Relationship between maximum principle and dynamic programming principle for recursive optimal control problem of stochastic evolution equations

Ying Hu, Guomin Liu, Shanjian Tang

Relationship between maximum principle and dynamic programming principle for recursive optimal control problem of stochastic evolution equations

Ying Hu* Guomin Liu† Shanjian Tang‡

Abstract

This paper aims to study the relationship between the maximum principle (MP) and the dynamic programming principle (DPP) for recursive optimal control problem of stochastic evolution equations, where the control domain is not necessarily convex and the value function can be non-smooth. By making use the notions of the super-, sub-differentials and the conditional expected operator-valued BSIEs, we establish the connection between the first and second-order adjoint variables in MP and the value funtion in DPP. Moreover, the discussions in the smooth case are also presented.

Keywords. Stochastic evolution equations, nonconvex control domain, recursive optimal control, maximum principle, dynamic programming principle.

AMS 2020 Subject Classifications. 93E20, 60H15, 49K27.

1 Introduction

Pontryagin's maximum principle (MP) and Bellman's dynamic programming principle (DPP) are two main approaches in solving optimal control problems. So besides studying them separately, it is important to ask what is the relationship between them (mainly the relationship between the adjoint variable in MP and the value funtion in DPP, which play the key roles in these two results, respectively).

The relationship between the MP and the DPP for controlled ordinary differential equations was first studied by Pontryagin, Boltyanskii, Gamkrelidze and Mishchenko [43], under the assumption that the value function is continuous differentiable. By making use the notion of viscosity solutions, the relationship without smoothness assumption on the value function was established by a series of works; see Barron and Jensen [1], Clarke and Vinter [8] and Zhou [51]. After that, Cannarsa and Frankowska [3, 4] and Cernea and Frankowska [5] generalized this result to control systems governed by partial differential equations. As far as for controlled stochastic differential equations, Peng [38] first present a general (i.e., when the control domain is nonconvex) MP for conventional utility case, by introducing a second-order adjoint process which solves a matrix-valued BSDE, and the relationship was established in the smooth case by Bensoussan [2] and in the nonsmooth case by Zhou [52, 53]. On the other hand, for the on the infinite dimensional stochastic systems for nonconvex control domain under the conventional (non-recursive) utilities, [11, 17, 30] studied the MP, and Chen and Lü [6] recently established the relationship between the MP and the DPP for SEEs.

*Institut de Recherche Mathématique de Rennes, Université Rennes 1, 35042 Rennes Cedex, France. ying.hu@univ-rennes1.fr. This author's research is partially supported by Lebesgue Center of Mathematics "Investissements d'avenir" Program (No. ANR-11-LABX-0020-01), by ANR CAESARS (No. ANR-15-CE05-0024) and by ANR MFG (No. ANR-16-CE40-0015-01).

†School of Mathematical Sciences, Nankai University, Tianjin, China. gmliu@nankai.edu.cn. Research supported by National Natural Science Foundation of China (No. 12201315) and the Fundamental Research Funds for the Central Universities, Nankai University (No. 63221036).

‡School of Mathematical Sciences, Fudan University, Shanghai, China. sjtang@fudan.edu.cn. Research supported by National Key R&D Program of China (No. 2018YFA0703900) and National Natural Science Foundation of China (No. 11631004 and No. 12031009).

The main objective of this paper is to establish the connection between the MP and the DPP for the following controlled SEE with different initial time and values:

$$\begin{cases} dX(t) &= \Big[A(s)X(s) + a(s, X(s), u(s))\Big]ds + \Big[B(s)X(s) + b(s, X(s), u(s))\Big]dw(s), \\ X(0) &= x, \end{cases} \tag{1.1}$$

where $w(\cdot)$ is a cylindrical Q-Brownian motion, (A, B) are unbounded linear operators (a, b) are nonlinear functions, and $u(\cdot)$ is a control process taking values in a given metric space. The diffusion coefficient depends on the control variable and the control domain is not necessarily convex, and the value function is not assumed to be smooth. The cost functional is defined by

$$J(x; u(\cdot)) := Y(t),$$

where $Y^{t,x;u}$ is the recursive utility subject to a BSDE:

$$Y(s) = h(X(T)) + \int_s^T k(r, X(r), Y(r), Z(r), u(r))dr - \int_s^T Z(r)dw(r), \quad s \in [t, T]. \tag{1.2}$$

The notion of a recursive utility in continuous time was introduced by Duffie and Epstein [12] and generalized to the form of (1.2) in Peng [40] and El Karoui, Peng and Quenez [14]. It represents a stochastic differential formulation of recursive utility which is an extension of the standard additive utility with the instantaneous utility. Stochastic recursive optimal control problems have found important applications in mathematical economics, mathematical finance and engineering (see, e.g., El Karoui, Peng and Quenez [14]). When k is invariant with (y, z), by taking expectation on both sides of (1.2), we get

$$J(x; u(\cdot)) = \mathbb{E}\Big[h(X(T)) + \int_t^T k(t, X(s), u(s))ds\Big],$$

and the stochastic recursive optimal control problem is reduced to the conventional one studied in [6].

As for the stochastic recursive optimal control problems for finite dimensional systems, Peng [40] first obtained a local MP when the control domain is convex. Recently, Hu [22] obtained a general MP for the stochastic recursive optimal control problem, which solves a long-standing open problem proposed by Peng [41] in the stochastic control theory; on the other hand, Peng [39, 49] obtained the DPP for the recursive controlled systems and introduced the generalized Hamilton-Jacobi-Bellman (HJB) equation. Concerning the connection between the MP and the DPP for stochastic recursive optimal control problem of stochastic differential equations, Shi [45] and Shi and Yu [46] investigated the local case in which the control domain is convex and the value function is smooth; within the framework of viscosity solution, Nie, Shi and Wu [33] investigated the local case when the domain of the control is convex; Nie, Shi and Wu [34] studied the general case when the domain of the control is non-convex.

Recently, two of the authors [27] established the MP for the recursive optimal control problem of SEEs, where the control domain is a general metric space (not necessarily convex). The present paper is a consequentive work of that, which studies the connection between the MP and the DPP for the control system (1.1). We consider this problem under the framework of variational frameworks, which are introduced in Pardoux [35, 36] and further developed by Krylov and Rozovskii [26]) and Gyöngy [20]. A main tool we use in this paper is the operator-valued conditionally BSIE introduced in [27], which served as the second-order adjoint equation. To carry out our purpose, we shall derive an Itô's formula for the second-order conditionally BSIE and the variational equations of the state equations with initial value perturbations, which plays an important role in the derivation of the relationship. Apart from the fact that we consider the more general recursive utility case, compared with the one in [6], our equation has an extra unobunded operator in the diffusion term and the unbounded operators (in both drift and diffusion terms) can be time-varying.

The rest of this paper is organized as follows. In Section 2, we recall the MP for recursive control problem of SEEs. In Section 3, we derive the DPP and present the relationship for MP and DPP non-smooth case. We discuss the special smooth case in Section 4. In the appendix, some technical results are proved.

2 Preliminaries and problem formulation

In this section, we recall the maximum principle (MP for short) for the recursive optimal control problem for stochastic evolution equations (SEEs for short); more details can be found in [27].

2.1 Conditionally expected BSIEs

We first present the notion of conditionally expected BSIEs, which will served as the second-adjoint equations for optimally controlled stochastic evolution equations (SEEs).

2.1.1 Spaces and norms

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Fix a terminal time $T > 0$, let $\mathbb{F} := \{\mathcal{F}_t\}_{0\leq t\leq T}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions. We denote by $\|\cdot\|_X$ the norm on a Banach space X. By $\mathfrak{L}(X;Y)$, we denote the space of all bounded linear operators from X to another Banach space Y, equipped with the operator norm. We write $\mathfrak{L}(X)$ for $\mathfrak{L}(X;X)$.

Let H be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle$. We adopt the standard identification viewpoint of $\mathfrak{L}(H;\mathbb{R}) = H$. By M^*, we denote the adjoint of an operator M. We denote by I_d the identity operator on H.

Given a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. For $\alpha \geq 1$, we denote by $L^\alpha(\mathcal{G}, H)$ the space of H-valued $\mathcal{G}$-measurable mapping y with norm $\|y\|_{L^\alpha(\mathcal{G},H)} = \{\mathbb{E}[\|y\|_H^\alpha]\}^{\frac{1}{\alpha}}$, and by $L^\alpha_{\mathbb{F}}(0,T;H)$ (resp. $L^{2,\alpha}_{\mathbb{F}}(0,T;H)$, $L^{1,2\alpha}_{\mathbb{F}}(0,T;H)$) the space of H-valued progressively measurable processes $y(\cdot)$ with norm $\|y\|_{L^\alpha_{\mathbb{F}}(0,T;H)} = \{\mathbb{E}[\int_0^T \|y(t)\|_H^\alpha dt]\}^{\frac{1}{\alpha}}$ (resp. $\|y\|_{L^{2,\alpha}_{\mathbb{F}}(0,T;H)} = \{\mathbb{E}[(\int_0^T \|y(t)\|_H^2 dt)^{\frac{\alpha}{2}}]\}^{\frac{1}{\alpha}}$, $\|y\|_{L^{1,2\alpha}_{\mathbb{F}}(0,T;H)} = \{\mathbb{E}[(\int_0^T \|y(t)\|_H dt)^{2\alpha}]\}^{\frac{1}{2\alpha}}$). We write $L^\alpha(\mathcal{G})$, $L^\alpha_{\mathbb{F}}(0,T)$, $L^{2,\alpha}_{\mathbb{F}}(0,T)$ and $L^{1,2\alpha}_{\mathbb{F}}(0,T;H)$ for $L^\alpha(\mathcal{G},\mathbb{R})$, $L^\alpha_{\mathbb{F}}(0,T;\mathbb{R})$, $L^{2,\alpha}_{\mathbb{F}}(0,T;\mathbb{R})$ and $L^{1,2\alpha}_{\mathbb{F}}(0,T;\mathbb{R})$, respectively.

Since the operator space $\mathfrak{L}(H)$ is not separable (see [21, Solution 99]), we make use of the following weak measurability notion (see, e.g., [25]) for random variables taking values in it. We say a mapping $Z : \Omega \to \mathfrak{L}(H)$ is weakly $\mathcal{G}$-measurable if for each $(u,v) \in H \times H$, $\langle Zu, v\rangle : \Omega \to \mathbb{R}$ is $\mathcal{G}$-measurable. A process $Y : \Omega \times [0,T] \to \mathfrak{L}(H)$ is said to be weakly progressively measurable (resp. weakly adapted) if for each $(u,v) \in H \times H$, the process $\langle Yu, v\rangle : \Omega \times [0,T] \to \mathbb{R}$ is progressively measurable (resp. adapted).

By $L^\alpha_w(\mathcal{G}, \mathfrak{L}(H))$, we denote the space of $\mathfrak{L}(H)$-valued weakly $\mathcal{G}$-measurable mapping F with norm $\|F\|_{L^\alpha_w(\mathcal{G},\mathfrak{L}(H))} = \{\mathbb{E}[\|F\|^\alpha_{\mathfrak{L}(H)}]\}^{\frac{1}{\alpha}}$. Similarly, we denote by $L^\alpha_{\mathbb{F},w}(0,T;\mathfrak{L}(H))$ (resp. $L^{2,\alpha}_{\mathbb{F},w}(0,T;\mathfrak{L}(H))$) the space of $\mathfrak{L}(H)$-valued weakly progressively measurable processes $F(\cdot)$ with the norm $\|F\|_{L^\alpha_{\mathbb{F},w}(0,T;\mathfrak{L}(H))} = \{\mathbb{E}[\int_0^T \|F(t)\|^\alpha_{\mathfrak{L}(H)} dt]\}^{\frac{1}{\alpha}}$ (resp. $\|F\|_{L^{2,\alpha}_{\mathbb{F},w}(0,T;\mathfrak{L}(H))} = \{\mathbb{E}[(\int_0^T \|F(t)\|^2_{\mathfrak{L}(H)} dt)^{\frac{\alpha}{2}}]\}^{\frac{1}{\alpha}}$). In the following, we shall not distinguish two random variables if they coincide P-a.s. and two processes if one is a modification of the other, unless otherwise stated.

Denote by L_w the weak σ-algebra on $\mathfrak{L}(H)$ generated by all the sets in the form of

$$\Big\{z \in \mathfrak{L}(H) : \langle zu, v\rangle \in A\Big\}, \quad u, v \in H,\ A \in \mathcal{B}(\mathbb{R}).$$

Then $Z : \Omega \to \mathfrak{L}(H)$ is weakly $\mathcal{G}$-measurable if and only if it is measurable from $(\Omega, \mathcal{G})$ to $(\mathfrak{L}(H), L_w)$. Similarly, $Y : \Omega \times [0,T] \to \mathfrak{L}(H)$ is weakly progressively measurable if and only if it is measurable from $(\Omega \times [0,T], \mathcal{P})$ to $(\mathfrak{L}(H), L_w)$, where $\mathcal{P}$ is the progressive σ-algebra on $\Omega \times [0,T]$.

2.1.2 Operator-valued conditional expectations

For $Y \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}))$, we call an $\mathfrak{L}_2(H \times H)$-valued weakly $\mathcal{G}$-measurable mapping Z the conditional expectation of Y with respect to $\mathcal{G}$, denoting it by $\mathbb{E}[Y|\mathcal{G}]$, if for each $(u, v) \in H \times H$,

$$Z(u, v) = \mathbb{E}[Y(u, v)|\mathcal{G}], \quad P\text{-a.s.}, \tag{2.1}$$

meaning that Z coincides with the classical conditional expectation at all the test points (u, v).

Theorem 2.1 *Let $Y \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}))$. Then the conditional expectation $\mathbb{E}[Y|\mathcal{G}]$ exists and is integrable (i.e., $\mathbb{E}[Y|\mathcal{G}] \in L^1_w(\mathcal{G}, \mathfrak{L}_2(H \times H))$) if and only if the mapping $(u, v) \longmapsto \mathbb{E}[Y(u, v)|\mathcal{G}] \in \mathfrak{L}_2(H \times H; L^1(\mathcal{G}))$ satisfies the domination condition*

$$\Big|\mathbb{E}[Y(u, v)|\mathcal{G}]\Big| \leq g\|u\|_H\|v\|_H, \quad P\text{-a.s., } \forall (u, v) \in H \times H, \tag{2.2}$$

for some $0 \leq g \in L^1(\mathcal{G})$. Moreover, such an $\mathbb{E}[Y|\mathcal{G}]$ is unique (up to P-a.s. equality) and satisfies

$$\Big\|\mathbb{E}[Y|\mathcal{G}]\Big\|_{\mathfrak{L}_2(H\times H)} \leq g, \quad P\text{-a.s.} \tag{2.3}$$

From $\mathfrak{L}_2(H \times H) = \mathfrak{L}(H)$, we can also write (2.1) as, for weakly $\mathcal{G}$-measurable Z taking values in $\mathfrak{L}(H) = \mathfrak{L}_2(H \times H)$ and $Y \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}))$,

$$\langle Zu, v\rangle = Z(u, v) = \mathbb{E}[Y(u, v)|\mathcal{G}], \quad P\text{-a.s., } \forall (u, v) \in H \times H. \tag{2.4}$$

2.1.3 Operator-valued BSIEs

By a stochastic evolution operator on H, we mean a family of mappings

$$\Big\{L(t, s) \in \mathfrak{L}(L^2(\mathcal{F}_t, H); L^2(\mathcal{F}_s, H)) : (t, s) \in \Delta\Big\}$$

with $\Delta = \{(t, s) : 0 \leq t \leq s \leq T\}$. We adopt a definition of the following formal adjoint L^* for L: For any fixed $(t, s) \in \Delta$ and $u \in L^1(\mathcal{F}_s, H)$, define $L^*(t, s)u$ by

$$\big(L^*(t, s)u\big)(v) := \langle u, L(t, s)v\rangle \quad P\text{-a.s.}, \quad \text{for each } v \in L^2(\mathcal{F}_t, H).$$

We consider a conditionally expected $\mathfrak{L}(H)$-valued BSIE (i.e., $\mathfrak{L}(H)$-valued BSIE in the conditional expectation form):

$$P(t) = \mathbb{E}\Big[L^*(t, T)\xi L(t, T) + \int_t^T L^*(t, s)f(s, P(s))L(t, s)ds\Big|\mathcal{F}_t\Big], \quad t \in [0, T], \tag{2.5}$$

where the coefficients ξ, f and L are given and subject to the following assumptions:

$(H1)$ There exists some constant $\Lambda \geq 0$ such that for each $(t, s) \in \Delta$ and $u \in L^4(\mathcal{F}_t, H)$, it holds that $L(t, s)u \in L^4(\mathcal{F}_s, H)$,

$$\mathbb{E}\Big[\ \|L(t, s)u\|_H^4\ \Big|\mathcal{F}_t\Big] \leq \Lambda\|u\|_H^4, \quad P\text{-a.s.},$$

and $(\omega, t, s) \mapsto (L(t, s)u)(\omega)$ admits a jointly measurable version.

$(H2)$ $\xi \in L^2_w(\mathcal{F}_T, \mathfrak{L}(H))$; the function $f(w, t, p) : \Omega \times [0, T] \times \mathfrak{L}(H) \to \mathfrak{L}(H)$ is $\mathcal{P} \otimes L_w/L_w$-measurable and satisfies the Lipschitz condition in p with constant $\lambda \geq 0$; $f(\cdot, \cdot, 0) \in L^2_{\mathbb{F},w}(0, T; \mathfrak{L}(H))$.

Under the assumption $(H1)$, for any $\eta \in L^2_w(\mathcal{F}_s, \mathfrak{L}(H))$ and $(u,v) \in H \times H$, we have $L^*(t,s)\eta L(t,s) \in \mathfrak{L}(H; \mathfrak{L}(H; L^1(\mathcal{F}_s))) = \mathfrak{L}_2(H \times H; L^1(\mathcal{F}_s))$ and we can also write $(L^*(t,s)\eta L(t,s)u)(v) = L^*(t,s)\eta L(t,s)(u,v)$. In particular, $L^*(t,T)\xi L(t,T) \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}_T))$.

For a $g \in \mathfrak{L}_2(H \times H; L^1_{\mathbb{F}}(t,T))$, we define its integral $\int_t^T g(s)ds$ in a weak sense by

$$\Big(\int_t^T g(s)ds\Big)(u,v) := \int_t^T g(s)(u,v)ds \quad P\text{-a.s.}, \quad \forall (u,v) \in H \times H.$$

Then $\int_t^T g(s)ds \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}_T))$. Note that for any $h \in L^2_{\mathbb{F},w}(t,T; \mathfrak{L}(H))$, we have $[t,T] \ni s \mapsto L^*(t,s)h(s)L(t,s) \in \mathfrak{L}_2(H \times H; L^1_{\mathbb{F}}(t,T))$ and the integral $\int_t^T L^*(t,s)h(s)L(t,s)ds \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}_T))$ is defined. Given any $P \in L^2_{\mathbb{F},w}(0,T; \mathfrak{L}(H))$. we have $f(\cdot, P(\cdot)) \in L^2_{\mathbb{F},w}(0,T; \mathfrak{L}(H))$ and thus,

$$L^*(t,T)\xi L(t,T) + \int_t^T L^*(t,s)f(s,P(s))L(t,s)ds \in \mathfrak{L}_2(H \times H; L^1(\mathcal{F}_T)). \tag{2.6}$$

Definition 2.2 *A process $P \in L^2_{\mathbb{F},w}(0,T; \mathfrak{L}(H))$ is called a solution of (2.5) if for each $0 \le t \le T$,*

$$P(t) = \mathbb{E}\Big[L^*(t,T)\xi L(t,T) + \int_t^T L^*(t,s)f(s,P(s))L(t,s)ds\Big|\mathcal{F}_t\Big], \quad P\text{-a.s.} \tag{2.7}$$

We have the following well-posedness result on BSIEs. In the following of this paper, the constant C may change from line to line.

Theorem 2.3 *Let Assumptions $(H1)$ and $(H2)$ be satisfied. Then there exists a unique solution P to BSIE (2.5). Moreover, for each $t \in [0,T]$,*

$$\|P(t)\|^2_{\mathfrak{L}(H)} \le C\mathbb{E}\Big[\|\xi\|^2_{\mathfrak{L}(H)} + \int_t^T \|f(s,0)\|^2_{\mathfrak{L}(H)}ds\Big|\mathcal{F}_t\Big], \quad P\text{-a.s.}, \tag{2.8}$$

for some constant C depending on Λ and λ.

Proposition 2.4 *For some $\alpha \ge 1$, suppose $(H1)$, $(H2)$ and*

$(H3)$ *$(\xi, f(\cdot,\cdot,0)) \in L^{2\alpha}_w(\mathcal{F}_T, \mathfrak{L}(H)) \times L^{2,2\alpha}_{\mathbb{F},w}(0,T; \mathfrak{L}(H))$ and there exists some constant $\Lambda_\alpha \ge 0$ such that for each $0 \le t \le r \le s \le T$ and $u \in L^{4\alpha}(\mathcal{F}_t, H)$, it holds that $L(t,s) = L(t,r)L(r,s)$,*

$$\mathbb{E}[\|L(t,s)u\|^{4\alpha}_H |\mathcal{F}_t] \le \Lambda_\alpha \|u\|^{4\alpha}_H \ P\text{-a.s.} \ \text{and} \ [t,T] \ni s \mapsto L(t,s)u \ \text{is strongly continuous in} \ L^{4\alpha}(\mathcal{F}_T, H).$$

Let P be the solution of (2.5). Then, for each $t \in [0,T)$ and $u, v \in L^{4\alpha}(\mathcal{F}_t, H)$, we have

$$\lim_{\delta \downarrow 0} \mathbb{E}\Big[\Big|\langle P(t+\delta)u, v\rangle - \langle P(t)u, v\rangle\Big|^\alpha \Big|\mathcal{F}_t\Big] = 0, \ P\text{-a.s.}$$

Proof. The proof is similar as that for Proposition 2.11 in [27] and we shall not repeat that. □

2.1.4 Evolution operators associated to forward SEEs

The formal solution of forward operator-valued SEEs is a typical example of stochastic evolution operator.

Let V be a separable Hilbert space densely embedded in H. Denote $V^* := \mathfrak{L}(V; \mathbb{R})$, then $V \subset H \subset V^*$ form a Gelfand triple. We denote the dualization between V^* and V by $\langle \cdot, \cdot \rangle_*$.

Let $w := (w(t))_{t\geq 0}$ be a one-dimensional standard Brownian motion with respect to $\mathbb{F}$. Consider the following linear homogeneous SEE on $[t,T]$:

$$\begin{cases} du^{t,u_0}(s) &= A(s)u^{t,u_0}(s)dt + B(s)u^{t,u_0}(s)dw(s), \quad s\in[t,T], \\ u^{t,u_0}(t) &= u_0, \end{cases} \tag{2.9}$$

where $u_0 \in L^2(\mathcal{F}_t, H)$ and $(A,B) : [0,T]\times\Omega \to \mathfrak{L}(V; V^*\times H)$.

Remark 2.5 We only write the one-dimensional Brownian motion case for simplicity of presentation. With direct modifications, the results throughout this paper still hold for the more general case that w is a Hilbert space K-valued cylindrical Q-Brownian motion (including multi-dimensional Brownian motion, finite-trace Q-Brownian motion, cylinderical Brownian motion as special cases) and the integrands f takes valued in the Hilbert-Schmidt space $\mathcal{L}_2(Q^{\frac{1}{2}}(K), H)$; see [29, 28, 32] for more discussions on this direction.

We make the following assumption.

$(H4)$ For each $u\in V$, $A(t,\omega)u$ and $B(t,\omega)u$ are progressively measurable and satisfying: There exist some constants $\delta > 0$ and $K \geq 0$ such that the following two assertions hold: for each t,ω and $u\in V$,

(i) coercivity condition:

$$2\langle A(t,\omega)u, u\rangle_* + \|B(t,\omega)u\|_H^2 \leq -\delta\|u\|_V^2 + K\|u\|_H^2 \quad \text{and} \quad \|A(t,\omega)u\|_{V^*} \leq K\|u\|_V;$$

(ii) quasi-skew-symmetry condition:

$$\Big|\langle B(t,\omega)u, u\rangle\Big| \leq K\|u\|_H^2.$$

From [26], Equation (2.9) has a unique solution $u^{t,u_0}(\cdot) \in L^2_{\mathbb{F}}(t,T;V)\cap S^2_{\mathbb{F}}(t,T;H)$, where $S^2_{\mathbb{F}}(t,T;H)$ is the space of adapted H-valued processes y with continuous paths such that $\mathbb{E}[\sup_{t\leq s\leq T}\|y(s)\|_H^2] < \infty$. Through this solution, we define a stochastic evolution operator $L_{A,B}$ as follows:

$$L_{A,B}(t,s)(u_0) := u^{t,u_0}(s) \in L^2(\mathcal{F}_s, H), \quad \text{for } t\leq s\leq T \text{ and } u_0\in L^2(\mathcal{F}_t,H). \tag{2.10}$$

From the basic estimates for SEEs, it satisfies the assumptions $(H1)$ and $(H3)$, for any $\alpha \geq 1$. Thus, in virtue of Theorem 2.3, the $\mathfrak{L}(H)$-valued BSIE

$$P(t) = \mathbb{E}\Big[L^*_{A,B}(t,T)\xi L_{A,B}(t,T) + \int_t^T L^*_{A,B}(t,s)f(s,P(s))L_{A,B}(t,s)ds\Big|\mathcal{F}_t\Big], \quad t\in[0,T], \tag{2.11}$$

has a unique solution $P\in L^2_{\mathbb{F},w}(0,T;\mathfrak{L}(H))$.

In the following, we shall always assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ is the augmented natural filtration of Brownian motion $(w(t))_{t\geq 0}$.

2.2 A priori estimates for SEEs

Consider the following SEE

$$\begin{cases} dz(s) &= [A(s)z(s) + \tilde{a}(s,z(s))]ds + [B(s)z(s) + \tilde{b}(s,z(s))]dw(s), \quad s\in[t,T], \\ z(t) &= z_0, \end{cases} \tag{2.12}$$

With $\alpha \geq 1$, we made the following assumptions.

(**A1**) The function $(\tilde{a},\tilde{b}):[0,T]\times\Omega\times H\to V^*\times H$ satisfies, for each $z\in H$, $\tilde{a}(\cdot,\cdot,z)$, $\tilde{b}(\cdot,\cdot,z)$ are progressively measurable, $\tilde{a}(\cdot,\cdot,0)\in L^{2,2\alpha}_{\mathbb{F}}(t,T;V^*)$, $\tilde{b}(\cdot,\cdot,0)\in L^{2,2\alpha}_{\mathbb{F}}(t,T;H)$, and there exists a constant $K>0$ such that for almost all $(t,\omega)\in[0,T]\times\Omega$ and all $z,z'\in H$,

$$\|\tilde{a}(t,z)-\tilde{a}(t,z')\|_{V^*}+\|\tilde{b}(t,z)-\tilde{b}(t,z')\|_H\le K\|z-z'\|_H.$$

The proof of the following lemma is similar to Lemma 4.4 in [27] and is put in the appendix.

Lemma 2.6 *Assume (H4), (A1) hold and $z_0\in L^{2\alpha}(\mathcal{F}_t,H)$. Then there is a constant $C>0$ depending on δ, K and α such that*

$$\begin{aligned}&\mathbb{E}\Big[\sup_{s\in[t,T]}\|z(s)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|z(s)\|_V^2\,ds\Big)^{\alpha}\Big|\mathcal{F}_t\Big]\\&\le C\Big\{\|z_0\|_H^{2\alpha}+\mathbb{E}\Big[\Big(\int_t^T\|\tilde{a}(s,0)\|_{V^*}^2ds\Big)^{\alpha}+\Big(\int_t^T\|\tilde{b}(s,0)\|_H^2ds\Big)^{\alpha}\Big|\mathcal{F}_t\Big]\Big\},\ P\text{-a.s.}\end{aligned}$$

We first present an continuity estimate of SEEs in t, which will be needed later.

(**A2**) The function $(\tilde{a},\tilde{b}):[0,T]\times\Omega\times H\to V^*\times H$ satisfies, for each $z\in H$, $\tilde{a}(\cdot,\cdot,z)$, $\tilde{b}(\cdot,\cdot,z)$ are progressively measurable, $\tilde{a}(\cdot,\cdot,0)$, $\tilde{b}(\cdot,\cdot,0)$ are bounded, and there exists a constant $K>0$ such that for almost all $(t,\omega)\in[0,T]\times\Omega$ and all $z,z'\in H$,

$$\|\tilde{a}(t,z)-\tilde{a}(t,z')\|_{V^*}+\|\tilde{b}(t,z)-\tilde{b}(t,z')\|_H\le K\|z-z'\|_H.$$

Lemma 2.7 *Suppose (H4), (A2) and $z_0\in L^{2\alpha}(\mathcal{F}_t,V)$. Then for $\rho\le T-t$,*

$$\mathbb{E}\Big[\sup_{t\le s\le t+\rho}\|z(s)-z_0\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]\le C\Big(1+\|z_0\|_V^{2\alpha}\Big)\rho^{\alpha},\ P\text{-a.s.},$$

where $C>0$ is a constant depends on δ, K and α.

Proof. We denote

$$\hat{z}(s):=z(s)-z_0,\ s\in[t,T].$$

Then on $[t,T]$,

$$\begin{aligned}\hat{z}(s)&=\int_t^s[A(r)z(r)+\tilde{a}(r,z(r))]dr+\int_t^s[B(r)z(r)+\tilde{b}(r,z(r))]dw(r)\\&=\int_t^s[A(r)\hat{z}(r)+Az_0+\tilde{a}(r,\hat{x}(r)+z_0)]dr+\int_t^s[B(r)\hat{z}(r)+Bz_0+\tilde{b}(r,\hat{z}(r)+z_0)]dw(r).\end{aligned}$$

So from Lemma 2.6,

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\le s\le t+\rho}\|\hat{z}(s)\|_H^{2\alpha}\,ds\Big|\mathcal{F}_t\Big]&\le C\mathbb{E}\Big[\Big(\int_t^{t+\rho}\|A(r)z_0+\tilde{a}(r,z_0)\|_{V^*}^2dr\Big)^{\alpha}|\mathcal{F}_t\Big]\\&\quad+C\mathbb{E}\Big[\Big(\int_t^{t+\rho}\|B(r)z_0+\tilde{b}(r,z_0)\|_H^2dr\Big)^{\alpha}|\mathcal{F}_t\Big]\\&\le C\rho^{\alpha-1}\int_t^{t+\rho}(1+\|z_0\|_V^{2\alpha})dr\\&\le C(1+\|z_0\|_V^{2\alpha})\rho^{\alpha},\ P\text{-a.s.}\end{aligned}$$

□

Remark 2.8 *It seems hard to prove such a kind of estimate if $z_0 \in H$. At this case, we can only show that*

$$\mathbb{E}\Big[\sup_{t\le s\le t+\rho}\|z(s)-z_0\|_H^{2\alpha}\,\Big|\mathcal{F}_t\Big]\to 0,\ P\text{-a.s., as } \rho\downarrow 0.$$

This follows trivially from the continuity of z with respect to s in the H-norm and the dominated convergence theorem.

2.3 Stochastic maximum principle

Consider the following controlled SEE:

$$\begin{cases} dX(t) &= \Big[A(t)X(t)+a(t,X(t),u(t))\Big]dt+\Big[B(t)X(t)+b(t,X(t),u(t))\Big]dw(t),\\ X(0) &= x_0, \end{cases} \tag{2.13}$$

where $x_0 \in H$,

$$(A,B):[0,T]\to\mathfrak{L}(V;V^*\times H)$$

are unbounded linear operators satisfying the coercivity and quasi-skew-symmetry condition $(H4)$ and

$$(a,b):[0,T]\times H\times U\to H\times H$$

are nonlinear functions. Define the cost functional $J(\cdot)$ as

$$J(x,u(\cdot)):=y(0),$$

where y is the recursive utility subject to a BSDE:

$$Y(t)=h(X(T))+\int_t^T k(s,X(s),Y(s),Z(s),u(s))ds-\int_t^T Z(s)dw(s). \tag{2.14}$$

Here,

$$k:[0,T]\times H\times\mathbb{R}\times\mathbb{R}\times U\to\mathbb{R}\quad\text{and}\quad h:H\to\mathbb{R}.$$

The control domain U is a separable metric space with distance $d(\cdot,\cdot)$. By fixing an element 0 in U, we define the length $|u|_U:=d(u,0)$. We define the admissible control set

$$\mathcal{U}[0,T]:=\Big\{u:[0,T]\times\Omega\to U \text{ is progressively measurable and } \mathbb{E}\Big[\int_0^T|u(t)|_U^\alpha dt\Big]<\infty,\ \text{for each } \alpha\ge 1\Big\}.$$

For $t\in[0,T)$, we define similarly $\mathcal{U}[t,T]$.

The optimal control Problem (S_x) is to find an admissible control $\bar{u}(\cdot)$ such that the cost functional $J(x,u(\cdot))$ is minimized at $\bar{u}(\cdot)$ over the control set $\mathcal{U}[0,T]$:

$$J(x,\bar{u}(\cdot))=\inf_{u(\cdot)\in\mathcal{U}[0,T]}J(x,u(\cdot)).$$

We make the following assumption for a, b, h and k.

$(H5)$ a, b, h, k are twice Fréchet differentiable with respect to (x,y,z); a, b, k, a_x, b_x, Dk, a_{xx}, b_{xx}, D^2k are continuous in (x,y,z,u), where Dk and D^2k are the gradient and Hessian matrix of k with respect to (x,y,z), respectively; a_x, b_x, Dk, a_{xx}, b_{xx}, D^2k, h_{xx} are bounded; a, b are bounded by $C(1+\|x\|_H+|u|_U)$ and k is bounded by $C(1+\|x\|_H+|y|+|z|+|u|_U)$. $\|a(s,0,u_s)\|_H, \|b(s,0,u_s)\|_H$ are bounded by C.

We introduce the following simplified notations: for $\psi = a, b, a_x, b_x, a_{xx}, b_{xx}$ and $v \in U$, define

$$\bar{\psi}(t) := \psi(t, \bar{X}(t), \bar{u}(t)), \quad \delta\psi(t; v) := \psi(t, \bar{X}(t), v) - \bar{\psi}(t)$$

and

$$\bar{A} := A + \bar{a}_x, \quad \bar{B} := B + \bar{b}_x.$$

Consider the following first-order H-valued adjoint backward stochastic evolution equation (BSEE for short, and the well-posedness result is referred to [10]):

$$\begin{cases} -dp(t) = & \Big\{\Big[\bar{A}^*(t) + k_y(t) + k_z(t)\bar{B}^*(t)\Big]p(t) + \Big[\bar{B}^*(t) + k_z(t)\Big]q(t) + k_x(t)\Big\}dt - q(t)dw(t), \\ p(T) = & h_x(\bar{X}(T)), \end{cases} \tag{2.15}$$

and the following second-order $\mathfrak{L}(H)$-valued adjoint BSIE

$$P(t) = \mathbb{E}\Big[\tilde{L}^*(t,T)h_{xx}(\bar{X}(T))\tilde{L}(t,T) + \int_t^T \tilde{L}^*(t,s)(k_y(s)P(s) + G(s))\tilde{L}(t,s)ds\Big|\mathcal{F}_t\Big], \quad 0 \le t \le T, \tag{2.16}$$

where

$$\begin{aligned} &\phi(t) := \phi(t, \bar{X}(t), \bar{Y}(t), \bar{Z}(t), \bar{u}(t)), \quad \text{for } \phi = k_x, k_{y,}, k_z, D^2k, \\ &\tilde{L}(t,s) := L_{\bar{A},\bar{B}}(t,s), \quad \text{for } \tilde{A}(s) := \bar{A}(s) + \frac{k_z(s)}{2}\bar{B}(s) - \frac{(k_z(s))^2}{8}I_d \text{ and } \tilde{B}(s) := \bar{B}(s) + \frac{k_z(s)}{2}I_d, \\ &G(t) := D^2k(t)\Big(\big[I_d, p(t), \bar{B}^*(t)p(t) + q(t)\big], \big[I_d, p(t), \bar{B}^*(t)p(t) + q(t)\big]\Big) + \langle p(t), \bar{a}_{xx}(t)\rangle \\ &\qquad + k_z(t)\langle p(t), \bar{b}_{xx}(t)\rangle + \langle q(t), \bar{b}_{xx}(t)\rangle. \end{aligned}$$

The maximum principle is stated as follows.

Theorem 2.9 *Suppose (H4) and (H5). Assume that $\bar{X}(\cdot)$ and $(\bar{Y}(\cdot), \bar{Z}(\cdot))$ are the solutions of SEE (2.18) and BSDE (2.19) corresponding to the optimal control $\bar{u}(\cdot)$. Denote by processes $(p,q) \in L^2_{\mathbb{F}}(0,T;V \times H)$ and $P \in L^2_{\mathbb{F},w}(0,T;\mathfrak{L}(H))$ the solutions of BSEE (2.15) and BSIE (2.16), respectively. Then*

$$\begin{aligned} \inf_{v\in U}\Big\{&\mathcal{H}\big(t, \bar{X}(t), \bar{Y}(t), \bar{Z}(t), v, p(t), q(t)\big) - \mathcal{H}\big(t, \bar{X}(t), \bar{Y}(t), \bar{Z}(t), \bar{u}(t), p(t), q(t)\big) \\ &+ \frac{1}{2}\big\langle P(t)(b(t, \bar{X}(t), v) - b(t, \bar{x}(t), \bar{u}(t))), b(t, \bar{X}(t), v) - b(t, \bar{X}(t), \bar{u}(t))\big\rangle\Big\} = 0, \quad P\text{-a.s. a.e.,} \end{aligned} \tag{2.17}$$

where the Hamiltonian

$$\begin{aligned} \mathcal{H}(t,x,y,z,v,p,q) := \langle p, a(t,x,v)\rangle + \langle q, b(t,x,v)\rangle + k(t,x,y,z + \langle p, b(t,x,v) - b(t, \bar{X}(t), \bar{u}(t))\rangle, v), \\ (t,\omega,x,y,z,v,p,q) \in [0,T] \times \Omega \times H \times \mathbb{R} \times \mathbb{R} \times U \times H \times H. \end{aligned}$$

2.4 Dynamic programming principle

For any given $(t,x) \in [0,T] \times H$, we consider the following controlled SEE with different initial time and values:

$$\begin{cases} dX^{t,x;u}(t) &= \Big[A(s)X^{t,x;u}(s) + a(s, X^{t,x;u}(s), u(s))\Big]ds + \Big[B(s)X^{t,x;u}(s) + b(s, X^{t,x;u}(s), u(s))\Big]dw(s), \\ X^{t,x;u}(t) &= x. \end{cases} \tag{2.18}$$

The cost functional is defined by

$$J(t,x;u(\cdot)) := Y^{t,x;u}(t),$$

where $Y^{t,x;u}$ is the recursive utility subject to a BSDE:

$$Y^{t,x;u}(s) = h(X^{t,x;u}(T)) + \int_s^T k(r, X^{t,x;u}(r), Y^{t,x;u}(r), Z^{t,x;u}(r), u(r))dr - \int_s^T Z^{t,x;u}(r)dw(r), \quad s \in [t,T]. \tag{2.19}$$

The optimal control problem (Problem $(S_{t,x})$) is to find an admissible control $\bar{u}(\cdot)$ such that the cost functional $J(t,x;\bar{u}(\cdot))$ is minimized at $\bar{u}(\cdot)$ over the control set $\mathcal{U}[t,T]$:

$$J(t,x;\bar{u}(\cdot)) = \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} J(t,x;u(\cdot)).$$

Note that when $t=0$, $(S_{t,x})$ reduces to (S_x), and $(X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s)) = (X(s), Y(s), Z(s))$.

We define the value function

$$V(t,x) := \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} J(t,x;u(\cdot)),\ (t,x) \in [0,T]\times H. \tag{2.20}$$

We define $\mathcal{U}^t[t,T]$ the space of all U-valued $(\mathcal{F}_s^t)_{t\le s\le T}$-progressively measurable processes on $[t,T]$, where $\mathcal{F}_s^t$ is augmented natural filtration of $(w(s)-w(t))_{s\ge t}$. Then for each $u\in\mathcal{U}^t[t,T]$, it is easy to verify that the solution $(X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s))_{t\le s\le T}$ of the system is $(\mathcal{F}_s^t)_{t\le s\le T}$-adapted. In particular, $(X^{t,x;u}(t), Y^{t,x;u}(t), Z^{t,x;u}(t)) \in \mathcal{F}_t^t$, in particular of this, $Y^{t,x;u}(t)\in\mathcal{F}_t^t$, so it is deterministic. From standard argument (see Proposition 4.1), we can see that, $V(t,x)$ is a deterministic function.

Given any initial data (t,x), a positive constant $\delta \le T-t$ and control $u\in\mathcal{U}[t,t+\delta]$, for each $\eta\in L^2(\mathcal{F}_{t+\delta})$, we define the following backward semigroup

$$G_{t,t+\delta}^{t,x,u}[\eta] := Y(t),$$

where $(Y(s), Z(s))$ solves the following BSDE

$$Y(s) = \eta + \int_s^{t+\delta} k(r, X^{t,x;u}(r), Y(r), Z(r), u(r))dr - \int_s^{t+\delta} Z(r)dw(r),\ s\in[t,t+\delta]. \tag{2.21}$$

We have the following DPP for the above recursive optimal control problems $(S_{t,x})$ of SEEs. The proof is standard and is given in the Appendix.

Theorem 2.10 *For each $(t,x)\in[0,T]\times H$ and $0\le\delta\le T-t$, we have*

$$V(t,x) = \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,t+\delta]} G_{t,t+\delta}^{t,x;u}[V(t+\delta, X^{t,x;u}(t+\delta))] = \inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} G_{t,t+\delta}^{t,x;u}[V(t+\delta, X^{t,x;u}(t+\delta))].$$

Remark 2.11 *The DPP for nonrecursive case in the weak formulation can be found in [15]. We also note that the DPP for recursive optimal control problems in the strong formulation of SEEs under the mild solution framework is obtained in [6, 48]. Here we shall prove the strong formulation version of the variational solution framework for the completeness of the paper. Moreover, compared with that result, our state equation can have a unbounded oprator in the diffusion term and all the unbounded operators can be time-invariant, and some technical assumptions, e.g., the separablitiy of the probability space, are not needed.*

3 Main results

To carry out our purpose, we first derive an Itô's formula for the operator-valued conditionally expected BSIEs.

3.1 An Itô's formula for second adjoint equations

Given any $t \in [0,T]$ and $x_0 \in L^{2\alpha}(\mathcal{F}_t, H)$. Consider the operator-valued BSIE

$$P(s) = \mathbb{E}\Big[\tilde{L}^*(s,T)\xi\tilde{L}(s,T) + \int_s^T \tilde{L}^*(s,r)f(r,P(r))\tilde{L}(s,r)dr\Big|\mathcal{F}_s\Big], \quad s \in [t,T], \tag{3.1}$$

and two forward SEEs in the form of

$$\begin{cases} dx(s) &= \Big[A(s)x(s) + \gamma_1(s)\Big]ds + \Big[B(s)x(s) + \gamma_2(s)\Big]dw(s), \quad s \in [t,T], \\ x(t) &= x_0, \end{cases} \tag{3.2}$$

where, for some $\beta \in L^\infty_{\mathbb{F}}(t,T)$,

$$\tilde{L}(s,r) := L_{\tilde{A},\tilde{B}}(s,r) \quad \text{with} \quad \tilde{A}(s) := A(s) + \frac{\beta(s)}{2}B(s) - \frac{\beta^2(s)}{8}I_d \quad \text{and} \quad \tilde{B}(s) := B(s) + \frac{\beta(s)}{2}I_d,$$

and $\gamma_1 \in L^{1,2\alpha}_{\mathbb{F}}(t,T;H)$, $\gamma_2 \in L^{2,2\alpha}_{\mathbb{F}}(t,T;H)$.

Then we have the following Itô's formula.

Theorem 3.1 *Let Assumptions $(H2)$ and $(H4)$ be satisfied and for some $\alpha > 1$,*

$$(\xi, f(\cdot,\cdot,0), \zeta) \in L^{2\alpha}_w(\mathcal{F}_T, \mathfrak{L}(H)) \times L^{2,2\alpha}_{\mathbb{F},w}(t,T;\mathfrak{L}(H)) \times L^{4\alpha}_{\mathbb{F}}(t,T;H). \tag{3.3}$$

Then

$$\langle P(s)x(s),x(s)\rangle + \sigma(s) = \langle \xi x(T), x(T)\rangle + \int_s^T \Big[\langle f(r,P(r))x(r),x(r)\rangle + \beta(r)\mathcal{Z}(r)\Big]ds - \int_s^T \mathcal{Z}(r)dw(r), \quad s \in [t,T], \tag{3.4}$$

for a unique couple of processes $(\sigma,\mathcal{Z}) \in L^\alpha_{\mathbb{F}}(t,T) \times L^{2,\alpha}_{\mathbb{F}}(t,T)$ satisfying, for some constant C depend on Λ and λ such that

$$\sup_{s\geq t}\Big(\mathbb{E}\Big[|\sigma(s)|^\alpha\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}} \leq C\mu_1(t), \tag{3.5}$$

$$\Big\{\mathbb{E}\Big[\Big(\int_t^T |\mathcal{Z}(s)|^2 ds\Big)^{\frac{\alpha}{2}}\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{\alpha}} \leq C\mu_2(t), \tag{3.6}$$

where

$$\begin{aligned}\mu_1(t) := &\Big\{\mathbb{E}\Big[\Big(\int_t^T \|\gamma_1(s)\|_H\, ds\Big)^{4\alpha}\Big|\mathcal{F}_t\Big] + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_2(s)\|^2_H\, dt\Big)^{2\alpha}\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{2\alpha}} \\ &+ \Big\{\mathbb{E}\Big[\Big(\int_t^T \|\gamma_1(s)\|_H\, ds\Big)^{4\alpha}\Big|\mathcal{F}_t\Big] + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_2(s)\|^2_H\, ds\Big)^{2\alpha}\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}} \|x_0\|_H,\end{aligned}$$

and

$$\mu_2(t) := \|x_0\|^2_H + \Big\{\mathbb{E}\Big[\Big(\int_t^T \|\gamma_1(s)\|_H\, ds\Big)^{4\alpha}\Big|\mathcal{F}_t\Big] + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_2(s)\|^2_H\, ds\Big)^{2\alpha}\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{2\alpha}}$$

Proof. We only prove the case of $f \equiv 0$, and the general case can be treated similarly. We introduce a new SEE

$$\begin{cases} d\tilde{x}(s) &= A(s)\tilde{x}(s)ds + B(s)\tilde{x}(s)dw(s), \quad s \in [t,T], \\ x(t) &= x_0, \end{cases} \tag{3.7}$$

Then from Lemma 2.6, for any $\alpha \geq 1$,

$$\mathbb{E}\Big[\sup_{s\in[t,T]}\|x(s)-\tilde{x}(s)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big] \leq C\Big\{\mathbb{E}\Big[\Big(\int_t^T \|\gamma_1(s)\|_H\,ds\Big)^{2\alpha}\Big|\mathcal{F}_t\Big] + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_2(s)\|_H^2\,ds\Big)^{\alpha}\Big|\mathcal{F}_t\Big]\Big\}.$$

Moreover,

$$\mathbb{E}\Big[\sup_{s\in[t,T]}\|x(s)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big] \leq C\Big\{\|x_0\|_H^{2\alpha} + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_1(s)\|_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big] + \mathbb{E}\Big[\Big(\int_t^T \|\gamma_2(s)\|_H^2\,ds\Big)^{\alpha}|\mathcal{F}_t\Big]\Big\}.$$

and

$$\mathbb{E}\Big[\sup_{s\in[t,T]}\|\tilde{x}(s)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big] \leq C\,\|x_0\|_H^{2\alpha}.$$

We have from Lemma 4.3 in [27] that

$$\tilde{L}(s,r) = \frac{\lambda_1(r)}{\lambda_1(s)}L(s,r), \quad \text{for any } t\leq s\leq r\leq T,$$

with $L(s,r) := L_{A,B}(s,r)$ and $\lambda_1(s) := e^{\int_0^s -\frac{1}{4}\beta^2(r)dr+\frac{1}{2}\beta(r)dw(r)}$. We define

$$\lambda(s) := e^{\int_0^s -\frac{1}{2}\beta^2(r)dr+\beta(r)dw(r)}.$$

Noting that $\lambda = \lambda_1\cdot\lambda_1$, then

$$P(s) = \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}L^*(s,T)\xi L(s,T)\Big|\mathcal{F}_s\Big].$$

From the definition of $L(s,r)$, we have $L(s,r)\tilde{x}(s) = \tilde{x}(r)$. Thus,

$$\langle P(s)\tilde{x}(s),\tilde{x}(s)\rangle = \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}\langle \xi L(s,T)\tilde{x}(t), L(s,T)\tilde{x}(s)\rangle\Big|\mathcal{F}_s\Big] = \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}\langle \xi\tilde{x}(T),\tilde{x}(T)\rangle\Big|\mathcal{F}_s\Big]. \tag{3.8}$$

Then

$$\langle P(s)x(s),x(s)\rangle + \sigma(s) = \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}\langle \xi x(T),x(T)\rangle\Big|\mathcal{F}_s\Big],\ s\in[t,T], \tag{3.9}$$

where

$$\begin{aligned}\sigma(s) &= \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}\langle \xi x(T),x(T)\rangle\Big|\mathcal{F}_s\Big] - \mathbb{E}\Big[\frac{\lambda(T)}{\lambda(s)}\langle \xi\tilde{x}(T),\tilde{x}(T)\rangle\Big|\mathcal{F}_s\Big]\\ &\quad - \Big(\langle P(s)x(s),x(s)\rangle - \langle P(s)\tilde{x}(s),\tilde{x}(s)\rangle\Big)\\ &=: I_1(s) + I_2(s).\end{aligned}$$

Note that, with denoting by α' the Holder conjugate of α,

$$\begin{aligned}&\mathbb{E}\Big[\Big|\frac{\lambda(T)}{\lambda(s)}\Big||\langle \xi x(T),x(T)\rangle - \langle \xi\tilde{x}(T),\tilde{x}(T)\rangle|\Big|\mathcal{F}_s\Big]\\ &\leq \Big(\mathbb{E}\Big[\Big|\frac{\lambda(T)}{\lambda(s)}\Big|^{\alpha'}\Big|\mathcal{F}_s\Big]\Big)^{\frac{1}{\alpha'}}\Big(\mathbb{E}\Big[|\langle \xi x(T),x(T)\rangle - \langle \xi\tilde{x}(T),\tilde{x}(T)\rangle|^{\alpha}\Big|\mathcal{F}_s\Big]\Big)^{\frac{1}{\alpha}}\\ &\leq C_1\Big(\mathbb{E}\Big[|\langle \xi x(T),x(T)\rangle - \langle \xi\tilde{x}(T),\tilde{x}(T)\rangle|^{\alpha}|\mathcal{F}_s\Big]\Big)^{\frac{1}{\alpha}}.\end{aligned}$$

Then

$$\begin{aligned}
\Big(\mathbb{E}\Big[|I_1(s)|^\alpha\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}} &\le C_1\Big(\mathbb{E}\Big[|\langle\xi x(T),x(T)\rangle-\langle\xi\tilde{x}(T),\tilde{x}(T)\rangle|^\alpha\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}}\\
&\le C_1\Big(\mathbb{E}[\|\xi\|^{2\alpha}_{\mathfrak{L}(H)}|\mathcal{F}_t]\Big)^{\frac{1}{2\alpha}}\Big(\mathbb{E}[\|x(T)-\tilde{x}(T)\|^{4\alpha}_H|\mathcal{F}_t]\Big)^{\frac{1}{4\alpha}}\Big\{\Big(\mathbb{E}[\|x(T)\|^{4\alpha}_H|\mathcal{F}_t]\Big)^{\frac{1}{4\alpha}}+\Big(\mathbb{E}[\|\tilde{x}(T)\|^{4\alpha}_H\Big|\mathcal{F}_t]\Big)^{\frac{1}{4\alpha}}\Big\}\\
&\le C_1\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}}\\
&\cdot\Big\{\Big\{\|x_0\|^{4\alpha}_H+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}}+\|x_0\|_H\Big\}\\
&\le C_1\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}}\\
&\cdot\Big\{\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}}+\|x_0\|_H\Big\}\\
&\le C_1\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{2\alpha}}\\
&+C_1\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{4\alpha}}\|x_0\|_H\\
&=:C\mu_1(t).
\end{aligned}$$

Similarly but yet more simply, we also have

$$\Big(\mathbb{E}\Big[|I_2(s)|^\alpha\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}}\le C\mu_1(t).$$

Thus,

$$\Big(\mathbb{E}\Big[|\sigma(s)|^\alpha\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}}\le C\mu_1(t).$$

Note that equation (3.9) is the explicit formula of the linear BSDE (3.4) with solution $(\langle P(s)x(s),x(s)\rangle+\sigma(s),\mathcal{Z}(s))\in L^\alpha_{\mathbb{F}}(t,T)\times L^{2,\alpha}_{\mathbb{F}}(t,T)$. The uniqueness of $(\sigma,\mathcal{Z})$ in the equation (3.4) follows from the basic results of BSDEs. We also have

$$\begin{aligned}
\Big(\mathbb{E}\Big[|\langle\xi x(T),x(T)\rangle|^\alpha|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}} &\le C_1\Big(\mathbb{E}[\|\xi\|^{2\alpha}_{\mathfrak{L}(H)}|\mathcal{F}_t]\Big)^{\frac{1}{2\alpha}}\Big(\mathbb{E}[\|x(T)\|^{4\alpha}_H|\mathcal{F}_t]\Big)^{\frac{1}{2\alpha}}\\
&\le C\Big\{\|x_0\|^2_H+\Big\{\mathbb{E}\Big[\Big(\int_t^T\|\gamma_1(s)\|_H\,ds\Big)^{4\alpha}|\mathcal{F}_t\Big]+\mathbb{E}\Big[\Big(\int_t^T\|\gamma_2(s)\|^2_H\,ds\Big)^{2\alpha}|\mathcal{F}_t\Big]\Big\}^{\frac{1}{2\alpha}}\Big\}\\
&=:C\mu_2(t).
\end{aligned}$$

Thus, from the basic estimates of BSDEs, we obtain that

$$\Big\{\mathbb{E}\Big[\Big(\int_t^T|\mathcal{Z}(s)|^2ds\Big)^{\frac{\alpha}{2}}\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{\alpha}}\le C\mu_2(t).$$

□

We also need the following corollary of DPP.

Lemma 3.2 *If $(\bar{X}^{t,x;\bar{u}}(\cdot),\bar{Y}^{t,x;\bar{u}}(\cdot),\bar{Z}^{t,x;\bar{u}}(\cdot),\bar{u}(\cdot))$ are optimal for Problem (S_{tx}), then for any $\delta\in[0,T-t]$,*

$$V(t+\delta,\bar{X}^{t,x;\bar{u}}(t+\delta))=\bar{Y}^{t,x;\bar{u}}(t+\delta),\ P\text{-a.s.}$$

Proof. According to (4.10), we have

$$\bar{Y}^{t,x;\bar{u}}(s) = \bar{Y}^{t,x;\bar{u}}(t+\delta) + \int_s^{t+\delta} k(r, X^{t,x;\bar{u}}(r), Y^{t,x;\bar{u}}(r), Z^{t,x;\bar{u}}(r), \bar{u}(r))dr - \int_s^{t+\delta} Z^{t,x;\bar{u}}(r)dw(r),\ s \in [t, t+\delta].$$

We introduce a BSDE

$$y^{t,x;\bar{u}}(s) = V(t+\delta, \bar{X}^{t,x;\bar{u}}(t+\delta)) + \int_s^{t+\delta} k(r, X^{t,x;\bar{u}}(r), y^{t,x;\bar{u}}(r), z^{t,x;\bar{u}}(r), \bar{u}(r))dr - \int_s^{t+\delta} z^{t,x;\bar{u}}(r)dw(r),\ s \in [t, t+\delta].$$

We also know that, from Proposition 4.4,

$$\bar{Y}^{t,x;\bar{u}}(t+\delta) = \bar{Y}^{t+\delta, \bar{X}^{t,x;\bar{u}}(t+\delta);\bar{u}}(t+\delta) \geq V(t+\delta, \bar{X}^{t,x;\bar{u}}(t+\delta)).$$

Then from the DPP and comparison theorem of classical BSDEs,

$$V(t,x) \leq G_{t,t+\delta}^{t,x;\bar{u}}\big[V(t+\delta, X^{t,x;\bar{u}}(t+\delta))\big] = y^{t,x;\bar{u}}(t) \leq \bar{Y}^{t,x;\bar{u}}(t).$$

But $V(t,x) = \bar{Y}^{t,x;\bar{u}}(t)$. So, all the inequalities above are in fact equalities. Then from the strict comparison theorem of classical BSDEs, it must holds

$$\bar{Y}^{t,x;\bar{u}}(t+\delta) = V(t+\delta, \bar{X}^{t,x;\bar{u}}(t+\delta)).$$

This completes the proof. □

3.2 Differential in spatial variable

Before stating the main result of this subsection, let us recall the notion of super and subdifferentials.

For $v \in C([0,T] \times H)$ and $(t,x) \in [0,T) \times H$, the second-order parabolic partial superdifferential of v with respect to x is defined as follows:

$$D_x^{2,+}v(t,x) = \Big\{(p,P) \in H \times \mathcal{S}(H)\Big| \\ v(t,y) \leq v(t,x) + \langle p, y-x\rangle_H + \frac{1}{2}\langle P(y-x), y-x\rangle_H + o(|y-x|^2),\ \text{as } y \to x\Big\},$$

where $\mathcal{S}(H)$ is the space of all symmetric (self-adjoint) bounded linear operators on H. Similarly, the second-order parabolic partial subdifferential of v with respect to x is defined as follows:

$$D_{t,x}^{2,-}v(t,x) = \Big\{(p,P) \in H \times \mathcal{S}(H)\Big| \\ v(t,y) \geq v(t,x) + \langle p, y-x\rangle_H + \frac{1}{2}\langle P(y-x), y-x\rangle_H + o(|y-x|^2),\ \text{as } y \to x\Big\}.$$

We have the following result on the differential in the spatial variable.

Theorem 3.3 *Assume (H1). Suppose $(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{Z}(\cdot), \bar{u}(\cdot))$ are the optimal 4-tuple of Problem (S_x) and $p(\cdot), q(\cdot), P(\cdot)$ are the solutions of corresponding adjoint equations. Let $V \in C([0,T] \times H)$ be defined as in (2.20). Then*

$$\{p(t)\} \times [P(t), \infty) \subset D_x^{2,+}V(t, \bar{X}(t)), \quad t \in [0,T],\ P\text{-a.s.}$$

and

$$D_x^{2,-}V(t, \bar{X}(t)) \subset \{p(t)\} \times (-\infty, P(t)], \quad t \in [0,T],\ P\text{-a.s.}$$

Proof. *Step 1: Variational equations.* Fix any $t \in [0,T]$ and any $x^1 \in H$, let X^{x^1} be the solution of the following SEE on $[t,T]$:

$$\begin{cases} dX^{x^1}(s) &= [A(s)X^{x^1}(s) + a(s, X^{x^1}(s), \bar{u}(s))]ds + [B(s)X^{x^1}(s) + b(s, X^{x^1}(s), \bar{u}(s))]dw(s), \\ X^{x^1}(t) &= x^1. \end{cases} \tag{3.10}$$

We denote $\hat{x}(s) := X^{x^1}(s) - \bar{X}(s)$, $s \in [t,T]$ (In particular, $\hat{x}(t) = X^1(t) - \bar{X}(t) = x^1 - \bar{X}(t)$).Then

$$\begin{aligned} \hat{x}(s) &= \hat{x}(t) + \int_t^s \Big[A(r)\hat{x}(r) + a(r, X^{x^1}(r), \bar{u}(r)) - a(r, \bar{X}(r), \bar{u}(r))\Big]dr \\ &\quad + \int_t^s \Big[B(r)\hat{x}(r) + b(r, X^{x^1}(r), \bar{u}(r)) - b(r, \bar{X}(r), \bar{u}(r))\Big]dw(r) \\ &= \hat{x}(t) + \int_t^s \Big[A(r)\hat{x}(r) + a(r, \hat{x}(r) + \bar{X}(r), \bar{u}(r)) - a(r, \bar{X}(r), \bar{u}(r))\Big]dr \\ &\quad + \int_t^s \Big[B(r)\hat{x}(r) + b(r, \hat{x}(r) + \bar{X}(r), \bar{u}(r)) - b(r, \bar{X}(r), \bar{u}(r))\Big]dw(r). \end{aligned}$$

From Lemma 2.6, we first have

$$\mathbb{E}\left[\sup_{t\le s\le T} \|\hat{x}(s)\|_H^{2\alpha}\, dr|\mathcal{F}_t\right] \le C\,\|\hat{x}(t)\|_H^{2\alpha},\ P\text{-a.s.} \tag{3.11}$$

With $\bar{A}, \bar{B}$ being defined as in subsection 2.2, we can write the equation of $\hat{x}(s)$ as:

$$\hat{x}(s) = \hat{x}(t) + \int_t^s [\bar{A}(r)\hat{x}(r) + \varepsilon_1(r)]dr + \int_t^s [\bar{B}(r)\hat{x}(r) + \varepsilon_2(r)]dw(r) \tag{3.12}$$

and

$$\begin{aligned} \hat{x}(s) = \hat{x}(t) + &\int_t^s [\bar{A}(r)\hat{x}(r) + \frac{1}{2}\bar{a}_{xx}(r)(\hat{x}(r), \hat{x}(r)) + \varepsilon_3(r)]\, dr \\ &+ \int_t^s [\bar{B}(r)\hat{x}(r) + \frac{1}{2}\bar{b}_{xx}(r)(\hat{x}(r), \hat{x}(r)) + \varepsilon_4(r)]\, dw(r), \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} \varepsilon_1(r) &:= \int_0^1 \left\langle a_x\left(r, \bar{X}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_x\left(r\right), \hat{x}(r)\right\rangle d\mu, \\ \varepsilon_2(r) &:= \int_0^1 \left\langle b_x\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{b}_x\left(r\right), \hat{x}(r)\right\rangle d\mu, \\ \varepsilon_3(r) &:= \int_0^1 (1-\mu)\left[a_{xx}\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_{xx}\left(r\right)\right](\hat{x}(r), \hat{x}(r))\, d\mu, \\ \varepsilon_4(r) &:= \int_0^1 (1-\mu)\left[b_{xx}\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{b}_{xx}\left(r\right)\right](\hat{x}(r), \hat{x}(r))\, d\mu. \end{aligned}$$

Step 2: Estimates of remainder terms of SEEs. We have following estimates: for any $\alpha \geq 2$,

$$\begin{aligned}
&\mathbb{E}\Big[\int_t^T \|\varepsilon_1(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] = o\left(\|\hat{x}(t)\|_H^\alpha\right), \ P\text{-a.s.} \\
&\mathbb{E}\Big[\int_t^T \|\varepsilon_2(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] = o\left(\|\hat{x}(t)\|_H^\alpha\right), \ P\text{-a.s.} \\
&\mathbb{E}\Big[\int_t^T \|\varepsilon_3(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] = o(\|\hat{x}(t)\|_H^{2\alpha}), \ P\text{-a.s.} \\
&\mathbb{E}\Big[\int_t^T \|\varepsilon_4(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] = o(\|\hat{x}(t)\|_H^{2\alpha}), \ P\text{-a.s.}
\end{aligned} \tag{3.14}$$

In the above and the whole paper, we use $o(\rho)$ to denote a deterministic infinitesimal function as $\rho \to 0$ (which may be different according to the context), and similarly, use $O(\rho)$ to represent a deterministic function that has the same order as ρ. We only present the proofs for the first and the third ones, and the other two can be derived similarly. Applying (3.11),

$$\begin{aligned}
\mathbb{E}\Big[\int_t^T \|\varepsilon_1(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] &= \int_t^T \mathbb{E}\Big[\|\varepsilon_1(r)\|_H^\alpha \, |\mathcal{F}_t\Big] dr \\
&\leq \int_t^T \mathbb{E}\Big[\int_0^1 \|a_x\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_x\left(r\right)\|_H^\alpha \, d\mu \, \|\hat{x}(r)\|_H^\alpha \, |\mathcal{F}_t\Big] dr \\
&\leq \int_t^T \mathbb{E}[\|\hat{x}(r)\|_H^{2\alpha} \, |\mathcal{F}_t] dr \\
&\leq C \, \|\hat{x}(t)\|_H^{2\alpha} \\
&= o(\|\hat{x}(t)\|_H^\alpha), \ P\text{-a.s.}
\end{aligned}$$

and (maybe we need to impose the γ-Hölder continuity assumption on a_{xx} here and in the proof of Theorem 3.4)

$$\begin{aligned}
&\mathbb{E}\Big[\int_t^T \|\varepsilon_3(r)\|_H^\alpha \, dr|\mathcal{F}_t\Big] \\
&= \int_t^T \mathbb{E}\left[\|\varepsilon_3(r)\|_H^\alpha \, |\mathcal{F}_t\right] dr \\
&\leq \int_t^T \mathbb{E}\left[\int_0^1 |(1-\mu)\left[a_{xx}\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_{xx}\left(r\right)\right]\left(\hat{x}(r), \hat{x}(r)\right)|^\alpha d\mu|\mathcal{F}_t\right] dr \\
&\leq \Big(\int_t^T \mathbb{E}\Big[\int_0^1 (1-\mu)^{2\alpha} \|a_{xx}\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_{xx}\left(r\right)\|_{\mathfrak{L}(H)}^{2\alpha} d\mu|\mathcal{F}_t\Big] dr\Big)^{\frac{1}{2}} \Big(\int_t^T \mathbb{E}\Big[\int_0^1 \|\hat{x}(r)\|_H^{4\alpha} \, d\mu|\mathcal{F}_t\Big] dr\Big)^{\frac{1}{2}} \\
&\leq \Big(\int_t^T \mathbb{E}\Big[\int_0^1 \|a_{xx}\left(r, \bar{x}(r) + \mu\hat{x}(r), \bar{u}(r)\right) - \bar{a}_{xx}\left(r\right)\|_{\mathfrak{L}(H)}^{2\alpha} d\mu|\mathcal{F}_t\Big] dr\Big)^{\frac{1}{2}} \|\hat{x}(t)\|_H^{2\alpha} \\
&\leq \Big(\int_t^T \mathbb{E}\Big[\int_0^1 \|\hat{x}(r)\|_{\mathfrak{L}(H)}^{2\gamma\alpha} d\mu|\mathcal{F}_t\Big] dr\Big)^{\frac{1}{2}} \|\hat{x}(t)\|_H^{2\alpha} \\
&= o(\|\hat{x}(t)\|_H^{2\alpha}), \ P\text{-a.s.}
\end{aligned}$$

Step 3: Duality relationship. Applying Itô's formula to $\langle p(r), \hat{x}(r)\rangle$, from (3.13) we get

$$\langle p(s), \hat{x}(s)\rangle = \langle h_x(\bar{X}(T)), \hat{x}(T)\rangle + \int_s^T J_1(r)dr - \int_s^T J_2(r)dw(r), \ s \in [t, T]. \tag{3.15}$$

where

$$\begin{aligned}J_1(s) :=&\langle k_x(s)+k_y(s)p(s)+k_z(s)q(s),\hat{x}(s)\rangle+k_z(s)\langle p(s),\bar{B}(s)\hat{x}(s)\rangle-\langle p(s),\varepsilon_3(s)\rangle-\langle q(s),\varepsilon_4(s)\rangle\\&-\frac{1}{2}[\langle p(s),(\bar{a}_{xx}(s)(\hat{x}(s),\hat{x}(s))\rangle+\langle q(s),\bar{b}_{xx}(s)(\hat{x}(s),\hat{x}(s))\rangle],\\J_2(s) :=&\langle p(s),\bar{B}(s)(\hat{x}(s))\rangle+\langle q(s),\hat{x}(s)\rangle+\langle p(s),\varepsilon_4(s)\rangle\\&+\frac{1}{2}\langle p(s),\bar{b}_{xx}(s)(\hat{x}(s),\hat{x}(s))\rangle.\end{aligned}$$

Next, taking $\gamma_1=\varepsilon_1,\gamma_2=\varepsilon_2,x_0=\hat{x}(t)$ in Theorem 3.1, from Step 2 we have

$$\mu_1(t)=o(\|\hat{x}(t)\|_H^2)\text{ and }\mu_2(t)=O(\|\hat{x}(t)\|_H^2).$$

Thus, applying Theorem 3.1 to P and $\hat{x}$ in (3.12), we obtain

$$\begin{aligned}\langle P(s)\hat{x}(s),\hat{x}(s)\rangle+\sigma(s)=&\langle h_{xx}(\bar{X}(T))\hat{x}(T),\hat{x}(T)\rangle+\int_s^T[k_y(s)\langle P(s)\hat{x}(s),\hat{x}(s)\rangle\\&+k_z(s)\mathcal{Z}(s)+\langle G(s)\hat{x}(s),\hat{x}(s)\rangle]ds-\int_s^T\mathcal{Z}(s)dw(s),\end{aligned}\tag{3.16}$$

for some processes $(\sigma,\mathcal{Z})\in L^{\alpha}_{\mathbb{F}}(t,T)\times L^{2,\alpha}_{\mathbb{F}}(t,T)$ satisfying

$$\sup_{s\in[t,T]}\Big(\mathbb{E}\Big[|\sigma(s)|^{\alpha}\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{\alpha}}=o(\|\hat{x}(t)\|_H^2)\ \text{ and }\ \mathbb{E}\Big[\Big(\int_t^T|\mathcal{Z}(t)|^2dt\Big)^{\frac{\alpha}{2}}|\mathcal{F}_t\Big]=O(\|\hat{x}(t)\|_H^{2\alpha}),\quad\text{for any }\alpha\geq 2.\tag{3.17}$$

Therefore,

$$\begin{aligned}\langle p(t),\hat{x}(t)\rangle+\frac{1}{2}\langle P(t)\hat{x}(t),\hat{x}(t)\rangle+\frac{1}{2}\sigma(t)=&\langle h_x(\bar{x}(T)),\hat{x}(T)\rangle\\&+\frac{1}{2}\langle h_{xx}(\bar{x}(T))\hat{x}(T),\hat{x}(T)\rangle+\int_t^T I_1(s)ds-\int_t^T I_2(s)dw(s),\end{aligned}$$

where

$$\begin{aligned}I_1(s) :=&\langle k_x(s)+k_y(s)p(s)+k_z(s)q(s),\hat{x}(s)\rangle+k_z(s)\langle p(s),\bar{B}(s)\hat{x}(s)\rangle+\frac{1}{2}\langle\{k_y(s)P(s)\\&+D^2k(s)([I_d,p(s),\bar{B}^*(s)p(s)+q(s)],[I_d,p(s),\bar{B}^*(s)p(s)+q(s)])+k_z(s)\langle p(s),\bar{b}_{xx}(s)\rangle\}\hat{x}(s),\hat{x}(s)\rangle\\&+\frac{1}{2}k_z(s)\mathcal{Z}(s)-\langle p(s),\varepsilon_3(s)\rangle-\langle q(s),\varepsilon_4(s)\rangle,\\I_2(s) :=&\langle p(s),\bar{B}(s)\hat{x}(s)\rangle+\langle q(s),\hat{x}(s)\rangle+\langle p(s),\varepsilon_4(s)\rangle+\frac{1}{2}\langle p(s),\bar{b}_{xx}(s)(\hat{x}(s),\hat{x}(s))\rangle+\frac{1}{2}\mathcal{Z}(s).\end{aligned}$$

Step 4: Variational equation for BSDE. We denote, on $[t,T]$,

$$Y^{x^1}(s)=h(X^{x^1}(T))+\int_s^T k(r,X^{x^1}(r),Y^{x^1}(r),Z^{x^1}(r),\bar{u}(r))dr-\int_s^T Z^{x^1}(r)dw(r).$$

Then we have

$$\begin{aligned}\hat{y}(s)-\frac{1}{2}\sigma(s)=&h(X^{x^1}(T))-h(\bar{X}(T))-\langle h_x(\bar{X}(T)),\hat{x}(T)\rangle-\frac{1}{2}\langle h_{xx}(\bar{X}(T))\hat{x}(T),\hat{x}(T)\rangle\\&+\int_s^T\{k(r,X^{x^1}(r),Y^{x^1}(r),Z^{x^1}(r),\bar{u}(r))-k(r,\bar{X}(r),\bar{Y}(r),\bar{Z}(r),\bar{u}(r))-I_1(r)\}dr-\int_s^T\hat{z}(r)dw(r),\end{aligned}\tag{3.18}$$

where

$$\hat{y}(s) := Y^{x^1}(s) - \bar{Y}(s) - \langle p(s), \hat{x}(s)\rangle - \frac{1}{2}\langle P(s)\hat{x}(s), \hat{x}(s)\rangle,$$
$$\hat{z}(s) := Z^{x^1}(s) - \bar{Z}(s) - I_2(s).$$

We denote

$$I_3(s) := \langle p(s), \hat{x}(s)\rangle + \frac{1}{2}\langle P(s)\hat{x}(s), \hat{x}(s)\rangle$$

From the Taylor's expansion,

$$\begin{aligned}\hat{y}(s) - \frac{1}{2}\sigma(s) = J_4 + \int_s^T \Big\{\tilde{k}_y(r)(\hat{y}(r) - \frac{1}{2}\sigma(r)) + \tilde{k}_z(r)\hat{z}(r) + \frac{1}{2}J_5(r) + \frac{1}{2}\tilde{k}_y(r)\sigma(r)\\ + k_z(r)\langle p(r), \varepsilon_4(r)\rangle + \langle p(r), \varepsilon_3(r)\rangle + \langle q(r), \varepsilon_4(r)\rangle\Big\}dr - \int_s^T \hat{z}(r)dw(r),\end{aligned} \tag{3.19}$$

where

$$\begin{aligned}
&\tilde{k}_y(s) := \int_0^1 k_y(s, X^{x^1}(s), \bar{Y}(s) + I_3(s) + \mu\hat{y}(s), \bar{Z}(s) + I_2(s) + \mu\hat{z}(s), \bar{u}(s))d\mu,\\
&\tilde{k}_z(s) := \int_0^1 k_z(s, X^{x^1}(s), \bar{Y}(s) + I_3(s) + \mu\hat{y}(s), \bar{Z}(s) + I_2(s) + \mu\hat{z}(s), \bar{u}(s))d\mu,\\
&\tilde{D}^2k(s) := 2\int_0^1\int_0^1 \mu D^2k(s, \bar{x}(s) + \mu\nu\hat{x}(s), \bar{Y}(s) + \mu\nu I_3(s), \bar{Z}(s) + \mu\nu I_2(s), \bar{u}(s))d\mu d\nu,\\
&J_3(s) := k_z(s)\langle p(s), \varepsilon_4(s)\rangle + \langle p(s), \varepsilon_3(s)\rangle + \langle q(s), \varepsilon_4(s)\rangle,\\
&J_4 := h(X^{x^1}(T)) - h(\bar{X}(T)) - \langle h_x(\bar{X}(T)), \hat{x}(T)\rangle - \frac{1}{2}\langle h_{xx}(\bar{X}(T))\hat{x}(T), \hat{x}(T)\rangle,\\
&J_5(s) := \tilde{D}^2k(s)([\hat{x}(s), I_3(s), I_2(s)], [\hat{x}(s), I_3(s), I_2(s)])\\
&\qquad - \langle D^2k(s)([I_d, p(s), \bar{B}^*(s)p(s) + q(s)], [I_d, p(s), \bar{B}^*(s)p(s) + q(s)])\hat{x}(s), \hat{x}(s)\rangle.
\end{aligned}$$

First, we have

$$\begin{aligned}
&\mathbb{E}[h(X^{x^1}(T)) - h(\bar{X}(T)) - \langle h_x(\bar{X}(T)), \hat{x}(T)\rangle - \frac{1}{2}\langle h_{xx}(\bar{X}(T))\hat{x}(T), \hat{x}(T)\rangle|\mathcal{F}_t]\\
&= \mathbb{E}[\langle h_x(\bar{X}(T)), \hat{x}(T)\rangle + \frac{1}{2}\langle \tilde{h}_{xx}(T)\hat{x}(T), \hat{x}(T)\rangle - \langle h_x(\bar{X}(T)), \hat{x}(T)\rangle - \frac{1}{2}\langle h_{xx}(\bar{X}(T))\hat{x}(T), \hat{x}(T)\rangle|\mathcal{F}_t]\\
&= \frac{1}{2}\mathbb{E}[\langle(\tilde{h}_{xx}(T) - h_{xx}(\bar{X}(T)))\hat{x}(T), \hat{x}(T)\rangle|\mathcal{F}_t]\\
&\le \frac{1}{2}\Big(\mathbb{E}[|\tilde{h}_{xx}(T) - h_{xx}(\bar{X}(T))|^2_{\mathcal{L}_2(H\times H)}|\mathcal{F}_t]\Big)^{\frac{1}{2}}\Big(\mathbb{E}[|\hat{x}(T)|^4_H|\mathcal{F}_t]\Big)^{\frac{1}{2}}\\
&= o(\|\hat{x}(t)\|^2_H),
\end{aligned}$$

where

$$\tilde{h}_{xx}(T) = 2\int_0^1\int_0^1 \mu h_{xx}\big(\bar{X}(T) + \mu\nu\hat{x}(T)\big)d\mu d\nu.$$

Moreover, it is direct to check that $\mathbb{E}[\int_t^T |J_3(s)|ds|\mathcal{F}_t] = o(\|\hat{x}(t)\|^2_H)$. We write $J_5(s) = J_6(s) + J_7(s)$, where

$$\begin{aligned}
J_6(s) :=& \langle\tilde{D}^2k(s)([I_d, p(s), \bar{B}^*(s)p(s) + q(s)], [I_d, p(s), \bar{B}^*(s)p(s) + q(s)])\hat{x}(s), \hat{x}(s)\rangle\\
&- \langle D^2k(s)([I_d, p(s), \bar{B}^*(s)p(s) + q(s)], [I_d, p(s), \bar{B}^*(s)p(s) + q(s)])\hat{x}(s), \hat{x}(s)\rangle,\\
J_7(s) :=& \tilde{D}^2k(s)([\hat{x}(s), I_3(s), I_2(s)], [\hat{x}(s), I_3(s), I_2(s)])\\
&- \langle\tilde{D}^2k(s)([I_d, p(s), \bar{B}^*(s)p(s) + q(s)], [I_d, p(s), \bar{B}^*(s)p(s) + q(s)])\hat{x}(s), \hat{x}(s)\rangle.
\end{aligned}$$

We only estimate J_6 and the treatment for J_7 is similar. First, recall that (H4) implies (see the proof of Proposition 3.5 in [27])

$$|\langle v, B(s,\omega)w\rangle| = |\langle B^*(s,\omega)v, w\rangle| \le C(K)\|v\|_V\|w\|_H, \quad \text{for } v,w \in V \text{ and } (s,\omega)\in[0,T]\times\Omega. \tag{3.20}$$

By denoting $\|\tilde{D}^2k(s) - D^2k(s)\| := \|\tilde{D}^2k(s) - D^2k(s)\|_{\mathfrak{L}_2((H\times\mathbb{R}\times\mathbb{R})\times(H\times\mathbb{R}\times\mathbb{R});\mathbb{R})}$, we then have

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_t^T |J_6(s)|ds\Big)^{2\alpha}|\mathcal{F}_t\Big] &\le C\mathbb{E}\Big[\Big(\int_t^T \|\tilde{D}^2k(s) - D^2k(s)\|\big((1+\|p(s)\|_H^2)\|\hat{x}(s)\|_H^2 \\
&\quad + \|p(s)\|_V^2\|\hat{x}(s)\|_H^2 + \|q(s)\|_H^2\|\hat{x}(s)\|_H^2\big)ds\Big)^{2\alpha}\Big|\mathcal{F}_t\Big] \\
&\le C\Big(\mathbb{E}\Big[\int_t^T \|\tilde{D}^2k(s) - D^2k(s)\|^{4\alpha}(1+\|p(s)\|_H^{8\alpha})ds\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{2}}\Big|\mathcal{F}_t\Big(\mathbb{E}[\int_t^T \|\hat{x}(s)\|_H^{8\alpha}ds\Big|\mathcal{F}_t]\Big)^{\frac{1}{2}} \\
&\quad + C\Big(\mathbb{E}\Big[\Big(\int_t^T \|\tilde{D}^2k(s) - D^2k(s)\|\|p(s)\|_V^2ds\Big)^{4\alpha}|\mathcal{F}_t\Big]\Big)^{\frac{1}{2}}(\mathbb{E}[\sup_{s\in[t,T]}\|\hat{x}(s)\|_H^{8\alpha}|\mathcal{F}_t])^{\frac{1}{2}} \\
&\quad + C\Big(\mathbb{E}\Big[\Big(\int_t^T \|\tilde{D}^2k(s) - D^2k(s)\|\|q(s)\|_H^2ds\Big)^{4\alpha}\Big|\mathcal{F}_t\Big]\Big)^{\frac{1}{2}}(\mathbb{E}[\sup_{s\in[t,T]}\|\hat{x}(s)\|_H^{8\alpha}|\mathcal{F}_t])^{\frac{1}{2}} \\
&= o(\|\hat{x}(t)\|_H^{4\alpha}).
\end{aligned}$$

So, $\mathbb{E}[(\int_t^T |J_5(s)|ds)^{2\alpha}|\mathcal{F}_t] = o(\|\hat{x}(t)\|_H^{4\alpha})$. Then from the a priori estimae for classical BSDEs,

$$\sup_{s\in[t,T]}\mathbb{E}[|\hat{y}(s) - \frac{1}{2}\sigma(s)|^{2\alpha}|\mathcal{F}_t] + \mathbb{E}\Big[\Big(\int_t^T |\hat{z}(s)|^2ds\Big)^{\alpha}|\mathcal{F}_t\Big] = o(\|\hat{x}(t)\|_H^{4\alpha}).$$

Taking into account of (3.17) again,

$$\sup_{s\in[t,T]}\mathbb{E}[|\hat{y}(s)|^{2\alpha}|\mathcal{F}_t] + \mathbb{E}\Big[\Big(\int_t^T |\hat{z}(t)|^2dt\Big)^{\alpha}|\mathcal{F}_t\Big] = o(\|\hat{x}(t)\|_H^{4\alpha}).$$

In particular,

$$Y^{x^1}(t) - \bar{Y}(t) = \langle p(t), \hat{x}(t)\rangle + \frac{1}{2}\langle P(t)\hat{x}(t), \hat{x}(t)\rangle + o(\|\hat{x}(t)\|_H^2),\ P\text{-a.s.} \tag{3.21}$$

Step 5: Completion of the proof. Let M be a countable dense subset of H. We can find a subset $\Omega_0 \subset \Omega$ such that $P(\Omega_0) = 1$ and for each $\omega_0 \in \Omega_0$,

$$\begin{gathered}
V(t, \bar{X}(t,\omega_0)) = \bar{Y}(t,\omega_0),\ Y^{x^1}(t,\omega_0) \le V(t,x^1),\ (3.21) \text{ holds for all } x^1 \in M, \\
\text{and } p(s,\omega_0) \in H,\ P(s,\omega_0) \in \mathfrak{L}(H),\ \forall s \in [0,T].
\end{gathered}$$

Fix any $\omega_0 \in \Omega_0$. Then for any $x^1 \in M$,

$$Y^{x^1}(t,\omega_0) - \bar{Y}(t,\omega_0) = \langle p(t,\omega_0), \hat{x}(t,\omega_0)\rangle + \frac{1}{2}\langle P(t,\omega_0)\hat{x}(t,\omega_0), \hat{x}(t,\omega_0)\rangle + o(\|\hat{x}(t,\omega_0)\|_H^2).$$

Thus

$$V(t,x^1) - V(t,\bar{X}(t,\omega_0)) \le \langle p(t,\omega_0), \hat{x}(t,\omega_0)\rangle + \frac{1}{2}\langle P(t,\omega_0)\hat{x}(t,\omega_0), \hat{x}(t,\omega_0)\rangle + o(\|\hat{x}(t,\omega_0)\|_H^2),\ \text{for all } x^1 \in M.$$

Note that the term $o(|\hat{x}(t,\omega_0)|^2)$ in the above inequality depends only on the size $|\hat{x}(t,\omega_0)|^2$ and is is independent of x^1. Therefore, from the continuity of $V(t,\cdot)$, we obtain that

$$V(t,x^1) - V(t,\bar{X}(t,\omega_0)) \le \langle p(t,\omega_0), \hat{x}(t,\omega_0)\rangle + \frac{1}{2}\langle P(t,\omega_0)\hat{x}(t,\omega_0), \hat{x}(t,\omega_0)\rangle + o(\|\hat{x}(t,\omega_0)\|_H^2),\ \text{for all } x^1 \in H. \tag{3.22}$$

This proves that, from the definition of upper-differentials,

$$\{p(t,\omega_0)\}\times[P(t,,\omega_0),\infty)\subset D_x^{2,+}V(t,\bar X(t,\omega_0)).$$

Now we prove the second one. Fix any ω_0 such that (3.22) hold. For any $(\hat p,\hat P)\in D_x^{2,-}V(t,\bar X(t,\omega_0))$, we have from the definition of subdifferentials and (3.22) that

$$\begin{aligned}0&\le\liminf_{x^1\to\bar X(t,\omega_0)}\left\{\frac{V(t,x^1)-V(t,\bar X(t,\omega_0))-\langle\hat p,x^1-\bar X(t,\omega_0)\rangle-\frac12\langle\hat P(x^1-\bar X(t,\omega_0)),x^1-\bar X(t,\omega_0)\rangle}{\left\|x^1-\bar X(t,\omega_0)\right\|_H^2}\right\}\\&\le\liminf_{x^1\to\bar X(t,\omega_0)}\left\{\frac{\langle p(t,\omega_0)-\hat p,x^1-\bar X(t,\omega_0)\rangle+\frac12\langle(P(t,\omega_0)-\hat P)(x^1-\bar X(t,\omega_0)),x^1-\bar X(t,\omega_0)\rangle}{\left\|x^1-\bar X(t,\omega_0)\right\|_H^2}\right\}.\end{aligned}$$

Then it is necessary that

$$\hat p=p(t,\omega_0),\hat P\le P(t,\omega_0),$$

which implies the desired result. □

3.3 Differential in time variable

For $v\in C([0,T]\times H)$ and $(t,x)\in[0,T)\times H$, the second-order parabolic partial superdifferential of v with respect to t is defined as:

$$D_{t+}^{1,+}v(t,x)=\Big\{r\in H\Big|v(s,x)\le v(t,x)+r(s-t)+o(|s-t|),\ \text{as } s\downarrow t\Big\}.$$

The second-order parabolic partial subdifferential of v with respect to t is defined as:

$$D_{t+}^{1,-}v(t,x)=\Big\{r\in H\Big|v(s,x)\ge v(t,x)+r(s-t)+o(|s-t|),\ \text{as } s\downarrow t\Big\}.$$

Theorem 3.4 *Suppose $(\bar X(\cdot),\bar Y(\cdot),\bar Z(\cdot),\bar u(\cdot))$ are the optimal 4-tuple of Problem (S_x) and $p(\cdot),q(\cdot),P(\cdot)$ are the solutions of corresponding adjoint equations. Then*

$$\begin{cases}[-\langle p(t),A(t)\bar X(t)\rangle-\langle q(t),B(t)\bar X(t)\rangle+\mathcal{H}_1(t,\bar X(t),\bar Y(t),\bar Z(t)),\infty)\subseteq D_{t+}^{1,+}V(t,\bar X(t)),\\ D_{t+}^{1,-}V(t,\bar X(t))\subseteq(-\infty,-\langle p(t),A(t)\bar X(t)\rangle-\langle q(t),B(t)\bar X(t)\rangle+\mathcal{H}_1(t,\bar X(t),\bar Y(t),\bar Z(t))],\ \textit{a.e., P-a.s.},\end{cases}$$

where

$$\begin{aligned}&\mathcal{H}_1(t,\bar X(t),\bar Y(t),\bar Z(t))\\&\quad:=-\mathcal{H}(s,\bar X(s),\bar Y(s),\bar Z(s),\bar u(s),p(t),q(t))\\&\qquad+\langle P(t)[B(t)\bar X(t)+b(t,\bar X(t),\bar u(t))],B(t)\bar X(t)+b(t,\bar X(t),\bar u(t))\rangle.\end{aligned}$$

Proof. *Step 1.* Take any $\tau\in(t,T)$. Denote by X^τ the solution of the following SEE in $[\tau,T]$:

$$\begin{cases}dX^\tau(s)&=[A(s)X^\tau(s)+a(s,X^\tau(s),\bar u(s))]dr+[B(s)X^\tau(s)+b(s,X^\tau(s),\bar u(s))]dw(s),\\X^\tau(\tau)&=\bar X(t).\end{cases}\tag{3.23}$$

We define

$$\hat\xi_\tau(s):=X^\tau(s)-\bar X(s),\ s\in[\tau,T].$$

Then on $[\tau,T]$,

$$\begin{aligned}\hat\xi_\tau(s)=\hat\xi_\tau(\tau)&+\int_\tau^s[A(r)\hat\xi_\tau(r)+a(r,X^{x^1}(r),\bar u(r))-a(r,\bar X(r),\bar u(r))]dr\\&+\int_\tau^s[B(r)\hat\xi_\tau(r)+b(r,X^{x^1}(r),\bar u(r))-b(r,\bar X(r),\bar u(r))]dw(r).\end{aligned}$$

From Lemma 2.6,

$$\mathbb{E}\Big[\sup_{\tau\le s\le T}\|\hat{\xi}_\tau(s)\|_H^{2\alpha}\Big|\mathcal{F}_\tau\Big]\le C\|\hat{\xi}_\tau(\tau)\|_H^{2\alpha},\ P\text{-a.s.} \tag{3.24}$$

Moreover, note that

$$\hat{\xi}_\tau(\tau)=X^\tau(\tau)-\bar{X}(\tau)=\bar{X}(t)-\bar{X}(\tau).$$

We also note that since $\bar{X}\in L^{2,\alpha}_{\mathbb{F}}(0,T;V)$, so for a.e. t, $\bar{X}(t)\in L^{2\alpha}(\mathcal{F}_t,V)$. Then from Lemma 2.7, we have (for a.e. t)

$$\mathbb{E}\Big[\|\hat{\xi}_\tau(\tau)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]=\mathbb{E}\Big[\|\bar{X}(\tau)-\bar{X}(t)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]\le C(1+\big\|\bar{X}(t)\big\|_V^{2\alpha})|\tau-t|^\alpha,\ P\text{-a.s.},$$

and so

$$\mathbb{E}\Big[\sup_{\tau\le r\le T}\|\hat{\xi}_\tau(r)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]\le\mathbb{E}\Big[\mathbb{E}\Big[\sup_{\tau\le r\le T}\|\hat{\xi}_\tau(r)\|_H^{2\alpha}\Big|\mathcal{F}_\tau\Big]\Big|\mathcal{F}_t\Big]\le C\mathbb{E}\Big[\|\hat{\xi}_\tau(\tau)\|_H^{2\alpha}\Big|\mathcal{F}_t\Big]\le C_t|\tau-t|^\alpha,\ P\text{-a.s.} \tag{3.25}$$

We can write the equation of $\hat{\xi}_\tau(s)$ as

$$\hat{\xi}_\tau(s)=\hat{\xi}_\tau(\tau)+\int_\tau^s[\bar{A}(r)\hat{\xi}_\tau(r)+\varepsilon_1(r)]dr+\int_\tau^s[\bar{B}(r)\hat{\xi}_\tau(r)+\varepsilon_2(r)]dw(r) \tag{3.26}$$

and

$$\begin{aligned}\hat{\xi}_\tau(s)=\hat{\xi}_\tau(\tau)+&\int_\tau^s[\bar{A}(r)\hat{\xi}_\tau(r)+\frac{1}{2}\bar{a}_{xx}(r)(\hat{\xi}_\tau(r),\hat{\xi}_\tau(r))+\varepsilon_3(r)]\,dr\\&+\int_\tau^s[\bar{B}(r)\hat{\xi}_\tau(r)+\frac{1}{2}\bar{b}_{xx}(r)(\hat{\xi}_\tau(r),\hat{\xi}_\tau(r))+\varepsilon_4(r)]\,dw(r),\end{aligned} \tag{3.27}$$

where

$$\begin{aligned}\varepsilon_1(r)&:=\int_0^1\Big\langle a_x\Big(r,\bar{X}(r)+\mu\hat{\xi}_\tau(r),\bar{u}(r)\Big)-\bar{a}_x(r),\hat{\xi}_\tau(r)\Big\rangle d\mu,\\
\varepsilon_2(r)&:=\int_0^1\Big\langle b_x\Big(r,\bar{X}(r)+\mu\hat{\xi}_\tau(r),\bar{u}(r)\Big)-\bar{b}_x(r),\hat{\xi}_\tau(r)\Big\rangle d\mu,\\
\varepsilon_3(r)&:=\int_0^1(1-\mu)\Big[a_{xx}\Big(r,\bar{X}(r)+\mu\hat{\xi}_\tau(r),\bar{u}(r)\Big)-\bar{a}_{xx}(r)\Big]\Big(\hat{\xi}_\tau(r),\hat{\xi}_\tau(r)\Big)d\mu.\\
\varepsilon_4(r)&:=\int_0^1(1-\mu)\Big[b_{xx}\Big(r,\bar{X}(r)+\mu\hat{\xi}_\tau(r),\bar{u}(r)\Big)-\bar{b}_{xx}(r)\Big]\Big(\hat{\xi}_\tau(r),\hat{\xi}_\tau(r)\Big)d\mu.\end{aligned}$$

Step 2. We have, for any $\alpha\ge 2$,

$$\begin{aligned}&\mathbb{E}\Big[\int_\tau^T\|\varepsilon_1(r)\|_H^\alpha\,dr|\mathcal{F}_\tau\Big]\le C\|\hat{\xi}_\tau(\tau)\|_H^{2\alpha},\ P\text{-a.s.}\\
&\mathbb{E}\Big[\int_\tau^T\|\varepsilon_2(r)\|_H^\alpha\,dr|\mathcal{F}_\tau\Big]\le C\|\hat{\xi}_\tau(\tau)\|_H^{2\alpha},\ P\text{-a.s.}\\
&\mathbb{E}\Big[\int_\tau^T\|\varepsilon_3(r)\|_H^\alpha\,dr|\mathcal{F}_t\Big]=o(|\tau-t|^\alpha),\ P\text{-a.s.}\\
&\mathbb{E}\Big[\int_\tau^T\|\varepsilon_4(r)\|_H^\alpha\,dr|\mathcal{F}_t\Big]=o(|\tau-t|^\alpha),\ P\text{-a.s.}\end{aligned} \tag{3.28}$$

Indeed, from the a priori estimate for SEEs and (3.25), we obtain

$$\begin{aligned}\mathbb{E}\Big[\int_\tau^T \|\varepsilon_1(r)\|_H^\alpha\, dr|\mathcal{F}_\tau\Big] &= \int_\tau^T \mathbb{E}\Big[\|\varepsilon_1(r)\|_H^\alpha\Big|\mathcal{F}_\tau\Big]\, dr\\ &\le \int_\tau^T \mathbb{E}\Big[\int_0^1 \|a_x(r,\bar X(r)+\mu\hat\xi_\tau(r),\bar u(r))-\bar a_x(r)\|_H^{2\alpha} d\mu \|\hat\xi_\tau(r)\|_H^\alpha\Big|\mathcal{F}_\tau\Big] dr\\ &\le \int_\tau^T \mathbb{E}\Big[\|\hat\xi_\tau(r)\|_H^{2\alpha}\Big|\mathcal{F}_\tau\Big] dr\\ &\le C\|\hat\xi_\tau(\tau)\|_H^{2\alpha},\ P\text{-a.s.},\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}[\int_\tau^T \|\varepsilon_3(r)\|_H^\alpha\, dr|\mathcal{F}_t]\\ &= \int_\tau^T \mathbb{E}\Big[\|\varepsilon_3(r)\|_H^\alpha\Big|\mathcal{F}_t\Big] dr\\ &\le \int_\tau^T \mathbb{E}\Big[\int_0^1 (1-\mu)\Big|\Big[a_{xx}(r,\bar X(r)+\mu\hat\xi_\tau(r),\bar u(r))-\bar a_{xx}(r)\Big]\Big(\hat\xi_\tau(r),\hat\xi_\tau(r)\Big)\Big|^\alpha d\mu|\mathcal{F}_t\Big] dr\\ &\le \Big(\int_\tau^T \mathbb{E}\Big[\int_0^1 \Big\|a_{xx}\Big(r,\bar X(r)+\mu\hat\xi_\tau(r),\bar u(r)\Big)-\bar a_{xx}(r)\Big\|_{\mathfrak{L}(H)}^{2\alpha} d\mu\Big|\mathcal{F}_t\Big] dr\Big)^{\frac12}\Big(\int_t^T \mathbb{E}\Big[\int_0^1 \Big\|\hat\xi_\tau(r)\Big\|_H^{4\alpha} d\mu\Big|\mathcal{F}_t\Big] dr\Big)^{\frac12}\\ &\le o(|\tau-t|^\alpha),\ P\text{-a.s.}\end{aligned}$$

Step 3. Applying Itô's formula to $\langle p(r),\hat x(r)\rangle$, from (3.27) we get

$$\langle p(s),\hat\xi_\tau(s)\rangle = \langle h_x(\bar X(T)),\hat x(T)\rangle + \int_s^T J_1(r)dr - \int_s^T J_2(r)dw(r),\ s\in[\tau,T]. \tag{3.29}$$

where

$$\begin{aligned}J_1(s) :=&\langle k_x(s)+k_y(s)p(s)+k_z(s)q(s),\hat\xi_\tau(s)\rangle + k_z(s)\langle p(s),\bar B(s)\hat\xi_\tau(s)\rangle - \langle p(s),\varepsilon_3(s)\rangle - \langle q(s),\varepsilon_4(s)\rangle\\ &-\frac12[\langle p(s),(\bar a_{xx}(s)(\hat\xi_\tau(s),\hat\xi_\tau(s))\rangle + \langle q(s),\bar b_{xx}(s)(\hat\xi_\tau(s),\hat\xi_\tau(s))\rangle],\\ J_2(s) :=&\langle p(s),\bar B(s)(\hat\xi_\tau(s))\rangle + \langle q(s),\hat\xi_\tau(s)\rangle + \langle p(s),\varepsilon_4(s)\rangle\\ &+\frac12\langle p(s),\bar b_{xx}(s)(\hat\xi_\tau(s),\hat\xi_\tau(s))\rangle.\end{aligned}$$

Moreover, taking $\tilde t=\tau$, $\gamma_1=\varepsilon_1,\gamma_2=\varepsilon_2,x_0=\hat x(\tau)$ in Theorem 3.1 and according to Step 2,

$$\mu_1(\tau)=\|\hat\xi_\tau(\tau)\|_H^4 \text{ and } \mu_2(\tau)=\|\hat\xi_\tau(\tau)\|_H^2.$$

Thus applying Theorem 3.1 and (3.26), we have on $[\tau,T]$ that

$$\begin{aligned}\langle P(t)\hat x(t),\hat x(t)\rangle+\sigma(t) = \langle h_{xx}(\bar X(T))\hat x(T),\hat x(T)\rangle + \int_t^T [k_y(s)\langle P(s)\hat x(s),\hat x(s)\rangle\\ + k_z(s)\mathcal{Z}(s) + \langle G(s)\hat x(s),\hat x(s)\rangle]ds - \int_t^T \mathcal{Z}(s)dw(s),\end{aligned} \tag{3.30}$$

for some processes $(\sigma,\mathcal{Z})\in L^\alpha_{\mathbb{F}}(\tau,T)\times L^{2,\alpha}_{\mathbb{F}}(\tau,T)$ satisfying

$$\sup_{s\in[\tau,T]}\Big(\mathbb{E}\Big[|\sigma(s)|^\alpha\Big|\mathcal{F}_\tau\Big]\Big)^{\frac1\alpha} \le C\Big\|\hat\xi_\tau(\tau)\Big\|_H^4 \text{ and } \Big(\mathbb{E}\Big[\Big(\int_\tau^T|\mathcal{Z}(s)|^2ds\Big)^{\frac\alpha2}|\mathcal{F}_\tau\Big]\Big)^{\frac1\alpha}\le C\Big\|\hat\xi_\tau(\tau)\Big\|_H^2,\quad \text{for any } \alpha\ge2. \tag{3.31}$$

Consequently, on $[\tau, T]$

$$\langle p(t), \hat{\xi}_\tau(t)\rangle + \frac{1}{2}\langle P(t)\hat{\xi}_\tau(t), \hat{x}(t)\rangle + \frac{1}{2}\sigma(t) = \langle h_x(\bar{X}(T)), \hat{\xi}_\tau(T)\rangle + \frac{1}{2}\langle h_{xx}(\bar{x}(T))\hat{\xi}_\tau(T), \hat{\xi}_\tau(T)\rangle + \int_t^T I_1(s)ds - \int_t^T I_2(s)dw(s),$$

where

$$\begin{aligned} I_1(t) :=\ & \langle k_x(t) + k_y(t)p(t) + k_z(t)q(t), \hat{\xi}_\tau(t)\rangle + k_z(t)\langle p(t), \bar{B}(t)\hat{\xi}_\tau(t)\rangle + \frac{1}{2}\langle \{k_y(t)P(t) \\ & + D^2k(t)([I_d, p(t), \bar{B}^*(t)p(t) + q(t)], [I_d, p(t), \bar{B}^*(t)p(t) + q(t)]) + k_z(t)\langle p(t), \bar{b}_{xx}(t)\rangle\}\hat{\xi}_\tau(t), \hat{\xi}_\tau(t)\rangle \\ & + \frac{1}{2}k_z(t)\mathcal{Z}(t) - \langle p(t), \varepsilon_3(t)\rangle - \langle q(t), \varepsilon_4(t)\rangle, \\ I_2(t) :=\ & \langle p(t), \bar{B}(t)\hat{\xi}_\tau(t)\rangle + \langle q(t), \hat{\xi}_\tau(t)\rangle + \langle p(t), \varepsilon_4(t)\rangle \\ & + \frac{1}{2}\langle p(t), \bar{b}_{xx}(t)(\hat{\xi}_\tau(t), \hat{\xi}_\tau(t))\rangle + \frac{1}{2}\mathcal{Z}(t). \end{aligned}$$

Step 4. We denote, on $[\tau, T]$,

$$Y^\tau(s) = h(X^\tau(T)) + \int_s^T k(r, X^\tau(r), Y^\tau(r), Z^\tau(r), \bar{u}(r))dr - \int_s^T Z^\tau(r)dw(r).$$

Then we have

$$\begin{aligned} \hat{y}(s) - \frac{1}{2}\sigma(s) =\ & h(X^{x^1}(T)) - h(\bar{X}(T)) - \langle h_x(\bar{X}(T)), \hat{x}(T) + \hat{x}(T)\rangle - \frac{1}{2}\langle h_{xx}(\bar{X}(T))\hat{x}(T), \hat{x}(T)\rangle \\ & + \int_s^T \{k(r, X^\tau(r), Y^\tau(r), Z^\tau(r), \bar{u}(s)) - k(s, \bar{X}(s), \bar{Y}(s), \bar{Z}(s), \bar{u}(s)) - I_1(s)\}ds - \int_s^T \hat{z}(s)dw(s), \end{aligned} \tag{3.32}$$

where

$$\begin{aligned} \hat{y}(s) &:= Y^\tau(s) - \bar{Y}(s) - \langle p(s), \hat{\xi}_\tau(s)\rangle - \frac{1}{2}\langle P(s)\hat{\xi}_\tau(s), \hat{\xi}_\tau(s)\rangle, \\ \hat{z}(s) &:= Z^\tau(s) - \bar{Z}(s) - I_2(s). \end{aligned}$$

Then from a similar analysis as in Step 4 in the proof of Theorem 3.3, we obtain (for a.e. t)

$$\mathbb{E}\left[\hat{y}(\tau)|\mathcal{F}_t\right] = o(|\tau - t|),\ P\text{-a.s.}$$

This is also

$$\mathbb{E}\left[Y^\tau(\tau) - \bar{Y}(\tau)|\mathcal{F}_t\right] = \mathbb{E}\left[\left\langle p(\tau), \hat{\xi}_\tau(\tau)\right\rangle + \frac{1}{2}\langle P(\tau)\hat{\xi}_\tau(\tau), \hat{\xi}_\tau(\tau)\rangle|\mathcal{F}_t\right] + o(|\tau - t|),\ P\text{-a.s.} \tag{3.33}$$

On the other hand, note that, for any $\phi \in L^2_{\mathbb{F}}(0, T; V^*)$, from the Lebesgue differentiation theorem,

$$\begin{aligned} \mathbb{E}\Big[\Big\|\int_t^\tau \phi(r)dr\Big\|^2_{V^*}\Big|\mathcal{F}_t\Big] &\le (\tau - t)\int_t^\tau \mathbb{E}\Big[\|\phi(r)\|^2_{V^*}\Big|\mathcal{F}_t\Big]dr \\ &= O(|\tau - t|^2),\ \text{as } \tau \downarrow t,\ \text{for a.e. } t,\ P\text{-a.s.}, \end{aligned}$$

and for any $\psi \in L^2_{\mathbb{F}}(0, T; H)$,

$$\begin{aligned} \mathbb{E}\Big[\Big\|\int_t^\tau \psi(r)dw(r)\Big\|^2_{H}\Big|\mathcal{F}_t\Big] &\le \mathbb{E}\Big[\int_t^\tau \|\psi(r)\|^2_H dr\Big|\mathcal{F}_t\Big] \\ &= \int_t^\tau \mathbb{E}\Big[\|\psi(r)\|^2_H\Big|\mathcal{F}_t\Big]dr \\ &= O(|\tau - t|),\ \text{as } \tau \downarrow t,\ \text{for a.e. } t,\ P\text{-a.s.} \end{aligned}$$

Thus, from the formulas of $p(\tau)$ and $\xi_\tau(\tau)$, we obtain

$$\begin{aligned}
&\mathbb{E}\big[\langle p(\tau),\xi_\tau(\tau)\rangle\big|\mathcal{F}_t\big]=\mathbb{E}\Big[\langle p(t),\xi_\tau(\tau)\rangle\Big|\mathcal{F}_t\Big]+\mathbb{E}\Big[\langle p(\tau)-p(t),\xi_\tau(\tau)\rangle\Big|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[-\Big\langle p(t),\int_t^\tau[A(r)\bar X(r)+a(r,\bar X(r),\bar u(r))]dr\Big\rangle-\int_t^\tau\langle q(r),[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]\rangle dr\Big|\mathcal{F}_t\Big]\\
&+o(|\tau-t|),\ \text{as}\ \tau\downarrow t,\ \text{for a.e.}\ t,\ P\text{-a.s.}
\end{aligned}$$

Similarly, we also have

$$\begin{aligned}
&\mathbb{E}\Big[\langle P(\tau)\xi_\tau(\tau),\xi_\tau(\tau)\rangle\\
&-\langle P(\tau)\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r),\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle P(\tau)\xi_\tau(\tau),\xi_\tau(\tau)\rangle-\langle P(\tau)\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r),\xi_\tau(\tau)\rangle|\mathcal{F}_t\Big]\\
&+\mathbb{E}\Big[\langle P(\tau)\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r),\xi_\tau(\tau)\rangle\\
&-\langle P(\tau)\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r),\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle P(\tau)\int_t^\tau[A(r)\bar X(r)+a(r,\bar X(r),\bar u(r))]dr,\xi_\tau(\tau)\rangle|\mathcal{F}_t\Big]\\
&+\mathbb{E}\Big[\Big\langle P(\tau)\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r),\int_t^\tau[A(r)\bar X(r)+a(r,\bar X(r),\bar u(r))]dr\Big\rangle\Big|\mathcal{F}_t\Big]\\
&\le\Big(\mathbb{E}\Big[\|P(\tau)\|^4_{\mathfrak{L}(H)}\Big|\mathcal{F}_t\Big]\Big)^{\frac14}\Big(\mathbb{E}\Big[\Big\|\int_t^\tau[A(r)\bar X(r)+a(r,\bar X(r),\bar u(r))]dr\Big\|^2_H\Big|\mathcal{F}_t\Big]\Big)^{\frac12}\Big(\mathbb{E}\Big[\Big\|\xi_\tau(\tau)\Big\|^4_H\Big|\mathcal{F}_t\Big]\Big)^{\frac14}\\
&+\Big(\mathbb{E}\Big[\|P(\tau)\|^4_{\mathfrak{L}(H)}\Big|\mathcal{F}_t\Big]\Big)^{\frac14}\Big(\mathbb{E}\Big[\Big\|\int_t^\tau[B(r)\bar X(r)+b(r,\bar X(r),\bar u(r))]dw(r)\Big\|^4_H\Big|\mathcal{F}_t\Big]\Big)^{\frac14}\\
&\cdot\Big(\mathbb{E}\Big[\Big\|\int_t^\tau[A(r)\bar X(r)+a(r,\bar X(r),\bar u(r))]dr\Big\|^2_H\Big|\mathcal{F}_t\Big]\Big)^{\frac12}\\
&=o(|\tau-t|),\ \text{as}\ \tau\downarrow t,\ \text{for a.e.}\ t,\ P\text{-a.s.},
\end{aligned}$$

and

$$
\begin{aligned}
&\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r),\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r)\rangle\\
&-\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r),\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r)\rangle\\
&-\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&+\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r)\rangle\\
&-\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))-B(t)\bar{X}(t)-b(t,\bar{X}(t),\bar{u}(t))]dw(r),\\
&\qquad\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&+\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\\
&\qquad\int_t^\tau [B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))-B(t)\bar{X}(t)-b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&\le(\tau-t)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\|P(\tau)\|^2_{\mathfrak{L}(H)}\Big]\Big)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\int_t^\tau\Big\|B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))-B(t)\bar{X}(t)-b(t,\bar{X}(t),\bar{u}(t))\Big\|^4_H dr\Big]\Big)^{\frac{1}{4}}\\
&\cdot\Big(\mathbb{E}\Big[\int_t^\tau\Big\|B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))\Big\|^4_H dr\Big]\Big)^{\frac{1}{4}}\\
&+(\tau-t)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\|P(\tau)\|^2_{\mathfrak{L}(H)}\Big]\Big)^{\frac{1}{2}}\Big(\mathbb{E}\Big[\int_t^\tau\Big\|B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\Big\|^4_H dr\Big]\Big)^{\frac{1}{4}}\\
&\cdot\Big(\mathbb{E}\Big[\int_t^\tau\Big\|B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))-B(t)\bar{X}(t)-b(t,\bar{X}(t),\bar{u}(t))\Big\|^4_H dr\Big]\Big)^{\frac{1}{4}}\\
&=o(|\tau-t|),\ \text{as}\ \tau\downarrow t,\ \text{for a.e.}\ t,\ P\text{-a.s.},
\end{aligned}
$$

and from Proposition 2.4,

$$
\begin{aligned}
&\mathbb{E}\Big[\langle P(\tau)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle\\
&-\langle P(t)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle (P(\tau)-P(t))\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\Big]\\
&=\mathbb{E}\Big[\langle (P(\tau)-P(t))[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))](w(\tau)-w(t))^2|\mathcal{F}_t\Big]\\
&\le\big(\mathbb{E}\left[\langle (P(\tau)-P(t))[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\rangle^2|\mathcal{F}_t\right]\big)^{\frac{1}{2}}\left(\mathbb{E}\left[(w(\tau)-w(t))^4|\mathcal{F}_t\right]\right)^{\frac{1}{2}}\\
&=o(|\tau-t|),\ \text{as}\ \tau\downarrow t,\ P\text{-a.s., for a.e.}\ t.
\end{aligned}
$$

Moreover, from Itô's isometry, for any t,

$$\begin{aligned}&\mathbb{E}\left[\langle P(t)\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r),\int_t^\tau [B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))]dw(r)\rangle|\mathcal{F}_t\right]\\&=\langle P(t)[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\rangle(\tau-t),\ P\text{-a.s.}\end{aligned}$$

Thus,

$$\begin{aligned}&\mathbb{E}\left[\langle P(\tau)\xi_\tau(\tau),\xi_\tau(\tau)\rangle|\mathcal{F}_t\right]\\&=\langle P(t)[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\rangle(\tau-t)+o(|\tau-t|),\ \text{for a.e. } t,\ P\text{-a.s.}\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{E}\left[Y^\tau(\tau)-\bar{Y}(\tau)|\mathcal{F}_t\right]\\&=\mathbb{E}\Big[-\Big\langle p(t),\int_t^\tau [A(r)\bar{X}(r)+a(r,\bar{X}(r),\bar{u}(r))]dr\Big\rangle-\int_t^\tau\big\langle q(r),[B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]\big\rangle dr\Big|\mathcal{F}_t\Big]\\&+\frac{1}{2}\langle P(t)[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\rangle(\tau-t)+o(|\tau-t|),\ \text{for a.e. } t,\ P\text{-a.s.}\end{aligned}\tag{3.34}$$

Step 5. Let M be a countable dense subset of H. We can find a subset $\Omega_0\subset\Omega$ such that $P(\Omega_0)=1$ and for each $\omega_0\in\Omega_0$,

$V(t,\bar{X}(t,\omega_0))=\bar{Y}(t,\omega_0),\ Y^\tau(\tau,\omega_0)\geq V(\tau,\bar{X}(t,\omega_0)),$ (3.34) hold for all rational $\tau>t,$ for almost all $t,$ and $p(s,\omega_0)\in H,\ P(s,\omega_0)\in\mathfrak{L}(H),\forall s\in[0,T].$

Fix any $\omega_0\in\Omega_0$ (we shall only consider the corresponding t with full measure in the above, which may depend on ω_0). Then

$$\begin{aligned}&V(\tau,\bar{X}(t,\omega_0))-V(t,\bar{X}(t,\omega_0))=\mathbb{E}\left[V(\tau,\bar{X}(t))-V(t,\bar{X}(t,\omega_0))|\mathcal{F}_t\right]\\&\leq\mathbb{E}\left[Y^\tau(\tau)-\bar{Y}(t)|\mathcal{F}_t\right](\omega_0)\\&\leq\mathbb{E}\left[Y^\tau(\tau)-\bar{Y}(\tau)|\mathcal{F}_t\right](\omega_0)+\mathbb{E}\left[\bar{Y}(\tau)-\bar{Y}(t)|\mathcal{F}_t\right]\\&\leq\mathbb{E}\Big[-\Big\langle p(t),\int_t^\tau [A(r)\bar{X}(r)+a(r,\bar{X}(r),\bar{u}(r))]dr\Big\rangle-\int_t^\tau\big\langle q(r),[B(r)\bar{X}(r)+b(r,\bar{X}(r),\bar{u}(r))]\big\rangle dr\\&-\int_t^\tau k(r,\bar{X}(r),\bar{Y}(r),\bar{Z}(r),\bar{u}(r))dr\Big|\mathcal{F}_t\Big](\omega_0)\\&+\frac{1}{2}\langle P(t,\omega_0)[B(t,\omega_0)\bar{X}(t,\omega_0)+b(t,\bar{X}(t,\omega_0),\bar{u}(t,\omega_0))],B(t,\omega_0)\bar{X}(t,\omega_0)+b(t,\bar{X}(t,\omega_0),\bar{u}(t,\omega_0))\rangle(\tau-t)\\&+o(|\tau-t|),\ \text{along all rational } \tau>t.\end{aligned}$$

It follows that (we will omit ω_0 for notational simplicity in the following)

$$\begin{aligned}&V(\tau,\bar{X}(t))-V(t,\bar{X}(t))\leq\mathbb{E}\Big[-\Big\langle p(t),\int_t^\tau A(r)dr\Big\rangle-\int_t^\tau\big\langle q(r),B(r)\big\rangle dr\Big|\mathcal{F}_t\Big]\\&+\mathbb{E}\Big[-\Big\langle p(t),\int_t^\tau a(r,\bar{X}(r),\bar{u}(r))dr\Big\rangle-\int_t^\tau\big\langle q(r),b(r,\bar{X}(r),\bar{u}(r))\big\rangle dr\\&-\int_t^\tau k(r,\bar{X}(r),\bar{Y}(r),\bar{Z}(r),\bar{u}(r))dr\Big|\mathcal{F}_t\Big]\\&+\frac{1}{2}\langle P(t)[B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))],B(t)\bar{X}(t)+b(t,\bar{X}(t),\bar{u}(t))\rangle(\tau-t)+o(|\tau-t|)\\&=[-\langle p(t),A(t)\bar{X}(t)\rangle-\langle q(t),B(t)\bar{X}(t)\rangle+\mathcal{H}_1(t,\bar{X}(t),\bar{Y}(t),\bar{Z}(t))](\tau-t)+o(|\tau-t|).\end{aligned}\tag{3.35}$$

Then from the continuity of V, we obtain that this relationship holds for all $\tau \in (s,T]$. This implies

$$-\langle p(t), A(t)\bar{X}(t)\rangle - \langle q(t), B(t)\bar{X}(t)\rangle + \mathcal{H}_1(t,\bar{X}(t),\bar{Y}(t),\bar{Z}(t)),\infty) \subseteq D^{1,+}_{t+}V(t,\bar{X}(t)).$$

Now we prove the second part. Fix an $\omega \in \Omega$ such that (3.35) holds for any $\tau \in (s,T]$. Then for any $\hat{q} \in D^{1,-}_{t+}V(t,\bar{X}(t))$, we have

$$\begin{aligned} 0 &\leq \liminf_{\tau \downarrow t} \left\{ \frac{V(\tau,\bar{X}(t)) - V(t,\bar{X}(t)) - \hat{q}(\tau - t)}{|\tau - t|} \right\} \\ &\leq \liminf_{\tau \downarrow s} \left\{ \frac{(-\langle p(t), A(t)\bar{X}(t)\rangle - \langle q(t), B(t)\bar{X}(t)\rangle + \mathcal{H}_1(t,\bar{X}(t),\bar{Y}(t),\bar{Z}(t)) - \hat{q})(\tau - s)}{|\tau - s|} \right\} \end{aligned}$$

Then, it is necessary that $\hat{q} \leq -\langle p(t), A(t)\rangle - \langle q(t), B(t)\rangle + \mathcal{H}_1(t,\bar{X}(t),\bar{Y}(t),\bar{Z}(t))$. The proof is complete. □

3.4 Smooth case

In this section, we study the relationship between MP and DPP under the special smooth case, i.e., we assume that the value function $V(t,x)$ is sufficiently smooth.

We first present a result of Hamilton-Jacobi-Bellman equation, which is needed later. Consider the following Hamilton-Jacobi-Bellman (H-J-B) equation

$$\begin{cases} \partial_t V(t,x) & +\langle Ax, \partial_x V(t,x)\rangle_* + \frac{1}{2}\langle \partial^2_{xx} V(t,x)Bx, Bx\rangle + \frac{1}{2}\langle Pb(t,x,v), Bx\rangle + \frac{1}{2}\langle PBx, b(t,x,v)\rangle + \\ & \inf_{v\in U} G(t,x,v,V(t,x),\partial_x V(t,x),\partial^2_{xx}V(t,x)) = 0, \\ V(T,x) & = h(x). \end{cases} \tag{3.36}$$

where

$$\begin{aligned} G(t,x,v,r,p,P) &:= \frac{1}{2}\langle Pb(t,x,v), b(t,x,v)\rangle + \langle p, a(t,\eta,v)\rangle + k(t,x,r,b(t,x,v)^*p + pBx, v), \\ &(t,x,v,p,P) \in [0,T) \times H \times U \times H \times S(H). \end{aligned}$$

(It seems we cannot handle the term $\langle \partial^2_{xx} V(t,x)Bx, Bx\rangle$ in the representation of PDEs, so maybe we need to assume $B = 0$. Then the above PDE reads:

$$\begin{cases} \partial_t V(t,x) & +\langle A^*\partial_x V(t,x), x\rangle + \\ & \inf_{v\in U} G(t,x,v,V(t,x),\partial_x V(t,x),\partial^2_{xx}V(t,x)) = 0, \ (t,x) \in [0,T] \times H, \\ V(T,x) & = h(x), \ x \in H. \end{cases} \tag{3.37}$$

where

$$\begin{aligned} G(t,x,v,r,p,P) &:= \frac{1}{2}\langle Pb(t,x,v), b(t,x,v)\rangle + \langle p, a(t,\eta,v)\rangle + k(t,x,r,b(t,x,v)^*p, v), \\ &(t,x,v,p,P) \in [0,T) \times H \times U \times H \times S(H). \end{aligned}$$

In the following calculations, we shall assume $B = 0$.).

In the following of the paper, we assume that U is compact.

Proposition 3.5 *Suppose that the value function $V \in C^{1,2}([0,T] \times H)$, $\partial_x V(t,z) \in V$, for almost all $(t,z) \in [0,T] \times V$, and $A^*\partial_x V(t,z) \in C([0,T] \times H; H)$, then it satisfies the H-J-B equation (3.37).*

Remark 3.6 *In the recursive utility case, the corresponding H-J-B equation in the infinite dimension under the mild solution framework is presented in [48]. Here we shall present a result in the variational solution framework for the completeness of the paper, and the proof is given in the Appendix.*

Theorem 3.7 *Assume* $(H1)$ *and fix* $x \in H$. *Suppose* $(\bar{X}(\cdot), \bar{Y}(\cdot), \bar{Z}(\cdot), \bar{u}(\cdot))$ *are the optimal 4-tuple of Problem* (S_x) *and* $p(\cdot), q(\cdot), P(\cdot)$ *are the solutions of corresponding adjoint equations. Suppose that the value function* $V \in C^{1,2}([0,T] \times H)$ *and* $\partial_x V(t,z) \in V$, $A^* \partial_x V(t,z) \in H$, *for almost all* $(t,z) \in [0,T] \times H$. *Then*

$$
\begin{aligned}
&- \partial_t V(t, \bar{X}(t)) \\
&= \langle A\bar{X}(t), \partial_x V(t, \bar{X}(t)) \rangle_* + G(t, \bar{X}(t), V(t, \bar{X}(t)), \partial_x V(t, \bar{X}(t)), \partial_{xx} V(t, \bar{X}(t)), \bar{u}(t)) \\
&= \langle A\bar{X}(t), \partial_x V(t, \bar{X}(t)) \rangle_* + \inf_{v \in U} G(t, \bar{X}(t), V(t, \bar{X}(t)), \partial_x V(t, \bar{X}(t)), \partial_{xx} V(t, \bar{X}(t)), v), \quad P\text{-a.s. a.e.},
\end{aligned}
$$

If moreover, $V \in C^{1,3}([0,T] \times H)$ *with* $V_{tx} \in H$, *and* $\partial_{xx} V(t,z) \in V$, *for all* $(t,z) \in [0,T] \times V$. *Then*

$$
\begin{aligned}
p(t) &= \partial_x V(t, \bar{X}(t)), \quad P\text{-a.s. a.e.}, \\
q(t) &= \partial_{xx} V(t, \bar{X}(t)) b(t, \bar{X}(t), \bar{u}(t)), \quad P\text{-a.s. a.e.}
\end{aligned}
$$

Proof. From Lemma 3.2, we know that

$$
V(t, \bar{X}(t)) = \bar{Y}(t).
$$

Applying Itô's formula (see Lemma 2.15 in [37]), we obtain

$$
\begin{aligned}
dV(t, \bar{X}(t)) = [&\partial_t V(t, \bar{X}(t)) + \langle \partial_x V(t, \bar{X}(t)), a(t, \bar{X}(t), \bar{u}(t)) \rangle + \langle A^* \partial_x V(t, \bar{X}(t)), \bar{X}(t) \rangle \\
&+ \frac{1}{2} \langle \partial_{xx} V(t, \bar{X}(t)) b(t, \bar{X}(t), \bar{u}(t)), b(t, \bar{X}(t), \bar{u}(t)) \rangle] dt + \langle \partial_x V(t, \bar{X}(t)), b(t, \bar{X}(t), \bar{u}(t)) \rangle dw(t).
\end{aligned}
$$

Then from the uniqueness of Itô's composition, we get

$$
\begin{aligned}
&\partial_t V(t, \bar{X}(t)) + \langle \partial_x V(t, \bar{X}(t)), a(t, \bar{X}(t), \bar{u}(t)) \rangle + \langle A\bar{X}(t), \partial_x V(t, \bar{X}(t)) \rangle_* \\
&+ \frac{1}{2} \langle \partial_{xx} V(t, \bar{X}(t)) b(t, \bar{X}(t), \bar{u}(t)), b(t, \bar{X}(t), \bar{u}(t)) \rangle = -k(t, \bar{X}(t), \bar{Y}(t), \bar{Z}(t), \bar{u}(t))
\end{aligned}
$$

and

$$
\langle \partial_x V(t, \bar{X}(t)), b(t, \bar{X}(t), \bar{u}(t)) \rangle = \bar{Z}(t).
$$

From the first equality in the above, we get

$$
\begin{aligned}
&\partial_t V(t, \bar{X}(t)) + \langle A\bar{X}(t), \partial_x V(t, \bar{X}(t)) \rangle_* \\
&+ G(t, \bar{X}(t), \bar{u}(t), V(t, \bar{X}(t)), \partial_x V(t, \bar{X}(t)), \partial_{xx} V(t, \bar{X}(t))) = 0
\end{aligned}
\tag{3.38}
$$

This proves the first one. The second equality follows from the fact that V is the classical solution of HJB equation.

We then consider the second part. Taking into account (3.38) and the fact that V is the solution to HJB equation

$$
\partial_t V(t,x) + \langle A^* \partial_x V(t,x), x \rangle + G(t, x, V(s,x), \partial_x V(t,x), \partial_{xx} V(t,x), \bar{u}(t)) \geq 0,
$$

we derive that

$$
V_t(t,x) + \langle A^* \partial_x V(t,x), x \rangle + G(t, x, V(t,x), \partial_x V(t,x), \partial_{xx} V(t,x), \bar{u}(t))
$$

attains its minimum at $\bar{X}(t)$. Thus,

$$
\frac{\partial}{\partial x} \left\{ V_t(t,x) + \langle A^* \partial_x V(t,x), x \rangle + G(t, x, V(t,x), \partial_x V(t,x), \partial_{xx} V(t,x), \bar{u}(t)) \right\} \Big|_{x = \bar{X}(t)} = 0. \tag{3.39}
$$

From (3.39),

$$\begin{aligned}0 &= \partial_{tx}V(t,\bar{X}(t)) + A^*\partial_{xx}V(t,\bar{X}(t))\bar{X}(t) + A^*\partial_x V(t,\bar{X}(t))\\ &+ \partial_{xx}V(t,\bar{X}(t))a(t,\bar{X}(t),\bar{u}(t)) + \partial_x V(t,\bar{X}(t))a_x(t,\bar{X}(t),\bar{u}(t))\\ &+ \frac{1}{2}\partial_{xxx}V(t,\bar{X}(t))b^2(t,\bar{X}(t),\bar{u}(t)) + \partial_{xx}V(t,\bar{X}(t))b(t,\bar{X}(t),\bar{u}(t))b_x(t,\bar{X}(t),\bar{u}(t))\\ &+ f_x + f_y\partial_x V(t,\bar{X}(t)) + f_z(b_x(t,\bar{X}(t),\bar{u}(t))\partial_x V(t,\bar{X}(t)) + b(t,\bar{X}(t),\bar{u}(t))\partial_{xx}V(t,\bar{X}(t))).\end{aligned}$$

Then, applying Itô's formula to $\partial_x V(t,\bar{X}(t))$, we get

$$\begin{aligned}d\partial_x V(t,\bar{X}(t)) &= \partial_{xt}V(t,\bar{X}(t))dt + \partial_{xx}V(t,\bar{X}(t))d\bar{X}(t) + \frac{1}{2}\partial_{xxx}V(t,\bar{X}(t))b^2(t,\bar{X}(t),\bar{u}(t))dt\\ &= -[\partial_x V(t,\bar{X}(t))(A + a_x + f_y + f_z b_x) + \partial_{xx}V(t,\bar{X}(t))bb_x\\ &+ f_x + f_z b\partial_{xx}V(t,\bar{X}(t)))]ds + \partial_{xx}V(t,\bar{X}(t))b(t,\bar{X}(t),\bar{u}(t))dw.\end{aligned}$$

Moreover, from the boundary condition in the H-J-B equation,

$$\partial_x(T,\bar{X}) = h_x(\bar{X}).$$

So we can note that $\tilde{p}(t) = \partial_x V(t,\bar{X}(t))$, $\tilde{q}(t) = \partial_{xx}V(t,\bar{X}(t))b(t,\bar{X}(t),\bar{u}(t))$ also solves the first-order adjoint equation. Thus from the uniqueness of solutions, we obtain

$$p(t) = \partial_x V(t,\bar{X}(t)),\ q(t) = \partial_{xx}V(t,\bar{X}(t))b(t,\bar{X}(t),\bar{u}(t)).$$

The proof is complete. □

4 Appendix

4.1 Proof of Theorem 2.10

Proposition 4.1 *Suppose the assumptions* $(H4)$ *and* $(H5)$*, then*

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} Y^{t,x;u}(t).$$

Proof. Noting that $\mathcal{U}^t[t,T] \subset \mathcal{U}[t,T]$, so $V(t,x) \le \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} Y^{t,x;u}(t)$. On the other hand, for any $u(\cdot)\in\mathcal{U}[t,T]$, by Lemma 13 in [23], we know that there exists a sequence u^m taking the form

$$u^m(s) = \sum_{i=1}^{N_m} v^{mi}(s)I_{A_{mi}},\ s\in[t,T],$$

where $\{A_{mi}\}_{i=1}^{N_m}$ is a $\mathcal{F}_t$-partition of Ω and $v_s^{mi}\in\mathcal{U}^t[t,T]$, such that

$$\mathbb{E}\Big[\int_t^T |u^m(s) - u(s)|_U^2 dt\Big] \to 0,\ \text{as } m\to\infty.$$

From the a priori estimate of classical BSDEs,

$$\begin{aligned}&\mathbb{E}\Big[\Big|Y^{t,x;u^m}(t) - Y^{t,x;u}(t)\Big|^2\Big]\\ &\le C\mathbb{E}\Big[\int_t^T \big|g(s,X_s^{t,x;u},Y_s^{t,x;u},Z_s^{t,x;u},u_s^m) - g(s,X_s^{t,x;u},Y_s^{t,x;u},Z_s^{t,x;u},u_s)\big|^2 ds\Big]\\ &\le \mathbb{E}[\int_t^T |u^m(s) - u(s)|_U^2 dt]\\ &\to 0,\ \text{as } m\to\infty.\end{aligned} \tag{4.1}$$

Note that for $s \in [t,T]$,

$$\Big(X_s^{t,x;u^m}, Y_s^{t,x;u^m}, Z_s^{t,x;u^m}\Big) = \Big(\sum_{i=1}^m X_s^{t,x;v^{mi}} I_{A_i}, \sum_{i=1}^m Y_s^{t,x;v^{mi}} I_{A_i}, \sum_{i=1}^m Z_s^{t,x;v^{mi}} I_{A_i}\Big).$$

Then,

$$Y_t^{t,x;u^m} = \sum_{i=1}^m Y_t^{t,x;v^{mi}} I_{A_i} \geq \sum_{i=1}^m \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v} I_{A_i} = \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v}. \tag{4.2}$$

This, combining (4.1), implies

$$Y_t^{t,x;u} \geq \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v}.$$

Thus, $V(t,x) \geq \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v}$. Therefore, $V(t,x) = \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v}$. □

Lemma 4.2 *Assume (H4) and (H5). Then there exists a constant C depending on such that, for each $u \in \mathcal{U}[t,T]$ and $\xi, \xi' \in L^2(\mathcal{F}_t;H)$,*

$$\mathbb{E}\Big[\sup_{t\leq s\leq T}\Big(\|X_s^{t,\xi;u} - X_s^{t,\xi';u}\|_H^2 + |Y_s^{t,\xi;u} - Y_s^{t,\xi';u}|^2\Big) + \int_t^T |Z_s^{t,\xi;u} - Z_s^{t,\xi';u}|^2 ds\Big|\mathcal{F}_t\Big] \leq C\|\xi - \xi'\|_H^2; \tag{4.3}$$

$$\mathbb{E}\Big[\sup_{t\leq s\leq T}\big(\|X_s^{t,\xi;u}\|_H{}^2 + |Y_s^{t,\xi;u}|^2\big) + \int_t^T |Z_s^{t,\xi;u}|^2 ds\Big|\mathcal{F}_t\Big] \leq C\Big(1 + \|\xi\|_H^2\Big). \tag{4.4}$$

Proof. We only prove the second one, and the first one can be handled similarly. From the basic estimate for SEEs (see (2.24) and the inequality after it in [27]), we have

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\leq s\leq T}\|X_s^{t,\xi;u}\|_H{}^2\Big|\mathcal{F}_t\Big] &\leq C\mathbb{E}\Big[\|\xi\|_H^2 + \int_t^T \|a(s,0,u_s)\|_H^2 ds + \int_t^T \|b(s,0,u_s)\|_H^2 ds\Big|\mathcal{F}_t\Big]\\ &\leq C\Big(1 + \|\xi\|_H^2\Big).\end{aligned}$$

Then from the basic estimate of classical BSDEs, we obtain

$$\begin{aligned}&\mathbb{E}\Big[\sup_{t\leq s\leq T}|Y_s^{t,\xi;u}|^2 + \int_t^T |Z_s^{t,\xi;u}|^2 ds\Big|\mathcal{F}_t\Big]\\ &\leq C\mathbb{E}\Big[\|\xi\|_H^2 + \Big(\int_t^T |g|\left(s, X_s^{t,\xi;u}, 0, 0, u_s\right) ds\Big)^2\Big|\mathcal{F}_t\Big]\\ &\leq C\Big(1 + \|\xi\|_H^2\Big).\end{aligned}$$

This completes the proof. □

Lemma 4.3 *Under (H4) and (H5), we have for each $t \in [0,T]$ and $x, x' \in H$,*

$$|V(t,x) - V(t,x')| \leq C\|x - x'\|_H \text{ and } |V(t,x)| \leq C(1 + \|x\|_H).$$

Proof. Applying Proposition 4.1 and Lemma 4.2, we have

$$\begin{aligned}|V(t,x) - V(t,x')| &= \Big|\inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x;v} - \inf_{v\in\mathcal{U}^t[t,T]} Y_t^{t,x';v}\Big|\\ &\leq \sup_{v\in\mathcal{U}^t[t,T]}\Big|Y_t^{t,x;v} - Y_t^{t,x';v}\Big|\\ &\leq \sup_{v\in\mathcal{U}^t[t,T]}\Big\{\mathbb{E}\Big[\sup_{t\leq s\leq T}\Big|Y_s^{t,x;v} - Y_s^{t,x';v}\Big|^2\Big|\mathcal{F}_t\Big]\Big\}^{\frac{1}{2}}\\ &\leq C\|x - x'\|_H.\end{aligned}$$

The second inequality is obtained in the same manner. □

We first need the following preliminary result.

Proposition 4.4 *Suppose (H4) and (H5). Then for each $\xi \in L^2(\mathcal{F}_t; H)$, we have $V(t,\xi) = ess\inf_{u\in\mathcal{U}[t,T]} Y_t^{t,\xi;u}$. On the other hand, for each $\varepsilon > 0$, there exists an admissible control $u_\varepsilon(\cdot) \in \mathcal{U}[t,T]$ such that*

$$V(t,\xi) \geq Y_t^{t,\xi;u_\varepsilon} - \varepsilon, \tag{4.5}$$

Proof. We take a sequence $\xi^m = \sum_{i=1}^m x_i^m I_{A_i^m}$ such that $\mathbb{E}\left[\|\xi^m - \xi\|_H^2\right] \to 0$ as $m \to \infty$, where $\{A_i^m\}_{i=1}^m$ is a $\mathcal{F}_t$-partition of Ω and $x_i^m \in H$. Note that for $s \in [t,T]$,

$$\Big(X_s^{t,\xi_m;u}, Y_s^{t,\xi_m;u}, Z_s^{t,\xi_m;u}\Big) = \Big(\sum_{i=1}^m X_s^{t,x_i^m;u} I_{A_i^m}, \sum_{i=1}^m Y_s^{t,x_i^m;u} I_{A_i^m}, \sum_{i=1}^m Z_s^{t,x_i^m;u} I_{A_i^m}\Big).$$

Then we have

$$\begin{aligned}
\underset{u\in\mathcal{U}[t,T]}{ess\inf} Y_t^{t,\xi_m;u} &= \underset{u\in\mathcal{U}[t,T]}{ess\inf} \sum_{i=1}^m Y_t^{t,x_i^m;u} I_{A_i^m} \\
&= \sum_{i=1}^m \Big(\underset{u\in\mathcal{U}[t,T]}{ess\inf} Y_t^{t,x_i^m;u}\Big) I_{A_i^m} \\
&= \sum_{i=1}^m V\left(t,x_i^m\right) I_{A_i^m} \\
&= V\left(t,\xi^m\right).
\end{aligned} \tag{4.6}$$

By Lemmas 4.2 and 4.3, we have

$$\begin{aligned}
\Big|\underset{u\in\mathcal{U}[t,T]}{ess\inf} Y_t^{t,\xi_m;u} - \underset{u\in\mathcal{U}[t,T]}{ess\inf} Y_t^{t,\xi;u}\Big| &\leq \underset{u\in\mathcal{U}[t,T]}{ess\sup} \Big|Y_t^{t,\xi_m;u} - Y_t^{t,\xi;u}\Big| \\
&\leq C\|\xi^m - \xi\|_H,
\end{aligned} \tag{4.7}$$

$$\Big|V(t,\xi^m) - V\left(t,\xi\right)\Big| \leq C. \tag{4.8}$$

Combining (4.6), (4.7) and (4.8), we get

$$\Big|\underset{u\in\mathcal{U}[t,T]}{ess\inf} Y_t^{t,\xi;u} - V\left(t,\xi\right)\Big| \leq 2C\|\xi^m - \xi\|_H.$$

Then the desired result is deduced by letting $m \to \infty$.

Then we consider (4.5). From Proposition 11 in Chapter 1 of [9], we can find elementary function $\zeta' = \sum_{i=1}^{\infty} 1_{A_i} x_i$ such that

$$\|\zeta' - \zeta\|_H \leq \varepsilon$$

Then from Lemma 4.2 and Lemma 4.3, we have

$$\left|Y_t^{t,\zeta;u} - Y_t^{t,\zeta';u}\right| \leq C\varepsilon, \quad \left|V(t,\zeta) - V(t,\zeta')\right| \leq C\varepsilon.$$

For each x_i, by Proposition 4.1, we can take $u^i \in \mathcal{U}^t[t,T]$ such that

$$V(t,x_i) \geq Y_t^{t,x_i;u^i} - \varepsilon.$$

Let $u(\cdot)=\sum\limits_{i=1}^{\infty}1_{A_i}u^i(\cdot)$. Then

$$\begin{aligned}
Y_t^{t,\zeta;u} &= Y_t^{t,\zeta';u}+Y_t^{t,\zeta;u}-Y_t^{t,\zeta';u}\\
&\le Y_t^{t,\zeta';u}+C\varepsilon\\
&=\sum_{i=1}^{\infty}1_{A_i}Y_t^{t,x_i;u_i}+C\varepsilon\\
&\le \sum_{i=1}^{\infty}1_{A_i}(V(t,x_i)+\varepsilon)+C\varepsilon\\
&\le V(t,\zeta')+C\varepsilon\\
&\le V(t,\zeta)+C\varepsilon.
\end{aligned}$$

This completes the proof by noting that ε can be arbitrary. □

Proof of Theorem 2.10. We first prove $V(t,x)\ge \inf_{u(\cdot)\in\mathcal{U}^t[t,T]}G_{t,t+\delta}^{t,x;u}\big[V(t+\delta,X^{t,x;u}(t+\delta))\big]$. Given any $u\in\mathcal{U}[t,T]$. Note that

$$\begin{aligned}
Y^{t,x;u}(s) &= h(X^{t,x;u}(T))+\int_s^T k(r,X^{t,x;u}(r),Y^{t,x;u}(r),X^{t,x;u}(r),u(r))dr-\int_s^T X^{t,x;u}(r)dw(r)\\
&= Y^{t,x;u}(t+\delta)+\int_s^{t+\delta} k(r,X^{t,x;u}(r),Y^{t,x;u}(r),X^{t,x;u}(r),u(r))dr-\int_s^{t+\delta} X^{t,x;u}(r)dw(r),\quad s\in[t,t+\delta].
\end{aligned}$$

Thus,

$$G_{t,T}^{t,x;u}\big[h(X^{t,x;u}(T))\big]=G_{t,t+\delta}^{t,x;u}\big[Y^{t,x;u}(t+\delta)\big]. \tag{4.9}$$

On the other hand, by the uniqueness of the solution to (2.18), we have

$$X^{t,x;u}(s)=X^{t+\delta,X^{t,x;u}(t+\delta);u}(s),\quad s\in[t+\delta,T].$$

Thus, on $[t+\delta,T]$,

$$\begin{aligned}
Y^{t,x;u}(s) &= h(X^{t,x;u}(T))+\int_s^T k(r,X^{t,x;u}(r),Y^{t,x;u}(r),X^{t,x;u}(r),u(r))dr-\int_s^T Z^{t,x;u}(r)dw(r)\\
&= h(X^{t+\delta,X^{t,x;u}(t+\delta);u}(T))+\int_s^T k(r,X^{t+\delta,X^{t,x;u}(t+\delta);u}(s),Y^{t,x;u}(r),Z^{t,x;u}(r),u(r))dr-\int_s^T Z^{t,x;u}(r)dw(r).
\end{aligned}$$

So, from the uniqueness of solutions of BSDEs,

$$Y^{t,x;u}(s)=Y^{t+\delta,X^{t,x;u}(t+\delta);u}(s),\quad s\in[t+\delta,T]. \tag{4.10}$$

Consequently,

$$G_{t,t+\delta}^{t,x;u}\big[Y^{t,x;u}(t+\delta)\big]=G_{t,t+\delta}^{t,x;u}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta)\big]. \tag{4.11}$$

From (4.9) and (4.11), we see that

$$\begin{aligned}
V(t,x) &= \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]}G_{t,T}^{t,x;u}\big[h(X^{t,x;u}(T))\big]=\operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]}G_{t,t+\delta}^{t,x;u}\big[Y^{t,x;u}(t+\delta)\big]\\
&= \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]}G_{t,t+\delta}^{t,x;u}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta)\big].
\end{aligned} \tag{4.12}$$

and

$$\begin{aligned}
V(t,x) &= \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} Y^{t,x;u}(t) \\
&= \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} G^{t,x;u}_{t,T}\big[h(X^{t,x;u}(T))\big] \\
&= \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} G^{t,x;u}_{t,t+\delta}\big[Y^{t,x;u}(t+\delta)\big] \\
&= \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} G^{t,x;u}_{t,t+\delta}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta)\big].
\end{aligned}$$

By Proposition 4.4,

$$Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta) \geq V(t+\delta, X^{t,x;u}(t+\delta)),$$

From the comparison theorem of the classical BSDE, we have

$$G^{t,x;u}_{t,t+\delta}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta)\big] \geq G^{t,x;u}_{t,t+\delta}\big[V(t+\delta, X^{t,x;u}(t+\delta))\big].$$

Thus taking infimum over $u(\cdot)\in\mathcal{U}^t[t,T]$ on the both sides, we get

$$V(t,x) \geq \inf_{u(\cdot)\in\mathcal{U}^t[t,T]} G^{t,x;u}_{t,t+\delta}\big[V(t+\delta, X^{t,x;u}(t+\delta))\big].$$

Then we prove $V(t,x) \leq \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,t+\delta]} G^{t,x;u}_{t,t+\delta}[V(t+\delta, X^{t,x;u}(t+\delta))]$. Fix any $u(\cdot)\in\mathcal{U}[t,t+\delta]$. From Proposition 4.4, for any $\varepsilon>0$, we can find an admissible control $\bar{u}(\cdot)\in\mathcal{U}[t+\delta,T]$ such that

$$V(t+\delta, X^{t,x;u}(t+\delta)) \geq Y^{t+\delta,X^{t,x;u}(t+\delta);\bar{u}}(t+\delta) - \varepsilon.$$

Since $\tilde{u}(\cdot) := u(\cdot)I_{[t,t+\delta]} + \bar{u}(\cdot)I_{(t+\delta,T]} \in \mathcal{U}[t,T]$, from (4.12) and the comparison theorem of classical BSDE, we get

$$\begin{aligned}
V(t,x) &= \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,x;u}_{t,t+\delta}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);u}(t+\delta)\big] \\
&\leq \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,x;\tilde{u}}_{t,t+\delta}\big[Y^{t+\delta,X^{t,x;\tilde{u}}(t+\delta);\tilde{u}}(t+\delta)\big] \\
&= \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,x;u}_{t,t+\delta}\big[Y^{t+\delta,X^{t,x;u}(t+\delta);\bar{u}}(t+\delta)\big] \\
&\leq \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,\eta;u}_{t,t+\delta}\big[V(t+\delta, X^{t,x;u}(t+\delta)) + \varepsilon\big].
\end{aligned}$$

Then from the a priori estimate of BSDEs,

$$V(t,x) \leq \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,x;u}_{t,t+\delta}\big[V(t+\delta, X^{t,x;u}(t+\delta))\big] + C\varepsilon.$$

Letting $\varepsilon\to 0$, we obtain

$$V(t,x) \leq \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,T]} G^{t,x;u}_{t,t+\delta}\big[V(t+\delta, X^{t,x;u}(t+\delta))\big].$$

Combining the above analysis, we obtain the desired result. □

Now we state the continuity property of V in t.

Proposition 4.5 *Assume* $(H4)$ *and* $(H5)$ *hold. Then* V *is continuous in* t.

Proof. For each $(t,x) \in [0,T) \times H$ and $\delta \in (0, T-t]$, from Theorem 2.10, we have

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} G^{t,x;u}_{t,t+\delta}[V(t+\delta, X^{t,x;u}(t+\delta))].$$

Then

$$|V(t,x) - V(t+\delta,x)| \le \sup_{u\in\mathcal{U}^t[t,t+\delta]} \left| G^{t,x;u}_{t,t+\delta}[V(t+\delta, X^{t,x;u}(t+\delta))] - V(t+\delta,x) \right|.$$

For any $u \in \mathcal{U}^t[t,t+\delta]$, from the definition of $G^{t,x;u}_{t,t+\delta}[\cdot]$, we have

$$G^{t,x;u}_{t,t+\delta}\left[V(t+\delta, X^{t,x;u}(t+\delta))\right] = \mathbb{E}\Big[V(t+\delta, X^{t,x;u}(t+\delta)) + \int_t^{t+\delta} g\left(s, X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s), u(s)\right) ds\Big].$$

Then applying Lemma 4.3,

$$\begin{aligned}
&\left| G^{t,x;u}_{t,t+\delta}\Big[V(t+\delta, X^{t,x;u}(t+\delta))\Big] - V(t+\delta,x) \right| \\
&\le \mathbb{E}\Big[\left| V(t+\delta, X^{t,x;u}(t+\delta)) - V(t+\delta,x) \right| + \int_t^{t+\delta} \left| g\left(s, X^{t,x;u}(s), Y^{t,x;u}(s), Z^{t,x;u}(s), u(s)\right) \right| ds\Big] \qquad (4.13)\\
&\le C\mathbb{E}\Big[\left| X^{t,x;u}(t+\delta) - x \right| + \int_t^{t+\delta} \Big(1 + |X^{t,x;u}(s)| + |Y^{t,x;u}(s)| + |Z^{t,x;u}(s)|\Big) ds\Big]
\end{aligned}$$

Noting that

$$\mathbb{E}\Big[\sup_{t\le s\le t+\delta} \left(|X^{t,x;u}(s)|^2 + |Y^{t,x;u}(s)|^2\right) + \int_t^{t+\delta} |Z^{t,x;u}(s)|^2 ds\Big] \le C(1+\|x\|_H^2),$$

we have by the Cauchy-Schwarz inequality that

$$\mathbb{E}\Big[\int_t^{t+\delta} \Big(1 + \left|X^{t,x;u}(s)\right| + \left|Y^{t,x;u}(s)\right| + \left|Z^{t,x;u}(s)\right|\Big) ds\Big] \le C\left(1 + \|x\|_H\right)\delta^{\frac{1}{2}}. \qquad (4.14)$$

Moreover, $\mathbb{E}\Big[\left| X^{t,x;u}(t+\delta) - x \right|\Big] \to 0$ due to Remark 2.8. This completes the proof. $\square$

4.2 Proof of Lemma 2.6

In the proof, we denote by $\varepsilon > 0$ and $\gamma > 0$ two undetermined constants and for the sake of notation simplicity, we use C_1 to denote a generic constant independent of ε and γ, which may be different from line to line. By the coercivity condition,

$$\|Bu\|_H \le C_1\|u\|_V, \text{ for } u \in V.$$

Then

$$\begin{aligned}
&2\langle Az(s) + a(s,z(s)), z(s)\rangle_* + \|Bz(s) + b(s,z(s))\|_H^2 \\
&\le 2\langle Az(s), z(s)\rangle_* + \|Bz(s)\|_H^2 + 2\langle Bz(s), b(s,z(s))\rangle + 2\langle a(s,z(s)), z(s)\rangle_* \\
&\le -\delta\|z(s)\|_V^2 + K\|z(s)\|_H^2 + C(K)\|z(s)\|_V\|b(s,z(s))\|_H + 2\|a(s,z(s))\|_{V^*}\|z(s)\|_V \\
&\le -\delta\|z(s)\|_V^2 + K\|z(s)\|_H^2 + \frac{\delta}{2}\|z(s)\|_V^2 + C(\delta)\|b(s,z(s))\|_H^2 + C(\delta)\|a(s,z(s))\|_{V^*}^2 \\
&\le -\frac{\delta}{2}\|z(s)\|_V^2 + C_1\|z(s)\|_H^2 + C_1\|b(s,0)\|_H^2 + C_1\|a(s,0)\|_{V^*}^2,
\end{aligned}$$

and from the quasi-skew-symmetry condition

$$|\langle Bz(s)+b(s,z(s)),z(s)\rangle|^2 \le 2|\langle Bz(s),z(s)\rangle|^2+2|\langle b(s,z(s)),z(s)\rangle|^2 \le C_1\|z(s)\|_H^4+2\|b(s,0)\|_H^2\|z(s)\|_H^2$$

We have by the Hölder inequality and the Young's inequality that

$$\begin{aligned}
&\mathbb{E}[\int_t^T e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\|a(s,z(s))\|_{V^*}^2 ds|\mathcal{F}_t]\\
&\le \varepsilon^2\mathbb{E}[\sup_{t\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T e^{-\frac{\gamma s}{\alpha}}\|a(s,z(s))\|_{V^*}^2 ds)^{\alpha}|\mathcal{F}_t]\\
&\le \varepsilon^2\mathbb{E}[\sup_{t\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}|\mathcal{F}_t]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T e^{-\frac{\gamma s}{\alpha}}\|a(s,0)\|_{V^*}^2 ds)^{\alpha}|\mathcal{F}_t]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T e^{-\frac{\gamma s}{\alpha}}\|z(s)\|_H^2 ds)^{\alpha}|\mathcal{F}_t],
\end{aligned}$$

and similarly,

$$\mathbb{E}[\int_t^T e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\|b(s,z(s))\|_H^2)ds|\mathcal{F}_t] \le \varepsilon^2\mathbb{E}[\sup_{t\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}|\mathcal{F}_t]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T \|b(s,z(s))\|_H^2 dt)^{\alpha}|\mathcal{F}_t]$$

$$\le \varepsilon^2\mathbb{E}[\sup_{s\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}|\mathcal{F}_t]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T \|b(s,0)\|_H^2 ds)^{\alpha}|\mathcal{F}_t]+\frac{C_1}{\varepsilon^2}\mathbb{E}[(\int_t^T \|z(s)\|_H^2 ds)^{\alpha}|\mathcal{F}_t].$$

We can calculate

$$\begin{aligned}
&\mathbb{E}[\sup_{s\in[t,T]}|\int_s^T e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\langle Bz(s)+b(s),z(s)\rangle\, dw(s)|]\\
&\le C_1\mathbb{E}[(\int_t^T e^{-2\gamma s}\|z(s)\|_H^{4\alpha-4}|\langle Bz(s)+b(s),z(s)\rangle|^2 ds)^{\frac{1}{2}}]\\
&\le C_1\mathbb{E}[\sup_{s\in[t,T]} e^{-\frac{\gamma s}{2}}\|z(s)\|_H^{\alpha}(\int_t^T e^{-\gamma s}(\|z(s)\|_H^{2\alpha}+\|z(s)\|_H^{2\alpha-2}\|b(s,z(s))\|_H^2)ds)^{\frac{1}{2}}]\\
&\le \varepsilon\mathbb{E}[\sup_{s\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}]+\frac{C_1}{\varepsilon}\mathbb{E}[\int_t^T e^{-\gamma s}(\|z(s)\|_H^{2\alpha}+\|z(s)\|_H^{2\alpha-2}\|b(s,z(s))\|_H^2)ds]\\
&\le 2\varepsilon\mathbb{E}[\sup_{s\in[t,T]} e^{-\gamma s}\|z(s)\|_H^{2\alpha}]+\frac{C_1}{\varepsilon}\mathbb{E}[\int_t^T e^{-\gamma s}\|z(s)\|_H^{2\alpha}\, ds]+\frac{C_1}{\varepsilon}\mathbb{E}[(\int_t^T \|b(s)\|_H^2 ds)^{\alpha}].
\end{aligned}$$

Applying Itô formula to $e^{-\gamma s}\|z(s)\|_H^{2\alpha}$ on $[t,T]$, we obtain

$$\begin{aligned}
&e^{-\gamma s}\|z(s)\|_H^{2\alpha}+\gamma\int_t^s e^{-\gamma u}\|z(u)\|_H^{2\alpha}\, du\\
&=\|z_0\|_H^{2\alpha}+\alpha\int_t^s e^{-\gamma r}\|z(r)\|_H^{2(\alpha-1)}(2\langle Az(r)+a(r,z(r)),z(r)\rangle_*+\|Bz(r)+b(r)\|_H^2)\, dr\\
&+2\alpha(\alpha-1)\int_t^s e^{-\gamma r}\|z(r)\|_H^{2(\alpha-2)}|\langle Bz(r)+b(r,z(r)),z(r)\rangle|^2\, dr\\
&+2\alpha\int_t^s e^{-\gamma r}\|z(r)\|_H^{2(\alpha-1)}\langle Bz(r)+b(s,z(s)),z(s)\rangle\, dw(s)\\
&\le\|z_0\|_H^{2\alpha}+C_1\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}(-\frac{\delta}{2}\|z(s)\|_V^2+C_1\|z(s)\|_H^2+C_1\|b(s,0)\|_H^2+C_1\|a(s,0)\|_{V^*}^2)\, ds\\
&+C_1\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-2)}(\|z(s)\|_H^4+\|b(s,0)\|_H^2\|z(s)\|_H^2)\, ds\\
&+2\alpha\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\langle Bz(s)+b(s),z(s)\rangle\, dw(s).
\end{aligned}$$

Thus,

$$
\begin{aligned}
&e^{-\gamma t}\|z(t)\|_H^{2\alpha} + \gamma\int_0^t e^{-\gamma s}\|z(s)\|_H^{2\alpha}\,ds + \frac{\delta}{2}C_1\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\|z(s)\|_V^2\,ds\\
&\le \|z_0\|_H^{2\alpha} + C_1\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}(\|z(s)\|_H^2 + \|b(s,0)\|_H^2 + \|a(s,0)\|_{V^*}^2)\,ds\\
&+ C_1\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-2)}(\|z(s)\|_H^4 + \|b(s,0)\|_H^2\|z(s)\|_H^2)\,ds\\
&+ 2\alpha\int_0^t e^{-\gamma s}\|z(s)\|_H^{2(\alpha-1)}\langle Bz(s)+b(s), z(s)\rangle\,dw(s).
\end{aligned}
\tag{4.15}
$$

Taking supremum and expectation on both sides, we get

$$
\begin{aligned}
&\mathbb{E}[\sup_{t\in[t,T]} e^{-\gamma t}\|z(t)\|_H^{2\alpha}] + \gamma\mathbb{E}[\int_t^T e^{-\gamma t}\|z(t)\|_H^{2\alpha}\,dt]\\
&\le C_1(\varepsilon+\varepsilon^2)\mathbb{E}[\sup_{t\in[0,T]} e^{-\gamma t}\|z(t)\|_H^{2\alpha}] + \mathbb{E}[\|z_0\|_H^{2\alpha}] + C_1(1+\frac{1}{\varepsilon}+\frac{1}{\varepsilon^2})\{\mathbb{E}[\int_0^t e^{-\gamma t}\|z(t)\|_H^{2\alpha}\,ds]\\
&+ \mathbb{E}[\int_t^T (\|a(s,0)\|_{V^*}^2)^{\,\alpha}dt] + \mathbb{E}[\int_t^T (\|b(s,0)\|_H^2)^{\,\alpha}dt]\}.
\end{aligned}
$$

Choosing ε small and then γ large, we obtain

$$
\mathbb{E}[\sup_{t\in[t,T]}\|z(t)\|_H^{2\alpha}] \le C\,\mathbb{E}[\|z_0\|_H^{2\alpha}] + \mathbb{E}[\int_t^T (\|a(s,0)\|_{V^*}^2)^{\,\alpha}dt] + \mathbb{E}[\int_t^T (\|b(s,0)\|_H^2)^{\,\alpha}dt]. \tag{4.16}
$$

Now let $\alpha = 1$ in (4.15), we obtain

$$
\begin{aligned}
&\|z(t)\|_H^2 + \int_0^t \|z(s)\|_H^2\,ds + \int_0^t e^{-\gamma s}\|z(s)\|_V^2\,ds\\
&\le \|z_0\|_H^2 + C\int_0^t \|z(s)\|_H^2\,ds\\
&+ C_1\int_0^t (\|a(s,0)\|_{V^*}^2 + \|b(s,0)\|_H^2)\,ds + C(\alpha)\int_0^t e^{-\gamma s}\langle Bz(s)+b(s), z(s)\rangle\,dw(s).
\end{aligned}
$$

Then

$$
\begin{aligned}
&\mathbb{E}[(\int_t^T \|z(s)\|_V^2\,ds)^{\alpha}]\\
&\le \|z_0\|_H^{2\alpha} + C\mathbb{E}\int_0^T \|z(s)\|_H^{2\alpha}\,ds\\
&+ C\mathbb{E}\int_0^T (\|a(s,0)\|_{V^*}^2 + \|b(s,0)\|_H^2)^{\alpha}\,ds.
\end{aligned}
$$

Combining with (4.16), we get the desired result.

4.3 Proof of Proof of Proposition 3.5

Proof. Fix any $(t,x)\in[0,T]\times H$ and $\delta\in(0,T-t]$. From Theorem 2.10, we know that

$$
V(t,x) = \operatorname*{ess\,inf}_{u(\cdot)\in\mathcal{U}[t,t+\delta]} G_{t,t+\delta}^{t,x;u}[V(t+\delta, X^{t,x;u}(t+\delta))] = \inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} G_{t,t+\delta}^{t,x;u}[V(t+\delta, X^{t,x;u}(t+\delta))]
$$

For any fixed control $u \in \mathcal{U}^t[t, t+\delta]$, let $X(s) := X^{t,x;u}(s)$ and let (Y, Z) be the solution of (2.21). Applying Itô's formula, we have that

$$V(s, X_s) = V(t+\delta, X_{t+\delta}) - \int_s^{t+\delta} [\partial_s V(r, X_r) + \langle \partial_x V(r, X_r), a(r, X_r, u(r))\rangle + \langle A^* \partial_x V(r, X_r), X_r\rangle$$
$$+ \frac{1}{2}\langle \partial^2_{xx} V(r, X_r) b(r, X_r, u(r)), b(r, X_r, u(r))\rangle] dr - \int_s^{t+\delta} \langle \partial_x V(r, X_r), b(r, X_r, u(r))\rangle dw(r), \ s \in [t, t+\delta].$$

Thus,

$$\begin{aligned} Y(s) - V(s, X_s) = &\int_s^{t+\delta} [k(r, X(r), Y(r), Z(r), u(r)) + \partial_s V(r, X_r) + \langle \partial_x V(r, X_r), a(r, X_r, u(r))\rangle \\ &+ \langle A^* \partial_x V(r, X_r), X_r\rangle + \frac{1}{2}\langle \partial^2_{xx} V(r, X_r) b(r, X_r, u(r)), b(r, X_r, u(r))\rangle] dr \\ &- \int_s^{t+\delta} [Z(r) - \langle \partial_x V(r, X_r), b(r, X_r, u(r))\rangle] dw(r), \ s \in [t, t+\delta]. \end{aligned} \tag{4.17}$$

We denote

$$\hat{Y}(s) := Y(s) - V(s, X_s) \text{ and } \hat{Z}(s) := Z(s) - \langle \partial_x V(s, X_s), b(s, X_s, u(s))\rangle.$$

Then (4.17) can be written as

$$\begin{aligned} \hat{Y}(s) = &\int_s^T [k(r, X(r), \hat{Y}(r) + V(r, X_r), \hat{Z}(r) + \langle \partial_x V(r, X_r), b(r, X_r, u(r))\rangle, u(r)) \\ &+ \partial_s V(r, X_r) + \langle \partial_x V(r, X_r), a(r, X_r, u(r))\rangle + \langle A^* \partial_x V(r, X_r), X_r\rangle \\ &+ \frac{1}{2}\langle \partial^2_{xx} V(r, X_r) b(r, X_r, u(r)), b(r, X_r, u(r))\rangle] dr - \int_s^{t+\delta} \hat{Z}(r) dw(r), \ s \in [t, t+\delta], \end{aligned}$$

which is a (real-valued) BSDE with $\hat{Y}(s), \hat{Z}(s)$ being the solutions.

We consider another BSDE

$$\begin{aligned} \hat{Y}^1(s) = &\int_s^{t+\delta} [k(r, X(r), \hat{Y}^1(r) + V(r, x), \hat{Z}(r) + \langle \partial_x V(r, x), b(r, x, u(r))\rangle, u(r)) \\ &+ \partial_s V(r, x) + \langle \partial_x V(r, x), a(r, x, u(r))\rangle + \langle A^* \partial_x V(r, x), x\rangle \\ &+ \frac{1}{2}\langle \partial^2_{xx} V(r, x) b(r, x, u(r)), b(r, x, u(r))\rangle] dr - \int_s^{t+\delta} \hat{Z}^1(r) dw(r), \ s \in [t, t+\delta]. \end{aligned} \tag{4.18}$$

From the a priori estimate for classical BSDEs, we have

$$\begin{aligned} |\hat{Y}(t) - \hat{Y}^1(t)|^2 &\le \delta \int_t^{t+\delta} \mathbb{E}\big[\sup_{t \le s \le t+\rho} \|X(r) - x\|^2_H \,\big|\mathcal{F}_t\big] dr \\ &\le \delta^2 \mathbb{E}\big[\sup_{t \le s \le t+\rho} \|X(r) - x\|^2_H \,\big|\mathcal{F}_t\big] \\ &= o(\delta^2). \end{aligned}$$

We denote the generator in BSDE (4.18) by

$$\begin{aligned} F(s, x, y, z, v) := &k(s, x, y + V(s, x), z + \langle \partial_x V(r, x), b(r, x, v)\rangle, v) \\ &+ \partial_s V(s, x) + \langle \partial_x V(s, x), a(s, x, v)\rangle + \langle A^* \partial_x V(s, x), x\rangle \\ &+ \frac{1}{2}\langle \partial^2_{xx} V(s, x) b(s, x, v), b(s, x, v)\rangle. \end{aligned}$$

We consider the backward ODE (which is also a BSDE with solutions $(\hat{Y}^0, 0)$ since the coefficients of it is deterministic)

$$\hat{Y}^0(s) = \int_s^{t+\delta} F_0(r, x, \hat{Y}^0(r), 0)dr, \ s \in [t, t+\delta], \tag{4.19}$$

where

$$\begin{aligned} F_0(s,x,y,z) &= \inf_{v\in U} F(s,x,y,z,v) \\ &= \partial_s V(s,x) + \langle A^*\partial_x V(s,x), x\rangle + \inf_{v\in U}[\langle \partial_x V(s,x), a(s,x,v)\rangle + \frac{1}{2}\langle \partial^2_{xx} V(s,x) b(s,x,v), b(s,x,v)\rangle \\ &+ k(r,x,y+V(s,x), z+\langle \partial_x V(r,x), b(r,x,v)\rangle, v)]. \end{aligned}$$

From the DPP, we have the following semegroup property:

$$V(t,x) = \inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} Y(t).$$

Thus,

$$\inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} \hat{Y}(t) = 0.$$

So, from Lemma 4.6,

$$\hat{Y}^0(t) = \inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} \hat{Y}^1(t) = o(\delta).$$

Letting $s = t$, and then divide by $\delta > 0$ on both sides of (4.19) and letting $\delta \downarrow 0$, from the formula for uppper limit integral, we get

$$F_0(t,x,0,0) = 0,$$

which is just HJB eqation. The proof is complete. □

Lemma 4.6 *For $\hat{Y}^1(t)$ and $\hat{Y}^0(t)$ defined as in the proof of Proposition 3.5. We have*

$$\inf_{u(\cdot)\in\mathcal{U}^t[t,t+\delta]} \hat{Y}^1(t) = \hat{Y}^0(t)$$

Proof. For each given $u(\cdot) \in \mathcal{U}^t[t, t+\delta]$, $F(s,x,0,0,u(s)) \geq F_0(s,x,0,0)$, so by the comparison theorem of classical BSDEs, $\hat{Y}^1(t) \geq Y(t)$. On the other hand, by the measurable selection theorem and the compactness of U, there exists (deterministic) measurable function $\bar{a} : [0,T] \times H \times \mathbb{R} \times \mathbb{R} \to \mathbb{U}$ such that

$$F_0(s,x,y,z) = F_1(s,x,y,z,\bar{a}(s,x,y,z)), \ (s,x,y,z) \in [0,T] \times H \times \mathbb{R} \times \mathbb{R}.$$

We define $u(s) = \bar{a}(s, x, \hat{Y}^0(s), 0)$. At this case, from the uniqueness of solutions of BSDEs, we have $\hat{Y}^{1,u}(s) = \hat{Y}^0(s), s \in [t, t+\delta]$, in particular, $\hat{Y}^{1,u}(t) = \hat{Y}^0(t)$. □

Acknowledgement.

References

[1] E.N. Barron and R. Jensen, The Pontryagin maximum principle from dynamic programming and viscosity solutions to first-order partial differential equations, Trans. Am. Math. Soc. 298 (1986) 635-641.

[2] A. Bensoussan, Lectures on stochastic control, Lecture Notes in Mathematics, vol. 972, Springer-Verlag, Berlin, 1982.

[3] P. Cannarsa and H. Frankowska, Value function and optimality conditions for semilinear control problems, Appl. Math Optim., 26 (1992), 139-169.

[4] P. Cannarsa and H. Frankowska, Value function and optimal conditions for semilinear coontrol problems. II: parabolic case, Appl. Math. Optim., 33 (1996), 1-33.

[5] A. Cernea, H. Frankowska, A connection between the maximum principle and dynamic programming for constrained control problems, SIAM J. Control Optim. 44 (2005) 673-703.

[6] L. Chen, Q. Lü, Relationships between the maximum principle and dynamic programming for infinite dimensional stochastic control systems. Journal of Differential Equations, 2023, 358: 103-146.

[7] J. Diestel and J. J. Uhl, Vector Measures. Mathematical Surveys, No. 15. American Mathematical Society, Providence, R.I., 1977.

[8] F. H. Clarke, R.B. Vinter, The relationship between the maximum principle and dynamic programming, SIAM J. Control Optim. 25 (1987) 1291-1311.

[9] N. Dinculeanu, Vector integration and stochastic integration in Banach spaces. Pure Appl. Math. Wiley-Interscience, New York, 2000.

[10] K. Du and Q. Meng, A revisit to W_2^n-theory of super-parabolic backward stochastic partial differential equations in $\mathbb{R}^d$. Stochastic Process. Appl. 120 (2010), no. 10, 1996-2015.

[11] K. Du and Q. Meng, A maximum principle for optimal control of stochastic evolution equations. SIAM J. Control Optim. 51 (2013), no. 6, 4343-4362.

[12] D. Duffie and L. G. Epstein, Stochastic differential utility. Econometrica 60 (1992), no. 2, 353-394.

[13] N. El Karoui, Les aspects probabilistes du contrôle stochastique. Ninth Saint Flour Probability Summer School-1979, pp. 73-238, Lecture Notes in Math., 876, Springer, Berlin-New York, 1981.

[14] N. El Karoui, S. Peng and M. C. Quenez, Backward stochastic differential equations in finance. Math. Finance 7 (1997), no. 1, 1-71.

[15] G. Fabbri, F. Gozzi and A. Swiech, Stochastic optimal control in infinite dimension. Dynamic programming and HJB equations. With a contribution by Marco Fuhrman and Gianmario Tessitore Probab. Theory Stoch. Model., 82 Springer, Cham, 2017.

[16] M. Fuhrman, Y. Hu and G. Tessitore, Stochastic maximum principle for optimal control of SPDEs. C. R. Math. Acad. Sci. Paris 350 (2012), no. 13-14, 683-688.

[17] M. Fuhrman, Y. Hu and G. Tessitore, Stochastic maximum principle for optimal control of SPDEs. Appl. Math. Optim. 68 (2013), no. 2, 181-217.

[18] G. Guatteri and G. Tessitore, On the backward stochastic Riccati equation in infinite dimensions. SIAM J. Control Optim. 44 (2005), no. 1, 159-194.

[19] G. Guatteri and G. Tessitore, Well posedness of operator valued backward stochastic Riccati equations in infinite dimensional spaces. SIAM J. Control Optim. 52 (2014), no. 6, 3776-3806.

[20] I. Gyöngy. On stochastic equations with respect to semimartingales III. Stochastics, 7(4): 231-254, 1982.

[21] P. Halmos, A Hilbert Space Problem Book. Second edition. Encyclopedia of Mathematics and its Applications, 17. Graduate Texts in Mathematics, 19. Springer-Verlag, New York-Berlin, 1982.

[22] M. Hu, Stochastic global maximum principle for optimization with recursive utilities. Probab. Uncertain. Quant. Risk 2 (2017), Paper No. 1, 20 pp.

[23] M. Hu and S. Ji, Stochastic maximum principle for stochastic recursive optimal control problem under volatility ambiguity. SIAM J. Control Optim. 54 (2016), no. 2, 918-945.

[24] J. Nedoma, Note on the generalized random variables. Trans. of the first Prague conference on information theory, etc. 1957, 139-141.

[25] J. van Neerven, Stochastic Evolution Equations. ISEM Lecture Notes. University of Delft, Delft, 2007.

[26] N. V. Krylov and B. L. Rozovskii, Stochastic evolution equations. J. Sov. Math. 16 (1981), no. 4, 1233-1277.

[27] G. Liu and S. Tang. Maximum principle for optimal control of stochastic evolution equations with recursive utilities. SIAM J. Control Optim., 61(6):3467–3500, 2023.

[28] G. Liu and S. Tang, Maximum principle for optimal control of stochastic evolution equations with recursive utilities. arXiv:2112.03165, 2021.

[29] G. Liu, J. Song, and M. Wang. Maximum principle for recursive optimal control problem of stochastic delay evolution equations. *arXiv preprint arXiv:2310.11376*, 2023.

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

[31] Q. Lü and X. Zhang, Operator-valued backward stochastic Lyapunov equations in infinite dimensions, and its application. Math. Control Relat. Fields 8 (2018), no. 1, 337-381.

[32] Q. Lü and X. Zhang, Mathematical control theory for stochastic partial differential equations. Probab. Theory Stoch. Model., 101 Springer, Cham, 2021.

[33] T. Nie, J. Shi and Z. Wu, Connection between MP and DPP for Stochastic Recursive Optimal Control Problems: Viscosity Solution Framework in Local Case, in Proceedings of the 2016 American Control Conference, Boston, 2016, pp. 7225-7230.

[34] T. Nie, J. Shi and Z. Wu, Connection between MP and DPP for Stochastic Recursive Optimal Control Problems: Viscosity Solution Framework in the general Case, SIAM J. Control Optim., 55 (2017), pp. 3258-3294.

[35] E. Pardoux, Sur des équations aux dérivées partielles stochastiques monotones. C. R. Acad. Sci., 275: A101–A103, 1972.

[36] E. Pardoux, Équations aux dérivées partielles stochastiques non linéaires monotones. Ph.D. thesis, Université Paris XI, 1975.

[37] E. Pardoux, Stochastic partial differential equations-an introduction. SpringerBriefs in Mathematics. Springer, Cham, 2021.

[38] S. Peng, A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. 28 (1990), no. 4, 966-979.

[39] S. Peng, A generalized dynamic programming principle and Hamilton-Jacobi-Bellman equation, Stoch. & Stoch. Rep., 38 (1992), pp. 119-134.

[40] S. Peng, Backward stochastic differential equations and applications to optimal control. Appl. Math. Optim. 27 (1993), no. 2, 125-144.

[41] S. Peng, Open problems on backward stochastic differential equations. Control of distributed parameter and stochastic systems (Hangzhou, 1998), 265-273, Kluwer Acad. Publ., Boston, MA, 1999.

[42] B. J. Pettis, On integration in vector spaces. Trans. Amer. Math. Soc. 44 (1938), no. 2, 277-304.

[43] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The mathematical theory of optimal processes. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt Interscience Publishers John Wiley & Sons, Inc., New York-London 1962.

[44] B. L. Rozovsky and S. V. Lototsky, Stochastic Evolution Systems. Linear Theory and Applications to Non-linear Filtering. Second edition. Probability Theory and Stochastic Modelling. Springer, Cham, 2018.

[45] J. Shi, The relationship between maximum principle and dynamic programming principle for stochastic recursive optimal control problems and applications to finance, in Proceedings of the 29th Chinese Control Conference, Beijing, China, 2010, pp. 1535-1540.

[46] J. Shi and Y. Yu, Relationship between maximum principle and dynamic programming for stochastic recursive optimal control problems and applications, Math. Probl. Eng., 2013, 285241.

[47] W. Stannat and L. Wessels, Peng's maximum principle for stochastic partial differential equations. SIAM J. Control Optim. 59 (2021), no. 5, 3552-3573.

[48] S. Tang , J. Zhou, Optimal Control of Unbounded Functional Stochastic Evolution Systems in Hilbert Spaces: Second-Order Path-dependent HJB Equation. arXiv preprint arXiv:2402.15998, 2024.

[49] J. Yan, S. Peng, S. Fang and L. Wu, Topics on Stochastic Analysis, Science Press, Beijing, 1997 (in Chinese).

[50] J. Yong and X. Zhou, Stochastic controls. Hamiltonian systems and HJB equations. Applications of Mathematics (New York), 43. Springer-Verlag, New York, 1999.

[51] X.Y. Zhou, Maximum principle, dynamic programming, and their connection in deterministic control, J. Optim. Theory Appl. 65 (1990) 363-373.

[52] X. Y. Zhou, The connection between the maximum principle and dynamic programming in stochastic control, Stoch. & Stoch. Rep., 31 (1990), pp. 1-13.

[53] X.Y. Zhou, A unified treatment of maximum principle and dynamic programming in stochastic control, Stoch. Stoch. Rep. 36 (1991) 137-161.