Numerical analysis of a Neumann boundary control problem with a stochastic parabolic equation

Qin Zhou¹,² & Binjie Li¹,*

¹School of Mathematics, Sichuan University, Chengdu 610064, China;
²School of Mathematics & Information, China West Normal University, Nanchong 637002, China

Email: zhouqinmath@stu.scu.edu.cn, libinjie@scu.edu.cn

Received May 9, 2021; accepted August 30, 2022; published online March 14, 2023

Abstract We analyze the discretization of a Neumann boundary control problem with a stochastic parabolic equation, where additive noise occurs in the Neumann boundary condition. The convergence is established for general filtration, and the convergence rate $O(\tau^{1/4-\epsilon} + h^{1/2-\epsilon})$ is derived for the natural filtration of the $Q$-Wiener process.

Keywords Neumann boundary control, stochastic parabolic equation, $Q$-Wiener process, boundary noise, discretization, convergence

MSC(2020) 65C30, 65K10

1 Introduction

Let $(\Omega, \mathcal{F}, P)$ be a given complete probability space with a normal filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ (i.e., $\mathbb{F}$ is right-continuous and $\mathcal{F}_0$ contains all the $P$-null sets in $\mathcal{F}$). Assume that $\mathcal{O} \subset \mathbb{R}^d$ ($d = 2, 3$) is a convex polygonal domain with the boundary $\Gamma$. Let $W$ be a $Q$-Wiener process in $L^2(\Gamma)$ of the form

$$W(t) = \sum_{n=0}^{\infty} \sqrt{\lambda_n} \beta_n(t) \phi_n, \quad t \geq 0,$$

where the following conditions hold: $\{\lambda_n\}_{n=0}^{\infty}$ is a sequence of strictly positive numbers satisfying $\sum_{n=0}^{\infty} \lambda_n < \infty$; $\{\beta_n\}_{n=0}^{\infty}$ is a sequence of real-valued Brownian motions defined on $(\Omega, \mathcal{F}, P)$, mutually independent and $\mathbb{F}$-adapted; $\{\phi_n\}_{n=0}^{\infty}$ is an orthonormal basis of $L^2(\Gamma)$. We consider the following model problem:

$$\min_{u \in U_{ad}} \mathcal{J}(u, y) := \frac{1}{2} \|y - y_d\|^2_{L^2(\Omega; L^2(0, T; L^2(\mathcal{O})))} + \frac{\nu}{2} \|u\|^2_{L^2(\Omega; L^2(0, T; L^2(\Gamma)))}$$

*Corresponding author
subject to the state equation

\[
\begin{align*}
\frac{\partial}{\partial t} y(t, x) - \Delta y(t, x) + y(t, x) &= 0, \quad 0 \leq t \leq T, \quad x \in \mathcal{O}, \\
\partial_n y(t, x) &= u(t, x) + \sum_{n=0}^{\infty} \lambda_n^{1/2} (\sigma(t) \phi_n)(x) \frac{d\beta_n(t)}{dt}, \quad 0 \leq t \leq T, \quad x \in \Gamma, \\
y(0, x) &= 0, \quad x \in \mathcal{O},
\end{align*}
\]

where \(0 < \nu, T < \infty, y_d\) and \(\sigma\) are two given processes, and \(n\) is the outward unit normal vector to \(\Gamma\).

The above admissible space \(U_{ad}\) is defined by

\[
U_{ad} := \{ u \in L^2(\Omega; L^2(0, T; L^2(\Gamma))) \mid u_* \leq u \leq u^* \},
\]

where \(u_* < u^*\) are two given constants.

In the past four decades, a considerable number of papers have been published in the field of optimal control problems of stochastic partial differential equations (see [2,8,9,11,12,17,18,21], [27,28,34,38] and the references therein). By now it is still a highly active research area. However, the numerical analysis of stochastic parabolic optimal control problems is quite rare. The numerical analysis of stochastic parabolic optimal control problems consists of two main challenges. The first is to derive the convergence rate for the rough data. To keep the co-state \(F\)-adapted, the adjoint equation has a correction term \(z(t) dW(t)\).

The process \(z\) is of low temporal regularity for the rough data, and this makes the convergence rate difficult to derive. Here, we especially emphasize the case where the filtration is not the natural one of the Wiener process. Since in this case, the martingale representation theorem cannot be utilized, generally there is essential difficulty in analyzing the adjoint equation of the stochastic optimal control problems (see [27]). The second is to design efficient fully implementable algorithms. Generally, the discretization of the adjoint equation will be used for the construction of an efficient algorithm for a parabolic optimal control problem. The adjoint equation of a stochastic optimal control problem is a backward stochastic parabolic equation. To solve a backward parabolic equation requires computing a conditional expectation at each time step. However, computing conditional expectations is a challenging problem (see [1,3,14,26] and the references therein).

We briefly summarize the numerical analysis of stochastic parabolic optimal control problems in the literature as follows. Dunst and Prohl [10] analyzed a spatial semi-discretization of a forward-backward stochastic heat equation, and using the least squares Monte-Carlo method to compute the conditional expectations, they constructed two fully implementable algorithms. Levajković et al. [22] presented an approximation framework for computing the solution of the stochastic linear quadratic control problem on Hilbert spaces; however, the time variable is not discretized in this framework. Recently, Prohl and Wang [32,33] numerically analyzed the distributed optimal control problems of the stochastic heat equation driven by additive and multiplicative noise, and Li and Zhou [25] analyzed a distributed optimal control problem of a stochastic parabolic equation driven by the multiplicative noise. We refer the readers to [15] and the references therein for the numerical analysis of optimal control problems governed by stochastic differential equations, and refer the readers to [5,19] and the references therein for the numerical analysis of optimal control problems governed by random partial differential equations. The above works of stochastic parabolic optimal control problems generally require that \(\mathcal{F}\) should be the natural filtration of the Wiener process, and to the best of our knowledge, no numerical result is available for the stochastic parabolic optimal control problems with boundary noise.

Since the noise is additive and the state equation is linear, we essentially use a backward parabolic equation parameterized by the argument \(\omega \in \Omega\) as the adjoint equation, instead of using the backward stochastic parabolic equation. This enables us to establish the convergence of the discrete stochastic optimal control problem for general filtration, which is the main novelty of our theoretical results. Furthermore, under the condition that \(\mathcal{F}\) is the natural filtration of \(W(\cdot)\), we mainly use the mild solution theory of the backward stochastic parabolic equations to derive the convergence rate \(O(\tau^{1/4-\epsilon} + h^{1/2-\epsilon})\)
for the rough data $y_d \in L^2_γ(\Omega; L^2(0, T; L^2(\mathcal{O})))$ and
\[
\sigma \in L^2_γ(\Omega; L^2(0, T; \mathcal{L}^2_0)) \cap L^\infty(0, T; L^2(\Omega; \mathcal{L}^2_0)),
\]
where $\epsilon$ is a sufficiently small number. Our numerical analysis can be easily extended to the distributed optimal control problems governed by linear stochastic parabolic equations with the additive noise and general filtration. When the filtration is the natural one of the $Q$-Wiener process, our analysis can remove the restrictions on the data imposed in [32].

The rest of this paper is organized as follows. In Section 2, we introduce some notations and the first-order optimality condition of the problem (1.1). In Section 3, we present a discrete stochastic optimal control problem. In Section 4, we prove the convergence of this discrete stochastic optimal control problem. In Section 5, we conclude this paper. In Appendix A, we give an interpolation space result.

## 2 Preliminaries

**Conventions.** For any random variable $v$ defined on $(\Omega, \mathcal{F}, P)$, $E_v$ denotes the expectation of $v$, and $E_vv$ denotes the conditional expectation of $v$ with respect to $\mathcal{F}_t$ for any $t > 0$. We use $\langle \cdot, \cdot \rangle_\mathcal{O}$ and $\langle \cdot, \cdot \rangle_\Gamma$ to denote the inner products of $L^2(\Omega)$ and $L^2(\Gamma)$, respectively, and use $[\cdot, \cdot]$ and $\langle \cdot, \cdot \rangle$ to denote the inner products of $L^2(\Omega; L^2(\mathcal{O}))$ and $L^2(\Omega; L^2(\Gamma))$, respectively. The notation $I$ denotes the identity mapping. For any Hilbert space $X$, we simply use $\|\cdot\|_X$ to denote the norm of the Hilbert space $L^2(\Omega; X)$. For any two separable Hilbert spaces $X_1$ and $X_2$, $\mathcal{L}(X_1, X_2)$ denotes the set of all the bounded linear operators from $X_1$ to $X_2$, and $\mathcal{L}_2(X_1, X_2)$ denotes the set of all the Hilbert-Schmidt operators from $X_1$ to $X_2$. Define

$$L^2_\lambda := \left\{ \sum_{n=0}^{\infty} c_n \sqrt{\lambda_n} \phi_n \mid \sum_{n=0}^{\infty} c_n^2 < \infty \right\}$$

and endow this space with the inner product

$$\left( \sum_{n=0}^{\infty} c_n \sqrt{\lambda_n} \phi_n, \sum_{n=0}^{\infty} d_n \sqrt{\lambda_n} \phi_n \right)_{L^2_\lambda} = \sum_{n=0}^{\infty} c_n d_n$$

for all $\sum_{n=0}^{\infty} c_n \sqrt{\lambda_n} \phi_n \in L^2_\lambda$ and $\sum_{n=0}^{\infty} d_n \sqrt{\lambda_n} \phi_n \in L^2_\lambda$. In particular, $\mathcal{L}_2(L^2_\lambda, L^2(\Gamma))$ is abbreviated to $\mathcal{L}_2(\mathcal{O})$. For any separable Hilbert space $X$, define

$$L^2_\mathcal{O}(\Omega; L^2(0, T; X)) := \{ \varphi : [0, T] \times \Omega \to X \mid \varphi \text{ is } \mathcal{F} \text{-progressively measurable and } \|\varphi\|_{L^2(0, T; X)} < \infty \},$$

and let $\mathcal{E}_\mathcal{O}$ be the $L^2(\Omega; L^2(0, T; X))$-orthogonal projection onto $L^2_\mathcal{O}(\Omega; L^2(0, T; X))$. In addition, let $L^2_\mathcal{O}(\Omega; C([0, T]; X))$ be the space of all the $X$-valued and $\mathcal{F}$-adapted continuous processes. By saying that $\mathcal{F}$ is the natural filtration of $W(\cdot)$, we mean that $\mathcal{F}$ is generated by $W(\cdot)$ and augmented by

$$\{ \mathcal{N} \in \mathcal{F} \mid P(\mathcal{N}) = 0 \}.$$

**Definition of $\hat{H}^\gamma$.** Let $A$ be the realization of the partial differential operator $\Delta - I$ in $L^2(\mathcal{O})$ with the homogeneous Neumann boundary condition. More precisely,

$$\text{Domain}(A) := \{ v \in H^2(\mathcal{O}) \mid \partial_n v = 0 \text{ on } \Gamma \} \quad \text{and} \quad Av := \Delta v - v \quad \text{for all } v \in \text{Domain}(A).$$

For any $\gamma \geq 0$, define

$$\hat{\mathcal{H}}^\gamma := \{ (-A)^{-\gamma/2} v \mid v \in L^2(\mathcal{O}) \}$$

and endow this space with the norm

$$\| v \|_{\hat{\mathcal{H}}^\gamma} := \| (-A)^{\gamma/2} v \|_{L^2(\mathcal{O})}, \quad \forall v \in \hat{\mathcal{H}}^\gamma.$$
For each $\gamma > 0$, we use $\dot{H}^{-\gamma}$ to denote the dual space of $\dot{H}^\gamma$, and use $\langle \cdot, \cdot \rangle_{\dot{H}^\gamma}$ to denote the duality pairing between $\dot{H}^{-\gamma}$ and $\dot{H}^\gamma$. In the sequel, we mainly use the notation $\dot{H}^0$ instead of $L^2(O)$ for convenience.

**Definitions of $S_0$ and $S_1$.** Extend $A$ as a bounded linear operator from $\dot{H}^0$ to $\dot{H}^{-2}$ by

$$\langle Av, \varphi \rangle_{\dot{H}^2} := \langle v, A\varphi \rangle_0$$

for all $v \in \dot{H}^0$ and $\varphi \in \dot{H}^2$. The operator $A$ then generates an analytic semigroup $\{e^{tA}\}_{t \geq 0}$ in $\dot{H}^{-2}$ (see [36, Theorems 2.5 and 2.8]), possessing the following property (see [31, Theorem 6.13, Chapter 2]): for any $-2 \leq \beta \leq \gamma \leq 2$ and $t > 0$,

$$\|e^{tA}\|_{\mathcal{L}(\dot{H}^\beta, \dot{H}^\gamma)} \leq Ce^{(\beta - \gamma)/2},$$

(2.1)

where $C$ is a positive constant depending only on $\beta$ and $\gamma$. For any $g \in L^2(0,T; \dot{H}^{-2})$, the equation

$$\begin{cases}
y'(t) - Ay(t) = g(t), & \forall 0 \leq t \leq T, \\
y(0) = 0
\end{cases}$$

(2.2)

admits a unique mild solution

$$(S_0 g)(t) := \int_0^t e^{(t-s)A}g(s)ds, \quad 0 \leq t \leq T.$$ 

(2.3)

Symmetrically, for any $g \in L^2(0,T; \dot{H}^{-2})$, the equation

$$\begin{cases}
y'(t) - Ay(t) = g(t), & \forall 0 \leq t \leq T, \\
y(T) = 0
\end{cases}$$

(2.4)

admits a unique mild solution

$$(S_1 g)(t) := \int_0^T e^{(s-t)A}g(s)ds, \quad 0 \leq t \leq T.$$ 

(2.5)

For any $f, g \in L^2(0,T; \dot{H}^{-1})$, a routine energy argument yields

$$\|S_0 f\|_{L^2(0,T; \dot{H}^1)} \leq \|f\|_{L^2(0,T; \dot{H}^{-1})},$$

(2.6)

$$\|S_1 g\|_{L^2(0,T; \dot{H}^1)} \leq \|g\|_{L^2(0,T; \dot{H}^{-1})},$$

(2.7)

and it is easily verified that

$$\int_0^T \langle g(t), (S_0 f)(t) \rangle_{\dot{H}^1} dt = \int_0^T \langle f(t), (S_1 g)(t) \rangle_{\dot{H}^1} dt.$$ 

(2.8)

**Definition of $G$.** Assume that $\sigma \in L^2(\Omega; L^2(0,T; \dot{L}_2^0))$. Let $G$ be the mild solution of the stochastic evolution equation

$$\begin{cases}
dG(t) = AG(t)dt + \mathcal{R}\sigma(t)dW(t), & \forall 0 \leq t \leq T, \\
G(0) = 0
\end{cases}$$

(2.9)

where $\mathcal{R} \in \mathcal{L}(L^2(\Gamma), \dot{H}^{-1/2-\epsilon})$ $(0 < \epsilon \leq 1/2)$ is defined by

$$\langle \mathcal{R}v, \varphi \rangle_{\dot{H}^{1/2+\epsilon}} := \langle v, \varphi \rangle_{\dot{H}^{1/2+\epsilon}}$$

for all $v \in L^2(\Gamma)$ and $\varphi \in \dot{H}^{1/2+\epsilon}$.

(2.10)

For any $0 \leq t \leq T$, we have (see [13, Chapter 3] and [7, Chapter 5])

$$G(t) = \int_0^t e^{(t-s)A}\mathcal{R}\sigma(s)dW(s) \quad \text{P.-a.s.},$$

(2.11)
If \( \sigma \in L^2_T(\Omega; L^2(0,T; \mathcal{L}_F^2)) \cap L^\infty(0,T; L^2(\Omega; \mathcal{L}_F^2)) \), then for any \( 0 < t \leq T \) and \( 0 \leq \gamma < (1/2 - \epsilon)/2 \),

\[
\| G(t) \|_{\mathcal{H}^{2\gamma}}^2 = \int_0^t \| e^{(t-s)A} \mathcal{R} \sigma(s) \|_{\mathcal{L}_F(L^2(\mathcal{H}^{1/2-\epsilon})))}^2 ds
\]

\[
\leq C \int_0^t (t-s)^{-2\gamma-1/2-\epsilon} \| R \|_{\mathcal{L}(L^2(\mathcal{H}^{1/2-\epsilon}-1/2-\gamma}) \| \sigma(s) \|_{\mathcal{L}_F(L^2(\mathcal{H}^{1/2-\epsilon}-1/2-\gamma})}^2 ds (by (2.1))
\]

\[
\leq C t^{1/2-2\gamma-\epsilon} \| R \|^2_{\mathcal{L}(L^2(\mathcal{H}^{1/2-\epsilon}-1/2-\gamma}} \| \sigma \|^2_{\mathcal{L}^\infty(0,T; L^2(\Omega; \mathcal{L}_F^2))}(by (2.1)).
\]

so that by the fact that \( \mathcal{R} \in \mathcal{L}(L^2(\Gamma), \mathcal{H}^{1/2-\epsilon}) \), we obtain

\[
\| G(t) \|_{\mathcal{H}^{2\gamma}} \leq C t^{1/4-\gamma-\epsilon/2} \| \sigma \|_{\mathcal{L}^\infty(0,T; L^2(\Omega; \mathcal{L}_F^2))}.
\]

The above \( C \) denotes a positive constant depending only on \( \epsilon, \gamma \) and \( \mathcal{O} \), and its value may differ in different places. For more theoretical results, we refer the readers to [6].

**Remark 2.1.** By [16, Theorem 1.5.1.2], we see that \( \mathcal{H}^{1/2+\epsilon} \) is continuously embedded into \( L^2(\Gamma) \), so the above operator \( \mathcal{R} \) is well defined.

**First-order optimality condition.** We call \( \bar{u} \in U_{ad} \) a solution to the problem (1.1) if \( \bar{u} \) minimizes the cost functional

\[
\mathcal{J}(u) := \frac{1}{2} \| S_0 \mathcal{R} u + G - y_d \|^2_{L^2(0,T; L^2(\Omega))} + \nu u^2_{L^2(0,T; L^2(\Gamma))}, \quad u \in U_{ad}.
\]

It is standard that the problem (1.1) admits a unique solution \( \bar{u} \). Moreover, by (2.8) and (2.10), we have the first-order optimality condition

\[
\int_0^T \langle \bar{p} + \nu \bar{u}, u - \bar{u} \rangle dt \geq 0 \quad \text{for all } u \in U_{ad}.
\]

where

\[
\bar{p} := S_0 \bar{R} \bar{u} + G,
\]

\[
\bar{p} := S_1 (\bar{y} - y_d).
\]

It follows that

\[
\bar{u} = \mathcal{P}_{[u_*, u^*]} \left( -\frac{1}{\nu} \mathcal{E}_2(\text{tr} \bar{p}) \right),
\]

where \( \text{tr} \) is the trace operator from \( \mathcal{H}^1 \) to \( L^2(\Gamma) \) and

\[
\mathcal{P}_{[u_*, u^*]}(r) := \begin{cases} u_*, & \text{if } r < u_*, \\ r, & \text{if } u_* \leq r \leq u^*, \\ u^*, & \text{if } r > u^*. \end{cases}
\]

**Remark 2.2.** Assume that \( \mathbb{F} \) is the natural filtration of \( \mathcal{W}(\cdot) \). The usual first-order optimality condition of the problem (1.1) is that (see [2,8,9,12,17,18])

\[
\int_0^T \langle \bar{p} + \nu \bar{u}, u - \bar{u} \rangle dt \geq 0 \quad \text{for all } u \in U_{ad}.
\]

The above \( \bar{p} \) is the first component of the solution \( (\bar{p}, \bar{z}) \) to the backward stochastic parabolic equation

\[
\begin{cases}
\frac{d\bar{p}(t)}{dt} = -\langle Ap + S_0 \mathcal{R} \bar{u} + G - y_d \rangle(t) dt + \bar{z}(t) d\mathcal{W}(t), & 0 \leq t \leq T, \\
\bar{p}(T) = 0,
\end{cases}
\]
where $\bar{W}$ is a cylindrical Wiener process in $L^2(\mathcal{O})$ defined later by (4.33). Moreover, following the proof of Lemma 4.15, we can obtain

$$\bar{u} \in L^2_F(\Omega; C([0, T]; H^{1/2}(\Gamma))),$$

where $H^{1/2}(\Gamma)$ is a standard fractional order Sobolev space on $\Gamma$.

### 3 Discrete stochastic optimal control problems

Let $J > 0$ be an integer and define $t_j := j \tau$ for each $0 \leq j \leq J$, where $\tau := T/J$. Let $\mathcal{K}_h$ be a conventional conforming, shape regular and quasi-uniform triangulation of $\mathcal{O}$ consisting of $d$-simplexes, and we use $h$ to denote the maximum diameter of the elements in $\mathcal{K}_h$. Define

$$\mathcal{V}_h := \{ v_h \in C(\overline{\mathcal{O}}) \mid v_h \text{ is linear on each } K \in \mathcal{K}_h \},$$

$$\mathcal{X}_{h, \tau} := \{ v : [0, T] \times \Omega \to \mathcal{V}_h \mid V(t_j) \in L^2(\Omega; \mathcal{V}_h) \text{ and } V \text{ is constant on } [t_j, t_{j+1}] \text{ for each } 0 \leq j < J \}.$$

For any $V \in \mathcal{X}_{h, \tau}$, by $V_j$ we mean $V(t_j)$ for each $0 \leq j \leq J$. Define $Q_h : \dot{H}^{-1} \to \mathcal{V}_h$ by

$$\langle Q_h v, v_h \rangle_\mathcal{O} = \langle v, v_h \rangle_{H^1} \quad \text{for all } v \in \dot{H}^{-1} \text{ and } v_h \in \mathcal{V}_h,$$

and define $A_h : \mathcal{V}_h \to \mathcal{V}_h$ by

$$\langle A_h v_h, w_h \rangle_\mathcal{O} = -\int_{\mathcal{O}} \nabla v_h \cdot \nabla w_h + v_h w_h \quad \text{for all } v_h, w_h \in \mathcal{V}_h.$$

For any $g \in L^2(\Omega; L^2(0, T; H^{-1}))$, define $S_{0}^{h, \tau} g \in \mathcal{X}_{h, \tau}$, an approximation of $S_0 g$, by

$$\begin{aligned}
(S_{0}^{h, \tau} g)_0 &= 0, \\
(S_{0}^{h, \tau} g)_{j+1} - (S_{0}^{h, \tau} g)_j &= \tau A_h (S_{0}^{h, \tau} g)_{j+1} + \int_{t_j}^{t_{j+1}} Q_h g(t) dt, \quad 0 \leq j < J. 
\end{aligned} \quad (3.1)$$

Define $G_{h, \tau} \in \mathcal{X}_{h, \tau}$, an approximation of $G$, by

$$\begin{aligned}
(G_{h, \tau})_0 &= 0, \\
(G_{h, \tau})_{j+1} - (G_{h, \tau})_j &= \tau A_h (G_{h, \tau})_{j+1} + \int_{t_j}^{t_{j+1}} Q_h R \sigma(t) dW(t), \quad 0 \leq j < J. 
\end{aligned} \quad (3.2)$$

**Remark 3.1.** For any $g \in L^2(\Omega; L^2(0, T; L^2(\Gamma)))$, a routine energy argument (see [35, Chapter 12] and [30]) yields that

$$\max_{0 \leq j \leq J} \| (S_{0}^{h, \tau} R g)_j \|_{H^0} \leq C \| g \|_{L^2(0, T; L^2(\Gamma))}, \quad (3.3)$$

where $C$ is a positive constant depending only on $\mathcal{O}$.

**Remark 3.2.** Although $\mathcal{X}_{h, \tau} \not\subset L^2_F(\Omega; L^2(0, T; \dot{H}^0))$, we have

$$G_{h, \tau} \in L^2_F(\Omega; L^2(0, T; \mathcal{V}_h)).$$

Moreover, $S_{0}^{h, \tau} g \in L^2_F(\Omega; L^2(0, T; \mathcal{V}_h))$ for all $g \in L^2_F(\Omega; L^2(0, T; \dot{H}^{-1}))$.

The discrete stochastic optimal control problem is to seek a discrete control $\bar{U} \in U_{ad, \tau}$ such that

$$J_{h, \tau}(\bar{U}) = \min_{U \in U_{ad, \tau}} J_{h, \tau}(U), \quad \text{ (3.4)}$$

where the discrete admissible control space is defined by

$$U_{ad, \tau} : = \{ U \in U_{ad} \mid U \text{ is constant on } [t_j, t_{j+1}], \forall 0 \leq j < J \}$$

and the discrete cost functional is defined by

$$J_{h, \tau}(U) := \frac{1}{2} \| S_{0}^{h, \tau} RU + G_{h, \tau} - y \|_{L^2(0, T; H^0)}^2 + \nu \| U \|_{L^2(0, T; L^2(\Gamma))}^2, \quad \forall U \in U_{ad, \tau}.$$
Remark 3.3. Note that the discrete control is not discretized in space. This is just for the sake of numerical analysis, and there is no essential difficulty in extending the numerical analysis of this paper to the case where the discrete control is discretized by the standard continuous piecewise linear element method in space.

Then let us present the first-order optimality condition for the above discrete problem. For any \( g \in L^2(\Omega; L^2(0, T; \tilde{H}^{-1})) \), define \( S^h_{1, \tau} g \in X_{h, \tau} \) by

\[
\begin{aligned}
(S^h_{1, \tau} g)_j &= 0, \\
(S^h_{1, \tau} g)_j - (S^h_{1, \tau} g)_{j+1} &= \tau A_h(S^h_{1, \tau} g)_j + \int_{t_j}^{t_{j+1}} Q_h g(t) dt, \quad 0 \leq j < J.
\end{aligned}
\]  

(3.5)

In view of (3.1) and (3.5), a straightforward calculation yields that for any \( f \in L^2(\Omega; L^2(0, T; L^2(\Gamma))) \) and \( g \in L^2(\Omega; L^2(0, T; \tilde{H}^0)) \),

\[
\int_0^T [S^h_{0, \tau} R f, g] dt = \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (f(t), (S^h_{1, \tau} g)_{j+1}) dt.
\]  

(3.6)

Using this equality, we readily conclude that

\[
\sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (\tilde{P}_{j+1} + \nu \tilde{U}, U - \tilde{U}) dt \geq 0 \quad \text{for all } U \in U^h_{ad},
\]  

(3.7)

where \( \tilde{U} \) is the unique solution of the problem (3.4) and

\[
\tilde{P} := S^h_{1, \tau} (S^h_{0, \tau} R \tilde{U} + G_{h, \tau} - y_d).
\]  

(3.8)

It is evident from (3.7) that for any \( 0 \leq j < J \),

\[
\tilde{U}_j = \mathcal{P}_{[u_0, u^*]} \left( -\frac{1}{\rho} E_{t_j} \text{tr} \tilde{P}_{j+1} \right).
\]

Remark 3.4. Let \( \tilde{U} \) be the solution to the problem (3.4), and let

\[
\tilde{Y} := S^h_{0, \tau} R \tilde{U} + G_{h, \tau}, \quad \tilde{P} := S^h_{1, \tau} (\tilde{Y} - y_d).
\]  

Assume that there exist deterministic functions \( \mathcal{U} \) and \( \mathcal{P} \) such that

\[
\tilde{U} = \mathcal{U}(\tilde{Y}) \quad \text{and} \quad \tilde{P} = \mathcal{P}(\tilde{Y}).
\]  

(3.9)

Assume that \( \sigma(t) = 0 \) for all \( 0 \leq t \leq T \). From the fact that \( \tilde{U}(t_0) \) is deterministic, it follows that \( \tilde{Y}(t_1) \) is deterministic. Hence, according to the assumption, \( \tilde{U}(t_1) \) and \( \tilde{P}(t_1) \) are deterministic. Similarly, \( \tilde{Y}(t_j) \) and \( \tilde{P}(t_j) \) are deterministic for all \( 2 \leq j \leq J \). It follows that \( \tilde{Y}, \tilde{U} \) and \( \tilde{P} \) are deterministic. However, in general this is evidently incorrect, since \( y_d \) is a stochastic process. The above example illustrates that for our model problem, generally we cannot expect the existence of deterministic functions \( \mathcal{U} \) and \( \mathcal{P} \) such that (3.9) holds. Since (3.9) is crucial for the application of the algorithms in [10] to the problem (3.4), these algorithms are not applicable.

Finally, we present the convergence of the discrete stochastic optimal control problem (3.4). For any \( v \in L^2_\mathcal{F}(\Omega; L^2(0, T; X)) \) with \( X \) being a separable Hilbert space, define \( \mathcal{P}_\tau v \) by

\[
(\mathcal{P}_\tau v)(t) = \frac{1}{\tau} E_{t_j} \int_{t_j}^{t_{j+1}} v(s) ds
\]  

(10.30)

for all \( t_j \leq t < t_{j+1} \) with \( 0 \leq j < J \).
We also have the following two well-known estimates (see [4, Theorems 4.4.4, 4.5.11 and 5.7.6]):

For each $\sigma \in L^2_\varrho(\Omega; L^2(0, T; \hat{H}^0))$ and $\sigma \in L^2_\varrho(\Omega; L^2(0, T; \hat{L}^0)) \cap L^\infty(0, T; L^2(\Omega; \hat{L}_2^0))$.

Let $\bar{u}$ be the solution of (1.1), $p$ be defined by (2.16), and $\bar{U}$ be the solution of (3.4). Then

$$
\|\bar{u} - \bar{U}\|_{L^2(0, T; \hat{H}^0)} + \|S_0 R \bar{u} - S_0 h^{-1} R \bar{U}\|_{L^2(0, T; \hat{H}^0)} \leq C(\tau^{1/4}\epsilon + h^{1/2}\epsilon + \|I - P_\tau\|_{\hat{L}^2(0, T; \hat{L}^2(\Gamma))} + \|I - P_\tau\|_{\hat{L}^2(\Omega; \hat{L}^2(\Gamma))}).
$$

Moreover, if $F$ is the natural filtration of $W(\cdot)$, then

$$
\|\bar{u} - \bar{U}\|_{L^2(0, T; \hat{H}^0)} + \|S_0 R \bar{u} - S_0 h^{-1} R \bar{U}\|_{L^2(0, T; \hat{H}^0)} \leq C(\tau^{1/4}\epsilon + h^{1/2}\epsilon).
$$

The above $\epsilon$ is a sufficiently small positive number and the above $C$ is a positive constant depending only on $u_\ast$, $u^\ast$, $y_d$, $\sigma$, $\nu$, $\mathcal{O}$, $T$ and the regularity parameters of $K_h$.

4 Proofs

Throughout this section, we use the following conventions: $\epsilon > 0$ denotes a sufficiently small number, and its value may differ in different places; $a \lesssim b$ means $a \leq C b$, where $C$ is a positive constant depending only on $\epsilon$, $\mathcal{O}$, $T$, the regularity parameters of $K_h$ and the indexes of the Sobolev spaces where the underlying functions belong.

4.1 Preliminary estimates

For each $\gamma \in \mathbb{R}$, let $\hat{H}^\gamma_h$ be the space $\mathcal{V}_h$ endowed with the norm

$$
\|v_h\|_{\hat{H}^\gamma_h} := \|(-A_h)^{\gamma/2} v_h\|_{\hat{H}^\gamma}, \quad \forall v_h \in \mathcal{V}_h.
$$

By the fact that $R \in \mathcal{L}(L^2(\Gamma), \hat{H}^{-1/2-\epsilon})$ and

$$
\|Q_h\|_{\mathcal{L}(\hat{H}^{-1/2-\epsilon}, \hat{H}^{-1/2-\epsilon})} \lesssim 1,
$$

we obtain

$$
\|Q_h R\|_{\mathcal{L}(L^2(\Gamma), \hat{H}^{1/2+\epsilon})} \lesssim 1.
$$

We also have the following two well-known estimates (see [4, Theorems 4.4.4, 4.5.11 and 5.7.6]):

$$
\|I - Q_h\|_{\mathcal{L}(\hat{H}^1, \hat{H}^0)} \lesssim h,
$$

$$
\|I - A_h^{-1} Q_h A\|_{\mathcal{L}(\hat{H}^1, \hat{H}^0)} \lesssim h.
$$

Remark 4.1. For any $v \in \hat{H}^{-1}$ we have

$$
\|Q_h v\|_{\hat{H}^{-1}_h} = \sup_{0 \neq \varphi_h \in \mathcal{V}_h} \frac{(Q_h v, \varphi_h)_\mathcal{O}}{\|\varphi_h\|_{\hat{H}^{-1}_h}} = \sup_{0 \neq \varphi_h \in \mathcal{V}_h} \frac{\langle v, \varphi_h \rangle_{\hat{H}^{-1}_h}}{\|\varphi_h\|_{\hat{H}^{-1}_h}} \quad \text{(by the definition of } Q_h) \leq \sup_{0 \neq \varphi_h \in \mathcal{V}_h} \frac{\|v\|_{\hat{H}^{-1}} \|\varphi_h\|_{\hat{H}^{1}_h}}{\|\varphi_h\|_{\hat{H}^{1}_h}}.
$$

Noting the fact that $\|\varphi_h\|_{\hat{H}^{1}_h} = \|\varphi_h\|_{\hat{H}^{1}_h}, \forall \varphi_h \in \mathcal{V}_h$, we readily conclude that

$$
\|Q_h v\|_{\mathcal{L}(\hat{H}^{-1}, \hat{H}^{-1}_h)} \leq 1.
$$

In addition, by definition, the restriction of $Q_h$ to $\hat{H}^0$ is the $L^2(\mathcal{O})$-orthogonal projection onto $\mathcal{V}_h$, and hence,

$$
\|Q_h\|_{\mathcal{L}(\hat{H}^0, \hat{H}^0_h)} \leq 1.
$$

Using the above two estimates, by interpolation (see [27, Theorems 2.6 and 4.36]), we readily obtain (4.1).
By the techniques in the proofs of [35, Lemmas 3.2 and 7.3], a straightforward computation yields the following lemma.

**Lemma 4.2.** For any $-2 \leq \beta \leq \gamma \leq 2$ and $t > 0$, we have
\[
\|e^{tA_h}\|_{L(H_h^\beta, H_h^\gamma)} \lesssim t^{(\beta-\gamma)/2}.
\]
(4.5)

For any $0 \leq \beta \leq \gamma \leq 2$ and $t > 0$,
\[
\|I - e^{tA_h}\|_{L(H_h^\beta, H_h^\gamma)} \lesssim t^{(\gamma-\beta)/2}.
\]
(4.6)

For any $0 \leq \gamma \leq 2$ and $j > 1$,
\[
\|(I - \tau A_h)^{-j}\|_{L(H_h^\gamma, H_h^\gamma)} < (j\tau)^{-\gamma/2}.
\]
(4.7)

**Lemma 4.3.** Assume that $0 \leq \gamma \leq 1$, $0 \leq k \leq j < J$ and $t_k \leq t < t_{k+1}$. Then
\[
\|e^{(t_{j+1}-t)A_h} - e^{(t_{j+1}-t_k)A_h}\|_{L(H_h^{\gamma}, H_h^{\gamma})} \lesssim \tau^{(1-\gamma)/2}((t_j+1-t) - t)^{-1/2}.
\]
(4.8)

**Proof.** We have
\[
\|e^{(t_{j+1}-t)A_h} - e^{(t_{j+1}-t_k)A_h}\|_{L(H_h^{\gamma}, H_h^{\gamma})}
= \|(I - e^{(t-t_k)A_h})e^{(t_{j+1}-t)A_h}\|_{L(H_h^{\gamma}, H_h^{\gamma})}
\leq \|I - e^{(t-t_k)A_h}\|_{L(H_h^{\gamma}, H_h^{\gamma})}\|e^{(t_{j+1}-t)A_h}\|_{L(H_h^{\gamma}, H_h^{\gamma})}.
\]
so (4.8) follows from (4.5) and (4.6).

**Lemma 4.4.** For any $0 \leq \gamma \leq 2$ and $1 \leq j \leq J$,
\[
\|e^{j\tau A_h} - (I - \tau A_h)^{-j}\|_{L(H_h^{\gamma}, H_h^{\gamma})} \lesssim \tau^{-\gamma/2}j^{-1}.
\]
(4.9)

**Proof.** By [35, Theorem 7.2], we have
\[
\|e^{j\tau A_h} - (I - \tau A_h)^{-j}\|_{L(H_h^{\gamma}, H_h^{\gamma})} \lesssim j^{-1}.
\]

Also, by (4.5) and (4.7), we have
\[
\|e^{j\tau A_h} - (I - \tau A_h)^{-j}\|_{L(H_h^{\gamma}, H_h^{\gamma})} \lesssim (j\tau)^{-1}.
\]

Hence, by interpolation (see [29, Theorems 2.6 and 4.36]), we obtain (4.9).

4.2 Convergence of $G_{h,\tau}$

Let $G_h$ be the mild solution of the equation
\[
\begin{cases}
    dG_h(t) = A_h G_h(t) dt + Q_h R \sigma(t) dW(t), & 0 \leq t \leq T, \\
    G_h(0) = 0.
\end{cases}
\]
(4.10)

Similar to (2.11), we see that for each $0 \leq t \leq T$,
\[
G_h(t) = \int_0^t e^{(t-s)A_h} Q_h R \sigma(s) dW(s) \quad \text{P-a.s.}
\]
(4.11)

Let us first estimate $\|G - G_h\|_{L^2(0,T;H^0)}$.

**Lemma 4.5.** If $\sigma \in L^2_s(\Omega; L^2(0,T; L^2_0))$, then
\[
\|G - G_h\|_{L^2(0,T;H^0)} \lesssim h^{1/2-\epsilon}\|\sigma\|_{L^2(0,T;L^2_0)}.
\]
(4.12)
Proof. For any $g \in L^2_0(\Omega; L^2(0, T; L_2(\tilde{A}, \tilde{H}^{-1})))$, let $\Pi g$ and $\Pi_h g$ be the mild solutions of the equations
\[
\begin{aligned}
    d(\Pi g)(t) &= A(\Pi g)(t)dt + g(t)dW(t), \quad 0 \leq t \leq T, \\
    (\Pi g)(0) &= 0 
\end{aligned}
\]
and
\[
\begin{aligned}
    d(\Pi_h g)(t) &= A_h(\Pi_h g)(t)dt + Q_h g(t)dW(t), \quad 0 \leq t \leq T, \\
    (\Pi_h g)(0) &= 0, 
\end{aligned}
\]
respectively. By the celebrated Itô’s formula, we have the following standard estimates:
\[
\begin{aligned}
    \|\Pi g\|_{L^2(0, T; H^{\alpha+1})} &\lesssim \|g\|_{L^2(0, T; L_2(\tilde{A}, H^{\alpha}))}, \\
    \|\Pi_h g\|_{L^2(0, T; H^{\alpha+1})} &\lesssim \|Q_h g\|_{L^2(0, T; L_2(\tilde{A}, H^{\alpha}))} 
\end{aligned}
\]
for any $g \in L^2_0(\Omega; L^2(0, T; L_2(\tilde{A}, \tilde{H}^{\alpha})))$ with $\alpha \geq -1$. Using the above estimates with $\alpha = -1$ and the fact that $\|Q_h\|_{L(H^{-1}, H^{\alpha})} \leq 1$, we then obtain
\[
\|\Pi - \Pi_h\|_{L^2(\tilde{A}(\Omega; L^2(0, T; L_2(\tilde{A}, \tilde{H}^{-1}))), H^{\alpha})} \lesssim 1. 
\]
Let $g \in L^2_0(\Omega; L^2(0, T; L_2(\tilde{A}, \tilde{H}^{0})))$. For any $0 \leq t \leq T$, we have (see [13, Chapter 3])
\[
(\Pi g)(t) = \int_0^t A(\Pi g)(s)ds + \int_0^t g(s)dW(s) \quad \text{P-a.s.,}
\]
so
\[
(Q_h \Pi g)(t) = \int_0^t Q_h A(\Pi g)(s)ds + \int_0^t Q_h g(s)dW(s) \quad \text{P-a.s.}
\]
It follows that
\[
\begin{aligned}
    d(Q_h \Pi g)(t) &= Q_h A(\Pi g)(t)dt + Q_h g(t)dW(t), \quad 0 \leq t \leq T, \\
    (Q_h \Pi g)(0) &= 0. 
\end{aligned}
\]
Letting $e_h := \Pi_h g - Q_h \Pi g$, by (4.13) and (4.17), we get
\[
\begin{aligned}
    de_h(t) &= A_h e_h(t)dt + (A_h Q_h g - A_h \Pi g)(t)dt, \quad 0 \leq t \leq T, \\
    e_h(0) &= 0. 
\end{aligned}
\]
It follows that
\[
\|e_h\|_{L^2(0, T; H^{\alpha})} \lesssim \|(A_h Q_h - A_h \Pi g)\|_{L^2(0, T; H^{-2})} = \|(Q_h - A_h^{-1} Q_h A) \Pi g\|_{L^2(0, T; H^\alpha)},
\]
so
\[
\|(\Pi - \Pi_h) g\|_{L^2(0, T; H^{\alpha})}
\leq \|e_h\|_{L^2(0, T; H^{\alpha})} + \|(I - Q_h) \Pi g\|_{L^2(0, T; H^{\alpha})} 
\lesssim \|(Q_h - A_h^{-1} Q_h A) \Pi g\|_{L^2(0, T; H^{\alpha})} + \|(I - Q_h) \Pi g\|_{L^2(0, T; H^{\alpha})} 
\lesssim \|(I - A_h^{-1} Q_h A) \Pi g\|_{L^2(0, T; H^{\alpha})} + \|(I - Q_h) \Pi g\|_{L^2(0, T; H^{\alpha})} 
\lesssim h \|\Pi g\|_{L^2(0, T; H^{\alpha})} \quad \text{by (4.3) and (4.4))} 
\lesssim h \|g\|_{L^2(0, T; L_2(\tilde{A}, \tilde{H}^{0})))} \quad \text{(by (4.14) with $\alpha = 0$)}.
\]
This implies that
\[
\|\Pi - \Pi_h\|_{L^2(\tilde{A}(\Omega; L^2(0, T; L_2(\tilde{A}, \tilde{H}^{0}))), H^{\alpha})} \lesssim h.
\]
In view of (4.16), (4.19) and Lemma A.1, by interpolation we obtain
\[
\|\mathbb{I} - \Pi_h\|_{\mathcal{L}(L^2(\Omega;L^2(0,T;L^2_h(t_k^2,\hat{H}^{-1/2-\epsilon})))},H^0) \lesssim h^{1/2-\epsilon}.
\]
Therefore, the desired estimate (4.12) follows from the following facts:
\[
\begin{align*}
G &= \Pi R\sigma, \\
G_h &= \Pi_h R\sigma, \\
\mathcal{R} &\in \mathcal{L}(L^2(\Gamma),\hat{H}^{-1/2-\epsilon}), \\
\sigma &\in L^2(\Omega;L^2(0,T;L^2(0,T;L^2_h(0,T;L^2)))).
\end{align*}
\]
This completes the proof. \(\square\)

Then let us estimate \(\|G_h - G_h,\tau\|_{L^2(0,T;H^0_h)}\).

**Lemma 4.6.** If \(\sigma \in L^2(\Omega;L^2(0,T;L^2_h(0,T;L^2_h(0,T;L^2)))\), then
\[
\max_{0 \leq j \leq J} \| (G_h - G_h,\tau)(t_j) \|_{H^0_h} \lesssim \tau^{1/4-\epsilon} \| \sigma \|_{L^\infty(0,T;L^2(\Omega;L^2_h(0,T;L^2)))}.
\]

**Proof.** Let \(0 < j < J\) be arbitrary but fixed. By (3.2) and (4.11), we obtain P-a.s.
\[
(G_h - G_h,\tau)(t_{j+1}) = \sum_{k=0}^j \int_{t_k}^{t_{k+1}} (e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)} - 1) Q_h \mathcal{R}\sigma(t) dW(t)
\]
so that
\[
\begin{align*}
\| (G_h - G_h,\tau)(t_{j+1}) \|_{H^0_h}^2 &\leq \sum_{k=0}^j \left\| \int_{t_k}^{t_{k+1}} (e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)+1}) Q_h \mathcal{R}\sigma(t) dW(t) \right\|_{H^0_h}^2 \\
&\leq \sum_{k=0}^j \int_{t_k}^{t_{k+1}} \| e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)+1} \|_{\mathcal{L}(H^{-1/2-\epsilon},H^0_h)}^2 \\
&\quad \times \| Q_h \mathcal{R}\sigma(t) \|_{\mathcal{L}(L^2(\Gamma),H^{-1/2-\epsilon})}^2 \| \sigma(t) \|_{L^2}^2 dt.
\end{align*}
\]

By (4.2), we then conclude that
\[
\| (G_h - G_h,\tau)(t_{j+1}) \|_{H^0_h}^2 \lesssim \| \sigma \|_{L^\infty(0,T;L^2(\Omega;L^2_h(0,T;L^2)))}^2 \sum_{k=0}^j \int_{t_k}^{t_{k+1}} \| e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)+1} \|_{\mathcal{L}(H^{-1/2-\epsilon},H^0_h)}^2 dt.
\]

By (4.5) and (4.7), we have
\[
\int_{t_j}^{t_{j+1}} \| e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)+1} \|_{\mathcal{L}(H^{-1/2-\epsilon},H^0_h)}^2 dt \lesssim \tau^{1/2-\epsilon},
\]
and by (4.8) and (4.9), we have
\[
\begin{align*}
\sum_{k=0}^{j-1} \int_{t_k}^{t_{k+1}} \| e^{(t_{j+1}-t)A_h} - (I - \tau A_h)^{-(j-k)+1} \|_{\mathcal{L}(H^{-1/2-\epsilon},H^0_h)}^2 dt \\
&\lesssim \sum_{k=0}^{j-1} \int_{t_k}^{t_{k+1}} \| e^{(t_{j+1}-t)A_h} - e^{(t_{j+1}-t_k)A_h} \|_{\mathcal{L}(H^{-1/2-\epsilon},H^0_h)}^2 dt
\end{align*}
\]
Hence, by (4.21) we get

\[ \tau^{1/2-\varepsilon} \sum_{k=0}^{j-1} \int_{t_k}^{t_{k+1}} (t_{j+1} - t)^{-1} dt + \tau^{1/2-\varepsilon} \sum_{k=0}^{j-1} (j - k + 1)^{-1} \]

\[ \lesssim \tau^{1/2-\varepsilon} (\ln(1/\tau) + \ln(j + 1)) \]

Combining the above two estimates yields

\[ \sum_{k=0}^{j} \int_{t_k}^{t_{k+1}} \|e^{(t_{j+1} - t)A_h} - (I - \tau A_h)^{(j-k+1)}}^{2} \|_{L(L^2(\Omega; L^2(\mathbb{C}^2)))} \lesssim \tau^{1/2-\varepsilon}. \]

Hence, by (4.21) we get

\[ \|G_h - G_{h,\tau})(t_{j+1})\|_{H^0_h(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))} \lesssim \tau^{1/4-\varepsilon} \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}. \]

Hence, as 0 \leq j < J is arbitrary, (4.20) follows from the fact that

\[ (G_h - G_{h,\tau})(0) = 0. \]

This completes the proof.

**Lemma 4.7.** Under the condition of Lemma 4.6, we have

\[ \|G_h - G_{h,\tau}\|_{L^2(0, T; H^0_h(\Omega; L^2(\mathbb{C}^2)))} \lesssim \tau^{1/4-\varepsilon} \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}. \]

**Proof.** By (4.20), we only need to prove

\[ \sum_{j=0}^{J-1} \|G_h - G_{h}(t_j)\|_{L^2(t_j, t_{j+1}; H^0_h)}^2 \lesssim \tau^{1/2-\varepsilon} \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}. \]

(4.23)

To this end, we proceed as follows. For any \( t_j \leq t \leq t_{j+1} \) with 0 \leq j < J, we have

\[ G_h(t) - e^{(t-t_j)A_h} G_h(t_j) = \int_{t_j}^{t} e^{(t-s)A_h} Q_h \mathcal{R} \sigma(s) dW(s) \quad \text{P-a.s.,} \]

so

\[ \|G_h(t) - e^{(t-t_j)A_h} G_h(t_j)\|_{H^0_h}^2 \]

\[ = \int_{t_j}^{t} \|e^{(t-s)A_h} Q_h \mathcal{R} \sigma(s)\|_{L^2(L^2(\mathbb{C}^2))}^2 ds \]

\[ \leq \int_{t_j}^{t} \|e^{(t-s)A_h}\|_{L(L^2(\mathbb{C}^2))}^2 \|Q_h \mathcal{R}\|_{L^2(\mathbb{C}^2)}^2 \|\sigma(s)\|_{L^2(\mathbb{C}^2)}^2 ds \]

\[ \lesssim \int_{t_j}^{t} (t - s)^{-1/2-\varepsilon} ds \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}^2 \] (by (4.2) and (4.5))

\[ \lesssim (t - t_j)^{1/2-\varepsilon} \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}^2. \]

It follows that

\[ \int_{t_j}^{t_{j+1}} \|G_h(t) - e^{(t-t_j)A_h} G_h(t_j)\|_{H^0_h}^2 dt \lesssim \tau^{3/2-\varepsilon} \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}^2. \]

(4.24)

Following the proof of (2.13), we obtain

\[ \|G_h(t_j)\|_{H^{1/2-\varepsilon}} \lesssim \|\sigma\|_{L^\infty(0, T; L^2(\Omega; L^2(\mathbb{C}^2)))}. \]
so that
\[
\int_{t_j}^{t_{j+1}} \left\| (I - e^{(t-t_j)A_h})G_h(t_j) \right\|^2_{H^p_h} dt \\
\lesssim 
\int_{t_j}^{t_{j+1}} (t - t_j)^{1/2-\epsilon} \left\| G_h(t_j) \right\|^2_{H^{1/2-h}} dt \quad \text{(by (4.6))}
\lesssim \tau^{3/2-\epsilon} \left\| G_h(t_j) \right\|^2_{H^{1/2-h}} \\
\lesssim \tau^{3/2-\epsilon} \| \sigma \|^2_{L^\infty(0,T;L^2(\Omega;L^2_\gamma))}.
\]

Combining the above estimate with (4.24) yields
\[
\left\| G_h - G_h(t_j) \right\|^2_{L^2(t_j,t_{j+1};H^p_h)} \lesssim \tau^{3/2-\epsilon} \| \sigma \|^2_{L^\infty(0,T;L^2(\Omega;L^2_\gamma))},
\]
and hence,
\[
\sum_{j=0}^{J-1} \left\| G_h - G_h(t_j) \right\|^2_{L^2(t_j,t_{j+1};H^p_h)} \lesssim \sum_{j=0}^{J-1} \tau^{3/2-\epsilon} \| \sigma \|^2_{L^\infty(0,T;L^2(\Omega;L^2_\gamma))} \\
\lesssim \tau^{1/2-\epsilon} \| \sigma \|^2_{L^\infty(0,T;L^2(\Omega;L^2_\gamma))}.
\]

This proves (4.23) and thus completes the proof. \(\square\)

Finally, combining Lemmas 4.5 and 4.7, we readily conclude the following error estimate.

**Lemma 4.8.** If \( \sigma \in L^2_h(\Omega;L^2(0,T;L^2_\gamma)) \cap L^\infty(0,T;L^2(\Omega;L^2_\gamma)) \), then
\[
\| G - G_h,\tau \|_{L^2(0,T;H^0)} \lesssim (\tau^{1/4-\epsilon} + h^{1/2-\epsilon}) \| \sigma \|_{L^\infty(0,T;L^2(\Omega;L^2_\gamma))}.
\] (4.25)

### 4.3 Convergence of \( S_0^{h,\tau} \) and \( S_1^{h,\tau} \)

**Lemma 4.9.** For any \( g_h \in L^2(0,T;V_h) \), define \( \{ Y_j \}_{j=0}^J \subset V_h \) by

\[
\begin{dcases}
Y_0 = 0, \\
Y_{j+1} - Y_j = \tau A_h Y_j + \int_{t_j}^{t_{j+1}} g_h(t) dt, \quad 0 \leq j < J.
\end{dcases}
\] (4.26)

Then for any \( 0 \leq \gamma \leq 1 \),
\[
\left( \sum_{j=0}^{J-1} \| Y_{j+1} - Y_j \|^2_{H^p_h} \right)^{1/2} \lesssim \tau (1-\gamma)^{1/2} \| g_h \|_{L^2(0,T;H^{-\gamma}_h)}.
\] (4.27)

**Proof.** Following the proof of [30, Theorem 4.6], we obtain
\[
\left( \sum_{j=0}^{J-1} \| Y_{j+1} - Y_j \|^2_{H^p_h} \right)^{1/2} \lesssim \tau^{1/2} \| g_h \|_{L^2(0,T;H^{-\gamma}_h)},
\] (4.28)
\[
\left( \sum_{j=0}^{J-1} \tau \| Y_j \|^2_{H^p_h} \right)^{1/2} \lesssim \| g_h \|_{L^2(0,T;H^{-\gamma}_h)}.
\] (4.29)

From (4.29), it follows that
\[
\left( \sum_{j=0}^{J-1} \| Y_{j+1} - Y_j \|^2_{H^p_h} \right)^{1/2} \lesssim \tau^{-1/2} \| g_h \|_{L^2(0,T;H^{-\gamma}_h)}.
\] (4.30)

In view of (4.28) and (4.30), by interpolation we obtain (4.27). This completes the proof. \(\square\)
Lemma 4.10. For any \( g \in L^2_2(\Omega; L^2(0, T; L^2(\Gamma))) \), we have
\[
\|(S_0 - S_0^{h, \tau})Rg\|_{L^2(0, T; H^0)} \lesssim (\tau^{3/4 - \epsilon} + h^{3/2 - \epsilon})\|g\|_{L^2(0, T; L^2(\Gamma))}. \tag{4.31}
\]

Proof. A routine energy argument (see [35, Chapter 12] and [30, Theorems 5.1 and 5.5]) gives that
\[
\left( \sum_{j=0}^{J-1} \left\| S_0 Rg - (S_0^{h, \tau} Rg)_{j+1} \right\|_{L^2(t_j, t_{j+1}; H^0)}^2 \right)^{1/2} \lesssim (\tau^{3/4 - \epsilon} + h^{3/2 - \epsilon})\|g\|_{L^2(0, T; L^2(\Gamma))}.
\]

We also have
\[
\left( \sum_{j=0}^{J-1} \left\| (S_0^{h, \tau} Rg)_{j+1} - (S_0^{h, \tau} Rg)_{j+1} \right\|_{L^2(t_j, t_{j+1}; H^0)}^2 \right)^{1/2} \lesssim \tau^{1/4 - \epsilon} \|Q_n \|_{L^2(0, T; H^{-1/2, -\epsilon})} \quad \text{(by (4.27))}
\]
\[
\lesssim \tau^{1/4 - \epsilon} \|g\|_{L^2(0, T; L^2(\Gamma))} \quad \text{(by (4.2)).}
\]

Hence,
\[
\|(S_0 - S_0^{h, \tau}) Rg\|_{L^2(0, T; H^0)}
= \left( \sum_{j=0}^{J-1} \left\| S_0 Rg - (S_0^{h, \tau} Rg)_{j+1} \right\|_{L^2(t_j, t_{j+1}; H^0)}^2 \right)^{1/2}
\leq \left( \sum_{j=0}^{J-1} \left\| (S_0 Rg - (S_0^{h, \tau} Rg)_{j+1} \right\|_{L^2(t_j, t_{j+1}; H^0)}^2 \right)^{1/2} + \left( \sum_{j=0}^{J-1} \tau \| (S_0^{h, \tau} Rg)_{j+1} - (S_0^{h, \tau} Rg)_{j+1} \|_{H^0}^2 \right)^{1/2}
\lesssim (\tau^{3/4 - \epsilon} + h^{3/2 - \epsilon})\|g\|_{L^2(0, T; L^2(\Gamma)).}
\]

This proves (4.31) and thus completes the proof. \(\square\)

Lemma 4.11. For any \( g \in L^2_2(\Omega; L^2(0, T; H^0)) \), we have
\[
\left( \sum_{j=0}^{J-1} \left\| S_1 g - (S_1^{h, \tau} g)_{j+1} \right\|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2} \lesssim (\tau^{3/4 - \epsilon} + h^{3/2 - \epsilon})\|g\|_{L^2(0, T; H^0)}. \tag{4.32}
\]

Since this lemma can be proved by using similar techniques to those in the proof of Lemma 4.10, its proof is omitted here.

4.4 Proof of Theorem 3.5

Let \( \{ \varphi_n : n \in \mathbb{N} \} \subset H^2 \) be an orthonormal basis of \( H^0 \) such that (see [16, Theorem 3.2.1.3] and [37, Theorem 4.A and Subsection 4.5])
\[-A \varphi_n = r_n \varphi_n \quad \text{for each } n \in \mathbb{N},
\]
where \( \{ r_n : n \in \mathbb{N} \} \) is a nondecreasing sequence of strictly positive numbers with limit \(+\infty\). Let \( \tilde{W}(t) \) \((t \geq 0)\) be a cylindrical Wiener process in \( L^2(\Omega) \) defined by
\[
\tilde{W}(t)(v) = \sum_{n=0}^{\infty} \langle v, \varphi_n \rangle_{\mathcal{O}} \beta_n(t) \quad \text{for all } t \geq 0 \text{ and } v \in H^0. \tag{4.33}
\]

Under the condition that \( F \) is the natural filtration of \( W(\cdot) \), for any \( g \in L^2_2(\Omega; L^2(0, T; H^0)) \), the backward stochastic parabolic equation
\[
\left\{ \begin{array}{l}
 dp(t) = -(Ap + g)(t)dt + z(t)d\tilde{W}(t), \quad \forall 0 \leq t \leq T, \\
p(T) = 0
\end{array} \right. \tag{4.34}
\]
admits a unique strong solution \((p, z)\), and
\[
\|p\|_{L^2(0,T;H^\gamma)} + \|p\|_{C([0,T];H^{1})} + \|z\|_{L^2(0,T;\mathcal{L}_2(\bar{H}^\alpha,H^\gamma))} \leq C \|g\|_{L^2(0,T,H^n)}, \tag{4.35}
\]
where \(C\) is a positive constant independent of \(g\) and \(T\). Moreover, for any \(0 \leq s < t \leq T\),
\[
p(s) - e^{(t-s)A}p(t) = \int_s^t e^{(r-s)A}g(r)dr - \int_s^t e^{(t-r)A}z(r)d\bar{W}(r) \quad \text{P-a.s.} \tag{4.36}
\]
Using the fact \(p(T) = 0\), by (4.36) we obtain that for any \(0 \leq s < T\),
\[
p(s) = \int_s^T e^{(r-s)A}g(r)dr - \int_s^T e^{(t-r)A}z(r)d\bar{W}(r) \quad \text{P-a.s.} \tag{4.37}
\]

**Remark 4.12.** For the above theoretical results of (4.34), we refer the readers to [21] and [28, Chapter 4].

**Lemma 4.13.** Assume that \(F\) is the natural filtration of \(W(\cdot)\). For any \(g \in L^2_2(\Omega; L^2(0,T;H^0))\), we have
\[
\|I - \mathcal{P}_T\mathcal{E}_F S_1 g\|_{L^2(0,T;L^2(\Gamma))} \lesssim \tau^{1/2} \|g\|_{L^2(0,T;H^0)}. \tag{4.38}
\]

**Proof.** Let \((p, z)\) be the strong solution of (4.34). Since (4.37) implies that for each \(0 \leq t \leq T\),
\[
p(t) = \mathcal{E}_F S_1 g.
\]
by (2.5) we then obtain
\[
p = \mathcal{E}_F S_1 g. \tag{4.39}
\]
Therefore, it remains to prove
\[
\|I - \mathcal{P}_T\| p\|_{L^2(0,T;L^2(\Gamma))} \lesssim \tau^{1/2} \|g\|_{L^2(0,T;H^0)}. \tag{4.40}
\]
To this end, we proceed as follows. Let \(1/2 < \gamma < 1\) be arbitrary but fixed. For any \(t_j \leq t \leq t_{j+1}\) with \(0 \leq j < J\), by (4.36) we have
\[
p(t_j) - e^{(t-t_j)A}p(t_j) = \int_{t_j}^t e^{(s-t_j)A}g(s)ds - \int_{t_j}^t e^{(s-t_j)A}z(s)d\bar{W}(s) \quad \text{P-a.s.}
\]
so
\[
\|p(t_j) - e^{(t-t_j)A}p(t_j)\|_{H^\gamma}^2 \leq 2 \left\| \int_{t_j}^t e^{(s-t_j)A}g(s)ds \right\|_{H^\gamma}^2 + 2 \left\| \int_{t_j}^t e^{(s-t_j)A}z(s)d\bar{W}(s) \right\|_{H^\gamma}^2
\]
\[
= 2 \left\| \int_{t_j}^t e^{(s-t_j)A}g(s)ds \right\|_{H^\gamma}^2 + 2 \left\| \int_{t_j}^t e^{(s-t_j)A}z(s)ds \right\|^2_{\mathcal{L}_2(\bar{H}^\alpha,H^\gamma)} ds
\]
\[
\leq 2 \left( \int_{t_j}^t \left\| e^{(s-t_j)A}g(s) \right\|_{H^\gamma}^2 ds \right)^2 + 2 \left( \int_{t_j}^t \left\| e^{(s-t_j)A}z(s) \right\|^2_{\mathcal{L}_2(\bar{H}^\alpha,H^\gamma)} ds \right)^2
\]
\[
\lesssim (t - t_j)^{1-\gamma} \|g\|^2_{L^2(t_j,t;H^\gamma)} + \|z\|^2_{L^2(t_j,t;\mathcal{L}_2(\bar{H}^\alpha,H^\gamma))},
\]
It follows that for any \(0 \leq j < J\),
\[
\int_{t_j}^{t_{j+1}} \left\| p(t_j) - e^{(t-t_j)A}p(t_j) \right\|^2_{H^\gamma} dt \lesssim \tau^{2-\gamma} \|g\|^2_{L^2(t_j,t_{j+1};\bar{H}^\alpha)} + \|z\|^2_{L^2(t_j,t_{j+1};\mathcal{L}_2(\bar{H}^\alpha,H^\gamma))}.
\]
Hence,
\[
\sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \|p(t_j) - e^{(t-t_j)A} p(t)\|_{H^s}^2 dt \lesssim \tau^{2-\gamma} \|g\|_{L^2(0,T;H^0)}^2 + \tau \|\tau\|_{L^2(0,T;\mathcal{L}_2(H^0,H^\gamma))}^2.
\]
We also have
\[
\sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \|I - e^{(t-t_j)A} p(t)\|_{H^s}^2 dt \lesssim \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (t-t_j)^{2-\gamma} \|p(t)\|_{H^2}^2 dt \lesssim \tau^{2-\gamma} \|p\|_{L^2(0,T;H^2)},
\]
by the evident estimate (see (4.6))
\[
\|I - e^{tA}\|_{\mathcal{L}(H^2,H^\gamma)} \lesssim t^{1-\gamma/2}.
\]
Consequently,
\[
\sum_{j=0}^{J-1} \|p - p(t_j)\|_{L^2(t_j,t_{j+1};H^\gamma)}^2 \lesssim \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \|p(t_j) - e^{(t-t_j)A} p(t)\|_{H^s}^2 dt + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \|I - e^{(t-t_j)A} p(t)\|_{H^s}^2 dt \lesssim \tau^{2-\gamma} \|g\|_{L^2(0,T;H^0)}^2 + \tau \|\tau\|_{L^2(0,T;\mathcal{L}_2(H^0,H^\gamma))}^2 + \tau^{2-\gamma} \|p\|_{L^2(0,T;H^2)}^2 \lesssim \tau \|g\|_{L^2(0,T;H^0)}^2
\]
by (4.35). Therefore, from the inequality
\[
\|(I - P_\tau)p\|_{L^2(t_j,t_{j+1};H^\gamma)} \lesssim \|p - p(t_j)\|_{L^2(t_j,t_{j+1};H^\gamma)}, \quad \forall 0 \leq j < J,
\]
we conclude that
\[
\|(I - P_\tau)p\|_{L^2(0,T;H^\gamma)} \lesssim \tau^{1/2} \|g\|_{L^2(0,T;H^0)}.
\]
The desired estimate (4.40) then follows from the trace inequality (see [16, Theorem 1.5.1.2])
\[
\|v\|_{L^2(\Gamma)} \lesssim \|v\|_{H^\gamma}, \quad \forall v \in H^\gamma.
\]
This completes the proof.
\[\]
**Lemma 4.14.** For any \(v \in L^2(\Omega; L^2(0,T;H^1))\), we have
\[
\mathcal{E}_F \text{tr} v = \text{tr} \mathcal{E}_F v.
\]
**Proof.** Let \(\text{tr}^* : L^2(\Gamma) \rightarrow \hat{H}^{-1}\) be the adjoint of \(\text{tr}\). By definition, \(\mathcal{E}_F \text{tr} v\) is the \(L^2(\Omega; L^2(0,T;L^2(\Gamma)))\)-orthogonal projection of \(\text{tr} v\) onto \(L^2(\Omega; L^2(0,T;L^2(\Gamma)))\), and \(\mathcal{E}_F v\) is the \(L^2(\Omega; L^2(0,T;\hat{H}^1))\)-orthogonal projection of \(v\) onto \(L^2(\Omega; L^2(0,T;\hat{H}^1))\). For any \(\varphi \in L^2(\Omega; L^2(0,T;L^2(\Gamma)))\), we have
\[
\int_0^T \langle \mathcal{E}_F \text{tr} v, \varphi \rangle dt = \int_0^T \langle \text{tr} v, \varphi \rangle dt = \int_0^T E(\text{tr}^* \varphi, v)_{H^1} dt = \int_0^T E(\text{tr}^* \varphi, E_F v)_{H^1} dt = \int_0^T \langle \text{tr} \mathcal{E}_F v, \varphi \rangle dt,
\]
which implies (4.43).
\[\]
**Lemma 4.15.** Assume that \(F\) is the natural filtration of \(W(\cdot)\). For any \(g \in L^2(\Omega; L^2(0,T;H^0))\), we have
\[
\|(I - P_\tau)u\|_{L^2(0,T;L^2(\Gamma))} \lesssim \tau^{1/2} \nu^{-1} \|g\|_{L^2(0,T;H^0)},
\]
where
\[
u := P_{[u_\tau,u_\tau]}(-\nu^{-1} \mathcal{E}_F (\text{tr} S g)).
\]
Proof. Let \((p, z)\) be the solution of (4.34). As stated in the proof of Lemma 4.13, we have

\[ p = E_p S_t g \]

so that by (4.43) we get

\[ u = \mathcal{P}_{[u^*, u^*]}(-\nu^{-1} \text{tr} p). \]

From the fact that \( p \in L^2_0(\Omega; C([0, T]; \dot{H}^1)) \), we then conclude that

\[ u \in L^2_0(\Omega; C([0, T]; L^2(\Gamma))). \]

Moreover,

\[
\sum_{j=0}^{J-1} \left\| u - u(t_j) \right\|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \\
= \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \left\| \mathcal{P}_{[u^*, u^*]}(-\nu^{-1} \text{tr} p(t)) - \mathcal{P}_{[u^*, u^*]}(-\nu^{-1} \text{tr} p(t_j)) \right\|_{L^2(\Gamma)}^2 dt \\
\leq \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \left\| (p(t) - p(t_j))/\nu \right\|_{L^2(\Gamma)}^2 dt \\
\leq \tau \nu^{-2} \| g \|_{L^2(0, T; H^0)}^2,
\]

by (4.41) and the trace inequality

\[ \| v \|_{L^2(\Gamma)} \lesssim \| v \|_{H^\gamma}, \quad \forall \, v \in \dot{H}^\gamma. \]

Therefore, (4.44) follows from the inequality

\[
\left\| (I - \mathcal{P}_\tau)u \right\|_{L^2(0, T; L^2(\Gamma))} \leq \left( \sum_{j=0}^{J-1} \left\| u - u(t_j) \right\|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2}.
\]

This completes the proof. \(\square\)

Finally, we use the argument in the numerical analysis of optimal control problems with partial differential equation (PDE) constraints (see [20, Theorem 3.4]) to prove Theorem 3.5.

Proof of Theorem 3.5. In this proof, \(C\) is a positive constant depending only on \(u^*, \, u^*, \, y_d, \, \epsilon, \, \sigma, \, \nu, \, \mathcal{O}, \, T\) and the regularity parameters of \(K_h\), and its value may differ in different places. Let \(\hat{P}\) be defined by (3.8), and define

\[ P := S_1^{h, \tau}(S_0 \mathcal{R} \hat{u} + G - y_d). \]  

(4.45)

By (3.31) and the fact that \(\hat{u} \in U_{ad}\), we obtain

\[ \left\| (S_0 - S_0^{h, \tau}) \mathcal{R} \hat{u} \right\|_{L^2(0, T; H^0)} \leq C(\tau^{3/2} + h^{3/2}). \]  

(4.46)

By (2.8) and (2.13), we have

\[ \| S_0 \mathcal{R} \hat{u} + G - y_d \|_{L^2(0, T; H^0)} \leq C \]

so that (4.32) implies

\[
\left( \sum_{j=0}^{J-1} \left\| \hat{P} - P_{j+1} \right\|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2} \leq C(\tau^{3/2} + h^{3/2}).
\]

(4.47)

By the definition of \(\mathcal{P}_\tau\), we have the following two evident equalities:

\[ \int_0^T \langle \hat{U}, \mathcal{P}_\tau \hat{u} \rangle dt = \int_0^T \langle \hat{U}, \hat{u} \rangle dt, \]

(4.48)
\[
\| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))}^2 = \| \bar{u} - \mathcal{P}_r \bar{u} \|_{L^2(0,T;L^2(\Gamma))}^2 + \| \bar{U} - \mathcal{P}_r \bar{u} \|_{L^2(0,T;L^2(\Gamma))}^2.
\]

Since (3.12) follows from (3.11) and Lemmas 4.13 and 4.15, we only need to prove (3.11). The rest of this proof is divided into the following three steps.

**Step 1.** Let us prove that
\[
\nu \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))}^2 + \frac{1}{2} \| S_0 \bar{u} - S_0 \mathcal{P}_r \bar{u} \|_{L^2(0,T;H^\nu)}^2 \leq I_0 + I_1 + I_2,
\]
where
\[
I_0 := \frac{1}{2} \| S_0 \bar{u} - S_0 \mathcal{P}_r \bar{u} \|_{L^2(0,T;H^\nu)}^2,
\]
\[
I_1 := \int_0^T [G - G_{h,\tau}, S_0 \mathcal{P}_r (\bar{U} - \mathcal{P}_r \bar{u})] dt,
\]
\[
I_2 := \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{p} - E_{t_j} P_{j+1}, \bar{U} - \bar{u} \rangle dt.
\]

Inserting \( u := \bar{U} \) into (2.14) yields
\[
\int_0^T \langle \bar{p} + \nu \bar{u}, \bar{U} - \bar{u} \rangle dt \geq 0,
\]
which implies
\[
\nu \int_0^T \langle \bar{u}, \bar{U} - \bar{u} \rangle dt \leq \int_0^T \langle \bar{p}, \bar{U} - \bar{u} \rangle dt.
\]

Inserting \( U := \mathcal{P}_r \bar{u} \) into (3.7) gives
\[
\sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{P}_{j+1} + \nu \bar{U}, \mathcal{P}_r \bar{u} - \bar{U} \rangle dt \geq 0,
\]
which together with (4.48) implies
\[
- \nu \int_0^T \langle \bar{U}, \bar{u} - \bar{U} \rangle dt \leq \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{P}_{j+1}, \mathcal{P}_r \bar{u} - \bar{U} \rangle dt.
\]

From (4.54) and (4.55), we conclude that
\[
\nu \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))}^2 \leq \int_0^T \langle \bar{p}, \bar{U} - \bar{u} \rangle dt + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{P}_{j+1}, \mathcal{P}_r \bar{u} - \bar{U} \rangle dt
\]
\[
= \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{p} - E_{t_j} P_{j+1}, \bar{U} - \bar{u} \rangle dt + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle E_{t_j} P_{j+1}, \bar{U} - \bar{u} \rangle dt
\]
\[
+ \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{P}_{j+1}, \mathcal{P}_r \bar{u} - \bar{U} \rangle dt
\]
\[
= I_2 + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle E_{t_j} P_{j+1}, \bar{U} - \bar{u} \rangle dt + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle \bar{P}_{j+1}, \mathcal{P}_r \bar{u} - \bar{U} \rangle dt \text{ (by (4.53))}
\]
\[
= I_2 + \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle P_{j+1} - \bar{P}_{j+1}, \bar{U} - \mathcal{P}_r \bar{u} \rangle dt.
\]
by the definition of $P_\tau$. Since

$$
\sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (P_{j+1} - P_j, \bar{U} - P_\tau \bar{u})dt
= \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} \langle (S_1^{h,\tau} (S_0 R \bar{u} - S_0^{h,\tau} R \bar{U} + G - G_{h,\tau}))_{j+1}, \bar{U} - P_\tau \bar{u} \rangle dt \text{ (by (3.8) and (4.45))}
= \int_0^T [S_0 R \bar{u} - S_0^{h,\tau} R \bar{U} + G - G_{h,\tau}, S_0^{h,\tau} R (\bar{U} - P_\tau \bar{u})]dt \text{ (by (3.6))}
= I_1 + \int_0^T [S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}, S_0^{h,\tau} R (\bar{U} - P_\tau \bar{u})]dt \text{ (by (4.52)),}
$$

it follows that

$$
\nu \|\bar{u} - \bar{U}\|^2_{L^2(0,T;L^2(\Gamma))} \leq I_1 + I_2 + \int_0^T [S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}, S_0^{h,\tau} R (\bar{U} - P_\tau \bar{u})]dt. \tag{4.56}
$$

We also have

$$
\int_0^T [S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}, S_0^{h,\tau} R (\bar{U} - P_\tau \bar{u})]dt
= -\|S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}\|^2_{L^2(0,T;H^v)} + \int_0^T [S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}, S_0 R \bar{u} - S_0^{h,\tau} R P_\tau \bar{u}]dt
\leq -\frac{1}{2} \|S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}\|^2_{L^2(0,T;H^v)} + \frac{1}{2} \|S_0 R \bar{u} - S_0^{h,\tau} R P_\tau \bar{u}\|^2_{L^2(0,T;H^v)}
= -\frac{1}{2} \|S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}\|^2_{L^2(0,T;H^v)} + I_0 \text{ (by (4.51)).}
$$

Inserting the above inequality into (4.56) yields

$$
\nu \|\bar{u} - \bar{U}\|^2_{L^2(0,T;H^v)} \leq I_1 + I_2 - \frac{1}{2} \|S_0 R \bar{u} - S_0^{h,\tau} R \bar{U}\|^2_{L^2(0,T;H^v)} + I_0,
$$

and then a simple calculation proves (4.50).

**Step 2.** Let us estimate $I_0$, $I_1$ and $I_2$. For $I_0$, we have

$$
I_0 = \frac{1}{2} \| (S_0 - S_0^{h,\tau}) R \bar{u} + S_0^{h,\tau} R (I - P_\tau) \bar{u} \|^2_{L^2(0,T;H^v)} \text{ (by (4.51))}
\leq \| (S_0 - S_0^{h,\tau}) R \bar{u} \|^2_{L^2(0,T;H^v)} + \| S_0^{h,\tau} R (I - P_\tau) \bar{u} \|^2_{L^2(0,T;H^v)}
\leq \| (S_0 - S_0^{h,\tau}) R \bar{u} \|^2_{L^2(0,T;H^v)} + C \| (I - P_\tau) \bar{u} \|^2_{L^2(0,T;\dot{L}^2(\Gamma))} \text{ (by (3.3))}
$$

so that (4.46) implies

$$
I_0 \leq C (\tau^{3/4-\epsilon} + h^{3/2-\epsilon})^2 + C \| (I - P_\tau) \bar{u} \|^2_{L^2(0,T;\dot{L}^2(\Gamma))}. \tag{4.57}
$$

For $I_1$, we have

$$
I_1 \leq \| G - G_{h,\tau} \|^2_{L^2(0,T;H^v)} \| S_0^{h,\tau} R (\bar{U} - P_\tau \bar{u}) \|_{L^2(0,T;H^v)} \text{ (by (4.52))}
\leq C \| G - G_{h,\tau} \|^2_{L^2(0,T;H^v)} \| \bar{U} - P_\tau \bar{u} \|_{L^2(0,T;\dot{L}^2(\Gamma))} \text{ (by (3.3))}
\leq C \| G - G_{h,\tau} \|^2_{L^2(0,T;H^v)} \| \bar{u} - \bar{U} \|_{L^2(0,T;\dot{L}^2(\Gamma))} \text{ (by (4.49))}
$$

so that by (4.25), we obtain

$$
I_1 \leq C (\tau^{3/4-\epsilon} + h^{1/2-\epsilon}) \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))}. \tag{4.58}
$$
Now let us estimate $I_2$. By definition, we have

$$(P_\tau \tilde p)(t_j) = (P_\tau \mathcal{E}_\tau \tilde p)(t_j), \quad 0 \leq j < J,$$

so

$$
\sum_{j=0}^{J-1} \| \mathcal{E}_\tau \tilde p - (P_\tau \tilde p)(t_j) \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 = \| (I - P_\tau) \mathcal{E}_\tau \tilde p \|_{L^2(0, T; L^2(\Gamma))}. \tag{4.59}
$$

We also have

$$
\sum_{j=0}^{J-1} \| (P_\tau \tilde p)(t_j) - E_{t_j} P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2
= \sum_{j=0}^{J-1} \tau \| (P_\tau \tilde p)(t_j) - E_{t_j} P_{j+1} \|_{L^2(\Gamma)}^2
= \sum_{j=0}^{J-1} \tau^{-1} \left\| E_{t_j} \int_{t_j}^{t_{j+1}} (\tilde p(t) - P_{j+1}) dt \right\|_{L^2(\Gamma)}^2
\leq \sum_{j=0}^{J-1} \tau^{-1} \left\| \int_{t_j}^{t_{j+1}} (\tilde p(t) - P_{j+1}) dt \right\|_{L^2(\Gamma)}^2
\leq \sum_{j=0}^{J-1} \| \tilde p - P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2.
$$

Combining the above estimate with (4.59) yields

$$
\left( \sum_{j=0}^{J-1} \| \mathcal{E}_\tau \tilde p - E_{t_j} P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2}
\leq \left( \sum_{j=0}^{J-1} \| \mathcal{E}_\tau \tilde p - (P_\tau \tilde p)(t_j) \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2}
+ \left( \sum_{j=0}^{J-1} \| (P_\tau \tilde p)(t_j) - E_{t_j} P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2}
\leq \| (I - P_\tau) \mathcal{E}_\tau \tilde p \|_{L^2(0, T; L^2(\Gamma))} + \left( \sum_{j=0}^{J-1} \| \tilde p - P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2}.
$$

Hence,

$$
I_2 = \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (\mathcal{E}_\tau \tilde p - E_{t_j} P_{j+1}, \tilde u - \bar u) dt \quad \text{(by (4.53))}
= \sum_{j=0}^{J-1} \int_{t_j}^{t_{j+1}} (\mathcal{E}_\tau \tilde p - E_{t_j} P_{j+1}, \tilde u - \bar u) dt \quad \text{(by (4.43))}
\leq \left( \sum_{j=0}^{J-1} \| \mathcal{E}_\tau \tilde p - E_{t_j} P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2} \| \tilde u - \bar u \|_{L^2(0, T; L^2(\Gamma))}
\leq \left( \| (I - P_\tau) \mathcal{E}_\tau \tilde p \|_{L^2(0, T; L^2(\Gamma))} + \left( \sum_{j=0}^{J-1} \| \tilde p - P_{j+1} \|_{L^2(t_j, t_{j+1}; L^2(\Gamma))}^2 \right)^{1/2} \right)
\times \| \tilde u - \bar u \|_{L^2(0, T; L^2(\Gamma))},
$$

which together with (4.47) implies

$$
I_2 \leq \left( C(\tau^{3/4-\epsilon} + h^{3/2-\epsilon}) + \| (I - P_\tau) \mathcal{E}_\tau \tilde p \|_{L^2(0, T; L^2(\Gamma))} \right) \| \tilde u - \bar u \|_{L^2(0, T; L^2(\Gamma))}. \tag{4.60}
$$
Step 3. Combining (4.50), (4.57), (4.58) and (4.60), we conclude that

\[
\nu \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))} + \| S_0 \mathcal{R} \bar{u} - S_0^{h,\tau} \mathcal{R} \bar{U} \|_{L^2(0,T;H^0)} \\
\leq C (\tau^{3/4-\epsilon} + h^{3/2-\epsilon})^2 + \| (I - P_\tau) \bar{u} \|_{L^2(0,T;L^2(\Gamma))} \\
+ (\tau^{1/4-\epsilon} + h^{1/2-\epsilon}) \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))} \\
+ \| (I - P_\tau) \mathcal{E}_p \bar{p} \|_{L^2(0,T;L^2(\Gamma))} \| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))}).
\]

Using Young’s inequality with \(\epsilon\), we then obtain

\[
\| \bar{u} - \bar{U} \|_{L^2(0,T;L^2(\Gamma))} + \| S_0 \mathcal{R} \bar{u} - S_0^{h,\tau} \mathcal{R} \bar{U} \|_{L^2(0,T;H^0)} \\
\leq C (\tau^{1/4-\epsilon} + h^{1/2-\epsilon}) + \| (I - P_\tau) \bar{u} \|_{L^2(0,T;L^2(\Gamma))} + \| (I - P_\tau) \mathcal{E}_p \bar{p} \|_{L^2(0,T;L^2(\Gamma))},
\]

namely, the inequality (3.11). This completes the proof. \(\square\)

5 Conclusion

In this paper, we analyze the discretization of a Neumann boundary control problem, where the state equation is a linear stochastic parabolic equation with the boundary noise. With the rough data, we establish the convergence for general filtration, and we derive the convergence rate \(O(\tau^{1/4-\epsilon} + h^{1/2-\epsilon})\) for the natural filtration of the \(Q\)-Wiener process, where \(0 < \epsilon < 1/4\) can be arbitrarily small.

The numerical analysis in this paper can be extended to the distributed optimal control problems governed by linear stochastic parabolic equations with the additive noise and general filtration. For the case where the noise is multiplicative or the coefficient before the \(Q\)-Wiener process contains the control variable, there will be essential difficulty in the numerical analysis, especially for general filtration. To solve these issues is our ongoing work.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 11901410) and the Fundamental Research Funds for the Central Universities in China (Grant No. 2020SCU12063).

References

1. Bender C, Denk R. A forward scheme for backward SDEs. Stochastic Process Appl, 2007, 117: 1793–1812
2. Bensoussan A, Stochastic maximum principle for distributed parameter systems. J Franklin Inst, 1983, 315: 387–406
3. Bouchard B, Touzi N. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. Stochastic Process Appl, 2004, 111: 175–205
4. Brenner S C, Scott R. The Mathematical Theory of Finite Element Methods. New York: Springer-Verlag, 2008
5. Choi Y, Lee H C. Error analysis of finite element approximations of the optimal control problem for stochastic Stokes equations with additive white noise. Appl Numer Math, 2018, 133: 144–160
6. Da Prato G, Zabczyk J. Evolution equations with white-noise boundary conditions. Stochastics Stochastics Rep, 1993, 42: 167–182
7. Da Prato G, Zabczyk J. Stochastic Equations in Infinite Dimensions, 2nd ed. Cambridge: Cambridge University Press, 2014
8. Debussche A, Fuhrman M, Tessitore G. Optimal control of a stochastic heat equation with boundary-noise and boundary-control. ESAIM Control Optim Calc Var, 2007, 13: 178–205
9. Du K, Meng Q X. A maximum principle for optimal control of stochastic evolution equations. SIAM J Control Optim, 2013, 51: 4343–4362
10. Dunst T, Prohl A. The forward-backward stochastic heat equation: Numerical analysis and simulation. SIAM J Sci Comput, 2016, 38: A2725–A2755
11. Fuhrman M, Hu Y, Tessitore G. Stochastic maximum principle for optimal control of SPDEs. C R Math Acad Sci Paris, 2012, 350: 683–688
12. Fuhrman M, Hu Y, Tessitore G. Stochastic maximum principle for optimal control of SPDEs. Appl Math Optim, 2013, 68: 181–217
13. Gawalecki L, Mandrekar V. Stochastic Differential Equations in Infinite Dimensions with Applications to Stochastic Partial Differential Equations. Berlin-Heidelberg: Springer-Verlag, 2011
Appendix A An interpolation space result

Throughout this appendix, let \((\cdot, \cdot)_{1-\theta, 2}\) \((0 < \theta < 1)\) denote the interpolation space defined by the \(K\)-method (see [29, Chapter 1]).

**Lemma A.1.** The space \(L^2_\theta(\Omega; L^2_\lambda(0, T; L^2_\lambda, \hat{H}^{-\theta}))\) is continuously embedded into
\[
(L^2_\theta(\Omega; L^2_\lambda(0, T; L^2_\lambda, \hat{H}^{-1}))), (L^2_\theta(\Omega; L^2_\lambda(0, T; L^2_\lambda, \hat{H}^0)))_{1-\theta, 2},
\]
where \(0 < \theta < 1\).

**Proof.** For any \(B \in \mathcal{L}_2(L^2_\lambda, (\hat{H}^{-1}, \hat{H}^0)_{1-\theta, 2})\), by definition we have
\[
\|B\|^2_{(\mathcal{L}_2(L^2_\lambda, \hat{H}^{-1}), \mathcal{L}_2(L^2_\lambda, \hat{H}^0))_{1-\theta, 2}}
\]
Proof. With equivalent norms. This completes the proof.

Therefore, the desired claim follows from the result that (see [29, p. 38])

\[
\begin{align*}
B \in L^2(B^0, H^{-1}) \\
B_0 \in L^2(B^0, H^{-1}) \\
B_1 \in L^2(B^0, H^0)
\end{align*}
\]

\[
\inf_{B = B_0 + B_1} \left( \|B_0\|_{L^2(B^0, H^{-1})} + t\|B_1\|_{L^2(B^0, H^0)} \right)^2 dt
\]

\[
\leq 2 \int_0^\infty t^{2\theta - 3} \inf_{B = B_0 + B_1} \left( \|B_0\|_{L^2(B^0, H^{-1})} + t^2\|B_1\|_{L^2(B^0, H^0)} \right) dt
\]

\[
= 2 \int_0^\infty t^{2\theta - 3} \inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t^2\|B_1 \phi_n\|_{H^0}) dt
\]

\[
= 2 \int_0^\infty t^{2\theta - 3} \inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t\|B_1 \phi_n\|_{H^0}) dt
\]

\[
= 2 \int_0^\infty t^{2\theta - 3} \inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t\|B_1 \phi_n\|_{H^0}) dt
\]

This implies that \(L^2(B^0, H^{-1}, H^0)_{1-\theta, 2}\) is continuously embedded into \((L^2(B^0, H^{-1}), L^2(B^0, H^0))_{1-\theta, 2}\).

Since \(H^{-\theta} = (H^{-1}, H^0)_{1-\theta, 2}\) with equivalent norms, we readily conclude that \(L^2(B^0, H^{-\theta})\) is continuously embedded into \((L^2(B^0, H^{-1}), L^2(B^0, H^0))_{1-\theta, 2}\).

Therefore, the desired claim follows from the result that (see [29, p. 38])

\[
L^2(B^0, H^{-1}, H^0)_{1-\theta, 2}
\]

is identical to

\[
(L^2(\Omega; L^2(0, T; L^2(B^0, H^{-1}))), L^2(\Omega; L^2(0, T; L^2(B^0, H^0))))_{1-\theta, 2}
\]

with equivalent norms. This completes the proof. \(\square\)

**Lemma A.2.** For any \(t > 0\) and \(B \in L^2(B^0, (H^{-1}, H^0)_{1-\theta, 2})\) with \(0 < \theta < 1\), we have

\[
\inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t^2\|B_1 \phi_n\|_{H^0}) = \sum_{n=0}^{\infty} \lambda_n \inf_{B = B_0 + B_1} \left( \|v_0\|_{H^{-1}} + t^2\|v_1\|_{H^0} \right).
\]

**Proof.** By the evident inequality

\[
\inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t^2\|B_1 \phi_n\|_{H^0}) \geq \sum_{n=0}^{\infty} \lambda_n \inf_{B = B_0 + B_1} \left( \|v_0\|_{H^{-1}} + t^2\|v_1\|_{H^0} \right),
\]

\[
\inf_{B = B_0 + B_1} \sum_{n=0}^{\infty} \lambda_n (\|B_0 \phi_n\|_{H^{-1}} + t^2\|B_1 \phi_n\|_{H^0}) \leq \sum_{n=0}^{\infty} \lambda_n \inf_{B = B_0 + B_1} \left( \|v_0\|_{H^{-1}} + t^2\|v_1\|_{H^0} \right).
\]
it remains only to prove
\[
\inf_{B=B_0+B_1} \sum_{n=0}^{\infty} \lambda_n (\| B_0 \phi_n \|_{H^{-1}}^2 + t^2 \| B_1 \phi_n \|_{H^0}^2) \leq \sum_{n=0}^{\infty} \lambda_n \inf_{B_0=v_0, B_1=v_1} (\| v_0 \|_{H^{-1}}^2 + t^2 \| v_1 \|_{H^0}^2). \tag{A.2}
\]

To this end, we proceed as follows. Let \( \{ v_{0,n} \mid n \in \mathbb{N} \} \subset \dot{H}^{-1} \) and \( \{ v_{1,n} \mid n \in \mathbb{N} \} \subset \dot{H}^0 \) be arbitrary such that
\[
B \phi_n = v_{0,n} + v_{1,n}, \quad \forall \ n \in \mathbb{N}
\]
and
\[
\inf_{\| v_0 \|_{H^{-1}} = \| v_1 \|_{H^0}} \sum_{n=0}^{\infty} \lambda_n (\| v_{0,n} \|_{H^{-1}}^2 + t^2 \| v_{1,n} \|_{H^0}^2) < \infty. \tag{A.3}
\]

Define \( B_0 \in \mathcal{L}_2(L^2, \dot{H}^{-1}) \) by \( B_0 \phi_n := v_{0,n}, \forall \ n \in \mathbb{N} \), and define \( B_1 \in \mathcal{L}_2(L^2, \dot{H}^0) \) by \( B_1 \phi_n := v_{1,n}, \forall \ n \in \mathbb{N} \). By (A.3), it is easily verified that the above \( B_0 \) and \( B_1 \) are both well defined. Clearly, we have
\[
B = B_0 + B_1 \quad \text{and} \quad \sum_{n=0}^{\infty} \lambda_n (\| B_0 \phi_n \|_{H^{-1}}^2 + t^2 \| B_1 \phi_n \|_{H^0}^2) = \sum_{n=0}^{\infty} \lambda_n (\| v_{0,n} \|_{H^{-1}}^2 + t^2 \| v_{1,n} \|_{H^0}^2).
\]

Because of the arbitrary choices of \( \{ v_{0,n} \mid n \in \mathbb{N} \} \) and \( \{ v_{1,n} \mid n \in \mathbb{N} \} \), we readily conclude (A.2). This completes the proof. \( \square \)