Optimal variational principle for backward stochastic control systems associated with Lévy processes

TANG MaoNing\textsuperscript{1} \& ZHANG Qi\textsuperscript{2, *}

\textsuperscript{1}Department of Mathematical Sciences, Huzhou University, Huzhou 313000, China;
\textsuperscript{2}School of Mathematical Sciences, Fudan University, Shanghai 200433, China

Email: tmorning@hutc.zj.cn, qzh@fudan.edu.cn

Received December 27, 2010; accepted May 10, 2011; published online February 27, 2012

Abstract The paper is concerned with optimal control of backward stochastic differential equation (BSDE) driven by Teugel's martingales and an independent multi-dimensional Brownian motion, where Teugel's martingales are a family of pairwise strongly orthonormal martingales associated with Lévy processes (see e.g., Nualart and Schoutens’ paper in 2000). We derive the necessary and sufficient conditions for the existence of the optimal control by means of convex variation methods and duality techniques. As an application, the optimal control problem of linear backward stochastic differential equation with a quadratic cost criteria (or backward linear-quadratic problem, or BLQ problem for short) is discussed and characterized by a stochastic Hamilton system.

Keywords stochastic control, stochastic maximum principle, Lévy processes, Teugel’s martingales, backward stochastic differential equations

MSC(2010) 93E20, 60H10, 60H30

Citation: Tang M N, Zhang Q. Optimal variational principle for backward stochastic control systems associated with Lévy processes. Sci China Math, 2012, 55(4): 745–761, doi: 10.1007/s11425-012-4370-6

\textbf{1 Introduction}

It is well known that the maximum principle for a stochastic optimal control problem involves the so-called adjoint process which solves the corresponding adjoint equation. In fact, the adjoint equation is, in general, a linear backward stochastic differential equation (BSDE) with a specified random terminal condition on the state. Unlike a forward stochastic differential equation, the solution of a BSDE is a pair of adapted solutions. Thus, in order to obtain the maximum principle, we need first obtain the existence and uniqueness theorem for the pair of adapted solutions of the adjoint equation.

The linear BSDE was first proposed by Bismut [4] in 1973. This research field developed fast after the pioneer work of Pardoux and Peng [16] in 1990 who got the existence and uniqueness theorem for the solution of nonlinear BSDE driven by Brownian motion. Now, BSDE theory has been playing an important role not only in dealing with stochastic optimal control problems, but in mathematical finance, particularly in hedging theory and nonlinear pricing theory for imperfect market (see e.g., [7]).

As for BSDE driven by the non-continuous martingale, Tang and Li [21] first discussed the existence and uniqueness theorem of the solution of BSDE driven by Poisson point process and consequently proved the maximum principle for optimal control of stochastic systems with random jumps. In 2000,
Nualart and Schoutens [14] got a martingale representation theorem for a type of Lévy processes through Teugel’s martingales which are a family of pairwise strongly orthonormal martingales associated with Lévy processes. Later, they proved in [15] the existence and uniqueness theory of BSDE driven by Teugel’s martingales. The above results are further extended to the one-dimensional BSDE driven by Teugel’s martingales and an independent multi-dimensional Brownian motion by Bahlali, Eddahbi and Essaky [1]. One can refer to [8,9,18,19] for more results of such kind of BSDEs.

In the mean time, the stochastic optimal control problems related to Teugel’s martingales were studied. In 2008, a stochastic linear-quadratic problem with Lévy processes was considered by Mitsui and Tabata [13], in which they established the closeness property of multi-dimensional backward stochastic Riccati differential equation (BSRDE) with Teugel’s martingales and proved the existence and uniqueness of the solution to such kind of one-dimensional BSRDE, and then an application of BSDE to a financial problem with full and partial observations was demonstrated. In 2009, Tang and Wu [20] also studied a class of stochastic linear-quadratic optimal control problem with Teugel’s martingales, where the cost weighting matrices of the state and control are allowed to be indefinite, and one kind of new stochastic Riccati equation was derived as well as the existence and uniqueness of its solution were obtained in some special cases. Motivated by [13,20], Meng and Tang [12] studied the general stochastic optimal control problem for the forward stochastic systems driven by Teugel’s martingales and an independent multi-dimensional Brownian motion, of which the necessary and sufficient optimality conditions in the form of stochastic maximum principle with the convex control domain were obtained.

However, [12,13,20] are only concerned with the optimal control problem of the forward controlled stochastic system. Since a BSDE is a well-defined dynamic system itself and has important applications in mathematical finance, it is necessary and natural to consider the optimal control problem of BSDE. Actually, there has been much literature on BSDE control system driven by Brownian motion (see e.g., [2,3,5,10,11]). But to our best knowledge, no discussion is given on the optimal control problem of BSDE driven by Teugel martingales and an independent Brownian motion.

In this paper, by means of convex variation methods and duality techniques, we will give the necessary and sufficient conditions for the existence of the optimal control of BSDE system driven by Teugel martingales and an independent multi-dimensional Brownian motion. As an application, the optimal control of linear backward stochastic differential equation with a quadratic cost criteria, i.e., the optimal control of the so-called backward linear-quadratic (BLQ) problem, is discussed in detail. Moreover, the optimal control of BLQ problem will be characterized by a stochastic Hamilton system. In this case, the stochastic Hamilton system is a linear forward-backward stochastic differential equation driven by Teugel’s martingales and an independent multi-dimensional Brownian motion, consisting of the state equation, the adjoint equation and the dual presentation of the optimal control.

The rest of this paper is organized as follows. In Section 2, we introduce useful notation and some existing results of stochastic differential equations (SDEs) and BSDEs driven by Teugel’s martingales. In Section 3, we indicate the optimal control problem we study, give needed assumptions and prove some preliminary results of variational equation and variational inequality. In Section 4, we prove the necessary and sufficient optimality conditions for the optimal control problem put forward in Section 3. As an application, the optimal control for BLQ problem is discussed in Section 5.

2 Notation and preliminaries

Let \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, P)\) be a complete probability space. The filtration \(\{\mathcal{F}_t\}_{0 \leq t \leq T}\) is right-continuous and generated by a \(d\)-dimensional standard Brownian motion \(\{W(t), 0 \leq t \leq T\}\) and a one-dimensional Lévy process \(\{L(t), 0 \leq t \leq T\}\) which are mutually independent. It is known that \(L\) has a characteristic function of the form

\[
E e^{i\theta L(t)} = \exp \left[ i\theta t - \frac{1}{2} \sigma^2 \theta^2 t + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x I_{\{|x|<1\}}) \nu(dx) \right],
\]
where $a \in \mathbb{R}^1$, $\sigma > 0$ and $v$ is a measure on $\mathbb{R}^1$ satisfying $(i) \int_{\mathbb{R}} (1 \wedge x^2) v(dx) < \infty$ and (ii) there exist $\varepsilon > 0$ and $\lambda > 0$, s.t. $\int_{\{|x| > \varepsilon\}} |v|(dx) < \infty$. These settings imply that the random variable $L(t)$ has moments of all orders for each $t \in [0, T]$. Denote by $\mathcal{P}$ the predictable sub-$\sigma$ field of $\mathcal{B}([0, T]) \times \mathcal{F}$, then we introduce the following notation used throughout this paper.

- $H$: a Hilbert space with norm $\| \cdot \|_H$.
- $(\alpha, \beta)$: the inner product in $\mathbb{R}^n$, for any $\alpha, \beta \in \mathbb{R}^n$.
- $|\alpha| = \sqrt{\langle \alpha, \alpha \rangle}$: the norm of $\mathbb{R}^n$, for any $\alpha \in \mathbb{R}^n$.
- $(A, B) = \text{tr}(AB^T)$: the inner product in $\mathbb{R}^{n \times m}$, for any $A, B \in \mathbb{R}^{n \times m}$.
- $|A| = \sqrt{\text{tr}(AA^T)}$: the norm of $\mathbb{R}^{n \times m}$, for any $A \in \mathbb{R}^{n \times m}$.
- $l^2$: the space of all real-valued sequences $x = \{x_n\}_{n \geq 0}$ satisfying

$$\|x\|_{l^2} \triangleq \left\{ \sum_{i=1}^{\infty} x_i^2 < +\infty \right\}.$$

- $l^2(H)$: the space of all $H$-valued sequence $f = \{f^i\}_{i \geq 1}$ satisfying

$$\|f\|_{l^2(H)} \triangleq \left\{ \sum_{i=1}^{\infty} \|f^i\|_H^2 < +\infty \right\}.$$

- $l^2_{\mathcal{P}}(0, T; H)$: the space of all $l^2(H)$-valued and $\mathcal{F}_t$-predictable processes $f = \{f^i(\omega, t), (\omega, t) \in \Omega \times [0, T]\}_{i \geq 1}$ satisfying

$$\|f\|_{l^2_{\mathcal{P}}(0, T; H)} \triangleq \left\{ E \int_0^T \sum_{i=1}^{\infty} \|f^i(t)\|_H^2 dt < \infty \right\}.$$

- $M^2_\mathcal{F}(0, T; H)$: the space of all $H$-valued and $\mathcal{F}_t$-adapted processes $f = \{f(\omega, t), (\omega, t) \in \Omega \times [0, T]\}$ satisfying

$$\|f\|_{M^2_\mathcal{F}(0, T; H)} \triangleq \left\{ E \int_0^T \|f(t)\|_H^2 dt < \infty \right\}.$$

- $S^2_\mathcal{F}(0, T; H)$: the space of all $H$-valued and $\mathcal{F}_t$-adapted cadlag processes $f = \{f(\omega, t), (\omega, t) \in \Omega \times [0, T]\}$ satisfying

$$\|f\|_{S^2_\mathcal{F}(0, T; H)} \triangleq \left\{ E \sup_{0 \leq t \leq T} \|f(t)\|_H^2 dt < \infty \right\}.$$

- $L^2(\Omega, \mathcal{F}, P; H)$: the space of all $H$-valued random variables $\xi$ satisfying

$$\|\xi\|_{L^2(\Omega, \mathcal{F}, P; H)} \triangleq \left\{ E\|\xi\|_H^2 < \infty \right\}.$$

We denote by $\{H^i(t), 0 \leq t \leq T\}_{i=1}^\infty$ the Teugel’s martingales associated with the Lévy process $\{L(t), 0 \leq t \leq T\}$. Here $H^i(t)$ is given by

$$H^i(t) = c_{i,1} Y^{(1)}(t) + c_{i,2} Y^{(2)}(t) + \cdots + c_{i,1} Y^{(1)}(t),$$

where $Y^{(0)}(t) = L^{(0)}(t) - E[L^{(i)}(t)]$ for all $i \geq 1$, $L^{(0)}(t)$ are so called power-jump processes with $L^{(1)}(t) = L(t), L^{(i)}(t) = \sum_{0 \leq s \leq t} (\Delta L(s))^i$ for $i \geq 2$ and the coefficients $c_{i,j}$ correspond to the orthonormalization of polynomials $1, x, x^2, \ldots$ w.r.t. the measure $\mu(dx) = x^2 v(dx) + \sigma^2 \delta_0(dx)$. The Teugel’s martingales $\{H^i(t)\}_{i=1}^\infty$ are pathwise strongly orthogonal and their predictable quadratic variation processes are given by

$$\langle H^{(1)}(t), H^{(1)}(t) \rangle = \delta_{1,j} t.$$

For more details of Teugel’s martingales, we invite the reader to consult Nualart and Schoutens [14,15].

In what follows, we will state some basic results of SDE and BSDE driven by Teugel’s martingales $\{H^i(t), 0 \leq t \leq T\}_{i=1}^\infty$ and the $d$-dimensional Brownian motion $\{W(t), 0 \leq t \leq T\}$.
Consider SDE:

\[ X(t) = a + \int_0^t b(s, X(s))ds + \sum_{i=1}^d \int_0^t g^i(s, X(s))dW^i(s) + \sum_{i=1}^{\infty} \int_0^t \sigma^i(s, X(s-))dH^i(s), \quad t \in [0, T], \quad (2.1) \]

where \((a, b, g, \sigma)\) are given mappings satisfying the assumptions below.

**Assumption 2.1.** Random variable \(a\) is \(\mathcal{F}_0\)-measurable and \((b, g, \sigma)\) are three random mappings

\[
\begin{align*}
    b & : \Omega \times [0, T] \times \mathbb{R}^n \to \mathbb{R}^n, \\
    g & : \Omega \times [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}, \\
    \sigma & : (\sigma^i)_{i=1}^\infty : \Omega \times [0, T] \times \mathbb{R}^n \to l^2(\mathbb{R}^n),
\end{align*}
\]

satisfying

(i) \(b, g\) and \(\sigma\) are \(\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n)\) measurable with \(b(\cdot, 0) \in M^2_{\mathcal{F}}(0; \mathbb{R}^n)\), \(g(\cdot, 0) \in M^2_{\mathcal{F}}(0; \mathbb{R}^{n \times d})\) and \(\sigma(\cdot, 0) \in l^2_{\mathcal{F}}(0; \mathbb{R}^n)\);

(ii) \(b, g\) and \(\sigma\) are uniformly Lipschitz continuous w.r.t. \(x\), i.e., there exists a constant \(C > 0\) such that for any \((t, x, \bar{x}) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n\) and a.s. \(\omega \in \Omega\),

\[
|b(t, x) - b(t, \bar{x})| + |g(t, x) - g(t, \bar{x})| + ||\sigma(t, x) - \sigma(t, \bar{x})||_{l^2(\mathbb{R}^n)} \leq C|x - \bar{x}|.
\]

**Lemma 2.1** (Existence and uniqueness theorem of SDE [20]). If coefficients \((a, b, g, \sigma)\) satisfy Assumption 2.1, then SDE (2.1) has a unique solution \(x(\cdot) \in S^2_{\mathcal{F}}(0; \mathbb{R}^n)\).

**Lemma 2.2** (Continuous dependence theorem of SDE [12]). Assume that coefficients \((a, b, g, \sigma)\) and \((\bar{a}, \bar{b}, \bar{g}, \bar{\sigma})\) satisfy Assumption 2.1. If \(x(\cdot)\) and \(\bar{x}(\cdot)\) are the solutions to SDE (2.1) corresponding to \((a, b, g, \sigma)\) and \((\bar{a}, \bar{b}, \bar{g}, \bar{\sigma})\), respectively, then we have

\[
E \sup_{0 \leq t \leq T} |x(t) - \bar{x}(t)|^2 \leq K \left[ |a - \bar{a}|^2 + E \int_0^T |b(t, \bar{x}(t)) - b(t, \bar{x}(t))|^2 dt \\
+ E \int_0^T |g(t, \bar{x}(t)) - g(t, \bar{x}(t))|^2 dt + E \int_0^T ||\sigma(t, \bar{x}(t)) - \sigma(t, \bar{x}(t))||_{l^2(\mathbb{R}^n)}^2 dt \right],
\]

where \(K\) is a positive constant depending only on \(T\) and the Lipschitz constant \(C\). In particular, when \((a, b, g, \sigma) = (0, 0, 0, 0)\), we have

\[
E \sup_{0 \leq t \leq T} |x(t)|^2 \leq K \left[ |a|^2 + E \int_0^T |b(t, 0)|^2 dt + E \int_0^T |g(t, 0)|^2 dt + E \int_0^T ||\sigma(t, 0)||_{l^2(\mathbb{R}^n)}^2 dt \right] < +\infty.
\]

Now we consider BSDE:

\[
y(t) = \xi + \int_t^T f(s, y(s), q(s), z(s))ds - \sum_{i=1}^d \int_t^T q^i(s)dW^i(s) - \sum_{i=1}^{\infty} \int_t^T z^i(s)dH^i(s), \quad t \in [0, T],
\]

where coefficients \((\xi, f)\) are given mappings satisfying the assumptions below.

**Assumption 2.2.** The terminal value \(\xi \in L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^n)\) and \(f\) is a random mapping

\[
f : \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n) \to \mathbb{R}^n
\]

satisfying

(i) \(f\) is \(\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^{n \times d}) \otimes \mathcal{B}(l^2(\mathbb{R}^n))\) measurable with \(f(\cdot, 0, 0, 0) \in M^2_{\mathcal{F}}(0; \mathbb{R}^n)\);

(ii) \(f\) is uniformly Lipschitz continuous w.r.t. \((y, q, z)\), i.e., there exists a constant \(C > 0\) such that for any \((t, y, q, z, \bar{y}, \bar{q}, \bar{z}) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n)\) and a.s. \(\omega \in \Omega\),

\[
|f(t, y, q, z) - f(t, \bar{y}, \bar{q}, \bar{z})| \leq C[|y - \bar{y}| + |q - \bar{q}| + |z - \bar{z}| ||2(\mathbb{R}^n)].
\]
Lemma 2.3 (Existence and uniqueness of BSDE [1]). If coefficients \((\xi, f)\) satisfy Assumption 2.2, then BSDE (2.2) has a unique solution
\[
(y(\cdot), q(\cdot), z(\cdot)) \in S^2_\mathcal{F}(0, T; \mathbb{R}^n) \times M^2_\mathcal{F}(0, T; \mathbb{R}^{n \times d}) \times \mathcal{L}^2_\mathcal{F}(0, T; \mathbb{R}^n).
\]

Lemma 2.4 (Continuous dependence theorem of BSDE [1]). Assume that coefficients \((\xi, f)\) and \((\bar{\xi}, \bar{f})\) satisfy Assumption 2.2. If \((y(\cdot), q(\cdot), z(\cdot))\) and \((\bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))\) are the solutions to BSDE (2.2) corresponding to \((\xi, f)\) and \((\bar{\xi}, \bar{f})\), respectively, then we have
\[
E \sup_{0 \leq t \leq T} |y(t) - \bar{y}(t)|^2 + E \int_0^T |q(t) - \bar{q}(t)|^2 dt + E \int_0^T \|z(t) - \bar{z}(t)\|^2 dt \leq K \left[ E|\xi - \bar{\xi}|^2 + E \int_0^T |f(t, \bar{y}(t), \bar{q}(t), \bar{z}(t)) - \bar{f}(t, \bar{y}(t), \bar{q}(t), \bar{z}(t))|^2 dt \right],
\]
where \(K\) is a positive constant depending only on \(T\) and the Lipschitz constant \(C\).

In particular, when \((\xi, f) = (0, 0)\), we have
\[
E \sup_{0 \leq t \leq T} |y(t)|^2 + E \int_0^T |q(t)|^2 dt + E \int_0^T \|z(t)\|^2 dt \leq K \left[ E|\xi|^2 + E \int_0^T |f(t, 0, 0, 0)|^2 dt \right].
\]

In view of Assumptions 2.1–2.2, Lemmas 2.1–2.4 follow from an application of Itô’s formula, Gronwall’s inequality and Burkholder-Davis-Gundy inequality. One can refer to [1], [12] and [20] for details.

3 Formulation of the problem and preliminary lemmas

Let the admissible control set \(U\) be a nonempty convex subset of \(\mathbb{R}^n\). An admissible control process \(u(\cdot)\) is defined as an \(\mathcal{F}_t\)-predictable process with values in \(U\) s.t. \(E \int_0^T |u(t)|^2 dt < +\infty\). We denote by \(\mathcal{A}\) the set including all admissible control processes.

For any given admissible control \(u(\cdot) \in \mathcal{A}\), we consider the following controlled nonlinear BSDE driven by multi-dimensional Brownian motion \(W\) and Teugel’s martingales \(\{H^i\}_{i=1}^\infty\):
\[
y(t) = \xi + \int_t^T f(s, y(s), q(s), z(s), u(s)) ds - \sum_{i=1}^d \int_t^T q_i(s) dW^i(s) - \sum_{i=1}^\infty \int_t^T z^i(s) dH^i(s), \ t \in [0, T], \tag{3.1}
\]
with the cost functional
\[
J(u(\cdot)) = E \left[ \int_0^T l(t, y(t), q(t), z(t), u(t)) dt + \Phi(y(0)) \right], \tag{3.2}
\]
where
\[
\begin{align*}
\xi & : \Omega \to \mathbb{R}^n, \\
f & : \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathcal{L}^2(\mathbb{R}^n) \times U \to \mathbb{R}^n, \\
l & : \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathcal{L}^2(\mathbb{R}^n) \times U \to \mathbb{R}^1 \\
\end{align*}
\]
and
\[
\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^1
\]
are given coefficients.

Throughout this paper, we introduce the following basic assumptions on coefficients \((\xi, f, l, \phi)\).

Assumption 3.1. The terminal value \(\xi \in L^2(\Omega, \mathcal{F}_T, P; \mathbb{R}^n)\) and the random mapping \(f\) in \(\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^{n \times d}) \otimes \mathcal{B}(\mathcal{L}^2(\mathbb{R}^n)) \otimes \mathcal{B}({\mathbb{R}})\) measurable with \(f(\cdot, 0, 0, 0, 0) \in M^2(0, T; \mathbb{R}^n)\). For almost all \((t, \omega) \in [0, T] \times \Omega\), \(f(t, \omega, y, p, z, u)\) is Gâteaux differentiable w.r.t. \((y, p, z, u)\) and the corresponding Gâteaux derivatives \((f_y, f_p, f_z, f_u)\) are continuous and uniformly bounded.
Assumption 3.2. The random mapping $l$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^{n \times d}) \otimes \mathcal{B}(\mathbb{P}(\mathbb{R}^n)) \otimes \mathcal{B}(U)$ measurable and for almost all $(t, \omega) \in [0, T] \times \Omega$, $l$ is Gâteaux differentiable w.r.t. $(y, p, z, u)$ with continuous Gâteaux derivatives $(l_y, l_p, l_z, l_u)$. The random mapping $\phi$ is $\mathcal{F}_0 \otimes \mathcal{B}(\mathbb{R}^n)$ measurable and for almost all $(\omega, t) \in \Omega \times [0, T]$, $\phi$ is Gâteaux differentiable w.r.t. $y$ with the continuous Gâteaux derivative $\phi_y$. Moreover, for almost all $(\omega, t) \in \Omega \times [0, T]$, there exists a constant $C$ s.t. for any $(p, q, z, u) \in \mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathbb{P}(\mathbb{R}^n) \times U$,

\[
|l| \leq C(1 + |y|^2 + |q|^2 + |z|^2 + |u|^2), \\
|\phi| \leq C(1 + |y|^2), \\
|l_y| + |l_q| + |l_z| + |l_u| \leq C(1 + |y| + |q| + |z| + |u|), \\
|\phi_y| \leq C(1 + |y|).
\]

Under Assumption 3.1, we know that the controlled system (3.1) satisfies Lipschitz condition, thus we can get from Lemma 2.3 that for each $u(\cdot) \in \mathcal{A}$, the system (3.1) admits a unique strong solution. We denote the strong solution of (3.1) by $(y^u(\cdot), q^u(\cdot), z^u(\cdot))$, or $(y(\cdot), q(\cdot), z(\cdot))$ if its dependence on admissible control $u(\cdot)$ is clear from the context. Then we call $(y(\cdot), q(\cdot), z(\cdot))$ the state process corresponding to the control process $u(\cdot)$ and call $(u(\cdot); y(\cdot), q(\cdot), z(\cdot))$ the admissible pair. Furthermore, by Assumption 3.2 and a priori estimate (2.3), it is easy to check that the optimal control problem is well-defined, i.e.,

\[
|J(u(\cdot))| < \infty.
\]

Then we put forward the optimal control problem we study.

**Problem 3.1.** Find an admissible control $\bar{u}(\cdot)$ such that

\[
J(\bar{u}(\cdot)) = \inf_{u(\cdot) \in \mathcal{A}} J(u(\cdot)).
\]

Any $\bar{u}(\cdot) \in \mathcal{A}$ satisfying the above is called an optimal control process of Problem 3.1 and the corresponding state process $(\bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))$ is called the optimal state process. Correspondingly, $(\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))$ is called an optimal pair of Problem 3.1.

Before we deduce the necessary and sufficient conditions for the optimal control of Problem 3.1, we need do some preparations. Since the control domain $U$ is convex, the classical method to get necessary conditions for optimal control processes is the so-called convex perturbation method. More precisely, assuming that $(\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))$ is an optimal pair of Problem 3.1, for any given admissible control $u(\cdot)$, we define an admissible control in the form of convex variation

\[
u^\varepsilon(\cdot) = \bar{u}(\cdot) + \varepsilon(u(\cdot) - \bar{u}(\cdot)),
\]

where $\varepsilon > 0$ can be chosen sufficiently small. Denoting by $(y^\varepsilon(\cdot), q^\varepsilon(\cdot), z^\varepsilon(\cdot))$ the state process of the control system (3.1) corresponding to the control process $\nu^\varepsilon(\cdot)$, we obtain the variational inequality

\[
J(\nu^\varepsilon(\cdot)) - J(\bar{u}(\cdot)) \geq 0.
\]

In what follows, we do some estimates on the optimal pair and the convex variable pair.

**Lemma 3.2.** Under Assumptions 3.1–3.2, we have

\[
E \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t)|^2 + E \int_0^T |q^\varepsilon(t) - \bar{q}(t)|^2 dt + E \int_0^T \|z^\varepsilon(t) - \bar{z}(t)\|^2_{\mathbb{P}(\mathbb{R}^n)} dt = O(\varepsilon^2).
\]

**Proof.** By the continuous dependence theorem of BSDE (Lemma 2.4) and the uniformly bounded
Lemma 3.3. Under Assumptions 3.1–3.2, it follows that

\[
E \sup_{0 \leq t \leq T} |y^\varepsilon(t) - \bar{y}(t) - \varepsilon Y(t)|^2 + E \int_0^T |q^\varepsilon(t) - \bar{q}(t) - \varepsilon Q(t)|^2 dt \\
+ E \int_0^T \|z^\varepsilon(t) - \bar{z}(t) - \varepsilon Z(t)\|^2_{\mathbb{L}^2(\mathbb{R}^n)} dt = o(\varepsilon^2).
\]

Proof. Firstly, one can check that

\[
y^\varepsilon(t) - \bar{y}(t) = \int_0^T [f_y(s) y^\varepsilon(s) - \bar{y}(s)] ds + \int_0^T [f_q(s) q^\varepsilon(s) - \bar{q}(s)] ds \\
+ \int_0^T [f_z(s) z^\varepsilon(s) - \bar{z}(s)] ds + \int_0^T [f_u(s) u^\varepsilon(s) - \bar{u}(s)] ds \\
- \sum_{i=1}^d \int_t^T (q^i(s) - \bar{q}(s)) dW^i(s) - \sum_{i=1}^d \int_t^T (z^i(s) - \bar{z}(s)) dH^i(s)
\]

and

\[
\varepsilon Y(t) = \int_t^T [f_y(s) \varepsilon Y(s) + f_q(s) \varepsilon Q(s) + f_z(s) \varepsilon Z(s) + f_u(s) \varepsilon (u(s) - \bar{u}(s))] ds \\
- \sum_{i=1}^d \int_t^T \varepsilon Q^i(s) dW^i(s) - \sum_{i=1}^d \int_t^T \varepsilon Z^i(s) dH^i(s).
\]
where we have used the abbreviations for $\varphi = f, l$, as follows:

\[
\begin{align*}
\varphi_y(t) &= \varphi_y(t, \dot{y}(t), \ddot{y}(t), \dot{z}(t), \ddot{u}(t)), \\
\varphi_z(t) &= \varphi_z(t, \dot{y}(t), \ddot{y}(t), \dot{z}(t), \ddot{u}(t)), \\
\varphi_q(t) &= \varphi_q(t, \dot{y}(t), \ddot{y}(t), \dot{z}(t), \ddot{u}(t)), \\
\varphi_u(t) &= \varphi_u(t, \dot{y}(t), \ddot{y}(t), \dot{z}(t), \ddot{u}(t)), \\
\varphi^x_y(t) &= \int_0^1 \varphi_y(t, \dot{y}(t) + \lambda(y^x(t) - \dot{y}(t)), \ddot{y}(t) + \lambda(q^x(t) - \ddot{y}(t)), \\
& \quad \dot{z}(t) + \lambda(z^x(t) - \dot{z}(t)), \ddot{u}(t) + \lambda(u^x(t) - \ddot{u}(t))d\lambda \\
\varphi^x_z(t) &= \int_0^1 \varphi_z(t, \dot{y}(t) + \lambda(y^x(t) - \dot{y}(t)), \ddot{y}(t) + \lambda(q^x(t) - \ddot{y}(t)), \\
& \quad \dot{z}(t) + \lambda(z^x(t) - \dot{z}(t)), \ddot{u}(t) + \lambda(u^x(t) - \ddot{u}(t))d\lambda \\
\varphi^x_q(t) &= \int_0^1 \varphi_q(t, \dot{y}(t) + \lambda(y^x(t) - \dot{y}(t)), \ddot{y}(t) + \lambda(q^x(t) - \ddot{y}(t)), \\
& \quad \dot{z}(t) + \lambda(z^x(t) - \dot{z}(t)), \ddot{u}(t) + \lambda(u^x(t) - \ddot{u}(t))d\lambda \\
\varphi^x_u(t) &= \int_0^1 \varphi_u(t, \dot{y}(t) + \lambda(y^x(t) - \dot{y}(t)), \ddot{y}(t) + \lambda(q^x(t) - \ddot{y}(t)), \\
& \quad \dot{z}(t) + \lambda(z^x(t) - \dot{z}(t)), \ddot{u}(t) + \lambda(u^x(t) - \ddot{u}(t))d\lambda.
\end{align*}
\]

Thus by Lemma 2.4 again, we get

\[
\begin{align*}
E & \sup_{0 \leq t \leq T} |g^x(t) - \bar{g}(t) - \varepsilon Y(t)|^2 + E \int_0^T |q^x(t) - \bar{q}(t) - \varepsilon Q(t)|^2 dt \\
& \quad + E \int_0^T \|z^x(t) - \bar{z}(t) - \varepsilon Z(t)\|^2 \|z(\mathbb{R}^n)\| dt \\
& \leq K\varepsilon^2 \left[ E \int_0^T (|\tilde{f}_y(t) - f_y(t)| Y(t) + (\tilde{f}_q(t) - f_q(t)) Q(t) + (\tilde{f}_z(t) - f_z(t)) Z(t) \\
& \quad \quad \quad + (\tilde{f}_u(t) - f_u(t))(u(t) - \ddot{u}(t))|^2 dt \right] \\
& \leq K\varepsilon^2 \cdot \alpha(\varepsilon),
\end{align*}
\]

where

\[
\alpha(\varepsilon) = E \int_0^T (|\tilde{f}_y(t) - f_y(t)| Y(t) + (\tilde{f}_q(t) - f_q(t)) Q(t) \\
& \quad + (\tilde{f}_z(t) - f_z(t)) Z(t) + (\tilde{f}_u(t) - f_u(t))(u(t) - \ddot{u}(t))|^2 dt.
\]

Consequently, using Lemma 3.2 and Assumption 3.1, by the dominated convergence theorem we can deduce

\[
\lim_{\varepsilon \to 0} \alpha(\varepsilon) = 0.
\]

Then the lemma follows from the above and (3.5).

**Lemma 3.4.** Under Assumptions 3.1–3.2, using the abbreviations in (3.4) we have

\[
J(u^x(\cdot)) - J(\bar{u}(\cdot)) = \varepsilon E\phi_y(\bar{y}(0)) Y(0) + \varepsilon E \int_0^T l_y(t) Y(t) dt + \varepsilon E \int_0^T l_q(t) Q(t) dt \\
& \quad + \varepsilon E \int_0^T l_z(t) Z(t) dt + \varepsilon E \int_0^T l_u(t)(u(t) - \ddot{u}(t)) dt + o(\varepsilon).
\]

**Proof.** After a first order development, we have

\[
J(u^x(\cdot)) - J(\bar{u}(\cdot)) = E \int_0^1 \phi_y(\bar{y}(0) + \lambda(y^x(0) - \bar{y}(0))(y^x(0) - \bar{y}(0))d\lambda
\]
Thus combining Lemma 3.2, Lemma 3.3 and Assumption 3.2, by the dominated convergence theorem, we have:

\[
\beta(\epsilon) = E\phi_y(\bar{y}(0))(y^*(0) - \bar{y}(0) - \epsilon Y(0)) \\
+ E \int_0^T [\phi_y(\bar{y}(t)) - \phi_y(\bar{y}(0))] \frac{dY(t)}{dt} dt \\
+ E \int_0^T l_y(t)(q^*(t) - \bar{q}(t)) - \epsilon Y(t) dt \\
+ E \int_0^T l_q(t)(y^*(t) - \bar{y}(t)) dt \\
+ E \int_0^T l_y(t)\frac{d}{dt}(z^*(t) - \bar{z}(t)) + E \int_0^T l_z(t)(z^*(t) - \bar{z}(t)) dt \\
+ E \int_0^T (\bar{l}_y(t) - l_y(t))\epsilon(u(t) - \bar{u}(t)) dt.
\]

where \( \beta(\epsilon) \) is given by

\[
\beta(\epsilon) = E\phi_y(\bar{y}(0))(y^*(0) - \bar{y}(0) - \epsilon Y(0)) \\
+ E \int_0^T [\phi_y(\bar{y}(t)) - \phi_y(\bar{y}(0))] \frac{dY(t)}{dt} dt \\
+ E \int_0^T l_y(t)(q^*(t) - \bar{q}(t)) - \epsilon Y(t) dt \\
+ E \int_0^T l_q(t)(y^*(t) - \bar{y}(t)) dt \\
+ E \int_0^T l_y(t)\frac{d}{dt}(z^*(t) - \bar{z}(t)) + E \int_0^T l_z(t)(z^*(t) - \bar{z}(t)) dt \\
+ E \int_0^T (\bar{l}_y(t) - l_y(t))\epsilon(u(t) - \bar{u}(t)) dt.
\]

Thus combining Lemma 3.2, Lemma 3.3 and Assumption 3.2, by the dominated convergence theorem, we conclude that \( \beta(\epsilon) = o(\epsilon) \).

By Lemma 3.3 and the fact that \( \lim_{\epsilon \to 0^+} \frac{J(w^*) - J(u^*)}{\epsilon} \geq 0 \), we can further deduce:

**Corollary 3.5.** Under Assumptions 3.1–3.2, we have the variation inequality below:

\[
E\phi_y(\bar{y}(0))Y(0) + E \int_0^T l_y(t)Y(t) dt + E \int_0^T l_q(t)Q(t) dt \\
+ E \int_0^T l_z(t)Z(t) dt + E \int_0^T l_u(t)(u(t) - \bar{u}(t)) dt \geq 0.
\]

(3.6)

## 4 Necessary and sufficient optimality conditions

We first introduce the adjoint equation corresponding to the variational equation (3.3):

\[
\begin{cases}
    dk(t) = -[f^*_y(t)k(t) + l_y(t)] dt - \sum_{i=1}^d [-f^*_{y_i}(t)k(t) + l_{y_i}(t)] dW^i(t) \\
    k(0) = -\phi_y(\bar{y}(0)), \quad 0 \leq t \leq T,
\end{cases}
\]

(4.1)

where \( f^*_y, f^*_{y_i} \) and \( f^*_z \) are the dual operators of \( f_y, f_{y_i} \) and \( f_z \), respectively.

Under Assumptions 3.1–3.2, the coefficients of above adjoint equation are square integrable. By Lemma 2.1, (4.1) has a unique solution \( k(\cdot) \in S^2(0, T; \mathbb{R}^n) \). Then we define the Hamiltonian function \( H : [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n) \times U \times \mathbb{R}^n \rightarrow \mathbb{R}^1 \) by

\[
H(t, y, q, z, u, k) = \langle k, -f(t, y, q, z, u) \rangle + l(t, y, q, z, u).
\]

(4.2)
and rewrite the adjoint equation in a Hamiltonian system form:

\[
\begin{aligned}
    dk(t) &= -H_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))dt \\
    &\quad - \sum_{i=1}^d H'_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))dW^i(t) \\
    &\quad - \sum_{i=1}^d H_z(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))dH^i(t), \\
    k(0) &= -\phi_y(\bar{y}(0)).
\end{aligned}
\]

(4.3)

Now, we are ready to give the necessary condition for an optimal control of Problem 3.1.

**Theorem 4.1.** Under Assumptions 3.1–3.2, if \((\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))\) is an optimal pair of Problem 3.1, then we have

\[
H_u(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))(u - \bar{u}(t)) \geq 0, \quad \text{for any } u \in U, \ a.e. \ a.s.,
\]

(4.4)

where \(k(\cdot)\) is the solution to the adjoint equation (4.1).

**Proof.** By (3.3) and (4.1), applying Itô formula to \(\langle Y(t), k(t) \rangle\) we have

\[
E\phi_y(\bar{y}(0))Y(0) + E \int_0^T l_y(t)Y(t)dt + E \int_0^T l_q(t)Q(t)dt \\
+ E \int_0^T l_z(t)Z(t)dt + E \int_0^T l_u(t)(u(t) - \bar{u}(t))dt \\
= -E \int_0^T (k(t), f_u(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t))(u(t) - \bar{u}(t)))dt + E \int_0^T l_u(t)(u(t) - \bar{u}(t))dt.
\]

Then noticing the definition of Hamilton function (4.2) and the variational inequality (3.6), for any \(u(\cdot) \in A\), we have

\[
E \int_0^T H_u(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))(u(t) - \bar{u}(t))dt \geq 0,
\]

which implies (4.4).

We next consider the sufficient condition for an optimal control of Problem 3.1.

**Theorem 4.2.** Under Assumptions 3.1–3.2, let \((\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))\) be the unique solution to the corresponding adjoint equation (4.3). Assume that for almost all \((\omega, t) \in \Omega \times [0,T]\), \(H(t, y, q, z, u, k(t))\) and \(\phi(y)\) are convex w.r.t. \((y, q, z, u)\) and \(y\), respectively, and the optimality condition

\[
H(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t)) = \min_{u \in U} H(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), u, k(t))
\]

holds, then \((\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot), \bar{z}(\cdot))\) is an optimal pair of Problem 3.1.

**Proof.** Let \((u(\cdot); y(\cdot), q(\cdot), z(\cdot))\) be an arbitrary admissible pair. It follows from the form of the cost functional (3.2) that

\[
J(u(\cdot)) - J(\bar{u}(\cdot)) = E \int_0^T [l(t, y(t), q(t), z(t), u(t)) - l(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t))]dt + E[\phi(y(0)) - \phi(\bar{y}(0))]
\]

\[
= I_1 + I_2,
\]

(4.5)

where

\[
I_1 = E \int_0^T [l(t, y(t), q(t), z(t), u(t)) - l(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t))]dt
\]

and

\[
I_2 = E[\phi(y(0)) - \phi(\bar{y}(0))].
\]
Due to the convexity of $\phi$, applying Itô formula to $\langle k(t), y(t) - \bar{y}(t) \rangle$, we have

$$I_2 = E[\phi(y(0)) - \phi(\bar{y}(0))]$$

$$\geq E[\phi(y(0)), y(0) - \bar{y}(0))$$

$$= -E[(k(0), y(0) - \bar{y}(0))]$$

$$= -E \int_0^T \langle H_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), y(t) - \bar{y}(t) \rangle \ dt$$

$$- \sum_{i=1}^d E \int_0^T \langle H^i_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), q^i(t) - \bar{q}^i(t) \rangle \ dt$$

$$- \sum_{i=1}^\infty E \int_0^T \langle H^i_z(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), z^i(t) - \bar{z}^i(t) \rangle \ dt$$

$$- E \int_0^T \langle f(t, y(t), q(t), z(t), u(t)) - f(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t)), k(t) \rangle \ dt$$

$$= -J_1 + J_2,$$  \hspace{1cm} (4.6)

where

$$J_1 = E \int_0^T \langle H_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), y(t) - \bar{y}(t) \rangle \ dt$$

$$+ \sum_{i=1}^d E \int_0^T \langle H^i_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), q^i(t) - \bar{q}^i(t) \rangle \ dt$$

$$+ \sum_{i=1}^\infty E \int_0^T \langle H^i_z(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), z^i(t) - \bar{z}^i(t) \rangle \ dt$$

and

$$J_2 = -E \int_0^T \langle f(t, y(t), q(t), z(t), u(t)) - f(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t)), k(t) \rangle \ dt.$$

Using the definition of the Hamiltonian function (4.2) again, we have

$$I_1 = E \int_0^T [l(t, y(t), q(t), z(t), u(t)) - l(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t))] \ dt$$

$$= E \int_0^T \langle H(t, y(t), q(t), z(t), u(t), k(t)) - H(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t)) \rangle \ dt$$

$$+ E \int_0^T \langle f(t, y(t), q(t), z(t), u(t)) - f(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t)), k(t) \rangle \ dt$$

$$= J_3 - J_2,$$  \hspace{1cm} (4.7)

where

$$J_3 = E \int_0^T \langle H(t, y(t), q(t), z(t), u(t), k(t)) - H(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t)) \rangle \ dt.$$  \hspace{1cm} (4.8)

Since $H(t, y, q, z, u, k(t))$ is convex w.r.t. $(y, q, z, u)$ for almost all $(\omega, t) \in \Omega \times [0, T]$, it turns out that

$$H(t, y(t), q(t), z(t), u(t), k(t)) - H(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{u}(t), k(t))$$

$$\geq \langle H_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), y(t) - \bar{y}(t) \rangle$$

$$+ \sum_{i=1}^d \langle H^i_y(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), q^i(t) - \bar{q}^i(t) \rangle$$

$$+ \sum_{i=1}^\infty \langle H^i_z(t, \bar{y}(t), \bar{q}(t), \bar{z}(t), \bar{\bar{u}}(t), k(t)), z^i(t) - \bar{z}^i(t) \rangle$$
Therefore, by (4.8)–(4.10) we first have
\[
J_3 \geq J_1.
\] (4.11)

By (4.11), together with (4.5)–(4.7), it also follows that
\[
J(u(\cdot)) - J(\bar{u}(\cdot)) = I_1 + I_2 = (J_3 - J_2) + (-J_1 + J_2) \geq (J_1 - J_2) + (-J_1 + J_2) = 0.
\]

Due to the arbitrariness of \(u(\cdot)\), we claim that \(\bar{u}(\cdot)\) is an optimal control process and thus \((\bar{u}(\cdot); \bar{y}(\cdot), \bar{q}(\cdot),
\bar{z}(\cdot))\) is an optimal pair. \qed

Remark 4.3. In this paper, we are concerned with the stochastic maximum principle for the optimal control of the backward stochastic systems associated with Lévy processes, when the control domain is a convex set. The stochastic maximum principle aims to establish the necessary and sufficient conditions for the optimal control, which are usually obtained by the convex perturbation method under the assumption of the convexity of control domain (see e.g., [17]). We would like to indicate that there is another important approach, i.e., the Bellman dynamic programming principle, to study the general optimal control problem. For the backward stochastic systems associated with Lévy processes, one can consider e.g., the following parameterized optimal control problem:

\[
\begin{aligned}
&dX(s) = b(s, X(s))ds + \sum_{i=1}^{d} \sigma(s, X(s))dW^i(s) + \sum_{i=1}^{\infty} c(s, X(s-))dH^i(s), \quad s > t, \\
&X(t) = x, \\
&dY(s) = -f(s, Y(s), Q(s), Z(s), u(s))ds + \sum_{i=1}^{d} Q(s)dW^i(s) + \sum_{i=1}^{\infty} Z(s)dH^i(s), \quad t \leq s < T, \\
&Y(T) = \psi(X(T)).
\end{aligned}
\]

In the simple case that the comparison theorem of the above BSDE is true, the standard procedure for this stochastic system is to establish the Bellman dynamic programming principle by the backward semigroup and derive Hamilton-Jacobi-Bellman (HJB) equation. If HJB equation is solvable, then the value function of optimal control is its unique solution. We will discuss this interesting problem in future work.

5 Applications in BLQ problem

In this section, we will apply our stochastic maximum principle to BLQ problem, i.e., minimizing the following quadratic cost functional over \(u(\cdot) \in \mathcal{A}\):

\[
J(u(\cdot)) := E\langle My(0), y(0) \rangle + E \int_0^T \langle E(s)g(s), y(s) \rangle ds + \sum_{i=1}^{d} E \int_0^T \langle F^i(s)q^i(s), q^i(s) \rangle ds
+ \sum_{i=1}^{\infty} E \int_0^T \langle G^i(s)z^i(s), z^i(s) \rangle ds + E \int_0^T \langle N(s)u(s), u(s) \rangle ds,
\] (5.1)
where the state process \((y(\cdot), q(\cdot), z(\cdot))\) is the solution to the controlled linear backward stochastic system below:

\[
\begin{aligned}
g y(t) &= - \left[ A(t) y(t) + \sum_{i=1}^{d} W^i(t) q^i(t) + \sum_{i=1}^{\infty} C^i(t) z^i(t) + D(t) u(t) \right] dt \\
&+ \sum_{i=1}^{d} q^i(t) dW^i(t) + \sum_{i=1}^{\infty} z^i(t) dH^i(t),
\end{aligned}
\tag{5.2}
\]

To study this problem, we need the assumptions on the coefficients as follows.

**Assumption 5.1.** The \(\mathcal{F}_t, 0 \leq t \leq T\)-predictable matrix processes \(A : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times n}, B^i : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times n}, i = 1, 2, \ldots, d, C^i : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times n}, i = 1, 2, \ldots, D : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times m}, E : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times m}, F^i : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times n}, i = 1, 2, \ldots, d, G^i : \Omega \times [0, T] \rightarrow \mathbb{R}^{n \times n}, i = 1, 2, \ldots, N : \Omega \times [0, T] \rightarrow \mathbb{R}^{m \times m}\) and the \(\mathcal{F}_T\)-measurable random matrix \(M : \Omega \rightarrow \mathbb{R}^{n \times n}\) are uniformly bounded.

**Assumption 5.2.** The state weighting matrix processes \(E, F^i, G^i\), the control weighting matrix process \(N\) and the random matrix \(M\) are a.e. a.s. symmetric and nonnegative. Moreover, \(N\) is a.e. a.s. uniformly positive, i.e., \(N \geq \delta I\) for some positive constant \(\delta\) a.e. a.s.

**Assumption 5.3.** There is no further constraint imposed on the control processes, i.e.,

\[\mathcal{A} = \left\{ u(\cdot) : u(\cdot) \text{ is } \mathcal{F}_t\text{-predictable with values in } \mathbb{R}^m \text{ and } E \int_0^T |u(t)|^2 dt < \infty \right\}.\]

By Assumption 5.3, \(\mathcal{A}\) is a Hilbert space. If we denote the norm of \(\mathcal{A}\) by \(\| \cdot \|_\mathcal{A}\), then for any control process \(u(\cdot) \in \mathcal{A}\),

\[\|u(\cdot)\|_\mathcal{A} = E \sqrt{\int_0^T |u(t)|^2 dt}.\]

Under Assumptions 5.1, by Lemma 2.3 we first know that the linear BSDE (5.2) in BLQ problem has a unique solution and thus the BLQ problem is well defined. Then, under Assumptions 5.1–5.3, we will demonstrate that BLQ problem has a unique optimal control.

**Lemma 5.1.** Under Assumptions 5.1–5.3, the cost functional \(J\) is strictly convex over \(\mathcal{A}\) and

\[\lim_{\|u(\cdot)\|_\mathcal{A} \to \infty} J(u(\cdot)) = \infty.\]

**Proof.** The convexity of the cost functional \(J\) over \(\mathcal{A}\) is obvious. Actually, since the weighting matrix process \(N\) is uniformly positive, \(J\) is strictly convex. In view of the nonnegative property of \(M, E, F^i, G^i\) and the strictly positive property of \(N\), we have

\[J(u(\cdot)) \geq \delta E \int_0^T |u(t)|^2 dt = \delta \|u(\cdot)\|^2_\mathcal{A}.\]

Therefore,

\[\lim_{\|u(\cdot)\|_\mathcal{A} \to \infty} J(u(\cdot)) = \infty.\]

**Lemma 5.2.** Under Assumptions 5.1–5.3, the cost functional \(J\) is Gâteaux differentiable over \(\mathcal{A}\) and its Gâteaux derivative \(J'\) at any admissible control process \(u(\cdot) \in \mathcal{A}\) is given by

\[
\langle J'(u(\cdot)), v(\cdot) \rangle = 2E \int_0^T \langle E(t)y^u(t), Y^v(t) \rangle dt + 2 \sum_{i=1}^{d} E \int_0^T \langle F^i(t)q^u(t), Q^{iv}(t) \rangle dt \\
+ 2 \sum_{i=1}^{\infty} E \int_0^T \langle G^i(t)z^{iu}(t), Z^{iv}(t) \rangle dt + 2E \int_0^T \langle N(t)u(t), v(t) \rangle dt \\
+ 2E \langle Mg^u(0), Y^v(0) \rangle,\]

\tag{5.3}
Then it follows from Assumption 5.1 and a priori estimate (2.3) that

Proof. For any $v(\cdot) \in \mathcal{A}$, we set

$$\Delta J = J(u(\cdot) + v(\cdot)) - J(u(\cdot)) - 2E \int_0^T \langle E(t)y^u(t), Y^v(t) \rangle dt$$

$$- 2 \sum_{i=1}^d E \int_0^T \langle F^i(t)q^u(t), Q^iv(t) \rangle dt - 2 \sum_{i=1}^d E \int_0^T \langle G^i(t)z^u(t), Z^iv(t) \rangle dt$$

$$- 2E \int_0^T \langle N(t)u(t), v(t) \rangle dt - 2E \langle My^u(0), Y^v(0) \rangle.$$

By the definition of cost functional (5.1), we have

$$\Delta J = E(MY^v(0), Y^v(0)) + E \int_0^T \langle E(s)Y^v(s), Y^v(s) \rangle ds + \sum_{i=1}^d E \int_0^T \langle F^i(s)Q^iv(s), Q^iv(s) \rangle ds$$

$$+ \sum_{i=1}^d E \int_0^T \langle G^i(s)Z^iv(s), Z^iv(s) \rangle ds + E \int_0^T \langle N(s)v(s), v(s) \rangle ds.$$

Then it follows from Assumption 5.1 and a priori estimate (2.3) that

$$|\Delta J| \leq K \left[ E \sup_{0 \leq t \leq T} |Y^v(t)|^2 + E \int_0^T |Q^v(t)|^2 dt + E \int_0^T ||Z^v(t)||^2_{L[\mathbb{R}^N]} dt + E \int_0^T |v(t)|^2 dt \right]$$

$$\leq KE \int_0^T |v(t)|^2 dt = K\|v(\cdot)\|^2_{\mathcal{A}}.$$

Consequently, we have

$$\lim_{\|v(\cdot)\|_{\mathcal{A}} \to 0} \frac{|\Delta J|}{\|v(\cdot)\|_{\mathcal{A}}} = 0,$$

which implies that $J$ is Gâteaux differentiable and its Gâteaux derivative $J'$ is given by (5.3). \hfill \qed

The strict convexity and the Gâteaux differentiability of $J$ deduced from Lemmas 5.1–5.2 lead to the lower semi-continuity of $J$, thus the following lemma is applicable to $J$ and $\mathcal{A}$ in our BLQ problem.

Lemma 5.3 (Proposition 1.2 of Chapter II in [6]). Let $\mathcal{A}$ be a reflexive Banach space and $J : \mathcal{A} \to \mathbb{R}^1$ be a convex function. Assume that $J$ is lower semi-continuous and proper, and consider the minimization problem

$$\inf_{u \in \mathcal{A}} J(u).$$

If the function $J$ is coercive over $\mathcal{A}$, i.e.,

$$\lim_{\|u\|_{\mathcal{A}} \to \infty} J(u) = \infty,$$

then the minimization problem has at least one solution. Moreover, if $J$ is strictly convex over $\mathcal{A}$, then the minimization problem has a unique solution.

By Lemma 5.3 we can immediately conclude

Theorem 5.4. Under Assumptions 5.1–5.3, BLQ problem has a unique optimal control.

In what follows, we will utilize the stochastic maximum principle to study the dual representation of the optimal control to BLQ problem and construct its stochastic Hamilton system. As in Section 4, we
first introduce the adjoint forward equation corresponding to an admissible pair \((u(\cdot); y(\cdot), q(\cdot), z(\cdot))\):

\[
\begin{align*}
dk(t) &= (A^*(t)k(t) - 2E(t)y(t))dt + \sum_{i=1}^{d} (B^{i*}(t)k^i(t) - 2F^i(t)q^i(t))dW^i(t) \\
& \quad + \sum_{i=1}^{\infty} (C^{i*}(t)k(t) - 2G^i(t)z^i(t))dH^i(t), \\
k(0) &= -2My(0).
\end{align*}
\]  

(5.4)

Also we define the Hamiltonian function \(H : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \times l^2(\mathbb{R}^n) \times U \times \mathbb{R}^n \rightarrow \mathbb{R}^1\) by

\[
H(t, y, q, z, u, k) = -\left\langle k, A(t)y + \sum_{i=1}^{d} W^i(t)q^i + \sum_{i=1}^{\infty} C^i(t)z^i + D(t)u \right\rangle \\
+ \langle E(t)y, q \rangle + \sum_{i=1}^{d} \langle F^i(t)q^i, q^i \rangle + \sum_{i=1}^{\infty} \langle G^i(t)z^i, z^i \rangle + (N(t)u, u).
\]  

(5.5)

Then the adjoint equation can be rewritten as a Hamiltonian form:

\[
\begin{align*}
dk(t) &= -H_y(t, y(t), q(t), z(t), u(t), k(t))dt - \sum_{i=1}^{d} H_q^i(t, y(t), q(t), z(t), u(t), k(t))dW^i(t) \\
& \quad - \sum_{i=1}^{\infty} H_z^i(t, y(t), q(t), z(t), u(t), k(t))dH^i(t), \\
k(0) &= -2My(0).
\end{align*}
\]  

(5.6)

Under Assumption 5.1, for any admissible pair \((u(\cdot); y(\cdot), q(\cdot), z(\cdot))\), the adjoint equation (5.6) has a unique solution \(k(\cdot)\) in view of Lemma 2.1.

It is time to give the the dual characterization of the optimal control.

**Theorem 5.5.** Under Assumptions 5.1–5.3, the BLQ problem has a unique optimal control and the optimal control is given by

\[
u(t) = -\frac{1}{2}N^{-1}(t)D^*(t)k(t), \quad \text{a.e. a.s.},
\]

(5.7)

where \(k(\cdot)\) is the unique solution to the adjoint equation (5.4) (or equivalently, (5.6)) corresponding to the optimal pair \((u(\cdot); y(\cdot), q(\cdot), z(\cdot))\).

**Proof.** By Theorem 5.4, we have the existence and uniqueness of optimal control to BLQ problem and denote the optimal control by \(u(\cdot)\). We only need to prove that \(u\) has an expression as (5.7). For this, let \((y(\cdot), q(\cdot), z(\cdot))\) be the optimal state process corresponding to \(u(\cdot)\) and \(k(\cdot)\) be the unique solution to the adjoint equation (5.6) corresponding to the optimal pair \((u(\cdot); y(\cdot), q(\cdot), z(\cdot))\). By the necessary optimality condition (4.4) and Assumption 5.3, we have

\[
H_u(t, y(t), q(t), z(t), u(t), k(t)) = 0, \quad \text{a.e. a.s.}
\]

Noticing the definition of \(H\) in (5.5), we get

\[
2N(t)u(t) + D^*(t)k(t) = 0, \quad \text{a.e. a.s.}
\]

Then the claim that the unique optimal control \(u(\cdot)\) satisfies (5.7) follows. \(\square\)

Finally, we introduce the so-called stochastic Hamilton system which consists of the state equation
Theorem 5.6. Under Assumptions (d-optimal control of BLQ problem and (Grant No. T200924), Natural Science Foundation of Zhejiang Province (Grant No. Y6110775, Y6110789) and the adjoint equation (5.4) (or equivalently, (5.6)) and the dual representation (5.7):

\[
\begin{aligned}
dy(t) &= -\left( A(t)y(t) + \sum_{i=1}^{d} B^i(t)q^i(t) + \sum_{i=1}^{\infty} C^i(t)z^i(t) + D(t)u(t) \right) dt \\
&\quad + \sum_{i=1}^{\infty} q^i dW^i(t) + \sum_{i=1}^{\infty} z^i dH^i(t), \\
y(T) &= \xi, \\
dk(t) &= (A^*(t)k(t) - 2E(t)y(t)) dt + \sum_{i=1}^{d} (B^{i*}(t)k^i(t) - 2F^i(t)q^i(t)) dW^i(t) \\
&\quad + \sum_{i=1}^{\infty} (C^{i*}(t)k(t) - 2G^i(t)z^i(t)) dH^i(t), \\
k(0) &= -2My(0), \\
u_t &= -\frac{1}{2} N^{-1}(t)D^*(t)k(t). 
\end{aligned}
\]  

(5.8)

Clearly this is a fully coupled forward-backward stochastic differential equation (FBSDE) driven by \( d \)-dimensional Brownian motion \( W \) and Teugel’s martingales \( \{H^i\}_{i=1}^{\infty} \), and its solution is a stochastic processes quaternary \( (k(\cdot), y(\cdot), q(\cdot), z(\cdot)) \).

**Theorem 5.6.** Under Assumptions 5.1–5.3, the stochastic Hamilton system (5.8) has a unique solution \( (k(\cdot), y(\cdot), q(\cdot), z(\cdot)) \in S^2_\mathbb{F}(0,T;\mathbb{R}^n) \times S^2_\mathbb{F}(0,T;\mathbb{R}^n) \times M^2_\mathbb{F}(0,T;\mathbb{R}^{n \times d}) \times l^2_\mathbb{F}(0,T;\mathbb{R}^n) \), where \( u(\cdot) \) is the optimal control of BLQ problem and \( (y(\cdot), q(\cdot), z(\cdot)) \) is its corresponding optimal state. Moreover,

\[
E \sup_{0 \leq t \leq T} |k(t)|^2 + E \sup_{0 \leq t \leq T} |y(t)|^2 + E \int_0^T |q(t)|^2 dt + E \int_0^T \| z(t) \|^2_{\mathbb{L}^2(\mathbb{R}^n)} dt \leq KE(\xi)^2. 
\]

(5.9)

**Proof.** The existence result follows from Theorem 5.5 and the uniqueness result is obvious once a priori estimate (5.9) holds. But noticing Assumptions 5.1–5.3 and using Lemmas 2.2 and 2.4, we can deduce (5.9) immediately.

In summary, the stochastic Hamilton system (5.8) completely characterizes the optimal control of BLQ problem in this section. Therefore, solving the BLQ problem is equivalent to solving the stochastic Hamilton system, and the unique optimal control of the stochastic Hamilton system can be given explicitly by (5.7) if exists.

**Acknowledgements** This work was supported by National Natural Science Foundation of China (Grant No. 11101090, 11101140, 10771122), Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20090071120002), Innovation Team Foundation of the Department of Education of Zhejiang Province (Grant No. T200924), Natural Science Foundation of Zhejiang Province (Grant No. Y6110775, Y6110789) and Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry. Also we are grateful to the referees for their constructive comments.

**References**

1. Bahali K, Eddahbi M, Essaky E. BSDE associated with Lévy processes and application to PDIE. J Appl Math Stoch Anal, 2003, 16: 1–17
2. Bahali K, Gherbal B, Mezerdi B. Existence and optimality conditions in stochastic control of linear BSDEs. Random Oper Stochastic Equations, 2010, 18: 185–197
3. Bahali S. Stochastic controls of backward systems. Random Oper Stochastic Equations, 2010, 18: 125–140
4. Bismut J M. Conjugate convex functions in optimal stochastic control. J Math Anal Appl, 1973, 44: 384–404
5. Dokuchaev N, Zhou X Y. Stochastic controls with terminal contingent conditions. J Math Anal Appl, 1999, 238: 143–165
6. Ekeland I, Témam R. Convex Analysis and Variational Problems. Amsterdam: North-Holland, 1976
7. El Karoui N, Peng S G, Quenez M C. Backward stochastic differential equations in finance. Math Finance, 1997, 7: 1–71
8 El Otmani M. Generalized BSDE driven by a Lévy process. J Appl Math Stoch Anal, 2006, Article ID 85407, 25
9 El Otmani M. Backward stochastic differential equations associated with Lévy processes and partial integro-differential equations. Commun Stoch Anal, 2008, 2: 277–288
10 Lim A E B, Zhou X Y. Linear-quadratic control of backward stochastic differential equations. SIAM J Control Optim, 2001, 40: 450–474
11 Lim A E B, Zhou X Y. Optimal control of linear backward stochastic differential equations with a quadratic cost criterion. Stochastic Theory and Control, Lecture Notes in Control and Information Sciences, vol. 280. New York: Springer, 2002, 301–317
12 Meng Q X, Tang M N. Necessary and sufficient conditions for optimal control of stochastic systems associated with Lévy processes. Sci China Ser F, 2009, 52: 1982–1992
13 Mitsui K, Tabata Y. A stochastic linear-quadratic problem with Lévy processes and its application to finance. Stochastic Process Appl, 2000, 90: 109–122
14 Nualart D, Schoutens W. Chaotic and predictable representation for Lévy processes. Stochastic Process Appl, 2000, 90: 109–122
15 Nualart D, Schoutens W. Backward stochastic differential equations and Feynman-Kac formula for Lévy processes with applications in finance. Bernoulli, 2001, 7: 761–776
16 Pardoux E, Peng S G. Adapted solution of a backward stochastic differential equation. Systems Control Lett, 1990, 14: 55–61
17 Peng S G. Backward stochastic differential equations and applications to optimal control. Appl Math Optim, 1993, 27: 125–144
18 Ren Y, Fan X L. Reflected backward stochastic differential equations driven by a Lévy process. ANZIAM J, 2009, 50: 486–500
19 Ren Y, El Otmani M. Generalized reflected BSDEs driven by a Lévy process and an obstacle problem for PDIEs with a nonlinear Neumann boundary condition. J Comput Appl Math, 2010, 233: 2027–2043
20 Tang H, Wu Z. Stochastic differential equations and stochastic linear quadratic optimal control problem with Lévy processes. J Syst Sci Complex, 2009, 22: 122–136
21 Tang S J, Li X J. Necessary conditions for optimal control of stochastic systems with random jumps. SIAM J Control Optim, 1994, 32: 1447–1475