RESEARCH   **Open Access**

Optimal control with delayed information flow of systems driven by *G*-Brownian motion

CrossMark

Francesca Biagini (iD) · **Thilo Meyer-Brandis** ·
Bernt Øksendal · **Krzysztof Paczka**

Abstract In this paper, we study strongly robust optimal control problems under
volatility uncertainty. In the *G*-framework, we adapt the stochastic maximum princi-
ple to find necessary and sufficient conditions for the existence of a strongly robust
optimal control.

1 Introduction

One of the motivations for this paper is to study the problem of optimal consumption
and optimal portfolio allocation in finance under model uncertainty. In particular, we
focus on volatility uncertainty, i.e., a situation where the volatility affecting the asset
price dynamics is unknown and we need to consider a family of different volatility
processes instead of just one fixed process (and hence also a family of models related
to them).

F. Biagini
Department of Mathematics, LMU Munich, Theresienstraße 39, 80333 Munich, Germany
e-mail: biagini@math.lmu.de

F. Biagini (✉)· B. Øksendal · K. Paczka
Department of Mathematics, University of Oslo, P.O.Box 1053 Blindern, N-0316 Oslo, Norway
e-mail: biagini@math.lmu.de; oksendal@math.uio.no; k.j.paczka@cma.uio.no

T. Meyer-Brandis
Department of Mathematics, University of Munich, Theresienstraße 39, 80333 Munich, Germany
e-mail: meyerbr@math.lmu.de

Volatility uncertainty has been investigated in the literature by following two approaches, i.e., by introducing an abstract sublinear expectation space with a special process called G-Brownian motion (see (Peng 2007), (Peng 2010)), or by capacity theory (see (Denis et al. 2011)). In (Denis et al. 2011), it is proven that these two methods are strongly related. The link between these two approaches is the representation of the sublinear expectation $\hat{\mathbb{E}}$ associated with the G-Brownian motion as a supremum of ordinary expectations over a tight family of probability measures $\mathcal{P}$, whose elements are mutually singular:

$$\hat{\mathbb{E}}[.] = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[.],$$

see (4) and Theorem 1 for more details.

In this paper, we work in a G-Brownian motion setting as in (Peng 2007) and use the related stochastic calculus, including the Itô formula, G-SDEs, martingale representation and G-BSDEs, as developed in (Peng 2007), (Peng 2010), (Soner et al. 2011a), (Song 2011), (Soner et al. 2011b), (Peng et al. 2014), (Hu et al. 2014c), (Hu et al. 2014a). It is important for understanding the nature of G-Brownian motion to note that its quadratic variation $\langle B \rangle$ is not deterministic, but it is absolutely continuous with the density taking value in a fixed set (for example, $[\underline{\sigma}^2, \bar{\sigma}^2]$ for $d = 1$). Each $\mathbb{P} \in \mathcal{P}$ can then be seen as a model with a different scenario for the quadratic variation. That justifies why G-Brownian motion is a good framework for investigating model uncertainty.

In a G-Brownian motion setting one considers the following stochastic optimal control problem: to find the control $\hat{u} \in \mathcal{A}$ such that

$$J(\hat{u}) = \sup_{u \in \mathcal{A}} J(u), \tag{1}$$

with

$$
\begin{aligned}
J(u) :&= \hat{\mathbb{E}}[\int_0^T f(t, X^u(t), u(t))dt + g(X^u(T))] \\
&= \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[\int_0^T f(t, X^u(t), u(t))dt + g(X^u(T))] =: \sup_{\mathbb{P} \in \mathcal{P}} J^{\mathbb{P}}(u),
\end{aligned}
\tag{2}
$$

where X^u is a controlled G-SDE, see (8). This problem has been studied in (Matoussi et al. 2013), (Hu et al. 2014b). In (Hu et al. 2014b), they show that the value function associated with such an optimal control problem satisfies the dynamic programming principle and is a viscosity solution of some HJB equation.[1] (Matoussi et al. 2013) investigates the robust investment problem for geometric G-Brownian motion, where 2BSDEs (which are closely related to G-BSDEs) are used to find an optimal solution. In both papers the optimal control is robust in a worst-case scenario sense.

It is interesting to note that in the simplest example of the optimal portfolio problem, which is the Merton problem with the logarithmic utility, one can easily prove that there exists a portfolio which is optimal not only in the worst-case scenario, but also for all probability measures $\mathbb{P}$ (with the optimality criterion $J^{\mathbb{P}}$). We call

[1]To be exact, the authors considered a more general problem of recursive utility.

this *a strongly robust control*. This strongly robust control is thus optimal in a much more robust sense than the worst-case scenario optimality. The new strongly robust optimality uses the fact that probability measures $\mathbb{P}$ are mutually singular. Informally speaking, one can therefore modify the $\mathbb{P}$-optimal control $\hat{u}^{\mathbb{P}}$ outside the support of a probability measure $\mathbb{P}$ without losing the $\mathbb{P}$-optimality. As a consequence, if the family $\{\hat{u}^{\mathbb{P}}\}_{\mathbb{P}\in\mathcal{P}}$ satisfies some consistency conditions, under a suitable choice of the underlying filtration the controls can be aggregated into a unique control $\hat{u}$, which is optimal under every probability measure $\mathbb{P}$. See (Soner et al. 2011b) for more details on aggregation.

In this paper, we study strongly robust optimal control problems. However, instead of checking the consistency condition for the family of controls and using the aggregation theory established in (Soner et al. 2011b), we adapt the stochastic maximum principle to the *G*-framework to find necessary and sufficient conditions for the existence of a strongly robust optimal control. We stress that this method has the clear advantage that we solve only one *G*-BSDE to produce the strongly robust optimal control instead of considering the optimal control problem for all $\mathbb{P} \in \mathcal{P}$ (which are usually not Markovian laws) and checking the consistency condition. Another advantage is that we work with the raw filtration instead of enlarging it.

In the recent paper (Hu and Ji 2016), they also study a stochastic maximum principle for stochastic recursive optimal control problems in the *G*-setting, but still using the worst-case approach. They use the Minimax Theorem to obtain the variational inequality under a reference probability P^*: the stochastic maximum principle holds then under such a P^*-a.s, which is the main difference with respect to our approach. They prove that this stochastic maximum principle is also a sufficient condition under some convex assumptions, but our control problem is different from the one in (Hu and Ji 2016) and considers delayed information.

The notion of the strongly robust optimal control also has a better financial interpretation than the standard robust optimality mentioned above. The main drawback of the classical robust optimal control is that it is a differential game from a mathematical point of view, where one player chooses the optimal control $\hat{u}$ and the other chooses the optimal volatility represented by the law $\hat{\mathbb{P}}$:

$$\sup_{u\in\mathcal{A}} \sup_{\mathbb{P}\in\mathcal{P}} J^{\mathbb{P}}(u) = J^{\hat{\mathbb{P}}}(\hat{u}).$$

Therefore, the optimal pair $(\hat{u}, \hat{\mathbb{P}})$ has the Nash equilibrium interpretation[2]. The problem is that in real life there is no reason to why we should assume that the worst case is true, as there are no players who try to maximize gains from choosing $\mathbb{P}$.

However, this is not the only problem with the standard robust optimality. Since the optimal probability measure $\hat{\mathbb{P}}$ is mutually singular to any other measure $\mathbb{Q} \in \mathcal{P}$, we can modify the control $\hat{u}$ outside the support of $\hat{\mathbb{P}}$ without losing the (classical) robust optimality. Since, as we noted above, usually the true probability will be different than $\hat{\mathbb{P}}$, the classical robust optimal control may have little sense for $\mathbb{Q}$. Moreover, in the standard robust optimality, the measure $\hat{\mathbb{P}}$ is chosen to be static and does not

---

[2]This is even more visible for minimization problems, where one has a saddle point.

change with time. As a result, not all available information is taken into consideration, as shown for the Merton problem with logarithmic utility in Section 5.

The paper is structured in the following way. In Section 2, we give a quick overview of the *G*-framework. Section 3 is devoted to a sufficient maximum principle in the partial-information case. In Section 4, we investigate the necessary maximum principle for the full-information case. In Section 5, we give four examples, including the Merton problem with logarithmic utility, mentioned earlier and an LQ-problem. In Section 6, we provide a counter-example and show that it is not possible to relax the crucial assumption of the sufficient maximum principle without losing the strongly robust sense of optimality.

2 Preliminaries

Let Ω be a given set and $\mathcal{H}$ be a vector lattice of real functions defined on Ω, i.e. a linear space containing 1 such that $X \in \mathcal{H}$ implies $|X| \in \mathcal{H}$. We will treat elements of $\mathcal{H}$ as random variables.

Definition 1 *A sublinear expectation $\mathbb{E}$ is a functional $\mathbb{E}\colon \mathcal{H} \to \mathbb{R}$ satisfying the following properties*

1. **Monotonicity:** *If $X, Y \in \mathcal{H}$ and $X \geq Y$, then $\mathbb{E}[X] \geq \mathbb{E}[Y]$.*
2. **Constant preserving:** *For all $c \in \mathbb{R}$, we have $\mathbb{E}[c] = c$.*
3. **Sub-additivity:** *For all $X, Y \in \mathcal{H}$, we have $\mathbb{E}[X] - \mathbb{E}[Y] \leq \mathbb{E}[X - Y]$.*
4. **Positive homogeneity:** *For all $X \in \mathcal{H}$, we have $\mathbb{E}[\lambda X] = \lambda \mathbb{E}[X]$ for all $\lambda \geq 0$.*

The triple $(\Omega, \mathcal{H}, \mathbb{E})$ is called a sublinear expectation space.

We will consider a space $\mathcal{H}$ of random variables having the following property: if $X_i \in \mathcal{H}$, $i = 1, \ldots n$, then

$$\phi(X_1, \ldots, X_n) \in \mathcal{H} \quad \forall \phi \in C_{b,Lip}(\mathbb{R}^n),$$

where $C_{b,Lip}(\mathbb{R}^n)$ is the space of all bounded Lipschitz continuous functions on $\mathbb{R}^n$.

Definition 2 *An m-dimensional random vector $Y = (Y_1, \ldots, Y_m)$ is said to be independent of an n-dimensional random vector $X = (X_1, \ldots, X_n)$ if for every $\phi \in C_{b,Lip}(R^n \times R^m)$*

$$\mathbb{E}[\phi(X, Y)] = \mathbb{E}[\mathbb{E}[\phi(x, Y)]_{x=X}].$$

Let X_1 and X_2 be n-dimensional random vectors defined on sublinear random spaces $(\Omega_1, \mathcal{H}_1, \mathbb{E}_1)$ and $(\Omega_2, \mathcal{H}_2, \mathbb{E}_2)$, respectively. We say that X_1 and X_2 are identically distributed and denote it by $X_1 \sim X_2$, if for each $\phi \in C_{b,Lip}(\mathbb{R}^n)$ one has

$$\mathbb{E}_1[\phi(X_1)] = \mathbb{E}_2[\phi(X_2)].$$

Definition 3 *A d-dimensional random vector $X = (X_1, \ldots, X_d)$ on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ is said to be G-normally distributed if for each $a, b \geq 0$*

and each $Y \in \mathcal{H}$ such that $X \sim Y$ and Y is independent of X, one has

$$aX + bY \sim \sqrt{a^2 + b^2} X.$$

The letter G denotes a function defined as

$$G(A) := \frac{1}{2} \mathbb{E}[(AX, X)] \colon \S_d \to \mathbb{R},$$

where $\S_d$ is the space of all $d \times d$ symmetric matrices. We assume that G is non-degenerate, i.e., $G(A) - G(B) \geq \beta \mathrm{tr}[A - B]$ for some $\beta > 0$.

It can be checked that *G* might be represented as

$$G(A) = \frac{1}{2} \sup_{\gamma \in \Theta} \mathrm{tr}\,(\gamma \gamma^T A), \tag{3}$$

where Θ is a non-empty bounded and closed subset of $\mathbb{R}^{d \times d}$.

Definition 4 *Let $G \colon \S_d \to \mathbb{R}$ be a given monotonic and sublinear function. A stochastic process $B = (B_t)_{t \geq 0}$ on a sublinear expectation space $(\Omega, \mathcal{H}, \mathbb{E})$ is called a G-Brownian motion if it satisfies the following conditions:*

1. $B_0 = 0$;
2. $B_t \in \mathcal{H}$ for each $t \geq 0$;
3. *For each $t, s \geq 0$ the increment $B_{t+s} - B_t$ is independent of $(B_{t_1}, \ldots, B_{t_n})$ for each $n \in \mathbb{N}$ and $0 \leq t_1 < \ldots < t_n \leq t$. Moreover, $(B_{t+s} - B_t)s^{-1/2}$ is G-normally distributed.*

Definition 5 *Let $\Omega = C_0(\mathbb{R}_+, \mathbb{R}^d)$, i.e., the space of all $\mathbb{R}^d$-valued continuous functions starting at 0. We equip this space with the uniform convergence on compact intervals topology and denote by $\mathcal{B}(\Omega)$ the Borel σ-algebra of Ω. Let*

$$\mathcal{H} = Lip(\Omega) := \left\{ \phi(\omega_{t_1}, \ldots, \omega_{t_n}) \colon \forall n \in \mathbb{N}, t_1, \ldots, t_n \in [0, \infty) \text{ and } \phi \in C_{b,Lip}\left(\mathbb{R}^{d \times n}\right) \right\}.$$

A G-expectation $\hat{\mathbb{E}}$ is a sublinear expectation on $(\Omega, \mathcal{H})$ defined as follows: for $X \in Lip(\Omega)$ of the form

$$X = \phi(\omega_{t_1} - \omega_{t_0}, \ldots, \omega_{t_n} - \omega_{t_{n-1}}), \quad 0 \leq t_0 < t_1 < \ldots < t_n,$$

we set

$$\hat{\mathbb{E}}[X] := \mathbb{E}\left[\phi\left(\xi_1\sqrt{t_1 - t_0}, \ldots, \xi_n\sqrt{t_n - t_{n-1}}\right)\right],$$

where $\xi_1, \ldots \xi_n$ are d-dimensional random variables on sublinear expectation space $(\tilde{\Omega}, \tilde{\mathcal{H}}, \mathbb{E})$ such that for each $i = 1, \ldots, n$, ξ_i is G-normally distributed and independent of $(\xi_1, \ldots, \xi_{i-1})$. We denote by $L^p_G(\Omega)$ the completion of $Lip(\Omega)$ under the norm $\|X\|_p := \hat{\mathbb{E}}[|X|^p]^{1/p}$, $p \geq 1$. Then it is easy to check that $\hat{\mathbb{E}}$ is also a sublinear expectation on the space $(\Omega, L^p_G(\Omega))$, $L^p_G(\Omega)$ is a Banach space, and the canonical process $B_t(\omega) := \omega_t$ is a G-Brownian motion.

Following (Peng 2010) and (Denis et al. 2011), we introduce the notation: for each $t \in [0, \infty)$

1. $\Omega_t := \{w_{.\wedge t} : \omega \in \Omega\}$, $\mathcal{F}_t := \mathcal{B}(\Omega_t)$,
2. $L^0(\Omega)$: the space of all $\mathcal{B}(\Omega)$-measurable real functions,
3. $L^0(\Omega_t)$: the space of all $\mathcal{B}(\Omega_t)$-measurable real functions,
4. $Lip(\Omega_t) := Lip(\Omega) \cap L^0(\Omega_t)$, $L^p_G(\Omega_t) := L^p_G(\Omega) \cap L^0(\Omega_t)$,
5. $M^2_G(0, T)$ is the completion of the set of elementary processes of the form

$$\eta(t) = \sum_{i=1}^{n-1} \xi_i \mathbb{1}_{[t_i, t_{i+1})}(s),$$

where $0 \leq t_1 < t_2 < \ldots < t_n \leq T$, $n \geq 1$ and $\xi_i \in Lip(\Omega_{t_i})$. The completion is taken under the norm

$$\|\eta\|^2_{M^2_G(0,T)} := \hat{\mathbb{E}} \left[\int_0^T |\eta(t)|^2 ds \right].$$

Definition 6 *Let $X \in Lip(\Omega)$ have the representation*

$$X = \phi \left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}} \right), \ \phi \in C_{b,Lip} \left(\mathbb{R}^{d \times n} \right), \ 0 \leq t_1 < \ldots < t_n < \infty.$$

We define the conditional G-expectation under $\mathcal{F}_{t_j}$ as

$$\hat{\mathbb{E}}[X|\mathcal{F}_{t_j}] := \psi \left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_j} - B_{t_{j-1}} \right),$$

where

$$\psi(x) := \hat{\mathbb{E}} \left[\phi \left(x, B_{t_{j+1}} - B_{t_j}, \ldots, B_{t_n} - B_{t_{n-1}} \right) \right].$$

Similarly to the *G*-expectation, the conditional *G*-expectation might also be extended to the sublinear operator $\hat{\mathbb{E}}[.|\mathcal{F}_t]: L^p_G(\Omega) \to L^p_G(\Omega_t)$ using the continuity argument. For more properties of the conditional G-expectation, see (Peng 2010).

    G-(conditional) expectation plays a crucial role in the stochastic calculus for *G*-Brownian motion. In (Denis et al. 2011), it was shown that the analysis of the *G*-expectation might be embedded in the theory of upper-expectations and capacities.

Theorem 1 ((Denis et al. 2011), Theorem 52 and 54) *Let $(\tilde{\Omega}, \mathcal{G}, \mathbb{P}_0)$ be a probability space carrying a standard d-dimensional Brownian motion W with respect to its natural filtration $\mathbb{G}$. Let Θ be a representation set defined as in (3) and denote by $\mathcal{A}^\Theta_{0,\infty}$ the set of all Θ-valued $\mathbb{G}$-adapted processes on an interval $[0, \infty)$. For each $\theta \in \mathcal{A}^\Theta_{0,\infty}$ define $\mathbb{P}^\theta$ as the law of a stochastic integral $\int_0^\cdot \theta_s dW_s$ on the canonical space $\Omega = C_0(\mathbb{R}_+, \mathbb{R}^d)$. We introduce the sets*

$$\mathcal{P}_1 := \left\{ \mathbb{P}^\theta : \theta \in \mathcal{A}^\Theta_{0,\infty} \right\}, \quad \text{and} \quad \mathcal{P} := \overline{\mathcal{P}_1}, \tag{4}$$

where the closure is taken in the weak topology. $\mathcal{P}_1$ is tight, so $\mathcal{P}$ is weakly compact. Moreover, one has the representation

$$\hat{\mathbb{E}}[X] = \sup_{\mathbb{P} \in \mathcal{P}_1} \mathbb{E}^{\mathbb{P}}[X] = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}}[X], \quad \text{for each } X \in L^1_G(\Omega). \tag{5}$$

For convenience, we always consider only a Brownian motion on the canonical space Ω with the Wiener measure $\mathbb{P}_0$.

Similarly, an analogous representation holds for the G-conditional expectation.

Proposition 1 ((Soner et al. 2011a), Proposition 3.4) *Let* $\mathcal{Q}(t, P) := \{\mathbb{P}' \in \mathcal{Q} \colon \mathbb{P} = \mathbb{P}'$ *on* $\mathcal{F}_t\}$, *where* $\mathcal{Q} = \mathcal{P}$ *or* $\mathcal{P}_1$. *Then, for any* $X \in L_G^1(\Omega)$ *and* $\mathbb{P} \in \mathcal{Q}$, $\mathcal{Q} = \mathcal{P}$ *or* $\mathcal{P}_1$, *one has*

$$\hat{\mathbb{E}}[X|\mathcal{F}_t] = \operatorname*{ess\,sup}_{\mathbb{P}' \in \mathcal{Q}(t, \mathbb{P})}{}^{\mathbb{P}} \mathbb{E}^{\mathbb{P}'}[X|\mathcal{F}_t], \ \mathbb{P} - a.s. \tag{6}$$

We now introduce the Choquet capacity (see (Denis et al. 2011)) related to $\mathcal{P}$

$$c(A) := \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}(A), \quad A \in \mathcal{B}(\Omega).$$

Definition 7 1. *A set A is said to be polar, if $c(A) = 0$. Let $\mathcal{N}$ be a collection of all polar sets. A property is said to hold quasi-surely (abbreviated to q.s.) if it holds outside a polar set.*
2. *We say that a random variable Y is a version of X if $X = Y$ q.s.*
3. *A random variable X is said to be quasi-continuous (q.c. in short), if for every $\varepsilon > 0$ there exists an open set O such that $c(O) < \varepsilon$ and $X|_{O^c}$ is continuous.*

We have the following characterization of spaces $L_G^p(\Omega)$. This characterization shows that $L_G^p(\Omega)$ is a rather small space.

Theorem 2 (Theorem 18 and 25 in (Denis et al. 2011)) *For each $p \geq 1$ one has*

$$L_G^p(\Omega) = \{X \in L^0(\Omega) \colon X \text{ has a q.c. version and } \lim_{n \to \infty} \hat{\mathbb{E}}\left[|X|^p \mathbb{1}_{\{|X|>n\}}\right] = 0\}.$$

G-expectation turns out to be a good framework to develop stochastic calculus of the Itô type. We can also use G-SDEs and a version of the backward SDEs. As backward equations are a key tool to consider the maximum principle, we now give a short introduction to G-BSDEs and their properties (for simplicity in a one-dimensional case).

Fix two functions $f, g \colon \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\xi \in L_G^p(\Omega_T)$, $p > 2$. We say that the triple $\left(p^G, q^G, K\right)$ is a solution of the G-BSDE with drivers f and g and terminal condition ξ if

$$dp^G(t) = -f\left(t, p^G(t), q^G(t)\right)dt - g\left(t, p^G(t), q^G(t)\right)d\langle B\rangle(t) + q^G(t)dB(t) + dK(t),$$

$$p^G(T) = \xi, \tag{7}$$

where K is a non-increasing G-martingale starting at 0. In (Hu et al. 2014c), the existence and uniqueness of such a G-BSDE are proved under some Lipschitz and regularity conditions on the driver.

Furthermore, under any $\mathbb{P} \in \mathcal{P}_1$, the process p^G is a supersolution of a classical BSDE with drivers f and g and terminal condition ξ on the probability space

$(\Omega, \mathcal{F}, \mathbb{P})$ (we will call such a BSDE a $\mathbb{P}$-BSDE). Hence, by the comparison theorem for supersolutions and solutions, we get

$$p^G(t) \geq p^{\mathbb{P}}(t) \quad \mathbb{P} - a.s.,$$

where $p^{\mathbb{P}}$ is a solution of a $\mathbb{P}$-BSDE. It can also be checked that p^G is minimal in the sense that

$$p^G(t) = \operatorname*{ess\,sup}_{\mathbb{Q} \in \mathcal{P}_1(t, \mathbb{P})}{}^{\mathbb{P}} \, p^{\mathbb{Q}}(t) \quad \mathbb{P} - a.s.,$$

see (Soner et al. 2011c) for this representation. From now on we drop the superscript G in the notation for G-BSDEs whenever this doesn't lead to confusion.

3 A sufficient maximum principle

Let $B(t)$ be a G-Brownian motion with associated sublinear expectation operator $\hat{\mathbb{E}}$. We consider controls u taking values in a closed convex set $U \subset \mathbb{R}$. Let $X(t) = X^u(t)$ be a controlled process of the form

$$dX(t) = b(t, X(t), u(t))dt + \mu(t, X(t), u(t))d\langle B \rangle_t + \sigma(t, X(t), u(t))dB(t); 0 \leq t \leq T,$$
$$X(0) = x \in \mathbb{R}. \tag{8}$$

We assume that the coefficients b, μ, σ are Lipschitz continuous w.r.t. the space variable uniformly in (t, u). Moreover, if the coefficients are not deterministic, they must belong to the space $M_G^2(0, T)$ for each $(x, u) \in \mathbb{R} \times U$.

Let $f : [0, T] \times \mathbb{R} \times U \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be two measurable functions such that f is $\mathcal{C}^1$ w.r.t. the second variable and g is a lower-bounded, differentiable function with quadratic growth such that there exists a constant $C > 0$ and $\epsilon > 0$ s.t.

$$|g'(x)| < C(1 + |x|)^{\frac{1}{1 + \epsilon/2}}.$$

We let $\mathcal{A}$ denote the set of all admissible controls. For u to be in $\mathcal{A}$ we require that $u(t)$ is quasi-continuous for all $t \in [0, T]$ and adapted to $(\mathcal{F}_{(t-\delta)^+})_{t \geq \delta}$, where $\delta \geq 0$ is a given constant. This means that our control u has access only to a delayed information flow. Moreover, we assume that for each $u \in \mathcal{A}$ the following integrability condition is satisfied

$$\hat{\mathbb{E}}\left[\int_0^T f(t, X(t), u(t))dt\right] < \infty.$$

Then, for each $\mathbb{P} \in \mathcal{P}$, the performance functional associated with $u \in \mathcal{A}$ is assumed to be of the form

$$J^{\mathbb{P}}(u) = \mathbb{E}^{\mathbb{P}}\left[\int_0^T f(t, X(t), u(t))dt + g(X(T))\right]. \tag{9}$$

We study the following strongly robust optimal control problem: find $\hat{u} \in \mathcal{A}$ such that

$$\sup_{u \in \mathcal{A}} J^{\mathbb{P}}(u) = J^{\mathbb{P}}(\hat{u}) \quad \forall \mathbb{P} \in \mathcal{P}, \tag{10}$$

where the set $\mathcal{P}$ is introduced in (4). To this end we define the Hamiltonian

$$H(t, x, u, p, q) = f(t, x, u) + \left[b(t, x, u) + \mu(t, x, u) \frac{d\langle B \rangle_t}{dt} \right] p + \sigma(t, x, u) \frac{d\langle B \rangle_t}{dt} q, \tag{11}$$

and the associated G-BSDE with adjoint processes $p(t), q(t), K(t)$ by

$$dp(t) = -\frac{\partial H}{\partial x}(t, X(t), u(t), p(t), q(t))dt + q(t)dB(t) + dK(t); \ 0 \le t \le T,$$
$$p(T) = g'(X(T)). \tag{12}$$

Note that the solution of such a G-BSDE exists thanks to the assumption on the functions f and g and on the definition of the admissible control (see (Hu et al. 2014c) for details).

Theorem 3 *Let $\hat{u} \in \mathcal{A}$ with corresponding solution $\hat{X}(t), \hat{p}(t), \hat{q}(t), \hat{K}(t)$ of (8) And (12) such that $\hat{K} \equiv 0$. Assume that:*

$$(x, u) \to H(t, x, u, \hat{p}(t), \hat{q}(t)) \text{ and } x \to g(x) \text{ are concave for all } t \text{ q.s.}, \tag{13}$$

and

$$\hat{\mathbb{E}} \left[\pm \frac{\partial}{\partial u} H(t, \hat{X}(t), u, \hat{p}(t), \hat{q}(t))|_{u=\hat{u}(t)}|\mathcal{F}_{(t-\delta)^+} \right] = 0 \tag{14}$$

for all t q.s. Then $\hat{u} = u$ is a strongly robust optimal control for the problem (10).

Proof For the sake of simplicity, in the following we adopt the concise notation $f(t) := f(t, X^u(t), u(t)), \hat{f}(t) = f(t, X^{\hat{u}}(t), \hat{u}(t)), X(T) = X^u(T), \hat{X}(T) = X^{\hat{u}}(T)$. Let $u \in \mathcal{A}$ be arbitrary and consider

$$\begin{aligned} \sup_{\mathbb{P} \in \mathcal{P}} \{ J^{\mathbb{P}}(u) - J^{\mathbb{P}}(\hat{u}) \} &= \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}} \left[\int_0^T \left(f(t) - \hat{f}(t) \right) dt + g(X(T)) - g\left(\hat{X}(T) \right) \right] \\ &= \hat{\mathbb{E}} \left[\int_0^T \left(f(t) - \hat{f}(t) \right) dt + g(X(T)) - g\left(\hat{X}(T) \right) \right] \\ &= \hat{\mathbb{E}}[I_1 + I_2], \end{aligned} \tag{15}$$

where J is introduced in (2) and

$$I_1 := \int_0^T \left(f(t) - \hat{f}(t) \right) dt, \quad I_2 := g(X(T)) - g\left(\hat{X}(T) \right).$$

By the definition of H, we can write

$$\begin{aligned} I_1 = \int_0^T \Bigg\{ H(t) - \hat{H}(t) - \left[b(t) - \hat{b}(t) + (\mu(t) - \hat{\mu}(t)) \frac{d\langle B \rangle_t}{dt} \right] \\ \hat{p}(t) - [\sigma(t) - \hat{\sigma}(t)] \frac{d\langle B \rangle_t}{dt} \hat{q}(t) \Bigg\} dt. \end{aligned} \tag{16}$$

By the concavity of g, (12), and the Itô formula, we have

$$I_2 \leq g'\left(\hat{X}(T)\right)\left(X(T) - \hat{X}(T)\right) = \hat{p}(T)\left(X(T) - \hat{X}(T)\right)$$

$$= \int_0^T \hat{p}(t)d\left(X(t) - \hat{X}(t)\right) + \int_0^T \left(X(t) - \hat{X}(t)\right)d\hat{p}(t) + \int_0^T d\left\langle \hat{p}, X - \hat{X}\right\rangle(t)$$

$$= \int_0^T \hat{p}(t)\left[b(t) - \hat{b}(t) + (\mu(t) - \hat{\mu}(t))\frac{d\langle B\rangle_t}{dt}\right]dt$$

$$+ \int_0^T \left(X(t) - \hat{X}(t)\right)\left(-\frac{\partial \hat{H}}{\partial x}(t)\right)dt + \int_0^T \left[\sigma(t) - \hat{\sigma}(t)\right]\frac{d\langle B\rangle_t}{dt}\hat{q}(t)dt$$

$$+ \int_0^T \hat{p}\left[\sigma(t) - \hat{\sigma}(t)\right]dB(t) + \int_0^T \left[X(t) - \hat{X}(t)\right]\hat{q}(t)dB(t).$$

$$(17)$$

Adding (16) and (17) and using the concavity of H we get, by the sublinearity of the G-expectation and by (15), that

$$\sup_{\mathbb{P} \in \mathcal{P}}\{J^{\mathbb{P}}(u) - J^{\mathbb{P}}(\hat{u})\} \leq \hat{\mathbb{E}}\left[\int_0^T \left(\hat{p}(t)\left[\sigma(t) - \hat{\sigma}(t)\right] + \left[X(t) - \hat{X}(t)\right]\hat{q}(t)\right)dB(t)\right]$$

$$+ \hat{\mathbb{E}}\left[\int_0^T \left[H(t) - \hat{H}(t) - \frac{\partial \hat{H}}{\partial x}(t)\left(X(t) - \hat{X}(t)\right)\right]dt\right]$$

$$\leq \hat{\mathbb{E}}\left[\int_0^T \frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t))dt\right]$$

$$\leq \int_0^T \hat{\mathbb{E}}\left[\frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t))\right]dt$$

$$\leq \int_0^T \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t))|\mathcal{F}_{(t-\delta)^+}\right]\right]dt$$

$$\leq \int_0^T \hat{\mathbb{E}}\left[\hat{\mathbb{E}}\left[\frac{\partial \hat{H}}{\partial u}(t)|\mathcal{F}_{(t-\delta)^+}\right](u(t) - \hat{u}(t))^+\right.$$

$$+ \hat{\mathbb{E}}\left[-\frac{\partial \hat{H}}{\partial u}(t)|\mathcal{F}_{(t-\delta)^+}\right](u(t) - \hat{u}(t))^-\right]dt = 0,$$

since $u = \hat{u}$ is a critical point of the Hamiltonian. This proves that $\hat{u} := \hat{u}$ is optimal.

$\square$

Remark 1 *Note that if $\delta = 0$, we can slightly relax the assumption in Eq. (14) by only requiring that*

$$\max_{v \in U} H(t, \hat{X}(t), v, \hat{p}(t), \hat{q}(t))] = H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t)).$$

4 A necessary maximum principle for the full-information case

A drawback of the previous result is that the concavity conditions are not satisfied in many applications. Therefore, it is of interest to have a maximum principle, which does not need this condition. Moreover, the requirement that the non-increasing G-martingale $\hat{K}$ disappears from the adjoint equation for the optimal control $\hat{u}$ is a very strong assumption, which, however, is crucial in the proof. In this section, we prove a result which doesn't depend on the concavity of the Hamiltonian. Moreover, in the Merton problem we show that the necessary maximum principle might be obtained without the assumption on the process $\hat{K}$. We make the following assumptions.

A1.   For all $u, \beta \in \mathcal{A}$ with β bounded, there exists $\delta > 0$ such that

$$u + a\beta \in \mathcal{A} \quad \text{for all } a \in (-\delta, \delta).$$

A2.   For all t, h such that $0 \leq t < t + h \leq T$ and all bounded random variables $\alpha \in L_G^1(\Omega_t)^3$, the control

$$\beta(s) := \alpha \mathbb{1}_{[t,t+h]}(s)$$

belongs to $\mathcal{A}$.

Remark 2 *Note that given $u, \beta \in \mathcal{A}$ with β bounded, the derivative process*

$$Y(t) := \frac{d}{da} X^{u+a\beta}(t)$$

exists, $Y(0) = 0$, and

$$
dY(t) = \left\{ \frac{\partial b}{\partial x}(t)Y(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right\} dt
$$
$$
+ \left\{ \frac{\partial \mu}{\partial x}(t)Y(t) + \frac{\partial \mu}{\partial u}(t)\beta(t) \right\} d\langle B \rangle_t + \left\{ \frac{\partial \sigma}{\partial x}(t)Y(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dB(t).
$$

This follows by the smoothness ($\mathcal{C}^1$) assumptions on the coefficients b, μ, σ and the Itô formula.

Before we give the necessary maximum principle for this problem, we will state the following remark showing that it is sufficient to consider just a control which is optimal for all $\mathbb{P} \in \mathcal{P}_1$.

Remark 3 *Note that if $\hat{u} \in \mathcal{A}$ is a strongly robust optimal control, it is of course the optimal control for the following problem:*

$$\sup_{u \in \mathcal{A}} J^{\mathbb{P}}(u) = J^{\mathbb{P}}(\hat{u}) \quad \forall \mathbb{P} \in \mathcal{P}_1. \tag{18}$$

However, we have also the opposite, thanks to the conditions on the set of admissible

---

[3]It is easy to see that for a fixed $\mathbb{P} \in \mathcal{P}$ the set of all bounded random variables from space $L_G^1(\Omega)$ is dense in the space $L_{\mathbb{P}}^p(\Omega_t)$ under the norm $(\mathbb{E}^{\mathbb{P}}[|.|^p])^{1/p}$ for any $p \geq 1$.

controls $\mathcal{A}$. Namely, if $\hat{u}$ satisfies (18), then we have that for any fixed $u \in \mathcal{A}$ and any $\mathbb{P} \in \mathcal{P}_1$

$$0 \geq \mathbb{E}^{\mathbb{P}} \left[\int_0^T f\left(t, X^u(t), u(t)\right) dt - \int_0^T f\left(t, X^{\hat{u}}(t), \hat{u}(t)\right) dt \right.$$
$$\left. + g(X^u(T)) - g\left(X^{\hat{u}}(T)\right) \right].$$

We use again the shortened notation $\hat{f}(t) := f\left(t, X^{\hat{u}}(t), \hat{u}(t)\right)$ and $f(t) := f(t, X^u(t), u(t))$ and conclude that

$$0 \geq \sup_{\mathbb{P} \in \mathcal{P}_1} \mathbb{E}^{\mathbb{P}} \left[\int_0^T f(t)dt - \int_0^T \hat{f}(t)dt + g(X^u(T)) - g\left(X^{\hat{u}}(T)\right) \right].$$

Note that due to the conditions on the admissible controls we know that the random variables: $\int_0^T f(t)dt$, $\int_0^T \hat{f}(t)dt$, $g(X^u(T))$, and $g\left(X^{\hat{u}}(T)\right)$ belong to $L_G^1(\Omega)$, hence by the representation of the G-expectation, we have

$$0 \geq \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}} \left[\int_0^T f(t)dt - \int_0^T \hat{f}(t)dt + g(X^u(T)) - g\left(X^{\hat{u}}(T)\right) \right]$$

and that implies by Proposition 1 that $\hat{u}$ is a strongly robust optimal control.

Lemma 1 *Assume that A1, A2 hold and that $\hat{u}$ is an optimal control for the performance functional*

$$u \to J^{\mathbb{P}}(u)$$

for some probability measure $\mathbb{P} \in \mathcal{P}_1$. Consider the adjoint equation as a BSDE under the probability measure $\mathbb{P}$:

$$d\hat{p}^{\mathbb{P}}(t) = -\frac{\partial H}{\partial x}\left(t, \hat{X}(t), \hat{u}(t), \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t)\right) dt + \hat{q}^{\mathbb{P}}(t)dB(t); \ 0 \leq t \leq T,$$
$$\hat{p}^{\mathbb{P}}(T) = g'\left(\hat{X}(T)\right) \quad \mathbb{P} - a.s. \tag{19}$$

Then

$$\frac{\partial \hat{H}^{\mathbb{P}}}{\partial u}(t) := \frac{\partial}{\partial u} H\left(t, \hat{X}(t), u, \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t)\right) |_{u=\hat{u}(t)} = 0.$$

Proof Consider

$$\frac{d}{da} J^{\mathbb{P}}(u + a\beta) = \frac{d}{da} \mathbb{E}^{\mathbb{P}} \left[\int_0^T f\left(t, X^{u+a\beta}(t), u(t)\right) dt + g\left(X^{u+a\beta}(T)\right) \right]$$

$$= \lim_{a \to 0} \frac{1}{a} \mathbb{E}^{\mathbb{P}} \left[\int_0^T f\left(t, X^{u+a\beta}(t), u(t)\right) dt + g(X^{u+a\beta}(T)) \right]$$

$$- \mathbb{E}^{\mathbb{P}} \left[\int_0^T f(t, X(t), u(t)) dt + g(X(T)) \right]$$

$$= \lim_{a \to 0} \mathbb{E}^{\mathbb{P}} \left[\int_0^T \frac{1}{a} \left\{ f(t, X^{u+a\beta}(t), u(t)) - f(t, X(t), u(t)) \right\} \right. \tag{20}$$

$$\left. dt + \frac{1}{a} \left\{ g(X^{u+a\beta}(T)) - g(X(T)) \right\} \right]$$

$$= \mathbb{E}^{\mathbb{P}} \left[\int_0^T \left(\frac{\partial f}{\partial x}(t, X(t), u(t)) Y(t) + \frac{\partial f}{\partial u}(t, X(t), u(t)) \beta(t) \right) \right.$$

$$\left. dt + g'(X(T)) Y(T) \right].$$

By the Itô formula,

$$\mathbb{E}^{\mathbb{P}} \left[g'(X(T)) Y(T) \right] = \mathbb{E}^{\mathbb{P}} \left[p(T) Y(T) \right]$$

$$= \mathbb{E}^{\mathbb{P}} \left[\int_0^T p^{\mathbb{P}}(t) dY(t) + \int_0^T Y(t) dp^{\mathbb{P}}(t) + \int_0^T q^{\mathbb{P}}(t) \left\{ \frac{\partial \sigma}{\partial x}(t) Y(t) + \frac{\partial \sigma}{\partial u}(t) \beta(t) \right\} d\langle B \rangle_t \right]$$

$$\leq \mathbb{E}^{\mathbb{P}} \left[\int_0^T p^{\mathbb{P}}(t) \left\{ \frac{\partial b}{\partial x}(t) Y(t) + \frac{\partial b}{\partial u}(t) \beta(t) \right\} dt + \int_0^T p^{\mathbb{P}}(t) \left\{ \frac{\partial \mu}{\partial x}(t) Y(t) + \frac{\partial \mu}{\partial u}(t) \beta(t) \right\} d\langle B \rangle_t \right.$$

$$+ \int_0^T Y(t) \left(-\frac{\partial \hat{H}^{\mathbb{P}}}{\partial x}(t) \right) dt + \int_0^T q^{\mathbb{P}}(t) \left\{ Y(t) \frac{\partial \sigma}{\partial x}(t) + \frac{\partial \sigma}{\partial u}(t) \beta(t) \right\} d\langle B \rangle_t \right]$$

$$= \mathbb{E}^{\mathbb{P}} \left[\int_0^T Y(t) \left\{ p^{\mathbb{P}}(t) \left(\frac{\partial b}{\partial x}(t) + \frac{\partial \mu}{\partial x}(t) \frac{d\langle B \rangle_t}{dt} \right) + q^{\mathbb{P}}(t) \frac{\partial \sigma}{\partial x}(t) \frac{d\langle B \rangle_t}{dt} - \frac{\partial H^{\mathbb{P}}}{\partial x}(t) \right\} dt \right.$$

$$\left. + \int_0^T \beta(t) \left\{ p^{\mathbb{P}}(t) \left(\frac{\partial b}{\partial u}(t) + \frac{\partial \mu}{\partial u}(t) \frac{d\langle B \rangle_t}{dt} \right) + q^{\mathbb{P}}(t) \frac{\partial \sigma}{\partial u}(t) \frac{d\langle B \rangle_t}{dt} \right\} dt \right]. \tag{21}$$

Adding (20) and (21), we get

$$\frac{d}{da} J^{\mathbb{P}}(u + a\beta) \leq \mathbb{E}^{\mathbb{P}} \left[\int_0^T \beta(t) \frac{\partial H^{\mathbb{P}}}{\partial u}(t) dt \right].$$

If $\hat{u}$ is an optimal control, then the above gives

$$0 = \frac{d}{da} J^{\mathbb{P}}(\hat{u} + a\beta) \leq \mathbb{E}^{\mathbb{P}} \left[\int_0^T \beta(t) \frac{\partial \hat{H}^{\mathbb{P}}}{\partial u}(t) dt \right]$$

for all bounded $\beta \in \mathcal{A}$. Applying this to both β and $-\beta$, we conclude that

$$\mathbb{E}^{\mathbb{P}} \left[\int_0^T \beta(t) \frac{\partial \hat{H}}{\partial u}(t) dt \right] = 0.$$

By A2 together with the footnote about the denseness, we can then proceed to deduce that

$$\frac{\partial \hat{H}^{\mathbb{P}}}{\partial u}(t) = 0 \quad \mathbb{P} - a.s.$$

$\square$

Using the lemma, we can easily get the following necessary maximum principle.

Theorem 4 *Assume that A1, A2 hold and that $\hat{u}$ is a strongly robust optimal control for the performance functional*

$$u \to J^{\mathbb{P}}(u)$$

for every probability measure $\mathbb{P} \in \mathcal{P}_1$. Consider the adjoint equation as a G-BSDE:

$$d\hat{p}^G(t) = -\frac{\partial H}{\partial x}(t, \hat{X}(t), \hat{u}(t), \hat{p}^G(t), \hat{q}^G(t))dt + \hat{q}^G(t)dB(t) + d\hat{K}(t); \ 0 \le t \le T,$$
$$\hat{p}^G(T) = g'(\hat{X}(T)) \quad q.s.$$

$$(22)$$

If $\hat{K} \equiv 0$ q.s., then

$$\frac{\partial \hat{H}^G}{\partial u}(t) := \frac{\partial}{\partial u} H(t, \hat{X}(t), u, \hat{p}^G(t), \hat{q}^G(t))|_{u=\hat{u}(t)} = 0, \ q.s. \quad (23)$$

Proof We now prove that the relation in (23) holds for every $\mathbb{P} \in \mathcal{P}_1$. Fix $\mathbb{P} \in \mathcal{P}_1$. If $\hat{K} \equiv 0$ q.s., then by the uniqueness of the solution of $\mathbb{P}$-BSDE, we get that $\hat{p}^G \equiv \hat{p}^{\mathbb{P}} \ \mathbb{P} - a.s.$ and $\hat{q}^G \equiv \hat{q}^{\mathbb{P}} \ \mathbb{P} - a.s.$ But by Lemma 1, we know that $\hat{u}$ is a $\mathbb{P} - a.s.$ critical point of $\hat{H}^{\mathbb{P}}(t)$ hence also $\hat{H}^G(t)$. By the arbitrariness of $\mathbb{P} \in \mathcal{P}_1$, we get

$$\frac{\partial}{\partial u} H(t, \hat{X}(t), u, \hat{p}^G(t), \hat{q}^G(t))|_{u=\hat{u}(t)} = 0 \ \forall \mathbb{P} \in \mathcal{P}_1.$$

We get the assertion of the theorem by stating a general fact that if $\xi, \eta \in L_G^1(\Omega)$ and $\xi = \eta \ \mathbb{P} - a.s.$ for all $\mathbb{P} \in \mathcal{P}_1$, then $\xi = \eta$ q.s. □

As we mentioned at the beginning of this section, the assumption on the process $\hat{K}$ is a big disadvantage. However, if we limit our considerations to Merton-type problems, we are able to show the necessary maximum principle without this assumption.

Theorem 5 *Assume that*

1. *A1, A2 hold,*
2. *$b \equiv 0$, $\mu(t, x, u) = \psi(x)l(u)m(t)$ and $\sigma(t, x, u) = \psi(x)h(u)s(t)$ for $\psi, l, h \in \mathcal{C}^1(\mathbb{R})$ and some bounded processes m and s such that for each $t \in [0, T]$ $m(t)$ and $s(t)$ are quasi-continuous. Moreover, let $c(s(t) = 0) = 0$ for all $t \in [0, T]$,*
3. *$f \equiv 0$,*
4. *$X(0) = x \neq 0$.*

Let $\hat{u}$ be a strongly robust optimal control for the performance functional

$$u \to J^{\mathbb{P}}(u)$$

for every probability measure $\mathbb{P} \in \mathcal{P}_1$. If $c(l(\hat{u}(t)) = 0) = 0$, $c(h(\hat{u}(t)) = 0) = 0$, $c(h'(\hat{u}(t)) = 0) = 0$, $c(\psi(\hat{X}(t)) = 0) = 0$, $c(\psi'(\hat{X}(t)) = 0) = 0$, and

$c(l(\hat{u}(t))h'(\hat{u}(t)) - l'(\hat{u}(t))h(\hat{u}(t)) \neq 0) = 0^4$ *for all* $t \in [0, T]$, *then*

$$\frac{\partial \hat{H}^G}{\partial u}(t) := \frac{\partial}{\partial u} H(t, \hat{X}(t), u, \hat{p}^G(t), \hat{q}^G(t)) = 0, \ q.s. \tag{24}$$

Proof Fix a probability measure $\mathbb{P} \in \mathcal{P}_1$. By Lemma 1, we know that $\hat{u}$ is a critical point ($\mathbb{P}$-a.s.) of the Hamiltonian

$$\frac{\partial}{\partial u} H(t, \hat{X}(t), \hat{u}, \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t)) = 0, \ \forall t \in [0, T].$$

Using this fact, we get

$$0 = \frac{\partial}{\partial u} H(t, \hat{X}(t), \hat{u}, \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t))$$

$$= \psi(\hat{X}(t)) \left[l'(\hat{u}(t))m(t)\hat{p}^{\mathbb{P}}(t) + h'(\hat{u}(t))s(t)\hat{q}^{\mathbb{P}}(t) \right] \frac{d\langle B \rangle(t)}{dt}.$$

By the assumption on process s and h', we compute that

$$\hat{q}^{\mathbb{P}}(t) = -\frac{m(t)}{s(t)} \frac{l'(\hat{u}(t))}{h'(\hat{u}(t))} \hat{p}^{\mathbb{P}}(t).$$

We have that

$$\frac{\partial}{\partial x} H(t, \hat{X}(t), \hat{u}(t), \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t)) = \psi'(\hat{X}(t)) \left[l(\hat{u}(t))m(t)\hat{p}^{\mathbb{P}}(t) \right.$$

$$\left. + h(\hat{u}(t))s(t)\hat{q}^{\mathbb{P}}(t) \right] \frac{d\langle B \rangle(t)}{dt}$$

$$= \psi'(\hat{X}(t))\hat{p}^{\mathbb{P}}(t)m(t) \left[l(\hat{u}(t)) \right.$$

$$\left. - h(\hat{u}(t)) \frac{l'(\hat{u}(t))}{h'(\hat{u}(t))} \right] \frac{d\langle B \rangle(t)}{dt}$$

$$= 0$$

since $c(l(\hat{u})h'(\hat{u}) - l'(\hat{u})h(\hat{u}) \neq 0) = 0$ by hypothesis. But then we see that $\hat{p}^{\mathbb{P}}$ has dynamics

$$d\hat{p}^{\mathbb{P}}(t) = -\frac{\partial}{\partial x} H(t, \hat{X}(t), \hat{u}(t), \hat{p}^{\mathbb{P}}(t), \hat{q}^{\mathbb{P}}(t))dt + \hat{q}^{\mathbb{P}}(t)dB(t)$$

$$= -\frac{m(t)}{s(t)} \frac{l'(\hat{u}(t))}{h'(\hat{u}(t))} \hat{p}^{\mathbb{P}}(t)dB(t).$$

Hence,

$$\hat{p}^{\mathbb{P}}(t) = \mathbb{E}^{\mathbb{P}}[g'(\hat{X}(T))|\mathcal{F}_t] \quad \mathbb{P} - a.s.$$

We also remember that

$$\hat{p}^G(t) = \underset{\mathbb{Q} \in \mathcal{P}_1(t, \mathbb{P})}{\text{ess sup}} {}^{\mathbb{P}}\hat{p}^{\mathbb{Q}}(t) = \underset{\mathbb{Q} \in \mathcal{P}_1(t, \mathbb{P})}{\text{ess sup}} {}^{\mathbb{P}}\mathbb{E}^{\mathbb{Q}}[g'(\hat{X}(T))|\mathcal{F}_t] \quad \mathbb{P} - a.s.$$

---

[4]Note that this condition is satisfied if $l(x) = ah(x)$ for some constant $a \in \mathbb{R}$. However, it may also include other cases, since it could be that the span of $\hat{u}(t)$, $t \in [0, T]$, is not necessarily the entire $\mathbb{R}$.

Thus, by the characterization of the conditional G-expectation in (6), we obtain that $\hat{p}^G(t)$ is a G-martingale with representation

$$\hat{p}^G(t) = \hat{\mathbb{E}}[g'(\hat{X}(T))|\mathcal{F}_t] = \hat{\mathbb{E}}[g'(\hat{X}(T))] + \int_0^t \hat{q}^G(s)dB(s) + \hat{K}(t) \quad q.s.$$

and consequently it has dynamics

$$d\hat{p}^G(t) = \hat{q}^G(t)dB(t) + d\hat{K}(t).$$

But in that case we know that for almost all $t \in [0, T]$, we must have that

$$
\begin{aligned}
0 &= \frac{\partial}{\partial x} H\left(t, \hat{X}(t), \hat{u}(t), \hat{p}^G(t), \hat{q}^G(t)\right) \\
&= \psi'\left(\hat{X}(t)\right)\left[l(\hat{u})m(t)\hat{p}^G(t) + h(\hat{u})s(t)\hat{q}^G(t)\right]\frac{d\langle B\rangle(t)}{dt}
\end{aligned}
\tag{25}
$$

q.s. By assumption on $\psi'(\hat{X})$, we conclude that

$$l(\hat{u})m(t)\hat{p}^G(t) + h(\hat{u})s(t)\hat{q}^G(t) = 0 \; q.s.$$

Hence,

$$\hat{q}^G(t) = -\frac{m(t)}{s(t)}\frac{l(\hat{u}(t))}{h(\hat{u}(t))}\hat{p}^G(t),$$

since $c(h(\hat{u}(t)) = 0) = 0$ for all $t \in [0, T]$, and we can easily check then that

$$\frac{\partial}{\partial u} H\left(t, \hat{X}(t), \hat{u}, \hat{p}^G(t), \hat{q}^G(t)\right) = 0.$$

$\square$

5 Examples

We now consider some examples to illustrate the previous results. In the following, we assume to work with a one-dimensional G-Brownian motion with operator G of the form

$$G(a) := \frac{1}{2}(a^+ - \underline{\sigma}^2 a^-), \quad \underline{\sigma}^2 > 0,
\tag{26}$$

i.e., with quadratic variation $\langle B\rangle(t)$ lying within the bounds $\underline{\sigma}^2 t$ and t.

5.1 Example I

Consider

$$dX(t) = dB(t) - c(t)dt,
\tag{27}$$

where $c(t)$, $t \in [0, T]$, is stochastic process such that $c(t) \in L_G^1(\Omega_t)$ for all $t \in [0, T]$. We wish to solve the optimal control problem for every $\mathbb{P} \in \mathcal{P}$ under the performance criterion

$$J^{\mathbb{P}}(c) = \mathbb{E}^{\mathbb{P}}\left[\int_0^T \ln c(t)dt + X(T)\right].
\tag{28}$$

In the notation of Section 3, we have chosen here $f(t, x, c) = \ln c$ and $g(x) = x$, i.e., $g'(x) = 1$. Then, the Hamiltonian is given by

$$H(t, x, c, p, q) = \ln c + q \frac{d\langle B \rangle_t}{dt} - cp, \tag{29}$$

and by (12), we obtain

$$\begin{aligned} dp(t) &= q(t)dB(t); \ 0 \le t \le T, \\ p(T) &= g'(X(T)) = 1, \end{aligned} \tag{30}$$

i.e., $q = 0$, $p = 1$. Furthermore, by (29), we have

$$\frac{\partial H}{\partial c} = \frac{\partial}{\partial c}[\ln c - cp] = \frac{1}{c} - p,$$

i.e., $\hat{c}(t) = 1$, $t \in [0, T]$, is a strongly robust optimal control by Theorem 3.

Note that by the proof we could choose a general utility function instead of a logarithmic utility function without losing the existence of the strongly robust optimal control.

5.2 Example II

Consider

$$dX(t) = X(t)[b(t)dt + dB(t)] - c(t)dt, \tag{31}$$

and problem (28). Here, $b(t)$ is a deterministic measurable function. Then, the Hamiltonian is given by

$$H(t, x, c, p, q) = \ln c + xq \frac{d\langle B \rangle_t}{dt} + (xb(t) - c))p. \tag{32}$$

Here,

$$\begin{aligned} dp(t) &= -\left(b(t)p(t) + q(t)\frac{d\langle B \rangle_t}{dt} \right)dt + q(t)dB(t); \ 0 \le t \le T, \\ p(T) &= g'(X(T)) = 1. \end{aligned} \tag{33}$$

Let $q = 0$, then

$$\begin{aligned} dp(t) &= -b(t)p(t)dt, \\ p(T) &= 1, \end{aligned}$$

i.e., $p(t) = \exp \int_t^T b(s)ds$ and $\hat{c}(t) = \frac{1}{p(t)}$ is a strongly robust optimal control by Theorem 3.

5.3 Example III

Consider the Merton-type problem with the logarithmic utility, i.e., let

$$dX^u(t) = X^u(t)\left[m(t)u(t)d\langle B \rangle(t) + s(t)u(t)dB(t)\right],$$

where $u(t) \in L_G^2(\Omega_t)$ for all $t \in [0, T]$ and m and s are two deterministic functions. Assume that $s(t) \ne 0$ for all $t \in [0.T]$. We are interested in finding a strongly

robust optimal control problem for the family of probability measures $\mathcal{P}$ with the performance criterion given by

$$J^{\mathbb{P}}(u) := \mathbb{E}^{\mathbb{P}}[\ln X^u(T)].$$

The Hamiltonian associated with this problem is given by

$$H(t, x, u, p, q) = xu[m(t)p + s(t)q]\frac{d\langle B\rangle}{dt}(t) \tag{34}$$

and for each admissible control u we consider adjoint G-BSDE of the form

$$dp(t) = -u(t)[m(t)p(t) + s(t)q(t)]d\langle B\rangle(t) + g(t)dB(t) + dK(t)$$
$$p(T) = X^{-1}(T).$$

Note that the adjoint equation is linear, hence by Remark 3.3 in (Hu et al. 2014a), we obtain the representation formula for the solution

$$p(t) = X^{-1}(t)\hat{\mathbb{E}}\left[X(T)X^{-1}(T)|\mathcal{F}_t\right] = X^{-1}(t).$$

Moreover, by the dynamics of X^{-1}, we deduce that

$$q(t) = -u(t)s(t)p(t), \quad K \equiv 0.$$

Plugging this solution into the Hamiltonian (34), we get that

$$H(t, X^u(t), v, p(t), q(t)) = X^u(t)v\left[m(t) - u(t)s^2(t)\right]p(t)\frac{\langle B\rangle}{dt}(t),$$

hence the critical point of the Hamiltonian must satisfy

$$\hat{u}(t) = \frac{m(t)}{s^2(t)}$$

and this is our strongly robust optimal control.

Note that we can also solve this problem directly by omega-wise maximization, without using the maximum principle and G-BSDEs. In fact, we may consider more general dynamics in X

$$dX^u(t) = X^u(t)[b(t)u(t)dt + m(t)u(t)d\langle B\rangle(t) + s(t)u(t)dB(t)]$$

and by direct computation it might be checked that the strongly robust optimal control takes the form

$$\hat{u}(t) = \frac{b(t) + m(t)\frac{d\langle B\rangle}{dt}(t)}{s^2(t)\frac{d\langle B\rangle}{dt}(t)}.$$

However, it is important to note that this control is no longer quasi-continuous (see (Song 2012)) and it doesn't make sense to consider G-BSDEs associated with such a control.

Finally, note that the classical robust optimal control for this problem would be $u^*(t) = m(t)/\underline{\sigma}^2$. It is clear that this control ignores the flow of information about the volatility path and instead it just sticks to its worst-case scenario assumption. It makes sense to assume the worst-case scenario at time 0, but later one should rather update its view about the past volatility, which is not done for the classical robust optimal controls.

5.4 Example IV

As another example of a problem which admits a strongly robust optimal control, consider an LQ-problem, in which the state equation has the linear dynamics

$$dX(t) = (F(t)X(t)+G(t)u(t)+\mu(t))dt+\sigma(t)dB(t); \ 0 \le t \le T; \quad X(0) = x \in \mathbb{R};$$
(35)

for $u \in \mathcal{A}$ as described in the beginning of Section 3. The performance functional is quadratic

$$J^{\mathbb{P}}(u) := \frac{1}{2}\mathbb{E}^{\mathbb{P}}\left[\int_0^T (Q(t)X^2(t) + R(t)u^2(t))dt + LX^2(T)\right].$$
(36)

Here, F, G, μ, σ, Q, R are continuous deterministic functions on $[0, T]$, $Q(t) > 0$, $R(t) > 0$ and $L > 0$ is a constant.

We want to find $\hat{u} \in \mathcal{A}$ (as described in Section 3) which maximizes $J^{\mathbb{P}}(u)$ over all $u \in \mathcal{A}$ for all $\mathbb{P} \in \mathcal{P}$.

In this case, the Hamiltonian in (11) gets the form

$$H(t, x, u, p, q) = \frac{1}{2}Q(t)x^2 + \frac{1}{2}R(t)u^2 + [F(t)x+G(t)u+\mu(t)]p + \sigma(t)\frac{d\langle B\rangle(t)}{dt}q$$
(37)

and the adjoint BSDE (12) becomes

$$dp(t) = -[Q(t)X(t)+F(t)]\,dt + q(t)dB(t) + dK(t), \ 0 \le t \le T, \quad p(T) = LX(T).$$
(38)

We intend to apply Theorem 3 and note that

$$\frac{\partial}{\partial u}H(t, \hat{X}(t), u, \hat{p}(t), \hat{q}(t))|_{u=\hat{u}(t)} = R(t)\hat{u}(t) + G(t)\hat{p}(t),$$
(39)

which is 0 when

$$\hat{u}(t) = -\frac{G(t)\hat{p}(t)}{R(t)},$$
(40)

where $\hat{p}(t)$ refers to the solution of (38) when $u = \hat{u}$ is applied to the BSDE.

Let us guess that (38) with $u = \hat{u}$ admits the solution of the form

$$\hat{p}(t) = S(t)\hat{X}(t) + Z(t), \ \hat{q}(t) = S(t)\sigma(t), \ dK(t) = 0$$
(41)

for some deterministic functions $S, Z \in \mathcal{C}^1(\mathbb{R}_+)$ to be determined.

We apply the Itô formula to the equation for $\hat{p}$ in (41) and plug in the candidate for optimal control from (40). By comparison with (38), after some simple computations we get that

$$\hat{X}(t)[S'(t) - G^2(t)S^2(t)/R(t) + S(t)F(t) + Q(t)]$$
$$+ Z'(t) - G^2(t)S(t)Z(t)/R(t) + S(t)\mu(t) + F(t) = 0,$$

i.e., (41) is indeed the solution of the adjoint Eq. 38 if S satisfies the Riccati equation

$$S'(t) - \frac{G^2(t)}{R(t)}S^2(t) + F(t)S(t) + Q(t) = 0, \ 0 \le t \le T; \quad S(T) = L,$$
(42)

and Z satisfies the linear differential equation

$$Z'(t) - \frac{G^2(t)}{R(t)} S(t) Z(t) + S(t)\mu(t) + F(t) = 0, \quad 0 \leq t \leq T; \quad Z(T) = 0. \quad (43)$$

By Theorem 3, we conclude that

$$\hat{u}(t) = -\frac{G(t)}{R(t)} \left(S(t)\hat{X}(t) + Z(t) \right) \quad (44)$$

is a strongly optimal control with S and Z given by (42) and (43), respectively.

6 Counterexample: the Merton problem with the power utility

In this example, we consider the Merton problem with the power utility and show that generally we cannot drop the assumption $\hat{K} \equiv 0$ without losing the strong sense of the optimality. First, we solve the classical robust utility maximization problem and then we prove that the optimal control for that problem is optimal usually only in a weaker sense, i.e., there exists a probability measure $\mathbb{P} \in \mathcal{P}$ such that the control is not optimal under $\mathbb{P}$, even though the control satisfies all the conditions of the sufficient maximum principle with the exception of $\hat{K} \equiv 0$.

Consider first the classical robust utility maximization problem

$$u \mapsto \hat{J}(u) := \hat{\mathbb{E}} \left[\int_0^T f(t, X(t), u(t))dt + g(X(T)) \right],$$

where X has dynamics for any $u \in \mathcal{A}$

$$dX(t) = m(t)X(t)u(t)d\langle B \rangle(t) + s(t)X(t)u(t)dB(t).$$

Then,

$$X(t) = x \exp \left\{ \int_0^t s(r)u(r)dB(r) + \int_0^t \left[m(r)u(r) - \frac{1}{2}s^2(r)u^2(r) \right] d\langle B \rangle(r) \right\}.$$

We assume that m and s are bounded and deterministic and $s \neq 0$. Put $f \equiv 0$ and $g(x) = \frac{1}{\alpha}x^\alpha$, $\alpha \in]0, 1[$. Hence,

$$\hat{J}(u) = \frac{x}{\alpha}\hat{\mathbb{E}} \left[\exp \left\{ \alpha \int_0^T s(r)u(r)dB(r) + \alpha \int_0^T \left[m(r)u(r) - \frac{1}{2}s^2(r)u^2(r) \right] d\langle B \rangle(r) \right\} \right]$$

$$= \frac{x}{\alpha}\hat{\mathbb{E}} \left[\exp\{\alpha \int_0^T s(r)u(r)dB(r) - \frac{\alpha^2}{2} \int_0^T s^2(r)u^2(r)d\langle B \rangle(r)\} \cdot \right.$$

$$\left. \cdot \exp \left\{ \int_0^T \left[\alpha m(r)u(r) + \frac{\alpha^2 - \alpha}{2}s^2(r)u^2(r) \right] d\langle B \rangle(r) \right\} \right].$$

We now use the Girsanov theorem for G-expectation and the G-martingale

$$M(t) := \exp\{\alpha \int_0^t s(r)u(r)dB(r) - \frac{\alpha^2}{2} \int_0^t s^2(r)u^2(r)d\langle B \rangle(r)\},$$

see Section 5.2 in (Hu et al. 2014a). We get the sublinear expectation $\hat{\mathbb{E}}^u$ under which the process $B^u(t) := B(t) - \int_0^t s(r)u(r)d\langle B\rangle(r)$ is a G-Brownian motion. Note that

$$\langle B^u\rangle(t) = \langle B\rangle(t) \tag{45}$$

q.s. Moreover, it is easy to check that the deterministic control

$$\hat{u}(r) = \frac{m(r)}{(1-\alpha)s^2(r)}$$

is a maximizer of the following function

$$u \mapsto \alpha m(r)u + \frac{\alpha^2-\alpha}{2}s^2(r)u^2.$$

Hence, we get that

$$
\begin{aligned}
\hat{J}(u) &= \frac{x}{\alpha}\hat{\mathbb{E}}^u\left[\exp\left\{\int_0^T\left[\alpha m(r)u(r)+\frac{\alpha^2-\alpha}{2}s^2(r)u^2(r)\right]d\langle B\rangle(r)\right\}\right] \\
&\leq \frac{x}{\alpha}\hat{\mathbb{E}}^u\left[\exp\left\{\int_0^T\left[\alpha m(r)\hat{u}(r)+\frac{\alpha^2-\alpha}{2}s^2(r)(\hat{u})^2(r)\right]d\langle B^u\rangle(r)\right\}\right] \\
&= \frac{x}{\alpha}\hat{\mathbb{E}}^{\hat{u}}\left[\exp\left\{\int_0^T\left[\alpha m(r)\hat{u}(r)+\frac{\alpha^2-\alpha}{2}s^2(r)(\hat{u})^2(r)\right]d\langle B^{\hat{u}}\rangle(r)\right\}\right] = \hat{J}(\hat{u}).
\end{aligned}
\tag{46}
$$

The last equalities are a consequence of (45) and of the fact that the integrand is deterministic and that B^u and $B^{\hat{u}}$ are G-Brownian motions under $\mathbb{E}^u$ and $\mathbb{E}^{\hat{u}}$, respectively. Eq. 46 shows then that $\hat{u}$ is an optimal control for this weaker optimization problem.

Now, consider the adjoint equation related to $\hat{u}$ in terms of a G-BSDE. The backward equation is linear due to linearity of the Hamiltonian, hence we may use the conditional expectation representation of linear G-BSDEs (compare with Remark 3.3 in (Hu et al. 2014a)):

$$
\begin{aligned}
\hat{p}^G(t) &= \frac{1}{\hat{X}(t)}\hat{\mathbb{E}}\left[(\hat{X}(T))^{\alpha-1}\hat{X}(T)|\mathcal{F}_t\right] \\
&= (\hat{X}(t))^{\alpha-1}\hat{\mathbb{E}}\left[\exp\left\{\alpha\int_t^T s(r)\hat{u}(r)dB(r)-\frac{\alpha^2}{2}\int_t^T s^2(r)\hat{u}^2(r)d\langle B\rangle(r)\right\}\cdot\right. \\
&\quad \left.\cdot\exp\left\{\int_t^T\left[\alpha m(r)\hat{u}(r)+\frac{\alpha^2-\alpha}{2}s^2(r)\hat{u}^2(r)\right]d\langle B\rangle(r)\right\}|\mathcal{F}_t\right].
\end{aligned}
$$

Springer Open

Applying the Girsanov theorem and the same reasoning as in (46), we easily get that

$$
\hat{p}^G(t) = \frac{1}{\hat{X}(t)} \hat{\mathbb{E}} \left[(\hat{X}(T))^{\alpha-1} \hat{X}(T) | \mathcal{F}_t \right]
$$

$$
= (\hat{X}(t))^{\alpha-1} \hat{\mathbb{E}}^{\hat{u}} \left[\exp\left\{ \int_t^T \left[\alpha m(r)\hat{u}(r) + \frac{\alpha^2 - \alpha}{2} s^2(r)\hat{u}^2(r) \right] d\langle B\rangle(r) \right\} | \mathcal{F}_t \right]
$$

$$
= (\hat{X}(t))^{\alpha-1} \hat{\mathbb{E}}^{\hat{u}} \left[\exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} d\langle B\rangle(r) \right\} | \mathcal{F}_t \right]
$$

$$
= (\hat{X}(t))^{\alpha-1} \hat{\mathbb{E}}^{\hat{u}} \left[\exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} d\langle B^{\hat{u}}\rangle(r) \right\} | \mathcal{F}_t \right]
$$

$$
= (\hat{X}(t))^{\alpha-1} \hat{\mathbb{E}} \left[\exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} d\langle B\rangle(r) \right\} | \mathcal{F}_t \right].
$$

Furthermore, we also know that the integrand is always positive by the assumption $\alpha \in]0, 1[$, hence we get by the representation of the conditional G-expectation (6) that for every $\mathbb{P} \in \mathcal{P}$ and by (26) that

$$
\hat{\mathbb{E}} \left[\exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} d\langle B\rangle(r) \right\} | \mathcal{F}_t \right]
$$

$$
= \operatorname*{ess\,sup}_{\mathbb{P}' \in \mathcal{P}(t,\mathbb{P})} {}^{\mathbb{P}}\mathbb{E}^{\mathbb{P}'} \left[\exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} d\langle B\rangle(r) \right\} | \mathcal{F}_t \right]
$$

$$
= \exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} dr \right\} \quad \mathbb{P} - a.s.
$$

Hence,

$$
\hat{p}^G(t) = (\hat{X}(t))^{\alpha-1} \exp\left\{ \int_t^T \frac{\alpha}{2(1-\alpha)} \frac{m^2(r)}{s^2(r)} dr \right\} =: (\hat{X}(t))^{\alpha-1} \cdot Z(t).
$$

By integration by parts for $\hat{X}^{-1}$ and Z, one can compute that

$$
d\hat{p}^G(t) = -\frac{m(t)}{s(t)} \hat{p}^G(t) dB(t) + \frac{\alpha m^2(t)}{2(1-\alpha)s^2(t)} \hat{p}^G(t)(d\langle B\rangle(t) - dt). \tag{47}
$$

By comparing Eq. 47 with the adjoint Eq. 12, we first obtain that

$$
\hat{q}^G(t) = -\frac{m(t)}{s(t)} \hat{p}^G(t)
$$

and hence that $\hat{u}$ is a maximizer of the function $u \mapsto H\left(t, \hat{X}(t), u, \hat{p}^G(t), \hat{q}^G(t)\right)$. Secondly, we get that the process $\hat{K}$ has the explicit form

$$
\hat{K}(t) = \int_0^t \frac{\alpha m^2(r)}{2(1-\alpha)s^2(r)} \hat{p}^G(r)(d\langle B\rangle(r) - dr)
$$

and consequently is a non-trivial process.

To summarize the example so far: we have shown that $\hat{u}$ is optimal in a weaker sense. We also showed that it satisfies the assumption for the necessary maximum

principle for strongly robust optimality and that all assumptions of the sufficient maximum principle are satisfied, with the exception of the vanishing of process $\hat{K}$. Now, we prove that $\hat{u}$ is not optimal in the stronger sense, hence the assumption on the process $\hat{K}$ is really crucial for our result and cannot be dropped.

Fix $\mathbb{P} \in \mathcal{P}_1$ and assume that $\hat{u}$ is optimal under $\mathbb{P}$. By Lemma 1, we know that $\hat{u}$ is a critical point of the Hamiltonian evaluated in $\hat{p}^{\mathbb{P}}$ and $\hat{q}^{\mathbb{P}}$. Hence, by the same analysis as in Theorem 5, we see that

$$d\hat{p}^{\mathbb{P}}(t) = -\frac{m(t)}{s(t)}\hat{p}^{\mathbb{P}}(t)dB(t),$$

therefore

$$\hat{p}^{\mathbb{P}}(T) = \hat{p}^{\mathbb{P}}(0) \exp\left\{-\int_0^T \frac{m(t)}{s(t)}dB(t) - \frac{1}{2}\int_0^T \frac{m^2(t)}{s^2(t)}d\langle B\rangle(t)\right\}. \quad (48)$$

However, we know by the dynamics of $\hat{X}$ and the terminal condition of $\mathbb{P}$-BSDE that

$$
\begin{aligned}
\hat{p}^{\mathbb{P}}(T) &= (\hat{X}(T))^{\alpha-1} \\
&= x^{\alpha-1} \exp\left\{(\alpha-1)\left[\int_0^T \hat{u}(t)s(t)dB(t)\right.\right. \\
&\quad \left.\left. + \int_0^T \left(\hat{u}(t)m(t) - \frac{1}{2}\hat{u}^2(t)s^2(t)\right) d\langle B\rangle(t)\right]\right\} \\
&= x^{\alpha-1} \exp\left\{-\int_0^T \frac{m(t)}{s(t)}dB(t) - \frac{1}{2}\int_0^T \frac{m^2(t)(1-2\alpha)}{s^2(t)(1-\alpha)}d\langle B\rangle(t)\right\}.
\end{aligned}
\quad (49)
$$

Dividing (48) by (49), we get that

$$1 = \frac{\hat{p}^{\mathbb{P}}(0)}{x^{\alpha-1}} \exp\left\{\int_0^T \frac{\alpha m^2(t)}{2s^2(t)(\alpha-1)}d\langle B\rangle(t)\right\}.$$

The equalities here are $\mathbb{P}$-a.s. so we get that the integral $\int_0^T \frac{\alpha m^2(t)}{2s^2(t)(\alpha-1)}d\langle B\rangle(t)$ must be equal $\mathbb{P}$-a.s. to a constant. However, the quadratic variation of the canonical process under $\mathbb{P}$ is generally a non-deterministic stochastic process, hence also the integral is a random variable, in general non-constant. This shows that $\hat{u}$ is optimal under $\mathbb{P}$ only for very specific probability measures such as the Wiener measure.

To conclude, $\hat{u}$ is not optimal for every probability measure $\mathbb{P} \in \mathcal{P}$ even though it is a maximizer of the Hamiltonian related to $\hat{u}$. This example shows that the new strong notion of optimality is rather restricted and we may expect it only in very special cases when the process $\hat{K}$ vanishes.

Funding

Availability of data and materials

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Authors' contributions

The four authors worked together on the manuscript and approved its final version.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

Denis, L., Hu, M., Peng, S.: Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths. Potential Anal. **34**, 139–161 (2011)

Hu, M., Ji, S.: Stochastic maximum principle for stochastic recursive optimal control problem under volatility ambiguity. SIAM J. Control Optim. **54**(2), 918–945 (2016)

Hu, M., Ji, S., Peng, S.: Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion. Stoch. Process. Appl. **124**, 1170–1195 (2014a)

Hu, M., Ji, S., Yang, S.: A stochastic recursive optimal control problem under the G-expectation framework. Appl. Math. Optim. **70**, 253–278 (2014b)

Hu, M., Ji, S., Peng, S., Song, Y.: Backward stochastic differential equations driven by G-Brownian motion. Stoch. Process. Appl. **124**, 759–784 (2014c)

Matoussi, A., Possamai, D., Zhou, C.: Robust utility maximization in non-dominated models with 2BSDEs. Math. Financ. (2013). https://doi.org/10.1111/mafi.12031

Peng, S.: G-expectation, G-Brownian motion and related stochastic calculus of Itô type. Stoch. Anal. Appl. **2**, 541–567 (2007)

Peng, S.: Nonlinear expectations and stochastic calculus under uncertainty. Preprint, arXiv1002.4546v1 (2010)

Peng, S., Song, Y., Zhang, J.: A complete representation theorem for G-martingales. Stochastics. **86**, 609–631 (2014)

Soner, M., Touzi, N., Zhang, J.: Martingale representation theorem for the G-expectation. Stoch. Anal. Appl. **121**, 265–287 (2011a)

Soner, M., Touzi, N., Zhang, J.: Quasi-sure stochastic analysis through aggregation. Electron. J. Probab. **16**, 1844–1879 (2011b)

Soner, M., Touzi, N., Zhang, J.: Wellposedness of second order backward SDEs. Probab. Theory Relat. Fields. **153**, 149–190 (2011c)

Song, Y.: Some properties on G-evaluation and its applications to G-martingale decomposition. Sci. China. **54**, 287–300 (2011)

Song, Y.: Uniqueness of the representation for G-martingales with finite variation. Electron. J. Probab. **17**, 1–15 (2012)