Probability, Uncertainty and Quantitative Risk
Vol. 10, No. 1, 2025, 67–102

https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2025005

Optimal control of a class of fully coupled forward-backward stochastic partial differential equations

Suya Zhang[1,2], Maozhong Xu[2], Qingxin Meng[3,*]

[1]*School of Mathematical Sciences, South China Normal University, Guangzhou* 510631, *China*
[2]*School of Mathematical Sciences, Zhejiang Normal University, Jinhua* 321004, *China*
[3]*School of Sciences, Huzhou University, Huzhou* 313000, *China*

Email: zsy530188498@163.com, xumaozhong1130@163.com, mqx@zjhu.edu.cn

Abstract This paper investigates the optimal control problem for a class of fully coupled forward-backward stochastic partial differential equations (FBSPDEs). Based on the existence of a unique solution to such equations, we formulated the associated optimal control problem within a convex control domain. By employing the convex variational method, we derive the associated stochastic maximum principle (SMP) for the optimal control problem intrinsic to this system. Finally, to demonstrate the applicability of our theoretical results, we apply SMP to a class of linear quadratic problems and obtain explicit expressions for the unique optimal control.

Keywords Forward-backward stochastic partial differential equation, Monotonicity condition, Stochastic maximum principle, Convex domain, Linear quadratic problem

2020 Mathematics Subject Classification 93E20, 49N10, 49K27

1. Introduction

A forward stochastic differential equation (SDE) coupled with a backward stochastic differential equation (BSDE) is called a forward backward stochastic differential equation (FBSDEs). The FBSDEs mainly originates from the study of stochastic control problems. When the stochastic maximum principle (SMP) is applied to address the stochastic optimal control problem for forward or backward systems, the system equations and their dual equations form a fully coupled FBSDEs, which is often referred to as a stochastic Hamiltonian system. In financial mathematics, fully coupled FBSDEs also arise in modeling scenarios such as those involving large investors. The current research on FBSDEs focuses on the following three aspects. The first is to study the existence and uniqueness of solutions and the properties of fully coupled FBSDEs under various conditions, and relevant literature is available at [2, 5, 13, 20, 25, 29, 30, 38, 42, 43, 55, 60, 63]. The second aspect focuses on the optimal control and differential strategy problems of controlled FBSDEs by establishing the corresponding SMP using the verification theorem and

Received 9 March 2024; Accepted 10 December 2024; Early access 20 February 2025
*Corresponding author

dynamic programming principle with viscosity solution theory. Key relevant literature include [16, 21, 23, 26, 27, 33, 41, 45, 47, 49, 52, 53, 61, 62, 64]. The third is to research the applications of FBSDEs in financial mathematics, such as the random recursive effect problem [7, 10], optimal control problem under nonlinear expectation [59], and optimal investment risk minimization problem [37]. Accordingly, the study of FBSDEs holds significant theoretical and practical value, particularly in financial engineering.

The current research on FBSDEs mainly concentrates on the finite dimensional case, whereas the research on the infinite dimensional FBSDEs is still in its early stages, offering significant potential for exploration. For example, Guatteri (2007) [14] investigated a specific class of infinite-dimensional FBSDEs known as forward-backward stochastic evolution equations (FBSEEs) characterized by the fact that the second-order differential operator in the equations is nonrandom and time independent, which is assumed to be an infinitesimal generator that can generate a class of semigroup operators. In the sense of mild solutions, Guatteri utilized the basic theory of infinite dimensional SEEs and BSEEs, along with the contraction mapping method, to establish the existence and uniqueness of local solutions under the Lipschitz condition. However, they did not establish the existence and uniqueness of global solutions in general cases, which are limited to practical applications. In the same year, Yin and Wang [58] studied a class of FBSDEs that take values in Hilbert space and establish the existence and uniqueness of the solution under the assumption of monotone coefficients, characterized by the fact that the equations do not contain second-order differential operators, i.e., they are first-order SDEs in the sense of partial differential equations (PDEs).

With the need for finer descriptions, higher applicability, richer solutions, and wider applications in describing phenomena in physics, biology, finance, etc., FBSPDE, as a natural extension of FBSDE, introduces partial derivatives on space, such as gradient and Laplace operator, which can better describe the interaction of multiple stochastic processes in space, as well as dealing with nonlinear problems. For specific FBSPDEs, Yin [56] first studied the Cauchy problem for a class of super parabolic FBSPDEs in 2014. They first established the existence and uniqueness of the local weak solutions using the contraction mapping principle under the Lipschitz condition, and then established the existence and uniqueness of global weak solutions using the continuation method and contraction mapping principle under some monotonicity assumptions on the coefficients of the equations. Yin [57] further investigated the solvability of FBSPDEs with nonmonotonic coefficients. In this paper, this topic is further investigated to consider the optimal control problem for stochastic systems described by fully coupled FBSPDEs. In 2018, Feng et al. [11] obtained an approximation procedure by studying the solutions of FBSDEs with smooth coefficients and their relation to the classical solutions of quasi-linear elliptic PDEs by establishing a connection with the weak solutions of quasi-linear elliptic PDEs. They further investigated the unique weak solution of the proposed linear elliptic PDEs by means of the solution of its quasi-linear elliptic PDEs on the infinite horizon. Cardaliaguet et al. (2019) [4] studied a class of FBSPDEs with linear coefficients of σ as well as periodic boundary conditions under specific strong assumptions and obtained wellposedness in Hölder space. Furthermore, one can refer to [8, 36] for the study of numerical approximation of fully coupled FBSPDEs. Among them, [8], as a pioneering work in this area, proposed a numerical method for solving a special class of coupled linear FBSPDEs. Recently, Molla and Qiu [36] investigated a class of coupled FBSPDEs with homogeneous Dirichlet boundary conditions and nonlinear coefficients. They presented an approach that combines finite element methods and machine learning techniques to solve such coupled FBSPDEs.

Since the 1970s, the optimal control problem for stochastic systems has been extensively studied, yielding a wealth of fundamental and significant results. A classical approach to solving the optimal control problem is to establish the necessary conditions satisfied by the optimal control, which is commonly referred to as Pontryagin's maximum principle. Regarding the SMP for the general case of finite-dimensional stochastic systems, Peng (1990) [40] conducted a related study and developed a global SMP by introducing a second-order dual equation through second-order variational differentiation of the state equation. The general case is the case where the control variable assumes a nonconvex domain, and the diffusion term of the state equation contains the control variable. Wu [54] and Yong [61] investigated the nonfully coupled and fully coupled cases of forward backward stochastic controlled systems, respectively. And they established the global form of the maximum principle for the general case via transforming the original problem into an optimal control problem for a terminal-constrained forward stochastic systemsexploiting Eckland' s variational principle and second-order dual methods. In contrast to these approaches [54, 61], Hu [17] obtained a global MP for stochastic recursive systems in the general case by introducing two new dual equations for the second-order variations of the backward component of the system. For the optimal control problem of infinite-dimensional forward stochastic systems, detailed investigations have also been conducted by numerous researchers, including Bensoussan [3], Peng and Hu [19], Zhou [67], Li and Tang [48] and so on. Nevertheless, these results are largely subject to one of the following assumptions: (i) the control domain is a convex set; (ii) the diffusion term of the state equation does not depend on the control process; (iii) both the state equation and cost functional are linear with respect to the state variables. Thus, for the general case of optimal control problems of infinite dimensional forward stochastic evolution systems, i.e., the case where the control variable takes on a nonconvex domain, and the diffusion term of the state equation contains a control variable, its SMP has been an open problem that has been unresolved for many years.

In view of the second-order variational methods developed by Peng in the finite dimensional case, the main difficulty in the infinite dimensional case lies in performing second-order dual analysis, i. e., conducting a dual characterization of the quadratic terms of the variational inequalities. In the finite dimensional case, second-order dual analysis can be characterized by an adapted solution of a matrix-valued BSDE, where the matrix-valued BSDE is called a second-order dual equation, and its adapted solution is called a second-order dual process. In the infinite dimensional case, the second-order dual equation transforms into an operator-valued infinite-dimensional BSDE, and there is no universal solvability theory. Zhang and Lü [28] made significant contribution to this problem by introducing a relaxation transpose solution to characterize operator-valued infinite-dimensional BSDEs. Consequently, they obtained the global SMP of the optimal control. Similarly, Fuhrman et al. [12] studied the optimal control problem for a specific class of stochastic parabolic PDEs with deterministic coefficients. They derived the corresponding SMP by leveraging the compactness and Markov structure of the system. In contrast to [12, 28], Meng and Du [6] established a global form of the SMP for the general case by utilizing the Lebesgue differential theorem and the Reisz representation theorem for stochastic bilinear functionals for a second-order dual representation of non-Markov state equations. Lenhart et al. [24] established the existence of optimal control and the necessary conditions for the existence of optimal control via dual principles and backward stochastic partial differential equations (BSPDEs) for stochastic partial differential equations (SPDEs) systems in which the noise is a space-time Gaussian random quantity with nonlocal terms. Stannat [46] extended the SMP of SPDEs in nonconvex control domains by representing the second order adjoint state as a solution to the function-valued BSPDE. For the stochastic optimal control problem for infinite dimensional backward systems, Mahmudov and McKibben (2007) [32] investigated the specific case in which the second-order differential operator of the system is non-stochastic and time-invariant by assuming that it is an infinitesimal generator that generates a class of semigroup operators. They established the corresponding SMP in the sense of mild solutions. Different from [32], Meng, et al. (2013, 2014) [34, 35] studied the optimal control problem for infinite dimensional backward stochastic systems with random coefficients, where the second-order differential operator of the system is stochastic and time-dependent. They established the local and global forms of SMP under convex control and nonconvex control domains, respectively, in the sense of the weak solution of the PDE. Notably, Al-Hussein and Gherbal (2014) [1] studied the optimal control problem for a class of forward backward doubly SDEs taking values in a Hilbert space. They established the corresponding SMP, which is characterized by the fact that the system does not contain second-order differential operators, i.e., it is a first-order differential equation in the sense of a PDE.

Similar to Yin [56], we use a set of non-degeneracy conditions under monotonicity assumptions. This approach was originally introduced by [20] for the infinite dimensional case and generalized by [15], which considers the existence of infinitesimal generators and unbounded operators. Therefore, the objective of this paper is to investigate Kolmogorov-type PDEs in infinite dimensional space with spatial variables, i.e., fully coupled FBSPDEs of parabolic type:

$$\begin{cases} \mathrm{d}X(t,x) = \Big[\mathfrak{L}_F(t,x,X(t,x),Y(t,x),Z(t,x)) + f(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x), \\ \qquad\qquad Z(t,x))\Big]\mathrm{d}t + \Big\langle \sigma(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x),Z(t,x)),\mathrm{d}W(t)\Big\rangle, \\ \mathrm{d}Y(t,x) = -\Big[\mathfrak{L}_B(t,x,X(t,x),Y(t,x),Z(t,x)) + b(t,x,X(t,x),\nabla X(t,x), \\ \qquad\qquad Y(t,x),\nabla Y(t,x),Z(t,x))\Big]\mathrm{d}t + \Big\langle Z(t,x),\mathrm{d}W(t)\Big\rangle,\ (t,x)\in[0,T]\times\mathbb{R}^n, \\ X(0,x) = \phi(x,Y_0(x)),\ Y(T,x) = \psi(x,X_T(x)),\ x\in\mathbb{R}^n, \end{cases} \tag{1.1}$$

where we denote

$$\begin{cases} \mathfrak{L}_F(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq \nabla\cdot\big(A\nabla X(t,x) + B\nabla Y(t,x) + CZ(t,x)\big), \\ \mathfrak{L}_B(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq \nabla\cdot\big(-D\nabla X(t,x) + E\nabla Y(t,x) + FZ(t,x)\big), \end{cases} \tag{1.2}$$

where $A=A(t,x)$, $B=B(t,x)$, $D=D(t,x)$, $E=E(t,x)\in\mathbb{S}^n$ are random fields, $C=C(t)$ and $F=F(t)$ are also matrix-valued random fields with corresponding dimensions. $T\in(0,\infty)$ is a finite deterministic time. Note that if A,B,D, and E are differentiable in x, then the divergence form of the operators $\mathfrak{L}_F$ and $\mathfrak{L}_B$ would be equivalent to the standard form with some corresponding changes in the functions f and b:

$$\begin{cases} \widetilde{\mathfrak{L}}_F(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq tr\big\{A\mathscr{D}^2X(t,x) + B\mathscr{D}^2Y(t,x) + C^\top\nabla Z(t,x)\big\}, \\ \widetilde{\mathfrak{L}}_B(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq tr\big\{-D\mathscr{D}^2X(t,x) + E\mathscr{D}^2Y(t,x) + F^\top\nabla Z(t,x)\big\}, \end{cases} \tag{1.3}$$

where $\mathscr{D}^2$ denotes the Hessian operator.

This study focuses on an innovative class of fully coupled forward-backward stochastic partial differential control systems, aiming to establish a necessary and sufficient maximum principle utilizing SPDE techniques. To achieve this objective, we first employ Itô's formula to demonstrate the continuous dependence theorem for SPDEs and BSPDEs within Gelfand triples under appropriate assumptions. The necessary maximum principle is rigorously established under the assumption of convexity in the control domain. We also establish a sufficient SMP by incorporating extra convexity assumptions on the initial cost, terminal cost, and Hamiltonian. To illustrate our results, we apply them to a linear quadratic (LQ) control problem. By employing sufficient and necessary maximum principles, we derive the optimal control strategy under dual representation. The novelty of this work lies in the use of PDE techniques. By constructing a Hamiltonian and Taylor expansion, we streamline the computation of performance metrics, thereby avoiding the complexities associated with traditional variational methods. Importantly, our study is not a mere extension of the results from finite-dimensional FBSDEs to FBSEEs, where coefficients are stochastic functions of space and time. Instead, our focus is on FBSDEs containing real partial derivatives, employing PDE techniques for estimation and solving optimal control problems, distinguishing our approach in terms of methodology and assumptions from existing methods.

The remainder of this paper is organized as follows. Section 2 introduces the notations in Hilbert spaces and their Itô's formula. Section 3 presents two lemmas on SPDE and BSPDE. Section 4 presents an important estimate of FBSPDE based on a Gelfand triple. In Section 5, we establish sufficient and necessary SMP for the optimal control problem of FBSPDE. Finally, in Section 6, we investigate a kind of LQ problem and obtain the expression for the optimal control using the previous results.

2. Notations and preliminaries

Let $T>0$ be given a positive number. Consider a filtered complete probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ with a filtration $\mathbb{F}=\{\mathcal{F}_t : t\in[0,T]\}$ satisfying the usual conditions of right-continuity and $\mathbb{P}$-completeness, which is generated by a d-dimensional standard Brownian motion $\{W(t)=(W_1(t),W_2(t),\ldots,W_d(t))^\top : t\in[0,T]\}$. Let $\mathscr{P}_{[m,n]}$ be the predictable σ-algebra on $[m,n]\times\Omega$ and, in particular, $\mathscr{P}$ be the predictable σ -algebra on $[0,T]\times\Omega$. We denote by $\mathscr{B}(\Lambda)$ the Borel σ-algebra of any topological space Λ. For any $m\in\mathbb{N}, l\in\mathbb{R}$, let $H^l(\mathbb{R}^n;\mathbb{R}^m)$ be the Sobolev space $W_2^l(\mathbb{R}^n;\mathbb{R}^m)$, where we denote $H^0(\mathbb{R}^n;\mathbb{R}^m)=L^2(\mathbb{R}^n;\mathbb{R}^m)$. Additionally, let $|\cdot|$ be the standard norm of $\mathbb{R}^m$, $\|\cdot\|_{L^2}$ be the norm of $L^2(\mathbb{R}^n;\mathbb{R}^m)$, $\|\cdot\|_{H^{-1}}$ be the norm of $H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$ and $\|\cdot\|$, $\|\cdot\|_{H^1}$ be norms of $H^1(\mathbb{R}^n;\mathbb{R}^m)$. They are given as follows:

$$\|X\|_{L^2}^2 \triangleq \int_{\mathbb{R}^n}|X|^2\mathrm{d}x,\quad \|X\|^2 \triangleq \int_{\mathbb{R}^n}|\nabla X|^2\mathrm{d}x,\quad \text{and } \|X\|_{H^1}^2 \triangleq \|X\|_{L^2}^2+\|X\|^2.$$

Denote by $H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$ the dual space of $H^1(\mathbb{R}^n;\mathbb{R}^m)$. Denote by $\langle\cdot,\cdot\rangle_{L^2}$, $\langle\cdot,\cdot\rangle_{H^1}$ the inner product of $L^2(\mathbb{R}^n;\mathbb{R}^m)$ and $H^1(\mathbb{R}^n;\mathbb{R}^m)$, respectively, and by $\langle\cdot,\cdot\rangle_{H^{-1},H^1}$ the duality product between $H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$ and $H^1(\mathbb{R}^n;\mathbb{R}^m)$. For $\forall X\in L^2(\mathbb{R}^n;\mathbb{R}^m)$, there exists an $X'\in H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$, s.t. $\langle X',Y\rangle_{H^{-1},H^1}=\langle X,Y\rangle_{L^2}, \forall Y\in H^1(\mathbb{R}^n;\mathbb{R}^m)$. The mapping $X\to X'$ is linear, injective, compact and continuous, and we can identify X' with X. In this sense, we identify $(L^2(\mathbb{R}^n;\mathbb{R}^m))^{-1}$ with $L^2(\mathbb{R}^n;\mathbb{R}^m)$ and then $L^2(\mathbb{R}^n;\mathbb{R}^m)$ is a dense subset of $H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$. Thus, we get a Gelfand triple $H^1(\mathbb{R}^n;\mathbb{R}^m)\subset L^2(\mathbb{R}^n;\mathbb{R}^m)\subset H^{-1}(\mathbb{R}^n;\mathbb{R}^m)$.

In the following, we introduce some notations for later use.

- $\mathbb{R}^n$: the n-dimensional Euclidean space.
- $\mathbb{R}^{n\times m}$: the collection of all $n\times m$ matrices.
- $A^\top$: the dual operator of the operator A.
- A^{-1}: the inverse operator of the operator A.
- $\mathbb{S}^n\subset\mathbb{R}^{n\times n}$: the collection of all $n\times n$ symmetric matrices.

- $\mathbb{S}^n_+ \subset \mathbb{S}^n$: the collection of all $n \times n$ non-negative definite symmetric matrices.
- $\nabla\xi = (\partial_{x_1}\xi, \partial_{x_2}\xi, \ldots, \partial_{x_n}\xi)^\top, \quad \forall \xi \in C^1(\mathbb{R}^n;\mathbb{R})$.
- $\nabla\zeta = (\nabla\zeta_1, \nabla\zeta_2, \ldots, \nabla\zeta_m), \quad \forall \zeta \in C^1(\mathbb{R}^n;\mathbb{R}^m)$.
- $\nabla\cdot\xi = \sum\limits_{i=1}^{n} \frac{\partial\xi_i}{\partial x_i}, \quad \forall \xi \in C^1(\mathbb{R}^n;\mathbb{R}^n)$.
- $\nabla\cdot\zeta = (\nabla\cdot\zeta_1, \nabla\cdot\zeta_2, \ldots, \nabla\cdot\zeta_m)^\top, \quad \forall \zeta \in C^1(\mathbb{R}^n;\mathbb{R}^{n\times m})$.
- $L^2_{\mathcal{F}_t}(\Omega;\mathbb{H})$: the space of all $\mathcal{F}_t$-strongly measurable random variables $\xi:\Omega\to\mathbb{H}$ satisfying

$$\|\xi\|^2_{L^2_{\mathcal{F}_t}(\Omega;\mathbb{H})} = \mathbb{E}\|\xi\|^2_{\mathbb{H}} < \infty.$$

- $C_{\mathbb{F}}\big(0,T;L^2(\Omega;\mathbb{H})\big)$: the space of all $\mathbb{H}$-valued and $\mathbb{F}$-adapted continuous processes $f:[0,T]\times\Omega \to \mathbb{H}$ satisfying

$$\|f(\cdot)\|^2_{C_{\mathbb{F}}(0,T;L^2(\Omega;H))} = \sup_{0\leqslant t\leqslant T}\Big[\mathbb{E}\|f(t)\|^2_{\mathbb{H}}\Big] < \infty.$$

- $L^2_{\mathbb{F}}\big(\Omega;C(0,T;\mathbb{H})\big)$: the space of all $\mathbb{H}$-valued and $\mathbb{F}$-adapted continuous processes $f:[0,T]\times\Omega \to \mathbb{H}$ satisfying

$$\|f(\cdot)\|^2_{L^2_{\mathbb{F}}(\Omega;C(0,T;\mathbb{H}))} = \mathbb{E}\Big[\sup_{t\in[0,T]}\|f(t)\|^2_{\mathbb{H}}\Big] < \infty.$$

- $L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{H}))$: the space of all $\mathbb{H}$-valued and $\mathbb{F}$-adapted processes $f:[0,T]\times\Omega\to\mathbb{H}$ satisfying

$$\|f(\cdot)\|^2_{L^2_{\mathbb{F}}(\Omega;L^2(0,T;\mathbb{H}))} = \mathbb{E}\Big[\int_0^T \|f(t)\|^2_{\mathbb{H}}\mathrm{d}t\Big] < \infty.$$

- $L^\infty_{\mathbb{F}}(0,T;\mathbb{H})$: the space of all $\mathbb{H}$-valued and $\mathbb{F}$-adapted essentially bounded processes.
- $\mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}) \triangleq L^2_{\mathbb{F}}\big(\Omega;\ C(0,T;\ L^2(\mathbb{R}^n;\mathbb{R}))\big) \cap\ L^2_{\mathbb{F}}\big(\Omega;\ L^2(0,T;H^1(\mathbb{R}^n;\ \mathbb{R}))\big)$. For any $X(\cdot,\cdot)\in\mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R})$, its norm is given by

$$\|X(\cdot,\cdot)\|^2_{\mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R})} = \mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t,\cdot)\|^2_{L^2} + \int_0^T \|X(t,\cdot)\|^2_{H^1}\mathrm{d}t\Big].$$

- $\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^d) \triangleq L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^d)))$. For any $Z(\cdot,\cdot)\in\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^d)$, its norm is given by

$$\|Z(\cdot,\cdot)\|^2_{\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^d)} = \mathbb{E}\Big[\int_0^T \|Z(t,\cdot)\|^2_{L^2}\mathrm{d}t\Big].$$

- $\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R}^{1+d}) \triangleq \mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R}) \times \mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R}^d)$.Foranу$\Theta(\cdot,\cdot) = (Y(\cdot,\cdot),Z(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{1+d})$, its norm is given by

$$\|\Theta(\cdot,\cdot)\|^2_{\mathscr{N}^2_{\mathbb{F}}(0,T;\mathbb{R}^{1+d})} = \mathbb{E}\Big[\sup_{t\in[0,T]}\|Y(t,\cdot)\|^2_{L^2} + \int_0^T \|Y(t,\cdot)\|^2_{H^1}\mathrm{d}t + \int_0^T \|Z(t,\cdot)\|^2_{L^2}\mathrm{d}t\Big].$$

- $\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R}^{2+d}) \triangleq \mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R})^2 \times \mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\ \mathbb{R}^d)$. For any $\Phi(\cdot,\cdot) = (X(\cdot,\cdot),Y(\cdot,\cdot),\ Z(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$, its norm is given by

$$\begin{aligned}\|\Phi(\cdot,\cdot)\|^2_{\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})} = \mathbb{E}\Big[&\sup_{t\in[0,T]}\|X(t,\cdot)\|^2_{L^2} + \int_0^T \|X(t,\cdot)\|^2_{L^2}\mathrm{d}t + \sup_{t\in[0,T]}\|Y(t,\cdot)\|^2_{L^2}\\ &+ \int_0^T \|Y(t,\cdot)\|^2_{L^2}\mathrm{d}t + \int_0^T \|Z(t,\cdot)\|^2_{L^2}\mathrm{d}t\Big].\end{aligned}$$

In what follows, we denote K to be a positive constant, which may differ from line to line.

Next, we propose the following form of Itô's formula.

Lemma 2.1 *Suppose that* $X(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega\,;\,L^2(0,T;H^1(\mathbb{R}^n\,;\,\mathbb{R})))$, $\alpha(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega\,;\,L^2(0,T;H^{-1}(\mathbb{R}^n;\mathbb{R})))$ *and* $\beta(\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^d)))$, *with the following relation satisfied*

$$\mathrm{d}X(s)=\alpha(s)\mathrm{d}s+\big\langle\beta(s),\mathrm{d}W(s)\big\rangle,\quad s\in[0,T]. \tag{2.1}$$

Then Itô's formula holds as follows:

$$\begin{aligned}\big\|X(t)\big\|^2_{L^2}=\big\|X(0)\big\|^2_{L^2}&+\int_0^t\Big[2\big\langle\alpha(s),X(s)\big\rangle_{H^{-1},H^1}+\big\|\beta(s)\big\|^2_{L^2}\Big]\mathrm{d}s\\&+2\int_0^t\Big\langle\int_{\mathbb{R}^n}X(s,x)\beta(s,x)\mathrm{d}x,\mathrm{d}W(s)\Big\rangle,\ s\in[0,T].\end{aligned} \tag{2.2}$$

Proof This result is classic, and the proof can be found in [39] and [44]. □

3. SPDE and BSPDE

In what follows, we present some basic results on SPDE and BSPDE. First, consider the SPDE of the following form:

$$\begin{cases}\mathrm{d}X(t,x)=\Big[\nabla\cdot\big(A\nabla X(t,x)+\theta(t,x)\big)+f(t,x,X(t,x),\nabla X(t,x))\Big]\mathrm{d}t\\ \qquad\qquad+\Big\langle\sigma(t,x,X(t,x),\nabla X(t,x)),\mathrm{d}W(t)\Big\rangle,\\ X(0,x)=\phi(x),\ (t,x)\in[0,T]\times\mathbb{R}^n.\end{cases} \tag{3.1}$$

Moreover, the coefficients (A,θ,f,σ,ϕ) are assumed to satisfy the following conditions:

Assumption 3.1 (i) *The random mapping* $A:[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{S}^n$ *is* $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)$*−measurable. Suppose that* A *is uniformly bounded and there exists a constant* $c_1>0$ *such that*

$$A\geqslant c_1I,$$

where I *is an identity matrix with the* $n\times n$ *dimension.*

(ii) *The mapping*

$$\theta(t,\omega,x):[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{R}^n$$

is $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)$*-measurable .*

(iii) *The mappings*

$$\begin{aligned}f(t,\omega,x,\gamma_1,\gamma_2)&:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\to\mathbb{R},\\ \sigma(t,\omega,x,\gamma_1,\gamma_2)&:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^d\end{aligned}$$

are $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)$*-measurable. There exist constants* $L^f,L^\sigma_1,L^\sigma_2>0$, *such that for all* $(\gamma_1,\gamma_2),(\bar\gamma_1,\bar\gamma_2)\in\mathbb{R}\times\mathbb{R}^n$, $\forall(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$,

$$\begin{aligned}\big|f(t,\omega,x,\gamma_1,\gamma_2)-f(t,\omega,x,\bar\gamma_1,\bar\gamma_2)\big|&\leqslant L^f\Big(\big|\gamma_1-\bar\gamma_1\big|+\big|\gamma_2-\bar\gamma_2\big|\Big),\\ \big|\sigma(t,\omega,x,\gamma_1,\gamma_2)-\sigma(t,\omega,x,\bar\gamma_1,\bar\gamma_2)\big|&\leqslant L^\sigma_1\big|\gamma_1-\bar\gamma_1\big|+L^\sigma_2\big|\gamma_2-\bar\gamma_2\big|.\end{aligned}$$

(iv) $\phi(\cdot,\cdot)\in L^2_{\mathcal{F}_0}(\Omega;L^2(\mathbb{R}^n;\mathbb{R}))$, $\theta(\cdot,\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^n)))$, $f(\cdot,\cdot,\cdot,0,0)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R})))$ *and* $\sigma(\cdot,\cdot,\cdot,0,0)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^d)))$.

Before discussing the a priori estimation of the SPDE, we present the definition of a weak solution.

Definition 3.1 *The random function $X(\cdot,\cdot) \in \mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R})$ is called a weak solution to the SPDE (3.1) if for each $\eta \in H^1(\mathbb{R}^n;\mathbb{R})$ and almost every $(t,\omega) \in [0,T]\times\Omega$, it holds that*

$$\begin{aligned}\Big\langle X(t,\cdot),\eta(\cdot)\Big\rangle_{L^2} = \Big\langle \phi(\cdot),\eta(\cdot)\Big\rangle_{L^2} &- \int_0^t \Big\langle A(s,\cdot)\nabla X(s,\cdot)+\theta(s,\cdot),\nabla\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s\\ &+\int_0^t \Big\langle f(s,\cdot,X(s,\cdot),\nabla X(s,\cdot)),\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s\\ &+\int_0^t \Big\langle \sigma(s,\cdot,X(s,\cdot),\nabla X(s,\cdot)),\eta(\cdot)\mathrm{d}W(s)\Big\rangle_{L^2}. \end{aligned}\tag{3.2}$$

Lemma 3.1 (Continuous Dependence Theorem of SPDE) *Under Assumption 3.1, we further assume that $2c_1 > |L_2^\sigma|^2$. Let (A,θ,f,σ,ϕ) and $(A,\bar\theta,\bar f,\bar\sigma,\bar\phi)$ be two sets of coefficients for the SPDE* (3.1) *and suppose that $X(\cdot,\cdot)$, $\bar X(\cdot,\cdot)\in\mathscr{R}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R})$ are the solutions of the SPDE* (3.1) *corresponding to (A,θ,f,σ,ϕ) and $(A,\bar\theta,\bar f,\bar\sigma,\bar\phi)$, respectively. Then, there exists a constant K depending only on c_1, T, L^f, L_1^σ and L_2^σ such that*

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t)-\bar X(t)\|^2_{L^2}\Big]+\mathbb{E}\Big[\int_0^T\|X(t)-\bar X(t)\|^2_{H^1}\mathrm{d}s\Big]\\ &\leqslant K\mathbb{E}\Big[\|\phi-\bar\phi\|^2_{L^2}+\int_0^T\big\|\theta(t)-\bar\theta(t)\big\|^2_{L^2}\mathrm{d}t+\int_0^T\big\|f(t,\bar X(t),\nabla\bar X(t))-\bar f(t,\bar X(t),\nabla\bar X(t))\big\|^2_{L^2}\mathrm{d}t\\ &\quad+\int_0^T\big\|\sigma(t,\bar X(t),\nabla\bar X(t))-\bar\sigma(t,\bar X(t),\nabla\bar X(t))\big\|^2_{L^2}\mathrm{d}t\Big].\end{aligned}\tag{3.3}$$

In particular, for $(A,\bar\theta,\bar f,\bar\sigma,\bar\phi)=(A,0,0,0,0)$, we have the following priori estimate:

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t)\|^2_{L^2}\Big]+\mathbb{E}\Big[\int_0^T\|X(t)\|^2_{H^1}\mathrm{d}s\Big]\\ &\leqslant K\mathbb{E}\Big[\|\phi\|^2_{L^2}+\int_0^T\big\|\theta(t)\big\|^2_{L^2}\mathrm{d}t+\int_0^T\big\|f(t,0,0)\big\|^2_{L^2}\mathrm{d}t+\int_0^T\big\|\sigma(t,0,0)\big\|^2_{L^2}\mathrm{d}t\Big].\end{aligned}\tag{3.4}$$

Proof First, for simplicity, we denote by

$$\begin{cases}\widehat X = X(t,x)-\bar X(t,x), \quad \widehat\theta=\theta(t,x)-\bar\theta(t,x), \quad \widehat\phi=\phi(x)-\bar\phi(x),\\ \widehat f(t,X,\nabla X)=f(t,X,\nabla X)-\bar f(t,X,\nabla X), \quad \widehat\sigma(t,X,\nabla X)=\sigma(t,X,\nabla X)-\bar\sigma(t,X,\nabla X).\end{cases}$$

Using Itô's formula to $\|\widehat X(s)\|^2_{L^2}$ and integration by parts technique, we obtain

$$\begin{aligned}&\|\widehat X(t)\|^2_{L^2}\\ &=\|\widehat\phi\|^2_{L^2}-2\int_0^t\langle\nabla\widehat X,A\nabla\widehat X+\widehat\theta\rangle_{L^2}\mathrm{d}s+2\int_0^t\langle\widehat X,f(s,X,\nabla X)-\bar f(s,\bar X,\nabla\bar X)\rangle_{L^2}\mathrm{d}s\\ &\quad+\int_0^t\big\|\sigma(s,X,\nabla X)-\bar\sigma(s,\bar X,\nabla\bar X)\big\|^2_{L^2}\mathrm{d}s+2\int_0^t\langle\widehat X,\big(\sigma(s,X,\nabla X)-\bar\sigma(s,\bar X,\nabla\bar X)\big)\mathrm{d}W_s\rangle_{L^2}.\end{aligned}\tag{3.5}$$

It is easy to verify that

$$-2\int_0^t \langle \nabla \widehat{X}, A\nabla \widehat{X}+\widehat{\theta}\rangle_{L^2}\mathrm{d}s \leqslant -2c_1\int_0^t \|\nabla \widehat{X}\|^2\mathrm{d}s+\epsilon_1\int_0^t \|\nabla \widehat{X}\|^2\mathrm{d}s+\frac{1}{\epsilon_1}\int_0^t \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s$$
$$=-\big(2c_1-\epsilon_1\big)\int_0^t \big(\|\widehat{X}\|_{H^1}^2-\|\widehat{X}\|_{L^2}^2\big)\mathrm{d}s+\frac{1}{\epsilon_1}\int_0^t \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s. \tag{3.6}$$

Substituting (3.6) into (3.5), we derive

$$\begin{aligned}
&\|\widehat{X}(t)\|_{L^2}^2+\big(2c_1-\epsilon_1\big)\int_0^t \|\widehat{X}\|_{H^1}^2\mathrm{d}s\\
\leqslant{}& \|\widehat{\phi}\|_{L^2}^2+\big(2c_1-\epsilon_1\big)\int_0^t \|\widehat{X}\|_{L^2}^2\mathrm{d}s+\frac{1}{\epsilon_1}\int_0^t \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s\\
&+2\int_0^t \langle \widehat{X}, f(s,X,\nabla X)-\bar{f}(s,\bar{X},\nabla\bar{X})\rangle_{L^2}\mathrm{d}s\\
&+\int_0^t \big\|\sigma(s,X,\nabla X)-\bar{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s\\
&+2\int_0^t \langle \widehat{X}, \big(\sigma(s,X,\nabla X)-\bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\mathrm{d}W_s\rangle_{L^2}.
\end{aligned} \tag{3.7}$$

Then, by taking expectation on both sides of (3.7), we can get

$$\begin{aligned}
&\mathbb{E}\Big[\|\widehat{X}(t)\|_{L^2}^2\Big]+\big(2c_1-\epsilon_1\big)\mathbb{E}\Big[\int_0^t \|\widehat{X}\|_{H^1}^2\mathrm{d}s\Big]\\
\leqslant{}& \mathbb{E}\Big[\big\|\widehat{\phi}\big\|_{L^2}^2+\big(2c_1-\epsilon_1\big)\int_0^t \|\widehat{X}\|_{L^2}^2\mathrm{d}s+\frac{1}{\epsilon_1}\int_0^t \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s\\
&+2\int_0^t \langle \widehat{X}, f(s,X,\nabla X)-\bar{f}(s,\bar{X},\nabla\bar{X})\rangle_{L^2}\mathrm{d}s\\
&+\int_0^t \big\|\sigma(s,X,\nabla X)-\bar{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s\Big].
\end{aligned} \tag{3.8}$$

From the Lipschitz condition in Assumption 3.1 and Young's inequality $ab\leqslant \epsilon a^2+\frac{1}{4\epsilon}b^2, \forall \epsilon>0$, we obtain

$$\begin{aligned}
&2\int_0^t \Big|\langle \widehat{X}, f(s,X,\nabla X)-\bar{f}(s,\bar{X},\nabla\bar{X})\rangle_{L^2}\Big|\mathrm{d}s\\
\leqslant{}& 2\int_0^t \big\|\widehat{X}\big\|_{L^2}\big\|f(s,X,\nabla X)-f(s,\bar{X},\nabla\bar{X})+f(s,\bar{X},\nabla\bar{X})-\bar{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s\\
\leqslant{}& 2\int_0^t \big\|\widehat{X}\big\|_{L^2}\big\|f(s,X,\nabla X)-f(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s+2\int_0^t \big\|\widehat{X}\big\|_{L^2}\big\|f(s,\bar{X},\nabla\bar{X})-\bar{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s\\
\leqslant{}& 2\big|L^f\big|\int_0^t \Big(\big\|\widehat{X}\big\|_{L^2}^2+\big\|\widehat{X}\big\|_{L^2}\big\|\nabla\widehat{X}\big\|\Big)\mathrm{d}s+2\int_0^t \big\|\widehat{X}\big\|_{L^2}\big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s\\
\leqslant{}& \left(2\big|L^f\big|+\frac{\big|L^f\big|^2}{\epsilon_2}+\epsilon_3\right)\int_0^t \big\|\widehat{X}\big\|_{L^2}^2\mathrm{d}s+\epsilon_2\int_0^t \big\|\nabla\widehat{X}\big\|\mathrm{d}s+\frac{1}{\epsilon_3}\int_0^t \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s\\
={}& \left(2\big|L^f\big|+\frac{\big|L^f\big|^2}{\epsilon_2}+\epsilon_3-\epsilon_2\right)\int_0^t \big\|\widehat{X}\big\|_{L^2}^2\mathrm{d}s+\epsilon_2\int_0^t \big\|\widehat{X}\big\|_{H^1}^2\mathrm{d}s+\frac{1}{\epsilon_3}\int_0^t \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}\mathrm{d}s,
\end{aligned} \tag{3.9}$$

likewise,

$$
\begin{aligned}
&\int_0^t \left\|\sigma(s,X,\nabla X)-\bar\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
=&\int_0^t \left\|\sigma(s,X,\nabla X)-\sigma(s,\bar X,\nabla\bar X)+\sigma(s,\bar X,\nabla\bar X)-\bar\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
\leqslant&\ (1+\epsilon_4)\int_0^t \left\|\sigma(s,X,\nabla X)-\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
&+\left(1+\frac{1}{\epsilon_4}\right)\int_0^t \left\|\sigma(s,\bar X,\nabla\bar X)-\bar\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
\leqslant&\ (1+\epsilon_4)\int_0^t \left(\left|L_1^\sigma\right|\|\widehat X\|_{L^2}+\left|L_2^\sigma\right|\|\nabla\widehat X\|_{L^2}\right)^2\mathrm{d}s+\left(1+\frac{1}{\epsilon_4}\right)\int_0^t \left\|\widehat\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
\leqslant&\ (1+\epsilon_4)\left(1+\frac{1}{\epsilon_5}\right)\left|L_1^\sigma\right|^2\int_0^t \|\widehat X\|_{L^2}\mathrm{d}s+(1+\epsilon_4)(1+\epsilon_5)\left|L_2^\sigma\right|^2\int_0^t \|\nabla\widehat X\|_{L^2}\mathrm{d}s\\
&+\left(1+\frac{1}{\epsilon_4}\right)\int_0^t \left\|\widehat\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
=&\ (1+\epsilon_4)\left(\left(1+\frac{1}{\epsilon_5}\right)\left|L_1^\sigma\right|^2-(1+\epsilon_5)\left|L_2^\sigma\right|^2\right)\int_0^t \|\widehat X\|_{L^2}^2\mathrm{d}s\\
&+(1+\epsilon_4)(1+\epsilon_5)\left|L_2^\sigma\right|^2\int_0^t \|\widehat X\|_{H^1}^2\mathrm{d}s\ +\left(1+\frac{1}{\epsilon_4}\right)\int_0^t \left\|\widehat\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s.
\end{aligned}
\tag{3.10}
$$

Combining (3.8), (3.9) and (3.10), we derive

$$
\begin{aligned}
&\mathbb{E}\Big[\|\widehat X(t)\|_{L^2}^2\Big]+\Big(2c_1-\epsilon_1-\epsilon_2-\big(1+\epsilon_4\big)\big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2\Big)\mathbb{E}\Big[\int_0^t \|\widehat X\|_{H^1}^2\mathrm{d}s\Big]\\
\leqslant&\ \mathbb{E}\Big[\|\widehat\phi\|_{L^2}^2+\Big(2c_1-\epsilon_1+2\big|L^f\big|+\frac{\left|L^f\right|^2}{\epsilon_2}+\epsilon_3-\epsilon_2\\
&+\big(1+\epsilon_4\big)\left(1+\frac{1}{\epsilon_5}\right)\left|L_1^\sigma\right|^2-\big(1+\epsilon_4\big)\big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2\Big)\int_0^t \|\widehat X\|_{L^2}^2\mathrm{d}s\\
&+\frac{1}{\epsilon_1}\int_0^t \|\widehat\theta\|_{L^2}^2\mathrm{d}s+\frac{1}{\epsilon_3}\int_0^t \left\|\widehat f(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s+\left(1+\frac{1}{\epsilon_4}\right)\int_0^t \left\|\widehat\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\Big].
\end{aligned}
\tag{3.11}
$$

Noting that $2c_1>|L_2^\sigma|^2$, for any $\epsilon_3>0$, choose sufficiently small $\epsilon_1,\epsilon_2,\epsilon_4,\epsilon_5>0$, such that $K_2=K_2(c_1,L_2^\sigma)=2c_1-\epsilon_1-\epsilon_2-\big(1+\epsilon_4\big)\big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2>0$ holds, then applying Gronwall's inequality gives

$$
\begin{aligned}
\sup_{t\in[0,T]}\mathbb{E}\Big[\|\widehat X(t)\|_{L^2}^2\Big]\leqslant&\ \exp\{TK_1\}\mathbb{E}\Big[\|\widehat\phi\|_{L^2}^2+\frac{1}{\epsilon_1}\int_0^t \|\widehat\theta\|_{L^2}^2\mathrm{d}s+\frac{1}{\epsilon_3}\int_0^T \left\|\widehat f(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\\
&+\left(1+\frac{1}{\epsilon_4}\right)\int_0^T \left\|\widehat\sigma(s,\bar X,\nabla\bar X)\right\|_{L^2}^2\mathrm{d}s\Big],
\end{aligned}
\tag{3.12}
$$

where $K_1=K_1(c_1,L^f,L_1^\sigma,L_2^\sigma)=2c_1-\epsilon_1+2\big|L^f\big|+\frac{|L^f|^2}{\epsilon_2}+\epsilon_3-\epsilon_2+\big(1+\epsilon_4\big)\left(1+\frac{1}{\epsilon_5}\right)\left|L_1^\sigma\right|^2-\big(1+\epsilon_4\big)\big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2$. Furthermore, according to (3.11) and (3.12), we get

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T \|\widehat{X}\|_{H^1}^2 \mathrm{d}s\Big] \\
&\leqslant \frac{1}{K_2}\Big\{\mathbb{E}\Big[\big\|\widehat{\phi}\big\|_{L^2}^2\Big] + K_1\mathbb{E}\Big[\int_0^T \big\|\widehat{X}\big\|_{L^2}^2 \mathrm{d}s\Big] + \frac{1}{\epsilon_1}\int_0^T \|\widehat{\theta}\|_{L^2}^2 \mathrm{d}s + \frac{1}{\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big] \\
&\quad + \Big(1+\frac{1}{\epsilon_4}\Big)\mathbb{E}\Big[\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big]\Big\} \\
&\leqslant \frac{1}{K_2}\Big\{\mathbb{E}\Big[\big\|\widehat{\phi}\big\|_{L^2}^2\Big] + TK_1 \sup_{t\in[0,T]}\mathbb{E}\Big[\big\|\widehat{X}\big\|_{L^2}^2\Big] + \frac{1}{\epsilon_1}\int_0^T \|\widehat{\theta}\|_{L^2}^2 \mathrm{d}s + \frac{1}{\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big] \\
&\quad + \Big(1+\frac{1}{\epsilon_4}\Big)\mathbb{E}\Big[\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big]\Big\} \\
&\leqslant \frac{1+TK_1\exp\{TK_1\}}{K_2}\Big\{\mathbb{E}\Big[\big\|\widehat{\phi}\big\|_{L^2}^2\Big] + \frac{1}{\epsilon_1}\int_0^T \|\widehat{\theta}\|_{L^2}^2 \mathrm{d}s + \frac{1}{\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big] \\
&\quad + \Big(1+\frac{1}{\epsilon_4}\Big)\mathbb{E}\Big[\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big]\Big\}. \qquad (3.13)
\end{aligned}
$$

From (3.7), (3.9) and (3.10), and taking supremum over $t \in [0,T]$, then we have

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \|\widehat{X}(t)\|_{L^2}^2\Big] + K_2\mathbb{E}\Big[\int_0^T \|\widehat{X}\|_{H^1}^2 \mathrm{d}s\Big] \\
&\leqslant \mathbb{E}\Big[\|\widehat{\phi}\|_{L^2}^2 + K_1\int_0^T \|\widehat{X}\|_{L^2}^2 \mathrm{d}s + \frac{1}{\epsilon_1}\int_0^T \|\widehat{\theta}\|_{L^2}^2 \mathrm{d}s + \frac{1}{\epsilon_3}\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s \\
&\quad + \Big(1+\frac{1}{\epsilon_4}\Big)\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s + 2\sup_{t\in[0,T]}\Big|\int_0^t \langle \widehat{X}, \big(\sigma(s,X,\nabla X) - \bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\mathrm{d}W_s\rangle_{L^2}\Big|\Big]. \qquad (3.14)
\end{aligned}
$$

Subsequently, we deal with the stochastic integrals in (3.14) using the Burkholder-Davis-Gundy inequality and Lipschitz continuity condition together with (3.12) and (3.13). As a result, we obtain the following equation

$$
\begin{aligned}
&2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \langle \widehat{X}, \big(\sigma(s,X,\nabla X) - \bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\mathrm{d}W_s\rangle_{L^2}\Big|\Big] \\
&\leqslant 8\sqrt{2}\mathbb{E}\Big[\Big(\int_0^T \Big|\langle \widehat{X}, \big(\sigma(s,X,\nabla X) - \bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\rangle_{L^2}\Big|^2 \mathrm{d}s\Big)^{\frac{1}{2}}\Big] \\
&\leqslant 8\sqrt{2}\mathbb{E}\Big[\Big(\int_0^T \big\|\widehat{X}\big\|_{L^2}^2 \cdot \big\|\big(\sigma(s,X,\nabla X) - \bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\big\|_{L^2}^2 \mathrm{d}s\Big)^{\frac{1}{2}}\Big] \\
&\leqslant 8\sqrt{2}\mathbb{E}\Big[\Big(\sup_{t\in[0,T]}\big\|\widehat{X}(t)\big\|_{L^2}^2 \int_0^T \big\|\big(\sigma(s,X,\nabla X) - \bar{\sigma}(s,\bar{X},\nabla\bar{X})\big)\big\|_{L^2}^2 \mathrm{d}s\Big)^{\frac{1}{2}}\Big] \\
&\leqslant \mathbb{E}\Big[\epsilon_6\sup_{t\in[0,T]}\big\|\widehat{X}(t)\big\|_{L^2}^2 + \frac{32}{\epsilon_6}(1+\epsilon_4)\Big(\Big(1+\frac{1}{\epsilon_5}\Big)\big|L_1^\sigma\big|^2 - (1+\epsilon_5)\big|L_2^\sigma\big|^2\Big)\int_0^T \big\|\widehat{X}\big\|_{L^2}^2 \mathrm{d}s \\
&\quad + \frac{32}{\epsilon_6}(1+\epsilon_4)(1+\epsilon_5)\big|L_2^\sigma\big|^2\int_0^T \big\|\widehat{X}\big\|_{H^1}^2 \mathrm{d}s + \frac{32}{\epsilon_6}\Big(1+\frac{1}{\epsilon_4}\Big)\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big] \\
&\leqslant \epsilon_6\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|\widehat{X}(t)\big\|_{L^2}^2\Big] + K_3\mathbb{E}\Big[\big\|\widehat{\phi}\big\|_{L^2}^2\Big] + \frac{K_3}{\epsilon_1}\mathbb{E}\Big[\int_0^T \|\widehat{\theta}\|_{L^2}^2 \mathrm{d}s\Big] + \frac{K_3}{\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big] \\
&\quad + \Big(K_3 + \frac{32}{\epsilon_6}\Big)\Big(1+\frac{1}{\epsilon_4}\Big)\mathbb{E}\Big[\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2 \mathrm{d}s\Big], \qquad (3.15)
\end{aligned}
$$

where $K_3 = K_3(T, c_1, L^f, L_1^\sigma, L_2^\sigma) = \frac{32}{\epsilon_6}\big(1+\epsilon_4\big)\Big(\Big(1+\frac{1}{\epsilon_5}\Big)\,\big|L_1^\sigma\big|^2 - \big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2\Big)T\exp\{TK_1\} + \frac{32}{\epsilon_6}\big(1+\epsilon_4\big)\big(1+\epsilon_5\big)\big|L_2^\sigma\big|^2\frac{1+TK_1\exp\{TK_1\}}{K_2}$.

Bringing (3.15) and (3.12) into (3.14) gives

$$
\begin{aligned}
&\big(1-\epsilon_6\big)\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{X}(t)\|_{L^2}^2\Big] + K_2\mathbb{E}\Big[\int_0^T \|\widehat{X}\|_{H^1}^2 \mathrm{d}s\Big] \\
\leqslant\ & K_4\mathbb{E}\Big[\|\widehat{\phi}\|_{L^2}^2\Big] + \frac{K_4}{\epsilon_1}\mathbb{E}\Big[\int_0^T \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s\Big] + \frac{K_4}{\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s\Big] \\
&+ K_4\left(1+\frac{1}{\epsilon_4}\right)\mathbb{E}\Big[\int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s\Big],
\end{aligned}
\tag{3.16}
$$

where $K_4 = K_4(T, c_1, L^f, L_1^\sigma, L_2^\sigma) = 1 + K_3 + K_1 T\exp\{TK_1\}$. Then, we choose $\epsilon_6 > 0$ small enough that $1-\epsilon_6 > 0$ holds, which means that (3.16) turns into the following form

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{X}(t)\|_{L^2}^2\Big] + \mathbb{E}\Big[\int_0^T \|\widehat{X}\|_{H^1}^2 \mathrm{d}s\Big] \\
\leqslant\ & K_5\mathbb{E}\Big[\|\widehat{\phi}\|_{L^2}^2 + \int_0^T \|\widehat{\theta}\|_{L^2}^2\mathrm{d}s + \int_0^T \big\|\widehat{f}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s + \int_0^T \big\|\widehat{\sigma}(s,\bar{X},\nabla\bar{X})\big\|_{L^2}^2\mathrm{d}s\Big],
\end{aligned}
\tag{3.17}
$$

where $K_5 = K_5(T, c_1, L^f, L_1^\sigma, L_2^\sigma) = \frac{K_4\max\{\frac{1}{\epsilon_1},\frac{1}{\epsilon_3},1+\frac{1}{\epsilon_4}\}}{\min\{1-\epsilon_6, K_2\}}$. Thus, the proof of the desired estimate (3.3) is completed. Next, we prove the estimate (3.4) simply by making $(A,\bar{\theta},\bar{f},\bar{\sigma},\bar{\phi}) = (A,0,0,0,0)$.

□

Lemma 3.2 (Existence and Uniqueness of SPDE) *For any generator (A,θ,f,σ,ϕ) satisfying Assumption 3.1 and the condition $2c_1 > |L_2^\sigma|^2$, SPDE (3.1) has a unique solution $X(\cdot,\cdot)\in\mathscr{R}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R})$.*

Proof For relevant proofs, we can refer to [56] and omit the detailed proofs. □

Secondly, we consider the BSPDE as follows:

$$
\begin{cases}
\mathrm{d}Y(t,x) = -\Big[\nabla\cdot\big(E\nabla Y(t,x) + FZ(t,x) + \vartheta(t,x)\big) + b(t,x,Y(t,x),\nabla Y(t,x),Z(t,x))\Big]\mathrm{d}t \\
\qquad\qquad + \Big\langle Z(t,x),\mathrm{d}W(t)\Big\rangle, \\
Y(T,x) = \psi(x),\ (t,x)\in[0,T]\times\mathbb{R}^n.
\end{cases}
\tag{3.18}
$$

Furthermore, the coefficients (E,F,ϑ,b,ψ) are assumed to satisfy the following assumptions:

Assumption 3.2 (i) *The random mappings $E:[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{S}^n$ is $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)$-measurable and $F:[0,T]\times\Omega\to\mathbb{R}^{n\times d}$ is* $\mathscr{P}$-measurable. *Suppose that E,F are uniformly bounded. Assume there exists a constant $c_2>0$ such that*

$$
E - FF^\top \geqslant c_2 I,
$$

where I is an identity matrix with the $n\times n$ dimension.

(ii) *The mapping*

$$\vartheta(t,\omega,x):[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{R}^n$$

is $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)$*-measurable* .

(iii) *The mappings*

$$b(t,\omega,x,\gamma_1,\gamma_2,\gamma_3):[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^d\to\mathbb{R}$$

is $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R}^d)$*-measurable. There exist constants* $L^b>0$*, such that for all* $(\gamma_1,\gamma_2,\gamma_3),(\bar\gamma_1,\bar\gamma_2,\bar\gamma_3)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^d,\forall(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$,

$$\big|b(t,\omega,x,\gamma_1,\gamma_2,\gamma_3)-b(t,\omega,x,\bar\gamma_1,\bar\gamma_2,\bar\gamma_3)\big|\leqslant L^b\Big(\big|\gamma_1-\bar\gamma_1\big|+\big|\gamma_2-\bar\gamma_2\big|+\big|\gamma_3-\bar\gamma_3\big|\Big).$$

(iv) $\psi(\cdot,\cdot)\in L^2_{\mathcal{F}_T}(\Omega;L^2(\mathbb{R}^n;\mathbb{R}))$, $\vartheta(\cdot,\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^n)))$, $b(\cdot,\cdot,\cdot,0,0,0)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R})))$.

Before discussing the a priori estimation of the BSPDE, we present the definition of a weak solution.

Definition 3.2 *The pair* $(Y(\cdot,\cdot),Z(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{1+d})$ *is called a weak solution to the BSPDE (3.18), if for each* $\eta\in H^1(\mathbb{R}^n;\mathbb{R})$ *and almost every* $(t,\omega)\in[0,T]\times\Omega$*, it holds that*

$$\begin{aligned}\Big\langle Y(t,\cdot),\eta(\cdot)\Big\rangle_{L^2}&=\Big\langle\psi(\cdot),\eta(\cdot)\Big\rangle_{L^2}-\int_t^T\Big\langle E(s,\cdot)\nabla Y(s,\cdot)+F(s)Z(s,\cdot)+\vartheta(s,\cdot),\nabla\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s\\&\quad+\int_t^T\Big\langle b(s,\cdot,Y(s,\cdot),\nabla Y(s,\cdot),Z(s,\cdot)),\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s-\int_t^T\Big\langle Z(s,\cdot),\eta(\cdot)\mathrm{d}W(s)\Big\rangle_{L^2}.\end{aligned}\tag{3.19}$$

Lemma 3.3 (Continuous Dependence Theorem of BSPDE) *Let Assumption 3.2 hold. Suppose that* (E,F,ϑ,b,ψ) *and* $(E,F,\bar\vartheta,\bar b,\bar\psi)$ *are two sets of coefficients for the BSPDE* (3.18)*. Moreover, we assume that* $(Y(\cdot,\cdot),Z(\cdot,\cdot)),(\bar Y(\cdot,\cdot),\bar Z(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{1+d})$ *are the solutions of the BSPDE* (3.18) *corresponding to* (E,F,ϑ,b,ψ) *and* $(E,F,\bar\vartheta,\bar b,\bar\psi)$ *, respectively. Then the following estimate holds:*

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]}\|Y(t)-\bar Y(t)\|^2_{L^2}\Big]+\mathbb{E}\Big[\int_0^T\|Y(t)-\bar Y(t)\|^2_{H^1}\mathrm{d}t\Big]+\mathbb{E}\Big[\int_0^T\|Z(t)-\bar Z(t)\|^2_{L^2}\mathrm{d}t\Big]\\&\leqslant K\mathbb{E}\Big[\|\psi-\bar\psi\|^2_{L^2}+\int_0^T\big\|\vartheta(t)-\bar\vartheta(t)\big\|^2_{L^2}\mathrm{d}t\\&\quad+\int_0^T\|b(t,\bar Y(t),\nabla\bar Y(t),\bar Z(t))-\bar b(t,\bar Y(t),\nabla\bar Y(t),\bar Z(t))\|^2_{L^2}\mathrm{d}t\Big],\end{aligned}\tag{3.20}$$

where K *is a positive constant depending on* c_2,T *and the Lipschitz constant* L^b *of the mapping* b*. Furthermore, if* $(E,F,\bar\vartheta,\bar b,\bar\psi)=(E,F,0,0,0)$ *, we have the following estimate:*

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\in[0,T]}\|Y(t)\|^2_{L^2}\Big]+\mathbb{E}\Big[\int_0^T\|Y(t)\|^2_{H^1}\mathrm{d}t\Big]+\mathbb{E}\Big[\int_0^T\|Z(t)\|^2_{L^2}\mathrm{d}t\Big]\\&\leqslant K\mathbb{E}\Big[\|\psi\|^2_{L^2}+\int_0^T\big\|\vartheta(t)\big\|^2_{L^2}\mathrm{d}t+\int_0^T\big\|b(t,0,0,0)\big\|^2_{L^2}\mathrm{d}t\Big],\end{aligned}\tag{3.21}$$

where the positive constant K *is similar to the previous one.*

Proof Similarly, to simplify our notation, we denote by

$$\begin{cases}\widehat{Y} = Y(t,x) - \bar{Y}(t,x), \quad \widehat{Z} = Z(t,x) - \bar{Z}(t,x), \\ \widehat{\vartheta} = \vartheta(t,x) - \bar{\vartheta}(t,x), \quad \widehat{\psi} = \psi(x) - \bar{\psi}(x), \\ \widehat{b}(t,\bar{Y},\nabla\bar{Y},\bar{Z}) = b(t,\bar{Y},\nabla\bar{Y},\bar{Z}) - \bar{b}(t,\bar{Y},\nabla\bar{Y},\bar{Z}).\end{cases}$$

Applying Itô's formula to $\|\widehat{Y}(s)\|_{L^2}^2$ first and using Assumption 3.2, we obtain

$$\begin{aligned}&\|\widehat{Y}(t)\|_{L^2}^2 + \int_t^T \|\widehat{Z}\|_{L^2}^2 \mathrm{d}s \\ &= \big\|\widehat{\psi}\big\|_{L^2}^2 - 2\int_t^T \langle E\nabla\widehat{Y}, \nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle Z, F^\top\nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle\vartheta, \nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s \\ &\quad + 2\int_t^T \langle b(s,Y,\nabla Y,Z) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z}), \widehat{Y}\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s. \end{aligned} \tag{3.22}$$

It is easy to verify that

$$\begin{aligned}&-2\int_t^T \langle E\nabla\widehat{Y}, \nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle Z, F^\top\nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle\vartheta, \nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s \\ &\leqslant -2\int_t^T \langle\big(E - FF^\top\big)\nabla\widehat{Y}, \nabla\widehat{Y}\rangle_{L^2}\mathrm{d}s + \frac{1}{2}\int_t^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s + \epsilon_1\int_t^T \|\nabla\widehat{Y}\|^2\mathrm{d}s + \frac{1}{\epsilon_1}\int_t^T \|\vartheta\|_{L^2}^2\mathrm{d}s \\ &\leqslant -\big(2c_2 - \epsilon_1\big)\int_t^T \Big(\|\widehat{Y}\|_{H^1}^2 - \|\widehat{Y}\|_{L^2}^2\Big)\mathrm{d}s + \frac{1}{2}\int_t^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s + \frac{1}{\epsilon_1}\int_t^T \|\vartheta\|_{L^2}^2\mathrm{d}s. \end{aligned} \tag{3.23}$$

Substituting (3.23) into (3.22), we derive

$$\begin{aligned}&\|\widehat{Y}(t)\|_{L^2}^2 + \big(2c_2 - \epsilon_1\big)\int_t^T \|\widehat{Y}\|_{H^1}^2\mathrm{d}s + \frac{1}{2}\int_t^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s \\ &\leqslant \big\|\widehat{\psi}\big\|_{L^2}^2 + \big(2c_2 - \epsilon_1\big)\int_t^T \|\widehat{Y}\|_{L^2}^2\mathrm{d}s + \frac{1}{\epsilon_1}\int_t^T \|\vartheta\|_{L^2}^2\mathrm{d}s \\ &\quad + 2\int_t^T \langle\widehat{Y}, b(s,Y,\nabla Y,Z) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\rangle_{L^2}\mathrm{d}s - 2\int_t^T \langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s. \end{aligned} \tag{3.24}$$

Then, by taking expectation on both sides of (3.24), we obtain

$$\begin{aligned}&\mathbb{E}\Big[\|\widehat{Y}(t)\|_{L^2}^2\Big] + \big(2c_2 - \epsilon_1\big)\mathbb{E}\Big[\int_t^T \|\widehat{Y}\|_{H^1}^2\mathrm{d}s\Big] + \frac{1}{2}\mathbb{E}\Big[\int_t^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big] \\ &\leqslant \mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big] + \big(2c_2 - \epsilon_1\big)\mathbb{E}\Big[\int_t^T \|\widehat{Y}\|_{L^2}^2\mathrm{d}s\Big] + \frac{1}{\epsilon_1}\mathbb{E}\Big[\int_t^T \|\vartheta\|_{L^2}^2\mathrm{d}s\Big] \\ &\quad + 2\mathbb{E}\Big[\int_t^T \langle\widehat{Y}, b(s,Y,\nabla Y,Z) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\rangle_{L^2}\mathrm{d}s\Big]. \end{aligned} \tag{3.25}$$

With the Lipschitz condition in Assumption 3.2 and Young's inequality $ab \leqslant \epsilon a^2 + \frac{1}{4\epsilon}b^2, \forall\epsilon > 0$, we obtain

$$
\begin{aligned}
& 2\int_t^T \left|\langle \widehat{Y}, b(s,Y,\nabla Y,Z) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\rangle_{L^2}\right| \mathrm{d}s \\
\leqslant & 2\int_t^T \big\|\widehat{Y}\big\|_{L^2} \big\|b(s,Y,\nabla Y,Z) - b(s,\bar{Y},\nabla\bar{Y},\bar{Z}) + b(s,\bar{Y},\nabla\bar{Y},\bar{Z}) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2} \mathrm{d}s \\
\leqslant & 2\int_t^T \big\|\widehat{Y}\big\|_{L^2} \big\|b(s,Y,\nabla Y,Z) - b(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2} \mathrm{d}s \\
& + 2\int_t^T \big\|\widehat{Y}\big\|_{L^2} \big\|b(s,\bar{Y},\nabla\bar{Y},\bar{Z}) - \bar{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2} \mathrm{d}s \\
\leqslant & 2\big|L^b\big| \int_t^T \Big(\big\|\widehat{Y}\big\|_{L^2}^2 + \big\|\widehat{Y}\big\|_{L^2}\big\|\nabla\widehat{Y}\big\| + \big\|\widehat{Y}\big\|_{L^2}\big\|\widehat{Z}\big\|_{L^2}\Big)\mathrm{d}s + 2\int_t^T \big\|\widehat{Y}\big\|_{L^2}\big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}\mathrm{d}s \\
\leqslant & \left(2\big|L^b\big| + \frac{\big|L^b\big|^2}{\epsilon_2} + \frac{\big|L^b\big|^2}{\epsilon_3} + \epsilon_4\right)\int_t^T \big\|\widehat{Y}\big\|_{L^2}^2\mathrm{d}s + \epsilon_2\int_t^T \big\|\nabla\widehat{Y}\big\|^2\mathrm{d}s + \epsilon_3\int_t^T \big\|\widehat{Z}\big\|_{L^2}^2\mathrm{d}s \\
& + \frac{1}{\epsilon_4}\int_t^T \big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s \\
= & \left(2\big|L^b\big| + \frac{\big|L^b\big|^2}{\epsilon_2} + \frac{\big|L^b\big|^2}{\epsilon_3} + \epsilon_4 - \epsilon_2\right)\int_t^T \big\|\widehat{Y}\big\|_{L^2}^2\mathrm{d}s + \epsilon_2\int_t^T \big\|\widehat{Y}\big\|_{H^1}^2\mathrm{d}s \\
& + \epsilon_3\int_t^T \big\|\widehat{Z}\big\|_{L^2}^2\mathrm{d}s + \frac{1}{\epsilon_4}\int_t^T \big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s. \qquad (3.26)
\end{aligned}
$$

Combining (3.25)–(3.26), we have

$$
\begin{aligned}
& \mathbb{E}\Big[\|\widehat{Y}(t)\|_{L^2}^2\Big] + \big(2c_2 - \epsilon_1 - \epsilon_2\big)\mathbb{E}\Big[\int_t^T \|\widehat{Y}\|_{H^1}^2\mathrm{d}s\Big] + \big(\frac{1}{2} - \epsilon_3\big)\mathbb{E}\Big[\int_t^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big] \\
\leqslant & \mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big] + \Big(2c_2 - \epsilon_1 + 2\big|L^b\big| + \frac{\big|L^b\big|^2}{\epsilon_2} + \frac{\big|L^b\big|^2}{\epsilon_3} + \epsilon_4 - \epsilon_2\Big)\mathbb{E}\Big[\int_t^T \|\widehat{Y}\|_{L^2}^2\mathrm{d}s\Big] \\
& + \frac{1}{\epsilon_4}\mathbb{E}\Big[\int_t^T \big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big] + \frac{1}{\epsilon_1}\mathbb{E}\Big[\int_t^T \|\vartheta\|_{L^2}^2\mathrm{d}s\Big]. \qquad (3.27)
\end{aligned}
$$

For any $\epsilon_4 > 0$, select $\epsilon_1, \epsilon_2, \epsilon_3 > 0$ small enough such that $2c_2 - \epsilon_1 - \epsilon_2 > 0$ and $\frac{1}{2} - \epsilon_3 > 0$, then applying Gronwall's inequality, we obtain

$$
\begin{aligned}
& \sup_{t\in[0,T]} \mathbb{E}\Big[\|\widehat{Y}(t)\|_{L^2}^2\Big] \\
\leqslant & \exp\{TK_1\}\mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big] + \frac{\exp\{TK_1\}}{\epsilon_4}\mathbb{E}\Big[\int_0^T \big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big] + \frac{\exp\{TK_1\}}{\epsilon_1}\mathbb{E}\Big[\int_0^T \|\vartheta\|_{L^2}^2\mathrm{d}s\Big], \qquad (3.28)
\end{aligned}
$$

where $K_1 = K_1(c_2, |L^b|) = 2c_2 - \epsilon_1 + 2\big|L^b\big| + \frac{|L^b|^2}{\epsilon_2} + \frac{|L^b|^2}{\epsilon_3} + \epsilon_4 - \epsilon_2$. Furthermore, according to (3.27) and (3.28), we obtain

$$
\mathbb{E}\Big[\int_0^T \|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big] \leqslant K_2\mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big] + \frac{K_2}{\epsilon_4}\mathbb{E}\Big[\int_0^T \big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big] + \frac{K_2}{\epsilon_1}\mathbb{E}\Big[\int_0^T \|\vartheta\|_{L^2}^2\mathrm{d}s\Big], \qquad (3.29)
$$

where $K_2 = K_2(T, c_2, |L^b|) = \frac{1 + K_1 T\exp\{TK_1\}}{\frac{1}{2} - \epsilon_3}$.

Combining (3.24) and (3.26), then taking supremum on both sides and expectation yields

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{Y}(t)\|_{L^2}^2\Big]+\big(2c_2-\epsilon_1-\epsilon_2\big)\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{H^1}^2\mathrm{d}s\Big]+\big(\frac{1}{2}-\epsilon_3\big)\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big]\\
\leqslant&\ \mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big]+K_1\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{L^2}^2\mathrm{d}s\Big]+\frac{1}{\epsilon_4}\mathbb{E}\Big[\int_0^T\big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big]+\frac{1}{\epsilon_1}\mathbb{E}\Big[\int_0^T\|\vartheta\|_{L^2}^2\mathrm{d}s\Big]\\
&+2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_t^T\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|\Big].
\end{aligned}\tag{3.30}$$

Subsequently, we deal with the stochastic integrals in (3.30) by means of the Burkholder-Davis-Gundy inequality and the Lipschitz continuity condition together with (3.29), and obtain the following result

$$\begin{aligned}
&2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_t^T\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|\Big]=2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^T\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s-\int_0^t\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|\Big]\\
\leqslant&\ 2\mathbb{E}\Big[\Big|\int_0^T\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|+\sup_{t\in[0,T]}\Big|\int_0^t\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|\Big]\\
\leqslant&\ 2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|+\sup_{t\in[0,T]}\Big|\int_0^t\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\mathrm{d}W_s\Big|\Big]\\
\leqslant&\ 16\sqrt{2}\mathbb{E}\Big[\int_0^T\Big|\langle\widehat{Y},\widehat{Z}\rangle_{L^2}\Big|^2\mathrm{d}s\Big]^{\frac{1}{2}}\leqslant16\sqrt{2}\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|\widehat{Y}(t)\big\|_{L^2}^2\int_0^T\big\|\widehat{Z}\big\|_{L^2}^2\mathrm{d}s\Big]^{\frac{1}{2}}\\
\leqslant&\ \epsilon_5\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|\widehat{Y}(t)\big\|_{L^2}^2\Big]+\frac{128}{\epsilon_5}\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big]\\
\leqslant&\ \epsilon_5\mathbb{E}\Big[\sup_{t\in[0,T]}\big\|\widehat{Y}(t)\big\|_{L^2}^2\Big]+\frac{128K_2}{\epsilon_5}\mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big]+\frac{128K_2}{\epsilon_4\epsilon_5}\mathbb{E}\Big[\int_0^T\big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big]\\
&+\frac{128K_2}{\epsilon_1\epsilon_5}\mathbb{E}\Big[\int_0^T\|\vartheta\|_{L^2}^2\mathrm{d}s\Big].
\end{aligned}\tag{3.31}$$

Substituting (3.28) and (3.31) into (3.30), we obtain

$$\begin{aligned}
&\big(1-\epsilon_5\big)\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{Y}(t)\|_{L^2}^2\Big]+\big(2c_2-\epsilon_1-\epsilon_2\big)\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{H^1}^2\mathrm{d}s\Big]+\Big(\frac{1}{2}-\epsilon_3\Big)\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big]\\
\leqslant&\ \Big(1+\frac{128K_2}{\epsilon_5}+TK_1\exp\{TK_1\}\Big)\mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big]\\
&+\frac{1}{\epsilon_4}\Big(1+\frac{128K_2}{\epsilon_5}+TK_1\exp\{TK_1\}\Big)\mathbb{E}\Big[\int_0^T\big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big]\\
&+\frac{1}{\epsilon_1}\Big(1+\frac{128K_2}{\epsilon_5}+TK_1\exp\{TK_1\}\Big)\mathbb{E}\Big[\int_0^T\|\vartheta\|_{L^2}^2\mathrm{d}s\Big],
\end{aligned}\tag{3.32}$$

which implies

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{Y}(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{H^1}^2\mathrm{d}s\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}s\Big]\\
\leqslant&\ K_3\Big\{\mathbb{E}\Big[\big\|\widehat{\psi}\big\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\big\|\widehat{b}(s,\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\mathrm{d}s\Big]+\mathbb{E}\Big[\int_0^T\|\vartheta\|_{L^2}^2\mathrm{d}s\Big]\Big\},
\end{aligned}\tag{3.33}$$

where $K_3 = K_3(T, c_2, |L^b|) = \frac{\max\{1, \frac{1}{\epsilon_1}, \frac{1}{\epsilon_4}\}\left(1+\frac{128K_2}{\epsilon_5}+TK_1\exp\{TK_1\}\right)}{\min\{1-\epsilon_5, 2c_2-\epsilon_1-\epsilon_2, \frac{1}{2}-\epsilon_3\}}$. Hence, we have completed the proof of estimate (3.20), and next, it is convenient to prove estimate (3.21) simply by making $(E, F, \bar{\vartheta}, \bar{b}, \bar{\psi}) = (E, F, 0, 0, 0)$.

□

Lemma 3.4 (Existence and Uniqueness of BSPDE) *For any generator* $(E, F, \vartheta, b, \psi)$ *satisfying Assumption 3.2, BSPDE* (3.18) *admits a unique solution* $(Y(\cdot,\cdot), Z(\cdot,\cdot)) \in \mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n; \mathbb{R}^{1+d})$.

Proof Similarly, this proof can be found in [18, 31, 56]; so, we omit the detailed proofs. □

4. FBSPDE

In this section, we study FBSPDE (1.1). Revisiting the FBSPDEs we studied:

$$\begin{cases} \mathrm{d}X(t,x) = \Big[\mathfrak{L}_F(t,x,X(t,x),Y(t,x),Z(t,x)) + f(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x), \\ \qquad\qquad Z(t,x))\Big]\mathrm{d}t + \Big\langle \sigma(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x),Z(t,x)), \mathrm{d}W(t)\Big\rangle, \\ \mathrm{d}Y(t,x) = -\Big[\mathfrak{L}_B(t,x,X(t,x),Y(t,x),Z(t,x)) + b(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x), \\ \qquad\qquad Z(t,x))\Big]\mathrm{d}t + \Big\langle Z(t,x), \mathrm{d}W(t)\Big\rangle, \quad (t,x) \in [0,T]\times\mathbb{R}^n, \\ X(0,x) = \phi(x, Y_0), \quad Y(T,x) = \psi(x, X_T), \quad x \in \mathbb{R}^n, \end{cases} \tag{4.1}$$

where we denote

$$\begin{cases} \mathfrak{L}_F(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq \nabla\cdot\big(A\nabla X(t,x) + B\nabla Y(t,x) + CZ(t,x)\big), \\ \mathfrak{L}_B(t,x,X(t,x),Y(t,x),Z(t,x)) \triangleq \nabla\cdot\big(-D\nabla X(t,x) + E\nabla Y(t,x) + FZ(t,x)\big). \end{cases} \tag{4.2}$$

Similar to the cases of SPDE (3.1) and BSPDE (3.18), we still have to make the following assumptions for the coefficients $(A, B, C, D, E, F, f, \sigma, b, \phi, \psi)$ of FBSPDE (1.1).

Assumption 4.1 (i) *The random mappings* $A, B, D, E : [0,T]\times\Omega\times\mathbb{R}^n \to \mathbb{S}^n$ *is* $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)$ *-measurable and* $C, F : [0,T]\times\Omega \to \mathbb{R}^{n\times d}$ *is* $\mathscr{P}$*-measurable. Suppose that* A, B, C, D, E, F *are uniformly bounded. Assumer there exists constants* $c_0, c_1, c_2, C_1, C_2, C_3 > 0$ *such that*

$$\begin{pmatrix} AA & 0 \\ 0 & EE \end{pmatrix} \leqslant C_1 I, \quad \begin{pmatrix} BB & 0 \\ 0 & DD \end{pmatrix} \leqslant C_2 I, \quad \begin{pmatrix} CC^\top & 0 \\ 0 & FF^\top \end{pmatrix} \leqslant C_3 I,$$

and

$$\begin{pmatrix} B - CC^\top & 0 \\ 0 & D - FF^\top \end{pmatrix} \geqslant c_0 I, \quad A \geqslant c_1 I, \quad E - FF^\top \geqslant c_2 I,$$

where I *is an identity matrix with the* $n\times n$ *dimension.*

(ii) *The mappings*

$$f(t,\omega,x,\gamma) : [0,T]\times\Omega\times\mathbb{R}^n\times\mathfrak{R} \to \mathbb{R},$$
$$b(t,\omega,x,\gamma) : [0,T]\times\Omega\times\mathbb{R}^n\times\mathfrak{R} \to \mathbb{R},$$
$$\sigma(t,\omega,x,\gamma) : [0,T]\times\Omega\times\mathbb{R}^n\times\mathfrak{R} \to \mathbb{R}^d$$

are $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R}^d)$*-measurable, where* $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5) \in \mathfrak{R}$ *and* $\mathfrak{R} \triangleq \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^d$. *There exist constants* $L^f, L^\sigma_1, L^\sigma_2, L^b > 0$, *such that for all* $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5), \bar{\gamma} = (\bar{\gamma}_1, \bar{\gamma}_2, \bar{\gamma}_3, \bar{\gamma}_4, \bar{\gamma}_5) \in \mathfrak{R}$, $\forall (t,\omega,x) \in [0,T]\times\Omega\times\mathbb{R}^n$,

$$\begin{aligned}
&\big|f(t,\omega,x,\gamma)-f(t,\omega,x,\bar{\gamma})\big| \leqslant L^f\Big(\big|\widehat{\gamma}_1\big|+\big|\widehat{\gamma}_2\big|+\big|\widehat{\gamma}_3\big|+\big|\widehat{\gamma}_4\big|+\big|\widehat{\gamma}_5\big|\Big),\\
&\big|\sigma(t,\omega,x,\gamma)-\sigma(t,\omega,x,\bar{\gamma})\big| \leqslant L_1^\sigma\Big(\big|\widehat{\gamma}_1\big|+\big|\widehat{\gamma}_3\big|+\big|\widehat{\gamma}_4\big|+\big|\widehat{\gamma}_5\big|\Big)+L_2^\sigma\big|\widehat{\gamma}_2\big|,\\
&\big|b(t,\omega,x,\gamma)-b(t,\omega,x,\bar{\gamma})\big| \leqslant L^b\Big(\big|\widehat{\gamma}_1\big|+\big|\widehat{\gamma}_2\big|+\big|\widehat{\gamma}_3\big|+\big|\widehat{\gamma}_4\big|+\big|\widehat{\gamma}_5\big|\Big),
\end{aligned}$$

where $\widehat{\gamma}_i=\gamma_i-\bar{\gamma}_i$ *with* $i=1,2,3,4,5$.

(iii) *The mappings*

$$\begin{aligned}
&\phi(\omega,x,\gamma_3):\Omega\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R},\\
&\psi(\omega,x,\gamma_1):\Omega\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}
\end{aligned}$$

are $\mathcal{F}_0\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})$*-measurable and* $\mathcal{F}_T\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})$*-measurable, respectively. There exist constants* $L^\phi, L^\psi>0$*, such that for all* $\gamma_1,\gamma_3,\bar{\gamma}_1,\bar{\gamma}_3\in\mathbb{R}$, $\forall(\omega,x)\in\Omega\times\mathbb{R}^n$,

$$\begin{aligned}
&\big|\phi(\omega,x,\gamma_3)-\phi(\omega,x,\bar{\gamma}_3)\big|\leqslant L^\phi\big|\widehat{\gamma}_3\big|,\\
&\big|\psi(\omega,x,\gamma_1)-\psi(\omega,x,\bar{\gamma}_1)\big|\leqslant L^\psi\big|\widehat{\gamma}_1\big|,
\end{aligned}$$

where $\widehat{\gamma}_i=\gamma_i-\bar{\gamma}_i$ *with* $i=1,3$.

(iv) $\phi(\cdot,0)\in L^2_{\mathcal{F}_0}(\Omega;H^1(\mathbb{R}^n;\mathbb{R}))$, $\psi(\cdot,T)\in L^2_{\mathcal{F}_T}(\Omega;H^1(\mathbb{R}^n;\mathbb{R}))$, $\vartheta(\cdot,\cdot,\cdot)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^n)))$, $f(\cdot,\cdot,\cdot,0,0,0,0,0), b(\cdot,\cdot,\cdot,0,0,0,0,0)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R})))$ $\sigma(\cdot,\cdot,\cdot,0,0,0,0,0)\in L^2_{\mathbb{F}}(\Omega;L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^d)))$.

(v) *For simplicity of notation, define*

$$\Lambda(t,x,\gamma):=\begin{pmatrix}-b(t,x,\gamma)\\ f(t,x,\gamma)\\ \sigma(t,x,\gamma)\end{pmatrix},\quad \gamma:=\begin{pmatrix}\gamma_1\\ \gamma_2\\ \gamma_3\\ \gamma_4\\ \gamma_5\end{pmatrix},\quad \gamma':=\begin{pmatrix}\gamma_1\\ \gamma_3\\ \gamma_5\end{pmatrix},$$

and

$$\Big[\Lambda(t,x,\gamma),\gamma'\Big]:=-b(t,x,\gamma)\gamma_1+f(t,x,\gamma)\gamma_3+\big\langle\sigma(t,x,\gamma),\gamma_5\big\rangle.$$

There exist four constants $\beta_1\geqslant 0,\beta_2\geqslant\frac{1}{2},\mu_1\geqslant 0,\mu_2\geqslant 0$ *such that the following conditions hold:*

(a) *One of the following two cases holds true. Case 1:* $\beta_1>0,\mu_1>0,\beta_2\geqslant\frac{1}{2},\mu_2\geqslant 0$. *Case 2:* $\beta_2>\frac{1}{2},\mu_2>0,\beta_1\geqslant 0,\mu_1\geqslant 0$.

(b) *For each* $\gamma=(\gamma_1,\gamma_2,\gamma_3,\gamma_4,\gamma_5),\bar{\gamma}=(\bar{\gamma}_1,\bar{\gamma}_2,\bar{\gamma}_3,\bar{\gamma}_4,\bar{\gamma}_5)\in\mathfrak{R}$, $\gamma'=(\gamma_1,\gamma_3,\gamma_5),\bar{\gamma}'=(\bar{\gamma}_1,\bar{\gamma}_3,\bar{\gamma}_5)\in\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d,\forall(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$,

$$\begin{cases}
\Big[\Lambda(t,x,\gamma)-\Lambda(t,x,\bar{\gamma}),\widehat{\gamma}'\Big]\leqslant-\beta_1\big|\widehat{\gamma}_1\big|^2-\beta_2\Big[\big|\widehat{\gamma}_3\big|^2+\big|\widehat{\gamma}_5\big|^2\Big],\\
\big(\psi(x,\gamma_1)-\psi(x,\bar{\gamma}_1)\big)\widehat{\gamma}_1\geqslant\mu_1\big|\widehat{\gamma}_1\big|^2,\\
\big(\phi(x,\gamma_3)-\phi(x,\bar{\gamma}_3)\big)\widehat{\gamma}_3\leqslant-\mu_2\big|\widehat{\gamma}_3\big|^2,
\end{cases}$$

where $\widehat{\gamma}_i=\gamma_i-\bar{\gamma}_i$ *with* $i=1,2,3,4,5$.

If the coefficients $(A,B,C,D,E,F,f,\sigma,b,\phi,\psi)$ satisfy Assumption 4.1, it is called a generator of FBSPDE. Before discussing the a priori estimation of the FBSPDE, we present the definition of a weak solution.

Definition 4.1 *The triple* $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$ *is called a weak solution to the FBSPDE (1.1) if for each* $\eta\in H^1(\mathbb{R}^n;\mathbb{R})$ *and almost every* $(t,\omega)\in[0,T]\times\Omega$*, it holds that*

$$\Big\langle X(t,\cdot),\eta(\cdot)\Big\rangle_{L^2}=\Big\langle \phi(\cdot),\eta(\cdot)\Big\rangle_{L^2}-\int_0^t \Big\langle A(s,\cdot)\nabla X(s,\cdot)+B(s,\cdot)\nabla Y(s,\cdot)+C(s)Z(s,\cdot),\nabla\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s$$
$$+\int_0^t \Big\langle f(s,\cdot,X(s,\cdot),\nabla X(s,\cdot),Y(s,\cdot),\nabla Y(s,\cdot),Z(s,\cdot)),\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s$$
$$+\int_0^t \Big\langle \sigma(s,\cdot,X(s,\cdot),\nabla X(s,\cdot),Y(s,\cdot),\nabla Y(s,\cdot),Z(s,\cdot)),\eta(\cdot)\mathrm{d}W(s)\Big\rangle_{L^2} \tag{4.3}$$

and

$$\Big\langle Y(t,\cdot),\eta(\cdot)\Big\rangle_{L^2}=\Big\langle \psi(\cdot),\eta(\cdot)\Big\rangle_{L^2}-\int_t^T \Big\langle D(s,\cdot)\nabla X(s,\cdot)+E(s,\cdot)\nabla Y(s,\cdot)+F(s)Z(s,\cdot),\nabla\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s$$
$$+\int_t^T \Big\langle b(s,\cdot,X(s,\cdot),\nabla X(s,\cdot),Y(s,\cdot),\nabla Y(s,\cdot),Z(s,\cdot)),\eta(\cdot)\Big\rangle_{L^2}\mathrm{d}s$$
$$-\int_t^T \Big\langle Z(s,\cdot),\eta(\cdot)\mathrm{d}W(s)\Big\rangle_{L^2}. \tag{4.4}$$

Lemma 4.1 (Existence and Uniqueness of FBSPDE) *Let Assumption 4.1,* $c_0 \geqslant \sqrt{C_1}$ *and* $2c_1 > |L_2^\sigma|^2$ *be satisfied. For any generator* $(A,B,C,D,E,F,f,\sigma,b,\phi,\psi)$ *, FBSPDEs* (1.1) *has a unique solution* $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot)\big)\in\mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$.

Proof In fact, Yin [56] has studied this problem and gave a detailed proof, and the conditions we propose satisfy those in [56, Theorem 4.4]. Therefore, we do not repeat it here. □

Theorem 4.1 (Continuous Dependence Theorem of FBSPDE) *Let Assumption 4.1 hold. Additionally, we assume that* $c_0 \geqslant \sqrt{C_1}$ *and* $2c_1 > |L_2^\sigma|^2$. *Suppose that* $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot)\big)$, $\big(\bar X(\cdot,\cdot),\bar Y(\cdot,\cdot),\bar Z(\cdot,\cdot)\big)\in\mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$ *are the solutions of the FBSPDEs* (1.1) *corresponding to two given generators* $(A,B,C,D,E,F,f,\sigma,b,\phi,\psi)$ *and* $(A,B,C,D,E,F,\bar f,\bar\sigma,\bar b,\bar\phi,\bar\psi)$ *, respectively. Then, we have the following estimate:*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat X(t)\|_{L^2}^2+\sup_{t\in[0,T]}\|\widehat Y(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\Big(\|\widehat X(t)\|_{H^1}^2+\|\widehat Y(t)\|_{H^1}^2+\|\widehat Z(t)\|_{L^2}^2\Big)\mathrm{d}t\Big]$$
$$\leqslant K\mathbb{E}\Big[\int_0^T\Big(\big\|f(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))-\bar f(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))\big\|_{L^2}^2$$
$$+\big\|\sigma(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))-\bar\sigma(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))\big\|_{L^2}^2$$
$$+\big\|b(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))-\bar b(t,\bar X(t),\nabla\bar X(t),\bar Y(t),\nabla\bar Y(t),\bar Z(t))\big\|_{L^2}^2\Big)\mathrm{d}t\Big]$$
$$+\mathbb{E}\Big[\big\|\phi(\bar Y(0))-\bar\phi(\bar Y(0))\big\|_{L^2}^2+\big\|\psi(\bar X(T))-\bar\psi(\bar X(T))\big\|_{L^2}^2\Big], \tag{4.5}$$

where K *is a positive constant depending on* $c_1, C_1, c_2, T, \mu_1, \mu_2, \beta$ *and the Lipschitz constants* $L^f, L_1^\sigma, L_1^\sigma, L^b$. *Furthermore, if* $(A,B,C,D,E,F,\bar f,\bar\sigma,\bar b,\bar\phi,\bar\psi)=(A,B,C,D,E,F,0,0,0,0,0)$ *, we have the following estimate:*

$$\mathbb{E}\Big[\sup_{t\in[0,T]}\|X(t)\|_{L^2}^2+\sup_{t\in[0,T]}\|Y(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\Big(\|X(t)\|_{H^1}^2+\|Y(t)\|_{H^1}^2+\|Z(t)\|_{L^2}^2\Big)\mathrm{d}t\Big]$$
$$\leqslant K\mathbb{E}\Big[\int_0^T\Big(\big\|f(t,0,0,0,0,0)\big\|_{L^2}^2+\big\|\sigma(t,0,0,0,0,0)\big\|_{L^2}^2+\big\|b(t,0,0,0,0,0)\big\|_{L^2}^2\Big)\mathrm{d}t$$
$$+\big\|\phi(0)\big\|_{L^2}^2+\big\|\psi(0)\big\|_{L^2}^2\Big], \tag{4.6}$$

where the positive constant K is similar to the previous one.

Proof Similarly, we denote by

$$\begin{cases} \widehat{X}=X(t,x)-\bar{X}(t,x), \quad \widehat{Y}=Y(t,x)-\bar{Y}(t,x), \quad \widehat{Z}=Z(t,x)-\bar{Z}(t,x), \\ \widehat{\varpi}(t,X,\nabla X,Y,\nabla Y,Z)=\varpi(t,X,\nabla X,Y,\nabla Y,Z)-\bar{\varpi}(t,X,\nabla X,Y,\nabla Y,Z), \quad \varpi=f,\sigma,b, \\ \widehat{\phi}(Y_0)=\phi(Y_0)-\bar{\phi}(Y_0), \quad \widehat{\psi}(X_T)=\psi(X_T)-\bar{\psi}(X_T). \end{cases}$$

On the one hand, the continuous-dependence theorem of SPDE (see Lemma 3.1) leads to

$$\begin{aligned} &\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{X}(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{X}\|_{H^1}^2\mathrm{d}t\Big] \\ \leqslant{}& K\mathbb{E}\Big[\int_0^T\Big(\big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2+\big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2+\|\widehat{Y}\|_{L^2}^2 \\ &+\|\nabla\widehat{Y}\|^2+\|\widehat{Z}\|_{L^2}^2\Big)\mathrm{d}t+\|\widehat{\phi}(\bar{Y}_0)\|_{L^2}^2+\|\widehat{Y}_0\|_{L^2}^2\Big]. \end{aligned} \tag{4.7}$$

On the other hand, under Assumption 4.1, the continuous-dependence theorem of BSPDE (see Lemma 3.3) yields

$$\begin{aligned} &\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{Y}(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{H^1}^2\mathrm{d}t\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}t\Big] \\ \leqslant{}& K\mathbb{E}\Big[\int_0^T\Big(\big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2+\|\widehat{X}\|_{L^2}^2+\|\nabla\widehat{X}\|^2\Big)\mathrm{d}t+\|\widehat{\psi}(\bar{X}_T)\|_{L^2}^2+\|\widehat{X}_T\|_{L^2}^2\Big]. \end{aligned} \tag{4.8}$$

Furthermore, by applying Itô's formula and the definition of $\Lambda(t,x,\gamma)$ to $\Big\langle\widehat{X}(s),\widehat{Y}(s)\Big\rangle_{L^2}$, we have

$$\begin{aligned} &\mathbb{E}\Big[\langle\widehat{X}_T,\psi(X_T)-\bar{\psi}(\bar{X}_T)\rangle_{L^2}\Big] \\ ={}&\mathbb{E}\Big[\langle\widehat{Y}_0,\phi(Y_0)-\bar{\phi}(\bar{Y}_0)\rangle_{L^2}\Big]+\mathbb{E}\Big[\int_0^T\langle\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z}),\widehat{Z}\rangle_{L^2}\mathrm{d}t \\ &+\int_0^T\langle\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z}),\widehat{Y}\rangle_{L^2}\mathrm{d}t+\int_0^T\langle-\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z}),\widehat{X}\rangle_{L^2}\mathrm{d}t\Big] \\ &+\mathbb{E}\Big[\int_0^T\Big[\Lambda(t,X,\nabla X,Y,\nabla Y,Z)-\Lambda(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z}),(\widehat{X},\widehat{Y},\widehat{Z})^\top\Big]\mathrm{d}t\Big] \\ &+\mathbb{E}\Big[\int_0^T\langle\nabla\cdot\big(A\nabla\widehat{X}+B\nabla\widehat{Y}+C\widehat{Z}\big),\widehat{Y}\rangle_{H^{-1},H^1}\mathrm{d}t \\ &+\int_0^T\Big\langle-\Big(\nabla\cdot\big(-D\nabla\widehat{X}+E\nabla\widehat{Y}+F\widehat{Z}\big)\Big),\widehat{X}\Big\rangle_{H^{-1},H^1}\mathrm{d}t\Big]. \end{aligned} \tag{4.9}$$

It is easy to verify that

$$\begin{aligned} &\Big\langle\nabla\cdot\big(A\nabla\widehat{X}+B\nabla\widehat{Y}+C\widehat{Z}\big),\widehat{Y}\Big\rangle_{H^{-1},H^1}+\Big\langle-\Big(\nabla\cdot\big(-D\nabla\widehat{X}+E\nabla\widehat{Y}+F\widehat{Z}\big)\Big),\widehat{X}\Big\rangle_{H^{-1},H^1} \\ ={}&-\Big\langle A\nabla\widehat{X}+B\nabla\widehat{Y}+C\widehat{Z},\nabla\widehat{Y}\Big\rangle_{L^2}+\Big\langle-D\nabla\widehat{X}+E\nabla\widehat{Y}+F\widehat{Z},\nabla\widehat{X}\Big\rangle_{L^2} \\ \leqslant{}&\frac{1}{2\epsilon_1}\Big\langle AA\nabla\widehat{Y},\nabla\widehat{Y}\Big\rangle_{L^2}+\frac{\epsilon_1}{2}\big\|\nabla\widehat{X}\big\|^2+\frac{1}{2\epsilon_2}\Big\langle EE\nabla\widehat{Y},\nabla\widehat{Y}\Big\rangle_{L^2}+\frac{\epsilon_2}{2}\big\|\nabla\widehat{X}\big\|^2+\frac{1}{2}\big\|\widehat{Z}\big\|_{L^2}^2 \\ &-\Big\langle\big(B-CC^\top\big)\nabla\widehat{Y},\nabla\widehat{Y}\Big\rangle_{L^2}-\Big\langle\big(D-FF^\top\big)\nabla\widehat{X},\nabla\widehat{X}\Big\rangle_{L^2}. \end{aligned} \tag{4.10}$$

From the monotonicity condition and the elementary equality $2ab \leqslant \frac{1}{\epsilon}a^2 + \epsilon b^2, \forall \epsilon > 0$, and taking it here as $\epsilon_1 = \epsilon_2 = \sqrt{C_1}$, then we derive

$$\begin{aligned}
&(\mu_1 - \epsilon_5)\mathbb{E}\Big[\big\|\widehat{X}_T\big\|_{L^2}^2\Big] + \mu_2\mathbb{E}\Big[\big\|\widehat{Y}_0\big\|_{L^2}^2\Big] \\
&+ (\beta_1 - \epsilon_3)\mathbb{E}\Big[\int_0^T \big\|\widehat{X}\big\|_{L^2}^2 \mathrm{d}t\Big] + (c_0 - \sqrt{C_1})\mathbb{E}\Big[\int_0^T \big\|\nabla\widehat{X}\big\|^2 \mathrm{d}t\Big] \\
&+ \beta_2\mathbb{E}\Big[\int_0^T \big\|\widehat{Y}\big\|_{L^2}^2 \mathrm{d}t\Big] + (c_0 - \sqrt{C_1})\mathbb{E}\Big[\int_0^T \big\|\nabla\widehat{Y}\big\|^2 \mathrm{d}t\Big] + (\beta_2 - \frac{1}{2})\mathbb{E}\Big[\int_0^T \big\|\widehat{Z}\big\|_{L^2}^2 \mathrm{d}t\Big] \\
\leqslant &\frac{1}{4\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \mathrm{d}t\Big] + \frac{1}{4\epsilon_4}\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \\
&+ \big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \frac{1}{4\epsilon_5}\mathbb{E}\Big[\big\|\widehat{\psi}(\bar{X}_T)\big\|_{L^2}^2\Big] \\
&+ \frac{1}{4\epsilon_6}\mathbb{E}\Big[\big\|\widehat{\phi}(\bar{Y}_0)\big\|_{L^2}^2\Big] + \epsilon_4\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{Y}\big\|_{L^2}^2 + \big\|\widehat{Z}\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \epsilon_6\mathbb{E}\Big[\big\|\widehat{Y}_0\big\|_{L^2}^2\Big].
\end{aligned} \tag{4.11}$$

For the case $\beta_1 > 0, \mu_1 > 0, \beta_2 \geqslant \frac{1}{2}, \mu_2 \geqslant 0$, notice that $c_0 \geqslant \sqrt{C_1}$, where we take ϵ_3 and ϵ_5 sufficiently small to make $\beta_1 - \epsilon_3 > 0$ and $\mu_1 - \epsilon_5 > 0$. Then letting $K_1 = \max\{\beta_1 - \epsilon_3, c_0 - \sqrt{C_1}\}$, we get

$$\begin{aligned}
&(\mu_1 - \epsilon_5)\mathbb{E}\big\|\widehat{X}_T\big\|_{L^2}^2 + K_1\mathbb{E}\int_0^T \big\|\widehat{X}\big\|_{H^1}^2 \mathrm{d}t \\
\leqslant &\frac{1}{4\epsilon_3}\mathbb{E}\Big[\int_0^T \big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \mathrm{d}t\Big] + \frac{1}{4\epsilon_4}\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \\
&+ \big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \frac{1}{4\epsilon_5}\mathbb{E}\Big[\big\|\widehat{\psi}(\bar{X}_T)\big\|_{L^2}^2\Big] + \frac{1}{4\epsilon_6}\mathbb{E}\Big[\big\|\widehat{\phi}(\bar{Y}_0)\big\|_{L^2}^2\Big] \\
&+ \epsilon_4\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{Y}\big\|_{L^2}^2 + \big\|\widehat{Z}\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \epsilon_6\mathbb{E}\Big[\big\|\widehat{Y}_0\big\|_{L^2}^2\Big].
\end{aligned} \tag{4.12}$$

Combining (4.12) and (4.8), and choosing ϵ_4 and ϵ_6 small enough, we have

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \|\widehat{Y}(t)\|_{L^2}^2\Big] + \mathbb{E}\Big[\int_0^T \|\widehat{Y}\|_{H^1}^2 \mathrm{d}t\Big] + \mathbb{E}\Big[\int_0^T \|\widehat{Z}\|_{L^2}^2 \mathrm{d}t\Big] \\
\leqslant &K\Big\{\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 + \big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \\
&+ \big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \mathbb{E}\Big[\big\|\widehat{\psi}(\bar{X}_T)\big\|_{L^2}^2\Big] + \mathbb{E}\Big[\big\|\widehat{\phi}(\bar{Y}_0)\big\|_{L^2}^2\Big]\Big\}.
\end{aligned} \tag{4.13}$$

Furthermore, substituting (4.13) into (4.7) yields

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]} \|\widehat{X}(t)\|_{L^2}^2\Big] + \mathbb{E}\Big[\int_0^T \|\widehat{X}\|_{H^1}^2 \mathrm{d}t\Big] \\
\leqslant &K\Big\{\mathbb{E}\Big[\int_0^T \Big(\big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 + \big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2 \\
&+ \big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\Big)\mathrm{d}t\Big] + \mathbb{E}\Big[\big\|\widehat{\psi}(\bar{X}_T)\big\|_{L^2}^2\Big] + \mathbb{E}\Big[\big\|\widehat{\phi}(\bar{Y}_0)\big\|_{L^2}^2\Big]\Big\}.
\end{aligned} \tag{4.14}$$

Then, combining (4.13) and (4.14), we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{X}(t)\|_{L^2}^2\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{X}\|_{H^1}^2\mathrm{d}t\Big]+\mathbb{E}\Big[\sup_{t\in[0,T]}\|\widehat{Y}(t)\|_{L^2}^2\Big]\\
&+\mathbb{E}\Big[\int_0^T\|\widehat{Y}\|_{H^1}^2\mathrm{d}t\Big]+\mathbb{E}\Big[\int_0^T\|\widehat{Z}\|_{L^2}^2\mathrm{d}t\Big]\\
\leqslant K\Big\{&\mathbb{E}\Big[\int_0^T\Big(\big\|\widehat{b}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2+\big\|\widehat{f}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\\
&+\big\|\widehat{\sigma}(t,\bar{X},\nabla\bar{X},\bar{Y},\nabla\bar{Y},\bar{Z})\big\|_{L^2}^2\Big)\mathrm{d}t\Big]+\mathbb{E}\Big[\big\|\widehat{\psi}(\bar{X}_T)\big\|_{L^2}^2\Big]+\mathbb{E}\Big[\big\|\widehat{\phi}(\bar{Y}_0)\big\|_{L^2}^2\Big]\Big\}.
\end{aligned}\tag{4.15}
$$

For the case $\beta_2>\frac{1}{2},\mu_2>0,\beta_1\geqslant 0,\mu_1\geqslant 0$, the estimate (4.5) can be proved similarly. Moreover, by letting $(A,B,C,D,E,F,\bar{f},\bar{\sigma},\bar{b},\bar{\phi},\bar{\psi})=(A,B,C,D,E,F,0,0,0,0,0)$, we obtain the estimate (4.6). The proof is completed. □

5. SMP for forward-backward control systems

This section presents a class of fully coupled forward-backward stochastic partial differential optimal control problems. Let $\mathscr{U}\subset\mathbb{R}^k$ be a nonempty convex subset, known as the control domain.

Definition 5.1 *A control process $u(\cdot,\cdot)$ is called admissible if $u(\cdot,\cdot)\in\mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)$ and $u(t,x)\in\mathscr{U}$, a.e., $t\in[0,T]$, $x\in\mathbb{R}^n$, $\mathbb{P}$-a.s.. The set of all admissible control is denoted by $\mathscr{U}_{ad}$.*

For any given admissible control $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, we consider the following control systems driven by FBSPDE:

$$
\begin{cases}
\mathrm{d}X(t,x)=\Big[\mathfrak{L}_F(t,x,X(t,x),Y(t,x),Z(t,x))+f(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x),Z(t,x),\\
\qquad\qquad u(t,x))\Big]\mathrm{d}t+\Big\langle\sigma(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x),Z(t,x),u(t,x)),\mathrm{d}W(t)\Big\rangle,\\
\mathrm{d}Y(t,x)=-\Big[\mathfrak{L}_B(t,x,X(t,x),Y(t,x),Z(t,x))+b(t,x,X(t,x),\nabla X(t,x),Y(t,x),\nabla Y(t,x),\\
\qquad\qquad Z(t,x),u(t,x))\Big]\mathrm{d}t+\Big\langle Z(t,x),\mathrm{d}W(t)\Big\rangle,\\
X(0,x)=\phi(x,Y_0(x)),\ Y(T,x)=\psi(x,X_T(x)),\ (t,x)\in[0,T]\times\mathbb{R}^n,
\end{cases}\tag{5.1}
$$

where we use the notations in Section 4 and extend the definitions of functions (b,σ,f) to the control set $\mathscr{U}$. The corresponding solution $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot)\big)$ is called the state under a given control $u(\cdot,\cdot)$. Next, we denote the cost functional as follows:

$$
\begin{aligned}
J(u(\cdot,\cdot))\triangleq\mathbb{E}\Big[&\int_0^T\int_{\mathbb{R}^n}g(t,x,X(t,x),Y(t,x),Z(t,x),u(t,x))\mathrm{d}x\mathrm{d}t\\
&+\int_{\mathbb{R}^n}h(x,X_T(x))\mathrm{d}x+\int_{\mathbb{R}^n}\lambda(x,Y_0(x))\mathrm{d}x\Big].
\end{aligned}\tag{5.2}
$$

Besides, we assume the following holds.

Assumption 5.1 (i) *The mappings*

$$
\begin{aligned}
&f(t,\omega,\gamma,u):[0,T]\times\Omega\times\mathfrak{R}\times\mathscr{U}\to\mathbb{R},\\
&b(t,\omega,\gamma,u):[0,T]\times\Omega\times\mathfrak{R}\times\mathscr{U}\to\mathbb{R},\\
&\sigma(t,\omega,\gamma,u):[0,T]\times\Omega\times\mathfrak{R}\times\mathscr{U}\to\mathbb{R}^d
\end{aligned}
$$

are $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R}^d)\otimes\mathscr{B}(\mathscr{U})$*-measurable, where* $\gamma=(\gamma_1,\gamma_2,\gamma_3,\gamma_4,\gamma_5)\in\mathfrak{R}$ *and* $\mathfrak{R}\triangleq\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^d$*. For almost all* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$*,* $\varpi(t,\omega,x,\gamma,u)$ *is continuously differentiable in* $(\gamma_1,\gamma_3,\gamma_5,u)$ *and the corresponding partial derivatives* $\varpi_{\gamma_1},\varpi_{\gamma_3},\varpi_{\gamma_5},\varpi_u$ *are continuous and uniformly bounded, where* $\varpi=f,\sigma,b$ *.*

(ii) *The mappings*

$$\begin{aligned}&\phi(\omega,x,\gamma_3),\lambda(\omega,x,\gamma_3):\Omega\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R},\\&\psi(\omega,x,\gamma_1),h(\omega,x,\gamma_1):\Omega\times\mathbb{R}^n\times\mathbb{R}\to\mathbb{R}\end{aligned}$$

are $\mathcal{F}_0\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})$*-measurable and* $\mathcal{F}_T\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})$*-measurable, respectively. For almost all* $(\omega,x)\in\Omega\times\mathbb{R}^n$*,* $\psi(\omega,x,\gamma_1),h(\omega,x,\gamma_1)$ *and* $\phi(\omega,x,\gamma_3),\lambda(\omega,x,\gamma_3)$ *are continuously differentiable in* γ_1 *and* γ_3*, respectively, and the corresponding partial derivatives* $\psi_{\gamma_1},\phi_{\gamma_3}$ *are continuous and uniformly bounded, and* $h_{\gamma_1},\lambda_{\gamma_3}$ *are continuous. Furthermore, for almost all* $(\omega,x)\in\Omega\times\mathbb{R}^n$*, there exist constants* $L^h,\tilde{L}^h,L^\lambda,\tilde{L}^\lambda>0$*, such that for all* $\gamma_1(\cdot),\gamma_3(\cdot)\in L^2(\mathbb{R}^n;\mathbb{R})$ *,*

$$\begin{aligned}&\left|\int_{\mathbb{R}^n}h(x,\gamma_1(x))\mathrm{d}x\right|\leqslant L^h\Big(1+\|\gamma_1\|_{L^2}^2\Big),\quad\left|\int_{\mathbb{R}^n}h_{\gamma_1}(x,\gamma_1(x))\mathrm{d}x\right|\leqslant\tilde{L}^h\Big(1+\|\gamma_1\|_{L^2}\Big),\\&\left|\int_{\mathbb{R}^n}\lambda(x,\gamma_3(x))\mathrm{d}x\right|\leqslant L^\lambda\Big(1+\|\gamma_3\|_{L^2}^2\Big),\quad\left|\int_{\mathbb{R}^n}\lambda_{\gamma_3}(x,\gamma_3(x))\mathrm{d}x\right|\leqslant\tilde{L}^\lambda\Big(1+\|\gamma_3\|_{L^2}\Big).\end{aligned}$$

(iii) *The mappings*

$$g(t,\omega,x,\gamma_1,\gamma_3,\gamma_5,u):[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\times\mathscr{U}\to\mathbb{R},$$

are $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^n)\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R})\otimes\mathscr{B}(\mathbb{R}^d)\otimes\mathscr{B}(\mathscr{U})$*-measurable. For almost all* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$*,* $g(t,\omega,x,\gamma_1,\gamma_3,\gamma_5,u)$ *is continuously differentiable in* $(\gamma_1,\gamma_3,\gamma_5,u)$ *and the corresponding partial derivatives* $g_{\gamma_1},g_{\gamma_3},g_{\gamma_5},g_u$ *are continuous. Furthermore, for almost all* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$*, there exist constants* $L^g,\tilde{L}^g>0$*, such that for all* $(\gamma_1(\cdot),\gamma_3(\cdot),\gamma_5(\cdot),u(\cdot))\in L^2(\mathbb{R}^n;\mathbb{R})\times L^2(\mathbb{R}^n;\mathbb{R})\times L^2(\mathbb{R}^n;\mathbb{R}^d)\times L^2(\mathbb{R}^n;\mathscr{U})$*,*

$$\begin{aligned}&\left|\int_{\mathbb{R}^n}g(t,x,\gamma_1(x),\gamma_3(x),\gamma_5(x),u(x))\mathrm{d}x\right|\leqslant L^g\Big(1+\|\gamma_1\|_{L^2}^2+\|\gamma_3\|_{L^2}^2+\|\gamma_5\|_{L^2}^2+\|u\|_{L^2}^2\Big),\\&\left|\int_{\mathbb{R}^n}g_{\gamma_1}(t,x,\gamma_1(x),\gamma_3(x),\gamma_5(x),u(x))\mathrm{d}x\right|+\left|\int_{\mathbb{R}^n}g_{\gamma_3}(t,x,\gamma_1(x),\gamma_3(x),\gamma_5(x),u(x))\mathrm{d}x\right|\\&+\left|\int_{\mathbb{R}^n}g_{\gamma_5}(t,x,\gamma_1(x),\gamma_3(x),\gamma_5(x),u(x))\mathrm{d}x\right|+\left|\int_{\mathbb{R}^n}g_{\gamma_u}(t,x,\gamma_1(x),\gamma_3(x),\gamma_5(x),u(x))\mathrm{d}x\right|\\&\leqslant\tilde{L}^g\Big(1+\|\gamma_1\|_{L^2}+\|\gamma_3\|_{L^2}+\|\gamma_5\|_{L^2}+\|u\|_{L^2}\Big).\end{aligned}$$

(iv) *For any given control* $u(\cdot,\cdot)\in\mathscr{U}_{ad}$*, the coefficients of the equation (5.1) satisfy Assumption 4.1.*

By Lemma 4.1, we know that FBSPDE (5.1) admits a unique solution $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot))\in\mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$, $\forall u(\cdot,\cdot)\in\mathscr{U}_{ad}$. In addition, (ii) in Assumption 5.1 and the a priori estimate (4.6) indicate that

$$\big|J\big(u(\cdot,\cdot)\big)\big|<\infty.$$

Next, we propose the optimal control problem as follows:

Problem 5.1 *Find an admissible control* $u(\cdot,\cdot)\in\mathscr{U}_{ad}$ *such that*

$$J\big(u^*(\cdot,\cdot)\big)=\inf_{u(\cdot,\cdot)\in\mathscr{U}_{ad}}J\big(u(\cdot,\cdot)\big).\tag{5.3}$$

Such an admissible control $u^*(\cdot,\cdot)$ is called an optimal control when (5.3) *is satisfied, and* $(X^*(\cdot,\cdot),Y^*(\cdot,\cdot),Z^*(\cdot,\cdot))$ *is called the corresponding optimal trajectory, which solves* (5.1) *under Assumptions 4.1 and 5.1.*

The Hamiltonian $\mathscr{H}:[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\times\mathscr{U}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ is formulated formally as

$$\begin{aligned}\mathscr{H}(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u,p,q,\kappa)=&-b(t,x,\mathfrak{x},\nabla\mathfrak{x},\mathfrak{y},\nabla\mathfrak{y},\mathfrak{z},u)p(t,x)+f(t,x,\mathfrak{x},\nabla\mathfrak{x},\mathfrak{y},\nabla\mathfrak{y},\mathfrak{z},u)q(t,x)\\&+\big\langle\sigma(t,x,\mathfrak{x},\nabla\mathfrak{x},\mathfrak{y},\nabla\mathfrak{y},\mathfrak{z},u),\kappa(t,x)\big\rangle+g(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u).\end{aligned}\tag{5.4}$$

It is obvious that the Hamiltonian $\mathscr{H}$ is also continuously differentiable for $(\mathfrak{x},\mathfrak{y},\mathfrak{z},u)$ from Assumption 4.1 and 5.1. Substituting $\mathscr{H}_\mathfrak{x},\mathscr{H}_\mathfrak{y},\mathscr{H}_\mathfrak{z},\mathscr{H}_u$ represent the corresponding partial derivatives. For any control process $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, let $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))$ be the corresponding state process (5.1) and the adjoint processes (5.6), respectively. In particular, we denote by $(X^*(\cdot,\cdot),Y^*(\cdot,\cdot),Z^*(\cdot,\cdot),p^*(\cdot,\cdot),q^*(\cdot,\cdot),\kappa^*(\cdot,\cdot))$ the state and adjoint processes corresponding to the optimal control $u^*(\cdot,\cdot)$. For simplicity, the argument (t,x) is suppressed, and we have the following notations for subsequent use:

$$\begin{cases}\bar{\mathfrak{L}}_F(t,x):=\mathfrak{L}_F(t,x,\bar X,\bar Y,\bar Z)=\nabla\cdot(A\nabla\bar X+B\nabla\bar Y+C\bar Z),\\ \bar{\mathfrak{L}}_B(t,x):=\mathfrak{L}_B(t,x,\bar X,\bar Y,\bar Z)=\nabla\cdot(-D\nabla\bar X+E\nabla\bar Y+F\bar Z),\\ \tilde{\mathscr{H}_i}(t,x):=\mathscr{H}_i(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u,p^*(t,x),q^*(t,x),\kappa^*(t,x)),\\ \mathscr{H}^\rho(t,x):=\mathscr{H}(t,x,\mathfrak{x}^\rho,\mathfrak{y}^\rho,\mathfrak{z}^\rho,u^\rho,p^*(t,x),q^*(t,x),\kappa^*(t,x)),\\ \mathscr{H}_i^*(t,x):=\mathscr{H}_i(t,x,\mathfrak{x}^*(t,x),\mathfrak{y}^*(t,x),\mathfrak{z}^*(t,x),u^*(t,x),p^*(t,x),q^*(t,x),\kappa^*(t,x)),\\ \varpi(t,x):=\varpi(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u),\ \varpi^\rho(t,x):=\varpi(t,x,\mathfrak{x}^\rho,\mathfrak{y}^\rho,\mathfrak{z}^\rho,u^\rho),\\ \varpi_i^*(t,x):=\varpi_i(t,x,\mathfrak{x}^*(t,x),\mathfrak{y}^*(t,x),\mathfrak{z}^*(t,x),u^*(t,x)),\\ \widehat{\phi}(x,Y_0)=\phi(x,Y_0(x))-\bar\phi(x,Y_0(x)),\ \widehat{\psi}(x,X_T)=\psi(x,X_T(x))-\bar\psi(x,X_T(x)),\\ p_0(x)=p(0,x),\ q_T(x)=q(T,x),\ i=\mathfrak{x},\mathfrak{y},\mathfrak{z},u,\ \varpi=f,\sigma,b,p,q,\kappa,g.\end{cases}\tag{5.5}$$

To derive the maximum principle, for any admissible pair $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot)\big)$, the following adjoint equations for the controlled system (5.1) and (5.2), which control the unknown $\mathbb{F}$-adapted processes $p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot)$ are introduced

$$\begin{cases}\mathrm{d}p(t,x)=-\Big[\nabla\cdot\big(B^\top\nabla q(t,x)-E^\top\nabla p(t,x)\big)+f_\mathfrak{y}(t,x)q(t,x)+\big\langle\sigma_\mathfrak{y}(t,x),\kappa(t,x)\big\rangle\\ \qquad-b_\mathfrak{y}(t,x)p(t,x)+g_\mathfrak{y}(t,x)\Big]\mathrm{d}t-\Big[-C^\top\nabla q(t,x)+F^\top\nabla p(t,x)\\ \qquad+f_\mathfrak{z}(t,x)q(t,x)+\big\langle\sigma_\mathfrak{z}(t,x),\kappa(t,x)\big\rangle-b_\mathfrak{z}(t,x)p(t,x)+g_\mathfrak{z}(t,x)\Big]\mathrm{d}W(t),\\ \mathrm{d}q(t,x)=-\Big[\nabla\cdot\big(A^\top\nabla q(t,x)+D^\top\nabla p(t,x)\big)+f_\mathfrak{x}(t,x)q(t,x)+\big\langle\sigma_\mathfrak{x}(t,x),\kappa(t,x)\big\rangle\\ \qquad-b_\mathfrak{x}(t,x)p(t,x)+g_\mathfrak{x}(t,x)\Big]\mathrm{d}t+\kappa(t,x)\mathrm{d}W(t),\quad(t,x)\in[0,T]\times\mathbb{R}^n,\\ p_0(x)=-\lambda_\mathfrak{y}(x,Y_0(x))-\psi_\mathfrak{y}(x,Y_0(x))q_0(x),\quad q_T(x)=h_\mathfrak{x}(x,X_T(x))-\psi_\mathfrak{x}(x,X_T(x))p_T(x).\end{cases}\tag{5.6}$$

Meanwhile, we can rewrite the adjoint equation (5.6) into the following form:

$$
\begin{cases}
\mathrm{d}p(t,x) = & -\Big[\nabla\cdot\big(B^\top\nabla q(t,x)-E^\top\nabla p(t,x)\big)+\mathscr{H}_{\mathfrak{y}}(t,x)\Big]\mathrm{d}t \\
& -\Big[-C^\top\nabla q(t,x)+F^\top\nabla p(t,x)+\mathscr{H}_{\mathfrak{z}}(t,x)\Big]\mathrm{d}W(t), \\
\mathrm{d}q(t,x) = & -\Big[\nabla\cdot\big(A^\top\nabla q(t,x)+D^\top\nabla p(t,x)+\mathscr{H}_{\mathfrak{x}}(t,x)\Big]\mathrm{d}t+\kappa(t,x)\mathrm{d}W(t),\ (t,x)\in[0,T]\times\mathbb{R}^n, \\
p_0(x) = & -\lambda_{\mathfrak{y}}(x,Y_0(x))-\psi_{\mathfrak{y}}(x,Y_0(x))q_0(x),\quad q_T(x)=h_{\mathfrak{x}}(x,X_T(x))-\psi_{\mathfrak{x}}(x,X_T(x))p_T(x).
\end{cases}
\tag{5.7}
$$

As a result, we have the following lemma.

Lemma 5.1 *Let Assumptions 4.1 and 5.1 hold. Then, there exists a unique triple* $(p(\cdot,\cdot), q(\cdot,\cdot),\kappa(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$ *, which solves the adjoint equation* (5.6).

Proof We can verify that the coefficients of equation (5.6) satisfy Assumption 4.1, and the unique solution exists from Lemma 4.1. Thus, we complete the proof. □

Next, we present the main results of this paper, i.e., the sufficient and necessary SMP for Problem 5.1.

5.1 Sufficient maximum principle

In this subsection, we obtained one of the main results of this paper, i.e., a sufficient condition for the optimality of Problem 5.1, which can also be called SMP.

For any control process $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, let $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})\times\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$ be the corresponding solutions to the state equations (5.1) and the adjoint equations (5.6), respectively.

First, we introduce the following lemma.

Lemma 5.2 *Suppose Assumptions 4.1 and 5.1 hold. For any* $(t,x)\in[0,T]\times\mathbb{R}^n$ *, we have*

$$
\begin{aligned}
&J(u(\cdot,\cdot))-J(u^*(\cdot,\cdot)) \\
=&\,\mathbb{E}\Big[\int_0^T\int_{\mathbb{R}^n}\Big(\tilde{\mathscr{H}}(t,x)-\mathscr{H}^*(t,x)-\mathscr{H}^*_{\mathfrak{x}}(t,x)\big(X(t,x)-X^*(t,x)\big) \\
&-\mathscr{H}^*_{\mathfrak{y}}(t,x)\big(Y(t,x)-Y^*(t,x)\big)-\big\langle\mathscr{H}^*_{\mathfrak{z}}(t,x),Z(t,x)-Z^*(t,x)\big\rangle\Big)\mathrm{d}x\mathrm{d}t\Big] \\
&+\mathbb{E}\Big[\int_{\mathbb{R}^n}\Big(h(x,X_T)-h(x,X^*_T)+\lambda(x,Y_0)-\lambda(x,Y^*_0)\Big)\mathrm{d}x\Big] \\
&+\mathbb{E}\Big[\int_{\mathbb{R}^n}\Big(q^*_0(x)\big(\phi(x,Y_0(x))-\phi(x,Y^*_0(x))\big)-\Big(h_{\mathfrak{x}}(x,X^*_T(x)) \\
&-\psi_{\mathfrak{x}}(x,X^*_T(x))p^*_T(x)\Big)\big(X_T(x)-X^*_T(x)\big)+\Big(-\lambda_{\mathfrak{y}}(x,Y^*_0(x)) \\
&-\phi_{\mathfrak{y}}(x,Y^*_0(x))q^*_0(x)\Big)\big(Y_0(x)-Y^*_0(x)\big)-p^*_T(x)\big(\psi(x,X_T(x))-\psi(x,X^*_T(x))\big)\Big)\mathrm{d}x\Big].
\end{aligned}
\tag{5.8}
$$

Proof Assume that $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot)\big)$ is an arbitrary admissible pair. With the definitions of the Hamiltonian (5.4) and cost functional (5.2), it is clear that one can obtain

$$
\begin{aligned}
&J(u(\cdot,\cdot))-J(u^*(\cdot,\cdot)) \\
=&\,\mathbb{E}\Big[\int_0^T\int_{\mathbb{R}^n}\Big(\tilde{\mathscr{H}}(t,x)-\mathscr{H}^*(t,x)+p^*(t,x)\big(b(t,x)-b^*(t,x)\big)-q^*(t,x)\big(f(t,x)-f^*(t,x)\big) \\
&-\big\langle\kappa^*(t,x),\sigma(t,x)-\sigma^*(t,x)\big\rangle\Big)\mathrm{d}x\mathrm{d}t\Big]+\mathbb{E}\Big[\int_{\mathbb{R}^n}\Big(h(x,X_T)-h(x,X^*_T)+\lambda(x,Y_0)-\lambda(x,Y^*_0)\Big)\mathrm{d}x\Big].
\end{aligned}
\tag{5.9}
$$

By applying Itô's formula to $\Big\langle q^*(t), X(t)-X^*(t)\Big\rangle_{L^2}+\Big\langle p^*(t), Y(t)-Y^*(t)\Big\rangle_{L^2}$, we have

$$\begin{aligned}
&\mathrm{d}\Big\langle q^*(t), X(t)-X^*(t)\Big\rangle_{L^2}+\mathrm{d}\Big\langle p^*(t), Y(t)-Y^*(t)\Big\rangle_{L^2}\\
=&-\Big\langle \nabla\cdot\big(A^\top\nabla q^*(t)+D^\top\nabla p^*(t)\big), X(t)-X^*(t)\Big\rangle_{H^{-1},H^1}\mathrm{d}t-\Big\langle \mathsf{H}^*_{\mathfrak{x}}(t), X(t)-X^*(t)\Big\rangle_{L^2}\mathrm{d}t\\
&+\Big\langle \nabla\cdot\big(A\nabla\big(X(t)-X^*(t)\big)+B\nabla\big(Y(t)-Y^*(t)\big)+C\big(Z(t)-Z^*(t)\big)\big), q^*(t)\Big\rangle_{H^{-1},H^1}\mathrm{d}t\\
&-\Big\langle \nabla\cdot\big(B^\top\nabla q^*(t)-E^\top\nabla p^*(t)\big), Y(t)-Y^*(t)\Big\rangle_{H^{-1},H^1}\mathrm{d}t-\Big\langle \mathsf{H}^*_{\mathfrak{y}}(t), Y(t)-Y^*(t)\Big\rangle_{L^2}\mathrm{d}t\\
&-\Big\langle \nabla\cdot\big(-D\nabla\big(X(t)-X^*(t)\big)+E\nabla\big(Y(t)-Y^*(t)\big)+F\big(Z(t)-Z^*(t)\big)\big), p^*(t)\Big\rangle_{H^{-1},H^1}\mathrm{d}t\\
&-\Big\langle -C^\top\nabla q^*(t)+F^\top\nabla p^*(t), Z(t)-Z^*(t)\Big\rangle_{H^{-1},H^1}\mathrm{d}t-\Big\langle \mathsf{H}^*_{\mathfrak{z}}(t), Z(t)-Z^*(t)\Big\rangle_{L^2}\mathrm{d}t\\
&+\Big\langle q^*(t), f(t)-f^*(t)\Big\rangle_{L^2}\mathrm{d}t-\Big\langle p^*(t), b(t)-b^*(t)\Big\rangle_{L^2}\mathrm{d}t+\Big\langle \kappa^*(t), \sigma(t)-\sigma^*(t)\Big\rangle_{L^2}\mathrm{d}t, \quad (5.10)
\end{aligned}$$

then taking (5.10) into (5.9), it is quite easy to get (5.8). □

Theorem 5.1 *Let Assumptions 4.1 and 5.1 hold. Suppose that* $\big(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot), u^*(\cdot,\cdot)\big)$ *is an admissible pair and* $(p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot))$ *is the adjoint process of the corresponding adjoint equation* (5.7) *with* $\psi(\cdot, X_T)=\psi X_T$ *and* $\phi(\cdot, Y_0)=\phi Y_0$ *, where* ψ *and* ϕ *are* $\mathcal{F}_T$*-measurable bounded random variable and* $\mathcal{F}_0$*-measurable bounded random variable, respectively. Moreover, if the following conditions hold:*

(1) $\mathscr{H}(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u,p^*,q^*,\kappa^*)$, $h(\omega,x,\mathfrak{x})$, $\lambda(\omega,x,\mathfrak{y})$ *are convex in* $(\mathfrak{x},\mathfrak{y},\mathfrak{z},u)$, $\mathfrak{x}$ *and* $\mathfrak{y}$*, respectively;*

(2) $\mathscr{H}^*(t,x)=\min_{u\in\mathscr{U}}\mathscr{H}(t,x,\mathfrak{x}^*,\mathfrak{y}^*,\mathfrak{z}^*,u,p^*,q^*,\kappa^*)$, *for almost all* $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$ *, then* $\big(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot), u^*(\cdot,\cdot)\big)$ *is an optimal pair of Problem 5.1.*

Proof Assume that $\big(X(\cdot,\cdot), Y(\cdot,\cdot), Z(\cdot,\cdot), u(\cdot,\cdot)\big)$ is an arbitrary admissible pair. According to Lemma 5.2, the difference of $J(u(\cdot,\cdot))-J(u^*(\cdot,\cdot))$ becomes (5.8). As $\mathscr{H}(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u,p^*,q^*,\kappa^*)$ is convex in $(\mathfrak{x},\mathfrak{y},\mathfrak{z},u)$, for almost all $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$, by virtue of Proposition 5.4 in Chapter 1 of Ekeland and Temam [9], we have

$$\begin{aligned}
\int_{\mathbb{R}^n}\Big(\tilde{\mathscr{H}}(t,x)-\mathscr{H}^*(t,x)\Big)\mathrm{d}x \geqslant &\Big\langle \mathscr{H}^*_{\mathfrak{x}}(t), X(t)-X^*(t)\Big\rangle_{L^2}+\Big\langle \mathscr{H}^*_{\mathfrak{y}}(t), Y(t)-Y^*(t)\Big\rangle_{L^2}\\
&+\Big\langle \mathscr{H}^*_{\mathfrak{z}}(t), Z(t)-Z^*(t)\Big\rangle_{L^2}+\Big\langle \mathscr{H}^*_{u}(t), u(t)-u^*(t)\Big\rangle_{L^2}, \quad (5.11)
\end{aligned}$$

and

$$\begin{aligned}
&\int_{\mathbb{R}^n}\Big(h(x,X_T(x))-h(x,X_T^*(x))+\lambda(x,Y_0(x))-\lambda(x,Y_0^*(x))\Big)\mathrm{d}x\\
\geqslant &\Big\langle h_{\mathfrak{x}}(X_T^*), X_T-X_T^*\Big\rangle_{L^2}+\Big\langle \lambda_{\mathfrak{y}}(Y_0^*), Y_0-Y_0^*\Big\rangle_{L^2}. \quad (5.12)
\end{aligned}$$

Additionally, as $\mathscr{H}^*(t,x)=\min_{u\in\mathscr{U}}\mathscr{H}(t,x,\mathfrak{x}^*,\mathfrak{y}^*,\mathfrak{z}^*,u,p^*,q^*,\kappa^*)$ and the convex optimization principle, by virtue of Proposition 2.1 in Chapter 2 of Ekeland and Temam [9], for almost all $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$, we have

$$\Big\langle \mathscr{H}^*_u(t), u(t)-u^*(t)\Big\rangle_{L^2}\geqslant 0. \quad (5.13)$$

Meanwhile,

$$\begin{aligned}&\Big\langle q_0^*, \phi(Y_0)-\phi(Y_0^*)\Big\rangle_{L^2} - \Big\langle p_T^*, \psi(X_T)-\psi(X_T^*)\Big\rangle_{L^2}\\&+\Big\langle \psi_{\mathfrak{x}}(X_T^*)p_T^*, X_T-X_T^*\Big\rangle_{L^2} - \Big\langle \phi_{\mathfrak{y}}(Y_0^*)q_0^*, Y_0-Y_0^*\Big\rangle_{L^2}=0. \end{aligned}\tag{5.14}$$

Thus, combining (5.11), (5.12), (5.13) and (5.14), it is easy to see that

$$J(u(\cdot,\cdot))-J(u^*(\cdot,\cdot))\geqslant 0. \tag{5.15}$$

□

5.2 Necessary maximum principle

This subsection presents one of the main results of this paper, i.e., a stochastic necessary condition for the optimality of Problem 5.1, which can also be called the necessary SMP.

For any given admissible control $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, since $\mathscr{U}$ is convex, we have the following perturbed control process $u^\rho(\cdot,\cdot):=u^*(\cdot,\cdot)+\rho\big(u(\cdot,\cdot)-u^*(\cdot,\cdot)\big)\in\mathscr{U}_{ad}, 0\leqslant\rho\leqslant 1$. Let $(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot))$ and $(X^\rho(\cdot,\cdot), Y^\rho(\cdot,\cdot), Z^\rho(\cdot,\cdot))$ be the state processes corresponding to $u^*(\cdot,\cdot)$ and $u^\rho(\cdot,\cdot)$, respectively. Meanwhile, $(p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot))$ is the adjoint process corresponding to the optimal control $u^*(\cdot,\cdot)$. Then, we have the following convergence result.

Lemma 5.3 *Suppose Assumptions 4.1 and 5.1 hold. Then,*

$$\begin{aligned}&\mathbb{E}\bigg[\sup_{t\in[0,T]}\big\|X^\rho(t)-X^*(t)\big\|_{L^2}^2+\sup_{t\in[0,T]}\big\|Y^\rho(t)-Y^*(t)\big\|_{L^2}^2\bigg]\\&+\ \mathbb{E}\bigg[\int_0^T\Big(\big\|X^\rho(t)-X^*(t)\big\|_{H^1}^2+\big\|Y^\rho(t)-Y^*(t)\big\|_{H^1}^2+\big\|Z^\rho(t)-Z^*(t)\big\|_{L^2}^2\Big)\mathrm{d}t\bigg]=O(\rho^2).\end{aligned}\tag{5.16}$$

Proof By a priori estimation (see Theorem 4.1) and the boundedness of the partial derivatives, we have the following inequality

$$\begin{aligned}&\mathbb{E}\bigg[\sup_{t\in[0,T]}\big\|X^\rho(t)-X^*(t)\big\|_{L^2}^2+\sup_{t\in[0,T]}\big\|Y^\rho(t)-Y^*(t)\big\|_{L^2}^2\bigg]\\&+\mathbb{E}\bigg[\int_0^T\Big(\big\|X^\rho(t)-X^*(t)\big\|_{H^1}^2+\big\|Y^\rho(t)-Y^*(t)\big\|_{H^1}^2+\big\|Z^\rho(t)-Z^*(t)\big\|_{L^2}^2\Big)\mathrm{d}t\bigg]\\\leqslant{}& K\mathbb{E}\bigg[\int_0^T\Big(\big\|u^\rho(t)-u^*(t)\big\|_{L^2}^2\Big)\mathrm{d}t\bigg]\\={}&K\rho^2\mathbb{E}\bigg[\int_0^T\Big(\big\|u(t)-u^*(t)\big\|_{L^2}^2\Big)\mathrm{d}t\bigg]\\={}&O(\rho^2).\end{aligned}\tag{5.17}$$

This proof is complete. □

Next, we present the variational formula for the cost functional $J(u(\cdot,\cdot))$ with Hamiltonian $\mathscr{H}$ and state process $(X(\cdot,\cdot), Y(\cdot,\cdot), Z(\cdot,\cdot))$.

Lemma 5.4 *Suppose Assumptions 4.1 and 5.1 hold. Then, for any admissible control* $u(\cdot,\cdot) \in \mathscr{U}_{ad}$ *, the directional derivative of the cost functional* $J(u(\cdot,\cdot))$ *at* $u^*(\cdot,\cdot)$ *in the direction* $u(\cdot,\cdot) - u^*(\cdot,\cdot)$ *is as follows*

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\rho} J\big(u^*(\cdot,\cdot) + \rho\big(u(\cdot,\cdot) - u^*(\cdot,\cdot)\big)\big)|_{\rho=0}\\ &= \lim_{\rho\to 0} \frac{J\big(u^*(\cdot,\cdot) + \rho\big(u(\cdot,\cdot) - u^*(\cdot,\cdot)\big)\big) - J\big(u^*(\cdot,\cdot)\big)}{\rho}\\ &= \mathbb{E}\Big[\int_0^T \Big\langle \mathscr{H}_u(t, X^*(t), Y^*(t), Z^*(t), u^*(t), p^*(t), q^*(t), \kappa^*(t)), u(t) - u^*(t)\Big\rangle_{L^2} \mathrm{d}t\Big]. \qquad (5.18)\end{aligned}$$

Proof Using (5.8), we obtain

$$\begin{aligned}&J\big(u^\rho(\cdot,\cdot)\big) - J\big(u^*(\cdot,\cdot)\big)\\ =\ &\mathbb{E}\Big[\int_0^T \int_{\mathbb{R}^n} \Big(\mathscr{H}^\rho(t,x) - \mathscr{H}^*(t,x) - \mathscr{H}_{\mathfrak{x}}^*(t,x)\big(X^\rho(t,x) - X^*(t,x)\big) - \mathscr{H}_{\mathfrak{y}}^*(t,x)\big(Y^\rho(t,x)\\ &- Y^*(t,x)\big) - \big\langle \mathscr{H}_{\mathfrak{z}}^*(t,x), Z^\rho(t,x) - Z^*(t,x)\big\rangle - \big\langle \mathscr{H}_u^*(t,x), u^\rho(t,x) - u^*(t,x)\big\rangle\Big)\mathrm{d}x\mathrm{d}t\Big]\\ &+ \mathbb{E}\Big[\int_0^T \int_{\mathbb{R}^n} \langle \mathscr{H}_u^*(t,x), u^\rho(t,x) - u^*(t,x)\rangle \mathrm{d}x\mathrm{d}t\Big]\\ &+ \mathbb{E}\Big[\int_{\mathbb{R}^n} \Big(h(x, X_T^\rho) - h(x, X_T^*) + \lambda(x, Y_0^\rho) - \lambda(x, Y_0^*)\Big)\mathrm{d}x\Big]\\ &+ \mathbb{E}\Big[\int_{\mathbb{R}^n} \Big(q_0^*(x)\big(\phi(x, Y_0^\rho(x)) - \phi(x, Y_0^*(x))\big)\\ &- \Big(h_{\mathfrak{x}}(x, X_T^*(x)) - \psi_{\mathfrak{x}}(x, X_T^*(x)) p_T^*(x)\Big)\big(X_T^\rho(x) - X_T^*(x)\big)\\ &+ \Big(-\lambda_{\mathfrak{y}}(x, Y_0^*(x)) - \phi_{\mathfrak{y}}(x, Y_0^*(x)) q_0^*(x)\Big)\big(Y_0^\rho(x) - Y_0^*(x)\big)\\ &- p_T^*(x)\big(\psi(x, X_T^\rho(x)) - \psi(x, X_T^*(x))\big)\Big)\mathrm{d}x\Big]. \qquad (5.19)\end{aligned}$$

By means of Taylor series expansion and change of variables, we have

$$\begin{aligned}&\mathbb{E}\Big[\int_0^T \int_{\mathbb{R}^n} \Big(\mathscr{H}^\rho(t,x) - \mathscr{H}^*(t,x)\Big)\mathrm{d}x\mathrm{d}t\Big]\\ =\ &\mathbb{E}\Big[\int_0^T \int_{\mathbb{R}^n} \int_0^1 \Big(\mathscr{H}_{\mathfrak{x}}^{\rho,\lambda}(t,x)\big(X^\rho(t,x) - X^*(t,x)\big) + \mathscr{H}_{\mathfrak{y}}^{\rho,\lambda}(t,x)\big(Y^\rho(t,x) - Y^*(t,x)\big)\\ &+ \big\langle \mathscr{H}_{\mathfrak{z}}^{\rho,\lambda}(t,x), Z^\rho(t,x) - Z^*(t,x)\big\rangle + \big\langle \mathscr{H}_u^{\rho,\lambda}(t,x), u^\rho(t,x) - u^*(t,x)\big\rangle\Big)\mathrm{d}\lambda\mathrm{d}x\mathrm{d}t\Big], \qquad (5.20)\end{aligned}$$

where $\mathscr{H}^{\rho,\lambda}(t,x) := \mathscr{H}(t, x, \mathfrak{x}^{\rho,\lambda}(t,x), \mathfrak{y}^{\rho,\lambda}(t,x), \mathfrak{z}^{\rho,\lambda}(t,x), u^{\rho,\lambda}(t,x), p^*(t,x), q^*(t,x), \kappa^*(t,x))$ with $\mathfrak{x}^{\rho,\lambda}(t,x) := \mathfrak{x}^*(t,x) + \lambda\mathfrak{x}^\rho(t,x)$ and $u^{\rho,\lambda}(t,x) := u^*(t,x) + \lambda u^\rho(t,x)$. Thus, combining (5.20) and (5.16) and using the dominated convergence theorem gives

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T\int_{\mathbb{R}^n}\Big(\mathscr{H}^\rho(t,x)-\mathscr{H}^*(t,x)-\mathscr{H}_{\mathfrak{x}}^*(t,x)\big(X^\rho(t,x)-X^*(t,x)\big)\\
&-\mathscr{H}_{\mathfrak{y}}^*(t,x)\big(Y^\rho(t,x)-Y^*(t,x)\big)-\big\langle\mathscr{H}_{\mathfrak{z}}^*(t,x),Z^\rho(t,x)-Z^*(t,x)\big\rangle\\
&-\big\langle\mathscr{H}_u^*(t,x),u^\rho(t,x)-u^*(t,x)\big\rangle\Big)\mathrm{d}x\mathrm{d}t\Big]\\
=\;&o(\rho).
\end{aligned}\tag{5.21}
$$

Similarly, we can conclude that

$$
\begin{aligned}
&\mathbb{E}\Big[\int_{\mathbb{R}^n}\Big(h(x,X_T^\rho)-h(x,X_T^*)+\lambda(x,Y_0^\rho)-\lambda(x,Y_0^*)\Big)\mathrm{d}x\Big]\\
&+\mathbb{E}\Big[\int_{\mathbb{R}^n}\Big(q_0^*(x)\big(\phi(x,Y_0^\rho(x))-\phi(x,Y_0^*(x))\big)\\
&-\Big(h_{\mathfrak{x}}(x,X_T^*(x))-\psi_{\mathfrak{x}}(x,X_T^*(x))p_T^*(x)\Big)\big(X_T^\rho(x)-X_T^*(x)\big)\\
&+\Big(-\lambda_{\mathfrak{y}}(x,Y_0^*(x))-\phi_{\mathfrak{y}}(x,Y_0^*(x))q_0^*(x)\Big)\big(Y_0^\rho(x)-Y_0^*(x)\big)\\
&-p_T^*(x)\big(\psi(x,X_T^\rho(x))-\psi(x,X_T^*(x))\big)\Big)\mathrm{d}x\Big]\\
=\;&o(\rho).
\end{aligned}\tag{5.22}
$$

Consequently, combining (5.21), (5.22) and (5.19) yields

$$
\begin{aligned}
&\lim_{\rho\to0}\rho^{-1}\Big(J\big(u^*(\cdot,\cdot)+\rho\big(u(\cdot,\cdot)-u^*(\cdot,\cdot)\big)\big)-J\big(u^*(\cdot,\cdot)\big)\Big)\\
=\;&\mathbb{E}\Big[\int_0^T\int_{\mathbb{R}^n}\Big\langle\mathscr{H}_u^*(t,x),u(t,x)-u^*(t,x)\Big\rangle\mathrm{d}x\mathrm{d}t\Big].
\end{aligned}
$$

The proof is complete. □

We now show the main results of this section.

Theorem 5.2 *Let Assumptions 4.1 and 5.1 hold. Assume that $u^*(\cdot,\cdot)$ is the optimal control with the corresponding trajectory $(X^*(\cdot,\cdot),Y^*(\cdot,\cdot),Z^*(\cdot,\cdot))$ and the corresponding adjoint process $(p^*(\cdot,\cdot),q^*(\cdot,\cdot),\kappa^*(\cdot,\cdot))$. Then, for any admissible control $u(\cdot,\cdot)\in\mathscr{U}$, and almost all $(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$, we have the following result:*

$$
\Big\langle\mathscr{H}_u(t,x,X^*(t,x),Y^*(t,x),Z^*(t,x),u^*(t,x),p^*(t,x),q^*(t,x),\kappa^*(t,x)),u(t,x)-u^*(t,x)\Big\rangle\geqslant0.\tag{5.23}
$$

Proof Based on Lemma 5.4 and the optimality of $u^*(\cdot,\cdot)$, we deduce that

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T\Big\langle\mathscr{H}_u(t,X^*(t),Y^*(t),Z^*(t),u^*(t),p^*(t),q^*(t),\kappa^*(t)),u(t)-u^*(t)\Big\rangle_{L^2}\mathrm{d}t\Big]\\
=\;&\lim_{\rho\to0}\frac{J\big(u^*(\cdot,\cdot)+\rho\big(u(\cdot,\cdot)-u^*(\cdot,\cdot)\big)\big)-J\big(u^*(\cdot,\cdot)\big)}{\rho}\geqslant0,\quad\forall u(\cdot,\cdot)\in\mathscr{U}_{ad}.
\end{aligned}\tag{5.24}
$$

The proof is complete. □

Based on the minimum condition (5.23) of Theorem 5.2, combined with the state equation (5.1) and the adjoint equation (5.7) corresponding to the optimal control $u^*(\cdot,\cdot)$, we derive the following stochastic system:

$$
\begin{cases}
\mathrm{d}X^*(t,x) = \Big[\nabla\cdot\big(A\nabla X^*(t,x)+B\nabla Y^*(t,x)+CZ^*(t,x)\big)+f^*(t,x)\Big]\mathrm{d}t+\sigma^*(t,x)\mathrm{d}W(t),\\
\mathrm{d}Y^*(t,x) = -\Big[\nabla\cdot\big(-D\nabla X^*(t,x)+E\nabla Y^*(t,x)+FZ^*(t,x)\big)+b^*(t,x)\Big]\mathrm{d}t+Z^*(t,x)\mathrm{d}W(t),\\
\mathrm{d}p^*(t,x) = -\Big[\nabla\cdot\big(B^\top\nabla q^*(t,x)-E^\top\nabla p^*(t,x)\big)+\mathscr{H}^*_{\mathfrak{y}}(t,x)\Big]\mathrm{d}t\\
\qquad\qquad\qquad -\Big[-C^\top\nabla q^*(t,x)+F^\top\nabla p^*(t,x)+\mathscr{H}^*_{\mathfrak{z}}(t,x)\Big]\mathrm{d}W(t),\\
\mathrm{d}q^*(t,x) = -\Big[\nabla\cdot\big(A^\top\nabla q^*(t,x)+D^\top\nabla p^*(t,x)\big)+\mathscr{H}^*_{\mathfrak{x}}(t,x)\Big]\mathrm{d}t+\kappa^*(t,x)\mathrm{d}W(t),\\
X^*(0,x)=\phi(x,Y_0^*(x)),\quad Y^*(T,x)=\psi(x,X_T^*(x)),\\
\quad p_0^*(x)=-\lambda_{\mathfrak{y}}(x,Y_0^*(x))-\phi_{\mathfrak{y}}(x,Y_0^*(x))q_0^*(x),\quad q_T^*(x)=h_{\mathfrak{x}}(x,X_T^*(x))-\psi_{\mathfrak{x}}(x,X_T^*(x))p_T^*(x),\\
\Big\langle \mathscr{H}^*_u(t,x),u(t,x)-u^*(t,x)\Big\rangle \geqslant 0,\quad \forall u\in\mathscr{U},\quad (t,x)\in[0,T]\times\mathbb{R}^n.
\end{cases}
\tag{5.25}
$$

Obviously, this is a kind of fully coupled FBSPDEs. The last inequality denote the minimum condition, which is the so-called optimal system or stochastic Hamiltonian system of Problem 5.1, whose solution is a 7-tuple stochastic process $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot),p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})\times\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^k)\times\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$. Under appropriate assumptions, we have the following equivalent relation between the solvability of the stochastic Hamiltonian system (5.25) and the existence of the optimal control of Problem 5.1.

Corollary 5.1 *Suppose Assumptions 4.1, 5.1, and condition (i) in Theorem 5.1 hold. Then, the existence of solution to the stochastic Hamiltonian system* (5.25) *is equivalent to the existence of the optimal control of Problem 5.1.*

Proof First, we prove the existence and uniqueness of an optimal control of Problem 5.1 when the adapted solution $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot),p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))\in\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})\times\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^k)\times\mathscr{N}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^{2+d})$ exists for the stochastic Hamiltonian system (5.25). As $\mathscr{H}(t,x,X^*,Y^*,Z^*,u,p^*,q^*,\kappa^*)$ is convex on u and based on the minimum condition of the Hamiltonian system (5.25), the following formula holds

$$
\mathscr{H}(t,x,X^*,Y^*,Z^*,u^*,p^*,q^*,\kappa^*)=\min_{u\in\mathscr{U}}\mathscr{H}(t,x,X^*,Y^*,Z^*,u,p^*,q^*,\kappa^*),\quad (t,x)\in[0,T]\times\mathbb{R}^n.
\tag{5.26}
$$

Therefore, $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot))$ is an optimal pair according to Theorem 5.1.

On the contrary, if $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot))$ is an optimal pair with corresponding adjoint process $(p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))$, then by Theorem 5.2, $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot),p(\cdot,\cdot),q(\cdot,\cdot),\kappa(\cdot,\cdot))$ is an adapted solution of the stochastic Hamiltonian system (5.25). Thus, the proof is completed. □

6. Application in LQ optimal control problem

In this section, we delve into a class of LQ control problems. The LQ control problem is a special form of the optimal control problem that assumes linear relationships between the state and control variables with a quadratic cost functional. Due to its well-structured mathematical framework, the LQ control problem has been widely used in the fields of control engineering, automated systems, robot control, economics, finance, and so on. These applications underscore its pivotal role in the field of optimal control. For readers interested in recent advances in LQ problems driven by FBSDEs, references include [22, 49–51, 65, 66].

Let the control domain $\mathscr{U}=\mathbb{R}^k$, then the admissible control set is denoted by $\mathscr{U}_{ad}=\mathscr{L}^2_{\mathbb{F}}(\mathbb{R}^n;\mathbb{R}^k)$. For any $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, we consider the following control system driven by a linear FBSPDE:

$$\begin{cases}\mathrm{d}X(t,x)=\Big[\nabla\cdot\Big(\mathcal{A}_3\nabla X(t,x)+\mathcal{B}_3\nabla Y(t,x)+\mathcal{C}_2Z(t,x)\Big)+\mathcal{A}_1X(t,x)+\mathcal{A}_2\nabla X(t,x)+\mathcal{B}_1Y(t,x)\\ \qquad\qquad +\mathcal{B}_2\nabla Y(t,x)+\mathcal{C}_1Z(t,x)+\mathcal{J}u(t,x)\Big]\mathrm{d}t+\Big[\mathcal{G}_1X(t,x)+\mathcal{G}_2\nabla X(t,x)+\mathcal{H}_1Y(t,x)\\ \qquad\qquad +\mathcal{H}_2\nabla Y(t,x)+\mathcal{I}_1Z(t,x)+\mathcal{K}u(t,x)\Big]\mathrm{d}W_t,\\ \mathrm{d}Y(t,x)=-\Big[\nabla\cdot\Big(-\mathcal{D}_3\nabla X(t,x)+\mathcal{E}_3\nabla Y(t,x)+\mathcal{F}_2Z(t,x)\Big)+\mathcal{D}_1X(t,x)+\mathcal{D}_2\nabla X(t,x)\\ \qquad\qquad +\mathcal{E}_1Y(t,x)+\mathcal{E}_2\nabla Y(t,x)+\mathcal{F}_1Z(t,x)+\mathcal{L}u(t,x)\Big]\mathrm{d}t+Z(t,x)\mathrm{d}W_t,\\ X(0,x)=\phi Y(0,x),\ Y(T,x)=\psi X(T,x),\ (t,x)\in[0,T]\times\mathbb{R}^n,\end{cases}\tag{6.1}$$

with the quadratic cost functional as follows:

$$J(u(\cdot,\cdot))=\mathbb{E}\Big[\int_0^T\Big(\big\langle\mathcal{Q}X(t,\cdot),X(t,\cdot)\big\rangle_{L^2}+\big\langle\mathcal{R}u(t,\cdot),u(t,\cdot)\big\rangle_{L^2}+\big\langle\mathcal{S}Y(t,\cdot),Y(t,\cdot)\big\rangle_{L^2}\\ +\big\langle\mathcal{V}Z(t,\cdot),Z(t,\cdot)\big\rangle_{L^2}\Big)\mathrm{d}t+\big\langle\mathcal{M}X_{T,\cdot},X_{T,\cdot}\big\rangle_{L^2}+\big\langle\mathcal{N}Y_{0,\cdot},Y_{0,\cdot}\big\rangle_{L^2}\Big],\tag{6.2}$$

where $\mathcal{A}_3=\mathcal{A}_3(t,x)$, $\mathcal{B}_3=\mathcal{B}_3(t,x)$, $\mathcal{D}_3=\mathcal{D}_3(t,x)$, $\mathcal{E}_3=\mathcal{E}_3(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{S}^n))$. Similarly, $\mathcal{A}_2=\mathcal{A}_2(t,x)$, $\mathcal{B}_2=\mathcal{B}_2(t,x)$, $\mathcal{D}_2=\mathcal{D}_2(t,x)$, $\mathcal{E}_2=\mathcal{E}_2(t,x)$, $\mathcal{G}_2=\mathcal{G}_2(t,x)$, $\mathcal{H}_2=\mathcal{H}_2(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}^{n\times n}))$. $\mathcal{C}_2=\mathcal{C}_2(t)$, $\mathcal{F}_2=\mathcal{F}_2(t)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}^{n\times d}))$. $\mathcal{A}_1=\mathcal{A}_1(t,x)$, $\mathcal{B}_1=\mathcal{B}_1(t,x)$, $\mathcal{D}_1=\mathcal{D}_1(t,x)$, $\mathcal{E}_1=\mathcal{E}_1(t,x)$, $\mathcal{G}_1=\mathcal{G}_1(t,x)$, $\mathcal{H}_1=\mathcal{H}_1(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}))$. $\mathcal{C}_1=\mathcal{C}_1(t)$, $\mathcal{I}_1=\mathcal{I}_1(t)$, $\mathcal{F}_1=\mathcal{F}_1(t)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}^d))$. $\mathcal{J}=\mathcal{J}(t,x)$, $\mathcal{K}=\mathcal{K}(t,x)$, $\mathcal{L}=\mathcal{L}(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}^k))$. $\mathcal{Q}=\mathcal{Q}(t,x),\mathcal{S}=\mathcal{S}(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{R}))$, $\mathcal{R}=\mathcal{R}(t,x)\in L^\infty_{\mathbb{F}}(0,T;C_b(\mathbb{R}^n;\mathbb{S}^k))$, $\mathcal{V}=\mathcal{V}(t)\in L^\infty_{\mathbb{F}}(0,T;\mathbb{R})$, $\psi,\mathcal{M}\in L^\infty_{\mathcal{F}_T}(\Omega;C_b(\mathbb{R}^n;\mathbb{R}))$, $\phi,\mathcal{N}\in L^\infty_{\mathcal{F}_0}(\Omega;C_b(\mathbb{R}^n;\mathbb{R}))$.

Moreover, we impose the following assumption:

Assumption 6.1 (i) $\forall(\omega,x)\in\Omega\times\mathbb{R}^n$, *the random variables $\mathcal{M},\mathcal{N}$ are nonnegative definite.*

(ii) $\forall(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$, *the stochastic processes $\mathcal{Q}$ and $\mathcal{S}$ are nonnegative definite.*

(iii) $\forall(t,\omega)\in[0,T]\times\Omega$, *the stochastic process $\mathcal{V}$ is nonnegative definite.*

(iv) $\forall(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n$, *the stochastic process $\mathcal{R}$ is positive definite. Besides, there exists a constant $\delta>0$ such that*

$$\big\langle\mathcal{R}u(t,\cdot),u(t,\cdot)\big\rangle_{L^2}\geqslant\delta\big\langle u(t,\cdot),u(t,\cdot)\big\rangle_{L^2},\quad\forall u\in\mathscr{U}.$$

(v) *The super-parabolic condition holds, i.e.,*

$$\begin{pmatrix}\mathcal{A}_3\mathcal{A}_3 & 0\\ 0 & \mathcal{E}_3\mathcal{E}_3\end{pmatrix}\leqslant C_1I,\quad\begin{pmatrix}\mathcal{B}_3\mathcal{B}_3 & 0\\ 0 & \mathcal{D}_3\mathcal{D}_3\end{pmatrix}\leqslant C_2I,\quad\begin{pmatrix}\mathcal{C}_2\mathcal{C}_2^\top & 0\\ 0 & \mathcal{F}_2\mathcal{F}_2^\top\end{pmatrix}\leqslant C_3I,$$

and let the following conditions hold,

$$\begin{pmatrix}\mathcal{B}_3-\mathcal{C}_2\mathcal{C}_2^\top & 0\\ 0 & \mathcal{D}_3-\mathcal{F}_2\mathcal{F}_2^\top\end{pmatrix}\geqslant c_0I,\quad\mathcal{A}_3\geqslant c_1I,\quad\mathcal{E}_3-\mathcal{F}_2\mathcal{F}_2^\top\geqslant c_2I,$$

where the constant $c_0,c_1,c_2,C_1,C_2,C_3>0$ with $c_0\geqslant\sqrt{C_1}$ and $2c_1>\operatorname*{ess\,sup}_{(t,x)\in[0,T]\times\mathbb{R}^n}|\mathcal{G}_2\mathcal{G}_2^\top|$ and I is an identity matrix with $n\times n$ dimension.

(vi) *For any given control $u(\cdot,\cdot)\in\mathscr{U}_{ad}$, the coefficients of the equation (5.1) satisfies Assumption 4.1.*

Now, we propose the optimal control problem as follows:

Problem 6.1 *Find an admissible control $u(\cdot,\cdot)\in\mathscr{U}_{ad}$ such that*

$$J\big(u^*(\cdot,\cdot)\big)=\inf_{u(\cdot,\cdot)\in\mathscr{U}_{ad}} J\big(u(\cdot,\cdot)\big). \tag{6.3}$$

Such an admissible control $u^*(\cdot,\cdot)$ is called an optimal control when (6.3) *is satisfied, and $(X^*(\cdot,\cdot),Y^*(\cdot,\cdot),Z^*(\cdot,\cdot))$ is called the corresponding optimal trajectory, which solves* (6.1) *under Assumptions 6.1.*

By Lemma 4.1, we know that FBSPDE (5.1) admits a unique solution $(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot))\in\mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$, $\forall u(\cdot,\cdot)\in\mathscr{U}_{ad}$. In addition, Assumption 6.1 and the a priori estimate (4.6) indicate that $\big|J\big(u(\cdot,\cdot)\big)\big|<\infty$. Therefore, Problem 6.1 is well-defined. The coefficients (f,σ,b,g,h,λ) in Problem 5.1 are specified as

$$\begin{aligned}
&f(t,x,X,\nabla X,Y,\nabla Y,Z,u)=\mathcal{A}_1X(t,x)+\mathcal{A}_2\nabla X(t,x)+\mathcal{B}_1Y(t,x)+\mathcal{B}_2\nabla Y(t,x)+\mathcal{C}_1Z(t,x)+\mathcal{J}u(t,x),\\
&\sigma(t,x,X,\nabla X,Y,\nabla Y,Z,u)=\mathcal{G}_1X(t,x)+\mathcal{G}_2\nabla X(t,x)+\mathcal{H}_1Y(t,x)+\mathcal{H}_2\nabla Y(t,x)+\mathcal{I}_1Z(t,x)+\mathcal{K}u(t,x),\\
&b(t,x,X,\nabla X,Y,\nabla Y,Z,u)=\mathcal{D}_1X(t,x)+\mathcal{D}_2\nabla X(t,x)+\mathcal{E}_1Y(t,x)+\mathcal{E}_2\nabla Y(t,x)+\mathcal{F}_1Z(t,x)+\mathcal{L}u(t,x),\\
&g(t,x,X,Y,Z,u)=\big\langle\mathcal{Q}X(t,x),X(t,x)\big\rangle+\big\langle\mathcal{R}u(t,x),u(t,x)\big\rangle+\big\langle\mathcal{S}Y(t,x),Y(t,x)\big\rangle+\big\langle\mathcal{V}Z(t,x),Z(t,x)\big\rangle,\\
&h(x,X_T)=\big\langle\mathcal{M}X(T,x),X(T,x)\big\rangle,\\
&\lambda(x,Y_0)=\big\langle\mathcal{N}Y(0,x),Y(0,x)\big\rangle.
\end{aligned}$$

From Assumption 6.1, we observe that the coefficients (f,σ,b,g,h,λ) satisfy Assumption 5.1 and 4.1. It is then possible to apply SMP to Problem 6.1. Likewise, the Hamiltonian $\mathscr{H}:[0,T]\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\times\mathscr{U}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ becomes

$$\begin{aligned}
&\mathscr{H}(t,x,\mathfrak{x},\mathfrak{y},\mathfrak{z},u,p,q,\kappa)\\
=&-p\Big(\mathcal{D}_1\mathfrak{x}+\mathcal{D}_2\nabla\mathfrak{x}+\mathcal{E}_1\mathfrak{y}+\mathcal{E}_2\nabla\mathfrak{y}+\mathcal{F}_1\mathfrak{z}+\mathcal{L}u\Big)\\
&+q\Big(\mathcal{A}_1\mathfrak{x}+\mathcal{A}_2\nabla\mathfrak{x}+\mathcal{B}_1\mathfrak{y}+\mathcal{B}_2\nabla\mathfrak{y}+\mathcal{C}_1\mathfrak{z}+\mathcal{J}u\Big)\\
&+\Big\langle\kappa,\mathcal{G}_1\mathfrak{x}+\mathcal{G}_2\nabla\mathfrak{x}+\mathcal{H}_1\mathfrak{y}+\mathcal{H}_2\nabla\mathfrak{y}+\mathcal{I}_1\mathfrak{z}+\mathcal{K}u\Big\rangle\\
&+\Big\langle\mathcal{Q}\mathfrak{x},\mathfrak{x}\Big\rangle+\Big\langle\mathcal{R}u,u\Big\rangle+\Big\langle\mathcal{S}\mathfrak{y},\mathfrak{y}\Big\rangle+\Big\langle\mathcal{V}\mathfrak{z},\mathfrak{z}\Big\rangle.
\end{aligned}\tag{6.4}$$

Therefore, for any given admissible control pair $\big(X(\cdot,\cdot),Y(\cdot,\cdot),Z(\cdot,\cdot),u(\cdot,\cdot)\big)$, the dual equation can be written as follows:

$$\begin{cases}
\mathrm{d}p(t,x)=-\Big[-\mathcal{E}_1p(t,x)+\mathcal{E}_2^\top\nabla p(t,x)-\mathcal{E}_3\Delta p(t,x)+\mathcal{B}_1q(t,x)-\mathcal{B}_2^\top\nabla q(t,x)+\mathcal{B}_3\Delta q(t,x)\\
\qquad\qquad+\mathcal{H}_1\kappa(t,x)-\mathcal{H}_2^\top\nabla\kappa(t,x)+2\mathcal{S}Y(t,x)\Big]\mathrm{d}t-\Big[-\mathcal{F}_1^\top p(t,x)+\mathcal{F}_2^\top\nabla p(t,x)\\
\qquad\qquad+\mathcal{C}_1^\top q(t,x)-\mathcal{C}_2^\top\nabla q(t,x)+\mathcal{I}_1^\top\kappa(t,x)+2\mathcal{V}Z(t,x)\Big]\mathrm{d}W_t,\\
\mathrm{d}q(t,x)=-\Big[-\mathcal{D}_1p(t,x)+\mathcal{D}_2^\top\nabla p(t,x)+\mathcal{D}_3\Delta p(t,x)+\mathcal{A}_1q(t,x)-\mathcal{A}_2^\top\nabla q(t,x)\\
\qquad\qquad+\mathcal{A}_3\Delta q(t,x)+\mathcal{G}_1\kappa(t,x)-\mathcal{G}_2^\top\nabla\kappa(t,x)+2\mathcal{Q}X(t,x)\Big]\mathrm{d}t+\kappa(t,x)\mathrm{d}W_t,\\
p(0,x)=-2\mathcal{N}Y(0,x)-\phi q(0,x),\quad q(T,x)=2\mathcal{M}X(T,x)-\psi p(T,x).
\end{cases}\tag{6.5}$$

As $\mathscr{U}=\mathbb{R}^k$, the control is unconstrained, and then the minimum condition in Theorem 5.2 is as follows:

$$\mathscr{H}_u(t,x,X,Y,Z,u,p,q,\kappa)=0,\quad a.e.,(t,\omega,x)\in[0,T]\times\Omega\times\mathbb{R}^n,\quad u\in\mathscr{U}. \tag{6.6}$$

Based on the above, the stochastic Hamiltonian system is rewritten in linear form as

$$
\begin{cases}
\mathrm{d}X(t,x) = \Big[\mathcal{A}_1 X(t,x) + \mathcal{A}_2 \nabla X(t,x) + \mathcal{A}_3 \Delta X(t,x) + \mathcal{B}_1 Y(t,x) + \mathcal{B}_2 \nabla Y(t,x) + \mathcal{B}_3 \Delta Y(t,x) \\
\qquad\qquad + \mathcal{C}_1 Z(t,x) + \mathcal{C}_2 \nabla Z(t,x) + \mathcal{J} u(t,x)\Big]\mathrm{d}t + \Big[\mathcal{G}_1 X(t,x) + \mathcal{G}_2 \nabla X(t,x) + \mathcal{H}_1 Y(t,x) \\
\qquad\qquad + \mathcal{H}_2 \nabla Y(t,x) + \mathcal{I}_1 Z(t,x) + \mathcal{K} u(t,x)\Big]\mathrm{d}W_t, \\
\mathrm{d}Y(t,x) = -\Big[\mathcal{D}_1 X(t,x) + \mathcal{D}_2 \nabla X(t,x) - \mathcal{D}_3 \Delta X(t,x) + \mathcal{E}_1 Y(t,x) + \mathcal{E}_2 \nabla Y(t,x) + \mathcal{E}_3 \Delta Y(t,x) \\
\qquad\qquad + \mathcal{F}_1 Z(t,x) + \mathcal{F}_2 \nabla Z(t,x) + \mathcal{L} u(t,x)\Big]\mathrm{d}t + Z(t,x)\mathrm{d}W_t, \\
\mathrm{d}p(t,x) = -\Big[- \mathcal{E}_1 p(t,x) + \mathcal{E}_2^\top \nabla p(t,x) - \mathcal{E}_3 \Delta p(t,x) + \mathcal{B}_1 q(t,x) - \mathcal{B}_2^\top \nabla q(t,x) + \mathcal{B}_3 \Delta q(t,x) \\
\qquad\qquad + \mathcal{H}_1 \kappa(t,x) - \mathcal{H}_2^\top \nabla \kappa(t,x) + 2\mathcal{S} Y(t,x)\Big]\mathrm{d}t - \Big[- \mathcal{F}_1^\top p(t,x) + \mathcal{F}_2^\top \nabla p(t,x) \\
\qquad\qquad + \mathcal{C}_1^\top q(t,x) - \mathcal{C}_2^\top \nabla q(t,x) + \mathcal{I}_1^\top \kappa(t,x) + 2\mathcal{V} Z(t,x)\Big]\mathrm{d}W_t, \\
\mathrm{d}q(t,x) = -\Big[- \mathcal{D}_1 p(t,x) + \mathcal{D}_2^\top \nabla p(t,x) + \mathcal{D}_3 \Delta p(t,x) + \mathcal{A}_1 q(t,x) - \mathcal{A}_2^\top \nabla q(t,x) \\
\qquad\qquad + \mathcal{A}_3 \Delta q(t,x) + \mathcal{G}_1 \kappa(t,x) - \mathcal{G}_2^\top \nabla \kappa(t,x) + 2\mathcal{Q} X(t,x)\Big]\mathrm{d}t + \kappa(t,x)\mathrm{d}W_t, \\
X(0,x) = \phi Y(0,x), \quad Y(T,x) = \psi X(T,x), \\
p(0,x) = -2\mathcal{N} Y(0,x) - \phi q(0,x), \quad q(T,x) = 2\mathcal{M} X(T,x) - \psi p(T,x), \\
\mathscr{H}_u(t,x,X,Y,Z,u,p,q,\kappa) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}^n.
\end{cases}
\tag{6.7}
$$

The most important theorem of this section, which is concerned with the representation of optimal control, is given as follows:

Theorem 6.1 *Suppose Assumption 6.1 holds. Then, there exists a unique adapted solution* $(X(\cdot,\cdot), Y(\cdot,\cdot), Z(\cdot,\cdot), u(\cdot,\cdot), p(\cdot,\cdot), q(\cdot,\cdot), \kappa(\cdot,\cdot)) \in \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d}) \times \mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k) \times \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$ *to the stochastic Hamiltonian system and there is a unique optimal control* $u^*(\cdot,\cdot)$ *of Problem 6.1, with the following dual representation:*

$$
u^*(t,x) = -\frac{1}{2}\mathcal{R}^{-1}\Big(- \mathcal{L}^\top p^*(t,x) + \mathcal{J}^\top q^*(t,x) + \mathcal{K}^\top \kappa^*(t,x)\Big). \tag{6.8}
$$

Proof First, we prove the existence and uniqueness of the optimal control of Problem 6.1. It follows from Theorem 4.6 that the cost functional $J\big(u(\cdot,\cdot)\big)$ is continuous and consequently lower-semi continuous on $\mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)$. Since the control weight matrix process $\mathcal{R}$ is uniformly positive definite, $J\big(u(\cdot,\cdot)\big)$ is strictly convex. In addition, based on the nonnegative property of the operators $\mathcal{Q}, \mathcal{S}, \mathcal{V}, \mathcal{M}$ and $\mathcal{N}$ and the uniformly positive property of $\mathcal{R}$, we have

$$
J\big(u(\cdot,\cdot)\big) \geqslant L\mathbb{E}\int_0^T \int_{\mathbb{R}^n} \big|u(t,x)\big|^2 \mathrm{d}x\mathrm{d}t = L\big\|u(\cdot,\cdot)\big\|^2_{\mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)}. \tag{6.9}
$$

Therefore, we deduce that the cost functional $J(u(\cdot,\cdot))$ satisfies the coercive condition:

$$
\lim_{\|u(\cdot,\cdot)\|_{\mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)} \to \infty} J(u(\cdot,\cdot)) = \infty. \tag{6.10}
$$

In summary, we have verified that the cost functional $J(u(\cdot,\cdot))$ satisfies coercive, strictly convex, lower-semi continuous, and that the space $\mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)$ is clearly a reflexive Banach space. Then, by Proposition 2.12 of [9], we have the existence and uniqueness of the optimal control $u^*(\cdot,\cdot) \in \mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k)$ of Problem 6.1.

Next, we establish that there exists a unique adapted solution to the stochastic Hamiltonian system (6.7). As a consequence of Corollary 5.1, as long as the optimal control $u^*(\cdot,\cdot)$ in Problem 6.1 exists, then there exists a solution $(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot), u^*(\cdot,\cdot), p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot)) \in \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d}) \times \mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k) \times \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$ to the stochastic Hamiltonian system (6.7), with $(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot))$ being the optimal state, and $(p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot))$ being the adjoint process corresponding to the optimal control $u^*(\cdot,\cdot)$. According to Corollary 5.1, if there exists another adapted solution $(\bar{X}^*(\cdot,\cdot), \bar{Y}^*(\cdot,\cdot), \bar{Z}^*(\cdot,\cdot), \bar{u}^*(\cdot,\cdot), \bar{p}^*(\cdot,\cdot), \bar{q}^*(\cdot,\cdot), \bar{\kappa}^*(\cdot,\cdot)) \in \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d}) \times \mathscr{L}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^k) \times \mathscr{N}_{\mathbb{F}}^2(\mathbb{R}^n;\mathbb{R}^{2+d})$ to the stochastic Hamiltonian system (6.7), then $\bar{u}^*(\cdot,\cdot)$ must be the optimal control of Problem 6.1. However, due to the uniqueness of the optimal control, the assertion $\bar{u}^*(\cdot,\cdot) = u^*(\cdot,\cdot)$ is obtained. Moreover, in view of the uniqueness of the solution of the FBSPDE (see Lemma 5.1), we give the conclusion $(\bar{X}^*(\cdot,\cdot), \bar{Y}^*(\cdot,\cdot), \bar{Z}^*(\cdot,\cdot), \bar{p}^*(\cdot,\cdot), \bar{q}^*(\cdot,\cdot), \bar{\kappa}^*(\cdot,\cdot)) = (X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot), p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot))$. In summary, there exists a unique solution to the stochastic Hamiltonian system (6.7).

In the following, we derive the dual characterization of the optimal control. On the basis of Corollary 5.1, assume that the unique solution to the stochastic Hamiltonian system is $(X^*(\cdot,\cdot), Y^*(\cdot,\cdot), Z^*(\cdot,\cdot), p^*(\cdot,\cdot), q^*(\cdot,\cdot), \kappa^*(\cdot,\cdot))$, then $u^*(\cdot,\cdot)$ is the unique optimal control of Problem 6.1 while satisfying the following minimum condition

$$\mathscr{H}_u(t,x,X^*,Y^*,Z^*,u^*,p^*,q^*,\kappa^*) = 0, \text{ a.e. } t \in [0,T], \ x \in \mathbb{R}^n, \ \mathbb{P}\text{-a.s.} \tag{6.11}$$

After taking equation (6.4) into the above equation, one gets

$$2\mathcal{R}u^*(t,x) - \mathcal{L}^\top p^*(t,x) + \mathcal{J}^\top q^*(t,x) + \mathcal{K}^\top \kappa^*(t,x) = 0. \tag{6.12}$$

To this point, it is easy to derive (6.8). We have completed the proof. □

Acknowledgements

Qingxin Meng was supported by the Key Projects of Natural Science Foundation of Zhejiang Province (Grant No. LZ22A010005) and the National Natural Science Foundation of China (Grant Nos. 12271158 and 11871121).

References

[1] Al-Hussein, A. and Gherbal, B., Stochastic maximum principle for Hilbert space valued forward-backward doubly SDEs with Poisson jumps, In: Pötzsche, C., Heuberger, C., Kaltenbacher, B. and Rendl, F.(eds.), System Modeling and Optimization: 26th IFIP TC 7 Conference, CSMO 2013, Springer, Berlin, Heidelberg, 2014: 1–10.

[2] Antonelli, F., Backward-forward stochastic differential equations, The Annals of Applied Probability, 1993, 3(3): 777–793.

[3] Bensoussan, A., Stochastic maximum principle for distributed parameter systems, Journal of the Franklin Institute, 1983, 315(5–6): 387–406.

[4] Cardaliaguet, P., Delarue, F., Lasry, J.-M. and Lions, P.-L., The master equation and the convergence problem in mean field games: (AMS-201), Princeton University Press, 2019.

[5] Delarue, F. and Guatteri, G., Weak existence and uniqueness for forward-backward SDEs, Stochastic Processes and their Applications, 2006, 116(12): 1712–1742.

[6] Du, K. and Meng, Q., A maximum principle for optimal control of stochastic evolution equations, SIAM Journal on Control and Optimization, 2013, 51(6): 4343–4362.

[7] Duffie, D. and Epstein, L. G., Stochastic differential utility, Econometrica, 1992, 60(2): 353–394.

[8] Dunst, T. and Prohl, A., The forward-backward stochastic heat equation: Numerical analysis and simulation, SIAM Journal on Scientific Computing, 2016, 38(5): A2725–A2755.

[9] Ekeland, I. and Temam, R., Convex Analysis and Variational Problems, SIAM, Philadelphia, 1999.

[10] El Karoui, N., Peng, S. and Quenez, M. C., A dynamic maximum principle for the optimization of recursive utilities under constraints, The Annals of Applied Probability, 2001, 11(3): 664–693.

[11] Feng, C., Wang, X. and Zhao, H., Quasi-linear PDEs and forward-backward stochastic differential equations: Weak solutions, Journal of Differential Equations, 2018, 264(2): 959–1018.

[12] Fuhrman, M., Hu, Y. and Tessitore, G., Stochastic maximum principle for optimal control of SPDEs, Applied Mathematics & Optimization, 2013, 68: 181−217.

[13] Fuhrman, M. and Tessitore, G., Existence of optimal stochastic controls and global solutions of forward backward stochastic differential equations, SIAM Journal on Control and Optimization, 2004, 43(3): 813−830.

[14] Guatteri, G., On a class of forward-backward stochastic differential systems in infinite dimensions, Journal of Applied Mathematics and Stochastic Analysis, 2007, 2007: 042640.

[15] Guatteri, G. and Masiero, F., On the existence of optimal controls for SPDEs with boundary noise and boundary control, SIAM Journal on Control and Optimization, 2013, 51(3): 1909−1939.

[16] Hao, T. and Meng, Q., A global maximum principle for optimal control of general mean-field forward-backward stochastic systems with jumps, ESAIM: Control, Optimisation and Calculus of Variations, 2020, 26: 87.

[17] Hu, M., Stochastic global maximum principle for optimization with recursive utilities, Probability, Uncertainty and Quantitative Risk, 2017, 2(1): 1−20.

[18] Hu, Y., Ma, J. and Yong, J., On semi-linear degenerate backward stochastic partial differential equations, Probability Theory and Related Fields, 2002, 123(3): 381−411.

[19] Hu, Y. and Peng, S., Maximum principle for semilinear stochastic evolution control systems, Stochastics and Stochastic Reports, 1990, 33(3−4): 159−180.

[20] Hu, Y. and Peng, S., Solution of forward-backward stochastic differential equations, Probability Theory and Related Fields, 1995, 103: 273−283.

[21] Hui, E. C. M. and Xiao, H., Maximum principle for differential games of forward-backward stochastic systems with applications, Journal of Mathematical Analysis and Applications, 2012, 386(1): 412−427.

[22] Ji, S. and Liu, H., Maximum principle for stochastic optimal control problem of forward-backward stochastic difference systems, International Journal of control, 2022, 95(7): 1979−1992.

[23] Li, J. and Wei, Q., Stochastic differential games for fully coupled FBSDEs with jumps, Applied Mathematics & Optimization, 2015, 71(3): 411−448.

[24] Lenhart, S., Xiong, J. and Yong, J., Optimal controls for stochastic partial differential equations with an application in population modeling, SIAM Journal on Control and Optimization, 2016, 54(2): 495−535.

[25] Li, J., Fully coupled forward-backward stochastic differential equations with general martingale, Acta Mathematica Scientia, 2006, 26(3): 443−450.

[26] Li, J. and Peng, S., Stochastic optimization theory of backward stochastic differential equations with jumps and viscosity solutions of Hamilton-Jacobi-Bellman equations, Nonlinear Analysis: Theory, Methods & Applications, 2009, 70(4): 1776−1796.

[27] Li, N. and Yu, Z., Forward-backward stochastic differential equations and linear-quadratic generalized stackelberg games, SIAM Journal on Control and Optimization, 2018, 56(6): 4148−4180.

[28] Lü, Q. and Zhang, X., General Pontryagin-Type Stochastic Maximum Principle and Backward Stochastic Evolution Equations in Infinite Dimensions, Springer, Cham, 2014.

[29] Ma, J., Protter, P. and Yong, J., Solving forward-backward stochastic differential equations explicitly−a four step scheme, Probability Theory and Related Fields, 1994, 98(3): 339−359.

[30] Ma, J., Wu, Z., Zhang, D. and Zhang, J., On well-posedness of forward-backward SDEs–a unified approach, The Annals of Applied Probability, 2015, 25(4): 2168−2214.

[31] Ma, J. and Yong, J., On linear, degenerate backward stochastic partial differential equations, Probability Theory and Related Fields, 1999, 113: 135−170.

[32] Mahmudov, N. I. and McKibben, M. A., On backward stochastic evolution equations in Hilbert spaces and optimal control, Nonlinear Analysis: Theory, Methods & Applications, 2007, 67(4): 1260−1274.

[33] Meng, Q., A maximum principle for optimal control problem of fully coupled forward-backward stochastic systems with partial information, Science in China Series A: Mathematics, 2009, 52(7): 1579−1588.

[34] Meng, Q. and Shi, P., Stochastic optimal control for backward stochastic partial differential systems, Journal of Mathematical Analysis and Applications, 2013, 402(2): 758−771.

[35] Meng, Q. and Tang, M., A variational formula for controlled backward stochastic partial differential equations and some application, Applied Mathematics−A Journal of Chinese Universities, 2014, 29(3): 295−306.

[36] Molla, H. U. and Qiu, J., Numerical approximations of coupled forward-backward SPDEs, Stochastic Analysis and Applications, 2023, 41(2): 291−326.

[37] Øksendal, B. and Sulem, A., Maximum principles for optimal control of forward-backward stochastic differential equations with jumps, SIAM Journal on Control and Optimization, 2010, 48(5): 2945−2976.

[38] Pardoux, E. and Tang, S., Forward-backward stochastic differential equations and quasilinear parabolic PDEs, Probability Theory and Related Fields, 1999, 114: 123−150.

[39] Pardouxt, E., Stochastic partial differential equations and filtering of diffusion processes, Stochastics, 1980, 3(1−4): 127−167.

[40] Peng, S., A general stochastic maximum principle for optimal control problems, SIAM Journal on Control and Optimization, 1990, 28(4): 966−979.

[41] Peng, S., Backward stochastic differential equations and applications to optimal control, Applied Mathematics and Optimization, 1993, 27(2): 125–144.

[42] Peng, S. and Shi, Y., Infinite horizon forward-backward stochastic differential equations, Stochastic Processes and their Applications, 2000, 85(1): 75–92.

[43] Peng, S. and Wu, Z., Fully coupled forward-backward stochastic differential equations and applications to optimal control, SIAM Journal on Control and Optimization, 1999, 37(3): 825–843.

[44] Prévôt, C. and Röckner, M., A Concise Course on Stochastic Partial Differential Equations, Springer, Berlin, Heidelberg, 2007.

[45] Shi, J. and Wu, Z., The maximum principle for fully coupled forward-backward stochastic control system, Acta Automatica Sinica, 2006, 32(2): 161.

[46] Stannat, W. and Wessels, L., Peng's maximum principle for stochastic partial differential equations, SIAM Journal on Control and Optimization, 2021, 59(5): 3552–3573.

[47] Tang, M., Meng, Q. and Wang, M., Forward and backward mean-field stochastic partial differential equation and optimal control, Chinese Annals of Mathematics, Series B, 2019, 40(4): 515–540.

[48] Tang, S. and Li, X., Maximum principle for optimal control of distributed parameter stochastic systems with random jumps, Lecture Notes in Pure and Applied Mathematics, 1994, 152: 863–891.

[49] Wang, G., Wu, Z. and Xiong, J., Maximum principles for forward-backward stochastic control systems with correlated state and observation noises, SIAM Journal on Control and Optimization, 2013, 51(1): 491–524.

[50] Wang, G., Wu, Z. and Xiong, J., A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information, IEEE Transactions on Automatic Control, 2015, 60(11): 2904–2916.

[51] Wang, G., Wu, Z. and Xiong, J., An Introduction to Optimal Control of FBSDE with Incomplete Information, Springer, Cham, 2018.

[52] Wang, G. and Xiao, H., Arrow sufficient conditions for optimality of fully coupled forward-backward stochastic differential equations with applications to finance, Journal of Optimization Theory and Applications, 2015, 165: 639–656.

[53] Wu, Z., Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems, Systems Science and Mathematical Sciences, 1998, 11(3): 249–259.

[54] Wu, Z., A general maximum principle for optimal control of forward-backward stochastic systems, Automatica, 2013, 49(5): 1473–1480.

[55] Wu, Z. and Xu, M., Comparison theorems for forward backward SDEs, Statistics & Probability Letters, 2009, 79(4): 426–435.

[56] Yin, H., Solvability of forward-backward stochastic partial differential equations, Stochastic Processes and their Applications, 2014, 124(8): 2583–2604.

[57] Yin, H., Forward-backward stochastic partial differential equations with non-monotonic coefficients, Stochastics and Dynamics, 2016, 16(6): 1650025.

[58] Yin, J. and Wang, Y., Hilbert space-valued forward-backward stochastic differential equations with poisson jumps and applications, Journal of Mathematical Analysis and Applications, 2007, 328(1): 438–451.

[59] Yong, J., A stochastic linear quadratic optimal control problem with generalized expectation, Stochastic Analysis and Applications, 2008, 26(6): 1136–1160.

[60] Yong, J., Forward-backward stochastic differential equations with mixed initial-terminal conditions, Transactions of the American Mathematical Society, 2010, 362(2): 1047–1096.

[61] Yong, J., Optimality variational principle for controlled forward-backward stochastic differential equations with mixed initial-terminal conditions, SIAM Journal on Control and Optimization, 2010, 48(6): 4119–4156.

[62] Yu, Z., Linear-quadratic optimal control and nonzero-sum differential game of forward-backward stochastic system, Asian Journal of Control, 2012, 14(1): 173–185.

[63] Yu, Z., On forward-backward stochastic differential equations in a domination-monotonicity framework, Applied Mathematics & Optimization, 2022, 85(1): 5.

[64] Zhang, L., Singular optimal controls of stochastic recursive systems and Hamilton-Jacobi-Bellman inequality, Journal of Differential Equations, 2019, 266(10): 6383–6425.

[65] Zhang, S., Xiong, J. and Liu, X., Stochastic maximum principle for partially observed forward-backward stochastic differential equations with jumps and regime switching, Science China Information Sciences, 2018, 61: 1–13.

[66] Zheng, Y. and Shi, J., The maximum principle for discounted optimal control of partially observed forward backward stochastic systems with jumps on infinite horizon, ESAIM: Control, Optimisation and Calculus of Variations, 2023, 29: 34.

[67] Zhou, X., On the necessary conditions of optimal controls for stochastic partial differential equations, SIAM Journal on Control and Optimization, 1993, 31(6): 1462–1478.