SIAM J. CONTROL AND OPTIMIZATION
Vol. 31, No. 6, pp. 1462–1478, November 1993

ON THE NECESSARY CONDITIONS OF OPTIMAL CONTROLS FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS*

XUN YU ZHOU†

Abstract. This paper concerns optimal control of systems governed by stochastic partial differential equations in which drift and diffusion terms are second- and first-order differential operators, respectively. Necessary conditions for an optimal control are derived for both nondegenerate and degenerate systems, and all the coefficients appearing in the equations are allowed to depend on the control variables. Furthermore, the results obtained in the paper can also be used to derive necessary conditions of optimality for partially observed diffusions with correlation between the signals and the observation noises.

Key words. stochastic partial differential equations (SPDEs), optimal control, adjoint equations, necessary conditions of optimality, partially observed diffusions.

AMS subject classifications. 60H15, 93E20

1. Introduction.

Let us consider a stochastic control problem, in which the state equation is a linear stochastic partial differential equation (SPDE):

$$
\begin{aligned}
dq(t,x) &= [\partial_i(a^{ij}(t,x,U(t))\partial_j q(t,x)) + b^i(t,x,U(t))\partial_i q(t,x) + c(t,x,U(t))q(t,x) \\
&\quad + f(t,x,U(t))]dt + [\sigma^{ik}(t,x,U(t))\partial_i q(t,x) + h^k(t,x,U(t))q(t,x) \\
&\quad + g^k(t,x,U(t))]dW_k(t), \quad x \in R^d, \quad t \in [0,1], \\
q(0,x) &= q_0(x), \quad x \in R^d,
\end{aligned} \tag{1.1}
$$

where $W := (W_1, W_2, \ldots, W_{d'})$ is a d'-dimensional standard Brownian motion with $W(0) = 0, \{U(t) : 0 \le t \le 1\}$ is an admissible control (the precise definition will be given later) and $\partial_i := \partial/\partial x_i, i = 1, 2, \ldots, d$. Throughout the paper, the conventional repeated indices for summation are used.

The optimal control problem is to minimize a given cost functional over the set of admissible controls.

The purpose of this paper is to study necessary conditions of an optimal control for the controlled system (1.1). It is well known that the so-called adjoint equations play a key role in dealing with the problem. Adjoint equations are, in general, *backward* equations with given terminal states. In stochastic problems, adjoint equations cannot be obtained simply by inversing the time, because the adaptness must be considered. For stochastic differential equations (SDEs), Bismut [5] introduced an adjoint equation with an additional martingale term. His method is based on the invertibility of certain fundamental matrices in finite dimensions and cannot be carried over to SPDEs whose state spaces are *infinite* dimensions. Using a finite-dimensional approximation method, Bensoussan [3] derived an adjoint equation of a nondegenerate SPDE (in a form more abstract than (1.1)) with its diffusion being a bounded operator. In the present paper, we solve the adjoint equation of (1.1) with $\sigma^{ik} \neq 0$ (i.e., the diffusion operator is unbounded), which is important in both theory and application. To handle the problem, the basic aproach we employ is still the finite-dimensional approximation. It should be noted that, while this is a very natural approach to infinite-dimensional problems, the difficulty is how to obtain certain *compactness* of the approximate solutions, which may vary from case to case. For explicit equations like

* Received by the editors May 28, 1991; accepted for publication (in revised form) May 1, 1992. This research was supported partially by the Monbusho Scholarship of Japanese Government and the National Natural Science Fundation of China.

† Department of Mathematics, Fudan University, Shanghai, China. This research was completed during a visit by the author to Department of Mathematics, Faculty of Science, Kobe University, Kobe 657, Japan.

(1.1), some delicate estimates of differential operators, originally due to Krylov and Rozovskii [8]–[10], will be used in this paper to show the compactness. Therefore, an adjoint equation of (1.1) as well as existence and uniqueness of its solutions will be derived.

Having obtained the adjoint equation, we derive necessary conditions of optimality for system (1.1) in which *all* the coefficients are allowed to depend on the control variable. Furthermore, the results obtained can be applied directly to a general model of partially observed diffusions with *correlation* between the signals and the observation noises. Therefore the existing results of Bensoussan [3], [4], Haussmann [7], and Baras, Elliott, and Kohlmann [1] are improved and extended considerably.

It should be mentioned that, in addition to the finite-dimensional approximation approach, there is a "time change" technique, which Bensoussan [4] used to solve the adjoint equation and study the necessity of optimality for the SPDE (1.1) with $\sigma^{ik} = 0$. Its main idea is to turn (1.1) into a P-almost surely *deterministic* PDE, basing on a transformation $e^{M^k W_k(t)}$, where M^k represents the diffusion operator (see (2.2), below). When $\sigma^{ik} \neq 0$, transformations of the same kind are also available, provided that the Brownian motion involved is one-dimensional; see [6], [15]. However, this method fails to work in general when $\sigma^{ik} \neq 0$ and the Brownian motion is multidimensional; refer to [15, §5.1] for a detailed discussion on this point.

The paper is organized as follows. In §2 we formulate our problem and introduce some basic notation and assumptions. In §3 we state a few fundamental results of SPDEs in the form convenient for us to use in the paper. Sections 4 and 5 are the main part of the paper. In §4 we derive an adjoint equation and prove existence and uniqueness of its solutions, and in §5 we investigate necessary conditions of optimality for system (1.1). The results in these two sections are valid, assuming that the system is nondegenerate (i.e., $S := (a^{ij} - \frac{1}{2}\sum_{k=1}^{d'} \sigma^{ik}\sigma^{jk})$ is uniformly positive definite). In §6 we discuss the degenerate case (i.e., S is nonnegative definite) and apply the main results to partially observed diffusions. Finally, §7 concludes the paper.

2. Problem formulation.

Let Γ be a Borel set in some Euclidean space R^M. We define a family of second-order differential operators $\{A(t,u) : t \in [0,1], u \in \Gamma\}$ and a family of first-order differential operators $\{M^k(t,u) : t \in [0,1], u \in \Gamma, k = 1,2,\dots,d'\}$ by

$$A(t,u)\phi(x) := \partial_i(a^{ij}(t,x,u)\partial_j\phi(x)) + b^i(t,x,u)\partial_i\phi(x) + c(t,x,u)\phi(x) \tag{2.1}$$

and

$$M^k(t,u)\phi(x) := \sigma^{ik}(t,x,u)\partial_i\phi(x) + h^k(t,x,u)\phi(x) \quad \text{for } x \in R^d, \phi \in C_0^\infty(R^d), \tag{2.2}$$

where $a^{ij}, b^i, c, \sigma^{ik}$, and h^k are given real-valued functions, $i,j = 1,2,\dots,d$ and $k = 1,2,\dots,d'$.

We also consider the formal adjoints of the operators $A(t,u)$ and $M^k(t,u)$

$$\begin{aligned} A^*(t,u)\phi(x) &:= \partial_i(a^{ij}(t,x,u)\partial_j\phi(x)) - b^i(t,x,u)\partial_i\phi(x) \\ &\quad + [c(t,x,u) - \partial_i b^i(t,x,u)]\phi(x), \end{aligned} \tag{2.3}$$

$$M^{k*}(t,u)\phi(x) := -\sigma^{ik}(t,x,u)\partial_i\phi(x) + [h^k(t,x,u) - \partial_i\sigma^{ik}(t,x,u)]\phi(x). \tag{2.4}$$

Let us now recall the definition of the Sobolev spaces. For $m = 0,1,2,\dots$, define $H^m := \{\phi : D^\alpha\phi \in L^2(R^d)$ for any $\alpha := (\alpha_1,\dots,\alpha_d)$ with $|\alpha| := |\alpha_1| + \cdots + |\alpha_d| \leq m\}$, with the norm

$$\|\phi\|_m := \left\{ \sum_{|\alpha|\le m} \int_{R^d} |D^\alpha \phi(x)|^2 dx \right\}^{1/2}.$$

For $m = -1, -2, \ldots$, define $H^m := (H^{-m})^*$. The Hilbert space H^m for any integer m is called a Sobolev space. For any integer m, let us consider the Gelfand triple $H^{m+1} \hookrightarrow H^m \hookrightarrow H^{m-1}$. We denote by $\langle \cdot, \cdot \rangle_m$ the duality pairing between H^{m-1} and H^{m+1}, and by $(\cdot, \cdot)_m$ the inner product in H^m.

For any second-order differential operator L that has the same form as (2.1), if we write $\langle L\phi, \psi \rangle_m$, then L is understood to be an operator from H^{m+1} to H^{m-1} by formally using Green's formula. For example, for the operator $A(t,u)$ defined by (2.1), we have

$$\begin{aligned} \langle A(t,u)\phi, \psi \rangle_m := &-(a^{ij}(t,\cdot,u)\partial_j\phi, \partial_i\psi)_m + (b^i(t,\cdot,u)\partial_i\phi, \psi)_m \\ &+ (c(t,\cdot,u)\phi,\psi)_m \qquad \text{for } \phi,\psi \in H^{m+1}. \end{aligned} \tag{2.5}$$

Remark 2.1. It is clear that $\langle A(t,u)\phi, \psi \rangle_0 = \langle \phi, A^*(t,u)\psi \rangle_0$ and $(M(t,u)\phi,\psi)_0 = (\phi, M^*(t,u)\psi)_0$ hold for $\phi, \psi \in H^1$. However, neither $\langle A(t,u)\phi,\psi \rangle_m = \langle \phi, A^*(t,u)\psi \rangle_m$ nor $(M(t,u)\phi,\psi)_m = (\phi, M^*(t,u)\psi)_m$ holds when $m \ge 1$.

For $\alpha, \beta \in (-\infty, +\infty)$ with $\alpha < \beta$, we are given a filtered probablity space $(\Omega, \mathcal{F}, P, \mathcal{F}_t : \alpha \le t \le \beta)$ and a Hilbert space X. For $p \in [1, +\infty]$, define $L^p_{\mathcal{F}}(\alpha, \beta, X) := \{\phi : \phi$ is an X-valued, $\mathcal{F}_t$-adapted process on $[\alpha, \beta]$, and $\phi \in L^p([\alpha,\beta] \times \Omega; X)\}$. We identify ϕ and ϕ' in $L^p_{\mathcal{F}}(\alpha,\beta;X)$ whenever $E\int_\alpha^\beta \|\phi(t) - \phi'(t)\|_x^p dt = 0$.

Now we recall the definition of admissible controls. By the set U_{ad} of admissible controls, we mean the collection of (i) standard probability spaces $(\Omega, \mathcal{F}, P)$ and d'-dimensional Brownian motions $\{W(t) : 0 \le t \le 1\}$ with $W(0) = 0$; (ii) Γ-valued, $\mathcal{F}_t$-adapted measurable processes $\{U(t) : 0 \le t \le 1\}$, where $\mathcal{F}_t := \sigma\{W(s) : 0 \le s \le t\}$. We denote $(\Omega, \mathcal{F}, P, W, U) \in U_{ad}$, but on occasion we will only write $U \in U_{ad}$ if no ambiguity arises.

Given $(\Omega, \mathcal{F}, P, W, U) \in U_{ad}$, we rewrite (1.1) in the following form, omitting the variable x:

$$\begin{aligned} dq(t) &= [A(t,U(t))q(t) + f(t,U(t))]dt \\ &\quad + [M^k(t,U(t))q(t) + g^k(t,U(t))]dW_k(t), \quad t \in [0,1], \\ q(0) &= q_0. \end{aligned} \tag{2.6}$$

A process $q = q^U \in L^2_{\mathcal{F}}(0,1;H^1)$ is called a solution of (2.6) or a response for the control U if, for each $\eta \in C_0^\infty(R^d)$ and almost all $(t,\omega) \in [0,1] \times \Omega$,

$$\begin{aligned} (q(t),\eta)_0 = (q_0,\eta)_0 &+ \int_0^t \langle A(s,U(s))q(s) + f(s,U(s)), \eta \rangle_0 \, ds \\ &+ \int_0^t (M^k(s,U(s))q(s) + g^k(s,U(s)), \eta)_0 dW_k(s). \end{aligned} \tag{2.7}$$

The optimal control problem is to choose $(\Omega, \mathcal{F}, P, W, U) \in U_{ad}$ to minimize the following cost functional:

$$J(U) := E\left[\int_0^1 \langle F(t,U(t)), q^U(t) \rangle_0 dt + (G, q^U(1))_0 \right], \tag{2.8}$$

where $F : [0,1] \times \Gamma \to H^{-1}$ and $G \in H^0$ are given.

Remark 2.2. The cost functional (2.8) includes the following one as a special case:

$$J(U) := E\left\{\int_0^1 [(F^0(t,U(t)),q^U(t))_0 + (F^i(t,U(t)),\partial_i q^U(t))_0]dt + (G,q^U(1))_0\right\}.$$

Let us fix an integer $m \geq 0$ and two positive constants K and δ. We introduce the following conditions:

$(A1)_m$ $a^{ij}, b^i, c, \sigma^{ik}, h^k : [0,1]\times R^d \times \Gamma \to R^1$ are measurable in (t,x,u) and continuous in u. Furthermore, the functions $a^{ij}, b^i, c, \sigma^{ik}, h^k, \partial_j\sigma^{ik}$, and $\partial_j h^k$ and their derivatives in x up to the order max$(2,m)$ do not exceed K in absolute value;

(A2) $a^{ij} = a^{ji}, i,j = 1,2,\ldots,d$, and the matrix $S := (a^{ij} - \frac{1}{2}\sum_{k=1}^{d'}\sigma^{ik}\sigma^{jk}) \geq 0$ for all (t,x,u);

(A2)$'$ $a^{ij} = a^{ji}, i,j = 1,2,\ldots,d$, and the matrix S is uniformly positive definite:

$$\xi^T S\xi \geq \delta|\xi|^2 \quad \text{for any } (t,x,u) \text{ and any } \xi \in R^d;$$

(A3) $F(t,u) \in H^{-1}, G \in H^0$ and $\|F(t,u)\|_{-1} + \|G\|_0 \leq K$;

(A4) $f, g^k : [0,1]\times R^d\times\Gamma \to R^1$ are measurable in (t,x,u) and continuous in $u, k = 1,2,\ldots,d'$. Furthermore, $f(t,\cdot,u) \in H^1, g^k(t,\cdot,u) \in H^2$, and

$$|f(t,x,u)| + |g^k(t,x,u)| + \|f(t,\cdot,u)\|_1 + \|g^k(t,\cdot,u)\|_2 \leq K, \quad k = 1,2,\ldots,d';$$

(A5) $q_0 \in H^1$.

3. Fundamental theory of SPDEs. In this section, we recall some fundamental facts about SPDEs, which are originally due to Krylov and Rozovskii [8]–[10]. We state the results and give some variants that are convenient to our later discussion.

PROPOSITION 3.1. (Krylov and Rozovskii [9]). *Let $\tilde{A}$ be any second-order differential operator having the same form as* (2.1), *and let $\tilde{M}^k$ be any first-order differential operator having the same form as* (2.2). *Assume that the coefficients of $\tilde{A}$ and $\tilde{M}^k$ satisfy* $(A1)_m$ *and* (A2). *Then there is a constant N_1 that depends only on K and m such that*

$$2[\langle\tilde{A}\phi,\phi\rangle_{\bar{m}} + (\tilde{f},\phi)_{\bar{m}}] + \sum_{k=1}^{d'}\|\tilde{M}^k\phi + \tilde{g}^k\|_{\bar{m}}^2 \leq N_1\left(\|\phi\|_{\bar{m}}^2 + \|\tilde{f}\|_{\bar{m}}^2 + \sum_{k=1}^{d'}\|\tilde{g}^k\|_{\bar{m}+1}^2\right)$$

(3.1) $\qquad$ *for any* $\phi \in H^{\bar{m}+1}, \tilde{f} \in H^{\bar{m}}$, *and* $\tilde{g}^k \in H^{\bar{m}+1}, \bar{m} = 0,\ldots,m.$

Remark 3.1. Estimate (3.1) is slightly different in form from the original result of Krylov and Rozovskii [9]. We may consult Nagase and Nisio [12, §3 and Appendix] for an explicit proof of (3.1).

COROLLARY 3.1. *In Proposition* 3.1., *let us assume, in addition, that the operator $\tilde{M}^k$ is of order zero. Then there is a constant N_2 that depends only on K and m such that*

$$2[\langle\tilde{A}\phi,\phi\rangle_{\bar{m}} + (\tilde{f},\phi)_{\bar{m}}] + \sum_{k=1}^{d'}\|\tilde{M}^k\phi + \tilde{g}^k\|_{\bar{m}}^2 \leq N_2\left(\|\phi\|_{\bar{m}}^2 + \|\tilde{f}\|_{\bar{m}}^2 + \sum_{k=1}^{d'}\|\tilde{g}^k\|_{\bar{m}}^2\right)$$

(3.2) $\qquad$ *for any* $\phi \in H^{\bar{m}+1}$, *and* $\tilde{f}, \tilde{g}^k \in H^{\bar{m}}, \bar{m} = 0,1,\ldots,m.$

Proof. From (3.1), it follows that

$$2[\langle\tilde{A}\phi,\phi\rangle_{\bar{m}} + (\tilde{f},\phi)_{\bar{m}}] \leq N_1(\|\phi\|_{\bar{m}}^2 + \|\tilde{f}\|_{\bar{m}}^2). \tag{3.3}$$

By the hypotheses of the corollary, $\bar{M}^k$ becomes a bounded operator on $H^{\bar{m}}$. Hence

$$\|\bar{M}^k\phi + \bar{g}^k\|_{\bar{m}}^2 \le 2(K^2\|\phi\|_{\bar{m}}^2 + \|\bar{g}^k\|_{\bar{m}}^2). \tag{3.4}$$

Therefore, estimate (3.2) follows from (3.3) and (3.4). □

COROLLARY 3.2. *In Proposition* 3.1, *let us assume, in addition, that* (A2)′ *is satisfied. Then there is a constant* N_3 *that depends only on* K, m, *and* δ *such that*

$$\begin{aligned} &2[\langle \bar{A}\phi, \phi\rangle_{\bar{m}} + \langle \bar{f}, \phi\rangle_{\bar{m}}] + \sum_{k=1}^{d'} \|\bar{M}^k\phi + \bar{g}^k\|_{\bar{m}}^2 \\ &\le -\delta\|\phi\|_{\bar{m}+1}^2 + N_3\left(\|\phi\|_{\bar{m}}^2 + \|\bar{f}\|_{\bar{m}-1}^2 + \sum_{k=1}^{d'}\|\bar{g}^k\|_{\bar{m}}^2\right) \end{aligned} \tag{3.5}$$

for $\phi \in H^{\bar{m}+1}, \bar{f} \in H^{\bar{m}-1}$ *and* $\bar{g}^k \in H^{\bar{m}}, \bar{m} = 0, 1, \ldots, m.$

Proof. The result can be easily derived by applying Proposition 3.1 to the operator $A - \delta\Delta$, where Δ is the Laplacian. □

Remark 3.2. When $m = 0$ (3.5) holds even if all the coefficients of $\bar{A}$ and $\bar{M}^k$ are only bounded measurable in x-variable; see Pardoux [13].

Let $\tilde{F}, \tilde{G}^k : [0, 1] \times R^d \times \Omega \to R^1, \xi_0 : R^d \times \Omega \to R^1$ be given. We now consider the following SPDE on an interval $[\alpha, \beta] \subset [0, 1]$:

$$\begin{aligned} d\xi(t) &= [A(t, U(t))\xi(t) + \tilde{F}(t)]dt \\ &\quad + [M^k(t, U(t))\xi(t) + \tilde{G}^k(t)]dW_k(t), \quad t \in [\alpha, \beta], \\ \xi(\alpha) &= \xi_0, \end{aligned} \tag{3.6}$$

PROPOSITION 3.2 (Krylov and Rozovskii [9], [10]). *Assume that* $(A1)_m$ *and* (A2) $(m \ge 1)$ *are satisfied and that* $\tilde{F} \in L^2_{\mathcal{F}}(0, 1; H^m), \tilde{G}^k \in L^2_{\mathcal{F}}(0, 1; H^{m+1})$, *and* $\xi_0 \in L^2(\Omega, \mathcal{F}_\alpha, P; H^m)$. *Then* (3.6) *has a unique solution* $\xi \in L^2_{\mathcal{F}}(\alpha, \beta; H^m) \cap L^2(\Omega; C(\alpha, \beta; H^{m-1}))$, *and there is a constant* N_4 *that depends only on* K *and* m *such that*

$$E\left(\sup_{\alpha\le t\le\beta} \|\xi(t)\|_{\bar{m}}^2\right) \le N_4 E\left\{\|\xi_0\|_{\bar{m}}^2 + \int_\alpha^\beta \left[\|\tilde{F}(t)\|_{\bar{m}}^2 + \sum_{k=1}^{d'}\|\tilde{G}^k(t)\|_{\bar{m}+1}^2\right]dt\right\}, \quad \bar{m} = 0, 1, \ldots, m. \tag{3.7}$$

COROLLARY 3.3. *Assume that* $(A1)_m$ *and* (A2) $(m \ge 1)$ *are satisfied with* $\sigma^{ik} = 0$ *and that* $\tilde{F}, \tilde{G}^k \in L^2_{\mathcal{F}}(0, 1; H^m)$ *and* $\xi_0 \in L^2(\Omega, \mathcal{F}_\alpha, P; H^m)$. *Then* (3.6) *has a unique solution* $\xi \in L^2_{\mathcal{F}}(\alpha, \beta; H^m) \cap L^2(\Omega; C(\alpha, \beta; H^{m-1}))$, *and estimate* (3.7) *can be strengthened to*

$$E\left(\sup_{\alpha\le t\le\beta} \|\xi(t)\|_{\bar{m}}^2\right) \le N_5 E\left\{\|\xi_0\|_{\bar{m}}^2 + \int_\alpha^\beta \left[\|\tilde{F}(t)\|_{\bar{m}}^2 + \sum_{k=1}^{d'}\|\tilde{G}^k(t)\|_{\bar{m}}^2\right]dt\right\}, \quad \bar{m} = 0, 1, \ldots, m. \tag{3.8}$$

Proof. Note that estimate (3.2) is available as $\sigma^{ik} = 0$; hence the result can be proved in a way analogous to that in Krylov and Rozovskii [9]. □

PROPOSITION (3.3) (Krylov and Rozovskii [8]). *Assume that* $(A1)_m$ *and* $(A2)'(m \geq 0)$ *are satisfied and that* $\tilde{F} \in L^2_{\mathcal{F}}(0,1;H^{m-1})$, $\tilde{G}^k \in L^2_{\mathcal{F}}(0,1;H^m)$, *and* $\xi_0 \in L^2(\Omega, \mathcal{F}_\alpha, P; H^m)$. *Then* (3.6) *has a unique solution* $\xi \in L^2_{\mathcal{F}}(\alpha,\beta;H^{m+1}) \cap L^2(\Omega; C(\alpha,\beta;H^m))$, *and there is a constant* N_6 *that depends only on* K, m, *and* δ *such that*

$$E\left(\sup_{\alpha\leq t\leq\beta}\|\xi(t)\|^2_{\bar{m}}\right) \leq N_6 E\left\{\|\xi_0\|^2_{\bar{m}} + \int_\alpha^\beta \left[\|\tilde{F}(t)\|^2_{\bar{m}-1} + \sum_{k=1}^{d'}\|\tilde{G}^k(t)\|^2_{\bar{m}}\right]dt\right\},$$

$$\bar{m} = 0,1,\ldots,m. \tag{3.9}$$

4. Adjoint equation. In this section, we derive an adjoint equation and study existence and uniqueness of its solutions. We fix an admissible control $(\Omega, \mathcal{F}, P, W, U)$ throughout this section.

THEOREM 4.1. *Assume that* $(A1)_0$, $(A2)'$, *and* (A3) *are satisfied. Then there exists a unique solution pair* $(\lambda, r) \in L^2_{\mathcal{F}}(0,1;H^1) \times [L^2_{\mathcal{F}}(0,1;H^0)]^{d'}$, *where* $r := (r^1, r^2, \ldots, r^{d'})$, *of the following backward SPDE:*

$$\begin{aligned} d\lambda(t) &= -\left[A^*(t,U(t))\lambda(t) + \sum_{k=1}^{d'} M^{k*}(t,U(t))r^k(t) + F(t,U(t))\right]dt \\ &\quad + \sum_{k=1}^{d'} r^k(t)dW_k(t), \quad t\in[0,1], \\ \lambda(1) &= G. \end{aligned} \tag{4.1}$$

Moreover, there is a constant N_7 *that depend only on* K *and* δ *such that*

$$E\int_0^1\left[\|\lambda(t)\|_1^2 + \sum_{k=1}^{d'}\|r^k(t)\|_0^2\right]dt \leq N_7 E\left[\int_0^1 \|F(t,U(t))\|^2_{-1}dt + \|G\|_0^2\right]. \tag{4.2}$$

Remark 4.1. Solutions of (4.1) are defined in a way similar to those of the SPDE (2.6). More precisely, $(\lambda, r) \in L^2_{\mathcal{F}}(0,1;H^1) \times [L^2_{\mathcal{F}}(0,1;H^0)]^{d'}$ is called a solution pair of (4.1) if, for each $\eta \in C_0^\infty(R^d)$ and almost all $(t,\omega) \in [0,1]\times\Omega$,

$$\begin{aligned} (\lambda(t),\eta)_0 = (G,\eta)_0 + \int_t^1 &\left[\langle\lambda(s), A(s,U(s))\eta\rangle_0 + \sum_{k=1}^{d'}(r^k(s), M^k(s,U(s))\eta)_0\right. \\ &\left. + \langle F(s,U(s)),\eta\rangle_0\right]ds - \sum_{k=1}^{d'}\int_t^1 (r^k(s),\eta)_0 dW_k(s). \end{aligned}$$

So (4.1) can be regarded as in the H^{-1} space.

Proof of Theorem 4.1. To avoid notational complexity, we prove the theorem for $d' = 1$ (there is no essential difficulty when $d' > 1$). Thus the index "k" will be dropped. On the other hand, we also omit to write the control $U(t)$ since it is fixed.

Uniqueness. Suppose that $(\lambda, r) \in L^2_{\mathcal{F}}(0,1;H^1) \times L^2_{\mathcal{F}}(0,1;H^0)$ satisfies

$$\begin{aligned} d\lambda(t) &= -[A^*(t)\lambda(t) + M^*(t)r(t)]dt + r(t)dW(t), \quad t\in[0,1], \\ \lambda(1) &= 0. \end{aligned} \tag{4.3}$$

Consider the following auxiliary SPDE, which admits a unique solution $\rho \in L^2_{\mathcal{F}}(0,1;H^1) \cap L^2(\Omega; C(0,1;H^0))$ by virtue of Proposition 3.3:

$$\begin{aligned} d\rho(t) &= [A(t)\rho(t) + \lambda(t)]dt + [M(t)\rho(t) + r(t)]dW(t), \quad t \in [0,1], \\ \rho(0) &= 0. \end{aligned}$$

Applying Ito's formula, we obtain

$$\begin{aligned} d(\rho(t), \lambda(t))_0 = {} & [\|\lambda(t)\|_0^2 + \|r(t)\|_0^2]dt \\ & + [(M(t)\rho(t) + r(t), \lambda(t))_0 + (\rho(t), r(t))_0]dW(t). \end{aligned}$$

Hence, $E\int_0^1 [\|\lambda(t)\|_0^2 + \|r(t)\|_0^2]dt = 0$. This implies the uniqueness.

Before proceeding to the proof of the existence, we introduce an adjoint equation in finite dimension, which was originally obtained by Bismut [5]; see also [2], [3].

LEMMA 4.1. *Let n be a fixed positive integer. Suppose that we are given $A_n, M_n \in L^\infty_{\mathcal{F}}(0,1;R^{n\times n})$, $F_n \in L^2_{\mathcal{F}}(0,1;R^n)$, and $G_n \in R^n$. Then there exists uniquely a pair $(\lambda_n, r_n) \in L^2_{\mathcal{F}}(0,1;R^n) \times L^2_{\mathcal{F}}(0,1;R^n)$ satisfying the following backward* SDE:

$$(4.4) \quad \begin{aligned} d\lambda_n(t) &= -[A_n^T(t)\lambda_n(t) + M_n^T(t)r_n(t) + F_n(t)]dt + r_n(t)dW(t), \quad t \in [0,1], \\ \lambda_n(1) &= G_n. \end{aligned}$$

Let us now continue proving Theorem 4.1.

Existence. Consider the Gelfand triple $H^1 \hookrightarrow H^0 \hookrightarrow H^{-1}$. Let $e_1, e_2, \ldots, e_n, \ldots$, be a Hilbert basis of H^1, which is orthonormal as a basis of H^0.

Fix a positive integer n. By Lemma 4.1, there is a unique pair

$$\bar{\lambda}_n := (\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{nn})^T \in L^2_{\mathcal{F}}(0,1;R^n)$$

and

$$\bar{r}_n := (r_{n1}, r_{n2}, \ldots, r_{nn})^T \in L^2_{\mathcal{F}}(0,1;R^n)$$

satisfying

$$(4.5) \quad \begin{aligned} d\lambda_{ni}(t) = {} & -\left[\sum_{j=1}^n (e_j, A(t)e_i)_0 \lambda_{nj}(t) + \sum_{j=1}^n (e_j, M(t)e_i)_0 r_{nj}(t) \right. \\ & \left. + (F(t), e_i)_0 \right] dt + r_{ni}(t)dW(t), \quad t \in [0,1], \\ \lambda_{ni}(1) = {} & G_{ni}, \quad i = 1, 2, \ldots, n, \end{aligned}$$

where $\sum_{i=1}^n G_{ni}e_i := G_n \to G$ in H^0 as $n \to \infty$. Define

$$(4.6) \quad \lambda_n := \sum_{i=1}^n \lambda_{ni}e_i \in L^2_{\mathcal{F}}(0,1;H^1)$$

and

$$(4.7) \quad r_n := \sum_{i=1}^n r_{ni}e_i \in L^2_{\mathcal{F}}(0,1;H^1).$$

Applying Ito's formula to (4.5) and adding up in i from 1 to n, we have

$$\begin{aligned} d\|\lambda_n(t)\|_0^2 = &-2[\langle\lambda_n(t), A(t)\lambda_n(t)\rangle_0 + (r_n(t), M(t)\lambda_n(t))_0 + \langle F(t), \lambda_n(t)\rangle_0]dt \\ &+ 2(r_n(t), \lambda_n(t))_0 dW(t) + \|r_n(t)\|_0^2 dt. \end{aligned} \tag{4.8}$$

Hence,

$$\begin{aligned} &E\|\lambda_n(t)\|_0^2 \\ &= E\|G_n\|_0^2 + 2E\int_t^1 [\langle A(s)\lambda_n(s), \lambda_n(s)\rangle_0 + (r_n(s), M(s)\lambda_n(s))_0 \\ &\quad +\langle F(s), \lambda_n(s)\rangle_0 - \frac{1}{2}\cdot\|r_n(s)\|_0^2]ds \\ &\le E\|G_n\|_0^2 + E\int_t^1 [2\langle A(s)\lambda_n(s), \lambda_n(s)\rangle_0 + \|M(s)\lambda_n(s)\|_0^2 + 2\langle F(s), \lambda_n(s)\rangle_0]ds \\ &\le E\|G_n\|_0^2 + E\int_t^1 [-\delta\|\lambda_n(s)\|_1^2 + N_3\|\lambda_n(s)\|_0^2 + 2/\delta\cdot\|F(s)\|_{-1}^2 + \delta/2\cdot\|\lambda_n(s)\|_1^2]ds \\ &\le -\delta/2\cdot E\int_t^1 \|\lambda_n(s)\|_1^2 ds + E\|G_n\|_0^2 + N_8 E\int_t^1 [\|\lambda_n(s)\|_0^2 + \|F(s)\|_{-1}^2]ds. \end{aligned} \tag{4.9}$$

So Gronwall's inequality yields

$$\sup_{0\le t\le 1} E\|\lambda_n(t)\|_0^2 + \delta/2\cdot E\int_0^1 \|\lambda_n(t)\|_1^2 dt \le N_9 E\left(\int_0^1 \|F(t)\|_{-1}^2 dt + \|G_n\|_0^2\right), \tag{4.10}$$

where N_9 depends only on K and δ.

Now let $\bar{\rho}_n := (\rho_{n1}, \rho_{n2}, \dots, \rho_{nn})^T \in L^2_{\mathcal{F}}(0,1;R^n)$ be the solution of the following SDE in R^n:

$$\begin{aligned} d\rho_{ni}(t) &= \sum_{j=1}^n \langle A(t)e_j, e_i\rangle_0 \rho_{nj}(t)dt + \left[\sum_{j=1}^n (M(t)e_j, e_i)_0\rho_{nj}(t) + r_{ni}(t)\right] dW(t), \\ \rho_{ni}(0) &= 0, \quad i = 1, 2, \dots, n. \end{aligned} \tag{4.11}$$

Define $\rho_n := \sum_{i=1}^n \rho_{ni}e_i \in L^2_{\mathcal{F}}(0,1;H^1)$. By a calculation similar to the above, we have

$$\begin{aligned} E\|\rho_n(t)\|_0^2 &= E\int_0^t [2\langle A(s)\rho_n(s), \rho_n(s)\rangle_0 + \|M(s)\rho_n(s) + r_n(s)\|_0^2]ds \\ &\le -\delta E\int_0^t \|\rho_n(s)\|_1^2 ds + N_3 E\int_0^t [\|\rho_n(s)\|_0^2 + \|r_n(s)\|_0^2]ds. \end{aligned} \tag{4.12}$$

Applying again Gronwall's inequality, we obtain

$$\sup_{0\le t\le 1} E\|\rho_n(t)\|_0^2 + \delta E\int_0^1 \|\rho_n(t)\|_1^2 dt \le N_3\exp(N_3)E\int_0^1 \|r_n(t)\|_0^2 dt. \tag{4.13}$$

On the other hand, Ito's formula gives

$$\begin{aligned} d\sum_{i=1}^n \lambda_{ni}(t)\rho_{ni}(t) &= \sum_{i=1}^n \{-[\langle A^*(t)\lambda_n(t), e_i\rangle_0 + (M^*(t)r_n(t), e_i)_0 \\ &\quad +\langle F(t), e_i\rangle_0]\rho_{ni}(t) + \lambda_{ni}(t)\langle A(t)\rho_n(t), e_i\rangle_0 \\ &\quad +r_{ni}(t)[(M(t)\rho_n(t), e_i)_0 + r_{ni}(t)]\}dt + \{\cdots\}dw(t) \\ &= [\|r_n(t)\|_0^2 - \langle F(t), \rho_n(t)\rangle_0]dt + \{\cdots\}dW(t). \end{aligned} \tag{4.14}$$

Integrating from 0 to 1 and taking expectation, we have

$$
\begin{aligned}
E\int_0^1 \|r_n(t)\|_0^2 dt &= E\left[\int_0^1 \langle F(t), \rho_n(t)\rangle_0 dt + (G_n, \rho_n(1))_0\right] \\
&\le \left(E\int_0^1 \|F(t)\|_{-1}^2 dt\right)^{1/2}\left(E\int_0^1 \|\rho_n(t)\|_1^2 dt\right)^{1/2} \\
&\quad + (E\|G_n\|_0^2)^{1/2}(E\|\rho_n(1)\|_0^2)^{1/2} \\
&\le \left(N_{10} E\int_0^1 \|r_n(t)\|_0^2 dt\right)^{1/2} \\
&\quad \cdot\left\{\left(E\int_0^1 \|F(t)\|_{-1}^2 dt\right)^{1/2} + (E\|G_n\|_0^2)^{1/2}\right\},
\end{aligned}
\tag{4.15}
$$

where $N_{10} = \max\{N_3 \exp(N_3), 1/\delta \cdot N_3 \exp(N_3)\}$. Consequently,

$$
E\int_0^1 \|r_n(t)\|_0^2 dt \le 2N_{10} E\left[\int_0^1 \|F(t)\|_{-1}^2 dt + \|G_n\|_0^2\right]. \tag{4.16}
$$

By (4.10) and (4.16), there exist a subsequence $\{n'\}$ of $\{n\}$ and a pair $(\lambda, r) \in L^2_{\mathcal{F}}(0,1; H^1) \times L^2_{\mathcal{F}}(0,1; H^0)$ such that

$$
\lambda_{n'} \to \lambda \text{ weakly in } L^2([0,1]\times\Omega; H^1) \tag{4.17}
$$

and

$$
r_{n'} \to r \text{ weakly in } L^2([0,1]\times\Omega; H^0) \text{ as } n' \to \infty. \tag{4.18}
$$

We show that (λ, r) satisfies (4.1). To this end, let γ be an absolutely continuous function from [0,1] to R^1 with $\dot\gamma := d\gamma/dt \in L^2[0,1]$ and $\gamma(0) = 0$. Set $\gamma_i(t) := \gamma(t)e_i$. Multiplying (4.5) by $\gamma_i(t)$ and using Ito's formula, we have

$$
\begin{aligned}
&\int_0^t (\lambda_n(t), \dot\gamma_i(t))_0 dt + \int_0^1 (r_n(t), \gamma_i(t))_0 dW(t) \\
&\quad = (G_n, \gamma_i(1))_0 + \int_0^1 [\langle \lambda_n(t), A(t)\gamma_i(t)\rangle_0 + (r_n(t), M(t)\gamma_i(t))_0 + \langle F(t), \gamma_i(t)\rangle_0] dt.
\end{aligned}
\tag{4.19}
$$

Letting n' go to infinity and observing (4.17) and (4.18), we conclude that

$$
\begin{aligned}
&\int_0^t (\lambda(t), \phi)_0 \dot\gamma(t) dt + \int_0^1 (r(t), \phi)_0 \gamma(t) dW(t) \\
&\quad = (G, \phi)_0 \gamma(1) + \int_0^1 [\langle \lambda(t), A(t)\phi\rangle_0 + (r(t), M(t)\phi)_0 + \langle F(t), \phi\rangle_0]\gamma(t) dt
\end{aligned}
\tag{4.20}
$$

for any $\phi \in H^1$.

For any $t \in (0,1)$, let

$$
\gamma_\varepsilon(s) := \begin{cases} 0 & \text{if } s \le t - \varepsilon/2, \\ 1/\varepsilon \cdot (s - t + \varepsilon/2) & \text{if } t - \varepsilon/2 < s < t + \varepsilon/2, \\ 1 & \text{if } s \ge t + \varepsilon/2. \end{cases}
$$

Substituting (4.20) with γ_ε and letting $\varepsilon \to 0$, we arrive at

$$(\lambda(t),\phi)_0 + \int_t^1 (r(s),\phi)_0 dW(s) = (G,\phi)_0 + \int_t^1 [\langle\lambda(s),A(s)\phi\rangle_0 + (r(s),M(s)\phi)_0 + \langle F(s),\phi\rangle_0]ds$$

$$\text{for any } \phi \in H^1, \text{ a.e. } t\in[0,1].$$

This implies that (λ, r) satisfies (4.1). Finally, (4.2) is obtained by letting $n' \to \infty$ in (4.10) and (4.16). □

COROLLARY 4.1. *Let the same assumptions as in Theorem* 4.1 *be satisfied. Given* $\tilde{f} \in L^2_{\mathcal{F}}(0,1;H^{-1})$ *and* $\tilde{g}^k \in L^2_{\mathcal{F}}(0,1;H^0)$, $k=1,2,\ldots,d'$, *suppose that*

$$\xi \in L^2_{\mathcal{F}}(0,1;H^1)\cap L^2(\Omega;C(0,1;H^0))$$

satisfies

$$d\xi(t) = [A(t,U(t))\xi(t) + \tilde{f}(t)]dt + [M^k(t,U(t))\xi(t) + \tilde{g}^k(t)]dW_k(t), \quad t\in[0,1], \tag{4.21}$$

and that (λ, r) *satisfies* (4.1). *Then, for any* $[\alpha,\beta]\subset[0,1]$,

$$E\left[\int_\alpha^\beta \langle F(t,U(t)),\xi(t)\rangle_0 dt + (\lambda(\beta),\xi(\beta))_0\right] = E\left\{\int_\alpha^\beta \left[\langle\lambda(t),f(t)\rangle_0 + \sum_{k=1}^{d'}(r^k(t),g^k(t))_0\right]dt + (\lambda(\alpha),\xi(\alpha))_0\right\}. \tag{4.22}$$

Proof. The result is easily derived by applying Ito's formula to $(\lambda(t),\xi(t))_0$. □

Remark 4.2. In Theorem 4.1 and Corollary 4.1, condition $(\mathrm{A1})_0$ can be weakened by assuming that all the coefficients a^{ij}, and so forth, are only bounded measurable in x. However, under such a weaker condition, the adjoint operators A^* and M^{k*} can be no longer written explicitly as (2.3) and (2.4), respectively. Instead, as operators mapping H^1 into H^{-1}, they can be determined by the following formulae:

$$\langle A^*(t,u)\phi,\psi\rangle_0 := \langle\phi,A(t,u)\psi\rangle_0 \text{ and}$$

$$\langle M^{k*}(t,u)\phi,\psi\rangle_0 := \langle\phi,M^k(t,u)\psi\rangle_0 \text{ for } \phi,\psi\in H^1.$$

It should be also noted that the definition of solutions of (4.1) does not require explicit expressions of A^* and M^{k*}; see Remark 4.1.

5. Necessary conditions of optimality. We study in this section necessary conditions of an optimal control for the general system (2.6) with the cost functional (2.8).

THEOREM 5.1. *Assume that* $(\mathrm{A1})_2$, $(\mathrm{A2})'$, (A3), (A4), *and* (A5) *are satisfied and that* $(\Omega,\mathcal{F},P,W,\hat{U})$ *is an optimal control along with the corresponding optimal state* $\hat{q}$. *Then, for almost every* $t\in[0,1]$, *we have the maximum condition*

$$H(t,\hat{q}(t),\hat{U}(t),\lambda(t),r(t)) = \max_{u\in\Gamma'} H(t,\hat{q}(t),u,\lambda(t),r(t)), \quad P-\text{a.s.}, \tag{5.1}$$

where (λ, r) *is the solution pair of* (4.1) *with* $U(t) \equiv \hat{U}(t)$, *and the Hamiltonian* H *is defined by*

$$H(t,\phi,u,\zeta,\eta) := -\langle A(t,u)\phi,\zeta\rangle_0 - (f(t,u),\zeta)_0 - \sum_{k=1}^{d'}[(M^k(t,u)\phi,\eta^k)_0 + (g^k(t,u),\eta^k)_0] - \langle F(t,u),\phi\rangle_0,$$
$$\text{for } (t,\phi,u,\zeta,\eta) \in [0,1]\times H^1\times\Gamma\times H^1\times(H^0)^{d'}. \tag{5.2}$$

Proof. We assume that $d' = 1$ and set $\hat{A}(t) := A(t,\hat{U}(t))$, and so on, for simplicity. From Proposition 3.3, it follows that $\hat{q} \in L^2_{\mathcal{F}}(0,1;H^2)\cap L^2(\Omega;C(0,1;H^1))$. Hence $\hat{q}(t) \in L^2(\Omega;H^2)$ for almost every $t \in [0,1]$. Fix a time $\bar{t} \in [0,1)$ such that $\hat{q}(\bar{t}) \in L^2(\Omega;H^2)$, along with a Γ-valued, $\mathcal{F}_{\bar{t}}$-measurable random variable u. For any $\varepsilon \in (0, 1-\bar{t})$, define $U_\varepsilon \in U_{ad}$ by

$$U_\varepsilon(t) := \begin{cases} u, & t \in [\bar{t},\bar{t}+\varepsilon], \\ \hat{U}(t), & t \in [0,1]\backslash[\bar{t},\bar{t}+\varepsilon]. \end{cases}$$

Let q_ε be the response for U_ε, namely,

$$q_\varepsilon(t) = \hat{q}(t), \quad t \in [0,\bar{t}], \tag{5.3}$$

$$q_\varepsilon(t) = \hat{q}(\bar{t}) + \int_{\bar{t}}^t [A(s,u)q_\varepsilon(s) + f(s,u)]ds + \int_{\bar{t}}^t [M(s,u)q_\varepsilon(s) + g(s,u)]dW(s),$$
$$t \in [\bar{t},\bar{t}+\varepsilon], \tag{5.4}$$

and

$$q_\varepsilon(t) = q_\varepsilon(\bar{t}+\varepsilon) + \int_{\bar{t}+\varepsilon}^t [\hat{A}(s)q_\varepsilon(s) + \hat{f}(s)]ds + \int_{\bar{t}+\varepsilon}^t [\hat{M}(s)q_\varepsilon(s) + \hat{g}(s)]dW(s),$$
$$t \in [\bar{t}+\varepsilon],1]. \tag{5.5}$$

It follows from Proposition 3.3 that $q_\varepsilon \in L^2(\Omega;C(\bar{t},1;H^2))$ and

$$E\left(\sup_{\bar{t}\le t\le\bar{t}+\varepsilon}\|q_\varepsilon(t)\|_2^2\right) \le N_6E\left\{\|\hat{q}(\bar{t})\|_2^2 + \int_{\bar{t}}^{\bar{t}+\varepsilon}[\|f(t,u)\|_1^2 + \|g(t,u)\|_2^2]dt\right\} \le N_6E(\|\hat{q}(\bar{t})\|_2^2 + 2K^2\varepsilon). \tag{5.6}$$

Define $\xi_\varepsilon(t) := q_\varepsilon(t) - \hat{q}(t)$ for $t \in [0.1]$. Then ξ_ε satisfies

$$d\xi_\varepsilon(t) = [\hat{A}(t)\xi_\varepsilon(t) + (A(t,u) - \hat{A}(t))q_\varepsilon(t) + f(t,u) - \hat{f}(t)]dt + [\hat{M}(t)\xi_\varepsilon(t) + (M(t,u) - \hat{M}(t))q_\varepsilon(t) + g(t,u) - \hat{g}(t)]dW(t),$$
$$t \in [\bar{t},\bar{t}+\varepsilon] \tag{5.7}$$

and

$$d\xi_\varepsilon(t) = \hat{A}(t)\xi_\varepsilon(t)dt + \hat{M}(t)\xi_\varepsilon(t)dW(t), \qquad t \in [\bar{t}+\varepsilon, 1]. \tag{5.8}$$

Since $\hat{U}$ is optimal, we have

$$\begin{aligned} 0 \le J(U_\varepsilon) - J(\hat{U}) &= E\left\{ \int_{\bar{t}}^1 [\langle F(t, U_\varepsilon(t)), q_\varepsilon(t)\rangle_0 - \langle \hat{F}(t), \hat{q}(t)\rangle_0] dt \right. \\ &\quad \left. + (G, q_\varepsilon(1) - \hat{q}(1))_0 \right\} \\ &= E\int_{\bar{t}}^{\bar{t}+\varepsilon} [\langle F(t,u) - \hat{F}(t), q_\varepsilon(t)\rangle_0 + \langle \hat{F}(t), \xi_\varepsilon(t)\rangle_0] dt \\ &\quad + E\left[\int_{\bar{t}+\varepsilon}^1 \langle \hat{F}(t), \xi_\varepsilon(t)\rangle_0 dt + (G, \xi_\varepsilon(1))_0\right] = I_1 + I_2. \end{aligned} \tag{5.9}$$

Applying Corollary 4.1 to (5.8) and (4.1), we have $I_2 = E(\lambda(\bar{t}+\varepsilon), \xi_\varepsilon(\bar{t}+\varepsilon))_0$. Thus we can rewrite (5.9) as

$$\begin{aligned} 0 \le E &\int_{\bar{t}}^{\bar{t}+\varepsilon} \langle F(t,u) - \hat{F}(t), q_\varepsilon(t)\rangle_0 dt \\ &+ E\left[\int_{\bar{t}}^{\bar{t}+\varepsilon} \langle \hat{F}(t), \xi_\varepsilon(t)\rangle_0 dt + (\lambda(\bar{t}+\varepsilon), \xi_\varepsilon(\bar{t}+\varepsilon))_0\right]. \end{aligned} \tag{5.10}$$

Applying Corollary 4.1 again to (5.7) and (4.1), we obtain

$$\begin{aligned} 0 \le E \int_{\bar{t}}^{\bar{t}+\varepsilon} &[\langle F(t,u) - \hat{F}(t), q_\varepsilon(t)\rangle_0 + \langle \lambda(t), (A(t,u) - \hat{A}(t)) q_\varepsilon(t)\rangle_0 \\ &+ \langle \lambda(t), f(t,u) - \hat{f}(t)\rangle_0 \\ &+ (r(t), (M(t,u) - \hat{M}(t)) q_\varepsilon(t) + g(t,u) - \hat{g}(t))_0] dt. \end{aligned} \tag{5.11}$$

On the other hand, we have

$$\begin{aligned} (1/\varepsilon) &\int_{\bar{t}}^{\bar{t}+\varepsilon} E\langle \lambda(t), (A(t,u) - \hat{A}(t))(q_\varepsilon(t) - \hat{q}(t))\rangle_0 dt \\ &\le \text{const } (1/\varepsilon) \int_{\bar{t}}^{\bar{t}+\varepsilon} E(\|\lambda(t)\|_1 \|\xi_\varepsilon(t)\|_1) dt. \end{aligned} \tag{5.12}$$

Recall that ξ_ε satisfies (5.7) on $[\bar{t}, \bar{t}+\varepsilon]$ with $\xi_\varepsilon(\bar{t}) = 0$. Hence Proposition 3.3 gives

$$\begin{aligned} E &\left(\sup_{\bar{t} \le t \le \bar{t}+\varepsilon} \|\xi_\varepsilon(t)\|_1^2 \right) \\ &\le N_6 E \int_{\bar{t}}^{\bar{t}+\varepsilon} [\|(A(t,u) - \hat{A}(t)) q_\varepsilon(t) + f(t,u) - \hat{f}(t)\|_0^2 \\ &\quad + \|(M(t,u) - \hat{M}(t)) q_\varepsilon(t) + g(t,u) - \hat{g}(t)\|_1^2] dt \\ &\le \text{const } E \int_{\bar{t}}^{\bar{t}+\varepsilon} (\|q_\varepsilon(t)\|_2^2 + 1) dt \\ &\le \text{const } (E\|\hat{q}(\bar{t})\|_2^2 + 1)\varepsilon \quad \text{(by (5.6))} \\ &\le \text{const } \varepsilon \quad \text{(Note that we have fixed } \bar{t} \text{ such that } E\|\hat{q}(\bar{t}\|_2^2 < +\infty). \end{aligned} \tag{5.13}$$

Thus, (5.12) reduces to

$$\begin{aligned} (1/\varepsilon) &\int_{\bar{t}}^{\bar{t}+\varepsilon} E\langle \lambda(t), (A(t,u) - \hat{A}(t))(q_\varepsilon(t) - \hat{q}(t))\rangle_0 dt \\ &\le \text{const } (1/\varepsilon) \int_{\bar{t}}^{\bar{t}+\varepsilon} \{(\varepsilon^{1/3}/2) E\|\lambda(t)\|_1^2 + [1/(2\varepsilon^{1/3})] E\|\xi_\varepsilon(t)\|_1^2\} dt \\ &\le \text{const } \left[\varepsilon^{1/3} \cdot (1/\varepsilon) \int_{\bar{t}}^{\bar{t}+\varepsilon} E\|\lambda(t)\|_1^2 dt + (1/\varepsilon)\varepsilon(1/\varepsilon^{1/3})\varepsilon\right] \\ &\to 0 \text{ as } \varepsilon \to 0, \end{aligned} \tag{5.14}$$

provided that $\bar{t}$ is a Lebesgue point of the function $t \to E\|\lambda(t)\|_1^2$. Similarly, we have

$$(5.15)\quad \begin{aligned}&(1/\varepsilon)E\int_{\bar{t}}^{\bar{t}+\varepsilon}[\langle F(t,u)-\hat{F}(t),q_\varepsilon(t)-\hat{q}(t)\rangle_0 \\ &\qquad +\langle r(t),(M(t,u)-\hat{M}(t))(q_\varepsilon(t)-\hat{q}(t))\rangle_0]dt \\ &\to 0 \text{ as } \varepsilon\to 0,\end{aligned}$$

provided that $\bar{t}$ is a Lebesgue point of the function $t \to E\|r(t)\|_0^2$. Thus (5.11) becomes

$$(5.16)\quad \begin{aligned}0\le E\int_{\bar{t}}^{\bar{t}+\varepsilon}&[\langle F(t,u)-\hat{F}(t),\hat{q}(t)\rangle_0+\langle\lambda(t),(A(t,u)-\hat{A}(t))\hat{q}(t)\rangle_0 \\ &+\langle\lambda(t),f(t,u)-\hat{f}(t)\rangle_0+\langle r(t),(M(t,u)-\hat{M}(t))\hat{q}(t) \\ &+g(t,u)-\hat{g}(t)\rangle_0]dt+o(\varepsilon).\end{aligned}$$

Dividing (5.16) by ε and letting $\varepsilon \to 0$, we obtain

$$(5.17)\quad EH(\bar{t},\hat{q}(\bar{t}),\hat{U}(\bar{t}),\lambda(\bar{t}),r(\bar{t}))\ge EH(\bar{t},\hat{q}(\bar{t}),u,\lambda(\bar{t}),r(\bar{t})).$$

Therefore, the desired result (5.1) follows from a standard argument; see, for example, [11]. This concludes the theorem. □

Remark 5.1. In Theorem 5.1, all the coefficients appearing in the control problem are allowed to depend on the control variable.

Remark 5.2. By the above proof (especially the argument between (5.12) and (5.17)), it is easy to see that, if a^{ij} and σ^{ik} contain no control variable, then the assumptions of Theorem 5.1 can be considerably relaxed. More precisely, $(A1)_2$ can be replaced by the assumption that all the coefficients a^{ij}, and so forth, are bounded measurable in x, and (A4) and (A5) together can be replaced by the assumptions that $q_0 \in H^0, f(t,u) \in H^{-1}, g^k(t,u) \in H^0$, and that their respective norms are bounded. Therefore, all the regularity restrictions imposed in Bensoussan [4] can be removed. It should be also noted that M^k is allowed to be unbounded in our results, compared with [4].

6. Discussion and application.

6.1. Degenerate cases. The main results in §§4 and 5 are derived under assumption (A2)′; namely, system (2.6) is nondegenerate. There seems to be some essential difficulties in treating possibly degenerate systems, and the adjoint equation (4.1) may no longer admit a solution pair, no matter how high regularity assumptions to be imposed on the functions F and G. At least the method of proving Theorem 4.1 no longer applies. Indeed, the basic idea behind the proof of Theorem 4.1 is to approximate the infinite-dimensional equation by certain finite-dimensional equations and to use estimate (3.5) to prove that the approximate solutions are weakly compact. However, when the systems are possibly degenerate, we have only estimate (3.1) available, which could not ensure the above compactness. On the other hand, estimate (3.1) is the "best one" in that it cannot be further improved.

However, our approach still works in the degenerate case if M^k is of order zero, for which a stronger estimate (3.2) is available.

Let us now introduce the following assumptions:

(B1) $F(t,u)\in H^1, G\in H^1$, and $\|F(t,u)\|_1+\|G\|_1\le K$;

(B2) $f, g^k : [0,1]\times R^d\times\Gamma\to R^1$ are measurable in (t,x,u) and continuous in $u, k=1,2,\ldots,d'$. Furthermore, $f(t,\cdot,u), g^k(t,\cdot,u)\in H^3$; and

$$|f(t,x,u)| + |g^k(t,x,u)| + \|f(t,\cdot,u)\|_3 + \|g^k(t,\cdot,u)\|_3 \le K, k = 1,2,\ldots,d';$$

(B3) $q_0 \in H^3$.

THEOREM 6.1. *Assume that* (B1), $(A1)_1$, *and* (A2) *are satisfied with* $\sigma^{ik} = 0$. *Then, for any* $(\Omega, \mathcal{F}, P, W, U) \in U_{ad}$, *there exists a unique solution pair* $(\lambda, r) \in L^2_{\mathcal{F}}(0,1;H^1) \times [L^2_{\mathcal{F}}(0,1;H^1)]^{d'}$ *of the adjoint equation* (4.1). *Moreover, there is a constant* N_{11} *that depends only on* K *such that*

$$(6.1)\quad E\int_0^1 \left[\|\lambda(t)\|_1^2 + \sum_{k=1}^{d'} \|r^k(t)\|_1^2\right] dt \le N_{11} E\left[\int_0^1 \|F(t,U(t))\|_1^2 dt + \|G\|_1^2\right].$$

Proof. Once estimate (3.2) is observed, the theorem can be proved in a way similar to that of Theorem 4.1. The proof is left to the reader. □

THEOREM 6.2. *Assume that* (B1)–(B3), $(A1)_3$, *and* (A2) *are satisfied with* $\sigma^{ik} = 0$ *and that* $(\Omega, \mathcal{F}, P, W, \hat{U}) \in U_{ad}$ *is an optimal control along with the optimal state* $\hat{q}$. *Then for almost every* $t \in [0,1]$,

$$(6.2)\qquad H(t,\hat{q}(t),\hat{U}(t),\lambda(t),r(t)) = \max_{u\in\Gamma} H(t,\hat{q}(t),u,\lambda(t),r(t)), \quad P-\text{a.s.},$$

where (λ, r) *is the solution pair of* (4.1) *with the control* $U = \hat{U}$, *and the Hamiltonian* H *is defined by* (5.2).

Proof. Basing on Theorem 6.1, we only need to slightly modify the proof of Theorem 5.1. From Corollary 3.3, it follows that $\hat{q} \in L^2_{\mathcal{F}}(0,1;H^3) \cap L^2(\Omega; C(0,1;H^2))$. Hence $\hat{q}(t) \in L^2(\Omega;H^3)$ for almost every $t \in [0,1]$. Fix $\bar{t} \in [0,1)$ such that $\hat{q}(\bar{t}) \in L^2(\Omega;H^3)$. For $\varepsilon \in (0, 1-t)$, define q_ε by (5.3)–(5.5). By virtue of Corollary 3.3, we have

$$\begin{aligned} E\left(\sup_{\bar{t}\le t\le\bar{t}+\varepsilon} \|q_\varepsilon(t)\|_3^2\right) &\le N_4 E\left\{\|\hat{q}(\bar{t})\|_3^2 + \int_{\bar{t}}^{\bar{t}+\varepsilon} [\|f(s,u)\|_3^2 + \|g(s,u)\|_3^2] ds\right\} \\ &\le N_4 E(\|\hat{q}(\bar{t})\|_3^2 + 2K^2\varepsilon). \end{aligned}$$

Define $\xi_\varepsilon(t) := q_\varepsilon(t) - \hat{q}(t)$. Then ξ_ε satisfies (5.7) on $[\bar{t}, \bar{t}+\varepsilon]$. Hence,

$$\begin{aligned} E\left(\sup_{\bar{t}\le t\le\bar{t}+\varepsilon} \|\xi_\varepsilon(t)\|_1^2\right) &\le N_4 E\int_{\bar{t}}^{\bar{t}+\varepsilon} [\|(A(t,u) - \hat{A}(t))q_\varepsilon(t) + f(t,u) - \hat{f}(t)\|_1^2 \\ &\qquad + \|(M(t,u) - \hat{M}(t))q_\varepsilon(t) + g(t,u) - \hat{g}(t)\|_1^2) dt \\ &\le \text{const } E\int_{\bar{t}}^{\bar{t}+\varepsilon} (\|q_\varepsilon(t)\|_3^2 + 1) dt \\ &\le \text{const } (E\|\hat{q}(\bar{t})\|_3^2 + 1)\varepsilon \le \text{const } \varepsilon. \end{aligned}$$

In what follows, we must only repeat the arguments in the proof of Theorem 5.1 to obtain the desired result. □

Remark 6.1. Theorem 6.2 requires the higher regularity on the coefficients a^{ij}, f, and so forth, as well as on the initial state q_0. However, if the second-order coefficients of the operator $A(t,u)$ contain no control variable, then it is easy to verify by the proof of Theorem 6.2 that all the regularity conditions can be reduced by two orders. It should be also noted that Theorem 6.2 extends the results of Bensoussan [4] to possibly degenerate systems.

6.2. Application to partially observed diffusions. The optimal control problem of partially observed diffusions with general nonlinear cost functionals can be formulated as a control problem of linear SPDEs (Zakai's equations) with linear cost functionals. Hence the results obtained in the previous sections can be applied directly to partially observed diffusions.

First, let us remark that the results in §§5 and 6 can be easily extended to the following class of systems:

$$\begin{aligned} dq(t) &= [A(t,W(t),U(t))q(t) + f(t,W(t),U(t))]dt \\ &\quad + [M^k(t,W(t),U(t))q(t) + g^k(t,W(t),U(t))]dW_k(t), \quad t \in [0,1], \\ q(0) &= q_0, \end{aligned} \tag{6.3}$$

where

$$\begin{aligned} A(t,w,u)\phi(x) &:= \partial_i(a^{ij}(t,x,w,u)\partial_j\phi(x)) + b^i(t,x,w,u)\partial_i\phi(x) + c(t,x,w,u)\phi(x), \\ M^k(t,w,u)\phi(x) &:= \sigma^{ik}(t,x,w,u)\partial_i\phi(x) + h^k(t,x,w,u)\phi(x), \quad k = 1,2,\ldots,d'. \end{aligned}$$

Note only the continuity of the coefficients a^{ij} and so forth in $w \in R^{d'}$ will be needed later.

Let W and $\tilde{W}$ be two independent Brownian motions on a probability space $(\Omega, \mathcal{F}, P)$, with values in R^d and $R^{d'}$, respectively. Consider the following SDE in R^d:

$$\begin{aligned} dX(t) &= \gamma(t,X(t),Y(t),U(t))dt + \alpha(t,X(t),Y(t),U(t))dW(t) \\ &\quad + \sigma(t,X(t),Y(t),U(t))d\tilde{W}(t), \quad t \in [0,1], \\ X(0) &= \xi, \end{aligned} \tag{6.4}$$

with the observation

$$\begin{aligned} dY(t) &= \kappa(t,X(t),Y(t),U(t))dt + d\tilde{W}(t), \quad t \in [0,1], \\ Y(0) &= 0, \end{aligned} \tag{6.5}$$

where U is an admissible control, namely, $\{U(t) : 0 \le t \le 1\}$ is a Γ-valued, $\sigma\{Y(s) : 0 \le s \le t\}$-adapted measurable process. Note σ is the correlation between the state and the observation noises.

Let $F : [0,1] \times R^d \times R^{d'} \times \Gamma \to R^1, G : R^d \times R^{d'} \to R^1$ be given. The objective is to minimize the cost functional defined by

$$J(U) := E\left[\int_0^1 F(t,X(t),Y(t),U(t))dt + G(X(1),Y(1))\right] \tag{6.6}$$

over the set of admissible controls.

Note that W and Y are independent Brownian motions under a new probability $\tilde{P}$ defined by $d\tilde{P} := \rho^{-1}(1)dP$, where

$$\rho(t) := \exp\left[\int_0^t \kappa(s,X(s),Y(s),U(s))dY(s) - \frac{1}{2}\int_0^t |\kappa(s,X(s),Y(s),U(s))|^2 ds\right].$$

Consider the following SPDE:

$$\begin{aligned} dq(t) &= A(t,Y(t),U(t))q(t)dt + M^k(t,Y(t),U(t))q(t)dY_k(t), \quad t \in [0,1], \\ q(0) &= q_0, \end{aligned} \tag{6.7}$$

where $Y_k, k = 1, 2, \ldots, d'$, are the components of Y, and

$$A(t, y, u)\phi(x) := \partial_i(a^{ij}(t, x, y, u)\partial_j\phi(x)) - \partial_i(a^i(t, x, y, u)\phi(x)),$$
$$M^k(t, y, u)\phi(x) := -\sigma^{ik}(t, x, y, u)\partial_i\phi(x) + h^k(t, x, y, u)\phi(x),$$
$$(a^{ij}(t, x, y, u))_{ij} \equiv a(t, x, y, u) := [\sigma\sigma^T(t, x, y, u) + \alpha\alpha^T(t, x, y, u)]/2,$$
$$a^i(t, x, y, u) := \gamma^i(t, x, y, u) - \partial_j a^{ij}(t, x, y, u),$$
$$h^k(t, x, y, u) := \kappa^k(t, x, y, u) - \partial_i\sigma^{ik}(t, x, y, u), \quad i, j = 1, 2, \ldots, d \; k = 1, 2, \ldots, d',$$
$$q_0 := \text{the density of } \xi.$$

Then $q(t)$ proves to be the unnormalized conditional probability density of the state process $X(t)$, and the cost functional (6.6) reduces to

$$J(U) = \tilde{E}\left[\int_0^1 (F(t, \cdot, Y(t), U(t)), q(t))_0 dt + (G(\cdot Y(1)), q(1))_0\right]; \tag{6.8}$$

see Pardoux [13], Rozovskii [14], and Nagase and Nisio [12] for details.

Now we have turned problem (6.4)–(6.6) to the one already solved in §§5 and 6, observing the remark at the beginning of this section. So we can obtain the necessary conditions for an optimal solution to problem (6.4)–(6.6) by appropriately interpreting the assumptions and conclusions. Here we omit this simple interpretative work and only remark that, in the present case, (A2) is satisfied automatically; (A2)′ is also satisfied, provided that $\alpha\alpha^T$ is uniformly positive definite.

7. Concluding remarks. In this paper, a finite-dimensional approximation approach and some a prior estimates of differential operators are employed to solve the adjoint equations of linear SPDEs, based on which necessary conditions of optimality for controlled SPDEs are derived. For some sufficient conditions for the existence of an optimal control, refer to [12], [15].

The adjoint equations of SPDEs may also be applied to some other important problems besides the necessity of optimality. For example, a relationship between the adjoint process λ and the value function is given in [16], which enables us to gain some insight into the deep connection between the adjoint equations and the dynamic programming equations (Hamilton-Jacobi-Bellman equations). Further research on the analytical and qualitative properties of solutions of the adjoint equations together with their applications to linear SPDEs is carried out in [17].

Let us conclude the paper by noting that our results on the adjoint equations are by no means optimal. In view of the observation at the beginning of §6.1, it remains a challenging open problem to solve the adjoint equations of degerate SPDEs in which diffusion terms contain first-order differential operators.

Acknowledgments. This research was carried out when I was visiting Kobe University. I would like to express my hearty thanks to Professor M. Nisio for her constant encouragement and kind hospitality. Thanks are also due to the referees for their helpful comments and suggestions.

REFERENCES

[1] J. S. Baras, R. J. Elliott, and M. Kohlmann, *The partially observed stochastic minimum principle*, SIAM J. Control Optim., 27 (1989), pp. 1279–1292.

[2] A. Bensoussan, *Lecture on stochastic control, Part* I, Lecture Notes in Math., 972 (1983), pp. 1–39.

[3] ———, *Stochastic maximum principle for distributed parameter system*, J. Franklin Inst., 315 (1983), pp. 387–406.

[4] ———, *Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions*, Stochastics, 9 (1983), pp. 169–222.

[5] J. M. Bismut, *Analyse convexte et probabilités*, Thèse, Faculté des Sciences de Paris, 1973.

[6] G. Da Prato, M. Iannelli, and L. Tubaro, *Some results on linear stochastic differential equations in Hilbert spaces*, Stochastics, 6 (1982), pp. 105–116.

[7] U. G. Haussmann, *The maximum principle for optimal control of diffusions with partial information*, SIAM. J. Control Optim., 25 (1987), pp. 341–361.

[8] N. V. Krylov and B. L. Rozovskii, *On the Cauchy problem for linear stochastic partial differential equations*, Isz. Akad. Nauk SSSR Ser. Mat., 41 (1977), pp. 1329–1347; Mat. USSR-Izv., 11 (1977), pp. 1267–1284.

[9] ———, *On characteristics of the degenerate parabolic Ito equations of the second order*, Proc. Petrovskii Sem., 8 (1982), pp. 153–168. (In Russian.)

[10] ———, *Stochastic partial differential equations and diffusion processes*, Uspekhi Mat. Nauk, 37:6 (1982), pp. 75–95; Russian Math. Surveys, 37:6 (1982), pp. 81–105.

[11] H. J. Kushner, *Necessary conditions for continuous parameter stochastic optimization problems*, SIAM J. Control, 10 (1972), pp. 550–565.

[12] N. Nagase and M. Nisio, *Optimal controls for stochastic partial differential equations*, SIAM J. Control Optim., 28 (1990), pp. 186–213.

[13] E. Pardoux, *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics, 3 (1979), pp. 127–167.

[14] B. L. Rozovskii, *Nonnegative L^1-solutions of second order stochastic parabolic equations with random coefficients*, Steklov Seminar, Stat. Control of Stoch. Proc., 1984; Transl. Math. English, 1985, pp. 410–427.

[15] X. Y. Zhou, *On the existence of optimal relaxed controls for stochastic partial differential equations*, SIAM J. Control Optim., 30 (1992), pp. 247–261.

[16] ———, *Remarks on optimal controls of stochastic partial differential equations*, System Control Lett., 16 (1991), pp. 465–472.

[17] ———, *A duality analysis on stochastic partial differential equations*, J. Funct. Anal., 103 (1992), pp. 275–293.