Optimal control of SDEs with expected path constraints and related constrained FBSDEs*

Ying Hu† Shanjian Tang‡ Zuo Quan Xu§

August 1, 2022

Abstract

In this paper, we consider optimal control of stochastic differential equations subject to an expected path constraint. The stochastic maximum principle is given for a general optimal stochastic control in terms of constrained FBSDEs. In particular, the compensated process in our adjoint equation is deterministic, which seems to be new in the literature. For the typical case of linear stochastic systems and quadratic cost functionals (i.e., the so-called LQ optimal stochastic control), a verification theorem is established, and the existence and uniqueness of the constrained reflected FBSDEs are also given.

MSC2020 subject classifications: Primary 93E20, 60H30; secondary 60H10, 93E03.

Keywords: optimal stochastic control, stochastic maximum principle, expected path constraint, reflected FBSDE.

---

*Ying Hu is partially supported by Lebesgue Center of Mathematics "Investissements d'avenir" program ANR-11-LABX-0020-01, by CAESARS-ANR-15-CE05-0024, and by MFG-ANR-16-CE40-0015-01. Shanjian Tang is partially supported by the National Science Foundation of China (Nos. 11631004 and 12031009). Zuo Quan Xu is partially supported by NSFC (No.11971409), the Research Grants Council of Hong Kong (GRF No.15202421), the PolyU-SDU Joint Research Center on Financial Mathematics, the CAS AMSS-POLYU Joint Laboratory of Applied Mathematics, and the Hong Kong Polytechnic University.

†Univ. Rennes, CNRS, IRMAR-UMR6625, F-35000, Rennes, France; School of Mathematical Sciences, Fudan University, Shanghai 200433, China (*E-mail: ying.hu@univ-rennes1.fr*).

‡Department of Finance and Control Sciences, School of Mathematical Sciences, Fudan University, Shanghai 200433, China (*e-mail: sjtang@fudan.edu.cn*).

§Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China (*e-mail: maxu@polyu.edu.hk*).

1 Introduction

In this paper, we consider the following real-valued controlled stochastic differential equation (SDE):

$$X_t = x + \int_0^t b(s, X_s, u_s)\, ds + \int_0^t \sigma(s, X_s, u_s)^\intercal\, dW_s, \quad t \in [0, T] \tag{1.1}$$

with the pointwise-in-time expected path constraint

$$\mathbb{E}[f(t, X_t)] \geqslant 0, \quad t \in [0, T]. \tag{1.2}$$

The cost functional reads

$$J(u) := \mathbb{E}\left[\int_0^T \ell(t, X_t, u_t)\, dt + g(X_T)\right].$$

The study of Stochastic Maximum Principles (SMPs) is traced back to Bismut [1, 2, 3], who introduced the notion of backward stochastic differential equations (BSDEs) to formulate the adjoint process and the stochastic Riccati equation, and was subsequently developed by Kushner [11] and Haussmann [10]. Initially, SMPs concerned only the stochastic systems where the control domain is convex or the diffusion coefficient does not contain a control variable, and the proof only involves the first-order expansion. Peng [12] established the SMP for the general stochastic optimal control problem where the control domain does not need to be convex and the diffusion coefficient can contain the control variable, where the second-order expansion and second-order backward stochastic differential equation are introduced. An extensive account of progress on SMPs is available in Yong and Zhou [14]. Recently, SMP found wide application in probabilistic analysis of mean-field games, see [7].

Our optimal stochastic control is featured by the inclusion of the expected path constraint. Our first aim is to establish a necessary condition (i.e., SMP) for this type of stochastic control problem, where the adjoint equation is a mean-reflected BSDE with the reflection being the consequence of the expected system path constraint. We note that a similar SMP was already established by Frankowska et al. [15]; in contrast to theirs, our compensated process μ_t is deterministic, which carries more information on the optimal control. For related results on optimal control of ordinary differential equations, see Dmitruk and Osmolovskii [8] and Bourdin [4]

While applying this SMP to a stochastic control problem, a new type of coupled reflected forward-backward stochastic differential equation (FBSDE) appears:

$$\begin{cases} dX_t = (A_t X_t - B_t^\intercal R_t^{-1}(B_t Y_t + D_t^\intercal Z_t))\, dt \\ \qquad\qquad + (C_t X_t - D_t R_t^{-1}(B_t Y_t + D_t^\intercal Z_t))^\intercal\, dW_t, \\ dY_t = -(Q_t X_t + A_t Y_t + C_t^\intercal Z_t)\, dt + d\mu_t + Z_t^\intercal\, dW_t, \\ \mathbb{E}[X_t] \geqslant L_t, \quad \int_0^T (\mathbb{E}[X_t] - L_t)\, d\mu_t = 0, \\ X_0 = x, \quad Y_T = G X_T, \quad \mu_T = 0. \end{cases}$$

This type of equation can be considered as an FBSDE counterpart of BSDEs with mean reflection introduced by Briand et al. in [6] and further studied by [5]. We will give a verification theorem and some well-solvability result concerning this new type of FBSDE.

The paper is organized as follows. After introducing some notation in the next subsection, we give the formulation of the problem in section 2. In section 3, we apply Ekeland's variational principle to deduce the stochastic maximum principle for the stochastic control problem. In section 4, we introduce the reflected FBSDE and show the verification theorem. The last two sections are devoted to the proof of uniqueness (section 5) and of existence (section 6).

1.1 Notation

Let $(W_t)_{0\leqslant t\leqslant T} = (W_t^1, \cdots, W_t^m)_{0\leqslant t\leqslant T}$ be an m-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote by $\{\mathcal{F}_t\}_{t\in[0,T]}$ the augmented filtration generated by (W_t). Let $\mathbb{R}^+$ and $\mathbb{R}^-$, respectively, denote the sets of nonnegative and nonpositive real numbers. We write $x_+ = \max\{x,\ 0\}$ and $x_- = \max\{-x,\ 0\}$ for $x \in \mathbb{R}$, the set of real numbers.

We often use vectors and matrices in this paper, where all vectors are column vectors. For a vector or matrix M, denote by $M^\intercal$ the transpose of M, and by $|M| = \sqrt{\sum_{i,j} m_{ij}^2}$ the Frobenius norm.

We use the following notation.

- U: a given closed convex subset of $\mathbb{R}^l$.
- $\mathcal{U}[0,T]$: the set of $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted functions $u(\cdot) : [0,T]\times\Omega \to U$ such that $\mathbb{E}\left[\int_0^T |u_t|^2\, dt\right] < \infty$.
- $L^p_{\mathcal{F}}([0,\ T];\ \mathbb{R}^k)$: the set of $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted processes $f = (f_t^1, \cdots, f_t^k)_{0\leqslant t\leqslant T}$ with $\mathbb{E}\left[\int_0^T |f_t|^p\, dt\right] < \infty$.
- $L^\infty_{\mathcal{F}}([0,\ T];\ \mathbb{R}^k)$: the set of essentially bounded $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted $\mathbb{R}^k$-valued processes on $[0,T]$.
- $L^\infty([0,\ T];\ \mathbb{R}^k)$: the set of essentially bounded deterministic measurable $\mathbb{R}^k$-valued functions on $[0,T]$.
- $C_{\mathcal{F}}([0,T];\ \mathbb{R}^k)$: the Banach space of all continuous $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted $\mathbb{R}^k$-valued processes f on $[0,T]$ with a finite squared norm $\mathbb{E}\left[\max_{t\in[0,T]} |f_t|^2\right]$.
- $\mathcal{M}^-([0,T])$: the set of all nonpositive Radon measures on $[0,T]$.
- $\mathcal{M}^+([0,T])$: the set of all nonnegative Radon measures on $[0,T]$.

For $\mu \in \mathcal{M}^+([0,T]) \cup \mathcal{M}^-([0,T])$, we write

$$\mu_t = \mu([0,t]) - \mu([0,T]).$$

Then the map $t \mapsto \mu_t$ is a càdlàg function on $[0,T]$ with $\mu_T = 0$.

2 Problem formulation

Consider the following $\mathbb{R}$-valued controlled SDE:

$$X_t = x + \int_0^t b(s, X_s, u_s)\, ds + \int_0^t \sigma(s, X_s, u_s)^\intercal\, dW_s, \quad t \in [0,T] \tag{2.1}$$

with the pointwise-in-time expected path constraint

$$\mathbb{E}[f(t, X_t)] \geqslant 0, \quad t \in [0,T]. \tag{2.2}$$

The cost functional reads

$$J(u) := \mathbb{E}\left[\int_0^T \ell(t, X_t, u_t)\, dt + g(X_T)\right].$$

In the above, $(b,\sigma) : [0,T] \times \mathbb{R} \times U \times \Omega \to \mathbb{R} \times \mathbb{R}^m$, $f : [0,T] \times \mathbb{R} \to \mathbb{R}$, $\ell : [0,T] \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$, and $g : \mathbb{R} \times \Omega \to \mathbb{R}$.

Let us assume the following conditions:

(H1) The maps b, σ, f, ℓ, and g are measurable. They are all continuously differentiable w.r.t. (x,v).

(H2) There exists a constant $L > 0$ such that

$$\begin{cases} |b_x(t,x,v)| + |b_v(t,x,v)| + |\sigma_x(t,x,v)| + |\sigma_v(t,x,v)| \leqslant L, \\ |\ell_x(t,x,v)| + |\ell_v(t,x,v)| \leqslant L(1 + |x| + |v|), \\ |g_x(x)| + |f_x(t,x)| \leqslant L(1 + |x|), \\ |b(t,0,0)| + |\sigma(t,0,0)| + |\ell(t,0,0)| + |g(0)| + |f(t,0)| \leqslant L, \end{cases}$$

for any $(t,x,v,\omega) \in [0,T] \times \mathbb{R} \times U \times \Omega$.

We call a control $u \in \mathcal{U}[0,T]$ admissible if the SDE (2.1) admits a unique strong solution $X(\cdot)$ such that the constraint (2.2) is satisfied. The set of all admissible controls is denoted by $\mathcal{U}_{ad}[0,T]$. We study the following optimal stochastic control problem

$$\min_{u \in \mathcal{U}_{ad}[0,T]} J(u). \tag{2.3}$$

3 Ekeland's variational principle and stochastic maximum principle

We use Ekeland's variational principle to study the optimization problem (2.3). Before proceeding, we first present two technical lemmas. Denote by $C([0,T];I)$ the set of all continuous functions $f:[0,T]\to I$ with $I=\mathbb{R},\mathbb{R}^+$. Set $\mathbb{K}:=C([0,T];\mathbb{R}^+)$, and define the distance function

$$d_{\mathbb{K}}(X):=\inf_{Y\in\mathbb{K}}\|X-Y\|_\infty,\quad X\in C([0,T];\mathbb{R}),$$

with $\|\cdot\|_\infty$ being the maximal norm in $C([0,T];\mathbb{R})$.

Lemma 3.1. *For any $X\in C([0,T];\mathbb{R})$, we have*

$$d_{\mathbb{K}}(X)=\max_{t\in[0,T]}X_-(t)$$

Proof. First, since $X_+\in\mathbb{K}$, we have

$$d_{\mathbb{K}}(X)=\inf_{Y\in\mathbb{K}}\|X-Y\|_\infty\leqslant\|X-X_+\|_\infty=\|X_-\|_\infty=\max_{t\in[0,T]}X_-(t).$$

It only remains to show the reverse inequality $d_{\mathbb{K}}(X)\geqslant\|X_-\|_\infty$. If $\|X_-\|_\infty=0$, then this inequality holds trivially. Otherwise, there is $t^*\in[0,T]$ such that

$$X(t^*)=\min_{t\in[0,T]}X(t)=-\|X_-\|_\infty<0.$$

So we have

$$d_{\mathbb{K}}(X)\geqslant\inf_{Y\in\mathbb{K}}|X(t^*)-Y(t^*)|\geqslant|X(t^*)|=\|X_-\|_\infty.$$

□

The subdifferential of the function $d_{\mathbb{K}}$ at X, denoted by $\partial d_{\mathbb{K}}(X)$, is defined to be the set of $\mathbb{R}$-valued Radon measures K on $[0,T]$ such that

$$\langle K,f\rangle:=\int_{[0,T]}f(t)K(dt)\leqslant d_{\mathbb{K}}(X+f)-d_{\mathbb{K}}(X),\quad\forall f\in\mathbb{K}.$$

Lemma 3.2. *For any $X\in C([0,T];\mathbb{R})$, the set $\partial d_{\mathbb{K}}(X)$ is not empty and*

$$\partial d_{\mathbb{K}}(X)\subseteq\mathcal{M}^-([0,T])$$

with

$$\text{supp}\ \partial d_{\mathbb{K}}(X)\subseteq\text{argmin}\ X.$$

Furthermore, if $X\notin\mathbb{K}$, we have $|K([0,T])|=1$ for any $K\in\partial d_{\mathbb{K}}(X)$.

Proof. We now show $\partial d_{\mathbb{K}}(X)$ is not empty. If $d_{\mathbb{K}}(X) = 0$, then trivially $K \equiv 0 \in \partial d_{\mathbb{K}}(X)$. Otherwise $d_{\mathbb{K}}(X) = -X(t^*) > 0$ for some $t^* \in [0, T]$. Let $-K$ be the Dirac measure at t^*. Then, by Lemma 3.1, for any $h \in \mathbb{K}$,

$$\begin{aligned} d_{\mathbb{K}}(X+h) - d_{\mathbb{K}}(X) &= \max_{t\in[0,T]} (X(t)+h(t))_- + X(t^*) \\ &\geqslant \max_{t\in[0,T]} (-(X(t)+h(t))) + X(t^*) \geqslant -h(t^*) = \langle K, h\rangle. \end{aligned}$$

Therefore, $K \in \partial d_{\mathbb{K}}(X)$ and hence $\partial d_{\mathbb{K}}(X)$ is not empty.

For any $(K, h) \in \partial d_{\mathbb{K}}(X) \times \mathbb{K}$, by Lemma 3.1,

$$\langle K, h\rangle \leqslant \lim_{\alpha\downarrow 0} \frac{d_{\mathbb{K}}(X+\alpha h) - d_{\mathbb{K}}(X)}{\alpha} = \lim_{\alpha\downarrow 0} \frac{\|(X+\alpha h)_-\|_\infty - \|X_-\|_\infty}{\alpha} \leqslant 0.$$

Hence, K is nonpositive.

If t_0 is not a minimum point of X, then X has no minimum point on $[t_0 - \varepsilon, t_0 + \varepsilon]$ for sufficiently small $\varepsilon > 0$. For any $h \in C([0, T]; \mathbb{R})$ with $\operatorname{supp} h \subset (t_0 - \varepsilon/2, t_0 + \varepsilon/2)$, we have that t_0 is not a minimum point of $X \pm \alpha h$ for sufficiently small $\alpha > 0$, so

$$\|(X \pm \alpha h)_-\|_\infty = \|X_-\|_\infty,$$

which by definition implies $\langle K, h\rangle = 0$. Hence, $\operatorname{supp} \partial d_{\mathbb{K}}(X) \subseteq \operatorname{argmin} X$.

The last assertion is referred to in [13, Proposition 3.11, p. 146]. □

Let us first recall Ekeland's variational principle (see [9, Theorem 1.1]).

Lemma 3.3 (Ekeland's variational principle)**.** *Let $(V, d(\cdot,\cdot))$ be a complete metric space and $F(\cdot) : V \to \mathbb{R}$ be a lower semi-continuous function, bounded from below. Suppose there exist $u \in V$ and $\varepsilon > 0$ such that*

$$F(u) \leqslant \inf_{v\in V} F(v) + \varepsilon.$$

Then, there exists $u_\varepsilon \in V$ such that

(i) $F(u_\varepsilon) \leqslant F(u)$,

(ii) $d(u, u_\varepsilon) \leqslant \sqrt{\varepsilon}$, *and*

(iii) $F(v) + \sqrt{\varepsilon} d(v, u_\varepsilon) \geqslant F(u_\varepsilon)$ *for all* $v \in V$.

We will work on the space $\mathcal{U}[0, T]$. To apply Ekeland's variational principle, we define a metric d such that $(\mathcal{U}[0, T], d)$ is a complete metric space. For this, set

$$d(v, u) = \left(\mathbb{E}\left[\int_0^T |v(t) - u(t)|^2 dt\right]\right)^{1/2}.$$

Then $(\mathcal{U}[0, T], d(\cdot,\cdot))$ forms a complete metric space.

Let $u^* \in \mathcal{U}_{ad}[0,T]$ be an optimal control for problem (2.3). For $\varepsilon > 0$ and $u \in \mathcal{U}[0,T]$, define the functional

$$\begin{aligned} J_\varepsilon(u) &= \left(\left([J(u) - J(u^*) + \varepsilon]_+ \right)^2 + \left(\max_{t\in[0,T]} \left(\mathbb{E}[f(t, X_t^u)] \right)_- \right)^2 \right)^{\frac{1}{2}} \\ &= \left(\left([J(u) - J(u^*) + \varepsilon]_+ \right)^2 + d_{\mathbb{K}}^2 \left(\mathbb{E}[f(\cdot, X_\cdot^u)] \right) \right)^{\frac{1}{2}}, \end{aligned}$$

where the second equation is due to Lemma 3.1. If $J_\varepsilon(u^\varepsilon) = 0$, then $[J(u^\varepsilon) - J(u^*) + \varepsilon]_+ = 0$ and $u^\varepsilon \in \mathcal{U}_{ad}[0,T]$, contradicting the optimality of u^* to problem (2.3). So we have $J_\varepsilon(u^\varepsilon) > 0$.

Since

$$J_\varepsilon(u^*) = \varepsilon \leqslant \inf_{u\in\mathcal{U}[0,T]} J_\varepsilon(u) + \varepsilon,$$

by Ekeland's variational principle, Lemma 3.3, there is $u^\varepsilon \in \mathcal{U}[0,T]$ such that

(i) $J_\varepsilon(u^\varepsilon) \leqslant J_\varepsilon(u^*)$,

(ii) $d(u^*, u^\varepsilon) \leqslant \sqrt{\varepsilon}$, and

(iii) $J_\varepsilon(v) + \sqrt{\varepsilon} d(v, u^\varepsilon) \geqslant J_\varepsilon(u^\varepsilon)$ for all $v \in \mathcal{U}[0,T]$.

The last assertion reads

$$J_\varepsilon(u^\varepsilon) = \min_{v\in\mathcal{U}[0,T]} (J_\varepsilon(v) + \sqrt{\varepsilon} d(v, u^\varepsilon)). \tag{3.1}$$

Let us establish the necessary condition for the optimization problem (3.1). For any $v \in \mathcal{U}[0,T]$, $0 < \alpha < 1$, and $0 < \varepsilon < 1$, define

$$u^{\varepsilon,\alpha} = \alpha v + (1-\alpha) u^\varepsilon.$$

Denote by X^ε and $X^{\varepsilon,\alpha}$ the trajectories corresponding to the controls u^ε and $u^{\varepsilon,\alpha}$, respectively. By the Taylor expansion, we identify δX^ε and $\delta J(u^\varepsilon)$, which are independent of α, such that for each fixed ε,

$$X^{\varepsilon,\alpha} = X^\varepsilon + \alpha \delta X^\varepsilon + o(\alpha)$$

and

$$\left([J(u^{\varepsilon,\alpha}) - J(u^*) + \varepsilon]_+ \right)^2 = \left([J(u^\varepsilon) - J(u^*) + \varepsilon]_+ \right)^2 + 2\alpha [J(u^\varepsilon) - J(u^*) + \varepsilon]_+ \delta J(u^\varepsilon) + o(\alpha)$$

as $\alpha \to 0$. Then, by (3.1),

$$\frac{(J_\varepsilon(u^{\varepsilon,\alpha}))^2 - (J_\varepsilon(u^\varepsilon))^2}{J_\varepsilon(u^{\varepsilon,\alpha}) + J_\varepsilon(u^\varepsilon)} = J_\varepsilon(u^{\varepsilon,\alpha}) - J_\varepsilon(u^\varepsilon) \geqslant -\sqrt{\varepsilon} d(u^{\varepsilon,\alpha}, u^\varepsilon) \geqslant -C\alpha\sqrt{\varepsilon}, \tag{3.2}$$

where $C = d(v, u^*) + 1$ and the last inequality is due to

$$\frac{1}{\alpha} d(u^{\varepsilon,\alpha}, u^\varepsilon) = d(v, u^\varepsilon) \leqslant d(v, u^*) + d(u^*, u^\varepsilon) \leqslant d(v, u^*) + \sqrt{\varepsilon} \leqslant d(v, u^*) + 1 = C.$$

Therefore,

$$\frac{(J_\varepsilon(u^{\varepsilon,\alpha}))^2 - (J_\varepsilon(u^\varepsilon))^2}{\alpha\,(J_\varepsilon(u^{\varepsilon,\alpha}) + J_\varepsilon(u^\varepsilon))} \geqslant -C\sqrt{\varepsilon}.$$

Since $(C([0,T];\mathbb{R}), \|\cdot\|_\infty)$ is a separable Banach space, we know from [13] that there exists an equivalent norm, denoted by $\|\cdot\|_0$, such that the dual of $(C([0,T];\mathbb{R}), \|\cdot\|_0)$ is strictly convex. Any element μ of $(C([0,T];\mathbb{R}), \|\cdot\|_0)^*$ can still be identified with a Radon measure on $[0,T]$. Since $(C([0,T];\mathbb{R}), \|\cdot\|_0)^*$ is strictly convex, $\partial d_{\mathbb{K}}(X)$ is a singleton for any $X \notin \mathbb{K}$. Furthermore, $d_{\mathbb{K}}$ is Gâteaux differentiable at any $X \notin \mathbb{K}$. As C does not depend on α, letting $\alpha \downarrow 0$ in the last inequality, we obtain

$$\frac{[J(u^\varepsilon) - J(u^*) + \varepsilon]_+ \delta J(u^\varepsilon) + d_{\mathbb{K}}(\mathbb{E}[f(\cdot, X^\varepsilon_\cdot)]) \int_0^T \mathbb{E}[f_x(t, X_t^\varepsilon)\delta X_t^\varepsilon]\, K^\varepsilon(dt)}{J_\varepsilon(u^\varepsilon)} \geqslant -C\sqrt{\varepsilon}, \tag{3.3}$$

where $K^\varepsilon \in \partial d_{\mathbb{K}}(\mathbb{E}[f(\cdot, X^\varepsilon_\cdot)]) \subseteq \mathcal{M}^-([0,T])$. Define

$$\lambda^\varepsilon := \frac{[J(u^\varepsilon) - J(u^*) + \varepsilon]_+}{J_\varepsilon(u^\varepsilon)} \geqslant 0, \quad \mu_t^\varepsilon := -\frac{d_{\mathbb{K}}(\mathbb{E}[f(\cdot, X^\varepsilon_\cdot)])K^\varepsilon([0,t])}{J_\varepsilon(u^\varepsilon)}.$$

By Lemma 3.2, $|K^\varepsilon([0,T])| = 1$ if $\mathbb{E}[f(\cdot, X^\varepsilon_\cdot)] \notin \mathbb{K}$, and $\mu^\varepsilon \equiv 0$, otherwise. Therefore, we have

$$|\lambda^\varepsilon|^2 + |\mu_T^\varepsilon|^2 = 1.$$

Thus, there is a subsequence $\varepsilon_n \downarrow 0$ such that

$$\lambda^{\varepsilon_n} \to \lambda \geqslant 0 \quad \text{and} \quad \mu^{\varepsilon_n} \to \mu, \quad \star\text{-weakly in } C^*([0,T];\mathbb{R}).$$

Since $\mathbb{K}$ is obviously of finite-dimensional co-dimension in $C([0,T];\mathbb{R})$, in view of Lemma 3.2, we have

$$\lambda^\varepsilon \cdot 0 + \langle \mu^\varepsilon, f\rangle \geqslant 0, \quad \forall f \in \mathbb{K}.$$

By [13, Lemma 3.6, p. 142], we have that $(\lambda, \mu) \neq (0,0)$.

Set

$$\begin{aligned}
b_x^\varepsilon(s) &:= b_x(s, X_s^\varepsilon, u_s^\varepsilon), & b_v^\varepsilon(s) &:= b_v(s, X_s^\varepsilon, u_s^\varepsilon),\\
b_x^*(s) &:= b_x(s, X_s^*, u_s^*), & b_v^*(s) &:= b_v(s, X_s^*, u_s^*),\\
\sigma_x^\varepsilon(s) &:= \sigma_x(s, X_s^\varepsilon, u_s^\varepsilon), & \sigma_v^\varepsilon(s) &:= \sigma_v(s, X_s^\varepsilon, u_s^\varepsilon),\\
\sigma_x^*(s) &:= \sigma_x(s, X_s^*, u_s^*), & \sigma_v^*(s) &:= \sigma_v(s, X_s^*, u_s^*),\\
\ell_x^*(s) &:= \ell_x(s, X_s^*, u_s^*), & \ell_v^*(s) &:= \ell_v(s, X_s^*, u_s^*),\\
\delta u_s^\varepsilon &:= v_s - u_s^\varepsilon, & \delta u_s^* &:= v_s - u_s^*.
\end{aligned}$$

Then

$$\delta X_t^\varepsilon = \int_0^t (b_x^\varepsilon(s)\delta X_s^\varepsilon + b_v^\varepsilon(s)\delta u_s^\varepsilon)\,ds + \int_0^t (\sigma_x^\varepsilon(s)\delta X_s^\varepsilon + \sigma_v^\varepsilon(s)\delta u_s^\varepsilon)^\intercal\,dW_s,$$

$$\delta J(u^\varepsilon) = \mathbb{E}\left[\int_0^T (\ell_x^\varepsilon(s)\delta X_s^\varepsilon + \ell_v^\varepsilon(s)\delta u_s^\varepsilon)\,ds\right] + \mathbb{E}[g_x(X_T^\varepsilon)\delta X_T^\varepsilon],$$

$$\delta X_t^* = \int_0^t (b_x^*(s)\delta X_s^* + b_v^*(s)\delta u_s^*)\,ds + \int_0^t (\sigma_x^*(s)\delta X_s^* + \sigma_v^*(s)\delta u_s^*)^\intercal\,dW_s,$$

and

$$\delta J(u^*) = \mathbb{E}\left[\int_0^T (\ell_x^*(s)\delta X_s^* + \ell_v^*(s)\delta u_s^*)ds\right] + \mathbb{E}[g_x(X_T^*)\delta X_T^*].$$

By (3.3),

$$\lambda^\varepsilon \delta J(u^\varepsilon) + \int_0^T \mathbb{E}[f_x(t, X_t^\varepsilon)\delta X_t^\varepsilon]\,d\mu_t^\varepsilon \geqslant -C\sqrt{\varepsilon}.$$

As C does not depend on ε, letting $\varepsilon \downarrow 0$,

$$\lambda \delta J(u^*) + \int_0^T \mathbb{E}[f_x(t, X_t^*)\delta X_t^*]\,d\mu_t \geqslant 0. \tag{3.4}$$

Denote by (Y, Z) the unique solution of the BSDE

$$\begin{aligned} Y_t = \ & \lambda g_x(X_T^*) + \int_t^T (b_x^*(s)Y_s + \sigma_x^*(s)^\intercal Z_s + \lambda \ell_x^*(s))\,ds \\ & + \int_t^T f_x(s, X_s^*)d\mu_s - \int_t^T Z_s^\intercal\,dW_s. \end{aligned} \tag{3.5}$$

We have the following stochastic maximum principle.

Theorem 3.4. *Let $u^* \in \mathcal{U}_{ad}[0,T]$ be an optimal control for problem (2.3). Then, there is $(\lambda, \mu) \in [0,1] \times \mathcal{M}^+([0,T])$ such that (i) $(\lambda, \mu) \neq (0,0)$ and (ii) the following maximum condition is satisfied:*

$$\min_{v \in U} \{\langle Y_t, b_v^*(t)(v - u_t^*)\rangle + \langle Z_t, \sigma_v(t)(v - u_t^*)\rangle + \lambda \ell_v^*(t)(v - u_t^*)\} = 0, \quad \text{a.e. } t \in [0,T],$$

where the pair (Y, Z) is the unique solution of BSDE (3.5).

Proof. By (3.4), we have

$$\begin{aligned} 0 &\leqslant \mathbb{E}\left[\int_0^T \lambda(\ell_x^*(s)\delta X_s^* + \ell_v^*(s)\delta u_s^*)ds + \lambda g_x(X_T^*)\delta X_T^* + \int_0^T f_x(t, X_t^*)\delta X_t^*\,d\mu_t\right] \\ &= \mathbb{E}\left[\int_0^T \lambda \ell_v^*(s)\delta u_s^* ds - \int_0^T \delta X_s^* dY_s + Y_T \delta X_T^* - \int_0^T (b_x^*(s)Y_s + \sigma_x^*(s)^\intercal Z_s)\delta X_s^* ds\right] \\ &= \mathbb{E}\bigg[\int_0^T \lambda \ell_v^*(s)\delta u_s^* ds + \int_0^T Y_s(b_x^*(s)\delta X_s^* + b_v^*(s)\delta u_s^*)\,ds \\ &\qquad + \int_0^T Z_s^\intercal(\sigma_x^*(s)\delta X_s^* + \sigma_v^*(s)\delta u_s^*)\,ds - \int_0^T (b_x^*(s)Y_s + \sigma_x^*(s)^\intercal Z_s)\delta X_s^* ds\bigg] \\ &= \mathbb{E}\bigg[\int_0^T \lambda \ell_v^*(s)\delta u_s^* + Y_s b_v^*(s)\delta u_s^* + Z_s^\intercal \sigma_v^*(s)\delta u_s^* ds\bigg]. \end{aligned}$$

This implies the desired result. □

4 LQ stochastic control problem with expected path constraints

We now study an LQ stochastic control problem with an expected path constraint. The dynamic of the state process is governed by the SDE

$$dX_t = (A_t X_t + B_t^\intercal u_t)dt + (C_t X_t + D_t u_t)^\intercal dW_t. \tag{4.1}$$

Here the state process X is one-dimensional and the control $u \in \mathcal{U}[0,T]$ is l-dimensional. The coefficient matrices A, B, C, D are essentially bounded adapted processes of proper sizes.

Let $\mathcal{U}_{ad}[0,T]$ be the set of all controls $u \in \mathcal{U}[0,T]$ such that the pair (u, X) solves equation (4.1) with the initial value $X(0) = x$, and satisfies the expected path constraint

$$\mathbb{E}[X_t] \geqslant L_t, \quad \forall\, t \in [0,T]. \tag{4.2}$$

Here, L is a given deterministic continuous function. Introduce the constrained problem

$$\min_{u\in\mathcal{U}_{ad}[0,T]} J(u) := \frac{1}{2}\mathbb{E}\left[\int_0^T (Q_t X_t^2 + u_t^\intercal R_t u_t)\, dt + G X_T^2\right], \tag{4.3}$$

and denote by $V(x)$ its optimal value.

Assumption 4.1. *We have $Q \geqslant 0, G \geqslant 0,$ and $R \geqslant \delta I_l$ uniformly in (t,ω) for some $\delta > 0$.*

To guarantee that the admissible set $\mathcal{U}_{ad}[0,T]$ is not empty, we put forth the following assumption.

Assumption 4.2. *There exist a control $u^a \in \mathcal{U}[0,T]$ and a constant $\varepsilon > 0$ such that (u^a, X^a), which solves equation (4.1) with $X_0^a = x$, satisfies $\mathbb{E}[X_t^a] > L_t + \varepsilon$ for all $t \in [0,T]$.*

The last assumption holds true if $L_t < -\varepsilon$ for all $t \in [0,T]$ and $x \geqslant 0$. In fact, it suffices to choose $u^a = 0$.

Remark 4.1. *Suppose there is $u^a \in \mathcal{U}[0,T]$ such that (u^a, X^a) solves equation (4.1) with $X_0^a < x$ and satisfies $\mathbb{E}[X_t^a] \geqslant L_t$ for all $t \in [0,T]$. Then, Assumption 4.2 holds. In fact, suppose (u^a, X) solves equation (4.1) with $X_0 = x$. Then, by the strict monotonicity of the SDE (4.1) with respect to the initial value, we have $\mathbb{E}[X_t - X_t^a] > \varepsilon$ for some $\varepsilon > 0$. So $\mathbb{E}[X_t] > L_t + \varepsilon$ for all $t \in [0,T]$.*

We assume that Assumptions 4.1 and 4.2 hold in the rest of the paper.

4.1 Existence, uniqueness, and approximation of the optimal control

Lemma 4.2. *Let Assumptions 4.1 and 4.2 be satisfied. Then, Problem* (4.3) *has a unique optimal control.*

Proof. We first show that problem (4.3) has an optimal solution. In fact, from Assumption 4.2, we see that there is a minimizing sequence $\{v^n, n = 1, 2, \cdots\}$ in the set $\mathcal{U}_{ad}[0, T]$. It suffices to prove that $\{v^n, n = 1, 2, \cdots\}$ is a Cauchy sequence in the Banach space $L^2_{\mathcal{F}}([0,\,T];\,\mathbb{R}^l)$, since its limit still lies in $\mathcal{U}_{ad}[0, T]$. We have

$$\lim_{n\to\infty} J(v^n) = V(x), \quad v^{n,k} := \frac{1}{2}(v^n + v^k) \in \mathcal{U}_{ad}[0, T], \quad X^{n,k} = \frac{1}{2}(X^n + X^k),$$

where X^n and $X^{n,k}$ are the state processes under the admissible controls v^n and $v^{n,k}$, respectively. Therefore, $J(v^{n,k}) \geqslant V(x)$, and the parallelogram rule holds:

$$\begin{aligned}&\frac{1}{4}\mathbb{E}\left[\int_0^T [Q_t(X_t^n - X_t^k)^2 + (v_t^n - v_t^k)^\intercal R_t(v_t^n - v_t^k)]\, dt + G(X_T^n - X_T^k)^2\right]\\ &= J(v^n) + J(v^k) - 2J(v^{n,k}) \leqslant J(v^n) + J(v^k) - 2V(x).\end{aligned} \tag{4.4}$$

Hence, we have

$$\frac{1}{4}\delta\|v^n - v^k\|^2 \leqslant J(v^n) + J(v^k) - 2V(x) \to 0, \qquad \text{as } n, k \to \infty,$$

and then $\{v^n, n = 1, 2, \cdots\}$ is a Cauchy sequence in the Banach space $L^2_{\mathcal{F}}([0,\,T];\,\mathbb{R}^l)$.

Uniqueness of optimal control can be proved in a similar way via the parallelogram rule. □

On the other hand, consider the following unconstrained problem for each $n > 0$:

$$\min_{u\in\mathcal{U}[0,T]} J_n(u) := \frac{1}{2}\mathbb{E}\left[\int_0^T Q_tX_t^2 + u_t^\intercal R_t u_t\, dt + GX_T^2\right] + \frac{1}{2}n\int_0^T [(\mathbb{E}[X_t] - L_t)_-]^2\, dt, \tag{4.5}$$

where the state process X solves equation (4.1). Problem (4.5) is a stochastic linear-convex optimal control problem, which admits a unique solution (see a similar proof of Yong and Zhou [14, Theorem 5.2, page 68]).

Let $V_n(x)$ be the optimal value function of (4.5). Then, for any control $u \in \mathcal{U}_{ad}[0, T]$, we have

$$J_n(u) = J(u), \tag{4.6}$$

which leads to

$$V_n(x) \leqslant V(x). \tag{4.7}$$

Lemma 4.3. *Let $(\bar{u}^n, \overline{X}^n)$ be the optimal pair of the unconstrained problem* (4.5)*. Then, $\bar{u}^n$ converges strongly to the optimal control of the constrained problem* (4.3)*.*

Proof. We have

$$J_n(\bar u^n) = \frac{1}{2}\mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^n)^2 + (\bar u_t^n)^\intercal R_t \bar u_t^n\, dt + G(\overline{X}_T^n)^2\right] \\ + \frac{1}{2}n\int_0^T [(\mathbb{E}[\overline{X}_t^n] - L_t)_-]^2 dt \leqslant J_n(u^a) = J(u^a), \tag{4.8}$$

where u^a is given in Assumption 4.2. As $R \geqslant \delta I_l$, it follows that the sequence $\bar u^n$ is bounded in $L^2_{\mathcal{F}}(0, T; \mathbb{R}^l)$. Consequently, it has a subsequence (still denoted by $\bar u^n$) which weakly converges to some control $\bar u^\infty \in L^2_{\mathcal{F}}(0, T; \mathbb{R}^l)$.

By Mazur's theorem, there exist real numbers $\epsilon_{n,k} \geqslant 0$ such that $\sum_{k\geqslant 0}\epsilon_{n,k} = 1$ for every n, and the sequence

$$\bar v^n = \sum_{k\geqslant 0}\epsilon_{n,k}\bar u^{n+k}$$

strongly converges to $\bar u^\infty$. Let $X^{\bar v^n}$ and $\overline{X}^\infty$ denote, respectively, the trajectories under the controls $\bar v^n$ and $\bar u^\infty$. Then the sequence $X^{\bar v^n}$ converges to $\overline{X}^\infty$ strongly in $C_{\mathcal{F}}([0,T];\mathbb{R})$. In particular,

$$\overline{X}_t^\infty = \lim_n X_t^{\bar v^n} = \lim_n \sum_{k\geqslant 0}\epsilon_{n,k}\overline{X}_t^{n+k}, \quad t \in [0, T].$$

Dividing both sides of (4.8) by $\frac{n}{2}$ and letting n go to ∞, we deduce from the convexity of the map $x \mapsto (x_-)^2$ and Fatou's lemma that

$$\int_0^T [(\mathbb{E}[\overline{X}_t^\infty] - L_t)_-]^2 dt \leqslant \liminf_{n\to\infty}\sum_{k\geqslant 0}\epsilon_{n,k}\int_0^T [(\mathbb{E}[\overline{X}_t^{n+k}] - L_t)_-]^2 dt \\ \leqslant \liminf_{n\to\infty}\sum_{k\geqslant 0}\epsilon_{n,k}\frac{2}{n}J(u^a) = \liminf_{n\to\infty}\frac{2}{n}J(u^a) = 0.$$

Since $\mathbb{E}[\overline{X}_\cdot^\infty]$ and $L_\cdot$ are continuous, we conclude $\mathbb{E}[\overline{X}_t^\infty] \geqslant L_t$ holds for all $t \in [0, T]$. This means $(\bar u^\infty, \overline{X}^\infty)$ is an admissible pair for the constrained problem (4.3), so

$$J(\bar u^\infty) \geqslant V(x). \tag{4.9}$$

The convexity of the map $x \mapsto (x_-)^2$ and Fatou's lemma also give

$$\limsup_{n\to\infty}\mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^n)^2\, dt + G(\overline{X}_T^n)^2\right] \\ \geqslant \liminf_{n\to\infty}\sum_{k\geqslant 0}\epsilon_{n,k}\mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^{n+k})^2\, dt + G(\overline{X}_T^{n+k})^2\right] \\ \geqslant \mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^\infty)^2\, dt + G(\overline{X}_T^\infty)^2\right]. \tag{4.10}$$

Thanks to the weak convergence of $\bar u^n$ to $\bar u^\infty$,

$$\liminf_{n\to\infty}\mathbb{E}\left[\int_0^T (\bar u_t^n)^\intercal R_t \bar u_t^n\, dt\right] \geqslant \mathbb{E}\left[\int_0^T Q_t(\bar u_t^\infty)^\intercal R_t \bar u_t^\infty\, dt\right]. \tag{4.11}$$

The above estimates yield

$$\begin{aligned}\lim_{n\to\infty} J_n(\bar{u}^n) &\geqslant \limsup_{n\to\infty} \mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^n)^2 + (\bar{u}_t^n)^\intercal R_t \bar{u}_t^n\, dt + G(\overline{X}_T^n)^2\right]\\ &\geqslant \mathbb{E}\left[\int_0^T Q_t(\overline{X}_t^\infty)^2 + (\bar{u}_t^\infty)^\intercal R_t \bar{u}_t^\infty\, dt + G(\overline{X}_T^\infty)^2\right]\\ &= J(\bar{u}^\infty) \geqslant V(x). \end{aligned}\tag{4.12}$$

But (4.7) gives

$$V(x) \geqslant V_n(x) = J_n(\bar{u}^n),$$

so all the inequalities in (4.9)–(4.12) are equations. In particular, $(\bar{u}^\infty, \overline{X}^\infty)$ is the optimal pair of the constrained problem (4.3) since (4.9) is an equation. By the weak convergence and norm convergence (4.11), we conclude $R^{\frac{1}{2}}\bar{u}^n$ strongly converges to $R^{\frac{1}{2}}\bar{u}^\infty$ in the space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$. As $R \geqslant \delta I_l$, $\bar{u}^n$ strongly converges to $\bar{u}^\infty$ in the space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$. Consequently, $\overline{X}^n$ strongly converges to $\overline{X}^\infty$ in the space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$. As a byproduct of (4.12), we have

$$\lim_{n\to\infty} n\int_0^T [(\mathbb{E}[\overline{X}_t^n] - L_t)_-]^2 dt = 0. \tag{4.13}$$

Finally, we note that as the optimal control is unique, the whole sequence $\bar{u}^n$ strongly converges to $\bar{u}^\infty$ in the space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$. □

4.2 Verification theorem

In this section, we express the unique optimal control for problem (4.3) with the solution of a reflected FBSDEs.

We say that $(X,Y,Z,\mu) \in L^2_{\mathcal{F}}([0,\ T];\ \mathbb{R}^3)\times\mathcal{M}^+([0,T])$ is a solution of the following reflected FBSDEs

$$\begin{cases} dX_t = (A_tX_t - B_t^\intercal R_t^{-1}(B_tY_t + D_t^\intercal Z_t))\, dt\\ \qquad\qquad +(C_tX_t - D_tR_t^{-1}(B_tY_t + D_t^\intercal Z_t))^\intercal\, dW_t,\\ dY_t = -(Q_tX_t + A_tY_t + C_t^\intercal Z_t)\, dt + d\mu_t + Z_t^\intercal\, dW_t,\\ \mathbb{E}[X_t] \geqslant L_t, \quad \int_0^T (\mathbb{E}[X_t] - L_t)\, d\mu_t = 0,\\ X_0 = x, \quad Y_T = GX_T, \quad \mu_T = 0, \end{cases}\tag{4.14}$$

if it satisfies the above FBSDEs.

Theorem 4.4. *Suppose that $(\overline{X}, \overline{Y}, \overline{Z}, \overline{\mu})$ is a solution of the reflected FBSDEs (4.14). Then*

$$\overline{u} := -R^{-1}(B\overline{Y} + D^\intercal\overline{Z})$$

is an optimal control for problem (4.3), and the optimal value is

$$J(\overline{u}) = \frac{1}{2}\overline{Y}_0 x + \frac{1}{2}\int_0^T L_t\, d\overline{\mu}_t. \tag{4.15}$$

Proof. Note that $(\overline{X},\overline{u})$ solves equation (4.1) with the initial value x. So, for any $u\in\mathcal{U}_{ad}[0,T]$,

$$
\begin{aligned}
&J(u)-J(\overline{u})\\
&=\frac{1}{2}\mathbb{E}[GX_T^2-G\overline{X}_T^2]+\frac{1}{2}\mathbb{E}\int_0^T[Q_tX_t^2-Q_t\overline{X}_t^2]dt+\frac{1}{2}\mathbb{E}\int_0^T[u_t^\intercal R_tu_t-\overline{u}_t^\intercal R_t\overline{u}_t]dt\\
&=\mathbb{E}[G\overline{X}_T(X_T-\overline{X}_T)]+\mathbb{E}\int_0^T Q_t\overline{X}_t(X_t-\overline{X}_t)\,dt+\mathbb{E}\int_0^T\overline{u}_t^\intercal R_t(u_t-\overline{u}_t)\,dt\\
&\quad+\frac{1}{2}\mathbb{E}[G(X_T-\overline{X}_T)^2]+\frac{1}{2}\mathbb{E}\int_0^T Q_t(X_t-\overline{X}_t)^2\,dt+\frac{1}{2}\mathbb{E}\int_0^T(u_t-\overline{u}_t)^\intercal R_t(u_t-\overline{u}_t)\,dt.\\
&\geqslant\mathbb{E}[G\overline{X}_T(X_T-\overline{X}_T)]+\mathbb{E}\int_0^T Q_t\overline{X}_t(X_t-\overline{X}_t)\,dt+\mathbb{E}\int_0^T\overline{u}_t^\intercal R_t(u_t-\overline{u}_t)\,dt. \qquad (4.16)
\end{aligned}
$$

Applying Itô's formula, we have

$$
\begin{aligned}
d(\overline{Y}_t(X_t-\overline{X}_t))=&-Q_t\overline{X}_t(X_t-\overline{X}_t)\,dt+(X_t-\overline{X}_t)\,d\overline{\mu}_t+\overline{Y}_tB_t^\intercal(u_t-\overline{u}_t)\,dt\\
&+\overline{Z}_t^\intercal[D(u_t-\overline{u}_t)]\,dt+(X_t-\overline{X}_t)\overline{Z}_t^\intercal dW_t\\
&+\overline{Y}_t[C_t(X_t-\overline{X}_t)+D(u_t-\overline{u}_t)]^\intercal dW_t.
\end{aligned}
$$

Integrating both sides and taking the expectation (also noting that the local martingale is in fact a martingale (see Bismut [1, Proposition I-1, p. 387])), we have

$$
\begin{aligned}
&\mathbb{E}[G\overline{X}_T(X_T-\overline{X}_T)]+\mathbb{E}\int_0^T[Q_t\overline{X}_t(X_t-\overline{X}_t)]\,dt\\
&=\int_0^T(\mathbb{E}[X_t]-\mathbb{E}[\overline{X}_t])\,d\overline{\mu}_t+E\left[\int_0^T\langle B_t\overline{Y}_t+D_t^\intercal\overline{Z}_t,u_t-\overline{u}_t\rangle dt\right].
\end{aligned}
$$

Thanks to $\overline{u}=-R^{-1}(B\overline{Y}+D^\intercal\overline{Z})$ and (4.16),

$$
\begin{aligned}
&J(u)-J(\overline{u})\\
&\geqslant\mathbb{E}[G\overline{X}_T(X_T-\overline{X}_T)]+\mathbb{E}\int_0^T[Q_t\overline{X}_t(X_t-\overline{X}_t)]\,dt+\mathbb{E}\left[\int_0^T\langle R_t\overline{u}_t,u_t-\overline{u}_t\rangle dt\right]\\
&=\int_0^T(\mathbb{E}[X_t]-L_t)\,d\overline{\mu}_t-\int_0^T(\mathbb{E}[\overline{X}_t]-L_t)\,d\overline{\mu}_t=\int_0^T(\mathbb{E}[X_t]-L_t)\,d\overline{\mu}_t\geqslant 0,
\end{aligned}
$$

where the last inequality is due to constraint (4.2) and $\overline{\mu}\in\mathcal{M}^+([0,T])$.

Again, using Itô's formula, we have

$$
\begin{aligned}
d(\overline{Y}_t\overline{X}_t)=&-Q_t\overline{X}_t^2\,dt+\overline{X}_t\,d\overline{\mu}_t+\overline{Y}_tB_t^\intercal\overline{u}_t\,dt+\overline{Z}_t^\intercal D_t\overline{u}_t\,dt\\
&+\overline{X}_t\overline{Z}_t^\intercal dW_t+\overline{Y}_t(C_t\overline{X}_t+D\overline{u}_t)^\intercal dW_t.
\end{aligned}
$$

Note that the local martingale is in fact a martingale (see Bismut [1, Proposition I-1, p. 387]). Therefore, integrating both sides yields

$$
\begin{aligned}
\mathbb{E}[G\overline{X}_T^2]-\overline{Y}_0x&=-\mathbb{E}\left[\int_0^T Q_t\overline{X}_t^2\,dt\right]+\int_0^T\mathbb{E}[\overline{X}_t]\,d\overline{\mu}_t+\mathbb{E}\left[\int_0^T\langle B_t^\intercal\overline{Y}_t+D_t^\intercal\overline{Z}_t,\overline{u}_t\rangle dt\right]\\
&=-\mathbb{E}\left[\int_0^T Q_t\overline{X}_t^2\,dt\right]+\int_0^T L_t\,d\overline{\mu}_t-\mathbb{E}\left[\int_0^T\langle R_t\overline{u}_t,\overline{u}_t\rangle dt\right].
\end{aligned}
$$

Thus, we proved the desired expression for the optimal value $J(\overline{u})$. □

In the rest of the paper, we focus on solution of the reflected FBSDEs (4.14). The main result is stated as follows.

Theorem 4.5. *If A is deterministic, $B^\intercal B$ is invertible, and $(B^\intercal B)^{-1}$ is bounded, then the reflected FBSDEs (4.14) admit a unique solution.*

Proof. This is an immediate consequence of Propositions 5.1 and 6.1 in the subsequent sections. □

We prove the uniqueness and existence in the following two sections, respectively.

5 Uniqueness of the solution for the reflected FBSDEs (4.14)

Proposition 5.1. *Let (X, Y, Z, μ) and $(\widehat{X}, \widehat{Y}, \widehat{Z}, \widehat{\mu})$ be two solutions for the reflected FBSDEs (4.14). Then $X = \widehat{X}$ and $BY + D^\intercal Z = B\widehat{Y} + D^\intercal \widehat{Z}$. Furthermore, $(X, Y, Z, \mu) = (\widehat{X}, \widehat{Y}, \widehat{Z}, \widehat{\mu})$ if A is deterministic and $\mathbb{E}[B^\intercal B] > 0$.*

Proof. We denote by $(\widetilde{X}, \widetilde{Y}, \widetilde{Z}, \widetilde{\mu})$ the difference of (X, Y, Z, μ) and $(\widehat{X}, \widehat{Y}, \widehat{Z}, \widehat{\mu})$. Set

$$\overline{u} := -R^{-1}(BY + D^\intercal Z), \quad \widehat{u} := -R^{-1}(B\widehat{Y} + D^\intercal \widehat{Z}), \quad \widetilde{u} := \overline{u} - \widehat{u}.$$

We now show the first assertion, that is, $\widetilde{X} = 0$ and $\widetilde{u} = 0$. By (4.14),

$$\begin{aligned} d\widetilde{X} &= (A\widetilde{X} + B^\intercal \widetilde{u})\, dt + (C\widetilde{X} + D\widetilde{u})^\intercal dW_t, \\ d\widetilde{Y} &= -(Q\widetilde{X} + A\widetilde{Y} + C^\intercal \widetilde{Z})\, dt + d\widetilde{\mu}_t + \widetilde{Z}_t^\intercal dW_t. \end{aligned} \tag{5.1}$$

Using Itô's formula, we have

$$\begin{aligned} d(\widetilde{X}_t\widetilde{Y}_t) = {} & B^\intercal \widetilde{u}\widetilde{Y}\, dt + (C\widetilde{X} + D\widetilde{u})^\intercal \widetilde{Y}\, dW_t \\ & - \widetilde{X}(Q\widetilde{X}\, dt - d\widetilde{\mu}_t - \widetilde{Z}^\intercal dW_t) + \widetilde{Z}_t^\intercal D\widetilde{u} dt. \end{aligned}$$

Integrating both sides and taking the expectation, since the local martingale is in fact a martingale (see Bismut [1, Proposition I-1, p. 387]), we have the duality formula

$$\begin{aligned} & \mathbb{E}[G\widetilde{X}_T^2] + \mathbb{E}\int_0^T Q\widetilde{X}_t^2 dt \\ = {} & \mathbb{E}\int_0^T \widetilde{X} d\widetilde{\mu}_t + \mathbb{E}\int_0^T \langle B\widetilde{Y} + D^\intercal \widetilde{Z}, \widetilde{u}\rangle dt \\ = {} & \int_0^T (\mathbb{E}[X_t] - L_t) d\widetilde{\mu}_t - \int_0^T (\mathbb{E}[\widehat{X}_t] - L_t) d\widetilde{\mu}_t - \mathbb{E}\int_0^T \langle R\widetilde{u}_t, \widetilde{u}_t\rangle dt. \end{aligned}$$

Hence,

$$\begin{aligned} & \mathbb{E}[G\widetilde{X}_T^2] + \mathbb{E}\int_0^T Q\widetilde{X}_t^2 dt + \mathbb{E}\int_0^T \langle R\widetilde{u}, \widetilde{u}\rangle dt \\ = {} & -\int_0^T (\mathbb{E}[X_t] - L_t) d\widehat{\mu}_t - \int_0^T (\mathbb{E}[\widehat{X}_t] - L_t) d\mu_t \leqslant 0. \end{aligned}$$

Because $G \geqslant 0$, $Q \geqslant 0$, $R \geqslant \delta I_l$, it follows that $\widetilde{u} = 0$. Consequently, (5.1) reduces to $d\widetilde{X} = A\widetilde{X}\,dt + (C\widetilde{X})^\intercal dW_t$. Together with $\widetilde{X}_0 = 0$, we infer that $\widetilde{X} = 0$. This completes the proof of the first assertion.

Now suppose A is deterministic. Let $U_t = e^{\int_0^t A_r dr} > 0$. Then U is deterministic and $dU = AUdt$. By (5.1),

$$d(U\widetilde{Y}) = -UC^\intercal \widetilde{Z}\,dt + Ud\widetilde{\mu}_t + U\widetilde{Z}^\intercal dW_t = U\widetilde{Z}^\intercal d\widetilde{W}_t + Ud\widetilde{\mu}_t,$$

where $\widetilde{W}_t = W_t - \int_0^t C_s ds$ is a Brownian motion under some probability measure $\widetilde{\mathbb{P}} \sim \mathbb{P}$. This means

$$U_t\widetilde{Y}_t - \int_0^t U_s d\widetilde{\mu}_s = U_0\widetilde{Y}_0 + \int_0^t U_s \widetilde{Z}_s^\intercal d\widetilde{W}_s$$

is a martingale under $\widetilde{\mathbb{P}}$. But the value of this martingale at $t = T$ is

$$U_T\widetilde{Y}_T - \int_0^T U_s d\widetilde{\mu}_s = U_T G\widetilde{X}_T - \int_0^T U_s d\widetilde{\mu}_s = -\int_0^T U_s d\widetilde{\mu}_s,$$

a constant, so it is a constant martingale. Hence, $\widetilde{Z} = 0$ and, consequently,

$$\widetilde{Y}_t = U_t^{-1}\left(U_0\widetilde{Y}_0 + \int_0^t U_s d\widetilde{\mu}_s\right) \tag{5.2}$$

is a deterministic function. From $\widetilde{u} = 0$ and $\widetilde{Z} = 0$, we get $B\widetilde{Y} = 0$. Thus

$$0 = \mathbb{E}[B^\intercal B\widetilde{Y}] = \mathbb{E}[B^\intercal B]\widetilde{Y}.$$

If $\mathbb{E}[B^\intercal B] > 0$, then $\widetilde{Y} = 0$ and consequently by (5.2), $\widetilde{\mu} = 0$. The second assertion is thus proved. □

Remark 5.2. *If the last condition in the above theorem does not hold, then the uniqueness can fail. For instance, when A is deterministic and $B = 0$, we may get infinitely many solutions $(Y + (k-1)\widetilde{Y},\ k\mu)$ from a solution (Y, μ) by setting $k > 0$ and*

$$\widetilde{Y}_t = -\int_t^T e^{\int_t^s A_r dr} d\mu_s.$$

6 Existence of the solution for the reflected FBSDEs (4.14)

Proposition 6.1. *If $B^\intercal B$ is invertible and $(B^\intercal B)^{-1}$ is bounded. Then the reflected FBSDEs (4.14) has a solution.*

We use the penalization method to prove the existence. The proof is given in the subsequent two subsections.

6.1 Approximation

For any $n \in \mathbb{N}$, consider the following penalized FBSDEs:

$$\begin{cases} dX^n = (AX^n - B^\intercal R^{-1}(BY^n + D^\intercal Z^n))\,dt \\ \qquad\qquad +(CX^n - DR^{-1}(BY^n + D^\intercal Z^n))^\intercal dW_t, \\ dY^n = -(QX^n + AY^n + C^\intercal Z^n)\,dt + n(\mathbb{E}[X^n] - L)_- dt + (Z^n)^\intercal dW_t, \\ X_0^n = x, \quad Y_T^n = GX_T^n; \end{cases} \tag{6.1}$$

It is a Mckean–Vlasov FBSDEs, and is actually the Hamiltonian system of the optimal control of Problem (4.5). We call $(X^n, Y^n, Z^n) \in L^2_{\mathcal{F}}([0,T];\mathbb{R}^3)$ a solution to the FBSDEs (6.1) if it satisfies (6.1).

Lemma 6.2. *The penalized FBSDEs (6.1) admit a unique solution.*

Proof. Since the optimal control problem (4.5) has an optimal control $\bar{u}$, we have the existence of FBSDEs (6.1) immediately from the stochastic maximum principle for optimally controlled Mckean–Vlasov SDEs.

We now turn to the proof of uniqueness. We suppress the superscript n here for simplicity. We denote by $(\widetilde{X}, \widetilde{Y}, \widetilde{Z})$ the difference of two solutions (X, Y, Z) and $(\widehat{X}, \widehat{Y}, \widehat{Z})$ to (6.1). Set

$$\overline{u} := -R^{-1}(BY + D^\intercal Z), \quad \widehat{u} := -R^{-1}(B\widehat{Y} + D^\intercal \widehat{Z}), \quad \widetilde{u} := \overline{u} - \widehat{u}.$$

Then, by (6.1),

$$d\widetilde{X} = (A\widetilde{X} + B^\intercal \widetilde{u})\,dt + (C\widetilde{X} + D\widetilde{u})^\intercal dW_t, \tag{6.2}$$

$$d\widetilde{Y} = -[(Q\widetilde{X} + A\widetilde{Y} + C^\intercal \widetilde{Z}) - n((\mathbb{E}[X] - L)_- - (\mathbb{E}[\widehat{X}] - L)_-)]dt + \widetilde{Z}_t^\intercal dW_t. \tag{6.3}$$

Using Itô's formula, we have

$$\begin{aligned} d(\widetilde{X}_t\widetilde{Y}_t) &= B^\intercal \widetilde{u}\widetilde{Y}\,dt + (C\widetilde{X} + D\widetilde{u})^\intercal \widetilde{Y}\,dW_t \\ &\quad - \widetilde{X}(Q\widetilde{X} - n((\mathbb{E}[X] - L)_- - (\mathbb{E}[\widehat{X}] - L)_-))dt - \widetilde{Z}^\intercal dW_t) + \widetilde{Z}_t^\intercal D\widetilde{u}dt. \end{aligned}$$

Integrating both sides yields

$$\begin{aligned} &\mathbb{E}[G\widetilde{X}_T^2] + \mathbb{E}\int_0^T Q\widetilde{X}_t^2 dt \\ &= \int_0^T \mathbb{E}[\widetilde{X}](n(\mathbb{E}[X] - L)_- - n(\mathbb{E}[\widehat{X}] - L)_-)dt - \mathbb{E}\int_0^T \langle R\widetilde{u}_t, \widetilde{u}_t\rangle dt. \end{aligned}$$

Hence,

$$\mathbb{E}[G\widetilde{X}_T^2] + \mathbb{E}\int_0^T Q\widetilde{X}_t^2 dt + \mathbb{E}\int_0^T \langle R\widetilde{u}, \widetilde{u}\rangle dt \leqslant 0,$$

from which we deduce $\widetilde{u} = 0$. We have $\widetilde{X} = 0$ from $\widetilde{X}_0 = 0$ and (6.2). This, in particular, implies $\widetilde{Y}_T = G\widetilde{X}_T = 0$. Together with (6.3), we have $\widetilde{Y} = 0$ and $\widetilde{Z} = 0$.

□

Lemma 6.3. *Suppose that* (X^n, Y^n, Z^n) *is a solution of the penalized FBSDEs* (6.1). *Then*

$$u^n := -R^{-1}(BY^n + D^\intercal Z^n)$$

is the optimal control for the unconstrained problem (4.5). *And the optimal value is*

$$J_n(u^n) = \frac{1}{2}Y_0^n x + \frac{n}{2}\int_0^T (\mathbb{E}[X_t^n] - L_t)_- L_t dt. \tag{6.4}$$

Proof. The proof is similar to that of Theorem 4.4. We leave the details to the interested readers. □

6.2 Convergence

We next show that the solutions of the penalized FBSDEs (6.1) have a limit, which turns out to be a solution of the reflected FBSDEs (4.14). In the following arguments, we choose a subsequence when necessary. Also, the constant $M \in \mathbb{R}^+$ might vary from line to line, but does not depend on n, k, or t.

Let (X^n, Y^n, Z^n, u^n) be given as in Lemma 6.3. Then u^n is the optimal control for the unconstrained problem (4.5). By section 4.1, we conclude that the sequence u^n strongly converges to u^∞ in the space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$ and X^n converges to X^∞ strongly in $C_{\mathcal{F}}([0,T];\mathbb{R})$, where (u^∞, X^∞) is the optimal pair of the constrained problem (4.3). Moreover,

$$0 \leqslant J(u^n) \leqslant J_n(u^n) \leqslant J_n(u^a) = J(u^a) \leqslant M, \tag{6.5}$$

where u^a is given in Assumption 4.2.

We now show that

$$\mu_t^n := -n\int_t^T (\mathbb{E}[X_s^n] - L_s)_- ds, \quad t \in [0,T]$$

is a uniformly bounded sequence in $L^\infty([0,\,T];\,\mathbb{R})$ and Y_0^n is a uniformly bounded sequence in $\mathbb{R}$.

To this end, let $\beta \in [x, x+1]$ and $(u^a, X^{\beta,a})$ evolve according to equation (4.1) with $X_0^{\beta,a} = \beta$. By Assumption 4.2 and monotonicity of SDE, we have

$$\mathbb{E}\left[X_t^{\beta,a}\right] > L_t, \quad t \in [0,T].$$

Applying Ito's formula to $X_t^{\beta,a}Y_t^n$, we get

$$\begin{aligned}
d(X^{\beta,a}Y^n) &= Y^n(AX^{\beta,a} + B^\intercal u^a)\,dt + Y^n(CX^{\beta,a} + Du^a)^\intercal dW_t \\
&\qquad - X^{\beta,a}(QX^n + AY^n + C^\intercal Z^n)\,dt + X^{\beta,a}d\mu_t^n + X^{\beta,a}(Z^n)^\intercal dW_t, \\
&\qquad + (CX^{\beta,a} + Du^a)^\intercal Z^n dt \\
&= (-QX^{\beta,a}X^n + (u^a)^\intercal(BY^n + D^\intercal Z^n))dt + X^{\beta,a}d\mu_t^n \\
&\qquad + (Y^nCX^{\beta,a} + Y^nDu^a + X^{\beta,a}Z^n)^\intercal dW_t \\
&= (-QX^{\beta,a}X^n - (u^a)^\intercal Ru^n)dt + X^{\beta,a}d\mu_t^n \\
&\qquad + (Y^nCX^{\beta,a} + Y^nDu^a + X^{\beta,a}Z^n)^\intercal dW_t.
\end{aligned}$$

Integrating on both sides, we have

$$\beta Y_0^n = GX_T^{\beta,a}X_T^n + \int_0^T (QX^{\beta,a}X^n + (u^a)^\intercal Ru^n)\,dt - \int_0^T X^{\beta,a}d\mu_t^n$$
$$- \int_0^T (Y^n CX^{\beta,a} + Y^n Du^a + X^{\beta,a}Z^n)^\intercal dW_t,$$

and the local martingale is in fact a martingale (see Bismut [1, Proposition I-1, p. 387]). By the elementary inequality $|\langle a,b\rangle| \leqslant \frac{1}{2}|a|^2 + \frac{1}{2}|b|^2$ and (6.5), we have

$$\begin{aligned}
\left|\beta Y_0^n + \int_0^T \mathbb{E}[X_t^{\beta,a}]d\mu_t^n\right| &\leqslant \mathbb{E}[G|X_T^{\beta,a}X_T^n|] + \int_0^T \mathbb{E}[Q_t|X_t^{\beta,a}X_t^n| + (u_t^a)^\intercal R_t u_t^n]\,dt \\
&\leqslant \frac{1}{2}\mathbb{E}[G(X_T^{\beta,a})^2] + \frac{1}{2}\int_0^T \mathbb{E}[Q_t(X_t^{\beta,a})^2 + (u_t^a)^\intercal R_t u_t^a]dt \\
&\quad + \frac{1}{2}\mathbb{E}[G(X_T^n)^2] + \frac{1}{2}\int_0^T \mathbb{E}[Q_t(X_t^n)^2 + (u_t^n)^\intercal R_t u_t^n]\,dt \\
&\leqslant \frac{1}{2}(M + J(u^n)) \leqslant M,
\end{aligned} \tag{6.6}$$

where M does not depend on $\beta \in [x, x+1]$. For the case of $\beta = x$, we have

$$\left|xY_0^n + \int_0^T \mathbb{E}[X^a]d\mu_t^n\right| \leqslant M.$$

By (6.4) and (6.5),

$$0 \leqslant xY_0^n + \int_0^T L d\mu_t^n = 2J_n(u^n) \leqslant M.$$

Comparing the above two inequalities, we get

$$\left|\int_0^T (\mathbb{E}[X^a] - L)d\mu_t^n\right| \leqslant M.$$

From Assumption 4.2 and the monotonicity of μ^n, we see that μ^n is a uniformly bounded sequence in $L^\infty([0,T];\mathbb{R})$. Consequently, choosing $0 \neq \beta \in [x, x+1]$ in (6.6), we see that Y_0^n is a uniformly bounded sequence in $\mathbb{R}$.

By (4.13), we have

$$\lim_{n\to\infty} \int_0^T (\mathbb{E}[X_t^n] - L_t)_- d\mu_t^n = 0.$$

Also trivially,

$$\int_0^T (\mathbb{E}[X_t^n] - L_t)_+ d\mu_t^n = n\int_0^T (\mathbb{E}[X_t^n] - L_t)_+(\mathbb{E}[X_t^n] - L_t)_- dt = 0,$$

so

$$\lim_{n\to\infty} \int_0^T |\mathbb{E}[X_t^n] - L_t|\, d\mu_t^n = 0.$$

As μ^n is a uniformly bounded sequence, it has a $\star$-weak limit $\mu^\infty \in \mathcal{M}^+([0,T])$. Hence,

$$\begin{aligned}\int_0^T |\mathbb{E}[X_t^\infty]-L_t|\,d\mu_t^\infty &= \lim_{n\to\infty}\int_0^T |\mathbb{E}[X_t^\infty]-L_t|\,d\mu_t^n\\ \leqslant \limsup_{n\to\infty}\int_0^T |\mathbb{E}[X_t^\infty]-\mathbb{E}[X_t^n]|\,d\mu_t^n &+ \limsup_{n\to\infty}\int_0^T |\mathbb{E}[X_t^n]-L_t|\,d\mu_t^n\\ \leqslant \limsup_{n\to\infty}\max_t |\mathbb{E}[X_t^\infty]-\mathbb{E}[X_t^n]||\mu_0^n| &= 0,\end{aligned}$$

where the last equation is due to the fact that X^n converges to X^∞ strongly in $C_{\mathcal{F}}([0,T];\mathbb{R})$. Therefore,

$$\int_0^T (\mathbb{E}[X_t^\infty]-L_t)d\mu_t^\infty = 0.$$

From the strong convergence of $X^n \to X^\infty$ and equality (4.13), we have

$$\int_0^T ((\mathbb{E}[X_t^\infty]-L_t)_-)^2 dt = \liminf_{n\to\infty}\int_0^T ((\mathbb{E}[X_t^n]-L_t)_-)^2 dt = 0.$$

We conclude that $\mathbb{E}[X_t^\infty] \geqslant L_t$ for all $t\in[0,T]$ by the continuity of $\mathbb{E}[X_\cdot^\infty]$ and $L_\cdot$.

Applying the standard estimate for SDE to (6.1), we have

$$\mathbb{E}\left[\sup_{t\leqslant T}|X_t^n|^2 + \int_0^T |u_t^n|^2 dt\right] \leqslant M,$$
$$\mathbb{E}\left[\sup_{t\leqslant T}|X_t^n - X_t^k|^2 + \int_0^T |u_t^n - u_t^k|^2 dt\right] \leqslant M,$$

uniformly for n, $k\in\mathbb{N}$. We use these estimates for linear SDEs frequently in the subsequent argument without claim.

We notee that

$$\lim_{n,k\to\infty}\mathbb{E}\left[\sup_{t\leqslant T}|X_t^n - X_t^k|^2 + \int_0^T |u_t^n - u_t^k|^2 dt\right] = 0.$$

By Hölder's inequality,

$$\begin{aligned}&\limsup_{n,k\to\infty}\overline{\mathbb{E}}\left[\left(\int_0^T |u_t^n-u_t^k|^2 dt\right)^{\frac{p}{2}}\right]\\ &\qquad\leqslant \limsup_{n,k\to\infty}\left(\mathbb{E}\left[\int_0^T |u_t^n-u_t^k|^2 dt\right]\right)^{\frac{p}{2}}\left(\mathbb{E}\left[\left(\frac{d\overline{\mathbb{P}}}{d\mathbb{P}}\right)^{\frac{2}{2-p}}\right]\right)^{\frac{2-p}{2}} = 0, \quad p\in(1,2).\end{aligned}$$

Thus,

$$\lim_{n,k\to\infty}\overline{\mathbb{E}}\left[\left(\int_0^T |u_t^n-u_t^k|^2 dt\right)^{\frac{p}{2}}\right] = 0, \quad p\in(1,2). \tag{6.7}$$

Similarly, we have

$$\lim_{n,k\to\infty}\overline{\mathbb{E}}\left[\sup_{t\leqslant T}|X_t^n - X_t^k|^p\right] = 0, \quad p\in(1,2). \tag{6.8}$$

Let $U_t = e^{\int_0^t A_r dr} > 0$ and $\overline{W}_t = W_t - \int_0^t C_s\, ds$. Then U is positive and uniformly bounded. By (6.1),

$$\begin{cases} d(UY^n) = -UQX^n\, dt + U d\mu_t^n + U(Z^n)^\intercal d\overline{W}_t, \\ Y_T^n = GX_T^n; \end{cases} \tag{6.9}$$

Integrating yields

$$\int_0^T U(Z^n)^\intercal d\overline{W}_t = U_T G X_T^n - Y_0^n + \int_0^T UQX^n\, dt - \int_0^T U d\mu_t^n.$$

Because Y_0^n, μ^n and X^n are convergent and U is uniformly bounded, by (6.8), we have

$$\lim_{n,k\to\infty} \mathbb{E}\left[\left|\int_0^T U_t(Z_t^n - Z_t^k)^\intercal d\overline{W}_t\right|^p\right] = 0, \quad p\in(1,2).$$

By Doob's martingale inequality,

$$\begin{aligned}&\limsup_{n,k\to\infty} \mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t U_s(Z_s^n - Z_s^k)^\intercal d\overline{W}_s\right|^p\right]\\ &\leqslant \left(\frac{p}{p-1}\right)^p \lim_{n,k\to\infty} \mathbb{E}\left[\left|\int_0^T U_t(Z_t^n - Z_t^k)^\intercal d\overline{W}_t\right|^p\right] = 0, \quad p\in(1,2),\end{aligned}$$

which gives

$$\lim_{n,k\to\infty} \mathbb{E}\left[\sup_{t\leqslant T}\left|\int_0^t U_s(Z_s^n - Z_s^k)^\intercal d\overline{W}_s\right|^p\right] = 0, \quad p\in(1,2).$$

Applying the Burkholder–Davis–Gundy inequality, we conclude

$$\lim_{n,k\to\infty} \mathbb{E}\left[\left(\int_0^T U_s^2 |Z_s^n - Z_s^k|^2\, ds\right)^{\frac{p}{2}}\right] = 0, \quad p\in(1,2).$$

In turn, by the boundedness of U and Hölder's inequality,

$$\lim_{n,k\to\infty} \mathbb{E}\left[\left(\int_0^T |Z_s^n - Z_s^k|^2\, ds\right)^{\frac{p}{2}}\right] = 0.$$

As

$$Y^n = -(B^\intercal B)^{-1} B^\intercal (Ru^n + D^\intercal Z^n),$$

we also deduce

$$\lim_{n,k\to\infty} \mathbb{E}\left[\left(\int_0^T |Y_s^n - Y_s^k|^2\, ds\right)^{\frac{p}{2}}\right] = 0. \tag{6.10}$$

Now, we fix one $p\in(1,2)$. Then there exists a unique $(Y^\infty, Z^\infty) \in L_{\mathcal{F}}^p([0,\,T];\,\mathbb{R}) \times L_{\mathcal{F}}^p([0,\,T];\,\mathbb{R}^m)$ such that

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\int_0^T |Z_s^n - Z_s^\infty|^2\, ds\right)^{\frac{p}{2}}\right] = 0$$

and

$$\lim_{n\to\infty}\mathbb{E}\left[\left(\int_0^T |Y_s^n - Y_s^\infty|^2\,ds\right)^{\frac{p}{2}}\right] = 0.$$

Let $\widehat{Y}^n = Y^n - \mu^n$. Thanks to (6.1) and $\mu_T^n = 0$,

$$\begin{cases} dX^n = (AX^n + B^\intercal u^n)\,dt + (CX^n + Du^n)^\intercal dW_t, \\ d\widehat{Y}^n = -(QX^n + A\widehat{Y}^n + A\mu^n + C^\intercal Z^n)\,dt + (Z^n)^\intercal dW_t, \\ X_0^n = x, \quad \widehat{Y}_T^n = GX_T^n. \end{cases}$$

By the standard estimate for BSDEs and the boundedness of the sequence μ^n,

$$\mathbb{E}\left[\int_0^T |\widehat{Y}_t^n|^2\,dt + \int_0^T |Z_t^n|^2\,dt\right] \leqslant M\Big(\mathbb{E}\left[(GX_T^n)^2\right] + \int_0^T (QX^n + A\mu^n)^2\,dt\Big) \leqslant M.$$

Since μ^n is a bounded sequence, it immediately yields that

$$\mathbb{E}\left[\int_0^T |Y_t^n|^2\,dt + \int_0^T |Z_t^n|^2\,dt\right] \leqslant M.$$

Taking lower limits, it follows from Fatou's lemma that

$$\mathbb{E}\left[\int_0^T |Y_t^\infty|^2\,dt + \int_0^T |Z_t^\infty|^2\,dt\right] < \infty.$$

So we conclude that $(Y^\infty, Z^\infty) \in L_{\mathcal{F}}^2([0,T];\mathbb{R}) \times L_{\mathcal{F}}^2([0,T];\mathbb{R}^m)$.

Note that μ_0^n is bounded, we denote its limit (along a subsequence) by μ_0. Setting

$$\overline{\mu}_t := Y_t^\infty - Y_0^\infty + \int_0^t (QX_s^\infty + AY_s^\infty + C^\intercal Z_s^\infty)ds + \mu_0 - \int_0^t (Z_s^\infty)^\intercal dW_s,$$

we have

$$\lim_{n\to\infty}\mathbb{E}\left[\left(\int_0^T |\mu_t^n - \overline{\mu}_t|^2\,dt\right)^{p/2}\right] = 0, \quad p \in (1,2).$$

Let $\varphi \in L_{\mathcal{F}}^\infty([0,T])$ and $\Phi = \int_0^t \varphi(s)ds$, then

$$\begin{aligned} \mathbb{E}\int_0^T \varphi(t)\overline{\mu}_t dt &= \lim_n \mathbb{E}\left[\int_0^T \varphi(t)\mu_t^n dt\right] = -\lim_n \mathbb{E}\left[\int_0^T \Phi(t)d\mu_t^n\right] \\ &= -\mathbb{E}\left[\int_0^T \Phi(t)d\mu_t^\infty\right] = \mathbb{E}\left[\int_0^T \varphi(t)\mu_t^\infty dt\right]. \end{aligned}$$

Hence $\overline{\mu} = \mu^\infty$, and consequently $(X^\infty, Y^\infty, Z^\infty, \mu^\infty)$ is a solution to (4.14).

7 Acknowledgments

The three authors would like to thank both referees for their careful reading and helpful comments.

References

[1] J. M. Bismut, Conjugate convex functions in optimal stochastic control, J. Math. Anal. Appl. 44 (1973) 384404.

[2] J. M. Bismut. Linear quadratic optimal stochastic control with random coefficients. SIAM J. Control Optimization 14 (1976), no. 3, 419–444.

[3] J. M. Bismut. An introductory approach to duality in optimal stochastic control. SIAM Rev. 20 (1978), no. 1, 62–78.

[4] L. Bourdin, Note on Pontryagin maximum principle with running state constraints and smooth dynamics – Proof based on the Ekeland variational principle, arXiv:1604.04051v1 [math.OC], 2016.

[5] P. Briand, P. Cardaliaguet, P. E. Chaudru de Raynal and Y. Hu, Forward and backward stochastic differential equations with normal constraint in law. Stochastic Process. Appl. 130 (2020), 7021–7097.

[6] P. Briand, R. Elie and Y. Hu, BSDEs with mean reflection. Ann. Appl. Probab. 28 (2018), 482–510.

[7] R. Carmona, F. Delarue. Probabilistic theory of mean field games with applications. I. Mean field FBSDEs, control, and games. Probability Theory and Stochastic Modelling, 83. Springer, Cham, 2018.

[8] A. V. Dmitruk, N. P. Osmolovskii, Necessary conditions for a weak minimum in optimal control problems with integral equations subject to state and mixed constraints. SIAM J. Control Optimization 52 (2014), no. 6, 3437–3462.

[9] I. Ekeland, On the variational principle, *J. Math. Anal. Appl.* **47** (1974), pp. 324-353.

[10] U. G. Haussmann. A stochastic maximum principle for optimal control of diffusions. Pitman Research Notes in Mathematics Series, 151. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1986.

[11] H. J. Kushner. Necessary conditions for continuous parameter stochastic optimization problems. SIAM J. Control 10 (1972), 550565.

[12] S. Peng, A general stochastic maximum principle for optimal control problems. SIAM Journal on control and optimization, 28:966-979, 1990.

[13] X. Li and J. Yong, Optimal control theory for infinite-dimensional systems. Systems & Control: Foundations & Applications. Birkh auser Boston, Inc., Boston, MA, 1995.

[14] J. Yong and X. Zhou (1999): Stochastic Controls, Hamiltonian Systems and HJB Equations, Springer-Verlag, New York.

[15] Frankowska H., Zhang H., and Zhang X.: Necessary optimality conditions for local minimizers of stochastic optimal control problems with state constraints. Trans. Amer. Math. Soc., 2019, 372 (2):12891331.