BULLETIN *of the*
MALAYSIAN MATHEMATICAL
SCIENCES SOCIETY
http://math.usm.my/bulletin

Bull. Malays. Math. Sci. Soc. (2) **37**(3) (2014), 797–811

Maximum Principle for Optimal Control of Stochastic Partial Differential Equations

ABDULRAHMAN AL-HUSSEIN
Department of Mathematics, College of Science, Qassim University
P. O. Box 6644, Buraydah 51452, Saudi Arabia
alhusseinqu@hotmail.com, hsien@qu.edu.sa

Abstract. We consider a stochastic maximum principle of optimal control for a control problem associated with a stochastic partial differential equation driven by a continuous martingale, which takes its values in a separable Hilbert space, and a random unbounded linear operator. We derive necessary conditions of optimality for this control problem without a convexity assumption on the control domain, and also when the control variable is allowed to enter in the martingale part of the equation. Linear and nonlinear equations are considered in this study.

2010 Mathematics Subject Classification: 60H15, 93E20, 35B50, 60G44

Keywords and phrases: Martingale, stochastic partial differential equation, optimal control, stochastic maximum principle, adjoint equation, backward stochastic partial differential equation.

1. Introduction

Consider the following stochastic partial differential equation (SPDE for short):

$$\begin{cases} dx(t) = (A(t)x(t) + a(t,u(t))x(t) + b(t,u(t)))dt \\ \qquad\qquad + [\langle \sigma(t,u(t)), x(t) \rangle_K + g(t,u(t))]dM(t), \quad 0 < t \leq T, \\ x(0) = x_0 \in K, \end{cases} \tag{1.1}$$

where $A(t), t \in [0,T]$, is a random, possibly unbounded, closed linear operator on a separable real Hilbert space K. The noise is modelled by a continuous martingale M in K, and a, b, σ and g are suitable predictable bounded mappings, while $u(\cdot)$ is a control process. This equation will be studied over a Gelfand triple (V, K, V'). That is V is a separable Hilbert space embedded continuously and densely in K, and V' is the dual space of V. Given a bounded measurable mapping $\ell : [0,T] \times \mathscr{O} \to K$ and a fixed element G of K, we shall be interested in minimizing the cost functional:

$$J(u(\cdot)) = \mathbb{E}\Big[\int_0^T \big\langle \ell(t,u(t)), x(t) \big\rangle_K dt + \big\langle G, x(T) \big\rangle_K \Big],$$

Communicated by Lee See Keong.
Received: May 25, 2012; *Revised:* July 30, 2012.

over the set of admissible controls (precise definitions to be given in Section 3 below). We will approach this by using the adjoint equation of SPDE (1.1), which is a backward stochastic partial differential equation (BSPDE for short) driven by an infinite dimensional martingale, and derive in particular a stochastic maximum principle for this optimal control problem. Let us recall that such BSPDEs (or even BSDEs) have their importance shown in applications in control theory like [5] and in some financial applications as in [23]. For more applications we refer the reader to Bally *et al.* [9], Imkeller *et al.* [18], and Fuhrman and Tessitore [14].

It is known that a Wiener filtration is usually required to deal with BSPDEs that arise as adjoint equations of controlled SPDEs. This is indeed a restriction insisted on for instance in [34] and [35]. Øksendal *et al.* in [26] and some other recent works have now considered the adjoint equation of a controlled BSPDE with a filtration generated by a Wiener process and a Poisson random measure. In our work here we can consider an arbitrary continuous filtration thanks to a result established in [3] giving the existence and uniqueness of solutions to BSPDEs driven by martingales. In this respect we refer the reader also to Imkeller *et al.* [18], where a filtration is being taken similarly to the one used here. The reader can also see [5, 15–17, 19, 20, 29, 30, 33] for SDEs and SPDEs with martingale noises. In fact in [5] we derived necessary conditions for optimality of stochastic systems similar to (1.1), but the result there describes the maximum principle only in a local form and requires, moreover, the convexity of the control domain U. In the present work we shall derive the maximum principle in its global form for our optimal control problem and, in particular, we shall not require the convexity of U. On the other hand, our results here generalize those in [35] and [11] and can be applied to the optimal control problem of partial observations with a given general nonlinear cost functional as done particularly in [35, Section 6]. The idea of reducing such a control problem to a control problem for a linear SPDE (Zakai's equation) was discussed also there. This is similar to (1.1).

We shall include also a more general result covering a control problem associated with an SPDE of the type:

$$
\begin{cases} dx(t) = (A(t)x(t) + F(t,x(t),u(t)))dt \\ \qquad\qquad + [\langle \sigma(t,u(t)), x(t)\rangle_K + g(t,u(t))]dM(t), \quad 0 < t \leq T, \\ x(0) = x_0 \in K, \end{cases} \tag{1.2}
$$

with respect to a cost functional:

$$
J(u(\cdot)) = \mathbb{E}\Big[\int_0^T \rho(u(t),x(t))\,dt + \langle G, x(T)\rangle_K\Big]. \tag{1.3}
$$

The main new features here are the driving noise is allowed to be an infinite dimensional martingale (as in Tudor [33] and Al-Hussein [4]), the control domain U need not be convex, and the control variable itself is allowed to enter in the martingale part of the equation as in SPDEs (1.1) and (1.2). To our best knowledge our results here are the first results that study such control systems and are much more general than those in literature so far, one can see now Remark 6.1 in Section 6.

The present paper is organized as follows. In Section 2 we introduce some definitions and notation that will be used throughout the paper. In Section 3 our main stochastic control problem is introduced. Section 4 is devoted to the adjoint equation of SPDE (1.1) as well as the existence and uniqueness of its solution. We state and establish the proof of our main result in Section 5. In Section 6 we generalize the main theorem of Section 5 to cover in particular the control problem associated with SPDE (1.2) and its cost functional (1.3). Further discussion is given in Remark 6.1 in Section 6.

2. Basic definitions and notation

We assume that $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ is a complete filtered probability space, such that $\{\mathscr{F}_t\}_{t\geq 0}$ is a continuous filtration, in the sense that every square integrable K-valued martingale with respect to $\{\mathscr{F}_t\,,\ 0\leq t\leq T\}$ has a continuous version. Let $\mathscr{P}$ denote the predictable σ-algebra of subsets of $\Omega\times[0,T]$. A K-valued process is said to be predictable if it is $\mathscr{P}/\mathscr{B}(K)$ measurable. Let $\mathscr{M}^{2,c}_{[0,T]}(K)$ be the space of all square integrable continuous martingales in K. We say that two elements M and N of $\mathscr{M}^{2,c}_{[0,T]}(K)$ are *very strongly orthogonal* (VSO) if $\mathbb{E}\,[M(\tau)\otimes N(\tau)] = \mathbb{E}\,[M(0)\otimes N(0)]$, for all $[0,T]$-valued stopping times τ. Here $\otimes$ denotes the tensor product so that $M\otimes N$ takes values in the space $K\otimes K$.

For $M\in\mathscr{M}^{2,c}_{[0,T]}(K)$ let $<< M >>$ be its *tensor quadratic variation* taking its values in the space $L_1(K)$, where $L_1(K)$ is the space of all nuclear operators on K, i.e. $<< M >>$ is the unique predictable process in $L_1(K)$ which vanishes at 0 and satisfies $M\otimes M - << M >> \in \mathscr{M}^{2,c}_{[0,T]}(L_1(K))$. Denote by $< M >$ the quadratic variation of M, which is defined similarly. It is known (see [25]) that there exist a predictable process $\tilde{\mathscr{Q}}_M(s,\omega)$ in $L_1(K)$ such that $<< M >>_t = \int_0^t \tilde{\mathscr{Q}}_M(s,\omega)\, d < M >_s$.

For (t,ω) if $\tilde{\mathscr{Q}}(t,\omega)$ is any symmetric, positive definite nuclear operator on K, we shall denote by $L_{\tilde{\mathscr{Q}}(t,\omega)}(K)$ the set of all linear (not necessarily bounded) operators Φ which map $\tilde{\mathscr{Q}}^{1/2}(t,\omega)(K)$ into K such that $\Phi\tilde{\mathscr{Q}}^{1/2}(t,\omega)\in L_2(K)$, the space of all Hilbert-Schmidt operators from K into itself. The inner product and norm in $L_2(K)$ will be denoted respectively by $\langle\cdot,\cdot\rangle_2$ and $||\cdot||_2$.

We recall that the stochastic integral $\int_0^{\cdot}\Phi(s)dM(s)$ is defined for mappings Φ such that for each (t,ω), $\Phi(t,\omega)\in L_{\tilde{\mathscr{Q}}_M(t,\omega)}(K)$, for every $h\in K$ the K-valued process $(t,\omega)\mapsto \big(\Phi(t,\omega)\tilde{\mathscr{Q}}_M^{1/2}(t,\omega)\big)(h)$ is predictable, and $\mathbb{E}\,[\int_0^T ||(\Phi\tilde{\mathscr{Q}}_M^{1/2})(t)||_2^2\, d < M >_t] < \infty$.

The space of such integrands is a Hilbert space with respect to the scalar product $(\Phi_1,\Phi_2)\mapsto \mathbb{E}\,[\int_0^T \langle \Phi_1\tilde{\mathscr{Q}}_M^{1/2}\,, \Phi_2\tilde{\mathscr{Q}}_M^{1/2}\rangle\, d < M >_t]$. Simple processes in $L(K)$ are examples of such integrands. Hence the closure of the set of simple processes in this Hilbert space is itself a Hilbert subspace. We denote it as in [25] by $\Lambda^2(K;\mathscr{P},M)$. More details and proofs can be found in [24] or [25].

In this paper we shall assume that there exists a measurable mapping $\mathscr{Q}(\cdot):[0,T]\times\Omega\to L_1(K)$ such that $\mathscr{Q}(t)$ is symmetric, positive definite, $<< M >>_t = \int_0^t \mathscr{Q}(s)\,ds$, and $\mathscr{Q}(t)\leq\mathscr{Q}$ for some positive definite nuclear operator $\mathscr{Q}$ on K. Thus $\tilde{\mathscr{Q}}_M(t) = \frac{\mathscr{Q}(t)}{q(t)}$ and $< M >_t = \int_0^t q(s)ds$, with $q(t) = \mathrm{tr}\,(\mathscr{Q}(t))$. Thus, if $\Phi\in\Lambda^2(K;\mathscr{P},M)$,

$$\mathbb{E}\Big[\,|\int_0^T \Phi(s)dM(s)|^2\Big] = \mathbb{E}\Big[\int_0^T ||\Phi(s)\mathscr{Q}^{1/2}(s)||_2^2\, ds\Big].$$

This equality will be used frequently in the proofs given in Section 5. The process $\mathscr{Q}(\cdot)$ will play an essential role in deriving the adjoint equation of the SPDE (1.1), as appearing in the equation (4.1) in Section 4; see in particular the discussion following equation (4.5).

On the other hand, by starting first with this nuclear covariance process $\mathscr{Q}(\cdot)$, one can also work more directly to define the above stochastic integrals.

3. Statement of the control problem

For a separable Hilbert space E, let us consider the following space:

$$L^2_{\mathscr{F}}(0,T;E) := \{\psi : [0,T]\times\Omega \to E \text{ (up to equivalence classes)}, \\ \text{predictable and } \mathbb{E}\Big[\int_0^T |\psi(t)|^2 dt\Big] < \infty\}.$$

Suppose that $\mathscr{O}$ is a separable Hilbert space with an inner product $\langle \cdot,\cdot\rangle_{\mathscr{O}}$, and U is a nonempty subset of $\mathscr{O}$. Denote by

$$\mathscr{U}_{ad} = \{u(\cdot) : [0,T]\times\Omega \to \mathscr{O} \text{ s.t. } u(\cdot) \in L^2_{\mathscr{F}}(0,T;\mathscr{O}), \text{ and } u(t)\in U \;\; a.e., \; a.s.\}.$$

This set is called the *set of admissible controls* and its elements are called *admissible controls*.

Now let us recall our first SPDE:

$$\begin{cases} dx(t) = (A(t)x(t) + a(t,u(t))x(t) + b(t,u(t)))dt \\ \qquad\qquad + [\langle \sigma(t,u(t)), x(t)\rangle_K + g(t,u(t))]dM(t), \\ x(0) = x_0 \in K, \end{cases} \tag{3.1}$$

and impose on it the following assumptions:
(i) $A(t,\omega)$ is a linear operator on K, $\mathscr{P}$-measurable, belongs to $L(V;V')$ uniformly in (t,ω) and satisfies the following two conditions.

(1) $A(t,\omega)$ satisfies the coercivity condition:

$$2\,\langle A(t,\omega)y, y\rangle_{V',V} + \alpha\,|y|_V^2 \le \lambda\,|y|_K^2 \quad a.e.\ t\in[0,T],\ \ a.s.\ \ \forall\, y\in V,$$

for some $\alpha,\lambda>0$, where $\langle A(t,\omega)y,y\rangle_{V',V}$ denotes the action of $A(t,\omega)y \in V'$ on $y\in V$.

(2) $\exists\,\beta \ge 0$ such that for all (t,ω)

$$|A(t,\omega)y|_{V'} \le \beta\,|y|_V \ \ \forall\, y\in V.$$

(ii) $a : \Omega\times[0,T]\times\mathscr{O} \to \mathbb{R}$, $b : \Omega\times[0,T]\times\mathscr{O} \to K$, $\sigma : \Omega\times[0,T]\times\mathscr{O} \to K$ and $g : \Omega\times[0,T]\times\mathscr{O} \to L_{\mathscr{Q}}(K)$ are predictable and bounded mappings.

Definition 3.1. *We say that* $x = x^{u(\cdot)} \in L^2_{\mathscr{F}}(0,T;V)$ *is a* solution *of (3.1) if* $\forall\,\eta\in V$ *(or any dense subset of* V*) and for almost all* $(t,\omega)\in[0,T]\times\Omega$

$$\langle x(t),\eta\rangle_K = \langle x_0,\eta\rangle_K + \int_0^t \langle A(s)x(s),\eta\rangle_{V',V}\,ds + \int_0^t \langle a(s,u(s))x(s) + b(s,u(s)),\eta\rangle_K\,ds \\ + \int_0^t \langle \eta, [\langle \sigma(s,u(s)), x(s)\rangle_K + g(s,u(s))]dM(s)\rangle_K.$$

Given a bounded measurable mapping $\ell : [0,T] \times \mathscr{O} \to K$ and a fixed element G of K, we define the *cost functional* by:

$$J(u(\cdot)) := \mathbb{E}\Big[\int_0^T \big\langle \ell(t,u(t)), x^{u(\cdot)}(t) \big\rangle_K dt + \big\langle G, x^{u(\cdot)}(T) \big\rangle_K \Big],\ u(\cdot) \in \mathscr{U}_{ad}. \tag{3.2}$$

It is easy to realize that under assumptions (i) and (ii) there exists a unique solution to (3.1) in $L^2_{\mathscr{F}}(0,T;K)$. This fact can be found in [15, Theorem 4.1, P. 105], [17, Theorem 2.10] or [4, Theorem 3.2], and also can be gleaned from [30]. Itô's formula for such SPDEs can be found in [16, Theorems 1, 2].

Our control problem is to minimize (3.2) over $\mathscr{U}_{ad}$. Any $u^*(\cdot) \in \mathscr{U}_{ad}$ satisfying

$$J(u^*(\cdot)) = \inf\{ J(u(\cdot)) :\ u(\cdot) \in \mathscr{U}_{ad} \} \tag{3.3}$$

is called an *optimal control*. The corresponding solution $x^{u^*(\cdot)}$ of (3.1), which we denote briefly by x^* and $(x^*, u^*(\cdot))$ are called respectively an *optimal solution* and an *optimal pair* of the stochastic optimal control problem (3.1)–(3.3).

The control problem associated with the SPDE (1.2) and (1.3) will be introduced and established in Section 6.

The existence problem of optimal control can be developed from the works of [1, 2] and [33]. However, a special case can be found in [4]. We refer the reader also to Remark 5.1 below.

4. Adjoint equation

Recall the SPDE (3.1) and the mappings in (3.2), and define the *Hamiltonian* $H : [0,T] \times \Omega \times K \times \mathscr{O} \times K \times L_2(K) \to \mathbb{R}$ for $(t,\omega,x,v,y,z) \in [0,T] \times \Omega \times K \times \mathscr{O} \times K \times L_2(K)$ by

$$\begin{aligned} H(t,\omega,x,v,y,z) :=& - \big\langle \ell(t,v), x \big\rangle_V - a(t,\omega,v) \big\langle x, y \big\rangle_K \\ & - \big\langle b(t,\omega,v), y \big\rangle_K - \big\langle \tilde{\sigma}(t,\omega,x,v) \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2, \end{aligned} \tag{4.1}$$

where $\tilde{\sigma} : [0,T] \times \Omega \times K \times \mathscr{O} \to L_{\mathscr{Q}}(K)$ is defined by

$$\tilde{\sigma}(t,\omega,x,v) = \big\langle \sigma(t,\omega,v), x \big\rangle_K \Phi(x) + g(t,\omega,v) \tag{4.2}$$

with Φ being the constant mapping $\Phi : K \to L_{\mathscr{Q}}(K), x \mapsto \Phi(x) = \mathrm{id}_K$. Then

$$\begin{aligned} \big\langle \tilde{\sigma}(t,\omega,x,v) \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2 &= \big\langle \big(\big\langle \sigma(t,\omega,v), x \big\rangle_K \Phi(x) + g(t,\omega,v) \big) \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2 \\ &= \big\langle \big\langle \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2 \sigma(t,\omega,v), x \big\rangle_K + \big\langle g(t,\omega,v) \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2 \\ &= \big\langle B(t,\omega,v) z, x \big\rangle_K + \big\langle g(t,\omega,v) \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2, \end{aligned}$$

where $B : [0,T] \times \Omega \times \mathscr{O} \to L(L_2(K),K)$ is defined such that

$$B(t,\omega,v) z = \big\langle \mathscr{Q}^{1/2}(t,\omega), z \big\rangle_2 \, \sigma(t,\omega,v). \tag{4.3}$$

Moreover,

$$\nabla_x H(t,\omega,x,v,y,z) = -\ell(t,v) - a(t,\omega,v) y - B(t,\omega,v) z. \tag{4.4}$$

Now we consider the adjoint equation of (3.1):

$$(4.5)\quad \begin{cases} dy^{u(\cdot)}(t) = -\Big[A^*(t)y^{u(\cdot)}(t) - \nabla_x H(t, x^{u(\cdot)}(t), u(t), y^{u(\cdot)}(t), z^{u(\cdot)}(t)\mathscr{Q}^{1/2}(t))\Big]dt \\ \qquad\qquad + z^{u(\cdot)}(t)\, dM(t) + dN^{u(\cdot)}(t), \quad 0 \le t < T, \\ y^{u(\cdot)}(T) = G, \end{cases}$$

where $A^*(t)$ is the adjoint operator of $A(t)$. A *solution* of this BSPDE (4.5) is a triple $(y^{u(\cdot)}, z^{u(\cdot)}, N^{u(\cdot)})$ in $L^2_{\mathscr{F}}(0,T;K) \times \Lambda^2(K;\mathscr{P},M) \times \mathscr{M}^{2,c}_{[0,T]}(K)$ such that the following equality holds *a.s.* for all $t \in [0,T]$, $N^{u(\cdot)}(0) = 0$ and $N^{u(\cdot)}$ is VSO to M:

$$\begin{aligned} y^{u(\cdot)}(t) =& \xi + \int_t^T \nabla_x H(s, x^{u(\cdot)}(s), u(s), y^{u(\cdot)}(s), z^{u(\cdot)}(s)\mathscr{Q}^{1/2}(s))\, ds \\ & - \int_t^T z^{u(\cdot)}(s)\, dM(s) - \int_t^T dN^{u(\cdot)}(s). \end{aligned}$$

The following theorem gives the solution to this BSPDE (4.5).

Theorem 4.1. *Assume that* (i)–(ii) *hold. Then there exists a unique solution* $(y^{u(\cdot)}, z^{u(\cdot)}, N^{u(\cdot)})$ *of the BSDE (4.5).*

The proof of this theorem can be found in [3].

We shall denote briefly the solution of (4.5) corresponding to an optimal control $u^*(\cdot)$ by (y^*, z^*, N^*).

Remark 4.1. (i) It is important to realize that the presence of the process $\mathscr{Q}^{1/2}(\cdot)$ in equation (4.5) is crucial in order for the mapping $\nabla_x H$ to be defined on the space $L_2(K)$, since the process $z^{u(\cdot)}$ need not be bounded as it is discussed in Section 2. This has to be taken always into account when we study BSPDEs or even BSDEs in infinite dimensions; one can see also [6].

(ii) We note that our assumption on the filtration $\{\mathscr{F}_t\}_{t\ge 0}$ to be a continuous filtration, i.e. every square integrable K-valued martingale with respect to $\{\mathscr{F}_t,\ 0 \le t \le T\}$ has a continuous version, is only needed to get the existence and uniqueness of the solution of the adjoint BSPDE (4.5), as it can be gleaned from [3], and guarantees the continuity of the martingale $N^{u(\cdot)}$ in (4.5). This assumption on the filtration is not needed when M is a Brownian motion or even a combination of a Brownian motion and a Poisson process because the martingale $N^{u(\cdot)}$ will vanish then; cf. for example [7]. Thus to our best knowledge our results here are much more general than those in literature so far.

5. Main results

In this section we shall derive and prove our main result on the maximum principle for optimal control of the SPDE (3.1) and its associated cost functional (3.2) and value function (3.3). We shall try to make use of the results of the previous section on the adjoint equation (BSPDE). Before doing so, let us mention that the relationship between BSPDEs and the maximum principle for some SPDEs is developed in several works, among them for instance are [28] and [34] and the references of Zhou cited therein. Other discussions in this respect can be found in [32] and [35] as well. Bensoussan in [12, Chapter 8] presents a stochastic maximum principle approach to the problem of stochastic control with partial information treating a general infinite dimensional setting, and the adjoint equation is derived also there.

Another work on the maximum principle that is connected to BSDEs can be found also in [8]. For an expanded discussion on the history of maximum principle we refer the reader to [34, P. 153–156]. One can find also useful information in Bensoussan's lecture notes [10] and Li and Yong [22].

Our main theorem is the following.

Theorem 5.1. *Suppose* (i)–(ii). *Assume moreover that U is compact and ℓ, a, b, σ, g are continuous as mappings in the control argument v a.s. If $(x^*, u^*(\cdot))$ is an optimal pair for the problem* (3.1)–(3.3)*, then there exists a unique solution (y^*, z^*, N^*) to the corresponding BSPDE* (4.5) *such that the following inequality holds:*

$$H(t, x^*(t), u, y^*(t), z^*(t)\mathscr{Q}^{1/2}(t)) \leq H(t, x^*(t), u^*(t), y^*(t), z^*(t)\mathscr{Q}^{1/2}(t)) \quad \text{a.e. } t \in [0,T], \text{ a.s. } \forall\, u \in U. \tag{5.1}$$

Remark 5.1. (i) The compactness assumption of U and the continuity of ℓ, a, b, σ, g in v in the above theorem are not actually needed in all the proofs that follow, however such assumptions are needed in order for the maximum of the mapping $v \mapsto H(t, X^*(t), v, Y^*(t), Z^*(t)\mathscr{Q}^{1/2}(t))$ required in (5.1) to exist in U.
(ii) A measurable selection theorem due to Ekeland and Temam, [13], can be applied to select an admissible control satisfying (5.1), one can see also [21, Theorem 3.2, p. 169] for the same purpose.

To start proving Theorem 5.1 we need to develop some necessary estimates using the so-called spike variation method. For this we let $(x^*, u^*(\cdot))$ be the given optimal pair. Let $0 \leq t_0 < T$ be fixed such that $\mathbb{E}\,[|x^*(t_0)|_K^2] < \infty$ and $0 \leq \varepsilon < T - t_0$. Let u be a random variable taking its values in U, $\mathscr{F}_{t_0}$-measurable and $\sup_{\omega\in\Omega} |u(\omega)| < \infty$. Consider the following spike variation of the control $u^*(\cdot)$:

$$u_\varepsilon(t) = \begin{cases} u^*(t) & \text{if } t \in [0,T] \backslash [t_0, t_0 + \varepsilon] \\ u & \text{if } t \in [t_0, t_0 + \varepsilon]. \end{cases}$$

We can consider the $x^{u_\varepsilon(\cdot)}$ as the solution of the SPDE (3.1) corresponding to $u_\varepsilon(\cdot)$. We shall denote it briefly by x_ε. Note that $x_\varepsilon(t) = x^*(t)$ for all $0 \leq t \leq t_0$.

The proof of this theorem will be divided into several lemmas as follows.

Lemma 5.1. *Suppose* (i)–(ii). *Then*

$$\sup_{t_0 \leq t \leq t_0+\varepsilon} \mathbb{E}\,[|x_\varepsilon(t)|_K^2] \leq C_1 \big(\mathbb{E}\,[|x^*(t_0)|_K^2 + C_2\, \varepsilon \big) \tag{5.2}$$

for some positive constants C_1 and C_2.

Proof. Observe first from (3.1) and (5.2) that, for $t_0 \leq t \leq t_0 + \varepsilon$,

$$\begin{aligned} x_\varepsilon(t) = & x^*(t_0) + \int_{t_0}^t \big(A(s) x_\varepsilon(s) + a(s,u) x_\varepsilon(s) + b(s,u) \big) ds \\ & + \int_{t_0}^t \big[\big\langle \sigma(s,u), x_\varepsilon(s) \big\rangle_K + g(s,u) \big] dM(s). \end{aligned} \tag{5.3}$$

Then, by Itô's formula, assumption (i), Cauchy-Schwartz inequality and assumption (ii) we get

$$
\begin{aligned}
\mathbb{E}[\,|x_\varepsilon(t)\,|_K^2\,] &+ \alpha\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_V^2\,ds\Big] \le \mathbb{E}[\,|x^*(t_0)\,|_K^2\,] + \lambda\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_K^2\,ds\Big] \\
&+ 2\mathbb{E}\Big[\int_{t_0}^t \big\langle a(s,u)\,x_\varepsilon(s)\,,x_\varepsilon(s)\big\rangle_K ds\Big] + 2\mathbb{E}\Big[\int_{t_0}^t \big\langle x_\varepsilon(s)\,,b(s,u)\big\rangle_K ds\Big] \\
&+ 2\mathbb{E}\Big[\int_{t_0}^t ||\big\langle \sigma(s,u)\,,x_\varepsilon(s)\big\rangle_K \,\mathrm{id}_K\,\mathscr{Q}^{1/2}(s)\,||_2^2 ds\Big] \\
&+ 2\mathbb{E}\Big[\int_{t_0}^t ||\,g(s,u)\mathscr{Q}^{1/2}(s)\,||_2^2 ds\Big] \\
\le \mathbb{E}[\,|x^*(t_0)\,|_K^2\,] &+ \lambda\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_K^2\,ds\Big] + 2\,k_1\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_K^2\,ds\Big] \\
&+ k_2^2\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_K^2\,ds\Big] + (t-t_0) + 2k_3^2\,||\mathscr{Q}^{1/2}||_2^2\,\mathbb{E}\Big[\int_{t_0}^t |x_\varepsilon(s)\,|_K^2\,ds\Big] \\
&+ 2k_4^2\varepsilon\,||\mathscr{Q}^{1/2}||_2^2\,(t-t_0) \\
\le \big(\lambda + 2k_1 + k_2^2 &+ 2k_3^2\,||\mathscr{Q}^{1/2}||_2^2\big)\int_{t_0}^t \mathbb{E}[\,|x_\varepsilon(s)|_K^2\,]\,ds \\
&+ (1+2k_4^2\,||\mathscr{Q}^{1/2}||_2^2)\,\varepsilon + \mathbb{E}[\,|x^*(t_0)\,|_K^2\,]. \qquad (5.4)
\end{aligned}
$$

In the last part of this inequality we have used the boundedness in assumption (ii) of the mappings a, b, σ, g respectively to get the constants $k_1, \dots, k_4$.

Thus, in particular, applying Gronwall's inequality to (5.4) gives (5.2) with

$$C_1 = e^{\varepsilon\,\big(\lambda+2k_1+k_2^2+2k_3^2\,||\mathscr{Q}^{1/2}||_2^2\big)}$$

and

$$C_2 = 1 + 2k_4^2\,||\mathscr{Q}^{1/2}||_2^2.$$

This completes the proof. ∎

Lemma 5.2. *Suppose* (i)–(ii). *Then*

$$\sup_{t_0+\varepsilon\le t\le T} \mathbb{E}[\,|x_\varepsilon(t)|_K^2\,] \le C_3\big(\mathbb{E}[\,|x^*(t_0)|_K^2\,] + C_4\,\varepsilon + 1\big) \qquad (5.5)$$

for some positive constants C_3 *and* C_4.

Proof. For $t_0 + \varepsilon \le t \le T$, it follows that

$$
\begin{aligned}
x_\varepsilon(t) =& x^*(t_0+\varepsilon) + \int_{t_0+\varepsilon}^t \big(A(s)x_\varepsilon(s) + a(s,u^*(s))x_\varepsilon(s) + b(s,u^*(s))\big)ds \\
&+ \int_{t_0+\varepsilon}^t \big[\big\langle \sigma(s,u^*(s)),x_\varepsilon(s)\big\rangle_K + g(s,u^*(s))\big]\,dM(s). \qquad (5.6)
\end{aligned}
$$

Thus mimicking the proof of Lemma 5.1 and then applying inequality (5.2) easily yield (5.5). ∎

Lemma 5.3. *Suppose* (i)–(ii). *Let* $\xi_\varepsilon(t) = x_\varepsilon(t) - x^*(t)$, *for* $t \in [0,T]$. *Then*

$$\sup_{t_0+\varepsilon\le t\le T} \mathbb{E}[\,|\xi_\varepsilon(t)|_K^2\,] = O(\varepsilon). \qquad (5.7)$$

Proof. It is easy to get for $t \in [t_0+\varepsilon, T]$,

$$\xi_\varepsilon(t) = \xi_\varepsilon(t_0+\varepsilon) + \int_{t_0+\varepsilon}^t \big(A(s)\xi_\varepsilon(s) + a(s,u^*(s))\,\xi_\varepsilon(s)\big)ds$$
$$+ \int_{t_0+\varepsilon}^t \big\langle \sigma(s,u^*(s)), \xi_\varepsilon(s)\big\rangle_K \, dM(s). \tag{5.8}$$

Hence, as done in the proof of Lemma 5.1, we get

$$\sup_{t_0+\varepsilon\le t\le T} \mathbb{E}\,[|\xi_\varepsilon(t)|_K^2] \le C_5\,\mathbb{E}\,[|\xi_\varepsilon(t_0+\varepsilon)|_K^2]. \tag{5.9}$$

On the other hand, for $t_0 \le t \le t_0+\varepsilon$ we have $\xi_\varepsilon(t_0)=0$ and

$$\xi_\varepsilon(t) = \int_{t_0}^t \Big[A(s)\xi_\varepsilon(s) + \big(a(s,u) - a(s,u^*(s))\big)\,x_\varepsilon(s)$$
$$+ \big(b(s,u) - b(s,u^*(s))\big) + a(s,u^*(s))\,\xi_\varepsilon(s)\Big]ds$$
$$+ \int_{t_0}^t \Big[\big\langle \sigma(s,u) - \sigma(s,u^*(s)), x_\varepsilon(s)\big\rangle_K$$
$$+ \big(g(s,u) - g(s,u^*(s))\big) + \big\langle \sigma(s,u^*(s)), \xi_\varepsilon(s)\big\rangle_K\Big] dM(s). \tag{5.10}$$

Hence by Itô's formula, assumption (i), Cauchy-Schwartz inequality and assumption (ii) it follows that

$$\mathbb{E}\,[\,|\,\xi_\varepsilon(t)\,|_K^2\,] + \alpha\,\mathbb{E}\,[\int_{t_0}^t |\,\xi_\varepsilon(s)\,|_V^2\,ds]$$
$$\le \lambda\,\mathbb{E}\Big[\int_{t_0}^t |\,\xi_\varepsilon(s)\,|_K^2\,ds\Big] + 2\,\mathbb{E}\Big[\int_{t_0}^t \big\langle \xi_\varepsilon(s), \big(a(s,u) - a(s,u^*(s))\big)\,x_\varepsilon(s)\big\rangle_K\,ds\Big]$$
$$+ 2\,\mathbb{E}\Big[\int_{t_0}^t \big\langle \xi_\varepsilon(s), b(s,u) - b(s,u^*(s))\big\rangle_K\,ds\Big]$$
$$+ 2\,\mathbb{E}\Big[\int_{t_0}^t \big\langle \xi_\varepsilon(s), a(s,u^*(s))\,\xi_\varepsilon(s)\big\rangle_K\,ds\Big]$$
$$+ 3\,\mathbb{E}\Big[\int_{t_0}^t ||\big\langle \sigma(s,u) - \sigma(s,u^*(s)), x_\varepsilon(s)\big\rangle_K \,\mathrm{id}_K\,\mathscr{Q}^{1/2}(s)\,||_2^2\,ds\Big]$$
$$+ 3\,\mathbb{E}\Big[\int_{t_0}^t ||\big\langle \sigma(s,u^*(s)), \xi_\varepsilon(s)\big\rangle_K \,\mathrm{id}_K\,\mathscr{Q}^{1/2}(s)\,||_2^2\,ds\Big]$$
$$+ 3\,\mathbb{E}\Big[\int_{t_0}^t ||\big(g(s,u) - g(s,u^*(s))\big)\,\mathscr{Q}^{1/2}(s)\,||_2^2\,ds\Big]$$
$$\le (\lambda + 4k_1^2 + 4k_2^2 + 2k_1 + 3k_3^2\,||\mathscr{Q}^{1/2}||_2^2) \int_{t_0}^t \mathbb{E}\,[\,|\,\xi_\varepsilon(s)|_K^2\,]\,ds$$
$$+ (12k_3^2\,||\mathscr{Q}^{1/2}||_2^2 + 1) \int_{t_0}^{t_0+\varepsilon} \mathbb{E}\,[|x_\varepsilon(s)|_K^2]\,ds + (1 + 12k_4^2\,||\mathscr{Q}^{1/2}||_2^2)\,\varepsilon$$
$$\le (\lambda + 4k_1^2 + 4k_2^2 + 2k_1 + 3k_3^2\,||\mathscr{Q}^{1/2}||_2^2) \int_{t_0}^t \mathbb{E}\,[\,|\,\xi_\varepsilon(s)|_K^2\,]\,ds$$
$$+ (12k_3^2\,||\mathscr{Q}^{1/2}||_2^2 + 1)\,C_1\cdot\big(\mathbb{E}\,[|x^*(t_0)|_K^2 + C_2\,\varepsilon\big)\,\varepsilon + (1 + 12k_4^2||\mathscr{Q}^{1/2}||_2^2)\,\varepsilon. \tag{5.11}$$

Therefore Gronwall's inequality gives

$$\sup_{t_0 \leq t_0+\varepsilon} \mathbb{E}\,[\,|\,\xi_\varepsilon(t)\,|_K^2\,] \leq C_6(\varepsilon)\cdot\varepsilon, \tag{5.12}$$

where

$$C_6(\varepsilon) = e^{(\lambda+4k_1^2+4k_2^2+2k_1+3k_3^2\,||\mathscr{Q}^{1/2}||_2^2)\,\varepsilon}\cdot\Big[\,(12k_3^2\,||\mathscr{Q}^{1/2}||_2^2+1)\,C_1\cdot\big(\mathbb{E}\,[\,|x^*(t_0)|_K^2\,] + C_2\,\varepsilon\big)+1+12k_4^2\,||\mathscr{Q}^{1/2}||_2^2\Big].$$

Now by applying (5.12) in (5.9) it yields eventually

$$\sup_{t_0+\varepsilon\leq t\leq T} \mathbb{E}\,[\,|\xi_\varepsilon(t)|_K^2\,] \leq C_5\,C_6(\varepsilon)\cdot\varepsilon. \tag{5.13}$$

Thus (5.7) follows. ∎

In the following result we shall try to compute $\mathbb{E}\,[\,\langle\, y^*(t_0+\varepsilon), \xi(t_0+\varepsilon)\,\rangle_K\,]$.

Lemma 5.4. *Suppose* (i)–(ii). *We have*

$$\begin{aligned}
&\mathbb{E}\Big[\,\big\langle\, y^*(t_0+\varepsilon)\,,\xi_\varepsilon(t_0+\varepsilon)\,\big\rangle_K + \int_{t_0}^{t_0+\varepsilon} \big\langle \ell(t,u^*(t))\,,\xi_\varepsilon(t)\big\rangle_K dt\Big]\\
&= \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,,\big(a(t,u)-a(t,u^*(t))\big)\,x_\varepsilon(t)\big\rangle_K dt\Big]\\
&\quad + \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,,b(t,u)-b(t,u^*(t))\big\rangle_K\,dt\Big]\\
&\quad + \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,u)-\sigma(t,u^*(t))\,,x_\varepsilon(t)\big\rangle_K\,\big\langle \mathscr{Q}^{1/2}(t)\,,z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\,dt\Big]\\
&\quad + \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \big(g(t,u)-g(t,u^*(t))\big)\mathscr{Q}^{1/2}(t)\,,z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\,dt\Big]
\end{aligned} \tag{5.14}$$

and

$$\mathbb{E}\,[\,\big\langle\, y^*(t_0+\varepsilon), \xi_\varepsilon(t_0+\varepsilon)\,\big\rangle_K\,] = \mathbb{E}\Big[\int_{t_0+\varepsilon}^{T} \big\langle \ell(t,u^*(t))\,,\xi_\varepsilon(t)\big\rangle_K dt\Big] + \mathbb{E}\,\big[\,\big\langle G,\xi_\varepsilon(T)\big\rangle_K\,\big]. \tag{5.15}$$

Proof. Note that for $t_0 \leq t \leq t_0+\varepsilon$ we have $\xi_\varepsilon(t_0)=0$ and (5.10). Therefore by using Itô's formula to (5.10) together with (4.5), (4.4) and (4.3) we get easily (5.14). The equality in (5.15) is proved similarly with the help of (5.8). ∎

Let us now record the following remark.

Remark 5.2. We note from Lebesgue differentiation theorem (e.g. [31]) that for *a.e.* $s \in [0,T]$,

$$\frac{1}{\varepsilon}\int_s^{s+\varepsilon} \big|\,\mathbb{E}\,[\,|y^*(t)|^2\,] - \mathbb{E}\,[\,|y^*(s)|^2\,]\,\big|\,dt \to 0, \text{ as } \varepsilon\to 0,$$

i.e. s is a Lebesgue point of the functions $t\mapsto \mathbb{E}\,[\,|y^*(t)\,|_K^2\,]$. This implies in particular that

$$\frac{1}{\varepsilon}\int_s^{s+\varepsilon} \mathbb{E}\,[\,|y^*(t)|^2\,]\,dt \to \mathbb{E}\,[\,|y^*(s)|^2\,], \text{ as } \varepsilon\to 0,$$

for *a.e.* s in $[0,T]$. Similarly, for the function $t\mapsto \mathbb{E}\,[\,||z^*(t)\mathscr{Q}^{1/2}(t)\,||_2^2\,]$.

So we can and will choose from here on $t_0 \in [0,T)$ so that t_0 is a Lebesgue point of the functions $t \mapsto \mathbb{E}[\,|y^*(t)\,|_K^2\,]$ and $t \mapsto \mathbb{E}[\,||z^*(t)\mathscr{Q}^{1/2}(t)\,||_2^2\,]$.

Lemma 5.5. *If* (i)–(ii) *hold, then*

$$
\begin{aligned}
0 \le\ & \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \ell(t,u) - \ell(t,u^*(t))\,, x^*(t)\big\rangle_K dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,, \big(a(t,u) - a(t,u^*(t))\big)\,x^*(t)\big\rangle_K dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,u) - \sigma(t,u^*(t))\,,\, x^*(t)\big\rangle_K \big\langle \mathscr{Q}^{1/2}(t)\,, z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,, b(t,u) - b(t,u^*(t))\big\rangle_K\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \big(g(t,u) - g(t,u^*(t)\big)\mathscr{Q}^{1/2}(t)\,, z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\, dt\Big] + o(\varepsilon).
\end{aligned}
\tag{5.16}
$$

Proof. Since $u^*(\cdot)$ is optimal, we have

$$
\begin{aligned}
0 \le J(u_\varepsilon(\cdot)) - J(u^*(\cdot)) =& \mathbb{E}\Big[\int_0^T \big(\big\langle \ell(t,u_\varepsilon(t))\,, x_\varepsilon(t)\big\rangle_K - \big\langle \ell(t,u^*(t))\,, x^*(t)\big\rangle_K\big)\, dt\Big] \\
&+ \mathbb{E}\,\Big[\big\langle G\,, x_\varepsilon(T)\big\rangle_K - \big\langle G\,, x^*(T)\big\rangle_K\Big] \\
=& \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big(\big\langle \ell(t,u) - \ell(t,u^*(t))\,, x_\varepsilon(t)\big\rangle_K + \big\langle \ell(t,u^*(t))\,, \xi_\varepsilon(t)\big\rangle_K\big)\, dt\Big] \\
&+ \mathbb{E}\,\Big[\int_{t_0+\varepsilon}^T \big\langle \ell(t,u^*(t))\,, \xi_\varepsilon(t)\big\rangle_K dt + \big\langle G, \xi_\varepsilon(T)\big\rangle_K\Big].
\end{aligned}
$$

Hence using Lemma 5.4 (5.15) in this inequality gives

$$
\begin{aligned}
0 \le& \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big(\big\langle \ell(t,u) - \ell(t,u^*(t))\,, x_\varepsilon(t)\big\rangle_K dt + \big\langle \ell(t,u^*(t))\,, \xi_\varepsilon(t)\big\rangle_K\big)\, dt\Big] \\
&+ \mathbb{E}\,\Big[\big\langle y^*(t_0+\varepsilon))\,, \xi_\varepsilon(t_0+\varepsilon)\big\rangle_K\Big].
\end{aligned}
\tag{5.17}
$$

Again by Lemma 5.4 (5.14) inequality (5.17) becomes

$$
\begin{aligned}
0 \le& \mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \ell(t,u) - \ell(t,u^*(t))\,, x_\varepsilon(t)\big\rangle_K\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,, \big(a(t,u) - a(t,u^*(t))\big)\, x_\varepsilon(t)\big\rangle_K\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle y^*(t)\,, b(t,u) - b(t,u^*(t))\big\rangle_K\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,u) - \sigma(t,u^*(t))\,, x_\varepsilon(t)\big\rangle_K \big\langle \mathscr{Q}^{1/2}(t)\,, z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\, dt\Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \big(g(t,u) - g(t,u^*(t))\big)\mathscr{Q}^{1/2}(t)\,, z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\, dt\Big].
\end{aligned}
\tag{5.18}
$$

On the other hand, assumption (ii) and Lemma 5.3 imply

$$\begin{aligned}
&\frac{1}{\varepsilon}\,\mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon}\big\langle y^*(t)\,,\big(a(t,u)-a(t,u^*(t))\big)\,\xi_\varepsilon(t)\big\rangle_K\,dt\Big]\\
&\leq C_7\,(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\mathbb{E}\Big(|y^*(t)|_K\cdot|\xi_\varepsilon(t)|_K\Big)\,dt\\
&\leq C_7\,(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\Big((\frac{\varepsilon^{1/3}}{2})\,\mathbb{E}\,[|y^*(t)|_K^2]+(\frac{1}{2\,\varepsilon^{1/3}})\,\mathbb{E}\,[|\xi_\varepsilon(t)|_K^2]\Big)\,dt\\
&\leq C_8\,\Big(\varepsilon^{1/3}\,(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\mathbb{E}\,[|y^*(t)|_K^2]dt+(\frac{1}{\varepsilon})\,\varepsilon\,(\frac{1}{\varepsilon^{1/3}})\,\varepsilon\Big)\to 0,
\end{aligned} \tag{5.19}$$

as $\varepsilon\to 0$, for some positive constants C_7 and C_8, since t_0 is a Lebesgue point of the function $t\mapsto\mathbb{E}\,[\,|y^*(t)\,|_K^2]$; cf. Remark 5.2.

Similarly,

$$\begin{aligned}
&\frac{1}{\varepsilon}\,\mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon}\Big(\big\langle\ell(t,u)-\ell(t,u^*(t))\,,\xi_\varepsilon(t)\big\rangle_K\\
&+\big\langle\sigma(t,u)-\sigma(t,u^*(t))\,,\xi_\varepsilon(t)\big\rangle_K\,\big\langle\mathscr{Q}^{1/2}(t)\,,z^*(t)\mathscr{Q}^{1/2}(t)\big\rangle_2\Big)\,dt\Big]\to 0,
\end{aligned} \tag{5.20}$$

as $\varepsilon\to 0$, since t_0 is a Lebesgue point of the function $t\mapsto\mathbb{E}\,[\,||z^*(t)\mathscr{Q}^{1/2}(t)\,||_2^2\,]$.

Therefore, by applying (5.19) and (5.20) in (5.18) we obtain (5.16). ∎

We are now ready to complete the proof of Theorem 5.1.

Proof of Theorem 5.1 Divide (5.16) in Lemma 5.5 by ε and let $\varepsilon\to 0$ to get

$$\begin{aligned}
&\mathbb{E}\,\Big[\big\langle\ell(t_0,u)-\ell(t_0,u^*(t_0))\,,x^*(t_0)\big\rangle_K+\big\langle y^*(t_0)\,,\big(a(t_0,u)-a(t_0,u^*(t_0))\big)\,x^*(t_0)\big\rangle_K\Big]\\
&+\mathbb{E}\,\big[\big\langle y^*(t_0)\,,b(t_0,u)-b(t_0,u^*(t_0))\big\rangle_K\big]\\
&+\mathbb{E}\,\big[\big\langle\sigma(t_0,u)-\sigma(t_0,u^*(t_0))\,,x^*(t_0)\big\rangle_K\,\big\langle\mathscr{Q}^{1/2}(t_0)\,,z^*(t_0)\mathscr{Q}^{1/2}(t_0)\big\rangle_2\big]\\
&+\mathbb{E}\,\big[\big\langle\big(g(t_0,u)-g(t_0,u^*(t_0))\big)\mathscr{Q}^{1/2}(t_0)\,,z^*(t_0)\mathscr{Q}^{1/2}(t_0)\big\rangle_2\big]\geq 0.
\end{aligned}$$

Consequently,

$$\mathbb{E}\,[H(t_0,x^*(t_0),u,y^*(t_0),z^*(t_0)\mathscr{Q}^{1/2}(t_0))]\leq\mathbb{E}\,[H(t_0,x^*(t_0),u^*(t_0),y^*(t_0),z^*(t_0)\mathscr{Q}^{1/2}(t_0))].$$

Hence (5.1) holds by a standard argument as for example in [34, Chapter 3], and the proof of Theorem 5.1 is then complete.

6. Nonlinear case

In Theorem 5.1 we derived the maximum principle for the stochastic optimal control problem, which is governed by the SDE (3.1). By using the proofs in Section 5 and our results in [6] we can generalize Theorem 5.1 to cover a control problem governed by the following SPDE:

$$\begin{cases}
dx^{u(\cdot)}(t)=(A(t)x^{u(\cdot)}(t)+F(t,x^{u(\cdot)}(t),u(t)))dt\\
\qquad\qquad+[\big\langle\sigma(t,u(t)),x^{u(\cdot)}(t)\big\rangle_K+g(t,u(t))]dM(t),\quad 0<t\leq T,\\
x^{u(\cdot)}(0)=x_0\in K,
\end{cases} \tag{6.1}$$

and a cost functional given by the following formula:

$$J(u(\cdot)) = \mathbb{E}\,[\int_0^T \rho(u(t),x^{u(\cdot)}(t))\,dt + \langle G, x(T)\rangle_K],\;\; u(\cdot)\in\mathscr{U}_{ad}, \tag{6.2}$$

where G is a fixed element of K as in (3.2) and $\rho : K\times\mathscr{O}\to\mathbb{R}$ is a measurable mapping satisfying the following assumption:

(E1) ρ is continuously Fréchet differentiable with respect to (ν,x) and the derivatives ρ_ν,ρ_x are uniformly bounded.

The operator $A(t)$ and the mappings σ, g in (6.1) are the same as in (3.1) which are assumed to satisfy in particular condition (i) in Section 3, while F satisfies the following condition:

(E2) $F : [0,T]\times\Omega\times K\times\mathscr{O}\to K$ is predictable, continuously Fréchet differentiable with respect to $x\in K$ and $\nu\in\mathscr{O}$, and the derivatives F_x, F_ν are uniformly bounded.

The value function is

$$J(u^*(\cdot)) = \inf\{J(u(\cdot)) :\; u(\cdot)\in\mathscr{U}_{ad}\}. \tag{6.3}$$

Theorem 6.1. *Suppose (i), (E1) and (E2). If $(x^*,u^*(\cdot))$ is an optimal pair for the problem (6.1)–(6.3), then there exists a unique solution (y^*,z^*,N^*) to the corresponding BSPDE:*

$$\begin{cases} dy^{u(\cdot)}(t) = -\Big[A^*(t)y^{u(\cdot)}(t) - \nabla_x H(t,x^{u(\cdot)}(t),u(t),y^{u(\cdot)}(t),z^{u(\cdot)}(t)\mathscr{Q}^{1/2}(t))\Big]dt \\ \qquad\qquad + z^{u(\cdot)}(t)\,dM(t) + dN^{u(\cdot)}(t), \quad 0\le t<T, \\ y^{u(\cdot)}(T) = G, \end{cases} \tag{6.4}$$

such that the following inequality holds:

$$H(t,x^*(t),u,y^*(t),z^*(t)\mathscr{Q}^{1/2}(t)) \le H(t,x^*(t),u^*(t),y^*(t),z^*(t)\mathscr{Q}^{1/2}(t)) \qquad a.e.\; t\in[0,T],\; a.s.\; \forall\, u\in U, \tag{6.5}$$

where

$$H(t,\omega,x,\nu,y,z) := -\rho(\nu,x) - \langle F(t,\omega,x,\nu),y\rangle_K - \langle \tilde\sigma(t,\omega,x,\nu)\mathscr{Q}^{1/2}(t,\omega),z\rangle_2,$$

with $\tilde\sigma$ being defined as in (4.2).

The proof can be done in parallel to the proofs of Theorem 5.1 and Theorem 5.3 in [6], so we omit it.

Remark 6.1. It would now be interesting to know if in the SPDE (6.1) a control process $u(\cdot)$ can be allowed to enter generally in a nonlinear way in the domain of the martingale M term, and in particular to consider the following SPDE:

$$\begin{cases} dx^{u(\cdot)}(t) = (A(t)x^{u(\cdot)}(t) + F(t,x^{u(\cdot)}(t),u(t)))dt + \kappa(t,x^{u(\cdot)}(t),u(t))dM(t), \\ x(0) = x_0, \end{cases} \tag{6.6}$$

with similar cost functional and value function to those in (6.2) and (6.3). In fact it is not obvious how to handle this optimal control problem because an approximation results to compute $\frac{d}{d\varepsilon}J(u_\varepsilon(\cdot))|_{\varepsilon=0}$ is needed to get a proper duality formula between the solution of (6.6) and its adjoint process. For a similar result, we refer the reader for example to Lemma 4.4, Corollary 4.6 and Lemma 5.2 in [6]. This new case requires new results concerning existence and uniqueness of the solutions to some certain "second-order" backward stochastic partial differential equations driven by martingales in order to solve the resulting adjoint system; Chapter 3 of [34] and [27] would be useful in this direction. However, the work of Peng in [27] does not apply here because the solutions of such so-called second-order adjoint system should consist of at least six processes $(y,z,N,\tilde{y},\tilde{z},\tilde{N})$ such that N and $\tilde{N}$ are martingales in $\mathscr{M}^{2,c}_{[0,T]}(K)$, for which we have to take into account that certain properties and may be notions of orthogonality between two martingales and a general martingale representation in parallel to those in [3] must also be present for this new research problem.

In Theorems (5.1, 6.1) we have succeeded to allow the martingale noise term to depend almost linearly on the control variable, and overcome this difficulty at least partially. The general result associated with SPDE (6.6), cost functional (6.2) and value function (6.3) is still an interesting open research problem. This problem is pointed out in [5, Remark 6.4] as well.

Acknowledgement. This work is supported by the Science College Research Center at Qassim University, project no. SR-D-011-724.

References

[1] N. U. Ahmed, Existence of optimal controls for a class of systems governed by differential inclusions on a Banach space, *J. Optim. Theory Appl.* **50** (1986), no. 2, 213–237.

[2] N. U. Ahmed, Relaxed controls for stochastic boundary value problems in infinite dimension, in *Optimal control of partial differential equations (Irsee, 1990)*, 1–10, Lecture Notes in Control and Inform. Sci., 149 Springer, Berlin.

[3] A. Al-Hussein, Backward stochastic partial differential equations driven by infinite-dimensional martingales and applications, *Stochastics* **81** (2009), no. 6, 601–626.

[4] A. Al-Hussein, Maximum principle for controlled stochastic evolution equations, *Int. J. Math. Anal. (Ruse)* **4** (2010), no. 29-32, 1447–1464.

[5] A. Al-Hussein, Necessary conditions for optimal control of stochastic evolution equations in Hilbert spaces, *Appl. Math. Optim.* **63** (2011), no. 3, 385–400.

[6] A. Al-Hussein, Necessary and sufficient conditions of optimal control for infinite dimensional SDEs, *Springer Proc. Math. Stat.*, to appear.

[7] F. Baghery and B. Øksendal, A maximum principle for stochastic control with partial information, *Stoch. Anal. Appl.* **25** (2007), no. 3, 705–717.

[8] S. Bahlali and B. Mezerdi, A general stochastic maximum principle for singular control problems, *Electron. J. Probab.* **10** (2005), no. 30, 988–1004 (electronic).

[9] V. Bally, E. Pardoux and L. Stoica, Backward stochastic differential equations associated to a symmetric Markov process, *Potential Anal*. **22** (2005), no. 1, 17–60.

[10] A. Bensoussan, Lectures on stochastic control, in *Nonlinear filtering and stochastic control (Cortona, 1981)*, 1–62, Lecture Notes in Math., 972 Springer, Berlin.

[11] A. Bensoussan, Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions, *Stochastics* **9** (1983), no. 3, 169–222.

[12] A. Bensoussan, *Stochastic control of partially observable systems*, Cambridge Univ. Press, Cambridge, 1992.

[13] I. Ekeland and R. Temam, *Convex analysis and variational problems*, translated from the French, North-Holland, Amsterdam, 1976.

[14] M. Fuhrman and G. Tessitore, Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control, Ann. Probab. **30** (2002), no. 3, 1397–1465.

[15] W. Grecksch and C. Tudor, *Stochastic evolution equations*, Mathematical Research, 85, Akademie Verlag, Berlin, 1995.

[16] I. Gyöngy and N. V. Krylov, On stochastics equations with respect to semimartingales. II. It*ô* formula in Banach spaces, Stochastics **6** (1981/82), no. 3-4, 153–173.

[17] I. Gyöngy, On stochastic equations with respect to semimartingales. III, Stochastics **7** (1982), no. 4, 231–254.

[18] P. Imkeller, A. Réveillac and A. Richter, Differentiability of quadratic BSDEs generated by continuous martingales, Ann. Appl. Probab. **22** (2012), no. 1, 285–336.

[19] P. Kotelenez, A stopped Doob inequality for stochastic convolution integrals and stochastic evolution equations, *Stochastic Anal. Appl.* **2** (1984), no. 3, 245–265.
[20] N. V. Krylov and B. L. Rozovskii, Stochastic evolution equations, in *Stochastic differential equations: theory and applications*, 1–69, Interdiscip. Math. Sci., 2 World Sci. Publ., Hackensack, NJ.
[21] H. C. Lai and J. C. Lee, Convergent theorems and L^p-selections for Banach-valued multifunctions, *Nihonkai Math. J.* **4** (1993), no. 2, 163–180.
[22] X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Boston, MA, 1995.
[23] M.-A. Morlais, Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem, *Finance Stoch.* **13** (2009), no. 1, 121–150.
[24] M. Métivier, *Semimartingales*, de Gruyter Studies in Mathematics, 2, de Gruyter, Berlin, 1982.
[25] M. Métivier and J. Pellaumail, *Stochastic integration*, Academic Press, New York, 1980.
[26] B. Øksendal, F. Proske and T. Zhang, Backward stochastic partial differential equations with jumps and application to optimal control of random jump fields, *Stochastics* **77** (2005), no. 5, 381–399.
[27] S. G. Peng, A general stochastic maximum principle for optimal control problems, *SIAM J. Control Optim.* **28** (1990), no. 4, 966–979.
[28] S. G. Peng, Backward stochastic differential equations and applications to optimal control, *Appl. Math. Optim.* **27** (1993), no. 2, 125–144.
[29] S. Peszat and J. Zabczyk, *Stochastic partial differential equations with Lévy noise*, Encyclopedia of Mathematics and its Applications, 113, Cambridge Univ. Press, Cambridge, 2007.
[30] B. L. Rozovskiĭ, *Stochastic evolution systems*, translated from the Russian by A. Yarkho, Mathematics and its Applications (Soviet Series), 35, Kluwer Acad. Publ., Dordrecht, 1990.
[31] W. Rudin, *Real and complex analysis*, third edition, McGraw-Hill, New York, 1987.
[32] S. J. Tang and X. J. Li, Maximum principle for optimal control of distributed parameter stochastic systems with random jumps, in *Differential equations, dynamical systems, and control science*, 867–890, Lecture Notes in Pure and Appl. Math., 152 Dekker, New York.
[33] C. Tudor, Optimal control for semilinear stochastic evolution equations, *Appl. Math. Optim.* **20** (1989), no. 3, 319–331.
[34] J. Yong and X. Y. Zhou, *Stochastic controls*, Applications of Mathematics (New York), 43, Springer, New York, 1999.
[35] X. Y. Zhou, On the necessary conditions of optimal controls for stochastic partial differential equations, *SIAM J. Control Optim.* **31** (1993), no. 6, 1462–1478.