Dynamic Systems and Applications 21 (2012) 205-218

MAXIMUM PRINCIPLE FOR OPTIMAL CONTROL OF INFINITE DIMENSIONAL STOCHASTIC DIFFERENTIAL EQUATIONS

ABDULRAHMAN AL-HUSSEIN

Department of Mathematics, College of Science, Qassim University
P. O. Box 6644, Buraydah 51452, Saudi Arabia

ABSTRACT. In this paper we provide necessary and sufficient conditions for optimality of a stochastic differential equation driven by an infinite dimensional martingale, and its solution takes its values in a separable Hilbert space. By using the adjoint equation, which is a backward stochastic differential equation, we derive the maximum principle in the sense of Pontryagin for this optimal control problem.

AMS (MOS) Subject Classification. 60H10, 60G44, 35B50

1. INTRODUCTION

We shall study a stochastic optimal control problem governed by the following stochastic differential equation (SDE shortly):

$$\begin{cases} dX(t) = (a(t,u(t))X(t) + b(t,u(t)))dt \\ \qquad\qquad\qquad + [\langle \sigma(t,u(t)), X(t)\rangle_K + g(t,u(t))]dM(t), \\ X(0) = x_0, \end{cases} \tag{1.1}$$

with bounded and predictable mappings a, b, σ and g, and an admissible control $\{u(t),\ t \geq 0\}$. The space K is a separable Hilbert space with inner product denoted by $\langle \cdot, \cdot \rangle_K$, or sometimes $\langle \cdot, \cdot \rangle$, if no ambiguity occurs. For this equation we shall consider a cost functional of the following type:

$$J(u(\cdot)) = \mathbb{E}\left[\int_0^T \langle \rho(t,u(t)), X(t)\rangle_K \, dt + \langle \theta, X(T)\rangle_K\right].$$

See Section 3 for more details.

Such SDEs with a martingale noise are studied for example in [13], [16], [26], [12] and [4]. The maximum principle for such problems was not studied in these works, except in [4], where we derived some necessary conditions for optimality of stochastic systems more general than (1.1), but the results there provide the maximum principle only in its local form. We also required there the control domain to be convex. In this paper we shall consider a suitable perturbation of an optimal control

Received November 15, 2011

by means of the spike variation method and derive the maximum principle in its global form. The convexity assumption on the control domain will not be required. Our approach here will be achieved by using adjoint equations, which are backward stochastic differential equations (BSDEs) driven by martingales. The main results are recorded in Theorems 4.2, 5.1. The ideas of Bensoussan in [7], [8] and Zhou in [28], [27] together with our earlier work in [3] will be very useful in our present study. Sufficient conditions for optimality for this control problem will be established in Section 5.

It is recorded in [4, Remark 6.4] that studying a controlled stochastic evolution equation, which is more general than (1.1), and when the control variable is allowed to enter in the noise term is still an open problem. This is due to the difficulty of handling the resulting adjoint equations, which are in this particular case BSDEs driven by martingales and contain both first-order and second-order adjoint processes. This raised from the fact that we were dealing there with a BSDE driven by a martingale M, as we are still considering here in equation (1.1), and not merely a Brownian motion. So the work in this paper can be considered as a progress in this direction.

Let us remark that the use of BSDEs for deriving the maximum principle for forward controlled stochastic equations was done first by Bismut in [9]; cf. also [7], [8]. In 1990 Pardoux & Peng, [22], initiated the theory of nonlinear BSDEs. Then Peng studied the stochastic maximum principle in [23] and [24]. Since then several works appeared consequently on the maximum principle and its relationship with BSDEs. For example we refer the reader to [14], and [27] and the references therein. On the other hand, the maximum principle in infinite dimensions started after the work of Pontryagin [25]. One can see also [18].

We shall start by giving some preliminary notation, and introduce in Section 3 our main control problem.

2. NOTATION

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, filtered by a continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$ in the sense that every square integrable K-valued martingale with respect to $\{\mathcal{F}_t\, ,\ 0 \leq t \leq T\}$ has a continuous version.

Denote by $\mathcal{P}$ the *predictable* σ-algebra of subsets of $\Omega \times [0,T]$. A K-valued process is said to be predictable if it is $\mathcal{P}/\mathcal{B}(K)$ measurable. Denote by $\mathcal{M}^2_{[0,T]}(K)$ the Hilbert space of cadlag square integrable martingales $\{M(t), 0 \leq t \leq T\}$ taking their values in K. Let $\mathcal{M}^{2,c}_{[0,T]}(K)$ be the subspace of $\mathcal{M}^2_{[0,T]}(K)$ consisting of all continuous square integrable martingales in K. We say that two elements M and N of $\mathcal{M}^2_{[0,T]}(K)$ are *very strongly orthogonal* (VSO shortly) if $\mathbb{E}\,[M(\tau) \otimes N(\tau)] = \mathbb{E}[M(0) \otimes N(0)]$, for all $[0,T]$-valued stopping times τ.

For $M \in \mathcal{M}^{2,c}_{[0,T]}(K)$ let $< M >$ denote the predictable quadratic variation of M and $<< M >>$ be the predictable tensor quadratic variation of M, which takes its values in the space $L_1(K)$ of all nuclear operators on K. Hence, $M \otimes M- << M >> \in \mathcal{M}^{2,c}_{[0,T]}(L_1(K))$. From here on we shall assume that, for a given fixed $M \in \mathcal{M}^{2,c}_{[0,T]}(K)$, there exists a measurable mapping $\mathcal{Q}(\cdot) : [0,T] \times \Omega \to L_1(K)$ such that $\mathcal{Q}(t)$ is symmetric, positive definite, $\mathcal{Q}(t) \leq \mathcal{Q}$ for some positive definite nuclear operator $\mathcal{Q}$ on K, and satisfies $\ll M \gg_t = \int_0^t \mathcal{Q}(s)\, ds$.

If for (t,ω), $\tilde{\mathcal{Q}}(t,\omega)$ is a symmetric, positive definite nuclear operator on K, we denote by $L_{\tilde{\mathcal{Q}}(t,\omega)}(K)$ to the set of all linear, not necessarily bounded operators Φ which map $\tilde{\mathcal{Q}}^{1/2}(t,\omega)(K)$ into K and satisfy $\Phi\tilde{\mathcal{Q}}^{1/2}(t,\omega) \in L_2(K)$. Here $L_2(K)$ is the space of all Hilbert-Schmidt operators from K into itself with inner product and norm denoted respectively by $\left\langle \cdot,\cdot \right\rangle_2$ and $||\cdot||_2$.

It is known (e.g. [21]) that the stochastic integral $\int_0^\cdot \Phi(s) dM(s)$ is defined for mappings Φ such that for each (t,ω), $\Phi(t,\omega) \in L_{\mathcal{Q}(t,\omega)}(K)$, $\Phi\mathcal{Q}^{1/2}(t,\omega)(h)$ is predictable $\forall\, h \in K$, and

$$\mathbb{E}\left[\int_0^T ||(\Phi\mathcal{Q}^{1/2})(t)||_2^2\, dt\right] < \infty.$$

The space of such integrands is a Hilbert space with respect to the scalar product $(\Phi_1,\Phi_2) \mapsto \mathbb{E}\left[\int_0^T \left\langle \Phi_1\mathcal{Q}^{1/2}(t), \Phi_2\mathcal{Q}^{1/2}(t)\right\rangle dt\right]$. Simple processes with values in $L(K)$ are examples of such integrands. Now denoting by $\Lambda^2(K;\mathcal{P},M)$ to the closure of the set of simple processes in this Hilbert space we obtain a Hilbert subspace. Now the following isometry property is expected:

$$\mathbb{E}\left[|\int_0^T \Phi(s)dM(s)|^2\right] \;=\; \mathbb{E}\left[\int_0^T ||\Phi(s)\mathcal{Q}^{1/2}(s)||_2^2\, ds\right] \tag{2.1}$$

for mappings $\Phi \in \Lambda^2(K;\mathcal{P},M)$. For more details and proofs we refer the reader to [21].

In the case where M is taken to be genuine Wiener process (or cylindrical as well) one should replace $\Lambda^2(K;\mathcal{P},M)$ by a space of the type:

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

where E is a separable Hilbert space.

3. PROBLEM FORMULATION

Let $\mathcal{O}$ be a separable Hilbert space equipped with an inner product $\left\langle \cdot,\cdot \right\rangle_{\mathcal{O}}$, and let U be a nonempty subset of $\mathcal{O}$. We say that $u(\cdot) : [0,T] \times \Omega \to \mathcal{O}$ is *admissible* if $u(\cdot) \in L^2_{\mathcal{F}}(0,T;\mathcal{O})$ and $u(t) \in U$ a.e., a.s. The set of admissible controls will be denoted by $\mathcal{U}_{ad}$.

Let θ be a fixed element of K and consider the following hypothesis:

(A) $a : \Omega \times [0,T] \times \mathcal{O} \to \mathbb{R}$, $b : \Omega \times [0,T] \times \mathcal{O} \to K$, $\sigma : \Omega \times [0,T] \times \mathcal{O} \to K$ and $g : \Omega \times [0,T] \times \mathcal{O} \to L_{\mathcal{Q}}(K)$ are predictable and bounded mappings, and $\rho : [0,T] \times \mathcal{O} \to K$ is a bounded measurable mapping.

Consider the following SDE:

$$\left\{ \begin{array}{l} dX(t) = (a(t,u(t))X(t) + b(t,u(t)))dt \\ \qquad\qquad + [\langle \sigma(t,u(t)), X(t) \rangle + g(t,u(t))]dM(t), \ 0 < t \leq T, \\ X(0) = x_0 \in K. \end{array} \right. \tag{3.1}$$

We shall be interested in minimizing the following cost functional over the set $\mathcal{U}_{ad}$:

$$J(u(\cdot)) = \mathbb{E} \left[\int_0^T \langle \rho(t,u(t)), X^{u(\cdot)}(t) \rangle \, dt + \langle \theta, X^{u(\cdot)}(T) \rangle \right]. \tag{3.2}$$

Any $u^*(\cdot) \in \mathcal{U}_{ad}$ satisfying

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

is called an *optimal control* of the stochastic optimal control problem (3.1)–(3.3), and its corresponding solution $X^* := X^{u^*(\cdot)}$ to (3.1) is called an *optimal solution* of this problem. In this case the pair $(X^*, u^*(\cdot))$ is called an *optimal pair* of the control problem (3.1)–(3.3).

We emphasize here that the case where M is a Wiener process (cylindrical or genuine) is rather much simpler and is treated in [28]. This special case can also be gleaned from the adjoint equation (4.4) in Section 4 by letting M in (3.1) be a cylindrical Wiener process for example and so $N^{u(\cdot)}$ in (4.4) vanishes. Indeed this comes from the martingale representation theorem and the construction of solutions of BSDEs driven by Wiener processes as in [1].

Since this control problem has no constraints we deal here with progressively measurable controls. However, for the case when there are final state constraints, one can mimic our results in Section 4, and use Ekeland's variational principle in a similar way to the works in [20], [23] or [27].

4. MAXIMUM PRINCIPLE

Let us first recall the optimal control problem (3.1)–(3.3). For this control problem we define the *Hamiltonian*

$$H : [0,T] \times \Omega \times K \times \mathcal{O} \times K \times L_2(K) \to \mathbb{R}$$

by

$$\begin{aligned} H(t,\omega,x,u,y,z) \; &= \; \langle \rho(t,u) \, , x \rangle + a(t,\omega,u) \langle x \, , y \rangle + \langle b(t,\omega,u) \, , y \rangle \\ &\quad + \langle B(t,\omega,u)z, x \rangle + \langle g(t,\omega,u) \mathcal{Q}^{1/2}(t,\omega) \, , z \rangle_2 \, , \end{aligned} \tag{4.1}$$

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

$$B(t, \omega, u)z = \left\langle \mathcal{Q}^{1/2}(t, \omega)\,, z \right\rangle_2 \sigma(t, \omega, u), \quad z \in L_2(K). \tag{4.2}$$

Hence it follows that

$$\nabla_x H(t, \omega, x, u, y, z) = \rho(t, u) + a(t, \omega, u)y + B(t, \omega, u)z. \tag{4.3}$$

Furthermore, the adjoint equation of (3.1) is the following BSDE:

$$\begin{cases} -dY^{u(\cdot)}(t) = \nabla_x H(t, X^{u(\cdot)}(t), u(t), Y^{u(\cdot)}(t), Z^{u(\cdot)}(t)\mathcal{Q}^{1/2}(t))dt \\ \qquad\qquad\qquad -Z^{u(\cdot)}(t)dM(t) - dN^{u(\cdot)}(t), \quad 0 \le t < T, \\ Y^{u(\cdot)}(T) = \theta. \end{cases} \tag{4.4}$$

This equation is a BSDE driven by the martingale M. Such types of equations are studied extensively in [2]. We refer the reader also to [19] for financial applications of these types of BSDEs, and to [6], [15], [11] and [5] for other applications.

It is important to recognize that the presence of the process $\mathcal{Q}^{1/2}(\cdot)$ in equation (4.4) 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 if we recall the fact that the integrand Φ in Section 2 does not have to be bounded. This always has to be taken into account when dealing with BSDEs driven by martingales in infinite dimensions.

The proceeding theorem gives the solution to BSDE (4.4) in the sense that there exists a triple $(Y^{u(\cdot)}, Z^{u(\cdot)}, N^{u(\cdot)})$ in $L^2_{\mathcal{F}}(0, T; K) \times \Lambda^2(K; \mathcal{P}, M) \times \mathcal{M}^{2,c}_{[0,T]}(K)$ such that the following equality holds $a.s.$ for all $t \in [0, T]$, $N(0) = 0$ and N is VSO to M:

$$\begin{aligned} Y^{u(\cdot)}(t) \;\; = \;\; & \xi + \int_t^T \nabla_x H(s, X^{u(\cdot)}(s), u(s), Y^{u(\cdot)}(s), Z^{u(\cdot)}(s)\mathcal{Q}^{1/2}(s))\, ds \\ & - \int_t^T Z^{u(\cdot)}(s)dM(s) - \int_t^T dN^{u(\cdot)}(s). \end{aligned}$$

Theorem 4.1. *Assume that (A) holds. Then there exists a unique solution* $(Y^{u(\cdot)}, Z^{u(\cdot)}, N^{u(\cdot)})$ *of BSDE* (4.4).

The proof of this theorem can be found in [2].

We shall denote briefly the solution of (4.4) corresponding to the optimal control $u^*(\cdot)$ by (Y^*, Z^*, N^*).

It would be useful to know that the adjoint equation of an SDE with a general filtration being larger than the Wiener filtration is a BSDE driven by a martingale. This is indeed the case even if the martingale M appearing in the equation (1.1) is a Brownian motion with respect to a right continuous filtration. Our earlier work in [2] on BSDEs will play an important role in deriving the stochastic maximum principle for the control problem (3.1)–(3.3).

The main theorem of this section is the following.

Theorem 4.2. *Suppose that (A) holds. Assume moreover that U is compact and ρ, a, b, σ, g are continuous as mappings in v a.s. If $(X^*, u^*(\cdot))$ is an optimal pair of the problem* (3.1)–(3.3)*, then there exists a unique solution (Y^*, Z^*, N^*) to the corresponding BSDE* (4.4) *such that the following inequality holds:*

$$
\begin{aligned}
(4.5) \qquad & H(t, X^*(t), v, Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)) \\
& \geq H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)) \\
& \qquad \qquad a.e.\ t \in [0, T],\ a.s.\ \forall\, v \in U.
\end{aligned}
$$

Remark 4.3. (i) The compactness assumption of U and the continuity of ρ, 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 minimum of the mapping $v \mapsto H(t, X^*(t), v, Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))$ required in (4.5) to exist in U. (ii) A measurable selection theorem due to Ekeland and Temam, [10], can be applied to select an admissible control satisfying (4.5), one can see also [17, Theorem 3.2, p. 169] for the same purpose.

We shall divide the proof of Theorem 4.2 into different parts. Let us first assume that $(X^*, u^*(\cdot))$ is the given optimal pair. Let $0 \leq t_0 < T$ be fixed such that $\mathbb{E}\,[\,|X^*(t_0)\,|^2] < \infty$, and $0 < \varepsilon < T - t_0$. Let v be a random variable taking its values in U, $\mathcal{F}_{t_0}$-measurable and $\sup_{\omega \in \Omega} |v(\omega)| < \infty$. Consider the following spike variation of the control $u^*(\cdot)$:

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

Let $X^{u_\varepsilon(\cdot)}$ denote the solution of SDE (3.1) corresponding to $u_\varepsilon(\cdot)$. We shall denote it briefly by X_ε. It is easy to see that $X_\varepsilon(t) = X^*(t)$ for all $0 \leq t \leq t_0$. This will be used in the following lemma.

Lemma 4.4. *Under (A), if $\varsigma_\varepsilon(t) = X_\varepsilon(t) - X^*(t),\ t \in [0, T]$, then*

$$
(4.6) \qquad \sup_{t_0 + \varepsilon \leq t \leq T} \mathbb{E}\,[|\varsigma_\varepsilon(t)|^2\,] = O(\varepsilon).
$$

Proof. For $t \in [t_0 + \varepsilon, T]$, we have

$$
\begin{aligned}
(4.7) \qquad \varsigma_\varepsilon(t) \;=\; & \varsigma_\varepsilon(t_0 + \varepsilon) + \int_{t_0+\varepsilon}^{t} a(s, u^*(s))\varsigma_\varepsilon(s) ds \\
& + \int_{t_0+\varepsilon}^{t} \big\langle \sigma(s, u^*(s))\,, \varsigma_\varepsilon(s) \big\rangle\, dM(s).
\end{aligned}
$$

Thus Itô's formula together with assumption (A) implies that

$$
(4.8) \qquad \sup_{t_0 + \varepsilon \leq t \leq T} \mathbb{E}\,[|\varsigma_\varepsilon(t)|^2\,] \leq C_1\, \mathbb{E}\,[|\varsigma_\varepsilon(t_0 + \varepsilon)|^2\,],
$$

for some positive constant C_1. But $\varsigma_\varepsilon(t_0) = 0$ and, for $t_0 \le t \le t_0 + \varepsilon$,

$$\begin{aligned}\varsigma_\varepsilon(t) = \int_{t_0}^{t} \Big[& \big(a(s,v) - a(s,u^*(s))\big) X_\varepsilon(s) \\ & + \big(b(s,v) - b(s,u^*(s))\big) + a(s,u^*(s))\, \varsigma_\varepsilon(s) \Big] ds \\ & + \int_{t_0}^{t} \Big[\big\langle \sigma(s,v) - \sigma(s,u^*(s))\,,\, X_\varepsilon(s) \big\rangle \\ & + \big(g(s,v) - g(s,u^*(s))\big) + \big\langle \sigma(s,u^*(s))\,,\, \varsigma_\varepsilon(s) \big\rangle \Big] dM(s). \end{aligned} \tag{4.9}$$

Hence, again by applying Itô's formula, Cauchy-Schwartz inequality and the boundedness properties in assumption (A) we get

$$\begin{aligned}\mathbb{E}\,[\,|\,\varsigma_\varepsilon(t)\,|^2\,] \le {} & C_2 \cdot (1 + ||\mathcal{Q}^{1/2}||_2^2) \int_{t_0}^{t} \mathbb{E}\,[|\,\varsigma_\varepsilon(s)|^2\,]\, ds \\ & + C_3 \cdot (1 + ||\mathcal{Q}^{1/2}||_2^2 + 1) \cdot \big(\mathbb{E}\,[|\,X^*(t_0)|^2\,] + C_4\, \varepsilon\big)\, \varepsilon + C_5\, \varepsilon, \end{aligned} \tag{4.10}$$

for some positive constants $C_i,\ i = 2, \dots, 5$. Therefore Gronwall's inequality gives

$$\sup_{t_0 \le t \le t_0 + \varepsilon} \mathbb{E}\,[\,|\,\varsigma_\varepsilon(t)\,|^2\,] \le C_6(\varepsilon) \cdot \varepsilon, \tag{4.11}$$

where

$$C_6(\varepsilon) = e^{C_2 \cdot (1 + \|\mathcal{Q}^{1/2}\|_2^2)\varepsilon} \Big(C_3 \cdot (1 + \|\mathcal{Q}^{1/2}\|_2^2) \cdot \big(\mathbb{E}\,[|\,X^*(t_0)|^2\,] + C_4\,\varepsilon\big) + C_5\,\varepsilon \Big).$$

Consequently, applying (4.11) in (4.8) yields (4.6). □

Our next step now is trying to derive a duality formula by computing $\mathbb{E}[\big\langle Y^*(t_0 + \varepsilon),\, \varsigma_\varepsilon(t_0+\varepsilon)\big\rangle]$.

Lemma 4.5. *If (A) holds, then*

$$\begin{aligned} & \mathbb{E}\Big[\big\langle\, Y^*(t_0 + \varepsilon)\,,\, \varsigma_\varepsilon(t_0 + \varepsilon)\,\big\rangle + \int_{t_0}^{t_0+\varepsilon} \big\langle \rho(t, u^*(t))\,,\, \varsigma_\varepsilon(t) \big\rangle dt \Big] \\ & = \mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t)\,,\, \big(a(t,v) - a(t,u^*(t))\big) X_\varepsilon(t) \big\rangle\, dt \Big] \\ & \quad + \mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t)\,,\, b(t,v) - b(t,u^*(t)) \big\rangle\, dt\,\Big] \\ & \quad + \mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,v) - \sigma(t,u^*(t))\,,\, X_\varepsilon(t) \big\rangle \big\langle \mathcal{Q}^{1/2}(t)\,,\, Z^*(t)\mathcal{Q}^{1/2}(t) \big\rangle_2\, dt\,\Big] \\ & \quad + \mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \big(g(t,v) - g(t,u^*(t))\big) \mathcal{Q}^{1/2}(t)\,,\, Z^*(t)\mathcal{Q}^{1/2}(t) \big\rangle_2\, dt\,\Big], \end{aligned} \tag{4.12}$$

and

$$\begin{aligned}\mathbb{E}\,[\,\big\langle\, Y^*(t_0 + \varepsilon), \varsigma_\varepsilon(t_0 + \varepsilon)\,\big\rangle\,] = {} & \mathbb{E}\,\big[\,\big\langle \theta\,,\, \varsigma_\varepsilon(T) \big\rangle\,\big] \\ & + \mathbb{E}\,\Big[\int_{t_0+\varepsilon}^{T} \big\langle \rho(t,u^*(t))\,,\, \varsigma_\varepsilon(t) \big\rangle dt\,\Big]. \end{aligned} \tag{4.13}$$

Proof. The proof of (4.12) follows from applying Itô's formula to (4.9) and (4.4), and then using (4.3) and (4.2). The equality in (4.13) is proved similarly but with the help of (4.7). ☐

Lemma 4.6. *Assume (A). Then*

$$
\begin{aligned}
(4.14)\quad 0 \le \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} & \big\langle \rho(t,v) - \rho(t,u^*(t)) \,, X^*(t) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t) \,, \big(a(t,v) - a(t,u^*(t)) \big) X^*(t) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t) \,, b(t,v) - b(t,u^*(t)) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,v) - \sigma(t,u^*(t)) \,, X^*(t) \big\rangle \big\langle \mathcal{Q}^{1/2}(t) \,, Z^*(t) \mathcal{Q}^{1/2}(t) \big\rangle_2 \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \big(g(t,v) - g(t,u^*(t)) \big) \mathcal{Q}^{1/2}(t) \,, Z^*(t) \mathcal{Q}^{1/2}(t) \big\rangle_2 \, dt \Big] + o(\varepsilon).
\end{aligned}
$$

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

$$
\begin{aligned}
0 &\le J(u_\varepsilon(\cdot)) - J(u^*(\cdot)) \\
&= \mathbb{E}\Big[\int_0^T \big(\big\langle \rho(t,u_\varepsilon(t)) \,, X_\varepsilon(t) \big\rangle - \big\langle \rho(t,u^*(t)) \,, X^*(t) \big\rangle \big) \, dt \Big] \\
&\qquad\qquad + \mathbb{E}\Big[\big\langle \theta \,, X_\varepsilon(T) \big\rangle - \big\langle \theta \,, X^*(T) \big\rangle \Big] \\
&= \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big(\big\langle \rho(t,v) - \rho(t,u^*(t)) \,, X_\varepsilon(t) \big\rangle + \big\langle \rho(t,u^*(t)) \,, \varsigma_\varepsilon(t) \big\rangle \big) \, dt \Big] \\
&\qquad\qquad + \mathbb{E}\Big[\int_{t_0+\varepsilon}^T \big\langle \rho(t,u^*(t)) \,, \varsigma_\varepsilon(t) \big\rangle \, dt + \big\langle \theta, \varsigma_\varepsilon(T) \big\rangle \Big].
\end{aligned}
$$

Thus applying Lemma 4.5 (4.13) in this inequality gives

$$
\begin{aligned}
0 \le \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} & \big(\big\langle \rho(t,v) - \rho(t,u^*(t)) \,, X_\varepsilon(t) \big\rangle + \big\langle \rho(t,u^*(t)) \,, \varsigma_\varepsilon(t) \big\rangle \big) \, dt \Big] \\
&+ \mathbb{E}\Big[\big\langle Y^*(t_0+\varepsilon)) \,, \varsigma_\varepsilon(t_0+\varepsilon) \big\rangle \Big].
\end{aligned}
$$

Also from Lemma 4.5 (4.12) it follows that

$$
\begin{aligned}
(4.15)\quad 0 \le \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} & \big\langle \rho(t,v) - \rho(t,u^*(t)) \,, X_\varepsilon(t) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t) \,, \big(a(t,v) - a(t,u^*(t)) \big) X_\varepsilon(t) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle Y^*(t) \,, b(t,v) - b(t,u^*(t)) \big\rangle \, dt \Big] \\
&+ \mathbb{E}\Big[\int_{t_0}^{t_0+\varepsilon} \big\langle \sigma(t,v) - \sigma(t,u^*(t)) \,, X_\varepsilon(t) \big\rangle \big\langle \mathcal{Q}^{1/2}(t) \,, Z^*(t) \mathcal{Q}^{1/2}(t) \big\rangle_2 \, dt \Big]
\end{aligned}
$$

$$+\,\mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon}\big\langle\big(g(t,v)-g(t,u^*(t))\big)\mathcal{Q}^{1/2}(t)\,,\,Z^*(t)\mathcal{Q}^{1/2}(t)\big\rangle_2\,dt\,\Big].$$

But from assumption (A) and Lemma 4.4 we know that

$$\begin{aligned}
&\frac{1}{\varepsilon}\,\mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon}\big\langle Y^*(t)\,,\,\big(a(t,v)-a(t,u^*(t))\big)\varsigma_\varepsilon(t)\big\rangle\,dt\,\Big] \\
&\le\; C_7\cdot(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\mathbb{E}\,\Big(|Y^*(t)|\cdot|\varsigma_\varepsilon(t)|\Big)\,dt \\
&\le\; C_7\cdot(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\Big((\frac{\varepsilon^{1/3}}{2})\,\mathbb{E}\,[|Y^*(t)|^2]+(\frac{1}{2\,\varepsilon^{1/3}})\,\mathbb{E}\,[|\varsigma_\varepsilon(t)|^2]\Big)\,dt \\
&\le\; C_8\cdot\Big(\varepsilon^{1/3}\,(\frac{1}{\varepsilon})\int_{t_0}^{t_0+\varepsilon}\mathbb{E}\,[|Y^*(t)|^2]dt+(\frac{1}{\varepsilon})\,\varepsilon\,(\frac{1}{\varepsilon^{1/3}})\,\varepsilon\Big)\to 0,
\end{aligned}\tag{4.16}$$

as $\varepsilon\to 0$, if t_0 is a Lebesgue point of the function $t\mapsto\mathbb{E}\,[\,|Y^*(t)\,|^2\,]$, where C_7 and C_8 are some positive constants. Similarly, one can find evidently that

$$\begin{aligned}
&\frac{1}{\varepsilon}\,\mathbb{E}\,\Big[\int_{t_0}^{t_0+\varepsilon}\Big(\big\langle\rho(t,v)-\rho(t,u^*(t))\,,\,\varsigma_\varepsilon(t)\big\rangle \\
&+\big\langle\sigma(t,v)-\sigma(t,u^*(t))\,,\,\varsigma_\varepsilon(t)\big\rangle\,\big\langle\mathcal{Q}^{1/2}(t)\,,\,Z^*(t)\mathcal{Q}^{1/2}(t)\big\rangle_2\Big)\,dt\,\Big]\to 0,
\end{aligned}\tag{4.17}$$

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

Now applying (4.16) and (4.17) in (4.15) yields (4.14). □

In the following we complete the proof of Theorem 4.2.

Proof of Theorem 4.2. Let us divide (4.14) in Lemma 4.6 by ε and let $\varepsilon\to 0$. We then get

$$\begin{aligned}
&\mathbb{E}\,\Big[\,\big\langle\rho(t_0,v)-\rho(t_0,u^*(t_0))\,,\,X^*(t_0)\big\rangle \\
&\qquad\qquad\qquad+\big\langle Y^*(t_0)\,,\,\big(a(t_0,v)-a(t_0,u^*(t_0))\big)X^*(t_0)\big\rangle\,\Big] \\
&+\mathbb{E}\,\big[\,\big\langle Y^*(t_0)\,,\,b(t_0,v)-b(t_0,u^*(t_0))\big\rangle\,\big] \\
&+\mathbb{E}\,\big[\,\big\langle\sigma(t_0,v)-\sigma(t_0,u^*(t_0))\,,\,X^*(t_0)\big\rangle\,\big\langle\mathcal{Q}^{1/2}(t_0)\,,\,Z^*(t_0)\mathcal{Q}^{1/2}(t_0)\big\rangle_2\,\big] \\
&+\mathbb{E}\,\big[\,\big\langle\big(g(t_0,v)-g(t_0,u^*(t_0))\big)\mathcal{Q}^{1/2}(t_0)\,,\,Z^*(t_0)\mathcal{Q}^{1/2}(t_0)\big\rangle_2\,\big]\ge 0.
\end{aligned}$$

So by recalling (4.1) and (4.2) this inequality reads as

$$\begin{aligned}
&\mathbb{E}\,[\,H(t_0,X^*(t_0),v,Y^*(t_0),Z^*(t_0)\mathcal{Q}^{1/2}(t_0))\,] \\
&\qquad\qquad\ge\;\mathbb{E}\,[\,H(t_0,X^*(t_0),u^*(t_0),Y^*(t_0),Z^*(t_0)\mathcal{Q}^{1/2}(t_0))\,].
\end{aligned}$$

Now by a standard argument as in [27, Chapet 3] for instance we deduce that (4.5) holds. The proof of Theorem 4.2 has then been completed. □

5. SUFFICIENT CONDITIONS FOR OPTIMALITY

In the previous section we derived the maximum principle for the optimal control problem (3.1)–(3.3), which gives us some necessary conditions for optimality. If we have also a convexity assumption on the control domain U, we shall also obtain sufficient conditions for optimality for this control problem. We record this result in the following theorem.

Theorem 5.1. *Assume that (A) holds. For a given $u^*(\cdot) \in \mathcal{U}_{ad}$ let X^* and (Y^*, Z^*, N^*) be respectively the corresponding solutions of equations* (3.1) *and* (4.4). *Suppose that the following conditions hold.*
(i) U is a convex domain in $\mathcal{O}$.
(ii) The mapping H satisfies the following two conditions:

1. *$H(t, \cdot, \cdot, Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))$ is convex for all $t \in [0, T]$ a.s.*
2. $$H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)) = \min_{v \in U} H(t, X^*(t), v, Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))$$

for a.e. $t \in [0, T]$ a.s.

Then $(X^*, u^*(\cdot))$ is an optimal pair for the control problem (3.1)–(3.3).

Proof. Let $u(\cdot) \in \mathcal{U}_{ad}$. Consider the following definitions:

$$I_1 := \mathbb{E}\,\Big[\int_0^T \big(\big\langle \rho(t, u^*(t))\,, X^*(t)\big\rangle - \big\langle \rho(t, u(t))\,, X^{u(\cdot)}(t)\big\rangle\big)dt\,\Big],$$

$$I_2 := \mathbb{E}\,\big[\big\langle \theta\,, X^*(T) - X^{u(\cdot)}(T)\big\rangle\big].$$

Then

$$J(u^*(\cdot)) - J(u(\cdot)) = I_1 + I_2 \tag{5.1}$$

and

$$I_2 = \mathbb{E}\,\big[\big\langle\, Y^*(T)\,, X^*(T) - X^u(T)\,\big\rangle\big]. \tag{5.2}$$

Let us next define

$$I_3 := \mathbb{E}\,\Big[\int_0^T \Big(H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)) - H(t, X^{u(\cdot)}(t), u(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))\Big)\,dt\,\Big],$$

$$I_4 := \mathbb{E}\,\Big[\int_0^T \Big(\big\langle a(t, u^*(t))X^*(t) - a(t, u(t))X^{u(\cdot)}(t)\,, Y^*(t)\big\rangle - \big\langle b(t, u^*(t)) - b(t, u(t))\,, Y^*(t)\big\rangle\Big)\,dt\Big]$$

and

$$I_5 := \mathbb{E}\,\Big[\int_0^T \Big(\big\langle B(t, u^*(t))Z^*(t), X^*(t)\big\rangle - \big\langle B(t, u(t))Z^*(t), X^{u(\cdot)}(t)\big\rangle$$

$$+\left\langle \big(g(t,u^*(t)) - g(t,u(t))\big)\mathcal{Q}^{1/2}(t)\,,Z^*(t)\mathcal{Q}^{1/2}(t)\right\rangle_2\Big)\,dt\Big].$$

Then from the definition of H in (4.1) we get easily

$$I_1 = I_3 - I_4 - I_5. \tag{5.3}$$

We next apply Itô's formula to compute $d\left\langle Y^*(t)\,,X^*(t) - X^{u(\cdot)}(t)\right\rangle$ by using equations (4.4) and (3.1) in order to find with the help of (5.2) that

$$I_2 = I_4 + I_5 - I_6, \tag{5.4}$$

where

$$I_6 := \mathbb{E}\Big[\int_0^T \big\langle \nabla_x H(t, X^*(t), u^*(t), Y^*(t), Z^{u^*(\cdot)}(t)\mathcal{Q}^{1/2}(t))\,, \\ X^*(t) - X^{u(\cdot)}(t)\big\rangle\,dt\Big].$$

Now by considering (5.1), (5.3) and (5.4) it follows that

$$J(u^*(\cdot)) - J(u(\cdot)) = I_3 - I_6. \tag{5.5}$$

On the other hand, from the convexity property of the mapping $(x,v)\mapsto H(t,x,u, Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))$ in assumption (ii)(1) the following inequality holds a.s.:

$$\begin{aligned}&\int_0^T \Big(H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)) \\ &\qquad\qquad - H(t, X^{u(\cdot)}(t), u(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))\Big)dt \\ &\le \int_0^T \big\langle\, \nabla_x H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))\,, \\ &\qquad\qquad X^*(t) - X^{u(\cdot)}(t)\,\big\rangle\,dt \\ &+\int_0^T \big\langle\, \nabla_u H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))\,,\, u^*(t) - u(t)\,\big\rangle_{\mathcal{O}}\,dt.\end{aligned}$$

Consequently

$$I_3 \le I_6 + I_7, \tag{5.6}$$

where

$$I_7 = \mathbb{E}\Big[\int_0^T \big\langle\, \nabla_u H(t, X^*(t), u^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t))\,, \\ u^*(t) - u(t)\,\big\rangle_{\mathcal{O}}\,dt\Big].$$

But the minimum condition (ii)(2) shows that $I_7 \le 0$. Hence (5.6) implies that $I_3 - I_6 \le 0$, which together with (5.5) shows that

$$J(u^*(\cdot)) - J(u(\cdot)) \le 0.$$

Finally, since $u(\cdot)\in\mathcal{U}_{ad}$ is arbitrary, $(X^*, u^*(\cdot))$ is an optimal pair for the control problem (3.1)–(3.3). □

Acknowledgement. This author would like to thank Professor N. U. Ahmed for reading the first draft of the paper and providing useful remarks.

REFERENCES

[1] A. Al-Hussein, Backward stochastic differential equations in infinite dimensions and applications, *Arab J. Math. Sc.*, 10, 2:1–42, 2004.

[2] A. Al-Hussein, Backward stochastic partial differential equations driven by infinite dimensional martingales and applications, *Stochastics*, 81, 6:601–626, 2009.

[3] A. Al-Hussein, Sufficient conditions of optimality for backward stochastic evolution equations, *Commun. Stoch. Anal.*, 4, 3:433–442, 2010.

[4] A. Al-Hussein, Necessary conditions for optimal control of stochastic evolution equations in Hilbert spaces, *Appl. Math. Optim.*, 63, 3:385–400, 2011.

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

[6] V. Bally, V., É. Pardoux L. and Stoica, Backward stochastic differential equations associated to a symmetric Markov process, *Potential Anal.*, 22, 1:17–60, 2005.

[7] A. Bensoussan, Lectures on stochastic control. Nonlinear filtering and stochastic control (Cortona, 1981), 1–62, *Lecture Notes in Math.*, 972, Springer, Berlin-New York, 1982.

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

[9] J.-M. Bismut, *Théorie probabiliste du contrôle des diffusions*, *Mem. Amer. Math. Soc.*, 4, 167, 1976.

[10] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems, Amsterdam*, North-Holland, 1976.

[11] 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, 3:1397–1465, 2002.

[12] W. Grecksch and C. Tudor, *Stochastic evolution equations. A Hilbert space approach*, *Mathematical Research*, 85, Akademie-Verlag, Berlin, 1995.

[13] I. Gyöngy and N. V. Krylov, On stochastics equations with respect to semimartingales. I., *Stochastics*, 4, 1:1–21, 1980.

[14] Y. Hu and S.-G. Peng, Maximum principle for optimal control of stochastic system of functional type, *Stochastic Anal. Appl.*, 14, 3:283–301, 1996.

[15] P. Imkeller, A. Reveillac and A. Richter, Differentiability of quadratic BSDEs generated by continuous martingales, *arXiv:0907.0941 [math.PR]*, 2010.

[16] N. V. Krylov and B. Rozovskii, Stochastic evolution equations, Stochastic differential equations: theory and applications, 1–69, *Interdiscip. Math. Sci.*, 2, World Sci. Publ., NJ, Hackensack, 2007.

[17] H. C. Lai and J. C. Lee, Convergent Theorems and L^P-selections for Banach-valued Multifunctions, *Nihonkai Math.J.*, 4:163–180, 1993.

[18] X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems. Systems & Control: Foundations & Applications*, Birkhauser Boston, Inc., Boston, MA, 1995.

[19] M.-A. Morlais, Quadratic BSDEs driven by a continuous martingale and applications to the utility maximization problem, *Finance Stoch.*, 13, 1:121–150, 2009.

[20] L. Mazliak, The maximum principle in stochastic control and backward equations, Backward stochastic differential equations (Paris, 19951996), 101–113, *Pitman Res. Notes Math. Ser.*, 364, Longman, Harlow, 1997.

[21] M. Métivier and J. Pellaumail, *Stochastic integration, Probability and Mathematical Statistics*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London-Toronto, 1980.

[22] É. Pardoux and S.-G. Peng, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, 14, 1:55–61, 1990.

[23] S.-G. Peng, A general stochastic maximum principle for optimal control problems, *SIAM J. Control Optim*, 28, 4:966–979, 1990.

[24] S.-G. Peng, Backward stochastic differential equations and applications to optimal control, *Appl. Math. Optim.*, 27, 2:125–144, 1993.

[25] L. S. Pontryagin, Optimal regulation processes, *Amer. Math. Soc. Transl.*, 18, 2:321–339, 1961.

[26] C. Tudor, Optimal control for semilinear stochastic evolution equations, *Appl. Math. Optim.*, 20, 3:319–331, 1989.

[27] J. Yong and X. Y. Zhou, *Stochastic controls. Hamiltonian systems and HJB equations*, Springer-Verlag, New-York, 1999.

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