Backward Stochastic Differential Equations with constraints on the gains-process *

Jakša Cvitanić
Department of Statistics
Columbia University
New York, NY 10027
cj@stat.columbia.edu

Ioannis Karatzas
Departments of Mathematics
and Statistics
Columbia University
New York, NY 10027
ik@math.columbia.edu

H. Mete Soner †
Department of Mathematics
Bogazici University
Bebek 80815
Istanbul, Turkey
sonermet@boun.edu.tr

October 13, 1997

Abstract

We consider Backward Stochastic Differential Equations with convex constraints on the gains (or intensity-of-noise) process. Existence and uniqueness of a *minimal* solution are established in the case of a drift coefficient which is Lipschitz-continuous in the state- and gains-processes, and convex in the gains-process. It is also shown that the minimal solution can be characterized as the unique solution of a functional stochastic control-type equation. This representation is related to the penalization method for constructing solutions of stochastic differential equations, involves change of measure techniques, and employs notions and results from convex analysis, such as the *support function* of the convex set of constraints and its various properties.

Key words: Backward SDEs, convex constraints, stochastic control.
AMS 1991 subject classifications: Primary 60H10, 93E20; secondary 60G40.

*Supported in part by the U.S. Army Research Office, under Grants DAAH 04-95-1-0528 (Cvitanić and Karatzas) and DAAH-04-95-1-0226 (Soner).

† On leave from Carnegie Mellon University.

1 Introduction

The standard theory for Stochastic Differential Equations (SDE) of the type

$$dX(t) = -f(t, X(t))dt + \sigma'(t, X(t))dB(t), \quad 0 \le t \le T \tag{1.1}$$

with *initial condition* $X(0) = x \in \mathbb{R}$, driven by the d-dimensional Brownian motion $B(\cdot)$, was developed by Itô (1942, 1946, 1951). It asserts that the equation (1.1) has a pathwise-unique solution $X(\cdot)$, a measurable process on the given probability space $(\Omega, \mathcal{F}, P)$ that satisfies

$$E[\sup_{0 \le t \le T} |X(t)|^2] < \infty \tag{1.2}$$

and is *adapted to the filtration* $\mathbf{F}$ *generated by the driving Brownian motion* $B(\cdot)$, provided that the drift $f : [0, T] \times \mathbb{R} \to \mathbb{R}$ and dispersion $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}^d$ coefficients satisfy appropriate Lipschitz and growth conditions; see, for instance, Karatzas & Shreve (1991), section 5.2.

In a very interesting paper, Pardoux & Peng (1990) developed recently a similar theory for equations analogous to (1.1), but in which one specifies a *terminal* rather than initial condition. More precisely, with $f(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ as above and with ξ a square-integrable and $\mathcal{F}(T)-$measurabe random variable, they showed that there exists a unique *pair of* $\mathbf{F}-$*adapted processes* $(X(\cdot), Y(\cdot))$ that satisfy (1.2),

$$E\int_0^T \|Y(t)\|^2 dt < \infty, \tag{1.3}$$

as well as the *Backwards Stochastic Differential Equation (BSDE)*

$$X(t) = \xi + \int_t^T f(s, X(s))ds - \int_t^T [\sigma(s, X(s)) + Y(s)]' dB(s), \quad 0 \le t \le T. \tag{1.4}$$

In other words, one tries to "steer" the state-process $X : [0, T] \times \Omega \to \mathbb{R}$ to the specified terminal condition $X(T) = \xi$ at time $t = T$, while keeping it adapted to the filtration $\mathbf{F}$ generated by the driving Brownian motion $B(\cdot)$. The abilility to accomplish this depends crucially on the freedom to choose the "gains", or intensity-of-noise, process $Y : [0, T] \times \Omega \to \mathbb{R}^d$, again in a non-anticipative manner. Indeed, one could try to solve the SDE (1.1) using a time-reversal, that is, for the process $\tilde{X}(s) := X(T - s), 0 \le s \le T$ starting with the condition $\tilde{X}(0) = X(T) = \xi$; but the resulting state-process $X(\cdot)$ would then be adapted to the "reversed-time" filtration $\tilde{\mathcal{F}}(s) := \sigma(W(u) - W(s), s \le u \le T),\ 0 \le s \le T$, *not* to $\mathbf{F}$.

The freedom to choose the "gains" process $Y(\cdot)$ as an element of control, is the crucial difference between the theory for BSDEs and the more classical Itô theory for SDEs. Suppose, however, that the controller's ability to choose this gains-process $Y(\cdot)$ is limited, say by the requirement that $Y(\cdot)$ take values in a given nonempty, closed convex set K of $\mathbb{R}^d$. Then it is, generally, no longer possible to find a pair of $\mathbf{F}-$adapted processes $(X(\cdot), Y(\cdot))$ that satisfy this requirement, in addition (1.2)-(1.4). One needs to give the controller freedom to take more swift "corrective action", captured by an $\mathbf{F}-$adapted processes $C:[0,T]\times\Omega\to[0,\infty)$ with increasing, right-continuous paths and

$$E(C(T))^2 < \infty; \tag{1.5}$$

here $C(t)$ represents the cumulative effect of his corrective actions up to time $t\in[0,T]$. More precisely, one seeks a triple of $\mathbf{F}-$adapted processes $(X(\cdot), Y(\cdot), C(\cdot))$ as above that satisfies almost surely the analogue

$$X(t) = \xi + \int_t^T f(s, X(s))ds - \int_t^T [\sigma(s, X(s)) + Y(s)]'dB(s) + C(T) - C(t), \quad 0 \le t \le T \tag{1.6}$$

of the BSDE (1.4), the conditions (1.2), (1.3), (1.5), as well as the *constraint*

$$Y(t) \in K, \quad 0 \le t \le T, \tag{1.7}$$

and is the *minimal solution* of (1.6) with these properties (meaning that for any other such triple $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{K}(\cdot))$ that satisfies the system (1.2), (1.3), (1.5)-(1.7) we have $X(\cdot) \le \tilde{X}(\cdot)$, a.s.). The Constrained Backwards Stochastic Differential Equation (CBSDE) of (1.6), (1.7) is the focus of this paper. In order to simplify things and help focus attention on the constraint (1.7), we have taken $\sigma \equiv 0$ throughout. Using notions, tools and results from convex analysis, and ideas from our earlier papers Cvitanić & Karatzas (1992, 1993) that dealt with constrained optimization and hedging problems in the special context of mathematical finance, we discuss first the case of Constrained Backwards Stochastic Equations (CBSE), that is, with $\sigma \equiv 0$ and $f(\cdot,\cdot)$ replaced by an $\mathbf{F}-$adapted process $g(\cdot)$ in (1.6) (section 2). Next, we develop in section 3 the solvability and properties of the "penalized" version

$$X_n(t) = \xi + \int_t^T f(s, X(s))ds - \int_t^T Y_n(s)'dB(s) + C_n(T) - C_n(t), \quad 0 \le t \le T \tag{1.8}$$

of (1.6) with $\sigma \equiv 0$ and

$$C_n(t) := n\int_t^T \rho(Y_n(s))ds, \quad \rho(y) := dist(y, K),$$

again with the help of tools from convex analysis. We put then together the theory of section 2 and the properties of the penalization scheme (1.8), to study the CBSDE (1.6) in the case of general Lipschitz-continuous drift function $f(t,\omega,\cdot)$ via martingale and stochastic-control methods. A crucial element of our approach, developed in section 4, is the *functional stochastic-control-type equation*

$$X^*(t) = ess\sup_{\nu\in\mathcal{D}} E^\nu\left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u))]du \,\middle|\, \mathcal{F}(t)\right], \quad 0 \le t \le T, \tag{1.9}$$

which seems to be encountered and studied in this paper for the first time. Here $\delta(z) = \sup_{y\in K}(z'y)$ is the support function of the set K of (1.7), $\mathcal{D}$ is the class of bounded, $\mathbf{F}-$adapted processes $\nu(\cdot)$ with values in the effective domain $\tilde{K} := \{x \in \mathbb{R}^d \ / \ \delta(x) < \infty\}$ of $\delta(\cdot)$, and E^ν denotes expectation with respect to the auxiliary probability measure $P^\nu(A) := E[\exp\{\int_0^T \nu'(s)dB(s) - \frac{1}{2}\int_0^T \|\nu(s)\|^2 ds\} \cdot 1_A], A \in \mathcal{F}(T)$ for every "adjoint variable" process $\nu(\cdot)$ in $\mathcal{D}$. We show in section 4 that the equation (1.9) admits a unique solution $X^*(\cdot)$ with the property (1.2); this process is dominated by the state-process of *any* solution to the constrained BSDE of (1.6), (1.7) leading, as we demonstrate, to the *minimal solution* of this equation. In sections 5 and 6 we show how to extend those results to the case of a drift coefficient $f(t,x,y)$ which depends also on the current value $Y(t) = y$ of the gains process, but in a convex fashion, and to the case of a reflecting lower-barrier for the state-process $X(\cdot)$; each of these cases necessitates the introduction of an additional "adjoint variable" (a process $\mu(\cdot)$, or a stopping time τ, respectively). In subsequent work we expect to be able to extend the methodology of this paper, to cover the case of general dispersion $\sigma(t,x)$ and drift $f(t,x,y)$ coefficients.

Related existence results are obtained by Buckdahn & Hu (1997) for the special, one-dimensional case ($d = 1$), but in a more general context of BSDEs with a lower-barrier process, driven by both a Brownian motion and a Poisson random measure. These authors do not use a stochastic control approach, or representations of the type (1.9).

Backwards Stochastic Differential Equations were apparently first studied in the context of the stochastic version of Pontryagin's "maximum principle" for the *optimal control of diffusions* (see Saksonov (1989), Arkin & Saksonov (1979), Peng (1990, 1993), Elliott (1990); as well as Haussmann (1986), Bensoussan (1981), Bismut (1978), and the references therein, for earlier work). They also arose in the context of "recursive utility" for mathematical economics, in the work of Duffie & Epstein (1992). Since their formal and systematic study by Pardoux & Peng (1990) in a general framework, they have found an enormous range of applications in such diverse fields as *partial differential equations* (cf. Peng (1991), Barles, Buckdahn & Pardoux (1997), Darling & Pardoux (1997)), *variational inequalities* and *obstacle problems* (cf. Pardoux & Tang (1996), El Karoui, Kapoudjian, Pardoux, Peng & Quenez (abbreviated [EKPPQ]) (1997), Ma & Cvitanić (1997)), *stochastic PDEs* (Pardoux & Peng (1994)), *stochastic control* (cf. Peng (1990, 1993), Hamadène & Lepeltier (1995a)), *stochastic games* (cf. Hamadène & Lepeltier (1995b), Cvitanić & Karatzas (1996)), and *mathematical finance* (cf. Cvitanić & Karatzas (1993), ElKaroui, Peng & Quenez (1997), Buckdahn & Hu (1996, 1997)).

2 Backward Stochastic Equations with constraints

On a given, complete probability space $(\Omega, \mathcal{F}, P)$, let $B(\cdot) = (B_1(\cdot), \ldots, B_d(\cdot))'$ be a standard d-dimensional Brownian motion over the finite interval $[0, T]$, and denote by $\mathbf{F} = \{\mathcal{F}(t)\}_{0 \le t \le T}$ the augmentation of the natural filtration $\mathbf{F}^B$ generated by $B(\cdot)$, namely $\mathcal{F}^B(t) = \sigma(B(s), 0 \le s \le t)$, $0 \le t \le T$. We shall need the following notation : For any given $n \in \mathbb{N}$, let us introduce the spaces

$\mathbf{L}^2_n$ of $\mathcal{F}(T)$-measurable random variables $\xi : \Omega \to \mathbb{R}^n$ with $E(\|\xi\|^2) < \infty$;

$\mathbf{H}^2_n$ of $\mathbf{F}$-predictable processes $\varphi : [0, T] \times \Omega \to \mathbb{R}^n$ with $\int_0^T E\|\varphi(t)\|^2 dt < \infty$;

$\mathbf{S}^k_n$ of $\mathbf{F}$-progressively measurable processes $\varphi : [0, T] \times \Omega \to \mathbb{R}^n$ with the property $E(\sup_{0 \le t \le T} \|\varphi(t)\|^k) < \infty$, $k \in \mathbb{N}$;

$\mathbf{A}^2_i$ of RCLL, $\mathbf{F}$-adapted, predictable increasing processes $A : [0, T] \times \Omega \to [0, \infty)$ with $A(0) = 0$, $E(A^2(T)) < \infty$.

Finally, we shall denote by $\mathcal{P}$ the $\sigma-$algebra of predictable sets in $[0, T] \times \Omega$.

Suppose now that we are given a random variable $\xi : \Omega \to \mathbb{R}$ in the space $\mathbf{L}^2_1$, as well as a process $g : [0, T] \times \Omega \to \mathbb{R}$ in the space $\mathbf{H}^2_1$. Suppose also that we are given a closed, convex set $K \subset \mathbb{R}^d$ which contains the origin, and whose *support function*

$$\delta(z) := \sup_{y \in K} (y'z), \ \ z \in \mathbb{R}^d \tag{2.1}$$

is continuous on its effective domain

$$\tilde{K} := \{x \in \mathbb{R}^d \ / \ \delta(x) < \infty\}, \tag{2.2}$$

the "barrier cone" of the set K. Here, $y'x$ denotes the inner product of the vectors y and x. It is shown in Rockafellar (1970) that $\delta(\cdot)$ is indeed continuous on $\tilde{K}$, if this latter set is locally simplicial.

We shall denote by $\mathcal{H}$ the class of $\mathbf{F}$-progressively measurable processes $\nu : [0, T] \times \Omega \to \tilde{K}$ with $E\int_0^T \|\nu(t)\|^2 dt < \infty$; for every $\nu(\cdot) \in \mathcal{H}$, the exponential process

$$Z_\nu(t) := \exp\{\int_0^t \nu'(s) dB(s) - \frac{1}{2}\int_0^t \|\nu(s)\|^2 ds\}, \quad 0 \le t \le T \tag{2.3}$$

is a local martingale and a supermartingale; it is a martingale if and only if $EZ_\nu(T) = 1$, in which case

$$P^\nu(A) := E[Z_\nu(T)\mathbf{1}_A], \ A \in \mathcal{F}(T) \tag{2.4}$$

is a probability measure. In particular, this is the case for every process $\nu(\cdot)$ in the space

$$\mathcal{D} = \cup_{n=1}^\infty \mathcal{D}_n \ , \ \mathcal{D}_n := \{\nu \in \mathcal{H} \ / \ \|\nu(t, \omega)\| \le n, \ \text{for a.e. } (t, \omega) \in [0, T] \times \Omega\} \tag{2.5}$$

of *bounded* processes in $\mathcal{H}$. (For the unconstrained case $K = \mathbb{R}^d$ we have trivially $\tilde{K} = \{0\}$; then $\mathcal{D}$ contains only the evanescent processes $\nu(\cdot) \equiv 0$, a.e. on $[0,T] \times \Omega$, and $P^0 = P$.)

We first consider the problem of a Backward Stochastic Equation (BSE) with constraints on the "gains", or "intensity-of-noise", process; the solution for this problem was provided, in a slightly different context, by Cvitanić & Karatzas (1993), hereafter abbreviated [CK'93].

Problem 2.1: Find a triple of $\mathbf{F}$-progressively measurable processes $(X(\cdot), Y(\cdot), C(\cdot))$ with $X(\cdot) \in \mathbf{S}_1^2$, $Y(\cdot) \in \mathbf{H}_d^2$, $C(\cdot) \in \mathbf{A}_i^2$, such that the *Backwards Stochastic Equation (BSE)*

$$X(t) = \xi + \int_t^T g(u)du - \int_t^T Y'(u)dB(u) + C(T) - C(t), \quad 0 \le t \le T \tag{2.6}$$

and the *constraint*

$$Y(t) \in K, \quad 0 \le t \le T \tag{2.7}$$

hold almost surely, and such that for any other triple $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot)) \in \mathbf{S}_1^2 \times \mathbf{H}_d^2 \times \mathbf{A}_i^2$ that satisfies (2.6) and (2.7) we have

$$X(t) \le \tilde{X}(t), \quad 0 \le t \le T$$

almost surely. □

In the interest of readability and completeness, we recall here the main results from [CK'93] related to this problem, modified and adapted to our framework. First, we notice that for any solution to the BSE of (2.6), we have

$$\begin{aligned} X(t) &= E[\xi + \textstyle\int_t^T g(u)du + C(T) - C(t) \mid \mathcal{F}(t)] - E[\textstyle\int_t^T Y'(u)dB(u) \mid \mathcal{F}(t)] \\ &\ge E[\xi + \textstyle\int_t^T g(u)du \mid \mathcal{F}(t)] =: X_0(t), \quad 0 \le t \le T. \end{aligned} \tag{2.8}$$

This process $X_0(\cdot)$ is the solution of the *unconstrained* version

$$X_0(t) = \xi + \int_t^T g(u)du - \int_t^T Y_0'(u)dB(u), \quad 0 \le t \le T$$

of (2.6), with $C_0(\cdot) \equiv 0$ and with a suitable process $Y_0(\cdot) \in \mathbf{H}_d^2$ that takes values in $\mathbb{R}^d$ (unconstrained); the existence and uniqueness of such a process $Y_0(\cdot)$ follows from the integral representation property for square-integrable martingales of the Brownian filtration (cf. Karatzas & Shreve (1991), pp. 182-184). Furthermore, let us notice that the process $X(\cdot) + \int_0^\cdot g(s)ds$ dominates the square-integrable, P-martingale

$$\begin{aligned} X_0(t) + \int_0^t g(u)du &= E\Big[\xi + \int_0^T g(u)du \,\Big|\, \mathcal{F}(t)\Big] \\ &= E[\xi + \int_0^T g(u)du] + \int_0^t Y_0'(u)dB(u), \quad 0 \le t \le T. \end{aligned} \tag{2.9}$$

Moreover, for every $\nu(\cdot) \in \mathcal{D}$ we know from Girsanov's theorem (e.g. Karatzas & Shreve (1991), section 3.5) that the process

$$B_\nu(t) := B(t) - \int_0^t \nu(s)ds, \quad 0 \le t \le T \tag{2.10}$$

is Brownian motion under the probability measure P^ν of (2.4).

Proposition 2.1 *For any triple $(X(\cdot), Y(\cdot), C(\cdot))$ that solves the constrained BSE of Problem 2.1, the process*

$$X(t) + \int_0^t [g(u) - \delta(\nu(u))]du, \quad 0 \le t \le T \tag{2.11}$$

is a $P^\nu-$supermartingale with RCLL paths.

Proof: It is easily seen from (2.6) and (2.10) that

$$X(t) + \int_0^t [g(u) - \delta(\nu(u)]du + [C(t) + \int_0^t (\delta(\nu(u)) - \nu'(u)Y(u))du] = X(0) + \int_0^t Y'(u)dB_\nu(u), \tag{2.12}$$

for all $0 \le t \le T$. The stochastic integral on the right-hand side is a $P^\nu-$martingale, since we have

$$E^\nu \left(\int_0^T \|Y(u)\|^2 du \right)^{\frac{1}{2}} \le \left(EZ_\nu^2(T) \cdot E \int_0^T \|Y(u)\|^2 du \right)^{\frac{1}{2}} < \infty \ ;$$

we are using here the boundedness of the process $\nu(\cdot)$, the assumption $Y(\cdot) \in \mathbf{H}_d^2$, and the Cauchy-Schwarz inequality. Here and in the sequel, E^ν denotes the expectation operator under the probability measure P^ν of (2.4). The statement of the proposition follows then from (2.12), after noting that $C(\cdot) + \int_0^\cdot [\delta(\nu(u)) - \nu'(u)Y(u)]du$ is an increasing process. □

Proposition 2.2 *For any triple $(X(\cdot), Y(\cdot), C(\cdot))$ that solves the constrained BSE of Problem 2.1, we have*

$$X(t) \ge \hat{X}(t) := ess \sup_{\nu \in \mathcal{D}} E^\nu[\xi + \int_t^T [g(u) - \delta(\nu(u)]du \,\Big|\, \mathcal{F}(t)], \quad a.s. \tag{2.13}$$

for every $t \in [0, T]$.

Proof: From Proposition 2.1 we have

$$X(t) \ge E^\nu \left[X(T) + \int_t^T [g(u) - \delta(\nu(u))]du \,\Big|\, \mathcal{F}(t) \right], \ a.s.$$

for every $\nu(\cdot) \in \mathcal{D}$, and we are done, because $X(T) = \xi$. □

It is clear now that, in order to find the minimal solution to the constrained BSE of Problem 2.1, *it suffices to show that there exist processes* $\hat{Y}(\cdot) \in \mathbf{H}_d^2$ and $\hat{C}(\cdot) \in \mathbf{A}_i^2$ *such* *that* $(\hat{X}(\cdot), \hat{Y}(\cdot), \hat{C}(\cdot))$ *is a solution.* Then this triple has to be the minimal solution, and the processes $\nu(\cdot) \in \mathcal{D}$ are seen (by comparing (2.13) with (2.9)) to play the role of "adjoint variables" that enforce the constraint $\hat{Y}(\cdot) \in K$. We shall do this by imposing the following, very mild assumption.

Assumption 2.1 *There exists at least one solution* $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot)) \in \mathbf{S}_1^2 \times \mathbf{H}_d^2 \times \mathbf{A}_i^2$ *to the constrained BSE of Problem 2.1; or equivalently, we have*

$$\xi + \int_0^T g(u)du \leq \eta, \quad a.s. \tag{2.14}$$

for some random variable $\eta \in \mathbf{L}_1^2(\Omega)$ *that can be represented in the form* $\eta = c + \int_0^T Y_\eta'(u)dB(u)$ *for suitable* $c \in \mathbb{R}$ *and* $Y_\eta(\cdot) \in \mathbf{H}_d^2$ *(thus* $c = E\eta$*) such that* $P[Y_\eta(t) \in K, \forall 0 \leq t \leq T] = 1$.

Let us show that the two assumptions are indeed equivalent: If $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot))$ is a solution of Problem 2.1, then we can take $\eta := \tilde{X}(0) + \int_0^T \tilde{Y}'(u)dB(u)$ and obtain the inequality (2.14) from (2.6) with $t = 0$. Conversely, given η as in the inequality (2.14), we can define $\tilde{X}(t) := E\eta + \int_0^t Y_\eta'(u)dB(u) - \int_0^t g(u)du$ and $\tilde{C}(t) := 0$ for $0 \leq t < T$, as well as $\tilde{X}(T) := \xi$ and $\tilde{C}(T) := \tilde{X}(T-) - \xi \geq 0$; it is easily seen that $(\tilde{X}(\cdot), Y_\eta(\cdot), \tilde{C}(\cdot))$ is then a solution of Problem 2.1.

Assumption 2.1 is satisfied, in its form (2.14), for example if both ξ and $g(\cdot)$ are bounded. Many more examples can be found in [CK'93] and in Broadie, Cvitanić & Soner (1996).

We state now a result which is analogous to Proposition 6.3 of [CK'93], and has a similar proof (sketched in the Appendix).

Proposition 2.3 *The process* $\hat{X}(\cdot)$ *of* (2.13) *can be considered in its RCLL modification; then, the process* $\hat{X}(t) + \int_0^t [g(u) - \delta(\nu(u))]du$, $0 \leq t \leq T$ *is a* P^ν*−supermartingale with RCLL paths for every* $\nu(\cdot) \in \mathcal{D}$*, and we have the stronger version*

$$P[X(t) \geq \hat{X}(t), \ \forall \ 0 \leq t \leq T] = 1 \tag{2.15}$$

of the result in Proposition 2.2.

Next, we have the following result.

Proposition 2.4 *The process*

$$\hat{Q}(t) := \hat{X}(t) + \int_0^t g(u)du, \ 0 \leq t \leq T \tag{2.16}$$

belongs to the space $\mathbf{S}_1^2$*, i.e.,* $E[\sup_{0 \leq t \leq T} (\hat{Q}(t))^2] < \infty$.

Proof: From (2.13) we have

$$\hat{Q}(t) \geq E\left[\xi + \int_0^T g(u)du \,\Big|\, \mathcal{F}(t)\right], \quad 0 \leq t \leq T.$$

The process on the right-hand side is a martingale in the space $\mathbf{S}_1^2$, by Doob's maximal inequality. On the other hand, (2.13) and Assumption 2.1 imply

$$\hat{Q}(t) \leq \tilde{X}(t) + \int_0^t g(u)du, \quad 0 \leq t \leq T.$$

The process on the right-hand side is also in $\mathbf{S}_1^2$, and we are done. □

Corollary 2.1 *For every given process $\nu(\cdot) \in \mathcal{D}$, the $P^\nu-$supermartingale*

$$\hat{X}(t) + \int_0^t [g(u) - \delta(\nu(u))]du = \hat{Q}(t) - \int_0^t \delta(\nu(u))du =: \hat{Q}_\nu(t), \quad 0 \leq t \leq T$$

is of class $\mathcal{D}([0,T])$ under P^ν; in other words, the family $\{\hat{Q}_\nu(\tau)\}_{\tau \in \mathcal{S}_{0,T}}$ is $P^\nu-$uniformly integrable, where $\mathcal{S}_{0,T}$ is the set of all $\mathbf{F}-$stopping times $\tau : \Omega \to [0,T]$.

Proof: Since the support function $\delta(\cdot)$ is continuous on its effective domain $\tilde{K}$, and the process $\nu(\cdot)$ is bounded, it suffices to show $E^\nu[\sup_{0 \leq t \leq T} |\hat{Q}(t)|] < \infty$. But this follows from Proposition 2.4, the Cauchy-Schwarz inequality, and the boundedness of the process $\nu(\cdot)$.
□

From Corollary 2.1, we have the *Doob-Meyer decomposition*

$$\hat{X}(t) + \int_0^t (g(u) - \delta(\nu(u))du = \hat{Q}(t) - \int_0^t \delta(\nu(u))du = \hat{X}(0) + M^{(\nu)}(t) - A^{(\nu)}(t), \quad 0 \leq t \leq T. \tag{2.17}$$

Here $A^{(\nu)}(\cdot)$ is an $\mathbf{F}-$predictable process with increasing, right-continuous paths and $A^{(\nu)}(0) = 0$, $E^\nu A^{(\nu)}(T) < \infty$. On the other hand, $M^{(\nu)}(\cdot)$ is a uniformly integrable P^ν-martingale of the Brownian filtration $\mathbf{F}$, and as such can be represented in the form

$$M^{(\nu)}(t) = \int_0^t \left(Y^{(\nu)}(u)\right)' dB_\nu(u), \quad 0 \leq t \leq T \tag{2.18}$$

for some process $Y^{(\nu)} : [0,T] \times \Omega \to \mathbb{R}^d$ which is $\mathbf{F}-$progressively measurable and satisfies $\int_0^T \|Y^{(\nu)}(t)\|^2 dt < \infty$, a.s.; cf. Karatzas & Shreve (1991), p. 375.

The proof of the following proposition proceeds along lines similar to those in the proof of Theorem 6.4 in [CK'93]; we sketch its main arguments in the Appendix.

Proposition 2.5 *The process*

$$\hat{Y}(\cdot) := Y^{(0)}(\cdot) \equiv Y^{(\nu)}(\cdot) \tag{2.19}$$

does not depend on the process $\nu(\cdot) \in \mathcal{D}$, *and neither does the predictable increasing, right-continuous process*

$$\hat{C}(\cdot) := A^{(0)}(\cdot) \equiv A^{(\nu)}(\cdot) - \int_0^{\cdot} [\delta(\nu(u)) - \nu'(u)\hat{Y}(u)]du. \tag{2.20}$$

Furthermore, we have

$$\hat{X}(t) = \xi + \int_t^T g(u)du - \int_t^T \hat{Y}'(u)dB(u) + \hat{C}(T) - \hat{C}(t), \quad 0 \le t \le T \tag{2.21}$$

and

$$\hat{Y}(t) \in K, \quad 0 \le t \le T \tag{2.22}$$

almost surely.

Finally, we have the following identification of the minimal solution.

Theorem 2.1 *Under Assumption 2.1, the triple* $(\hat{X}(\cdot), \hat{Y}(\cdot), \hat{C}(\cdot))$, *as defined in* (2.13), (2.19) *and* (2.20) *provides the minimal solution of the constrained BSE of Problem 2.1.*

Proof: It remains to prove

$$E[\sup_{0 \le t \le T} (\hat{X}(t))^2] < \infty, \tag{2.23}$$

$$E[\hat{C}(T)]^2 < \infty, \tag{2.24}$$

and

$$E\int_0^T \|\hat{Y}(t)\|^2 dt < \infty. \tag{2.25}$$

The inequality (2.23) follows from Proposition 2.4. The inequality (2.25) will follow from (2.24), because (2.17) with $\nu(\cdot) \equiv 0$ implies then that $M^{(0)}(\cdot)$ is a square-integrable martingale. Thus, it remains to show (2.24).

Let $Q_* := \sup_{0 \le t \le T} |\hat{Q}(t)|$, $q(t) := E[Q_* | \mathcal{F}(t)]$. Moreover, for every $k \in \mathbb{N}$, let $\rho_k := \inf\{t \in [0,T)/\hat{C}(t) \ge k\} \wedge T$. These are $\mathbf{F}$−stopping times, and we have $\rho_k \uparrow T$ as $k \to \infty$, a.s. Clearly,

$$\begin{aligned}
E[\hat{C}(\rho_k)]^2 &= 2E\int_0^{\rho_k} [\hat{C}(\rho_k) - \hat{C}(t)]d\hat{C}(t) \\
&= 2E\int_0^{\rho_k} E[\hat{C}(\rho_k) - \hat{C}(t) \mid \mathcal{F}(t)]d\hat{C}(t) \\
&= 2E\int_0^{\rho_k} E[\hat{Q}(t) - \hat{Q}(\rho_k) + M^{(0)}(\rho_k) - M^{(0)}(t) \mid \mathcal{F}(t)]d\hat{C}(t) \\
&= 2E\int_0^{\rho_k} E[\hat{Q}(t) - \hat{Q}(\rho_k) \mid \mathcal{F}(t)]d\hat{C}(t) \\
&\le 4E\int_0^{\rho_k} q(t)d\hat{C}(t) \le 4E\left[\sup_{0 \le t \le T} (q(t)) \cdot \hat{C}(\rho_k)\right] \\
&\le 4\left(E[\sup_{0 \le t \le T} q^2(t)] \cdot E[\hat{C}(\rho_k)]^2\right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore, we have

$$E[\hat{C}(\rho_k)]^2 \leq 16 \cdot E[\sup_{0\leq t\leq T} q^2(t)],$$

for all $k \in \mathbb{N}$. Furthermore, by Doob's Maximal Inequality and Proposition 2.4,

$$E[\sup_{0\leq t\leq T} q^2(t)] \leq 4Eq^2(T) = 4E[Q_*^2] = 4E[\sup_{0\leq t\leq T} (\hat{Q}(t))^2] < \infty.$$

Thus, letting $k \uparrow \infty$, we obtain

$$E[\hat{C}(T-)]^2 \leq 64 \cdot E[\sup_{0\leq t\leq T} (\hat{Q}(t))^2] < \infty.$$

On the other hand, since $\int_0^{\cdot} \hat{Y}'(s)dB(s)$ is continuous, (2.21) implies

$$\hat{C}(T) - \hat{C}(T-) = \hat{Q}(T-) - \hat{Q}(T) \ \in \ \mathbf{L}_1^2(\Omega)$$

thus $\hat{C}(T) \in \mathbf{L}_1^2(\Omega)$ as well, and we are done. □

3 Penalization and BSDEs with constraints

Suppose now that the process $g(\cdot) \in \mathbf{H}_1^2$ is replaced by the random field $f : [0,T]\times\Omega\times\mathbb{R} \to \mathbb{R}$, a given $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})-$measurable mapping that satisfies

$$E\int_0^T f^2(t,\omega,0)dt < \infty, \tag{3.1}$$

as well as

$$|f(t,\omega,x) - f(t,\omega,x')| \leq \kappa|x - x'| \tag{3.2}$$

for all $(t,\omega) \in [0,T] \times \Omega$ and $(x,x') \in \mathbb{R}^2$, for some $0 < \kappa < \infty$. Thus, instead of the constrained BSE of Problem 2.1, our focus now is the following Constrained Backwards Stochastic Differential Equation (CBSDE) problem.

Problem 3.1: *Find a triple of* $\mathbf{F}$*-progressively measurable processes* $(X(\cdot), Y(\cdot), C(\cdot))$ *with* $X(\cdot) \in \mathbf{S}_1^2$, $Y(\cdot) \in \mathbf{H}_d^2$, $C(\cdot) \in \mathbf{A}_i^2$, *such that the Backwards Stochastic Differential Equation (BSDE)*

$$X(t) = \xi + \int_t^T f(u, X(u))du - \int_t^T Y'(u)dB(u) + C(T) - C(t), \ 0 \leq t \leq T \tag{3.3}$$

and the constraint

$$Y(t) \in K, \ 0 \leq t \leq T \tag{3.4}$$

hold almost surely, and such that for any other solution $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot)) \in \mathbf{S}_1^2 \times \mathbf{H}_d^2 \times \mathbf{A}_i^2$ *satisfying* (3.3) *and* (3.4), *we have*

$$X(t) \leq \tilde{X}(t), \quad 0 \leq t \leq T$$

almost surely. $\square$

In order to solve Problem 3.1, we introduce the *penalized* BSDE

$$X_n(t) = \xi + \int_t^T [f(u, X_n(u)) + n\rho(Y_n(u))]du - \int_t^T Y_n'(u)dB(u),\ 0 \le t \le T \tag{3.5}$$

for every $n \in \mathbb{N}$, where $\rho(y)$ denotes the distance of the vector $y \in \mathbb{R}^d$ to the set K. Since the function $y \mapsto \rho(y)$ satisfies the Lipschitz condition

$$|\rho(y) - \rho(z)| \le |y - z|, \quad \forall (y, z) \in (\mathbb{R}^d)^2,$$

the equation (3.5) has a unique solution $(X_n(\cdot), Y_n(\cdot)) \in \mathbf{S}_1^2 \times \mathbf{H}_d^2$, by the standard theory of Pardoux & Peng (1990). We have the following characterization of this solution.

Proposition 3.1 *The solution $X_n(\cdot)$ of the penalized BSDE (3.5) satisfies the following stochastic equation*

$$X_n(t) = ess \sup_{\nu \in \mathcal{D}_n} E^\nu \left[\xi + \int_t^T [f(u, X_n(u)) - \delta(\nu(u))du \,\middle|\, \mathcal{F}(t) \right],\ 0 \le t \le T \tag{3.6}$$

almost surely.

In order to prove this result, we need a property of the support function $\delta(\cdot)$ in (2.1).

Lemma 3.1

$$\sup_{y \in \mathbb{R}^d} [\nu' y - n\rho(y)] = \begin{Bmatrix} \delta(\nu) & , & \nu \in \tilde{K} \cap B_n \\ \infty\ , & & \nu \notin \tilde{K} \cap B_n \end{Bmatrix}, \tag{3.7}$$

where $B_n := \{\nu \in \mathbb{R}^d;\ \|\nu\| \le n\}$.

Proof: For every $\nu \in \tilde{K}$, we have

$$\delta(\nu) = \sup_{y \in K} (\nu' y) = \sup_{y \in K} (\nu' y - n\rho(y)) \le \sup_{y \in \mathbb{R}^d} (\nu' y - n\rho(y)).$$

If, moreover, $\|\nu\| \le n$, and we denote by y_K the projection of y on K (i.e., $\rho(y) = \|y - y_K\|$), we get

$$\begin{aligned} \nu' y - n\rho(y) &= \nu' y_K + \nu'(y - y_K) - n\|y - y_K\| \\ &\le \delta(\nu) + \|y - y_K\|(\|\nu\| - n) \le \delta(\nu) \end{aligned}$$

for all $y \notin K$. For $y \in K$, we have clearly $\nu' y - n\rho(y) \le \delta(\nu)$ again, and thus

$$\delta(\nu) = \sup_{y \in \mathbb{R}^d} [\nu' y - n\rho(y)], \quad \text{for } \nu \in \tilde{K} \cap B_n.$$

Next, for any $\nu \in \mathbb{R}^d$ and $k \in \mathbb{N}$ with $\|\nu\| > n + \varepsilon$ for some $\varepsilon > 0$, there exists $y \in \mathbb{R}^d$, such that $\nu' \frac{y}{\|y\|} \geq n + \varepsilon$ and $\|y\| \geq k$. Thus,

$$\begin{aligned} \nu' y - n\rho(y) &= \|y\| \left[\nu' \frac{y}{\|y\|} - n \frac{\rho(y)}{\|y\|} \right] \\ &\geq \|y\| \left[\varepsilon + n \left(1 - \frac{\rho(y)}{\|y\|} \right) \right] \geq \varepsilon \|y\| \geq \varepsilon k, \end{aligned}$$

and letting $k \uparrow \infty$ we obtain $\sup_{y \in \mathbb{R}^d} [\nu' y - n\rho(y)] = \infty$, for all $\nu \notin B_n$.

Finally, for $\nu \notin \tilde{K}$, we have $\sup_{y \in \mathbb{R}^d} [\nu' y - n\rho(y)] \geq \sup_{y \in K} (\nu' y) = \delta(\nu) = \infty$. □

Proof of Proposition 3.1: Let $\nu(\cdot) \in \mathcal{D}_n$ and $t \in [0, T]$. From the BSDE (3.5) and Lemma 3.1, we have

$$\begin{aligned} X_n(t) + \int_t^T \delta(\nu(s)) ds &= \xi + \int_t^T f(s, X_n(s)) ds - \int_t^T Y_n'(s) dB_\nu(s) \\ &\quad + \int_t^T [n\rho(Y_n(s)) - Y_n'(s)\nu(s) + \delta(\nu(s))] ds \\ &\geq \xi + \int_t^T f(s, X_n(s)) ds - \int_t^T Y_n'(s) dB_\nu(s). \end{aligned} \tag{3.8}$$

By analogy with the proof of Proposition 2.1, the stochastic integrand in the last expression is a $P^\nu-$martingale. Hence, after taking conditional expectations, we obtain

$$X_n(t) \geq E^\nu \left[\xi + \int_t^T [f(s, X_n(s)) - \delta(\nu(s))] ds \;\middle|\; \mathcal{F}(t) \right]$$

almost surely. On the other hand, because the function $n\rho(\cdot)$ is Lipschitz-continuous and convex, we conclude as in p.36 of ElKaroui, Peng and Quenez (1997) (hereafter abbreviated [EPQ]), that there exists a process $\hat{\nu}_n(\cdot) \in \mathcal{D}_n$ such that $n\rho(Y_n) + Y_n' \hat{\nu}_n + \delta(\hat{\nu}_n) \equiv 0$, a.e. on $[0, T] \times \Omega$. Setting $\nu(\cdot) = \hat{\nu}_n(\cdot)$ in (3.8) we get equality there, and therefore also

$$X_n(t) = E^{\hat{\nu}_n} \left[\xi + \int_t^T [f(s, X_n(s)) - \delta(\hat{\nu}_n(s))] ds \;\middle|\; \mathcal{F}(t) \right],$$

almost surely. Thus we obtain the a.s. equality of (3.6), first for fixed $t \in [0, T]$, and then for all $0 \leq t \leq T$ simultaneously, from the continuity of its left-hand-side $X_n(\cdot)$ and the right-continuity of its right-hand-side (recall (3.5) and Proposition 2.3, respectively). □

We now embark on the problem of finding and characterizing the limit of the sequence $\{X_n(\cdot)\}_{n \in \mathbb{N}}$. The standard comparison theorem for BSDEs (see [EPQ], p. 23) implies that

$$X_n(t) \leq X_{n+1}(t), \quad 0 \leq t \leq T \tag{3.9}$$

holds almost surely for all $n \in \mathbb{N}$, since $n\rho(\cdot) \leq (n+1)\rho(\cdot)$. We also impose the following analogue of Assumption 2.1:

Assumption 3.1 *There exists at least one solution* $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot))$ *to the constrained BSDE of Problem 3.1.*

Lemma 3.2 *Let Assumption 3.1 hold and* $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot))$ *be any solution to the constrained BSDE of Problem 3.1. Then, we have*

$$X_n(t) \le \tilde{X}(t), \quad 0 \le t \le T$$

almost surely, for every $n \in \mathbb{N}$.

Proof: Choose $\hat{\nu}_n(\cdot)$ as in the proof of Proposition 3.1 so that, by (3.8), the process $X_n(\cdot)$ satisfies the BSDE

$$X_n(t) = \xi + \int_t^T [f(s, X_n(s)) - \delta(\hat{\nu}_n(s))] ds - \int_t^T Y_n'(s) dB_{\hat{\nu}_n}(s), \ 0 \le t \le T.$$

We also observe from (3.3), (2.10) that $\tilde{X}(\cdot)$ satisfies the BSDE

$$\tilde{X}(t) = \xi + \int_t^T [f(s, \tilde{X}(s)) - \tilde{Y}'(s)\hat{\nu}_n(s)] ds + \tilde{C}(T) - \tilde{C}(t) - \int_t^T \tilde{Y}'(s) dB_{\hat{\nu}_n}(s), \ 0 \le t \le T.$$

However, $0 \le \tilde{C}(T) - \tilde{C}(\cdot)$ and $-\delta(\hat{\nu}_n(\cdot)) \le -\tilde{Y}'(\cdot)\hat{\nu}_n(\cdot)$, so that the comparison theorem for BSDEs ([EPQ], p. 23) applies again, to give $X_n(\cdot) \le \tilde{X}(\cdot)$. (Note that, even though these BSDEs are driven by $B_{\hat{\nu}_n}(\cdot)$ rather than by $B(\cdot)$, the comparison theorem cited earlier is still valid because the stochastic integrals $\int_0^t Y_n'(s) dB_{\hat{\nu}_n}(s)$, $\int_0^t \tilde{Y}'(s) dB_{\hat{\nu}_n}(s)$, $0 \le t \le T$ are $P^{\hat{\nu}_n}-$martingales.)

□

We conclude from (3.9) and Lemma 3.2 that the limit

$$X^*(t) := \lim_{n \to \infty} X_n(t), \ 0 \le t \le T \tag{3.10}$$

exists almost surely. In the next section we prove that the limit-process $X^*(\cdot)$ leads to the minimal solution of the constrained BSDE of Problem 3.1.

4 Constrained BSDE and a stochastic equation

We shall impose throughout this section the Assumption 3.1, and establish with its help the following main result.

Theorem 4.1 *The process* $X^*(\cdot)$ *of* (3.10) *is the unique solution, in the space* $\mathbf{S}_1^2$, *of the stochastic equation*

$$X^*(t) = ess \sup_{\nu \in \mathcal{D}} E^{\nu} \left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u)) du \,\middle|\, \mathcal{F}(t) \right], \quad 0 \le t \le T. \tag{4.1}$$

Corollary 4.1 *(Existence and Uniqueness for Problem 3.1): There exist processes* $Y^*(\cdot) \in \mathbf{H}_d^2$ *and* $C^*(\cdot) \in \mathbf{A}_i^2$ *such that the triple* $(X^*(\cdot), Y^*(\cdot), C^*(\cdot))$ *is the minimal solution to the constrained BSDE of Problem 3.1.*

Proof of Corollary 4.1: Since $X_n(\cdot) \le X^*(\cdot) \le \tilde{X}(\cdot)$, we have $X^*(\cdot) \in \mathbf{S}_1^2$. From this, and from Theorem 4.1, it is easily checked that the analogue of Proposition 2.4 holds, with $\hat{X}(\cdot)$ replaced by $X^*(\cdot)$ and $g(\cdot)$ replaced by $f(\cdot, X^*(\cdot))$. Then, using the theory developed in section 2, one constructs processes $Y^*(\cdot) \in \mathbf{H}_d^2$ and $C^*(\cdot) \in \mathbf{A}_i^2$ such that the triple $(X^*(\cdot), Y^*(\cdot), C^*(\cdot))$ is a solution to the constrained BSDE of Problem 3.1. By Lemma 3.2 we also conclude that this solution is minimal. □

The following "change of variable" result will be needed in the proof of Theorem 4.1.

Proposition 4.1 *For a given process* $g(\cdot) \in \mathbf{H}_1^2$ *and random variable* $\xi \in \mathbf{L}_1^2$*, let*

$$\hat{X}(t) := ess \sup_{\nu \in \mathcal{D}} E^{\nu} \left[\xi + \int_t^T [g(u) - \delta(\nu(u))]du \,\middle|\, \mathcal{F}(t) \right], \quad 0 \le t \le T$$

as in (2.13)*. Then, for any* $\lambda \in \mathbb{R}$*, we have*

$$e^{\lambda t}\hat{X}(t) = ess \sup_{\nu \in \mathcal{D}} E^{\nu} \left[\xi e^{\lambda T} + \int_t^T e^{\lambda u}[G(u) - \delta(\nu(u))]du \,\middle|\, \mathcal{F}(t) \right], \ 0 \le t \le T \tag{4.2}$$

almost surely, where $G(u) := g(u) - \lambda \hat{X}(u)$.

Proof: We recall from (2.21) that the equation

$$\begin{aligned}\hat{X}(t) &= \xi + \int_t^T g(u)du - \int_t^T \hat{Y}'(u)dB(u) + \hat{C}(T) - \hat{C}(t) \\ &= \xi + \int_t^T [g(u) - \delta(\nu(u))]du - \int_t^T \hat{Y}'(u)dB_{\nu}(u) + m(t,T;\nu)\end{aligned}$$

holds almost surely for every process $\nu(\cdot)$ in $\mathcal{D}$, where we have set

$$m(t,r;\nu) := \hat{C}(r) - \hat{C}(t) + \int_t^r [\delta(\nu(u)) - \hat{Y}'(u)\nu(u)]du, \quad 0 \le t \le r \le T.$$

Since $\hat{Y}(\cdot) \in K$, the nonnegative random field $(t,r) \mapsto m(t,r;\nu)$ is nonincreasing in the first variable (t), and nondecreasing in the second variable (r). As in [CK '93], p. 677, there exists a sequence of processes $\{\nu_n(\cdot)\}_{n\in\mathbb{N}} \subseteq \mathcal{D}$ such that

$$\hat{X}(t) = \lim_{n\to\infty} E^{\nu_n} \left[\xi + \int_t^T [g(u) - \delta(\nu_n(u))]du \,\middle|\, \mathcal{F}(t) \right], \ 0 \le t \le T$$

holds almost surely (in fact, one can take $\nu_n(\cdot) \equiv \hat{\nu}_n(\cdot)$ as selected in the proof of Proposition 3.1). Recalling that $\int_0^\cdot \hat{Y}'(u)dB_\nu(u)$ is a $P^\nu-$martingale, we have then

$$E^{\nu_n}\left[\xi + \int_t^T [g(u) - \delta(\nu_n(u))]du \,\middle|\, \mathcal{F}(t)\right] = \hat{X}(t) - M_n(t), \quad a.s. \tag{4.3}$$

and

$$\lim_{n\to\infty} M_n(t) = 0, \qquad a.s.$$

for every fixed $t \in [0,T]$, where

$$\begin{aligned} M_n(t) \;&:=\; E^{\nu_n}\left[m(t,T;\nu_n) \,|\, \mathcal{F}(t)\right] \\ &=\; E^{\nu_n}[m(0,T;\nu_n) \,|\, \mathcal{F}(t)] - m(0,t;\nu_n), \quad 0 \le t \le T \end{aligned} \tag{4.4}$$

is a nonnegative $P^{\nu_n}-$supermartingale with RCLL paths (recall Theorem 1.1.13 in Karatzas & Shreve (1991)).

We deduce from (4.3) that the process $\hat{X}(t) - M_n(t) + \int_0^t [g(u) - \delta(\nu_n(u))]du$ is a $P^{\nu_n}-$martingale. Therefore, by Itô's rule on $e^{\lambda t}\hat{X}(t)$, the process

$$e^{\lambda t}\hat{X}(t) - \int_0^t e^{\lambda u} dM_n(u) + \int_0^t e^{\lambda u}[g(u) - \lambda\hat{X}(u) - \delta(\nu(u))]du, \quad 0 \le t \le T$$

is also a $P^{\nu_n}-$martingale. This implies the equation

$$E^{\nu_n}\left[e^{\lambda T}\xi + \int_t^T e^{\lambda u}[G(u) - \delta(\nu_n(u))]du \middle| \mathcal{F}(t)\right] = e^{\lambda t}\hat{X}(t) + E^{\nu_n}\left[\int_t^T e^{\lambda u} dM_n(u) \middle| \mathcal{F}(t)\right]. \tag{4.5}$$

We want to show that the last term on the right-hand side of (4.5) tends to zero, as $n \to \infty$. First, recall that $M_n(\cdot)$ of (4.4) is an $(\mathbf{F}, P^{\nu_n})-$supermartingale, and integrate by parts to obtain

$$0 \le -E^{\nu_n}\left[\int_t^T e^{\lambda u} dM_n(u) \,\middle|\, \mathcal{F}(t)\right] = e^{\lambda t}M_n(t) + \lambda E^{\nu_n}\left[\int_t^T e^{\lambda u} M_n(u)du \,\middle|\, \mathcal{F}(t)\right]. \tag{4.6}$$

Suppose now that $\lambda \le 0$; since $M_n(\cdot)$ is nonnegative, the right-hand side of (4.6) is bounded from above by $e^{\lambda t}M_n(t)$, which converges to zero as $n \to \infty$. If, on the other hand, $\lambda > 0$, then we have

$$\lambda E^{\nu_n}\left[\int_t^T e^{\lambda u} M_n(u)du \,\middle|\, \mathcal{F}(t)\right] \le M_n(t) \cdot \int_t^T \lambda e^{\lambda u} du \le (e^{\lambda T} - 1) \cdot M_n(t).$$

In conjunction with (4.6), and letting n tend to infinity, we conclude that

$$\lim_{n\to\infty} E^{\nu_n}\left[\int_t^T e^{\lambda u} dM_n(u) \,\middle|\, \mathcal{F}(t)\right] = 0$$

holds almost surely, for every $t \in [0,T]$ fixed.

Returning to (4.5), we obtain from this representation

$$\lim_{n\to\infty} E^{\nu_n}\left[\xi e^{\lambda T} + \int_t^T e^{\lambda u}[G(u) - \delta(\nu_n(u))]du \;\middle|\; \mathcal{F}(t)\right] = e^{\lambda t}\hat{X}(t)$$

and thus also

$$e^{\lambda t}\hat{X}(t) \le ess\sup_{\nu\in\mathcal{D}} E^{\nu}\left[\xi e^{\lambda T} + \int_t^T e^{\lambda u}[G(u) - \delta(\nu(u))]du \middle| \mathcal{F}(t)\right],$$

almost surely. The reverse inequality follows as in the previous section (first part in the proof of Proposition 3.1), after noting that the triple $(e^{\lambda t}\hat{X}(t), \int_0^t e^{\lambda u}d\hat{C}(u), e^{\lambda t}\hat{Y}(t))$ solves the BSDE (3.3), with the terminal condition ξ replaced by $\xi e^{\lambda T}$, with $f(t,\hat{X}(t)))$ replaced by $e^{\lambda t}G(t)$, and with the constraint $\hat{Y}(t) \in K$ replaced by $e^{\lambda t}\hat{Y}(t) \in e^{\lambda t}K$, $0 \le t \le T$.

We conclude that the representation (4.2) holds almost surely, first for $t \in [0,T]$ fixed, and then for all $0 \le t \le T$ simultaneously, thanks to the RCLL regularity of both sides in (4.2) (recall Proposition (2.3)).

□

Proof of Theorem 4.1:

Existence: We have to show that the process $X^*(\cdot)$ of (3.10) solves the stochastic equation (4.1). Fix a process $\nu(\cdot)$ in $\mathcal{D}$, and select an integer n sufficiently large, so that $\nu(\cdot)$ belongs to $\mathcal{D}_n$. From Proposition 3.1 we get

$$X^*(t) \;\; \ge \;\; X_n(t) \ge E^{\nu}\left[\xi + \int_t^T [f(u, X_n(u)) - \delta(\nu(u))]du \;\middle|\; \mathcal{F}(t)\right], \;\; 0 \le t \le T.$$

The comparison theorem ([EPQ], p.23) implies $X^{(0)}(\cdot) \le X_n(\cdot)$, for all $n \in \mathbb{N}$, where $X^{(0)}(\cdot) \in \mathbf{S}_1^2$ is the state-process in the solution $(X^{(0)}(\cdot), Y^{(0)}(\cdot), 0)$ to the unconstrained version

$$X^{(0)}(t) = \xi + \int_t^T f(u, X^{(0)}(u))du - \int_t^T (Y^{(0)}(u))'dB(u), \quad 0 \le t \le T$$

of the BSDE (3.3). Since we also have $X_n(\cdot) \le X^*(\cdot) \le \tilde{X}(\cdot) \in \mathbf{S}_1^2$, by the Lipschitz property of f, we can use the dominated convergence theorem for conditional expectations to conclude that

$$X^*(t) \ge E^{\nu}\left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u))]du \;\middle|\; \mathcal{F}(t)\right]$$

holds almost surely for all $\nu(\cdot) \in \mathcal{D}$, thus

$$X^*(t) \ge ess\sup_{\nu\in\mathcal{D}} E^{\nu}\left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u))du \;\middle|\; \mathcal{F}(t)\right], \;\; 0 \le t \le T.$$

In order to prove the reverse inequality, let us observe that the function

$$F(s,x) := -\lambda x + e^{\lambda s} f(s, e^{-\lambda s} x), \quad 0 \le s \le T, \; x \in \mathbb{R} \tag{4.7}$$

is nondecreasing in the variable x, provided we select $\lambda = -\kappa$, where κ is the Lipschitz constant of the function f as in (3.2). Then, using Proposition 3.1 and the analogue of Proposition 4.1, we get

$$\begin{aligned} e^{\lambda t} X_n(t) &= \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}_n} E^\nu \left[\xi e^{\lambda T} + \int_t^T [F(u, e^{\lambda u} X_n(u)) - e^{\lambda u} \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \\ &\le \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi e^{\lambda T} + \int_t^T [F(u, e^{\lambda u} X^*(u)) - e^{\lambda u} \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \\ &= \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi e^{\lambda T} + \int_t^T e^{\lambda u} [f(u, X^*(u)) - \lambda X^*(u) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] =: X^{(\lambda)}(t). \end{aligned}$$

Therefore, letting $n \to \infty$ leads to $X^*(t) \le e^{-\lambda t} X^{(\lambda)}(t)$, $0 \le t \le T$, and another application of Proposition 4.1, this time to the process $e^{-\lambda t} X^{(\lambda)}(t)$, $0 \le t \le T$, implies

$$\begin{aligned} X^*(t) &\le \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u))] du + \int_t^T \lambda [e^{-\lambda u} X^{(\lambda)}(u) - X^*(u)] du \,\middle|\, \mathcal{F}(t) \right] \\ &\le \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi + \int_t^T [f(u, X^*(u)) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right], \; 0 \le t \le T. \end{aligned}$$

□

Uniqueness: Let $\tilde{X}(\cdot) \in \mathbf{S}_1^2$ be another solution to the stochastic equation (4.1). As in Corollary 4.1, there exist processes $\tilde{C}(\cdot)$ and $\tilde{Y}(\cdot)$, such that $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot))$ is a solution to the BSDE (3.3). In particular then, $\tilde{X}(\cdot)$ has RCLL paths, and Lemma 3.2 implies $X^*(\cdot) \le \tilde{X}(\cdot)$ a.s. In order to prove the reverse inequality, let $\lambda = \kappa$, where again κ is the Lipschitz constant of f as in (3.2), and observe that the function $x \mapsto F(s,x) = -\lambda x + e^{\lambda s} f(s, e^{-\lambda s} x)$ of (4.7) is then nonincreasing. Using Proposition 4.1, we obtain

$$\begin{aligned} e^{\lambda t} \tilde{X}(t) &\ge e^{\lambda t} X^*(t) = \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi e^{\lambda T} + \int_t^T [F(u, e^{\lambda u} X^*(u)) - e^{\lambda u} \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \\ &\ge \operatorname*{\mathit{ess}\,sup}_{\nu \in \mathcal{D}} E^\nu \left[\xi e^{\lambda T} + \int_t^T [F(u, e^{\lambda u} \tilde{X}(u)) - e^{\lambda u} \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \\ &= e^{\lambda t} \tilde{X}(t), \; 0 \le t \le T \end{aligned}$$

almost surely, and uniqueness follows.

□

5 The case of convex drift $f(t, \omega, x, \cdot)$

In this section we study the case of a drift random field f which is also a function of the gains process $Y(\cdot)$. More precisely, we consider a random field $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ which is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)/\mathcal{B}(\mathbb{R})-$measurable and satisfies

$$E \int_0^T f^2(t, \omega, 0, 0) dt < \infty, \tag{5.1}$$

as well as

$$|f(t, \omega, x, y) - f(t, \omega, x', y')| \leq \kappa(|x - x'| + |y - y'|) \tag{5.2}$$

for all $(t, \omega) \in [0, T] \times \Omega$, $(x, x') \in \mathbb{R}^2$ and $(y, y') \in \mathbb{R}^{2d}$, for some $0 < \kappa < \infty$. Our aim is to study the analogue of Problem 3.1, in which the equation (3.3) is replaced by

$$X(t) = \xi + \int_t^T f(u, X(u), Y(u)) du - \int_t^T Y'(u) dB(u) + C(T) - C(t), \; 0 \leq t \leq T. \tag{5.3}$$

We shall refer to this modified problem as *Problem 3.1′*. We shall be able to study the modified problem with minimal extra effort, but under the following assumption.

Assumption 5.1 *The function $y \mapsto f(t, \omega, x, y)$ is convex on $\mathbb{R}^d$, for every $(t, \omega, x) \in [0, T] \times \Omega \times \mathbb{R}$.*

Following [EPQ], we introduce the dual function $\tilde{f}(t, \omega, x, \cdot)$ of the convex function $f(t, \omega, x, \cdot)$ by

$$\tilde{f}(t, \omega, x, \mu) := \sup_{y \in \mathbb{R}^d} [\mu' y - f(t, \omega, x, y)], \; \mu \in \mathbb{R}^d \tag{5.4}$$

for every fixed $(t, \omega, x) \in [0, T] \times \Omega \times \mathbb{R}$, as well as its *effective domain*

$$\tilde{O} := \{(t, \omega, x, \mu) \in [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \; / \; \tilde{f}(t, \omega, x, \mu) < \infty\}. \tag{5.5}$$

As in [EPQ], one can show that each $(t, \omega, x)-$section of $\tilde{O}$, denoted as $\tilde{O}^{t,\omega,x}$, is included in a *bounded* set $\tilde{R}$ in $\mathbb{R}^d$, independent of (t, ω, x). Moreover, we have the following result.

Lemma 5.1 *For any given $(t, \omega) \in [0, T] \times \Omega$, the set $\tilde{O}^{t,\omega,x}$ does not depend on x.*

Proof: Let $\mu \in \tilde{O}^{t,\omega,x}$ for some $(t, \omega, x) \in [0, T] \times \Omega \times \mathbb{R}$. Let $x' \in \mathbb{R}$ be arbitrary. There exists a sequence $\{y_n\}_{n \in \mathbb{N}} \in \mathbb{R}^d$ attaining the (possibly infinite) supremum in the definition of $\tilde{f}(t, \omega, x', \mu)$. We have

$$\begin{aligned} \tilde{f}(t, \omega, x', \mu) - \tilde{f}(t, \omega, x, \mu) &\leq \lim_n [\mu' y_n - f(t, \omega, x', y_n)] + \lim_n \inf [f(t, \omega, x, y_n) - \mu' y_n] \\ &\leq \kappa |x - x'| \end{aligned}$$

and thus $\tilde{f}(t,\omega,x',\mu)<\infty$. $\square$

Consequently, we may omit x in the notation $\tilde{O}^{t,\omega,x}$, and write $\tilde{O}^{t,\omega}$ instead. Let us also introduce the class $\mathcal{A}$ of $\mathbf{F}-$progressively measurable processes $\mu(\cdot):[0,T]\times\Omega\to\tilde{R}$ which satisfy $E\int_0^T\tilde{f}^2(t,0,\mu(t))dt<\infty$.

Lemma 5.2 *For any pair of processes* $(X(\cdot),Y(\cdot))\in\mathbf{S}_1^2\times\mathbf{H}_d^2$, *there exists a process* $\mu(\cdot)\in\mathcal{A}$ *such that*

$$f(t,X(t),Y(t))=\mu'(t)Y(t)-\tilde{f}(t,X(t),\mu(t)),\ 0\le t\le T \tag{5.6}$$

holds almost surely.

This result is proved in $[EPQ]$. We shall also need the following.

Lemma 5.3 *The function* $\tilde{f}(t,\omega,\cdot,\mu)$ *is uniformly Lipschitz in* x*; more precisely, there exists a constant* $C>0$ *such that, for any given* $(t,\omega)\in[0,T]\times\Omega$, $(x,x')\in\mathbb{R}^2$, *and* $\mu\in\tilde{O}^{t,\omega}$, *we have*

$$|\tilde{f}(t,\omega,x,\mu)-\tilde{f}(t,\omega,x',\mu)|\le C|x-x'|.$$

Proof: Fix (t,ω,μ) and (x,x') as above. There exists a sequence $\{y_n\}_{n\in\mathbb{N}}\in\mathbb{R}^d$ such that for any $\varepsilon>0$ and n large enough we have

$$\begin{aligned}\tilde{f}(t,\omega,x,\mu)-\tilde{f}(t,\omega,x',\mu) &\le [\mu'y_n-f(t,\omega,x,y_n)+\varepsilon]-[\mu'y_n-f(t,\omega,x',y_n)]\\ &\le \kappa|x-x'|+\varepsilon.\end{aligned}$$

The inequality with the roles of x,x' interchanged is obtained in a similar fashion, and we conclude. $\square$

For any given pair of processes $(\nu(\cdot),\mu(\cdot))\in\mathcal{D}\times\mathcal{A}$, let us introduce now the exponential martingale

$$Z_{\nu,\mu}(t)=\exp\left\{\int_0^t(\nu(s)+\mu(s))'dB(s)-\frac{1}{2}\int_0^t\|\nu(s)+\mu(s)\|^2ds\right\},\ 0\le t\le T, \tag{5.7}$$

as well as the probability measure

$$P^{\nu,\mu}(A):=E[Z_{\nu,\mu}(T)\mathbf{1}_A],\ A\in\mathcal{F}(T), \tag{5.8}$$

under which the process

$$B_{\nu,\mu}(t):=B(t)-\int_0^t[\nu(s)+\mu(s)]ds,\ 0\le t\le T \tag{5.9}$$

is Brownian motion. We also denote by $E^{\nu,\mu}$ the expectation with respect to the probability measure of (5.8). Moreover, we introduce the *penalized* BSDEs

$$X_n(t)=\xi+\int_t^T[f(u,X_n(u),Y_n(u))+n\rho(Y_n(u))]du-\int_t^T Y_n'(u)dB(u),\ 0\le t\le T, \tag{5.10}$$

for every $n\in\mathbb{N}$, by analogy with (3.5).

Proposition 5.1 *The solution $X_n(\cdot)$ of the penalized BSDE* (5.10) *satisfies the stochastic equation*

$$X_n(t) = ess \sup_{(\nu,\mu)\in\mathcal{D}_n\times\mathcal{A}} E^{\nu,\mu}\left[\xi - \int_t^T [\tilde{f}(u, X_n(u), \mu(u)) + \delta(\nu(u))]du \,\Big|\, \mathcal{F}(t)\right], \; 0 \le t \le T \tag{5.11}$$

almost surely.

The proof is completely analogous to that of Proposition 3.1, and uses Lemma 5.2. In particular, to show that the supremum of (5.11) is attained, we choose $(\nu_n(\cdot), \mu_n(\cdot)) \in \mathcal{D}_n \times \mathcal{A}$ as to have $n\rho(Y_n(\cdot)) + Y_n'(\cdot)\nu_n(\cdot) + \delta(\nu_n(\cdot)) \equiv 0$ and $\tilde{f}(\cdot, X_n(\cdot), \mu_n(\cdot)) \equiv -f(\cdot, X_n(\cdot), Y_n(\cdot)) + \mu_n'(\cdot)Y_n(\cdot)$, a.e. on $[0,T]\times\Omega$.

Assumption 5.2 *There exists at least one solution $(\tilde{X}(\cdot), \tilde{Y}(\cdot), \tilde{C}(\cdot))$ to the constrained BSDE* (5.3) *of Problem 3.1′.*

Under this assumption, one shows as before that the limit

$$X^*(t) := \lim_{n\to\infty} X_n(t), \; 0 \le t \le T \tag{5.12}$$

exists almost surely, and establishes the following analogues of Theorem 4.1 and Corollary 4.1.

Theorem 5.1 *Under the Assumption 5.2, the process $X^*(\cdot)$ of* (5.12) *is the unique solution, in the space $\mathbf{S}_1^2$, of the stochastic equation*

$$X^*(t) = ess \sup_{(\nu,\mu)\in\mathcal{D}\times\mathcal{A}} E^{\nu,\mu}\left[\xi - \int_t^T [\tilde{f}(u, X^*(u), \mu(u)) + \delta(\nu(u))]du \,\Big|\, \mathcal{F}(t)\right], \quad 0 \le t \le T. \tag{5.13}$$

Corollary 5.1 *There exist processes $Y^*(\cdot) \in \mathbf{H}_d^2$ and $C^*(\cdot) \in \mathbf{A}_i^2$ such that the triple $(X^*(\cdot), Y^*(\cdot), C^*(\cdot))$ is the minimal solution to the constrained BSDE* (5.3) *of Problem 3.1′.*

The proofs of these results are parallel to those of Theorem 4.1 and Corollary 4.1, with the help of Lemma 5.2. In particular, the proof of Theorem 5.1 uses the following analogue of Proposition 4.1.

Proposition 5.2 *For a given process $W(\cdot) \in \mathbf{S}_1^2$ and a random variable variable $\xi \in \mathbf{L}_1^2$, let*

$$\hat{X}(t) := ess \sup_{(\nu,\mu)\in\mathcal{D}\times\mathcal{A}} E^{\nu,\mu}\left[\xi - \int_t^T [\tilde{f}(u, W(u), \mu(u)) + \delta(\nu(u))]du \,\Big|\, \mathcal{F}(t)\right], \; 0 \le t \le T.$$

Then, for any $\lambda \in \mathbb{R}$, we have

$$e^{\lambda t}\hat{X}(t) = ess \sup_{(\nu,\mu)\in\mathcal{D}\times\mathcal{A}} E^{\nu,\mu}\left[\xi e^{\lambda T} - \int_t^T e^{\lambda u}[\tilde{f}(u, W(u), \mu(u)) + \lambda\hat{X}(u) + \delta(\nu(u))]du \,\middle|\, \mathcal{F}(t)\right] \tag{5.14}$$

for all $0 \le t \le T$, almost surely.

We only sketch the beginning of the proof of this result, since the rest is similar to that of Proposition 4.1. By analogy with the proof of Proposition 2.5 in the Appendix (and using Lemma 5.2), one shows that the following analogue of (2.21)

$$\begin{aligned}\hat{X}(t) &= \xi + \int_t^T f(u, W(u), \hat{Y}'(u))du - \int_t^T \hat{Y}'(u)dB(u) + \hat{C}(T) - \hat{C}(t) \\ &= \xi - \int_t^T [\tilde{f}(u, W(u), \mu(u)) + \delta(\nu(u))]du - \int_t^T \hat{Y}'(u)dB_{\nu,\mu}(u) + m(t, T; \nu, \mu), \;\; 0 \le t \le T\end{aligned}$$

holds almost surely, for some process $\hat{Y}(\cdot) \in \mathbf{H}^2_d$ taking values in K, some $\hat{C}(\cdot) \in \mathbf{A}^2_i$, and for every pair of processes $(\nu(\cdot), \mu(\cdot))$ in $\mathcal{D} \times \mathcal{A}$. Here we have set

$$\begin{aligned}m(t, r; \nu, \mu) \;&:=\; \hat{C}(r) - \hat{C}(t) + \int_t^r [\delta(\nu(u)) + \tilde{f}(u, W(u), \mu(u)) + f(u, W(u), \hat{Y}(u)) \\ &\qquad - \hat{Y}'(u)(\nu(u) + \mu(u))]du, \;\;\; 0 \le t \le r \le T.\end{aligned}$$

By the definitions of the functions δ in (2.1) and $\tilde{f}$ in (5.4), the nonnegative random field $(t, r) \mapsto m(t, r; \nu, \mu)$ is nonincreasing in the first variable (t), and nondecreasing in the second variable (r). Moreover, there is a sequence $\{\nu_n(\cdot), \mu_n(\cdot)\}_{n\in\mathbb{N}} \subseteq \mathcal{D} \times \mathcal{A}$ such that

$$\hat{X}(t) = \lim_{n\to\infty} E^{\nu_n,\mu_n}\left[\xi - \int_t^T [\tilde{f}(u, W(u), \mu_n(u)) + \delta(\nu_n(u))]du \,\middle|\, \mathcal{F}(t)\right], \;\; 0 \le t \le T$$

holds almost surely. One can take $\nu_n(\cdot) \equiv \hat{\nu}_n(\cdot)$, as in the proof of Proposition 3.1, while $\mu_n(\cdot)$ is selected as in Lemma 5.2, so that $\tilde{f}(\cdot, X_n(\cdot), \mu_n(\cdot)) \equiv \mu_n'(\cdot)Y_n(\cdot) - f(\cdot, X_n(\cdot), Y_n(\cdot))$, a.e. on $[0, T] \times \Omega$. The rest of the proof is similar to that of Proposition 4.1.

□

6 The case of a lower-barrier

Let us suppose now that we are given a process $L(\cdot) \in \mathbf{S}^2_1$ with continuous paths and $L(T) \le \xi$ almost surely, and consider *Problem 2.1* with the requirement

$$X(t) \ge L(t), \;\; 0 \le t \le T \tag{6.1}$$

on its state-process, in addition to (2.6) and (2.7). Similarly, consider the analogue of *Problem 3.1* where, along with (3.3) and (3.4), we impose the lower-bound (6.1) on the state-process.

In both these so-modified problems, denoted henceforth as *Problem 2.1''* and *Problem 3.1''*, respectively, we treat $L(\cdot)$ as a lower-barrier that the state-process $X(\cdot)$ is not allowed to cross on its way to the terminal condition $X(T) = \xi \geq L(T)$. As before, we seek a minimal solution to each of these problems (assuming, of course, that at least one solution exists).

For the unconstrained case $K = \mathbb{R}^d$, these problems were discussed thoroughly by [EKPPQ]. In our setting, it is not hard to modify the theory developed in sections 2-4 in order to take into account the imposition of the lower bound (6.1). For instance, the minimal solution to *Problem 2.1''* is given as

$$\hat{X}(t) = ess \sup_{\substack{\nu \in \mathcal{D} \\ \tau \in \mathcal{S}_{t,T}}} E^{\nu} \left[\xi \mathbf{1}_{\{\tau = T\}} + L(\tau) \mathbf{1}_{\{\tau < T\}} + \int_t^{\tau} [g(u) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \tag{2.13$''$}$$

for $0 \leq t \leq T$, by analogy with Theorem 4.1, where $\mathcal{S}_{t,T}$ denotes the class of $\mathbf{F}-$*stopping times* τ with values in the interval $[t, T]$.

Notice here the need to introduce a *double* optimization problem, of mixed stochastic control/stopping type, in order to represent this minimal solution. The maximization over control processes $\nu(\cdot)$ ensures that the constraint (2.7) on the gains-process is observed; whereas the optimization over stopping times τ guarantees that the state-process $X(\cdot)$ satisfies the constraint (6.1). In other words, $\nu(\cdot)$ and τ play the roles of "dual (adjoint) variables" that enforce the constraints (2.7) and (6.1), respectively.

By analogy with Theorem 4.1 and its Corollary, there is now a unique process $X^*(\cdot)$ in the space $\mathbf{S}_1^2$ that solves the *stochastic functional equation*

$$X^*(t) = ess \sup_{\substack{\nu \in \mathcal{D} \\ \tau \in \mathcal{S}_{t,T}}} E^{\nu} \left[\xi \mathbf{1}_{\{\tau = T\}} + L(\tau) \mathbf{1}_{\{\tau < T\}} + \int_t^{\tau} [f(u, X^*(u)) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \tag{4.1$''$}$$

for $0 \leq t \leq T$, and this $X^*(\cdot)$ is the state-process of the minimal solution to *Problem 3.1''*. As in section 3, it is constructed through a penalization scheme which now takes a more complicated form due to the presence of the "reflecting lower-barrier", namely:

$$\left\{ \begin{array}{c} X_n(t) = \xi + \int_t^T [f(u, X_n(u)) + n\rho(Y_n(u))] du - \int_t^T Y_n'(u) dB(u) + C_n(T) - C_n(t) \\ X_n(t) \geq L(t), \ 0 \leq t \leq T \\ C_n(\cdot) \text{ continuous, increasing and } \int_0^T [X_n(t) - L(t)] dC(t) = 0 \end{array} \right\} \tag{3.5$''$}$$

almost surely, for a suitable triple $(X_n(\cdot), Y_n \cdot), C_n(\cdot)) \in \mathbf{S}_1^2 \times \mathbf{H}_d^2 \times \mathbf{A}_i^2, \ n \in \mathbb{N}$.

The solvability of the system (3.5)$''$, and the a.s. comparison $X_n(\cdot) \leq X_{n+1}(\cdot)$, $n \in \mathbb{N}$, are consequences of Theorems 4.1, 5.2 in [EKPPQ]. The state-process of the (unique) solution to (3.5)$''$ satisfies the equation

$$X_n(t) = ess \sup_{\substack{\nu \in \mathcal{D} \\ \tau \in \mathcal{S}_{t,T}}} E^{\nu} \left[\xi \mathbf{1}_{\{\tau = T\}} + L(\tau)\mathbf{1}_{\{\tau < T\}} + \int_t^{\tau} [f(u, X_n(u)) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right] \tag{3.6$''$}$$

for $0 \leq t \leq T$. This supremum is attained by the pair $(\nu, \tau) = (\nu_n(\cdot), \tau_n(\cdot))$, where $\nu_n(\cdot)$ satisfies $\rho(Y_n(\cdot)) + Y_n'(\cdot)\nu_n(\cdot) + \delta(\nu_n(\cdot)) = 0$ a.e. on $[0, T]$ as in the proof of Proposition (3.1), and

$$\tau_n(t) := \inf\{u \in [t, T) \ / \ X_n(u) = L(u)\} \wedge T, \tag{6.2}$$

namely

$$X_n(t) = E^{\nu_n} \left[\xi \mathbf{1}_{\{\tau_n(t) = T\}} + L(\tau_n(t))\mathbf{1}_{\{\tau_n(t) < T\}} + \int_t^{\tau_n(t)} [f(u, X_n(u)) - \delta(\nu_n(u))] du \,\middle|\, \mathcal{F}(t) \right].$$

One can also show that the limit-process $X^*(t) := \lim \uparrow X_n(t)$, $0 \leq t \leq T$ is the minimal solution of *Problem 3.1*$''$.

The details of these derivations are more-or-less straightforward, with the possible exception of the proof of the change-of-variable formula

$$e^{\lambda t} \hat{X}(t) = ess \sup_{\substack{\nu \in \mathcal{D} \\ \tau \in \mathcal{S}_{t,T}}} E^{\nu} \left[\xi \mathbf{1}_{\{\tau = T\}} + L(\tau)\mathbf{1}_{\{\tau < T\}} + \int_t^{\tau} e^{\lambda u} [g(u) - \lambda \hat{X}(u) - \delta(\nu(u))] du \,\middle|\, \mathcal{F}(t) \right], \tag{4.2$''$}$$

valid for every $\lambda \in \mathbb{R}$, for the process $\hat{X}(\cdot)$ of (2.13)$''$ (analogue of Proposition (4.1)). This formula plays again a crucial role in establishing the existence and uniqueness of solution to the stochastic functional equation (4.1)$''$. We shall leave these details to the care of the diligent reader.

□

7 Appendix

In this section, we sketch the proofs of Propositions 2.3 and 2.5, by adapting to our current situation the techniques developed in [CK '93] and ElKaroui & Quenez (1995).

Proof of Proposition 2.3: With the notation $g_\nu(\cdot) := g(\cdot) - \delta(\nu(\cdot))$, let us start by establishing the *equation*

$$\hat{X}(t) = ess \sup_{\nu \in \mathcal{D}_{t,\theta}} E^\nu \left[\hat{X}(\theta) + \int_t^\theta g_\nu(u) du \,\Big|\, \mathcal{F}(t) \right], \quad a.s. \tag{7.1}$$

of Dynamic Programming, for every $0 \le t \le \theta \le T$. We have denoted by $\mathcal{D}_{t,\theta}$ the restriction of $\mathcal{D}$ to the set $[t, \theta] \times \Omega$; note that (7.1) with $\theta = T$ becomes just the definition of $\hat{X}(t)$ in (2.13), since $\hat{X}(T) = \xi$. Let us observe also that, for any $\nu(\cdot) \in \mathcal{D}$ and with the notation $Z_\nu(t, \theta) := Z_\nu(\theta)/Z_\nu(t)$ as in (2.3), the random variable

$$J_\nu(\theta) := E^\nu \left[\xi + \int_\theta^T g_\nu(u) du \,\Big|\, \mathcal{F}(\theta) \right] = E \left[Z_\nu(\theta, T) \left\{ \xi + \int_\theta^T g_\nu(u) du \right\} \,\Big|\, \mathcal{F}(\theta) \right] \tag{7.2}$$

depends only on the restriction of the process $\nu(\cdot)$ to $[\theta, T] \times \Omega$. In particular, we obtain from (2.13) written in the form

$$\hat{X}(\theta) = ess \sup_{\nu \in \mathcal{D}} J_\nu(\theta), \tag{7.3}$$

that

$$\hat{X}(t) = ess \sup_{\nu \in \mathcal{D}} E^\nu \left[J_\nu(\theta) + \int_t^\theta g_\nu(u) du \,\Big|\, \mathcal{F}(t) \right] \le ess \sup_{\nu \in \mathcal{D}_{t,\theta}} E^\nu \left[\hat{X}(\theta) + \int_t^\theta g_\nu(u) du \,\Big|\, \mathcal{F}(t) \right],$$

holds almost surely. In order to prove the reverse inequality, it suffices to fix an arbitrary process $\mu(\cdot)$ in $\mathcal{D}$ and show that

$$\hat{X}(t) \ge E^\mu \left[\hat{X}(\theta) + \int_t^\theta g_\mu(u) du \,\Big|\, \mathcal{F}(t) \right] \tag{7.4}$$

holds almost surely, for any $0 \le t \le \theta \le T$. To this end, notice that the family $\{J_\nu(\theta)\}_{\nu \in \mathcal{D}}$ as in (7.2), is *directed upwards*: for any two processes $\mu(\cdot)$ and $\nu(\cdot)$ in $\mathcal{D}$, there exists a third process $\lambda(\cdot) \in \mathcal{D}$, such that $J_\lambda(\theta) \ge max(J_\mu(\theta), J_\nu(\theta))$ holds almost surely. Thus (e.g. Neveu (1975)) we can write the essential supremum of (7.3) in the form

$$\hat{X}(\theta) = \lim_{k \to \infty} \uparrow J_{\nu_k}(\theta), \quad a.s. \tag{7.5}$$

of an increasing limit, for some sequence $\{\nu_k(\cdot)\}_{k \in \mathbb{N}}$ of processes in $\mathcal{D}_{\theta,T}$; and without loss of generality, this sequence can be selected from the class $\mathcal{M}_{t,\theta} := \{\nu(\cdot) \in \mathcal{D}/\nu(\cdot) \equiv \mu(\cdot) \text{ on } [t, \theta] \times \Omega\}$. Now we have

$$\hat{X}(t) \ge E^\nu \left[J_\nu(\theta) + \int_t^\theta g_\nu(u) du \,\Big|\, \mathcal{F}(t) \right] = E^\mu \left[J_\nu(\theta) + \int_t^\theta g_\mu(u) du \,\Big|\, \mathcal{F}(t) \right], \quad a.s.$$

for every process $\nu(\cdot)$ in $\mathcal{M}_{t,\theta}$; thus, by (7.5) and the monotone convergence theorem, we obtain

$$\begin{aligned}
\hat{X}(t) &\geq \lim_{k\to\infty}\uparrow \; E^{\mu}\left[J_{\nu_k}(\theta)+\int_t^{\theta} g_{\mu}(u)du \,\middle|\, \mathcal{F}(t)\right] \\
&= E^{\mu}\left[\lim_{k\to\infty}\uparrow \; J_{\nu_k}(\theta)+\int_t^{\theta} g_{\mu}(u)du \,\middle|\, \mathcal{F}(t)\right] \\
&= E^{\mu}\left[\hat{X}(\theta)+\int_t^{\theta} g_{\mu}(u)du \,\middle|\, \mathcal{F}(t)\right], \quad a.s.
\end{aligned}$$

This proves (7.4), and thus also the $P^{\mu}-$supermartingale property of the process $\hat{X}(t) + \int_0^t g_{\mu}(u)du,\ 0 \leq t \leq T$. The RCLL regularity of the process $\hat{X}(\cdot)$ is then argued as in [CK '93], pp. 679-680.

Proof of Proposition 2.5: For any process $\mu(\cdot)$ in the class $\mathcal{D}$ of (2.5), we have from (2.16)-(2.18) and (2.10):

$$\begin{aligned}
\hat{Q}(t) &= \hat{X}(t)+\int_0^t g(u)du \\
&= \hat{X}(0)+\int_0^t \delta(\nu(u))du+\int_0^t Y_{\nu}'(u)[dB_{\mu}(u)+(\mu(u)-\nu(u))du]-A^{(\nu)}(t) \\
&= \hat{X}(0)+\int_0^t Y_{\nu}'(u)dB_{\mu}(u)+\int_0^t [\delta(\nu(u))+(\mu(u)-\nu(u))'Y_{\nu}(u)]du-A^{(\nu)}(t)
\end{aligned}$$

for $0 \leq t \leq T$. But again from (2.17), now read with $\nu(\cdot)$ replaced by $\mu(\cdot)$, the process $\hat{Q}(\cdot)$ has the $P^{\mu}-$ supermartingale representation

$$\hat{Q}(t)=\hat{X}(0)+\int_0^t Y_{\mu}'(u)dB_{\mu}(u)+\int_0^t \delta(\mu(u))du-A^{(\mu)}(t), \quad 0\leq t\leq T.$$

The equality of these two decompositions leads to the identities of (2.19) and (2.20), whereas (2.21) follows from the P^{o}-decomposition of $\hat{Q}(\cdot)$.

Consider now the product set $F_{\nu} := \{(t,\omega)/\delta(\nu(t,\omega) < \nu'(t,\omega)\hat{Y}(t,\omega)\}$, and suppose that, for some process $\nu(\cdot)$ in $\mathcal{D}$, we have $(\lambda \otimes P)(F_{\nu}) > 0$. Then, for any real constant $k > 0$, the process

$$\mu(\cdot) := \nu(\cdot)\cdot 1_{F_{\nu}^c} + k\nu(\cdot)\cdot 1_{F_{\nu}}$$

belongs to $\mathcal{D}$, and we have

$$EA^{(\mu)}(T)=E\hat{C}(T)+\int\int_{F_{\nu}^c}(\delta(\nu))-\nu'\hat{Y})dtdP+k\int\int_{F_{\nu}}(\delta(\nu))-\nu'\hat{Y})dtdP<0,$$

for $k > 0$ sufficiently large. This contradicts the fact that $A^{(\mu)}(T) \geq 0$ holds with probability one. Therefore, $(\lambda \otimes P)(F_\nu) = 0$ holds for every $\nu(\cdot) \in \mathcal{D}$. In particular, for all (t, ω) in a set of full product-measure, we have

$$x'\hat{Y}(t, \omega) \leq \delta(x), \quad \forall x \in \tilde{K}$$

(observe that both sides are continuous on $\tilde{K}$, as functions of x); but this leads to (2.22), since the set K is closed, as in Theorem 13.1 of Rockafellar (1971).

References

[1] Arkin, V. & Saxonov, M. (1979) Necessary optimality conditions for stochastic diferential equations. *Soviet Mathematics Doklady* **20**, 1-15.

[2] Barles, G., Buckdahn, R. & Pardoux, E. (1997) Backward stochastic differential equations and integral-partial differential equations. *Stochastics & Stochastics Reports* **60**, 57-83.

[3] Bensoussan, A. (1981) Lectures on Stochastic Control. *Lecture Notes in Mathematics* **972**, 1-62.

[4] Bismut, J.M. (1978) An introductory approach to duality in optimal stochastic control. *SIAM Review* **20**, 62-78.

[5] Broadie, M., Cvitanić, J. & Soner, H.M. (1996) Optimal replication of contingent claims under portfolio constraints. To appear in *Review of Financial Studies.*

[6] Buckdahn, R. & Hu, Y. (1996) Hedging contingent claims for a large investor in an incomplete market. Preprint.

[7] Buckdahn, R. & Hu, Y. (1997) Pricing of American contingent claims with jump stock price and constrained portfolios. Preprint.

[8] Cvitanić, J. & Karatzas, I. (1992) Convex duality for constrained porfolio optimization. *Ann. Appl. Probab.* **2**, 767-818.

[9] Cvitanić, J. & Karatzas, I. (1993) Hedging contingent claims with constrained portfolios. *Ann. Appl. Probab.* **3**, 652-681.

[10] Cvitanić, J. & Karatzas, I. (1996) Backwards stochastic differential equations with reflection and Dynkin games. *Ann. Probab.* **24**, 2024-2056.

[11] Darling, R.W.R. & Pardoux, E. (1997) Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *Ann. Probab* **25**, 1135-1159.

[12] Duffie, D. & Epstein, L. (1992) Stochastic differential utility. *Econometrica* **60**, 353-394.

[13] Elliott, R.J. (1990) The optimal control of diffusions. *Appl. Math. & Optimiz.* **22**, 229-240.

[14] El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S. & Quenez, M.C. (1997) Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *Ann. Probab.* **25**, 702-737.

[15] El Karoui, N., Peng, S. & Quenez, M.C. (1997) Backward Stochastic Differential Equations in finance. *Math. Finance*, **7**, 1-71.

[16] El Karoui, N. & Quenez, M.C. (1995) Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control & Optimization* **33**, 29-66.

[17] Hamadène, S. & Lepeltier, J.M. (1995a) Backward equations, stochastic control, and zero-sum stochastic differential games. *Stochastics & Stochastics Reports* **54**, 221-231.

[18] Hamadène, S. & Lepeltier, J.M. (1995b) Zero-sum stochastic differential games and backward equations. *Systems & Control Letters* **24**, 259-263.

[19] Haussmann, U.G. (1986) *A Stochastic Maximum Principle for Optimal Control of Diffusions.* Longman, Essex, UK.

[20] Itô, K. (1942) Differential equations determining Markov processes (in Japanese). *Zenkoku Shijō Sūgaku Danwakai* **1077**, 1352-1400.

[21] Itô, K. (1946) On a stochastic integral equation. *Proc. Imperial Acad. Tokyo* **22**, 32-35.

[22] Itô, K. (1951) On stochastic differential equations. *Mem. Amer. Math. Soc.* **4**, 1-51.

[23] Karatzas, I. & Shreve, S.E. (1991) *Brownian Motion and Stochastic Calculus*, 2nd edition. Springer-Verlag, New York.

[24] Ma, J. & Cvitanić, J. (1997) Reflected Forward-Backward SDEs and obstacle problems with boundary conditions. Preprint.

[25] Pardoux, E. & Peng, S. (1990) Adapted solution of a backward stochastic differential equation. *Systems & Control Letters* **14**, 55-61.

[26] Pardoux, E. & Peng, S. (1994) Backward doubly stochastic differential equations and systems of quasilinear SPDEs. *Probab. Theory & Related Fields* **98**, 209-227.

[27] Pardoux, E. & Tang, S. (1996) The study of Forward-Backward Stochastic Differential Equation and its application in quasilinear PDEs. Preprint.

[28] Peng, S. (1990) A general stochastic maximum principle for optimal control problems. *SIAM J. Control & Optimization* **28**, 966-979.

[29] Peng, S. (1991) Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics & Stochastics Reports* **37**, 61-74.

[30] Peng, S. (1993) Backwards stochastic differential equations and applications to optimal control. *Appl. Math. & Optimization* **27**, 125-144.

[31] Rockafellar, R.T. (1970) *Convex Analysis.* Princeton University Press, Princeton.

[32] Saksonov, M. (1989) *The Variational Principles in Stochastic Control.* Preprint, Dushanbe State Educational Institute.