J. Math. Anal. Appl. 395 (2012) 654–672

Contents lists available at SciVerse ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Optimality conditions for stochastic boundary control problems governed by semilinear parabolic equations[☆]

Huaiqiang Yu, Bin Liu[*]

School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, Hubei, PR China

ARTICLE INFO

Article history:
Received 11 February 2011
Available online 7 June 2012
Submitted by Martin Hairer

Keywords:
Stochastic partial differential equations
Boundary control
Necessary conditions
Convex state constraints
BSPDE with non-homogeneous double boundary conditions

ABSTRACT

We study the boundary control problems for stochastic parabolic equations with Neumann boundary conditions. Imposing super-parabolic conditions, we establish the existence and uniqueness of the solution of state and adjoint equations with non-homogeneous boundary conditions by the Galerkin approximations method. We also find that, in this case, the adjoint equation (BSPDE) has two boundary conditions (one is non-homogeneous, the other is homogeneous). By these results we derive necessary optimality conditions for the control systems under convex state constraints by the convex perturbation method.

1. Introduction

Let $(\Omega, \mathscr{F}, \mathscr{P})$ be a complete probability space equipped with a right-continuous filtration $\{\mathscr{F}_t\}_{t\geq 0}$, on which a d'-dimensional mutually independent standard Brownian motion $B = \{w(t)\}_{t\geq 0} = \{w_1(t), \ldots, w_{d'}(t)\}$ is defined. For fixed $0 < T < \infty$, denote by $\mathscr{P}$ the σ-algebra of predictable sets on $\Omega \times (0, T)$ associated with $\{\mathscr{F}_t\}$. Let $\mathcal{O}$ be an open subset in $\mathbb{R}^n$ with a smooth boundary Γ. Let $T < +\infty$ and $Q = \mathcal{O} \times (0, T)$, $\Sigma = \Gamma \times (0, T)$. When $\mathcal{X}$ is a topological space, $\mathcal{B}(\mathcal{X})$ stands for the σ-algebra of all Borel subsets of $\mathcal{X}$.

In the present paper, we consider the following stochastic distributed optimal control system with $\mathbb{R}$-value control processes in $\Omega \times \Sigma$:

$$\begin{cases} dy(x,t) = [Ay(x,t) + f(x,t,y(x,t))]dt + [By(x,t) + g(x,t,y(x,t))]dw(t) & \text{in } Q, \\ \partial_{\nu_A} y(x,t) = h(x,t,y(x,t),u(x,t)) & \text{on } \Sigma, \\ y(x,0) = y_0(x) & \text{in } \mathcal{O} \end{cases} \tag{1.1}$$

where $y_0 \in L^2(\mathcal{O})$ (a.s.) is a random function and adapted to $\mathscr{F}_0$, A, B are defined in the next section. The free terms f, h are $\mathbb{R}$-value functions and g is an $\mathbb{R}^{d'}$-value function. An adapted solution of the state equation (1.1) is a $\mathcal{B}(\mathbb{R}^n) \times \mathcal{B}([0,T])$-measurable and $\{\mathscr{F}_t\}_{t\geq 0}$-adapted function satisfying (1.1) under some appropriate sense.

We introduce the following cost functional

$$J(y^u, u) = E\int_Q L(x,t,y^u(x,t))dxdt + E\int_\Sigma l(x,t,y^u(x,t),u(x,t))d\rho(x)dt + E\int_{\mathcal{O}} r(x,y^u(x,T))dx, \tag{1.2}$$

where y^u is the adapted solution of (1.1) associated with u, ρ denotes the usual $(n-1)$-dimensional measure on Γ, and L, l and r are $\mathbb{R}$-value functions.

☆ This work was partially supported by NNSF of China (Grant No. 11171122).

* Corresponding author.
E-mail addresses: bingliug@public.wh.hb.cn, binliu@mail.hust.edu.cn (B. Liu).

0022-247X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmaa.2012.05.071

The space of feasible controls $\mathcal{U}$ is the set of u satisfying (1) $u : \Omega \times (0, T) \times \Gamma \to \mathbb{R}$ is $\mathcal{B}(\mathbb{R}^{n-1}) \times \mathcal{B}([0, T])$-measurable and $\{\mathscr{F}_t\}_{t\geq 0}$-adapted, (2) $E \int_\Sigma |u(x, t)|^2 d\rho(x)dt < \infty$.

The set of admissible controls $\mathcal{U}_{ad}$ is a closed, convex subset of $\mathcal{U}$.

Our optimal control problem can be stated as follows:

Problem (C). Find a $\bar{u} \in \mathcal{U}_{ad}$ such that

$$J(y^{\bar{u}}, \bar{u}) = \inf_{u\in\mathcal{U}_{ad}} J(y^u, u). \tag{1.3}$$

Any $\bar{u}(\cdot, \cdot) \in \mathcal{U}_{ad}$ satisfying the above equality is called an optimal control, and the corresponding state $y^{\bar{u}}(\cdot, \cdot) \equiv y(\cdot, \cdot; y_0, \bar{u})$ is called an optimal state. The pair $(y^{\bar{u}}(\cdot, \cdot), \bar{u}(\cdot, \cdot))$ is called an optimal pair.

The boundary control problem (1.1)–(1.3) is extensively studied in many papers in the case of determinate control systems, for example, we can see [1,2] and the reference in that papers. But in random cases, there are very few papers concerning this problem. One of the difficulties is that properties of the solution of state equation (1.1) seems not clear compared to the determinate cases. Another difficulty is the existence and uniqueness of the backward partial differential equations. The purpose of this paper is to study necessary conditions of an optimal boundary control problem for the control systems governed by stochastic parabolic equations with Neumann boundary conditions through convex perturbation. We know that the main method is spike perturbation in similar problems, but if we use this method, stronger regularities of the state are necessary. So the conditions we need must be stronger than we will give in the next section.

The optimality conditions for stochastic control systems have been considered in many papers. In finite dimensional spaces, see [3–14], etc. For the cases of infinite dimensional spaces, to our knowledge, in general cases, there are very few papers to treat this problem. In [15], the author considered the case in which the system is a semilinear evolution control system and the infinitesimal operator is strongly elliptic by the variational method. For a general C_0-semigroup, the stochastic maximum principle has been established in [16]. Zhou [17] studied this problem for a linear stochastic partial differential equation, and obtained the corresponding maximum principle without state constraints. In [15], more than half of the paper is devoted to obtaining the existence and uniqueness of the adjoint equation. By introducing an extension of the martingale representation theorem, in [16], the authors derived the existence and uniqueness of the adjoint equations for the C_0-semigroup. This method is extensively applied to many backward differential systems, for instance, see [18]. In [17], the author used the Galerkin approximation to get the existence and uniqueness of the adjoint equations.

The paper is organized as follows: In Section 2, we state various assumptions and the main results. Sections 3 and 4 are devoted to the proofs of the existence and uniqueness of the state and adjoint equations. In Section 5, we will give the proof of the main results that are given in Section 2.

2. Notation and statement of the main results

Let $H^1(\mathcal{O})$ be the Sobolev space $W^{1,2}(\mathcal{O})$. We denote

$$H^0(\mathcal{O}) = L^2(\mathcal{O}) \quad \text{and} \quad \mathbb{H}^1 = L^2_{\mathscr{F}}(\Omega \times (0, T); H^1(\mathcal{O})).$$

In addition, we denote the norm of the spaces H^0 and H^1 by $\|\cdot\|_0 = \|\cdot\|_{H^0}$ and $\|\cdot\|_1 = \|\cdot\|_{H^1}$, respectively. Moreover, for a function $y \in \mathbb{H}^1$, we denote

$$|||y|||_1^2 = E \int_0^T \|y(\cdot, t)\|_1^2 dt.$$

Similarly, we can define $\mathbb{H}^0$ and $|||y|||_0$. Throughout this paper, unless a special explanation is given, the positive constant C will have different values on different occasions.

$A, B = (B^1, \ldots, B^{d'})$ are linear operators defined by

$$Ay(x, t) = \sum_{i,j=1}^n \partial_{x_j}[a^{ij}(x, t)\partial_{x_i} y(x, t) + b^j(x, t)y(x, t)] + \sum_{j=1}^n d^j(x, t)\partial_{x_j} y(x, t) + c(x, t)y(x, t), \tag{2.1}$$

$$B^k y(x, t) = \sum_{i=1}^n \sigma^{ik}(x, t)\partial_{x_i} y(x, t) + \eta^k(x, t)y(x, t), \quad k = 1, \ldots, d' \tag{2.2}$$

and

$$\partial_{\nu_A} y(x, t) = \sum_{j=1}^n \left\{ \sum_{i=1}^n [a^{ij}(x, t)\partial_{x_i} y(x, t)] + b^j(x, t)y(x, t) \right\} \nu_j(x), \tag{2.3}$$

$\nu(x)$ being the outward unit normal vector to Γ at the point x. For an interpretation of this Neumann condition in a trace sense we can see [1].

The results of this paper rely on the following assumptions:

Hypothesis 2.1. The coefficients $a = (a^{ij})$, $b = (b^j)$, $d = (d^j)$, c, $\sigma = (\sigma^{ik})$, and $\eta = (\eta^k)$ are measurable in (x, t) with values in the set of real symmetric $n \times n$ matrices, $\mathbb{R}^n$, $\mathbb{R}^{n\times n}$, $\mathbb{R}$, $\mathbb{R}^{n\times d}$ and $\mathbb{R}^d$, respectively. The real function y_0 is $\mathscr{F}_0 \times \mathcal{B}(\mathbb{R}^n)$-measurable and $E\|y_0(x)\|_0^2 \le K$, where K is a constant. Furthermore, the functions a^{ij}, b^j, c, d^j, σ^{ik} and η^k do not exceed K in absolute value. Also the matrix $S = (a^{ij} - \frac{1}{2}\sum_{k=1}^{d'} \sigma^{ik}\sigma^{jk})$ is uniformly positive definite:

$$\xi^T S \xi \ge \lambda |\xi|^2 \quad \text{for any } (t, x) \text{ and any } \xi \in \mathbb{R}^n,$$

for a fixed constant $\lambda > 0$.

Hypothesis 2.2. $f : \Omega \times \mathbb{R}^n \times [0, T] \times \mathbb{R} \to \mathbb{R}$, $g = (g_1, \dots, g_{d'}) : \Omega \times \mathbb{R}^n \times [0, T] \times \mathbb{R} \to \mathbb{R}^{d'}$ and $h : \Omega \times \mathbb{R}^{n-1} \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are mappings such that the following properties are satisfied:

(i) f is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable, g is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R}^{d'})$-measurable and h is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{n-1}) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable.

(ii) $f(\cdot, \cdot, 0), g_k(\cdot, \cdot, 0) \in \mathbb{H}^0$ and $E\int_0^T \int_\Gamma |h(x, t, 0, u(x, t))|^2 dxdt < \infty$ ($k = 1, \dots, d'$, $\forall u(\cdot, \cdot) \in \mathcal{U}$).

(iii) There exists an $M > 0$ such that $\forall y, y' \in \mathbb{R}$

$$|f(x, t, y) - f(x, t, y')| + \sum_{k=1}^{d'} |g_k(x, t, y) - g_k(x, t, y')| \le M|y - y'| \quad \text{uniformly in } (\omega, t, x) \in \Omega \times Q$$

for every $k \in (1, \dots, d')$ and

$$\left|\frac{\partial h}{\partial y}(x, t, y, u)\right| \le M \quad \text{uniformly in } (\omega, t, x, u) \in \Omega \times \Sigma \times \mathbb{R}.$$

(iv)

$$\frac{\partial h}{\partial y}(x, t, y, u) \le 0 \quad \text{for almost every } (\omega, t, x, u) \in \Omega \times \Sigma \times \mathbb{R}.$$

Hypothesis 2.3. $f : \Omega \times \mathbb{R}^n \times [0, T] \times \mathbb{R} \to \mathbb{R}$, $g = (g_1, \dots, g_{d'}) : \Omega \times \mathbb{R}^n \times [0, T] \times \mathbb{R} \to \mathbb{R}^{d'}$ and $h : \mathbb{R}^{n-1} \times [0, T] \times \mathbb{R} \to \mathbb{R}$ are mappings such that the following properties are satisfied:

(i) f is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable, g is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R}^{d'})$-measurable and h is $\mathcal{B}(\mathbb{R}^{n-1}) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable.

(ii) $f(\cdot, \cdot, 0), g_k(\cdot, \cdot, 0) \in \mathbb{H}^0$. For almost every $(x, t) \in \Sigma$, $h(x, t, y, \cdot)$ is a continuously differentiable function for every fixed $(x, t, y) \in \Sigma \times \mathbb{R}$ and $\int_0^T \int_\Gamma |h(x, t, 0, 0)|^2 dxdt < \infty$.

(iii) f, g, h are continuously differentiable functions with respect to state. Also there exists an $M > 0$ such that

$$\left|\frac{\partial f}{\partial y}(x, t, y)\right| + \sum_{k=1}^{d'} \left|\frac{\partial g_k}{\partial y}(x, t, y)\right| \le M \quad \text{uniformly in } (\omega, t, x) \in \Omega \times Q$$

for every $k \in (1, \dots, d')$ and

$$\left|\frac{\partial h}{\partial y}(x, t, y, u)\right| + \left|\frac{\partial h}{\partial u}(x, t, y, u)\right| \le M \quad \text{uniformly in } (t, x, u) \in \Sigma \times \mathbb{R}.$$

(iv)

$$\frac{\partial h}{\partial y}(x, t, y, u) \le 0 \quad \text{for almost every } (t, x, u) \in \Sigma \times \mathbb{R}.$$

Hypothesis 2.4. $L : \mathbb{R}^n \times [0, T] \times \mathbb{R} \to \mathbb{R}$, $l : \mathbb{R}^{n-1} \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $r : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ are mappings such that the following conditions are satisfied:

(i) L is $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable, l is $\mathcal{B}(\mathbb{R}^{n-1}) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable and r is $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})/\mathcal{B}(\mathbb{R})$-measurable.

(ii) $L(\cdot, \cdot, 0) \in L^2(Q)$, $l(\cdot, \cdot, 0, u(\cdot, \cdot)) \in L^2(\Sigma)$ for every $u(\cdot, \cdot) \in \mathcal{U}$ and $r(\cdot, 0) \in L^2(\mathcal{O})$.

(iii) L, l, r are continuously differentiable functions with respect to state. Also there exists an $M > 0$ such that

$$\left|\frac{\partial L}{\partial y}(x, t, y)\right| + \left|\frac{\partial r}{\partial y}(x, y)\right| \le M \quad \text{uniformly in } (t, x) \in Q$$

and

$$\left|\frac{\partial l}{\partial y}(x, t, y, u)\right| \le M \quad \text{uniformly in } (t, x, u) \in \Sigma \times \mathbb{R}.$$

Next, we give the necessary condition of control systems (1.1)–(1.3):

Theorem 2.5. *Let Hypotheses 2.1, 2.3 and 2.4 hold. Let $(\bar{y}(\cdot,\cdot),\bar{u}(\cdot,\cdot))$ be an optimal pair of Problem (C). Let $(\bar{p}(\cdot,\cdot),\bar{q}(\cdot,\cdot))\in\mathbb{H}^1\times[\mathbb{H}^0]^{d'}$ be the solution of the following adjoint equation:*

$$\begin{cases} d\bar{p}(x,t) = -\left[A^*\bar{p}(x,t)+\sum_{k=1}^{d'}B_k^*\bar{q}_k(x,t)+\frac{\partial f}{\partial y}(x,t,\bar{y}(x,t))\bar{p}(x,t)+\frac{\partial L}{\partial y}(x,t,\bar{y}(x,t))\right. \\ \qquad \left.+\sum_{k=1}^{d'}\frac{\partial g_k}{\partial y}(x,t,\bar{y}(x,t))\bar{q}_k(x,t)\right]dt+\bar{q}(x,t)dw(t) & \text{in } Q, \\ \partial_{\nu_{A^*}}\bar{p}(x,t)=\frac{\partial h}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t))\bar{p}(x,t)+\frac{\partial l}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t)) & \text{on } \Sigma, \\ \partial_{\nu_{B_k^*}}\bar{q}_k(x,t)=0,\quad k\in\{1,2,\ldots,d'\} & \text{on } \Sigma, \\ \bar{p}(x,T)=\frac{\partial r}{\partial y}(x,\bar{y}(x,T)) & \text{in } \mathcal{O} \end{cases} \tag{2.4}$$

where

$$q(x,t)=(q_1(x,t),\ldots,q_{d'}(x,t)),$$

$$A^*p(x,t)=\sum_{i=1}^n\partial_{x_i}\left\{\sum_{j=1}^n[a^{i,j}(x,t)\partial_{x_j}p(x,t)]-d^i(x,t)p(x,t)\right\}-\sum_{j=1}^n b^j(x,t)\partial_{x_j}p(x,t)+c(x,t)p(x,t),$$

$$B_k^*q_k(x,t)=-\sum_{i=1}^n\partial_{x_i}[\sigma^{ik}(x,t)q_k(x,t)]+\eta^k(x,t)q_k(x,t),\quad k\in\{1,2,\ldots,d'\},$$

$$\partial_{\nu_{A^*}}p(x,t)=\sum_{i=1}^n\left\{\sum_{j=1}^n[a^{i,j}(x,t)\partial_{x_j}p(x,t)]-d^i(x,t)p(x,t)\right\}\nu_i(x),$$

and

$$\partial_{\nu_{B_k^*}}q_k(x,t)=-\sum_{i=1}^n[\sigma^{ik}(x,t)q_k(x,t)]\nu_i(x)\quad k\in\{1,2,\ldots,d'\}.$$

Then

$$\begin{aligned} &E\int_\Sigma H_w(x,t,\bar{p}(x,t),\bar{y}(x,t),\bar{u}(x,t),\bar{u}(x,t))d\rho(x)dt \\ &\quad=\min_{u(\cdot,\cdot)\in U}E\int_\Sigma H_w(x,t,\bar{p}(x,t),\bar{y}(x,t),\bar{u}(x,t),u(x,t))d\rho(x)dt \end{aligned} \tag{2.5}$$

where the Hamiltonian $H_w:\Sigma\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is defined by

$$H_w(x,t,p,y,u,v)=\left[p\frac{\partial h}{\partial u}(x,t,y,u)+\frac{\partial l}{\partial u}(x,t,y,u)\right]v.$$

In the following, we apply this result to an example.

Example 2.1. Let the cost function be

$$T(u)=E\int_Q(y^u(x,t)-z_d(x,t))^2dxdt+E\int_\Sigma|u(x,t)|^2d\rho(x)dt,\quad z_d\in L^2(Q). \tag{2.6}$$

Then the adjoint state $(\bar{p},\bar{q})$ is given by

$$\begin{cases} d\bar{p}(x,t) = -\left[A^*\bar{p}(x,t)+\sum_{k=1}^{d'}B_k^*\bar{q}_k(x,t)+\frac{\partial f}{\partial y}(x,t,\bar{y}(x,t))\bar{p}(x,t)+2(\bar{y}(x,t)-z_d(x,t))\right. \\ \qquad \left.+\sum_{k=1}^{d'}\frac{\partial g_k}{\partial y}(x,t,\bar{y}(x,t))\bar{q}_k(x,t)\right]dt+\bar{q}(x,t)dw(t) & \text{in } Q, \\ \partial_{\nu_{A^*}}\bar{p}(x,t)=\frac{\partial h}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t))\bar{p}(x,t) & \text{on } \Sigma, \\ \partial_{\nu_{B_k^*}}\bar{q}_k(x,t)=0,\quad k\in\{1,2,\ldots,d'\} & \text{on } \Sigma, \\ \bar{p}(x,T)=0 & \text{in } \mathcal{O} \end{cases} \tag{2.7}$$

and the optimality condition is

$$E\int_{\Sigma}\left[\bar{p}(x,t)\frac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)\right](u(x,t)-\bar{u}(x,t))d\rho(x)dt\geq 0\quad \forall u\in\mathcal{U}_{ad}. \tag{2.8}$$

(i) The case where there are no constraints ($\mathcal{U}_{ad}=\mathcal{U}$).
Then (2.8) reduces to the equation

$$\bar{p}(x,t)\frac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)=0\quad \text{on } \Sigma \text{ a.s.} \tag{2.9}$$

(ii) $\mathcal{U}_{ad}=\{v|v\geq 0 \text{ almost everywhere on } \Sigma \text{ a.s.}\}$.
Then (2.8) reduces

$$\begin{cases}\bar{p}(x,t)\dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)\geq 0 & \text{on } \Sigma \text{ a.s.,}\\ u(x,t)\geq 0 & \text{on } \Sigma \text{ a.s.,}\\ \left[\bar{p}(x,t)\dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)\right]\bar{u}(x,t)=0 & \text{on } \Sigma \text{ a.s.}\end{cases} \tag{2.10}$$

(iii) $\mathcal{U}_{ad}=\{v|\xi_0(x,t)\leq v(x,t)\leq \xi_1(x,t) \text{ almost everywhere on } \Sigma \text{ a.s., } \xi_i\in L^{\infty}(\Sigma)\}$.
From (2.8) we then deduce

$$\begin{cases}\bar{p}(x,t)\dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)>0, & \bar{u}(x,t)=\xi_0(x,t),\\ \bar{p}(x,t)\dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)<0, & \bar{u}(x,t)=\xi_1(x,t),\\ \bar{p}(x,t)\dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))+2\bar{u}(x,t)=0. & \end{cases} \tag{2.11}$$

3. State equations

In this section we will give the existence of unique solutions of state equations for a fixed control $u(x,t)$ (we will omit it for simplicity). The main method is *Galerkin* approximations. To avoid notational complexity, we prove the results for $d'=1$ (there is no essential difficulty when $d'>1$). We assume the functions $e_k=e_k(x)(k=1,2,\ldots)\in H^1(\mathcal{O})$, and, for the sake of nonessential simplifications, we let

$$\{e_k\}_{k=1}^{\infty} \text{ is a basis of } H^1(\mathcal{O}), \tag{3.1}$$

and

$$\{e_k\}_{k=1}^{\infty} \text{ is an orthogonal basis of } H^0(\mathcal{O}). \tag{3.2}$$

First, the definition of solution of the state equation is given as follows:

Definition 3.1. We say a function $y\in\mathbb{H}^1(\mathcal{O})$ is a weak solution of the non-homogeneous Neumann boundary value problem (1.1), if it satisfies

$$\begin{aligned}\int_{\mathcal{O}}y(x,t)\phi(x)dx=&\int_{\mathcal{O}}y_0(x)\phi(x)dx-\sum_{i,j=1}^{n}\int_0^t\int_{\mathcal{O}}a^{ij}(x,s)\partial_{x_i}y(x,s)\partial_{x_j}\phi(x)dxds\\&-\sum_{j=1}^{n}\int_0^t\int_{\mathcal{O}}b^j(x,s)y(x,s)\partial_{x_j}\phi(x)dxds+\sum_{j=1}^{n}\int_0^t\int_{\mathcal{O}}d^j(x,s)\partial_{x_j}y(x,s)\phi(x)dxds\\&+\int_0^t\int_{\mathcal{O}}c(x,s)y(x,s)\phi(x)dxds+\int_0^t\int_{\mathcal{O}}f(x,s,y(x,s))\phi(x)dxds\\&+\int_0^t\int_{\Gamma}h(x,s,y(x,s))\phi(x)d\rho(x)ds+\sum_{i=1}^{n}\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\partial_{x_i}y(x,t)\phi(x)dxdw(s)\\&+\int_0^t\int_{\mathcal{O}}\eta(x,s)y(x,s)\phi(x)dxdw(s)+\int_0^t\int_{\mathcal{O}}g(x,s,y(x,s))\phi(x)dxdw(s),\ \text{a.s.}\end{aligned} \tag{3.3}$$

for every $\phi\in H^1(\mathcal{O})$ and almost every $t\in(0,T]$.

Our main aim now is to prove the following theorem.

Theorem 3.1 (*Existence and Uniqueness of State Equation*). *Under Hypotheses 2.1 and 2.2, state equation (1.1) has a unique (weak) solution y in $\mathbb{H}^1$ for every control $u \in \mathcal{U}$. Moreover, there exists a constant $C > 0$ such that*

$$\sup_{0\le t\le T} E\|y(t)\|_0^2 + \|\!|y|\!\|_1^2 \le C. \tag{3.4}$$

Remark 3.1. By this theorem we can denote $J(y^u, u)$ by $J(u)$.

We shall divide the proof of Theorem 3.1 into two cases which considers a simple version of Eq. (1.1).

Lemma 3.1. *Suppose that $f, g \in L^2_{\mathscr{F}}(\Omega \times (0, T); H^0(\mathcal{O}))$, $h \in L^2_{\mathscr{F}}(\Omega \times (0, T); H^0(\Gamma))$ and Hypothesis 2.1 holds. Then,*

$$\begin{cases} dy(x,t) = [Ay(x,t) + f(x,t)]dt + [By(x,t) + g(x,t)]dw(t) & \text{in } Q, \\ \partial_{\nu_A} y(x,t) = h(x,t) & \text{on } \Sigma, \\ y(x,0) = y_0(x) & \text{in } \mathcal{O} \end{cases} \tag{3.5}$$

have a unique solution y in $\mathbb{H}^1$ satisfying (3.4).

Proof. Fix now a positive integer m, we look for a function $y_m : \Omega \times \mathbb{R}^n \times [0, T] \to \mathbb{R}$ of the form

$$y_m(x,t) = \sum_{k=1}^m r_m^k(t) e_k(x). \tag{3.6}$$

We hope to select the coefficients $r_m^k(t)$ of y_m such that

$$r_m^k(0) = \int_{\mathcal{O}} y_0(x) e_k(x) dx \quad (k = 1, \ldots, m) \tag{3.7}$$

and

$$\begin{aligned} r_m^k(t) &= r_m^k(0) - \sum_{i,j=1}^n \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} a^{ij}(x,s) \partial_{x_i} e_p(x) \partial_{x_j} e_k(x) dx r_m^p(s) ds \\ &\quad - \sum_{j=1}^n \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} b^j(x,s) e_p(x) \partial_{x_j} e_k(x) dx r_m^p(s) ds \\ &\quad + \sum_{j=1}^n \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} d^j(x,s) \partial_{x_j} e_p(x) e_k(x) dx r_m^p(s) ds + \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} c(x,s) e_p(x) e_k(x) dx r_m^p(s) ds \\ &\quad + \int_0^t \int_{\mathcal{O}} f(x,s) e_k(x) dx ds + \int_0^t \int_{\Gamma} h(x,s) e_k(x) d\rho(x) ds \\ &\quad + \sum_{i=1}^n \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} \sigma^i(x,s) \partial_{x_i} e_p(x) e_k(x) dx r_m^p(s) dw(s) \\ &\quad + \sum_{p=1}^m \int_0^t \int_{\mathcal{O}} \eta(x,s) e_p(x) e_k(x) dx r_m^p(s) dw(s) + \int_0^t \int_{\mathcal{O}} g(x,s) e_k(x) dx dw(s) \end{aligned} \tag{3.8}$$

for $k = 1, \ldots, m$ and almost every $(\omega, t) \in \Omega \times [0, T]$.

By the results of [13], we can get that the Eq. (3.8) have a unique $\{\mathscr{F}_t\}_{t\ge0}$-adapted continuous solution $(r_m^1, \ldots, r_m^m)$ for every $m \in \mathbb{R}$.

Next, we shall give the energy estimates of y_m. Applying Ito's formula to $|r_m^k(t)|^2$, and adding up in k from 1 to m, we have

$$\begin{aligned} E\int_{\mathcal{O}} |y_m(x,t)|^2 dx &= E\int_{\mathcal{O}} |y_m(x,0)|^2 dx - 2E\sum_{i,j=1}^n \int_0^t \int_{\mathcal{O}} a^{ij}(x,s) \partial_{x_i} y_m(x,s) \partial_{x_j} y_m(x,s) dx ds \\ &\quad + 2E\sum_{j=1}^n \int_0^t \int_{\mathcal{O}} [d^j(x,s) - b^j(x,s)] y_m(x,s) \partial_{x_j} y_m(x,s) dx ds \\ &\quad + 2E\int_0^t \int_{\mathcal{O}} c(x,s) |y_m(x,s)|^2 dx ds + 2E\int_0^t \int_{\mathcal{O}} f(x,s) y_m(x,s) dx ds \end{aligned}$$

$$
\begin{aligned}
&+2E\int_0^t\int_\Gamma h(x,s)y_m(x,s)d\rho(x)ds\\
&+E\int_0^t\sum_{k=1}^m\left[\sum_{i=1}^n\int_\mathcal{O}\sigma^i(x,s)\partial_{x_i}y_m(x,s)e_k(x)dx\right]^2ds\\
&+E\int_0^t\sum_{k=1}^m\left[\int_\mathcal{O}\eta(x,s)y_m(x,s)e_k(x)dx\right]^2ds+E\int_0^t\sum_{k=1}^m\left[\int_\mathcal{O}g(x,s)e_k(x)dx\right]^2ds\\
&+2E\int_0^t\sum_{k=1}^m\left[\left(\sum_{i=1}^n\int_\mathcal{O}\sigma^i(x,s)\partial_{x_i}y_m(x,s)e_k(x)dx\right)\int_\mathcal{O}\eta(x,s)y_m(x,s)e_k(x)\right]dxds\\
&+2E\int_0^t\sum_{k=1}^m\left[\left(\sum_{i=1}^n\int_\mathcal{O}\sigma^i(x,s)\partial_{x_i}y_m(x,s)e_k(x)dx\right)\int_\mathcal{O}g(x,s)e_k(x)dx\right]dxds\\
&+2E\int_0^t\sum_{k=1}^m\left[\int_\mathcal{O}\eta(x,s)y_m(x,s)e_k(x)dx\int_\mathcal{O}g(x,s)e_k(x)dx\right]ds.
\end{aligned}\tag{3.9}
$$

By Yong's inequality and Hypothesis 2.1, we get

$$
E\|y_m(t)\|_0^2\le C(\varepsilon)-2\lambda E\int_0^t\|y_m(s)\|_1^2ds+C\varepsilon E\int_0^t\|y_m(s)\|_1^2ds+CE\int_0^t\|y_m(s)\|_0^2ds,\tag{3.10}
$$

where the constant C and $C(\varepsilon)$ depend on f,g,h,ε,K and the imbedding theorem constants. Let $\varepsilon=\lambda/C$ in (3.10), we get

$$
E\|y_m(t)\|_0^2+\lambda E\int_0^t\|y_m(s)\|_1^2ds\le C+CE\int_0^t\|y_m(s)\|_0^2ds.\tag{3.11}
$$

Hence, Gronwall's inequality yields

$$
\sup_{0\le t\le T}E\|y_m(t)\|_0^2+\|\!|y_m|\!\|_1^2\le C.\tag{3.12}
$$

So the sequence $\{y_m\}_{m=1}^\infty$ is bounded in $\mathbb{H}^1$. Consequently, there exists a subsequence (still denoted by itself) and a function $y\in\mathbb{H}^1$ such that

$$
y_m\to y\quad\text{weakly in }\mathbb{H}^1.\tag{3.13}
$$

Next, we prove that y satisfies (3.5). For this end, we take ψ to be an absolutely continuous function mapping $[0,T]$ to $\mathbb{R}$ with $\psi'=d\psi/dt\in L^2([0,T])$ and $\psi(T)=0$. Let $\psi^k(t)=\psi(t)e_k$ multiply y_m by $\psi^k(t)$ and using Ito's formula, we get

$$
\begin{aligned}
-\int_0^T\int_\mathcal{O}y_m(x,s)e_k(x)\psi'(s)dxds=\;&\psi^k(0)y_m(0)-\sum_{i,j=1}^n\int_0^T\int_\mathcal{O}a^{ij}(x,s)\partial_{x_i}y_m(x,s)\partial_{x_j}e_k(x)\psi(s)dxds\\
&-\sum_{j=1}^n\int_0^T\int_\mathcal{O}b^j(x,s)y_m(x,s)\partial_{x_j}e_k(x)\psi(s)ds\\
&+\sum_{j=1}^n\int_0^T\int_\mathcal{O}d^j(x,s)\partial_{x_j}y_m(x,s)e_k(x)\psi(s)dxds\\
&+\int_0^T\int_\mathcal{O}[c(x,s)y_m(x,s)+f(x,s)]e_k(x)\psi(s)dxds\\
&+\int_0^T\int_\Gamma h(x,s)e_k(x)\psi(s)d\rho(x)ds\\
&+\sum_{i=1}^n\int_0^T\int_\mathcal{O}\sigma^i(x,s)\partial_{x_i}y_m(x,s)e_k(x)\psi(s)dxdw(s)\\
&+\int_0^T\int_\mathcal{O}[\eta(x,s)y_m(x,s)+g(x,s)]e_k(x)\psi(s)dxdw(s).
\end{aligned}\tag{3.14}
$$

By (3.13), letting m go to infinity, we conclude that

$$
\begin{aligned}
-\int_0^T\int_{\mathcal{O}} y(x,s)\phi(x)\psi'(s)dxds = {} & \psi(0)\int_{\mathcal{O}}\phi(x)y(x,0)dx \\
& -\sum_{i,j=1}^n\int_0^T\int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_i}y(x,s)\partial_{x_j}\phi(x)\psi(s)dxds \\
& -\sum_{j=1}^n\int_0^T\int_{\mathcal{O}} b^j(x,s)y(x,s)\partial_{x_j}\phi(x)\psi(s)ds \\
& +\sum_{j=1}^n\int_0^T\int_{\mathcal{O}} d^j(x,s)\partial_{x_j}y(x,s)\phi(x)\psi(s)dxds \\
& +\int_0^T\int_{\mathcal{O}}[c(x,s)y(x,s)+f(x,s)]\phi(x)\psi(s)dxds \\
& +\int_0^T\int_{\Gamma} h(x,s)\phi(x)\psi(s)d\rho(x)ds \\
& +\sum_{i=1}^n\int_0^T\int_{\mathcal{O}}\sigma^i(x,s)\partial_{x_i}y(x,s)\phi(x)\psi(s)dxdw(s) \\
& +\int_0^T\int_{\mathcal{O}}[\eta(x,s)y(x,s)+g(x,s)]\phi(x)\psi(s)dxdw(s)
\end{aligned}
\tag{3.15}
$$

for any $\phi \in H^1$. For any $t \in (0,T)$, let

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

Substituting (3.15) with ψ_ε and letting $\varepsilon \to 0$, we get that $y \in \mathbb{H}^1$ is a solution of (3.5).

Next we shall prove the uniqueness of solution of (3.5). It is easy obtained by the linearity of (3.5) and Gronwall's inequality. So we omit it. □

In the following, we complete the proof of Theorem 3.1:

Proof. (1) Uniqueness. We suppose that y^1 and y^2 are two solutions of state equation (1.1). Let $\bar{y} = y_1 - y_2$. Then applying Ito's formula to $\|\bar{y}\|_0^2$, we get

$$
\begin{aligned}
E\int_{\mathcal{O}}|\bar{y}(x,t)|^2dx = {} & -2E\sum_{i,j=1}^n\int_0^t\int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_i}\bar{y}(x,s)\partial_{x_j}\bar{y}(x,s)dxds \\
& +2E\sum_{j=1}^n\int_0^t\int_{\mathcal{O}}[d^j(x,s)-b^j(x,s)]\partial_{x_j}\bar{y}(x,s)\bar{y}(x,s)dxds \\
& +2E\int_0^t\int_{\mathcal{O}}\{c(x,s)|\bar{y}(x,s)|^2+[f(x,s,y^1(x,s))-f(x,s,y^2(x,s))]\bar{y}(x,s)\}dxds \\
& +2E\int_0^t\int_{\Gamma}[h(x,s,y^1(x,s))-h(x,s,y^2(x,s))]\bar{y}(x,s)d\rho(x)ds \\
& +E\sum_{i,j=1}^n\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\sigma^j(x,s)\partial_{x_i}\bar{y}(x,s)\partial_{x_j}\bar{y}(x,s)dxds \\
& +E\int_0^t\int_{\mathcal{O}}[\eta(x,s)\bar{y}(x,s)]^2+[g(x,s,y^1(x,s))-g(x,s,y^2(x,s))]^2dxds \\
& +2E\sum_{i=1}^n\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\eta(x,s)\partial_{x_i}\bar{y}(x,s)\bar{y}(x,s)dxds
\end{aligned}
$$

$$
\begin{aligned}
&+2E\sum_{i=1}^{n}\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\partial_{x_i}\bar{y}(x,s)(g(x,s,y^1(x,s))-g(x,s,y^2(x,s)))dxds\\
&+2E\int_0^t\int_{\mathcal{O}}\eta(x,s)\bar{y}(x,s)(g(x,s,y^1(x,s))-g(x,s,y^2(x,s)))dxds. \qquad (3.16)
\end{aligned}
$$

Using the Hypotheses 2.1 and 2.2, and Yong's inequality, similar to Lemma 3.1, we obtain that

$$E\|\bar{y}(t)\|_0^2+\lambda E\int_0^t\|\bar{y}(s)\|_1^2ds\le CE\int_0^t\|\bar{y}(s)\|_0^2ds. \qquad (3.17)$$

Hence Gronwall's inequality yields the uniqueness of state equation (1.1).

(2) Existence. Let $y^0\equiv 0$ and define recursively by using Lemma 3.1 the following equation:

$$
\begin{cases}
dy^n(x,t)=[Ay^n(x,t)+f(x,t,y^{n-1}(x,t))]dt+[By^n(x,t)+g(x,t,y^{n-1}(x,t))]dw(t) & \text{in } Q,\\
\partial_{\nu_A}y^n(x,t)=h(x,t,y^{n-1}(x,t)) & \text{on } \Sigma,\\
y^n(x,0)=y_0(x) & \text{in } \mathcal{O}
\end{cases} \qquad (3.18)
$$

for $n\ge 1$. By using Lemma 3.1 we know that the solution y^n lies in $\mathbb{H}^1$ for each $n\ge 1$ and satisfies (3.4).

By applying Ito's formula, denote by $Y^{n+1}=y^{n+1}-y^n$, it follows that

$$
\begin{aligned}
E\int_{\mathcal{O}}|Y^{n+1}(x,t)|^2dx=&-2E\sum_{i,j=1}^{n}\int_0^t\int_{\mathcal{O}}a^{ij}(x,s)\partial_{x_i}Y^{n+1}(x,s)\partial_{x_j}Y^{n+1}(x,s)dxds\\
&+2E\sum_{j=1}^{n}\int_0^t\int_{\mathcal{O}}[d^j(x,s)-b^j(x,s)]\partial_{x_j}Y^{n+1}(x,s)Y^{n+1}(x,s)dxds\\
&+2E\int_0^t\int_{\mathcal{O}}\{c(x,s)|Y^{n+1}(x,s)|^2+[f(x,s,y^n(x,s))-f(x,s,y^{n-1}(x,s))]Y^{n+1}(x,s)\}dxds\\
&+2E\int_0^t\int_{\Gamma}[h(x,s,y^n(x,s))-h(x,s,y^{n-1}(x,s))]Y^{n+1}(x,s)d\rho(x)ds\\
&+E\sum_{i,j=1}^{n}\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\sigma^j(x,s)\partial_{x_i}Y^{n+1}(x,s)\partial_{x_j}Y^{n+1}(x,s)dxds\\
&+E\int_0^t\int_{\mathcal{O}}[\eta(x,s)Y^{n+1}(x,s)]^2+[g(x,s,Y^n(x,s))-g(x,s,Y^{n-1}(x,s))]^2dxds\\
&+2E\sum_{i=1}^{n}\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\eta(x,s)\partial_{x_i}Y^{n+1}(x,s)Y^{n+1}(x,s)dxds\\
&+2E\sum_{i=1}^{n}\int_0^t\int_{\mathcal{O}}\sigma^i(x,s)\partial_{x_i}Y^{n+1}(x,s)(g(x,s,Y^n(x,s))-g(x,s,Y^{n-1}(x,s)))dxds\\
&+2E\int_0^t\int_{\mathcal{O}}\eta(x,s)Y^{n+1}(x,s)(g(x,s,Y^n(x,s))-g(x,s,Y^{n-1}(x,s)))dxds. \qquad (3.19)
\end{aligned}
$$

Using Hypotheses 2.1 and 2.2, and Yong's inequality, we have

$$E\|Y^{n+1}(t)\|_0^2+\lambda E\int_0^t\|Y^{n+1}(s)\|_1^2ds\le CE\int_0^t\|Y^n(s)\|_0^2ds+CE\int_0^t\|Y^{n+1}(s)\|_0^2ds. \qquad (3.20)$$

Let

$$P^{n+1}(t)=\int_0^t E\|Y^{n+1}(s)\|_0^2ds.$$

Observe from (3.20) that

$$\frac{d}{dt}P^{n+1}(t)\le CP^n(t)+CP^{n+1}(t),$$

for each t, this yields that

$$P^{n+1}(t)\le C\int_0^t e^{C(t-s)}P^n(s)ds.$$

Hence iterating this inequality gives

$$P^{n+1}(t) \leq (Ce^{CT})^n P^1(T)\frac{t^n}{n!}.$$

This implies $\sum_{n=1}^{\infty} P^{n+1}(T)$ is convergent and as a result of this, the definition of P^{n+1} and (3.20), we conclude that $\{y^n\}$ is a Cauchy sequence in $\mathbb{H}^1$, and so it is convergent. Let y denote the limit of this sequence. This convergence together with Hypotheses 2.1 and 2.2 allows us to let $n \to \infty$ in the weak solution form of (3.18), in other words, y satisfies (3.3). It follows that y is a solution of (1.1). □

4. Adjoint equations (BSPDE with non-homogeneous double boundary conditions)

In this subsection, we will give the existence and uniqueness of the adjoint Eq. (2.4). For simplicity, it is denoted by

$$\begin{cases} dp(x,t) = -\left[A^* p(x,t) + \sum_{k=1}^{d'} B_k^* q_k(x,t) + \bar{f}(x,t)p(x,t) + \bar{L}(x,t) + \sum_{k=1}^{d'} \bar{g}_k(x,t)q_k(x,t)\right] dt & \\ \quad + q(x,t)dw(t) & \text{in } Q, \\ \partial_{\nu_{A^*}} p(x,t) = \bar{h}(x,t)p(x,t) + \bar{l}(x,t) & \text{on } \Sigma, \\ \partial_{\nu_{B_k^*}} q_k(x,t) = 0, \quad k \in \{1,2,\ldots,d'\} & \text{on } \Sigma, \\ p(x,T) = \bar{r}(x,T) & \text{in } \mathcal{O} \end{cases} \tag{4.1}$$

where

$$\bar{f}(x,t) = \frac{\partial f}{\partial y}(x,t,\bar{y}(x,t)), \qquad \bar{L}(x,t) = \frac{\partial L}{\partial y}(x,t,\bar{y}(x,t)),$$

$$\bar{g}_k(x,t) = \frac{\partial g_k}{\partial y}(x,t,\bar{y}(x,t)), \qquad \bar{h}(x,t) = \frac{\partial h}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t)),$$

and

$$\bar{l}(x,t) = \frac{\partial l}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t)), \qquad \bar{r}(x,T) = \frac{\partial r}{\partial y}(x,\bar{y}(x,T)).$$

First we give a definition of backward stochastic partial differential equations with non-homogeneous boundary conditions. It is important for the research of our problems. As we know, this is new in this area.

Definition 4.1. $(p,q) \in \mathbb{H}^1 \times [\mathbb{H}^0]^{d'}$ is called a weak solution pair of (4.1) if it satisfies

$$\begin{aligned} \int_{\mathcal{O}} p(x,t)\phi(x)dx = {} & \int_{\mathcal{O}} \bar{r}(x,T)\phi(x)dx - \sum_{i,j=1}^{n}\int_t^T\int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_i}p(x,s)\partial_{x_j}\phi(x)dxds \\ & + \sum_{i=1}^{n}\int_t^T\int_{\mathcal{O}} [d^i(x,s)p(x,s)\partial_{x_i}\phi(x) - b^i(x,s)\partial_{x_i}p(x,s)\phi(x)]dxds \\ & + \int_t^T\int_{\mathcal{O}} c(x,s)p(x,s)\phi(x)dxds + \sum_{k=1}^{d'}\int_t^T\int_{\mathcal{O}} \eta^k(x,s)p(x,s)\phi(x)dxds \\ & + \sum_{k=1}^{d'}\sum_{i=1}^{n}\int_t^T\int_{\mathcal{O}} \sigma^{ik}(x,s)q_k(x,s)\partial_{x_i}\phi(x)dxds \\ & + \int_t^T\int_{\mathcal{O}} \left[\bar{f}(x,s)p(x,s) + \bar{L}(x,s) + \sum_{k=1}^{d'} \bar{g}_k(x,s)q_k(x,s)\right]\phi(x)dxds \\ & + \int_t^T\int_{\Gamma} [\bar{h}(x,s)p(x,s) + \bar{l}(x,s)]\phi(x)d\rho(x)ds - \sum_{k=1}^{d'}\int_t^T\int_{\mathcal{O}} q_k(x,s)\phi(x)dxdw_k(s) \end{aligned} \tag{4.2}$$

for every $\phi \in H^1(\mathcal{O})$.

In the following we also assume that $d' = 1$. Now, we give a result of existence and uniqueness of the backward differential equation in the finite dimensional case, which was originally obtained by Bismut [19], see also [17,7,15];

Lemma 4.1. *Let m be a fixed positive integer. Let $A_m, M_m \in L^\infty_{\mathcal{F}}(\Omega \times [0,T]; \mathbb{R}^{m\times m})$, $F_m \in L^2_{\mathcal{F}}(\Omega \times [0,T]; \mathbb{R}^m)$, and $r_m \in L^2(\Omega, \mathcal{F}_T, \mathbb{R}^m)$. Then there exists uniquely a pair $(p^m, q^m) \in L^2_{\mathcal{F}}(\Omega \times [0,T]; \mathbb{R}^m) \times L^2_{\mathcal{F}}(\Omega \times [0,T]; \mathbb{R}^m)$ satisfying the following* *backward SDE:*

$$\begin{cases} dp^m(t) = -[A_m^T p^m(t) + M_m^T(t)q^m(t) + F_m(t)]dt + q^m(t)dw(t), \quad t \in [0, T), \\ p^m(x, T) = r_m. \end{cases} \tag{4.3}$$

Theorem 4.1. *Under Hypotheses 2.1, 2.4 and 2.4, there exists a unique solution pair $(p, q) \in \mathbb{H}^1 \times \mathbb{H}^0$ satisfying BSPDE (4.1).*

Proof. (1) Uniqueness. Assume that $(p, q) \in \mathbb{H}^1 \times \mathbb{H}^0$ satisfies

$$\begin{cases} dp(x, t) = -[A^* p(x, t) + B^* q(x, t) + \bar{f}(x, t)p(x, t) + \bar{g}(x, t)q(x, t)]dt + q(x, t)dw(t) & \text{in } Q, \\ \partial_{\nu_{A^*}} p(x, t) = \bar{h}(x, t)p(x, t) & \text{on } \Sigma, \\ p(x, T) = 0 & \text{in } \mathcal{O}. \end{cases} \tag{4.4}$$

Consider the following equation, which admits a unique solution $\xi \in \mathbb{H}^1$ by Lemma 3.1:

$$\begin{cases} d\xi(x, t) = [A\xi(x, t) + \bar{f}(x, t)\xi(x, t) + p(x, t)]dt + [B\xi(x, t) + \bar{g}(x, t)\xi(x, t) + q(x, t)]dw(t) & \text{in } Q, \\ \partial_{\nu_A}\xi(x, t) = \bar{h}(x, t)\xi(x, t) & \text{on } \Sigma, \\ \xi(x, 0) = 0 & \text{in } \mathcal{O}. \end{cases} \tag{4.5}$$

Applying Ito's formula to $\langle p(t), \xi(t)\rangle_{H^0}$, we obtain

$$E\int_0^T \|p(t)\|_0^2 + \|q(t)\|_0^2 dt = 0.$$

This implies

$$|||p|||_1^2 = 0.$$

So we have the uniqueness.

(2) Existence. Also let $\{e_k\}_{k=1}^\infty$ be a basis of $H^1(\mathcal{O})$, which is orthonormal as a basis of $H^0(\mathcal{O})$. Fix a positive integer m. Let

$$p^m(t) = \sum_{k=1}^m p_k^m(t)e_k \quad \text{and} \quad q^m(t) = \sum_{k=1}^m q_k^m(t)e_k,$$

such that

$$\begin{aligned} p_k^m(t) = {} & \bar{r}_k^m(T) - \sum_{i,j=1}^n \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} a^{ij}(x, s)\partial_{x_j} e_\tau(x)\partial_{x_i} e_k(x)dx p_\tau^m(s)ds \\ & + \sum_{i=1}^n \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} [d^i(x, s)e_\tau(x)\partial_{x_i} e_k(x) - b^j(x, s)\partial_{x_i} e_\tau(x)e_k(x)]dx p_\tau^m(s)ds \\ & + \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} c(x, s)e_\tau(x)e_k(x)dx p_\tau^m(s)ds + \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} \eta(x, s)e_\tau(x)e_k(x)dx q_\tau^m(s)ds \\ & + \sum_{i=1}^n \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} \sigma^i(x, s)e_\tau(x)\partial_{x_i} e_k(x)dx q_\tau^m(s)ds \\ & + \int_t^T \int_{\mathcal{O}} \bar{L}(x, s)e_k(x)dxds + \sum_{\tau=1}^m \int_t^T \int_{\mathcal{O}} \bar{g}(x, s)e_\tau(x)e_k(x)dx q_\tau^m(s)ds \\ & + \sum_{\tau=1}^m \int_t^T \int_{\Gamma} \bar{h}(x, s)e_\tau(x)e_k(x)d\rho(x)p_\tau^m(s)ds \\ & + \int_t^T \int_{\Gamma} \bar{l}(x, s)e_k(x)d\rho(x)ds - \int_t^T q_k^m(s)dw(s), \quad k = 1, 2, \ldots, m, \end{aligned} \tag{4.6}$$

where $\sum_{k=1}^m \bar{r}_k^m(T)e_k(\cdot) = \bar{r}^m(\cdot, T) \to \bar{r}(\cdot, T)$ in $H^0(\mathcal{O})$ as $m \to \infty$. By Lemma 4.1, there exists a unique pair

$$(p_1^m, \ldots, p_m^m)^T \in L^2_{\mathcal{F}}(\Omega \times [0, T]; \mathbb{R}^m)$$

and

$$(q_1^m, \ldots, q_m^m)^T \in L^2_{\mathcal{F}}(\Omega \times [0, T]; \mathbb{R}^m)$$

satisfying (4.6). Applying Ito's formula to (4.6) and adding in k from 1 to m, we get

$$\begin{aligned}
E\int_{\mathcal{O}} |p^m(x,t)|^2 dx = {} & E\int_{\mathcal{O}} |\bar{r}^m(x,T)|^2 dx - 2E\sum_{i,j=1}^{n}\int_t^T\int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_j}p^m(x,s)\partial_{x_i}p^m(x,s)dxds \\
& + 2E\sum_{i=1}^{n}\int_t^T\int_{\mathcal{O}} [d^i(x,s)-b^i(x,s)]\partial_{x_i}p^m(x,s)p^m(x,s)dxds \\
& + 2E\int_t^T\int_{\mathcal{O}} (c(x,s)+\eta(x,s))|p^m(x,s)|^2 dxds \\
& + 2E\sum_{i=1}^{n}\int_t^T\int_{\mathcal{O}} \sigma^i(x,s)q^m(x,s)\partial_{x_i}p^m(x,s)dxds \\
& + 2E\int_t^T\int_{\mathcal{O}} \bar{L}(x,s)p^m(x,s)dxds + 2E\int_t^T\int_{\mathcal{O}} \bar{g}(x,s)p^m(x,s)q^m(x,s)dxds \\
& + 2E\int_t^T\int_{\Gamma} \bar{h}(x,s)|p^m(x,s)|^2 d\rho(x)ds + 2E\int_t^T\int_{\Gamma} \bar{l}(x,s)p^m(x,s)d\rho(x)ds \\
& - E\int_t^T\int_{\mathcal{O}} |q^m(x,s)|^2 dxds. \qquad (4.7)
\end{aligned}$$

Hence by Yong's inequality we have that

$$E\|p^m(t)\|_0^2 + \lambda E\int_t^T \|p^m(s)\|_1^2 ds \le C + CE\int_t^T \|p^m(s)\|_0^2 ds \qquad (4.8)$$

where C depends only on K, M, λ and the constant in the imbedding theorem. So Gronwall's inequality yields

$$\sup_{t\in[0,T]} E\|p^m(t)\|_0^2 + \|\!|p^m|\!\|_1^2 \le C. \qquad (4.9)$$

Now let $\xi^m = (\xi_1^m, \xi_2^m, \dots, \xi_m^m)^T \in L^2_{\mathscr{F}}(\Omega\times[0,T];\mathbb{R}^m)$ be the solution of the following equation:

$$\begin{aligned}
d\xi_k^m(t) = {} & -\sum_{i,j=1}^{n}\sum_{\tau=1}^{m}\int_{\mathcal{O}} a^{ij}(x,t)\partial_{x_i}e_\tau(x)\partial_{x_j}e_k(x)dx\xi_\tau^m(t)dt \\
& + \sum_{j=1}^{n}\sum_{\tau=1}^{m}\int_{\mathcal{O}} [d^j(x,t)\partial_{x_j}e_\tau(x)e_k(x) - b^j(x,t)e_\tau(x)\partial_{x_j}e_k(x)]dx\xi_\tau^m(t)dt \\
& + \sum_{\tau=1}^{m}\int_{\mathcal{O}} c(x,t)e_\tau(x)e_k(x)dx\xi_\tau^m(t)dt + \sum_{\tau=1}^{m}\int_{\Gamma} \bar{h}(x,t)e_\tau(x)e_k(x)d\rho(x)\xi_\tau^m(t)dt \\
& + \int_{\Gamma} \bar{l}(x,t)e_k(x)d\rho(x)dt + \sum_{i=1}^{n}\sum_{\tau=1}^{m}\int_{\mathcal{O}} \sigma^i(x,t)\partial_{x_i}e_\tau(x)e_k(x)dx\xi_\tau^m(t)dw(t) \\
& + \sum_{\tau=1}^{m}\int_{\mathcal{O}} \eta(x,t)e_\tau(x)e_k(x)dx\xi_\tau^m(t)dw(t) \\
& + \sum_{\tau=1}^{m}\int_{\mathcal{O}} \bar{g}(x,t)e_\tau(x)e_k(x)dx\xi_\tau^m(t)dw(t) + q_k^m(t)dw(t), \\
\xi_k^m(0) = {} & 0, \quad k = 1,2,\dots,m. \qquad (4.10)
\end{aligned}$$

By a calculation similar to the above, we get

$$\sup_{t\in[0,T]} E\|\xi^m(t)\|_0^2 + E\int_0^T \|\xi^m(t)\|_1^2 dt \le CE\int_0^T \|q^m(t)\|_0^2 dt. \qquad (4.11)$$

On the other hand, Ito's formula gives

$$d\sum_{k=1}^{m} p_k^m(t)\xi_k^m(t) = -\int_{\mathcal{O}} \bar{L}(x,t)\xi^m(x,t)dxdt + \|q^m(t)\|_0^2 dt + \{\cdots\}dw(t). \qquad (4.12)$$

Integrating from 0 to T and taking the expectation, we get

$$E\int_0^T \|q^m(t)\|_0^2 dt = E\int_Q \bar{L}(x,t)\xi^m(x,t)dxdt + E\int_{\mathcal{O}} p^m(x,T)\xi^m(x,T)dx$$
$$\le C\left(E\int_0^T \|\xi(t)\|_0^2 dt\right)^{1/2} + (E\|\bar{r}^m(T)\|_0^2)^{1/2}(E\|\xi^m(T)\|_0^2)^{1/2}$$
$$\le C\left(E\int_0^T \|q^m(t)\|_0^2 dt\right)^{1/2} \tag{4.13}$$

where the constant C depends on M. Consequently

$$E\int_0^T \|q^m(t)\|_0^2 dt \le C. \tag{4.14}$$

By (4.9) and (4.14), there exists a subsequence (also denoted by itself) and a pair $(p,q) \in \mathbb{H}^1 \times \mathbb{H}^0$ such that

$$p^m \to p \quad \text{weakly in } \mathbb{H}^1 \tag{4.15}$$

and

$$q^m \to q \quad \text{weakly in } \mathbb{H}^0 \text{ as } m \to \infty. \tag{4.16}$$

Next we show that (p,q) satisfies (4.1). To this end, similar to Theorem 3.1, let ψ be an absolutely continuous function from $[0,T]$ to $\mathbb{R}$ with $\psi' = d\psi/dt \in L^2(0,T)$, $\psi(0) = 0$ and

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

Then the proof is completed. □

5. The proof of necessary conditions

In this section, we present the proof of our necessary conditions. The proof is divided into two steps. Let $\bar{u} \in \mathcal{U}_{ad}$ be an optimal control and let $\bar{y}(\cdot,\cdot)$ be the corresponding optimal state. And $u \in \mathcal{U}_{ad}$ be fixed. Because of the admissible controls set $\mathcal{U}_{ad}$ is a convex set, then $u_\theta = \bar{u} + \theta(u - \bar{u})(\theta \in (0,1))$ also belongs to $\mathcal{U}_{ad}$ and let $y_\theta(\cdot,\cdot)$ be the corresponding state. In this section, we using the convex perturbation to obtain the main results.

First we give some lemmas for our needs.

5.1. Some lemmas

Lemma 5.1. *Under Hypotheses 2.1 and 2.3, there exists a constant C such that*

$$\sup_{t\in[0,T]} E\|y_\theta(t) - \bar{y}(t)\|_0^2 + |||y_\theta - \bar{y}|||_1^2 \le C\theta^2. \tag{5.1}$$

Proof. Let $\xi = y_\theta - \bar{y}$, then by the state equation we have

$$\begin{cases} d\xi(x,t) = [A\xi(x,t) + f(x,t,y_\theta(x,t)) - f(x,t,\bar{y}(x,t))]dt \\ \quad + [B\xi(x,t) + g(x,t,y_\theta(x,t)) - g(x,t,\bar{y}(x,t))]dw(t) & \text{in } Q, \\ \partial_{\nu_A}\xi(x,t) = h(x,t,y_\theta(x,t),u_\theta(x,t)) - h(x,t,\bar{y}(x,t),\bar{u}(x,t)) & \text{on } \Sigma, \\ \xi(x,0) = 0 & \text{in } \mathcal{O}. \end{cases} \tag{5.2}$$

By Lemma 3.1, we know that (5.1) has a solution in $\mathbb{H}^1$, and

$$E\int_{\mathcal{O}} |\xi(x,t)|^2 dx = -2E\sum_{i,j=1}^n \int_0^t \int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_i}\xi(x,s)\partial_{x_j}\xi(x,s)dxds$$
$$+ 2E\sum_{j=1}^n \int_0^t \int_{\mathcal{O}} [d^j(x,s) - b^j(x,s)]\xi(x,s)\partial_{x_j}\xi(x,s)dxds$$
$$+ 2E\int_0^t \int_{\mathcal{O}} c(x,s)|\xi(x,s)|^2 dxds + 2E\int_0^t \int_{\mathcal{O}} [f(x,s,y_\theta(x,s))$$

$$
\begin{aligned}
&- f(x,s,\bar{y}(x,s))]\xi(x,s)dxds + 2E\int_0^t \int_\Gamma C[h(x,s,y_\theta(x,s),u_\theta(x,s)) \\
&- h(x,s,\bar{y}(x,s),\bar{u}(x,s))]\xi(x,s)d\rho(x)ds + E\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} [\eta^k(x,s)\xi(x,s)]^2 dxds \\
&+ E\sum_{i,j=1}^{n}\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} \sigma^{ik}(x,s)\sigma^{jk}(x,s)\partial_{x_i}\xi(x,s)\partial_{x_j}\xi(x,s)dxds \\
&+ E\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} [g_k(x,s,y_\theta(x,s)) - g_k(x,s,\bar{y}(x,s))]^2 dxds \\
&+ 2E\sum_{i=1}^{n}\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} \sigma^{ik}(x,s)\eta^k(x,s)\partial_{x_i}\xi(x,s)\xi(x,s)dxds \\
&+ 2E\sum_{i=1}^{n}\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} \sigma^{ik}(x,s)\partial_{x_i}\xi(x,s)(g_k(x,s,y_\theta(x,s)) - g_k(x,s,\bar{y}(x,s)))dxds \\
&+ 2E\sum_{k=1}^{d'}\int_0^t \int_\mathcal{O} \eta^k(x,s)\xi(x,s)(g_k(x,s,y_\theta(x,s)) - g_k(x,s,\bar{y}(x,s)))dxds. \qquad (5.3)
\end{aligned}
$$

Using the Hypotheses 2.1 and 2.3, and Yong's inequality, we obtain that

$$
\begin{aligned}
E\|\xi(t)\|_0^2 \le\ & (-2\lambda + C\varepsilon)E\int_0^t \|\xi(s)\|_1^2 ds + C(\varepsilon)E\int_0^t \|\xi(s)\|_0^2 ds \\
&+ 2E\int_0^t \int_\Gamma [h(x,s,y_\theta(x,s),u_\theta(x,s)) - h(x,s,\bar{y}(x,s),\bar{u}(x,s))]\xi(x,s)d\rho(x)ds \\
=\ & (-2\lambda + C\varepsilon)E\int_0^t \|\xi(s)\|_1^2 ds + C(\varepsilon)E\int_0^t \|\xi(s)\|_0^2 ds \\
&+ 2E\int_0^t \int_\Gamma [h(x,s,y_\theta(x,s),u_\theta(x,s)) - h(x,s,\bar{y}(x,s),u_\theta(x,s))]\xi(x,s)d\rho(x)ds \\
&+ 2E\int_0^t \int_\Gamma [h(x,s,\bar{y}(x,s),u_\theta(x,s)) - h(x,s,\bar{y}(x,s),\bar{u}(x,s))]\xi(x,s)d\rho(x)ds \\
\le\ & (-2\lambda + C\varepsilon)E\int_0^t \|\xi(s)\|_1^2 ds + C(\varepsilon)\int_0^t \|\xi(s)\|_0^2 ds \\
&+ \frac{1}{\varepsilon}E\int_0^t \int_\Gamma \Bigg[\int_0^1 \frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s) + \tau(u_\theta(x,s) \\
&- \bar{u}(x,s)))d\tau(u_\theta(x,s) - \bar{u}(x,s))\Bigg]^2 d\rho(x)ds \\
\le\ & (-2\lambda + C\varepsilon)E\int_0^t \|\xi(s)\|_1^2 ds + C(\varepsilon)\int_0^t \|\xi(s)\|_0^2 ds + C\frac{1}{\varepsilon}\theta^2. \qquad (5.4)
\end{aligned}
$$

Letting $\varepsilon = \lambda/C$ in the above inequality, we get

$$
E\|\xi(t)\|_0^2 + \lambda E\int_0^t \|\xi(s)\|_1^2 ds \le C\theta^2 + CE\int_0^t \|\xi(s)\|_0^2 ds.
$$

Hence by Gronwall's inequality we get that

$$
\sup_{t\in[0,T]} E\|\xi(t)\|_0^2 + E\int_0^T \|\xi(t)\|_1^2 ds \le C\theta^2,
$$

where constant C depends on λ, M, K. The proof is completed. □

Lemma 5.2. *Suppose Hypotheses 2.1 and 2.3 are satisfied. Let $z(\cdot,\cdot)$ be the solution of the following equation:*

$$\begin{cases} dz(x,t) = \left[Az(x,t) + \dfrac{\partial f}{\partial y}(x,t,\bar{y}(x,t))z(x,t)\right]dt \\ \quad + \left[Bz(x,t) + \dfrac{\partial g}{\partial y}(x,t,\bar{y}(x,t))z(x,t)\right]dw(t) & \text{in } Q, \\ \partial_{\nu_A} z(x,t) = \dfrac{\partial h}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t))z(x,t) \\ \quad + \dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))(u(x,t)-\bar{u}(x,t)) & \text{on } \Sigma, \\ z(x,0) = 0 & \text{in } \mathcal{O}. \end{cases} \tag{5.5}$$

Then we have

$$\lim_{\theta\to 0}\left\{\sup_{t\in[0,T]} E\left\|\frac{1}{\theta}(y_\theta(t)-\bar{y}(t)) - z(t)\right\|_0^2 + \left\|\left|\frac{1}{\theta}(y_\theta-\bar{y}) - z\right|\right\|_1^2\right\} = 0. \tag{5.6}$$

Proof. We denote

$$\phi_\theta(x,t) = \frac{1}{\theta}(y_\theta(x,t)-\bar{y}(x,t)) - z(x,t).$$

Then by the state equation and (5.5), we have

$$\begin{cases} d\phi_\theta(x,t) = \left[A\phi_\theta(x,t) + \dfrac{1}{\theta}(f(x,t,y_\theta(x,t)) - f(x,t,\bar{y}(x,t))) - \dfrac{\partial f}{\partial y}(x,t,\bar{y}(x,t))z(x,t)\right]dt \\ \quad + \left[B\phi_\theta(x,t) + \dfrac{1}{\theta}(g(x,t,y_\theta(x,t)) - g(x,t,\bar{y}(x,t))) - \dfrac{\partial g}{\partial y}(x,t,\bar{y}(x,t))z(x,t)\right]dw(t) & \text{in } Q, \\ \partial_{\nu_A}\phi_\theta(x,t) = \dfrac{1}{\theta}[h(x,t,y_\theta(x,t),u_\theta(x,t)) - h(x,t,\bar{y}(x,t),\bar{u}(x,t))] \\ \quad - \dfrac{\partial h}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t))z(x,t) - \dfrac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))(u(x,t)-\bar{u}(x,t)) & \text{on } \Sigma, \\ \phi_\theta(x,0) = 0 & \text{in } \mathcal{O}. \end{cases} \tag{5.7}$$

Hence, similar to Lemma 5.1, we obtain

$$\begin{aligned} E\int_{\mathcal{O}} |\phi_\theta(x,t)|^2 dx = {} & -2E\sum_{i,j=1}^n \int_0^t \int_{\mathcal{O}} a^{ij}(x,s)\partial_{x_i}\phi_\theta(x,s)\partial_{x_j}\phi_\theta(x,s)dxds \\ & + 2E\sum_{j=1}^n \int_0^t \int_{\mathcal{O}} [d^j(x,s) - b^j(x,s)]\phi_\theta(x,s)\partial_{x_j}\phi_\theta(x,s)dxds \\ & + 2E\int_0^t \int_{\mathcal{O}} c(x,s)|\phi_\theta(x,s)|^2 dxds + 2E\int_0^t \int_{\mathcal{O}} \left[\frac{1}{\theta}(f(x,s,y_\theta(x,s))\right. \\ & \left. - f(x,s,\bar{y}(x,s))) - \frac{\partial f}{\partial y}(x,s,\bar{y})z(x,s)\right]\phi_\theta(x,s)dxds \\ & + 2E\int_0^t \int_{\Gamma} \phi_\theta(x,s)\left\{\frac{1}{\theta}[h(x,s,y_\theta(x,s),u_\theta(x,s)) - h(x,s,\bar{y}(x,s),\bar{u}(x,s))]\right. \\ & - \frac{\partial h}{\partial y}(x,s,\bar{y}(x,s),\bar{u}(x,s))z(x,s) - \frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s))(u(x,s) \\ & \left. - \bar{u}(x,s))\right\}d\rho(x)ds + E\sum_{i,j=1}^n \sum_{k=1}^{d'} \int_0^t \int_{\mathcal{O}} \sigma^{ik}(x,s)\sigma^{jk}(x,s)\partial_{x_i}\phi_\theta \end{aligned}$$

$$\times (x,s)\partial_{x_j}\phi_\theta(x,s)dxds + E\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}[\eta^k(x,s)\phi_\theta(x,s)]^2dxds$$
$$+E\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}\Bigg[\frac{1}{\theta}(g_k(x,s,y_\theta(x,s))-g_k(x,s,\bar{y}(x,s)))$$
$$-\frac{\partial g}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\Bigg]^2dxds + 2E\sum_{i=1}^{n}\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}\sigma^{ik}(x,s)\eta^k(x,s)\partial_{x_i}\phi_\theta$$
$$\times (x,s)\phi_\theta(x,s)dxds$$
$$+2E\sum_{i=1}^{n}\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}\sigma^{ik}(x,s)\partial_{x_i}\phi_\theta(x,s)\Bigg[\frac{1}{\theta}g_k(x,s,y_\theta(x,s))-g_k(x,s,\bar{y}(x,s))$$
$$-\frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\Bigg]dxds + 2E\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}\eta^k(x,s)\phi_\theta(x,s)$$
$$\times\Bigg[\frac{1}{\theta}g_k(x,s,y_\theta(x,s))-g_k(x,s,\bar{y}(x,s))-\frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\Bigg]dxds. \tag{5.8}$$

Hence, we have

$$E\|\phi_\theta(t)\|_0^2+\lambda E\int_0^t\|\phi_\theta(s)\|_1^2ds\le CE\int_0^t\|\phi_\theta(s)\|_0^2ds+\varrho(\theta,t), \tag{5.9}$$

where

$$\varrho(\theta,t) = CE\int_0^t\int_{\mathcal{O}}\Bigg[\frac{1}{\theta}(f(x,s,y_\theta(x,s))-f(x,s,\bar{y}(x,s)))-\frac{\partial f}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\Bigg]^2dxds$$
$$+CE\sum_{k=1}^{d'}\int_0^t\int_{\mathcal{O}}\Bigg[\frac{1}{\theta}(g_k(x,s,y_\theta(x,s))-g_k(x,s,\bar{y}(x,s)))-\frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\Bigg]^2dxds$$
$$+CE\int_0^t\int_{\Gamma}\phi_\theta(x,s)\Bigg\{\frac{1}{\theta}[h(x,s,y_\theta(x,s),u_\theta(x,s))-h(x,s,\bar{y}(x,s),\bar{u}(x,s))]$$
$$-\frac{\partial h}{\partial y}(x,s,\bar{y}(x,s),\bar{u}(x,s))z(x,s)-\frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s))(u(x,s)-\bar{u}(x,s))\Bigg\}d\rho(x)ds.$$

We define $\varrho_h(\theta,t)$ by

$$\varrho_h(\theta,t) = CE\int_0^t\int_{\Gamma}\phi_\theta(x,s)\Bigg\{\frac{1}{\theta}[h(x,s,y_\theta(x,s),u_\theta(x,s))-h(x,s,\bar{y}(x,s),\bar{u}(x,s))]$$
$$-\frac{\partial h}{\partial y}(x,s,\bar{y}(x,s),\bar{u}(x,s))z(x,s)-\frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s))(u(x,s)-\bar{u}(x,s))\Bigg\}d\rho(x)ds.$$

By Hypothesis 2.3 and some simple computations we get

$$\varrho_h(\theta,t)\le C\varepsilon E\int_0^t\|\phi_\theta(s)\|_1^2ds+C\frac{1}{\varepsilon}E\int_0^t\int_{\Gamma}\Bigg[\Bigg(\int_0^1\frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s)$$
$$+\tau(u_\theta(x,s)-\bar{u}(x,s)))d\tau-\frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s))\Bigg)(u(x,s)-\bar{u}(x,s))\Bigg]^2d\rho(x)ds$$
$$+C\frac{1}{\varepsilon}E\int_0^t\int_{\Gamma}\Bigg[\Bigg(\frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s))-\frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s))\Bigg)$$

$$\times (u(x,s)-\bar{u}(x,s))\Bigg]^2 d\rho(x)ds + C\frac{1}{\varepsilon}E\int_0^t\int_\Gamma\Bigg[\Bigg(\int_0^1 \frac{\partial h}{\partial y}(x,s,\bar{y}(x,s) + \tau(y_\theta(x,s)-\bar{y}(x,s)),\bar{u}(x,s))d\tau - \frac{\partial h}{\partial y}(x,s,\bar{y}(x,s),\bar{u}(x,s))\Bigg)z(x,s)\Bigg]^2 d\rho(x)ds.$$

Let $\varepsilon = 2C/\lambda$, we get

$$\varrho_h(\theta,t) \le \frac{\lambda}{2}E\int_0^t \|\phi_\theta(s)\|_1^2 ds + \varrho_h^1(\theta,t), \tag{5.10}$$

where

$$\begin{aligned}\varrho_h^1(\theta,t) = {} & CE\int_0^t\int_\Gamma\Bigg[\Bigg(\int_0^1 \frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s)+\tau(u_\theta(x,s)-\bar{u}(x,s)))d\tau \\ & - \frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s))\Bigg)(u(x,s)-\bar{u}(x,s))\Bigg]^2 d\rho(x)ds \\ & + CE\int_0^t\int_\Gamma\Bigg[\Bigg(\frac{\partial h}{\partial u}(x,s,y_\theta(x,s),\bar{u}(x,s)) - \frac{\partial h}{\partial u}(x,s,\bar{y}(x,s),\bar{u}(x,s))\Bigg) \\ & \times (u(x,s)-\bar{u}(x,s))\Bigg]^2 d\rho(x)ds + CE\int_0^t\int_\Gamma\Bigg[\Bigg(\int_0^1 \frac{\partial h}{\partial y}(x,s,\bar{y}(x,s) \\ & + \tau(y_\theta(x,s)-\bar{y}(x,s)),\bar{u}(x,s))d\tau - \frac{\partial h}{\partial y}(x,s,\bar{y}(x,s),\bar{u}(x,s))\Bigg)z(x,s)\Bigg]^2 d\rho(x)ds.\end{aligned}$$

Next, we let

$$\varrho_f(\theta,t) = CE\int_0^t\int_\Theta\left[\frac{1}{\theta}(f(x,s,y_\theta(x,s)) - f(x,s,\bar{y}(x,s))) - \frac{\partial f}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\right]^2 dxds.$$

Then, we have

$$\varrho_f(\theta,t) \le CE\int_0^t \|\phi_\theta(s)\|_0^2 ds + \varrho_f^1(\theta,t), \tag{5.11}$$

where

$$\begin{aligned}\varrho_f^1(\theta,t) = {} & CE\int_0^t\int_\Theta\Bigg[\Bigg(\int_0^1 \frac{\partial f}{\partial y}(x,s,\bar{y}(x,s)+\tau(y_\theta(x,s)-\bar{y}(x,s))) \\ & - \frac{\partial f}{\partial y}(x,s,\bar{y}(x,s,\bar{y}(x,s)))\Bigg)z(x,s)\Bigg]^2 dxds.\end{aligned}$$

Similarly we also have

$$\begin{aligned}\varrho_g(\theta,t) = {} & CE\sum_{k=1}^{d'}\int_0^t\int_\Theta\left[\frac{1}{\theta}(g_k(x,s,y_\theta(x,s)) - g_k(x,s,\bar{y}(x,s))) - \frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s))z(x,s)\right]^2 dxds \\ \le {} & CE\int_0^t \|\phi_\theta(s)\|_0^2 ds + \varrho_g^1(\theta,t),\end{aligned} \tag{5.12}$$

where

$$\begin{aligned}\varrho_g^1(\theta,t) = {} & CE\sum_{k=1}^{d'}\int_0^t\int_\Theta\Bigg[\Bigg(\int_0^1 \frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s)+\tau(y_\theta(x,s)-\bar{y}(x,s))) \\ & - \frac{\partial g_k}{\partial y}(x,s,\bar{y}(x,s,\bar{y}(x,s)))\Bigg)z(x,s)\Bigg]^2 dxds.\end{aligned}$$

Hence by (5.9)–(5.12) we get that

$$E\|\phi_\theta(t)\|_0^2 + \frac{\lambda}{2}E\int_0^t \|\phi_\theta(s)\|_1^2 ds \le CE\int_0^t \|\phi_\theta(s)\|_0^2 ds + \varrho_h^1(\theta,t) + \varrho_f^1(\theta,t) + \varrho_g^1(\theta,t). \tag{5.13}$$

So by Gronwall's inequality, we know that

$$\sup_{t\in[0,T]} E\|\phi_\theta(t)\|_0^2 + E\int_0^T \|\phi_\theta(t)\|_1^2 dt \le C[\varrho_h^1(\theta,T) + \varrho_f^1(\theta,T) + \varrho_g^1(\theta,T)]. \tag{5.14}$$

Clearly, $\lim_{\theta\to 0}\varrho_h^1(\theta,T)=0$, $\lim_{\theta\to 0}\varrho_f^1(\theta,T)=0$ and $\lim_{\theta\to 0}\varrho_g^1(\theta,T)=0$, so by (5.14), we complete the proof. □

Lemma 5.3. *Suppose Hypotheses 2.1, 2.3 and 2.4 are satisfied. Let $z(\cdot,\cdot)$ be the solution of Eq. (5.5). Then we have the following inequality:*

$$\begin{aligned} 0 \le\ & E\int_Q \frac{\partial L}{\partial y}(x,t,\bar y(x,t))z(x,t)dxdt + E\int_\Sigma \frac{\partial l}{\partial u}(x,t,\bar y(x,t),\bar u(x,t))(u(x,t)-\bar u(x,t))d\rho(x)dt \\ & + E\int_\Sigma \frac{\partial l}{\partial y}(x,t,\bar y(x,t),\bar u(x,t))z(x,t)d\rho(x)dt + E\int_\Theta \frac{\partial r}{\partial y}(x,\bar y(x,T))z(x,T)dx. \end{aligned} \tag{5.15}$$

Proof. Also let

$$\phi_\theta(x,t) = \frac{1}{\theta}(y_\theta(x,t)-\bar y(x,t)) - z(x,t).$$

Because $(\bar y,\bar u)$ is an optimal pair, then we have

$$\begin{aligned} 0 \le\ & \frac{1}{\theta}[J(u_\theta)-J(\bar u)] \\ =\ & \frac{1}{\theta}E\int_Q [L(x,t,y_\theta(x,t)) - L(x,t,\bar y(x,t))]dxdt \\ & + \frac{1}{\theta}E\int_\Sigma [l(x,t,y_\theta(x,t),u_\theta(x,t)) - l(x,t,\bar y(x,t),\bar u(x,t))]d\rho(x)dt \\ & + \frac{1}{\theta}E\int_\Theta [r(x,y_\theta(x,T)) - r(x,\bar y(x,T))]dx \\ =\ & E\int_Q\int_0^1 \frac{\partial L}{\partial y}(x,t,\bar y(x,t)+\tau(y_\theta(x,t)-\bar y(x,t)))d\tau\,\phi_\theta(x,t)dxdt \\ & + E\int_Q\int_0^1 \frac{\partial L}{\partial y}(x,t,\bar y(x,t)+\tau(y_\theta(x,t)-\bar y(x,t)))d\tau\, z(x,t)dxdt \\ & + E\int_\Sigma\int_0^1 \frac{\partial l}{\partial u}(x,t,y_\theta(x,t),\bar u(x,t)+\tau(u_\theta(x,t)-\bar u(x,t)))d\tau(u(x,t)-\bar u(x,t))d\rho(x)dt \\ & + E\int_\Sigma\int_0^1 \frac{\partial l}{\partial y}(x,t,\bar y(x,t)+\tau(y_\theta(x,t)-\bar y(x,t)),\bar u(x,t))d\tau\,\phi_\theta(x,t)d\rho(x)dt \\ & + E\int_\Sigma\int_0^1 \frac{\partial l}{\partial y}(x,t,\bar y(x,t)+\tau(y_\theta(x,t)-\bar y(x,t)),\bar u(x,t))d\tau\, z(x,t)d\rho(x)dt \\ & + E\int_\Theta\int_0^1 \frac{\partial r}{\partial y}(x,\bar y(x,T)+\tau(y_\theta(x,T)-\bar y(x,T)))d\tau\,\phi_\theta(x,T)dx \\ & + E\int_\Theta\int_0^1 \frac{\partial r}{\partial y}(x,\bar y(x,T)+\tau(y_\theta(x,T)-\bar y(x,T)))d\tau\, z(x,T)dx. \end{aligned} \tag{5.16}$$

By Hypothesis 2.4, Lemmas 5.1 and 5.2 we can deduce (5.15). □

5.2. Duality analysis

In this subsection, we complete the proof of the Theorem 2.5.

Let $z(\cdot,\cdot) \in \mathbb{H}^1$ be the solution of Eq. (5.5) and $(\bar{p}(\cdot,\cdot), \bar{q}(\cdot,\cdot)) \in \mathbb{H}^1 \times [\mathbb{H}^0]^{d'}$ be the solution of the adjoint equation (2.4). Then using Ito's formula for $\langle z(t), \bar{p}(t)\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the inner product in $\mathbb{H}^1$, we get

$$\begin{aligned} E\int_{\mathcal{O}} z(x,T)\bar{p}(x,T)dx = {} & E\int_{\Sigma} \frac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))[u(x,t)-\bar{u}(x,t)]\bar{p}(x,t)d\rho(x)dt \\ & - E\int_{Q} \frac{\partial L}{\partial y}(x,t,\bar{y}(x,t))z(x,t)dxdt \\ & - E\int_{\Sigma} \frac{\partial l}{\partial y}(x,t,\bar{y}(x,t),\bar{u}(x,t))z(x,t)d\rho(x)dt. \end{aligned} \tag{5.17}$$

By $\bar{p}(x,T) = \partial r/\partial y(x,\bar{y}(x,T))$ and (5.15) we obtain

$$\begin{aligned} & E\int_{\Sigma} \left[\frac{\partial l}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t)) + \frac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))\bar{p}(x,t)\right] u(x,t)d\rho(x)dt \\ & \quad \geq E\int_{\Sigma} \left[\frac{\partial l}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t)) + \frac{\partial h}{\partial u}(x,t,\bar{y}(x,t),\bar{u}(x,t))\bar{p}(x,t)\right] \bar{u}(x,t)d\rho(x)dt. \end{aligned} \tag{5.18}$$

Then we get the results of Theorem 2.5.

References

[1] E. Casas, Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optim. 35 (1997) 1297–1327.
[2] E. Casas, J.P. Raymond, H. Zidani, Pontryagin's principle for local solutions of control problems with mixed control-state constraints, SIAM J. Control Optim. 39 (2000) 1182–1203.
[3] H.J. Kushner, Necessary conditions for continuous parameter stochstic optimization problems, SIAM J. Control Optim. 10 (1973) 550–565.
[4] U.G. Haussmann, General necessary conditions for optimal control of stochastic systems, Math. Program. Stud. 6 (1976) 30–48.
[5] U.G. Haussmann, A Stochstic Maximum Principle for Optimal Control of Diffusions, in: Pitman Research Notes in Math Series, vol. 151, John Wiley & Sons, New York, 1986.
[6] V.I. Arkin, M.T. Saksonov, Necessary optimality conditions for stochastic differential equations, Sov. Math. Dokl. 20 (1979) 1–5.
[7] A. Bensoussan, Lecture on stochastic control Part I, Lecture Notes in Math. 972 (1983) 1–39.
[8] J.M. Bismut, An introductory approach to duality in optimal stochastic control, SIAM Rev. 20 (1978) 61–78.
[9] A. Cadenillas, I. Karatzas, The stochastic maximum principle for linear convex optimal control with random coefficients, SIAM J. Control Optim. 33 (1995) 590–624.
[10] R.J. Elliott, The optimal control of diffusions, Appl. Math. Optim. 22 (1990) 229–240.
[11] R.J. Elliott, M. Kohlmann, The second order minimum principle and adjoint process, Stoch. Stoch. Rep. 46 (1994) 15–39.
[12] S. Peng, A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim. 28 (1990) 966–979.
[13] J. Yong, X. Zhou, Stochastic Control, Hamilton Systems and HJB Equations, in: Appl. Math., vol. 43, Springer-Verlag, New York, 1999.
[14] S. Bahlali, Necessary and sufficient optimality conditions for relaxed and strict control problems, SIAM J. Control Optim. 47 (2008) 2078–2095.
[15] A. Bensoussan, Stochastic maximum principle for distributed parameter system, J. Franklin Inst. 315 (1983) 387–406.
[16] Y. Hu, S. Peng, Maximum principle for semilinear stochastic evolution controls systems, Stochastics 33 (1990) 159–180.
[17] X.Y. Zhou, On the necessary conditions of optimal controls for stochastic partial differential equations, SIAM J. Control Optim. 31 (1993) 1462–1478.
[18] Y. Hu, S. Peng, Adapted solution of a backward semilinear stochastic evolution equation, Stoch. Anal. Appl. 9 (1991) 445–459.
[19] J.M. Bismut, Analyse convexte et probabilités, thèse, faculté des sciences de paris, 1973.