



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Faculty of Engineering and Information Sciences -
Papers: Part A

Faculty of Engineering and Information Sciences

2013

A maximum principle for optimal control of stochastic evolution equations

Kai Du

University of Wollongong, kaid@uow.edu.au

Qingxin Meng

Huzhou University, mqx@hutczj.cn

Publication Details

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

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:
research-pubs@uow.edu.au

A maximum principle for optimal control of stochastic evolution equations

Abstract

A general maximum principle is proved for optimal controls of abstract semilinear stochastic evolution equations. The control variable and linear unbounded operators act in both drift and diffusion terms, and the control set need not be convex. Read More: <http://epubs.siam.org/doi/abs/10.1137/120882433>

Keywords

evolution, stochastic, control, equations, optimal, maximum, principle

Disciplines

Engineering | Science and Technology Studies

Publication Details

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

A MAXIMUM PRINCIPLE FOR OPTIMAL CONTROL OF STOCHASTIC EVOLUTION EQUATIONS

KAI DU* AND QINGXIN MENG†

Abstract. A general maximum principle is proved for optimal controls of abstract semilinear stochastic evolution equations. The control variable, as well as linear unbounded operators, acts in both drift and diffusion terms, and the control set need not be convex.

Key words. Maximum principle, stochastic evolution equation, L^p estimate, stochastic bilinear functional, operator-valued stochastic process

AMS subject classifications. 49K27, 93E20, 60H25

1. Introduction. In this paper, we study an abstract infinite-dimensional stochastic control problem whose purpose is to minimize a cost functional

$$J(u) = \mathbf{E} \int_0^1 l(t, x(t), u(t)) dt + \mathbf{E}h(x(1))$$

subject to the semilinear stochastic evolution equation (SEE)

$$(1.1) \quad \begin{cases} dx(t) = [A(t)x(t) + f(t, x(t), u(t))]dt + [B(t)x(t) + g(t, x(t), u(t))]dW_t, \\ x(0) = x_0, \end{cases}$$

where the state process $x(\cdot)$ and control $u(\cdot)$ take values in infinite dimensional spaces, $A(t)$ and $B(t)$ are both random unbounded operators, f, g, l and h are given random functions, and W is a standard Wiener process.

A classical approach for optimal control problems is to derive necessary conditions satisfied by an optimum, such as Pongtryagin's maximum principle (cf. [12]). Since 1970s, the maximum principle was extensively studied for stochastic control systems: in the finite dimensional case, it has been solved by Peng [11] in a general setting where the control was allowed to take values in a nonconvex set and enter into the diffusion, while in the infinite-dimensional case, the existing literature, e.g. [1, 8, 14, 15, 17], required at least one of the three assumptions that 1) the control domain was convex, 2) the diffusion did not depend on the control, and 3) the state equation and cost functional were both linear in the state variable. So far the general maximum principle for infinite-dimensional stochastic control systems, i.e., the counterpart of Peng's result, remained open for a long time. In this paper we attempt to fill this gap.

In view of the second-order variation method developed by Peng [11], the main difficulty in the infinite-dimensional case lies in the step of second-order duality analysis, i.e., analyzing the quadratic term in a variational inequality, which, in the finite-dimensional case, can be worked out by means of the fact that the auxiliary process,

*Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland (kai.du@math.ethz.ch). Financial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management) is gratefully acknowledged. The NCCR FINRISK is a research instrument of the Swiss National Science Foundation.

†Department of Mathematical Sciences, Huzhou University, Zhejiang 313000, China (mqx@hutc.zj.cn). Financial support by the National Natural Science Foundation of China (Grant No. 11101140, No.11301177), China Postdoctoral Science Foundation (2011M500721, 2012T50391), and the Natural Science Foundation of Zhejiang Province (Grant No. Y6110775, Y6110789) is gratefully acknowledged.

called the second-order adjoint process, satisfies a well-solved backward stochastic differential equation (BSDE). However, in the case here, the corresponding BSDE is operator-valued which seems rather difficult to solve. In this paper, we develop a new procedure to do the second-order duality analysis: by virtue of the Lebesgue differentiation theorem and an approximation argument, we prove that, when the perturbation tends to zero, the quadratic term in the variational inequality converges to a bilinear functional which can be represented by an operator-valued process — this is the second-order adjoint process. Our approach, we believe, could work not only in the abstract framework but also in many concrete cases.

Very recently, two other works [10, 5] besides ours were also concerned with the general stochastic maximum principle in infinite dimensions, while the three ones differ in both forms of state equations and key approaches. The preprint [10], within an abstract framework, focused on how to solve the operator-valued BSDE properly, and to this end, introduced a notion called “relaxed transposition solution”. Their approach and result looked restrictive due to some technical assumptions. In [5] as well as its long version [6], the authors considered a concrete stochastic parabolic PDE with deterministic coefficients. The approach there, including a compactness argument, required the Markov structure of the system.

The rest of this paper is organized as follows. In Section 2, we formulate the problem in abstract form, and state our main results. Section 3 is devoted to derive the basic L^p -estimate for SEEs. In Section 4, we study the representation and properties of a stochastic bilinear function. The maximum principle is proved in Section 5. In the final section, we discuss two examples.

We finish the introduction with several notations. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ be a filtered probability space with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ generated by a 1-dimensional¹ standard Wiener process $\{W_t; t \geq 0\}$ and satisfying the usual conditions. Denote by \mathcal{P} the predictable σ -field. Let H be a separable real Hilbert space and $\mathcal{B}(H)$ be its Borel σ -field, $p \in [1, \infty]$. The following classes of processes will be used in this article.

- $L^p(\Omega \times [0, 1], \mathcal{P}, H)$: the space of equivalence classes of H -valued $\mathcal{F} \times \mathcal{B}([0, 1])$ -measurable processes $x(\cdot)$ admitting a predictable version such that $\mathbf{E} \int_0^1 \|x(t)\|_H^p dt < \infty$.
- $L_{\mathbb{F}}^p(\Omega, C([0, 1], H))$: the space of H -valued adapted processes $x(\cdot)$ with continuous paths such that $\mathbf{E} \sup_{t \in [0, 1]} \|x(t)\|_H^p < \infty$. Elements of this space are defined up to indistinguishability.
- $L_{\mathbb{F}}^p([0, 1], L^p(\Omega, H))$: the space of equivalence classes of H -valued adapted processes $x(\cdot)$ such that $x(\cdot) : [0, 1] \rightarrow L^p(\Omega, H)$ is $\mathcal{B}([0, 1])$ -measurable and $\int_0^1 \mathbf{E} \|x(t)\|_H^p dt < \infty$.
- $C_{\mathbb{F}}([0, 1], L^p(\Omega, H))$: the space of H -valued adapted processes $x(\cdot)$ such that $x(\cdot) : [0, 1] \rightarrow L^p(\Omega, H)$ is strongly continuous and $\sup_{t \in [0, 1]} \mathbf{E} \|x(t)\|_H^p < \infty$. Elements of this space are defined up to modification.

Moreover, denote by $\mathfrak{B}(H)$ the Banach space of all bounded linear operators from H to itself endowed with the norm $\|T\|_{\mathfrak{B}(H)} = \sup\{\|Tx\|_H : \|x\|_H = 1\}$. We shall define the following spaces with respect to $\mathfrak{B}(H)$ -valued processes and random variables.

- $L_{\mathbb{W}}^p(\Omega \times [0, 1], \mathcal{P}, \mathfrak{B}(H))$: the space of equivalence classes of $\mathfrak{B}(H)$ -valued processes $T(\cdot)$ such that $\langle x, T(\cdot)y \rangle \in L^p(\Omega \times [0, 1], \mathcal{P}, \mathbb{R})$ for any $x, y \in H$, and

¹This restriction is just for simplicity. The approaches and results in this paper can be extended without any nontrivial difficulty to the system driving by a multi-dimensional Wiener process.

$\mathbf{E} \int_0^1 \|T(t)\|_{\mathfrak{B}(H)}^p dt < \infty$. Here the subscript “w” stands for “weakly measurable”.

- $C_{\mathbb{F}}([0, 1], L_w^p(\Omega, \mathfrak{B}(H)))$: the space of $\mathfrak{B}(H)$ -valued processes $T(\cdot)$ such that $\langle x, T(\cdot)y \rangle \in C_{\mathbb{F}}([0, 1], L^p(\Omega, \mathbb{R}))$ for any $x, y \in H$, and $\sup_{t \in [0, 1]} \mathbf{E} \|T(t)\|_{\mathfrak{B}(H)}^p < \infty$. Elements of this space are defined up to modification.

- $L_w^p(\Omega, \mathcal{F}_t, \mathfrak{B}(H))$: the space of equivalence classes of $\mathfrak{B}(H)$ -valued random variable T such that $\langle x, Ty \rangle \in L^p(\Omega, \mathcal{F}_t, \mathbb{R})$ for any $x, y \in H$, and $\mathbf{E} \|T\|_{\mathfrak{B}(H)}^p < \infty$.

2. Formulation and main results.

2.1. Problem formulation. Let H and V be two separable real Hilbert spaces such that V is densely embedded in H . We identify H with its dual space, and denote by V^* the dual of V . Then we have $V \subset H \subset V^*$. Denote by $\|\cdot\|_H$ the norms of H , by $\langle \cdot, \cdot \rangle$ the inner product in H , and by $\langle \cdot, \cdot \rangle_*$ the duality product between V and V^* . The notation $\mathfrak{B}(X; Y)$ stands for the usual Banach space of all bounded linear operators from Banach space X to Banach space Y , and simply $\mathfrak{B}(X) = \mathfrak{B}(X; X)$.

Now we consider the controlled stochastic evolution system (1.1) in an abstract way:

$$\begin{cases} dx(t) = [A(t)x(t) + f(t, x(t), u(t))]dt + [B(t)x(t) + g(t, x(t), u(t))]dW_t, \\ x(0) = x_0, \end{cases}$$

with the control process $u(\cdot)$ taking values in a set U , given stochastic evolution operators

$$A : \Omega \times [0, 1] \rightarrow \mathfrak{B}(V; V^*) \quad \text{and} \quad B : \Omega \times [0, 1] \rightarrow \mathfrak{B}(V; H),$$

and nonlinear terms

$$f, g : \Omega \times [0, 1] \times H \times U \rightarrow H.$$

Here the *control set* U is a nonempty Borel-measurable subset of a metric space whose metric is denoted by $\text{dist}(\cdot, \cdot)$. Fix an element (denoted by 0) in U , and then define $|u|_U = \text{dist}(u, 0)$. A U -valued predictable process $u(\cdot)$ is *admissible* if there exists a number $\delta_u > 0$ such that

$$\sup \{ \mathbf{E} |u(t)|_U^{4+\delta_u} : t \in [0, 1] \} < \infty.$$

Denote by U_{ad} the set of all admissible controls.

Our optimal control problem is to find $u(\cdot) \in U_{\text{ad}}$ to minimize the cost functional

$$J(u(\cdot)) = \mathbf{E} \int_0^1 l(t, x(t), u(t)) dt + \mathbf{E} h(x(1))$$

with given functions

$$l : \Omega \times [0, 1] \times H \times U \rightarrow \mathbb{R} \quad \text{and} \quad h : \Omega \times H \rightarrow \mathbb{R}.$$

Throughout this paper, the SEE of form (1.1) is read in a weak sense, i.e., for each $v \in V$,

$$\begin{aligned} \langle v, x(t) \rangle &= \langle v, x_0 \rangle + \int_0^t [\langle v, A(t)x(t) \rangle_* + \langle v, f(t, x(t), u(t)) \rangle] dt \\ &\quad + \int_0^t \langle v, B(t)x(t) + g(t, x(t), u(t)) \rangle dW_t, \quad \text{a.e. } (\omega, t). \end{aligned}$$

We make the following assumptions. Fix some constants $\kappa \in (0, 1)$ and $K \in (0, \infty)$.

ASSUMPTION 2.1. *The operator processes A and B are weakly predictable, i.e., $\langle x, A(\cdot)y \rangle_*$ and $\langle x, B(\cdot)y \rangle$ are both predictable processes for any $x, y \in V$; and for each $(t, \omega) \in [0, 1] \times \Omega$,*

$$\begin{aligned} \langle x, A(t)x \rangle_* + \|B(t)x\|_H^2 &\leq -\kappa \|x\|_V^2 + K \|x\|_H^2, \\ \|A(t)x\|_{V^*}^2 &\leq K \|x\|_V^2, \quad \forall x \in V. \end{aligned}$$

ASSUMPTION 2.2. *For each $(x, u) \in H \times U$, $f(\cdot, x, u)$, $g(\cdot, x, u)$ and $l(\cdot, x, u)$ are all predictable processes, $h(x)$ is \mathcal{F}_1 -measurable random variable; for each $(t, u, \omega) \in [0, 1] \times U \times \Omega$, f, g, l and h are globally twice Fréchet differentiable with respect to x . The functions $f_x, g_x, f_{xx}, g_{xx}, l_{xx}, h_{xx}$ are continuous in x and dominated by the constant K ; f, g, l_x, h_x are dominated by $K(1 + \|x\|_H + |u|_U)$; l, h are dominated by $K(1 + \|x\|_H^2 + |u|_U^2)$.*

ASSUMPTION 2.3. *For each (t, ω) ,*

$$|\langle x, B(t)x \rangle| \leq K \|x\|_H^2, \quad \forall x \in V.$$

Here we call this condition the quasi-skew-symmetry.

REMARK 2.4. (1) Assumption 2.1 is a kind of *coercivity condition* (cf. [9]), which ensures the solvability of SEEs of form (1.1). Indeed, in view of a well-known result ([9]), SEE (1.1) has a unique solution $x(\cdot) \in L_{\mathbb{F}}^2(\Omega; C([0, 1]; H)) \cap L^2(\Omega \times [0, 1], \mathcal{P}, V)$ for each $u(\cdot) \in U_{\text{ad}}$ under Assumptions 2.1 and 2.2.

(2) In this paper, the *quasi-skew-symmetric condition* is used to establish the L^p -estimate ($p > 2$) for solutions to stochastic evolution equations. Such a condition is refined from many concrete examples, for instance, the nonlinear filtering equation, and other stochastic parabolic PDEs (cf. [2, 13]). One can give other characterizations of this condition. Indeed, for any given $B \in \mathfrak{B}(V; H)$, the following statements are equivalent: (i) for any $x \in V$, $|\langle x, Bx \rangle| \leq K \|x\|_H^2$; (ii) $B + B^* \in \mathfrak{B}(H)$, where B^* is the dual operator of B ; and (iii) there are a skew-symmetric operator $S \in \mathfrak{B}(V; H)$ and a symmetric operator $T \in \mathfrak{B}(H)$ such that $B = S + T$.

2.2. Main results. The main theorem in the paper is Theorem 2.6 — the maximum principle. As a preliminary, we first state the following result. Hereafter we denote $\mathbf{E}_t[\cdot] = \mathbf{E}[\cdot | \mathcal{F}_t]$.

THEOREM 2.5. *Let A and B satisfy Assumptions 2.1 and 2.3, $M \in L_{\mathbb{w}}^2(\Omega, \mathcal{F}_1, \mathfrak{B}(H))$ and $N(\cdot) \in L_{\mathbb{w}}^2(\Omega \times [0, 1], \mathcal{P}, \mathfrak{B}(H))$. In this case, we say the four-tuple $(A, B; M, N)$ is “appropriate”. We formally define, for each $\tau \in [0, 1)$, a stochastic bilinear function on the Banach space $L^4(\Omega, \mathcal{F}_\tau, H)$, associated with the four-tuple $(A, B; M, N)$, as the form:*

$$(2.1) \quad [T_\tau(A, B; M, N)](\xi_1, \xi_2) := \mathbf{E}_\tau \langle z^{\tau, \xi_1}(1), M z^{\tau, \xi_2}(1) \rangle + \mathbf{E}_\tau \int_\tau^1 \langle z^{\tau, \xi_1}(t), N(t) z^{\tau, \xi_2}(t) \rangle dt$$

with the processes $z^{\tau, \xi_i}(\cdot)$ ($i = 1, 2$) satisfying

$$(2.2) \quad z^{\tau, \xi_i}(t) = \xi_i + \int_\tau^t A(s) z^{\tau, \xi_i}(s) ds + \int_\tau^t B(s) z^{\tau, \xi_i}(s) dW_s, \quad t \in [\tau, 1].$$

Under the above setting, we have the following assertions:

a) For each $\tau \in [0, 1]$ and any $\xi, \zeta \in L^4(\Omega, \mathcal{F}_\tau, H)$, $[T_\tau(A, B; M, N)](\xi, \zeta)$ is uniquely determined and belongs to $L^1(\Omega, \mathcal{F}_\tau, \mathbb{R})$.

b) There exists a unique

$$P \in L^2_{\mathbb{W}}(\Omega \times [0, 1], \mathcal{P}, \mathfrak{B}(H)) \cap C_{\mathbb{F}}([0, 1], L^2_{\mathbb{W}}(\Omega, \mathfrak{B}(H)))$$

such that for each $\tau \in [0, 1]$ and any $\xi, \zeta \in L^4(\Omega, \mathcal{F}_\tau, H)$,

$$(2.3) \quad \langle \xi, P_\tau \zeta \rangle = [T_\tau(A, B; M, N)](\xi, \zeta) \quad (\text{a.s.}).$$

We call P the representation of $T.(A, B; M, N)$.

The proof of this theorem is placed in Subsection 4.1.

Our main result is the following

THEOREM 2.6 (Maximum Principle). *Let Assumptions 2.1–2.3 be satisfied. Define the Hamiltonian $\mathcal{H} : [0, 1] \times H \times U \times H \times H \rightarrow \mathbb{R}$ as the form*

$$(2.4) \quad \mathcal{H}(t, x, u, p, q) := l(t, x, u) + \langle p, f(t, x, u) \rangle + \langle q, g(t, x, u) \rangle.$$

Suppose $\bar{x}(\cdot)$ is the state process with respect to an optimal control $\bar{u}(\cdot)$. Then

i) (first-order adjoint process) the backward stochastic evolution equation (BSEE)

$$(2.5) \quad \begin{cases} dp(t) = -[A^*(t)p(t) + B^*(t)q(t) + \mathcal{H}_x(t, \bar{x}(t), \bar{u}(t), p(t), q(t))] dt + q(t) dW_t, \\ p(1) = h_x(\bar{x}(1)) \end{cases}$$

has a unique (weak) solution² $(p(\cdot), q(\cdot))$;

ii) (second-order adjoint process) the four-tuple $(\tilde{A}, \tilde{B}; \tilde{M}, \tilde{N})$ with

$$\begin{aligned} \tilde{A}(t) &:= A(t) + f_x(t, \bar{x}(t), \bar{u}(t)), & \tilde{B}(t) &:= B(t) + g_x(t, \bar{x}(t), \bar{u}(t)), \\ \tilde{M} &:= h_{xx}(\bar{x}(1)), & \tilde{N}(t) &:= \mathcal{H}_{xx}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)) \end{aligned}$$

is “appropriate”; then by Theorem 2.5 there exists a unique weakly predictable operator-valued process P as the representation of $T.(\tilde{A}, \tilde{B}; \tilde{M}, \tilde{N})$;

iii) (maximum condition) for each $u \in U$, the inequality

$$(2.6) \quad \begin{aligned} &\mathcal{H}(t, \bar{x}(t), u, p(t), q(t)) - \mathcal{H}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)) \\ &+ \frac{1}{2} \langle g(t, \bar{x}(t), u) - g(t, \bar{x}(t), \bar{u}(t)), P_t [g(t, \bar{x}(t), u) - g(t, \bar{x}(t), \bar{u}(t))] \rangle \geq 0 \end{aligned}$$

holds for a.e. $(t, \omega) \in [0, 1] \times \Omega$.

The proof of this theorem will be completed in Section 5.

REMARK 2.7. i) The inequality (2.6) holds almost surely on the product space $[0, 1] \times \Omega$, while the predictable version of process P insures the (t, ω) -joint measurability of the left-hand side of this inequality.

ii) Let us single out a couple of important special cases: 1) The diffusion does not contain the control variable, i.e., $g(t, x, u) \equiv g(t, x)$. In this case, (2.6) becomes

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)) = \min_{u \in U} \mathcal{H}(t, \bar{x}(t), u, p(t), q(t)).$$

This is a well known result, cf. [8]. 2) The control domain (a subset of a separable Hilbert space U) is convex and all the coefficients are C^1 in u . Then from (2.6) we can deduce

$$\langle \mathcal{H}_u(t, \bar{x}(t), \bar{u}(t), p(t), q(t)), u - \bar{u}(t) \rangle_U \geq 0, \quad \forall u \in U, \text{ a.e. } (t, \omega).$$

²For the definition of (weak) solutions to BSEEs, we refer to [4, Def. 3.1]

This is called a local form of the maximum principle, coinciding with the result of Bensoussan [1]. In both cases, Assumption 2.2 can be weakened, i.e., only the first Gâteaux differentiability of the coefficients is in force.

3. L^p -estimates for stochastic evolution equations. The L^p -estimate of the solutions to stochastic evolution equations plays a basic role in our approach. Now we consider the following linear equation

$$(3.1) \quad \begin{cases} dy(t) = [\tilde{A}(t)y(t) + a(t)] dt + [\tilde{B}(t)y(t) + b(t)]dW_t, \\ y(0) = y_0 \in H. \end{cases}$$

Under Assumption 2.1, the above equation has a unique solution $y(\cdot) \in L^2_{\mathbb{F}}(\Omega, C([0, 1], H))$ providing $a(\cdot), b(\cdot) \in L^2(\Omega \times [0, 1], \mathcal{P}, H)$, see [9]. If, in addition, the operator \tilde{B} satisfies *the quasi-skew-symmetric condition*, then we have

LEMMA 3.1. *Let \tilde{A} and \tilde{B} satisfy Assumptions 2.1 and 2.3, and $p \geq 1$. Then the solution to equation (3.1) satisfies*

$$\mathbf{E} \sup_{t \in [0, 1]} \|y(t)\|_H^{2p} \leq C(\kappa, K, p) \mathbf{E} \left[\|y_0\|_H^{2p} + \left(\int_0^1 \|a(t)\|_H dt \right)^{2p} + \left(\int_0^1 \|b(t)\|_H^2 dt \right)^p \right],$$

provided the right-hand side is finite.

Proof. In view of Assumptions 2.1 and 2.3, we have

$$(3.2) \quad \|\tilde{B}y(t)\|_H^2 \leq C(K)\|y(t)\|_V^2, \quad \langle y(t), \tilde{B}y(t) \rangle \leq K\|y(t)\|_H^2,$$

and furthermore,

$$(3.3) \quad \begin{aligned} & 2\langle y(t), \tilde{A}y(t) + a(t) \rangle_* + \|\tilde{B}y(t) + b(t)\|_H^2 \\ &= 2\langle y(t), \tilde{A}y(t) \rangle_* + \|\tilde{B}y(t)\|_H^2 + 2\langle \tilde{B}y(t), b(t) \rangle + \|b(t)\|_H^2 + 2\langle y(t), a(t) \rangle \\ &\leq -\kappa\|y(t)\|_V^2 + K\|y(t)\|_H^2 + C(K)\|y(t)\|_V\|b(t)\|_H + \|b(t)\|_H^2 + 2\langle y(t), a(t) \rangle \\ &\leq K\|y(t)\|_H^2 + C(\kappa, K)\|b(t)\|_H^2 + 2\|y(t)\|_H\|a(t)\|_H, \end{aligned}$$

$$(3.4) \quad \begin{aligned} & |\langle y(t), \tilde{B}y(t) + b(t) \rangle|^2 \leq 2|\langle y(t), \tilde{B}y(t) \rangle|^2 + 2|\langle y(t), b(t) \rangle|^2 \\ &\leq 2K^2\|y(t)\|_H^4 + 2\|y(t)\|_H^2\|b(t)\|_H^2. \end{aligned}$$

Define stopping time $\tau_k := \inf\{t \in [0, 1] : \|y(t)\|_H^{2p} > k\} \wedge 1$. With $\lambda, \epsilon > 0$, it follows from the Burkholder-Davis-Gundy and Hölder inequalities that

(3.5)

$$\begin{aligned} & \mathbf{E} \sup_{t \in [0, \tau_k]} \left| \int_0^t e^{-\lambda s} \|y(s)\|_H^{2(p-1)} \langle y(s), \tilde{B}y(s) + b(s) \rangle dW_s \right| \\ & \leq C \mathbf{E} \left[\int_0^{\tau_k} e^{-2\lambda s} \|y(s)\|_H^{4(p-1)} |\langle y(s), \tilde{B}y(s) + b(s) \rangle|^2 ds \right]^{1/2} \\ & \leq C \mathbf{E} \left[\left(\sup_{s \in [0, \tau_k]} e^{-\lambda s} \|y(s)\|_H^{2p} \int_0^{\tau_k} e^{-\lambda s} \left(\|y(s)\|_H^{2p} + \|y(s)\|_H^{2p-2} \|b(s)\|_H^2 \right) ds \right)^{1/2} \right] \\ & \leq \epsilon \mathbf{E} \sup_{s \in [0, \tau_k]} e^{-\lambda s} \|y(s)\|_H^{2p} + \frac{C}{\epsilon} \mathbf{E} \int_0^{\tau_k} e^{-\lambda s} \left(\|y(s)\|_H^{2p} + \|y(s)\|_H^{2p-2} \|b(s)\|_H^2 \right) ds \end{aligned}$$

(3.6)

$$\begin{aligned} & \mathbf{E} \int_0^{\tau_k} e^{-\lambda s} \left(\|y(s)\|_H^{2p-1} \|a(s)\|_H + \|y(s)\|_H^{2p-2} \|b(s)\|_H^2 \right) ds \\ & \leq \epsilon \mathbf{E} \sup_{s \in [0, \tau_k]} e^{-\lambda s} \|y(s)\|_H^{2p} + \frac{C}{\epsilon} \mathbf{E} \left[\left(\int_0^{\tau_k} e^{-\frac{\lambda s}{2p}} \|a(s)\|_H ds \right)^{2p} + \left(\int_0^{\tau_k} e^{-\frac{\lambda s}{p}} \|b(s)\|_H^2 ds \right)^p \right] \\ & \leq \epsilon \mathbf{E} \sup_{s \in [0, \tau_k]} e^{-\lambda s} \|y(s)\|_H^{2p} + \frac{C}{\epsilon} \mathbf{E} \left[\left(\int_0^1 \|a(s)\|_H ds \right)^{2p} + \left(\int_0^1 \|b(s)\|_H^2 ds \right)^p \right]. \end{aligned}$$

On the other hand, applying the Itô formula (see [9]) to $e^{-\lambda t} \|y(t)\|_H^{2p}$, we have

$$\begin{aligned} (3.7) \quad & e^{-\lambda(t \wedge \tau_k)} \|y(t \wedge \tau_k)\|_H^{2p} - \|y_0\|_H^{2p} + \lambda \int_0^{t \wedge \tau_k} e^{-\lambda s} \|y(s)\|_H^{2p} ds \\ & = p \int_0^{t \wedge \tau_k} e^{-\lambda s} \|y(s)\|_H^{2(p-1)} (2 \langle y(s), \tilde{A}y(s) + a(s) \rangle_* + \|\tilde{B}y(s) + b(s)\|_H^2) ds \\ & \quad + 2p(p-1) \int_0^{t \wedge \tau_k} e^{-\lambda s} \|y(s)\|_H^{2(p-2)} |\langle y(s), \tilde{B}y(s) + b(s) \rangle|^2 ds \\ & \quad + 2p \int_0^{t \wedge \tau_k} e^{-\lambda s} \|y(s)\|_H^{2(p-1)} \langle y(s), \tilde{B}y(s) + b(s) \rangle dW_s \end{aligned}$$

Now we take $\mathbf{E} \sup_{t \in [0, \tau_k]}$ on both sides of the above equality. Then by virtue of (3.3)–(3.7) with sufficiently large λ and small ϵ , we obtain

$$\mathbf{E} \sup_{t \in [0, \tau_k]} e^{-\lambda t} \|y(t)\|_H^{2p} \leq C(\kappa, K, p) \mathbf{E} \left[\|y_0\|_H^{2p} + \left(\int_0^1 \|a(t)\|_H dt \right)^{2p} + \left(\int_0^1 \|b(t)\|_H^2 dt \right)^p \right].$$

Sending k to infinity, we can easily conclude the lemma. \square

REMARK 3.2. The *quasi-skew-symmetric condition* is unnecessary in the case of $p = 1$, since the term $2p(p-1) \|y(t)\|_H^{2(p-2)} |\langle y(t), \tilde{B}y(t) + b(t) \rangle|^2 dt$ does not appear in this case.

Proceeding a similar argument, we have the L^p -estimate for SEE (1.1).

COROLLARY 3.3. *Under Assumptions 2.1–2.3, the solution $x(\cdot)$ to SEE (1.1) satisfies*

$$\mathbf{E} \sup_{t \in [0, 1]} \|x(t)\|_H^p \leq C(\kappa, K) \sup_{t \in [0, 1]} \mathbf{E}(1 + |u(t)|_U^p)$$

with $p \in [2, 4 + \delta]$.

4. Investigation into a stochastic bilinear function. The purpose of this section is to prove Theorem 2.5 and another important property of the representation process P .

4.1. Proof of Theorem 2.5. For convenience, we write $T_\tau(A, B; M, N)$ as T_τ to the end of this section, and hereafter denote $C = C(\kappa, K)$.

Step 1. In view of a known result (cf. [9]), equation (2.2) has a unique solution for each $\xi \in L^4(\Omega, \mathcal{F}_\tau, H)$; moreover, it follows from Lemma 3.1 that

$$(4.1) \quad \mathbf{E}_\tau \sup_{t \in [\tau, 1]} \|z^{\tau, \xi}(t)\|_H^4 \leq C \|\xi\|_H^4.$$

Indeed, for any set $E \in \mathcal{F}_\tau$, we obtain from Lemma 3.1 that

$$\mathbf{E} \left(\mathbf{1}_E \cdot \sup_{t \in [\tau, 1]} \|z^{\tau, \xi}(t)\|_H^4 \right) = \mathbf{E} \sup_{t \in [\tau, 1]} \|z^{\tau, \mathbf{1}_E \xi}(t)\|_H^4 \leq C \mathbf{E} \|\mathbf{1}_E \xi\|_H^4 = C \mathbf{E} (\mathbf{1}_E \cdot \|\xi\|_H^4),$$

which implies (4.1). In what follows, we define

$$(4.2) \quad \Lambda := \|M\|_{\mathfrak{B}(H)}^2 + \int_0^1 \|N(t)\|_{\mathfrak{B}(H)}^2 dt \in L^1(\Omega, \mathcal{F}_1, \mathbb{R}),$$

and $\Lambda_t = \mathbf{E}_t[\Lambda]$. The process $(\Lambda_t)_{t \geq 0}$ is a Doob's martingale and has a continuous version. Then we know, for each $\tau \in [0, 1]$ and any $\xi, \zeta \in L^4(\Omega, \mathcal{F}_\tau, H)$,

$$(4.3) \quad |T_\tau(\xi, \zeta)| \leq C \sqrt{\Lambda_\tau} \|\xi\|_H \|\zeta\|_H \quad (\text{a.s.}).$$

Therefore, for any $\xi, \zeta \in L^4(\Omega, \mathcal{F}_\tau, H)$, $T_\tau(\xi, \zeta)$ is uniquely determined and, by the Hölder inequality, belongs to $L^1(\Omega, \mathcal{F}_\tau, \mathbb{R})$. The assertion (a) is proved.

Step 2. Next we shall prove

$$(4.4) \quad T_\tau(x, y) \in C_{\mathbb{F}}([0, 1], L^2(\Omega, \mathbb{R})), \quad \forall x, y \in H.$$

For convenience, we denote

$$(4.5) \quad Y_t^{x, y} := \langle z^{t, x}(1), M z^{t, y}(1) \rangle + \int_t^1 \langle z^{t, x}(r), N(r) z^{t, y}(r) \rangle dr, \quad \text{for } x, y \in H,$$

with $t \in [0, 1)$, and then $T_t(x, y) = \mathbf{E}_t[Y_t^{x, y}]$. Now we have

$$\begin{aligned} \lim_{s \rightarrow t} \mathbf{E} |\mathbf{E}_s[Y_s^{x, y}] - \mathbf{E}_t[Y_t^{x, y}]| &= \lim_{s \rightarrow t} \mathbf{E} |\mathbf{E}_s[Y_s^{x, y} - Y_t^{x, y}] - (\mathbf{E}_t[Y_t^{x, y}] - \mathbf{E}_s[Y_t^{x, y}])| \\ &\leq \lim_{s \rightarrow t} \mathbf{E} |Y_s^{x, y} - Y_t^{x, y}| + \lim_{s \rightarrow t} \mathbf{E} |\mathbf{E}_t[Y_t^{x, y}] - \mathbf{E}_s[Y_t^{x, y}]|. \end{aligned}$$

Without loss of generality, we assume $t < s$. On the one hand, the process $(\mathbf{E}_r[Y_r^{x, y}])_{r \geq 0}$ is a uniformly integrable martingale, thus it follows from the Doob martingale convergence theorem (cf. [3]) that

$$\lim_{s \rightarrow t} \mathbf{E} |\mathbf{E}_t[Y_t^{x, y}] - \mathbf{E}_s[Y_t^{x, y}]| = 0.$$

On the other hand, note that

$$\begin{aligned}
|Y_s^{x,y} - Y_t^{x,y}| &\leq |\langle z^{s,x}(1), Mz^{s,y}(1) \rangle - \langle z^{t,x}(1), Mz^{t,y}(1) \rangle| \\
&\quad + \int_s^1 |\langle z^{s,x}(r), N(r)z^{s,y}(r) \rangle - \langle z^{t,x}(r), N(r)z^{t,y}(r) \rangle| dr \\
&\quad + \int_t^s |\langle z^{t,x}(r), N(r)z^{t,y}(r) \rangle| dr \\
&=: I_1 + I_2 + I_3.
\end{aligned}$$

First, it follows from (4.1) and the Hölder inequality that

$$(4.6) \quad |\mathbf{E}I_3|^2 \leq C\Lambda_0 |t-s| \|x\|_H^2 \|y\|_H^2 \rightarrow 0,$$

as $|s-t| \rightarrow 0$. Next, from (4.1) and the continuity of the solution to (2.2), we have

$$\begin{aligned}
|\mathbf{E}I_2|^2 &\leq C\Lambda_0 \left(\|x\|_H^2 \sqrt{\mathbf{E} \|y - z^{t,y}(s)\|_H^4} \right. \\
&\quad \left. + \|y\|_H^2 \sqrt{\mathbf{E} \|x - z^{t,x}(s)\|_H^4} \right) \rightarrow 0, \quad \text{as } |s-t| \rightarrow 0.
\end{aligned}$$

Similarly, we can show

$$|\mathbf{E}I_1|^2 \rightarrow 0, \quad \text{as } |s-t| \rightarrow 0.$$

Therefore, we have

$$(4.7) \quad \lim_{s \rightarrow t} \mathbf{E} |T_s(x,y) - T_t(x,y)| = 0.$$

This implies $T(x,y) \in C_{\mathbb{F}}([0,1], L^1(\Omega, \mathbb{R}))$. Next, it follows from (4.3) and the Doob martingale convergence theorem that

$$\begin{aligned}
|T_s(x,y) - T_t(x,y)|^2 &\leq C \|x\|_H^2 \|y\|_H^2 \cdot (\Lambda_s + \Lambda_t) \\
&\xrightarrow{L^1} C \|x\|_H^2 \|y\|_H^2 \cdot 2\Lambda_t \quad \text{as } s \rightarrow t.
\end{aligned}$$

This along with (4.7) and the Lebesgue dominated convergence theorem yields

$$(4.8) \quad \lim_{s \rightarrow t} \mathbf{E} |T_s(x,y) - T_t(x,y)|^2 = 0.$$

Thus we have $T(x,y) \in C_{\mathbb{F}}([0,1], L^2(\Omega, \mathbb{R}))$.

*Step 3.*³ Now we shall prove that, for any $x, y \in H$, there is a predictable modification of the process $T(x,y)$.

Recalling (4.5), let $Y_t^{x,y}(\omega) = Y_1^{x,y}(\omega)$ when $t > 1$. Then, for any $x, y \in H$, the mapping $(t, \omega) \mapsto Y_t^{x,y}(\omega)$ is $\mathcal{B}([0, \infty)) \times \mathcal{F}$ -measurable.

Let $(Y_t^{x,y})^+ = Y_t^{x,y} \vee 0$ and $(Y_t^{x,y})^- = (-Y_t^{x,y}) \vee 0$. In view of a well-known result (c.f. [7, Theorem 5.2]), there a unique predicable projection of $(Y_t^{x,y})^\pm$, denoted by $\mathbf{P}Y^{x,y,\pm}$, such that for every predictable time σ ,

$$\mathbf{E}[(Y_\sigma^{x,y})^\pm 1_{\{\sigma < \infty\}} | \mathcal{F}_{\sigma-}] = \mathbf{P}Y_\sigma^{x,y,\pm} 1_{\{\sigma < \infty\}} \quad (\text{a.s.}).$$

³The argument in this step is borrowed from [6].

With the continuity of filtration \mathbb{F} in mind, we have for every $t \in [0, 1]$,

$$\mathbf{P}Y_t^{x,y,\pm} = \mathbf{E}[(Y_t^{x,y})^\pm | \mathcal{F}_{t-}] = \mathbf{E}_t[(Y_t^{x,y})^\pm] \quad (\text{a.s.}).$$

Thus we obtain

$$\mathbf{P}Y_t^{x,y} := \mathbf{P}Y_t^{x,y,+} - \mathbf{P}Y_t^{x,y,-} = \mathbf{E}_t[Y_t^{x,y}] = T_t(x, y) \quad (\text{a.s.}),$$

which implies $\mathbf{P}Y_t^{x,y}$ is a predictable version of $T_t(x, y)$.

Step 4. The construction of process P .

Take a standard complete orthonormal basis $\{e_i\}$ in H , and a predictable version of $T_t(e_i, e_j)$ for each $i, j \in \mathbb{N}$. Set

$$\Gamma_{ij} = \{(t, \omega) \in [0, 1] \times \Omega : T_t(e_i, e_j)(\omega) \leq C\sqrt{\Lambda_t(\omega)}\},$$

where the constant C is taken from (4.3). Then Γ_{ij} is a predictable set with full measure, and so $\Gamma := \bigcap_{i,j \in \mathbb{N}} \Gamma_{ij}$ is also a predictable set with full measure; moreover, in view of (4.3), the section $\Gamma(t) := \{\omega : (t, \omega) \in \Gamma\}$ is a set of probability 1 for each $t \in [0, 1]$.

Thanks to the Riesz representation theorem, there is a unique $P_t(\omega) \in \mathfrak{B}(H)$ for each $(t, \omega) \in \Gamma$ such that

$$\langle e_i, P_t(\omega)e_j \rangle_H = T_t(e_i, e_j)(\omega)$$

and $\|P_t(\omega)\|_{\mathfrak{B}(H)} \leq C\sqrt{\Lambda_t(\omega)}$. Let $P_t(\omega) = 0$ for $(t, \omega) \in \Gamma^c$. Then $\langle e_i, P_t e_j \rangle$ belongs to $L^2(\Omega \times [0, 1], \mathcal{P}, \mathbb{R})$. Since for each (t, ω) and $x, y \in H$,

$$\langle x, P_t(\omega)y \rangle = \lim_{n \rightarrow \infty} \langle x_n, P_t(\omega)y_n \rangle$$

with $x_n = \sum_{i=1}^n \langle x, e_i \rangle e_i$ and $y_n = \sum_{i=1}^n \langle y, e_i \rangle e_i$, we have $\langle x, P_t y \rangle \in L^2(\Omega \times [0, 1], \mathcal{P}, \mathbb{R})$; moreover, from the bilinearity of $T_t(\cdot, \cdot)$, we know that $\langle x, P_t y \rangle = T_t(x, y)$ a.s. for each $t \in [0, 1]$ and any $x, y \in H$. Recalling (4.4), we have

$$P \in L^2_{\mathbb{w}}(\Omega \times [0, 1], \mathcal{P}, \mathfrak{B}(H)) \cap C_{\mathbb{F}}([0, 1], L^2_{\mathbb{w}}(\Omega, \mathfrak{B}(H))).$$

It remains to show the relation (2.3). Fix arbitrary $t \in [0, 1]$. It follows from the definition of $T_t(\cdot, \cdot)$ that (i) for any $E \in \mathcal{F}_t$,

$$T_t(x\mathbf{1}_E, y\mathbf{1}_E) = \mathbf{1}_E T_t(x, y) \quad (\text{a.s.}) \quad \forall x, y \in H;$$

and (ii) for any $E_1, E_2 \in \mathcal{F}_t$ with $E_1 \cap E_2 = \emptyset$,

$$T_t(x\mathbf{1}_{E_1}, y\mathbf{1}_{E_2}) = 0 \quad (\text{a.s.}) \quad \forall x, y \in H.$$

This means that, for any simple H -valued \mathcal{F}_t -measurable random variables ξ and ζ , we have

$$\langle \xi, P_t \zeta \rangle = T_t(\xi, \zeta) \quad (\text{a.s.}),$$

which along with a standard argument of approximation yields the relation (2.3).

The uniqueness of the representation is obvious. The proof of Theorem 2.5 is complete.

4.2. An important property of P . Give the same four-tuple $(A, B; M, N)$ as in Theorem 2.5. Let $\varepsilon \in (0, 1 - \tau)$. For each $\xi \in L^4(\Omega, \mathcal{F}_\tau, H)$, consider the equation

$$(4.9) \quad z_\varepsilon^{\tau, \xi}(t) = \int_\tau^t A z_\varepsilon^{\tau, \xi}(s) ds + \int_\tau^t [B z_\varepsilon^{\tau, \xi}(s) + \varepsilon^{-\frac{1}{2}} \mathbf{1}_{[\tau, \tau + \varepsilon)} \xi] dW_s, \quad t \in [\tau, 1]$$

and the following bilinear functional on $L^4(\Omega, \mathcal{F}_\tau, H)$ with parameter ε :

$$[T_\tau^\varepsilon(A, B; M, N)](\xi, \zeta) := \mathbf{E} \langle z_\varepsilon^{\tau, \xi}(1), M z_\varepsilon^{\tau, \zeta}(1) \rangle + \mathbf{E} \int_\tau^1 \langle z_\varepsilon^{\tau, \xi}(t), N(t) z_\varepsilon^{\tau, \zeta}(t) \rangle dt.$$

In view of Lemma 3.1, equation (4.9) has a unique solution $z_\varepsilon^{\tau, \xi}(\cdot)$ such that

$$\mathbf{E} \sup_{t \in [0, 1]} \|z_\varepsilon^{\tau, \xi}(t)\|_H^4 \leq C \mathbf{E} \|\xi\|_H^4.$$

Thus $T_\tau^\varepsilon = T_\tau^\varepsilon(A, B; M, N)$ is well-defined.

Next, we shall prove a result concerning the relation between T_τ^ε and P_τ , which plays a key role in the proof of the maximum principle.

PROPOSITION 4.1. *Under the above setting, we have*

$$(4.10) \quad \mathbf{E} \langle \xi, P_\tau \zeta \rangle = \lim_{\varepsilon \rightarrow 0} [T_\tau^\varepsilon(A, B; M, N)](\xi, \zeta)$$

for each $\tau \in [0, 1)$ and any $\xi, \zeta \in L^4(\Omega, \mathcal{F}_\tau, H)$.

Proof. First of all, we claim: the assertion holds true if it does in a dense subset D_τ of $L^4(\Omega, \mathcal{F}_\tau, H)$, i.e., the relation (4.10) holds for any $\xi, \zeta \in D_\tau$.

Indeed, from the density, for arbitrary $\eta > 0$, we can find $\xi^\eta, \zeta^\eta \in D_\tau$ such that

$$\mathbf{E} \|\xi - \xi^\eta\|_H^4 + \mathbf{E} \|\zeta - \zeta^\eta\|_H^4 < \eta^4.$$

Bearing in mind (4.3), one can show

$$|\mathbf{E} \langle \xi, P_\tau \zeta \rangle - \mathbf{E} \langle \xi^\eta, P_\tau \zeta^\eta \rangle| + |T_\tau^\varepsilon(\xi, \zeta) - T_\tau^\varepsilon(\xi^\eta, \zeta^\eta)| < C \eta \sqrt{\Lambda_0} [\mathbf{E}(\|\xi\|_H^4 + \|\zeta\|_H^4)]^{\frac{1}{4}}.$$

So, if the assertion holds in D_τ , i.e., $\mathbf{E} \langle \xi^\eta, P_\tau \zeta^\eta \rangle = \lim_{\varepsilon \rightarrow 0} T_\tau^\varepsilon(\xi^\eta, \zeta^\eta)$, then

$$\limsup_{\varepsilon \rightarrow 0} |\mathbf{E} \langle \xi, P_\tau \zeta \rangle - T_\tau^\varepsilon(\xi, \zeta)| \leq C \eta \sqrt{\Lambda_0} [\mathbf{E}(\|\xi\|_H^4 + \|\zeta\|_H^4)]^{\frac{1}{4}}.$$

From the arbitrariness of η , we prove the claim.

Now we define

$$D_\tau := \{\xi \in \mathcal{F}_\tau : \xi \text{ is a simple random variable with values in } V\}.$$

Obviously, D_τ is dense in $L^4(\Omega, \mathcal{F}_\tau, H)$.

The next result is the key-point in the proof, which gives a simple asymptotic alternative of $z_\varepsilon^{\tau, \xi}(\tau + \varepsilon)$, independent of the operators A and B .

LEMMA 4.2. *Define $\xi_t = \varepsilon^{-\frac{1}{2}}(W_t - W_\tau)\xi$ for $\xi \in D_\tau$. Then*

$$\mathbf{E} \|z_\varepsilon^{\tau, \xi}(\tau + \varepsilon) - \xi_{\tau + \varepsilon}\|_H^4 \leq C \varepsilon^2 \mathbf{E} \|\xi\|_V^4.$$

Proof. Denoting $y_t = z_\varepsilon^{\tau, \xi}(t) - \xi_t$, we can write down the equation

$$dy_t = (Ay_t + A\xi_t) dt + (By_t + B\xi_t) dW_t, \quad y_\tau = 0$$

with $t \in [\tau, \tau + \varepsilon]$. Inspired by the proof of Lemma 3.1, we can deduce

$$\begin{aligned} \mathbf{E} \|y_t\|_H^4 &\leq 2 \mathbf{E} \int_\tau^t \|y_s\|_H^2 \left[-\kappa \|y_s\|_V^2 + C(K) \|y_s\|_H^2 + \langle y_s, A\xi_s \rangle_* + \|B\xi_s\|_H^2 \right] ds \\ &\leq C(\kappa, K) \mathbf{E} \int_\tau^t \|y_s\|_H^2 \left[\|y_s\|_H^2 + \|A\xi_s\|_{V^*}^2 + \|B\xi_s\|_H^2 \right] ds \end{aligned}$$

with $t \in [\tau, \tau + \varepsilon]$. Note that $\|A\xi_s\|_{V^*} + \|B\xi_s\|_H \leq C(K)\|\xi_s\|_V$. Then by means of the Fubini theorem, the Gronwall and Young inequalities, we have

$$\begin{aligned} \sup_{t \in [\tau, \tau + \varepsilon]} \mathbf{E} \|y_t\|_H^4 &\leq C \left(\int_\tau^{\tau + \varepsilon} \mathbf{E} (\|A\xi_s\|_{V^*}^2 + \|B\xi_s\|_H^2) ds \right)^2 \\ &\leq C \left(\int_\tau^{\tau + \varepsilon} \mathbf{E} \|\xi_s\|_V^2 ds \right)^2 = C \left(\frac{\varepsilon}{2} \mathbf{E} \|\xi\|_V^2 \right)^2. \end{aligned}$$

This concludes the lemma. \square

Let us move on the proof of Proposition 4.1. For $\xi, \zeta \in D_\tau$, define

$$\xi_t := \varepsilon^{-\frac{1}{2}}(W_t - W_\tau)\xi \quad \text{and} \quad \zeta_t := \varepsilon^{-\frac{1}{2}}(W_t - W_\tau)\zeta.$$

Notice the fact that for any $\xi, \zeta \in D_\tau$,

$$\begin{aligned} T_\tau^\varepsilon(\xi, \zeta) &= \mathbf{E} \int_\tau^{\tau + \varepsilon} \langle z_\varepsilon^{\tau, \xi}(t), N(t)z_\varepsilon^{\tau, \zeta}(t) \rangle dt + \mathbf{E} \langle z_\varepsilon^{\tau, \xi}(\tau + \varepsilon), P_{\tau + \varepsilon} z_\varepsilon^{\tau, \zeta}(\tau + \varepsilon) \rangle \\ &=: I_1 + I_2. \end{aligned}$$

Now we let ε tend to 0. First, one can show, just like (4.6), the term I_1 tends to 0; Next, by means of Lemma 4.2 and the relation (4.3), we have

$$\begin{aligned} \mathbf{E} |\langle \xi_{\tau + \varepsilon}, P_{\tau + \varepsilon} \zeta_{\tau + \varepsilon} \rangle - I_2| &\leq C \sqrt{\mathbf{E} \Lambda_1} (\mathbf{E} \|\xi\|_H^4)^{\frac{1}{4}} (\mathbf{E} \|z_\varepsilon^{\tau, \zeta}(\tau + \varepsilon) - \zeta_{\tau + \varepsilon}\|_H^4)^{\frac{1}{4}} \\ &\quad + C \sqrt{\mathbf{E} \Lambda_1} (\mathbf{E} \|\zeta\|_H^4)^{\frac{1}{4}} (\mathbf{E} \|z_\varepsilon^{\tau, \xi}(\tau + \varepsilon) - \xi_{\tau + \varepsilon}\|_H^4)^{\frac{1}{4}} \\ &\rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0. \end{aligned}$$

Thus we obtain

$$(4.11) \quad \lim_{\varepsilon \rightarrow 0} |\mathbf{E} \langle \xi_{\tau + \varepsilon}, P_{\tau + \varepsilon} \zeta_{\tau + \varepsilon} \rangle - T_\tau^\varepsilon(\xi, \zeta)| = 0, \quad \forall \xi, \zeta \in D_\tau.$$

On the other hand, for $\xi, \zeta \in D_\tau$, we deduce

$$\begin{aligned} |\mathbf{E} \langle \xi_{\tau + \varepsilon}, (P_{\tau + \varepsilon} - P_\tau) \zeta_{\tau + \varepsilon} \rangle|^2 &= |\mathbf{E} [\varepsilon^{-1} |W_{\tau + \varepsilon} - W_\tau|^2 \langle \xi, (P_{\tau + \varepsilon} - P_\tau) \zeta \rangle]|^2 \\ &\leq [\mathbf{E} (\varepsilon^{-2} |W_{\tau + \varepsilon} - W_\tau|^4)] \cdot [\mathbf{E} |\langle \xi, (P_{\tau + \varepsilon} - P_\tau) \zeta \rangle|^2] \leq 3 \mathbf{E} |\langle \xi, (P_{\tau + \varepsilon} - P_\tau) \zeta \rangle|^2; \end{aligned}$$

since ξ and ζ are simple random variables, recalling (4.8), we know that the last term in the above relation tends to 0 when $\varepsilon \rightarrow 0$. Note that $\mathbf{E} \langle \xi_{\tau + \varepsilon}, P_\tau \zeta_{\tau + \varepsilon} \rangle = \mathbf{E} \langle \xi, P_\tau \zeta \rangle$. Hence, we get

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \langle \xi_{\tau + \varepsilon}, P_{\tau + \varepsilon} \zeta_{\tau + \varepsilon} \rangle = \mathbf{E} \langle \xi, P_\tau \zeta \rangle, \quad \forall \xi, \zeta \in D_\tau.$$

This along with (4.11) yields

$$\lim_{\varepsilon \rightarrow 0} T_\tau^\varepsilon(\xi, \zeta) = \mathbf{E} \langle \xi, P_\tau \zeta \rangle, \quad \forall \xi, \zeta \in D_\tau,$$

which completes the proof of Proposition 4.1. \square

5. Proof of Theorem 2.6. In this section, we shall prove the maximum principle.

5.1. Second-order expansion of spike variation. Assume $\bar{x}(\cdot)$ is the state process with respect to an optimal control $\bar{u}(\cdot)$. We fix a $\tau \in [0, 1)$ in this subsection.

Following a classical technique in the optimal control problem, we construct a perturbed admissible control in the following way (named *spike variation*)

$$u^\varepsilon(t) := \begin{cases} u(t), & \text{if } t \in [\tau, \tau + \varepsilon], \\ \bar{u}(t), & \text{otherwise,} \end{cases}$$

with fixed $\tau \in [0, 1)$, sufficiently small positive ε , and any given admissible control $u(\cdot)$.

Let $x^\varepsilon(\cdot)$ be the state process with respect to the control $u^\varepsilon(\cdot)$. For the sake of convenience, we denote for $\varphi = f, g, l, f_x, g_x, l_x, f_{xx}, g_{xx}, l_{xx}$,

$$\begin{aligned} \bar{\varphi}(t) &:= \varphi(t, \bar{x}(t), \bar{u}(t)), \\ \varphi^\Delta(t) &:= \varphi(t, \bar{x}(t), u(t)) - \bar{\varphi}(t), \\ \varphi^{\Delta, \varepsilon}(t) &:= \varphi^\Delta(t) \cdot \mathbf{1}_{[\tau, \tau + \varepsilon]}(t), \\ \tilde{\varphi}_{xx}^\varepsilon(t) &:= 2 \int_0^1 \lambda \varphi_{xx}(t, \lambda \bar{x}(t) + (1 - \lambda)x^\varepsilon(t), u^\varepsilon(t)) d\lambda. \end{aligned}$$

By means of the basic L^p estimate, we have

LEMMA 5.1. *Under Assumptions 2.1–2.3, we have*

$$\mathbf{E} \sup_{t \in [0, 1]} \|\Xi^\varepsilon(t)\|_H^2 := \mathbf{E} \sup_{t \in [0, 1]} \|x^\varepsilon(t) - \bar{x}(t) - x_1^\varepsilon(t) - x_2^\varepsilon(t)\|_H^2 = o(\varepsilon^2),$$

where $x_1^\varepsilon(\cdot)$ and $x_2^\varepsilon(\cdot)$ are the solutions respectively to

(5.1)

$$\begin{aligned} x_1^\varepsilon(t) &= \int_0^t [A(s)x_1^\varepsilon(s) + \bar{f}_x(s)x_1^\varepsilon(s)] ds \\ &+ \int_0^t [B(s)x_1^\varepsilon(s) + \bar{g}_x(s)x_1^\varepsilon(s) + g^{\Delta, \varepsilon}(s)] dW_s, \end{aligned}$$

(5.2)

$$\begin{aligned} x_2^\varepsilon(t) &= \int_0^t [A(s)x_2^\varepsilon(s) + \bar{f}_x(s)x_2^\varepsilon(s) + \frac{1}{2}\bar{f}_{xx}(s)(x_1^\varepsilon \otimes x_1^\varepsilon)(s) + f^{\Delta, \varepsilon}(s)] ds \\ &+ \int_0^t [B(s)x_2^\varepsilon(s) + \bar{g}_x(s)x_2^\varepsilon(s) + \frac{1}{2}\bar{g}_{xx}(s)(x_1^\varepsilon \otimes x_1^\varepsilon)(s) + g_x^{\Delta, \varepsilon}(s)x_1^\varepsilon(s)] dW_s. \end{aligned}$$

Proof. The proof is rather standard (cf. [16]). The L^p -estimate for SEEs plays a key role here. Indeed, by means of Lemma 3.1, and keeping in mind Assumption 2.2 and Corollary 3.3, we deduce that

$$\begin{aligned} (5.3) \quad \mathbf{E} \sup_{t \in [0, 1]} \|x_1^\varepsilon(t)\|_H^{2p} &\leq C \mathbf{E} \left[\int_\tau^{\tau + \varepsilon} \|g^\Delta(t)\|_H^2 \cdot \mathbf{1}_{[\tau, \tau + \varepsilon]}(t) dt \right]^p \\ &\leq C \varepsilon^{p-1} \mathbf{E} \int_\tau^{\tau + \varepsilon} \|g^\Delta(t)\|_H^{2p} dt \leq C \varepsilon^p \sup_{t \in [0, 1]} \mathbf{E} \|g^\Delta(t)\|_H^{2p} \\ &\leq C \varepsilon^p \sup_{t \in [0, 1]} \mathbf{E} \left(1 + |u(t)|_U^{2p} + |\bar{u}(t)|_U^{2p} \right) \end{aligned}$$

with $p \in [1, 2 + \frac{1}{2}(\delta_u \wedge \delta_{\bar{u}})]$; moreover, by similar arguments we have the following estimates:

$$(5.4) \quad \mathbf{E} \left\{ \varepsilon^{-p} \sup_t \|x^\varepsilon(t) - \bar{x}(t)\|_H^{2p} + \varepsilon^{-2} \sup_t \|x_2^\varepsilon(t)\|_H^2 + \varepsilon^{-2} \sup_t \|x^\varepsilon(t) - \bar{x}(t) - x_1^\varepsilon(t)\|_H^2 \right\} \leq C.$$

On the other hand, a direct calculation gives

$$(5.5) \quad \begin{aligned} \Xi^\varepsilon(t) &= \int_0^t [A(s)\Xi^\varepsilon(s) + \bar{f}_x(s)\Xi^\varepsilon(s) + \alpha^\varepsilon(s)] ds \\ &\quad + \int_0^t [B(s)\Xi^\varepsilon(s) + \bar{g}_x(s)\Xi^\varepsilon(s) + \beta^\varepsilon(s)] dW_s, \end{aligned}$$

where

$$\begin{aligned} \alpha^\varepsilon(s) &:= f_x^{\Delta, \varepsilon}(s)(x^\varepsilon(s) - \bar{x}(s)) + \frac{1}{2}(\bar{f}_{xx}^\varepsilon(s) - \bar{f}_{xx}(s))(x^\varepsilon(s) - \bar{x}(s)) \otimes (x^\varepsilon(s) - \bar{x}(s)) \\ &\quad + \frac{1}{2}\bar{f}_{xx}(s)[(x^\varepsilon(s) - \bar{x}(s)) \otimes (x^\varepsilon(s) - \bar{x}(s)) - (x_1^\varepsilon \otimes x_1^\varepsilon)(s)], \\ \beta^\varepsilon(s) &:= g_x^{\Delta, \varepsilon}(s)(x^\varepsilon(s) - \bar{x}(s) - x_1^\varepsilon(s)) + \frac{1}{2}(\bar{g}_{xx}^\varepsilon(s) - \bar{g}_{xx}(s))(x^\varepsilon(s) - \bar{x}(s)) \otimes (x^\varepsilon(s) - \bar{x}(s)) \\ &\quad + \frac{1}{2}\bar{g}_{xx}(s)[(x^\varepsilon(s) - \bar{x}(s)) \otimes (x^\varepsilon(s) - \bar{x}(s)) - (x_1^\varepsilon \otimes x_1^\varepsilon)(s)]. \end{aligned}$$

Now apply Lemma 3.1 to (5.5). Keeping in mind (5.3) and (5.4), and by means of the Hölder inequality and the Lebesgue dominated convergence theorem, we conclude

$$\mathbf{E} \sup_{t \in [0,1]} \|\Xi^\varepsilon(t)\|_H^2 \leq \mathbf{E} \left[\int_0^1 \|\alpha^\varepsilon(s)\|_H ds \right]^2 + \mathbf{E} \int_0^1 \|\beta^\varepsilon(s)\|_H^2 ds = o(\varepsilon^2).$$

The lemma is proved. \square

With the aid of the above lemma and by the fact

$$J(u^\varepsilon(\cdot)) - J(\bar{u}(\cdot)) \geq 0,$$

we can prove the following result.

LEMMA 5.2. *Under Assumptions 2.1–2.3, we have*

$$(5.6) \quad \begin{aligned} o(\varepsilon) &\leq \mathbf{E} \int_0^1 \left[l^{\Delta, \varepsilon}(t) + \langle \bar{l}_x(t), x_1^\varepsilon(t) + x_2^\varepsilon(t) \rangle + \frac{1}{2} \langle x_1^\varepsilon(t), \bar{l}_{xx}(t)x_1^\varepsilon(t) \rangle \right] dt \\ &\quad + \mathbf{E} \left[\langle h_x(\bar{x}(1)), x_1^\varepsilon(1) + x_2^\varepsilon(1) \rangle + \frac{1}{2} \langle x_1^\varepsilon(1), h_{xx}(\bar{x}(1))x_1^\varepsilon(1) \rangle \right]. \end{aligned}$$

Proof. The proof is also standard (cf. [16]), we give a sketch here. A direct calculation shows that

$$\begin{aligned} 0 &\leq J(u^\varepsilon(\cdot)) - J(\bar{u}(\cdot)) \\ &= \mathbf{E} \int_0^1 \left[l^{\Delta, \varepsilon}(t) + \langle \bar{l}_x(t), x_1^\varepsilon(1) + x_2^\varepsilon(1) \rangle + \frac{1}{2} \langle x_1^\varepsilon(t), \bar{l}_{xx}(t)x_1^\varepsilon(t) \rangle \right] dt \\ &\quad + \mathbf{E} \left[\langle h_x(\bar{x}(1)), x_1^\varepsilon(1) + x_2^\varepsilon(1) \rangle + \frac{1}{2} \langle x_1^\varepsilon(1), h_{xx}(\bar{x}(1))x_1^\varepsilon(1) \rangle \right] + \gamma(\varepsilon), \end{aligned}$$

where

$$\begin{aligned}
\gamma(\varepsilon) &:= \mathbf{E} \langle h_x(\bar{x}(1)), \Xi^\varepsilon(1) \rangle + \mathbf{E} \int_0^1 \langle \bar{l}_x(t), \Xi^\varepsilon(t) \rangle dt \\
&+ \frac{1}{2} \mathbf{E} \left\langle \left[\tilde{h}_{xx}^\varepsilon - h_{xx}(\bar{x}(1)) \right] (x^\varepsilon(1) - \bar{x}(1)), x^\varepsilon(1) - \bar{x}(1) \right\rangle \\
&+ \frac{1}{2} \mathbf{E} \langle h_{xx}(\bar{x}(1)) (x^\varepsilon(1) - \bar{x}(1)), x^\varepsilon(1) - \bar{x}(1) - x_1^\varepsilon(1) \rangle \\
&+ \frac{1}{2} \mathbf{E} \langle h_{xx}(\bar{x}(1)) (x^\varepsilon(1) - \bar{x}(1) - x_1^\varepsilon(1)), x_1^\varepsilon(1) \rangle \\
&+ \mathbf{E} \int_0^1 \langle l_x^{\Delta, \varepsilon}(t), x^\varepsilon(t) - \bar{x}(t) \rangle dt \\
&+ \frac{1}{2} \mathbf{E} \int_0^1 \left\langle \left[\tilde{l}_{xx}^\varepsilon(t) - \bar{l}_{xx}(t) \right] (x^\varepsilon(t) - \bar{x}(t)), x^\varepsilon(t) - \bar{x}(t) \right\rangle dt \\
&+ \frac{1}{2} \mathbf{E} \int_0^1 \langle \bar{l}_{xx}(t) (x^\varepsilon(t) - \bar{x}(t)), x^\varepsilon(t) - \bar{x}(t) - x_1^\varepsilon(t) \rangle dt \\
&+ \frac{1}{2} \mathbf{E} \int_0^1 \langle \bar{l}_{xx}(t) (x^\varepsilon(t) - \bar{x}(t) - x_1^\varepsilon(t)), x_1^\varepsilon(t) \rangle dt
\end{aligned}$$

with

$$\tilde{h}_{xx}^\varepsilon := 2 \int_0^1 \lambda h_{xx}(\lambda \bar{x}(1) + (1 - \lambda)x^\varepsilon(1)) d\lambda.$$

Consequently, by virtue of (5.3), (5.4) and the Lebesgue dominated convergence theorem, we can deduce $|\gamma(\varepsilon)| = o(\varepsilon)$, which implies the lemma. \square

5.2. First-order duality analysis. We need do some duality analysis in order to get the maximum condition (2.6) by sending ε to 0 in inequality (5.6). In this subsection, we still fix the $\tau \in [0, 1)$. Recall the Hamiltonian

$$\mathcal{H}(t, x, u, p, q) = l(t, x, u) + \langle p, f(t, x, u) \rangle + \langle q, g(t, x, u) \rangle,$$

and BSEE (2.5). Under Assumptions 2.1 and 2.2, it follows from Du-Meng [4, Proposition 3.2] that equation (2.5) has a unique weak solution $(p(\cdot), q(\cdot))$ with the estimate

$$(5.7) \quad \mathbf{E} \sup_{t \in [0, 1]} \|p(t)\|_H^2 + \mathbf{E} \int_0^1 \|q(t)\|_H^2 dt \leq C(\kappa, K) \sup_{t \in [0, 1]} \mathbf{E}(1 + |\bar{u}(t)|_U^2).$$

Thus the assertion (i) of Theorem 2.6 holds true. Furthermore, from Lemma 5.2 we have

LEMMA 5.3. *Under Assumptions 2.1–2.3, we have*

$$\begin{aligned}
(5.8) \quad o(1) &\leq \varepsilon^{-1} \mathbf{E} \int_\tau^{\tau+\varepsilon} [\mathcal{H}(t, \bar{x}(t), u(t), p(t), q(t)) - \mathcal{H}(t, \bar{x}(t), \bar{u}(t), p(t), q(t))] dt \\
&+ \frac{1}{2} \varepsilon^{-1} \mathbf{E} \int_0^1 \langle x_1^\varepsilon(t), \mathcal{H}_{xx}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)) x_1^\varepsilon(t) \rangle dt \\
&+ \frac{1}{2} \varepsilon^{-1} \mathbf{E} \langle x_1^\varepsilon(1), h_{xx}(\bar{x}(1)) x_1^\varepsilon(1) \rangle,
\end{aligned}$$

where $(p(\cdot), q(\cdot))$ is the solution to equation (2.5).

Proof. In view of the dual relation between the SEE and BSEE (or the Itô formula), and by (5.3) and (5.7), we have

$$\begin{aligned}
& \mathbf{E} \int_0^1 [\langle \bar{l}_x(t), x_1^\varepsilon(t) + x_2^\varepsilon(t) \rangle] dt + \mathbf{E} \langle h_x(\bar{x}(1)), x_1^\varepsilon(1) + x_2^\varepsilon(1) \rangle \\
&= \mathbf{E} \int_0^1 \left[\langle p(t), f^{\Delta, \varepsilon}(t) + \frac{1}{2} \bar{f}_{xx}(t) (x_1^\varepsilon \otimes x_1^\varepsilon)(t) \rangle \right] dt \\
&\quad + \mathbf{E} \int_0^1 \left[\langle q(t), g^{\Delta, \varepsilon}(t) + \frac{1}{2} \bar{g}_{xx}(t) (x_1^\varepsilon \otimes x_1^\varepsilon)(t) + g_x^{\Delta, \varepsilon}(t) x_1^\varepsilon(t) \rangle \right] dt \\
&= o(\varepsilon) + \mathbf{E} \int_0^1 [\langle p(t), f^{\Delta, \varepsilon}(t) \rangle + \langle q(t), g^{\Delta, \varepsilon}(t) \rangle] dt \\
&\quad + \frac{1}{2} \mathbf{E} \int_0^1 [\langle p(t), \bar{f}_{xx}(t) (x_1^\varepsilon \otimes x_1^\varepsilon)(t) \rangle + \langle q(t), \bar{g}_{xx}(t) (x_1^\varepsilon \otimes x_1^\varepsilon)(t) \rangle] dt,
\end{aligned}$$

this along with Lemma 5.2 yields

$$\begin{aligned}
o(1) &\leq \varepsilon^{-1} \mathbf{E} \int_0^1 [\mathcal{H}(t, \bar{x}(t), u^\varepsilon(t), p(t), q(t)) - \mathcal{H}(t, \bar{x}(t), \bar{u}(t), p(t), q(t))] dt \\
&\quad + \frac{1}{2} \varepsilon^{-1} \mathbf{E} \int_0^1 \langle x_1^\varepsilon(t), \mathcal{H}_{xx}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)) x_1^\varepsilon(t) \rangle dt \\
&\quad + \frac{1}{2} \varepsilon^{-1} \mathbf{E} \langle x_1^\varepsilon(1), h_{xx}(\bar{x}(1)) x_1^\varepsilon(1) \rangle.
\end{aligned}$$

Recalling the definition of $u^\varepsilon(\cdot)$, we conclude the lemma. \square

5.3. Second-order duality analysis and completion of the proof. In this subsection, we deal with the second-order expansion part, i.e., the second and third terms on the right-hand side of inequality (5.8), and complete the proof of Theorem 2.6.

Recall the four-tuple $(\tilde{A}, \tilde{B}; \tilde{M}, \tilde{N})$ with

$$\begin{aligned}
\tilde{A}(t) &:= A(t) + \bar{f}_x(t), \quad \tilde{B}(t) := B(t) + \bar{g}_x(t), \\
\tilde{M} &:= h_{xx}(\bar{x}(1)), \quad \tilde{N}(t) := \mathcal{H}_{xx}(t, \bar{x}(t), \bar{u}(t), p(t), q(t)).
\end{aligned}$$

Bearing in mind Assumptions 2.1–2.3 and the estimate (5.7), we can easily obtain that the four-tuple $(\tilde{A}, \tilde{B}; \tilde{M}, \tilde{N})$ is “appropriate”, and then from Theorem 2.5 there exists a unique representation

$$P \in L_w^2(\Omega \times [0, 1], \mathcal{P}, \mathfrak{B}(H)) \cap C_{\mathbb{F}}([0, 1], L_w^2(\Omega, \mathfrak{B}(H)))$$

with respect to $T.(\tilde{A}, \tilde{B}; \tilde{M}, \tilde{N})$. Therefore, the assertion (ii) of Theorem 2.6 is proved.

From now on we shall let τ be variable, and, to be more clarified, write $x_1^{\tau, \varepsilon} = x_1^\varepsilon$ and $g^{\Delta, \tau, \varepsilon} = g^{\Delta, \varepsilon}$ to indicate the dependence of τ .

Now we fix a predictable version $\tilde{g}^\Delta(\cdot)$ of $g^\Delta(\cdot)$ such that $\mathbf{E} \|\tilde{g}^\Delta(\tau)\|_H^4 < \infty$ for each $\tau \in [0, 1]$. We introduce the following equation

$$z^{\tau, \varepsilon}(t) = \int_\tau^t \tilde{A}(s) z^{\tau, \varepsilon}(s) ds + \int_\tau^t [\tilde{B}(s) z^{\tau, \varepsilon}(s) + \varepsilon^{-\frac{1}{2}} \tilde{g}^\Delta(\tau) \mathbf{1}_{[\tau, \tau + \varepsilon]}(s)] dW_s, \quad t \in [\tau, 1].$$

Then we have

LEMMA 5.4. *For a.e. $\tau \in [0, 1)$,*

$$\lim_{\varepsilon \downarrow 0} \mathbf{E} \sup_{t \in [\tau, 1]} \left\| \varepsilon^{-\frac{1}{2}} x_1^{\tau, \varepsilon}(t) - z^{\tau, \varepsilon}(t) \right\|_H^4 = 0.$$

Proof. Recall that $g^{\Delta, \tau, \varepsilon}(t) = g^{\Delta, \varepsilon}(t) = g^{\Delta}(t) \cdot \mathbf{1}_{[\tau, \tau + \varepsilon]}(t)$. In view of Lemma 3.1, the Hölder inequality and the Fubini theorem, we have that, for each $\tau \in [0, 1)$,

$$\begin{aligned} \mathbf{E} \sup_{t \in [\tau, 1]} \left\| \varepsilon^{-\frac{1}{2}} x_1^{\tau, \varepsilon}(t) - z^{\tau, \varepsilon}(t) \right\|_H^4 &\leq \frac{C}{\varepsilon^2} \mathbf{E} \left(\int_{\tau}^{\tau + \varepsilon} \|\tilde{g}^{\Delta}(t) - \tilde{g}^{\Delta}(\tau)\|_H^2 \cdot \mathbf{1}_{[\tau, \tau + \varepsilon]}(t) dt \right)^2 \\ &\leq \frac{C}{\varepsilon} \mathbf{E} \int_{\tau}^{\tau + \varepsilon} \|\tilde{g}^{\Delta}(t) - \tilde{g}^{\Delta}(\tau)\|_H^4 dt \leq \frac{C}{\varepsilon} \int_{\tau}^{\tau + \varepsilon} \mathbf{E} \|\tilde{g}^{\Delta}(t) - \tilde{g}^{\Delta}(\tau)\|_H^4 dt. \end{aligned}$$

By the Lebesgue differentiation theorem, we have for each $X \in L^4(\Omega, \mathcal{F}_1, H)$,

$$(5.9) \quad \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{\tau}^{\tau + \varepsilon} \mathbf{E} \|\tilde{g}^{\Delta}(t) - X\|_H^4 dt = \mathbf{E} \|\tilde{g}^{\Delta}(\tau) - X\|_H^4, \quad \text{for a.e. } \tau \in [0, 1).$$

Since $L^4(\Omega, \mathcal{F}_1, H)$ is separable, let X run through a countable dense subset Q of $L^4(\Omega, \mathcal{F}_1, H)$, and denote

$$E := \bigcup_{X \in Q} E_X := \bigcup_{X \in Q} \{\tau : \text{relation (5.9) does not hold for } X\}.$$

Then $\text{meas}(E) = 0$. For each $\tau \in [0, 1) \setminus E$ and any positive η , we can take an $X_{\tau, \eta} \in Q$ such that

$$\mathbf{E} \|\tilde{g}^{\Delta}(\tau) - X_{\tau, \eta}\|_H^4 < \eta,$$

then we have

$$\begin{aligned} &\lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \int_{\tau}^{\tau + \varepsilon} \mathbf{E} \|\tilde{g}^{\Delta}(t) - \tilde{g}^{\Delta}(\tau)\|_H^4 dt \\ &\leq \lim_{\varepsilon \downarrow 0} \frac{8}{\varepsilon} \int_{\tau}^{\tau + \varepsilon} \mathbf{E} \|\tilde{g}^{\Delta}(t) - X_{\tau, \eta}\|_H^4 dt + 8 \mathbf{E} \|\tilde{g}^{\Delta}(\tau) - X_{\tau, \eta}\|_H^4 \\ &\leq 16 \mathbf{E} \|\tilde{g}^{\Delta}(\tau) - X_{\tau, \eta}\|_H^4 < 16\eta. \end{aligned}$$

From the arbitrariness of η , we conclude this lemma. \square

Thanks to the the above lemma, we have

$$(5.10) \quad \begin{aligned} &\varepsilon^{-1} \mathbf{E} \int_0^1 \left\langle x_1^{\tau, \varepsilon}(t), \tilde{N}(t) x_1^{\tau, \varepsilon}(t) \right\rangle dt + \varepsilon^{-1} \mathbf{E} \left\langle x_1^{\tau, \varepsilon}(1), \tilde{M} x_1^{\tau, \varepsilon}(1) \right\rangle \\ &= o(1) + \mathbf{E} \int_{\tau}^1 \left\langle z^{\tau, \varepsilon}(t), \tilde{N}(t) z^{\tau, \varepsilon}(t) \right\rangle dt + \mathbf{E} \left\langle z^{\tau, \varepsilon}(1), \tilde{M} z^{\tau, \varepsilon}(1) \right\rangle, \quad \forall \tau \in [0, 1) \setminus E. \end{aligned}$$

Keeping in mind the above relation, and applying Proposition 4.1, we conclude for each $\tau \in [0, 1) \setminus E$,

$$(5.11) \quad \mathbf{E} \langle \tilde{g}^{\Delta}(\tau), P_{\tau} \tilde{g}^{\Delta}(\tau) \rangle = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \left\{ \mathbf{E} \int_0^1 \left\langle x_1^{\varepsilon}(t), \tilde{N}(t) x_1^{\varepsilon}(t) \right\rangle dt + \mathbf{E} \left\langle x_1^{\varepsilon}(1), \tilde{M} x_1^{\varepsilon}(1) \right\rangle \right\}.$$

In view of Lemma 5.3, by denoting

$$\mathcal{H}^\Delta(\tau) := \mathcal{H}(\tau, \bar{x}(\tau), u(\tau), p(\tau), q(\tau)) - \mathcal{H}(\tau, \bar{x}(\tau), \bar{u}(\tau), p(\tau), q(\tau)),$$

and using the Lebesgue differentiation theorem and (5.11), we obtain

$$(5.12) \quad 0 \leq \mathbf{E} \left[\mathcal{H}^\Delta(\tau) + \frac{1}{2} \langle \tilde{g}^\Delta(\tau), P_\tau \tilde{g}^\Delta(\tau) \rangle \right], \quad \text{a.e. } \tau \in [0, 1].$$

In view of the Fubini theorem, we have

$$\int_0^1 \mathbf{E} \langle \tilde{g}^\Delta(\tau), P_\tau \tilde{g}^\Delta(\tau) \rangle d\tau = \mathbf{E} \int_0^1 \langle \tilde{g}^\Delta(\tau), P_\tau \tilde{g}^\Delta(\tau) \rangle d\tau = \mathbf{E} \int_0^1 \langle g^\Delta(\tau), P_\tau g^\Delta(\tau) \rangle d\tau.$$

Combining with (5.12), we obtain

$$0 \leq \mathbf{E} \int_0^1 \left[\mathcal{H}^\Delta(\tau) + \frac{1}{2} \langle g^\Delta(\tau), P_\tau g^\Delta(\tau) \rangle \right] d\tau.$$

Therefore, the desired maximum condition (2.6) follows from the arbitrary choice of control $u(\cdot)$. This completes the proof of Theorem 2.6.

6. Examples. In the following let us discuss two examples which can be covered by the abstract results of the present paper.

EXAMPLE 6.1 (SPDE with controlled coefficients). Given a bounded domain $\mathcal{D} \subset \mathbb{R}^n$, we consider the following controlled stochastic PDE

$$(6.1) \quad dy(t, \xi) = [Ay(t, \xi) + f(t, \xi, u_t)y(t, \xi)] dt + [By(t, \xi) + g(t, \xi, u_t)y(t, \xi)] dW_t$$

with $(t, \xi) \in [0, 1] \times \mathcal{D}$, initial data $y(0, \xi) = y_0(\xi)$ and a proper boundary condition. Here the control u is a stochastic process with values in a set $U \subset \mathbb{R}$, and A, B are differential operators defined as

$$A(t, \xi) = \sum_{i,j=1}^n a^{ij}(t, \xi) \frac{\partial^2}{\partial \xi_i \partial \xi_j}, \quad B(t, \xi) = \sum_{i=1}^n \sigma^i(t, \xi) \frac{\partial}{\partial \xi_i}.$$

The equation is usually called super-parabolic SPDE if there is a constant $\kappa > 0$ such that

$$\kappa I_n \leq (2a^{ij} - \sigma^i \sigma^j)_{n \times n} \leq \kappa^{-1} I_n.$$

The fundamental theory about this kind of SPDEs can be found in [2, 13], etc. Here we consider minimizing the cost functional

$$(6.2) \quad J(u(\cdot)) = \mathbf{E} \int_{\mathcal{D}} |y(t, \xi)|^2 d\xi + \mathbf{E} \int_0^1 \int_{\mathcal{D}} l(t, \xi, u_t) |y(t, \xi)|^2 d\xi dt.$$

Provided some assumptions on coefficients, such as f, g and l are all $\mathcal{P} \times \mathcal{B}(\mathcal{D}) \times \mathcal{B}(U)$ -measurable functions and dominated by a given constant, one can verify the conditions of Theorem 2.6 and apply the abstract result directly in this case.

EXAMPLE 6.2. Given a bounded domain $\mathcal{D} \subset \mathbb{R}^n$, we consider the stochastic heat equation

$$\begin{cases} dy(t, \xi) = \Delta y(t, \xi) dt + b(\xi)u(t, \xi) dW_t, & (t, \xi) \in [0, 1] \times \mathcal{D}; \\ y(0, \xi) = y_0(\xi), \xi \in \mathcal{D}; \quad y(t, \xi) = 0, & (t, \xi) \in [0, 1] \times \partial\mathcal{D}. \end{cases}$$

where the control $u(\cdot, \cdot)$ is a random field with values in a set $E \subset \mathbb{R}$, and the coefficient b is a given bounded function. The objective of the control problem is to minimize the following cost functional

$$J(u(\cdot)) = \frac{1}{2} \mathbf{E} \int_{\mathcal{D}} |y(1, \xi)|^2 d\xi + \mathbf{E} \int_0^1 \int_{\mathcal{D}} c(\xi) |u(t, \xi)|^2 d\xi dt.$$

This problem can be covered by our result by taking

$$H = L^2(\mathcal{D}), \quad U = L^2(\mathcal{D}; E), \quad A(t) \equiv \Delta, \quad B(t) \equiv 0.$$

Let (\bar{y}, \bar{u}) be an optimal solution. Then the first-order adjoint process (p, q) is given by the following equation

$$\begin{cases} dp(t, \xi) = -\Delta p(t, \xi) dt + q(t, \xi) dW_t, & (t, \xi) \in [0, 1] \times \mathcal{D}; \\ p(1, x) = \bar{y}(1, x), \quad x \in \mathcal{D}; \quad p(t, \xi) = 0, & (t, \xi) \in [0, 1] \times \partial\mathcal{D}. \end{cases}$$

Moreover, it is easy to verify that the second-order adjoint process P satisfies

$$\langle f, P_t f \rangle = \int_{\mathcal{D}} |e^{(1-t)\Delta} f|^2(\xi) d\xi, \quad \forall f \in H$$

for each $t \in [0, 1]$, where $(e^{t\Delta})_{t \geq 0}$ is the semigroup generated by the Laplacian operator on H . In view of Theorem 2.6 we can write down the following necessary condition for optimal control: for any $u(\cdot, \cdot) \in U$,

$$\begin{aligned} \int_{\mathcal{D}} \left\{ b(\xi) q(t, \xi) [u(t, \xi) - \bar{u}(t, \xi)] + c(\xi) [|u(t, \xi)|^2 - |\bar{u}(t, \xi)|^2] \right. \\ \left. + \frac{1}{2} |e^{(1-t)\Delta} [bu(t, \cdot) - b\bar{u}(t, \cdot)]|^2(\xi) \right\} d\xi \geq 0 \end{aligned}$$

holds for a.e. (t, ω) .

REMARK 6.3. (1) The above examples cannot be covered by the results in either [10] or [5] due to the quadratic form of the cost functionals.

(2) Frankly speaking, the requirement of twice Fréchet differentiability of coefficients restricts the applicability of our abstract result; for instance, the SPDE with general Nemytskii-type coefficients rarely fits into this framework, and yet the result obtained in [5] is not covered by ours. Nevertheless, we believe that some key approaches in this paper, especially the techniques used in second-order duality analysis, could also apply to many other concrete problems. Further investigations thereof would be presented in future publications.

Acknowledgments. The authors would like to thank the referees and editor for their helpful comments and suggestions.

REFERENCES

- [1] A. BENSOUSSAN, *Stochastic maximum principle for distributed parameter systems*, Journal of the Franklin Institute, 315 (1983), pp. 387–406.
- [2] P.L. CHOW, *Stochastic partial differential equations*, vol. 11, CRC Press, 2007.
- [3] J.L. DOOB, *Stochastic processes*, Wiley, 1953.
- [4] K. DU AND Q. MENG, *A revisit to W_2^n -theory of super-parabolic backward stochastic partial differential equations in R^d* , Stoch. Proc. Appl., 120 (2010), pp. 1996–2015.

- [5] M. FUHRMAN, Y. HU, AND G. TESSITORE, *Stochastic maximum principle for optimal control of SPDEs*, Comptes Rendus Mathematique, (2012).
- [6] ———, *Stochastic maximum principle for optimal control of spdes*, Applied mathematics & optimization, DOI: 10.1007/s00245-013-9203-7, (2013).
- [7] SHENG-WU HE, CHIA-KANG WANG, AND JIA-AN YAN, *Semimartingale theory and stochastic calculus*, Taylor & Francis US, 1992.
- [8] Y. HU AND S. PENG, *Maximum principle for semilinear stochastic evolution control systems*, Stochastics: An International Journal of Probability and Stochastic Processes, 33 (1990), pp. 159–180.
- [9] N.V. KRYLOV AND B.L. ROZOVSKII, *Stochastic evolution equations*, Journal of Mathematical Sciences, 16 (1981), pp. 1233–1277.
- [10] Q. LÜ AND X. ZHANG, *General pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions*, Arxiv preprint arXiv:1204.3275v5, (2012).
- [11] S. PENG, *A general stochastic maximum principle for optimal control problems*, SIAM Journal on control and optimization, 28 (1990), pp. 966–979.
- [12] L.S. PONTRYAGIN, V.G. BOLTYANSKII, R.V. GAMKRELIDZE, AND E. MISHCHENKO, *The mathematical theory of optimal processes*, 1962.
- [13] BL ROZOVSKY, *Stochastic Evolution Systems*, Nauka, Moscow (in Russian-English Translation), 1990.
- [14] S. TANG AND X.J. LI, *Maximum principle for optimal control of distributed parameter stochastic systems with random jumps*, Lecture Notes in Pure and Applied Mathematics, (1993), pp. 867–890.
- [15] C. TUDOR, *Optimal control for semilinear stochastic evolution equations*, Applied mathematics & optimization, 20 (1989), pp. 319–331.
- [16] J. YONG AND X.Y. ZHOU, *Stochastic controls: Hamiltonian systems and HJB equations*, vol. 43, Springer Verlag, 1999.
- [17] X.Y. ZHOU, *On the necessary conditions of optimal controls for stochastic partial differential equations*, SIAM J. Cont. Optim, 31 (1993), pp. 1462–1478.