arXiv: arXiv:1507.07699

The Sharp Constant for the Burkholder-Davis-Gundy Inequality and Non-Smooth Pasting

Walter Schachermayer [1,*,*] **and Florian Stebegg** [2,**,†]

[1] *Fakultät für Mathematik, Universität Wien*
Oskar-Morgenstern-Platz 1, Room 06.131
A-1090 Wien.
e-mail: [*] `walter.schachermayer@univie.ac.at`

[2] *Department of Statistics, Columbia University*
1255 Amsterdam Avenue, Room 906
10027 New York, NY.
e-mail: [**] `florian.stebegg@columbia.edu`

Abstract: We revisit the celebrated family of BDG-inequalities introduced by Burkholder, Gundy [8] and Davis [10] for continuous martingales. For the inequalities $\mathbb{E}[\tau^{\frac{p}{2}}] \leqslant C_p \mathbb{E}[(B^*(\tau))^p]$ with $0 < p < 2$ we propose a connection of the optimal constant C_p with an ordinary integro-differential equation which gives rise to a numerical method of finding this constant. Based on numerical evidence we are able to calculate, for $p = 1$, the explicit value of the optimal constant C_1, namely $C_1 = 1,27267\ldots$. In the course of our analysis, we find a remarkable appearance of "non-smooth pasting" for a solution of a related ordinary integro-differential equation.

MSC 2010 subject classifications: Primary 62L15, 93E20; secondary 45J05.
Keywords and phrases: BDG inequality, Optimal Stopping, Non-Smooth Pasting, Ordinary Integro-Differential Equations.

1. Introduction

We consider the following version of the Burkholder-Davis-Gundy inequality [8], [10]:

Theorem 1.1. *There is a constant $C > 0$ such that, for every bounded stopping time τ, we have*

$$\mathbb{E}\left[\tau^{\frac{1}{2}}\right] \leqslant C\,\mathbb{E}[B^*(\tau)]. \tag{1}$$

Here $(B(t))_{t \geqslant 0}$ denotes a standard Brownian motion, starting at $B(0) = 0$.

---

[*] Supported in part by the Austrian Science Fund (FWF) under grant P25815 and P28661 and by the Vienna Science and Technology Fund (WWTF) under grant MA14-008.

[†] Supported in part by the Austrian Science Fund (FWF) under grants P26736 and Y782-N25.

By $B^*(t)$ we denote the corresponding running maximum of the absolute value

$$B^*(t) := \sup_{0 \leqslant u \leqslant t} |B(u)|.$$

It is obvious that the set of constants C which satisfy inequality (1) is a closed, unbounded interval in $\mathbb{R}_+$. By the results of [5] it is known that $C = \frac{3}{2}$ is contained in this set. To the best of our knowledge, this is the smallest constant known in the previous literature. In the present paper we establish the optimal value for this constant.

Theorem 1.2. *There is an ordinary integro-differential equation (see* (31) *below) depending on real parameters* $C > 0$ *and* $t_0 > 0$ *such that* C *satisfies* (1) *if and only if there is* t_0 *such that this equation has a well-defined solution.*

Numerical solutions of the equation (31) *reveal that the smallest such* C*, i.e. the optimal constant in the Burkholder-Davis-Gundy inequality* (1)*, equals*

$$\widehat{C} \approx 1,27267\ldots.$$

The paper is organised as follows. As usual in stochastic control theory, we first introduce the *value function* of the *optimal stopping* problem which corresponds to the inequality (1). After some structural facts about the stopping problem we turn to some analytic properties of the value function in Section 3. We deduce the OIDE (ordinary integro-differential equation) which is referred to in Theorem 1.2. The subsequent section is devoted to properties of solutions to the fundamental OIDE (31) which are needed to identify these solutions with the value function of the stopping problem in Section 5.

The critical $\widehat{t}_0 > 0$ associated to the optimal constant $\widehat{C}$ via (31) below also turns out to be of somewhat independent interest: if ρ denotes the first moment, say after $t = 1$, when t is bigger than $cB^*(t)^2$, then $\mathbb{E}[\rho^{\frac{1}{2}}]$ is finite or infinite depending on whether c is smaller or bigger than $\widehat{t}_0$ (Proposition 5.6 and 5.7). In Section 6 we state a pointwise version of the BDG inequalities and in Section 7 we briefly discuss the case of general $0 < p < 2$ without entering into a numerical analysis. Finally, in Section 8 we discuss the fact why the constant $\widehat{C} = \sqrt{3}$ which was established by D. Burkholder [7] as the optimal constant for (1) in the case of martingales which are not necessarily continuous, is different from the present constant $\widehat{C} = 1,27267\ldots$ which holds true for continuous processes. We relate this discrepancy with a certain lack of concavity of the value function.

2. The Value Function of an Optimal Stopping Problem

Fix a constant $C > 0$. Following a well-known path in optimal control theory we define the value function

$$V(t, b, b^*) := \sup_{\tau \in \mathcal{T}(t)} \mathbb{E}^{(t,b,b^*)}[\tau^{\frac{1}{2}} - CB^*(\tau)], \tag{2}$$

where $\mathcal{T}(t)$ denotes the set of bounded stopping times $\tau \geqslant t$ and $\mathbb{E}^{(t,b,b^*)}$ denotes the expectation conditionally on starting the Brownian motion B at time t with the values $B_t = b, B_t^* = b^*$. The domain of definition of V is

$$D = \{(t, b, b^*) : \quad 0 \leqslant t < \infty, \quad 0 \leqslant |b| \leqslant b^* < \infty\}. \tag{3}$$

Equivalently we can write

$$V(t, b, b^*) := \sup_{\tau \in \mathcal{T}} \mathbb{E}[\sqrt{t+\tau} - C(b^* \vee (b + B(\tau))^*], \tag{4}$$

which follows from the strong Markov property and stationarity of increments of Brownian motion.

Denote by $\widehat{C}$ the infimum of $C > 0$ such that (1) holds true. Clearly $\widehat{C}$ still satisfies (1). If $C < \widehat{C}$ then $V(t, b, b^*) \equiv \infty$, otherwise we have:

Lemma 2.1. *Let $C \geqslant \widehat{C}$ then V defined via (2) is*

(i) continuous,
(ii) finite-valued, and
(iii) $t \mapsto V(t, b, b^*) - \sqrt{t}$ is decreasing for fixed $|b| \leqslant b^*$.

In particular (ii) follows from the bounds

$$\sqrt{t} + Cb^* \geqslant V(t, b, b^*) \geqslant \sqrt{t} - Cb^*. \tag{5}$$

Proof. The lower bound of V follows from choosing the stopping time $t \in \mathcal{T}(t)$. For the upper bound observe that we can estimate for an arbitrary $\tau \in \mathcal{T}$

$$\begin{aligned} \sqrt{t+\tau} - C(b^* \vee (b + B(\tau))^* &\leqslant \sqrt{t} + \sqrt{\tau} - C(b + B(\tau))^* \\ &\leqslant \sqrt{t} + \sqrt{\tau} - CB(\tau)^* + C|b| \\ &\leqslant \sqrt{t} + Cb^* + \sqrt{\tau} - CB(\tau)^*. \end{aligned}$$

Taking expectations we get the upper bound for V from the representation in (4).

Next observe that for $\tau \in \mathcal{T}$ we have for $t < t'$

$$\sqrt{t+\tau} - C(b^* \vee (b + B(\tau))^* - \sqrt{t} \leqslant \sqrt{t'+\tau} - C(b^* \vee (b + B(\tau))^* - \sqrt{t'}$$

by concavity of the square root. Now (iii) follows by taking expectations suprema.

For (i), please refer to Sections 7 and 9.2 in [14]. □

To exclude the trivial case, we assume in the sequel that $C \geqslant \widehat{C}$.

For fixed C, the *stopping region* $S \subseteq D$ and the *non-stopping region* $NS \subseteq D$ are defined by

$$S := \{V = t^{\frac{1}{2}} - Cb^*\}, \quad NS := \{V > t^{\frac{1}{2}} - Cb^*\}. \tag{6}$$

To characterize the stopping region S first note that it is certainly not a good idea to stop when $|B(t)| < B^*(t)$.

Lemma 2.2. *Let* $(t, b, b^*) \in D$ *with* $|b| < b^*$. *Then* $(t, b, b^*) \in NS$.

Proof. Consider the first exit time of the interval $[-b^*, b^*]$. ☐

Next we observe a useful scaling property of V (compare Burkholder [7]).

Lemma 2.3. *For* $a > 0$ *and* $(t, b, b^*) \in D$, *we have*

$$V(a^2 t, ab, ab^*) = aV(t, b, b^*). \tag{7}$$

Proof. This follows directly from the scaling property of Brownian motion: if $(B_t)_{t \geqslant 0}$ is a standard Brownian motion, then $(a^{-1}B_{a^2 t})_{t \geqslant 0}$ again is a standard Brownian motion. Also, a random time $a^2\tau$ is a stopping time for the first process if and only if τ is a stopping time for the second process. ☐

This allows us to derive the following Lemma where the first part is a direct consequence of Lemmas 2.1 (iii) and 2.3 and the second part is technical and deferred to the appendix in Lemma B.3.

Lemma 2.4. *Let* $0 \leqslant t \leqslant t'$ *and* $b \in \mathbb{R}$. *Then* $(t, b, |b|) \in S$ *implies* $(t', b, |b|) \in S$.

Hence, for fixed $C \geqslant \widehat{C}$ *there is a smallest* $t_0 \in [0, \infty]$ *such that* $(t, b, |b|) \in S$ *if and only if*

$$\frac{t}{|b|^2} \geqslant t_0. \tag{8}$$

In fact, we have $t_0 \in (0, \infty)$.

The next result is a standard result in optimal control theory and also intuitively rather obvious. Again, the proof is deferred to the appendix.

Lemma 2.5. *Suppose* $C \geqslant \widehat{C}$ *and let* (t, b, b^*) *be in the non-stop region* NS. *Consider a Brownian motion* $(B(u))_{t \leqslant u}$ *starting at time* t *conditionally on* $B(t) = b$ *and* $B^*(t) = b^*$. *Let* τ *be the first hitting time of the stopping region* S, *i.e.*

$$\tau = \inf\{u \geqslant t : (u, B(u), B^*(u)) \in S\} \tag{9}$$

Then the value process stopped at time τ

$$M(u)^\tau := V(u \wedge \tau, B(u \wedge \tau), B^*(u \wedge \tau)), \qquad u \geqslant t, \tag{10}$$

is a martingale.

The unstopped value process

$$M(u) := V(u, B(u), B^*(u)), \qquad u \geqslant t, \tag{11}$$

still is a supermartingale.

We conclude this section with a minor technical remark. In the above statement, as well as in most of the paper, we follow the usual language of optimal control theory to condition on the event $\{B(t) = b, B^*(t) = b^*\}$. As this is a null set under $\mathbb{P}$ this procedure needs some proper interpretation in order to make it rigorous.

Let us now introduce some notation to make this a bit clearer.

We denote by $(\mathcal{F}(u))_{u \geqslant 0}$ the (right-continuous, saturated) filtration generated by the Brownian motion $(B(u))_{u \geqslant 0}$. Of course, in definition (2) the stopping time $\tau \in \mathcal{T}(t)$ is understood with respect to this filtration. But it is clear from the Markov property that, for fixed (t, b, b^*), we may assume that $\tau \in \mathcal{T}(t)$ depends only on the behavior of the Brownian motion $(B(u))_{u \geqslant t}$ after time t and not on the previous behavior of $(B(u))_{0 \leqslant u \leqslant t}$ (except for the requirements $B(t) = b$ and $B^*(t) = b^*$).

To formalize this fact, we denote by $(\mathcal{G}^{(t)}(u))_{u \geqslant t}$ the (right-continuous, saturated) filtration generated by $(B(u) - B(t))_{u \geqslant t}$. A stopping time $\tau \in \mathcal{T}(t)$ (i.e., with respect to the filtration $(\mathcal{F}(u))_{u \geqslant 0}$) then may also be considered as a randomized stopping time with respect to the filtration $(\mathcal{G}^{(t)}(u))_{u \geqslant t}$, the randomization given by the trajectories of $(B(u))_{0 \leqslant u \leqslant t}$. As $(B(u))_{0 \leqslant u \leqslant t}$ is independent of the filtration $(\mathcal{G}^{(t)}(u))_{u \geqslant t}$, we conclude that the value of (2) does not change whether we optimize over the randomized or the non-randomized stopping times with respect to the filtration $(\mathcal{G}^{(t)}(u))_{u \geqslant t}$. For an introduction to the notion of randomized stopping times, please refer to [3]

The bottom line of these considerations is that we may assume w.l.o.g. in (2) that $\tau \in \mathcal{T}(t)$ is a stopping time with respect to the filtration $(\mathcal{G}^{(t)}(u))_{u \geqslant t}$.

Now, the statement Lemma 2.5 could be rephrased without referring to conditioning on a null set, by noting that τ is a stopping time with respect to the filtration $(\mathcal{G}^{(t)}(u))_{t \leqslant u}$.

All other statements in the paper referring to conditioning on the values $B(t)$ and $B^*(t)$ could be made rigorous in an analogous way if the reader insists, but we do not further elaborate on these technicalities.

3. The Value Function from an Analytic Perspective

Again fix $C \geqslant \widehat{C}$. Differentiating the scaling equation (7) with respect to a and setting $a = 1$ we obtain, at least formally, the PDE

$$2tV_t + bV_b + b^* V_{b*} = V. \tag{12}$$

The optimal constant C for inequality (1) will be determined by analyzing whether this PDE has a reasonable solution for given $C > 0$ or not.

We need some preparation. For $0 < h < 1$ we denote by $f^h(s)$ the density of the distribution of the stopping time $\rho^h = \inf\{t : |B(t)| = 1\}$, where B is a Brownian motion starting at $B(0) = 1 - h$.

Define

$$g(s) = \lim_{h \searrow 0} \frac{f^h(s)}{h}, \qquad s > 0.$$

It is well-known (e.g. [11, Exercise 2.2.8.11]) that there is an explicit representation of $f^h(s)$ as an infinite sum. By differentiation of each summand we obtain an explicit infinite sum representation also for $g(s)$ (see the appendix below).

The function g appears in the formulation of the subsequent lemma which will turn out to be of crucial relevance for our analysis.

Lemma 3.1. *Let $W : D \to \mathbb{R}^+$ be a continuous function such that*

(a) $b^* \mapsto W(t, b, b^*)$ is Lipschitz continuous and
(b) $W(t, 1, 1) - \sqrt{t}$ is decreasing.

Furthermore let $S \subseteq D$ be defined by $S_W := \{(t, b, |b|) \in D : t/b^2 \geqslant t_W\}$ for some fixed t_W. Consider a standard Brownian Motion $B(t)$ and define τ_W to be the first hitting time of S_W. Suppose that $X(t) := W(t, B(t), B^*(t))$ is a supermartingale and $X(t \wedge \tau_W)$ is a martingale. Further assume that the process $X(t \wedge \tau_W \wedge \sigma^h)$ is uniformly integrable where σ^h is given by $\sigma^h := \inf\{s \geqslant t : |B(s)| = 1 + h\}$.

Then,

(i)

$$W_{b*}(t, 1, 1) := \lim_{h \searrow 0} \frac{1}{h}[W(t, 1, 1 + h) - W(t, 1, 1)] = 0 \quad \text{for } 0 < t < t_W \tag{13}$$

(ii)

$$\begin{aligned} W_b(t, 1, 1) &:= \lim_{h \searrow 0} \frac{1}{h}[W(t, 1, 1) - W(t, 1 - h, 1] \\ &= -\int_0^\infty [W(t + s, 1, 1) - W(t, 1, 1)]g(s)ds \end{aligned} \tag{14}$$

Observe that (i) in the above Lemma would follow directly from Ito's formula if we assume that W is sufficiently differentiable by considering

$$dX(t) = (W_t + \tfrac{1}{2}W_{bb})dt + W_b dB(t) + W_{b*} dB^*(t) \tag{15}$$

which is the increment of a martingale. The process $dB^*(t)$ is non-decreasing and its variation is a.s. singular with respect to Lebesgue measure. A necessary condition for $(W(t, B(t), B^*(t)))_{\tau_W \geqslant t}$ to be a martingale therefore is that W_{b*} vanishes a.s. with respect to the variation measure of dB^*. This indicates that $W_{b*}(t, b, b^*) = 0$ should hold true whenever $|b| = b^*$ and (t, b, b^*) is in the non-stop region NS. In particular, we should have $W_{b*}(t, 1, 1) = 0$, for $t < t_W$.

Proof of 3.1. (i) For $h > 0$ as in (13) define, conditionally on $B(t) = 1$ and $B^*(t) = 1$, the stopping times

$$\sigma^h := \inf\{u \geqslant t : |B(u)| = 1 + h\} \qquad \text{and} \tag{16}$$

$$\tau^h := \sigma^h \wedge \tau_W. \tag{17}$$

Recall that the random variable τ_W is a stopping time with respect to the filtration $(\mathcal{G}^{(t)}(u))_{t \leqslant u}$. Note that the process $(u, B(u), B^*(u))_{u \geqslant t}$, starting at $B(t) = 1$ and $B^*(t) = 1 + h$, also remains in $D \backslash S_W$ up to the stopping time τ_W. To see this, we distinguish two cases:

(i) $t \leqslant u < \tau^h$: Here we have $B(u) < 1+h = B^*(u)$ and thus $(u, B(u), B^*(u)) \in D\backslash S_W$.

(ii) $\tau^h \leqslant u < \tau_W$: In this case $B^*(u)$ is the same for $B^*(t) = 1$ and $B^*(t) = 1+h$.

As $t < t_0$ we have that τ is a.s. strictly positive. This implies that

$$\lim_{h \searrow 0} \mathbb{P}[\sigma^h = \tau^h] = 1, \tag{18}$$

as $\lim_{h \searrow 0} \sigma^h = t$, almost surely.

We may write

$$\begin{aligned} &W(t,1,1+h) - W(t,1,1) \\ &\quad = \mathbb{E}^{(t,1,1+h)}[W(\tau^h, B(\tau^h), B^*(\tau^h))] - \mathbb{E}^{(t,1,1)}[W(\tau^h, B(\tau^h), B^*(\tau^h))] \\ &\quad = \mathbb{E}^{(t,1,1+h)}[W(\tau^h, B(\tau^h), 1+h)] - \mathbb{E}^{(t,1,1)}[W(\tau^h, B(\tau^h), B^*(\tau^h))]. \end{aligned} \tag{19}$$

Here we use that $B^*(\tau^h) = 1+h$ conditionally on $B^*(t) = 1+h$ as $\tau^h \leqslant \sigma^h$ by definition. Furthermore we make use of the martingale property of $W(t, B(t), B^*(t))$ and the optional stopping theorem. For the use of the latter we need the assumption that $(W(u \wedge \tau^h, B(u \wedge \tau^h), B^*(u \wedge \tau^h)))_{u \geqslant t}$ is uniformly integrable.

On the set $\{\sigma^h = \tau^h\}$ we have $B^*(\tau^h) = B^*(\sigma^h) = 1+h$ for both initial conditions $B(t) = 1, B^*(t) = 1+h$ and $B(t) = 1, B^*(t) = 1$. Therefore, the value $W(\tau^h, B(\tau^h), B^*(\tau^h))$ is the same under both initial conditions. It follows that

$$\begin{aligned} &\mathbb{E}^{(t,1,1+h)}[W(\tau^h, B(\tau^h), 1+h)\mathbb{1}_{\{\sigma^h = \tau^h\}}] \\ &\qquad - \mathbb{E}^{(t,1,1)}[W(\tau^h, B(\tau^h), B^*(\tau^h))\mathbb{1}_{\{\sigma^h = \tau^h\}}] = 0. \end{aligned}$$

On the remaining set $\{\tau^h < \sigma^h\}$ we have $B^*(\tau^h) \in [1, 1+h]$. Because W is Lipschitz continuous in the variable b^* with some constant L we may estimate

$$\begin{aligned} &\mathbb{E}^{(t,1,1+h)}[W(\tau^h, B(\tau^h), 1+h)\mathbb{1}_{\{\tau^h < \sigma^h\}}] \\ &\qquad - \mathbb{E}^{(t,1,1)}[W(\tau^h, B(\tau^h), B^*(\tau^h)\mathbb{1}_{\{\tau^h < \sigma^h\}}] \leqslant Lh\mathbb{P}[\tau^h < \sigma^h]. \end{aligned}$$

Dividing (19) by h and passing to the limit we obtain from (18) that

$$W_{b^*}(t,1,1) := \lim_{h \searrow 0} \frac{1}{h}[W(t,1,1+h) - W(t,1,1)] = 0$$

(ii) As X is a martingale before hitting S_W we have for ρ^h as above that

$$W(t, 1-h, 1) = \mathbb{E}^{(t,1-h,1)}[W(\rho^h, 1, 1)] = \int_0^\infty W(t+s, 1, 1) f^h(s) ds,$$

where the density f^h is given by

$$f^h(s) := \frac{1}{\sqrt{2\pi s^3}} \sum_{n=-\infty}^{\infty} \left[(4n+h) e^{-\frac{(4n+h)^2}{2s}} + (4n+2-h) e^{-\frac{(4n+2-h)^2}{2s}} \right]. \tag{20}$$

We can use this relation to calculate the derivative w.r.t. the second component:

$$-W_b(t,1,1) := \lim_{h\searrow 0}\frac{1}{h}\int_0^\infty [W(t+s,1,1)-W(t,1,1)]f^h(s)ds$$

We split this integral into two parts at some point $\alpha > 0$ and observe that $f^h(s)$ is continuous and $f^h(s) \searrow 0$ for $h \searrow 0$ pointwise at $s > 0$ and thus by Dini's Theorem also uniformly (monotone) on any interval $[\alpha, K]$, for $0 < h < h_0(\alpha)$. Therefore we have

$$\begin{aligned}\lim_{h\searrow 0}\frac{1}{h}\int_\alpha^K &[W(t+s,1,1)-W(t,1,1)]f^h(s)ds \\ &= \int_\alpha^K [W(t+s,1,1)-W(t,1,1)]g(s)ds\end{aligned} \tag{21}$$

for g given by

$$g(s) := \lim_{h\searrow 0}\frac{f^h(s)}{h} = \frac{1}{\sqrt{2\pi s^3}}\left[1+2\cdot\sum_{n=1}^\infty(-1)^n\left(1-\frac{(2n)^2}{s}\right)e^{-\frac{(2n)^2}{2s}}\right]. \tag{22}$$

As before, because $u \mapsto W(u,1,1)-\sqrt{u}$ is decreasing and $s \mapsto \sqrt{t+s}-\sqrt{t}$ is concave, we have $W(t+s,1,1)-W(t,1,1) \leqslant \sqrt{t+s}-\sqrt{t} \leqslant s\frac{1}{2\sqrt{t}}$. Therefore the integrand is dominated by $\frac{sf^h(s)}{2\sqrt{t}h}$.

For $K > 0$ we can then estimate

$$\begin{aligned}\frac{1}{h}\int_K^\infty sf^h(s)ds &= \frac{1}{h}\mathbb{P}[\rho^h > K]\mathbb{E}[(B_{\rho^h}-B_0)^2|\rho^h > K] \\ &= \mathbb{P}[\rho^h > K](2-h) \leqslant 2\mathbb{P}[\rho^h > K].\end{aligned}$$

This probability tends to 0 uniformly in K, thus the integrals over $[K,\infty)$ can be neglected and we can replace K by ∞ in (21).

For the other part of the integral we first observe that

$$\begin{aligned}&\lim_{h\searrow 0}\frac{1}{h}\int_0^\alpha sf^h(s)ds \\ &\quad= \lim_{h\searrow 0}\frac{1}{h}\int_0^\infty sf^h(s)ds - \frac{1}{h}\int_\alpha^\infty sf^h(s)ds \\ &\quad= \lim_{h\searrow 0}\frac{1}{h}\mathbb{E}[\rho^h] - \frac{1}{h}\int_\alpha^\infty sf^h(s)ds \\ &\quad= \lim_{h\searrow 0} 2-h-\frac{1}{h}\int_\alpha^\infty sf^h(s)ds = 2-\int_\alpha^\infty s\,g(s)ds.\end{aligned}$$

The last integral converges to 2 for $\alpha \to 0$ by monotone convergence.

We conclude by setting $M \geqslant |W_t(u,1,1)|$ for $u \in [t, t+\alpha]$, making the estimate

$$\left|\frac{1}{h}\int_0^\alpha [W(t+s,1,1)-W(t,1,1)]f^h(s)ds\right| \leqslant \frac{M}{h}\int_0^\alpha sf^h(s)ds$$
$$\overset{h\to 0}{\to} M\left(2-\int_\alpha^\infty sg(s)ds\right),$$

and then taking the limit for $\alpha \to 0$. ☐

We can now apply this technical Lemma to our value function V by checking that the assumptions of the previous Lemma are satisfied by the value function V:

Lemma 3.2. *Let $V : D \to \mathbb{R}$ be the value function for* (2)*. Then,*

(i)

$$V_{b*}(t,1,1) := \lim_{h\searrow 0}\frac{1}{h}[V(t,1,1+h)-V(t,1,1)] = 0 \qquad \text{for } 0<t<t_0 \tag{23}$$

(ii)

$$\begin{aligned} V_b(t,1,1) &:= \lim_{h\searrow 0}\frac{1}{h}[V(t,1,1)-V(t,1-h,1] \\ &= -\int_0^\infty [V(t+s,1,1)-V(t,1,1)]g(s)ds \end{aligned} \tag{24}$$

Proof. V is continuous by Lemma 2.1 and Lipschitz-continuous in b^* by definition. We also have $V(t,1,1)-\sqrt{t}$ is decreasing by Lemma 2.1. Furthermore $V(t,B(t),B^*(t))$ is a martingale up to hitting S by Lemma 2.5. Thus, setting $W=V$, $S_W=S$ and $t_W=t_0$, it remains to check the required uniform integrability condition. We have

$$\begin{aligned} 0 \leqslant & V(u\wedge\tau^h, B(u\wedge\tau^h), B^*(u\wedge\tau^h)) \\ & \leqslant V(t, B(u\wedge\tau^h), B^*(u\wedge\tau^h)) + (\sqrt{u\wedge\tau^h}-\sqrt{t}) \\ & \leqslant V(t,0,1+h)+\sqrt{\tau^h} \leqslant V(t,0,1+h)+\frac{1}{4}+\tau^h. \end{aligned}$$

where the first estimate follows because the function $u \mapsto V(u,b,b^*)-\sqrt{u}$ is decreasing in u. The second inequality is due to the fact that V is decreasing in $|b|$ and increasing in b^* as well as $B^*(u\wedge\tau^h)\leqslant 1+h$. Now, $\tau^h \leqslant \sigma^h$ and σ^h has exponential moments and is therefore integrable which yields the desired uniform integrability.

Observe that τ^h is smaller than the first hitting time of the non-stop region NS no matter whether we condition on $(t,1,1)$ or $(t,1,1+h)$ which warrants the use of Lemma 2.5. ☐

The subsequent lemma shows that, for $t>t_0$, the behavior of V_b and V_{b*} follows a different pattern than the one given by Lemma 3.2. We find that

$$V_{b*}(t,1,1)+V_b(t,1,1) = -C \tag{25}$$

where we have to interpret this equation properly.

Lemma 3.3. *For $t > t_0$ we have*

$$\lim_{h \searrow 0} \frac{1}{h}[V(t, 1+h, 1+h) - V(t, 1, 1)] = -C. \tag{26}$$

Proof. For $t > t_0$ we have

$$V(t, b^*, b^*) = t^{\frac{1}{2}} - Cb^*,$$

for b^* in a neighbourhood of 1. □

To abbreviate notation we shall sometimes denote by $V(t)$ the function $V(t, 1, 1)$ (recall that we keep $C \geqslant \widehat{C}$ and the corresponding $t_0 = t_0(C)$ fixed). We thus obtain the following integro-differential equation for $V(t)$.

Lemma 3.4. *The function $V(t)$ satisfies the following equations*

$$2tV'(t) = V(t) + C, \qquad t > t_0, \tag{27}$$

$$2tV'(t) = V(t) + \int_0^\infty [V(t+s) - V(t)]g(s)ds, \qquad t < t_0. \tag{28}$$

Proof. The first assertion is obvious, as we have

$$V(t) = t^{\frac{1}{2}} - C, \quad \text{for} \quad t > t_0. \tag{29}$$

The second equation follows, at least formally, from (12), (23) and (24).

To justify (28) in a more pedantic way, note that for $a > 1$ we obtain from (7)

$$\begin{aligned} aV(t,1,1) - V(t,1,1) &= V(a^2t, a, a) - V(t,1,1) \\ &= (V(a^2t, a, a) - V(t, a, a)) \\ &+ (V(t, a, a) - V(t, 1, a)) \\ &+ (V(t, 1, a) - V(t, 1, 1)). \end{aligned}$$

Dividing by $a - 1$ and letting a decrease to 1, we deduce (28) from Lemma 3.2. □

Let us discuss the behaviour of the function $V(t)$ at $t = t_0$. As observed in the previous section, $V(t)$ is continuous so that we must have "continuous pasting" at t_0. It is the immediate reflex – at least it was so for the present authors – to expect *smooth pasting* of $V(t)$ at $t = t_0$ i.e. $\lim_{t \searrow t_0} V'(t) = \lim_{t \nearrow t_0} V'(t)$. By (27) and (28) this would result in determining t_0 by equating C with $\int_0^\infty (V(t_0 + s) - V(t_0))g(s)ds$. To our big surprise this turned out *not* to be the case; after some time of reconsidering we had to conclude that there is little reason why the smooth pasting principle should prevail in the present context. Here is one intuitive reason: for a fixed number $t_0 > 0$ we have that for almost all trajectories of a Brownian motion $B = (B(t))_{t \geqslant 0}$, starting at $B(0) = 0$, there is no $t > 0$ such that the two equalities $|B(t)| = B^*(t) = (\frac{t}{t_0})^{\frac{1}{2}}$ are simultaneously verified. By Lemma 2.4 we conclude that a discontinuity of the derivatives of $V(t, B(t), B^*(t))$ can only take place where these two equations are simultaneously satisfied. Roughly speaking: the Brownian motion B "*does not see*" a kink of the function $V(t)$ at $t = t_0$.

As a matter of fact, this natural example of a case of *non-smooth pasting* in the case of continuous martingales seems to us a remarkable feature of the present paper. The literature on *non-smooth pasting* is generally revolving around non-continuous processes. Some examples of non-smooth pasting for processes with jumps can be found in [1], [2], [6], [9] and [14].

4. The Integro-Differential-Equation

Fix the parameters $C > 0$ and $t_0 > 0$. We consider the ordinary integro-differential equation for the function $U = U^{C,t_0}$

$$U(t) = t^{\frac{1}{2}} - C, \qquad t \geqslant t_0, \tag{30}$$

$$2tU'(t) = U(t) + \int_0^\infty [U(t+s) - U(t)]g(s)ds, \qquad 0 < t \leqslant t_0, \tag{31}$$

where g is given by (22).

Here the fixed behaviour (30) of $U(t)$, for $t \geqslant t_0$, is considered as the initial condition, and subsequently the OIDE (ordinary integro-differential equation) (31) is solved by letting t decrease from t_0 to 0. For $t = t_0$, the derivative $U'(t_0)$ in (31) is understood as the *left* limit of $U'(t)$ when t increases to t_0.

It is standard to verify that, for $C > 0, t_0 > 0$, and $\epsilon > 0$ the solution U^{C,t_0} of (30) is well-defined for $t \in [\epsilon, \infty)$ and depends smoothly on the parameters C and t_0. On the other hand, the $2t$ term on the left hand side of (31) indicates that only for special cases of C and t_0 this solution can be extended to a continuous and finitely valued function $U(t)$, defined for all $t \in [0, \infty)$.

The evidence resulting from our numerical analysis of the solutions $(U^{C,t_0}(t))_{t>0}$, in dependence of C and t_0, can be resumed as follows:

Numerical Evidence 4.1.

(i) *There is a smallest number* $\check{C} > 0$ *as well as a unique number* $\check{t}_0 > 0$ *such that the OIDE (31) admits a solution* $U^{\check{C},\check{t}_0}(t)$ *which has a finite limit* $\lim_{t\to 0} U^{\check{C},\check{t}_0}(t)$. *This solution is monotone increasing and* $U^{\check{C},\check{t}_0}(t) \geqslant t^{\frac{1}{2}} - \check{C}$.

(ii) *For* $0 < C < \check{C}$ *there does not exist* $t_0 > 0$ *such that* $U(t) = U^{C,t_0}(t)$ *remains bounded from below as* $t \searrow 0$.

(iii) *For* $C > \check{C}$ *there are precisely two values* $t_1 = t_1(C), t_2 = t_2(C)$ *depending on* C *in a continuous and one-to-one way such that* $U^{C,t_i}(t)$ *has a finite limit, as* $t \searrow 0$, *for* $i = 1, 2$. *These solutions satisfy* $U^{C,t_i}(t) \geqslant t^{\frac{1}{2}} - C$.
For $t_0 \in (t_1, t_2)$, *the solutions of the OIDE (31) tend to* $+\infty$, *for* $t \searrow 0$, *while, for* $t_0 \notin [t_1, t_2]$, *the solutions tend to* $-\infty$, *for* $t \searrow 0$. *We have* $t_1(C) < \check{t}_0 < t_2(C)$ *and* $\lim_{C\to\check{C}} t_1(C) = \lim_{C\to\check{C}} t_2(C) = \check{t}_0$.

The functions U^{C,t_1} *and* U^{C,t_2} *are monotone increasing and*

$$U^{C,t_1}(t) < U^{\check{C},\check{t}_0}(t)\,, \quad \textit{for all} \quad t \in [0,\infty[. \tag{32}$$

Finally, we find the numerical values

$$\check{C} \approx 1.27267\ldots \quad \textit{and} \quad \check{t}_0 \approx 0.9036\ldots \tag{33}$$

We have not been able to provide a mathematically rigorous proof of the above assertions and only rely on the numerical evidence (which is based on Euler-type simulations in Python with variable step sizes). We therefore consider the above statements rather as hypotheses underlying our subsequent results and we shall carefully point out in the subsequent statements where we rely on this evidence.

For example, for $C = 1.25 < \check{C}$ which is case (ii) above we illustrate the situation by Figure 1.

FIG 1. *The subcritical case* $C < \check{C}$*: Numerical solutions for* $C = 1.25$ *and various values for the pasting position* t_0 *in the interval* $[0.8, 1]$*. The graph underneath is* $t^{1/2} - C$ *through which all solutions cut in the subcritical case when they get close to* 0.

For $C = 1.274 > \check{C}$, which is case (iii) above, we find $t_1 \approx 0.85\ldots$ and $t_2 \approx 0.95\ldots$, as illustrated in Figure 2.

When C decreases to the critical value $\check{C} \approx 1.27267\ldots$ the numerics suggest that the length $t_2 - t_1$ of the intervals (t_1, t_2) decreases to zero and that these intervals shrink to a single point $\check{t}_0 \in]0,\infty[$ for which we find $\check{t}_0 \approx 0.9036\ldots$. It is convincing from the numerics that the limiting solution $U^{\check{C},\check{t}_0}(t)$ then is well-defined for all $t \geqslant 0$ by letting $U^{\check{C},\check{t}_0}(0) := \lim_{t\searrow 0} U^{\check{C},\check{t}_0}(t)$. This function

Fig 2. *The supercritical case* $C > \check{C}$*: Numerical solutions for* $C = 1.274$ *and and various values for the pasting position* t_0 *in the interval* $[0.6, 1.2]$*. We find that while solutions for* t_0 *in an interval* (t_1, t_2) *with* $t_1 \approx 0.85$ *and* $t_2 \approx 0.95$ *stay above the graph of* $t^{1/2} - C$ *(in fact, they tend to* $+\infty$*, as* $t \to 0$*), the solutions fall to* $-\infty$ *for* $t \to 0$ *if* $t_0 \notin [t_1, t_2]$*. At the transition of these two regimes lie the bounded solutions* U^{C,t_1} *and* U^{C,t_2} *(which are not explicitly displayed in the above figure but are squeezed between the neighbouring solutions* $U^{C,t}$*).*

$U^{\check{C},\check{t}_0}(t)$ is monotone increasing and such that $U^{\check{C},\check{t}_0}(t)$ lies above the function $U^{\check{C},0}(t) = t^{\frac{1}{2}} - \check{C}$. Clearly we expect that $U^{\check{C},\check{t}_0}$ must be the "right" solution, which may be identified with the value function V defined in (2) for the optimal constant $\widehat{C} = \check{C}$ and, in particular, that $\check{C} \approx 1.27267\ldots$ equals the optimal constant $\widehat{C}$ in the Burkholder-Davis-Gundy inequality (1). We shall subsequently deduce this result more formally.

5. Identifying the Value-Function

Admitting the Numerical Evidence 4.1 we shall show that the function $U^{\check{C},\check{t}_0}(t)$, obtained above from the analysis of the OIDE (31), indeed determines the value function $V(t, b, b^*)$ as defined in (2) for the constant $\widehat{C} = \check{C}$ and that this constant is indeed the optimal Burkholder-Davis-Gundy constant in inequality (1).

Starting from a solution $U(t) = U^{C,t_0}(t)$ of the OIDE (31) for parameters $C > 0$ and $t_0 > 0$ such that $U(t)$ extends continuously to a finite value $U(0)$ we may extend this solution (by slight abuse of notation) to a function $U(t, b, b^*)$, defined on D, by first letting

$$U(t, b, 1) := \int_0^\infty U(t + s) f^{1-|b|}(s) ds, \quad \text{for} \quad 0 \leqslant |b| < 1, \tag{34}$$

where $f^h(s)$ is defined in Section 3. For general $(t,b,b^*) \in D$ we use (7) to define

$$U(t,b,b^*) = b^* U\Big(\frac{t}{(b^*)^2}, \frac{b}{b^*}, 1\Big). \tag{35}$$

For later reference, we note that $U(t) \geqslant \sqrt{t} - C$ implies

$$U(t,b,b^*) \geqslant \sqrt{t} - Cb^*, \text{for all } (t,b,b^*) \in D \tag{36}$$

Lemma 5.1. *Fix $C > 0$ and $t_0 > 0$ such that $U(t)$ extends continuously to a finite $U(0)$ and admit the Numerical Evidence 4.1 (i) and (iii). Let $(B(u))_{u \geqslant t}$ be a Brownian motion starting at some time $t > 0$ at $B(t) = b$ and $B^*(t) = b^*$. The process $U(u, B(u), B^*(u))_{u \geqslant t}$ is then a local super-martingale. It is a local martingale up to entering the stopping area $S := \{(t,b,b^*) : |b| = b^*, t/(b^*)^2 \geqslant t_0\}$.*

For the proof we need the following Lemma to justify the use of Ito's formula for a function that is not smooth everywhere but where the Brownian Motion hardly ever touches the set where it is not differentiable.

Lemma 5.2. *Let $W : D \to \mathbb{R}$ be a continuous function and $t_W > 0$ such that*

(a) the derivatives W_t and W_{bb} exist and are continuous on the interior of D, and
(b) $W_t + \frac{1}{2}W_{bb} = 0$,
(c) $W_{b*} \leqslant 0$, and
(d) $W_{b*}(t,b,|b|) = 0$ for $\frac{t}{b^2} < t_W$.

Define $S_W := \{(t,b,|b|) \in D : \frac{t}{b^2} \geqslant t_W\}$. For a standard Brownian Motion $B(t)$ let τ_W be the first hitting time of S for $(t, B(t), B^*(t))$. Then,

(i) $X(t) := W(t, B(t), B^*(t))$ is a local supermartingale, and
(ii) $X(t \wedge \tau_W)$ is a local martingale.

Proof. This follows at least formally from the assumptions and Ito's formula

$$dX(t) = (W_t + \tfrac{1}{2}W_{bb})dt + W_b dB(t) + W_{b*} dB^*(t). \tag{37}$$

To address this in a more formal way, let $\epsilon > 0$ and define the stopping times $(\rho_n^\epsilon)_{n=0}^\infty$ by $\rho_0^\epsilon = 0$ and

$$\rho_n^\epsilon = \inf\{t : t \geqslant \rho_{n-1}^\epsilon + \epsilon \quad \text{and} \quad |B(t)| = B^*(t)\}.$$

We also denote by A^ϵ the union $\bigcup_{n=0}^\infty [\![\rho_n^\epsilon, \rho_n^\epsilon + \epsilon]\!]$ which is a predictable subset of $\Omega \times \mathbb{R}_+$. Denoting by $A^0 = \bigcap_{n=0}^\infty A^{\frac{1}{n}}$, the set A^0 simply equals $\{|B(t)| = B^*(t)\}$. Fixing $T > 0$, the Lebesgue-measure of $\{\omega\} \times [0,T] \cap A^\epsilon$ tends to zero, for almost all $\omega \in \Omega$.

Fix a bounded stopping time τ such that $(B(t))_{0 \leqslant t \leqslant \tau}$ remains bounded. It follows that W_t as well as W_{bb} also remain bounded on $A^\epsilon \cap [\![0, \tau]\!]$ so that

$$X^\epsilon(t) := \int_0^t \mathbb{1}_{B^\epsilon} dX(s)$$

is a martingale and $B^\epsilon = (\Omega \times \mathbb{R}_+)\backslash A^\epsilon$ is the complement of A^ϵ. Indeed, it suffices to reason on the stochastic intervals $[\![\rho^\epsilon_{n-1} + \epsilon, \rho^\epsilon_n]\!]$ and to observe that $B^*(t)$ remains constant on these intervals.

Turning to the remaining part

$$Y^\epsilon(t) := \int_0^t \mathbb{1}_{A^\epsilon} dX(s) = X(t) - X^\epsilon(t)$$

we shall show that along a sequence these processes tend almost surely to the non-increasing process

$$\begin{aligned} Y^0(t) &= \int_0^t \mathbb{1}_{A^0} dX(s) \\ &= \int_0^t \mathbb{1}_{A^0} W_{b*}(t, B(t), B^*(t)) dB^*(t). \end{aligned}$$

Indeed, the dominated convergence theorem for Ito-Integrals yields convergence in probability and thus subsequence convergence almost surely. Fixing this sequence of ε's we can take the process to the appropriate limit. $\square$

Proof of Lemma 5.1. It follows from definition (34) that, for $(t, b, b^*) \in D$ such that $t > 0$ and $0 \leqslant |b| < b^*$, the heat equation

$$U_t + \frac{1}{2} U_{bb} = 0$$

is satisfied.

On the boundary, for $t < t_0$, we can apply the definition of U to obtain

$$\begin{aligned} U_b(t,1,1) + U_{b*}(t,1,1) &:= \lim_{h \searrow 0} \frac{1}{h}[U(t, 1+h, 1+h) - U(t,1,1)] \\ &= \lim_{h \searrow 0} \frac{1}{h}[U(t/(1+h)^2) - U(t)] + U(t/(1+h)^2) \\ &= -2tU'(t) + U(t) \\ &= -\int_0^\infty [U(t+s) - U(t)] g(s) ds \end{aligned}$$

The last equality is exactly the OIDE (31). Applying Lemma 3.1 (ii) one obtains that the last expression is equal to

$$U_b(t,1,1) := -\lim_{h \searrow 0} \frac{1}{h}[U(t, 1-h, 1) - U(t,1,1)],$$

where all the assumptions of this Lemma are easily checked. Clearly the left and right derivatives of U_b agree. It follows that $U_{b*}(t,1,1) = 0$ and more generally that $U_{b*}(t, b, |b|) = 0$ for $t/b^2 < t_0$. For $t \geqslant t_0$ we can derive as in Lemma 3.3 that

$$U_{b*}(t,1,1) = -U_b(t,1,1) - C = \int_0^\infty [(t+s)^{\frac{1}{2}} - t^{\frac{1}{2}}] g(s) ds - C.$$

This expression is monotone decreasing in t, and is necessarily non-positive at t_0 so that $U(t,1,1) \geqslant t^{\frac{1}{2}} - C$ holds. We conclude that, for arbitrary $t > 0$, $U_{b*} \leqslant 0$.

Having established that $U_{b*} \leqslant 0$ and $U_{b*}(t,b,|b|) = 0$ for $\frac{t}{b^2} < t_0$ we may derive, at least formally, the assertion of the present lemma from (57) and Ito's formula as in (15).

Now, we can conclude using Lemma 5.2 □

Let us now observe the following relations between value functions to the optimal stopping problem and solutions to the OIDE (31).

Lemma 5.3. *Let V^C be the value function as defined in (2) for a constant $C > 0$ that satisfies the inequality (1). Take $t_0 = t_0(C) \in (0,\infty)$ to be the corresponding point separating S from NS (see (6)). Then $V^C(t) := V^C(t,1,1)$ satisfies the OIDE (31) for this choice of C and t_0.*

Proof. For $t \geqslant t_0$, we have $V^C(t) = t^{\frac{1}{2}} - C$.

For $t < t_0$ denote by $\tau^h = \sigma^h \wedge \tau$ the stopping time as in (17) above, conditionally on $(t,1,1)$. Note that $(V^C(u,B(u),B^*(u)))_{t \leqslant u \leqslant \tau^h}$ then is a *uniformly integrable* martingale. To see this, we make a distinction for $u \leqslant \nu$ and $u > \nu$ where ν is the stopping time $\nu := \inf\{s : B^*(s)^2 t_0 \leqslant s\}$:

1. $u \leqslant \nu^h := \nu \wedge \tau^h$: The domain of $((u,B(u),B^*(u))$ is bounded by $|B(u)| \leqslant B^*(u) \leqslant 1+h$ and $u \leqslant \nu^h \leqslant B^*(\nu^h)^2 t_0 \leqslant (1+h)^2 t_0$. Clearly V^C is bounded on this domain.
2. $u > \nu^h$: Here the properties of V^C given in Lemma 2.5 allow us to see that one can rewrite the process as

$$V^C(u,B(u),B^*(u)) = \mathbb{E}^{(u,B(u),B^*(u))}[\sigma^{\frac{1}{2}}] - CB^*(u)$$

where σ is again the first time where $B(u) = B^*(u)$ holds. Note that this holds because $B^*(v)$ is now constant for $u \leqslant v \leqslant \sigma$, and it is clearly the definition of a uniformly integrable martingale, provided the conditional expectation is well defined, which it is by the estimate

$$\mathbb{E}^{(u,B(u),B^*(u))}[\sigma^{\frac{1}{2}}] \leqslant u^{\frac{1}{2}} + B^*(u).$$

Hence the formula

$$V^C(t,1,1) = \mathbb{E}^{(t,1,1)}[V^C(\tau^h,B(\tau^h),B^*(\tau^h)]$$

is justified. This was the only assumption of Lemma 3.1 which is not immediate from the definition of V^C. Now the claim follows from Lemma 3.1. □

Proposition 5.4. *Admitting the Numerical Evidence 4.1 (i) and (ii), the constant $\check{C}$ obtained in (33) equals the optimal constant $\widehat{C}$ for (1), and the function $U^{\check{C},\check{t}_0}(t,b,b^*)$ obtained in (34) and (35) equals the value function $V(t,b,b^*)$ as defined in (2) for the constant $\widehat{C} = \check{C}$.*

The value $\widehat{t}_0$ associated to $\widehat{C}$ by Lemma 2.4 equals the constant $\check{t}_0$ in (33).

Proof. To show $\widehat{C} \geqslant \check{C}$, suppose that $C > 0$ is a constant satisfying the Burkholder-Davis-Gundy inequality (1), i.e. suppose that $C \geqslant \widehat{C}$. Then by Lemma 5.3 we have that $V^C(t) := V^C(t,1,1)$ satifies the OIDE (31) for this choice of C and the corresponding $t_0(C)$ separating S from NS.

As $V^C(t)$ is increasing in t and satisfies $V^C(0) \geqslant -C$ we conclude from the *Numerical Evidence 4.1* (i) and (ii) that $C \geqslant \check{C}$. This yields $\widehat{C} \geqslant \check{C}$.

To show conversely that $\widehat{C} \leqslant \check{C}$ consider the function $U^{\check{C},\check{t}_0}(t,b,b^*)$. By Lemma 5.1 the process $(U^{\check{C},\check{t}_0}(t,B(t),B^*(t))_{t\geqslant 0}$ is a local supermartingale. Hence we have, conditionally on $(t,b,b^*) \in D$ and for each bounded stopping time $\tau \geqslant t$ and localizing sequence $(\tau_n)_{n=1}^{\infty}$.

$$\begin{aligned} U^{\check{C},\check{t}_0}(t,b,b^*) &\geqslant \mathbb{E}^{(t,b,b^*)}[U^{\check{C},\check{t}_0}(\tau\wedge\tau_n, B(\tau\wedge\tau_n), B^*(\tau\wedge\tau_n))] \\ &\geqslant \mathbb{E}^{(t,b,b^*)}[(\tau\wedge\tau_n)^{\frac{1}{2}} - \check{C}B^*(\tau\wedge\tau_n)] \end{aligned}$$

where the second inequality derives from (36). In the limit for $n \to \infty$ this yields $U^{\check{C},\check{t}_0}(t,b,b^*) \geqslant \mathbb{E}^{(t,b,b^*)}[\tau^{\frac{1}{2}} - \check{C}B^*(\tau)]$. Hence $U^{\check{C},\check{t}_0}(t,b,b^*)$ dominates the value function $V^{\check{C}}(t,b,b^*)$ as defined in (2) for the constant $\check{C}$. This shows $\check{C} \geqslant \widehat{C}$ as well as $U^{\check{C},\check{t}_0} \geqslant V^{\check{C}} = V^{\widehat{C}}$. ☐

We can finally summarize these results to proof the main theorem.

Proof of Theorem 1.2. Suppose that $C = \widehat{C}$ is the optimal constant for (1) and let $\widehat{t}_0 \in (0,\infty)$ the corresponding critical value given by Lemma 2.4. Then $V = V(t)$ satisfies the OIDE (30) and this solution is increasing in t and satisfies $V(0) \geqslant -C$.

As shown in the previous section there is a minimal C allowing for such a solution, for an appropriately chosen $t_0 \in (0,\infty)$. This value of C therefore must coincide with the optimal value $\widehat{C}$ for the Burkholder-Davis-Gundy inequality (1).

Conversely if C is chosen such that for some $t_0 \in (0,\infty)$ the OIDE (30),(31) has a solution on $[0,T]$ then the numerical evidence 4.1 (ii) gives that $C \geqslant \check{C} = \widehat{C}$ and so (1) holds. ☐

Remark 5.5. It is interesting to consider, for a fixed constant $C > \widehat{C}$, the relation between the value-function $V^C(t,b,b^*)$ defined in (2) and the corresponding solutions of the OIDE (31). In this case the numerical evidence 4.1 (iii) indicates that there are two bounded solutions $U^{C,t_1}(t)$ and $U^{C,t_2}(t)$. Which of the two is the “good one”, i.e. which one equals the value function $V^C(t,1,1)$?

To answer this question, first note that, for $C > \widehat{C}$, we clearly have the monotonicity relation $V^C(t,b,b^*) \leqslant V^{\widehat{C}}(t,b,b^*)$. It is also easy to see that $t_0(C) < t_0(\widehat{C}) = \widehat{t}_0 = \check{t}_0$, where $t_0(C)$ is associated to the value function $V^C(t,b,b^*)$ via Lemma 2.4. In other words, the stopping region S for the function $t^{\frac{1}{2}} - CB^*(t)$ in (2) is bigger than the stopping region S for the function $t^{\frac{1}{2}} - \widehat{C}B^*(t)$.

It follows from the numerical evidence that the value $t_1(C)$ for which we have $t_1(C) < \widehat{t}_0$ is the only candidate for the "good" solution while for $t_2(C)$ for which we have $t_2(C) > \widehat{t}_0$, we cannot have $U^{C,t_2(C)}(t) = V^C(t,1,1)$. We can conclude from Lemma 5.3 that the value function $V^C(t,1,1)$ indeed equals the solution $U^{C,t_1(C)}(t)$ of the OIDE (31).

The fact that $U^{C,t_2(C)}(t)$ cannot be the "good" solution has the following consequence which is interesting in its own right (compare [15]).

Proposition 5.6. *Admitting the Numerical Evidence 4.1 (iii) we have that, for* $t_2 > \widehat{t}_0$, *the stopping time*

$$\rho = \inf\left\{s \geqslant 1 : \frac{s}{(B^*(s))^2} \geqslant t_2\right\} \tag{38}$$

satisfies

$$\mathbb{E}[\rho^{\frac{1}{2}}] = \infty. \tag{39}$$

Proof. Define the stopping time τ by

$$\tau := \inf\left\{s \geqslant 1 : \frac{s}{B^*(s))^2} \geqslant t_2 \text{ and } |B(s)| = B^*(s)\right\}. \tag{40}$$

Clearly $\tau \geqslant \rho$, as we may equivalently define

$$\tau := \inf\{s \geqslant \rho : |B(s)| = B^*(s)\}. \tag{41}$$

We claim that $\mathbb{E}[\rho^{\frac{1}{2}}] < \infty$ if and only if $\mathbb{E}[\tau^{\frac{1}{2}}] < \infty$. Indeed, it follows from (41) that the law of $\tau - \rho$, conditionally on $(\rho, B(\rho), B^*(\rho))$, is that of the first hitting time σ of the level $B^*(\rho)$ by the absolute value of a Brownian motion $(W(u))_{u \geqslant 0}$ starting at $W(0) = B(\rho)$. We may (very crudely) estimate $\mathbb{E}[\sigma] \leqslant B^*(\rho)^2$.

Noting that at time ρ we have $B^*(\rho)^2 \leqslant \frac{\rho}{t_2}$ we may estimate

$$\mathbb{E}[\tau - \rho | \rho, B(\rho), B^*(\rho)] \leqslant \frac{\rho}{t_2}.$$

Hence we obtain

$$\begin{aligned} \mathbb{E}[\tau^{\frac{1}{2}}] &= \mathbb{E}[\rho^{\frac{1}{2}}] + \mathbb{E}[\mathbb{E}[\tau^{\frac{1}{2}} - \rho^{\frac{1}{2}} | \rho, B(\rho), B^*(\rho)]] \\ &\leqslant \mathbb{E}[\rho^{\frac{1}{2}}] + \mathbb{E}[\mathbb{E}[(\tau - \rho)^{\frac{1}{2}} | \rho, B(\rho), B^*(\rho)]] \\ &\leqslant \mathbb{E}[\rho^{\frac{1}{2}}] + \mathbb{E}[\mathbb{E}[\tau - \rho | \rho, B(\rho), B^*(\rho)]^{\frac{1}{2}}] \\ &\leqslant \left(1 + \frac{1}{\sqrt{t_2}}\right) \mathbb{E}[\rho^{\frac{1}{2}}], \end{aligned}$$

which readily shows that $\mathbb{E}[\rho^{\frac{1}{2}}] < \infty$. This implies that $\mathbb{E}[\tau^{\frac{1}{2}}] < \infty$.

So let us suppose that $\mathbb{E}[\tau^{\frac{1}{2}}] < \infty$ and work towards a contradiction.

Define the stopping region $S(t_2)$ relative to t_2 as

$$S(t_2) = \left\{(t, b, |b|) \in D : \frac{t}{b^2} \geqslant t_2\right\}.$$

and the corresponding non-stopping region by $NS(t_2) = D \backslash S(t_2)$.

We condition on some fixed $(1, b, b^*) \in NS(t_2)$. Note that τ is the first time when $(t, B(t), B^*(t))_{t \geqslant 1}$ leaves $NS(t_2)$.

Admitting the Numerical Evidence 4.1 (iii), associate to $t_2 > \widehat{t}_0$ the constant $C > \widehat{C}$ such that $U^{C,t_2}(t)$ is a solution of the OIDE (31) which remains bounded as $t \searrow 0$. We write $U^{C,t_2}(t, b, b^*)$ for its extension defined in (35). In contrast, we denote by $V^C(t, b, b^*)$ the value function as defined in (2) for the constant C.

The process $(U^{C,t_2}(t, B(t), B^*(t)))_{1 \leqslant t \leqslant \tau}$ is a local martingale by Lemma 5.1, where the present t_2 corresponds to t_0 in the statement of this lemma.

In addition we show that this local martingale is a uniformly integrable martingale up to time τ, i.e. the family of random variables $U^{C,t_2}(\sigma, B(\sigma), B^*(\sigma))$, where σ ranges in the stopping times $1 \leqslant \sigma \leqslant \tau$, is uniformly integrable. Recall the scaling relation

$$U^{C,t_2}(\sigma, B(\sigma), B^*(\sigma)) = B^*(\sigma) U^{C,t_2}\left(\frac{\sigma}{B^*(\sigma)^2}, \frac{B(\sigma)}{B^*(\sigma))}, 1\right),$$

and note that $\frac{\sigma}{B^*(\sigma)^2}$ remains in the interval $[0, t_2]$ so that, by compactness, the term $|U^{C,t_2}(\frac{\sigma}{B^*(\sigma)^2}, \frac{B(\sigma)}{B^*(\sigma))}, 1)|$ remains bounded by some constant $M > 0$. Therefore

$$|U^{C,t_2}(\sigma, B(\sigma), B^*(\sigma))| \leqslant M B^*(\sigma) \leqslant M B^*(\tau).$$

If $\mathbb{E}[\tau^{\frac{1}{2}}] < \infty$ we infer from the Burkholder-Davis-Gundy inequality (this time the reverse inequality to (1)) that the random variable $B^*(\tau)$ is integrable. Hence the family of random variables $U^{C,t_2}(\sigma, B(\sigma), B^*(\sigma))$ is dominated by the integrable random variable $MB^*(\tau)$ which shows that the local martingale $U^{C,t_2}(t, B(t), B^*(t))_{1 \leqslant t \leqslant \tau}$ is of class D and is thus a uniformly integrable martingale.

Hence, conditionally on $(1, b, b^*) \in NS(t_2)$ we obtain

$$U^{C,t_2}(1, b, b^*) = \mathbb{E}^{(1,b,b^*)}[U^{C,t_2}(\tau, B(\tau), B^*(\tau))]. \tag{42}$$

We now pass to the process $(V^C(t, B(t), B^*(t))_{t \geqslant 1}$ again conditionally on $(1, b, b^*) \in NS(t_2)$. By Lemma 2.5 we know that this process is a supermartingale. Repeating the above argument, we obtain that this supermartingale is uniformly integrable up to time τ. Hence

$$V^C(1, b, b^*) \geqslant \mathbb{E}^{(1,b,b^*)}[V^C(\tau, B(\tau), B^*(\tau))]. \tag{43}$$

Noting that at time τ we arrived in the stopping region $S(t_2)$ we obtain

$$U^{C,t_2}(\tau, B(\tau), B^*(\tau)) = V^C(\tau, B(\tau), B^*(\tau)) = \tau^{\frac{1}{2}} - CB^*(\tau).$$

Hence (42) and (43) yield

$$V^C(1, b, b^*) \geqslant U^{C,t_2}(1, b, b^*),$$

for all $(1, b, b^*) \in D$. As we have seen that $V^C = U^{C,t_1(C)} \leqslant U^{C,t_2(C)}$, and $U^{C,t_1(C)}$ is not equal to $U^{C,t_2(C)}$, we arrive at the desired contradiction. □

The above result is complemented by the following estimate in the reverse direction.

Proposition 5.7. *Admitting the Numerical Evidence 4.1 (iii), we have, for* $t_1 < \widehat{t}_0$, *that the stopping time*

$$\rho = \inf \left\{ s \geqslant 1 : \frac{s}{(B^*(s))^2} \geqslant t_1 \right\} \tag{44}$$

satisfies

$$\mathbb{E}[\rho^{\frac{1}{2}}] < \infty. \tag{45}$$

Proof. Similarly as in the proof of the previous proposition, we define

$$\tau := \inf\{s \geqslant 1 : \frac{s}{(B^*(s))^2} \geqslant t_1 \text{ and } |B(s)| = B^*(s)\}.$$

We shall show that

$$\mathbb{E}[\tau^{\frac{1}{2}}] < \infty \tag{46}$$

which will imply (45).

We condition on $(1, b, b^*)$ in the non-stop region $NS(t_1)$ defined as in the preceding proof, and will show

$$\mathbb{E}^{(1,b,b^*)}[\tau^{\frac{1}{2}}] \leqslant Kb^* \tag{47}$$

for some constant $K > 0$, which will imply (46) by integrating over the values $B(1) = b$ and $B^*(1) = b^*$.

We associate to $t_1 < \widehat{t}_0$ the corresponding $C > \widehat{C}$ such that the solution $U^{C,t_1}(t)$ of the OIDE (31) remains bounded (Numerical Evidence 4.1 (iii)).

Using the Numerical Evidence 4.1 (iii) there is some $\alpha > 0$ such that the solutions $U^{C,t_1}(t)$ of the OIDE (31) and the value function $V^{\widehat{C}}(t) = U^{\check{C},\check{t}_0}(t)$ for the optimal constant $\widehat{C} = \check{C}$ are separated by some $\alpha > 0$, i.e.

$$V^{\widehat{C}}(t) \geqslant U^{C,t_1}(t) + \alpha, \quad \text{for all} \quad t \geqslant 0. \tag{48}$$

Indeed, for $t \geqslant \widehat{t}_0$ we have $U^{C,t_1} = t^{\frac{1}{2}} - C$ and $U^{\widehat{C},\widehat{t}_0}(t) = t^{\frac{1}{2}} - \widehat{C}$. For $t \in [0, \widehat{t}_0]$ we have $U^{\widehat{C},\widehat{t}_0}(t) < U^{C,t_0}(t)$ by (32) so that by compactness we obtain a separating constant $\alpha > 0$.

More generally, we obtain from (34)

$$V^{\widehat{C}}(t, b, 1) \geqslant U^{C,t_1}(t, b, 1) + \alpha, \quad \text{for all} \quad (t, b, 1) \in D. \tag{49}$$

Similarly as in the above proof we consider, conditionally on $(1, b, b^*)$, the processes

$$(U^{C,t_1}(u, B(u), B^*(u)))_{1 \leqslant u \leqslant \tau} \text{ and } (V^{\widehat{C}}(u, B(u), B^*(u)))_{1 \leqslant u \leqslant \tau}.$$

Both are local martingales up to time τ. Let $(\tau_n)_{n=1}^{\infty}$ be a sequence of localizing, bounded stopping times, $\tau_n \geqslant 1$, increasing to τ.

$$\begin{aligned} V^{\widehat{C}}(1,b,b^*) &= \mathbb{E}^{(1,b,b^*)}[V^{\widehat{C}}(\tau_n, B(\tau_n), B^*(\tau_n)] \\ &= \mathbb{E}^{(1,b,b^*)}\left[B^*(\tau_n)V^{\widehat{C}}\left(\tfrac{\tau_n}{B^*(\tau_n)^2}, \tfrac{B(\tau_n)}{B^*(\tau_n)}, 1\right)\right] \\ &\geqslant \mathbb{E}^{(1,b,b^*)}\left[B^*(\tau_n)U^{C,t_1}\left(\tfrac{\tau_n}{B^*(\tau_n)^2}, \tfrac{B(\tau_n)}{B^*(\tau_n)}, 1\right)\right] + \alpha\mathbb{E}^{(1,b,b^*)}[B^*(\tau_n)] \\ &= U^{C,t_1}(1,b,b^*) + \alpha\mathbb{E}^{(1,b,b^*)}[B^*(\tau_n)]. \end{aligned}$$

Hence, letting $n \to \infty$, for each $(1,b,b^*)$

$$V^{\widehat{C}}(1,b,b^*) - U^{C,t_1}(1,b,b^*) \geqslant \alpha\mathbb{E}^{(1,b,b^*)}[B^*(\tau)].$$

Using the scaling relation again, we get

$$V^{\widehat{C}}\left(\frac{1}{(b^*)^2}, \frac{b}{b^*}, 1\right) - U^{C,t_1}\left(\frac{1}{(b^*)^2}, \frac{b}{b^*}, 1\right) \geqslant \frac{\alpha}{b^*}\mathbb{E}^{(1,b,b^*)}[B^*(\tau)]. \tag{50}$$

Observe that we can also find a bound β such that

$$\beta \geqslant V^{\widehat{C}}(t) - U^{C,t_1}(t), \quad \text{for all} \quad t \geqslant 0$$

as the difference equals $C - \widehat{C}$ for $t \geqslant \widehat{t_0}$ and is bounded for the compact interval $[0, \widehat{t_0}]$. This directly yields β as a bound on $V^{\widehat{C}}(t,b,1) - U^{C,t_1}(t,b,1)$ which shows that the left hand side of (50) remains uniformly bounded. This yields (47) with $K = \beta/\alpha$ and finishes the proof. □

Remark 5.8. As regards the limiting case when we define ρ in (44) by replacing t_1 by the critical value $\widehat{t_0}$, we conjecture that we obtain $\mathbb{E}[\rho^{\frac{1}{2}}] = \infty$. But we were not able to prove this result.

6. A pointwise version of one of Davis' inequalities

The value function V allows to derive a *pointwise* version of the Burkholder-Davis-Gundy inequality (1), which holds true in an almost sure sense rather than in expectation as stated in (1). This line of argument, inspired by the idea of robust superhedging from mathematical finance, is well-known (see e.g. [5] and [4]).

Theorem 6.1. *Denote by $V = V^{\widehat{C}}$ the value function (2) associated to the optimal constant $\widehat{C}$ and consider the Brownian motion $B = (B(t))_{t\geqslant 0}$ with its (right continuous, saturated) natural filtration $(\mathcal{F}(t))_{t\geqslant 0}$.*

There is a predictable process $H(t)$ satisfying $\mathbb{E}[\int_0^T H^2(t)dt] < \infty$, for each $T > 0$, given a.s. by

$$H(t) = V_b(t, B(t), B^*(t)) \tag{51}$$

for Lebesgue almost all $t > 0$, such that, for every bounded stopping time τ,

$$\tau^{\frac{1}{2}} - CB^*(\tau) \leqslant \int_0^\tau H(t)dB(t), \qquad a.s.. \tag{52}$$

Before giving the proof we observe the well-known fact that (52) trivially implies (2) by taking expectations on both sides of (52).

Proof. Lemma 2.5 states that the continuous process

$$X(t) = V(t, B(t), B^*(t))$$

is a super-martingale, starting at $X(0) = V(0,0,0) = 0$. sBy Doob-Meyer we may decompose X as

$$X = M - A \tag{53}$$

where M is a continuous local martingale and A is a continuous non-decreasing predictable process, and $M(0) = A(0) = 0$.

In fact, M is a square integrable martingale as we will show in Lemma A.1 in the appendix.

By martingale representation we may find a predictable process H with $\mathbb{E}[\int_0^T H(t)^2 dt] < \infty$, for each $T > 0$ such that

$$M(t) = \int_0^t H(u) dB(u).$$

By applying Ito to both sides of (53) we obtain the relation (51) which must hold true, for $\mathbb{P}$-almost each ω and for Lebesgue almost all t (the null set depending on ω). The formal application of Ito's formula can be justified using the result in Lemma 5.2. $\square$

7. BDG-Inequalities for general $0 < p < 2$

The above procedure can be easily modified to obtain similar results for the inequalities $\mathbb{E}[\tau^{\frac{p}{2}}] \leqslant C_p \mathbb{E}[(B^*(\tau))^p]$ with $0 < p < 2$. Lemma 2.2 and 2.4 stay essentially the same, with a different scaling given by

$$V(a^2 t, ab, ab^*) = a^p V(t, b, b^*).$$

This leads to the PDE

$$2tV_t + bV_b + b^* V_{b*} = pV$$

and the OIDE

$$2tV_t(t) = pV(t) + \int_0^\infty [V(t+s) - V(t)] g(s) ds$$

for $0 \leqslant t \leqslant t_0$ and the starting condition $V(t) = t^{\frac{p}{2}} - C$ for $t \geqslant t_0$.

In principle a similar analysis as in the present paper should provide explicit numerical values $\widehat{C}(p)$ and $\widehat{t}_0(p)$, in dependence of $0 < p < 2$. We leave this task to future research.

On the other hand, for $p > 2$ the present method does not seem to apply and some new idea is needed.

8. Relation to the Burkholder constant $\widehat{C} = \sqrt{3}$

In this section we consider martingales also allowing for jumps and we focus (w.l.o.g.) on martingales $(M_n)_{n=0}^N$ defined on a finite probability space Ω (see Lemma 8.2 below). The BDG inequality (1) reads in this context as

$$\mathbb{E}[[M,M]_N^{\frac{1}{2}}] \leqslant C\mathbb{E}[M_N^*], \tag{54}$$

where $[M,M]_n = \sum_{j=1}^n (M_j - M_{j-1})^2$ denotes the quadratic variation process. It was shown by D. Burkholder [7] that in this context the sharp constant $\widehat{C}$ equals $\widehat{C} = \sqrt{3}$.

One may ask for a deeper reason why we obtain a different sharp constant in (1) for continuous martingales as for martingales also having jumps. One reason is that the value function V fails to have a certain concavity property.

Fix a point $d = (t, b, b^*) \in D$ as well as $\alpha > 0, \beta > 0$. Define the points $d_\alpha, d_\beta \in D$ by

$$\begin{aligned} d_\alpha &= (t + \alpha^2, b + \alpha, \max(b^*, |b + \alpha|)), \\ d_\beta &= (t + \beta^2, b - \beta, \max(b^*, |b - \beta|)). \end{aligned}$$

We also define $p = \frac{\beta}{\alpha+\beta}$ and $q = \frac{\alpha}{\alpha+\beta}$.

Proposition 8.1. *There exist $x \in D$ as well as $\alpha > 0$, $\beta > 0$ such that*

$$V(d) < pV(d_\alpha) + qV(d_\beta). \tag{55}$$

Of course, we could verify the above proposition in a trivial way by numerically analyzing the function $V(t, b, b^*)$ and detecting explicitly some d, α and β. It is also clear where we should search for such a “bad” triple (x, α, β), namely in a neighborhood of the “kink” related to the “non-smooth pasting” (Figure 1 and 2) which displays a strong form of non-concavity.

But this is not our point. The purpose of the above statement is to show how the non-concavity (55) of the value function V is related to the difference between the case of continuous martingales and the case of martingales with jumps.

Also note that the equations $V_t + \frac{1}{2}V_{bb} = 0$ in the interior of D and $V_{b*} = 0$ on the non-stopping boundary of D (i.e. (23) and (15) above), imply that in the (properly interpreted) case of *infinitesimal* increments α and β we do have a “$\leqslant$” in (55) above. This is the message of Lemma 2.5.

Proof of Proposition 8.1. Admitting the subsequent lemma, we consider a dyadic martingale $(M_n)_{n=0}^N$ starting at $M_0 = 0$.

Let us fix some notation: The underlying probability space is given by

$$\Omega = \{(\omega_1, \ldots, \omega_N) : \omega_n \in \{-1, 1\}\}$$

and the filtration $(\mathcal{F}_n)_{n=0}^N$ is given by $\mathcal{F}_n = \sigma(\omega_1, \ldots, \omega_n)$. Consider the process

$$(X_n)_{n=1}^N = (V([M,M]_n, M_n, M_n^*))_{n=0}^N$$

where $V = V^{\widehat{C}}$ is the value function (2) associated to the optimal constant $\widehat{C} \approx 1,27267\dots$ for continuous processes.

It may happen that $(X_n)_{n=0}^N$ is a super-martingale. In this case

$$\begin{aligned} 0 = X_0 &\geqslant \mathbb{E}[V([M,M]_N, M_N, M_N^*)] \\ &\geqslant \mathbb{E}[[M,M]_N^{\frac{1}{2}} - \widehat{C} M_N^*], \end{aligned}$$

so that we obtain the inequality

$$\mathbb{E}[[M,M]_N^{\frac{1}{2}}] \leqslant \widehat{C}\mathbb{E}[M_N^*].$$

However, we know that $\widehat{C} \approx 1,27267\dots$ is smaller than the sharp constant $\widehat{C} = \sqrt{3}$ for martingales with jumps so that there must exist some dyadic martingale $(M_n)_{n=1}^N$ such that the corresponding process $(X_n)_{n=1}^N$ fails to be a supermartingale.

This means that there is some $0 \leqslant n \leqslant N-1$ and $\omega^{(n)} = (\omega_1, \dots, \omega_n)$ such that – with slight abuse of notation – we find

$$d := ([M,M]_n(\omega^{(n)}), M_n(\omega^{(n)}), M_n^*(\omega^{(n)}))$$

as well as

$$\begin{aligned} d_\alpha &= ([M,M]_{n+1}(\omega^{(n)},1), M_{n+1}(\omega^{(n)},1), M_{n+1}^*(\omega^{(n)},1)), \\ d_\beta &= ([M,M]_{n+1}(\omega^{(n)},-1), M_{n+1}(\omega^{(n)},-1), M_{n+1}^*(\omega^{(n)},-1)). \end{aligned}$$

such that inequality (55) holds true. □

For the following Lemma recall that a martingale is *dyadic* if the increment $M_{n+1} - M_n$ can attain at most two values, conditionally on $\sigma(M_1, \dots, M_n)$,.

Lemma 8.2. *For a constant $C > 0$ the following are equivalent:*

- *(i) Every dyadic martingale $(M_n)_{n=0}^N$ satisfies (54).*
- *(ii) Every martingale $(M_n)_{n=0}^N$ defined on a finite probability space satisfies (54).*
- *(iii) Every L^2 bounded martingale $(M_t)_{0\leqslant t \leqslant T}$ satisfies (54).*

Proof. The equivalence (ii) ⇔ (iii) is standard but for the convenience of the reader we will recall the argument for the non-trivial implication (ii) ⇒ (iii).

First, we can reduce the problem to discrete L^2-martingales: Fix an L^2-bounded martingale $M = (M_t)_{0\leqslant t\leqslant 1}$, based on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leqslant t\leqslant T}, \mathbb{P})$ and consider the martingales $(M_{k2^{-n}})_{k=0}^{2^n}$ for $n \in \mathbb{N}$. If they fulfill (54), then letting $n \to \infty$ yields that M satisfies (54).

Now, fix an L^2-bounded martingale $M = (M_n)_{0\leqslant n\leqslant N}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{0\leqslant n\leqslant N}, \mathbb{P})$. Consider the net of finite subfiltrations $(\mathcal{F}_n)_{0\leqslant n\leqslant N}$ of this filtration and their associated martingales $M^{\mathcal{G}} = (M_n^{\mathcal{G}})_{n=0}^N$ (i.e. $M_n^{\mathcal{G}} = \mathbb{E}[M_n|\mathcal{G}_n]$). By (ii), every $M^{\mathcal{G}}$ satisfies (54). The L^2 limit of $M^{\mathcal{G}}$ is M and (iii) follows.

The implication (ii) $\Rightarrow$ (i) is trivial.

To show (i) $\Rightarrow$ (ii) first observe that without loss of generality we can assume M_0 to be deterministic. First, we can translate the martingale such that it has mean 0. Second, if M_0 is random, then define a martingale $(M'_n)_{n=0}^{N+1}$ with $M'_0 := 0$ and $M'_i := M_{i-1}$ for $i > 0$. Then $[M', M']_{N+1} = [M, M]_N$ and $M'^*_{N+1} = M^*_N$.

Now, suppose first that $(M_n)_{n=0}^1$ is just a one step martingale on a finite probability space Ω with $M_0 = 0$. We then have that M_1 is a finitely valued random variable with $\mathbb{E}[M_1] = M_0 = 0$.

By possibly passing to a bigger (still finite) Ω we may find a partition $(A_1, \ldots, A_p)$ of Ω such that M_1 takes at most 2 values on each A_j and

$$\mathbb{E}[M_1 \mathbb{1}_{A_j}] = 0, \quad j = 1, \ldots, p.$$

We now define a dyadic martingale $(\tilde{M}_j)_{j=0}^p$ by $\tilde{M}_0 = 0$ and

$$\tilde{M}_j - \tilde{M}_{j-1} := M_1 \mathbb{1}_{A_j}, \quad j = 1, \ldots, p.$$

Clearly the variables $([M, M]_1, M_1^*)$ and $([\tilde{M}, \tilde{M}]_p, \tilde{M}_p^*)$ are equal in law. This shows that, at least for $N = 1$, we may associate to every finitely values martingale $(M_n)_{n=0}^1$ a dyadic martingale $(\tilde{M}_j)_{j=0}^p$ such that (54) holds true for M if and only if it does so for $\tilde{M}$.

It is rather obvious how to continue the above construction in an inductive way so that we may associate to each finitely valued martingale $(M_n)_{n=0}^N$ a dyadic martingale $(\tilde{M}_j)_{j=0}^p$ such that $([M, M]_N, M_N^*)$ and $([\tilde{M}, \tilde{M}]_p, \tilde{M}_p^*)$ are equal in law. This readily shows (i) $\Rightarrow$ (ii). ☐

Appendix A: The Martingale Property of the Value Process and Square Integrability

Proof of Lemma 2.5. We denote the function appearing on the right side of (5) by $v(t, b^*)$:

$$v(t, b^*) := t^{\frac{1}{2}} - Cb^*, \qquad t, b^* \geqslant 0.$$

For fixed $(t, b, b^*) \in D$ and $t \leqslant T$ we denote by $V^T(t, b, b^*)$ the value function defined similarly as in (2), but where we only allow for stopping times $\tau \in \mathcal{T}(t)$ which are bounded by T. Clearly $V^T(t, b, b^*)$ increases to $V(t, b, b^*)$, as $T \to \infty$, pointwise for $(t, b, b^*) \in D$. Also note that $V(T, b, b^*) = v(T, b^*)$. Fix (t, b, b^*) and a bounded stopping time $t \leqslant \sigma \leqslant \tau$. We have to show that

$$V(t, b, b^*) = \mathbb{E}^{(t,b,b^*)}[V(\sigma, B(\sigma), B^*(\sigma))] \tag{56}$$

By the monotone convergence theorem it will suffice to show that

$$\lim_{T\to\infty} V^T(t, b, b^*) = \lim_{T\to\infty} \mathbb{E}^{(t,b,b^*)}[V^T(\sigma^T, B(\sigma^T), B^*(\sigma^T))] \tag{57}$$

for $\sigma^T := \sigma \wedge \tau^T$ where τ^T is the stopping time defined conditionally on (t, b, b^*) by

$$\tau^T = \inf\{u \geqslant t : V^T(u, B(u), B^*(u)) = v(u, B^*(u))\}.$$

We then have that τ^T is bounded by T and increases a.s. to τ. The crucial property is

$$V^T(t, b, b^*) = \mathbb{E}^{(t,b,b^*)}[v(\tau^T, B^*(\tau^T))],$$

and, more generally, for any stopping time $t \leqslant \rho \leqslant \tau^T$,

$$V^T(\rho, B(\rho), B^*(\rho)) = \mathbb{E}[v(\tau^T, B^*(\tau^T))|\mathcal{F}(\rho)].$$

This classical result can be found in [13, Theorem 2.2]. Putting this together and taking $\rho = \sigma^T$, we obtain (57).

The proof of the supermartingale property which still holds true, after time τ is identical with an inequality instead of an equality.

□

We can even show that the value process is bounded in L^2 up to some fixed time T:

Lemma A.1. *The supermartingale* $(X(t))_{0 \leqslant t \leqslant T}$ *given by*

$$X(t) := V(t, B(t), B^*(t))$$

is uniformly bounded from above and bounded in L^2. *Furthermore the martingale component of its Doob-Meyer decomposition* $X(t) = M(t) - A(t)$ *is also bounded in* L^2 *and we obtain the following quantitative estimates for every stopping times* σ *with* $0 \leqslant \sigma \leqslant T$*:*

(i) $X(\sigma) \leqslant T^{\frac{1}{2}}$,
(ii) $\mathbb{E}[X(\sigma)^2] \leqslant KT$,
(iii) $\mathbb{E}[M(\sigma)^2] \leqslant \mathbb{E}[M(T)^2] < \infty$.

Proof. We first observe that $V(t, b, |b|) - (t^{\frac{1}{2}} - C|b|) \leqslant V(0, b, |b|) - (-C|b|) \leqslant C|b|$ holds as a consequence of the proof of Lemma 2.4 which gives us the estimate $V(t, b, |b|) \leqslant t^{\frac{1}{2}}$. Letting $t \to 0$ this also imples $V(t, 0, 0) \leqslant t^{\frac{1}{2}}$ and we can show (i) by

$$V(t, b, b^*) \leqslant V(t, 0, b^*) \leqslant V(t, 0, 0) \leqslant t^{\frac{1}{2}}.$$

V is monotone increasing in t and monotone decreasing in $|b|$ and b^*. So we can observe for the positive part of X that

$$\mathbb{E}[(X(\sigma)_+)^2] \leqslant (V(T, 0, 0)_+)^2 \leqslant T.$$

For the negative part $X(t)_-$ we can use $V(t, b, b^*) \geqslant t^{\frac{1}{2}} - Cb^* \geqslant -Cb^*$ and estimate

$$\mathbb{E}[(X(\sigma)_-)^2] \leqslant C^2\mathbb{E}[B^*(T)^2].$$

In summary we have

$$\mathbb{E}[X(\sigma)^2] \leqslant T + C^2 \mathbb{E}[B^*(T)^2].$$

To show the last assertion, we now split X into a sum of bounded processes in the following way. Define the stopping times $(\sigma_n)_{n=0}^{\infty}$ by

$$\sigma_n = \inf\{t : |B(t)| = 2^n\} \wedge T,$$

and define the processes X_n, obtained by starting X at time σ_{n-1} and stopping it at time σ_n:

$$X_n(t) = {}^{\sigma_{n-1}}X^{\sigma_n}(t) = (X_{t \wedge \sigma_n} - X_{\sigma_{n-1}}) \mathbb{1}_{[\![\sigma_{n-1}, T]\!]}(t).$$

Of course, we have $X = \sum_{n=1}^{\infty} X_n$ and the trajectories of $(X_n(t))_{0 \leqslant t \leqslant T}$ are only different from zero on the set $(\{\sigma_{n-1} < T\})_{n=1}^{\infty}$.

The probability of these events can be estimated by

$$\mathbb{P}[\sigma_{n-1} < T] \leqslant c_1 e^{-c_2 2^{2n}}, \tag{58}$$

for some constants $c_1 = c_1(T)$ and $c_2 = c_2(T)$.

Using a classical inequality on uniformly bounded supermartingales (apparently due to P. Meyer [12]) we obtain that each M_n is a square integrable martingale whose norm can be estimated by

$$||M_n(T)||_{L^2(\mathbb{P})}^2 \leqslant c_3 2^{c_4 n} \mathbb{P}[\sigma_{n-1} < T] \tag{59}$$

for some constants c_3, c_4 depending only on T.

Combining (58) and (59), we deduce that

$$||M(T)||_{L^2(\mathbb{P})}^2 = \sum_{n=1}^{\infty} ||M_n(T)||_{L^2(\mathbb{P})}^2 < \infty.$$

□

For the convenience of the reader we spell out the message of Meyer's Theorem [12, Theorem 46] in the present context.

Theorem A.2 (Meyer). *Let $X = (X(t))_{0 \leqslant t \leqslant T}$ be a uniformly bounded supermartingale*

$$||X||_{\infty} := \sup_{0 \leqslant t \leqslant T} ||X(t)||_{L^{\infty}} \leqslant c < \infty.$$

Denoting by $X = M - A$ its Doob-Meyer decomposition we get that M is a square integrable martingale whose norm can be estimated by $||M||_2^2 := ||M(T)||_{L^2}^2 \leqslant 18c^2$

Proof. By standard approximation results it will suffice to show the result for a super-martingale $X = (X(n))_{n=0}^{N}$ in finite discrete time. Note that in this case we have $A(0) = 0$ and

$$A(n+1) - A(n) = \mathbb{E}[X(n+1) - X(n) | \mathcal{F}(n)], \quad n = 1, \dots, N$$

so that

$$||\Delta A(n)||_{L^\infty} \leqslant 2c.$$

We may telescope $A(N) = \sum_{n=1}^{N}(A(n) - A(n-1))$ to obtain

$$\begin{aligned} A(N)^2 &= 2\sum_{n=1}^{N}(A(n) - A(n-1))\sum_{j=n}^{N}(A(j) - A(j-1)) \\ &= 2\sum_{n=1}^{N}(A(n) - A(n-1))(A(N) - A(n-1)). \end{aligned}$$

By taking expectations we get

$$\begin{aligned} \mathbb{E}[A(N)^2] &= 2\sum_{n=1}^{N}\mathbb{E}[\mathbb{E}[(A(n) - A(n-1))(A(N) - A(n-1))|\mathcal{F}(n-1)]] \\ &= 2\sum_{n=1}^{N}\mathbb{E}[(A(n) - A(n-1))\mathbb{E}[A(N) - A(n-1)|\mathcal{F}(n-1)]]. \end{aligned}$$

The final term is uniformly bounded as

$$\mathbb{E}[A(N) - A(n-1)|\mathcal{F}(n-1)] = \mathbb{E}[X(N) - X(n-1)|\mathcal{F}(n-1)] \leqslant 2c.$$

This yields

$$\begin{aligned} \mathbb{E}[A(N)^2] &\leqslant 4c\sum_{n=1}^{N}\mathbb{E}[A(n) - A(n-1)] \\ &= 4c\mathbb{E}[X(N) - X(0)] \leqslant 8c^2. \end{aligned}$$

To obtain a bound for $||M||_2$ we use the relation $M = X + A$ and $||X(N)||_{L^\infty} \leqslant c$ to get

$$\mathbb{E}[M(N)^2] \leqslant 2\mathbb{E}[A(N)^2] + 2c^2 \leqslant 18c^2.$$

□

Appendix B: Some facts on the stopping time of first leaving a corridor

We discuss the first exit time of the interval $[-h, 2+h]$ for some $h > 0$ for a standard Brownian motion B started at $B(0) = 0$.

$$\sigma^h := \inf\{t : |B(t) - 1| = 1 + h\}.$$

This stopping time has a well-known density and a well-known Laplace-Transform $\mathcal{L}$ (see e.g. [11, Section 2.2.8.C] given by

$$\mathcal{L}(\theta) = \frac{\cosh(\sqrt{2\theta})}{\cosh((1+h)\sqrt{2\theta})}.$$

One can calculate the expected value of the stopping time by noting that $\mathbb{E}[\sigma^h] = \mathbb{E}[B^2_{\sigma^h}]$ so that $\mathbb{E}[\sigma^h] = h(2+h)$. The Jensen-inequality directly implies that the fractional moments of order less than 1 also exist and one has the estimate $\mathbb{E}[(\sigma^h)^{\frac{1}{2}}] \leqslant (h(2+h))^{\frac{1}{2}}$. This motivates to conjecture that $\frac{\mathbb{E}[(\sigma^h)^{\frac{1}{2}}]}{h} \to \infty$ for $h \to 0$. We can get an expression for this moment using the Laplace-transform above by the formula $\mathbb{E}[(\sigma^h)^{\frac{1}{2}}] = \frac{-1}{\Gamma(\frac{1}{2})} \int_0^\infty \theta^{-\frac{1}{2}} \mathcal{L}'(\theta)\, d\theta$, which we can evaluate to

$$\mathbb{E}[(\sigma^h)^{\frac{1}{2}}] = \sqrt{\frac{2}{\pi}} \left[\int_0^\infty \frac{\sinh(hu)}{u \cosh((1+h)u)^2}\, du + h \int_0^\infty \frac{\cosh(u) \tanh((1+h)u)}{u \cosh((1+h)u)}\, du \right]$$

where a substitution $u = \sqrt{2\theta}$ was used. It can be shown that both of these integrals are in fact finite for positive h, but we are only interested in the above limiting behavior of this expression. To see this first note that the integrands of both integrals are always positive. Furthermore the hyperbolic tangent converges to 1. Therefore we can fix a constant K such that for $h < 1$ we have that $\tanh((1+h)u) \geqslant \frac{1}{2}$ for $u \geqslant K$. Putting this together we make the following estimate:

$$\begin{aligned} \frac{\mathbb{E}[(\sigma^h)^{\frac{1}{2}}]}{h} &\geqslant \frac{1}{\sqrt{2\pi}} \int_K^\infty \frac{\cosh(u)}{u \cosh((1+h)u)}\, du \\ &\geqslant \frac{1}{2\sqrt{2\pi}} \int_K^\infty \frac{e^{-hu}}{u}\, du = \frac{1}{2\sqrt{2\pi}} \int_{hK}^\infty \frac{e^{-u}}{u}\, du. \end{aligned}$$

The last expression now obviously diverges for $h \to 0$. The following Lemma which we will need later on uses the above observations:

Lemma B.1. *Let $K > 0$ be an arbitrary constant. There exist $t, h > 0$ such that*

$$\mathbb{E}[(t+\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] \geqslant Kh.$$

Proof. By the above observations we can choose h small enough to obtain $\mathbb{E}[(\sigma^h)^{\frac{1}{2}}] \geqslant Kh + 1$. We also have $\mathbb{E}[(\sigma^h)^{\frac{1}{2}}] - \mathbb{E}[(t+\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] < 1$ for small enough t by monotone convergence of $(t+s)^{\frac{1}{2}} - t^{\frac{1}{2}}$ to $s^{\frac{1}{2}}$ for $t \to 0$. □

We can now proceed to show the following facts about t_0 to prove the final statement of Lemma 2.4.

Lemma B.2. *Let $C \geqslant \widehat{C}$ and $V(t, b, b^*)$ the corresponding value function. The map $t \mapsto V(t,1,1) - (t^{\frac{1}{2}} - C)$ is decreasing and if $V(1,1,1) > 1 - C$ it follows that $V(1,1,1) - (1-C) < V(0,1,1) - (-C)$.*

Proof. We need a quantitative version of Lemma 2.1 (ii) which already shows that $t \mapsto V(t,1,1) - (t^{\frac{1}{2}} - C)$ is decreasing. First choose some $\nu \geqslant 1$ to be a bounded stopping time which achieves

$$V(1,1,1) - \mathbb{E}^{(1,1,1)}[\nu^{\frac{1}{2}} - CB^*(\nu)] < \varepsilon \tag{60}$$

such that

$$\mathbb{P}[\nu > 1 + 2\varepsilon] > (2\varepsilon)^{\frac{1}{2}} \tag{61}$$

for arbitrary $\varepsilon > 0$. It is clear by definition of V that there exists a stopping time which satisfies (60). Suppose there is no appropriate stopping time such that (61) is satisfied, then there is an optimizing sequence of bounded stopping times which converge to 1 in probability and thus a subsequence which converges almost surely. This would imply that $V(1,1,1) = 1 - C$ which contradicts the assumptions of the Lemma.

Now, we can consider the stopping time ν as a randomized stopping time with respect to the filtration $(\mathcal{G}^{(1)}(u))_{u\geqslant 1}$. The shifted stopping time $\nu' := \nu - 1$ is then a randomized stopping time with respect to $(\mathcal{G}^{(0)}(u))_{u\geqslant 0}$. We can now estimate

$$\begin{aligned}
&[V(0,1,1) - (-C)] - [V(1,1,1) - (1-C)] \\
&\qquad \geqslant \mathbb{E}^{(0,1,1)}[\nu'^{\frac{1}{2}} - CB^*(\nu')] - \mathbb{E}^{(1,1,1)}[\nu^{\frac{1}{2}} - CB^*(\nu)] + 1 - \varepsilon \\
&\qquad = \mathbb{E}[(\nu')^{\frac{1}{2}} - (1+\nu')^{\frac{1}{2}} + 1] - \varepsilon \\
&\qquad \geqslant \mathbb{E}[((\nu')^{\frac{1}{2}} - (1+\nu')^{\frac{1}{2}} + 1)\mathbb{1}_{\nu' > 2\varepsilon}] - \varepsilon \\
&\qquad \geqslant [(2\varepsilon)^{\frac{1}{2}} - ((1+2\varepsilon)^{\frac{1}{2}} - 1)](2\varepsilon)^{\frac{1}{2}} - \varepsilon \\
&\qquad > \varepsilon - \sqrt{2}\varepsilon^{\frac{3}{2}} > 0.
\end{aligned}$$

To get from the second to the third line, we used that by definition $\nu = \nu' + 1$ and $\mathbb{E}^{(0,1,1)}[B^*(\nu')] = \mathbb{E}^{(1,1,1)}[B^*(\nu)]$. We dropped the superscript to emphasize that we now view ν' as a stopping time with respect to the filtration $(\mathcal{G}^{(0)}(u))_{u\geqslant 0}$. To obtain the fourth and the fifth line in the derivation, we observe that the map $t \mapsto t^{\frac{1}{2}} - (1+t)^{\frac{1}{2}} + 1$ is non-negative and non-decreasing, and use (61). The sixth line can be derived by noting that $s \mapsto (1+s)^{\frac{1}{2}} - 1$ is concave and thus lies completely under its tangent at $s = 0$. The last inequality holds for ε small enough. In the same way one can actually show that as long as the spread $V(t,1,1) - (t^{\frac{1}{2}} - C)$ is strictly positive, it is also strictly decreasing. ☐

Lemma B.3. *Let $C \geqslant \widehat{C}$ and $t_0 = t_0(C)$ the critical point separating S from NS. Then*

1. *$t_0 > 0$ and*
2. *$t_0 < \infty$.*

Proof. (1): Assume $t_0 = 0$. This means that we actually have

$$V(t, b^*, b^*) = t^{\frac{1}{2}} - Cb^*$$

for all $t > 0$. We then obtain for arbitrary $t, h > 0$ that

$$\mathbb{E}^{(t,1,1)}[(\sigma^h)^{\frac{1}{2}} - C(1+h)] \leqslant (t)^{\frac{1}{2}} - C$$

by the supermartingale property of the value-process. This is a contradiction to Lemma B.1 for small enough t and h.

(2): Assume $t_0 = \infty$. This means that $V(t,1,1) > t^{\frac{1}{2}} - C$ everywhere. As $C \geqslant \widehat{C}$ we also have $V(0,1,1) \leqslant 0$. By Lemma B.2 we can set $\alpha := V(1,1,1) - 1 + C < C$. Now fix some $h > \frac{\alpha}{C-\alpha}$ and $t > (1+h)^2$. We can then make the following estimate, where we use twice the fact that the function $t \mapsto V(t,b,|b|) - (t^{\frac{1}{2}} - C|b|)$ is decreasing and σ^h is defined as before:

$$\begin{aligned} V(t,1,1) &= \mathbb{E}^{(t,1,1)}[V(\sigma^h, 1+h, 1+h)] \\ &\leqslant V(t,1+h,1+h) + \mathbb{E}^{(t,1,1)}[(\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] \\ &= (1+h)V\left(\frac{t}{(1+h)^2}, 1, 1\right) + \mathbb{E}^{(t,1,1)}[(\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] \\ &\leqslant t^{\frac{1}{2}} - (1+h)(C-\alpha) + \mathbb{E}^{(t,1,1)}[(\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] \\ &< V(t,1,1) - h(C-\alpha) + \alpha + \mathbb{E}^{(t,1,1)}[(\sigma^h)^{\frac{1}{2}} - t^{\frac{1}{2}}] \end{aligned}$$

Now we can eliminate $V(t,1,1)$ on both sides and note that h, C and α do not depend on t. The last term however goes to 0 for $t \to \infty$ by dominated convergence (since $0 \leqslant \sqrt{t+\sigma^h} - \sqrt{t} \leqslant \sqrt{\sigma^h}$ and $\sqrt{t+\sigma^h} - \sqrt{t} \searrow 0$). This leads to the desired contradiction. ☐

Acknowledgements

Thanks go to an extremely helpful referee who provided a very careful report. The paper in its present form owes much to his suggestions. We also thank Mathias Beiglböck for his insight and advice in the course of many discussions on the present paper and Josef Teichmann for pointing out to us Meyer's inequality (Theorem A.2).

References

[1] L. Alili and A. E. Kyprianou. Some remarks on first passage of Lévy processes, the American put and pasting principles. *Ann. Appl. Probab.*, 15(3):2062–2080, 2005.

[2] S. Asmussen, F. Avram, and M. R. Pistorius. Russian and american put options under exponential phase-type Lévy models. *Stochastic Processes and their Applications*, 109(1):79–111, 2004.

[3] J. R. Baxter and R. V. Chacon. Compactness of stopping times. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(3):169–181, 1977.

[4] M. Beiglböck and M. Nutz. Martingale inequalities and deterministic counterparts. *Electron. J. Probab.*, 2014.

[5] M. Beiglböck and P. Siorpaes. Pathwise versions of the Burkholder-Davis-Gundy inequality. *Bernoulli*, 21(1):360–373, 2015.

[6] S. I. Boyarchenko and S. Z. Levendorskii. Perpetual american options under lévy processes. *SIAM Journal on Control and Optimization*, 40(6):1663–1696, 2002.

[7] D. Burkholder. The best constant in the Davis inequality for the expectation of the martingale square function. *Transactions of the American Mathematical Society*, 354(1):91–105, 2002.

[8] D. L. Burkholder and R. F. Gundy. Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.*, 124:249–304, 1970.

[9] R. C. Dalang and M.-O. Hongler. The right time to sell a stock whose price is driven by markovian noise. *Annals of Applied Probability*, pages 2176–2201, 2004.

[10] B. Davis. On the intergrability of the martingale square function. *Israel Journal of Mathematics*, 8(2):187–190, 1970.

[11] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.

[12] P.-A. Meyer. *Martingales and Stochastic Integrals I.* Springer-Verlag, Berlin, 1972.

[13] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems.* Springer, 2006.

[14] G. Peskir and A. N. Shiryaev. Sequential testing problems for Poisson processes. *Annals of Statistics*, pages 837–859, 2000.

[15] L. Shepp. A first passage problem for the Wiener process. *The Annals of Mathematical Statistics*, pages 1912–1914, 1967.