Optimal Control with State Constraints for Stochastic Evolution Equation with Jumps in Hilbert Space *

Qingxin Meng† QiuHong Shi Maoning Tang

Department of Mathematics, Huzhou University, Zhejiang 313000, China

Abstract

This paper studies a stochastic optimal control problem with state constraint, where the state equation is described by a controlled stochastic evolution equation with jumps in Hilbert Space and the control domain is assumed to be convex. By means of Ekeland variational principle, combining the convex variation method and the duality technique, necessary conditions for optimality are derived in the form of stochastic maximum principles.

Keywords: Stochastic evolution equation; Backward stochastic evolution equation; Stochastic maximum principle; State constraint.

Introduction

In this paper, we study the optimal control for the following stochastic evolution equation with jumps

\[
\begin{aligned}
&dX(t) = [A(t)X(t) + b(t, X(t), u(t))]dt + [B(t)X(t) + g(t, X(t), u(t))]dW(t) \\
&\quad + \int_{E} \sigma(t, c, X(t-), u(t))\tilde{\mu}(de, dt), \\
&X(0) = x, \quad t \in [0, T],
\end{aligned}
\]

(1.1)

with the cost functional

\[
J(u(\cdot)) = \mathbb{E}\left[\int_{0}^{T} l(t, X(t), u(t))dt + \Phi(X(T))\right],
\]

(1.2)

and state constraint

\[
\mathbb{E}[\phi(X(T))] = 0,
\]

(1.3)

in the framework of a Gelfand triple \(V \subset H = H^* \subset V^*\), where \(H\) and \(V\) are two given Hilbert spaces. Here on a given filtrated probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, P)\), \(W\) is a one-dimensional Brownian motion and \(\tilde{\mu}\) is a Poisson random martingale measure on a fixed nonempty Borel measurable subset \(E\) of \(\mathbb{R}^1\), \(A : [0, T] \times \Omega \rightarrow \mathcal{L}(V, V^*)\), \(B : [0, T] \times \Omega \rightarrow \mathcal{L}(V, H)\), \(\sigma : [0, T] \times \Omega \times H \times U_{ad} \rightarrow H\), \(g : [0, T] \times \Omega \times H \times U_{ad} \rightarrow H\) and \(\phi : [0, T] \times \Omega \times E \times H \times U_{ad} \rightarrow H\) are given random mappings, where the control variable \(u\) takes value in a nonempty convex subset \(U_{ad}\) of a real Hilbert space \(U\). Here we denote by \(\mathcal{L}(V, V^*)\) the space of bounded linear transformations of \(V\) into \(V^*\), by \(\mathcal{L}(V, H)\) the space of bounded linear transformations of \(H\) into \(V\).

An adapted solution of (1.1) is a \(V\)-valued, \(\{\mathcal{F}_t\}_{0 \leq t \leq T}\)-adapted process \(X(\cdot)\) which satisfies (1.1) under some appropriate sense. The optimal control problem is to find an admissible control to minimize the cost functional (1.2) over the set of admissible controls.

One of the basic method to solve stochastic optimal control problems is the stochastic maximum principle whose objective is to establish necessary (as well as sufficient) optimality conditions of controls. For optimal control problems of infinite dimensional stochastic systems, many works are concerned with the stochastic systems and the corresponding stochastic maximum principles, see e.g. [9, 5, 14, 2, 1, 4, 3, 10, 8, 6].

In contrast, there have not been a number of results on the optimal control for stochastic partial differential equations driven by jump processes. In 2005, Øksendal, Proske, Zhang [12] studied the optimal control problem of quasilinear semielliptic SPDEs driven by Poisson random measure and gave sufficient maximum principle results, not necessary ones. In 2017, Tang and Meng [13] studied the optimal control problem for a controlled stochastic evolution equation (1.1) with the cost functional (1.2), where the control domain is assumed to be convex. [13] adopt the convex variation method and the first adjoint duality analysis to

*This work was supported by the Natural Science Foundation of Zhejiang Province for Distinguished Young Scholar (No.LR15A010001), and the National Natural Science Foundation of China (No.11471079, 11301177)

†Corresponding author. E-mail: mqx@zjhu.edu.cn
show a necessary maximum principle. And Under the convexity assumption of the Hamiltonian and the terminal cost, a sufficient maximum principle for this optimal problem which is the so-called verification theorem is obtained

The purpose of this paper is to establish the maximum principle for the optimal control problem where the state process is driven by a controlled stochastic evolution equation \([1.1]\) with the cost functional \([1.2]\) and the state constraint \([2.5]\) by Ekland variational principle, combining the convex variation method and the duality technique.

The paper is organized as follows. In section 2 we formulate the problem and give various assumptions used throughout the paper. In section 3, we present a penalized optimal control problem. Section 4 is devoted to derive necessary optimality conditions in the form of stochastic maximum principles in a unified way. Some basic results on the SEE and the BSEE with jump are given in the Appendix which will be used in this paper.

## 2 Problem formulation

In this section, we introduce basic notation and standing assumptions, and state an optimal control problem with state constraint under a stochastic evolution equation with jumps in Hilbert space, which was considered by Tang and Meng [3].

Let \((Ω, ℵ, P)\) be a complete probability space equipped with a one-dimensional standard Brownian motion \(\{W(t), 0 ≤ t ≤ T\}\) and a stationary Poisson point process \(\{η_t\}_{t ≥ 0}\) defined on a fixed nonempty Borel measurable subset \(E\) of \(\mathbb{R}^1\). Denote by \(E[·]\) the expectation under the probability \(P\). We denote by \(μ(de, dt)\) the counting measure induced by \(\{η_t\}_{t ≥ 0}\) and by \(ν(de)\) the corresponding characteristic measure. Then the compensatory random martingale measure is denoted by \(\overline{μ}(de, dt) := μ(de, dt) − ν(de)dt\) which is assumed to be independent of the Brownian motion \(\{W(t), 0 ≤ t ≤ T\}\). Furthermore, we assume that \(ν(E) < ∞\). Let \((ℵ_t)_{0 ≤ t ≤ T}\) be the \(P\)-augmentation of the natural filtration generated by \(\{W_t\}_{t ≥ 0}\) and \(\{η_t\}_{t ≥ 0}\). By \(ℬ\) we denote the predictable σ-field on \(Ω × [0, T]\) and by \(ℬ(Λ)\) the Borel σ-algebra of any topological space \(Λ\). Let \(X\) be a separable Hilbert space with norm \(∥·∥_X\). Denote by \(M^{2,2}(E; X)\) the set of all \(X\)-valued measurable functions \(r = \{r(e), e ∈ E\}\) defined on the measure space \((E, ℬ(E); ν)\) such that \(∥r∥_{M^{2,2}(E; X)} := \sqrt{\int_E ∥r(e)∥_X^2 ν(de)} < ∞\), by \(M^{2,2}_F(0, T; X)\) the set of all \(ℬ_X×ℬ(E)-measurable \(X\)-valued processes \(r = \{r(t, ω, e), (t, ω, e) ∈ [0, T] × Ω × E\}\) such that \(∥r∥_{M^{2,2}_F(0, T; E; X)} := \sqrt{E[\int_0^T ∥r(t, ω)∥^2_X ν(dt)]} < ∞\), by \(M^{2,2}_F(0, T; X)\) the set of all \(ℬ_t\)-adapted \(X\)-valued càdlàg processes \(f = \{f(t, ω), (t, ω) ∈ [0, T] × Ω\}\) such that \(∥f∥_{M^{2,2}_F(0, T; X)} := \sqrt{E[∫_0^T ∥f(t)∥^2_X dt]} < ∞\), by \(S^{2,2}_F(0, T; X)\) the set of all \(ℬ_t\)-adapted \(X\)-valued random variables \(ξ\) on \((Ω, ℵ, P)\) such that \(∥ξ∥_{L^2(Ω, ℬ; ℬ; X)} := \sqrt{E[∥ξ∥^2_X]} < ∞\). Throughout this paper, we let \(C\) and \(K\) be two generic positive constants, which may be different from line to line.

In what follows, we set up a Gelfand triple \((V, H, V^*)\), based on which the state process and the adjoint process is defined. Indeed, the state process is governed by a SEE with jumps, while the adjoint process is governed by a BSEE with jumps. We provide the existence, uniqueness and continuous dependence theorems for SEEs with jumps and BSEEs with jumps in the appendix.

Let \(V\) and \(H\) be two separable (real) Hilbert spaces such that \(V\) is densely embedded in \(H\). We identify \(H\) with its dual space by the Riesz mapping. Then we can take \(H\) as a pivot space and get a Gelfand triple \((V, H, V^*)\) such that \(V ⊂ H = H^* ⊂ V^*\). Let \((·, ·)_H\) denote the inner product in \(H\), and \((·, ·)^*\) denote the duality product between \(V\) and \(V^*\). Moreover, we write \(L(V, V^*)\) for the space of bounded linear transformations of \(V\) into \(V^*\).

The state process is governed by the following controlled SEE with jumps in the Gelfand triple \((V, H, V^*)\):

\[
\begin{aligned}
dX(t) &= [A(t)X(t) + b(t, X(t), u(t)) + H(t)]dt + [B(t)X(t) + g(t, X(t), u(t))]dW(t) \\
&+ \int_E σ(t, e, X(t), u(t))\overline{μ}(de, dt), \\
X(0) &= x, \quad t ∈ [0, T],
\end{aligned}
\]

(2.1)

where the space of controls \(U_{ad}\) is given by a nonempty closed convex subset of a separable real Hilbert space \(U\).

**Definition 2.1.** A stochastic process \(u(·)\) is an admissible control, if \(u(t) ∈ U_{ad}\) for almost \(t ∈ [0, T]\) and \(u(·) ∈ M^2_F(0, T; U)\). The set of all admissible controls is denoted by \(A\).

The cost functional is given by

\[
J(u(·)) = E[∫_0^T l(t, x(t), u(t))dt + Φ(x(T))].
\]

(2.2)
We assume that the control system \( X(t, x) \) is subject to the following state constraint

\[
E[\phi(X(T))] = 0.
\]

(2.3)

Here the coefficients \( (A, B, b, g, \sigma, l, \Phi, \phi) \) of the control system \( X(t, x) \) are

**Assumption 2.1.**

(i) The operator processes \( A : [0, T] \times \Omega \rightarrow \mathcal{L}(V, V^*) \) and \( B : [0, T] \times \Omega \rightarrow \mathcal{L}(V, H) \) are weakly predictable; i.e., \( \langle A(\cdot, x, y) \rangle \) and \( \langle B(\cdot, x, y) \rangle \) are both predictable process for every \( x, y \in V \), and satisfy the coercive condition, i.e., there exist some constants \( C, \alpha > 0 \) and \( \lambda \) such that for any \( x \in V \) and each \( (t, \omega) \in [0, T] \times \Omega \),

\[
-\langle A(t, x, x) \rangle + \lambda \|x\|_H^2 \geq \alpha \|x\|_{V^*}^2 + \|Bx\|_H^2,
\]

and

\[
\sup_{(t, \omega) \in [0, T] \times \Omega} \|A(t, \omega)\|_{\mathcal{L}(V, V^*)} + \sup_{(t, \omega) \in [0, T] \times \Omega} \|B(t, \omega)\|_{\mathcal{L}(V, H)} \leq C.
\]

(2.4)

(2.5)

(ii) \( b, g : [0, T] \times \Omega \times H \times \mathcal{U} \rightarrow H \) are \( \mathcal{P} \times \mathcal{B}(H) \times \mathcal{B}(\mathcal{U})/\mathcal{B}(H) \)-measurable mappings and \( \sigma : [0, T] \times \Omega \times E \times H \times \mathcal{U} \rightarrow H \) is a \( \mathcal{P} \times \mathcal{B}(E) \times \mathcal{B}(H) \times \mathcal{B}(\mathcal{U})/\mathcal{B}(H) \)-measurable mapping such that \( b(\cdot, 0, 0), g(\cdot, 0, 0) \in M^2_2([0, T]; H), \sigma(\cdot, \cdot, 0, 0) \in M^2_2([0, T] \times E; H) \). Moreover, for almost all \( (t, \omega, e) \in [0, T] \times \Omega \times E, b, g \) and \( \sigma \) are Gateaux differentiable in \( (x, u) \) with continuous bounded Gateaux derivatives \( b_x, g_x, \sigma_x, b_u, g_u \) and \( \sigma_u \);

(iii) \( I : [0, T] \times \Omega \times H \times \mathcal{U} \rightarrow \mathbb{R} \) is a \( \mathcal{P} \otimes \mathcal{B}(H) \otimes \mathcal{B}(\mathcal{U})/\mathcal{B}(\mathbb{R}) \)-measurable mapping and \( \Phi, \phi : \Omega \times H \rightarrow \mathbb{R} \) is a \( \mathcal{F}_T \otimes \mathcal{B}(H)/\mathcal{B}(\mathbb{R}) \)-measurable mapping. For almost all \( (t, \omega) \in [0, T] \times \Omega \), \( I \) is continuous Gateaux differentiable in \( (x, u) \) with continuous Gateaux derivatives \( I_x \) and \( I_u \), and \( \Phi \) and \( \phi \) are Gateaux differentiable in \( x \) with continuous Gateaux derivative \( \Phi_x \) and \( \phi_x \). Moreover, for almost all \( (t, \omega) \in [0, T] \times \Omega \), there exists a constant \( C \) such that for all \( (x, u) \in H \times \mathcal{U} \)

\[
|I(t, x, u)| \leq C(1 + \|x\|_H^2 + \|u\|_U^2),
\]

\[
\|I_x(t, x, u)\|_H + \|I_u(t, x, u)\|_U \leq C(1 + \|x\|_H + \|u\|_U),
\]

and

\[
|\Phi(x)| \leq C(1 + \|x\|_H^2), \quad |\phi(x)| \leq C(1 + \|x\|_H^2), \quad \|\Phi_x(x)\|_H \leq C(1 + \|x\|_H), \quad \|\phi_x(x)\|_H \leq C(1 + \|x\|_H).
\]

Under Assumption 2.1 it can be shown from Lemma A.6 that for any \( u(\cdot) \in A \), the state equation (2.1) admits a unique solution \( X(\cdot) \in M^2_2(0, T; V) \cap S^2_2(0, T; H) \). We also denote this solution as \( X^u(\cdot) \). Then we call \( X(\cdot) \) the state process corresponding to the control process \( u(\cdot) \) and \( (u(\cdot), X(\cdot)) \) the admissible pair. Furthermore, from Assumption 2.1 and the a priori estimate (A.7), we can easily validate that

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

Now we state formally the optimal control problem

**Problem 2.1.** Find an admissible control \( \bar{u}(\cdot) \) such that

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

subject to (2.1) and (2.3), where the cost functional is given by (2.2).

Any \( \bar{u}(\cdot) \in A \) satisfying the above is called an optimal control process of Problem 2.1, the corresponding state process \( \bar{X}(\cdot) \) is called an optimal state process; correspondingly, \( (\bar{u}(\cdot), \bar{X}(\cdot)) \) is called an optimal pair of Problem 2.1.

### 3 Penalized optimal control problem

In this section, we relate the original constrained control problem with one without state constraint.

The results relies on the following Ekeland’s principle.
Lemma 3.1 (Ekeland’s principle, [7]). Let $(S,d)$ be a complete metric space and $ho(\cdot) : S \to \mathbb{R}$ be lower-semicontinuous and bounded from below. For $\varepsilon \geq 0$, suppose $u^\varepsilon \in S$ satisfies

$$ \rho(u^\varepsilon) \leq \inf_{u \in S} \rho(u) + \varepsilon. $$

Then for any $\lambda > 0$, there exists $u^\lambda \in S$ such that

$$ \rho(u^\lambda) \leq \rho(u^\varepsilon), \quad d(u^\lambda, u^\varepsilon) \leq \lambda, $$

and

$$ \rho(u^\lambda) \leq \rho(u) + \frac{\varepsilon}{\lambda} d(u^\lambda, u), \text{ for all } u \in S. $$

Define a metric $d$ on the admissible controls set $A$ as

$$ d(u_1(\cdot), u_2(\cdot)) \triangleq \left\{ \mathbb{E} \left[ \int_0^T ||u_1(t) - u_2(t)||_2^2 \, dt \right] \right\}^{\frac{1}{2}}, \quad \forall u_1(\cdot), u_2(\cdot) \in A. \tag{3.1} $$

We can assume that $A$ is a bounded closed convex set in the sense of (3.1), the unbounded case can be reduced to the bounded case.

Under this assumption of boundedness and closedness of $A$, we have the following basic lemma which will be used in the sequence.

Lemma 3.2. $(\Lambda, d)$ is a complete metric space.

Proof. Since the control space $U$ is a Hilbert space $M_2^2(0, T; U)$ is also a Hilbert space under (3.1). Therefore, $A$ is complete under the distance defined by (3.1), since $A$ is a closed subset of $M_2^2(0, T; U)$. The proof is complete. \hfill $\Box$

The next lemma shows that a mapping from the control process in $A$ to the state process in $M_2^2(0, T)$, to be defined below, is bounded and continuous. To simplify our notation, we write

$$ M_2^2(0, T) \defeq \mathcal{S}_2^2(0, T; H) \cap M_2^2(0, T; V) \tag{3.2} $$

and

$$ ||X(\cdot)||_{M_2^2(0, T)} \defeq \sqrt{||X(\cdot)||_{\mathcal{S}_2^2(0, T; H)}^2 + ||X(\cdot)||_{M_2^2(0, T; V)}^2}. \tag{3.3} $$

The next lemma shows that a mapping from the control process in $A$ to the state process in $M_2^2(0, T; V)$ is bounded and continuous.

Lemma 3.3. Let Assumption 2.1 be satisfied. Then the mapping $I : (A, d) \to (M_2^2(0, T), ||\cdot||_{M_2^2(0, T)})$ defined by

$$ I(u(\cdot)) = X^u(\cdot) $$

is bounded and continuous.

Proof. By the a priori estimate of SEE (Lemma A.7), it can be shown that for any $u(\cdot) \in A$,

$$ ||X^u(\cdot)||_{M_2^2(0, T)}^2 \leq K \left\{ \mathbb{E}[||x||_{H}^2] + \mathbb{E} \left[ \int_0^T ||u(t)||_2^2 \, dt \right] + 1 \right\} \leq K. \tag{3.4} $$

Here $K$ is a positive constant independent of $u(\cdot)$ and may change from line to line.

On the other hand, let $\{v_n(\cdot)\}_{n \geq 1}$ be a sequence in $A$ such that it converges an admissible $v(\cdot) \in A$ under the metric $d$. Suppose that $X_n(\cdot)$, for each $n = 1, 2, \cdots$, and $X(\cdot)$ are the state processes corresponding to $v_n(\cdot)$ and $v(\cdot)$, respectively. By making use of the a priori estimate of SEE (Lemma A.7), we can deduce that

$$ ||X^{v_n}(\cdot) - X^v(\cdot)||_{M_2^2(0, T)}^2 \tag{3.5} $$

$$ \leq K \left\{ \mathbb{E} \left[ \int_0^T ||b(t, X^v(t), v_n(t)) - b(t, X^v(t), v(t))||_H^2 \, dt \right] + \mathbb{E} \left[ \int_0^T ||g(t, X^v(t), v_n(t)) - g(t, X^v(t), v(t))||_H^2 \, dt \right] \\
+ \mathbb{E} \left[ \int_0^T ||\sigma(t, X^v(t), v_n(t)) - \sigma(t, X^v(t), v(t))||_{M_2^{1, 2}(0, T)}^2 \, dt \right] \right\} $$
dominated convergence theorem, we obtain

\[ \lim_{n \to \infty} J(v_n) = J(v) \]

Sending \( n \to \infty \) in (3.6) yields

\[ \|X^{v_n}(\cdot) - X^v(\cdot)\|_{M_d^2(0,T)}^2 \to 0. \]  

This validates the continuity of \( \mathcal{I} \).

**Lemma 3.4.** Let Assumption 2.1 be satisfied. Then the cost functional \( J(u(\cdot)) \) is bounded and continuous on \( \mathcal{A} \) under the metric (3.1).

**Proof.** For any \( u(\cdot) \in \mathcal{A} \), under Assumption 2.1 and from Lemma 3.3, we have

\[
|J(u(\cdot))| \leq \mathbb{E} \left[ \int_0^T |l(t, X^n(t), u(t))| dt + |\phi(X^n(T))| \right] 
\leq K \left[ 1 + \|X^n(\cdot)\|_{M^2(0,T;V)}^2 + \|u(\cdot)\|_{M^2(0,T;U)}^2 + \|X(T)\|_{L^2(\Omega, \mathcal{F}_T, \mathbb{P}; H)}^2 \right] 
\leq K \left[ 1 + \|X^n(\cdot)\|_{M^2(0,T)}^2 + \|u(\cdot)\|_{M^2(0,T;U)}^2 \right] 
\leq K. \] 

(3.8)

Here \( K \) is a positive constant independent of \( u(\cdot) \) and may change from line to line. This implies the cost functional \( J(u(\cdot)) \) is bounded on \( \mathcal{A} \).

To show the continuity of the cost functional, as in the proof of Lemma 3.3, we pick up the sequence \( \{v_n(\cdot)\}_{n \geq 1} \) and its converging point \( v(\cdot) \) in \( \mathcal{A} \) as well as the corresponding state processes \( X_n(\cdot) \) and \( X(\cdot) \). Thus using Lemma 3.3 and the Lebesgue dominated convergence theorem, we obtain

\[ J(v_n(\cdot)) \to J(v(\cdot)), \quad \text{as } n \to \infty. \]  

(3.9)

The completes the proof.

Define a penalized cost functional associated with Problem (2.1) as

\[
J^\varepsilon(v(\cdot)) \triangleq \left\{ \left[ J(v(\cdot)) - J(\bar{u}(\cdot)) + \varepsilon^2 + \mathbb{E}[\phi(X^v(T))]^2 \right] \right\}^\frac{1}{2}, \quad \forall \varepsilon > 0. \] 

(3.10)

It is worthwhile to point out that we will study this functional over \( \mathcal{A} \).

**Lemma 3.5.** \( J^\varepsilon(v(\cdot)) \) is bounded and continuous on \( \mathcal{A} \) under the metric (3.1).

**Proof.** The proof can be obtained by Lemma 3.4 and Lemma 3.3 immediately.

Now we introduce an auxiliary optimal control problem without state constraint:

**Problem 3.6 ((SC)\*).** Find an admissible control such that

\[
\inf_{v(\cdot) \in \mathcal{A}} J^\varepsilon(v(\cdot)), \]  

(3.11)

where the state process is given by (2.1) and the cost functional \( J^\varepsilon(v(\cdot)) \) is given by (A.7).

From the definition of the penalized cost functional (3.10), we see that

\[ J^\varepsilon(\bar{u}(\cdot)) = \varepsilon \leq \inf_{v(\cdot) \in \mathcal{A}} J^\varepsilon(v(\cdot)) + \varepsilon. \] 

(3.12)

An application of Ekeland’s variational principle shows that there is a \( u^\varepsilon(\cdot) \in \mathcal{A} \) such that

\[
\begin{aligned}
 J^\varepsilon(u^\varepsilon(\cdot)) &\leq J^\varepsilon(\bar{u}(\cdot)) = \varepsilon, \\
 d(u^\varepsilon(\cdot), \bar{u}(\cdot)) &\leq \varepsilon^\frac{1}{2}, \\
 J^\varepsilon(v(\cdot)) - J^\varepsilon(u^\varepsilon(\cdot)) &\geq -\varepsilon^\frac{1}{2} d(u^\varepsilon(\cdot), v(\cdot)), \quad \forall v(\cdot) \in \mathcal{A}. 
\end{aligned} \] 

(3.13)
Define a convex perturbed control of \( \varepsilon(\cdot) \)

\[
\varepsilon^{\varepsilon, \rho}(\cdot) \triangleq \varepsilon(\cdot) + \rho(u(\cdot) - \varepsilon^{\varepsilon}(\cdot)),
\]

where \( u(\cdot) \) is an arbitrary admissible control in \( \mathcal{A} \) and \( 0 \leq \rho \leq 1 \). It is easy to verify that \( \varepsilon^{\varepsilon, \rho}(\cdot) \) is also in \( \mathcal{A} \). Suppose that \( X^{\varepsilon, \rho}(\cdot) \) and \( X^{\varepsilon}(\cdot) \) are the state processes corresponding to \( \varepsilon^{\varepsilon, \rho}(\cdot) \) and \( \varepsilon(\cdot) \), respectively. By (3.13) and the fact

\[
d(u^{\varepsilon, \rho}(\cdot), u^{\varepsilon}(\cdot)) \leq C\rho,
\]

we have

\[
J^{\varepsilon}(u^{\varepsilon, \rho}(\cdot)) - J^{\varepsilon}(u^{\varepsilon}(\cdot)) \geq -\varepsilon\frac{2}{\rho} d(u^{\varepsilon, \rho}(t), u^{\varepsilon}(t)) \geq -\varepsilon\frac{2}{\rho} C\rho.
\]

On the other hand, from the definition of \( J^{\varepsilon}(\bar{u}(\cdot)) \), we have

\[
J^{\varepsilon}(u^{\varepsilon, \rho}(\cdot)) - J^{\varepsilon}(u^{\varepsilon}(\cdot)) = \frac{\left[ J^{\varepsilon}(u^{\varepsilon, \rho}(\cdot)) \right]^2 - \left[ J^{\varepsilon}(u^{\varepsilon}(\cdot)) \right]^2}{J^{\varepsilon}(u^{\varepsilon, \rho}(\cdot)) + J^{\varepsilon}(u^{\varepsilon}(\cdot))}
\]

\[
= \frac{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot)) - 2J(\bar{u}(\cdot)) + 2\varepsilon}{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot))} \times \left[ J(u^{\varepsilon, \rho}(\cdot)) - J(u^{\varepsilon}(\cdot)) \right]
\]

\[
+ \frac{\mathbb{E}[\phi(X^{\varepsilon, \rho}(T))] + \mathbb{E}[\phi(X^{\varepsilon}(T))]}{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot))} \times \left\{ \mathbb{E}[\phi(X^{\varepsilon, \rho}(T))] - \mathbb{E}[\phi(X^{\varepsilon}(T))] \right\}
\]

\[
= \lambda^{\varepsilon, \rho} [J(u^{\varepsilon, \rho}(\cdot)) - J(u^{\varepsilon}(\cdot))] + \mu^{\varepsilon, \rho} \left\{ \mathbb{E}[\phi(X^{\varepsilon, \rho}(T))] - \mathbb{E}[\phi(X^{\varepsilon}(T))] \right\},
\]

where

\[
\lambda^{\varepsilon, \rho} \triangleq \frac{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot)) - 2J(\bar{u}(\cdot)) + 2\varepsilon}{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot))}
\]

and

\[
\mu^{\varepsilon, \rho} \triangleq \frac{\mathbb{E}[\phi(X^{\varepsilon, \rho}(T))] + \mathbb{E}[\phi(X^{\varepsilon}(T))]}{J(u^{\varepsilon, \rho}(\cdot)) + J(u^{\varepsilon}(\cdot))}.
\]

From (4.1), we have

\[
\lim_{\rho \to 0} d(u^{\varepsilon, \rho}(\cdot), u^{\varepsilon}(\cdot)) = 0
\]

Then it follows from Lemma 3.3 and Lemma 3.5 that

\[
\lim_{\rho \to 0} ||X^{\varepsilon, \rho}(\cdot) - X^{\varepsilon}(\cdot)||_{\mathcal{M}^2_{\varepsilon}(0,T)} = 0
\]

and

\[
\lim_{\rho \to 0} J^{\varepsilon}(u^{\varepsilon, \rho}(\cdot)) = J^{\varepsilon}(u^{\varepsilon}(\cdot)).
\]

Consequently,

\[
\lim_{\rho \to 0} \lambda^{\varepsilon, \rho} = \lambda^{\varepsilon}, \quad \lim_{\rho \to 0} \mu^{\varepsilon, \rho} = \mu^{\varepsilon},
\]

where

\[
\lambda^{\varepsilon} \triangleq \frac{J(u^{\varepsilon}(\cdot)) - J(\bar{u}(\cdot)) + \varepsilon}{J(u^{\varepsilon}(\cdot))}
\]

and

\[
\mu^{\varepsilon} \triangleq \frac{\mathbb{E}[\phi(X^{\varepsilon}(0))] + \mathbb{E}[\phi(X^{\varepsilon}(T))]}{J(u^{\varepsilon}(\cdot))}
\]

Note that

\[
|\lambda|^2 + |\mu|^2 = 1.
\]

Therefore, there exists a subsequence \( \{ (\lambda^\varepsilon, \mu^\varepsilon) \}_{\varepsilon > 0} \) such that

\[
\lim_{\varepsilon \to 0} \lambda^\varepsilon = \lambda, \quad \lim_{\varepsilon \to 0} \mu^\varepsilon = \mu,
\]

and

\[
|\lambda|^2 + |\mu|^2 = 1.
\]
4 Stochastic Maximum Principle

In this section, we first derive a variational formula for the penalized cost functional $J^\varepsilon(u(\cdot))$. To simplify our notation, we write partial derivatives of $b, g, \sigma$, and $l$ as

$$
\varphi^\varepsilon(t) \equiv \varphi(t, X^\varepsilon(t), u^\varepsilon(t)),
$$

$$
\varphi^x(t) \equiv \varphi_a(t, X^\varepsilon(t), u^\varepsilon(t)),
$$

$$
\varphi^\varepsilon_x(t) \equiv \varphi_a(t, X^\varepsilon(t), \bar{u}(t)),
$$

where $\varphi = b, g, \sigma$, and $l$.

Define the Hamiltonian $\mathcal{H} : [0, T] \times \Omega \times H \times \mathcal{W} \times H \times \mathcal{M}^{\nu/2}(E; H) \times \mathbb{R} \rightarrow \mathbb{R}$ by

$$
\mathcal{H}(t, x, u, p, q, r(\cdot), \lambda) := (b(t, x, u), p)_H + (g(t, x, u), q)_H + \int_E (\sigma(t, x, u, r(t, e)))_H \nu(de) + \lambda(t, x, u).
$$

(4.1)

Using Hamiltonian $\mathcal{H}$, the adjoint equation (4.3) can be written in the following form:

$$
\begin{cases}
\begin{aligned}
dp^\varepsilon(t) &= -A^*(t)p^\varepsilon(t) + B(t)q^\varepsilon(t) + \mathcal{H}^\varepsilon_x(t) dt + q^\varepsilon(t) dW(t) + \int_E \hat{r}(t, e) \mu(de, dt), \\
p^\varepsilon(T) &= \Phi(X(T)),
\end{aligned}
\end{cases}
$$

(4.2)

where we denote

$$
\mathcal{H}(t) \equiv \mathcal{H}(t, \bar{x}(t), \bar{u}(t), \bar{p}(t), \bar{q}(t), \bar{r}(t, \cdot)).
$$

(4.3)

Similarly, for notational simplify, we write partial derivatives of $H$ as

$$
\mathcal{H}_x^\varepsilon(t) \equiv \mathcal{H}_x(t, X^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), \lambda^\varepsilon(t)),
$$

$$
\mathcal{H}_a^\varepsilon(t) \equiv \mathcal{H}_a(t, X^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), \lambda^\varepsilon(t)),
$$

where $a = x$ or $u$.

For the admissible pair $(u^\varepsilon(\cdot); X^\varepsilon(\cdot))$ and $(u^\epsilon(\cdot); X^\epsilon(\cdot))$ and the optimal pair $(\bar{u}(\cdot); \bar{X}(\cdot))$, the corresponding adjoint processes are denoted by $\{p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), 0 \leq t \leq T\}$, $\{p^\epsilon(t), q^\epsilon(t), r^\epsilon(t, \cdot), 0 \leq t \leq T\}$ and $\{\bar{p}(t), \bar{q}(t), \bar{r}(t, \cdot), 0 \leq t \leq T\}$.

We now define the adjoint equations for $(p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), 0 \leq t \leq T)$, $(p^\epsilon(t), q^\epsilon(t), r^\epsilon(t, \cdot), 0 \leq t \leq T)$ and $\{\bar{p}(t), \bar{q}(t), \bar{r}(t, \cdot), 0 \leq t \leq T\}$ as

$$
\begin{cases}
\begin{aligned}
dp^\varepsilon(t) &= -A^*(t)p^\varepsilon(t) + B(t)q^\varepsilon(t) + \mathcal{H}^\varepsilon_x(t) dt + q^\varepsilon(t) dW(t) + \int_E \hat{r}(t, e) \mu(de, dt), \\
p^\varepsilon(T) &= \lambda^\varepsilon(\Phi(X^\varepsilon(T))) + \mu^\varepsilon\phi_y(X^\varepsilon(T)),
\end{aligned}
\end{cases}
$$

(4.4)

$$
\begin{cases}
\begin{aligned}
dp^\epsilon(t) &= -A^*(t)p^\epsilon(t) + B(t)q^\epsilon(t) + \mathcal{H}^\epsilon_x(t) dt + q^\epsilon(t) dW(t) + \int_E \hat{r}(t, e) \mu(de, dt), \\
p^\epsilon(T) &= \lambda^\epsilon(\Phi_x(X^\epsilon(T))) + \mu\phi_x(X^\epsilon(T)),
\end{aligned}
\end{cases}
$$

(4.5)

and

$$
\begin{cases}
\begin{aligned}
dp^\varepsilon(t) &= -A^*(t)p^\varepsilon(t) + B(t)q^\varepsilon(t) + \mathcal{H}^\varepsilon_x(t) dt + q^\varepsilon(t) dW(t) + \int_E \hat{r}(t, e) \mu(de, dt), \\
p^\varepsilon(T) &= \lambda\Phi_x(\bar{X}(T)) + \mu\phi_x(\bar{X}(T)),
\end{aligned}
\end{cases}
$$

(4.6)

respectively. In fact, the adjoint equations (4.4), (4.5) and (4.6) are three linear BSEEs satisfying Assumptions A.3 and A.4.

Hence by Lemma 3.3, it is easy to check that these three adjoint equations have unique solutions, respectively.

Lemma 4.1. Under Assumptions 2.1 the following convergence results hold

$$
\lim_{\rho \to 0} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \| p^\varepsilon(t) - \bar{p}(t) \|^2_H \right] + \mathbb{E} \left[ \int_0^T \| p^\varepsilon(t) - \bar{p}(t) \|^2 dt \right] + \mathbb{E} \left[ \int_0^T \| q^\varepsilon(t) - \bar{q}(t) \|^2_H dt \right] + \mathbb{E} \left[ \int_0^T \| r^\varepsilon(t, \cdot) - r^\epsilon(t, \cdot) \|^2_{M^{\nu/2}(E; H)} dt \right] = 0,
$$

(4.7)
Proof. By the continuous dependence theorem of BSEE (i.e., Lemma A.9), we derive

\[
\begin{align*}
\mathbb{E} \left[ \sup_{0 \leq t \leq T} \|p^\varepsilon(t) - p^\varepsilon(t)\|_H^2 \right] + \mathbb{E} \left[ \int_0^T \|p^\varepsilon(t) - p^\varepsilon(t)\|_V^2 dt + \mathbb{E} \left[ \int_0^T \|q^\varepsilon(t) - q^\varepsilon(t)\|_H^2 dt \right] \\
+ \mathbb{E} \left[ \int_0^T \|r^\varepsilon(t, \cdot) - \tilde{r}(t, \cdot)\|_{M^2_{\varepsilon, \rho}([0, T] \times E; H)}^2 dt \right] = 0
\end{align*}
\]  
(4.8)

Then using (3.21) and (3.23) gives the desired result (4.7). The proof of (4.8) is similar and omitted here. \(\square\)

In the next lemma, we give a representation of the difference \(J^\varepsilon(u^\varepsilon(\cdot)) - J^\varepsilon(u^\varepsilon(\cdot))\) in terms of the Hamiltonian \(H\), the adjoint process \((p^\varepsilon(\cdot), q^\varepsilon(\cdot), r^\varepsilon(\cdot, \cdot))\) and other relevant expressions associated with the admissible pair \((u^\varepsilon(\cdot), X^\varepsilon(\cdot))\).

**Lemma 4.2.** Under Assumptions (2.7) it holds

\[
J^\varepsilon(u^\varepsilon(\cdot)) - J^\varepsilon(u^\varepsilon(\cdot)) = \mathbb{E} \left[ \int_0^T \left\{ \mathcal{H}^\varepsilon(t) - \mathcal{H}(t, X^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), \lambda^\varepsilon) \\
- \lambda^\varepsilon \mathcal{P}^\varepsilon_T(t) \cdot (X^\varepsilon(t) - X^0) \right\} dt \right] + \lambda^\varepsilon \mathbb{E} \left[ \phi^\varepsilon(T) - \Phi^\varepsilon(T) \right].
\]  
(4.9)

**Proof.** From the definition of the Hamiltonian \(\mathcal{H}\) and \(J^\varepsilon(u(\cdot))\) (see (3.17)), we deduce

\[
J^\varepsilon(u^\varepsilon(\cdot)) - J^\varepsilon(u^\varepsilon(\cdot)) = \lambda^\varepsilon \mathbb{E} [J(u^\varepsilon(\cdot)) - J(u^\varepsilon(\cdot))] + \mu^\varepsilon \mathbb{E} [\phi^\varepsilon(T) - \phi^\varepsilon(T)]
\]

\[
= \mathbb{E} \left[ \int_0^T \left\{ \mathcal{H}^\varepsilon(t) - \mathcal{H}(t, X^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t), r^\varepsilon(t, \cdot), \lambda^\varepsilon) \\
- (p^\varepsilon(t, b^\varepsilon(t)) - b^\varepsilon(t))H - (q^\varepsilon(t, g^\varepsilon(t)) - g^\varepsilon(t))H \right\} dt \right] + \lambda^\varepsilon \mathbb{E} [\phi^\varepsilon(T) - \Phi^\varepsilon(T)].
\]  
(4.10)

On the other hand,

\[
\begin{align*}
d(X^\varepsilon(t) - X^0) &= [A(t)(X^\varepsilon(t) - X^0) + B(t)(X^\varepsilon(t), u^\varepsilon(t)) - b(t, X^\varepsilon(t), u^\varepsilon(t))] dt \\
&\quad + [B(t)(X^\varepsilon(t) - X^0) + g(t, X^\varepsilon(t), u^\varepsilon(t)) - g(t, X^\varepsilon(t), u^\varepsilon(t))] dw(t) \\
&\quad + \int_E \left[ \sigma(t, c, X^\varepsilon(t), u^\varepsilon(t)) - \sigma(t, c, X^\varepsilon(t), u^\varepsilon(t)) \right] \mu(c) dt,
\end{align*}
\]  
(4.11)

Then applying Itô formula to \((p^\varepsilon(t), X^\varepsilon(t) - X^0)\) gives

\[
\begin{align*}
&\mathbb{E} \left[ \int_0^T \left\{ (p^\varepsilon(t), b^\varepsilon(t))_H + (q^\varepsilon(t), g^\varepsilon(t))_H - \sigma^\varepsilon(t, c, X^\varepsilon(t))_H \right\} dt \right] \\
&\quad + \lambda^\varepsilon \mathbb{E} [\phi^\varepsilon(T) - \Phi^\varepsilon(T)].
\end{align*}
\]  
(4.12)

Putting (4.12) into (4.11) leads to the desired representation (4.9). \(\square\)
We have the following basic Lemma.

**Lemma 4.3.** Under Assumptions \[4.7\], it follows that
\[
\|X^{\varepsilon,\rho}(-) - X^{\varepsilon}(-)\|_{M_{\mathcal{F}}^2(0,T)}^2 = O(\rho^2), \tag{4.13}
\]
and
\[
\|X^{\varepsilon}(-) - \bar{X}(-)\|_{M_{\mathcal{F}}^2(0,T)}^2 = O(\varepsilon^2). \tag{4.14}
\]

**Proof.** By the continuous dependence theorem of BSEE (Lemma \[A.9\]) and the uniform boundedness of the Gâteaux derivative \(b_u\), we have
\[
\|X^{\varepsilon,\rho}(-) - X^{\varepsilon}(-)\|_{M_{\mathcal{F}}^2(0,T)}^2
\leq K \mathbb{E} \left[ \int_0^T \left\{ \|b(t, X^{\varepsilon}(t), u^{\varepsilon,\rho}(t)) - b^{\varepsilon}(t)\|_{L_2}^2 dt + \|g(t, y^{\varepsilon}(t), z^{\varepsilon}(t), u^{\varepsilon,\rho}(t)) - g^{\varepsilon}(t)\|_{H}^2 \right\} \nu(de) \right] dt
\leq K \rho^2 \mathbb{E} \left[ \int_0^T \|v(t) - u^{\varepsilon}(t)\|_{L_2}^2 dt \right]
= K \rho^2 \mathbb{E} \left[ \int_0^T \|v(t) - u^{\varepsilon}(t)\|_{L_2}^2 dt \right]
\leq K \rho^2
= O(\rho^2).
\]

Here \(K\) is a generic positive constant and might change from line to line.

In the same vein, we deduce
\[
\|X^{\varepsilon}(-) - \bar{X}(-)\|_{M_{\mathcal{F}}^2(0,T)}^2
\leq K \mathbb{E} \left[ \int_0^T \|u^{\varepsilon}(t) - \bar{u}(t)\|_{L_2}^2 dt \right]
= K d^2(u^{\varepsilon}(t), \bar{u}(t))
\leq K \varepsilon^2
= O(\varepsilon).
\]

The proof is complete. \(\Box\)

Now we state the variational formula for the cost functional \(J^\varepsilon(-)\).

**Theorem 4.4.** Under Assumptions \(2.1\), it follows that for any admissible control \(v(-)\), the cost functional \(J(u(-))\) is Gâteaux differentiable at \(u^\varepsilon(-)\) in the direction \(v(-) - u^\varepsilon(-)\) and the corresponding Gâteaux derivative \(J'\) is given by
\[
\frac{d}{d\rho} J^\varepsilon(u^\varepsilon(-) + \rho(v(-) - u^\varepsilon(-)))_{\rho=0} = \lim_{\rho \to 0} \frac{J^\varepsilon(u^\varepsilon(-) + \rho(v(-) - u^\varepsilon(-))) - J^\varepsilon(u^\varepsilon(-))}{\rho}
= \mathbb{E} \left[ \int_0^T (\mathcal{H}^\varepsilon(v(t) - u^\varepsilon(t))_{L_2} dt \right]
\geq -C \varepsilon^2. \tag{4.15}
\]

Here \(\rho > 0\) is a sufficiently small positive constant.

**Proof.** By \(1.9\), we have
\[
J^\varepsilon(u^\varepsilon(-) + \rho(v(-) - u^\varepsilon(-))) - J^\varepsilon(u^\varepsilon(-)) = I + II, \tag{4.16}
\]
where
\[
I \triangleq \mathbb{E} \left[ \int_0^T \left\{ \mathcal{H}^\varepsilon \rho(t) - \mathcal{H}(t, X^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t), y^\varepsilon(t), z^\varepsilon(t), X^\varepsilon) 
- \mathcal{H}^\varepsilon \rho(t) \cdot (X^\varepsilon(t) - X^\varepsilon(t)) - \mathcal{H}^\varepsilon \rho(t) \cdot (u^\varepsilon(t) - u^\varepsilon(t)) \right\} dt \right]
\]
Here convergence theorem, we conclude that 
\[ \varepsilon \]
From 4.1, Lemma 4.3 and Assumption 2.1 and (3.27), sending which implies that (4.20) holds. This completes the proof.

\[ \text{Proof.} \]
From (3.28), there exists a pair \((A, B, b, g)\) and \((\lambda, \mu(t))\) satisfying the following standard assumptions.

\[ \begin{aligned}
&\text{(A.1)}
\end{aligned} \]

Now we are ready to give the necessary condition of optimality for the existence of the optimal control of Problem 2.1.

**Theorem 4.5.** Let Assumptions 2.1 be satisfied. Let \((\tilde{u}(\cdot); \tilde{X}(\cdot))\) be an optimal pair of Problem 2.1. Then there exist a \((\lambda, \mu)\) satisfying \(|\lambda|^2 + |\mu|^2 = 1\) such that
\[ \forall u \in U_{ad}, \text{ a.e. a.s.} \]

\[ \begin{aligned}
&\text{Here} \{\tilde{p}(t), \tilde{q}(t), \tilde{r}(t, \cdot)\}, 0 \leq t \leq T \text{ be the solution of the corresponding adjoint equation (4.17) associated with} \ ((\tilde{u}(\cdot); \tilde{X}(\cdot)).
\end{aligned} \]

**Proof.** From (3.28), there exists a pair \((\lambda, \mu)\) satisfying \(|\lambda|^2 + |\mu|^2 = 1\). Note that
\[ \begin{aligned}
&\text{(4.21)}
\end{aligned} \]

From 4.1, Lemma 4.3 and Assumption 2.1 and (3.27), sending \(\varepsilon\) to 0 on the both sides of (4.15) and using the dominated convergence theorem, we conclude that
\[ \begin{aligned}
&\text{(A.1)}
\end{aligned} \]

which implies that (4.20) holds. This completes the proof.

**Appendix**

In this appendix, we introduce some preliminary results of SEEs and BSEEs, including existence, uniqueness and continuous dependence theorems.

Consider a SEE in the Gelfand triple \((V, H, V^*)\):
\[ \begin{aligned}
&\begin{cases}
\frac{dX(t)}{dt} = [A(t)X(t) + b(t, X(t))]dt + [B(t)X(t) + g(t, X(t))]dW(t) \\
+ \int_E \sigma(t, c, X(t-))\tilde{p}(dc, dt),
\end{cases}
\end{aligned} \]
\[ \begin{aligned}
&\begin{cases}
X(0) = x \in H, \quad t \in [0, T],
\end{cases}
\end{aligned} \]

where \(A, B, b, g\) and \(\sigma\) are given random mappings which satisfy the following standard assumptions.
Assumption A.1. The operator processes $A : [0, T] \times \Omega \rightarrow \mathcal{L}(V, V^*)$ and $B : [0, T] \times \Omega \rightarrow \mathcal{L}(V, H)$ are weakly predictable; i.e., $(A(\cdot), x, y)$ and $(B(\cdot), x, y)_{H}$ are both predictable process for every $x, y \in V$, and satisfy the coercive condition, i.e., there exist some constants $C, \alpha > 0$ and $\lambda$ such that for any $x \in V$ and each $(t, \omega) \in [0, T] \times \Omega$,

$$-(A(t)x, x) + \lambda||x||^2_H \geq \alpha||x||^2_H,$$

(A.2)

and

$$\sup_{(t, \omega) \in [0, T] \times \Omega} ||A(t, \omega)||_{\mathcal{L}(V, V^*)} + \sup_{(t, \omega) \in [0, T] \times \Omega} ||B(t, \omega)||_{\mathcal{L}(V, H)} \leq C.$$  

(A.3)

Assumption A.2. The mappings $b : [0, T] \times \Omega \times H \rightarrow H$ and $g : [0, T] \times \Omega \times H \rightarrow H$ are both $\mathcal{P} \times \mathcal{B}(H)/\mathcal{B}(H)$-measurable such that $b(0, \cdot), g(0, \cdot) \in M^2_{\mathbb{F}}(0, T; H)$; the mapping $\sigma : [0, T] \times \Omega \times E \times H \rightarrow H$ is $\mathcal{P} \times \mathcal{B}(E) \times \mathcal{B}(H)/\mathcal{B}(H)$-measurable such that $\sigma(\cdot, \cdot, \cdot, \cdot) \in M^2_{\mathbb{F}}(0, T] \times E; H)$. Now we state our main result.

To prove this theorem, we first show the following result on the continuous dependence of the solution to the SEE (A.1).

Lemma A.6. Let Assumptions A.3-A.4 be satisfied by any given coefficients $(A, B, b, g, \sigma)$ of the SEE (A.1). Then for any initial value $X(0) = x$, the SEE (A.1) has a unique solution $X(\cdot) \in M^2_{\mathbb{F}}(0, T; V)$ such that for every $\phi \in V$ and a.e. $(t, \omega) \in [0, T] \times \Omega$, it holds that

\[
\begin{align*}
(X(t), \phi)_H &= (x, \phi)_H + \int_0^t \langle A(s)X(s), \phi \rangle ds + \int_0^t \langle b(s, X(s)), \phi \rangle_H dW(s) \\
&\quad + \int_0^t \langle B(s)X(s) + g(s, X(s)), \phi \rangle_H dW(s) + \int_0^t \langle \sigma(s, e, X(s)), \phi \rangle_H d\tilde{\mu}(de, ds), \quad t \in [0, T],
\end{align*}
\]

or alternatively, $X(\cdot)$ satisfies the following Itô’s equation in $V^*$:

\[
\begin{align*}
X(t) &= x + \int_0^t A(s)X(s) ds + \int_0^t b(s, X(s)) ds + \int_0^t \int_E \sigma(s, e, X(s)) d\tilde{\mu}(de, ds), \quad t \in [0, T],
\end{align*}
\]

(A.6)

Now we state our main result.

Lemma A.7. Let $X(\cdot)$ be a solution to the SEE (A.1) with the initial value $X(0) = x$ and the coefficients $(A, B, b, g, \sigma)$ which satisfy Assumptions A.3-A.4. Then the following estimate holds:

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} ||X(t)||^2_H \right] + \mathbb{E}\left[\int_0^T ||X(t)||^2_H dt\right] \leq K\left\{||x||^2_H + \mathbb{E}\left[\int_0^T ||b(t, 0)||^2_H dt\right] + \mathbb{E}\left[\int_0^T ||g(t, 0)||^2_H dt\right] + \mathbb{E}\left[\int_0^T \int_E ||\sigma(t, e, 0)||^2_H d\nu(de) dt\right]\}.$$

(A.7)

Furthermore, suppose that $\tilde{X}(\cdot)$ is a solution to the SEE (A.1) with the initial value $\tilde{X}(0) = \tilde{x} \in H$ and the coefficients $(A, B, \tilde{b}, \tilde{g}, \tilde{\sigma})$ satisfying Assumptions A.3-A.4, then we have

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} ||X(t) - \tilde{X}(t)||^2_H \right] + \mathbb{E}\left[\int_0^T ||X(t) - \tilde{X}(t)||^2_H dt\right] \leq K\left\{||x - \tilde{x}||^2_H + \mathbb{E}\left[\int_0^T ||b(t, X(t)) - \tilde{b}(t, \tilde{X}(t))||^2_H dt\right]
$$

$$+ \mathbb{E}\left[\int_0^T ||g(t, X(t)) - \tilde{g}(t, \tilde{X}(t))||^2_H dt\right] + \mathbb{E}\left[\int_0^T \int_E ||\sigma(t, e, X(t)) - \sigma(t, e, \tilde{X}(t))||^2_H d\nu(de) dt\right]\}.$$ 

(A.8)
Next we consider a BSEE in the Gelfand triple \((V, H, V^*)\):

\[
\begin{aligned}
  dY(t) &= [A^*(t)Y(t) + B^*(t)Z(t) + f(t, Y(t), Z(t), R(t, \cdot))]dt + Z(t)dW(t) + \int_E R(t, e)\tilde{\mu}(dt, de), \\
  Y(T) &= \xi,
\end{aligned}
\]

where \((A^*, B^*, f, \xi)\) are given random mappings. Here \(A^*\) and \(B^*\) are the adjoint operators of \(A\) and \(B\), respectively. Furthermore, we assume that the coefficients \((A^*, B^*, f, \xi)\) satisfy the following conditions:

**Assumption A.3.** The operator processes \(A^* : [0, T] \times \Omega \to \mathcal{L}(V, V^*)\) and \(B^* : [0, T] \times \Omega \to \mathcal{L}(V, H)\) are weakly predictable, i.e., \(\langle A^*(\cdot) x, y \rangle\) and \(\langle B^*(\cdot) x, y \rangle_H\) are both predictable process for every \(x, y \in V\), and satisfy the coercive condition, i.e., there exist some constants \(C, \alpha > 0\) and \(\lambda\) such that for any \(x \in V\) and each \((t, \omega) \in [0, T] \times \Omega)\),

\[
-(A^*(t)x, x) + \lambda \|x\|_H \geq \alpha \|x\|_V + \|B^* x\|_H, \tag{A.10}
\]

and

\[
\sup_{(t, \omega) \in [0, T] \times \Omega} \|A^*(t, \omega)\|_{\mathcal{L}(V, V^*)} + \sup_{(t, \omega) \in [0, T] \times \Omega} \|B^*(t, \omega)\|_{\mathcal{L}(V, H)} \leq C. \tag{A.11}
\]

**Assumption A.4.** The mapping \(\xi : \Omega \to H\) is \(\mathcal{F}_T\)-measurable such that \(\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; H)\). The mappings \(f : [0, T] \times \Omega \times H \times M^{\nu, 2}(E; H) \to \mathcal{F} \times \mathcal{B}(H) \times \mathcal{B}(M^{\nu, 2}(E; H))/\mathcal{B}(H)\)-measurable such that \(f(\cdot, 0, 0, 0) \in M^{\nu, 2}_F(0, T; H)\). And there exists a constant \(C\) such that for all 

\[
(t, y, z, r, \bar{y}, \bar{z}, \bar{r}) \in [0, T] \times H \times M^{\nu, 2}(E; H) \times H \times H \times M^{\nu, 2}(E; H)
\]

and a.s. \((t, \omega) \in [0, T] \times \Omega\),

\[
\|f(t, y, z, r) - b(t, y, z, r)\|_H \leq C\left\{\|y - \bar{y}\|_H + \|z - \bar{z}\|_H + \|r - \bar{r}\|_{M^{\nu, 2}(E; H)}\right\}. \tag{A.12}
\]

If the coefficients \((A^*, B^*, f, \xi)\) satisfy Assumptions A.3 and A.4, they are said to be a generator of BSEE (A.9).

**Definition A.2.** A \((V \times H \times M^{\nu, 2}(E; H))\)-valued, \(\mathbb{F}\)-adapted process \((Y(\cdot), Z(\cdot), R(\cdot, \cdot))\) is called a solution to the BSEE (A.9), if \(Y(\cdot) \in M^F_2(0, T; V)\), \(Z(\cdot) \in M^{2}_F(0, T; H)\) and \(R(\cdot, \cdot) \in M^{\nu, 2}_F(0, T; H)\) such that for every \(\phi \in V\) and a.e. \((t, \omega) \in [0, T] \times \Omega\), it holds that

\[
\begin{aligned}
  (Y(t), \phi)_H &= (\xi, \phi)_H - \int_t^T \left(A^*(s)Y(s) + B^*(s)Z(s) + f(s, Y(s), Z(s), Y(s), R(s, \cdot)), \phi\right) dt \\
  &\quad - \int_t^T (Z(s), \phi)_H dW(s) - \int_t^T \int_E (R(s, e), \phi)_H \tilde{\mu}(ds, de), \quad t \in [0, T],
\end{aligned}
\]

or alternatively, \((Y(\cdot), Z(\cdot), R(\cdot, \cdot))\) satisfies the following Itô’s equation in \(V^*:\)

\[
\begin{aligned}
  Y(t) &= \xi - \int_t^T \left[A^*(s)Y(s) + B^*(s)Z(s) + f(t, Y(s), Z(s), R(s, \cdot))\right] ds \\
  &\quad - \int_t^T Z(s) dW(s) - \int_t^T \int_E R(s, e) d\tilde{\mu}(ds, de), \quad t \in [0, T].
\end{aligned}
\]

**Lemma A.8 (Existence and Uniqueness of BSEE [11]).** For any generator \((A^*, B^*, f, \xi)\), BSEE (A.9) has a unique solution \((Y(\cdot), Z(\cdot), R(\cdot, \cdot))\). Moreover, \(Y(\cdot) \in S^2_{\mathcal{F}}(0, T; H)\).

**Lemma A.9 (Continuous Dependence Theorem of BSEE).** Let \((A^*, B^*, f, \xi)\) and \((A^*, B^*, \tilde{f}, \tilde{\xi})\) be two generators of BSEE (A.9). Suppose that \((Y(\cdot), Z(\cdot), R(\cdot, \cdot))\) and \((\tilde{Y}(\cdot), \tilde{Z}(\cdot), \tilde{R}(\cdot, \cdot))\) are the solutions of BSEE (A.9) corresponding to \((A^*, B^*, f, \xi)\) and \((A^*, B^*, \tilde{f}, \tilde{\xi})\), respectively. Then

\[
\begin{aligned}
  &\mathbb{E}\left(\sup_{t \in [0, T]} \|Y(t) - \tilde{Y}(t)\|_H^2\right) + \mathbb{E}\left(\int_0^T \|Y(t) - \tilde{Y}(t)\|_H^2 dt\right) + \mathbb{E}\left(\int_0^T \|Z(t) - \tilde{Z}(t)\|_H^2 dt\right) \\
  &\quad + \mathbb{E}\left(\int_0^T \int_E \|R(t, e) - \tilde{R}(t, e)\|_H^2 \nu(de) dt\right) \\
  &\quad \leq K\left\{\mathbb{E}\|\xi - \tilde{\xi}\|_H^2 + \mathbb{E}\left(\int_0^T \|f(t, Y(t), \tilde{Z}(t), \tilde{R}(t, \cdot)) - \tilde{f}(t, \tilde{Y}(t), \tilde{Z}(t), \tilde{R}(t, \cdot))\|_H^2 dt\right)\right\}, \tag{A.15}
\end{aligned}
\]
where $K$ is a positive constant depending only on $T$ and the constants $C, \alpha, \lambda$ in Assumption A.3.

In particular, if $(A^*, B^*, \bar{f}, \bar{\xi}) = (A^*, B^*, 0, 0)$, the following a priori estimate holds

\[
\mathbb{E}\left[ \sup_{t \in [0,T]} \| Y(t) \|_{H}^2 \right] + \mathbb{E}\left[ \int_0^T \| Y(t) \|_{V}^2 dt \right] + \mathbb{E}\left[ \int_0^T \| Z(t) \|_{H}^2 dt \right] + \mathbb{E}\left[ \int_0^T \int_E \| R(t,e) \|_{H}^2 \nu(de) dt \right]
\leq K \left\{ \mathbb{E}\left[ \| \xi \|_{H}^2 \right] + \mathbb{E}\left[ \int_0^T \| f(t,0,0) \|_{H}^2 dt \right] \right\}.
\]

(A.16)

(A.17)

References

[1] A. Al-Hussein. Maximum principle for controlled stochastic evolution equations. *Int. J. Math. Anal.*, 4:1447–1464, 2010.

[2] A. Al-Hussein. Sufficient conditions of optimality for backward stochastic evolution equations. *Commun. Stoch. Anal.*, 4:433–442, 2010.

[3] A. Al-Hussein. Bsdes driven by infinite dimensional martingales and their applications to stochastic optimal control. *Random Oper. Stoch. Equ.*, 19:45–61, 2011.

[4] A. Al-Hussein. Erratum: Bsdes driven by infinite dimensional martingales and their applications to stochastic optimal control. *Random Oper. Stoch. Equ.*, 19:295–297, 2011.

[5] A. Bensoussan. Stochastic maximum principle for distributed parameter systems. *Journal of the Franklin Institute*, 315:387–406, 1983.

[6] K. Du and Q. Meng. (2013). A maximum principle for optimal control of stochastic evolution equations. *SIAM Journal on Control and Optimization*, 51(6), 4343-4362.

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

[8] G. Guatter. Stochastic maximum principle for spdes with noise and control on the boundary. *Systems and Control Letters*, 60:198–204, 2011.

[9] Y. Hu and S. Peng. Maximum principle for semi linear stochastic evolution control systems. *Stochastics*, 33:159–180, 1990.

[10] Q. Lü and X. Zhang. General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions. Springer, 2014.

[11] Q. Meng, & S. Tang. Stochastic Hamilton-Jacobi-Bellman Equations with Jumps and Random Coefficient, *preprint, 2017*

[12] Øksendal, F. Prosk and T. Zhang. Backward stochastic partial differential equations with jumps and application to optimal control of random jump fields. *Stochastics An International Journal of Probability and Stochastic Processes*, 77(5), 381-399, 2007.

[13] M. Tang and Q. Meng. Stochastic Evolution Equations of Jump Type with Random Coefficients: Existence, Uniqueness and Optimal Control. *SCIENCE CHINA Information Sciences 10.1007/s11432-016-9107-1*

[14] X. Zhou. On the necessary conditions of optimal controls for stochastic partial differential equations. *SIAM J. Control Optim.*, 31(6):1462–1478, 1993.