobson, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math.* **2003** 267–318. Berlin: Springer. [MR2762363](#)
- [23] Hobson, D. and Klimmek, M. (2012). Model-independent hedging strategies for variance swaps. *Finance Stoch.* **16** 611–649. [MR2972236](#)
- [24] Hobson, D. and Neuberger, A. (2012). Robust bounds for forward start options. *Math. Finance* **22** 31–56. [MR2881879](#)
- [25] Hobson, D.G. and Pedersen, J.L. (2002). The minimum maximum of a continuous martingale with given initial and terminal laws. *Ann. Probab.* **30** 978–999. [MR1906424](#)
- [26] Meyer, P.A. (1976). 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** 245–400. Berlin: Springer. [MR0501332](#)
- [27] Oblój, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1** 321–390. [MR2068476](#)
- [28] Oblój, J. and Yor, M. (2006). On local martingale and its supremum: Harmonic functions and beyond. In *From Stochastic Calculus to Mathematical Finance* 517–533. Berlin: Springer. [MR2234288](#)
- [29] Ośgowski, A. (2010). Sharp maximal inequalities for the martingale square bracket. *Stochastics* **82** 589–605. [MR2783421](#)
- [30] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Berlin: Springer. [MR1725357](#)
- [31] Rogers, L.C.G. (1993). The joint law of the maximum and terminal value of a martingale. *Probab. Theory Related Fields* **95** 451–466. [MR1217446](#)

- [32] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2, Itô Calculus. Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. Reprint of the second (1994) edition. [MR1780932](#)

Received June 2013