Appl Math Optim (2011) 63: 385–400
DOI 10.1007/s00245-010-9125-6

Necessary Conditions for Optimal Control of Stochastic Evolution Equations in Hilbert Spaces

AbdulRahman Al-Hussein

Published online: 6 November 2010

Abstract We consider a nonlinear stochastic optimal control problem associated with a stochastic evolution equation. This equation is driven by a continuous martingale in a separable Hilbert space and an unbounded time-dependent linear operator.

We derive a stochastic maximum principle for this optimal control problem. Our results are achieved by using the adjoint backward stochastic partial differential equation.

Keywords Martingale · Stochastic evolution equation · Stochastic maximum principle · Optimal control · Variational inequality · Adjoint equation · Backward stochastic partial differential equation

1 Introduction and Motivation

We consider a system governed by a stochastic evolution equation (SEE for short) given by

$$\begin{cases} dX(t) = (A(t)X(t) + F(X(t), \nu(t)))dt + G(X(t))dM(t), & t \in [0, T], \\ X(0) = x \in K, \end{cases} \tag{1.1}$$

where $A(t)$, $t \in [0, T]$, is an unbounded closed linear operator on a separable Hilbert space K, M is a continuous martingale in K, F and G are suitable functions to be

Communicating Editor: Alain Bensoussan.

This work was supported by the Science College Research Center at Qassim University, project No. SR-D-010-092.

A. Al-Hussein (✉)
Department of Mathematics, College of Science, Qassim University, P.O. Box 6644, Buraydah 51452, Saudi Arabia
e-mail: alhusseinqu@hotmail.com

introduced later and $\nu(\cdot)$ is a control. These types of SEEs have appeared in [15–17, 20, 21, 34].

We are concerned here with minimizing the cost functional over the set of admissible controls, see (3.3) and (3.4). We will be approaching this through using the theory of backward stochastic partial differential equations (BSPDEs for short) for deriving a stochastic maximum principle for this control problem; cf. the BSPDE (5.2) in Sect. 5. This is due to the fact that we can not use the dynamic programming as we shall shortly tell about it.

Since the existence problem of optimal control is a different problem, we do not include it in this paper. However, a special case can be found in [4]. Some results in this direction can also be developed from the works of [1, 2] and [34].

It is noted that the use of backward stochastic differential equations (BSDEs) in deriving the maximum principle for forward controlled stochastic equations was first discussed by Bismut in [9]. Actually, it was shown there that linear BSDEs may arise from some stochastic control problems and can then be regarded as their resulting adjoint equations. See also Kushner [22, 23] and [18]. Bensoussan in his lecture notes [7] studied similar cases particularly for controlled diffusions, and used somewhat different variational methods. Later Peng in [28, 29] studied the nonlinear BSDEs in order to study the stochastic maximum principle; one can see also [19]. The relationship between BSDEs and the maximum principle for some SDEs is shown in several works, among them for instance are [29] and [35] and the references of Zhou cited therein. Extensive discussions are found in [33] and [36]. Bensoussan in [8, Chap. 8] presents a stochastic maximum principle approach to the problem of stochastic control with partial information treating a general infinite dimensional setting. The adjoint equation as a BSPDE is derived also there. Another work on the maximum principle that is connected to BSDEs can be found also in [6]. An expanded discussion on the history of maximum principle can be found in [35, pp. 153–156].

To work with adjoint BSPDEs (or even with BSDEs if the above operator $A(t) = 0$) it is common to require the filtration furnished for such equations to be the Wiener (Brownian) filtration. This is the case of [35, pp. 114, 116], where such a restriction is made. One can see also our work in [5] in this respect. Recently we proved in [3] the existence and uniqueness of solutions to infinite dimensional BSPDEs, which are driven by infinite dimensional martingales. In our work here we shall drop that restriction and work with an arbitrary right continuous filtration that can also be larger than the Wiener filtration. The work in [3] will be our main tool to overcome the difficulties that may arise when working without such a restriction.

Another motivation is that the dynamic programming approach for a similar optimal control problems associated with a given controlled SDE requires usually a Markov property to be satisfied by its solution, which does not hold in general when the driving noise is a martingale as in the case of the SEE (1.1). So the maximum principle remains a good tool to study the optimality for such controlled SEEs.

On the other hand, SEEs and SPDEs arise naturally from studying various applications in a number of fields that include physics, biology, control theory and finance. We refer the reader to [11, 12, 27, 32] and the references therein and to [10] and [14] for some applications to control theory.

It may be useful to know that the maximum principle in infinite dimensions started after the work of Pontryagin [31]. One can see Li & Yong [24] and the references therein for a detailed description of theses aspects. Besides Tudor [34] and Al-Hussein [4] the work of this paper, where the driving noise is an infinite dimensional martingale, is among the first works and our results here are new.

2 Preliminaries

We will give the necessary background on infinite dimensional martingales and stochastic integration.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a right continuous filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Fix $0 < T < \infty$ and denote by $\mathcal{P}$ the predictable σ-algebra of subsets of $\Omega \times [0, T]$. For a K-valued process we say that it is 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, which is equipped with the inner product $(M, N) \mapsto \mathbb{E}[\langle M(T), N(T)\rangle_K]$. Let $\mathcal{M}^{2,c}_{[0,T]}(K)$ be the subspace of $\mathcal{M}^2_{[0,T]}(K)$ consisting of all square integrable continuous martingales in K. We say that two elements M and N of $\mathcal{M}^2_{[0,T]}(K)$ are *very strongly orthogonal* (*VSO*) if $\mathbb{E}[M(u) \otimes N(u)] = \mathbb{E}[M(0) \otimes N(0)]$, for all $[0, T]$-valued stopping times u.

For $M \in \mathcal{M}^{2,c}_{[0,T]}(K)$ let $\langle M\rangle$ denote the predictable quadratic variation of M and $\tilde{\mathcal{Q}}_M$ be the predictable process associated with the Doléans measure of $M \otimes M$ and takes its values in the space $L_1(K)$, where $L_1(K)$ is the space of nuclear operators on K. Let $\langle\langle M\rangle\rangle_t = \int_0^t \tilde{\mathcal{Q}}_M(s, \omega)\, d\langle M\rangle_s$. Then $M \otimes M - \langle\langle M\rangle\rangle \in \mathcal{M}^{2,c}_{[0,T]}(L_1(K))$. Moreover M is VSO to N if and only if $\langle\langle M, N\rangle\rangle = 0$. More details and proofs of these as well as the following facts can be found in [25] and [26].

For (t, ω) if $\mathcal{Q}(t, \omega)$ is a symmetric, positive definite nuclear operator on K, we denote by $L_{\mathcal{Q}(t,\omega)}(K)$ the set of all linear (not necessarily bounded) operators Φ which map $\mathcal{Q}^{1/2}(t, \omega)(K)$ into K such that $\Phi\mathcal{Q}^{1/2}(t, \omega) \in L_2(K)$, where $L_2(K; H)$ (or shortly $L_2(K)$ when $H = K$) denotes the space of Hilbert-Schmidt operators from K to the Hilbert space H. We shall denote the inner product in $L_2(K)$ by $\langle\cdot, \cdot\rangle_2$ and by $\|\cdot\|_2$ to its norm.

We define the stochastic integral $\int_0^\cdot \Phi(s)\, dM(s)$ for mappings Φ such that for each (t, ω), $\Phi(t, \omega) \in L_{\tilde{\mathcal{Q}}_M(t,\omega)}(K)$, for every $h \in K$ the K-valued process $\Phi \circ \tilde{\mathcal{Q}}_M^{1/2}(h)$ is predictable, and

$$\mathbb{E}\left[\int_0^T \|(\Phi \circ \tilde{\mathcal{Q}}_M^{1/2})(t)\|_2^2\, d\langle M\rangle_t\right] < \infty.$$

The space of such integrands is a Hilbert space with respect to the scalar product $(\Phi_1, \Phi_2) \mapsto \mathbb{E}[\int_0^T \langle\Phi_1 \circ \tilde{\mathcal{Q}}_M^{1/2}, \Phi_2 \circ \tilde{\mathcal{Q}}_M^{1/2}\rangle\, d\langle M\rangle_t]$. Simple processes 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.

Let us next consider the following space:

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

where $\tilde{K}$ is a separable Hilbert space.

We quote the following proposition from [4], which describes the integrands with respect to martingales as in this paper.

Proposition 2.1 *Let $M \in \mathcal{M}^{2,c}_{[0,T]}(K)$. Suppose that there exists a predictable process $\mathcal{Q}(\cdot)$ such that for each (t,ω), $\mathcal{Q}(t,\omega)$ is symmetric, positive definite nuclear operator on K and $\langle\langle M\rangle\rangle_t = \int_0^t \mathcal{Q}(s)\,ds$. If for some positive definite nuclear operator $\mathcal{Q}$ on K, $\mathcal{Q}(t) \le \mathcal{Q}$ for all $t \in [0,T]$ a.s., then*

$$\phi \in L^2_{\mathcal{F}}(0,T; L_2(\mathcal{Q}^{1/2}(K);K)) \quad \Rightarrow \quad \phi \in \Lambda^2(K;\mathcal{P},M)$$

and

$$\mathbb{E}\left[\int_0^T \|\phi \circ \tilde{\mathcal{Q}}_M^{1/2}\|_2^2\, d\langle M\rangle_t\right] \le \mathbb{E}\left[\int_0^T \|\phi\|^2_{\mathcal{Q}}\, dt\right]. \tag{2.1}$$

Proof It is known for example from [13] that for bounded linear mappings $\Psi_i : K \to K$, $i = 1,2$, it is true that $Image(\Psi_1) \subseteq Image(\Psi_2)$ if and only if $\exists c > 0$ such that

$$|\Psi_1^* x| \le c|\Psi_2^* x| \quad \forall x \in K. \tag{2.2}$$

Thus by taking $\Psi_1 = \mathcal{Q}^{1/2}(t)$ (for a fixed t) and $\Psi_2 = \mathcal{Q}^{1/2}$, we find that (2.2) holds with $c = 1$; see for instance [30, Lemma 8.15]. Consequently $Image(\mathcal{Q}^{1/2}(t)) \subseteq Image(\mathcal{Q}^{1/2})$ which implies since t is arbitrary that we can define $\phi(t)$ on any $(\mathcal{Q}^{1/2}(t))(K)$ for all t.

The inequality (2.1) is justified as follows. We observe that $\tilde{\mathcal{Q}}_M(t) = \frac{\mathcal{Q}(t)}{q(t)}$ and so $\langle M\rangle_t = \int_0^t q(s)\,ds$, where $q(t) := \mathrm{tr}(\mathcal{Q}(t))$. Hence

$$\begin{aligned}\mathbb{E}\left[\int_0^T \|\phi(s)\tilde{\mathcal{Q}}_M^{1/2}(t)\|_2^2\, d\langle M\rangle_t\right] &= \mathbb{E}\left[\int_0^T \|\phi(s)\mathcal{Q}^{1/2}(s)\|_2^2\, ds\right] \\ &\le \mathbb{E}\left[\int_0^T \|\phi(s)\mathcal{Q}^{1/2}\|_2^2\, ds\right] \\ &< \infty,\end{aligned}$$

as required. □

The inequality (2.1) and the isometry property (3.1) given below will be used frequently in the sequel.

We close this section by introducing the notion of a rigged Hilbert space (V,K,V'). That is V is a separable Hilbert space embedded continuously and densely in K. Hence by identifying K with its dual, we obtain the continuous and dense two inclusions: $V \subseteq K \subseteq V'$, where V' is the dual space of V. If $v \in V$ and $x \in V'$ we denote by $\langle x, v\rangle$ (or $\langle v, x\rangle$) the duality between V and V'. We observe that $|\langle v,x\rangle| \le \mathrm{const}\,|v|_V \cdot |x|_{V'}$, $\forall v \in V$, $\forall x \in V'$ and $\langle v,x\rangle = \langle v,x\rangle_K$ if $x \in K$.

3 Statement of the Control Problem

Let $\mathcal{O}$ be a separable Hilbert space (e.g. $\mathbb{R}^k$ as in [29]) equipped with an inner product $\langle\cdot,\cdot\rangle_{\mathcal{O}}$, and let U be a nonempty convex subset of $\mathcal{O}$. We say that $\nu(\cdot):[0,T]\times\Omega\to\mathcal{O}$ is *admissible* if $\nu(\cdot)\in L^2_{\mathcal{F}}(0,T;\mathcal{O})$ and $\nu(t)\in U$ *a.e.*, *a.s.* The set of admissible controls will be denoted by $\mathcal{U}_{ad}$.

Let $M\in\mathcal{M}^{2,c}_{[0,T]}(K)$ be such that $M(0)=0$, and there exists a measurable mapping[1] $\mathcal{Q}(\cdot):[0,T]\to L_1(K)$ for which $\mathcal{Q}(t)$ is symmetric, positive definite, $\langle\langle M\rangle\rangle_t=\int_0^t\mathcal{Q}(s)\,ds$, and $\mathcal{Q}(t)\le\mathcal{Q}$ for some positive definite nuclear operator $\mathcal{Q}$ on K. This process $\mathcal{Q}(t)$ is called the *local covariation operator* of the martingale $M(t)$. As in the proof of Proposition 2.1 it follows that $\tilde{\mathcal{Q}}_M(t)=\frac{\mathcal{Q}(t)}{q(t)}$ and $\langle M\rangle_t=\int_0^t q(s)\,ds$, where $q(t):=\operatorname{tr}(\mathcal{Q}(t))$. In particular if $\Phi\in\Lambda^2(K;\mathcal{P},M)$,

$$\mathbb{E}\left[\left|\int_0^T\Phi(s)\,dM(s)\right|^2\right]=\mathbb{E}\left[\int_0^T\|\Phi(s)\mathcal{Q}^{1/2}(s)\|_2^2\,ds\right].\tag{3.1}$$

We shall consider from here on the so-called continuous filtration $\{\mathcal{F}_t,0\le t\le T\}$ in the sense that every square integrable K-valued martingale with respect to it has a continuous version.

Let $F:K\times\mathcal{O}\to K$, $G:K\to L_{\mathcal{Q}}(K)$, $g:K\times\mathcal{O}\to\mathbb{R}$ and $\phi:K\to\mathbb{R}$ be measurable mappings.

Consider the following controlled SEE:

$$\begin{cases} dX(t)=(A(t)X(t)+F(X(t),\nu(t)))\,dt+G(X(t))\,dM(t), & t\in[0,T],\\ X(0)=x\in K, \end{cases}\tag{3.2}$$

for which we shall denote its solution by $X^{\nu(\cdot)}$. Consider also the *cost functional*:

$$J(\nu(\cdot)):=\mathbb{E}\left[\int_0^T g(X^{\nu(\cdot)}(t),\nu(t))\,dt+\phi(X^{\nu(\cdot)}(T))\right],\tag{3.3}$$

where $\nu(\cdot)\in\mathcal{U}_{ad}$.

We shall impose the following assumptions.

(H1) F,G,g,ϕ are continuously Fréchet differentiable with respect to x, F is continuously Fréchet differentiable with respect to ν and the derivatives F_x, F_ν, G_x, g_x are uniformly bounded. Also

$$|\phi_x|_K\le C(1+|x|_K)$$

for some constant $C>0$.

(H2) g_x satisfies Lipschitz condition with respect to ν uniformly in x.

(H3) $A(t,\omega)$ is a linear operator on K, $\mathcal{P}$-measurable, belongs to $L(V;V')$ uniformly in (t,ω) and satisfies the following two conditions.

[1] One can also let $\mathcal{Q}(\cdot):[0,T]\times\Omega\to L_1(K)$ be a predictable process and proceed as in this work without any major difficulties.

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

$$2\langle A(t,\omega)y, y\rangle + \alpha|y|_V^2 \le \lambda|y|^2 \quad a.e.\ t\in[0,T], \quad a.s.\ \forall y\in V,$$

for some $\alpha, \lambda > 0$.

- (ii) $A(t,\omega)$ is uniformly continuous, i.e. $\exists k_1 \ge 0$ such that for all (t,ω)

$$|A(t,\omega)y|_{V'} \le k_1|y|_V \quad \forall y\in V.$$

Under (H1) and (H3) there exists a unique solution to (3.2) in $L^2_{\mathcal{F}}(0,T;K)$. The proof of this fact can be found in [4, Theorem 3.2], [15, Theorem 4.1, p. 105] and can be gleaned from [16].

Our control problem is to minimize (3.3) over $\mathcal{U}_{ad}$. Any $\nu^*(\cdot)\in\mathcal{U}_{ad}$ satisfying

$$J(\nu^*(\cdot)) = \inf\{J(\nu(\cdot)) : \nu(\cdot)\in\mathcal{U}_{ad}\} \tag{3.4}$$

is called an *optimal control*. The corresponding solution $X^{\nu^*(\cdot)}$ of (3.2), which we denote briefly by X^* and $(X^*, \nu^*(\cdot))$ are called respectively an *optimal solution* and an *optimal pair* of the stochastic optimal control problem (3.2)–(3.4).

Since this control problem has no constraints we shall deal generally with progressively measurable controls. However for the case when there are final state constraints one can mimic our result here and use in addition Ekeland's variational principle (see [35]). For our approach here we shall use a certain variation method starting from the following section to derive the maximum principle for (3.2). This result will be established in Sect. 6 below, but we need first to prove some estimates.

4 Estimates

Let $\nu^*(\cdot)$ be an optimal control and let $X^{\nu^*(\cdot)}$ be the corresponding solution of (3.2). Let $\nu(\cdot)$ be such that $\nu^*(\cdot)+\nu(\cdot)\in\mathcal{U}_{ad}$. For a given $0\le\varepsilon\le 1$ consider the variational control:

$$\nu_\varepsilon(t) = \nu^*(t) + \varepsilon\nu(t), \quad t\in[0,T].$$

We note that the convexity of U implies that $\nu_\varepsilon(\cdot)\in\mathcal{U}_{ad}$.

Considering this control $\nu_\varepsilon(\cdot)$ we can consider the $X^{\nu_\varepsilon(\cdot)}$ as the solution of the SEE (3.2) corresponding to $\nu_\varepsilon(\cdot)$. We shall denote it briefly by X_ε. Let p be the solution of the following linear equation:

$$\begin{cases} dp(t) = (A(t)p(t) + F_x(X^*(t),\nu^*(t))p(t))\,dt \\ \quad + F_\nu(X^*(t),\nu^*(t))\nu(t)\,dt + G_x(X^*(t))p(t)\,dM(t), \\ p(0)=0. \end{cases} \tag{4.1}$$

The following three lemmas contain some important estimates that will play a vital role in deriving our desired maximum principle in Sect. 6.

Lemma 4.1 *Assume* (H1) *and* (H3). *Then*

$$\sup_{t\in[0,T]} \mathbb{E}[|p(t)|^2] < \infty.$$

Proof Using Itô's formula we find that

$$\begin{aligned}\mathbb{E}[|p(t)|^2] &= 2\mathbb{E}\left[\int_0^t \langle p(s), A(s)p(s)\rangle\, ds\right]\\ &\quad + 2\mathbb{E}\left[\int_0^t \langle p(s), F_x(X^*(s), \nu^*(s))p(s)\rangle\, ds\right]\\ &\quad + 2\mathbb{E}\left[\int_0^t \langle p(s), F_\nu(X^*(s), \nu^*(s))\nu(s)\rangle\, ds\right]\\ &\quad + \mathbb{E}\left[\int_0^t \|G_x(X^*(s))p(s)\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2\, d\langle M\rangle_s\right].\end{aligned}$$

By applying (H1), Cauchy-Schwartz inequality and Proposition 2.1 we deduce that

$$\begin{aligned}&\mathbb{E}[|p(t)|^2] + \alpha\mathbb{E}\left[\int_0^t |p(s)|_V^2\, ds\right]\\ &\quad \le (\lambda + C_1)\mathbb{E}\left[\int_0^t |p(s)|^2\, ds\right] + C_2\mathbb{E}\left[\int_0^t |\nu(s)|^2\, ds\right],\end{aligned}$$

where C_1 and C_2 are some positive constants.

Thus Gronwall's inequality gives

$$\sup_{t\in[0,T]} \mathbb{E}[|p(t)|^2] \le M_1 \tag{4.2}$$

for some universal constant $M_1 > 0$. □

Lemma 4.2 *Assuming* (H1) *and* (H3) *we have*

$$\sup_{t\in[0,T]} \mathbb{E}[|X_\varepsilon(t) - X^*(t)|^2] = O(\varepsilon^2).$$

Proof Applying Itô's formula, taking expectation and using (H1) yield

$$\begin{aligned}&\mathbb{E}[|X_\varepsilon(t) - X^*(t)|^2] + \alpha\mathbb{E}\left[\int_0^t |X_\varepsilon(s) - X^*(s)|_V^2\, ds\right]\\ &\quad \le (\lambda + 1)\mathbb{E}\left[\int_0^t |X_\varepsilon(s) - X^*(s)|^2\, ds\right]\\ &\qquad + \mathbb{E}\left[\int_0^t |F(X_\varepsilon(s), \nu_\varepsilon(s)) - F(X^*(s), \nu^*(s))|^2\, ds\right]\end{aligned}$$

$$+\mathbb{E}\bigg[\int_0^t \|(G(X_\varepsilon(s))-G(X^*(s)))\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2\,d\langle M\rangle_s\bigg]. \tag{4.3}$$

But from (H1) we get

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^t |F(X_\varepsilon(s),\nu_\varepsilon(s))-F(X^*(s),\nu^*(s))|^2\,ds\bigg]\\
&\quad\le 2\mathbb{E}\bigg[\int_0^t |F(X_\varepsilon(s),\nu_\varepsilon(s))-F(X^*(s),\nu_\varepsilon(s))|^2\,ds\bigg]\\
&\qquad+2\mathbb{E}\bigg[\int_0^t |F(X^*(s),\nu_\varepsilon(s))-F(X^*(s),\nu^*(s))|^2\,ds\bigg]\\
&\quad=2\mathbb{E}\bigg[\int_0^t |\tilde{F}_x(s,\varepsilon)(X_\varepsilon(s)-X^*(s))|^2\,ds\bigg]+2\mathbb{E}\bigg[\int_0^t |\delta_\varepsilon F(s)|^2\,ds\bigg]\\
&\quad\le 2C_4\mathbb{E}\bigg[\int_0^t |X_\varepsilon(s)-X^*(s)|^2\,ds\bigg]+2C_3\varepsilon^2,
\end{aligned} \tag{4.4}$$

where, for $y\in K$,

$$\tilde{F}_x(s,\varepsilon)(y)=\int_0^1 F_x(X^*(s)+\theta(X_\varepsilon(s)-X^*(s)),\nu_\varepsilon(s))(y)\,d\theta$$

and

$$\delta_\varepsilon F(s)=F(X^*(s),\nu_\varepsilon(s))-F(X^*(s),\nu^*(s)).$$

We have used here with the help of (H1) the following inequality:

$$\begin{aligned}
\mathbb{E}\bigg[\int_0^T |\delta_\varepsilon F(s)|^2\,ds\bigg]&=\mathbb{E}\bigg[\int_0^T |F(X^*(s),\nu_\varepsilon(s))-F(X^*(s),\nu^*(s))|^2\,ds\bigg]\\
&=\mathbb{E}\bigg[\int_0^T \bigg|\int_0^1 F_\nu(X^*(s),\nu^*(s)+\theta(\nu_\varepsilon(s)-\nu^*(s)))\\
&\qquad\times(\nu_\varepsilon(s)-\nu^*(s))\,d\theta\bigg|^2\,ds\bigg]\\
&\le\varepsilon^2 C_3.
\end{aligned} \tag{4.5}$$

Similarly Proposition 2.1 and (H1) imply

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^t \|(G(X_\varepsilon(s))-G(X^*(s)))\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2\,d\langle M\rangle_s\bigg]\\
&\quad=\mathbb{E}\bigg[\int_0^t \|\tilde{G}_x(s,\varepsilon)(X_\varepsilon(s)-X^*(s))\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2\,d\langle M\rangle_s\bigg]\\
&\quad\le C_5\mathbb{E}\bigg[\int_0^t |X_\varepsilon(s)-X^*(s)|^2\,ds\bigg],
\end{aligned} \tag{4.6}$$

where, for $y \in K$,

$$\tilde{G}_x(s,\varepsilon)(y) = \int_0^1 [G_x(X^*(s) + \theta(X_\varepsilon(s) - X^*(s)))(y)\tilde{\mathcal{Q}}_M^{1/2}(s)]\, d\theta.$$

It follows by applying (4.4), (4.6) in (4.3) and using Gronwall's inequality that

$$\mathbb{E}[|X_\varepsilon(t) - X^*(t)|^2] \leq M_2 \varepsilon^2 \tag{4.7}$$

for some universal constant $M_2 > 0$. □

Lemma 4.3 *Assume* (H1) *and* (H3). *Let* $\eta_\varepsilon(t) = \frac{X_\varepsilon(t) - X^*(t)}{\varepsilon} - p(t)$. *Then*

$$\lim_{\varepsilon \to 0^+} \sup_{t \in [0,T]} \mathbb{E}[|\eta_\varepsilon(t)|^2] = 0.$$

Proof Using the corresponding equations (3.2) and (4.1) we deduce as in the proof of Lemma 4.2 that

$$\begin{aligned}
\eta_\varepsilon(t) = \int_0^t & A(s)\eta_\varepsilon(s)\, ds \\
&+ \int_0^t [F(X_\varepsilon(s), \nu_\varepsilon(s)) - F(X^*(s), \nu^\varepsilon(s)) - F_x(X^*(s), \nu^*(s))p(s)]\, ds \\
&+ \int_0^t \left[\frac{1}{\varepsilon}\delta_\varepsilon F(s) - F_\nu(X^*(s), \nu^*(s))\nu(s)\right] ds \\
&+ \int_0^t [G(X_\varepsilon(s)) - G(X^*(s)) - G_x(X^*(s))p(s)]\, dM(s), \\
= \int_0^t & A(s)\eta_\varepsilon(s)\, ds \\
&+ \int_0^t [\tilde{F}_x(s,\varepsilon)\eta_\varepsilon(s) + (\tilde{F}_x(s,\varepsilon) - F_x(X^*(s), \nu^*(s)))p(s)]\, ds \\
&+ \int_0^t \left[\frac{1}{\varepsilon}\delta_\varepsilon F(s) - F_\nu(X^*(s), \nu^*(s))\nu(s)\right] ds \\
&+ \int_0^t [\tilde{G}_x(s,\varepsilon)\eta_\varepsilon(s) + (\tilde{G}_x(s,\varepsilon) - G_x(X^*(s)))p(s)]\, dM(s).
\end{aligned} \tag{4.8}$$

Consequently by using Itô's formula, (H1) and (H3) it follows that, for each $t \in [0, T]$,

$$\mathbb{E}[|\eta_\varepsilon(t)|^2] + \alpha\mathbb{E}\left[\int_0^t |\eta_\varepsilon(s)|_V^2\, ds\right] \leq (\lambda + C_6)\int_0^t \mathbb{E}[|\eta_\varepsilon(s)|^2]\, ds + \rho(\varepsilon), \tag{4.9}$$

where

$$\rho(\varepsilon) = 2\mathbb{E}\left[\int_0^T |(\tilde{F}_x(s,\varepsilon) - F_x(X^*(s), \nu^*(s)))p(s)|^2\, ds\right]$$

$$+ 2\mathbb{E}\bigg[\int_0^T \|(\tilde{G}_x(s,\varepsilon) - G_x(X^*(s)))p(s)\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2 \, d\langle M\rangle_s\bigg]$$
$$+ \mathbb{E}\bigg[\int_0^T \bigg|\frac{1}{\varepsilon}\delta_\varepsilon F(s) - F_\nu(X^*(s), \nu^*(s))\nu(s)\bigg|^2 ds\bigg]. \tag{4.10}$$

But (H1), (4.2) and the dominated convergence theorem give

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T |(\tilde{F}_x(s,\varepsilon) - F_x(X^*(s), \nu^*(s)))p(s)|^2 \, ds\bigg] \\
&\quad = \mathbb{E}\bigg[\int_0^T \bigg|\int_0^1 \big(F_x(X^*(s) + \theta(X_\varepsilon(s) - X^*(s)), \nu_\varepsilon(s)) \\
&\qquad - F_x(X^*(s), \nu^*(s))\big)p(s)\, d\theta\bigg|^2 ds\bigg] \\
&\quad \le \int_0^T \int_0^1 \mathbb{E}\Big[\Big|\big(F_x(X^*(s) + \theta(X_\varepsilon(s) - X^*(s)), \nu_\varepsilon(s)) \\
&\qquad - F_x(X^*(s), \nu^*(s))\big)p(s)\Big|^2\Big] d\theta \, ds \\
&\quad \to 0, \quad \text{as } \varepsilon \to 0^+.
\end{aligned}$$

Similarly we have

$$\mathbb{E}\bigg[\int_0^T \|(\tilde{G}_x(s,\varepsilon) - G_x(X^*(s)))p(s)\tilde{\mathcal{Q}}_M^{1/2}(s)\|_2^2 \, d\langle M\rangle_s\bigg] \to 0,$$

as $\varepsilon \to 0^+$.

On the other hand, as done for (4.5),

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T \bigg|\frac{1}{\varepsilon}\delta_\varepsilon F(s) - F_\nu(X^*(s), \nu^*(s))\nu(s)\bigg|^2 ds\bigg] \\
&\quad \le \int_0^T \int_0^1 \mathbb{E}\Big[\Big|\big(F_\nu(X^*(s), \nu^*(s) + \theta(\nu_\varepsilon(s) - \nu^*(s))) \\
&\qquad - F_\nu(X^*(s), \nu^*(s))\big)\nu(s)\Big|^2\Big] d\theta \, ds \to 0,
\end{aligned} \tag{4.11}$$

if $\varepsilon \to 0^+$, by using (H1) and the dominated convergence theorem.

Finally applying these three results in (4.10) shows that

$$\rho(\varepsilon) \to 0, \quad \text{as } \varepsilon \to 0^+,$$

whence from (4.9) and Gronwall's inequality we obtain

$$\sup_{t\in[0,T]} \mathbb{E}[|\eta_\varepsilon(t)|^2] \to 0, \quad \text{as } \varepsilon \to 0^+,$$

as required. □

Remark 4.4 From the preceding lemma one can obtain that

$$\sup_{t\in[0,T]} \mathbb{E}[|X_\varepsilon(t) - X^*(t) - \varepsilon p(t)|^2] = o(\varepsilon^2).$$

In the following we shall try to derive our variational inequality.

Theorem 4.5 *We suppose* (H1)–(H3). *Then for each* $\varepsilon > 0$,

$$\begin{aligned} J(v_\varepsilon(\cdot)) - J(v^*(\cdot)) = \varepsilon\mathbb{E}[\phi_x(X^*(T))p(T)] \\ + \varepsilon\mathbb{E}\bigg[\int_0^T g_x(X^*(s), v^*(s))p(s)\,ds\bigg] \\ + \mathbb{E}\bigg[\int_0^T \big(g(X^*(s), v_\varepsilon(s)) - g(X^*(s), v^*(s))\big)\,ds\bigg] + o(\varepsilon). \end{aligned} \tag{4.12}$$

Proof Note that

$$J(v_\varepsilon(\cdot)) - J(v^*(\cdot)) = I_1(\varepsilon) + I_2(\varepsilon), \tag{4.13}$$

where

$$I_1(\varepsilon) = \mathbb{E}[\phi(X_\varepsilon(T)) - \phi(X^*(T))]$$

and

$$I_2(\varepsilon) = \mathbb{E}\bigg[\int_0^T \big(g(X_\varepsilon(s), v_\varepsilon(s)) - g(X^*(s), v^*(s))\big)\,ds\bigg].$$

It is easy to see from Lemma 4.3, Lemma 4.2 and the dominated convergence theorem that

$$\begin{aligned} \frac{1}{\varepsilon}I_1(\varepsilon) &= \frac{1}{\varepsilon}\mathbb{E}\bigg[\int_0^1 \phi_x(X^*(T) + \theta(X_\varepsilon(T) - X^*(T))(X_\varepsilon(T) - X^*(T))\,d\theta\bigg] \\ &= \mathbb{E}\bigg[\int_0^1 \phi_x(X^*(T) + \theta(X_\varepsilon(T) - X^*(T))(p(T) + \eta_\varepsilon(T))\,d\theta\bigg] \\ &\to \mathbb{E}[\phi_x(X^*(T))p(T)], \quad \text{as } \varepsilon \to 0^+. \end{aligned}$$

In particular

$$I_1(\varepsilon) = \varepsilon\mathbb{E}[\phi_x(X^*(T))p(T)] + o(\varepsilon). \tag{4.14}$$

Similarly

$$\frac{1}{\varepsilon}I_2(\varepsilon) = \frac{1}{\varepsilon}\mathbb{E}\bigg[\int_0^T \big(g(X_\varepsilon(s), v_\varepsilon(s)) - g(X^*(s), v_\varepsilon(s))\big)\,ds\bigg]$$

$$+\frac{1}{\varepsilon}\mathbb{E}\bigg[\int_0^T \big(g(X^*(s), \nu_\varepsilon(s)) - g(X^*(s), \nu^*(s))\big)\,ds\bigg]$$
$$= \mathbb{E}\bigg[\int_0^T \int_0^1 g_x(X^*(s) + \theta(X_\varepsilon(s) - X^*(s)), \nu_\varepsilon(s))(p(s) + \eta_\varepsilon(s))\,d\theta\,ds\bigg]$$
$$+\frac{1}{\varepsilon}\mathbb{E}\bigg[\int_0^T \big(g(X^*(s), \nu_\varepsilon(s)) - g(X^*(s), \nu^*(s))\big)\,ds\bigg].$$

But by applying Lemma 4.3, Lemma 4.2, the continuity and boundedness of g_x in (H1) and (H2) and the dominated convergence theorem we can derive

$$\mathbb{E}\bigg[\int_0^T \int_0^1 g_x(X^*(s) + \theta(X_\varepsilon(s) - X^*(s)), \nu^*(s) + \varepsilon\nu(s))(p(s) + \eta_\varepsilon(s))\,d\theta\,ds\bigg]$$
$$\to \mathbb{E}\bigg[\int_0^T g_x(X^*(s), \nu^*(s))p(s)\,ds\bigg].$$

Thus in particular

$$I_2(\varepsilon) = \varepsilon\mathbb{E}\bigg[\int_0^T g_x(X^*(s), \nu^*(s))p(s)\,ds\bigg] + \mathbb{E}\bigg[\int_0^T \big(g(X^*(s), \nu_\varepsilon(s)) - g(X^*(s), \nu^*(s))\big)\,ds\bigg] + o(\varepsilon). \tag{4.15}$$

The theorem then follows from (4.13), (4.14) and (4.15). □

5 The Adjoint Equation

Recall the SEE (3.2) and the mappings in (3.3), and let the *Hamiltonian* $H : [0,T] \times K \times \mathcal{O} \times K \times L_2(K) \to \mathbb{R}$ be defined by

$$H(t, x, \nu, y, z) = -g(x, \nu) + \big\langle F(x, \nu), y\big\rangle + \big\langle G(x)Q^{1/2}(t), z\big\rangle_2, \tag{5.1}$$

where $(t, x, \nu, y, z) \in [0,T] \times K \times \mathcal{O} \times K \times L_2(K)$.

The adjoint equation of (3.2) is the following BSPDE:

$$\begin{cases} -dY^{\nu(\cdot)}(t) = [A^*(t)Y^{\nu(\cdot)}(t) \\ \qquad\qquad + \nabla_x H(X^{\nu(\cdot)}(t), \nu(t), Y^{\nu(\cdot)}(t), Z^{\nu(\cdot)}(t)Q^{1/2}(t))]\,dt \\ \qquad\qquad - Z^{\nu(\cdot)}(t)\,dM(t) - dN^{\nu(\cdot)}(t), \quad 0 \le t \le T, \\ Y^{\nu(\cdot)}(T) = -\nabla\phi(X^{\nu(\cdot)}(T)), \end{cases} \tag{5.2}$$

where $\nabla\phi$ denotes the gradient of ϕ, which is defined, by using the directional derivative $D\phi(x)(h)$ of ϕ at a point $x \in K$ in the direction of $h \in K$, as $\langle\nabla\phi(x), h\rangle_K = D\phi(x)(h)(= \phi_x(h))$. The operator $A^*(t)$ is the adjoint operator of $A(t)$.

The following theorem gives the unique solution to this BSPDE (5.2).

Theorem 5.1 *Assume that* (H1)–(H3) *hold. Then there exists a unique solution* $(Y^{\nu(\cdot)}, Z^{\nu(\cdot)}, N^{\nu(\cdot)})$ *of the BSPDE* (5.2) *in* $L^2_{\mathcal{F}}(0,T;V) \times \Lambda^2(K;\mathcal{P},M) \times \mathcal{M}^{2,c}_{[0,T]}(K)$, *which satisfies* $N^{\nu(\cdot)}(0) = 0$ *and* $N^{\nu(\cdot)}$ *is VSO to* M.

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

We shall denote briefly the solution of (5.2) corresponding to the optimal control $\nu^*(\cdot)$ by (Y^*, Z^*, N^*).

6 Maximum Principle

In Theorem 4.5 we derived our main variational inequality. This inequality will play a vital role in proving the maximum principle as we shall see in this section.

We start by giving another variational inequality.

Lemma 6.1 *Suppose that* (H1)–(H3) *hold. Then*

$$-\varepsilon\mathbb{E}\big\langle Y^*(T), p(T)\big\rangle + \varepsilon\mathbb{E}\bigg[\int_0^T g_x(X^*(s),\nu^*(s))p(s)\,ds\bigg] + \mathbb{E}\bigg[\int_0^T \big(-\delta_\varepsilon H(s) + \big\langle Y^*(s), \delta_\varepsilon F(s)\big\rangle\big)\,ds\bigg] \geq o(\varepsilon), \tag{6.1}$$

where

$$\delta_\varepsilon H(s) = H(X^*(s),\nu_\varepsilon(s),Y^*(s),Z^*(s)Q^{1/2}(s)) - H(X^*(s),\nu^*(s),Y^*(s),Z^*(s)Q^{1/2}(s)).$$

Proof Since $\nu^*(\cdot)$ is an optimal control, then $J(\nu_\varepsilon(\cdot)) - J(\nu^*(\cdot)) \geq 0$. Thus the rest follows from (4.12) and (5.1). □

We need the following relation.

Lemma 6.2

$$\mathbb{E}\big\langle Y^*(T), p(T)\big\rangle = \mathbb{E}\bigg[\int_0^T g_x(X^*(s),\nu^*(s))p(s)\,ds\bigg] + \mathbb{E}\bigg[\int_0^T \big\langle Y^*(s), F_\nu(X^*(s),\nu^*(s))\nu(s)\big\rangle\,ds\bigg]. \tag{6.2}$$

Proof Use Itô's formula together with the fact that

$$\begin{aligned}&\big\langle \nabla_x H(X^*(t),\nu^*(t),Y^*(t),Z^*(t)Q^{1/2}(t)), p(t)\big\rangle \\ &\quad = -g_x(X^*(t),\nu^*(t))p(t) + \big\langle F_x(X^*(t),\nu^*(t))p(t), Y^*(t)\big\rangle \\ &\qquad + \big\langle G(X^*(t))Q^{1/2}(t), Z^*(t)Q^{1/2}(t)\big\rangle_2 \quad \text{a.s. } \forall t\in[0,T]\end{aligned}$$

and

$$\int_0^T \big\langle G(X^*(s))dM(s), Z^*(s)\,dM(s)\big\rangle$$
$$= \int_0^T \big\langle G(X^*(s))\mathcal{Q}^{1/2}(s), Z^*(s)\mathcal{Q}^{1/2}(s)\big\rangle_2\,ds \quad \text{a.s.},$$

which can be gleaned from the proof of Proposition 2.1. □

We are now ready to state our main result.

Theorem 6.3 *Suppose* (H1)–(H3). *Assume that* $(X^*, \nu^*(\cdot))$ *is an optimal pair for the problem* (3.2)–(3.4). *Then there exists a unique triple* (Y^*, Z^*, N^*) *solving the corresponding adjoint BSPDE* (5.2) *such that the following variational inequality holds*:

$$\big\langle \nabla_\nu H(t, X^*(t), \nu^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)), \nu^*(t) - \nu\big\rangle_{\mathcal{O}} \geq 0,$$
$$\forall \nu \in U,\ a.e.\ t \in [0,T],\ a.s. \tag{6.3}$$

Proof Consider the BSPDE (5.2). With the help of Theorem 5.1 there exists a unique solution (Y^*, Z^*, N^*) to (5.2). Hence it remains to prove (6.3).

From (6.1) and (6.2) we derive

$$\mathbb{E}\bigg[\int_0^T \Big(\big\langle Y^*(s), \delta_\varepsilon F(s) - \varepsilon F_\nu(X^*(s), \nu^*(s))\nu(s)\big\rangle - \delta_\varepsilon H(s)\Big)\,ds\bigg] \geq o(\varepsilon).$$

Moreover as we did for (4.11), by using the continuity and boundedness of F_ν in (H1) and the dominated convergence theorem, it follows that

$$\frac{1}{\varepsilon}\mathbb{E}\bigg[\int_0^T \big\langle Y^*(s), \delta_\varepsilon F(s) - \varepsilon F_\nu(X^*(s), \nu^*(s))\nu(s)\big\rangle\bigg]\,ds$$
$$= \mathbb{E}\bigg[\int_0^T \Big\langle Y^*(s), \int_0^1 \Big(F_\nu(X^*(s), \nu^*(s) + \theta(\nu_\varepsilon(s) - \nu^*(s)))$$
$$- F_\nu(X^*(s), \nu^*(s))\Big)\nu(s)\,d\theta\Big\rangle\,ds\bigg] \quad \to \quad 0,$$

as $\varepsilon \to 0^+$. In particular

$$\mathbb{E}\bigg[\int_0^T \big\langle Y^*(s), \delta_\varepsilon F(s) - \varepsilon F_\nu(X^*(s), \nu^*(s))\nu(s)\big\rangle\,ds\bigg] = o(\varepsilon).$$

Hence

$$-\mathbb{E}\bigg[\int_0^T \delta_\varepsilon H(s)\,ds\bigg] \geq o(\varepsilon). \tag{6.4}$$

Therefore dividing (6.4) by ε and letting $\varepsilon \to 0^+$ yield:

$$\mathbb{E}\left[\int_0^T \left\langle \nabla_\nu H(t, X^*(t), \nu^*(t), Y^*(t), Z^*(t)\mathcal{Q}^{1/2}(t)), \nu(t)\right\rangle_{\mathcal{O}} dt\right] \le 0.$$

As a result by arguing as in [8, p. 280] we deduce (6.3). □

Remark 6.4 It would be intersecting to know if in the SEE (3.2) a control variable $\nu(\cdot)$ can be allowed to enter in the domain of the mapping G, which is the integrand of the stochastic integral with respect to the martingale M. This case would require new results concerning existence and uniqueness of 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 [35] and [28] would be useful in this direction. 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 $\mathcal{M}^2_{[0,T]}(K)$, for which we have to take into account that certain properties of a very strong orthogonality and a general martingale representation in parallel to those in [3] must also be present for this research problem.

Acknowledgement I would like to thank the referee(s) for the useful remarks and suggestions which helped in improving further the first version of the paper.

References

1. Ahmed, N.U.: Existence of optimal controls for a class of systems governed by differential inclusions on a Banach space. J. Optim. Theory Appl. **50**(2), 213–237 (1986)
2. Ahmed, N.U.: Relaxed controls for stochastic boundary value problems in infinite dimension. In: Optimal Control of Partial Differential Equations, Irsee, 1990. Lecture Notes in Control and Inform. Sci., vol. 149, pp. 1–10. Springer, Berlin (1991)
3. Al-Hussein, A.: Backward stochastic partial differential equations driven by infinite dimensional martingales and applications. Stochastics **81**(6), 601–626 (2009)
4. Al-Hussein, A.: Maximum principle for controlled stochastic evolution equations. Int. J. Math. Analysis **4**(30), 1447–1464 (2010)
5. Al-Hussein, A.: Sufficient conditions of optimality for backward stochastic evolution equations. Commun. Stoch Analysis **4**(3), 433–442 (2010)
6. Bahlali, S., Mezerdi, B.: A general stochastic maximum principle for singular control problems. Electron. J. Probab. **10**(30), 988–1004 (2005)
7. Bensoussan, A.: Lectures on stochastic control. In: Nonlinear Filtering and Stochastic Control, Cortona, 1981. Lecture Notes in Math., vol. 972, pp. 1–62. Springer, Berlin (1982)
8. Bensoussan, A.: Stochastic Control of Partially Observable Systems. Cambridge University Press, Cambridge (1992)
9. Bismut, J.-M.: Théorie probabiliste du contrôle des diffusions. Mem. Am. Math. Soc. **4**(167) (1976)
10. Cerrai, S.: Second Order PDE's in Finite and Infinite Dimension. A Probabilistic Approach. Lecture Notes in Mathematics, vol. 1762. Springer, Berlin (2001)
11. Chow, P.-L.: Stochastic Partial Differential Equations. Applied Mathematics and Nonlinear Science Series. Chapman & Hall/CRC, Boca Raton (2007)
12. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications, vol. 44. Cambridge University Press, Cambridge (1992)
13. Douglas, R.G.: On majorization, factorization, and range inclusion of operators on Hilbert space. Proc. Am. Math. Soc. **17**, 413–415 (1966)
14. Fuhrman, M., Tessitore, G.: 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)

15. Grecksch, W., Tudor, C.: Stochastic Evolution Equations. A Hilbert Space Approach. Mathematical Research, vol. 85. Akademie-Verlag, Berlin (1995)
16. Gyöngy, I., Krylov, N.V.: On stochastics equations with respect to semimartingales. I. Stochastics **4**(1), 1–21 (1980)
17. Gyöngy, I., Krylov, N.V.: On stochastics equations with respect to semimartingales. II. Itô formula in Banach spaces. Stochastics **6**(3–4), 153–173 (1982)
18. Haussmann, U.G.: A Stochastic Maximum Principle for Optimal Control of Diffusions. Pitman Research Notes in Mathematics Series, vol. 151. Longman Scientific & Technical, Harlow (1986)
19. Hu, Y., Peng, S.G.: Maximum principle for optimal control of stochastic system of functional type. Stoch. Anal. Appl. **14**(3), 283–301 (1996)
20. Kotelenez, P.: A stopped Doob inequality for stochastic convolution integrals and stochastic evolution equations. Stoch. Anal. Appl. **2**(3), 245–265 (1984)
21. Krylov, N.V., Rozovskii, B.: Stochastic evolution equations. In: Stochastic Differential Equations: Theory and Applications. Interdiscip. Math. Sci., vol. 2, pp. 1–69. World Sci., Hackensack (2007)
22. Kushner, H.J.: On the stochastic maximum principle: Fixed time of control. J. Math. Anal. Appl. **11**, 78–92 (1965)
23. Kushner, H.J.: Necessary conditions for continuous parameter stochastic optimization problems. SIAM J. Control **10**, 550–565 (1972)
24. Li, X.J., Yong, J.M.: Optimal control theory for infinite-dimensional systems. In: Systems & Control: Foundations & Applications. Birkhäuser Boston, Boston (1995)
25. Métivier, M.: Semimartingales. A Course on Stochastic Processes. de Gruyter Studies in Mathematics, vol. 2. Walter de Gruyter & Co., Berlin (1982)
26. Métivier, M., Pellaumail, J.: Stochastic Integration, Probability and Mathematical Statistics. Academic Press, New York (1980)
27. Métivier, M.: Stochastic Partial Differential Equations in Infinite-Dimensional Spaces. Scuola Normale Superiore, Pisa (1988)
28. Peng, S.G.: A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. **28**(4), 966–979 (1990)
29. Peng, S.G.: Backward stochastic differential equations and applications to optimal control. Appl. Math. Optim. **27**(2), 125–144 (1993)
30. Peszat, S., Zabczyk, J.: Stochastic Partial Differential Equations with Lévy Noise. An Evolution Equation Approach. Encyclopedia of Mathematics and Its Applications, vol. 113. Cambridge University Press, Cambridge (2007)
31. Pontryagin, L.S.: Optimal regulation processes. Am. Math. Soc. Transl. (2) **18**, 321–339 (1961)
32. Rozovskiĭ, B.L.: Stochastic Evolution Systems. Linear Theory and Applications to Nonlinear Filtering. Translated from the Russian by A. Yarkho. Mathematics and Its Applications (Soviet Series), vol. 35. Kluwer Academic, Dordrecht (1990)
33. Tang, S., Li, X.: Maximum principle for optimal control of distributed parameter stochastic systems with random jumps. In: Differential Equations, Dynamical Systems, and Control Science. Lecture Notes in Pure and Appl. Math., vol. 152, pp. 867–890. Dekker, New York (1994)
34. Tudor, C.: Optimal control for semilinear stochastic evolution equations. Appl. Math. Optim. **20**(3), 319–331 (1989)
35. Yong, J., Zhou, X.Y.: Stochastic Controls. Hamiltonian Systems and HJB Equations. Springer, New York (1999)
36. Zhou, X.Y.: On the necessary conditions of optimal controls for stochastic partial differential equations. SIAM J. Control Optim. **31**(6), 1462–1478 (1993)